# Phonon-limited transport in two-dimensional materials: A unified approach for *ab initio* mobility and current calculations

Jonathan Backman[ORCID],[*] Youseung Lee[ORCID], and Mathieu Luisier[ORCID]

*Integrated Systems Laboratory, ETH Zürich, Zürich 8092, Switzerland*

In this paper, we present an *ab initio* methodology to account for electron-phonon interactions in two-dimensional (2D) materials, focusing on transition-metal dichalcogenides (TMDCs). It combines density functional theory and maximally localized Wannier functions to acquire material data and relies on the linearized Boltzmann transport equation (LBTE) and the nonequilibrium Green's functions (NEGF) method to determine the transport properties of materials and devices, respectively. It is shown that both the LBTE and NEGF methods return very close mobility values, without the need to adjust any parameters. The excellent agreement between the two approaches results from the inclusion of nondiagonal entries in the electron-phonon scattering self-energies. The NEGF solver is then used to shed light on the "current versus voltage" characteristics of a monolayer $MoS_2$ transistor, highlighting how the interactions with phonons impact both the current magnitude and its distribution. The mobility of other TMDCs is considered as well, demonstrating the capabilities of the proposed technique to assess the potential of 2D channel materials in next-generation logic applications.

## I. INTRODUCTION

More than 60 years of electronic device miniaturization have pushed silicon, the semiconductor of reference, toward its intrinsic limit, thus calling for the emergence of alternative channel materials. Two-dimensional (2D) compounds such as transition-metal dichalcogenides (TMDCs) are seen as promising candidates to equip future electronic products due to their unique electronic [1], thermal [2], and mechanical [3] properties. The first demonstration of a single-layer molybdenum disulfide ($MoS_2$) transistor [4] moved the focus of the device community toward single- and few-layer TMDCs as channel of logic switches. Since then, motivated by technological advancements in the growth and exfoliation of 2D semiconductors [5,6], transistors based on a wide range of TMDCs, such as $WSe_2$ [7], $WS_2$ [8], $MoTe_2$ [9], $MoSe_2$ [10], $ReS_2$ [11], $HfSe_2$, or $ZrSe_2$ [12], have been reported. Notably, a recent work demonstrated a nanosheet field-effect transistor (FET) in a gate-all-around configuration with a monolayer $MoS_2$ channel [13]. Despite the progress that has been made in the fabrication of 2D FETs, these components remain far from reaching their ultimate potential in terms of carrier transport. The limits of the latter properties are very often assessed through theoretical investigations, which allows for the elimination of nonintrinsic factors, e.g., impurities or interface defects. What remains are carrier-phonon interactions, which play an important role at room temperature [14]. They therefore require a special treatment.

*Ab initio* methods—in particular, density functional theory (DFT)—have proven invaluable to predict the carrier-transport properties of solids in the presence of interactions with phonons. These models involve combining electronic band structures [15,16] with phonon dispersions [17,18] to obtain the required electron-phonon (*e*-ph) interaction elements [19–22]. The success of these *ab initio* techniques at predicting intrinsic carrier transport in bulk materials has motivated the exploration of phonon-limited mobility in a variety of 2D semiconductors [23–34]. It should be emphasized that these studies on 2D materials often report a broad spectrum of mobility values, revealing a lack of consensus and frequent discrepancies with experimental data. Several factors can explain the large range of mobility values, the different DFT settings that are used being probably the most important one.

The knowledge of the phonon-limited mobility provides insight into the potential of a given material as a transistor channel but it only tells one side of the story. To assess the suitability of integrating a material into next-generation logic devices, the scope of inquiry must be broadened. Ultimately, what matters is the influence of carrier-phonon interactions on the current that flows through the transistor, which requires full device simulations. While previous 2D device simulations have included electron-phonon interactions [35–37], they have often relied on simplified models where only a limited portion of these interactions is retained. The absence of a unified model that allows for the

_____

[*]Corresponding author: jbackman@iis.ee.ethz.ch

inclusion of electron-phonon scattering both in materials and device calculations motivated the current study. It aims to link the materials and device properties within a single framework to enable realistic investigations of transistors with a 2D channel material.

We introduce a fully atomistic approach to compute electron-phonon interactions that facilitates the incorporation of materials properties into device calculations with focus on 2D materials [38]. The method exploits DFT and maximally localized Wannier functions (MLWFs) to extract material data that can be fed either to a linearized Boltzmann transport equation (LBTE) or to a nonequilibrium Green's function (NEGF) solver, delivering a holistic perspective on phonon-limited transport properties. Through the deployment of this approach, the impact of electron-phonon scattering on the functionality of 2D devices can be accurately determined, enhancing our understanding of the mechanisms limiting the performance of these compounds. Furthermore, mobility serves as a benchmark to validate and cross correlate the results obtained with the LBTE and NEGF methods. Such comparisons ensure the robustness and reliability of the derived insights.

The paper is organized as follows. In Sec. II, we present the developed modeling techniques, starting with the approach to calculate electron-phonon coupling elements. This section also outlines the electronic structure method that we have adopted and elaborates on how electron-phonon scattering is treated within the LBTE and NEGF frameworks. Section III concentrates on the results of our research. First, electronic and phononic dispersions are reported for various TMDCs, before the mobility values obtained via both the LBTE and NEGF methodologies are compared with each other. Finally, the capability of the developed simulation framework is demonstrated through device simulations, highlighting the role of electron-phonon scattering in influencing both the current magnitude and its spectral distribution in a single-gate $MoS_2$ transistor. The paper is summarized and conclusions are drawn in Sec. IV.

## II. METHODS

In this section, the theoretical basis and computational frameworks pivotal to this study are laid out. We start by constructing an atomistic electron-phonon scattering model, specifically tailored for electronic transport calculations. A transformation from a plane-wave basis to localized electronic states is required for that purpose. Special attention is paid to the periodicity and confinement of the dimensional degrees of freedom pertinent to 2D materials and devices. Subsequent sections address the application of the LBTE and NEGF approaches to calculate the transport properties of the considered systems. The emphasis is set on the integration of the developed

electron-phonon scattering model. The goal here is to provide detailed explanation of all ingredients entering our approach, setting the stage for the upcoming discussion on our findings and their broader implications.

### A. Electronic and vibrational structure methods

Our method to compute electron-phonon coupling elements is outlined here. It relies on the harmonic approximation, DFT, MLWF [39], and the frozen-phonon approach [17]. Starting from the Born-Oppenheimer adiabatic approximation, which allows for the separation of electronic and vibrational time scales [40], the total Hamiltonian of the system of interest can be decomposed into three distinct components,

$$\hat{\mathbf{H}} = \hat{\mathbf{H}}_e^0 + \hat{\mathbf{H}}_{\text{ph}}^0 + \hat{\mathbf{H}}_{e\text{-ph}}. \tag{1}$$

In Eq. (1), $\hat{\mathbf{H}}_e^0$ is the single-particle mean-field Hamiltonian. It governs the behavior of electrons within a fixed ionic lattice. Conversely, $\hat{\mathbf{H}}_{\text{ph}}^0$ is the Hamiltonian describing independent noninteracting phonons. Finally, $\hat{\mathbf{H}}_{e\text{-ph}}$ denotes the electron-phonon Hamiltonian, which captures the interaction between these two populations.

#### 1. Electron-phonon coupling

The Hamiltonian encompassing the electron-phonon interactions in a periodic system can be generally written as [41,42]

$$\hat{\mathbf{H}}_{e\text{-ph}} = \sum_{mn\mathbf{k}} \sum_{\lambda\mathbf{q}} M_{mn}^{\lambda}(\mathbf{k}, \mathbf{q}) \hat{\mathbf{c}}_{m\mathbf{k}+\mathbf{q}}^{\dagger} \hat{\mathbf{c}}_{n\mathbf{k}} [\hat{\mathbf{b}}_{\lambda\text{-}\mathbf{q}}^{\dagger} + \hat{\mathbf{b}}_{\lambda\mathbf{q}}], \tag{2}$$

where $\hat{\mathbf{c}}_{m\mathbf{k}+\mathbf{q}}^{\dagger}(\hat{\mathbf{c}}_{n\mathbf{k}})$ and $\hat{\mathbf{b}}_{\lambda\text{-}\mathbf{q}}^{\dagger}(\hat{\mathbf{b}}_{\lambda\mathbf{q}})$ are the electron and phonon creation (annihilation) operators, respectively. The indices $m$ and $n$ correspond to the basis functions of the electrons, while $\lambda$ represents the phonon band index. The momentum vectors of electrons are indicated by $\mathbf{k}$, while those of phonons are given by $\mathbf{q}$. To understand the form of Eq. (2), we should go back to the Hamiltonian of the electronic system, $\hat{\mathbf{H}}_e$. Because electrons operate on a far shorter timescale than heavy nuclei, the adiabatic approximation suggests that the nuclear coordinates have a parametric influence on the electronic Hamiltonian, i.e., $\hat{\mathbf{H}}_e = \hat{\mathbf{H}}_e(\mathbf{Q})$, where $\mathbf{Q} = \boldsymbol{\tau} - \boldsymbol{\tau}^0$ is the displacement of an atom at position $\boldsymbol{\tau}$ from its static ionic lattice configuration $\boldsymbol{\tau}^0$. To account for the perturbation $\mathbf{Q}$, the electronic Hamiltonian can be expanded in a Taylor series around the equilibrium positions of the atoms. In the harmonic

approximation, only the first-order terms are retained [19]:

$$\hat{\mathbf{H}}_e \approx \hat{\mathbf{H}}_e^0 + \sum_{I\eta\gamma} \langle \hat{\Psi} | \frac{\partial \hat{\mathbf{H}}_e^0}{\partial Q_{I\eta\gamma}} | \hat{\Psi} \rangle \, \hat{\mathbf{Q}}_{I\eta\gamma} = \hat{\mathbf{H}}_e^0 + \hat{\mathbf{H}}_{e\text{-ph}}. \quad (3)$$

In this expression, the first term corresponds to the equilibrium electronic Hamiltonian $\hat{\mathbf{H}}_e^0$ in Eq. (1). The second term, which includes the derivative of the electronic Hamiltonian operator with respect to the displacement of atom $I$ in unit cell $\gamma$ along the Cartesian coordinate $\eta$, represents the interaction between electrons and phonons. It is projected onto a basis made of the wave functions $|\hat{\Psi}\rangle$. The displacement of the atoms from their equilibrium position due to lattice vibrations is given by the quantum operator [41,42]

$$\hat{\mathbf{Q}}_{I\eta\gamma} = \sum_{\lambda\mathbf{q}} \sqrt{\frac{\hbar}{2N_{\mathbf{q}} m_I \omega_{\lambda\mathbf{q}}}} f_{\lambda\mathbf{q}}^{I\eta} e^{i\mathbf{q}\cdot\mathbf{R}_\gamma} [\hat{b}_{\lambda\text{-}\mathbf{q}}^\dagger + \hat{b}_{\lambda\mathbf{q}}], \quad (4)$$

where $N_{\mathbf{q}}$ represents the number of $\mathbf{q}$ points, $m_I$ is the mass of the displaced atom, and $f_{\lambda\mathbf{q}}^{I\eta}$ is the ionic displacement vector of mode $(\lambda, \mathbf{q})$ with frequency $\omega_{\lambda\mathbf{q}}$. It is assumed that the perturbed ion is located in a unit cell $\gamma$ that has been displaced from the reference cell by a lattice vector $\mathbf{R}_\gamma$. Next, we expand the wave function $|\hat{\Psi}\rangle$ in terms of a localized basis set,

$$|\hat{\Psi}\rangle = \frac{1}{\sqrt{N}} \sum_{\alpha n \mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_\alpha} |n, \mathbf{R}_\alpha\rangle \, \hat{c}_{n\mathbf{k}}. \quad (5)$$

Here, $|n, \mathbf{R}_\alpha\rangle$ is a state localized at position $\mathbf{R}_\alpha$ in unit cell $\alpha$ and $N$ is the number of unit cells. Inserting Eqs. (4) and (5) into Eq. (3) now gives an expression for the electron-phonon interaction Hamiltonian:

$$\hat{\mathbf{H}}_{e\text{-ph}} = \frac{1}{N} \sum_{\alpha n \mathbf{k}} \sum_{\beta m \mathbf{k}'} \sum_{I\eta\gamma\lambda\mathbf{q}} e^{-i\mathbf{k}'\cdot\mathbf{R}_\beta} e^{i\mathbf{k}\cdot\mathbf{R}_\alpha} e^{i\mathbf{q}\cdot\mathbf{R}_\gamma} .$$

$$\langle m, \mathbf{R}_\beta | \frac{\partial \hat{\mathbf{H}}_e^0}{\partial Q_{I\eta\gamma}} | n, \mathbf{R}_\alpha \rangle \sqrt{\frac{\hbar}{2N_{\mathbf{q}} m_I \omega_{\lambda\mathbf{q}}}} f_{\lambda\mathbf{q}}^{I\eta} \hat{c}_{m\mathbf{k}'}^\dagger \hat{c}_{n\mathbf{k}} .$$

$$[\hat{b}_{\lambda\text{-}\mathbf{q}}^\dagger + \hat{b}_{\lambda\mathbf{q}}]. \quad (6)$$

Because of the periodicity of the system, the term containing the derivative of the Hamiltonian $\hat{\mathbf{H}}_e$ can be shifted to have $\mathbf{R}_\alpha$ as the reference unit cell, i.e.,

$$\langle m, \mathbf{R}_\beta | \frac{\partial \hat{\mathbf{H}}_e^0}{\partial Q_{I\eta\gamma}} | n, \mathbf{R}_\alpha \rangle = \langle m, \mathbf{R}_\beta - \mathbf{R}_\alpha | \frac{\partial \hat{\mathbf{H}}_e^0}{\partial Q_{I\eta\gamma}} | n, \mathbf{R}_\alpha - \mathbf{R}_\alpha \rangle .$$

By defining all lattice vectors relative to $\mathbf{R}_\alpha$, we can rewrite the phase factors in Eq. (6) as

$$e^{-i\mathbf{k}'\cdot\mathbf{R}_\beta} e^{i\mathbf{k}\cdot\mathbf{R}_\alpha} e^{i\mathbf{q}\cdot\mathbf{R}_\gamma} = e^{i(\mathbf{k}+\mathbf{q}-\mathbf{k}')\cdot\mathbf{R}_\alpha} e^{-i\mathbf{k}'\cdot\mathbf{R}_{\beta 0}} e^{i\mathbf{q}\cdot\mathbf{R}_{\gamma 0}},$$

where $\mathbf{R}_{\beta 0} = \mathbf{R}_\beta - \mathbf{R}_\alpha$ and $\mathbf{R}_{\gamma 0} = \mathbf{R}_\gamma - \mathbf{R}_\alpha$ are relative displacements with respect to $\mathbf{R}_\alpha$. The sum over $\alpha$ in Eq. (6) can now be analytically carried out, giving rise to a factor $\sum_\alpha e^{i(\mathbf{k}+\mathbf{q}-\mathbf{k}')\cdot\mathbf{R}_\alpha} = N\delta_{\mathbf{k}',\mathbf{k}+\mathbf{q}}$, which enforces momentum conservation. By applying these simplifications to Eq. (6) and comparing the result with Eq. (2), an expression for the electron-phonon coupling elements can be derived:

$$M_{mn}^{\lambda\eta}(\mathbf{k}, \mathbf{q}) = \sum_\beta \sum_{I\gamma} e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}_{\beta 0}} e^{i\mathbf{q}\cdot\mathbf{R}_{\gamma 0}} .$$

$$\langle m, \mathbf{R}_{\beta 0} | \frac{\partial \hat{\mathbf{H}}_e^0}{\partial Q_{I\eta\gamma}} | n, \mathbf{0} \rangle \sqrt{\frac{\hbar}{2N_{\mathbf{q}} m_I \omega_{\lambda\mathbf{q}}}} f_{\lambda\mathbf{q}}^{I\eta}, \quad (7)$$

where the dependence on the displacement direction $(\eta)$ is explicitly retained in the expression for the coupling elements.

### 2. Electronic Hamiltonian

The electronic Hamiltonian $\hat{\mathbf{H}}_e^0$ plays a key role in electronic transport, both directly and through the electron-phonon coupling elements. Here, it is calculated from first principles using plane-wave DFT, as implemented in the Vienna *ab initio* simulation package (VASP) [15]. The resulting Bloch wave functions are then transformed into a set of MLWFs using the WANNIER90 code [43]:

$$|n, \mathbf{R}\rangle = \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_l \mathbf{U}_{nl}(\mathbf{k}) |l, \mathbf{k}\rangle . \quad (8)$$

In Eq. (8), the $|l, \mathbf{k}\rangle$ are the Bloch wave functions with band index $l$ and momentum $\mathbf{k}$, whereas $V$ is the volume of the unit cell. Calculating the MLWFs primarily consists of finding the unitary transformation matrix $\mathbf{U}_{nl}(\mathbf{k})$ representing a rotation of the Bloch states that minimizes the quadratic spread of the Wannier functions. This is done through an iterative process where the electronic Bloch states projected onto atomic orbitals serve as an initial guess. Once the $\mathbf{U}_{nl}(\mathbf{k})$ matrix has been determined, it can be used to transform the Bloch Hamiltonian $H_{ij}(\mathbf{k})$ into the MLWF basis,

$$H_{mn}(\mathbf{R}) = \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_{i,j} \mathbf{U}_{mi}(\mathbf{k}) H_{ij}(\mathbf{k}) \mathbf{U}_{jn}^\dagger(\mathbf{k}), \quad (9)$$

where we define $H_{mn}(\mathbf{R}) = \langle m, \mathbf{R} | \hat{\mathbf{H}}_e^0 | n, \mathbf{0} \rangle$. The localized nature of the Wannier functions makes the Hamiltonian matrix sparse; the interactions between Wannier functions decrease rapidly as the distance between them increases. This makes the MLWF basis suitable for transport calculations at the unit cell or device level. By associating

each Wannier function with atoms located on a regular lattice, we can remap the primitive unit cell Hamiltonian into a larger orthorhombic cell that only includes interactions with first nearest neighbors (NNs) in **R** [44]. This remapping is depicted in Fig. 1. By applying this scheme, the same primitive unit cell DFT Hamiltonian can either be (i) interpolated onto a dense momentum grid $(k_x, k_y, k_z)$ for 3D or $(k_x, k_y)$ for 2D materials,

$$H_{mn}(\mathbf{k}) = \sum_{\mathbf{R}} H_{mn}(\mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}, \tag{10}$$

to calculate mobilities with a LBTE solver or (ii) extended in space to correspond to the device structure considered. In the latter case, the periodicity along the transport direction is broken and the resulting system of equations is solved with the NEGF method. For a 2D device, with transport along the $x$ axis, confinement along the $y$ axis, and with the $z$ axis periodic, the device Hamiltonian becomes $k_z$ dependent, $H_{mn}(k_z)$, and Eq. (10) reduces to

$$H_{mn}(k_z) = H_{mn}^0 + H_{mn}^- e^{-ik_z R_z} + H_{mn}^+ e^{ik_z R_z}. \tag{11}$$

Here, orthorhombic unit cells of width $R_z$ along $z$ are assumed. Interactions within this unit cell are cast into the matrix $H^0$, while connections to the NN cell at $+z$ $(-z)$ form the matrix $H^+$ $(H^-)$. As already mentioned, $R_z$ is chosen large enough so that only NN-cell interactions exist [35].

### 3. Hamiltonian derivative

Obtaining the electron-phonon coupling elements in Eq. (7) requires the computation of expectation values of the gradient of the Hamiltonian operator. Following the approach of Ref. [19], one can bypass this direct calculation by focusing on the derivative of the Hamiltonian matrix elements and taking into account the change of basis functions [45]. This results in a compact expression for the sought derivative,

$$\langle m| \frac{\partial \hat{\mathbf{H}}_e^0}{\partial Q_{I\eta\gamma}} |n\rangle = \frac{\partial \langle m| \hat{\mathbf{H}}_e^0 |n\rangle}{\partial Q_{I\eta\gamma}} - \left( \langle m'| \hat{\mathbf{H}}_e^0 |n\rangle + \langle m| \hat{\mathbf{H}}_e^0 |n'\rangle \right), \tag{12}$$

where $|n'\rangle \equiv \frac{\partial |n\rangle}{\partial Q_{I\eta\gamma}}$ represents the shift in basis orbital $|n\rangle$ associated with the displacement $Q_{I\eta\gamma}$. The dependence on **R** in Eq. (12) has been dropped for simplicity. The first part of this expression, the gradient of the Hamiltonian matrix elements, can be evaluated in a straightforward manner by
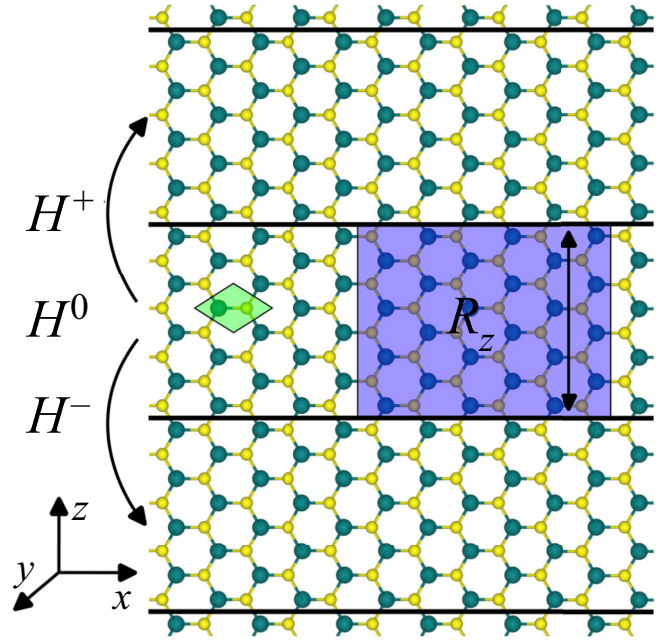


FIG. 1. Top view of a $MoS_2$ monolayer constructed by replicating either the primitive cell of $MoS_2$ (green-shaded area on the left) or the orthorhombic unit cell of width $R_z$ (blue-shaded area on the right). Electrons are confined along the $y$ axis (out of plane). The primitive hexagonal unit cell of $MoS_2$ is used to calculate the electronic structure of the material, generate MLWFs, and solve the LBTE system, while the orthorhombic unit cell is utilized to construct the full device Hamiltonian that enters the NEGF equations.

using a finite-difference scheme in the localized basis,

$$\frac{\partial \langle m| \hat{\mathbf{H}}_e^0 |n\rangle}{\partial Q_{I\eta\gamma}} \approx \frac{[H_{mn}(Q_{I\eta\gamma}) - H_{mn}(-Q_{I\eta\gamma})]}{2Q_{I\eta\gamma}}. \tag{13}$$

Here, $H_{mn}(Q_{I\eta\gamma})$ represents the Hamiltonian matrix element $\langle m| \hat{\mathbf{H}}_e^0 |n\rangle$, where atom $I$ in unit cell $\gamma$ has been displaced by a distance $Q$ along the direction $\eta$. To compute this derivative, a supercell is first built by repeating the primitive unit cell of the considered material along its periodic lattice vectors. Second, a collection of supercells is assembled by individually displacing the atoms situated within the central primitive unit cell. The size of the supercell must be chosen in such a way that all parameter combinations $(m, n, I, \gamma, \eta)$ that lead to a non-negligible derivative in Eq. (13) remain within the constructed domain. The number of displacements $(I, \eta)$ can be reduced by limiting the displaced atoms and directions to symmetrically independent perturbations. Finally, a MLWF Hamiltonian is obtained for each displacement, based on the method previously described. The localized nature of the MLWFs ensures that the derivative of the Hamiltonian is nonzero primarily in the vicinity of the displaced atom, so that the sums over $\beta$ and $\gamma$ in Eq.

(7) is limited to neighboring cells only. Although the calculations utilize supercells, the transformation of the Hamiltonian derivatives in Eq. (12)—and consequently the electron-phonon coupling elements in Eq. (7)—into a localized basis allows us to map them onto unit (LBTE) or device (NEGF) transport cells, as for the Hamiltonian entries. A key advantage of this approach resides in its ability to evaluate the derivative of the Hamiltonian operator regardless of the specificities of the selected electronic structure code, contrary to the requirements imposed by density functional perturbation theory (DFPT). Our method shows enhanced flexibility, making it compatible with various DFT codes, either through the transformation of plane waves into Wannier functions or by directly employing a localized basis set, e.g., linear combination of atomic orbitals (LCAOs) [46] or Gaussian-type orbitals (GTOs) [47].

The second part of Eq. (12) can be expanded by introducing the projection operator

$$\sum_{kl} |k\rangle \, (\mathbf{S}^{-1})_{kl} \, \langle l| = 1, \tag{14}$$

where $\mathbf{S}$ is the orbital overlap matrix. By inserting Eq. (14) into Eq. (12), we obtain a more practical form of the Hamiltonian derivative [19]:

$$\langle m| \frac{\partial \hat{\mathbf{H}}_e^0}{\partial Q_{I\eta\gamma}} |n\rangle = \frac{\partial \, \langle m| \hat{\mathbf{H}}_e^0 |n\rangle}{\partial Q_{I\eta\gamma}} - \\ \sum_{kl} \langle m'|k\rangle \, (\mathbf{S}^{-1})_{kl} \, \langle l| \hat{\mathbf{H}}_e^0 |n\rangle + \langle m| \hat{\mathbf{H}}_e^0 |k\rangle \, (\mathbf{S}^{-1})_{kl} \, \langle l|n'\rangle \,. \tag{15}$$

Here, $\langle m'|k\rangle$ and $\langle l|n'\rangle$ are derivatives of the orbital overlaps. They can be determined in the same way as the gradient of the Hamiltonian matrix elements in Eq. (13), i.e., through finite differences. The overlap matrix is calculated with the MLWFs obtained in Eq. (8). Wannier function perturbation theory (WFPT) [48] appears as a potential alternative approach to determine the second part of Eq. (12). Initially developed to integrate Wannier functions with linear response via DFPT, WFPT could possibly be extended to produce the quantities required in our formalism.

### 4. Phonon frequency and vibrational modes

The phonon frequencies and modes are critical ingredients of the electron-phonon coupling in Eq. (7). They can be obtained from the Hamiltonian $\hat{\mathbf{H}}_{ph}^0$ in Eq. (1), known as the dynamical matrix $\mathbf{\Phi}(\mathbf{q})$. Similar to the electronic Hamiltonian in Eq. (10), $\mathbf{\Phi}(\mathbf{q})$ is the Fourier transform of the real-space dynamical matrix, which is often referred to

as the mass-scaled interatomic force constants $C_{I\eta J\xi}$:

$$\Phi_{I\eta J\xi}(\mathbf{q}) = \sum_{\mathbf{R}} C_{I\eta J\xi}(\mathbf{R}) e^{i\mathbf{q}\cdot\mathbf{R}}, \tag{16}$$

where $\mathbf{R}$ points to the unit cell in which atom $J$ is located. $C_{I\eta J\xi}$ represents the derivative of the Born-Oppenheimer total energy surface $E(\mathbf{R})$ with respect to the atomic displacements $Q_{I\eta}$ and $Q_{J\xi}$, defined as in Eq. (13) [18],

$$C_{I\eta J\xi} = \frac{1}{\sqrt{M_I M_J}} \frac{\partial^2 E(\mathbf{R})}{\partial Q_{I\eta} \partial Q_{J\xi}}. \tag{17}$$

Since the force $F_{I\eta}$ acting on an atom $I$ along $\eta$ is readily available as an output of most electronic structure codes and because $\partial^2 E(\mathbf{R})/\left(\partial Q_{I\eta} \partial Q_{J\xi}\right) = -\partial F_{I\eta}/\partial Q_{J\xi}$, we can approximate the force constants as

$$C_{I\eta J\xi} \approx -\frac{1}{\sqrt{M_I M_J}} \frac{[F_{I\eta}(Q_{J\xi}) - F_{I\eta}(-Q_{J\xi})]}{2Q_{J\xi}}. \tag{18}$$

This so-called frozen-phonon approach allows for the simultaneous calculation of the interatomic force constants and Hamiltonian derivatives, as described in Sec. II A 3. To avoid numerical inaccuracies related to electronic structure calculations, we enforce the acoustic sum rule and apply a numerical symmetrization of $C_{I\eta J\xi}$ [49]. The PHONOPY software package lends itself optimally to the construction of the displaced supercells and the subsequent calculation of the interatomic force constants [50]. Although PHONOPY is capable of incorporating corrections to the produced dynamical matrix arising from long-range electrostatics, this feature has not been included here. As for the real-space MLWF Hamiltonian and the electron-phonon coupling elements, the force constants can be remapped into a larger device transport cell from the original DFT primitive unit cell. Finally, the ion displacement vectors $f_{\lambda\mathbf{q}}^{I\eta}$ (phonon mode) and frequency $\omega_{\lambda\mathbf{q}}$ are calculated by solving the following eigenvalue problem involving the dynamical matrix $\Phi_{I\eta J\xi}(\mathbf{q})$:

$$\sum_{J\xi} \Phi_{I\eta J\xi}(\mathbf{q}) f_{\lambda\mathbf{q}}^{J\xi} - \omega_{\lambda\mathbf{q}}^2 f_{\lambda\mathbf{q}}^{I\eta} = 0. \tag{19}$$

### B. Electronic transport

This section briefly summarizes the two transport methods used in this work, LBTE and NEGF. We show how the electron-phonon coupling elements that were derived in the previous section can be inserted into both formalisms to calculate phonon-limited mobility and extract electrical currents in the case of NEGF.

### 1. LBTE

In a uniform system, under a constant time-independent electric field, without any magnetic field, at steady state,

and in the presence of electron-phonon scattering, the Boltzmann transport equation (BTE) [51–53] reduces to

$$-\frac{e\mathbf{E}}{\hbar}\frac{\partial f_{i\mathbf{k}}}{\partial \mathbf{k}} = \sum_{j\mathbf{k}'\lambda\mathbf{q}} f_{j\mathbf{k}'}\left(1-f_{i\mathbf{k}}\right)\Gamma_{ji\mathbf{k}'}^{\lambda\mathbf{q}} - f_{i\mathbf{k}}\left(1-f_{j\mathbf{k}'}\right)\Gamma_{ij\mathbf{k}}^{\lambda\mathbf{q}}.$$
(20)

Here, $e$ stands for the elementary charge, $\hbar$ represents the reduced Planck's constant, and $\mathbf{E}$ denotes the electric field. The term $f_{i\mathbf{k}}$ ($f_{j\mathbf{k}'}$) refers to the electronic distribution function for a state $|i\mathbf{k}\rangle$ ($|j\mathbf{k}'\rangle$) with band index $i$ ($j$) and momenta $\mathbf{k}$ ($\mathbf{k}'$). The scattering rate, mediated by a phonon mode $\lambda$ and wave vector $\mathbf{q}$, is given by $\Gamma_{ij\mathbf{k}}^{\lambda\mathbf{q}}$ ($\Gamma_{ji\mathbf{k}'}^{\lambda\mathbf{q}}$). In the case of a weak electric field, as considered here, the solution of $f_{i\mathbf{k}}$ can be restricted to its linear response: $f_{i\mathbf{k}} \approx f_{i\mathbf{k}}^0 + \frac{\partial f_{i\mathbf{k}}^0}{\partial E_{i\mathbf{k}}} e\mathbf{E}\cdot\mathbf{F}_{i\mathbf{k}}$, where $f_{i\mathbf{k}}^0$ is the equilibrium Fermi distribution function and $\mathbf{F}_{i\mathbf{k}}$ the carrier mean free displacement vector [54]. It represents the quantity to be calculated. By applying the linearization of Eq. (20), together with the conditions of detailed balance [55] and momentum conservation, we arrive at the LBTE [56]:

$$\mathbf{v}_{i\mathbf{k}} = \sum_{j\lambda\mathbf{q}} \frac{1-f_{j\mathbf{k+q}}^0}{1-f_{i\mathbf{k}}^0}\Gamma_{ij\mathbf{k}}^{\lambda\mathbf{q}}\left(\mathbf{F}_{i\mathbf{k}}-\mathbf{F}_{j\mathbf{k+q}}\right),$$
(21)

where $\mathbf{v}_{i\mathbf{k}}$ is the carrier velocity of state $|i\mathbf{k}\rangle$.

The electronic state $|i\mathbf{k}\rangle$ with energy $E_{i\mathbf{k}}$ is a solution of the eigenvalue problem involving the Hamiltonian defined in Eq. (10),

$$\sum_n H_{mn}(\mathbf{k})c_{i\mathbf{k}}^n - E_{i\mathbf{k}}c_{i\mathbf{k}}^m = 0,$$
(22)

where the $c_{i\mathbf{k}}^n$ are the expansion coefficients for the electron wave function. The scattering rate $\Gamma_{ij\mathbf{k}}^{\lambda\mathbf{q}}$ from an initial electronic state $|i\mathbf{k}\rangle$ into a final state $|j\,k\pm q\rangle$ caused by phonon absorption $(+)$ or emission $(-)$ can be calculated with Fermi's golden rule [57]:

$$\Gamma_{ij\mathbf{k}}^{\lambda\mathbf{q}} = \frac{2\pi}{\hbar}\left|g_{ij}^\lambda(\mathbf{k},\mathbf{q})\right|^2 \left[N_{\lambda\mathbf{q}}^0\delta\left(E_{i\mathbf{k}}+\hbar\omega_{\lambda\mathbf{q}}-E_{j\mathbf{k+q}}\right)+\right.$$
$$\left.\left(1+N_{\lambda-\mathbf{q}}^0\right)\delta\left(E_{i\mathbf{k}}-\hbar\omega_{\lambda-\mathbf{q}}-E_{j\mathbf{k+q}}\right)\right].$$
(23)

Here, $N_{\lambda\mathbf{q}}^0$ is the equilibrium Bose-Einstein distribution, for which it holds that $N_{\lambda-\mathbf{q}}^0 = N_{\lambda\mathbf{q}}^0$ and $\omega_{\lambda-\mathbf{q}} = \omega_{\lambda\mathbf{q}}$. The probability amplitude of the scattering process $g_{ij}^\lambda(\mathbf{k},\mathbf{q})$ is obtained from the electron-phonon coupling elements in Eq. (7) as

$$g_{ij}^\lambda(\mathbf{k},\mathbf{q}) = \sum_{mn\eta} c_{i\mathbf{k+q}}^{m\dagger} M_{mn}^{\lambda\eta}(\mathbf{k},\mathbf{q})c_{j\mathbf{k}}^n.$$
(24)

Equation (21) forms a linear system of equations the size of which rapidly explodes as the number of $\mathbf{k}/\mathbf{q}$ points

increases. Given the necessity for dense Brillouin zone sampling to ensure precise results, solving the LBTE becomes a computationally demanding task. To address this issue, an iterative approach can be employed. We first consider a simpler solution where we neglect the $\mathbf{F}_{j\mathbf{k+q}}$ term in Eq. (21). This is referred to as the energy-relaxation-time approximation (ERTA) [51,52], which leads to a simple solution for Eq. (21), namely, $\mathbf{F}_{i\mathbf{k}}^{\mathrm{ERTA}} = \mathbf{v}_{i\mathbf{k}}\tau_{i\mathbf{k}}$, with the generalized relaxation time $\tau_{i\mathbf{k}}$ defined as [24,58]

$$\frac{1}{\tau_{i\mathbf{k}}} = \frac{2\pi}{\hbar}\sum_{j\lambda\mathbf{q}}\left|g_{ij}^\lambda(\mathbf{k},\mathbf{q})\right|^2 \cdot$$
$$\left[\left(N_{\lambda\mathbf{q}}^0 + f_{j\mathbf{k+q}}^0\right)\delta\left(E_{i\mathbf{k}}+\hbar\omega_{\lambda\mathbf{q}}-E_{j\mathbf{k+q}}\right)+\right.$$
$$\left.\left(1+N_{\lambda-\mathbf{q}}^0 - f_{j\mathbf{k-q}}^0\right)\delta\left(E_{i\mathbf{k}}-\hbar\omega_{\lambda-\mathbf{q}}-E_{j\mathbf{k+q}}\right)\right].$$
(25)

The exact solution of Eq. (21) can then be iteratively computed with $\mathbf{F}_{i\mathbf{k}}^{\mathrm{ERTA}}$ as an initial guess, i.e., $\mathbf{F}_{i\mathbf{k}}^0 \equiv \mathbf{F}_{i\mathbf{k}}^{\mathrm{ERTA}}$ [54]. The iterative procedure obeys the following equation:

$$\mathbf{F}_{i\mathbf{k}}^{\xi+1} = \mathbf{F}_{i\mathbf{k}}^0 + \tau_{i\mathbf{k}}\sum_{j\lambda\mathbf{q}}\frac{1-f_{j\mathbf{k+q}}^0}{1-f_{i\mathbf{k}}^0}\Gamma_{ij\mathbf{k}}^{\lambda\mathbf{q}}\mathbf{F}_{j\mathbf{k+q}}^\xi.$$
(26)

Convergence is achieved when the low-field conductivity,

$$\sigma_{\alpha\beta}^\xi = \frac{2e^2}{Vk_BT}\sum_{i\mathbf{k}}f_{i\mathbf{k}}^0\left(1-f_{i\mathbf{k}}^0\right)F_{i\mathbf{k}}^{\alpha\xi}v_{i\mathbf{k}}^\beta,$$
(27)

does not change by more than a predefined criterion between two consecutive iterations, ($\xi + 1$ and $\xi$). In Eq. (27), $\alpha$ and $\beta$ run over the three Cartesian directions, $k_B$ is Boltzmann's constant, $T$ is the temperature, and the factor of 2 accounts for spin degeneracy. Here, we restrict ourselves to the case in which the transport direction aligns with the electric field direction $\eta$. In this configuration, the conductivity $\sigma$ becomes a scalar and it corresponds to $\sigma_{\eta\eta}$. Finally, once the iterative solution has converged, the phonon-limited drift mobility $\mu$ is given by

$$\mu = \frac{\sigma}{en},$$
(28)

where $n$ is the electron density. It it calculated according to the equilibrium Fermi distribution function,

$$n = \frac{2}{V}\sum_{n\mathbf{k}}f_{n\mathbf{k}}^0.$$
(29)

While the iterative mobility (IBTE) provides the most accurate solution of the LBTE, it can also be of interest to calculate the mobility within the ERTA approximation

for comparison with NEGF and other published works. To do this, $F_{i\mathbf{k}}^{\alpha\xi}$ in Eq. (27) is replaced with $F_{i\mathbf{k}}^{\alpha\text{ERTA}}$.

The carrier velocity, which enters Eq. (27), is defined as [56]

$$\mathbf{v}_{i\mathbf{k}} = \frac{1}{\hbar} \sum_{mn} c_{i\mathbf{k}}^{m\dagger} \left[ \frac{\partial H_{mn}(\mathbf{k})}{\partial \mathbf{k}} \right] c_{i\mathbf{k}}^{n}, \tag{30}$$

where the $\mathbf{k}$ derivative of the Hamiltonian in Eq. (10) can be carried out analytically [59]:

$$\frac{\partial H_{mn}(\mathbf{k})}{\partial k_\eta} = \sum_{\mathbf{R}} i R_\eta e^{i\mathbf{k}\cdot\mathbf{R}} H_{mn}(\mathbf{R}). \tag{31}$$

In Eq. (31), $\eta$ indicates the direction along which the derivative is calculated.

Note that the delta function $\delta\left(E_{i\mathbf{k}} \pm \hbar\omega_{\lambda\pm\mathbf{q}} - E_{j\mathbf{k+q}}\right)$ in Eq. (23) is approximated as the limit of a Gaussian function whose smearing parameter approaches zero. We use an adaptive smearing method [24,60] to determine the best-suited value of this parameter for a given $\mathbf{k}/\mathbf{q}$ grid.

The LBTE transport solver used in this study has been implemented as a stand-alone C $+$ $+$ software tool.

### 2. NEGF

Next, we move beyond material properties under equilibrium conditions and target device simulation, where a system is driven out of equilibrium by the application of an external voltage. For this, the NEGF formalism has established itself as a robust computational methodology that is capable of addressing both coherent and dissipative quantum transport [61–65]. Specifically, this work uses NEGF to model transistors with a 2D channel material including electron-phonon scattering through the self-consistent Born approximation [35,56,66,67]. In 2D systems, as in the monolayer MoS$_2$ FET depicted in Fig. 2, electron transport occurs along one principal axis, designated here as the $x$ axis. Perpendicular to this, the $y$ axis serves as the direction of confinement, while the $z$ axis is assumed to be periodic. In such a configuration, the following system of equations must be solved to obtain all desired observables:

$$\sum_l \left\{ [\mathbf{E} - \mathbf{V}(\mathbf{r_m})]\delta_{lm} - \mathbf{H}_{ml}(k_z) - \mathbf{\Sigma}_{ml}^{RB}(E, k_z) - \right.$$
$$\left. \mathbf{\Sigma}_{ml}^{RS}(E, k_z) \right\} \mathbf{G}_{ln}^{R}(E, k_z) = \delta_{mn}, \tag{32}$$

$$\mathbf{G}_{mn}^{\gtrless}(E, k_z) = \sum_{l_1 l_2} \mathbf{G}_{ml_1}^{R}(E, k_z) \left[ \mathbf{\Sigma}_{l_1,l_2}^{\gtrless B}(E, k_z) + \right.$$
$$\left. \mathbf{\Sigma}_{l_1,l_2}^{\gtrless S}(E, k_z) \right] \mathbf{G}_{l_2 n}^{A}(E, k_z). \tag{33}$$

The indices $l$, $m$, and $n$ refer to the atomic positions $\mathbf{r}_l$, $\mathbf{r}_m$, and $\mathbf{r}_n$, respectively. The diagonal matrices $\mathbf{E}$ and $\mathbf{V}(\mathbf{r}_m)$
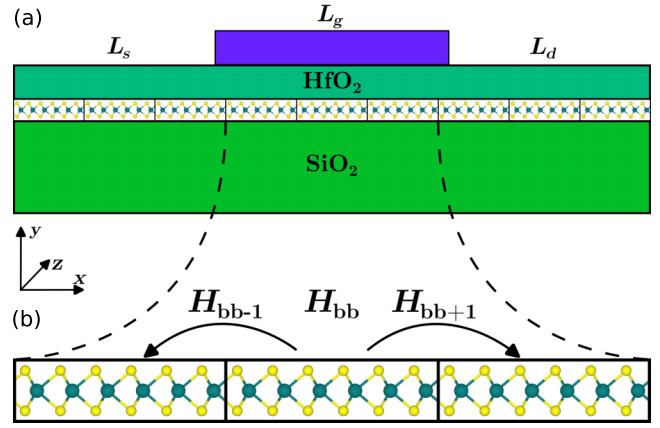


FIG. 2. (a) Schematic of a 2D single-gate transistor structure featuring a MoS$_2$ monolayer as the channel region. The lengths of the source, drain, and gate are denoted as $L_s$, $L_d$, and $L_g$, respectively. (b) Detailed view of the channel region, which is divided into orthorhombic unit cells. The corresponding Hamiltonian blocks are indicated; block $b$ is connected with itself ($\mathbf{H}_{bb}$) as well as its next- ($\mathbf{H}_{bb+1}$) and previous- ($\mathbf{H}_{bb-1}$) neighbor blocks.

contain the injection energy $E$ and the self-consistent electrostatic potential $V$ at position $\mathbf{r}_m$. The size of both matrices is $p_m \times p_m$, where $p_m$ is the number of Wannier functions located on the $m^{\text{th}}$ atom. The block Hamiltonian $\mathbf{H}_{mn}(k_z)$ is expressed in the MLWF basis constructed according to Eq. (11). Its entries are the matrix elements between atoms with index $m$ and $n$ and its size is $p_m \times p_n$. The scattering $\mathbf{\Sigma}_{mn}^{S}(E, k_z)$ and boundary $\mathbf{\Sigma}_{mn}^{B}(E, k_z)$ self-energies together with the retarded, $\mathbf{G}_{mn}^{R}(E, k_z)$, advanced, $\mathbf{G}_{mn}^{A}(E, k_z) = \mathbf{G}_{mn}^{R\dagger}(E, k_z)$, lesser, $\mathbf{G}_{mn}^{<}(E, k_z)$, and greater, $\mathbf{G}_{mn}^{>}(E, k_z)$, Green's functions are also of size $p_m \times p_n$. These quantities must be solved for each injection energy $E$ and transverse momentum $k_z$.

The boundary self-energies [66], which are different from zero only when atoms $m$ and $n$ are directly connected to the device contacts, can be computed through decimation techniques [68] or eigenvalue problems [69]. Their scattering counterparts, $\mathbf{\Sigma}_{mn}^{S}(E, k_z)$, are full matrices and consist here of electron-phonon interactions. They are obtained via perturbation theory by expanding the Green's function to the second order in the electron-phonon interaction Hamiltonian in Eq. (6), applying Wick's theorem, and writing down the corresponding Dyson equation [41, 61,70]. The resulting general form of the electron-phonon scattering self-energy in the time domain is

$$\mathbf{\Sigma}_{mn}^{e\text{-ph}}(tt', k_z) = i\hbar \sum_{l_1 l_2 \eta_1 \eta_2} \sum_{\lambda q_z} \mathbf{M}_{ml_1}^{\lambda\eta_1}(k_z - q_z, q_z) \cdot$$
$$\mathbf{G}_{l_1 l_2}(tt', k_z - q_z) \cdot \mathbf{M}_{l_2 n}^{\lambda\eta_2}(k_z, -q_z)\mathbf{D}^{\lambda}(tt', q_z). \tag{34}$$

In this expression, the elements of the electron-phonon coupling blocks $\mathbf{M}_{mn}^{\lambda\eta}(\mathbf{k},\mathbf{q})$ are defined as in Eq. (7) and $\mathbf{D}^{\lambda}(\mathbf{q},tt')$ represents the phonon Green's function. Applying the Langreth theorem, replacing the phonon Green's function by its unperturbed form at equilibrium, and assuming steady-state conditions ($tt' \to t - t'$), which allows us to Fourier transform the time difference into the energy domain, we obtain the energy-dependent greater and lesser self-energy expressions [41,71]

$$
\begin{aligned}
\mathbf{\Sigma}_{mn}^{\gtrless e\text{-ph}}(E,k_z) = \sum_{l_1 l_2 \eta_1 \eta_2} \sum_{\lambda q_z} \mathbf{M}_{ml_1}^{\lambda\eta_1}(k_z - q_z, q_z) \cdot \\
\{N_{\lambda q_z}^0 \mathbf{G}_{l_1 l_2}^{\gtrless}(E \pm \hbar\omega_{\lambda q_z}, k_z - q_z) + \\
[N_{\lambda q_z}^0 + 1]\mathbf{G}_{l_1 l_2}^{\gtrless}(E \mp \hbar\omega_{\lambda q_z}, k_z - q_z)\} \cdot \\
\mathbf{M}_{l_2 n}^{\lambda\eta_2}(k_z, -q_z),
\end{aligned} \tag{35}
$$

where $N_{\lambda q_z}^0$ is the equilibrium Bose-Einstein distribution of phonons with frequency $\omega_{\lambda q_z}$. By employing the greater and lesser electron-phonon scattering self-energies $\mathbf{\Sigma}_{mn}^{\gtrless e\text{-ph}}(E,k_z)$, we can derive an expression for the retarded scattering self-energy, $\mathbf{\Sigma}_{mn}^{RS}(E,k_z)$, in Eq. (32):

$$
\begin{aligned}
\mathbf{\Sigma}_{mn}^{RS}(E,k_z) = \frac{1}{2}\left(\mathbf{\Sigma}_{mn}^{> e\text{-ph}}(E,k_z) - \mathbf{\Sigma}_{mn}^{< e\text{-ph}}(E,k_z)\right) + \\
i\mathcal{P}\int \frac{dE'}{2\pi} \frac{(\mathbf{\Sigma}_{mn}^{> e\text{-ph}}(E,k_z) - \mathbf{\Sigma}_{mn}^{< e\text{-ph}}(E',k_z))}{E - E'},
\end{aligned} \tag{36}
$$

where $\mathcal{P}$ denotes the Cauchy principal integral [72].

Because of the interdependence between $\mathbf{G}_{mn}^{\gtrless}(E,k_z)$ and $\mathbf{\Sigma}_{mn}^{\gtrless e\text{-ph}}(E,k_z)$, Eqs. (32)–(33) and Eqs. (35)–(36) must be solved self-consistently within the Born approximation. The convergence of the loop between the Green's functions and the scattering self-energies is verified by monitoring the variations of the current $I_d$ and carrier density $n(\mathbf{r}_l)$ between consecutive iterations [44,67]. The said current $I_d$, calculated as

$$
\begin{aligned}
I_{d,b\to b+1} = \\
\frac{2e}{\hbar}\sum_{k_z}\sum_{m\in b}\sum_{n\in b+1}\int\frac{dE}{2\pi}\mathrm{tr}\left\{\mathbf{H}_{mn}(k_z)\mathbf{G}_{nm}^{<}(E,k_z) - \right. \\
\left. \mathbf{G}_{mn}^{<}(E,k_z)\mathbf{H}_{nm}(k_z)\right\},
\end{aligned} \tag{37}
$$

represents the flow of electrons from a block of atoms in the orthorhombic cell $b$ to its neighboring cell $b+1$. The position indices, $m$ and $n$, therefore encompass all atoms within the connected cells, as exemplified in Fig. 2. The trace operator "tr" runs over the orbitals of atoms $m$ and $n$. Finally, the factor of 2 accounts for spin. The carrier

density,

$$
n(\mathbf{r}_l) = -2i\sum_{k_z}\int\frac{dE}{2\pi}tr\left\{\mathbf{G}_{ll}^{<}(E,k_z)\right\}, \tag{38}
$$

at atomic position $\mathbf{r}_l$ is the trace of the lesser Green's function block corresponding to this atom.

As in the LBTE case, the phonon-limited mobility can be derived from the electrical conductivity $\sigma$ through Eq. (28). This can be done with the $dR/dL$ method [73], where the conductivity $\sigma$ is obtained by extracting the channel resistivity $\rho$

$$
\rho = \frac{dR}{dL} = \frac{\Delta V}{\Delta L}\left(\frac{1}{I_d(L + \Delta L)} - \frac{1}{I_d(L)}\right). \tag{39}
$$

The $dR/dL$ method requires calculation of the drain current $I_d(L)$ for samples of different lengths $L$. It is essential for $L$ to be sufficiently large to guarantee that the transport occurs within the diffusive regime, i.e., $L$ must be larger than the phonon-limited mean free path for scattering [74]. The simulations are carried out on structures that have a uniform charge density $n$, a small bias difference $\Delta V = 1$ mV between the contacts, and that include electron-phonon scattering.

To minimize the computational burden associated with the investigation of devices containing a large number of atoms, approximations must be applied to the calculation of the Green's functions and scattering self-energies. First, to obtain the retarded scattering self-energy in Eq. (36), the Cauchy principal integral term coupling all energies together is neglected. Since it only contributes to energy renormalization and not to relaxation or phase breaking events, leaving it out is not expected to significantly alter the device behavior, as demonstrated in previous studies [72,75]. Second, the phonon momentum points $\mathbf{q}$ are restricted to the $\Gamma$ point ($\mathbf{q} = \mathbf{0}$) of the orthorhombic transport cell. This simplification allows us to perform the sum over the displaced atoms ($I$ and $\gamma$), directly in Eq. (7). Consequently, the electron-phonon coupling element "only" depends on 5 ($\lambda, \mathbf{k}, m, n, \eta$) instead of 8 ($I, \gamma, \mathbf{q}, \lambda, \mathbf{k}, m, n$, and $\eta$) variables, thus significantly reducing the memory requirements and the time needed to evaluate Eq. (35). To compensate for the omission of the $\mathbf{q}$ dependence, a large orthorhombic transport cell is chosen ($R_z = 2.5$ nm for $MoS_2$ in Fig. 1), so that most band-structure features are folded back to the $\Gamma$ point. Note that while the number of phonon momentum points is limited, all phonon modes are taken into account individually, preserving the coupling to all phonon energies $\hbar\omega_{\lambda\mathbf{q}}$ at the $\Gamma$ point. Finally, it is not possible to treat the scattering self-energies as full matrices in large structures with thousands of atoms. A cutoff radius $r_{\text{cut}}$ is therefore introduced in Eq. (35) so that $|r_n - r_m| < r_{\text{cut}}$, as illustrated in Fig. 3. A convergence analysis with respect to $r_{\text{cut}}$ and comparisons with
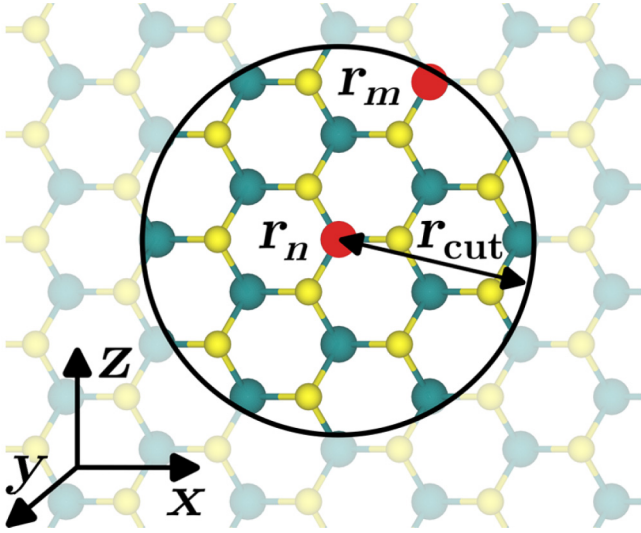
FIG. 3. Top view of a $MoS_2$ monolayer, illustrating the interaction range considered when constructing the scattering self-energies $\Sigma^{e\text{-}ph}_{mn}(E, k_z)$. Interactions between atoms $n$ and $m$ (highlighted in red) are included if their distance is smaller than a predefined cutoff radius $r_{cut}$—i.e., if $|r_n - r_m| < r_{cut}$. For distances exceeding this cutoff radius, the self-energy entries $\Sigma^{e\text{-}ph}_{mn}(E, k_z)$ are set to zero.

LBTE (see Sec. III and the Appendix) indicate that such an approach is sufficient to obtain meaningful results. No scaling of the electron-phonon interactions is needed, contrary to what was done in Ref. [36]. If all scattering events are treated properly, i.e., $r_{cut}$ is large enough so that all important interactions are captured and if the unit cell size is large enough so that $\Gamma$-point sampling is adequate, the NEGF mobility corresponds to that obtained with the IBTE approach. Despite all these simplifications and although the calculation of Eq. (35) has been ported to graphics processing units (GPUs), the computational intensity of such simulations remains gigantic, thus limiting the number of samples that could be considered.

The electron-phonon coupling approach used in this study has been implemented in the already existing NEGF transport solver OMEN [35].

### III. SIMULATION RESULTS

In this section, our simulation results are presented, starting with the lattice and band structure of selected 2D materials and silicon. We then analyze their mobility, using both LBTE and NEGF ($MoS_2$, $WS_2$, and $WSe_2$). Finally, we calculate the electrical current flowing through a $MoS_2$ transistor and assess the importance of employing an accurate electron-phonon scattering model.

Each electronic structure calculation is performed within the generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (PBE) [76], using projector augmented-wave (PAW) pseudopotentials [77,78], as

implemented in the VASP software [15]. A plane-wave cutoff energy of 500 eV is applied, while the total-energy-convergence criterion is set to less than $10^{-10}$ eV. The unit cells are relaxed to ensure that the residual forces acting on each atom are below $10^{-8}$ eV/Å. During this process, both the lattice constant and atomic positions are allowed to change. Gaussian smearing with a smearing width of 0.05 eV is employed. For the TMDCs, the electronic structure calculations are performed on a $27 \times 27 \times 1$ Monkhorst-Pack **k**-point grid, while a $15 \times 15 \times 15$ grid is used for silicon. Spin-orbit coupling (SOC) is neglected. In the case of TMDCs, a vacuum space of 20 Å is used along the $y$ axis (the direction of confinement in Fig. 1) to minimize interactions with periodic images. Displacement calculations are performed with $9 \times 9 \times 1$ TMDC supercells and $5 \times 5 \times 5$ silicon supercells. A displacement distance of 0.01 Å is employed to calculate all derivatives (Hamiltonian and forces). The Wannier Hamiltonians are obtained from the DFT results by considering 5 $d$-like orbitals on the transition-metal atoms and 3 $sp^3$-like orbitals on each of the chalcogen atoms, while $sp^3$-like orbitals are retained for each silicon atom as initial guesses for the minimization procedure. During the Wannier optimization process, a frozen-energy window [43], located 1 eV above the conduction-band edge and 1 eV below the valence band edge, is defined to ensure that all states relevant for transport calculations are accurately represented by the tight-binding-like Wannier Hamiltonian.

### A. Lattice and electronic band structures

It has been observed that small changes in simulation parameters—in particular, DFT settings—can lead to significant differences in calculated mobilities [29]. It is therefore important to precisely report the crystal lattice and electronic band structures that enter the developed modeling framework to ensure meaningful comparisons with other studies.

Starting with the crystal structure, the lattice constants and the layer thickness of the considered TMDCs, obtained through lattice relaxation, are summarized in Table I. The layer thickness is measured as the distance between the top and bottom chalcogen atoms. For $MoS_2$, our calculated lattice constant is equal to 3.183 Å, which agrees well with values reported by other groups utilizing VASP [79,80] or QUANTUM ESPRESSO [81]. A similar agreement is observed for the three other TMDCs analyzed here. It should, however, be noted that several studies based either on plane-wave or local-orbital methods report somewhat smaller lattice constants, ranging between 3.130 and 3.140 Å for $MoS_2$ [23,25,26,30–32]. The experimentally determined lattice constant of bulk $MoS_2$ is positioned in between and is equal to 3.15 Å [82]. For the remaining TMDCs, we calculate lattice constants of 3.318 Å for $MoSe_2$, 3.182 Å for $WS_2$, and 3.316 Å for $WSe_2$.

TABLE I. The lattice and band-structure parameters of selected TMDCs and silicon. The typical range of values reported in the literature is provided in parentheses. The lattice constant and distance between the upper and lower chalcogen atoms in the 2D TMDC layers are reported in the second and third columns, respectively. The variable $E_{BG}$ refers to the electronic band gap and $\Delta E_{KQ}$ to the energy separation between the $K$ and $Q$ valleys located in the conduction band of the TMDCs.

| Material | Lattice constant (Å) | Layer thickness (Å) | $E_{BG}$ (eV) | $\Delta E_{KQ}$ (meV) |
|---|---|---|---|---|
| $MoS_2$ | 3.183 (3.130–3.187) | 3.127 (3.097–3.177) | 1.67 (1.60–1.89) | 269 (70–310) |
| $MoSe_2$ | 3.318 (3.264–3.318) | 3.337 (3.337–3.342) | 1.44 (1.43–1.59) | 169 (18–169) |
| $WS_2$ | 3.182 (3.100–3.199) | 3.139 (3.131–3.166) | 1.81 (1.80–1.99) | 211 (22–211) |
| $WSe_2$ | 3.316 (3.250–3.317) | 3.354 (3.348–3.365) | 1.54 (1.54–1.71) | 131 (16–131) |
| Si | 5.430 (5.430) | $\cdots$ $\cdots$ | 0.57 $\cdots$ | $\cdots$ $\cdots$ |

Similar to $MoS_2$, we find a range of values reported in the literature, e.g., 3.264-3.318 Å for $MoSe_2$ [30–32,79–81] and 3.250–3.317 Å for $WSe_2$ [30–32,79–81]. The range is broadest for $WS_2$, with reported lattice constants from 3.100 to 3.199 Å [30–32,79–81]. Matching lattice parameters do not necessarily translate into identical layer thicknesses, which directly influence the vibrational modes of the material. For instance, a comparison of our results with those of Ref. [81] reveals a discrepancy of only 0.13% in the lattice constant of $MoS_2$, whereas the layer thickness diverges by 1.6%, a one-order-of-magnitude-larger difference. The reported layer thicknesses include 3.097–3.177 Å for $MoS_2$ [30,31,79–81], 3.337–3.342 Å for $MoSe_2$ [30,31,79–81], 3.131–3.166 Å for $WS_2$ [30,31, 79–81], and 3.348–3.365 Å for $WSe_2$ [30,31,79–81].

In the case of silicon, our calculations produce a lattice constant of 5.43 Å. This result is in agreement with the experimentally measured value of this semiconductor [83] and in the range of other simulation works; e.g., Ref. [58] has found values between 5.40 and 5.47 Å, depending on the exchange and correlation functional used in the DFT calculations.

Moving on to the electron and phonon dispersion, in Fig. 4, we plot the electronic and phonon band structures of the selected monolayer TMDCs, while in Table I, we list the extracted band gaps ($E_{BG}$) and the energy separation between the $K$ and $Q$ valleys ($\Delta E_{KQ}$). The exact location of the $Q$ valley is marked along the high-symmetry path of the electronic band structure. The accuracy of the Wannier transformations is demonstrated by comparing the MLWF and DFT band structures. All monolayer TMDCs are found to be direct-gap semiconductors, as expected, with band gaps ranging between 1.44 and 1.81 eV. Focusing on $MoS_2$, we obtain a band gap of 1.67 eV, in agreement with Refs. [79–81] but slightly below the experimental band gap of 1.8 eV [84]. Several simulation studies have

reported band gaps closer to that of experimental measurements [20,23,25,26,30–32], but also a smaller lattice constant. A previous study [35] using the same simulation code, pseudopotentials, and exchange and correlation functional as here obtained a band-gap value close to that of the experiments by fixing the lattice constant to that of bulk $MoS_2$. This trick also resulted in a reduction of $\Delta E_{KQ}$, a critical parameter for transport, to 48 meV, far below the 269 meV of the present study. It should, however, be noted that the band gap of $MoS_2$ remains an open issue. The accepted value of 1.8 eV corresponds to the optical band gap. The latter can be reproduced by DFT simulations when introducing GW corrections and accounting for excitonic effects [85]. Reported band gaps for the other TMDCs include 1.43–1.59 eV for $MoSe_2$ [31,79–81], 1.80–1.99 eV for $WS_2$ [31,79–81], and 1.54–1.71 eV for $WSe_2$ [31,79–81].

The $K$-$Q$ valley separation, $\Delta E_{KQ}$, in contrast to the band gap, directly influences the mobility of TMDCs through intervalley scattering. While we will delve deeper into this effect in Sec. III B, preliminary observations are made here. As already mentioned, $\Delta E_{KQ}$ is highly sensitive to even small variations in the simulation parameters, giving rise to a large range of reported values in the literature: 70–310 meV for $MoS_2$ [23–32,79,81], 18–169 meV for $MoSe_2$ [29,31,32,79,81], 22–211 meV for $WS_2$ [28,29,31,32,79,81], and 16–131 meV for $WSe_2$ [28,29, 31,32,79,81]. It can generally be inferred that methods employing the local density approximation (LDA) tend to predict a smaller valley separation than GGA PBE. Sensitivity to spin-orbit coupling has been shown to be TMDC dependent. For instance, in $WSe_2$, the value of $\Delta E_{KQ}$ drops from 117 meV to 41 meV when spin-orbit coupling is taken into account. For $MoSe_2$, the inclusion of spin-orbit coupling does not significantly affect $\Delta E_{KQ}$ [29].
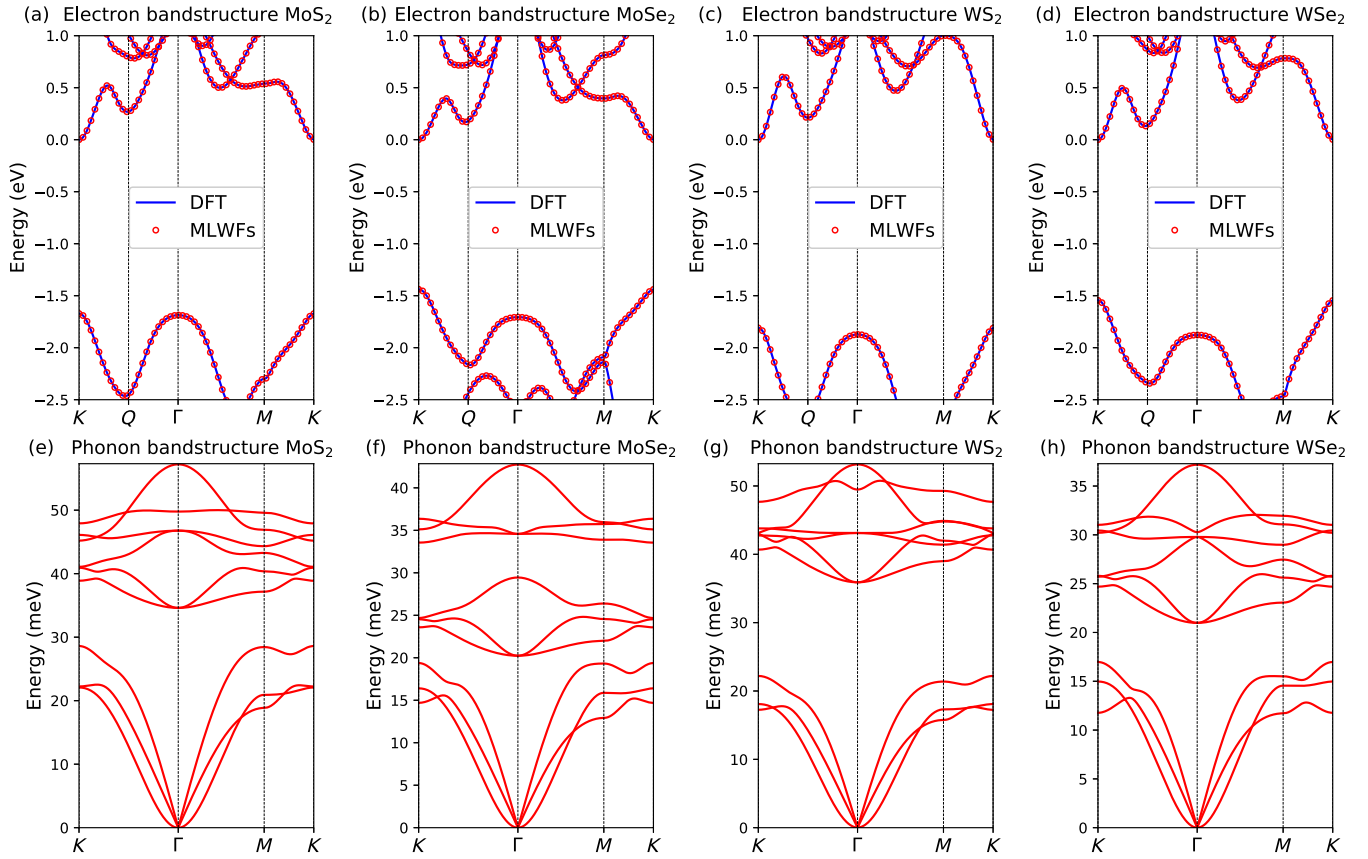
FIG. 4.   Electronic (a)–(d) and phonon (e)–(h) band structure of monolayer $MoS_2$, $MoSe_2$, $WS_2$, and $WSe_2$, respectively. In (a)–(d), the plane-wave DFT (solid lines) and MLWF (circles) band structures are compared to each other. The position of the second conduction-band minimum at $Q$ is indicated along the high-symmetry path. The zero energy level is set to the conduction-band minimum of each material.

When it comes to phonons, our results generally agree with studies reporting similar lattice parameters and based on GGA PBE, such as those in Refs. [29,81]. The same level of agreement remains when different pseudopotentials are used, with and without spin-orbit coupling. The latter effect is expected to be important for materials with a net magnetic moment [29]. We do note some small differences around the $\Gamma$ point of the upper optical bands of $WS_2$ when comparing to calculations performed with DFPT [29,81].

The calculated electron and phonon band structures of silicon are depicted in Fig. 5 and the corresponding band gap is given in Table I. As in the case of the TMDCs, the calculated band gap (0.57 eV) underestimates its experimental counterpart (1.12 eV) [83]. The silicon band structure is nevertheless consistent with other theoretical works [58,86,87] and its overall shape is in good agreement with the experimental measurements [88]. Improving the band-gap accuracy would require applying hybrid functionals or GW corrections, which is computationally very expensive and not necessary to extract mobility values.

## B. Phonon-limited mobility

### 1. LBTE

The mobility calculations were carried out with the iterative LBTE method, utilizing homogeneous and commensurate **k**/**q**-point grids across the entire Brillouin zone.
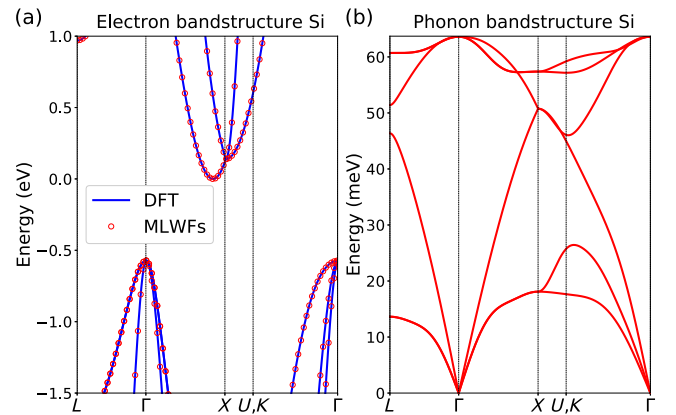


FIG. 5.   (a) Electronic and (b) phonon band structure of silicon. The same plotting conventions as in Fig. 4 are used.

TABLE II.   Transport properties of the selected TMDCs and silicon. The typical range of values reported in the literature is provided in parentheses. The intrinsic mobility ($\mu_{\text{int}}$), the mobility at an electron concentration of $5 \times 10^{13}$ cm$^{-2}$ ($\mu_{5 \times 10^{13}}$), the effective mass at the conduction-band minimum (m*) (longitudinal (*l*) and transverse (*t*) components for silicon), the longitudinal effective mass at the $Q$ valley (m$_Q^{*l}$), and the coefficient representing the temperature dependence of the mobility ($\gamma$) are reported.

| Material | $\mu_{\text{int}}$ (cm$^2$/Vs) | $\mu_{5 \times 10^{13}}$ (cm$^2$/Vs) | m* ($m_0$) | m$_Q^{*l}$ ($m_0$) | $\gamma$ |
|---|---|---|---|---|---|
| MoS$_2$ | 221 (47–410) | 130 (144–265) | 0.46 (0.45–0.51) | 0.61 ⋯ | 1.63 (1.35–1.69) |
| MoSe$_2$ | 89 (18–269) | 53 ⋯ | 0.53 (0.52–0.64) | 0.55 ⋯ | 1.54 ⋯ |
| WS$_2$ | 335 (37–1739) | 179 (60) | 0.31 (0.24–0.35) | 0.60 (0.48–0.60) | 1.35 ⋯ |
| WSe$_2$ | 255 (23–1083) | 83 (25) | 0.33 (0.31–0.39) | 0.51 ⋯ | 1.42 ⋯ |
| Si | 1452 (1080–1970) | ⋯ ⋯ | 0.96$^l$,0.19$^t$ (0.98$^l$,0.19$^t$) | ⋯ ⋯ | 2.28 (2.29) |

Convergence has been tested with respect to the grid density and interaction range, as shown in the Appendix. The results presented here have been obtained on a $301 \times 301 \times 1$ **k**/**q**-point grid for the TMDCs and on a $201 \times 201 \times 201$ **k**/**q**-point grid for silicon. Scattering states up to 350 meV above the conduction-band minimum were included, capturing the entire Fermi tail in case of high carrier concentrations. The TMDC mobilities correspond to transport occurring along the zigzag direction. Their values at 300 K are reported in Table II for two carrier concentrations, together with the extracted effective masses and the coefficient $\gamma$ representing the temperature dependence of the mobility, i.e., $\mu(T) \propto T^{-\gamma}$ [23].

We report intrinsic phonon-limited mobilities of 221, 89, 335, and 255 cm$^2$/Vs for MoS$_2$, MoSe$_2$, WS$_2$, and WSe$_2$, respectively. These results should be put in perspective with the large range of values reported in the literature: 47–410 cm$^2$/Vs for MoS$_2$ [20,23–27,29–32], 18–269 cm$^2$/Vs for MoSe$_2$ [29–32], 37–1739 cm$^2$/Vs for WS$_2$ [28–32], and 23–1083 cm$^2$/Vs for WSe$_2$ [28–32]. Although MoS$_2$ has probably been the most-studied TMDC, there is still no consensus regarding its true phonon-limited mobility value. A recent study accounting for the effects of long-range interpolation and including SOC narrowed down the range to 120–150 cm$^2$/Vs [34]. Issues still persist regarding the variability of results obtained using different pseudopotentials and exchange-correlation functionals.

The $K$-$Q$ valley separation has been identified as a possible source of variability, as it influences intervalley scattering. However, attempts to explain the observed discrepancies solely based on this parameter have led to contradictory results. In Ref. [25], a decrease in the mobility of MoS$_2$ from 320 cm$^2$/Vs down to 130 cm$^2$/Vs was found upon including intervalley scattering between the $K$ and $Q$ valleys separated by 70 meV, while in Ref.

[29], the mobility did not change when varying both the pseudopotentials and the exchange and correlation functional, although the valley separation increased from 100 to 270 meV. This indicates that the $K$-$Q$ valley separation, though important, is only one piece of a more complex interplay of parameters. This feature becomes even more evident when examining WS$_2$, which exhibits the broadest range of calculated mobility values. Part of this variation can be attributed to discrepancies in the effective mass. For example, the study reporting the highest mobility for WS$_2$ (1739 cm$^2$/Vs) relies on an effective mass of 0.26 $m_0$ [31]. This is nearly 20% lower than the values both our study and several others have extracted of 0.31 $m_0$ [28,29,32]. In Ref. [29], the effective mass at $K$ varies from 0.30 $m_0$ to 0.35 $m_0$, depending on the chosen DFT parameters. The authors also show that the heavier effective mass of the $Q$ valley can also change between 0.48 $m_0$ and 0.60 $m_0$, which is expected to impact the mobility. Hence, similarly to the $K$-$Q$ valley separation, the effective mass of WS$_2$ shows a high sensitivity to the simulation setup. Reported effective masses for all the studied TMDCs include 0.45–0.51 $m_0$ for MoS$_2$ [23,26,27,30–32], 0.52–0.64 $m_0$ for MoSe$_2$ [30–32], 0.24–0.35 $m_0$ for WS$_2$ [28–32], and 0.31–0.39 $m_0$ for WSe$_2$ [28,30–32].

In summary, sometimes even small changes in the calculated electron and phonon dispersions can have a profound impact on the mobility values. Yet, even with identical dispersions, drastically different mobilities can be obtained. A compelling example is given by Ref. [29]. Specifically, for WS$_2$, the authors demonstrated that even with comparable band structures and phonon dispersions, a tenfold difference in the electron-phonon matrix-elements can exist if the pseudopotentials differ. While the derived electron energies and phonon frequencies, i.e., their eigenvalues, remain very close, their eigenvectors can be quite different, which has a direct impact on the electron-phonon

coupling elements. It can thus be concluded that the selected methodology to determine the scattering rates (here, the frozen-phonon approach is used as opposed to DFPT) can lead to different results because of the underlying electron and phonon eigenvectors, among other factors. Keeping these facts in mind, we believe that our calculated TMDC mobilities agree well with the literature, especially when looking at some of the most recent works for $MoS_2$ [27], $MoSe_2$ [29], $WS_2$ [29], and $WSe_2$ [29].

We note that the frozen-phonon approach used here does not fully capture long-range interactions such as polar optical phonon scattering (Fröhlich interaction), because it is limited by the supercell size of the DFT calculations. Since the Fröhlich interaction is slow to converge with respect to the supercell size, it cannot be included [20]. The choice to partly ignore these effects in our mobility calculations comes from the fact they are difficult to integrate into a real-space NEGF solver, while they can be readily accounted for in LBTE. Since our goal consisted of comparing the results of both methods, we decided to use the same level of approximation in both cases.

Several DFPT-based approaches have been developed to account for the long-range contributions to electron-phonon interactions [34,89]. We have previously shown that for $MoS_2$, including the analytical 2D Fröhlich contribution of Ref. [89] has a negligible impact on the mobility of this material [38]. However, as demonstrated in Ref. [34], the method of Ref. [89] does not fully capture all aspects of long-range contributions, resulting in an overestimation of the mobility between 20% (IBTE) and 40% (ERTA). While the Fröhlich interactions may be different in other TMDCs, we generally observe a good agreement between our results and those from studies considering long-range contributions via DFPT [29]. It should, however, be noted that neither of these works makes use of the method proposed in Ref. [34].

Next, we place our $MoS_2$ simulation findings in the context of experimental measurements. The focus is set on this TMDC because more data are available. The measured mobility of monolayer $MoS_2$ ranges from 23 to 217 $cm^2/Vs$ [6,90–96]. In experiments too, multiple parameters can influence the mobility, from the samples themselves to the measurement equipment, the quality of the data, and the dielectric environment. The studies reporting the largest mobility values typically involve high-permittivity gate dielectrics such as $HfO_2$. High-$\kappa$ dielectrics can better suppress Coulomb interactions and, as a consequence, charged impurity scattering (CIS) [90]. Reducing scattering brings the mobility closer to its intrinsic phonon-limited value. Our simulations specifically target the mobility of freestanding monolayers and consider electron-phonon scattering—and no other mechanisms. Additional scattering processes, notably surface optical phonon (SOPS) and CIS, are known to have a negative impact on the transport properties of TMDCs [36].

Hence, the calculated $MoS_2$ mobilities accounting for electron-phonon scattering only should be larger than the experimental ones, which suffer from these different scattering sources. With a value of 221 $cm^2/Vs$ for monolayer $MoS_2$, we are in a range compatible with the experimental data, knowing that the inclusion of SOPS and CIS would decrease the mobility bellow 100 $cm^2/Vs$ [36]. This sanity check gives us confidence in the developed computational framework.

Another possibility to access the accuracy of our method consists of simulating materials for which there are less uncertainty in the mobility value. Silicon emerges as the ideal candidate to do that and to further validate our methodology. We obtain a phonon-limited electron mobility of 1452 $cm^2/Vs$ for bulk silicon at 300 K. This result aligns very well with the experimental values, which lie between 1300 and 1450 $cm^2/Vs$ [83,97,98]. On the modeling side, the range is broader and encompasses values between 1080 and 1970 $cm^2/Vs$ [24,56, 99,100]. Recent investigations have narrowed this range down to 1305–1555 $cm^2/Vs$ [58]. This latter study has meticulously analyzed the influence of various simulation parameters on the mobility. Different exchange and correlation functionals, spin-orbit coupling, and GW quasiparticle corrections have been tested. It has been found that the most reliable theoretical estimate for the mobility of bulk silicon is 1366 $cm^2/Vs$. Interestingly, when employing simulation parameters similar to ours—specifically, the use of GGA, no spin-orbit coupling, and no crystal relaxation (experimental lattice constant)—the authors obtained a mobility of 1457 $cm^2/Vs$, a value almost identical to ours. At lower temperatures of 200 and 100 K, we calculated mobilities of 3740 and 17831 $cm^2/Vs$, respectively. In Ref. [58] using their most accurate computational settings, values of $\sim$3460 and $\sim$16 890 $cm^2/Vs$ have been found, which aligns relatively well with our results. Our calculated longitudinal (0.96 $m_0$) and transverse (0.19 $m_0$) effective masses for silicon also closely match the experimental values of 0.98 and 0.19 $m_0$ [83]. This excellent agreement with both experimental data and the calculations of Ref. [58] reinforces our confidence in the accuracy of our modeling approach and demonstrates its application potential beyond just TMDCs.

So far, we have considered mobility calculations under intrinsic carrier concentrations, where the Fermi level is located in the middle of the TMDC band gap. Given the fact that semiconductors are doped in most applications, it is crucial to also explore their transport properties at higher carrier concentrations. To this end, we have computed with LBTE the phonon-limited electron mobility of the TMDCs as a function of the carrier density and temperature. Our findings are shown in Fig. 6. The mobilities have been determined as a function of the Fermi level, which has been shifted upward, from the middle of the band gap into
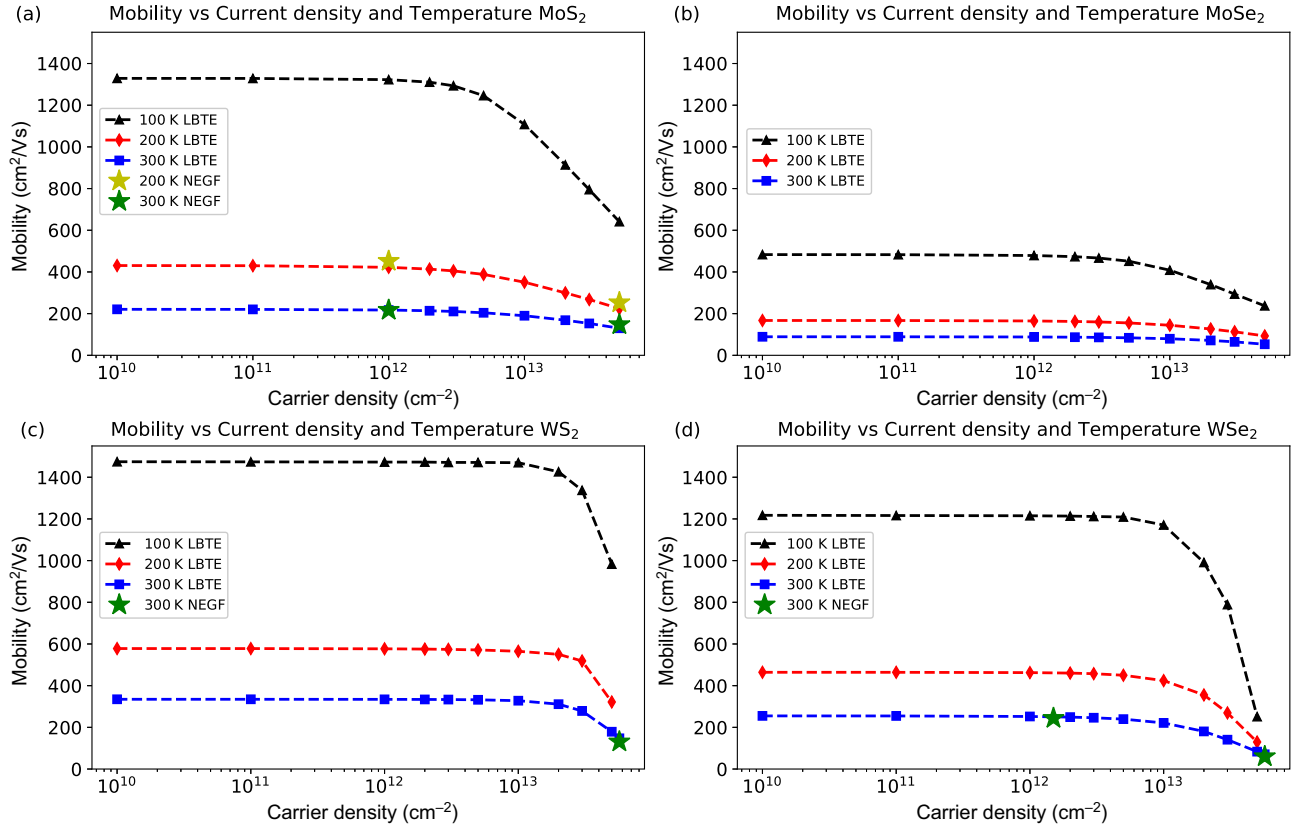
FIG. 6. Phonon-limited mobility of the considered TMDCs—(a) MoS$_2$, (b) MoSe$_2$, (c) WS$_2$, and (d) WSe$_2$—as a function of the carrier density for a temperature $T = 100$ K (dashed line with triangles), $T = 200$ K (dashed line with diamonds), and $T = 300$ K (dashed line with squares). All these results have been obtained with LBTE. The stars in subplots (a), (c), and (d) refer to NEGF calculations.

the conduction band. The corresponding carrier density has then subsequently been computed with Eq. (29). The calculated mobilities exhibit the expected behavior: they initially remain flat as the carrier concentration increases up to $10^{12}$ cm$^{-2}$. Above this point, higher-energy states in the conduction band become populated. The electrons occupying them can relax into low-energy states through optical phonon emission, which increases the scattering rate and decreases the mobility. The observed mobility drop accelerates when the $Q$ valley starts to fill up and intervalley scattering kicks in.

Turning to the temperature dependence of the mobility, we observe that the lower this parameter is, the smaller the phonon population. Hence, the scattering rate diminishes and the mobility is enhanced. In monolayer MoS$_2$, the intrinsic phonon-limited mobility rises from 221 cm$^2$/Vs at 300 K to 431 cm$^2$/Vs at 200 K and further to 1329 cm$^2$/Vs at 100 K. At the same time, the mobility still drops rapidly at high carrier concentrations, in line with the predictions of other studies [20,23]. Assuming that the mobility evolves as $\mu(T) \propto T^{-\gamma}$ [23], where $T$ is the temperature, we can extract the decay parameter $\gamma$ from our calculations. For MoS$_2$, we find that $\gamma = 1.63$,

a value similar to what other computational studies have reported (between 1.35 and 1.69 [20,22,23]). The calculated $\gamma$ values for the other TMDCs considered in this work are listed in Table II. For silicon, we obtain $\gamma = 2.28$, in good agreement with the value (2.29) extracted from Ref. [58].

For MoS$_2$, we have also calculated the intrinsic hole mobility, which comes out to be 19 cm$^2$/Vs. Contrary to the electron mobility, fewer simulation studies provide this quantity. For instance, in Ref. [33] a comparable hole mobility of 26 cm$^2$/Vs has been obtained, whereas a value of 73 cm$^2$/Vs is reported in Ref. [34] when incorporating spin-orbit coupling.

### 2. NEGF

LBTE offers a computationally efficient platform to determine the mobility of bulk and nanostructured materials but it does not provide information about electrical current or device characteristics. This is where the NEGF formalism can be of great help. Although both approaches take the same quantities as input, they rely on different systems of equations. To validate their implementation, it is

therefore useful to compare physical quantities that can be computed with both of them. This is the case of the phonon-limited mobility. In particular, since the scattering self-energies are truncated in NEGF ($\Gamma$-point calculation and introduction of a cutoff radius) while no such approximations are made in LBTE, comparisons between the mobility calculated with both methods allow us to assess the accuracy of the simplifications we have made to NEGF.

The phonon-limited mobility is computed in NEGF with the $dR/dL$ method [73], which enables consideration of different temperatures and carrier concentrations. We first construct $3 \times 1 \times 8$ supercells comprising 144 atoms that are derived from the smallest possible orthorhombic cell made of 6 atoms. These supercells are then replicated along the transport direction $x$ to give rise to devices of lengths roughly equal to 10, 17, and 30 nm. The periodic direction $z$ is modeled via 3 $k_z$-points at low carrier densities ($n = 1 \times 10^{12}$ cm$^{-2}$) and 5 $k_z$-points at $n = 5 \times 10^{13}$ cm$^{-2}$. We have verified that this number of $k_z$ points is sufficient to capture all band-structure features. Such small numbers of $k_z$ points are justified by the large dimension of the transport cell along $z$ ($R_z = 2.5$ nm for MoS$_2$). The phonon dispersion and eigenvectors are evaluated at the $\Gamma$ point and include a total of 432 phonon energies. In the calculation of the scattering self-energies, $\boldsymbol{\Sigma}_{mn}^{\gtrless e\text{-ph}}$, in Eq. (35), a cutoff radius $r_{cut} < 12$ Å (13.2 Å) for MoS$_2$ and WS$_2$ (WSe$_2$) is introduced to keep the computational burden and memory manageable. On average, each atom interacts with 155 neighbors. No larger value could be simulated on the available machine [101]. A detailed analysis of the $r_{cut}$-dependent convergence of the electrical current is provided in the Appendix. All NEGF-based mobility calculations are performed under small bias conditions ($\Delta V = 1$ mV), assuming a flat electrostatic potential.
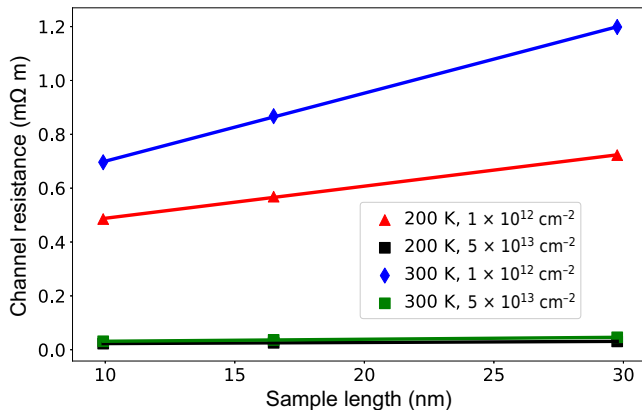


FIG. 7. The channel resistance of monolayer MoS$_2$ as a function of the sample length, at temperatures of 200 K and 300 K and for electron concentrations $n = 1 \times 10^{12}$ cm$^{-2}$ and $n = 5 \times 10^{13}$ cm$^{-2}$. The symbols represent simulation results, while the lines are used as fits and serve as inputs to the $dR/dL$ method [73].

TABLE III. Phonon-limited mobility of selected TMDCs at 300 K, calculated using LBTE [within the IBTE (ERTA) approximation] and NEGF at different carrier concentrations.

| Material | $1 \times 10^{12}$ cm$^{-2}$ | | $5 \times 10^{13}$ cm$^{-2}$ | |
|---|---|---|---|---|
| | LBTE (cm$^2$/Vs) | NEGF (cm$^2$/Vs) | LBTE (cm$^2$/Vs) | NEGF (cm$^2$/Vs) |
| MoS$_2$ | 217 (197) | 217 | 130 (139) | 148 |
| WS$_2$ | 334 (128) | $\cdots$ | 147 (51)[b] | 131[b] |
| WSe$_2$ | 251 (166)[a] | 244[a] | 71 (57)[b] | 60[b] |

[a]([b]) Actual carrier concentration of $1.5 \times 10^{12}$ cm$^{-2}$ ($5.7 \times 10^{13}$ cm$^{-2}$).

Focusing on MoS$_2$, in Fig. 7, we present the obtained MoS$_2$ channel resistances as a function of the device length. Due to the high computational intensity associated with such calculations, we have limited ourselves to two temperatures, 200 K and 300 K, and two electron concentrations, $n = 5 \times 10^{13}$ cm$^{-2}$ and $n = 1 \times 10^{12}$ cm$^{-2}$. Ideally, the sample lengths should be much larger then the mean free path for scattering, but for the same reason as already mentioned above, we have not been able to go beyond 30-nm-long devices. Nonetheless, it is obvious from the results in Fig. 7 that the resistance increases linearly with the sample length in all cases, underscoring the diffusive nature of transport. From these data, the mobility at 300 K and $n = 1 \times 10^{12}$ cm$^{-2}$ is found to be 217 cm$^2$/Vs, which is in perfect agreement with our LBTE finding under the same condition (217 cm$^2$/Vs). At a reduced temperature of 200 K, the NEGF mobility rises to 452 cm$^2$/Vs, thus exceeding the LBTE value of 422 cm$^2$/Vs. Meanwhile, increasing the carrier concentration to $n = 5 \times 10^{13}$ cm$^{-2}$ reduces the mobility to 148 cm$^2$/Vs at 300 K and 254 cm$^2$/Vs at 200 K in case of NEGF. These results also overestimate the LBTE values, which are equal to 130 cm$^2$/Vs and 224 cm$^2$/Vs at 300 K and 200 K, respectively.

To further validate our approach, we have calculated the mobility of monolayer WS$_2$ and WSe$_2$ with NEGF. Note that for WS$_2$, only the value at high carrier density could be obtained, as the maximum sample length that we can currently simulate (30 nm) is too short to ensure that transport is in the diffusive regime and that the $dR/dL$ approach is valid. The calculated mobilities, at 300 K, of all considered TMDC are reported in Table III, including the IBTE and ERTA values. The NEGF mobilities are also shown as stars in Fig. 6.

We believe that the relatively good agreement between LBTE and NEGF (with differences of less than 15% in all cases and under 3% at low carrier densities and a temperature of 300 K) validates our approximations to calculate the electron-phonon scattering self-energy. Focusing on MoS$_2$, the remaining discrepancies can possibly be explained as follows. Considering the lower-temperature

case first, the higher NEGF mobility can be attributed to the short sample lengths (10, 17, and 30 nm). From the calculated mobilities and resistivities ($dR/dL$), we can estimate the phonon-limited mean free path for carriers [74], which is equal to 9.8 nm at 300 K and 22.1 nm at 200 K for low carrier concentrations. Given that the mean free path at 200 K is almost on par with the maximum sample length of 30 nm, the channel might not be long enough to let the electron population fully relax due to phonon emission or absorption. An underestimation of the scattering rate is expected to lead to higher mobility values. This reasoning does not apply to the high-carrier-concentration case, where the mean free path is 6.7 nm at 300 K. In this configuration, the difference between LBTE and NEGF may be due to the lack of $k_z$-point coupling. While the size of the transport cell ensures that the most relevant states are projected to the $\Gamma$ point, or close to it, this might no longer be true at high carrier concentrations. For example, even in the chosen large cell, not all $Q$-valley states are projected to $\Gamma$. However, they start to play an important role when the electron density reaches $5 \times 10^{13}$ cm$^{-2}$. Our $\Gamma$-point calculations neglect parts of the coupling to the $Q$ valley, which might underestimate the scattering rate and artificially increase the mobility. The error (148 versus 130 cm$^2$/Vs and 254 versus 224 cm$^2$/Vs) remains acceptable and indicates that most scattering events are still accounted for.

### C. Device simulation

Next, after demonstrating that the NEGF approach can reproduce the LBTE mobilities fairly accurately, we move to full *ab initio* device simulations. The transfer characteristics of a single-gate monolayer MoS$_2$ FET, as depicted in Fig. 2, are investigated, assuming perfectly Ohmic contacts. The gate length, $L_g = 11.8$ nm, has been chosen according to the latest International Roadmap for Devices and Systems (IRDS) for the year 2028 [102]. The source and drain extensions measure $L_s = L_d = 9$ nm and are doped with a donor concentration $N_D = 5 \times 10^{13}$ cm$^{-2}$. The transistor structure is therefore composed of 2592 atoms. The MoS$_2$ channel is deposited on a SiO$_2$ "substrate" with a top HfO$_2$ dielectric layer of thickness $t_{ox} = 3$ nm and a relative permittivity $\epsilon_{ox} = 20$, resulting in an equivalent oxide thickness (EOT) of 0.58 nm. The OFF-state current $I_{OFF}$ is set to 0.1 $\mu$A/$\mu$m by adjusting the gate work function. The same simulation settings as for the mobility calculations at $n = 5 \times 10^{13}$ cm$^{-2}$ and $T = 300$ K are used (5 $k_z$-points, an energy resolution of 2 meV).

First, the $I_d$-$V_{gs}$ transfer characteristics of the MoS$_2$ transistors are simulated in the quasiballistic limit of transport. The phenomenological electron-phonon scattering model of Ref. [79] is recalled for that purpose with a phonon frequency $\hbar\omega = 40$ meV and a scattering strength $D_{e\text{-ph}} = 25$ (meV)$^2$. This model ensures that the
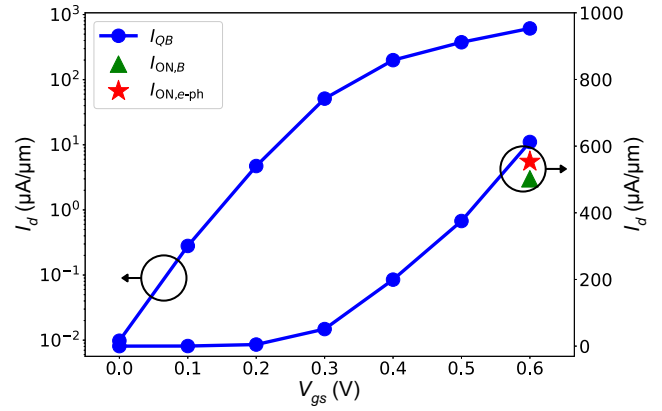


FIG. 8. Transfer characteristics $I_d$-$V_{gs}$ at $V_{ds} = 0.6$ V of a single-gate monolayer MoS$_2$ field-effect transistor in the quasiballistic limit of transport ($I_{QB}$, blue curve with circles). The ON-state current in the pure ballistic limit ($I_{ON,B}$, green triangle) and in the presence of electron-phonon interactions according to Eq. (35) ($I_{ON,e\text{-ph}}$, red star) are indicated. Both a linear and a log scale are provided for $I_{QB}$.

nonphysical negative differential behavior often observed in TMDC-based FET is eliminated. The resulting curve at $V_{ds} = 0.6$ V is depicted in Fig. 8. It can be seen that the device exhibits good electrostatic properties with a subthreshold swing of 69 mV/dec and a decent ON-state current ($I_d$ at $V_{ds} = V_{gs} = 0.6$ V) of 613 $\mu$A/$\mu$m. It is also found that the considered transistor does not suffer from the so-called density-of-states bottleneck [103], the extracted gate capacitance $C_g = \partial n/V_{gs} = 5.6$ $\mu$F/cm$^2$ approaching the value of the oxide capacitance ($C_{ox} = 5.9$ $\mu$F/cm$^2$).

As computing the entire $I_d$-$V_{gs}$ in the presence of electron-phonon scattering as described in Eq. (35) is computationally not feasible, we only include the developed electron-phonon scattering model when calculating the ON-state current of the transistor and evaluate it with the same electrostatic potential as in the case of the phenomenological model. As a reference, we have also determined the value of the pure ballistic current by turning off all scattering interactions. Both the pure ballistic and the dissipative currents are displayed in Fig. 8 as a triangle (pure ballistic current, $I_{ON,B} = 502$ $\mu$A/$\mu$m) and as a star (current with real electron-phonon scattering, $I_{ON,e\text{-ph}} = 554$ $\mu$A/$\mu$m). It can be noted that the $I_{ON,B}$ is smaller than the ON-state currents obtained in the presence of scattering. This phenomenon has been explained in Ref. [35]: phonons connect electronic bands with a narrow energy width that cannot carry current in the ballistic limit of transport. In this case, electron-phonon scattering can increase the current. There is, however, a competing effect that manifests itself at high scattering rates: back scattering or the decrease of the current caused by electron-phonon interactions. This is exactly what happens when the full
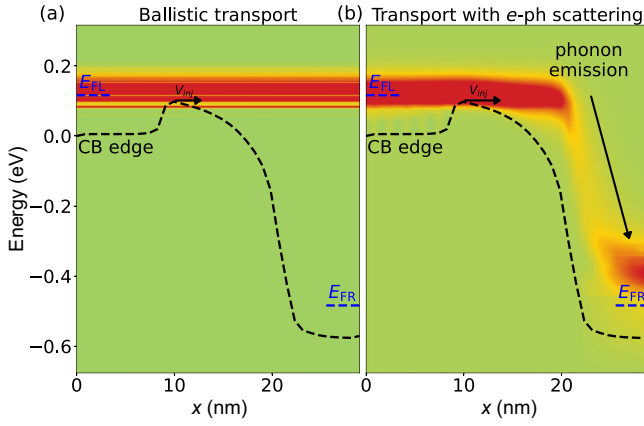
FIG. 9. Spectral ON-state current ($I_d$ as a function of the energy $E$ and position $x$) for a single-gate monolayer $MoS_2$ field-effect transistor in (a) the ballistic limit of transport and (b) when the proposed electron-phonon scattering model is turned on. Red indicates high current concentration, while no current flows through the green regions. The conduction-band edge is marked by a black dashed line. The Fermi levels of the left ($E_{FL}$) and right ($E_{FR}$) contacts are highlighted in blue. The location at which the injection velocity is extracted is marked with $v_{inj}$.

electron-phonon scattering model is turned on: back scattering starts to dominate, so that the current decreases as compared to the quasiballistic case. It nevertheless remains higher than in the pure ballistic case. Overall, the difference in current is rather small, $I_{ON,B} = 502$ μA/μm and $I_{ON,e\text{-}ph} = 554$ μA/μm but the electron behavior differs significantly.

This is exemplified in Fig. 9, where the energy- and position-resolved $I_{ON,B}$ and $I_{ON,e\text{-}ph}$ are plotted. In the ballistic case, the spectral current remains constant from the left side (source) to the right side (drain): the energy of the electrons injected into the simulation domain cannot vary in the absence of a dissipative scattering mechanism. When electron-phonon interactions are introduced, the electrons lose a substantial amount of their energy between the source and the drain of the transistor. Most of the energy relaxation occurs in the region located after the top of the potential barrier, where the electric field reaches its maximum.

The effect of back scattering, which causes electrons flowing from the source to the drain to be scattered back to their origin, cannot be directly visualized, but its influence can be indirectly measured when considering the injection velocity extracted at the top of the barrier. This velocity decreases from $3.6 \times 10^6$ cm/s in the quasiballistic case to $3.2 \times 10^6$ cm/s when electron-phonon scattering is fully accounted for. Note that in ultrascaled devices, the injection velocity and charge at the top of the potential barrier are not as clearly defined as in their long-channel counterparts. This is due to the absence of a plateau in the center of the channel. The extracted injection velocity values tend to

vary depending on the interpolation technique that is used to identify the exact location of the top of the potential barrier.

## IV. CONCLUSIONS AND OUTLOOK

This work has explored the transport properties of selected TMDCs and bulk silicon based on an *ab initio* description of their electron and phonon dispersions and of the interaction between these particles. The developed methodology relies on a combination of first-principles DFT, MLWF, LBTE, and NEGF. Mobility calculations reveal that the obtained results align very well with the current literature, despite the myriad of factors that may impact the calculated quantities. The sensitivity of the TMDC mobility to factors such as pseudopotentials, exchange-correlation functionals, and phonon calculation method has been highlighted. Hence, as a sanity check, the mobility of silicon, a well-characterized semiconductor, has been computed and compared to experimental data. While full inclusion of Fröhlich interactions is lacking in our model, the agreement between our calculations, measurements, and other computational studies validates our treatment of electron-phonon scattering within the considered framework. The latter has therefore been employed in NEGF to construct electron-phonon scattering self-energies going beyond the diagonal approximation but restricted to Γ-point calculations. Here again, good agreement between LBTE and NEGF indicates that the most important features of electron-phonon scattering are captured by our NEGF solver. As a consequence, *ab initio* device simulations of TMDC-based FETs incorporating the real mobility of these materials could be performed. Being able to precisely account for electron-phonon scattering in the investigation of TMDC devices is essential to shed light on their intrinsic transport properties. While the accuracy achieved by our NEGF approach for mobility calculations and device simulations is promising, further enhancements are required to make it readily applicable to any system. Its computational efficiency should be improved so that it can treat larger devices and structures embedded within a dielectric environment [37]. Such an extension is key to directly accounting for surface optical phonon scattering. Significant speed-ups might be possible by applying the mode-space approximation [104] and by projecting the Hamiltonian and electron-phonon scattering self-energies into this basis. As matrices with lower dimensions are obtained when going from real space to mode space, the gain in processing time and storage requirements could be leveraged to increase the cutoff radius applied to the scattering self-energies and/or to increase the size of the transport cell. Additionally, the possibility of going beyond the diagonal approximation for scattering self-energies can now be used to treat other mechanisms as well; e.g., (charged) impurity scattering, surface

roughness, or alloy disorder. The inclusion of nondiagonal scattering entries ensures a more detailed and accurate description of carrier interactions with their environment.

## ACKNOWLEDGMENTS

## APPENDIX: CONVERGENCE ANALYSIS

To ensure the accuracy of the electrical results presented in this paper, in this appendix we evaluate the convergence of the calculated mobility and electrical currents, using monolayer $MoS_2$ as a representative example, with respect to the $\mathbf{k}/\mathbf{q}$-point grid (LBTE) and interaction range (LBTE and NEGF).

The implemented LBTE solver enables us to sample the Brillouin zone with high resolution, extending up to $501 \times 501 \times 1$ $\mathbf{k}/\mathbf{q}$-point grids for TMDCs. The convergence of the phonon-limited mobility of $MoS_2$ is plotted in Fig. 10 as a function of the $\mathbf{k}/\mathbf{q}$-point density used in the LBTE calculations. It can be seen that a $301 \times 301 \times 1$ $\mathbf{k}/\mathbf{q}$-point grid is sufficient to produce converged results, the mobility at this grid size only displaying a 0.3% deviation from the densest grid.

In the case of LBTE, the interaction range used to construct the electron-phonon coupling elements can be effectively reduced by constraining the summation over neighboring cells, $(\beta, \gamma)$ in Eq. (7). In Fig. 11, we show the resulting mobility as a function of the number of
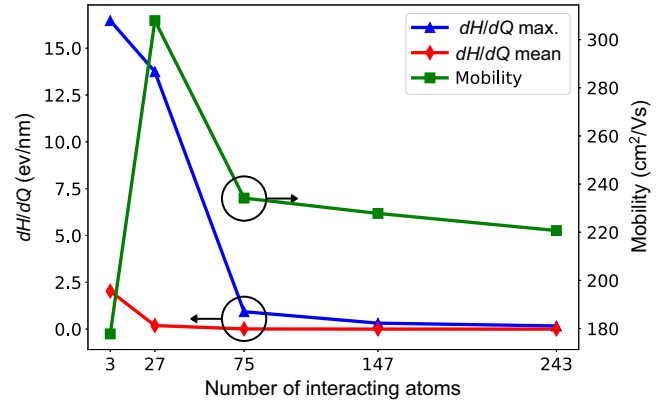


FIG. 11. Phonon-limited mobility of monolayer $MoS_2$ as calculated with LBTE (green line with squares), the maximum value of the Hamiltonian derivative $dH/dQ$ (blue line with triangles), and the mean value of $dH/dQ$ (red line with diamonds) are reported as a function of the number of atoms interacting with an arbitrary point of reference.

interacting atoms included. It can be observed that the mobility, with interactions limited to 75 atoms, already approaches its converged value with approximately 6% difference as compared to the case in which the complete data set is incorporated. This discrepancy diminishes to below 3% when interactions with up to 147 atoms are included. Although the slope of the mobility convergence with respect to the interaction range might appear linear when the number of interacting atoms increases from 75 to 243, it is not. The slope between 147 and 243 atoms is 20% smaller than between 75 and 147 atoms. Hence, the mobility converges, but slowly with respect to the DFT supercell size, as previously demonstrated [20]. In contrast, relatively rapid convergence is also observed in the localized Hamiltonian derivatives, $dH/dQ_{I\eta\gamma}$, in Eq. (7). Their maximum and mean values are reported in Fig. 11 as a function of the number of interacting atoms. As expected, these derivatives quickly decrease and remain small when more than 75 atoms are included. When the full data are considered, the magnitude of the additional long-range interactions becomes comparable to the noise floor of the MLWF calculations, which is defined as all numerical values below a certain magnitude (usually about 0.00001 eV), where the entries of the Wannier Hamiltonian are indistinguishable from computational noise. Hence, removing them has no impact on the resulting eigenvalues or eigenvectors.

Moving on to the NEGF approach, we have calculated the electrical current flowing through 10-nm-long $MoS_2$ and $WSe_2$ samples at charge densities of $1 \times 10^{12}$ and $5.7 \times 10^{13}$ $cm^{-2}$, respectively, as a function of the interaction range $r_{cut}$. Our interaction range refers to the maximum distance between two interacting atomic orbitals. Although beyond $r_{cut}$ the scattering self-energies
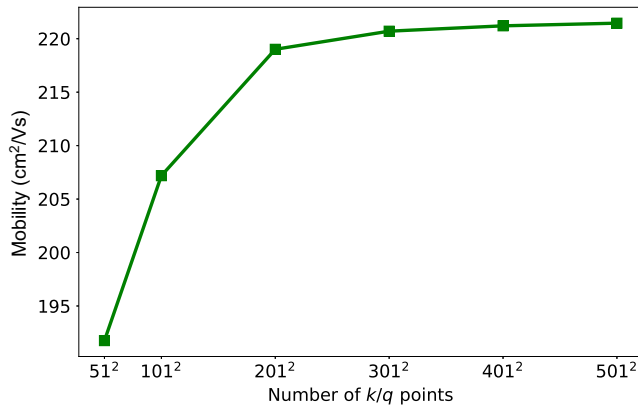


FIG. 10. Phonon-limited mobility of monolayer $MoS_2$ as a function of the $\mathbf{k}/\mathbf{q}$-point density. At a $301 \times 301 \times 1$ $\mathbf{k}/\mathbf{q}$-point grid, the phonon-limited mobility shows only a 0.3% deviation from the value obtained with a denser $501 \times 501 \times 1$ $\mathbf{k}/\mathbf{q}$-point grid.
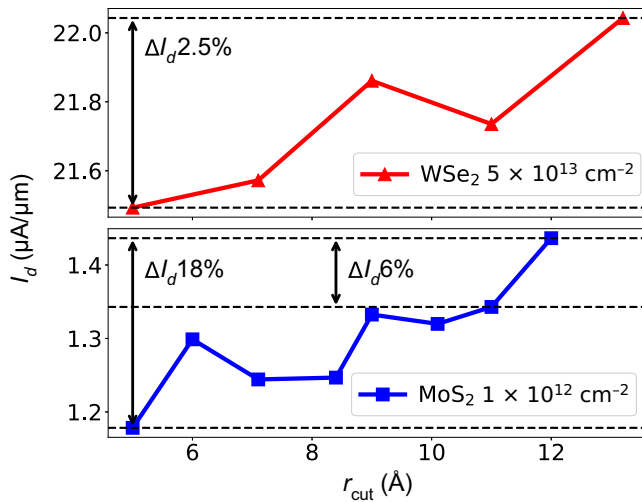
FIG. 12. Electrical current flowing through 10-nm-long monolayer $MoS_2$ and $WSe_2$ samples under a flat-band potential and an applied voltage of 1 mV as a function of the cutoff radius $r_{cut}$ applied to Eq. (35). The relative changes of the current between $r_{cut} = 5$ and 12 Å (18%), $r_{cut} = 11$ and 12 Å (6%) for $MoS_2$, and $r_{cut} = 5$ and 13.2 Å (2.5%) for $WSe_2$ are indicated.

are assumed to be zero, all available atomic displacements ($I\eta\gamma$) in Eq. (7) are included, regardless of $r_{cut}$. The convergence behavior is reported in Fig. 12 for $r_{cut}$ between 5 and 13.2 Å, the maximum value that could be reached on the available hardware. For $MoS_2$, the current changes by only 18% when $r_{cut}$ increases from 5 to 12 Å, yet it does not converge toward a fixed value. In contrast, the change in the $WSe_2$ current is much smaller, less than 2.5%. Currently, we do not have an explanation for the worse convergence of $MoS_2$. However, the fact that the mobilities calculated using the largest $r_{cut}$ values align fairly well with those from LBTE, regardless of the convergence behavior suggests that a sufficiently large number of nondiagonal entries in the scattering self-energies are considered in the present calculations. Further investigations will be conducted after getting access to a larger machine or after projecting the Hamiltonian and electron-phonon coupling elements onto a significantly smaller mode-space basis.

[1] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, Electronics and optoelectronics of two-dimensional transition metal dichalcogenides, Nat. Nanotechnol. **7**, 699 (2012).

[2] S. E. Kim, F. Mujid, A. Rai, F. Eriksson, J. Suh, P. Poddar, A. Ray, C. Park, E. Fransson, Y. Zhong, D. A. Muller, P. Erhart, D. G. Cahill, and J. Park, Extremely anisotropic van der Waals thermal conductors, Nature **597**, 660 (2021).

[3] D. Akinwande, et al., A review on mechanics and mechanical properties of 2D materials—graphene and beyond, Extreme Mech. Lett. **13**, 42 (2017).

[4] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, Single-layer $MoS_2$ transistors, Nat. Nanotechnol. **6**, 147 (2011).

[5] Y. Lee, S. Bae, H. Jang, S. Jang, S.-E. Zhu, S. H. Sim, Y. I. Song, B. H. Hong, and J.-H. Ahn, Wafer-scale synthesis and transfer of graphene films, Nano Lett. **10**, 490 (2010).

[6] K. Kang, S. Xie, L. Huang, Y. Han, P. Y. Huang, K. F. Mak, C.-J. Kim, D. Muller, and J. Park, High-mobility three-atom-thick semiconducting films with wafer-scale homogeneity, Nature **520**, 656 (2015).

[7] H. Fang, S. Chuang, T. C. Chang, K. Takei, T. Takahashi, and A. Javey, High-performance single layered $WSe_2$ p-FETs with chemically doped contacts, Nano Lett. **12**, 3788 (2012).

[8] D. Ovchinnikov, A. Allain, Y.-S. Huang, D. Dumcenco, and A. Kis, Electrical transport properties of single-layer $WS_2$, ACS Nano. **8**, 8174 (2014).

[9] S. Fathipour, N. Ma, W. S. Hwang, V. Protasenko, S. Vishwanath, H. G. Xing, H. Xu, D. Jena, J. Appenzeller, and A. Seabaugh, Exfoliated multilayer $MoTe_2$ field-effect transistors, Appl. Phys. Lett. **105**, 192101 (2014).

[10] Y. Meng, C. Ling, R. Xin, P. Wang, Y. Song, H. Bu, S. Gao, X. Wang, F. Song, J. Wang, X. Wang, B. Wang, and G. Wang, Repairing atomic vacancies in single-layer $MoSe_2$ field-effect transistor and its defect dynamics, npj Quantum Mater. **2**, 1 (2017).

[11] E. Liu, et al., Integrated digital inverters based on two-dimensional anisotropic $ReS_2$ field-effect transistors, Nat. Commun. **6**, 6991 (2015).

[12] M. J. Mleczko, C. Zhang, H. R. Lee, H.-H. Kuo, B. Magyari-Köpe, R. G. Moore, Z.-X. Shen, I. R. Fisher, Y. Nishi, and E. Pop, $HfSe_2$ and $ZrSe_2$: Two-dimensional semiconductors with native high-$\kappa$ oxides, Sci. Adv. **3**, e1700481 (2017).

[13] Y.-Y. Chung, et al., in *2022 International Electron Devices Meeting (IEDM)* (IEEE, 2022), pp. 34.5.1–34.5.4.

[14] Z. Bai, D. He, S. Fu, Q. Miao, S. Liu, M. Huang, K. Zhao, Y. Wang, and X. Zhang, Recent progress in electron-phonon interaction of two-dimensional materials, Nano Select **3**, 1112 (2022).

[15] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[16] P. Giannozzi, et al., QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, J. Phys.: Condens. Matter **21**, 395502 (2009).

[17] A. Togo and I. Tanaka, First principles phonon calculations in materials science, Scr. Mater. **108**, 1 (2015).

[18] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Phonons and related crystal properties from density-functional perturbation theory, Rev. Mod. Phys. **73**, 515 (2001).

[19] T. Frederiksen, M. Paulsson, M. Brandbyge, and A.-P. Jauho, Inelastic transport theory from first principles: Methodology and application to nanoscale devices, Phys. Rev. B **75**, 205413 (2007).

[20] T. Gunst, T. Markussen, K. Stokbro, and M. Brandbyge, First-principles method for electron-phonon coupling and electron mobility: Applications to two-dimensional materials, Phys. Rev. B **93,** 035414 (2016).

[21] S. Poncé, E. R. Margine, C. Verdi, and F. Giustino, EPW: Electron-phonon coupling, transport and superconducting properties using maximally localized Wannier functions, Comput. Phys. Commun. **209,** 116 (2016).

[22] J.-J. Zhou, J. Park, I.-T. Lu, I. Maliyov, X. Tong, and M. Bernardi, PERTURBO: A software package for *ab initio* electron-phonon interactions, charge transport and ultrafast dynamics, Comput. Phys. Commun. **264,** 107970 (2021).

[23] K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, Phonon-limited mobility in *n*-type single-layer $MoS_2$ from first principles, Phys. Rev. B **85,** 115317 (2012).

[24] W. Li, Electrical transport limited by electron-phonon coupling from Boltzmann transport equation: An *ab initio* study of Si, Al, and $MoS_2$, Phys. Rev. B **92,** 075405 (2015).

[25] X. Li, J. T. Mullen, Z. Jin, K. M. Borysenko, M. Buongiorno Nardelli, and K. W. Kim, Intrinsic electrical transport properties of monolayer silicene and $MoS_2$ from first principles, Phys. Rev. B **87,** 115418 (2013).

[26] Y. Zhao, Z. Dai, C. Zhang, C. Lian, S. Zeng, G. Li, S. Meng, and J. Ni, Intrinsic electronic transport and thermoelectric power factor in *n*-type doped monolayer $MoS_2$, New J. Phys. **20,** 043009 (2018).

[27] A. Pilotto, P. Khakbaz, P. Palestri, and D. Esseni, Semiclassical transport in $MoS_2$ and $MoS_2$ transistors by a Monte Carlo approach, Solid State Electron. **192,** 108295 (2022).

[28] T. Sohier, D. Campi, N. Marzari, and M. Gibertini, Mobility of two-dimensional materials from first principles in an accurate and automated framework, Phys. Rev. Mater. **2,** 114010 (2018).

[29] G. Gaddemane, S. Gopalan, M. L. Van de Put, and M. V. Fischetti, Limitations of ab initio methods to predict the electronic-transport properties of two-dimensional semiconductors: The computational example of 2H-phase transition metal dichalcogenides, J. Comput. Electron. **20,** 49 (2021).

[30] W. Zhang, Z. Huang, W. Zhang, and Y. Li, Two-dimensional semiconductors with possible high room temperature mobility, Nano Res. **7,** 1731 (2014).

[31] Z. Huang, W. Zhang, and W. Zhang, Computational search for two-dimensional MX2 semiconductors with possible high electron mobility at room temperature, Materials **9,** 716 (2016).

[32] Z. Jin, X. Li, J. T. Mullen, and K. W. Kim, Intrinsic transport properties of electrons and holes in monolayer transition-metal dichalcogenides, Phys. Rev. B **90,** 045422 (2014).

[33] F. Guo, Z. Liu, M. Zhu, and Y. Zheng, Electron-phonon scattering limited hole mobility at room temperature in a $MoS_2$ monolayer: First-principles calculations, Phys. Chem. Chem. Phys. **21,** 22879 (2019).

[34] S. Poncé, M. Royo, M. Stengel, N. Marzari, and M. Gibertini, Long-range electrostatic contribution to electron-phonon couplings and mobilities of two-

dimensional and bulk materials, Phys. Rev. B **107,** 155424 (2023).

[35] A. Szabó, R. Rhyner, and M. Luisier, *Ab initio* simulation of single- and few-layer $MoS_2$ transistors: Effect of electron-phonon scattering, Phys. Rev. B. **92,** 035435 (2015).

[36] Y. Lee, S. Fiore, and M. Luisier, in *2019 IEEE International Electron Devices Meeting (IEDM)* (IEEE, San Francisco, CA, USA, 2019), pp. 24.4.1–24.4.4.

[37] S. Fiore, C. Klinkert, F. Ducry, J. Backman, and M. Luisier, Influence of the hBN dielectric layers on the quantum transport properties of $MoS_2$ transistors, Materials **15,** 1062 (2022).

[38] J. Backman, Y. Lee, and M. Luisier, Electron-phonon calculations using a wannier-based supercell approach: Applications to the monolayer $MoS_2$ mobility, Solid State Electron. **198,** 108461 (2022).

[39] N. Marzari and D. Vanderbilt, Maximally localized generalized Wannier functions for composite energy bands, Phys. Rev. B **56,** 12847 (1997).

[40] A. P. Horsfield, D. R. Bowler, H. Ness, C. G. Sánchez, T. N. Todorov, and A. J. Fisher, The transfer of energy between electrons and ions in solids, Rep. Prog. Phys. **69,** 1195 (2006).

[41] G. D. Mahan, *Many-Particle Physics* (Springer Science & Business Media, New York, NY, USA, 2000).

[42] A. Wacker, Semiconductor superlattices: A model system for nonlinear transport, Phys. Rep. **357,** 1 (2002).

[43] G. Pizzi, *et al.*, WANNIER90 as a community code: New features and applications, J. Phys.: Condens. Matter **32,** 165902 (2020).

[44] C. Stieger and M. Luisier, WINTERFACE: An interface from WANNIER90 to quantum transport, arXiv:2007.04268.

[45] M. Head-Gordon and J. C. Tully, Vibrational relaxation on metal surfaces: Molecular-orbital theory and application to Co/Cu (100), J. Chem. Phys. **96,** 3939 (1992).

[46] E. Artacho, E. Anglada, O. Diéguez, J. D. Gale, A. García, J. Junquera, R. M. Martin, P. Ordejón, J. M. Pruneda, D. Sánchez-Portal, and J. M. Soler, The SIESTA method; developments and applicability, J. Phys.: Condens. Matter **20,** 064208 (2008).

[47] T. D. Kühne, *et al.*, CP2K: An electronic structure and molecular dynamics software package—QUICKSTEP: Efficient and accurate electronic structure calculations, J. Chem. Phys. **152,** 194103 (2020).

[48] J.-M. Lihm and C.-H. Park, Wannier function perturbation theory: Localized representation and interpolation of wave function perturbation, Phys. Rev. X **11,** 041053 (2021).

[49] G. J. Ackland, M. C. Warren, and S. J. Clark, Practical methods in *ab initio* lattice dynamics, J. Phys.: Condens. Matter **9,** 7861 (1997).

[50] A. Togo, First-principles phonon calculations with PHONOPY and PHONO3PY, J. Phys. Soc. Jpn. **92,** 012001 (2023).

[51] J. M. Ziman, *Electrons and Phonons: The Theory of Transport Phenomena in Solids* (Oxford University Press, Oxford, UK, 2001).

[52] M. Lundstrom, *Fundamentals of Carrier Transport* (Cambridge University Press, Cambridge, UK, 2000), 2nd ed.

[53] C. Jacoboni, *Theory of Electron Transport in Semiconductors: A Pathway from Elementary Physics to Nonequilibrium Green Functions* (Springer Berlin, Heidelberg, Berlin, Germany, 2010), 1st ed., pp. XXVI, 590.

[54] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, ShengBTE: A solver of the Boltzmann transport equation for phonons, Comput. Phys. Commun. **185,** 1747 (2014).

[55] T. Kawamura and S. D. Sarma, Phonon-scattering-limited electron mobilities in $Al_xGa_{1-x}As/GaAs$ heterojunctions, Phys. Rev. B **45,** 3612 (1992).

[56] R. Rhyner and M. Luisier, Phonon-limited low-field mobility in silicon: Quantum transport vs. linearized Boltzmann transport equation, J. Appl. Phys. **114,** 223708 (2013).

[57] P. A. M. Dirac, The quantum theory of the emission and absorption of radiation, Proc. R. Soc. Lond. Ser. A-Contain. Pap. Math. Phys. Character **114,** 243 (1927).

[58] S. Poncé, E. R. Margine, and F. Giustino, Towards predictive many-body calculations of phonon-limited carrier mobilities in semiconductors, Phys. Rev. B **97,** 121201 (2018).

[59] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Maximally localized Wannier functions: Theory and applications, Rev. Mod. Phys. **84,** 1419 (2012).

[60] W. Li, N. Mingo, L. Lindsay, D. A. Broido, D. A. Stewart, and N. A. Katcho, Thermal conductivity of diamond nanowires from first principles, Phys. Rev. B **85,** 195436 (2012).

[61] P. Danielewicz, Quantum theory of nonequilibrium processes, I, Ann. Phys. **152,** 239 (1984).

[62] S. Datta, A simple kinetic equation for steady-state quantum transport, J. Phys.: Condens. Matter **2,** 8023 (1990).

[63] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, Single and multiband modeling of quantum electron transport through layered semiconductor devices, J. Appl. Phys. **81,** 7845 (1997).

[64] S. Datta, Nanoscale device modeling: The Green's function method, Superlattices Microstruct. **28,** 253 (2000).

[65] A. Svizhenko, M. Anantram, T. Govindan, B. Biegel, and R. Venugopal, Two-dimensional quantum mechanical modeling of nanotransistors, J. Appl. Phys. **91,** 2343 (2002).

[66] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck, Atomistic simulation of nanowires in the $sp^3d^5s^*$ tight-binding formalism: From boundary conditions to strain calculations, Phys. Rev. B **74,** 205323 (2006).

[67] M. Luisier and G. Klimeck, Atomistic full-band simulations of silicon nanowire transistors: Effects of electron-phonon scattering, Phys. Rev. B **80,** 155430 (2009).

[68] M. P. L. Sancho, J. M. L. Sancho, J. M. L. Sancho, and J. Rubio, Highly convergent schemes for the calculation of bulk and surface Green functions, J. Phys. F: Metal Phys. **15,** 851 (1985).

[69] M. Calderara, S. Brück, A. Pedersen, M. H. Bani-Hashemian, J. VandeVondele, and M. Luisier, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis,* SC '15 (Association for Computing Machinery, 2015), pp. 1–12.

[70] R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem* (Dover Publications, New York, NY, USA, 1992), 2nd ed., p. 464.

[71] D. C. Langreth, in *Linear and Nonlinear Electron Transport in Solids,* edited by J. T. Devreese and V. E. van Doren (Springer US, Boston, MA, 1976), pp. 3–32.

[72] M. Frey, A. Esposito, and A. Schenk, in *ESSDERC 2008—38th European Solid-State Device Research Conference* (IEEE, Edinburgh, UK, 2008), pp. 258–261.

[73] K. Rim, S. Narasimha, M. Longstreet, A. Mocuta, and J. Cai, in *Digest. International Electron Devices Meeting* (IEEE, San Francisco, CA, USA, 2002), pp. 43–46.

[74] Y.-M. Niquet, V.-H. Nguyen, F. Triozon, I. Duchemin, O. Nier, and D. Rideau, Quantum calculations of the carrier mobility: Methodology, Matthiessen's rule, and comparison with semi-classical approaches, J. Appl. Phys. **115,** 054512 (2014).

[75] A. Svizhenko and M. Anantram, Role of scattering in nanotransistors, IEEE Trans. Electron Devices **50,** 1459 (2003).

[76] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77,** 3865 (1996).

[77] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50,** 17953 (1994).

[78] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59,** 1758 (1999).

[79] C. Klinkert, A. Szabó, C. Stieger, D. Campi, N. Marzari, and M. Luisier, 2-D materials for ultrascaled field-effect transistors: One hundred candidates under the *Ab Initio* microscope, ACS Nano **14,** 8605 (2020).

[80] A. Rawat, N. Jena, Dimple, and A. D. Sarkar, A comprehensive study on carrier mobility and artificial photosynthetic properties in group VI B transition metal dichalcogenide monolayers, J. Mater. Chem. A **6,** 8693 (2018).

[81] N. Mounet, M. Gibertini, P. Schwaller, D. Campi, A. Merkys, A. Marrazzo, T. Sohier, I. E. Castelli, A. Cepellotti, G. Pizzi, and N. Marzari, Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds, Nat. Nanotechnol. **13,** 246 (2018).

[82] N. Wakabayashi, H. G. Smith, and R. M. Nicklow, Lattice dynamics of hexagonal $MoS_2$ studied by neutron scattering, Phys. Rev. B **12,** 659 (1975).

[83] S. Sze and K. K. Ng, *Physics of Semiconductor Devices* (Wiley, New York, 2007), 3rd ed.

[84] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Atomically thin $MoS_2$: A new direct-gap semiconductor, Phys. Rev. Lett. **105,** 136805 (2010).

[85] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, Optical spectrum of $MoS_2$: Many-body effects and diversity of exciton states, Phys. Rev. Lett. **111,** 216805 (2013).

[86] J. Ma, A. S. Nissimagoudar, and W. Li, First-principles study of electron and hole mobilities of Si and GaAs, Phys. Rev. B **97,** 045201 (2018).

[87] S. Poncé, W. Li, S. Reichardt, and F. Giustino, First-principles calculations of charge carrier mobility and conductivity in bulk semiconductors and two-dimensional materials, Rep. Prog. Phys. **83,** 036501 (2020).

[88] S. Wei and M. Chou, Phonon dispersions of silicon and germanium from first-principles calculations, Phys. Rev. B **50**, 2221 (1994).

[89] T. Sohier, M. Calandra, and F. Mauri, Two-dimensional Fröhlich interaction in transition-metal dichalcogenide monolayers: Theoretical modeling and first-principles calculations, Phys. Rev. B **94**, 085415 (2016).

[90] Z. Yu, Z.-Y. Ong, Y. Pan, Y. Cui, R. Xin, Y. Shi, B. Wang, Y. Wu, T. Chen, Y.-W. Zhang, G. Zhang, and X. Wang, Realization of room-temperature phonon-limited carrier transport in monolayer $MoS_2$ by dielectric and carrier screening, Adv. Mater. **28**, 547 (2016).

[91] X. Cui, G.-H. Lee, Y. D. Kim, G. Arefe, P. Y. Huang, C.-H. Lee, D. A. Chenet, X. Zhang, L. Wang, F. Ye, F. Pizzocchero, B. S. Jessen, K. Watanabe, T. Taniguchi, D. A. Muller, T. Low, P. Kim, and J. Hone, Multi-terminal transport measurements of $MoS_2$ using a van der Waals heterostructure device platform, Nat. Nanotechnol. **10**, 534 (2015).

[92] Y. Liu, H. Wu, H.-C. Cheng, S. Yang, E. Zhu, Q. He, M. Ding, D. Li, J. Guo, N. O. Weiss, Y. Huang, and X. Duan, Toward barrier free contact to molybdenum disulfide using graphene electrodes, Nano Lett. **15**, 3030 (2015).

[93] Z. Yu, Y. Pan, Y. Shen, Z. Wang, Z.-Y. Ong, T. Xu, R. Xin, L. Pan, B. Wang, L. Sun, J. Wang, G. Zhang, Y. W. Zhang, Y. Shi, and X. Wang, Towards intrinsic charge transport in monolayer molybdenum disulfide by defect and interface engineering, Nat. Commun. **5**, 5290 (2014).

[94] A. Sanne, R. Ghosh, A. Rai, H. C. P. Movva, A. Sharma, R. Rao, L. Mathew, and S. K. Banerjee, Top-gated chemical vapor deposited $MoS_2$ field-effect transistors on $Si_3N_4$ substrates, Appl. Phys. Lett. **106**, 062101 (2015).

[95] B. Radisavljevic and A. Kis, Mobility engineering and a metal-insulator transition in monolayer $MoS_2$, Nat. Mater. **12**, 815 (2013).

[96] K. K. H. Smithe, C. D. English, S. V. Suryavanshi, and E. Pop, High-field transport and velocity saturation in synthetic monolayer $MoS_2$, Nano Lett. **18**, 4516 (2018).

[97] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta, Electron drift velocity in silicon, Phys. Rev. B **12**, 2265 (1975).

[98] P. Norton, T. Braggins, and H. Levinstein, Impurity and lattice scattering parameters as determined from Hall and mobility analysis in *n*-type silicon, Phys. Rev. B **8**, 5632 (1973).

[99] M. Fiorentini and N. Bonini, Thermoelectric coefficients of *n*-doped silicon from first principles via the solution of the Boltzmann transport equation, Phys. Rev. B **94**, 085204 (2016).

[100] O. Restrepo, K. Varga, and S. Pantelides, First-principles calculations of electron mobilities in silicon: Phonon and Coulomb scattering, Appl. Phys. Lett. **94**, 212103 (2009).

[101] Swiss National Supercomputing Centre (CSCS), Piz daint supercomputer, 2018, https://www.cscs.ch/computers/piz-daint.

[102] IEEE, IEEE international roadmap for devices and systems (IRDS), 2023, https://irds.ieee.org/.

[103] M. V. Fischetti, L. Wang, B. Yu, C. Sachs, P. M. Asbeck, Y. Taur, and M. Rodwell, in *2007 IEEE International Electron Devices Meeting* (IEEE, Washington, DC, USA, 2007), pp. 109–112.

[104] F. Ducry, M. H. Bani-Hashemian, and M. Luisier, Hybrid mode-space–real-space approximation for first-principles quantum transport simulation of inhomogeneous devices, Phys. Rev. Appl. **13**, 044067 (2020).