Carrier mobility is one of the defining properties of semiconductors. Significant progress on parameter-free calculations of carrier mobilities in real materials has been made during the past decade; however, the role of various approximations remains unclear and a unified methodology is lacking. Here, we present and analyse a comprehensive and efficient approach to compute the intrinsic, phonon-limited drift and Hall carrier mobilities of semiconductors, within the framework of the first-principles Boltzmann transport equation. The methodology exploits a novel approach for estimating quadrupole tensors and including them in the electron-phonon interactions, and capitalises on a rigorous and efficient procedure for numerical convergence. The accuracy reached in this work allows to assess common approximations, including the role of exchange and correlation functionals, spin-orbit coupling, pseudopotentials, Wannier interpolation, Brillouin-zone sampling, dipole and quadrupole corrections, and the relaxation-time approximation. A detailed analysis is showcased on ten prototypical semiconductors, namely diamond, silicon, GaAs, 3C-SiC, AlP, GaP, c-BN, AlAs, A1Sb, and SrO. By comparing this extensive dataset with available experimental data, we explore the intrinsic nature of phonon-limited carrier transport and magnetotransport phenomena in these compounds. We find that the most accurate calculations predict Hall mobilities up to a factor of two larger than experimental data; this could point to promising experimental improvements in the samples quality, or to the limitations of density-function theory in predicting the carrier effective masses and overshielding the electron-phonon matrix elements. By setting tight standards for reliability and reproducibility, the present work aims to facilitate validation and verification of data and software towards predictive calculations of transport phenomena in semiconductors.
this effect arises predominantly from the change of the vibrational eigenmodes induced by the quadrupole term of the dynamical matrix.

In this work we also find that Wannier-function interpolations (especially those associated with the conduction bands) converge more slowly than previously estimated with respect to the sampling of the coarse grid employed in the Wannier representation. As a result, in order to obtain carrier mobilities with an accuracy of 1% for fixed fine grids, it is often necessary to use coarse grids including up to $18^3$ points, still providing a speed-up of two to three orders of magnitude with respect to brute force sampling. We also find that a good indicator of convergence is provided by the mobility effective mass, as defined below. The sampling of Wannier-interpolated quantities requires between $80^3$ and $250^3$ $k$ and $q$ grids, with the exception of the electron mobility of GaAs which requires a $500^3$ grid due to its very small electron effective mass. We also show that the mobility and Hall factor converge linearly with grid spacing, allowing for a simple extrapolation which achieves an accuracy of 1%

This work answers crucial methodological and physical questions. On the technical side, it establishes well-defined criteria for high-accuracy calculations of transport coefficients, clarifying the role and the level of (i) the approximations in the underlying first-principles calculation to extract materials parameters (exchange-and-correlation functionals, pseudopotentials, lattice parameters and SOC), (ii) the interpolation procedures for quasiparticle dispersions and interactions, and (iii) the numerical techniques that ensure the efficient convergence of transport coefficients. On the physical side, the ability to describe with high accuracy transport phenomena provides detailed access to the carrier dynamics and to the inherent limits of the electrical transport properties of key prototypical semiconductors. Crucially, since usually it is the Hall mobility that is measured rather than the drift mobility, this work allows a detailed and unambiguous comparison with experimental data.

The manuscript is organized as follows. In Section II we briefly present the theory for computing phonon-limited drift mobilities. We then extend the theory to include a small finite external magnetic field, to access the Hall mobility. We also present the Wannier interpolation scheme employed here, and discuss how to obtain accurate electron velocities. We also discuss the multipole expansion of the long-range part of the dynamical matrix and electron-phonon matrix elements. Finally, we describe how we use adaptive broadening in practical calculations.

In Section III we report our interpolation of the electron-phonon matrix elements and deformation potential including dynamical quadrupole, and show that the interpolated values reproduce density functional perturbation theory calculations. We then discuss typical coarse and fine grid convergence rates. We conclude the section by showing that the mobility and the Hall factor converge linearly with the spacing of the fine grid, allowing for a linear extrapolation to the limit of exact momentum integration.

In Section IV we present and discuss our results, starting with analysis of the band structures, phonon dispersions, and effective masses. We then assess the quality of various popular approximations, such as the neglect of the non-local pseudopotential contribution to the electron velocity, the neglect of SOC, the neglect of quadrupole correction, the use of the relaxation time approximation, the effect of the isotropic Hall factor approximation, the effect of the exchange-correlation functional, the effect of lattice parameters and pseudopotential choice. In addition, we identify the dominant scattering mechanisms responsible for limiting the carrier mobility. Finally, we conduct an in-depth comparison with available experimental mobility data, and assess the overall predictive power of the BTE. We offer our conclusions in Section V.

II. THEORY

In this Section we first present the main equations for calculating the drift mobility, and their generalization to the case of vanishing magnetic field, yielding the Hall mobility. We then present the Wannier interpolation scheme employed here, and discuss how to obtain accurate electron velocities. We also discuss the multipole expansion of the long-range part of the dynamical matrix and electron-phonon matrix elements. Finally, we describe how we use adaptive broadening in practical calculations.

A. Drift mobility

The low-field phonon-limited charge carrier mobility is calculated as

$$\mu_{\alpha\beta} = \frac{-1}{V_{uc} n_e} \sum_{\alpha} \int \frac{d^3k}{\Omega_{BZ}} v_{nko} \partial E_{\beta} f_{nk}.$$  \hspace{1cm} (1)

Here $\alpha, \beta$ run over the three Cartesian directions and $\partial E_{\beta} f_{nk} = \langle \partial f_{nk} / \partial E_{\beta} \rangle_{E=0}$ is the linear variation of the electronic occupation function $f_{nk}$ in response to the electric field $E$. $V_{uc}$ is the unit cell volume, $\Omega_{BZ}$ the first Brillouin zone volume, $v_{nko} = \hbar^{-1} \partial \varepsilon_{nk} / \partial k_\alpha$ is the band velocity for the Kohn-Sham state $\varepsilon_{nk}$, and $n_e = (1/V_{uc}) \sum_{\alpha} \int (d^3k/\Omega_{BZ}) f_{nk}^0$ is the carrier concentration. $f_{nk}^0$ is the Fermi-Dirac occupation function at
equilibrium (in the absence of fields). We follow the notation of Refs. [14]. The BTE describes the detailed balance between the carriers’ populations moving through phase space under the action of the driving electric field $\mathbf{E}$ and the carriers scattered by phonons. It can be derived from the Kadanoff-Baym equation of motion by approximating the Hartree and exchange-correlation potential, taking the diagonal Bloch state projection and assuming a spatially homogenous electric field $\mathbf{E}$. One then obtains the quantum time-dependent BTE that can be further simplified by considering a time-independent external field (DC), static electron-one-phonon interactions and adiabatic phonons to give:

$$\epsilon \mathbf{E} \cdot \frac{1}{\hbar} \frac{\partial f_{nk}}{\partial \mathbf{k}} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int \frac{d^3q}{\Omega_{\text{BZ}}} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2$$

$$\times \left[ f_{nk}(1 - f_{mk+q}) \delta(\epsilon_{nk} - \epsilon_{mk+q} + \hbar \omega_{q}) n_{q} + f_{nk}(1 - f_{mk+q}) \delta(\epsilon_{nk} - \epsilon_{mk+q} - \hbar \omega_{q}) n_{q} + 1 \right]$$

$$- (1 - f_{nk}) f_{mk+q} \delta(\epsilon_{mk+q} - \epsilon_{nk} - \hbar \omega_{q}) n_{q} + 1$$

$$- (1 - f_{nk}) f_{mk+q} \delta(\epsilon_{mk+q} - \epsilon_{nk} + \hbar \omega_{q}) n_{q} \right],$$

where $n_{q}$ is the Bose-Einstein distribution. The electron-phonon matrix elements $g_{mn\nu}(\mathbf{k}, \mathbf{q})$ are the amplitudes for scattering from an initial state $nk$ to a final state $mk+q$ via the emission or absorption of a phonon of frequency $\omega_q$. We obtain $\partial E_{\beta} f_{nk}$ required for Eq. (1) by taking the field derivative of Eq. (2) at vanishing field. This yields the linearized Boltzmann transport equation (BTE) [11]:

$$\partial E_{\beta} f_{nk} = \epsilon v_{nk}\beta \frac{\partial f_{nk}}{\partial \epsilon_{nk}} \tau_{nk} + \frac{2\pi \tau_{nk}}{\hbar} \sum_{m,\nu} \int \frac{d^3q}{\Omega_{\text{BZ}}} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2$$

$$\times \left[ (n_{q} + 1 - f_{0}) \delta(\epsilon_{nk} - \epsilon_{mk+q} + \hbar \omega_{q}) \right]$$

$$+ (n_{q} + f_{0}) \delta(\epsilon_{nk} - \epsilon_{mk+q} - \hbar \omega_{q}) \right] \partial E_{\beta} f_{nk} + q.$$  \tag{3}

with $\tau_{nk}$ being the total scattering lifetime and the inverse $\tau_{nk}^{-1}$ is the scattering rate given by:

$$\tau_{nk}^{-1} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int \frac{d^3q}{\Omega_{\text{BZ}}} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2$$

$$\times \left[ (n_{q} + 1 - f_{0}) \delta(\epsilon_{nk} - \epsilon_{mk+q} + \hbar \omega_{q}) \right]$$

$$+ (n_{q} + f_{0}) \delta(\epsilon_{nk} - \epsilon_{mk+q} - \hbar \omega_{q}) \right].$$ \tag{4}

A common approximation that we refer to as the self-energy relaxation time approximation (SERTA) consists in neglecting the second term on the right-hand side of Eq. (3). The mobility then takes the simpler form [12]:

$$\text{SERTA:} \quad \mu_{\alpha\beta} = \frac{-e}{V_{uc} n_c} \sum_n \int \frac{d^3k}{\Omega_{\text{BZ}}} \frac{\partial f_{nk}}{\partial \epsilon_{nk}} \frac{\partial f_{nk}}{\partial \epsilon_{nk}} \tau_{nk}.$$ \tag{5}

In order to analyze the role of band curvature in mobility calculations, we introduce a “mobility effective mass” as follows:

$$\frac{1}{m_{mob}^*} = \frac{-1}{3V_{uc} n_c} \sum_{\alpha\beta} \int \frac{d^3k}{\Omega_{\text{BZ}}} \frac{\partial f_{nk}}{\partial \epsilon_{nk}} \partial_{\epsilon_{nk}}^2 \tau_{nk}. \tag{6}$$

This definition is obtained by setting $\tau_{nk}$ to a constant value $\tau$ in Eq. (5), and by requiring that the resulting mobility can be expressed in the standard Drude form with the same relaxation time $\tau$:

$$\mu = \frac{e\tau}{m_{mob}^*}. \tag{7}$$

Lastly, we introduce an approximate version of the SERTA mobility, which is useful to disentangle the convergence of electron band structures from that of the phonon energies and electron-phonon matrix elements. To this aim we define the “constant adiabatic relaxation time approximation” (CARTA) as the mobility obtained by taking the SERTA of Eq. (3) with a constant electron-phonon matrix elements and vanishing phonon frequencies: $|g_{mn\nu}(\mathbf{k}, \mathbf{q})| = g^2$, with $g$ an arbitrary constant with units of energy, and $\omega_q = 0$.

$$\mu_{\alpha\beta}^{\text{CARTA}} = \frac{-e^2 g^2}{2\pi V_{uc} n_c} \sum_n \int \frac{d^3k}{\Omega_{\text{BZ}}} \frac{\partial f_{nk}}{\partial \epsilon_{nk}} \partial_{\epsilon_{nk}}^2 \frac{1}{\tau}.$$ \tag{8}

B. Hall mobility

Calculation of the direct current Hall coefficient has seen a resurgence of interest [17–19, 51]. Experimentally, Hall mobility measurements are more common than time-of-flight measurements of drift mobilities due to their superior accuracy and simplicity. This is reflected in the literature by a significantly higher number of available Hall mobility data compared to drift mobilities. However, Hall measurements are performed in the presence of an external finite magnetic field, which introduces an additional Lorentz force on the carriers, thereby altering the mobility. To compare with experiment, we have to augment the BTE of Eq. (3) to account for the additional magnetic field B [17, 19, 47]:

$$\left[ 1 - \frac{e}{\hbar} \tau_{nk} (\mathbf{v}_{nk} \times \mathbf{B}) \cdot \nabla_{\mathbf{k}} \right] \partial E_{\beta} f_{nk} = \epsilon v_{nk}\beta \frac{\partial f_{nk}}{\partial \epsilon_{nk}} \tau_{nk}$$

$$+ \frac{2\pi \tau_{nk}}{\hbar} \sum_{m,\nu} \int \frac{d^3q}{\Omega_{\text{BZ}}} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2$$

$$\times \left[ (n_{q} + 1 - f_{0}) \delta(\epsilon_{nk} - \epsilon_{mk+q} + \hbar \omega_{q}) \right]$$

$$+ (n_{q} + f_{0}) \delta(\epsilon_{nk} - \epsilon_{mk+q} - \hbar \omega_{q}) \right] \partial E_{\beta} f_{nk} + q.$$ \tag{9}

with $\tau_{nk}$ being the total scattering lifetime defined in Eq. (4). Taking derivatives on both sides of Eq. (9) with respect to the Cartesian components of the magnetic field, $B_{\gamma}$, at zero field, yields an equation for the
linear response coefficients $\partial E_\beta f_{nk}$ and $\partial E_\beta \partial B_\gamma f_{nk}$:

$$\partial^2 E_\beta, B_\gamma f_{nk} = -\frac{e}{\hbar} \tau_{nk}(v_{nk} \times \nabla_k) \partial E_\beta f_{nk}$$

$$+ \frac{2\pi \tau_{nk}}{\hbar} \sum_{\mu\nu} \int \frac{d^3q}{\Omega_{BZ}} |g_{\mu\nu}(k, q)|^2$$

$$\times \left[ (n_{q\nu} + f_{nk}^0) \delta(\varepsilon_{nk} - \varepsilon_{nk+q} + \hbar\omega_{q\nu}) + (n_{q\nu} + f_{nk}^0) \delta(\varepsilon_{nk} - \varepsilon_{nk+q} - \hbar\omega_{q\nu}) \right] \partial^2 E_\beta, B_\gamma f_{nk+q}.$$  

(10)

The Hall conductivity tensor is obtained from the second derivatives of the current density with respect to electric and magnetic field for vanishing fields:

$$\sigma_{a\beta\gamma}^H = -\frac{e}{v_{nc}} \sum_n \int \frac{d^3k}{\Omega_{BZ}} \nu_{nk\alpha} B_\gamma \partial^2 E_\beta, B_\gamma f_{nk}.$$  

(11)

or equivalently directly from Eq. (9) as:

$$\sigma_{a\beta\gamma}^H = -\frac{e}{v_{nc}} \sum_n \int \frac{d^3k}{\Omega_{BZ}} \nu_{nk\alpha} \partial E_\beta f_{nk}(B_\gamma).$$  

(12)

Besides the drift and Hall conductivity and their mobility analogues, a commonly reported quantity is the dimensionless Hall tensor, which is defined as the ratio between the Hall conductivity and the drift conductivity:

$$r_{a\beta\gamma}^H = n_e e \sum_{\delta\epsilon} \sigma_{a\delta\epsilon}^{-1} \sigma_{\delta\epsilon\gamma}^H \sigma_{\epsilon\beta}^{-1} B_\gamma,$$  

(13)

$$= \sum_{\delta\epsilon} \mu_{a\delta\epsilon}^{-1} \mu_{\delta\epsilon\gamma}^H \mu_{\epsilon\beta}^{-1} B_\gamma,$$  

(14)

where Eq. (13) is the tensorial generalization of Eq. (1) from Ref. [52]. For the cubic materials that we studied here, Eq. (13) has only one non-equivalent component, $r_{123}^H$.

A popular approximation to Eq. (13) consists in assuming a parabolic and non-degenerate band extremum, following Ref. [54], p. 118 and Ref. [55], Eq. 3.12. Within this approximation, the isotropic and temperature-dependent Hall factor is given by:

$$r_{iso}^H = \langle \tau^2 \rangle / \langle \tau \rangle^2,$$  

(15)

with

$$\langle \tau^n \rangle = \int_0^\infty \tau^n (xk_B T)x^{3/2}e^{-x}dx / \int_0^\infty x^{3/2}e^{-x}dx.$$  

(16)

Here, $x = \varepsilon / (k_B T)$ and we introduced the distribution function of the total decay rate:

$$\tau(\varepsilon) = \sum_n \int \frac{d^3k}{\Omega_{BZ}} \delta(\varepsilon - \varepsilon_{nk}) \tau_{nk}.$$  

(17)

The numerical solution of a large number of electron-phonon matrix elements in order to converge the double momentum integrals ($k$ and $q$). Various schemes have been developed to deal with this challenge including models with parameters computed from first principles [57], direct evaluation of the electron-phonon matrix elements using density-functional perturbation theory (DFPT) [10], linear interpolation of the scattering rates [11,12], and local orbital implementations [13,14]. Fourier interpolation of the perturbed potential [15,16,17] and interpolation based on MLWFs [18,19]. The use of MLWFs makes the calculations affordable and is the method of choice in this work, where we rely on the EPW software [20,21]. Such interpolation implies subtleties that need to be dealt with carefully, as discussed in Section III C.

C. Interpolation of the electron bands, phonon dispersions, and electron-phonon matrix elements

The interpolation of the various quantities required to compute the drift mobility on ultra-dense grids relies on a discrete Fourier transform followed by a Fourier interpolation at arbitrary crystal momenta. Due to the fact that the discrete Fourier transform is in practice performed on a uniform finite grid, the quality of the interpolation depends on the localization in real space of each quantity.

In the case of the electronic Hamiltonian, we can leverage the phase freedom of the Bloch orbitals to create Wannier functions that are maximally localized in real space [22]. Since the interatomic force constants do not have such gauge freedom, we use direct Fourier interpolation. This interpolation requires some care in the case of polar materials, in order to correctly describe long-range interactions and the splitting of longitudinal-optical (LO) and transverse-optical (TO) modes [23,24]. Electronic Wannier and Bloch states are related by [25]

$$|\Psi_m R_p\rangle = \frac{1}{N_p} \sum_{nk} e^{-ik_\mu R_p n_{nk}} |\Psi_{nk}\rangle,$$  

(18)

$$|\Psi_{nk}\rangle = \sum_{mR_p} e^{ik_\mu R_p} U_{mnk}^\dagger |\Psi_{mR_p}\rangle,$$  

(19)

where $R_p$ is a lattice vector, $N_p$ is the number of unit cells in the Born-von Kármán supercell, corresponding to the number of $k$-points, and $U_{mnk}$ is a unitary rotation matrix that transforms the Bloch wavefunctions to a Wannier gauge:

$$|\Psi_{nk}^W\rangle = \sum_{U_{nmk}} |\Psi_{nk}\rangle.$$  

(20)

In particular, the unitary transformation $U$ can be chosen to obtain MLWFs that minimize the spatial spread functional [25].

In this context, the Hamiltonian $H$ and dynamical matrices $D$ can be transformed from the Bloch to Wannier
where \( N_p, N_{p'} \) are the number of unit cells in the Born-von Kármán supercells for the electrons and phonons, respectively and \( \epsilon_{\alpha \nu q} \) are the eigendisplacement vectors corresponding to the atom \( \kappa \) in the Cartesian direction \( \alpha \) for a collective phonon mode \( \nu \) of momentum \( \mathbf{q} \).

The extension of these two concepts to interpolate the electron-phonon matrix elements has been derived in Ref.\textsuperscript{65} and yields:

\[
g_{mn\kappa\alpha}(\mathbf{R}_p, \mathbf{R}_{p'}) = \frac{1}{N_p N_{p'}} \sum_{\mathbf{k}, \mathbf{q}} e^{-i(\mathbf{k} \cdot \mathbf{R}_p + \mathbf{q} \cdot \mathbf{R}_{p'})} \sum_{m' n' \nu} e^{i \epsilon_{\kappa \alpha} \nu (\mathbf{k} + \mathbf{q})} \langle \mathbf{k}, \mathbf{q} | H_{mn'k} U_{n'k} \rangle\]

\[
\times \sqrt{\frac{2 M_{\kappa \alpha} \hbar^2}{\hbar}} e^{i \epsilon_{\kappa \alpha} \nu (\mathbf{k} + \mathbf{q})} U_{mn'k} U_{n'k}^\dagger (\mathbf{k}, \mathbf{q}) U_{n'k}. \tag{23}
\]

In practice, we compute the electron-phonon matrix elements \( g_{mn\kappa\alpha}(\mathbf{k}, \mathbf{q}) \) on a coarse momentum grid and rely on crystal symmetries to reduce the number of calculations to be performed. To carry out the Wannier transformation from coarse reciprocal space to real space, we first evaluate the matrix elements on the entire Brillouin zone using symmetries. As pointed out in a recent preprint\textsuperscript{66} in the case of SOC the spinor wavefunctions need to be rotated using the SU(2) unitary group. There was a bug in version 5.2 of the EPW software\textsuperscript{67} related to this rotation. The bug has been eliminated in version 5.3, and for the systems tested so far we found a negligible impact. For example, we recalculated the mobilities of silicon, GaAs, and CsPbI\(_3\), and found differences smaller than 0.4% in all cases.

There are additional subtleties related to the Wannier interpolation of the various quantities required to compute the mobility. These aspects will be discussed in Section IV.

### D. Band velocities

A key ingredient required in the calculation of carrier mobility is the carrier velocity as can be seen in Eq. (1). The velocity operator can be expressed in terms of the commutator between the Hamiltonian of the system and the position operator, \( \mathbf{v} = (i/\hbar) [H, \mathbf{r}] \). The matrix elements of \( \mathbf{v} \) in the presence of non-local pseudopotentials are\textsuperscript{68,69}

\[
\mathbf{v}_{nmk} = \langle \psi_{nk} | \hat{p} / m_e + (i/\hbar) [\hat{V}_{NL}, \mathbf{r}] | \psi_{nk} \rangle, \tag{24}
\]

where \( \hat{p} = -i \hbar \partial / \partial \mathbf{k} \) is the momentum operator, \( m_e \) is the electron mass, and \( \hat{V}_{NL} \) is the non-local part of the pseudopotential. Neglecting the term with \( \hat{V}_{NL} \) in the case of a non-local pseudopotential is what we call the “local approximation”\textsuperscript{13}. Within this approximation, Eq. (24) reads:

\[
v_{nmk}^{loc} = \frac{\hbar k}{m_e} \delta_{mn} + \frac{\hbar}{m_e} \int dr u_{nk}^*(\mathbf{r}) \frac{\partial}{\partial k} u_{nk}(\mathbf{r}), \tag{25}
\]

where \( u_{nk} \) are the periodic part of the Bloch wavefunction. If in the second part of Eq. (25) the terms corresponding to the periodic part of the Bloch function are expanded into planewaves, \( u_{nk} = \sum_G c_{nk}(G) e^{iG \cdot \mathbf{r}} \), we obtain the simple velocity expression in the local approximation:

\[
v_{nmk}^{loc} = \frac{\hbar k}{m_e} \delta_{mn} + \sum_G \frac{\hbar}{m_e} c_{nk}(G) c^*_G(G). \tag{26}
\]

We note that this expression is only valid in the case of local norm-conserving pseudopotentials. In practice, we will show that this local approximation is too crude and one needs to compute the matrix elements of the commutator in Eq. (24). We carry out this procedure in the Wannier representation, as described in Refs. 72-75. In particular, we construct the velocities interpolated at an arbitrary momentum \( \mathbf{k}' \) following Eq. (31) of Ref. 73:

\[
v_{nmk'} = \frac{1}{\hbar} \frac{\partial}{\partial k} H_{nmk'} - i \frac{\hbar}{m_e} (\varepsilon_{nk'} - \varepsilon_{mk'}) A_{nmk'}, \tag{27}
\]

where \( A_{nmk'} \) is the Berry connection defined as:

\[
A_{nmk'} = i \langle m_{nk'} \mid \partial \psi_{nk'} / \partial k' \rangle. \tag{28}
\]

We construct the Berry connection in the Wannier basis [Eq. (29)] on the coarse \( \mathbf{k} \)-point grid and compute the derivative using finite differences\textsuperscript{63}:

\[
A_{nmk'} = i \sum_b w_b \langle m_{nk'} | \psi_{mk'+b}^W \rangle - \delta_{nm}, \tag{29}
\]

where \( \mathbf{b} \) are the vectors connecting \( \mathbf{k} \) to its nearest neighbours with \( w_b \) weighted associated with each shell of neighbours \( |\mathbf{b}| \). These vectors are constructed in such a way as to satisfy\textsuperscript{76}:

\[
\sum_b w_b a_\alpha b_\beta = \delta_{\alpha\beta}, \tag{30}
\]

using the smallest number of neighbours. The rotated overlap matrices (from Bloch to Wannier basis) are obtained as:

\[
\langle u_{nk}^W | u_{m'k+b}^W \rangle = \sum_{n' m'} U_{nm'k}^\dagger M_{n'm'}(\mathbf{k}, \mathbf{b}) U_{m'm'k+b}, \tag{31}
\]

where \( M_{nm}(\mathbf{k}, \mathbf{b}) = \langle u_{nk} | u_{m'k+b} \rangle \) is the overlap matrix between the cell-periodic Bloch eigenstates at neighboring points \( \mathbf{k} \) and \( \mathbf{k} + \mathbf{b} \).
We first Fourier transform the position matrix into real space following Eq. (43) of Ref. 73:

$$\langle \Psi_{n0}| \hat{r} | \Psi_{mR} \rangle = \frac{1}{N_p} \sum_{k} e^{-i k \cdot R} A_{nmk}^W, \quad (32)$$

where $k$ belongs to the same grid employed for the Wannier interpolation of the band structure. The quantity on the left hand side decreases rapidly with $R$ owing to the localization of MLWFs, therefore we can use these real-space quantities to interpolate back to arbitrary momenta $k'$:

$$A_{nmk',\alpha}^W = \sum_{R} e^{i k' \cdot R} \langle \Psi_{n0}| \hat{r} | \Psi_{mR} \rangle. \quad (33)$$

We do the same for the $k'$ derivative of the Hamiltonian following Eq. (38) of Ref. 73:

$$\frac{\partial}{\partial k'} H_{nmk'}^W = \sum_{R} e^{i k' \cdot R} \langle \Psi_{n0}| \hat{H} | \Psi_{mR} \rangle. \quad (34)$$
Finally we un-rotate from the Wannier basis to the Bloch state basis using the diagonalizers of the Hamiltonian:

$$\frac{\partial}{\partial k'} H^\dagger_{nmk} = \sum_{n'm'} U^\dagger_{nm'm'} k' \frac{\partial}{\partial k'} H^W_{n'm'm'} U_{m'mk'},$$  

(35)

$$A^\dagger_{nmk'} = \sum_{n'm'} U^\dagger_{nm'm'} A^W_{n'm'm} U_{m'mk'},$$  

(36)

which gives the desired Eq. (27).

We have implemented the interpolation of the velocity matrix elements including the correction for non-local potentials in the EPW software. Figure 1 shows that the local approximation yields significant errors in the calculated mobilities, ranging from -80% to +40% of the correct value. For this reason, we use the exact velocity matrix elements given by Eq. (27) throughout this study.

To evaluate the quality of the Wannier interpolation of velocities, we perform direct DFT calculations using central finite differences, with six neighbouring points spaced by 5 · 10^-2π/a.

E. Long range corrections: dipoles and quadrupoles

The analytic properties of the long wavelength limit of the electron-phonon matrix elements in bulk polar materials, resulting from the Fröhlich interaction, have been known for a long time, but the generalization to first-principles calculations using Wannier-Fourier interpolation appeared only recently [E28,E29]. Following a line of thinking closely related to the Fourier interpolation of the phonon frequencies [E36,E37,E38], one can decompose the matrix elements into a short- (S) and long-range (L) contribution [E28,E29]:

$$g^S_{mnν}(k, q) = g^L_{mnν}(k, q) + g^C_{mnν}(k, q).$$  

(37)

In Eq. (37), the long-range part is given by the infinite multipole expansion:

$$g^C_{mnν}(k, q) = g^C_{mnν}(k, q) + g^C_{mnν}(k, q) + g^C_{mnν}(k, q) + \ldots,$$  

(38)

where the first term comes from dipole contribution, then quadrupole, octopoles and higher. The analytic form of the long-range part of the dipole contribution is given as [E28].

$$g^C_{mnν}(k, q) = \frac{4\pi}{\Omega} \frac{e^2}{4\pi\varepsilon_0} \sum_{κ} \left[ \frac{\hbar}{2N_{p'} M_κ\omega_{qν}} \right] \sum_{G≠q}$$

× \((G + q) \cdot Z_κ \cdot |e_{qν}| \cdot e^{i(G + q) \cdot τ_κ} \cdot Ψ_{mnν|^r}|Ψ_{nk}|,$$  

(39)

where G is a reciprocal lattice vector, κ is the position of atom κ, Z_κ is the Born effective charge tensor of the atom κ of mass $M_κ$, $e_{qν}$ is the vibrational eigendisplacement vector normalized in the unit cell, $ε^0$ is the vacuum dielectric constant and $ε^\infty$ is the high-frequency dielectric tensor of the material. In the past, a Gaussian filter has been added to Eq. (39). Since this filter breaks the periodicity of the matrix element in reciprocal space, we recommend not to include it in future calculations. One can notice that Eq. (39) diverges as 1/|q| as we approach the zone center. This is an integrable singularity, so that upon performing integrations in reciprocal space one obtains finite values of physical properties.

The quadrupolar component of the matrix element at long wavelength is given by:

$$g^L_{mnν}(k, q) = \frac{4\pi}{\Omega} \frac{e^2}{4\pi\varepsilon_0} \sum_{κ} \left[ \frac{\hbar}{2N_{p'} M_κ\omega_{qν}} \right] \sum_{G≠q}$$

× \((G + q) \cdot (G + q) \cdot \cdot \cdot |e_{qν}| \cdot \cdot \cdot Ψ_{mnν|^r}|Ψ_{nk}|,$$  

(40)

where $Ψ_{mnν|^r}|Ψ_{nk}|$ is a third-rank effective quadrupole tensor defined as:

$$Ψ_{mnν|^r}|Ψ_{nk}| = \frac{1}{2} Q_{νκ|γ}(Ψ_{mk+q}|e^{i(q+G)\cdot r}|Ψ_{nk}),$$  

(41)

where $Q_{νκ|γ}$ is the dynamic quadrupole tensor [E21]. The third-rank quadrupole tensor should obey the same symmetry rules as other third-rank tensors such as the piezo-electric tensor or the second-order magnetoelectric tensor [E22]. One can therefore use tools such as the Bilbao Crystallographic Server [E22] to determine the non-zero components of the quadrupole tensor for a given space group, and their relations. In the case of all the polar materials investigated in this work, the quadrupole tensor takes the form:

$$Q_{νκ|γ} = Q_κ|ε_{νκ|γ}|,$$  

(42)

where $ε_{νκ|γ}$ is the Levi-Civita symbol and $Q_κ$ is an atom-dependent scalar. In the case of silicon and diamond, the effective tensor in Eq. (41) is completely specified by a single scalar [E21].

$$Ψ_{mnν|^r}|Ψ_{nk}| = \frac{1}{2} (-1)^{κ+1} Q_{νκ|γ}(Ψ_{mk+q}|e^{i(q+G)\cdot r}|Ψ_{nk}),$$  

(43)

The quantity $V_{Hxc,νκ}$ in Eq. (41) is the change of the Hartree and exchange-correlation potential with respect to the electric field $E$. This term is null in the case of materials with vanishing Born effective charges and has been shown to be small in the case of polar materials such as GaAs and GaP. Finally, we note that in the $q + G \rightarrow 0$ limit, the wavefunction overlap in
Eqs. [39],[41] can be written in terms of the Wannier rotation matrices $U_{nmk}$ as:
\begin{equation}
\langle \Psi_{nk+q}| e^{i(a+G) \cdot r} |\Psi_{nk} \rangle = |\langle k+q | U^{\dagger}_{kk} | m \rangle |^2.
\end{equation}
A similar approach can be used to enforce the correct behavior of the dynamical matrices at long wavelengths:

$$D_{\kappa \alpha, \kappa' \beta}(q) = D_{\kappa \alpha, \kappa' \beta}^{S}(q) + D_{\kappa \alpha, \kappa' \beta}^{C}(q),$$
where the nonanalytical, direction-dependent term, is given by the contribution of dipole-dipole, dipole-quadrupole and quadrupole-quadrupole terms as:

$$D_{\kappa \alpha, \kappa' \beta}^{C}(q) = \frac{4 \pi e^2}{\Omega} \sum_{G \neq 0} D_{\kappa \alpha, \kappa' \beta}(G + q)$$

$$\quad - \delta_{\kappa \kappa'} \sum_{\kappa''} \sum_{G \neq 0} D_{\kappa \alpha, \kappa' \beta}(G)$$

with:

$$D_{\kappa \alpha, \kappa' \beta}^{C,D}(q) = \frac{e^{i q \cdot (\tau_{\kappa} - \tau_{\kappa'})}}{q \cdot \epsilon_{\kappa} \cdot q} \left[ q \cdot Z_{\kappa}^* \cdot q \cdot Z_{\kappa'}^* \right]$$

$$+ \frac{1}{4} q \cdot q \cdot Q_{\kappa} \cdot q \cdot Q_{\kappa'} + \frac{i}{2} q \cdot Z_{\kappa}^* \cdot q \cdot q \cdot Q_{\kappa'}$$

$$\quad - q \cdot q \cdot Q_{\kappa} \cdot q \cdot Z_{\kappa'}^* \right].$$

In this expression we neglected octopoles and higher-order as well as fourth-order and higher dielectric functions. The sum over the lattice of charges is performed using the Ewald technique, with a parameter $\Lambda$. In our calculations we use $\Lambda = 1$ bohr$^{-1}$ so that the real-space contributions in the Ewald summation can be neglected.

## F. Adaptive broadening

The numerical evaluation of Eqs. (3) and (4) requires to approximate Dirac delta functions by Lorentzian or Gaussian functions of finite broadening. This means that the resulting mobility will depend on the value of the broadening used, and careful convergence tests are required. To circumvent this difficulty, various schemes have been proposed including adaptive broadening and linear tetrahedron method.

We implemented the adaptive broadening approach in the EPW software following Eq. (18) from Ref. 85, where the state-dependent broadening is given by:

$$\eta_{nk}(q) = \frac{\hbar}{\sqrt{12}} \sum_{\alpha} \left[ (v_{q\alpha} - v_{nk+q}) \frac{G_{\alpha}}{N_{\alpha}} \right]^2,$$

where $G_{\alpha}$ are the three primitive reciprocal lattice vectors, $\alpha$ the three crystallographic directions, $v_{nk+q}$ the electronic velocities defined by Eq. (27), $N_{\alpha}$ the $q$-point grid density in the three crystalline directions. Eq. (48) can be scaled arbitrarily, and for this work we choose empirically a factor 1/2 after a number of numerical tests. In addition, the phonon velocity $v_{q\nu}$ is obtained from the momentum derivative of the dynamical matrix, following an approach similar to Eq. (38) of Ref. 73:

$$v_{q\nu\beta} = \frac{1}{2 \omega_{q\nu}} \frac{\partial D_{\mu\nu}(q)}{\partial q_{\beta}} = \frac{1}{2 \omega_{q\nu}} \sum_{R} i R_{\beta} \epsilon_{q R} D_{\mu\nu}(R).$$

In the case of polar materials, the momentum derivative of the long-range dipole part is obtained by taking the derivative of Eq. (46) as:

$$\frac{\partial D_{\mu\nu}(q)}{\partial q_{\beta}} = \sum_{\alpha \neq -q} \frac{\partial D_{\mu\nu}^{c,D}(q+q)}{\partial q_{\beta}}$$

with:

$$\frac{\partial D_{\mu\nu}^{c,D}(q)}{\partial q_{\beta}} = \frac{1}{2 \omega_{q\nu}} \sum_{R} i R_{\beta} \epsilon_{q R} D_{\mu\nu}(R).$$

where we have considered only dipole-dipole interaction term for simplicity. This choice does not change any of the results presented in this work, because it only affects the broadening parameter. This point will be discussed further in Section III F. All results described below were obtained using this adaptive broadening scheme.

## III. COMPUTATIONAL METHODOLOGY

In this section we describe the software and computational parameters employed. We discuss the interpolation of electron-phonon matrix elements using the long-range dipole and quadrupole corrections. We look at the convergence rate with respect to the coarse and fine grid interpolation, and show that the drift mobility, Hall mobility, and Hall factor can be extrapolated to the limit of continuous Brillouin zone sampling. Finally, we propose in Appendix 1, 2 a way to construct an optimal Wigner-Seitz cell for the double three dimensional Fourier interpolation of the electron-phonon matrix elements.

### A. Density functional theory and density functional perturbation theory calculations

We use the Quantum ESPRESSO software distribution with the relativistic Perdew-Burke-Ernzerhof (PBE) parametrization of the generalized gradient approximation (GGA) to density-functional theory.
To ensure reproducibility and follow the FAIR principles (findable, accessible, interoperable, and reusable), all pseudopotentials are made available with the manuscript as well as a tagged release of the EPW software that can be accessed at [DOI: 10.24435/materialscloud:b2-j5].

### B. Construction of Wannier functions and spatial localization

| Material | # | Initial Wind. Froz. | k | # Spread |
|----------|---|---------------------|---|---------|
| C-e      | 8 | C1:sp3              | 203 | 210  | 2.40 |
| C-h      | 8 | C1:sp3              | 203 | 200  | 0.79 |
| Si-e     | 12 | Si1:d+Si1:s        | 163 | 1320 | 5.67 |
| Si-h     | 8 | Si1:sp3            | 203 | 210  | 2.22 |
| GaAs-e   | 8 | Ga:sp3             | 5.75 | 3750 | 8.85 |
| GaAs-h   | 16 | Ga+:As:sp3        | 5.75 | 170  | 3.04 |
| 3C-SiC-e | 8 | Si:sp3             | 5.85 | 234  | 3.96 |
| 3C-SiC-h | 8 | Si:sp3             | 163 | 160  | 1.21 |
| AlP-e    | 8 | Al:sp3             | 2.79 | 183  | 4.08 |
| AlP-h    | 6 | P:p                | 163 | 170  | 3.00 |
| GaP-e    | 8 | Ga:sp3             | 3.99 | 1490 | 8.21 |
| GaP-h    | 6 | P:p                | 163 | 205  | 3.65 |
| c-BN-e   | 8 | B:sp3              | 183 | 260  | 2.51 |
| c-BN-h   | 6 | N:p                | 143 | 205  | 1.22 |
| AlAs-e   | 8 | Al:sp3             | 2.34 | 163  | 8.51 |
| AlAs-h   | 6 | As:p               | 143 | 105  | 3.36 |
| AlSb-e   | 8 | Al:sp3             | 1.36 | 4160 | 9.80 |
| AlSb-h   | 6 | Sb:p               | 129 | 2260 | 4.46 |
| SrO-e    | 12 | Sr:d+Srs          | 5.06 | 7830 | 3.26 |
| SrO-h    | 6 | O:p                | 123 | 30   | 1.38 |

**TABLE II**: Data related to the wannierization of all band structures considered in this work. The valence and conduction manifold are computed separately. We indicate the initial projections and the number of iterations to reach a relative convergence of $10^{-12}$ in the spread, the total spread, and the average spread per Wannier function for the converged k-point grids. For the initial projection, a number next to the atom means that the initial projection is applied to only one of the two equivalent atoms in the primitive cell. The disentanglement window as well as the frozen window are given with respect to the band edge. All calculations include SOC.

When performing calculations of carrier mobility, and especially in the presence of magnetic fields, it is crucial to have symmetric and highly localized Wannier functions. In order to provide a benchmark for future work, here we describe the procedure followed to generate ML-WFs and the properties of the resulting matrix elements.

For diamond we choose 4 $sp^3$ orbitals (8 Wannier functions) located on each atom of the inversion-symmetric unit cell as initial projection for the valence and conduction manifold. This choice leads to a Wannier function spread $\sigma^2$ of 2.40 Å$^2$ per function for the converged coarse k-point grid, and 0.79 Å$^2$ for the valence bands, see Table II.
have been determined from a least-squares fitting of the form
\[ g(R_{\kappa 0}, \eta) \sim 0 \text{ or } \eta = 0 \]
and
\[ \exp(-\hbar x) \text{ in a } 10 \text{ Å range.} \]

The strong localization of these MLWFs leads to a fast spatial decay of the matrix elements in the Wannier representation of the electronic Hamiltonian, the electron velocity, the interatomic force constants, and the electron-phonon matrix elements, down to nine orders of magnitude, as seen in Fig. 2. In all the tests reported here, we find as expected that Wannier functions describing the valence bands are more localized than for the conduction bands.

The case of silicon, the Wannierization of the conduction band manifold is especially challenging. Using the wannier90 software we need 1320 iterations to converge to a relative accuracy of \(10^{-12} \text{ Å}^2\) in the spread (we note in passing the usefulness of the conjugate gradients algorithm such that we changed the default of resetting the search direction to steepest descents every 100 iterations, rather than 5, as default). For the conduction bands, we find that adding one \(d\) and one \(s\) orbital on one of the silicon atoms as initial projection works best, and yields a spread of 5.67 Å². The comparison between the electronic bandstructure of the conduction bands of silicon using this projection and an \(sp^3\) one is shown in Appendix, Fig. 2a. The challenge in obtaining carefully converged effective masses lies in the fact that the conduction band minimum of silicon does not lie at a high symmetry point and is therefore not included in the \(k\)-point grid used for interpolation. The calculation of the conduction band manifold is accelerated in silicon and diamond by excluding the valence bands. The correspond-

| Excluded bands | H | v | D | g(\(R_{\kappa 0}\)) | g(\(R_{\eta}\)) |
|----------------|---|---|---|-----------------|-----------------|
| C-e            | 1-8 | 1.057 | 1.784 | 0.859 | 1.524 | 1.487 |
| C-h            | -  | 1.028 | 1.224 | 0.859 | 1.172 | 1.156 |
| Si-e           | 1-8 | 1.478 | 2.101 | 1.290 | 2.133 | 2.034 |
| Si-h           | -  | 1.710 | 1.617 | 1.290 | 1.547 | 1.644 |
| GaAs-e         | 1-28 | 1.465 | 2.850 | 1.601 | 2.459 | 2.278 |
| GaAs-h         | 1-20 | 1.533 | 1.906 | 1.603 | 2.171 | 2.038 |
| SiC-e          | 1-8 | 1.187 | 1.909 | 1.214 | 1.751 | 1.765 |
| SiC-h          | -  | 1.245 | 1.362 | 1.215 | 1.381 | 1.719 |
| AlP-e          | 1-8 | 1.517 | 2.785 | 1.373 | 2.511 | 2.272 |
| AlP-h          | 1-2 | 9-40 | 1.328 | 1.328 | 1.373 | 1.736 | 2.060 |
| GaP-e          | 1-18 | 35-40 | 1.434 | 2.716 | 1.397 | 2.360 | 2.205 |
| GaP-h          | 1-12 | 19-40 | 1.293 | 1.490 | 1.397 | 1.903 | 1.985 |
| BN-e           | 1-8 | 17-40 | 1.057 | 2.057 | 0.994 | 1.532 | 1.507 |
| BN-h           | 1-2 | 9-40 | 0.976 | 1.091 | 1.000 | 1.254 | 1.539 |
| AlAs-e         | 1-18 | 1.509 | 2.647 | 1.450 | 2.456 | 2.267 |
| AlAs-h         | 1-12 | 19-40 | 1.247 | 1.375 | 1.450 | 1.720 | 2.057 |
| AlSb-e         | 1-26 | 1.642 | 2.712 | 1.536 | 2.751 | 2.437 |
| AlSb-h         | 27-50 | 1.305 | 1.638 | 1.537 | 2.041 | 2.182 |
| SrO-e          | 1-16 | 1.284 | 1.813 | 1.505 | 2.409 | 1.732 |
| SrO-h          | 1-10 | 17-40 | 0.952 | 1.035 | 1.505 | 1.326 | 2.076 |

TABLE III: Number of bands excluded from the Wannierization step, and the exponential decay length of the Hamiltonian, velocity matrix, dynamical matrix, and electron-phonon vertex in the limiting case of \(R = 0\) or \(R' = 0\). The decays have been determined from a least-squares fitting of the form \(\exp(-\hbar x)\) in a 10 Å range.

For GaAs we used 16 Wannier functions (8 Wannier functions times two due to spin-orbit coupling); on the Ga and As atoms, with \(sp^3\) character. These span both the valence and conduction manifolds, and are used to calculate hole mobilities. To reduce the computational cost for the electron mobility, we used only 8 Wannier functions centered on the Ga atom, with \(sp^3\) character, to describe the conduction manifold. Since we used semicon-
FIG. 2: The first two rows present the Wannier functions in real space for the conduction and valence bands, respectively; generated using the XCrySDen visualization program. In the case of conduction bands, we show the $\pm 1.0\,\text{Å}^{-3/2}$ isosurface everywhere, except in the case of SrO for which we show the isosurface at $\pm 2.0\,\text{Å}^{-3/2}$. In the case of valence bands, we show the isosurfaces at $\pm 0.2\,\text{Å}^{-3/2}$, $\pm 0.5\,\text{Å}^{-3/2}$, $\pm 1.0\,\text{Å}^{-3/2}$, and $\pm 1.0\,\text{Å}^{-3/2}$ for Si, diamond, 3C-SiC and GaAs, respectively; and the isosurface at $\pm 2.0\,\text{Å}^{-3/2}$ for all the other materials. By denoting the spinor components as $[\phi, \psi]$, in the figure we show $|\phi| \times \text{sign}(\Re\{\phi\})$. The next five rows show the spatial decays of the electronic Hamiltonian, velocity matrix, dynamical matrix and electron-phonon vertex in the Wannier representation, in the limiting case of $R = 0$ or $R' = 0$, as a function of $R = |R - R'|$. The data points correspond to the largest value taken over the Wannier functions in each of the two unit cells at a distance $R$, and are normalized such that the value is 1 at $R = 0$. The pink and purple colors indicate the Wannierization for the conduction and valence manifold, respectively. The size of the $k$- and $q$-points grids used for each manifold are also indicated in pink and purple colors, respectively.
the phonon part of the electron-phonon matrix elements we expect a power law decay. To characterize the decay using a simple unified descriptor, we fit our data with an exponential function for all cases. The resulting decay length is reported in Table III for all materials.

In the specific case of the valence band manifold of silicon, the Hamiltonian decay length is 1.710 Å, improving on earlier work which reported a decay length of 3.2 Å with an empirical pseudopotential starting from four bond-centered trial functions. The worst localization is found for the conduction band manifold of GaAs, GaP, and AlSb (> 2.7 Å), while the best localization is achieved for the valence manifold of c-BN, diamond, and SrO (~ 1 Å). We also note that the spread and number of iterations required to reach convergence systematically increase with k-point grid density. Furthermore, it has been reported that convergence fails for ultra-dense grids (80 × 80 × 80 and above).

We also tested two additional procedures for constructing Wannier functions, namely the selectively-localized Wannier functions (SLWFs) and the symmetry-adapted Wannier functions (SAWFs). With SLWFs one constrains the Wannier center \( r_n \) to be located at the position \( r_{on} \) using a Lagrange multiplier \( \lambda_c \) as \( \lambda_c \sum_n (r_n - r_{on})^2 \), where the summation can be restricted to selected Wannier functions. In this case we consider the entire manifold. The multiplier \( \lambda_c \) can be chosen arbitrarily; the larger the value, the more difficult the convergence, so that in practice the Wannier centers are close to but not exactly positioned at the target positions. The SAWFs guarantee that the Wannier functions respect the point group symmetry of the crystal. We note that the existing implementation in the wannier90 code does not currently support band disentanglement with a frozen window nor SOC. In both cases, we find that the minimum spread is larger than those reported in Table II. Therefore, we did not pursue these avenues further.

The MLWFs obtained as described in this section were used to calculate interpolated band structures, phonon dispersions, and electron-phonon matrix elements in the following sections.

### C. Dynamical quadrupoles

For practical calculations, we need a procedure for evaluating the quadrupole tensor introduced in Sec. IX. An efficient way to compute dynamical quadrupoles rests on perturbation theory and has been implemented in the Abinit software. However, this implementation is currently limited to the LDA exchange and correlation functional, and to pseudopotentials without non-linear core corrections. To circumvent these limitations, we propose an alternative strategy for estimating the quadrupole tensor, which can be used with any exchange and correlation functional and pseudopotential type. The general idea is to determine the quadrupole tensor by matching Eqs. (38)-(41) to explicit DFPT calculations at small \( q \).

| Compounds     | \( Q_{x_1} \) | \( Q_{x_2} \) | \( Q_{x_1} \) | \( Q_{x_2} \) |
|---------------|--------------|--------------|--------------|--------------|
| Diamond       | 3.2372       | -            | -            | -            |
| Si            | 11.8307      | 13.67        | -            | -            |
| GaAs          | 28.88        | -20.41       | 16.54        | -8.57        |
| 3C-SiC        | 7.41         | -2.63        | -            | -            |
| AlP           | 9.73         | -4.03        | -            | -            |
| GaP           | 13.72        | -6.92        | 12.73        | -5.79        |
| c-BN          | 4.50         | -0.82        | -            | -            |
| AlAs          | 12.25        | -6.18        | -            | -            |
| AlSb          | 14.61        | -8.70        | -            | -            |
| SrO           | 0.00         | 0.00         | -            | -            |

TABLE IV: Effective quadrupole tensor (e bohr) computed in this work. The complete tensors are given by \( Q_{\alpha \beta \gamma} = (\varepsilon_{\alpha \beta \gamma} + Q_{\alpha \beta} \varepsilon_{\gamma \alpha \beta}) - (\varepsilon_{\alpha \beta \gamma} + \kappa_{\alpha \beta} \kappa_{\gamma \alpha \beta}) \) for all other materials. We compare our values with previous calculations based on perturbation theory from Ref. 22.

The minimum number of calculations required is equal to the number of independent components of the tensor. In practice, to minimize numerical noise we compute the DFPT matrix elements for several points on a sphere with \( |q| = \) constant.

FIG. 3: Comparison between the deformation potential of the \( \Gamma_1 \) valence band of silicon from the direct DFPT (dots) and the Wannier interpolation with and without quadrupoles, for different k-point grids (the \( q \)-point grid is half of that). The components included in the dynamical matrix and in the electron-phonon matrix elements are labelled as “D” and “C”, respectively. The labels “DQ”, “Q”, and “QQ” indicate the dipole-dipole, dipole-quadrupole and quadrupole-quadrupole contributions, while “eD” and “eQ” denote the monopole-dipole and monopole-quadrupole interactions, respectively.
FIG. 4: Comparison between the deformation potential of the valence band manifold from the direct DFPT calculations (dots) and the Wannier interpolation including dipoles and quadrupoles (green lines). The k-point for the deformation potential is located at the valence band maximum.
To this aim, we use a Python script with the long-range analytic electron-phonon matrix elements from Eq. \(38\), and the phonon frequencies, Born effective charges, dielectric constant, eigendisplacements and wavefunction overlaps from the DFPT calculations. We then perform least squares optimization with respect to the components of the quadrupole tensor, and include a model-dependent offset as an additional minimization parameter to capture the \(\mathcal{O}(|q|^3)\) term of the matrix elements, following Eq. (3.17) of Ref. \[17\]. For the compounds considered in this work, we constrain the quadrupole tensor to have the form of Eq. \(12\) or Eq. \(43\) during the minimization, although the procedure holds for tensors of any symmetry.

We emphasize that this procedure for calculating the quadrupole tensor is inexpensive, as it requires the calculation of only a dozen \(q\)-points, see Fig. \[23\].

As a sanity check, we computed the quadrupoles of 3C-SiC using the Abinit software \[62,114\]. To use the Abinit implementation we considered LDA pseudopotentials without non-linear core correction and without SOC. We obtained \(Q_{3C-SiC} = 7.025 \; e\bohr\) and \(Q_{C} = -1.416 \; e\bohr\). Using the same pseudopotentials and the Quantum Espresso software \[55\], we obtained \(Q_{3C-SiC} = 7.392 \; e\bohr\) and \(Q_{C} = -2.483 \; e\bohr\). We attribute the small difference to the fact that the implementation of Ref. \[22\] does not include the second term of Eq. \(11\), and that our procedure does not capture the \(q \rightarrow 0\) limit with the same precision as in DFPT.

Using the above procedure, we computed the effective quadrupole tensor for all ten compounds. The results are reported in Table \[IV\]. Among the compounds considered here, GaAs has the quadrupole tensor with the largest elements, while this tensor vanishes in SrO due to the cubic \(Fm\overline{3}m\) space group symmetry. In Table \[IV\] we also report the calculated quadrupole terms from Refs. \[10,22\] for comparison, as obtained using perturbation theory, LDA without non-linear core correction, and neglecting SOC. Our calculations compare well with Refs. \[10,22\] although close agreement is not expected due to the aforementioned differences.

D. Interpolation of the electron-phonon matrix elements

To assess the quality of the Wannier interpolation, we compare the interpolated electron-phonon matrix element with those obtained from a direct DFPT calculation. For an easier comparison, we compute the total deformation potential \(D''(\Gamma, q)\),

\[
D''(\Gamma, q) = \frac{1}{\hbar N_w} \left[ 2\rho \Omega_\nu \omega_\nu \sum_{nm} |g_{m\nu}(\Gamma, q)|^2 \right]^{1/2},
\]

where the \(k = \Gamma\) point was chosen, the sum over bands is carried over the \(N_w\) states of the Wannier manifold, and \(\rho\) is mass density of the crystal. The advantage of examining the deformation potential rather than the matrix elements directly is that electronic degeneracies are traced out, and the frequency that enters the zero-point amplitude is removed, so that one can concentrate on the linear variation of the potential. To investigate specific states, we also define the band-resolved deformation potential \(D''_{mn}(\Gamma, q)\), where the band summation in Eq. \(51\) has been removed.

In Fig. \[5\] we compare the deformation potential of silicon for the lowest valence band state at \(\Gamma\), for the direct DFPT calculation (black dots) and the Wannier interpolation, with and without quadrupoles (lines).

In silicon there is no Fröhlich contribution to the matrix elements due to inversion symmetry. There are, however, quadrupole contributions. As can be seen in Fig. \[5\], the Wannier interpolation without quadrupole correction reproduces the direct DFPT calculation everywhere, except close to \(\Gamma\) in the \(IL\) and \(IK\) directions for the LO and TO modes. The discontinuity of the LO and TO modes at \(\Gamma\) poses a challenge to the interpolation without quadrupole correction, as shown by the solid lines in gray. Upon including quadrupole corrections, the deformation potential is found to be independent of grid size. We therefore only report the curve corresponding to the grid used throughout this manuscript (see Sec. \[III\]). In contrast, the longitudinal acoustic (LA) mode of silicon shown in Fig. \[5\] does not exhibit any discontinuities at the zone center, and is therefore well described without quadrupole correction. The deformation potential of the transverse-acoustic modes (TA) vanishes and is not shown. These data are consistent with Ref. \[15\].

Figure \[4\] shows a systematic comparison between the deformation potential computed via explicit DFPT calculations and the results of our Wannier interpolation including dipole, dipole-quadrupole and quadrupole corrections using Eq. \(51\). We note the sharp change in deformation potential occurring near the K point in diamond, and to a smaller extent in c-BN. This effect is due to the avoided crossing in the phonon bandstructure of diamond and c-BN at the same point in momentum space. Quantitatively, diamond and c-BN show the largest deformation potential, which is consistent with these compounds being the hardest materials in our set.

Overall, the agreement between the direct calculation and the Wannier interpolation with quadrupole correction is excellent, guaranteeing a good interpolation. The only exception is SrO which is the only material investigated here for which the quadrupole contribution vanishes by symmetry. This finding suggests that a very accurate interpolation for SrO might require the inclusion of octopole contributions. We leave the investigation of the effect of octopoles for further studies.
E. Convergence of the mobility with the coarse Brillouin zone grid

We present in Fig. 5 the convergence rate of the electron and hole carrier mobility with increasing coarse grid sizes. Overall, we find that using the same coarse \( k \) and \( q \) grid leads to the fastest convergence. On the other hand, using a coarse \( k \)-grid three times denser than the \( q \)-grid leads to similar results but at a much higher computational cost, as shown in Fig. 24 in the Appendix for c-BN. Therefore it is recommended that calculations be performed using the same \( k \) and \( q \) coarse grids. We note that the calculations presented below were performed using a \( k \)-grid twice as dense as the \( q \)-grid for historical reasons, but in future work equal grids will be employed.

To perform the convergence study, we use either a \( 60^3 \) or a \( 100^3 \) fine grids as indicated in Fig. 5. We verify that the cross convergence between coarse and fine grids is weak. As can be seen in Fig. 5, the interpolated mobility converges slowly with respect to the coarse grid size, despite dipole and quadrupole corrections being already included in the calculations. The origin of this slow convergence has not been investigated so far. To shed light on this effect, we compute the mobility effective mass defined in Eq. (8) and report the convergence of this purely electronic quantity with orange dots in Fig. 5. We note that the mobility effective mass reflects the quality of the Wannier interpolation of the band structures. Its value should converge to the mobility effective mass computed directly from DFT bands without interpolation. In Fig. 5, the reference DFT value is shown as horizontal dashed orange lines. Figure 6 shows that the mobility effective mass plays a significant role in the slow convergence of the mobility as a function of coarse grid size. However, the convergence of the mobility effective mass is a necessary but not sufficient condition for the overall convergence of the mobility with coarse grid size. Since computing the mobility effective mass is much cheaper than computing the mobility, this quantity can be used to estimate a minimum coarse grid size. In the cases of diamond, Si, 3C-SiC, AlP, GaP, AlAs, AlSb and SrO in Fig. 6, we see that the electron mobility effective mass does not behave consistently with the convergence of the mobility. It is also the case for the hole mobility effective mass of GaAs and AlSb. The reason for this deviation is that the effective mass does not fully capture the scattering phase space, since it neglects the energy conservation connected with the electron-phonon scattering amplitudes.

To incorporate these effects, we employ the CARTA approximation defined in Eq. (8). In this case, we have to compute the sum over \( q \)-points explicitly, making it more computationally involved, but still far cheaper than the SERTA since we do not have to interpolate the dynamical matrix and electron-phonon matrix elements. For the cases considered here the CARTA calculations are at least two orders of magnitude faster than the SERTA.
FIG. 6: Convergence of the CARTA mobility as a function of the coarse k-point grid size (magenta dots). We only show the relative convergence since the CARTA mobility contains the arbitrary parameter $g_\alpha^2$, see Eq. (5). The coarse q-point grid and the fine k and q-point grids are given in the inset. The vertical gray line indicates that convergence has been achieved, and represents the coarse grid that is used in the remainder of this work. The orange dots are the mobility effective masses, and are referred to the vertical axis on the right. The horizontal orange and magenta dashed lines are the DFT reference values.

Gaussian broadening for the Dirac deltas. The CARTA results are shown as magenta dots in Fig. 6 and show how the electronic degrees of freedom influence the convergence with respect to the coarse Brillouin zone sampling. Since these calculations do not carry information about the electron-phonon matrix element, we present the data as a deviation from the CARTA mobility computed without Wannier interpolation, by using directly DFT eigenvalues.

By comparing Figs. 5 and 6 we observe that the CARTA mobility converges at a similar rate as the BTE mobility. This observation suggests to speed up the coarse grid convergence by factoring out the electronic degrees of freedom. To this aim, we rescale our BTE mobilities $\mu_{\alpha\beta}$ using the following expression:

$$\mu_{\alpha\beta}^{\text{scaled}} = \frac{\mu_{\alpha\beta}^{\text{CARTA,DFT}}}{\mu_{\alpha\beta}^{\text{CARTA}}}.$$  \hspace{1cm} (52)

where $\mu_{\alpha\beta}^{\text{CARTA,DFT}}$ is the mobility obtained in the CARTA with DFT eigenvalues and velocities. Eq. (52) allows one to correct for the small off-set observed in some materials for the effective mass mobility. Moreover, computing $\mu_{\alpha\beta}^{\text{CARTA}}$ is over two orders of magnitude faster than $\mu_{\alpha\beta}$ since we only need to interpolate the eigenvalues and velocities. The results for the scaled BTE mobility are shown as orange dots in Fig. 6. Importantly, not only $\mu_{\alpha\beta}^{\text{scaled}}$ converges faster, but it also does so in a smoother and more systematic way.

The vertical thin gray line in Fig. 5 is the coarse grid that we consider converged and that is used in all subsequent calculations, unless stated otherwise. For the electron mobility of GaAs we do not report the mobility effective mass since we used the experimental value to describe the bands and velocities. From Fig. 5 we can note that the hole mobility converges in a smooth and systematic way while this is not necessarily the case for the electron mobility. This behavior reflects the better real-space localization of Wannier functions obtained for the valence manifold than the conduction manifold, see Fig. 2.

For the case of silicon, the electron mobility is significantly improved upon using the scaling of Eq. (52). However, we can notice that for our chosen converged grid, the calculated mobility does not converge to the correct limit without scaling (see Fig. 6). Specifically, the interpolated CARTA result overestimates the CARTA/DFT result by 1%. We expect a similar error to occur for the BTE mobility. We anticipate that the slight difference will disappear upon using denser grids or by using a more extended Wannier functions basis set, but we have not investigated this further due to the computational cost. An overestimation of 2% is also observed for the hole mobility of AlSb. In all other cases, the interpolated CARTA values converge smoothly to the CARTA/DFT result.

We are now in a position to answer the question on why is the convergence of the carrier mobility slow with re-
spect to coarse grid size even when dipole and quadrupole corrections are included. The answer is material dependent. In the case of the electron mobility of Si, 3C-SiC, AlP, GaP, AlAs, AlSb, and SrO, the slow convergence is mostly due to purely electronic effects as demonstrated by the CARTA results. Specifically, the Wannier-interpolated electron bands converge slowly with grid size. A possible way to accelerate convergence would be to increase the number of Wannier functions in order to improve the flexibility of the basis functions. In general, for the valence bands the convergence is smoother and more systematic, and, therefore, less can be gained by using the CARTA rescaling. Nevertheless, we noticed some improvements for diamond, AlP, GaP, c-BN, AlAs, and AlSb.

In the case of diamond and c-BN, very little improvement is noticed using the CARTA rescaling. This indicates that for these two materials the slow convergence can be attributed to a slow convergence of the phonon dispersions and electron-phonon matrix elements with grid size.

In Fig. [7] we compare our coarse grid convergence results with previous work for Si and GaP in Ref. 16. Since only the SERTA mobility is reported in Ref. 16, we compare with our results at the same approximation level. Overall our results agree quite well given that there are a number of differences in the calculations. The coarse k-point grid in Ref. 16 was fixed to 18\(^3\), while we used a coarse k-point grid that is twice the one reported in this figure for the q-point grid. We used an equal fine k- and q-point grids of 150\(^3\) for Si and 120\(^3\) for GaP, while Ref. 16 used a fine 72\(^3\) k-point grid and a 144\(^3\) q-point grid for Si and a 78\(^3\) k-point grid and a 156\(^3\) q-point grid for GaP. Different pseudopotentials and exchange correlation functionals were used as well. In addition, we included SOC, while Ref. 16 neglects it, but as shown in Fig. 7 for the electron mobility is negligible. The main difference between our results and those of Ref. 16 is that our calculations including dipole but not quadrupole corrections seem to converge rapidly, while in the earlier work there is a slow drift of in the SERTA mobility calculated with dipole correction. However, in agreement with Ref. 16, we observe that the inclusion of quadrupole corrections leads to a shift of the calculated mobility even at convergence.

F. Convergence of the mobility with the fine Brillouin zone grid

Using the converged coarse grids determined in the previous section, we can proceed to testing the convergence of the mobility with respect to the fine grid. The presence of the \(\partial f_{nk}/\partial \varepsilon_{nk}\) term in the mobility, see for example Eq. (5), implies that only those states close to the valence band or conduction band edge will contribute. To reduce the computational cost of these calculations we use the following scheme. We construct a fine k-point grid composed of only the points within a small energy window around the band edges, and rely on crystal symmetry to further reduce the number of points. For example in the case of the hole mobility of c-BN, using an energy window of 0.3 eV and an homogeneous fine grid of 250\(^3\) points, we only have 12,390 irreducible k-points. Importantly, at no point we store the entire uniform grid in memory. This reduction allowed us to use extremely dense grids for the electron mobility of GaAs, up to 800\(^3\) points in the first Brillouin zone.

Since all q-points can contribute (for example through inter-valley scattering), we do not use symmetry reduction, and keep all the q-points of the uniform grid that lie with the small energy window. For example, in the case of c-BN and a grid of 250\(^3\) points, we have to compute 814,981 q-points. To this aim we generate the q-points on the fly, so that we do not have to store the full grid in memory.

Finally, we interpolate the electron-phonon matrix elements for all these k- and q-points and store on disk only the k- and q-points identified above that contribute
FIG. 8: Convergence of the SERTA and BTE mobility and Hall factor as a function of the fine $k$- and $q$-point grid size, using the converged coarse grids from Fig. [5]. Both quantities converge linearly with the fine grid density, and can therefore be extrapolated to infinity. The shaded area denotes the interpolation range for the linear least squares fit of the Hall factor and mobility.

to the mobility. To retain only the information that will contribute to the mobility calculation, for each $q$-point we store the matrix elements for the $nk$ states that fulfill the following condition:

$$\hbar \frac{\partial F^0}{\partial n_k} \tau^{-1}_{nk}(q) > \frac{10^{-16}}{N_k N_q N_h},$$

(53)

where $N_k$, $N_q$, and $N_h$ are the total number of $k$-points, $q$-points, and Wannierized bands, respectively. The threshold in Eq. (53) guarantees that the cumulative absolute error made by neglecting very small matrix elements will be lower than machine precision, or $10^{-16}$. In the c-BN example mentioned above, this procedure required the storage of a binary file of approximately 40 GB instead of multiple TBs required to store all the matrix elements. Once all the scattering rates are written to disk, we read the file in parallel using the message passing interface input-output (MPI-IO) and solve the iterative BTE. The first step of the iterative solution yields the SERTA mobility. The iterative solution of the BTE is extremely fast (a few minutes) and converges typically within a few tens of iterations. For the compounds considered in this work, we determined that the optimal energy window is between 0.2 and 0.3 eV, as shown in Table VII of the Appendix.

The results for the SERTA and BTE mobility and Hall factor are shown for all materials in Fig. [8]. For similar reasons as in Ref. [17] we stress that we are showing the results as a function of inverse grid density and the fine $k$-point and $q$-point grids are always the same. For example, the label 1/100 on the horizontal axis corresponds to a $100^3$ fine grid, such that the left hand side of each figure corresponds to infinitely dense sampling.

Interestingly, both the mobility and the Hall factor appear to converge linearly with inverse grid density and can therefore be extrapolated. The linear interpolation range is depicted in Fig. [8] with a gray shaded region. We perform the linear extrapolation as soon as we enter the linear regime. This linearity comes from the fact that we compute the momentum integrals using the basic rectangular integration, therefore the error scales linearly with the grid spacing.

We notice that the electron and hole mobility of diamond, Si, c-BN as well as the electron mobility of GaAs and 3C-SiC show a relatively steep slope, therefore it is important to extrapolate the values for accurate results. In contrast, all the other cases considered here exhibit a milder slope, therefore the extrapolation is not necessary in these cases.

The linear behavior of the mobility is closely linked to the smearing used for the energy conserving Dirac deltas in Eq. (8). In Fig. [9] we focus on the electron mobility
of c-BN since the linear trend is very pronounced, and we show how the carrier mobility depends on the grid density for various Gaussian smearing parameters. We see that, the smaller the smearing, the more pronounced are the fluctuations, and finer grids are needed before the linear regime is reached. The extrapolated value depends on the smearing used.

The above discussion does not take into account the finite lifetimes of carriers in the calculations. In reality, a finite smearing in the range of 10-50 meV is to be expected as a result of the electronic line-widths. This effect could be incorporated by setting the smearing to $\hbar/\tau_{\alpha k}$ with $\tau_{\alpha k}$ given by Eq. (4), but we have not explored this direction.

Since the linear extrapolation at fixed smearing is rather tedious, in practice we proceed by using the adaptive smearing introduced in Eq. (48). As seen in Fig. 9, this approach offers a very good approximation to the zero-smearing extrapolation. All data presented in Fig. 8 have been generated using adaptive smearing. We also show a similar analysis for the case of the electron mobility of GaAs in Fig. 25 of the Appendix.

In Fig. 8 we also present the convergence of the SERTA and BTE Hall factor defined in Eq. (13). In this case, the Hall factor systematically converges linearly with grid density and can be extrapolated to infinite grid density. Also in this case we can either extrapolate the results to zero smearing, or use adaptive smearing. The results in either case are extremely close, therefore we proceed with adaptive smearing in the following.

IV. RESULTS

In this section we discuss the main results of our calculations for all the compounds considered in our work. We first describe the electronic and vibrational properties, then we discuss the Hall factor and its temperature dependence, as well as drift and Hall carrier mobilities. We also analyze in details the microscopic mechanisms responsible for limiting the carrier mobilities. Lastly, we compare our results to available experimental data.

A. Wannier interpolation of electron bands and phonon dispersions

The interpolated electronic band structures and phonon dispersions for the ten compounds considered in this work are presented in Fig. 10. The phonon dispersions are compared to neutron scattering data. Close agreement between theory and experiment is found in all cases except for SrO, where the theory underestimates measured frequencies. The associated Born effective charges, high-frequency dielectric constants and maximum phonon frequencies are reported in Table V. The average deviation for the highest phonon frequency between calculations and experiments is 1.6%.

Since DFT/PBE systematically underestimates the electronic bandgap, the dielectric constant is systematically overestimated with respect to experiment. For the systems considered here, the average overestimation of the dielectric constant is 10%, but the largest discrepancy is as large as 30%.

As expected, the band gaps are severely underestimated. The average deviation from experiments is of 41%. On the other hand, the spin-orbit splitting is in slightly better agreement with experiment, the average deviation being 18%. Given these data, we expect that the electron-phonon matrix elements will be over-screened, and we anticipate that our calculated mobilities will tend to overestimate experimental data. We will come back to this point in Sec. IV E.

The DFT mobility effective masses obtained without Wannier interpolation are reported in Table V. The val-
values range from 0.1 to 0.8 $m_e$. Direct comparison of such mobility effective mass with experiment is not possible, however we can compare the effective masses calculated using the standard definition with experiments. We report all the experimental and theoretical electron and hole effective masses in Table V. Whenever the experimental values were reported along the principal axis, we transform them to the density of state effective mass using for example $m_{hh}^* = \left[ (m_{110}^{hh})^2 m_{100}^{hh} \right]^{\frac{1}{3}}$, and similarly for the light-hole effective masses.

For diamond, the experimental longitudinal and transverse electron effective masses are 1.4 and 0.36 $m_e$, respectively. For the hole, we used the values from Willatzen and Cardona, which are obtained from linear muffin-tin orbital calculations corrected with measured energies, and are considered most reliable (experimental values obtained from the field dependence of the hole mobility show significant dispersion, with 1.10, 0.32, and 1.06 for the heavy-hole, light-hole and spin-orbit band, respectively).

For silicon, the experimental longitudinal and transverse electron effective masses are 0.92 and 0.19 $m_e$. The gray dots are experimental neutron scattering data for diamond, silicon, GaAs, 3C-SiC, AlP, GaP, c-BN, AlAs, AlSb, and SrO.
respectively, while the light and heavy hole masses are 0.16 and 0.5 \( m_{\text{hh}} \), respectively.

For GaAs, the electron effective mass is isotropic, with a recommended effective mass value of 0.067 \( m_{\text{ee}} \). The measured spin-orbit, light hole, and heavy hole effective masses of GaAs are 0.15, 0.082, and 0.51 \( m_{\text{ee}} \), respectively.

For 3C-SiC, the measured longitudinal and transverse electron masses are 0.68 and 0.25 \( m_{\text{ee}} \), respectively; for the holes we could find the density of states effective mass, 0.45 \( m_{\text{eh}} \). For AIP, we note a \( \mathbf{k} \cdot \mathbf{p} \) study in relatively good agreement with our results which reports 2.68 \( m_{\text{eh}} \) and 0.16 \( m_{\text{eh}} \) for the transverse and perpendicular electron effective masses, respectively. For the hole, the same study obtains 0.73, 0.19, and 0.3 \( m_{\text{eh}} \) for the heavy-hole, light-hole, and spin-orbit hole, respectively.

In the case of GaP, the low-temperature electron transverse and longitudinal effective masses were measured to be 0.252 and 0.87 \( m_{\text{ee}} \), respectively. The heavy and light hole effective masses were obtained by cyclotron resonance at 16 K, and are 0.54 and 0.16 \( m_{\text{eh}} \), respectively.

In the case of c-BN, the longitudinal and transverse electron effective masses are 1.2 and 0.26 \( m_{\text{ee}} \); the heavy hole masses were reported in the range 0.38-0.96 \( m_{\text{eh}} \), and the light-hole masses were reported in the range 0.11-0.15 \( m_{\text{eh}} \). In the case of AlAs, a compilation of experimental data from Ref. 161 shows that the electron transverse effective mass ranges between 0.124 (from optical absorption) and 0.19 \( m_{\text{ee}} \) (from quantum-well spectroscopic measurements).

For AlSb, the room temperature transverse and longitudinal electron effective masses were measured to be 0.26 and 1.0 \( m_{\text{ee}} \), using indirect exciton absorption measurements. For SrO we could not find reliable measurements of the effective masses.

Overall, the agreement between DFT effective masses and experiment is reasonable (within a factor 2-3) but there are a few large discrepancies. For the electron effective masses, the transverse mass of GaP and AlSb are overestimated with respect to the experimental values by 63 % and 82 %, respectively. In the case of silicon, the computed electron effective masses are very close to the experimental values. We note that our results are in line with previous calculations employing perturbation theory. Since DFT strongly underestimates the effective mass of GaAs, with values ranging from 0.03 \( m_{\text{ee}} \) to 0.053 \( m_{\text{ee}} \), we decided to use the experimental effective mass for the calculation of the electron mobility. To this aim we employed a parabolic band approximation to evaluate analytically the eigenenergies and electron velocities from the experimental effective mass. This semi-empirical procedure was used only for the electron mobility of GaAs.

In the case of hole effective masses, the largest discrepancies are found in the cases of c-BN (the mass is underestimated by almost a factor of three), GaAs, Si, and AlAs (all underestimate experimental values by 40-60%). The impact of these discrepancies on the calculated mobilities will be analyzed in Sec. [IV.E]
B. Impact of various approximations on the drift mobility

Having introduced all the required concepts and results, we are now able to discuss Fig. 1. For each of the approximations tested, we only change a single parameter, keeping everything else the same as in our most accurate calculations. The reference calculations have been performed using converged coarse and fine grids, velocities including non-local contributions, dipole and quadrupole corrections, SOC, BTE, adaptive broadening, and experimental lattice parameters. The following considerations also apply to the Hall factor.

We discuss the impact of the various approximations in order of importance. We do not discuss the constant relaxation time approximation (CRTA) which depends on the empirical choice of the scattering rates and whose predictive power is unclear.

The most severe approximation concerns the long-range treatment of the dynamical matrix and the electron-phonon matrix element. In Fig. 1 we break down the long-range treatment into four levels: (i) no long-range treatment; (ii) including Fröhlich dipoles; (iii) including dipole and quadrupole corrections in the dynamical matrix; and (iv) including dipole and quadrupole corrections in the electron-phonon matrix elements.

Completely neglecting the long-range behavior has been recognized in 2015 to be a severe limitation but its impact on the mobility has not yet been quantified. The consequence of neglecting long-range contributions during Fourier transformation of the matrix elements is that the Fröhlich divergence when approaching the zone centre is spuriously suppressed. As a result, the overall electron-phonon coupling, and therefore the scattering rates, will be strongly reduced. This leads to a strong overestimation of the carrier mobility, up to 220% in the case of the electron mobility of SrO. Surprisingly, the hole mobility of 3C-SiC is only slightly underestimated (8%) when ignoring long-range couplings. This error cancellation is due to incorrect eigendisplacement vectors and will be discussed shortly.

Next in order of significance, with a standard deviation of 30%, is the neglect of dynamical quadrupoles during the interpolation of the dynamical matrix and the electron-phonon matrix element. Neglecting quadrupoles typically leads to an underestimation of the mobility by an average of -5.6%, and up to -45% in the case of 3C-SiC. A notable exception is the case of c-BN, for which the neglect of quadrupole corrections leads to an overestimation of the mobility by 71%. This important effect has been discovered recently, and to our knowledge it has been taken into account only in a handful of works.

In the case of silicon, neglecting quadrupoles for the LO and TO modes in the $\Gamma - L$ and $\Gamma - K$ directions shown in Fig. 3 produces spurious oscillations, leading to a 4%/5% overestimation of the electron/hole mobility.

In the case of diamond, the neglect of quadrupole corrections yields an overestimation of less than 1% of the mobilities. Data for all compounds are collected in Fig. 1, and their magnitude is in line with prior work.

The effect of quadrupole correction can be substantial in some compounds. For example, as seen in Fig. 1, the hole mobility of c-BN is strongly overestimated. To better understand why this is the case, in Fig. 11(a) we compare the deformation potential obtained by including or neglecting the quadrupole term. We can see that the high-energy modes are unaffected by quadrupoles in this case. Conversely, the quadrupole correction becomes important for low-energy acoustic modes. The red lines in Fig. 11(a) shows the result of the interpolation performed by including only the dipole long-range correction. This result clearly deviates from the direct DFPT calculations,
shown as the black dots close to the zone center. In particular, the deformation potential corresponding to the LA mode in the \( L - \Gamma \) and \( K - \Gamma \) directions is strongly underestimated, leading to the observed large overestimation of the hole mobility of c-BN (in Fig. [1]).

Interestingly, we found that including quadrupole corrections to the electron-phonon matrix elements only was not sufficient to correctly describe the deformation potential, see the orange lines in Fig. [11]. Indeed, also in this case the deformation potential of the LA mode is overestimated, albeit the overestimation is less significant, and the hole mobility of c-BN is underestimated accordingly (23%, see fourth column of Fig. [1]). This result is counter-intuitive since the phonon frequencies are extremely well described by including dipole-only terms in the dynamical matrix, see Fig. [20] for example (in certain classes of materials such as piezoelectrics this is often not the case\(^3,13\)). A careful investigation shows that the root cause which explains the difference between the orange and green curves in Fig. [11] are the eigendisplacement vectors obtained by diagonalization of the dynamical matrix. The eigendisplacement vectors \( \epsilon_{\kappa \alpha q \nu} \) enter in Eq. (23) during the Fourier transformation of the electron-phonon matrix elements and within the long-range term in Eq. (40). It is therefore crucial to include at least dipole-dipole, dipole-quadrupole and quadrupole-quadrupole in the dynamical matrix, see Eq. (40) as well as dipole-dipole and quadrupole-quadrupole effects in the electron-phonon matrix elements, see Eq. (40) to achieve a reliable interpolation of the deformation potential. We also emphasize that this correction must be applied both to the dynamical matrix and the scattering matrix elements: a partial correction of the dynamical matrix only is not sufficient, as shown by the purple line in Fig. [11] and the third column of Fig. [1].

Another interesting situation is found for 3C-SiC. We show the deformation potential for the lowest energy band in Fig. [11(b)]. Again we can see that the deformation potential close to \( \Gamma \) along the \( K - \Gamma \) direction has a spurious finite value at the zone center. Including quadrupoles in the dynamical matrix and electron-phonon matrix elements fixes the problem. Among all materials, we note that including quadrupole correction in the electron-phonon matrix element but not in the dynamical matrix, yields an average underestimation of the mobility by 5%.

Beyond long-range dipole and quadrupole corrections, the most significant source of error in mobility calculations is the local velocity approximation, whereby the non-local part of the pseudopotential is neglected. This approximation yields a large standard deviation of 28% with values ranging from -71% in c-BN to +37% for 3C-SiC. We note that the local approximation was used by some of us for the calculation of the mobility of silicon\(^{14}\).

We are now able to state that this approximation yields an overestimation of the mobility by +17% and +7% for the electron and hole mobility of silicon, respectively. It is worth mentioning that the velocity including non-local contributions was implemented in the EPW software shortly after the publication of Ref. [14] so that subsequent work is not impacted by this approximation.

The next source of error in order of importance is the neglect of SOC. Without SOC the mobility is underestimated by up to 62%. For the compounds investigated here, this effect is almost negligible for electron mobility, while it is significant for hole mobility. The strongest underestimation occurs in AlSb with a -62% decrease in mobility. These results are in line with earlier findings\(^{13,19}\). The reason for this underestimation is that SOC only affects valence bands in the semiconductors investigated here, because the valence band maximum is predominantly of \( p \) character. The effect of SOC is to lower the spin-split band, thereby reducing the phase space available for scattering, which increases the mobility. For systems with \( p \) states at the conduction band bottom we expect a similar effect, for example in halide perovskites\(^{16}\).

Less significant than the above but still important is the error that one makes by using the popular SERTA method. This approximation systematically underestimates the mobility within the range 1% to 40%. At an intuitive level, the fact that the SERTA mobility underestimates the BTE mobility can be understood as follows: in the SERTA all scattering processes reduce the electron current in the same way, while in the BTE one takes into account the fact that forward scattering does not reduce the current as strongly as backward scattering. Therefore the BTE is expected to yield a higher mobility.

The last three effects shown in Fig. [1] have a smaller impact on mobility, with errors up to 20%. We go over these effects in order of appearance in the figure. First, we investigated the sensitivity of the mobility to the lattice parameter. To this aim we re-calculated all mobilities using the relaxed lattice parameters. As discussed in Sec. IIIA, the PBE lattice parameter is overestimated with respect to experiment by an average of +1.3%. This effect yields an average decrease of mobility by -6.5%. We then considered the effect of the exchange-correlation functional by using LDA instead of PBE but keeping all other parameters the same. In this case we find only small changes, with an average of -1.5%. Lastly, we explored the effect of pseudization parameters. To this aim we repeated the calculations using the SG15 ONCV library\(^{23}\) that was extended to fully relativistic potential\(^{25}\) instead of the PseudoDojo library\(^{29}\). In this case we find an average deviation of the order of 2%. The small impact of these effects on the calculated mobilities highlights the robustness of our approach.

We note that many other effects can impact the calculation of carrier mobilities, including electron-two-phonon interactions\(^{20}\), thermal lattice expansion\(^{14}\), and band structure renormalization due to electron-phonon coupling\(^{17\text{,}16\text{,}29}\). We expect these effects to be small in most materials, but it will be important in the future to perform a detailed analysis. The effect of quasiparticle corrections\(^{14\text{,}23}\) and many-body renormalization of the
FIG. 12: Mode decomposition and total scattering rates at room temperature as a function of energy from the band edges. The energy range covers the energy window used in the calculations.

FIG. 13: Spectral decomposition of the electron (orange) and hole (green) scattering rates as a function of phonon energy at 300 K. The rates are calculated as angular averages for carriers at an energy of $3k_BT/2 = 39$ meV away from the band edges. The dashed lines represent the cumulative integrals of the calculated rates, and add up to the carrier scattering rate $\tau_{sc}^{-1}$. The percentage indicates the contributions from acoustic modes.

electron-phonon matrix elements \cite{166,167} will be discussed in Sec. IV E in relation with experimental data.

C. Dominant scattering mechanisms

To gain a better understanding of the microscopic phenomena responsible for limiting the intrinsic transport properties of the compounds investigated here, we show in the Appendix Fig. 20 the electron and hole scattering rates $\tau_{nk}$ computed via Eq. (1) at 150 K, 300 K, and 500 K. Without surprises, the higher the temperature, the higher the scattering rates and the lower the mobility.

As expected, in most cases we find a mild increase of the scattering rates as we move away from the band edges, which corresponds to acoustic-phonon scattering. At the onset for the emission or absorption of optical phonons, the rates show a marked increase. For example in diamond this sharp increase occurs at around 160 meV, when optical phonon emission/absorption processes become allowed. We note a couple of exceptions to this general trend: the scattering rates of electrons in GaAs and SrO, and both the electron and hole scattering rates...
of c-BN first decrease when moving away from the band edges which is attributed to LO phonon absorption and was investigated in Ref. [19].

In Fig. [12] we present the mode decomposition of the room-temperature scattering rates, as well as the total scattering rates. Starting from the holes: For diamond, the onset of scattering is mostly related to transverse acoustic mode scattering, while for silicon the longitudinal acoustic mode also plays an important role. In the case of GaAs, AlP, GaP, AlAs, AlSb, and SrO, the hole scattering near the band edges is dominated by the longitudinal optical mode. For the electrons: The scattering of electrons near the band edge in GaAs, AlP, GaP, AlAs, AlSb, and SrO is dominated by longitudinal optical modes, while diamond, silicon, 3C-SiC and c-BN show significant contributions from LA and TA modes.

In 1966, Berman, Lax and Loudon have derived intervalley-scattering selection rules for III-V semiconductors. For example, in the case of electrons in GaP, X-X intervalley scattering is expected. The rule is that, if the mass of the group-V element is heavier than that of the group-III element, then LO scattering should be allowed by selection rules, otherwise the LA scattering would be allowed. Therefore, for GaP one would expect LA-phonon scattering, while for AlP, AlSb, and AlAs one would expect LO-phonon scattering to dominate the intervalley scattering. Looking at Fig. [12] confirms the dominance of LO scattering in AlP, AlSb, and AlAs. For the case of GaP, it seems the scattering at an energy $3/2k_B T$ away from the band edge has about equal contribution from LA and LO phonon modes.

A complementary viewpoint on the scattering rates is obtained by considering only carriers at a given energy, and resolving the rates by phonon energy. This analysis is shown in Fig. [13]. The calculations correspond to carriers at an energy of $3k_B T/2$ away from the band edge and room temperature, following previous work [19]. From this analysis, we can see that the scattering rate is overwhelmingly dominated by acoustic scattering in diamond and c-BN. In contrast, optical scattering dominates for electrons in GaAs, SrO, AlAs, AlP, and for holes in GaAs, AlAs, AlSb, GaP, and AlP.

We also mention a few counter examples to the common wisdom whereby if a material is dominated by deformation potential acoustic scattering, then the temperature exponent of the mobility is $T^{-3/2}$ [19]. Diamond, silicon, c-BN and 3C-SiC are all dominated by acoustic-phonon scattering, but the temperature exponents evaluated on the 150 K-500 K range are -1.81, -2.05, -1.33, and -2.86 for the electrons and -2.06, 2.93, -1.51, and -1.97 for the holes, respectively. These data suggest that one should exert some caution when applying simplified models based on the electron gas to real materials.

For completeness we also examine the predictive power of the classic Drude formula [Eq. (7)]. To this aim we compare our BTE calculations for all the compounds considered in this work with the Drude prediction. In the Drude formula we employ the mobility effective masses reported in Table V and the total scattering rates evaluated as in Fig. [13]. The resulting data are shown in Fig. [14]. The Drude formula reproduces the trend of ab initio BTE calculations very closely, although it tends to overestimate mobilities by a factor of 1.1 to 3.1. This favourable comparison implies that the materials descriptors that we introduced, namely the mobility effective mass and the scattering rates of carriers at $3k_B T/2$, provide a meaningful description of carrier transport in the systems considered here.

D. Temperature dependence of the Hall factor

In this section we investigate the temperature dependence of the Hall factor. We use Eqs. (10)-(13) to compute the Hall factors within both the SERTA and full BTE. For each temperature we extrapolate the fine grids as discussed in Sec. IIIF and we present the results in Fig. [15]. We find that the Hall factor can vary substantially with temperature, ranging from 0.7 to 1.9 at room temperature. These results suggest that caution should be used when comparing ab initio calculations of carrier mobilities to experiments, since most reported experimental data refer to Hall mobilities, not drift mobilities. As a technical note, we mention that below 150 K we could not converge the Brillouin zone grids due to excessive computational cost. Accordingly these data are not shown in Fig. [15].

Our calculations indicate that the Hall factor exhibits rather complex temperature trends, which reflects the variety of phonon energy scales involved.
Comparison with experimental data for the Hall factor is challenging given the scarcity of available data. We were only able to find experimental values for the hole Hall factor of diamond and for the electron Hall factor of GaAs, which we report in Fig. 15. In both cases we find reasonable agreement between our calculations and experiments, with the largest deviation being of the order of 30% (note the magnified scales in Fig. 15).

In a recent preprint, the authors calculated the electron Hall factor of silicon to be 1.15 at room temperature. These results are consistent with our Fig. 8 for the same fine grid where we have an electron Hall factor of 1.12. The authors of Ref. 51 also computed the electron Hall factor of GaAs, finding a BTE value ranging from 1.22 to 1.14 in the 250-400 K temperature range. Their results compare well with our calculations, yielding a Hall factor of 1.28 to 1.13 in the same temperature range.

In Fig. 15 we also show the Hall factor as computed using the standard isotropic approximation given by Eqs. (15)–(16) (gray dashed lines). For these calculations we evaluated the Dirac deltas using a Gaussian of width 1 meV. This approximation reproduces the BTE results quite well for diamond, GaAs, AlP(h), c-BN(e), AlAs, AlSb, and SrO. However, it fails even qualitatively in the case of silicon, AlP(h), and GaP(e). One important limitation of the isotropic approximation is that it cannot lead to a Hall factor smaller than unity.

E. Temperature dependence of the drift and Hall mobility and comparison to experiment

In this section we examine the accuracy of the various methods presented here in reproducing experimental carrier mobility. In Fig. 16 we show the dependence of the BTE drift and Hall mobility on temperature, and compare our calculations to experimental data whenever available. The calculations are presented in a log-log scale. We indicate with dashed lines the linear fits used to determine the temperature exponents in Table VIII of the Appendix. Our results span a large range of power law decays, from \( T^{−1.18} \) for c-BN to \( T^{−3.15} \) for AlP. In the temperature range considered here (150-500 K), we find that a single exponent captures fairly well the calculated trends for all materials. Some deviations are found to occur above 400 K in the case of GaAs, AlP, c-BN, AlAs, and SrO. These deviations are indicative of high-frequency optical phonons becoming increasingly important in carrier scattering at these temperatures. We note that this fitting procedure is done in order to extract a single exponent which can be compared to experimental data fitted with the same procedure and in the same temperature range, whenever available.

In Fig. 17(a) we compare the calculated and measured temperature exponents. We find that in general the measured exponents are well reproduced by the calculations.
with an error of less than 20%. The only exceptions are found for SiC and AlSb, for which very few experimental studies exist. Given the overall good agreement found for the other compounds, we expect that the agreement between theory and experiment for SiC and AlSb will improve as new experimental data will be produced in the future.

In Fig. 17(b) we show the comparison between calculated and measured mobilities, in a logarithmic scale. In all cases the experimental data lay below the predicted phonon-limited Hall mobility. The difference can be as severe as 100% in some cases. The calculated drift mobility is found to be closer to the experimental data, but we have to emphasize that most experimental data reported in this figure are for the Hall mobility, as indicated in Table VII. The SERTA drift mobility appears to be closer to experiments, since experimental data incorporate additional scattering mechanisms.

F. Corrections to band structures and matrix elements from experimental data

In this section we investigate the role of many-body renormalization of the band structures and electron-phonon matrix elements on the mobility. Since it is not currently possible to perform complete many-body calculations of the BTE, we estimate these effects by using the experimental effective masses and by rescaling the electron-phonon matrix elements via the experimental dielectric permittivities.

To this aim, starting from the experimental masses reported in Table V for the electrons we determine the conductivity effective mass defined as the harmonic mean of the transverse and longitudinal effective masses:

$$\frac{3}{m_{\text{cond}}^*} = \frac{1}{m_{\parallel}^*} + \frac{2}{m_{\perp}^*}. \quad (54)$$

For the holes we average the band contribution with the following weighted average formula which assumes band
FIG. 17: Comparison between theoretical and experimental (a) temperature exponents and (b) mobilities. The values are reported in Table VI. The mean absolute error (MAE) and mean absolute relative error (MARE) are indicated as inset in the bottom figure and do not include the electron mobility of GaAs.

parabolicity:

\[ m_{\text{cond}}^{*} = \frac{(m_{\text{hh}}^{*})^{5/2} + (m_{\text{lh}}^{*})^{5/2}}{(m_{\text{hh}}^{*})^{3/2} + (m_{\text{lh}}^{*})^{3/2}}. \]  

Using these experimentally-derived quantities, we rescale our calculated electron and hole Hall mobilities using:

\[ \mu_{\text{mass}} = \mu \left( \frac{m^{*}_{\text{DFT}}}{m^{*}_{\text{exp}}} \right). \]  

TABLE VI: Effect of rescaling the Hall mobilities obtained via the BTE using experimental effective masses and dielectric constants.

| Material   | C-e | C-h | Si-e | Si-h | GaAs-e | GaAs-h | SiC-e | SiC-h | AlP-e | AlP-h | GaP-e | GaP-h | BN-e | BN-h | AlAs-e | AlAs-h | A1Sb-e | A1Sb-h | SrO-e | SrO-h |
|------------|-----|-----|------|------|--------|--------|-------|-------|-------|-------|-------|-------|------|------|--------|--------|--------|--------|-------|-------|
|          | 1705 | 2467 | 1770 | 574  | 17860  | 1068   | 2257  | 111   | 698   | 92    | 442   | 353   | 956  | 281   | 640    | 357    | 431    | 1000   | 65    | 5     |
|          | -16% | -28% | +2%  | -47% | -42%   | -32%   | +3%   | +17%  | -14%  | -14%  | +56%  | -32%  | +11% | +57%  | +20%   | -45%   | +20%   | -38%  | -14%  | -14%  |
|          | -9%  | -9%  | -14% | -14% | -42%   | -42%   | -11%  | -11%  | -14%  | -14%  | -22%  | -22%  | -3%   | -3%   | -10%  | -10%   | -10%   | -14%  | -14%  | -14%  |
|          | -24% | -35% | -13% | -54% | -42%   | -63%   | -8%   | +4%   | -14%  | -14%  | +21%  | -47%  | +7%   | +52%  | +16%  | +59%   | +47%   | -63%   | -54%   | -14%  | -14%  | -14%  |

The relative change in mobility obtained by this procedure is reported in Table VI. In all materials considered here this rescaling decreases the hole mobility, with the exception of SiC and BN. For the electron mobility, in all cases but diamond the mobility increases upon rescaling.

Along the same line, we also proceed to renormalize the electron-phonon matrix elements by using the experimental dielectric screening. This strategy allows us to capture effects similar to the many-body renormalization of the electron-phonon coupling. Since the mobility contains the square moduli of the matrix elements, we rescale the mobility using:

\[ \mu_{\text{screen}} = \mu \left( \frac{\epsilon_{\text{DFT}}}{\epsilon_{\text{exp}}} \right)^2. \]  

In this operation we employ the dielectric constants reported in Table V.

In Fig. 18 we show the impact of rescaling the matrix elements (black triangles) on the mobility together with the rescaling of the matrix elements and effective masses (blue crosses). For all cases we also report the mean absolute error (MAE) and the mean absolute relative error (MARE) between the computed and experimental mobilities.

Our best first-principles BTE Hall mobilities have a large absolute error around 1000 cm²/Vs due to the strong overestimation of the high electron mobility of GaAs and with a MARE of 64%. If we remove GaAs from the dataset, which is the compound with the highest mobility and hence it weighs disproportionately in the calculation of the average, we obtain a maximum absolute error of 335 cm²/Vs, and a MARE of 62%. When we proceed to renormalize the electron-phonon matrix ele-
FIG. 18: Comparison between theoretical and experimental mobility with experiment. Black triangles indicate the impact of rescaling the matrix elements while blue crosses indicate additional effective masses rescaling. The mean absolute error (MAE) and mean absolute relative error (MARE) are indicated as inset in the bottom figure and do not include the electron mobility of GaAs.

ments using the experimental dielectric constant, the results improve drastically with a MAE below 205 cm²/Vs and MARE of 35%. However, this improvement is partially cancelled when we also use the experimental effective masses, leading to a MARE of 55%.

These results indicate that improving the accuracy of the band structures and electron-phonon matrix elements, for example by using hybrid functionals or many-body perturbation theory, will be important to achieve agreement with experiment. Whether these improvements will lead to close agreement with experiment, or else a more powerful theory beyond the BTE formalism will be needed in the future, remains an important open question.

Finally, we note that regardless of the level of approximation or re-scaling applied, the electron mobility of 3C-SiC and AlAs, and the hole mobility of GaP are significantly higher than experiment. This suggests that higher mobilities could be achieved in these materials, paving the way for improved electronic devices.

V. CONCLUSION

In this work we have carried out an extensive investigation of the carrier transport properties of 10 simple semiconductors including diamond, silicon, GaAs, 3C-SiC, AlP, GaP, c-BN, AlAs, AlSb, and SrO. Exploiting a precise interpolation approach and an efficient procedure for numerical convergence, we have been able to compute from first-principles transport coefficients with high accuracy. This key step has allowed us to directly quantify the impact of a wide range of approximations often used in transport calculations. In addition, we have been able to assess the role of a number of physical factors such as the contribution of long-range dipole and quadrupole interactions, the SOC, effective masses and lattice parameters as well as the effect of temperature and magnetic field. One of the most interesting result of this extensive study is that our most accurate Hall mobility calculations seem to systematically overestimate the experimental data. In a few cases the overestimation is as large as a factor of two. This is a quite promising outcome as in our calculation we only included electron-phonon scattering and our result should be regarded as an upper limit to the Hall mobility achievable in highly pure samples, suggesting the possibility of further improvement in the transport properties of semiconductors of technological relevance. Here it is also important to stress that our work has clearly shown that the Hall factor exhibits significant variations across materials and as a function of temperature, ranging between 0.7 and 2. This is a clear indication that the common practice of comparing computed drift mobilities with experimental Hall mobilities should be done with particular care. Finally, it is important to stress that in our analysis we have also showed that part of the difference between theory and experiment may be connected with the inaccuracy in the DFT effective masses and electron-phonon matrix elements, but further work using explicit many-body perturbation theory calculations will be needed to confirm this point.

Our work demonstrates that first-principles calculations of carrier transport are making considerable progress. We believe we are reaching a point where detailed comparison with high-quality experimental data on pure samples will be increasingly important. For example, we highlight the need for new accurate measurements of the carrier mobility of SiC, AlAs, and GaP, and we hope to reignite interest in the experimental community to generate modern transport datasets for theorists to compare with. This manuscript describes a comprehensive protocol to perform such delicate calculations, and as a blueprint for reporting first-principles transport data in a way that is accessible, reliable, and reproducible. In this spirit, we encourage the adoption of the FAIR data and software principles in the study first-principles transport.

Acknowledgments

The authors would like to thank C. Verdi, X. Gonze, Junfeng Qiao, Hyungjun Lee, Massimiliano Stengel,
Miquel Royo, for useful discussions. Computer time was provided by the PRACE-17 and PRACE-21 resources MareNostrum at BSC-CNS, and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. S.P. acknowledges support from the European Unions Horizon 2020 Research and Innovation Programme, under the Marie Skłodowska-Curie Grant Agreement SELPH2D No. 839217. F.M. and N.B. acknowledge the Cirus UK National Tier-2 HPC Service at EPCC. E.R.M. acknowledges support from the National Science Foundation (Award No. OAC-1740263). N.M. acknowledges support from the Swiss National Science Foundation (Award No. OAC-1740263). N.M. acknowledges support from the European Union’s Horizon 2020 Research and Innovation Programme, under the Marie Skłodowska-Curie Grant Agreement SEPLasH2D No. 839217. F.M. and N.B. acknowledge the Cirus UK National Tier-2 HPC Service at EPCC. E.R.M. acknowledges support from the National Science Foundation (Award No. OAC-1740263). N.M. acknowledges support from the Swiss National Science Foundation (Award No. OAC-1740263). N.M. acknowledges support from the European Union’s Horizon 2020 Research and Innovation Programme, under the Marie Skłodowska-Curie Grant Agreement SEPLasH2D No. 839217. F.M. and N.B. acknowledge the Cirus UK National Tier-2 HPC Service at EPCC. E.R.M. acknowledges support from the National Science Foundation (Award No. OAC-1740263). N.M. acknowledges support from the Swiss National Science Foundation (Award No. OAC-1740263).

Appendix A: Optimal Wigner-Seitz construction

The interpolation of the Hamiltonian, dynamical matrix, and electron-phonon matrix elements from real space to reciprocal space via Eqs. [21],[22] and [23] is usually performed by using a radial cutoff on the direct lattice vector $|\tau|$. While this is certainly the simplest possible choice, it is not the most efficient, especially in the case of unit cells with high aspect ratios.

In order to improve the accuracy of the interpolation, it is convenient to select the direct lattice vectors so that the distance between two Wannier centers or the distance between two atoms are within a cutoff radius. For example, in the interpolation of the Hamiltonian we construct a shifted Wigner-Seitz cell of lattice vectors $R_p$ with the requirement that this cell be centered on one of the two Wannier functions, say $r_n$. The radial cutoff is then imposed with respect to the distance $|R_p + r_m - r_n|$, where $r_m$ denotes the center of the second Wannier function.

In practice, we determine the set of Wigner-Seitz vectors to describe the electronic Hamiltonian:

$$H_{mn}(R_p) = \frac{1}{N_p} \sum_{k \in \{i_p \times j_p \times o_p\}} U_{mn'}^\dagger k H_{mn'} k U_{n'n} k e^{-ikR_p},$$

(A1)

such that $R_p + r_m - r_n \in \{i_p \times j_p \times o_p\}$ where $r_m$ is the origin vector pointing to the position of the $m$th Wannier center. A similar approach was recently implemented in the Wannier90 software for the interpolation of the Hamiltonians.\[33\]

Similarly, in the case of the dynamical matrix we consider pairs of atoms $\tau_n$ and $\tau_{n'}$. For each such pair, we construct a Wigner-Seitz cell centered at $\tau_{n'}$, and we impose a truncation based on the distance $|R_{p'} + \tau_{n'} - \tau_n|$. For the optimal choice of a Wigner-Seitz vector for a discretized Brillouin zone grid $i_p' \times j_p' \times o_p'$, the short-range part of the dynamical matrix takes the form:\[34\]

$$D_{k'n',\beta}(R_{p'}) = \sum_{q \in \{i_{p'} \times j_{p'} \times o_{p'}\}} e^{-ikR_{p'} q} D_{k'n',\beta}(q) e_{\kappa'\beta},$$

(A2)

for all direct lattice vectors $R_{p'}$, such that $R_{p'} + \tau_{n'} - \tau_n \in \{i_{p'} \times j_{p'} \times o_{p'}\}$ and 0 otherwise. In Eq. (A2) we have made a rigid shift to the central primitive cell such that it only depends on one lattice vector $R_{p'}$. As a result one needs to store a set of atom-dependent Wigner-Seitz cells. In first-principles software such as Quantum ESPRESSO\[33\], a Fourier-transform grid is constructed with weights centered on pairs of atoms and zeros elsewhere. In EPW, we slightly optimize the computational cost of the interpolation by only retaining the union of nonzero elements between all the Wigner-Seitz cells.

In the case of the electron-phonon matrix elements, we exploit the same ideas as above, this time for both the Wannier functions and the atomic positions. The Wigner-Seitz cells for electrons and phonons are chosen with the requirement that this cell be centered on one of the two Wannier functions, say $r_n$. The radial cutoff is then imposed with respect to the distance $|R_p + r_m - r_n|$, where $r_m$ denotes the center of the second Wannier function.

In order to improve the accuracy of the interpolation, it is convenient to select the direct lattice vectors so that the distance between two Wannier centers or the distance between two atoms are within a cutoff radius. For example, in the interpolation of the Hamiltonian we construct a shifted Wigner-Seitz cell of lattice vectors $R_p$ with the requirement that this cell be centered on one of the two Wannier functions, say $r_n$. The radial cutoff is then imposed with respect to the distance $|R_p + r_m - r_n|$, where $r_m$ denotes the center of the second Wannier function.

In practice, we determine the set of Wigner-Seitz vectors to describe the electronic Hamiltonian:

$$H_{mn}(R_p) = \frac{1}{N_p} \sum_{k \in \{i_p \times j_p \times o_p\}} U_{mn'}^\dagger k H_{mn'} k U_{n'n} k e^{-ikR_p},$$

(A1)

such that $R_p + r_m - r_n \in \{i_p \times j_p \times o_p\}$ where $r_m$ is the origin vector pointing to the position of the $m$th Wannier center. A similar approach was recently implemented in the Wannier90 software for the interpolation of the Hamiltonians.\[33\]

Similarly, in the case of the dynamical matrix we consider pairs of atoms $\tau_n$ and $\tau_{n'}$. For each such pair, we construct a Wigner-Seitz cell centered at $\tau_{n'}$, and we impose a truncation based on the distance $|R_{p'} + \tau_{n'} - \tau_n|$. For the optimal choice of a Wigner-Seitz vector for a discretized Brillouin zone grid $i_p' \times j_p' \times o_p'$, the short-range part of the dynamical matrix takes the form:\[34\]

$$D_{k'n',\beta}(R_{p'}) = \sum_{q \in \{i_{p'} \times j_{p'} \times o_{p'}\}} e^{-ikR_{p'} q} D_{k'n',\beta}(q) e_{\kappa'\beta},$$

(A2)

for all direct lattice vectors $R_{p'}$, such that $R_{p'} + \tau_{n'} - \tau_n \in \{i_{p'} \times j_{p'} \times o_{p'}\}$ and 0 otherwise. In Eq. (A2) we have made a rigid shift to the central primitive cell such that it only depends on one lattice vector $R_{p'}$. As a result one needs to store a set of atom-dependent Wigner-Seitz cells. In first-principles software such as Quantum ESPRESSO\[33\], a Fourier-transform grid is constructed with weights centered on pairs of atoms and zeros elsewhere. In EPW, we slightly optimize the computational cost of the interpolation by only retaining the union of nonzero elements between all the Wigner-Seitz cells.

In the case of the electron-phonon matrix elements, we exploit the same ideas as above, this time for both the Wannier functions and the atomic positions. The Wigner-Seitz cells for electrons and phonons are chosen so that both the $|R_p + r_m - r_n|$ and $|R_{p''} + \tau_n - r_n|$ distances are minimized:

$$g_{mn\alpha\beta}(R_p, R_{p''}) = \frac{1}{N_p N_{p''}} \sum_{k \in \{i_p \times j_p \times o_p\}} \sum_{q \in \{i_{p'} \times j_{p'} \times o_{p'}\}} e^{-ikR_p + q R_{p''}} \sum_{m'n''} U_{mn'k + q} U_{m'n''k}$$

\times \frac{2M_{\kappa\alpha\beta}}{\hbar} \frac{M_{\kappa'\beta}}{\hbar} e^{i\kappa\alpha\beta} U_{mn'k + q} U_{m'n''k}^{\dagger} (k, q) U_{n'n} k,$$

(A3)

where the lattice vectors are obtained such that $R_{p''} + \tau_n - r_n \in \{i_{p'} \times j_{p'} \times o_{p'}\}$ and $R_p + r_m - r_n \in \{i_p \times j_p \times o_p\}$. We note that the set of lattice vectors $R_{p''}$ used for the
interpolation of electron-phonon matrix elements can in general be different from the set employed for the interpolation of the dynamical matrix. This procedure used to determine the optimal set of vectors is sketched in Fig. 19. The set of direct lattice vectors \( \mathbf{R}_p \) and \( \mathbf{R}_p' \), are selected such that the trial point \( \mathbf{t} = \mathbf{R}_p + \mathbf{r}_m - \mathbf{r}_n \), constructed from the distance between two Wannier centers, or \( \mathbf{t}' = \mathbf{R}_p' + \mathbf{r}_m - \mathbf{r}_n \) constructed from a Wannier center and an atomic position in the central cell and any other cell, is closer to the supercell center \( \mathbf{0} \) than any other supercell centers \( \mathbf{S} \). In the case of a trial point at equal distance between \( \mathbf{0} \) and \( \mathbf{S} \) (i.e., on the surface of the supercell Wigner-Seitz cell) a weight proportional to the number of degeneracies is given. The optimal Wigner-Seitz construction is activated with the new \texttt{use.ws} input variable in EPW. As a sanity check, we verified that the mobility of 3C-SiC calculated on a 8×8×8 \( k \)-point and 4×4×4 \( q \)-point grid was the same as the one obtained from a 2×1×1 supercell with a 4×8×8 \( k \)-point and 2×4×4 \( q \)-point grid.

In Fig. 20 we report the calculated phonon dispersions of the three acoustic modes of 3C-SiC. The dashed gray line represents the interpolation from a coarse \( 4 \times 4 \times 4 \) \( q \)-point and \( q \)-point grid using a \( \Gamma \)-centered fixed Wigner-Seitz cell with dipole-dipole correction (dash gray line), an optimal Wigner-Seitz cell with dipole-dipole correction (red line), and an optimal Wigner-Seitz cell with dipole-dipole, dipole-quadrupole and quadrupole-quadrupole corrections (blue line). The black dots are reference datapoints obtained via direct DFPT calculations.

We also show in Fig. 21 a comparison between the deformation potential of 3C-SiC computed with a \( \Gamma \)-centered Wigner-Seitz cell or with the improved construction. The new construction produces results closer to those obtained directly with DFPT everywhere. We note that the results shown in Fig. 21 are for an un-converged small coarse grid of 3×3×3 for both \( k \) and \( q \)-points. When using denser coarse grids, these difference rapidly disappear (at least for the present case of cubic semiconductors). Given that we use much denser coarse grids in this work, for simplicity we employed the basic \( \Gamma \)-centered construction in all calculations reported in this work.

Appendix B: Additional figures and tables
| Compounds | Coarse Temp. | Wind. | Fine Hole mobility | Hall factor | Wind. | Fine Electron mobility | Hall factor |
|-----------|--------------|-------|-------------------|------------|-------|------------------------|------------|
|           | k q (K) | eV grids | SERTA | BTE | SERTA | BTE | SERTA | BTE | SERTA | BTE |
| Diamond   | 8° 4° | 100 0.30 100° | 6694.2 | 6700.1 | 0.721 | 0.749 | 0.40 80° | 11016 | 10675 | 0.656 | 0.662 |
|           | 300   | 1048.5 | 1051.3 | 0.755 | 0.792 |
|           | 300   | 1049.6 | 1052.4 | 0.755 | 0.793 |
| Silicon   | 8° 4° | 100 0.30 100° | 1715.4 | 1746.3 | 0.809 | 0.875 |
|           | 300   | 1110.6 | 10676 | 0.656 | 0.662 |
| GaAs      | 8° 4° | 100 0.20 100° | 7036.7 | 6709.5 | 0.720 | 0.748 | 0.20 80° | 11035 | 10693 | 0.655 | 0.661 |
| 3C-SiC    | 8° 4° | 100 0.30 60° | 7175.5 | 7070.4 | 0.129 | 0.135 | 0.30 60° | 12592 | 13167 | 0.638 | 0.693 |
|           | 300   | 691.43 | 708.94 | 0.292 | 0.335 |
| AIP       | 8° 4° | 100 0.20 100° | 2700.0 | 2728.0 | 0.622 | 0.693 |
|           | 300   | 82.11 | 91.82 | 0.532 | 0.583 |
| GaP       | 8° 4° | 100 0.15 60° | 1043.4 | 1078.0 | 0.416 | 0.461 |
|           | 300   | 136.15 | 171.62 | 0.412 | 0.456 |
| c-BN      | 8° 4° | 100 0.30 60° | 1363.0 | 1718.8 | 0.419 | 0.555 | 0.30 60° | 2347.9 | 2275.4 | 0.822 | 0.879 |
| AlAs      | 8° 4° | 100 0.15 60° | 1535.4 | 1595.0 | 0.674 | 0.865 |
|           | 300   | 1535.4 | 1595.0 | 0.674 | 0.865 |
| AISb      | 8° 4° | 100 0.30 60° | 1720.1 | 1785.6 | 0.346 | 0.515 | 0.30 60° | 1490.5 | 1394.0 | 0.645 | 0.703 |
| Sn        | 8° 4° | 100 0.15 60° | 1735.0 | 1801.1 | 0.343 | 0.511 | 0.15 60° | 1471.6 | 1404.9 | 0.660 | 0.698 |
| SrO       | 8° 4° | 100 0.30 60° | 319.9 | 331.0 | 0.795 | 0.794 | 0.30 60° | 166.9 | 152.3 | 0.761 | 0.793 |
|           | 300   | 5.89 | 7.64 | 0.940 | 0.931 |

TABLE VII: Sensitivity of the calculated mobility and Hall factor to the size of the energy window in the BTE. The calculations have been performed using adaptive smearing. The numbers indicated in bold represent the smallest window such that the results are unaffected. The same energy window has been used throughout this work.
| Materials | C-e | 1383 | 1613 | 1583 | 1634 | 1664 | 1705 |
|----------|-----|------|------|------|------|------|------|
|          | C-h | 2256 | 2290 | 2364 | 2410 | 2409 | 2467 |
|          | Si-e | 1885 | 2068 | 1803 | 1944 | 2204 | 2257 |
|          | Si-h | 623  | 627  | 600  | 601  | 572  | 574  |
|          | GaAs-e | 8743 | 14799| 8853 | 15099| 12354| 17860|
|          | GaAs-h | 392  | 500  | 389  | 500  | 988  | 1068 |
|          | SiC-e | 166  | 170  | 165  | 170  | 105  | 111  |
|          | SiC-h | 166  | 170  | 165  | 170  | 105  | 111  |
|          | AIP-e | 461  | 556  | 454  | 554  | 629  | 698  |
|          | AIP-h | 94   | 102  | 93   | 101  | 84   | 92   |
|          | GaP-e | 281  | 325  | 284  | 328  | 409  | 442  |
|          | GaP-h | 228  | 271  | 231  | 276  | 315  | 353  |
|          | BN-e  | 573  | 1002 | 470  | 967  | 456  | 956  |
|          | BN-h  | 196  | 321  | 170  | 319  | 124  | 281  |
|          | AlAs-e | 409  | 507  | 405  | 501  | 578  | 640  |
|          | AlAs-h | 181  | 217  | 180  | 215  | 337  | 357  |
|          | AlSb-e | 298  | 332  | 303  | 335  | 415  | 431  |
|          | AlSb-h | 491  | 598  | 490  | 600  | 918  | 1000 |
|          | SrO-e | 38   | 53   | 38   | 53   | 53   | 65   |
|          | SrO-h | 3.5  | 5.0  | 3.4  | 4.9  | 3.5  | 5.0  |

TABLE VIII: Room temperature drift SERTA and BTE mobility (cm²/Vs) as well as temperature exponent, as evaluated at the densest fine k/q-point grids from Fig. 8, and drift SERTA and BTE Hall mobility extrapolated to infinitely dense sampling. The experimental data in bold are the reference values used in Figs. 17 and 18. The † symbol indicates time-of-flight drift mobility experiments, while no symbols indicate Hall measurements.
FIG. 22: Conduction band edge of silicon – including spin-orbit coupling (SOC) – where the zero energy has been set at the X high-symmetry point and computed using DFT (black line). (a) Electronic bands obtained using Wannier interpolation with $k$-point grids ranging from $10^3$ to $16^3$ and constructed from an initial set of four $sp^3$ projections on each Si atom. Both the valence and conduction bands have been Wannierized, totalling 16 Wannier functions due to SOC. (b) Electronic bands obtained using Wannier interpolation with $k$-point grids ranging from $10^3$ to $16^3$ and constructed from an initial set of one $s$ and five $d$ projections on one Si atom only. Only the conduction bands have been Wannierized, totalling 12 Wannier functions due to SOC.

FIG. 23: Effective quadrupole term $\tilde{Q}^{\beta\gamma\kappa\alpha}_{\beta\gamma\kappa\alpha}$ obtained from the least-mean square fit minimization of direct DFPT calculations for 3C-SiC. In these calculations we used the LDA exchange and correlation functional and without nonlinear core correction, so that we can compare with Ref. [22]. The two non-equivalent values of the quadrupole tensor are obtained by using the phonon frequencies, Born-effective charges, dielectric constant, eigendisplacement vectors and wavefunction overlaps from explicit DFPT calculations, and by optimizing the quadrupole tensor using Eq. (40). We repeat the fitting for several sets of $q$-points on spherical surfaces of varying radii $R$. The radii are in units of $2\pi/a$.

FIG. 24: Calculated hole drift mobility of c-BN for various Brillouin zone grid sizes. The horizontal axis indicates the coarse $q$-point grid, and the lines correspond to different choices of the coarse $k$-point grid: same as the $q$-point grid (blue), twice as dense as the $q$-point grid (green), and three times as dense as the $q$-point grid (red). In all cases the fine grids were set to $60^3$ points for both $k$ and $q$. We also show with dashed lines the relative computational time needed in each case (normalized to the most time-consuming calculation). These tests show that the results obtained by using equal coarse grids for electrons and phonons are of the same accuracy as those with denser electron grids, but much cheaper computationally. Therefore we recommend to always use equal coarse grids.
FIG. 25: (a) Convergence of the calculated Hall mobility of GaAs with the fine Brillouin zone grid size. Both SERTA and BTE results are shown using the converged coarse grids from Fig. [x]. We consider several smearing parameters as well as the adaptive smearing. (b) Calculated electron mobility of GaAs, extrapolated to infinitely dense Brillouin zone sampling, for several smearing parameters. The extrapolation of these results at zero smearing agrees with calculations using adaptive smearing (stars).
FIG. 26: Calculated electron and hole scattering rates as a function of energy at 150 K, 300 K, and 500 K. The energy range displayed here encompasses the energy window used in the calculations of carrier mobilities.
