Atomistic T-matrix theory of disordered two-dimensional materials: Bound states, spectral properties, quasiparticle scattering, and transport

Kaasbjerg, Kristen

Published in:
Physical Review B (Condensed Matter and Materials Physics)

Link to article, DOI:
10.1103/PhysRevB.101.045433

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Atomistic $T$-matrix theory of disordered two-dimensional materials: Bound states, spectral properties, quasiparticle scattering, and transport

Kristen Kaasbjerg
Center for Nanostructured Graphene (CNG), Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

(Received 1 November 2019; published 31 January 2020)

In this work, we present an atomistic first-principles framework for modeling the low-temperature electronic and transport properties of disordered two-dimensional (2D) materials with randomly distributed point defects (impurities). The method is based on the $T$-matrix formalism in combination with realistic density-functional theory descriptions of the defects and their scattering matrix elements. From the $T$-matrix approximations to the disorder-averaged Green’s function and the collision integral in the Boltzmann transport equation, the method allows calculations of, e.g., the density of states including contributions from bound defect states, the quasiparticle spectrum and the spectral linewidth (scattering rate), and the conductivity/mobility of disordered 2D materials. We demonstrate the method by examining these quantities in monolayers of the archetypal 2D materials graphene and transition metal dichalcogenides contaminated with vacancy defects and substitutional impurity atoms. By comparing the Born and $T$-matrix approximations, we also demonstrate a strong breakdown of the Born approximation for defects in 2D materials manifested in a pronounced renormalization of, e.g., the scattering rate by the higher-order $T$-matrix method. As the $T$-matrix approximation is essentially exact for dilute disorder, i.e., low defect concentrations ($c_{\text{dis}} \ll 1$) or density ($n_{\text{dis}} \ll A_{\text{cell}}^{-1}$ where $A_{\text{cell}}$ is the unit-cell area), our first-principles method provides an excellent framework for modeling the properties of disordered 2D materials with defect concentrations relevant for devices.

DOI: 10.1103/PhysRevB.101.045433

I. INTRODUCTION

Over the past decade, there has been an explosive development in theoretical predictions [1–4] and experimental fabrication of new two-dimensional (2D) materials hosting exciting electronic properties. This holds great promise for novel applications in electronics, optoelectronics, and other emerging disciplines. However, atomic disorder which degrades the material properties is still a major hindrance, and fabrication platforms that can deliver high-quality materials with low disorder concentrations are needed. Recently, there have been advances in some of the most widely studied 2D materials such as, e.g., monolayers of graphene and transition metal dichalcogenides (TMDs), where devices based on high-quality materials encapsulated in ultraclean van der Waals (vdW) heterostructures have shown promising electrical and optical properties [5]. Such developments are essential for the realization of quantum devices based on 2D materials [6,7].

The initial characterization of atomic disorder due to point defects in 2D materials often proceeds by means of scanning tunneling microscopy/spectroscopy (STM/STS) which provides valuable insight into the defect type as well as the structural and electronic properties of the defects. In, for example, graphene [8–10] and TMDs [11–16] this has been useful for the identification of the most common types of defects as well as probing for bound defect states which lead to strong modifications of the electronic properties of the pristine material. In addition, measurements of quasiparticle interference in various 2D materials [8,17–24], i.e., spatial ripples in the local density of states (LDOS) in the vicinity of a defect, is a direct fingerprint of the defect-induced scattering processes which may govern the electron dynamics and limit the electrical and optical performance of materials at low temperatures. Theoretical methods which can supplement such experiments in predicting the impact of defects on the electron dynamics are of high value for the understanding of electrical and optical properties of new materials.

In this work, we introduce an atomistic first-principles method for modeling the electronic properties of disordered 2D materials. Our method is based on realistic density-functional theory (DFT) calculations of the defect scattering potential and matrix elements, in combination with the $T$-matrix formalism [25,26] for the description of the interaction with the random disorder potential. From the disorder-averaged Green’s function (GF), accurate descriptions of experimentally relevant quantities such as, e.g., the density of states, in-gap bound and resonant quasibound defect states, spectral properties, and the disorder-induced quasiparticle scattering rate/lifetime can be obtained. Furthermore, using the $T$-matrix scattering amplitude in the calculation of the momentum relaxation time in the Boltzmann transport equation allows for theoretical predictions of the disorder-limited low-temperature conductivity/mobility as well as its dependence on the Fermi energy (carrier density). This work is thus complementary to our previous first-principles $T$-matrix study of the LDOS and quasiparticle interference in 2D materials [27].
In comparison with analytic and tight-binding based $T$-matrix studies of defects in, e.g., graphene [28–35] and black phosphorus [36], our first-principles method permits for parameter-free modeling of realistic defects in disordered materials. It furthermore goes beyond other first-principles studies of defects and their transport-limiting effects based on the Born approximation [37–41], which we here demonstrate breaks down for point defects in 2D materials. The first-principles $T$-matrix method introduced in this work is therefore of high relevance for the further development of first-principles transport methodologies with high predictive power [41–47].

The details of our method, which is implemented in the GPAW electronic-structure code [48–50], are described in Secs. II and III. In Secs. IV and V, we demonstrate the power of our method on a series of timely problems in disordered monolayers of TMDs and graphene. For the TMDs MoS$_2$ and WSe$_2$ with vacancies and oxygen substitutional, we analyze (i) bound and quasibound defect states, respectively, in the gap and in the bands, (ii) the linewidth in the quasiparticle spectrum, the energy dependence of the scattering rate in the $K, K'$ valleys, and the complete suppression of intervalley scattering by the spin-orbit splitting and a symmetry-induced selection rule [27], and (iii) a prediction of unconventional transport characteristics in $p$-type WSe$_2$ with a mobility that decreases with increasing Fermi energy. In graphene we focus on vacancies and nitrogen substitutionals and examine (i) the position of quasibound defect states on the Dirac cone, (ii) their signature in the quasiparticle spectrum and the presence (absence) of a band-gap opening for sublattice asymmetric (symmetric) defect configurations, and (iii) the pronounced electron-hole asymmetry in the transport characteristics induced by strong resonant scattering.

An important finding of this work is that the Born approximation breaks down for point defects in both TMDs and graphene, and hence most likely also in other 2D materials. While it is well known that the description of quasibound states and resonant scattering in graphene is beyond the Born approximation, our finding that it severely overestimates the disorder-induced scattering rate in the 2D TMDs by up to several orders of magnitude is remarkable, and only emphasizes the high relevance of a first-principles $T$-matrix approach for the modeling of disordered 2D materials.

II. ATOMICISTIC DEFECT POTENTIALS

We start by introducing an atomistic first-principles method for the calculation of the single-defect (or impurity) potential $\hat{V}_d$ and its matrix elements which are the basic building block in the diagrammatic $T$-matrix formalism for disordered systems outlined in Sec. III. The method is analogous to the method for calculating the electron-phonon interaction [51–53], and is based on DFT within the projector augmented-wave (PAW) method [54], a linear combination of atomic orbitals (LCAO) supercell representation of the defect potential, and is implemented in the GPAW electronic-structure code [48–50].

In this work, we restrict the considerations to nonmagnetic spin-diagonal defects in which case the defect potential for a defect of type $i$ takes the form

$$\hat{V}_i = V_i(\textbf{r}) \otimes \hat{S}_0,$$

where $V_i(\textbf{r})$ is the scalar spin-independent defect potential and $\hat{S}_0$ is the identity operator in spin space. We thus neglect defect-induced changes in the spin-orbit interaction, which are, in general, small relative to the spin-independent potential. The spin dependence is, of course, important for spin relaxation and spin-orbit scattering, but this is outside the scope of this work.

Here, the spin-diagonal defect potential is defined as the change in the microscopic crystal potential induced by the defect, and is obtained from DFT as the difference in the crystal potential between the lattice with a defect and the pristine lattice, i.e.,

$$V_i(\textbf{r}) = V_{\text{def}}(\textbf{r}) - V_{\text{pris}}(\textbf{r}).$$

The two potentials have contributions from the atomic cores (ions), which define the overall potential landscape in the lattice, as well as from the valence electrons which describe interactions between the valence electrons at a mean-field level (see Sec. II B below). The defect potential in Eq. (2) thus carries information about (i) the defect-induced lattice imperfection (e.g., vacancy, substitutional, or impurity atom), and (ii) the electronic relaxation in the vicinity of the defect. Both are important for a quantitative description of the defect potential.

In practice, the defect potential is expressed in a basis of Bloch states $|\mathbf{n}\mathbf{k}\rangle$ of the pristine lattice, where $n$ is the band index, $\mathbf{k} \in 1$st Brillouin zone (BZ) is the electronic wave vector, and $s$ is the spin index. For brevity, we combine in the following the band and spin indices in a composite “band” index. The matrix elements of the defect potential become

$$V'_{\text{def}}^{mn}_{\mathbf{r},\mathbf{k}k} = \langle m\mathbf{k}|V_i|n\mathbf{k}'\rangle = \sum_{s_\mathbf{c}} \langle m\mathbf{k}; s_\mathbf{c}|V_i(\textbf{r})|n\mathbf{k}'; s_\mathbf{c}\rangle,$$

where as a consequence of the spin-orbit mixing of up and down spin ($s_\mathbf{c} = \pm 1$) in the Bloch states, the matrix elements, in general, have contributions from both spin components $|\cdot; s_\mathbf{c}\rangle$ in spite of the fact that the defect potential itself is spin diagonal.

The following two subsections summarize our DFT-based supercell method for the calculation of the defect matrix elements. The two main technical aspects of the method concern (i) the representation of the defect potential in an LCAO basis, and (ii) the calculation of the defect potential in the PAW formalism [54].

A. LCAO supercell representation

The numerical evaluation of the defect matrix element in Eq. (3) is based on an LCAO expansion of the Bloch functions of the pristine lattice $|\phi_{\mu \mathbf{k}}\rangle = \sum_{\alpha s_\mathbf{c}} c_{\mu s_\mathbf{c}}^{\mathbf{k}} |\phi_{\alpha s_\mathbf{c}}\rangle$, where $\mu = (\alpha, i)$ is a composite atomic ($\alpha$) and orbital index ($i$) and

$$|\phi_{\mu \mathbf{k}}\rangle = \frac{1}{\sqrt{N}} \sum_{i} e^{i \mathbf{k} \mathbf{R}_i} |\phi_{\mu i}\rangle$$

are Bloch expansions of the spin-independent LCAO basis orbitals $|\phi_{\mu i}\rangle$, where $N$ is the number of unit cells in the lattice.
and \( \mathbf{R}_i = l_1 \mathbf{a}_1 + l_2 \mathbf{a}_2, l_i \in \mathbb{Z} \), is the lattice vector to the \( i \text{th} \) unit cell with \( \mathbf{a}_i \) denoting the primitive lattice vectors.

Inserting in the expression for the matrix element in Eq. (3), we find

\[
V_{i\mathbf{k}k}^{mn} = \sum_{c_{\mathbf{mk}}} \sum_{\mu\nu} \left( e^{i\mathbf{k}_c \cdot \mathbf{R}_c} \right)^{\alpha_{\mathbf{mk}}} \left( e^{i\mathbf{k}_c \cdot \mathbf{R}_c} \right)^{\alpha_{\mathbf{mk}}} \langle \phi_{\mu \mathbf{k}} | V_i(\mathbf{r}) | \phi_{\nu \mathbf{k}} \rangle
\]

where the factor of \( 1/N \) stems from the normalization of the Bloch sum in Eq. (4) to the lattice, the last factor in the second equality is the LCAO representation of the defect potential \( V_i(\mathbf{r}) \) illustrated in Fig. 1, and the \( k, l \) sums run over the cells in the lattice.

In practice, the defect potential is calculated in a finite \( N_1 \times N_2 \) supercell constructed by repeating the primitive unit cell \( N_1 \) times in the direction of the \( i \text{th} \) primitive lattice vector, and with the defect site located at the center. Due to periodic boundary conditions in the in-plane directions, the supercell must be chosen large enough that defect sites in neighboring supercells do not interact. In the direction perpendicular to the material plane, the cell boundaries are imposed with Dirichlet boundary conditions, which ensures a common reference for the two potentials on the right-hand side of Eq. (2) and avoids spurious interactions between repetitions of the defect in the perpendicular direction.

In order to impose an isotropic range of the defect potential, matrix elements involving LCAO basis functions located beyond a cutoff distance \( r_{\text{cut}} \) from the defect site \( \mathbf{R}_0 \) are zeroed, i.e.,

\[
\langle \phi_{\mu \mathbf{k}} | V_i(\mathbf{r}) | \phi_{\nu \mathbf{k}} \rangle = 0 \quad \text{if} \quad | \mathbf{R}_{\mu \mathbf{k},\nu \mathbf{k}} - \mathbf{R}_0 | > r_{\text{cut}},
\]

where \( \mathbf{R}_{\mu \mathbf{k},\nu \mathbf{k}} = \mathbf{R}_k + \mathbf{R}_\alpha \) is the center of the LCAO orbital \( | \phi_{\mu \mathbf{k}} \rangle \) at the atomic site \( \alpha \) in unit cell \( k \). Once the LCAO representation of the defect potential in the supercell has been obtained, the defect matrix elements can be evaluated efficiently at arbitrary \( \mathbf{k}, \mathbf{k}' = \mathbf{k} + \mathbf{q} \) vectors using Eq. (5).

It should be noted that the LCAO procedure for the calculation of the defect matrix elements outlined here bears close resemblance to methods based on Wannier functions [47,55]. However, the use of a fixed LCAO basis has the advantage that the additional step for the generation of the Wannier functions, which is not always trivial, is avoided.

### B. PAW method

In the PAW formulation to DFT [54], the basic idea is to transform the all-electron Hamiltonian, whose eigenstates \( | \psi_{\alpha \mathbf{k}} \rangle \) oscillate strongly in the vicinity atomic cores, into an auxiliary Hamiltonian with smooth pseudoeigenstates \( | \tilde{\psi}_{\alpha \mathbf{k}} \rangle \), thereby eliminating the numerical complications associated with an accurate description of rapidly varying functions.

The physically relevant all-electron wave functions and the auxiliary pseudo-wave functions are connected via the transformation \( \tilde{T} \) defined as

\[
| \psi_{\alpha \mathbf{k}} \rangle = | \tilde{\psi}_{\alpha \mathbf{k}} \rangle + \sum_{a,i} \left[ | \tilde{\varphi}^{\alpha}_{\mu} \rangle - | \varphi^{\alpha}_{\mu} \rangle \right] | \tilde{p}^{\alpha}_{\mu} \rangle \tilde{T} | \tilde{\psi}_{\alpha \mathbf{k}} \rangle \equiv \tilde{T} | \tilde{\psi}_{\alpha \mathbf{k}} \rangle,
\]

where the terms in the sum over atomic sites \( a \), respectively, add and subtract expansions of the all-electron and pseudo-wave functions inside so-called augmentation spheres \( \Omega_a \) centered on the atoms. Here, \( | \varphi^{\alpha}_{\mu} \rangle \) are the correct all-electron wave functions inside the augmentation spheres, and the pseudopartial waves \( | \tilde{\varphi}^{\alpha}_{\mu} \rangle \) and projector functions \( | \tilde{p}^{\alpha}_{\mu} \rangle \) are constructed to obey the completeness relation \( \sum | \tilde{\varphi}^{\alpha}_{\mu} \rangle \langle \tilde{\varphi}^{\alpha}_{\mu} | = 1 \). This ensures the orthogonality of the all-electron wave functions

\[
\langle \psi_{\alpha \mathbf{k}} | \psi_{\alpha \mathbf{k}} \rangle = \langle \tilde{\psi}_{\alpha \mathbf{k}} | \tilde{T}^\dagger \tilde{T} | \tilde{\psi}_{\alpha \mathbf{k}} \rangle = \delta_{\alpha \mathbf{k}} \delta_{\alpha \mathbf{k}},
\]

via the operator \( \tilde{T}^\dagger \tilde{T} \).

Likewise, the all-electron matrix elements of the defect potential \( \tilde{V}_i \) can be expressed as a matrix element of a transformed operator with respect to the smooth wave functions \( | \tilde{\psi}_{\alpha \mathbf{k}} \rangle \), i.e.,

\[
V_{i\mathbf{k}k}^{mn} = \langle \psi_{\alpha \mathbf{k}} | \tilde{V}_i | \psi_{\alpha \mathbf{k}} \rangle = \langle \tilde{\psi}_{\alpha \mathbf{k}} | \tilde{T}^\dagger \tilde{V}_i \tilde{T} | \tilde{\psi}_{\alpha \mathbf{k}} \rangle \equiv \langle \tilde{\psi}_{\alpha \mathbf{k}} | \tilde{\hat{V}}_i | \tilde{\psi}_{\alpha \mathbf{k}} \rangle,
\]

where the transformed operator \( \tilde{V}_i \) in the last line is given by

\[
\tilde{V}_i = \tilde{T}^\dagger \tilde{T} \tilde{V}_i \tilde{T} = \left( 1 + \sum_{a,i} | \tilde{p}^a_{\mu i} \rangle \langle \tilde{p}^a_{\mu i} | \right) \tilde{V}_i \times \left( 1 + \sum_{a,i} | \tilde{\varphi}^a_{\mu i} \rangle \langle \tilde{\varphi}^a_{\mu i} | \right) \tilde{V}_i
\]

\[
\approx \tilde{V}_i + \sum_{a,i} | \tilde{p}^a_{\mu i} \rangle \Delta V_{i,i}^{n\alpha} | \tilde{p}^a_{\mu i} \rangle.
\]

and the atomic coefficients are defined as

\[
\Delta V_{i,i}^{n\alpha} = \langle \tilde{\varphi}^a_{\mu i} | \tilde{V}_i | \tilde{\varphi}^a_{\mu i} \rangle - \langle \tilde{\varphi}^a_{\mu i} | \tilde{V}_i | \tilde{\varphi}^a_{\mu i} \rangle.
\]

The last line in Eq. (10) holds for local operators and furthermore assumes that the bases are complete and that the atomic augmentation spheres \( \Omega_a \) do not overlap. However, in practical PAW calculations, finite bases and small overlaps between different augmentation spheres can be tolerated without substantial loss of accuracy.
The transformed operator, which incorporates the full details of the potential due to the all-electron density (frozen core + valence electrons), can be expressed as
\[
\hat{V} = \nu_{\text{eff}}(\hat{r}) + \sum_{\alpha} \sum_{i_1i_2} |\tilde{p}_{i_1}^{\alpha}| \Delta V_{i_1i_2}^{\alpha} / |\tilde{p}_{i_2}^{\alpha}|,
\]
(12)
where \(\nu_{\text{eff}} = \nu_0 + \nu_{\text{xc}}\) is the effective potential given by the sum of the electrostatic Hartree potential \(\nu_0\) (including the potential due to the atomic cores) and the exchange-correlation potential \(\nu_{\text{xc}}\), and the last term is the all-electron corrections given by the atomic coefficients \(\Delta V_{i_1i_2}^{\alpha}\) defined in Eq. (11). Full PAW expressions for \(\nu_{\text{eff}}\) and \(\Delta V_{i_1i_2}^{\alpha}\) can be found in, e.g., Ref. [50].

In the defect potential in Eq. (2), the two contributions to the potential in Eq. (12) describe, respectively, perturbations in the crystal potential and atomic-core states on the defect. The latter term can be regarded as the PAW analog of a Löwdin downfolding of atomic defect states onto the Bloch functions [34].

C. Examples

In this section, we show examples of matrix elements for the defects in 2D TMDs and graphene studied in Secs. IV and V below.

In order to relate the DFT-calculated matrix elements (which have units of energy) to the impurity strength \(V_i\) below.

\[
\langle nKs\hat{V}_i|nK's\rangle = 0,
\]
(14)
where \(\vec{s} \neq s\). Note that this holds even in the presence of spin-orbit coupling in the defect potential which does not break time-reversal symmetry.

2. Graphene

Graphene is a host of a large variety of defects ranging from vacancies and lattice reconstructions like, e.g., Stone-Wales defects, to adatoms and substitutional atoms.
in Fig. 2, there are no selection rules on the matrix elements for vacancies and substitutionals in graphene. Consequently, their matrix elements show less variation as a function of wave vectors compared to single carbon vacancies [74–77] and nitrogen substitutionals [9,78,79]

In Fig. 3 we show the valence and conduction band matrix elements for a carbon vacancy (left) and a nitrogen substitutional (right). In contrast to the matrix elements for TMDs in Fig. 2, there are no selection rules on the matrix elements for vacancies and substitutionals in graphene. Consequently, their matrix elements show less variation as a function of k'.

FIG. 2. Defect matrix elements for vacancies (V, i = W, Se) and substitutional oxygen (OSe) in 2D WSe2. The plots show the absolute value of the spin-diagonal intraband matrix elements Vmn i,kk' in the valence (bottom) and conduction (top) bands, with the initial state fixed to k = K (marked with dots o) and as a function of K'. The small (large) arrow in the lower left plot corresponds to intravalley (intervalley) couplings. DFT parameters: 11 × 11 supercell with 10 Å of vacuum to the cell boundaries in the vertical direction.

FIG. 3. Defect matrix elements for vacancies (V,A,B) and substitutional nitrogen (NSUB,A,B) in graphene. The plots show the absolute value of the spin-diagonal intraband matrix elements Vmn i,kk' in the valence (bottom) and conduction (top) bands, with the initial state fixed to k = K + δ, δ ≪ π/a (marked with dots o), and as a function of k'. DFT parameters: 11 × 11 supercell with 10 Å of vacuum to the cell boundaries in the vertical direction.

In Fig. 3 we show the valence and conduction band matrix elements for a carbon vacancy (left) and a nitrogen substitutional (right). In contrast to the matrix elements for TMDs in Fig. 2, there are no selection rules on the matrix elements for vacancies and substitutionals in graphene. Consequently, their matrix elements show less variation as a function of k'.

In Fig. 3 we show the valence and conduction band matrix elements for a carbon vacancy (left) and a nitrogen substitutional (right). In contrast to the matrix elements for TMDs in Fig. 2, there are no selection rules on the matrix elements for vacancies and substitutionals in graphene. Consequently, their matrix elements show less variation as a function of k'.

In this work, we restrict the considerations to single carbon vacancies [74–77] and nitrogen substitutionals [9,78,79].

It is instructive to analyze the matrix elements in Fig. 3 in terms of the standard tight-binding (TB) model for vacancies [28,29,33,35,80–82] and substitutional atoms [83–85], where the defect is often described by a change in the onsite energy of the defect site. In the A, B sublattice (pseudospin) basis, the defect potential can therefore be expressed as

$$\hat{V}_i = \frac{V_0}{2} (\delta_0 \pm \delta_i), \quad (15)$$

where $\delta_i = x, y, z$, are Pauli matrices ($i = 0$ denotes the identity matrix) in the pseudospin basis, and $\pm$ is for defects on the A and B sublattice, respectively.

For wave vectors in the vicinity of the K, K' points, the graphene TB Hamiltonian can be approximated by the Dirac model $\hat{H}_{k} = \hbar v_F \sigma r \cdot k$, where $r = \pm 1$ is the K, K' valley index, and $\sigma_r = (\tau \delta_x, \delta_y)$, with eigenstates $\chi_{k,n} = 1/2(1, n e^{i \pi n})^T$ and eigenenergies $E_{k,n} = n \hbar v_F k$, where $n = \pm 1$ is the band index for the conduction (c) and valence (v) bands, respectively.

Without loss of generality, we now for simplicity consider a defect on an A site. Performing a unitary transformation to the eigenstate basis, the matrix elements of the defect potential in Eq. (15) become

$$V_{r,kk'}^{m,n} = \frac{1}{2} V_0, \quad (16)$$

which are independent on the wave vectors k, k' and band indices m, n, and thus correspond to identical intravalley and intervalley as well as intraband and interband couplings.

For the vacancy defect in Fig. 3, Eq. (16) is seen to be a reasonable approximation in the vicinity of the K, K' points. From the value of the intravalley and intervalley matrix elements, $\sim 70 \text{ eV Å}^2$, we find via Eq. (13) that the energy shift at the vacancy site is $V_0 \approx +27 \text{ eV (A_{cell} = 5.24 Å^2)}$. 

DFT parameters: 11 × 11 supercell with 10 Å of vacuum to the cell boundaries in the vertical direction.
The positive sign can be attributed to the missing attractive core potential as well as unpaired $\sigma$ electrons left at the vacancy site which yield an overall repulsive defect potential in Eq. (12).

For the N substitutional in Fig. 3, the different values of the intravalley and intervalley matrix elements as well as the anisotropy of the intravalley matrix element indicate that Eq. (16) is a less good approximation. This is due to the fact that substitutional nitrogen donates a fraction of an electron to the graphene lattice and thereby ends up as a positively charged impurity characterized by a strong intravalley matrix element [86]. The defect potential due to chemical substitutions therefore presents both short-range (since there is a substantial onsite chemical energy shift) as well as some long-range features [83]. In Sec. V below, we find that an average value of $V_0 \approx -10$ eV ($\sim 26$ eV Å$^2$), where the minus sign is due to the partially positively charged N substitutional, yields good agreement with our full DFT-based results.

III. 7-MATRIX FORMALISM

In this section, we introduce the $T$-matrix formalism [25,26] for the description of (i) a single isolated defect at $R_0$, with defect potential $V_i(r - R_0)$, and (ii) disordered systems with a random configuration of defects and total disorder potential $V_{di}(r) = \sum_{i, R} V_i(r - R_i)$ where $R_i$ denotes the positions of defects of type $i$. The latter case is the main focus of this work.

A. Single defects

The situation of a single isolated defect in an otherwise perfect infinite lattice is relevant for, e.g., STM/STS studies which probe the LDOS in the vicinity of the defect site which can be obtained as

$$\rho(r, \epsilon) = -\frac{1}{\pi} \text{Im} G(r, r; \epsilon),$$

where $G(r, r; \epsilon) = \langle \epsilon | \hat{G}(\epsilon) | r \rangle$ is the real-space representation of the Green’s function (GF; all Green’s functions are assumed to be retarded in this work).

For the single-defect problem, the exact GF can be expressed in terms of the $T$ matrix which describes scattering off the defect to infinite order in the defect potential $T_i(\epsilon) = \hat{V}_i + \hat{V}_i \hat{G}^\sigma(\epsilon) \hat{T}_i(\epsilon)$, where $\hat{G}^\sigma(\epsilon) = 1/[\epsilon - H_0]$ is the GF of the pristine lattice. In the basis of the Bloch states $\{\psi_{nk}\}$ of the pristine lattice, the GF becomes

$$\hat{T}_{nk}(\epsilon) = \delta_{nk} \hat{G}_{nk}(\epsilon) + \hat{T}_{nk}(\epsilon) \hat{G}_{nk}(\epsilon),$$

where $\hat{G}_{nk}(\epsilon) = (\epsilon - \epsilon_{nk} + i\eta)^{-1}$, and

$$\hat{G}_{nk}(\epsilon) = \hat{V}_{nk} + \sum_{k'} \hat{V}_{nk'} \hat{G}_{n'k'}(\epsilon) \hat{T}_{nk'}(\epsilon).$$

Given the GF in Eq. (18), the real-space LDOS in Eq. (17) which contains information about the electronic properties of the defect can be obtained via a Fourier transform [27]. For example, defect-induced bound states manifest themselves in a high LDOS intensity at energies corresponding to the bound-state energy. They arise when the $T$ matrix introduces new poles in the GF via the correction $\delta G = \hat{G}^\sigma T \hat{G}$ in the last term of Eq. (18). From the full matrix form of the $T$ matrix

$$T(\epsilon) = [1 - \hat{V} \hat{G}^0(\epsilon)]^{-1} \hat{V},$$

where the boldface symbols denote matrices in the band $(n)$ and wave-vector $(k)$ indices, the poles of the $T$ matrix are seen to appear at energies where the determinant of the matrix in the square brackets vanishes, i.e.,

$$\text{det}[1 - \hat{V} \hat{G}^0(\epsilon)] = 0.$$}

Therefore, the positions of the bound states depend sensitively on the defect matrix elements and the band structure, and since also high-energy bands can be involved in the formation of bound states, their exact position can, in general, not be inferred from low-energy models. For in-gap bound states residing in the band gap of a semiconductor, the bound states form discrete energy levels and will be strongly localized to the defect site due to a weak interaction with the delocalized Bloch states. On the other hand, quasibound resonant states in the bands acquire a finite width and tend to be more delocalized.

An in-depth study of the LDOS and the associated quasiparticle interference is beyond the scope of this work and has been deferred to other works [27,87].

B. Disordered systems

In disordered systems with a random configuration of defects, experimental observables are often self-averaging and must be obtained on the basis of the disorder-averaged GF. In contrast to the single-defect problem discussed above, the problem for the disorder-averaged GF cannot be solved exactly.

The disorder-averaged GF is given by the Dyson equation

$$\hat{G}_{nk}(\epsilon) = \hat{G}_{nk}^0(\epsilon) + \hat{G}_{nk}^0(\epsilon) \hat{\Sigma}_k(\epsilon) \hat{G}_{nk}(\epsilon),$$

where the disorder self-energy $\hat{\Sigma}_k$ accounts for the interaction with the disorder potential, and introduces spectral shifts, broadening, and potentially bound defect states. As the disorder average restores translational symmetry, the GF is diagonal in $k$. However, the matrix structure in the band and spin indices is retained, and Eq. (22) must be solved by matrix inversion.

In this work, the self-energy is described at the level of the Born and $T$-matrix (full Born) approximations [25,26], which apply to dilute concentrations of defects. The two self-energies are illustrated with Feynman diagrams in Fig. 4, where the individual diagrams describe repeated scattering off single defects to different orders in the scattering potential. The $T$-matrix self-energy $\hat{\Sigma}^T_k$ in the bottom equation of Fig. 4 takes into account multiple scattering off defects to all orders in the defect potential, and is therefore exact to lowest
order in the disorder concentration \( c_i = N_i/N \) (or density \( n_i = N_i/A \)) where \( N_i \) is the number of defects of type \( i \). The self-energy is given by [25,26]

\[
\Sigma_{T,i,k}(\epsilon) = N_i \left[ \hat{V}_{i,kk} + \sum_{k'} \hat{V}_{i,kk'} \hat{G}^{0}_{kk'}(\epsilon) \hat{T}_{i,kk'}(\epsilon) \right],
\]

where \( \hat{T} \) is the \( T \) matrix in Eq. (19), and can be expressed in terms of the \( k \) diagonal of the \( T \) matrix as

\[
\Sigma_{T,i,k}(\epsilon) = N_i \hat{T}_{i,kk}(\epsilon) = c_i \hat{T}_{i,kk}(\epsilon) = n_i \hat{T}_{i,kk}.
\]

Here, the definitions of the symbols in the two last equalities are analogous to the ones for the defect matrix elements in Eq. (13), and express the self-energy in terms of the disorder densities where this is fulfilled in 2D materials, we consider graphene (\( \alpha_{\text{cell}} = 5.24 \text{ Å}^2 \)) where a disorder concentration of \( c_i = 1\% \) is equivalent to a density of \( n_i \sim 2 \times 10^{13} \text{ cm}^{-2} \). This corresponds to a rather poor material quality, why the \( T \)-matrix self-energy is an excellent approximation for most experimentally relevant disorder concentrations.

At the level of the two-particle GF for the conductivity there are, however, effects not captured by the \( T \)-matrix approximation even at low disorder concentrations. One well-known example is the weak-localization correction to the conductivity which arises due to interference between scattering processes at different defect sites [88,89]. Nevertheless, the \( T \)-matrix approach presented here provides a good compromise between wide applicability and practicality for applications with realistic defects and band structures.

1. Quasiparticle spectrum and scattering

As a consequence of disorder scattering, the pristine band structure is renormalized and broadened, yielding quasiparticle (QP) states with a finite lifetime which can be probed in, e.g., ARPES. The measured spectral function \( A_k(\epsilon) = \sum_n A_n(\epsilon) \) is given by the diagonal elements of the GF as

\[
A_k(\epsilon) = -2 \text{Im} \ G_{kk}(\epsilon),
\]

where \( A_n \) obeys the sum rule \( \int \frac{d\epsilon}{2\pi} A_n(\epsilon) = 1 \).

While our numerical calculations of the spectral function and DOS presented below are based on the full matrix form of the GF in Eq. (22), it is instructive to assume a diagonal form of the self-energy and GF,

\[
G_{kk}(\epsilon) = \frac{1}{\epsilon - \epsilon_{nk} - \Sigma_{kk}(\epsilon)},
\]

in order to analyze the effects of disorder scattering in closer detail.

In the diagonal approximation for the GF in Eq. (29), the spectral function can in the vicinity of the QP energies \( \tilde{\epsilon}_{nk} \) averaging over the self-energies of the different defect types, i.e.,

\[
\langle \Sigma_{kk} \rangle_{\text{dis}} = \sum_i \frac{N_{\text{dis}}}{N} \sum_{j} x_i \hat{T}_{i, kk},
\]
given by the solution to the QP equation
\[ \varepsilon - \varepsilon_{nk} - \text{Re} \Sigma_{nk}(\varepsilon) = 0 \] (30)
be approximated as
\[ A_{nk}(\varepsilon) \approx Z_{nk}\frac{\gamma_{nk}}{(\varepsilon - \varepsilon_{nk})^2 + (\gamma_{nk}/2)^2}, \] (31)
where the QP weight is given by \( Z_{nk} = [1 - \partial_\varepsilon \text{Re} \Sigma_{nk}(\varepsilon_{nk})]^{-1} \), and the linewidth broadening is given by the imaginary part of the self-energy
\[ \gamma_{nk} = -2Z_{nk}\text{Im} \Sigma_{nk}(\varepsilon_{nk}) = -2Z_{nk}N_1\text{Im} T_{nm,\text{kk}}^\text{nn}(\varepsilon_{nk}), \] (32)
evaluated at the on-shell QP energy \( \varepsilon = \varepsilon_{nk} \), and where the last equality holds for the \( T \)-matrix self-energy.

Via the optical theorem [25,26], the diagonal elements of the imaginary part of the \( T \) matrix can be expressed as
\[ -2 \text{Im} T_{i,kk}^\text{nn}(\varepsilon) = -2 \text{Im} \sum_{nk} \frac{|T_{nm,\text{kk}}(\varepsilon)|^2}{\varepsilon - \varepsilon_{nk} + i\eta} = 2\pi \sum_{nk} |T_{nm,\text{kk}}(\varepsilon)|^2 \delta(\varepsilon - \varepsilon_{nk}), \] (33)
and the lifetime broadening (or scattering rate) in Eq. (32) can be brought on a form which resembles the Born expression in Eq. (26). This allows to identify the elements \( T_{nm,\text{kk}} \) of the \( T \) matrix as the renormalized Born scattering amplitude given by the bare matrix element \( V_{nm,\text{kk}} \). Furthermore, the optical theorem in Eq. (33) can be used to separate out the contributions to the lifetime broadening from, e.g., intravalley and intervalley scattering by splitting the \( \varepsilon \) sum into sums over intravalley and intervalley processes. This may be desirable in order to extract, e.g., the disorder-limited valley lifetime.

In context of the discussion of scattering above, it is important to note that selection rules in the defect matrix elements \( V_{nm,\text{kk}} \) imposed by a symmetry \( \Theta \) common to the lattice and defect potential \( \hat{V} = \Theta \hat{V} \Theta^{-1} \) are transferred to the elements of the \( T \) matrix. This follows straightforwardly from the fact that the \( T \) matrix transforms as the defect potential \( \hat{T} = \Theta \hat{T} \Theta^{-1} \) under such symmetry transformations. Thus, scattering processes which are forbidden by symmetry due to vanishing matrix elements in the Born approximation are also forbidden in \( T \)-matrix approximations, in spite of the fact that scattering processes in the latter case proceed via virtual intermediate states.

### 2. Bound defect states

When bound states appear in the single-defect problem, it is interesting to ask how they manifest themselves in the spectral function and the DOS (here defined per unit cell) of the disordered system
\[ \rho(\varepsilon) = -\frac{1}{N\pi} \text{Im} \text{Tr} \hat{G}(\varepsilon), \] (34)
where \( N \) is the number of unit cells in the lattice. Naively, one would expect peaks at the bound-state energies of the isolated defect. To shed light on the the bound-state DOS of a dilute disordered system, we consider a situation where the single-defect GF in Eq. (18) has an in-gap bound state stemming from a pole of the \( T \) matrix with energy \( \varepsilon_b \).

To facilitate a simple analysis, we resort again to the diagonal form of the GF in Eq. (29). In this case, the self-energy is given by the diagonal elements of the \( T \) matrix, which in the vicinity of the pole can be approximated as
\[ T_{nk}(\varepsilon) = \frac{1}{N} \frac{a_{nk}}{\varepsilon - \varepsilon_b + i\eta}, \quad \varepsilon \approx \varepsilon_b, \] (35)
where \( a_{nk} \) (with unit \( eV^2 \)) is the strength of the pole for a given band \( n \) and \( k \) point, and the positive infinitesimal \( \eta \) ensures the correct analytic behavior of the \( T \) matrix.

We now demonstrate how the pole of the \( T \) matrix gives rise to a new bound-state pole in the disorder-averaged GF, whose energy we denote \( \bar{\varepsilon}_b \), and in order to distinguish it from the \( T \)-matrix pole at \( \varepsilon = \varepsilon_b \). The bound state, being a well-defined quasiparticle of the disordered system, emerges as a new solution to the QP equation in Eq. (30) with energy \( \varepsilon = \bar{\varepsilon}_b \) as sketched graphically in Fig. 5 for states at the band edges, i.e., \( \varepsilon_{nk} = \varepsilon_{v,c} \). Here, the generic shape of the real part of the \( T \) matrix (self-energy) originates from its pole form in Eq. (35).

Expanding the self-energy around \( \varepsilon = \bar{\varepsilon}_b \), \( \Sigma(\varepsilon) \approx \text{Re} \Sigma(\bar{\varepsilon}_b) + (\varepsilon - \bar{\varepsilon}_b)\partial_\varepsilon \text{Re} \Sigma|_{\varepsilon=\bar{\varepsilon}_b} \), the in-gap GF takes the form
\[ G_{nk}(\varepsilon) = \frac{Z_{nk}^b}{\varepsilon - \bar{\varepsilon}_b + i\eta}, \quad \varepsilon \approx \bar{\varepsilon}_b, \] (36)
where \( Z_{nk}^b = (1 - \partial_\varepsilon \text{Re} \Sigma_{nk}|_{\varepsilon=\bar{\varepsilon}_b})^{-1} \) is the contribution to the spectral weight of the bound states from the state \( nk \). Note that bound states may have contributions from several bands, and by virtue of the sum rule for the spectral function in Eq. (28), \( Z_{nk}^b \ll 1 \) in order for the pristine bands to remain well defined.

In the presence of bound states, the spectral function in Eq. (28) thus becomes a sum of two contributions \( A_{nk}^b = A_{nk}^{QP} + A_{nk}^b \), given, respectively, by (i) Eq. (31) describing the QPs associated with the pristine bands, and (ii) the imaginary part of Eq. (36) for the bound state. Note that the bound-state solution to the QP equation may, in fact, depend on both \( n \) and \( k \), such that dispersive defect bands, i.e., \( \bar{\varepsilon}_b \rightarrow \bar{\varepsilon}_{b,nk} \).
may arise even though hybridization between the defects is not accounted for in the $T$-matrix approximation. As we demonstrated in a recent work, this situation arises in, e.g., alkali-metal-decorated graphene [90].

The bound-state position $E_b$ naturally depends on the disorder concentration via Eq. (30). With the self-energy given by the pole form of the $T$ matrix in Eq. (35), it follows straightforwardly that $E_b \rightarrow E_i$ in the limit of $c_i \rightarrow 0$. This is not surprising as the $c_i \rightarrow 0$ limit is equivalent to setting $N_i = 1$ and letting $N \rightarrow \infty$, i.e., the single-defect limit. In addition, the quasiparticle weight vanishes as $\delta(N_0) \sim c_i^2/(\epsilon - \epsilon_N)^2$, implying that the in-gap form of the imaginary part of the GF becomes

$$\text{Im}G_N(\epsilon) \approx \frac{\pi}{N} \frac{a_{nk}}{(\epsilon - \epsilon_N)^2} \delta(\epsilon - E_b), \quad c_i \rightarrow 0, \quad (37)$$

which is identical to the in-gap form of the imaginary part of the single-defect GF (its diagonal elements) in Eq. (18). Thus, the bound-state DOS of a dilute system approximates the single-defect DOS for $c_i \rightarrow 0$ as anticipated, and its weight vanishes as $1/N$ compared to the DOS of the pristine lattice. At higher disorder concentrations (but still $c_i \ll 1$), the graphical solution of Eq. (30) in Fig. 5 indicates that the GF pole $E_b$ drifts away from the $T$-matrix pole at $E_b$

When the disorder concentration becomes so high that electronic states on neighboring defect sites start to interact and form impurity bands, the single-site $T$-matrix approximation considered here breaks down and more advanced methods are required [91,92].

3. Transport

At low temperatures where electron-phonon scattering is frozen out, the longitudinal conductivity is often limited by the intrinsic disorder of the material. Within the framework of Boltzmann transport theory, the disorder-limited longitudinal conductivity $\sigma$ can be obtained from the current density,

$$j = q \sum_{nk} v_{nk} \delta f_{nk} \equiv \sigma \mathbf{E}, \quad (38)$$

where $q$ is the charge of the carriers, $v_{nk} = 1/\hbar \nabla_k \epsilon_{nk}$ is the band velocity, and $\delta f_{nk} = f_{nk} - f_0^{nk}$ is the deviation of the distribution function away from the equilibrium Fermi-Dirac distribution $f_0^{nk} = f_0(\epsilon_{nk})$, to first order in the applied field $\mathbf{E}$.

The deviation function is given by the linearized Boltzmann equation which for elastic disorder scattering in a multiband system takes the form

$$q v_{nk} \cdot \mathbf{E} \frac{\partial f^0}{\partial \epsilon} \bigg|_{\epsilon = \epsilon_{nk}} = - \sum_{nk'} P_{nk,nk'} [\delta f_{nk} - \delta f_{nk'}], \quad (39)$$

where

$$P_{nk,nk'} = \frac{2\pi}{\hbar} N_i |T^0_{kk'}(\epsilon_{nk})|^2 \delta(\epsilon_{nk} - \epsilon_{nk'}) \quad (40)$$

is the transition rate in the $T$-matrix approximation, which follows from the optical theorem in Eq. (33).

In Appendix A we outline a least-square method for the solution of the Boltzmann equation (BE) (39) on general $k$-point grids and with first-principles inputs for the band structure, band velocities, and elastic scattering rate. The method does not rely on any assumptions about the functional form of the deviation function $\delta f_{nk}$ or a relaxation-time approximation. Other approaches for the solution of the BE based on first-principles input for inelastic electron-phonon scattering have been discussed in the literature [42–44]. Our method in Appendix A was recently applied in calculations of the transport properties of Li-doped graphene within a TB description of the graphene bands and the Li-induced carrier scattering [93].

In the calculations presented in this work, we restrict the discussion to transport involving a single band (spin-degenerate or spin-orbit split), and furthermore assume that the band structure is isotropic with a constant effective mass $m^*$ (or Fermi velocity $v_F$ in the case of graphene), which is a good approximation for the transport-relevant energy range close to the band edges. In this case, the conductivity can be expressed in terms of the relaxation time given by the $T$-matrix scattering amplitude as

$$\tau^{-1}_{kk'} = \frac{2\pi}{\hbar} N_i \sum_k |T^0_{kk'}(\epsilon_{nk})|^2 \times [1 - \cos \theta_{kk'}] \delta(\epsilon_{nk} - \epsilon_{nk'}), \quad (41)$$

where $\theta_{kk'} = \theta_k - \theta_{k'}$ is the scattering angle. The only difference between the QP scattering rate in Eqs. (32) and (33) and the inverse transport relaxation time is the factor $1 - \cos \theta_{kk'}$ in the square brackets, which accounts for the fact that the transport is insensitive to small-angle scattering while the QP lifetime is equally sensitive to all scattering processes. For isotropic scattering, the two scattering times become identical as the angular integral of the $\cos \theta_{kk'}$ term vanishes.

With the above-mentioned assumptions and considering the low-temperature limit ($k_B T \ll E_F$), the conductivity in Eq. (38) simplifies to the well-known Drude form given by

$$\sigma = \frac{n e^2 \tau(E_F)}{m^*} \quad \text{and} \quad \sigma = \frac{e^2 v_F^2}{2 \rho(E_F)} \tau(E_F), \quad (42)$$

in, respectively, a 2D semiconductor and graphene, where $n$ is the carrier density and $\rho$ is the density of states. In combination with first-principles calculations of the $T$-matrix transport relaxation time in Eq. (41), these expressions provide a simple and accurate framework for calculating the low-temperature conductivity in disordered 2D materials.

C. Numerical and calculational details

The calculation of the $T$ matrix in Eq. (19) is the most demanding step in the evaluation of the above-mentioned quantities. Rather than solving the equation by direct matrix inversion as in (20), it is numerically more stable to recast it as a system of coupled linear equation (one set of coupled equations for each column in $T$ and $V$)

$$[1 - VG^0(\epsilon)]T(\epsilon) = V, \quad (43)$$

and solve it with a standard linear solver. This requires one factorization followed by a matrix-vector multiplications and scales as $O(M^3)$ where $M$ denotes the matrix dimension.

The calculation of the $T$ matrix must be checked for convergence with respect to (i) the BZ sampling with $N_k$ $k$ points
on either $N_k \times N_k$ uniform BZ grids or nonuniform grids with a higher density of grid points in the vicinity of important high-symmetry points (see Fig. 6), and (ii) the number of bands $N_b$ included in the calculation of the $T$ matrix. In general, we have found that the convergence of the position of bound defect states in the gap of 2D semiconductors requires a large number of bands ($N_b > 50$) starting from the bottom of the spectrum and up to high energies, while only a moderate $k$-point sampling ($\sim 21 \times 21$) is required. On the other hand, the linewidth broadening in Eq. (32) [and Eq. (33)] requires only a few bands ($N_k \sim 2–4$) (as long as there are no bound states in the bands), but a dense $k$-point grid ($135 \times 135$) in order to sample the constant-energy surface on which the quasiparticle scattering takes place. In the case of graphene, quasibound resonant states are inherent to the Dirac cone dispersion, and hence both the resonant state and the scattering rate can be calculated with a dense $k$-point sampling including only a few bands.

With the above-mentioned number of bands and $k$-point samplings, the dimension $M$ of the matrices in Eqs. (43) becomes $M = N_k \times N_k \approx 20 000–50 000$. With the matrix elements represented as 128-bit complex floating-point numbers, the memory requirement for each of the dense complex matrices in Eq. (43) becomes $M^2 \times 128/8 \text{ bytes} \approx 10–30 \text{ GBs}$. To tackle the large matrix dimensions in the solution of the matrix equation in Eq. (43), we exploit the automatic openMP multithreading of the LAPACK linear solvers.

In the calculation of the DOS in Eq. (34), a very fine $k$-point sampling is needed to converge the DOS of the bands near the band edges. We therefore (i) first calculate the difference $\delta \rho = \rho_{\text{dis}} - \rho_0$ between the DOS of the disordered and pristine materials on a coarse $k$-point grid (in order to include enough bands to capture bound states), and (ii) subsequently add $\delta \rho$ to the DOS of the pristine material obtained on a fine $k$-point grid. In this way, we avoid spiky artifacts in the DOS of the bands due to insufficient $k$-point sampling, while at the same time capturing potential defect states in the bands.

For the results presented in the following sections, the band structures and defect matrix elements have been obtained with the GPAW electronic-structure code [48–50], using DFT-LDA within the projector augmented-wave (PAW) method, a double-zeta polarized (DZP) LCAO basis, and including spin-orbit interaction [94]. The parameters used in the individual calculations are listed in the figure captions.

IV. DISORDERED 2D TMDs

The experimental consensus on the prevalent types of defects in the 2D TMDs (see Sec. II C 1) has led to numerous theoretical DFT studies of their structural and electronic properties [27,61,62,95–101]. On the other hand, first-principles calculations of the impact on electron dynamics in disordered 2D TMDs remain few [102,103]. In the following subsections, we analyze in detail the electronic (DOS and quasiparticle spectrum) and transport (conductivity and mobility) properties.

A. DOS and in-gap bound states

We start by discussing the impact of defects on the DOS, and in particular the defect-induced in-gap states observed in various STM/STS experiments [11–16]. Figure 7 shows the DOS for disordered MoS$_2$ (top) and WSe$_2$ (bottom) with different types of defects. The dashed vertical lines mark the position of the valence and conduction band edges (black) as well as the Fermi energy ($E_F$; red dashed line). All the defects, in particular transition metal vacancies, introduce a series of defect-localized in-gap states at positions in good agreement with previously reported DFT supercell calculations [61,100,102,104].

Our results in Fig. 7 show that some of the orbitally degenerate bound states are subject to a notable spin-orbit-induced spin splitting. This is also in agreement with previous supercell calculations of defect-induced in-gap states taking into account spin-orbit interaction [14,102,104]. In MoS$_2$, the splitting is of the order of $\sim 50$ meV for the two V$_{Mo}$ states above the Fermi energy and the two V$_S$ top states, whereas a significantly larger splitting of $\sim 270$ meV is observed for the unoccupied V$_{Se}$ and O$_{Se}$ states in WSe$_2$. It should be noted that this effect is captured in spite of the fact that we here consider spin-independent defect potentials, i.e., the spin-orbit interaction enters only through the unperturbed band structure.

It is interesting to note that most of the considered defects introduce both occupied and unoccupied in-gap states occur, except for the V$_{Se}$ and O$_{Se}$ defects in WSe$_2$ which only introduce unoccupied in-gap states. For all the other defects, the shallow occupied states above the valence-band edge act as hole traps in the $p$-doped (gated) materials. The converse holds for the unoccupied in-gap states which act as deep electron traps in the $n$-doped materials. As we have recently demonstrated, the charging of the defect sites resulting from such carrier trapping has detrimental impact on the transport properties of gated 2D TMDs [103].

In addition to in-gap bound state, defects may also introduce quasibound states inside the bands as predicted in, e.g., MoSe$_2$ and WS$_2$ [14,102]. As witnessed by the $c_i = 1\%$ curves in Fig. 7 which show pronounced deviations from the pristine DOS inside the bands, this seems to be the case also in MoS$_2$ and WSe$_2$. However, we find that some of these features are artifacts from the procedure we have used to
calculate the DOS in the bands of the disordered system (see Sec. III C above). Only the features in the valence band for Mo vacancies in MoS$_2$ ($\sim$350 meV below the band edge), W ($\sim$150 meV below the band edge) and Se vacancies ($\sim$350 meV below the band edge) in WSe$_2$ correspond to true quasibound states. As their positions are reasonably far away from the band edges, resonant scattering off the quasibound states can be neglected in calculations of the disorder-limited transport properties [103]. By contrast, both bound and quasibound defect states have been demonstrated to alter the optical properties of 2D TMDs by binding the excitons in the defect states [63,66,67,102].

**B. Spectral function and quasiparticle scattering**

In Fig. 8, we show the valence and conduction band spectral functions (grayscale intensity plots) for disordered MoS$_2$ and WSe$_2$ with a $c_i = 1\%$ concentration of, respectively, Mo and Se vacancies together with the unperturbed band structure of the pristine materials (dashed lines). Although our DFT calculations indicate that the direct and indirect band gaps in MoS$_2$ and WSe$_2$ are almost identical, and that the band gap in some cases is indirect [105], recent microARPES experiments have given conclusive evidence that the band gap in monolayers of the semiconducting TMDs $MX_2$ with $M = Mo, W$ and $X = S, Se$ is direct [70].

Overall, the spectral functions overlap almost perfectly with the unperturbed band structures, indicating that disorder-induced renormalization of the bands is small at $c_i = 1\%$ in the $T$-matrix approximation. This is in stark contrast to the Born approximation (not shown), where the first term in Eq. (25) given by the bare defect matrix element gives rise to a giant shift of the bands. In the $T$-matrix approximation, this shift is strongly renormalized by the matrix inverse on the right-hand side of Eq. (20).

![Fig. 8](image-url) Spectral function of disordered 2D TMDs showing the two first spin-orbit-split valence and conduction bands for (left) MoS$_2$ with Mo vacancies and (right) WSe$_2$ with Se vacancies. The dashed lines show the bands of the pristine materials. Parameters: $c_i = 1\%$, $135 \times 135$ k points, 4 bands, and $\eta = 20$ meV.
The quasiparticle lifetime given by the broadening of the spectral function is difficult to infer from Fig. 8 due to the numerical broadening $\eta$. In Fig. 9 we therefore show the linewidth broadening obtained directly from the on-shell self-energy via Eq. (32). Overall, the linewidths show a pronounced dependence on $k$ with multiple peaks and dips along the considered path in the BZ. The strong increase in the linewidth in the vicinity of the $\Gamma$ point in the conduction bands (top plots) is due to overlap with higher-lying bands outside the energy range shown in Fig. 8. The sharp peak in the linewidth along the $K-M$ path in the lower right plot is due to resonant scattering off the quasibound defect state introduced by the Se vacancy in the valence band. For the defect concentration $c_i = 1\%$ considered here, the overall magnitude of the disorder-induced linewidth is comparable to the phonon-induced linewidth at elevated temperatures [106].

In the $\Gamma$ and $K$ valleys close to the band edges, the linewidths show characteristic dips with a particularly sharp shape. To analyze these features in closer detail, we show in Figs. 10 and 11 the linewidths for the spin-up and -down bands in $K$ valley of, respectively, the conduction band of MoS$_2$ (Mo and S vacancies) and the valence band of WSe$_2$ (W and Se vacancies) as a function of the band energy $\varepsilon = \varepsilon_{nk}$ (measured with respect to the band edges) instead of $k$. In the two figures, the left columns show a comparison between the Born [Eq. (26)] and $T$-matrix approximations, whereas the right columns show the contributions to linewidth from intravalley and intervalley scattering. Due to the large spin-orbit splitting in the valence band of WSe$_2$, only the linewidth for the spin-up band appears in Fig. 11. Note that different $k$-point samplings have been used in the two figures, hence the difference in energy resolution.

In Figs. 10 and 11, the sharp dips in the linewidths in Fig. 9 mentioned above are manifested in a strong energy dependence of the $T$-matrix linewidths close to the band edges. For comparison, the linewidths in the Born approximation exhibit a weaker energy dependence which can be traced back to the almost constant matrix elements in Fig. 2 inside the $K, K'$ valleys, implying that the Born linewidth given by Eq. (26) to a good approximation becomes proportional to the DOS in the $K, K'$ valleys. The energy dependence of the Born linewidths therefore reflects the gradual increase in the DOS in Fig. 7 at the band edges. On the contrary, the sharp drop in the $T$-matrix linewidths at the band edges is a consequence of higher-order renormalization of the Born scattering amplitude by multiple scattering processes. This makes the $T$-matrix amplitude strongly energy dependent, and can modeled quantitatively with a simple analytic $T$-matrix model as we have demonstrated in Ref. [103].

Another consequence of the higher-order renormalization is a strong reduction of the scattering amplitude between the Born and $T$-matrix approximations. This is evident from Fig. 10 where the former overestimates the scattering rates up to three orders of magnitude. The reduction is largest for $M$ vacancies which are strong defects (cf. the matrix elements in Fig. 2) for which the $T$ matrix leads to a giant renormalization of the Born scattering amplitude. For the weaker $X$ centered defects (vacancies and substitutional atoms), the matrix elements in Fig. 2 are smaller, but still large enough for the $T$ matrix to yield a non-negligible renormalization of the scattering amplitude. These observations point to a concomitant breakdown of the Born approximation and stress the importance of a $T$-matrix description of atomic defects in 2D TMDs.

In the plots in the right columns of Figs. 10 and 11, we have separated out the contributions to the linewidth...
broadening from intravalley (solid lines) and intervalley (dashed lines) scattering. Interestingly, the plots show that while the intravalley and intervalley contributions are of the same order of magnitude for Mo vacancies in MoS$_2$, the intervalley contribution is negligibly small for S vacancies as well as W and Se vacancies in WSe$_2$. In the almost spin-degenerate conduction band of MoS$_2$ [73], this is related to the symmetry-induced selection rules for the intervalley matrix elements discussed in Sec. II C 1 and Ref. [27], which strongly suppress $K \leftrightarrow K'$ intervalley scattering by $X$-centered defects in 2D TMDs. In the valence band of WSe$_2$, this as well as the large and opposite spin-orbit splitting in the $K, K'$ valleys [72] suppress intervalley scattering. Long valley lifetimes exceeding hundreds of ps are therefore achievable even in highly disordered 2D TMDs if $M$-centered defects can be eliminated.

Our finding for the suppression of intervalley scattering by chalcogen-centered defects is also relevant for studies of weak localization/antilocalization in 2D TMDs [107–111], where, e.g., S vacancies in $n$-doped MoS$_2$ have often been mentioned as a source pronounced intervalley scattering [109–111]. As we have demonstrated here, this is not the case and intervalley scattering must instead be attributed to the existence of other point defects.

C. Transport

In studies of the disorder-limited transport properties of 2D TMDs, it is important to consider the fact that defect-induced in-gap states can trap holes or electrons as extrinsic carriers (i.e., gate induced) are introduced into the bands. This holds, respectively, for occupied in-gap states in $p$-doped as well as unoccupied in-gap states in $n$-doped samples, and results in charging of the defect sites. A description of such charging effects within the framework of our method in Sec. II is beyond the scope of this work. Based on a simple model, we recently demonstrated that the charge-impurity scattering resulting from charging of defects has detrimental consequences for the carrier mobility in 2D TMDs [103], and is therefore unfavorable in order to realize high-mobility TMD samples.

In the following, we focus on cases where defect charging is not expected. As witnessed by Fig. 7, this situation is encountered in $p$-doped WSe$_2$ with Se vacancies or O substitutionals which only introduce unoccupied in-gap states above the intrinsic Fermi level. Upon $p$ doping the material, i.e., moving the Fermi level into the valence band with a gate voltage, the defects thus remain overall neutral as there are no occupied in-gap states to deplete. The defect potentials in Sec. II C 1 obtained for the charge-neutral defect supercells therefore give a realistic description of the defects in the $p$-doped material.

In the top plot of Fig. 12 we show the low-temperature transport characteristics of disordered $p$-doped WSe$_2$ with a concentration of Se vacancies of $c_v = 0.01\%$ ($n_i \approx 10^{11} \text{ cm}^{-2}$). (Top) Conductivity and mobility vs Fermi level. (Bottom) Ratio between the transport and quantum scattering times. Parameters: $m^* = 0.46$ and caption of Fig. 11.
In $p$-doped WSe$_2$, the carrier density $n$ scales with the Fermi level as $n \approx \frac{e \mu}{3 \text{meV} \times 10^{12} \text{cm}^{-2}}$. The energy range considered in Fig. 12 thus corresponds to typical values of the carrier densities accessible in experiments. The conductivity and mobility in Fig. 12 directly probe the energy dependence of the scattering rate shown in Fig. 11. Since the scattering rate decreases with energy, the conductivity exhibits an initial sublinear density dependence, which translates into a mobility that decreases with carrier density. The characteristic density scaling of the mobility in Fig. 12 is therefore a direct fingerprint of the inherent energy dependence of the $T$-matrix scattering amplitude for point defects in 2D TMDs.

Another important observation to make from Fig. 12 is the overall large magnitude of the carrier mobility $\mu \sim 15,000-35,000 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$, which exceeds all previously reported experimental values, so far not exceeding $\mu \sim 5000 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$ [7,109,111–120]. The large theoretical mobility predicted here is a hallmark of (i) the low defect density used in the calculation which corresponds to high-quality TMDs [5], and (ii) the absence of the above-mentioned defect charging which leads to a significant reduction of the mobility due to charged-impurity scattering [103]. Both factors are essential for the realization of high-mobility monolayer TMD samples.

Finally, the bottom plot in Fig. 12 shows the ratio between the transport scattering time and the quantum (quasiparticle) scattering time which is accessible from Shubnikov–de Haas oscillations in magnetotransport [121]. The close-to-unity value of the ratio is a direct manifestation of a weak $q = k - k'$ dependence of the $T$-matrix scattering amplitude in the $K, K'$ valleys which is inherited from the Se vacancy defect matrix element in Fig. 2. In this case, the $\cos \theta_{kk'}$ term in the transport relaxation time in Eq. (41) vanishes, and the two scattering times become identical.

Aside from the impact on the longitudinal conductivity considered here, other theoretical works have studied the effect of disorder in 2D TMDs on various other properties such as, e.g., the optical conductivity [122], excitons and optical absorption [123], localization [107,108,110], and spin and valley Hall effects [124,125]. Extensions to studies based on atomistic descriptions of the defect potential offer interesting perspectives for future developments.

V. DISORDERED GRAPHENE

Graphene is known to host a wide variety of atomic-scale point defects which are predicted to introduce resonant states on the Dirac cone associated with quasi-bound defect states [28,31,33,80]. The energy of such quasi-bound states depends on the interaction between the defect and the graphene lattice which is highly sensitive to the position of the defect [34,81,82]. For vacancies and substitutional atoms, quasi-bound states with energies in direct vicinity of the Dirac point arise in a robust manner. In transport, such defects act as resonant scatterers exhibiting a strong peak in the scattering cross section at the resonance energy which suppresses the conductivity [126,127] and affects electron cooling [128,129].

![Image](image_url)

**FIG. 13.** Density of states of disordered graphene with different defect concentrations of vacancies (top) and N substitutionals (bottom). Parameters: $99 \times 99^\circ k$ points ($300 \times 300$ for the pristine DOS), 2 bands, and $\eta = 50 \text{ meV}$ (15 meV for pristine).

In the following, we focus on monoatomic vacancies and nitrogen (N) substitutionals on the A and/or B sublattice.

A. DOS and spectral function

The DOS of disordered graphene has been studied in numerous works (see, e.g., Refs. [28–31]) addressing, e.g., defect-induced resonant states and band-gap openings. Below we separate the discussion in two cases: (1) sublattice-asymmetric disorder, i.e., defects located exclusively on one sublattice, and (2) sublattice-symmetric disorder where the defects are distributed equally between the $A$ and $B$ sublattices.

1. Sublattice-asymmetric disorder

In Fig. 13, we show the DOS for disordered graphene with vacancies (top) and N substitutionals (bottom) located on only the $A$ sublattice ($V_A$ and $N_A$). At low concentration, $c_V = 0.1\%$, the DOS is hardly discernible from the DOS of pristine graphene. By contrast, at $c_V = 1\%$ a clear peak in the DOS associated with a quasi-bound resonant state emerges, respectively, below and above the Dirac point for $V_A$ and $N_A$ defects. This is in agreement with STM studies of the LDOS [9,77,79]. Our finding for the position of the vacancy bound state contrasts previous theoretical studies based on tight-binding modeling [31–35,130], which predict the resonant
state to be at the Dirac point. In our analysis below, we comment on this discrepancy.

Figure 14 shows plots of the spectral functions corresponding to the different defects and concentrations in Fig. 13. The spectral function at \( c_1 = 0.1\% \) reflects the pristine bands (red dashed lines), although a finite broadening due to disorder scattering, here masked by the numerical broadening \( \eta \), is present (see also Fig. 17 below). For \( c_1 = 1\% \), the Dirac cone is strongly perturbed due to resonant scattering at the position of the quasibound defect states in the DOS. In addition to a pronounced broadening of the states which completely washes out the Dirac cone, this also produces a significant renormalization of the bands below and above the position of the resonance. Signatures of such effects in ARPES on nitrogen-doped graphene have so far not been observed [78,131], probably because the concentration of N substitutionals is too low.

By closer inspection of the spectral functions in Fig. 14, a concentration-dependent band-gap opening can be observed at the Dirac point which at \( c_1 = 1\% \) is \( \sim 100 \text{ meV} \). This is expected as defects located on a single sublattice break the sublattice symmetry, effectively turning the disordered system into gapped graphene [132] as also demonstrated in other theoretical works considering sublattice-asymmetric disorder [29,85,133]. Band-gap openings have also been reported experimentally in ARPES on nitrogen-doped graphene [78] and graphene with hydrogen adatoms [134], but the underlying mechanism is believed to be of a different nature, i.e., not associated with sublattice asymmetry.

To shed additional light on the band-gap opening as well as the resonant spectral features, the GF in the 2 \( \times \) 2 subspace spanned by the valence and conduction bands. In this subspace, the diagonal elements obtained by matrix inversion of the Dyson equation (22) take the form

\[
G^{\text{eff}}_{nn}(\epsilon) = \frac{1}{\epsilon - \epsilon_{nk} - \Sigma^{\text{eff}}_{nk}(\epsilon)} \quad \text{for A or B,}
\]

where the effective self-energy is given by

\[
\Sigma^{\text{eff}}_{nk}(\epsilon) = \Sigma_{nn}^{\text{eff}}(\epsilon) + \frac{\Sigma^{\text{eff}}_{nk}(\epsilon)\Sigma^{\text{eff}}_{nk}(\epsilon)}{\epsilon - \epsilon_{nk} - \eta^{\text{eff}}_{nk}(\epsilon)}, \quad \bar{n} \neq n.
\]

Here, the second term introduced by the matrix inversion describes a defect-induced coupling between the valence and conduction bands, and is responsible for the band-gap opening.

The effective self-energy in Eq. (45) based on the DFT-calculated \( T \) matrix for N substitutionals is shown in the left top plot of Fig. 15 for \( \mathbf{k} = \mathbf{K} \). Here, the solution to the transcendental equation for the QP equation in Eq. (30) corresponds to the intersection between the solid (green) and dashed lines. Clearly, the effective self-energy shows a feature just below the Dirac point which gives rise to two solutions of the QP equation, corresponding, respectively, to the top of the valence band and the bottom of the conduction band, and hence mark a band-gap opening.

The feature in the effective self-energy responsible for the band-gap opening can be analyzed further based on the TB defect model introduced in Sec. II C 2. In this case, we can solve Eq. (20) for the \( T \) matrix analytically, yielding a k-independent \( T \) matrix which in the pseudospin basis inherits the matrix structure of the defect potential in Eq. (15),

\[
\tilde{T}_l(\epsilon) = \frac{T_0(\epsilon)}{2}(\delta_0 \pm \delta_\epsilon),
\]
where the prefactor is given by

$$T_{0}(\varepsilon) = \frac{V_{0}}{1 - V_{0}\hat{G}(\varepsilon)}$$

and \(\hat{G}(\varepsilon) = \frac{1}{2} \text{Tr} \hat{G}^{(0)}(\varepsilon) = \frac{1}{2} \sum_{\mathbf{k}} \hat{G}_{\mathbf{k}}^{(0)}(\varepsilon)\) is the \(\mathbf{k}\)-summed GF for pristine graphene. In the Dirac model, it is given by

$$\hat{G}_{\mathbf{k}}^{(0)}(\varepsilon) = A_{\text{cell}} \hat{p}_{0} \left[ \ln \left| \frac{\varepsilon^{2} - \Lambda^{2}}{\varepsilon^{2} - \Lambda^{2}} \right| - i\pi |\varepsilon| \Theta(\Lambda - |\varepsilon|) \right],$$

where \(\hat{p}_{0} = g_{v}/2\pi(\hbar v_{f})^{2}\), \(g_{v} = 2\) is the valley degeneracy, and \(\Lambda\) is an ultraviolet cutoff.

Performing a unitary transformation to the eigenstate basis (as the results below are independent on the valley index we omit it here), the \(T\)-matrix self-energy in Eq. (24) becomes

$$\Sigma_{\mathbf{k}}^{T}(\varepsilon) = \Sigma_{0}(\varepsilon) \begin{pmatrix} 1 & \pm 1 \\ 1 & 1 \end{pmatrix}, \quad \Sigma_{0}(\varepsilon) = \frac{c_{1}T_{0}(\varepsilon)}{2},$$

where the \(\pm\) sign on the off-diagonal elements is for defects on the \(A, B\) sublattice. In either case, the diagonal elements of the GF again take the form in Eq. (44), with the effective self-energy now given by

$$\Sigma^{\text{eff}}_{\mathbf{k}}(\varepsilon) = \Sigma_{0}(\varepsilon) + \frac{[\Sigma_{0}(\varepsilon)]^{2}}{\varepsilon - \varepsilon_{\mathbf{k}} - \Sigma_{0}(\varepsilon)}, \quad \mathbf{n} \neq n.$$

To see the role of the second term for the band-gap opening, we note that \(\hat{G}_{0} \to 0\) in the vicinity of the Dirac point, i.e., \(|\varepsilon| \to 0\), and hence the \(T\) matrix in Eq. (47) can be approximated as \(T_{0}(\varepsilon) \approx |1 + i\delta(\varepsilon)|V_{0}\), where \(\delta(\varepsilon) = V_{0}\text{Im} \hat{G}(0)\). In the effective self-energy in Eq. (50) this leads to a pole in the second term which for the Dirac-point self-energy, i.e., \(\mathbf{k} = \mathbf{K}\), is located at \(\varepsilon_{0} = c_{1}V_{0}/2\). The right top plot in Fig. 15 shows the analytic TB self-energy in Eq. (50) for parameters corresponding to \(N\) substitutional (see caption). The parameters have been obtained by fitting to the DFT self-energy as follows: We first fix \(V_{0}\) to a value resembling the average value of the intravalley and intervalley matrix elements in Fig. 3, and then treat \(\Lambda\) as a fitting parameter in order to match the DFT self-energies in Fig. 15. This is in contrast to the Debye-model-inspired approach in Ref. [31], which is here found to be unable to yield a satisfactory description of the DFT self-energy.

Remarkably, the DFT- and TB-calculated self-energies are in almost perfect agreement. However, due to (i) the nontrivial \(\mathbf{q} = \mathbf{k}' - \mathbf{k}\) dependence of the matrix elements in Eq. 3, and (ii) the finite numerical broadening \(\eta\) used in the DFT calculation of the \(T\)-matrix self-energy, some quantitative differences arise.

Via the analytic self-energy in the top plot of Fig. 15, the feature in the DFT self-energy responsible for the band-gap opening can now be identified with the pole introduced by the second term in Eq. (50), which clearly emerges just below the Dirac point, and is seen to give rise to the two solutions to the QP equation also found in the DFT self-energy in Fig. 15. Note that the QP equations for the valence and conduction bands are identical at \(\mathbf{k} = \mathbf{K}\) (since \(\varepsilon_{\mathbf{k}} = 0\) for \(n = v, c\)), implying that the states at the band-gap opening are formed by a combination of the original valence and conduction band states as in conventional gapped graphene.

In the top plots of Fig. 15, the pole structure in the self-energy at positive energy stems from the pole in the \(T\) matrix associated with the quasibound defect state in the DOS in Fig. 13. For the \(T\) matrix in Eq. (47), the pole is positioned at the energy where \(1/V_{0} = \text{Re} \hat{G}_{0}\), and is thus located above (below) the Dirac point for \(V_{0} < 0\) \((V_{0} > 0)\) (see, e.g., Fig. 10 in Ref. [84]). This is in agreement with our discussion of the \(V_{0}\) parameter for the \(V_{A}\) and \(N_{A}\) defects in Sec. II C 2. Note that \(|V_{0}| \to \infty\) produces a bound state at the Dirac point as discussed in several TB studies of vacancy defects [31–35]. However, as our DFT calculations here demonstrate, this limiting value of \(V_{0}\) provides an unrealistic description of the vacancy potential and consequently also the position of the quasibound defect state.

In addition to introducing the defect state itself, poles in the \(T\) matrix also account for resonant scattering off the defect state which strongly perturbs the bands. As evident from the bottom plots in Fig. 14, this introduces a splitting of the conduction band at the resonance energy which resembles a gap opening, but the broadening of the states due to resonant scattering prevents the opening of a spectral gap.

2. Sublattice-symmetric disorder

In Fig. 16 we show the spectral function of disordered graphene \((c = 1\%)\) with defects distributed equally on the \(A\) and \(B\) sublattices for vacancies (top left), \(N\) substitutional (top right), and both vacancies and \(N\) substitutional defects (bottom). The two top plots to a large extent resemble the bottom plots for \(c_{1} = 1\%\) of \(A\)-only defects in Fig. 14, however, with the important difference that the spectral functions in Fig. 16 do not feature a band-gap opening. In the presence of both
$V_{A,B}$ and $N_{A,B}$ defects, the renormalization of the bands is strongly reduced at energies between the two resonant states in the DOS in Fig. 13 as if the two types of defects cancel the effect of each other. We can again rationalize these findings by considering the GF in the valence and conduction band subspace. As otherwise identical defects on the $A$ and the $B$ sublattices must be considered as different types of defects, the total self-energy is given by the average over the self-energies for the individual $A,B$ defects as in Eq. (27), which here amounts to averaging the matrix structure of the self-energies for $A$ and $B$ sublattice defects [126,127].

For the DFT self-energy, we find that the sublattice-averaged self-energy is almost perfectly diagonal (not shown), whereas the right plot shows the individual intravalley and intervalley matrix elements in Fig. 3. Our findings by considering the GF in the valence and conduction band subspace. As otherwise identical defects on the $A$ and the $B$ sublattices must be considered as different types of defects, the total self-energy is given by the average over the self-energies for the individual $A,B$ defects as in Eq. (27), which here amounts to averaging the matrix structure of the self-energies for $A$ and $B$ sublattice defects [126,127].

For the DFT self-energy, we find that the sublattice-averaged self-energy is almost perfectly diagonal (not shown), whereas the right plot shows the individual intravalley and intervalley matrix elements in Fig. 3. Our findings by considering the GF in the valence and conduction band subspace. As otherwise identical defects on the $A$ and the $B$ sublattices must be considered as different types of defects, the total self-energy is given by the average over the self-energies for the individual $A,B$ defects as in Eq. (27), which here amounts to averaging the matrix structure of the self-energies for $A$ and $B$ sublattice defects [126,127].

This could also have been anticipated from the TB self-energy where the sublattice average obviously eliminates the off-diagonal elements in Eq. (49) and $\Sigma^{m}_{k}(\epsilon)=\Sigma_{0}(\epsilon)$. Thus, for identical defects distributed equally on the $A$ and $B$ sublattices, overall sublattice symmetry and, hence the chirality of the graphene states, is conserved.

The two bottom plots in Fig. 15 show, respectively, the DFT and TB self-energies for N substitutions on both sublattices. Again, there is an excellent qualitative agreement between the TB and DFT self-energies, and quantitative differences can be attributed to the factors mentioned in Sec. V A 1 above. While the structure in the self-energy due to the pole in the $T$ matrix is retained, the form of $\text{Re} \Sigma$ in the vicinity of the Dirac point does evidently not give rise to a band-gap opening, but only a small downshift of the bands also visible in the right plot of Fig. 16.

In the case of both $V_{A,B}$ and $N_{A,B}$ defects, the reduction of the band renormalization at energies immediately above and below the Dirac point can be attributed to a partial cancellation between the real parts of the two self-energies in this energy range. This follows straightforwardly from the fact that the self-energy for $V_A+N_B$ defects resembles a shifted version of the self-energy for $N_A+N_B$ defects in the bottom plot of Fig. 16 with the pole structure centered around the position of the bound state in Fig. 13.

B. Quasiparticle scattering and transport

In this section, we study in further detail the disorder-induced quasiparticle scattering responsible for the spectral linewidth broadening in Figs. 14 and 16 as well as its impact on the transport properties of graphene. In order to avoid complicating the discussion with potential band-gap openings, we here focus on sublattice-symmetric disorder.

In Fig. 17 we show the linewidth broadening in graphene with $A+B$ nitrogen substitutionals ($c=0.1\%$) as a function of the on-shell energy on the Dirac cone. The left plot shows a comparison between the Born and $T$-matrix approximations, whereas the right plot shows the individual intravalley and intervalley contributions to the $T$-matrix linewidth in the left plot. Note that the energy dependence of the linewidth has been obtained from $k$ points along the $\Gamma-K-M$ path in the BZ, and the fact that it forms a single continuous curve along the two line segments shows that it is highly isotropic.

While the energy dependence of the linewidth broadening in the Born approximation reflects the energy dependence of the density of states of pristine graphene in Fig. 13, the $T$-matrix linewidth is strongly electron-hole asymmetric with a pronounced peak on the electron side due to resonant scattering. Also, on the hole side where the DOS of pristine and nitrogen-substituted graphene are almost identical (cf. Fig. 13), does the $T$ matrix yield a strong renormalization of the Born approximation, with an almost energy-independent linewidth broadening.

The separation into intravalley and intervalley scattering contributions in the right plot of Fig. 17 reveals that the two types of scattering processes contribute equally to the linewidth broadening at positive energies where resonant scattering dominates. This is in agreement with the TB model in Eqs. (51) and (49). On the contrary, this is not the case on the hole side where intravalley scattering is stronger than intervalley scattering, which can be attributed to the different intravalley and intervalley matrix elements in Fig. 3. Our finding for the strong electron-hole asymmetry in the intervalley scattering rate is in excellent qualitative agreement with recent magnetotransport measurements where the intervalley rate was extracted from the weak localization (WL) correction to the conductivity in nitrogen-doped graphene [135] and graphene with point defect created by ion bombardment [136].

In Fig. 18, the left and center plots show the low-temperature transport characteristics of disordered graphene with $a=0.01\%$ concentration of sublattice-symmetric vacancies ($V_A+V_B$) and N substitutions ($N_A+N_B$, respectively. The plots show the conductivity (left y axis) and mobility (right y axis) as a function of the Fermi level, with the corresponding carrier density scaling as $n \approx (E_F / 70 \text{ meV})^2 \times 10^{12} \text{ cm}^{-2}$. For both types of defects, the transport exhibits a strong electron-hole asymmetry in the conductivity/mobility which is inherited from the resonant-scattering-induced asymmetry in the underlying scattering rates (cf. Fig. 17). Similar electron-hole asymmetries in the transport characteristics of disordered graphene have been addressed in other theoretical works [83,85,127], and demonstrated experimentally in defective and nitrogen-doped graphene [74,135].
Due to the fact that the quasibound states for the two defects considered in Fig. 18 are, respectively, on the hole and electron sides of the Dirac point (cf. Fig. 13), the two sets of conductivities/mobilities are almost mirror-symmetric versions of each other. Considering the large difference in the value of the matrix elements for $V_A$ and $N_A$ defects in Fig. 3, it is perhaps surprising that the magnitude of the conductivities/mobilities are almost identical when comparing the electron (hole) side for $V_A + V_B$ with the hole (electron) side for $N_A + N_B$. However, as witnessed by the Born matrix element which together with the DOS determines the overall magnitude of the Born scattering rate, simply does not reflect the magnitude of the true scattering rate given by the $T$ matrix. This holds, in particular, for strong defects where the renormalization of the Born scattering amplitude is most significant.

In the right plot of Fig. 18 we show the ratio between the transport and quantum scattering times as a function of the Fermi level. (Right) Ratio between the quantum and transport scattering times. Parameters: $\eta = 25$ meV, $v_F = 10^6$ m/s, and caption of Fig. 13.

C. Adatoms and adsorbates

In addition to the in-plane defects in graphene considered above, adatoms and molecular adsorbates sitting on top of graphene are also of high relevance as they can be used to functionalize and dope graphene, but may at the same time dominate the transport properties due to resonant scattering off the adatom and adsorbate levels.

In a recent work on disordered Li-decorated graphene, we demonstrated that in $T$-matrix descriptions of adatoms (we expect that the same holds for other types of adatoms and adsorbates), it is essential to express the $T$ matrix and the Dyson equation in Eqs. (19) and (22) in a “complete” Bloch-state basis; i.e., the basis must include bands which describe the electronic structure of both graphene and the surface region where the adatoms are located.

From the point of view of first-principles calculations, the importance of using a complete basis is not a surprising observation. On the other hand, $T$-matrix studies of graphene with adatoms or adsorbates have consistently been based on simple TB models considering a single impurity level coupled to the $\pi$ bands of graphene. While this can be expected to capture, e.g., resonant scattering at a qualitative level, it does not account for the fact that the impurity level itself may depend on the impurity concentration via their coupling to so-called surface states. The two, i.e., resonant scattering and the position of the impurity level, are obviously interconnected and must hence be treated in a self-contained framework.

In the $T$-matrix approach outlined here in Secs. II and III, electronic levels of adatoms and adsorbates enter through the second term in the defect potential in Eq. (12) and emerge as poles in the $T$ matrix. In this respect, the method presented here must be expected to give a more complete description of the spectral properties and defect scattering in disordered 2D materials with adatoms and adsorbates.
on the Born approximation (see, e.g., Refs. [37–41]). As witnessed by the examples included here, this is a critical step in 2D materials where the Born approximation often breaks down and fails to capture even the qualitative picture.

Our work has identified some of the main technical challenges associated with first-principles T-matrix calculations as described in Sec. III C. For example, the large matrix dimensions and memory requirements encountered in the solution of the matrix equation in Eq. (43) call for careful parallelization considerations, beyond the simple multitreading/shared memory approach adopted here. Alternatively, the T-matrix equation can be solved in real-space as described in Appendix B by using the supercell LCAO basis \( \{ \phi_{\text{full}} \} \) in Eq. (4) and with a subsequent transformation to the Bloch-state basis as in Eq. (5). This has the immediate advantage that the dimensions of the matrices in Eq. (43) will be fixed to the number of LCAO basis functions in the supercell, which is rather low (~1000–10 000 for the supercell sizes considered here). However, this comes at the cost of having to perform the transformation in Eq. (5) to the desired \( k \)-point grid for each energy in the \( T \)-matrix, but this is manageable and can be more efficient when only the \( k \)-diagonal elements of the \( T \)-matrix (self-energy) are needed.

Irrespective of the strategy chosen for the solution of the \( T \)-matrix equation, it is essential to use nonuniform \( k \)-point samplings of the BZ in order to achieve a satisfactory energy resolution in subsequent calculations of, e.g., spectral properties, scattering rates, or transport properties; see, e.g., Refs. [41,44,45,144] for other recent developments in this direction. As this allows for energy resolutions of the order of meV, our \( T \)-matrix method is advantageous in comparison to Kubo-based approaches [83,85,92,122] for the calculation of the low-temperature longitudinal conductivity and its dependence on the Fermi energy (carrier density) in dilute, disordered materials.

In addition to the technical aspects of the implementation and the electronic properties discussed in this work, there are several interesting extensions to be considered in future works. For example, a generalization to spin-dependent defect potentials \( \hat{V}_i = \sum_{\mu} \hat{V}_{\mu,i}(\mathbf{r}) \otimes \hat{\sigma}_\mu \), accounting for the local change in the spin-orbit interaction around the defect is straightforward, and would allow to address spin-orbit and spin-flip scattering, and hence defect-mediated spin relaxation [39].

A generalization of our method to the treatment of charged defects due to filling of bound defect states by extrinsic carriers presents another highly relevant extension of this work. This requires a self-consistent framework as well as a proper treatment of the resulting long-range Coulomb contribution to the defect potential like in calculations of long-range electron-phonon interactions [145,146]. Preliminary steps for modeling charged defects have recently been reported [68].

Finally, extensions to other 2D materials and vdW multilayer structures [147] as well as to solutions of the Boltzmann equation based on first-principles inputs for the band structure, band velocities, and \( T \)-matrix scattering amplitude (cf. Appendix A) will be important for the future characterization of the electronic and transport properties of new 2D materials. Also, our method for calculating the defect matrix elements paves the way for new diagrammatic first-principles treatments of, e.g., excitons and optical properties [123], as well as transport phenomena such as, e.g., localization [88,89] and anomalous Hall [92,148–151] effects in disordered 2D materials.

### VII. Conclusions

In conclusion, we have presented a DFT-based first-principles method for the calculation of defect matrix elements for realistic descriptions of impurities, defects, substitutional, adatoms, adsorbates, etc., in 2D materials. In combination with a full first-principles-based evaluation of the \( T \)-matrix approximation for the disorder self-energy, we have developed a powerful parameter-free first-principles framework for the description of bound defect states, spectral properties, quasiparticle and carrier scattering, and transport in disordered 2D materials. In spite of the fact that the focus here has been on 2D materials, the method is completely general and can be applied also to 1D and 3D materials. The method is implemented in the GPAW electronic structure code [48–50].

We first applied the method to defects in the two monolayer TMDs MoS\(_2\) and WSe\(_2\). We demonstrated that both vacancies and substitutional oxygen give rise to a series of in-gap bound states with some of the states exhibiting a large spin-orbit-induced splitting. As we have discussed in a recent work [103], the presence of in-gap states leads to charging of the defect sites in the extrinsic (i.e., gated) materials, and the resulting charged-impurity scattering has detrimental consequences for the achievable mobility. However, interestingly we find that Se vacancies and oxygen substitutionals in WSe\(_2\) only introduce empty in-gap states above the intrinsic Fermi level, implying that these defects will remain charge neutral in extrinsic \( p \)-type WSe\(_2\). In the transport characteristics of high-quality vdW WSe\(_2\) devices \( (n_{\text{dis}} \sim 10^{10–10^{11}} \text{cm}^{-2}) \) free from charged impurities in the substrate [5], this manifests itself in a record-high low-temperature mobility \( \mu \sim 15 000–35 000 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1} \), which surprisingly decreases with the carrier density (Fermi energy). The unconventional density dependence of the mobility can be traced back to a strong renormalization of the Born scattering amplitude by multiple-scattering processes accounted for by the \( T \)-matrix. As a consequence, the quantum and transport scattering times become strongly energy dependent and increase away from the band edge. In conjunction with the severe overestimation of the scattering rate by the Born approximation, this underlines the importance of a \( T \)-matrix treatment of point defects in disordered 2D semiconductors.

We also discussed our previously reported symmetry-induced protection against intervalley scattering by defects in 2D TMDs [27], and showed that it completely suppresses intervalley scattering by, e.g., S vacancies in the conduction band of MoS\(_2\), which has often been suggested as the origin of the intervalley scattering extracted from WL/WAL in 2D MoS\(_2\) [109,111]. This finding furthermore points to the possibility of achieving extremely long valley lifetime even in disordered 2D TMDs.

In the last part, we studied the effect of carbon vacancies and nitrogen substitutionals on the electronic properties of graphene. Here, we found that the two types of defects give rise to quasibound resonant states, respectively, below and
above the Dirac point. While the latter is in agreement with experimental studies of nitrogen substitutionals in graphene [9,79], our finding for the position of the vacancy-induced resonant state below the Dirac point is in contrast to numerous tight-binding studies where it appears at the Dirac point [31–35]. For now, we can only speculate that this is due to an oversimplified treatment of the vacancy defect potential in the tight-binding models.

Studying the spectral properties of disordered graphene with, respectively, sublattice-asymmetric and sublattice-symmetric distributions of vacancies and nitrogen substitutionals, we demonstrated defect concentrations of the order of $c \sim 1\% \left( n_{\text{dis}} \sim 10^{13} \text{ cm}^{-2} \right)$ are required in order to see fingerprints of the resonant states in the spectral function measured in ARPES. We furthermore showed that sublattice-asymmetric disorder with the defects located exclusively on one of the sublattices opens a concentration-dependent band gap in graphene, which for the above-mentioned concentration is of the order of $\sim 100$ meV. For sublattice-symmetric disorder, the spectrum again becomes gapless, but retains its characteristic form at the position of the quasibound states caused by strong resonant scattering. In the presence of equal concentrations of sublattice-symmetric vacancies and nitrogen substitutionals, the band renormalizations due to the two types of defects cancel each other, which results in a less dramatic deformation of the Dirac cone in the vicinity of the quasibound states.

Finally, we demonstrated that the transport characteristics of disordered graphene become strongly electron-hole asymmetric in the presence of quasibound resonant states.

Altogether, our first-principles-based $T$-matrix method is an important step toward accurate modeling of realistic defects and their impact on the electronic properties of disordered materials.

ACKNOWLEDGMENTS

The author would like to thank J. H. J. Martiny, J. J. Mortensen, N. Papier, M. Brandbyge, T. Olsen, and A.-P. Jauho for stimulating discussions and comments on the manuscript. K.K. acknowledges support from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 713683 (COFUNDfellowsDTU). The Center for Nanostructured Graphene (CNG) is sponsored by the Danish National Research Foundation, Project No. DNRF103.

APPENDIX A: NUMERICAL SOLUTION OF THE BOLTZMANN EQUATION

The Boltzmann equation in Eq. (39) can be recast as a matrix equation in the composite band and wave-vector index $(n, k)$,

$$ C \bar{f} = b, \quad \bar{f}_{nk} = \frac{\delta f_{nk}}{q |E| \frac{\partial f_{nk}}{\partial \epsilon}|_{\epsilon=\epsilon_{nk}}}, \quad (A1) $$

which is solved for $\bar{f}_{nk}$. Here, the matrix elements of the collision matrix and the vector on the right-hand side are given, respectively, by

$$ C_{nk,n'k'} = -\delta_{nk,n'k'} \sum_{n''k''} P_{nk,n'k''} + P_{nk,n'k'}, \quad (A2) $$

$$ b_{nk} = v_{nk} \cdot \hat{E}, \quad (A3) $$

where $P_{nk,n'k'}$ is the $T$-matrix transition rate in Eq. (40), $v_{nk}$ is the band velocity, and $\hat{E} = E/|E|$ is a unit vector in the direction of the applied electric field. It should be noted that no approximations have been invoked in the solution of the Boltzmann equation outlined above which applies in the case of elastic scattering.

The matrix form of the Boltzmann equation in Eq. (A1), appended with the additional particle-conserving constraint $\sum_{nk} \delta f_{nk} = 0$ on the distribution function, can be solved with a standard least-squares method based on a singular-value decomposition of the collision matrix.

APPENDIX B: T-MATRIX EQUATION IN LCAO BASIS

As an alternative to the Bloch-state formulation of the $T$-matrix equation in Eq. (19) of the main text, it may for practical reasons be advantageous to use the LCAO supercell basis as discussed in Sec. VI.

In the nonorthogonal LCAO basis $|\phi_{nk}\rangle$, the completeness relation takes the form $\sum_{n'l'k'} |\phi_{nk}\rangle \langle \phi_{n'l'k'}| = \hat{1}$, where $S^{-1}$ is the inverse of the overlap matrix defined by $S_{n'l'}^{kl} = \langle \phi_{nk} | \phi_{n'l'} \rangle$. Considering the matrix elements $T_{n'l'}^{kl}(\epsilon) = \langle \phi_{nk} | \hat{T}(\epsilon) | \phi_{n'l'} \rangle$ and inserting completeness relations in the $T$-matrix equation $\hat{T}(\epsilon) = \hat{V} + \hat{\Sigma}(\epsilon) \hat{T}(\epsilon)$, the latter becomes

$$ \hat{T}_{i',kl}(\epsilon) = \hat{V}_{i',kl} + \sum_{k'l'} \hat{V}_{i',kl} \hat{\Sigma}_{k'l'}^{0}(\epsilon) \hat{T}_{i,l'j'}(\epsilon), \quad (B1) $$

where the orbital index has been omitted for brevity and $\hat{\Sigma}^{0} = S^{-1} \hat{\Sigma} S^{-1}$.

As in Sec. III, this can be recast as a matrix equation in the cell and orbital indices,

$$ [1 - \mathbf{V} \hat{\Sigma}_{\epsilon}^{0}(\epsilon)] \mathbf{T}(\epsilon) = \mathbf{V}, \quad (B2) $$

with the GF given by

$$ \hat{\Sigma}_{\epsilon}^{0}(\epsilon) = [\epsilon \mathbf{S} - \mathbf{H}_{0}]^{-1}. \quad (B3) $$

The dimension of the matrices is here given by the number of LCAO orbitals in the supercell.


KRISTEN KAASBJERG

PHYSICAL REVIEW B 101, 045433 (2020)


