Ab initio Methods for Electronic Transport in Semiconductors and Nanostructures

42

Massimo V. Fischetti ⓘ, William G. Vandenberghe ⓘ, Maarten L. Van de Put ⓘ, Gautam Gaddemane ⓘ, and Jingtian Fang ⓘ

Contents

Abstract

The past decade has seen the growing use of *ab initio* methods to study theoretically not only the atomic and electronic structure of semiconductors but also their charge-transport properties. This chapter focuses on "intrinsic" charge transport (limited by scattering with phonons) and starts with a brief historical overview of early work, mostly based at first on the deformation potential theorem and, later, on empirical pseudopotentials and on the rigid-ion approximation, to calculate electron-phonon matrix elements in semiconductors. This historical overview is followed by an outline of the theoretical framework employed when using density functional theory. Having described the full-band Monte Carlo method to solve the Boltzmann transport equation, the chapter presents examples of the use of *ab initio* methods to study the low-field mobility, high-field transport, and device performance in silicon, group-III nitrides, and two-dimensional materials. Throughout the discussion attention is paid to the limitations of *ab initio* methods. Finally, the chapter discusses how *ab initio* methods are used to study the dielectric response of solids and charge transport in the quantum limit.

Massimo V. Fischetti (✉) · W. G. Vandenberghe · M. L. Van de Put · G. Gaddemane
The University of Texas at Dallas, Department of Materials Science and Engineering, Richardson, TX, USA
e-mail: max.fischetti@utdallas.edu;
william.vandenberghe@utdallas.edu;
Maarten.VandePut@utdallas.edu; Gautam.Gaddemane@utdallas.edu

J. Fang
Synopsys, Inc., Silicon Engineering Group, Mountain View, CA, USA
e-mail: jingtian.fang@synopsys.com

Keywords

Density functional theory · Dielectric screening · Electronic transport · Electron-phonon interaction · Quantum transport · Two-dimensional semiconductors

42.1 Introduction

Until recently, the theoretical study of charge transport in semiconductors and semiconductor devices has been based on some level of "empiricism." The library of software tools that constitute the technology computer-aided design (TCAD) in this area may indeed range from simple – but computationally efficient – models based on a drift-and-diffusion approximation of the semiclassical Boltzmann equation (BTE) [1], to models based on a hydrodynamic approximation [2, 3] or even to full-band Monte Carlo methods [4]. Yet, all of these models require some degree of empirical information, such as models for the carrier mobility as a function of doping and temperature, for the energy and momentum relaxation times, or deformation potentials to describe the strength of the interaction between charge carriers and phonons.

This degree of empiricism has been possible – and highly successful – thanks to the wealth of experimental information that has become available, since the late 1940s, about charge transport in silicon, germanium, and III–V compound semiconductors, materials used in the bulk of the micro- and nanotechnology products in logic, memory, and optoelectronics offered by the very-large-scale integration industry (VLSI). However, in the past decade or so, research has shifted to novel materials that may possibly replace or augment these "conventional" semiconductors, such as wide band gap semiconductors for power electronics (diamond, silicon carbide, group-III nitrides, or gallia – Ga_2O_3), or the growing number of two-dimensional (2D) materials that have attracted attention since the advent of graphene [5–8]. In many cases, little information is available about these materials, often not even their stability, and even less information is available about their charge-transport properties (*e.g.*, band gap, effective mass, and carrier mobility).

The rise of these new materials has moved the old dream of predicting the performance of an electronic device exclusively from "first-principles" (or *ab initio*) from the realm of a cultural and academic exercise to a necessity. Therefore, it is perhaps a fortunate coincidence that at the same time theoretical progress and improvement of the computational hardware have rendered *ab initio* methods reliable and predictable *almost* to the point at which experimental information is required not to build new empirical models but to simply confirm the theoretical predictions. (The reasons why we wrote *almost* will be evident below in Sect. 42.7.)

In the past, the use of first-principles methods has been restricted to small systems, mainly focused on their structures, total and formation energies, and bonding configurations but had little or no connection to electronic transport. However, the recent progress we have mentioned above has broadened their range of applications, improved their accuracy, and extended their scope to electronic transport [9, 10].

Density functional theory (DFT) is now routinely used to predict the atomic and electronic structure of new materials, thanks to the wide availability of computer packages, such as the Vienna *Ab initio* Software Package (VASP) [11–14] or Quantum ESPRESSO (QE) [15, 16]. Even the strength of the electron-phonon interaction can now be calculated using DFT using either the finite ion displacements [17] or density function perturbation theory (DFPT) [18, 19]. This constitutes a dramatic improvement since the early "pioneering" days in which the rigid-ion approximation [20] and empirical pseudopotentials were painstakingly used to estimate deformation potentials in Si, intervalley deformation potentials in III-V compound semiconductors [21–24], and used in Monte Carlo transport studies [25–28]. Even transport in open systems has been studied using DFT [29], and such an *ab initio* formalism has also been used to study dissipative transport in the two-dimensional materials of current interest [30].

In this chapter we present a birds-eye view of the historical evolution of "first-principles" methods toward the present state of art. Keeping in mind the scope of this chapter, we shall limit our attention to the use of first-principles methods in the study of electronic transport. Therefore, in this historical overview, we shall pay attention to the deformation potential theorem and the rigid-(pseudo)ion approximation, approximations that are now replaced by genuine *ab initio* methods, but that still play a huge role in research, thanks to their computational efficiency and their sound physical basis. Restricting our attention to density functional theory, in Sect. 42.3 we shall briefly present a general formulation of DFT and its application to the study of the electron-phonon interaction. As important examples of application of the theory, in Sect. 42.4 we shall consider silicon (obviously), considering the electron-phonon scattering rates, both intraband and inter-band; the transport properties of group-III nitrides and selected 2D materials (to follow the present trend), employing DFT to compute the electron-phonon scattering rates and their transport properties (low-field mobility, velocity-field characteristics); and also device characteristics (in some cases) obtained from full-band Monte Carlo simulations based on the band structure and scattering rates obtained from these *ab initio* calculations. We shall then present, in Sect. 42.5, a different application of *ab initio* methods, namely, the calculation of the dielectric response of nanostructures, an important factor that controls the electrostatics of devices as well as dielectric screening of several perturbations. We shall present in Sect. 42.6 a brief overview of the use of pseudopotentials to deal with quantum transport in ultra-scaled devices, considering graphene nanoribbons as examples. In our final section, Sect. 42.7, having assessed the predictive power of *ab initio* methods, we shall consider also their limitations and speculate on what the future may promise.

Throughout this discussion, we consider only the electron-phonon interaction, since it is the subject that has received more attention as it is an intrinsic element in determining the charge-transport properties of electron devices. However, additional scattering processes, such as Coulomb scattering with ionized impurities and roughness at the semiconductor-insulator interface, have also been addressed using *ab initio* methods, as illustrated in Refs. [31, 32], for example. We also refrain from discussing molecular devices, although also in this case first-principles methods have been employed successfully (see Refs. [33, 34], for example).

42.2 Historical Overview

It is hard to define precisely the term "first principles" or *ab initio*. If we take a broad view and define a first-principles' approach as a way of expressing a transport-related quantity (such as the electron mobility, for example) in terms of "higher-level" physical quantities that are not directly related to transport (*e.g.*, a change of the band structure in the presence of vibrations of the crystal, phonons), probably the first example of such an *ab initio* approach is given by Bardeen's and Shockley's "deformation potential" theorem [35].

As we shall see below, the interaction between electrons and phonons can be expressed as the perturbation of the electronic motion caused by the displacements of the ions away from their equilibrium positions, displacements that are associated with the thermal vibrations of the lattice. Therefore, the Hamiltonian that describes this perturbation can be written as:

$$\widehat{H}_{ep} = \int_{\Omega} d\mathbf{r} \, \delta\widehat{U}_{tot}(\mathbf{r})\widehat{\rho}(\mathbf{r}) , \qquad (42.1)$$

having indicated operators with a "hat." The quantity $\delta\widehat{U}_{tot}(\mathbf{r})$ is the change of the total crystal potential at position $\mathbf{r}$ due to the presence of phonons and $\widehat{\rho}(\mathbf{r})$ is the electron density (that is, $\widehat{\psi}^{\dagger}(\mathbf{r})\widehat{\psi}(\mathbf{r})$ in terms of the (second-quantized) electron field $\widehat{\psi}$). The term $\delta\widehat{U}_{tot}(\mathbf{r})$ can be written explicitly in terms of the displacement, $\delta\widehat{\mathbf{R}}_{l,\alpha,\mathbf{q}}^{(\eta)}$ of each ion α in the unit cell l caused by a phonon of wave vector $\mathbf{q}$ of branch η (longitudinal or transverse, acoustic or optical):

$$\delta\widehat{U}_{tot}(\mathbf{r}) = \sum_{l,\alpha,\eta,\mathbf{q}} \frac{\delta E_{tot}(\mathbf{r})}{\delta\mathbf{R}_{l,\alpha}} \cdot \delta\widehat{\mathbf{R}}_{l,\alpha,\mathbf{q}}^{(\eta)} \, e^{i\mathbf{q}\cdot\mathbf{r}} . \qquad (42.2)$$

The quantity $\delta E_{tot}(\mathbf{r})/\delta\mathbf{R}_{l,\alpha}$ represents the change of the total energy of the crystal at position $\mathbf{r}$ when the ion α in cell l is displaced from its equilibrium position by a shift $\delta\mathbf{R}_{l,\alpha}$. (In the following we shall express $\mathbf{R}_{l,\alpha}$ as $\mathbf{R}_l + \boldsymbol{\tau}_{\alpha}$, where $\boldsymbol{\tau}_{\alpha}$ is the position of ion α in the unit cell.) In second quantization, the ionic displacement can be expressed in terms of the

phonon creation and annihilation operators $\widehat{b}_{\mathbf{q}}^{(\eta)\dagger}$ and $\widehat{b}_{\mathbf{q}}^{(\eta)}$, respectively:

$$\delta\widehat{\mathbf{R}}_{l,\alpha,\mathbf{q}}^{(\eta)} = \left(\frac{\hbar}{2\rho_x\omega_{\mathbf{q}}}\right)^{1/2} \mathbf{e}_{\mathbf{q}}^{(\eta)} \left(\widehat{b}_{\mathbf{q}}^{(\eta)\dagger} + \widehat{b}_{-\mathbf{q}}^{(\eta)}\right) , \qquad (42.3)$$

where $\hbar$ is the reduced Planck's constant and $\mathbf{e}_{\mathbf{q}}^{(\eta)}$ is the phonon polarization vector. The use of first-order perturbation theory is standard practice in our context. Therefore, the matrix element for this process is given by (see Eq. (42.29) below):

$$\langle n', \mathbf{k} + \mathbf{q} \, | \delta\widehat{U}_{tot}(\mathbf{r}) | \, n, \mathbf{k} \rangle , \qquad (42.4)$$

where for simplicity we have denoted by $|n, \mathbf{k}\rangle$ a state in Fock space that contains a (Bloch) electron in band n and with wave vector $\mathbf{k}$. We have also omitted to write explicitly the phonons states, as we shall do in the following, since we shall assume them to be at thermal equilibrium.

42.2.1 The Deformation Potential Theorem

In the early days of research of semiconductor physics (around the late 1940s or 1950s), it was impossible to calculate the matrix element given by Eq. (42.4). However, Bardeen and Shockley viewed the *nonpolar* scattering of electrons with acoustic phonons (the process that limits the carrier mobility in Si and Ge), for example as caused by the local change of the energy of the conduction band minimum (or valence band maximum for the hole-phonon process) due to the local strain wave associated with the propagation of a phonon [35]. For example, under an isotropic (hydrostatic) dilatation/compression of the lattice, the lattice constant changes locally from a to $a+u$. This shifts the energy of the conduction band minimum. Therefore, as a dilatation/compression wave propagates in the crystal (*i.e.*, an acoustic phonon), an electron sees a "scattering potential":

$$\Delta E_c = \frac{dE_c}{da}u . \qquad (42.5)$$

More generally, since a displacement $\mathbf{u}$ of the medium corresponds to a local change $\Delta\Omega \approx \Omega\nabla\cdot\mathbf{u}$, of the volume Ω, we have (omitting the operator "hats" for simplicity):

$$\delta U_{tot}(\mathbf{r}) \sim \Delta E_c(\mathbf{r})$$
$$= \Omega\frac{dE_c}{d\Omega}\nabla\cdot\mathbf{u}(\mathbf{r}) \sim \Delta_{ac} \, \mathbf{q}\cdot\mathbf{u}_{\mathbf{q}} \, e^{i\mathbf{q}\cdot\mathbf{r}} , \qquad (42.6)$$

where the last expression considers a single Fourier component $\mathbf{u}_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{r}}$ of the displacement $\mathbf{u}(\mathbf{r})$. The quantity $\Delta_{ac} = \Omega dE_c/d\Omega$ is called the "deformation potential". From Eq. (42.3) we see that the presence of an acoustic phonon

of wavelength $\lambda = 2\pi/q$ and frequency $\omega_{\mathbf{q}}$ (where $\mathbf{q}$ is the phonon wave vector) implies a shift of the atomic positions of magnitude:

$$u_{\mathbf{q}} \approx \left(\frac{\hbar}{2\rho_{\mathrm{x}}\omega_{\mathbf{q}}}\right)^{1/2} , \qquad (42.7)$$

(where ρ_{x} is the crystal mass density). Thus, the scattering potential becomes:

$$\Delta E_{\mathrm{c}} \, q u_{\mathbf{q}} \, \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \approx \Delta E_{\mathrm{c}} q \left(\frac{\hbar}{2\rho_{\mathrm{x}}\omega_{\mathbf{q}}}\right)^{1/2} \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} , \qquad (42.8)$$

so that, for longitudinal phonons, the squared electron-phonon matrix element can be approximated as:

$$|\langle \mathbf{k} + \mathbf{q}|\widehat{H}_{\mathrm{ac}}|\mathbf{k}\rangle|^2 \approx \frac{\hbar\Delta_{\mathrm{ac}}^2 q^2}{2\rho_{\mathrm{x}}\omega_{\mathbf{q}}} \left(N_{\mathbf{q}} + \frac{1}{2} \pm \frac{1}{2}\right) , \qquad (42.9)$$

where the plus (minus) sign applies to emission (absorption) processes and $N_{\mathbf{q}}$ is the Bose-Einstein population of phonons of momentum $\hbar\mathbf{q}$ at thermal equilibrium. Therefore, considering that for small energies electrons emit or absorb only small-q acoustic phonons, we can assume a linear phonon dispersion with a slope given by the sound velocity υ_{s}, that is, $\omega_{\mathbf{q}} = \upsilon_{\mathrm{s}}q$. We can also assume that the phonon energy is much smaller than the thermal energy $k_{\mathrm{B}}T$ (where k_{B} is the Boltzmann constant), so that $N_{\mathbf{q}} \approx k_{\mathrm{B}}T/(\hbar\upsilon_{\mathrm{s}}q) \gg 1$ and scattering can be treated as an elastic process. Thus, finally, the momentum relaxation rate becomes simply:

$$\frac{1}{\tau_{\mathrm{ac}}(E)} = \frac{2^{1/2}\Delta_{\mathrm{ac}}^2 m^{*3/2}k_{\mathrm{B}}T}{\pi\hbar^4\rho_{\mathrm{x}}\upsilon_{\mathrm{s}}^2}E^{1/2} , \qquad (42.10)$$

and the electron mobility can now be calculated with the only unknown parameter Δ_{ac} that can be obtained experimentally from measurements of the band gap of the crystal as a function of hydrostatic compression. Therefore, a transport quantity, the electron mobility, is now linked to a more fundamental quantity, the deformation potential, which is not directly related to charge transport.

In this oversimplified picture, electrons can couple only to longitudinal acoustic (LA) phonons, since the anisotropy of the band structure of silicon and germanium near the band extrema has been ignored. A few years later, Herring and Vogt [36] accounted for this effect and provided an approximated form for the deformation potentials for the interaction of electrons with LA and transverse acoustic (TA) phonons in terms of the dilatation deformation potential, Ξ_{d}, of the uniaxial shear deformation potential, Ξ_{u}, and of the angle $\theta_{\mathbf{q}}$ between the wave vector $\mathbf{q}$ of the phonon and the longitudinal axis of the ellipsoidal equi-energy surfaces:

$$\begin{aligned} \Delta_{\mathrm{LA}} &\approx \Xi_{\mathrm{d}} + \Xi_{\mathrm{u}} \cos^2 \theta_q \\ \Delta_{\mathrm{TA}} &\approx \Xi_{\mathrm{u}} \sin \theta_q \, \cos \theta_q . \end{aligned} \qquad (42.11)$$

42.2.2 The Rigid (Pseudo)ion

When Bardeen and Shockley presented their deformation potential theorem, there was only scant and often incorrect experimental information about both the band structure and the transport properties of silicon and germanium. However, in the mid-to-late 1960s, as more reliable experimental information became available and computer technology continued to progress, it became possible to perform full-band structure calculations, empirical pseudopotentials being at the time the preferred choice.

Developed for bulk semiconductors of the diamond and zinc blende structure at first [37,38], the pseudopotential only needs "form factors" for a small set of wave vectors coinciding with reciprocal lattice $\mathbf{G}$ vectors. For nanostructures, full functional forms $V_{\alpha q}$ were developed to study strained materials and nanostructures that require larger "supercells," since in both cases knowledge of $V_{\alpha q}$ is required for arbitrary values of q. Details about these functional forms for various semiconductors, as well as details about their applications to study electronic transport in nanostructures, can be found in Refs. [39,40].

These empirical pseudopotentials were used to calculate the phonon-induced narrowing of the band gap at high temperatures [41–44], finally giving a solid (if not *ab initio*) justification of the empirical models that were in use (see, *e.g.*, Ref. [45]). This problem has been revisited recently using methods that can be labeled genuine *ab initio* calculations [46]. However, this interesting problem falls outside the scope of this chapter.

Moreover, and more to the point of our discussion, these empirical pseudopotentials were used to improve upon the deformation potential approximation: once the empirical pseudopotential $V_\alpha(\mathbf{r})$ of an ion α is known, Eq. (42.2) can be re-expressed in terms of the Fourier components of the ionic pseudopotential, $V_{\alpha,\mathbf{p}}$, assuming that the ionic potential shifts rigidly:

$$\begin{aligned} &\delta\widehat{U}_{\mathrm{tot}}(\mathbf{r}) \\ &\approx \sum_{l,\alpha,\eta,\mathbf{q},\mathbf{p}} \mathrm{i}\, V_{\alpha,\mathbf{p}} \, \mathrm{e}^{\mathrm{i}\mathbf{p}\cdot(\mathbf{r}+\mathbf{R}_{l,\alpha})}\mathbf{p} \cdot \delta\widehat{\mathbf{R}}_{l,\alpha,\mathbf{q}}^{(\eta)} \, \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} . \end{aligned} \qquad (42.12)$$

Using some semiempirical models to obtain the phonon dispersion and polarization vectors, such as the valence shell model [47,48]), Zollner and Cardona were able to calculate the scattering rates for intervalley scattering in several III–V compound semiconductors [21–24], thus confirming the ex-

perimental results obtained by sub-picosecond pump-probe measurements [49].

At the same time, empirical pseudopotentials were beginning to be used not only to calculate the electron-phonon interaction (so, the electron dynamics) but also to extend the validity of transport models to a range of high energy in which the universally used effective mass approximation clearly failed. This was of tremendous practical interest because of the reliability issues caused by hot electrons in the ever-shrinking devices of the VLSI technology. The major breakthrough occurred in 1981, when Shichijo and Hess [50] developed a Monte Carlo approach to account for the electron kinematics (energy and group velocity) using the band structure of GaAs calculated using local empirical pseudopotentials. The electron-phonon scattering rates were treated empirically starting from the *Ansatz* that their energy dependence must reflect the density of (final) states, an *Ansatz* whose validity has been confirmed recently (see Ref. [51] and references therein). This "full-band" approach was later extended by Tang and Hess to silicon [52, 53], by Brennan and Hess to several other III–V compound semiconductors [54] and by Fischetti and Laux to simulate not only homogeneous transport but also inhomogeneous and realistic semiconductor devices [4].

Obviously, silicon has been the subject of studies performed by several groups also employing empirical pseudopotentials and empirical "deformation potentials" fitted to available experimental data on high-energy transport, such as velocity-field characteristics at high fields [55, 56], impact ionization [52, 55, 57–59], and injection into the SiO_2 gate insulator. Theoretical support for these semiempirical full-band models has been obtained by using empirical pseudopotentials (and Harris potentials) to calculate the electron-phonon scattering rates in Si and other semiconductors within the rigid-ion approximation [25–28], as discussed recently in Ref. [51]. In Fig. 42.1 we show the total electron-phonon scattering rates averaged over equi-energy surfaces and bands plotted as a function of electron kinetic energy as calculated using empirical deformation potentials [4, 55, 56] compared to quasi-*ab initio* calculations using Harris potentials [26–28, 56] and empirical pseudopotential [25].

More recently, the rigid-(pseudo)ion method has been extended to 2D materials, specifically to calculate the electron-phonon scattering rates in free-standing graphene [60]. Figure 42.2 shows the electron-phonon scattering rates calculated at 300 K for a free-standing sheet. The mobility and velocity-field characteristics obtained using these results (as reported in Ref. [60]) are in good agreement with available experimental data and show that even relatively poorly known materials (at least, less known than Si) can be studied successfully using these models.

The discussion presented so far can be viewed as a snapshot of the state of the art before DFT began to replace these
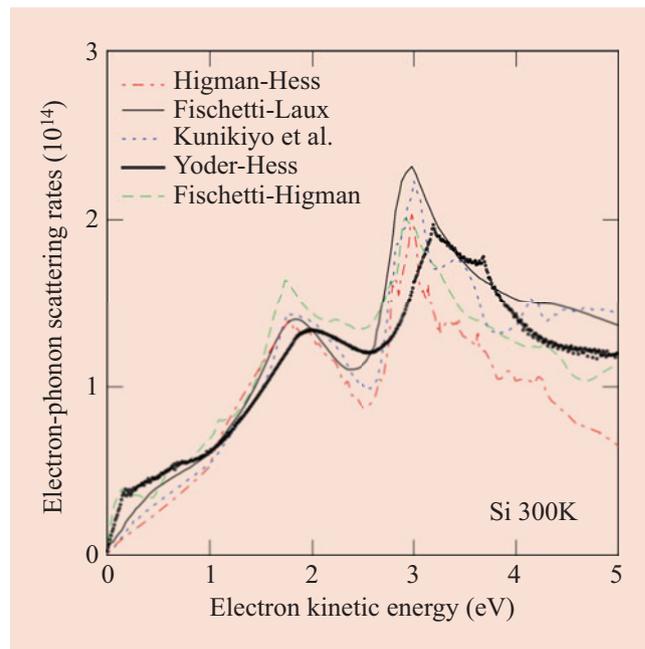
Fig. 42.1 Electron-phonon scattering rates calculated using the rigid-(pseudo)ion model and empirical pseudopotentials [25, 55] (labeled "Fischetti-Higman" and "Kunikiyo et al."), Harris potentials [27, 56] (labeled "Higman-Hess" and "Yoder-Hess") compared to rates calculated using empirically determined deformation potential calibrated to reproduce known experimental data [4] (labeled "Fischetti-Laux")

quasi-empirical methods. Most of the research has focused on electron transport in silicon, and undeniably the empirical nature of these models (via the calibrated pseudopotentials and simple models for the phonon spectra and polarization) limits their predictive power. Therefore, exploring novel materials requires the more fundamental approach inherent to genuine *ab initio* methods that we discuss next. Nevertheless, the very empirical nature of these pseudopotentials guarantees "by definition" an accurate description of the excited spectrum of semiconductors (viz., band gap and effective masses), as required when attempting to assess the performance of devices based on "known" materials. Considering also their relative numerical simplicity, these quasi-first-principles method are still used in useful applications.

42.3   Theoretical Framework

In this chapter we restrict our discussion to plane-wave DFT together with (self-consistent) pseudopotentials. However, we should at least mention the fact that ballistic electronic transport has been studied also using all-electron methods [62, 63] by Polizzi's group [64, 65], among others, thus avoiding the need to "pseudize" the system. Moreover, at least when studying relatively small molecules, all-electron methods have also been used without resorting to approxi-
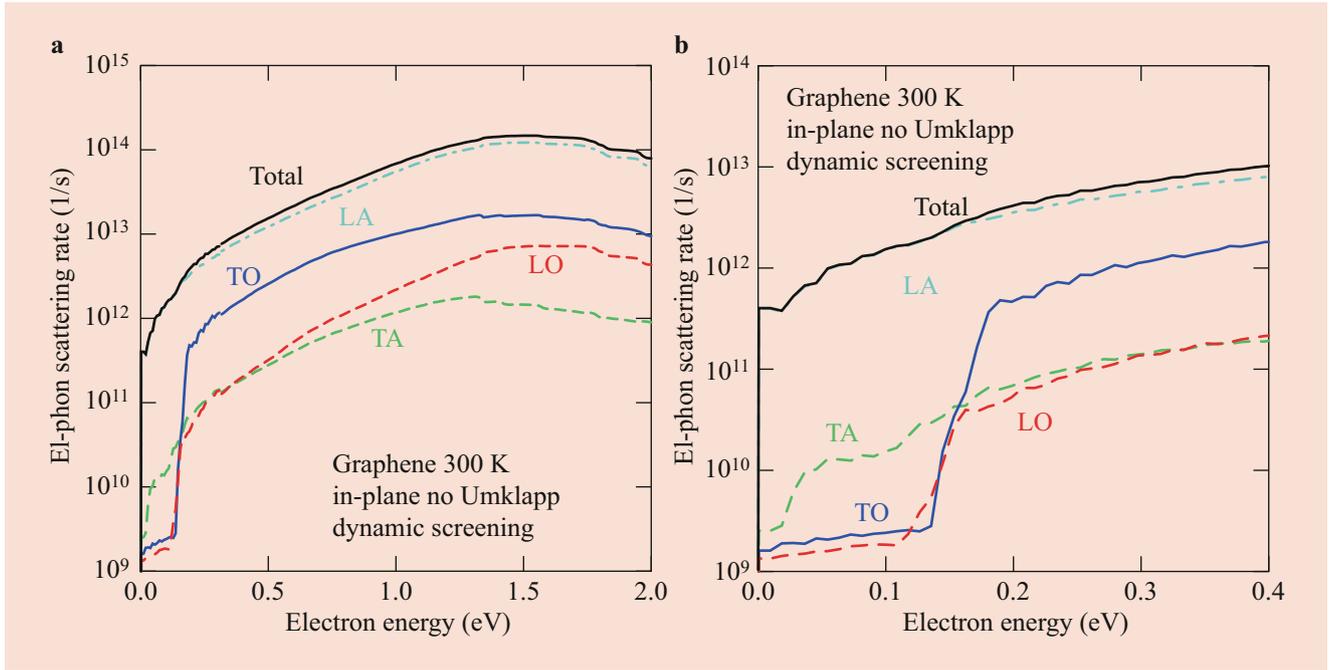
Fig. 42.2 (**a**) calculated total electron-phonon scattering rates in graphene at 300 K using the dynamic screening model by Wunsch et al. [61]. (**b**) as in the left frame but expanding the low-energy region.

(After M. V. Fischetti et al. [60], with permission from the Institute of Physics ©2013 Institute of Physics)

mated exchange and correlation functionals but employing, instead, the GW method (see Ref. [66], for example).

The maturity of DFT can be judged by the number of software packages that are presently available. Most of them are based on plane waves or projector-augmented plane waves (VASP [11–14], Quantum ESPRESSO [15, 16], GPAW [67, 68], CASTEP [69], ABINIT [70]), others on localized orbitals (such as SIESTA [71, 72], or ATK [73, 74]). As we had noted in the Introduction, DFT has been born with the goal of studying the atomic and electronic structure of solids. Only recently some of these computer packages have been extended to handle electronic transport (such as TranSIESTA [75] or Quantum Wise-ATK [73, 74] or EPW [76, 77]). Here we shall focus only in DFT based on a plane-wave basis and look exclusively at electronic transport.

Modern DFT implementations start from the self-consistent solution of the Kohn-Sham equation [14, 15, 78, 79]:

$$H_e \psi_i = \left[-\frac{\hbar^2}{2m} \nabla^2 + V_{eff}(\mathbf{r}) \right] \psi_i = E_i \psi_i(\mathbf{r}) , \quad (42.13)$$

where H_e is the electron Hamiltonian whose first term represents the kinetic energy and V_{eff} represents the effective Kohn-Sham potential:

$$V_{eff} = [V_{ion}(\mathbf{r}) + V_h(\mathbf{r}) + V_{xc}(\mathbf{r})] . \quad (42.14)$$

Here $V_{ion}(\mathbf{r})$ is the sum of the ionic potentials ($-Z_j e^2/(4\pi \epsilon_0 |\mathbf{r} - \mathbf{R}_j|)$) originating from all atoms, with valence Z_j and located at $\mathbf{R}_j$, in the system. The second term is the Hartree potential, capturing the effect of the interaction of all electrons with each other:

$$V_h = \int \frac{e^2}{4\pi \epsilon_0 |\mathbf{r} - \mathbf{r}'|} \rho(\mathbf{r}') d\mathbf{r}' , \quad (42.15)$$

using SI units (e is the magnitude of the electron charge and ϵ_0 is the vacuum permittivity). The Hartree potential does not fully capture the interaction of electrons with each other. This is where DFT comes with the final term on the left-hand side, the exchange-correlation potential:

$$V_{xc}(\mathbf{r}) = \frac{\partial E_{xc}[\rho]}{\partial \rho(\mathbf{r})} , \quad (42.16)$$

where $E_{xc}[\rho(\mathbf{r})]$ is the exchange-correlation functional. The final element required to generate a closed set of equations, is the charge density $e\rho(\mathbf{r})$ which is simply calculated by summing over all occupied Kohn-Sham wavefunctions:

$$\rho(\mathbf{r}) = 2 \sum_i^{N/2} |\psi_i(\mathbf{r})|^2 . \quad (42.17)$$

The prefactor 2 accounts for spin degeneracy, and the sum runs over $N/2$ wavefunctions where N is the total number of electrons in the system.

Equations (42.13–42.17) provide a set of equations which can be solved for any atomic configuration, provided an exchange-correlation functional is selected. Some minor and major numerical issues appear in a practical implementation. Several approximations are generally in order. We discuss two important approximations that we use to study the electron-phonon interaction in crystals.

First, we consider ideal crystals; this means that an infinite number of atoms, arranged in a periodic configuration, is present in the system under consideration. The periodic system is described by a set of three basis vectors, spanning the unit cell, and an additional set of vectors specifying the location of all composing atomic species inside the unit cell, $\boldsymbol{\tau}_\alpha$. In Eq. (42.17), we introduced i as a generic quantum number. In a crystal, the ionic potential is periodic, and the quantum number i actually consists of a wave vector $\mathbf{k}$ and a band index n. The charge density is then determined by summing over an appropriate k-space discretization, possibly exploiting crystal symmetry, and a number of bands defined by the number of electrons in the system. We also introduced a second generic index j for the position of each atom; however, from now on we will use $\mathbf{R}_{l,\alpha}$ to denote the position of ion α in unit cell l.

Second, a numerical all-electron solution, solving Eq. (42.14) directly, is possible but requires a very fine discretization of the wavefunctions. To deal with systems of sufficiently large size, an approximation needs to be introduced that significantly reduces the computational burden. The need for a very fine discretization stems from the $-Z_j e^2 / |\mathbf{r} - \mathbf{R}_{l,\alpha}|$ contributions to the ionic potential, which are singular at $\mathbf{R}_{l,\alpha}$. There are also a large number of core electronic states, which remain identical regardless of the crystal under consideration. A first simplification to reduce the computational burden is to compute only the charge density of the valence electrons, *i.e.*, the electrons which are not core electrons, and to incorporate the impact of core electrons through a modification of the (bare) ionic potential $V_{\text{ion}}(\mathbf{r})$. However, even when calculating only the valence states, rapid wavefunction oscillations remain near the center location of the ion $\mathbf{R}_{l,\alpha}$. To avoid these rapid oscillations, a pseudopotential approximation is introduced.

We omit the details of the implementation of modern pseudopotential-like approximations, such as the projector-augmented wavefunction approach [80, 81], but the general picture of these approaches is captured by a wavefunction composed of plane waves:

$$\psi_{\text{pw}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \qquad (42.18)$$

where $\mathbf{G}$ are reciprocal lattice vectors. The number of reciprocal lattice vectors $\mathbf{G}$ determines the size of the Hamiltonian in its numerical implementation. To realize maximal computational efficiency, approaches try to minimize the number of reciprocal lattice vectors while still faithfully capturing the electronic behavior of the system. The plane-wave wavefunctions are not actual wavefunctions and need to be augmented to accurately capture the rapidly oscillating nature of the actual wavefunction near the core. Overall, the pseudopotential approach gives excellent results and limits the number of plane waves required for an accurate solution quite efficiently.

Once the Kohn-Sham wavefunctions are found, the ground-state energy of the crystal is determined as:

$$\begin{aligned}
E = {} & \sum_i \langle \psi_i | - \frac{\hbar^2 \nabla^2}{2m} | \psi_i \rangle \\
& + \frac{e^2}{2} \int \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} \\
& + \int \mathrm{d}\mathbf{r}\, V_{\text{ion}}(\mathbf{r})\rho(\mathbf{r}) + E_{\text{xc}}[\rho] + E_{\text{ions}} \qquad (42.19)
\end{aligned}$$

where the sum over i runs over all valence band states. This ground-state energy is a function of the atomic positions and is a measure of the stability of the structure under study, *i.e.*, a lower energy means a more favorable atomic configuration. The favorable/unfavorable nature of the configuration can be quantified through the calculation of the force acting on each atom:

$$\mathbf{F}_{l,\alpha} = -\frac{\partial E}{\partial \mathbf{R}_{l,\alpha}}. \qquad (42.20)$$

This force can be evaluated from the Kohn-Sham wavefunctions using the Hellmann-Feynman theorem [82] and comes at minimal computational cost once the Kohn-Sham equation is solved.

This means that once an initial atomic configuration is provided, the energy of the atomic configuration can be computed as well as the force acting on each atom. Based on these forces, a new atomic configuration can be constructed and again the energy and the forces can be computed. Repeating this process and using an optimization algorithm like the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [83], the atomic configuration with minimal energy can be determined. This process is referred to as "structural relaxation," and at the end of relaxation, the crystal is in its minimal energy configuration and the force acting on each atom vanishes.

Next, we are interested in the possible lattice vibrations in the system: the phonons. To determine the phonon spectrum, the second-order force-constant tensor is required:

$$\Phi_{l\alpha l'\alpha'} = \frac{\partial \mathbf{F}_{l,\alpha}}{\partial \mathbf{R}_{l'\alpha'}} \qquad (42.21)$$

to construct the dynamical matrix. Unfortunately, the calculation of the second-order force constants is much more expensive compared to the calculation of the forces, *i.e.*, the first-

order force constants acting on each atom. The reason why the evaluation of the first-order constants is computationally cheap is because the ground-state energy minimizes the energy with respect to the charge density: $\partial E[\rho(\mathbf{r})]/\partial \rho(\mathbf{r}) = 0$. This enables the evaluation of the first-order force constants without computing $\Delta\rho(\mathbf{r}) = \partial\rho(\mathbf{r})/\partial\mathbf{R}_{l,\alpha}$. For the second-order force constants on the other hand, $\Delta\rho(\mathbf{r})$ has to be computed.

Similarly, the final quantity of interest for the electron-phonon interaction is the change in potential due to the displacement of an atom:

$$\frac{\partial(V_{\text{ion}} + V_{\text{h}} + V_{\text{xc}})}{\partial\mathbf{R}_{l,\alpha}}. \tag{42.22}$$

Just like for the second-order force constants, more information than is available from the solutions of the Kohn-Sham equation is required and $\Delta\rho(\mathbf{r})$ must be calculated.

Two methods are available to compute the second-order force constants and the change in potential due to atomic displacements: the density functional perturbation theory (DFPT) method and the finite-displacement method. Both methods are equivalent if implemented with sufficient accuracy; we give a brief overview of both methods next sections.

42.3.1 Density Functional Perturbation Theory

The density functional perturbation theory [16,84] approach introduces a new set of equations that have to be solved simultaneously to determine the change in the Kohn-Sham wavefunctions $\Delta\psi_n$ and the change in the energies ΔE_n under the application of a perturbation of the ionic potential $\Delta V_{\text{ion}}(\mathbf{r})$:

$$(H_{\text{e}} - E_n)\,|\Delta\psi_n\rangle = -(\Delta V_{\text{eff}} - \Delta E_n)\,|\psi_n\rangle, \tag{42.23}$$

$$\Delta V_{\text{eff}}(\mathbf{r}) = \Delta V_{\text{ion}}(\mathbf{r}) + e^2 \int \frac{\Delta\rho(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|}\mathrm{d}\mathbf{r}'$$
$$+ \left.\frac{\mathrm{d}v_{\text{xc}}(\rho)}{\mathrm{d}\rho}\right|_{\rho=\rho(\mathbf{r})}\Delta\rho(\mathbf{r}), \tag{42.24}$$

where:

$$\Delta\rho(\mathbf{r}) = \sum_i \psi_i^*(\mathbf{r})\Delta\psi_i(\mathbf{r}). \tag{42.25}$$

These three equations have to solved simultaneously for every displacement of every relevant perturbation of the ionic potential under consideration. Assuming the computational burden to solve the DFPT equations (42.23–42.25) is similar to that of solving the Kohn-Sham equation (42.13–42.17); computing the phonon spectrum and the electron-phonon interaction strength is nominally $3N_{\text{ion}}N_q$ more expensive

where N_{ion} counts the number of ions in each unit cell and N_q counts the number of phonon wave vectors of interest. Computing the phonon dispersion (or spectrum) is relatively straightforward for crystals with smaller unit cells, especially when symmetry further reduces the computational expense [85,86]. On the other hand, determining the phonon spectrum of crystals with large unit cells quickly becomes very expensive.

Having determined $\Delta\rho(\mathbf{r})$ and $\Delta V_{\text{eff}}(\mathbf{r})$, the second-order force constants as well as the change of the potential can be determined, and we can proceed with the calculation of the phonon spectrum or the electron-phonon matrix elements. To save computational resources, interpolation methods [76] can be used to interpolate calculated phonon energies and electron-phonon matrix elements onto a finer grid of phonon wave vectors.

42.3.2 Finite-Displacement Method

In the finite-displacement method [17,87], atoms are slightly displaced away from their equilibrium position $\mathbf{R}_{l,\alpha}^{(0)}$ to a new position $\mathbf{R}_{l,\alpha}^{(0)} + \delta\mathbf{R}$. The whole procedure of solving the Kohn-Sham equation (42.13–42.17) is repeated for a whole set of displacements, and for every displacement the forces are computed. The second-order force constants and the potential change due to atom displacement can then be found using the finite difference method.

For example, to determine the derivative of the total energy required to evaluate the electron-phonon matrix elements, the Kohn-Sham equation is solved when the atom $l\alpha$ is slightly displaced from its equilibrium position $\mathbf{R}_{l'\alpha'}^{(0)}$ along the positive and negative j direction, $\mathbf{R}_{l'\alpha'} = \mathbf{R}_{l'\alpha'}^{(0)} \pm \epsilon\mathbf{e}_j$, where $\mathbf{e}_j$ is a unit vector along the direction j and ϵ is a small number. This gives us the total crystal energies, E_{tot}^{j+} and E_{tot}^{j-}, after the positive and negative displacement along the j Cartesian direction. Repeating the procedure for the other two Cartesian directions and using the central difference scheme, the j component of the quantity $\delta E_{\text{tot}}/\delta\mathbf{R}_{l,\alpha}$ appearing in Eq. (42.2) (or ΔV_{eff} in Eq. (42.27) below) can be approximated by the expression:

$$\left.\frac{\delta E_{\text{tot}}}{\delta\mathbf{R}_{l,\alpha}}\right|_j \approx \frac{E_{\text{tot}}^{j+} - E_{\text{tot}}^{j-}}{2\epsilon}. \tag{42.26}$$

Moreover, the (j, j') element of the second-order force constant (i.e., the (j, j') element of the dynamical matrix) can be obtained in a similar way, although somewhat more laboriously, by solving the Kohn-Sham equation when both the atom $l\alpha$ and the atom $l'\alpha'$ are slightly displaced from their equilibrium position along the positive and negative Cartesian axes, thus obtaining the second-order change, $\Delta^{(2)}E_{\text{tot}ll'\alpha\alpha'}$, which is central difference approximation to the dynamical matrix $\Phi_{jl\alpha,j'l'\alpha'}$, Eq. (42.21). In practice, such

a cumbersome and time-consuming procedure is streamlined by using the Hellmann-Feynman theorem, as mentioned above.

One important consideration for the finite-displacement method is the interaction between periodic images. When displacing an atom in one unit cell in the simulation, the equivalent atom in every unit cell is displaced in exactly the same way in all other unit cells. This generates spurious interactions between neighboring cells. To minimize these spurious interactions, it is important to perform the calculations on a larger unit cell. For example, in the extreme case of a crystal which has only one atom in its primitive unit cell, no change in the energy will be observed when moving the atom in the primitive unit cell. Instead, the primitive unit cell has to be duplicated forming a supercell that is larger in size by an integer factor γ in each direction. The supercell now contains γ^3 atoms instead of 1. A larger γ will result in reduced spurious interaction but will unfortunately also rapidly increase computational memory requirements.

Both the finite-displacement method and the density functional perturbation approach will give the same quantities, provided calculations are performed with sufficient accuracy. The choice of one method over the other may be motivated by convenience or by a more efficient execution on the available computational resources. The finite-displacement method has the advantage that it only needs the solution of the Kohn-Sham equation and only relies on the most mature part of DFT codes. Parallelization is straightforward in the finite-displacement method, and the calculation over all displacements can be performed as parallel DFT calculations. Interpolation onto a finer grid may also be more straightforward using the finite difference method compared to the DFPT method. The DFPT approach may however be more resource efficient and require fewer iterations than a new solution of the Kohn-Sham equation for every displacement.

42.3.3 Electron-Phonon Interaction

The final object of interest is the electron-phonon interaction Hamiltonian which is computed from ΔV_{eff} (which we can identify with $\delta \widehat{U}_{\text{tot}}$; see Eqs. (42.1), (42.2), and (42.4)):

$$
\widehat{H}_{\text{e-ph}} = \int d\mathbf{r}\, \widehat{\rho}(\mathbf{r}) \Delta V_{\text{eff}}(\mathbf{r}) \approx
$$
$$
\times \int d\mathbf{r}\, \widehat{\rho}(\mathbf{r}) \sum_{l,\alpha,\eta,\mathbf{q}} \frac{\delta E_{\text{tot}}(\mathbf{r})}{\delta \mathbf{R}_{l,\alpha}} \cdot \delta \widehat{\mathbf{R}}^{(\eta)}_{l,\alpha,\mathbf{q}}\, e^{i\mathbf{q}\cdot\mathbf{r}},
$$
$$
(42.27)
$$

where $\delta \widehat{\mathbf{R}}^{(\eta)}_{l,\alpha,\mathbf{q}}$ is the ion displacement operator expressed in Eq. (42.3) in terms of the annihilation and creation operators, $\widehat{b}^{(\eta)}_{\mathbf{q}}$ and $\widehat{b}^{(\eta)\dagger}_{\mathbf{q}}$, for a phonon in branch η with wave vector $\mathbf{q}$ and

frequency $\omega^{(\eta)}_{\mathbf{q}}$. Using a finite-volume normalization, here we define the amplitude of the displacement $d^{(\eta)}_{\mathbf{q}\alpha}$ as:

$$
d^{(\eta)}_{\mathbf{q}\alpha} = \sqrt{\frac{\hbar}{2NM_\alpha \omega^{(\eta)}_{\mathbf{q}}}}
\qquad (42.28)
$$

where M_α is the mass of ion α and N is the number of unit cells in the lattice.

Consider an electron in band n with wave vector $\mathbf{k}$ that makes a transition to band m with wave vector $\mathbf{k}+\mathbf{q}$ while emitting a phonon. For this electron, the matrix element between states allowed by Pauli's principle reads:

$$
\langle m\mathbf{k}+\mathbf{q} \mid \widehat{H}_{\text{e-ph}}\, \widehat{b}^{(\eta)\dagger}_{\mathbf{q}} \mid n\mathbf{k}\rangle = \sum_{l,\alpha} d^{(\eta)}_{\mathbf{q}\alpha} e^{i\mathbf{q}\cdot\mathbf{R}_{l,\alpha}} \mathbf{e}^{(\eta)}_{\mathbf{q}\alpha}
$$
$$
\cdot \left(\int_\Omega d\mathbf{r}\, \psi^*_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}) \frac{\delta E_{\text{tot}}}{\delta \mathbf{R}_{l,\alpha}}\, \psi_{n\mathbf{k}}(\mathbf{r}) \right). \qquad (42.29)
$$

Invoking the Bloch theorem, $\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$, and for Eq. (42.29) to hold, the periodic part $u_{n\mathbf{k}}(\mathbf{r})$ must be normalized on the unit cell.

Equation (42.29) is proportional to the square root of the inverse of the phonon frequency $(\omega^\eta_{\mathbf{q}})^{-1/2}$ through the magnitude of the displacement. A smoother quantity measuring electron-phonon interaction strength is the deformation potential which can be implicitly defined using:

$$
\langle m\mathbf{k}+\mathbf{q} \mid \widehat{H}_{\text{e-ph}}\, \widehat{b}^{(\eta)\dagger}_{\mathbf{q}} \mid n\mathbf{k}\rangle
$$
$$
= \frac{1}{\sqrt{\Omega}} \sum_\eta DK^{(\eta)}_{nm\mathbf{k}\mathbf{q}} \sqrt{\frac{\hbar}{2\rho \omega^{(\eta)}_{\mathbf{q}}}}, \qquad (42.30)
$$

where we introduced the mass density $\rho_{\text{x}} = M_{\text{cell}}/\Omega_{\text{cell}}$ and the volume of the unit cell Ω_{cell}. The volume of the crystal Ω is then $N\Omega_{\text{cell}}$. Combining Eqs. (42.29) and (42.30), the deformation potential is:

$$
DK^{(\eta)}_{\mathbf{k}\mathbf{q}mn} = \sum_\alpha \mathbf{M}_{mn\mathbf{k}\mathbf{q}\alpha} \cdot \mathbf{e}^{(\eta)}_{\mathbf{q}\alpha} \sqrt{\frac{M_{\text{cell}}}{M_\alpha}}, \qquad (42.31)
$$

where

$$
\mathbf{M}_{mn\mathbf{k}\mathbf{q}\alpha} = \int_\Omega d\mathbf{r}\, u^*_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}) \Delta V_{\text{eff}} u^*_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{q}\cdot(\tau_\alpha - \mathbf{r})}, \quad (42.32)
$$

where we should recall that τ_α is the position of ion α in the unit cell.

One subtlety in the calculation and interpretation of deformation potentials arises when initial or final electronic states are degenerate or when the phonon branches are degenerate. In case of degeneracy, the band or branch index is

not sufficient to uniquely determine the state and denoting the deformation potential as "$D_{nm\mathbf{kq}}^{(\eta)}$" is ambiguous. For example, this is the case for the threefold degenerate valence band at Γ in many popular materials, *e.g.* silicon.

Similarly, many materials have twofold degenerate TA and TO phonons. In the case of degeneracy, depending on which linear combination of the valence band wavefunctions is taken and the direction of the phonon displacement vectors, different values for "$D_{nm\mathbf{kq}}^{(\eta)}$" are obtained. One simple way to remove the ambiguity is to take the square root of the sum of all squared deformation potentials for degenerate initial or final electronic states, or degenerate phonons. An alternative choice is the largest deformation potential upon rotation of any wavefunctions or phonon displacement vectors within the subspace of the degeneracy.

42.3.4 *Ab initio* Simulation of Electronic Transport: The Monte Carlo Method

The Monte Carlo method is the most accurate method to study electron transport at the semiclassical level. It is a well-established method, originally proposed by Kurosawa in the context of semiconductors [88] that has been described in detail in Refs. [89–91]. Its use in conjunction with a description of the band structure of semiconductors obtained from empirical pseudopotentials has been pioneered by Hess' group [50, 52, 53] and by the IBM group [4].

In the context of *ab initio* studies of charge transport, the same numerical methods used in these latter references can be used, the major difference consisting in the use of DFT to obtain the band structure, of the electron-phonon scattering rates obtained using the finite-displacement method or DFPT. Moreover, the "exact" matrix elements given by Eqs. (42.4) and (42.27) are used together with the full phonon dispersion to select the final state after each collision.

Whereas a fully detailed presentation of the full-band Monte Carlo method goes beyond the scope of this chapter, here we present a brief outline.

We have employed the synchronous ensemble Monte Carlo method since it is suitable to study time-dependent transport in inhomogeneous systems, such as electron devices. In this method, the motion of an ensemble of particles ($\approx$300 up to 50,000 particles, depending on whether one considers homogeneous transport or transport in a realistically large device), is simulated subjecting them to the externally applied electric field (possibly obtained self-consistently with Poisson equation) and to the given scattering mechanisms. An initial state $\mathbf{k}$, band n, and position $\mathbf{r}$ are assigned to each particle stochastically. The motion of each particle is then evolved for a constant duration of time Δt (free flight) conforming to the equations of motion: $\hbar d\mathbf{k}/dt = \mp e\mathbf{F}(\mathbf{r})$ (the sign depending on the type of carriers under study, electrons of holes) and $d\mathbf{r}/dt = \upsilon_n(\mathbf{k})$, where $\mathbf{F}(\mathbf{r})$ is the electric field at position $\mathbf{r}$ and $\upsilon_n(\mathbf{k})$ is the carrier group velocity in band n, $\nabla E_n(\mathbf{k})/\hbar$, $E_n(\mathbf{k})$ being the kinetic energy of the carrier. In synchronous simulations, the duration of the free flight, Δt, is chosen based on the maximum scattering rate in the energy interval under consideration. At the end of their free flight, a random scattering mechanism is chosen stochastically. Based on the selected scattering mechanism, a new $\mathbf{k}$ state is also chosen stochastically for the particles according to the probability distribution given by the density of final states (calculated using the Gilat-Raubenheimer algorithm [92, 93], obviously modified to deal with transport in 2D materials [39]) multiplied by the squared matrix element of the scattering process. This is then treated as the initial state for the next free flight.

Quantities of interest, such as average velocity, energy, carrier density, and currents at the contacts, are recorded at the end of each Monte Carlo iteration. This process is repeated until steady state is reached or until the end of a transient (such as the turn-on of a transistor). In full-band simulations, the energy and group velocity are calculated (in our case, using DFT) on a discretized Brillouin zone, and stored in look up tables, to solve the equations of motion in free flight. The scattering rates are calculated and tabulated on the same mesh.

Calculations performed assuming a homogeneous electric field are usually performed to obtain information about the "intrinsic" transport properties of a material, namely, the low-field mobility and velocity-field characteristics. In order to minimize the stochastic noise at low field (the drift velocity being of the same order of magnitude of – or even much smaller than – the thermal velocity), the low-field mobility, μ_θ, is calculated from the diffusion constant, D_θ, along the chosen direction θ, using the Einstein relation. We should note that in the case of 2D materials with a Dirac-like electron dispersion, μ_θ, has to be extracted from D_θ using a generalization of the Einstein relation [60]:

$$\mu_\theta = \frac{eD_\theta}{k_B T} \int_0^\infty d\eta \, \frac{\eta \, e^{\eta-\eta_F}}{[1 + e^{\eta-\eta_F}]^2}$$
$$\times \left(\int_0^\infty d\eta \, \frac{\eta}{[1 + e^{\eta-\eta_F}]} \right)^{-1} , \qquad (42.33)$$

where now $\eta_F = E_F/(k_B T)$ denotes the Fermi energy in thermal units. The diffusion constant is evaluated from the Monte Carlo simulator:

$$D_\theta = \frac{1}{2} \frac{d}{dt} \left\langle (x_\theta - \langle x_\theta \rangle)^2 \right\rangle , \qquad (42.34)$$

where $\langle x_\theta \rangle$ is the time-dependent ensemble-average position along the direction θ of electrons initially at the origin, $\mathbf{r} = 0$, diffusing in the absence of an electric field.

The simulation of electronic transport in inhomogeneous conditions, such as in a transistor, is performed in a similar way. In this case the electric field $\mathbf{F}(\mathbf{r})$ is obtained from a numerical solution of the Poisson equation. In "atomistic" simulations, like those considered below in Sect. 42.6, the problem arises of how to define dielectric interfaces (such as the channel/gate insulator interface). In Sect. 42.5 we shall provide an *ab initio* solution of this problem.

42.4 Silicon, Group-III Nitrides, and 2D Materials

42.4.1 Silicon

It is of the utmost interest to show how electron-phonon scattering rates calculated with the DFT method just discussed compare with the semiempirical rates shown in Fig. 42.1 and discussed in Sect. 42.2. Since these results have been verified against a wealth of experimental data (see Ref. [51]), this constitutes a stringent test for *ab initio* methods.

In Fig. 42.3 we show the results obtained using the Quantum ESPRESSO DFT computer program [15, 16] with a
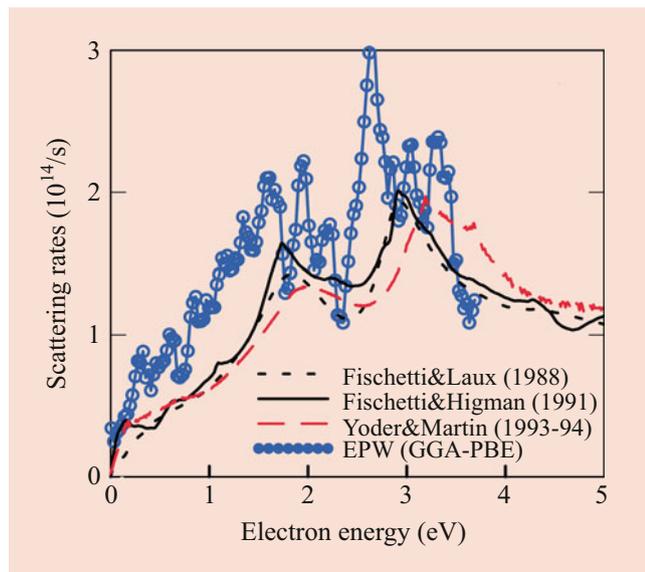
Fig. 42.3 Electron-phonon scattering rates at 300 K calculated following Refs. [76, 77] (EPW) compared to those obtained within the rigid-ion approximation [25–28] and those employed in the Monte Carlo simulations reported in Ref. [4]. The data represent an average over all "initial" wave vectors (distributed on a uniform mesh in the Brillouin zone) and bands at a given energy. The "noise" affecting the EPW data is the result of the coarser mesh used. The lines connecting the symbols are only a guide to the eye. (After M. V. Fischetti et al. [51], with permission from the American Institute of Physics ©2019 American Institute of Physics)

kinetic energy cutoff of 35 Ry using the Perdew-Burke-Ernzerhof (PBE) [94] generalized gradient approximation (GGA) exchange-correlation functionals, norm-conserving pseudopotentials [95], a uniform grid of $12 \times 12 \times 12$ **k**-points for the self-consistent charge density calculation, and of $6 \times 6 \times 6$ **q**-points for the calculation of the phonon spectra. The 300 K electron-phonon scattering rates have been calculated using the Wannier interpolation method implemented by the EPW package [76] on a finer grid of $18 \times 18 \times 18$ **k**-points with a Gaussian smearing of 10 meV.

When looking at Fig. 42.3, we should stress that the results of Refs. [4, 25–28] have been averaged over constant-energy shells with weight given by the density of initial states using the algorithm proposed by Gilat and Raubenheimer [92, 93]. On the contrary, the EPW scattering rates shown here have been "smoothed" by averaging the raw data over the **k**-mesh described above over energy bins $\sim$27 meV wide. This permits a straightforward reproduction of these results.

The overall agreement is quite remarkable. However, as we shall discuss in Sect. 42.7, electron transport is very sensitive to variations of the band structure (such as effective masses and energy of satellite valleys, when present) that are usually considered "small errors" in DFT. For example, an error of the order of $k_B T$ ($\approx$25 meV at room temperature) for the energy of satellite valleys in some 2D TMDs may be considered negligible when looking at optical absorption spectra but may cause very large errors in determining the electron mobility if intervalley scattering dominates transport.

To assess the sensitivity of the computed mobility and velocity-field characteristics to different choices of pseudopotentials and exchange-correlation functionals, we have calculated the mobility *(i)* using the GGA-PBE approximation and the less accurate local density approximation [95], both using the same norm-conserving pseudopotentials [95], and *(ii)* using the norm-conserving and ultrasoft pseudopotentials [96], both within the PBE-GGA approximation.

The results of this exercise are shown in Fig. 42.4. A significant spread of the data is quite evident. We shall discuss below, also in Sect. 42.7, how numerical issues may affect the results. This spread has been also discussed in Ref. [97] in the case of black phosphorus monolayers (phosphorene) and by Poncé et al. [77] in a general case. Accurate scattering rates can be obtained only when employing an extremely fine discretization of the Brillouin zone. This is certainly possible when considering only the "low-energy pockets" that are need to calculate the low-field carrier mobility, as indeed done in Ref. [77]. However, a fine discretization remains numerically challenging when considering the entire Brillouin zone as required when studying high-energy (high-field) transport. The practical need for a coarser discretization results in an additional and significant numerical uncertainty that affects the results. Therefore, at present we cannot yet expect a perfect quantitative or conclusive agreement with experimental data.
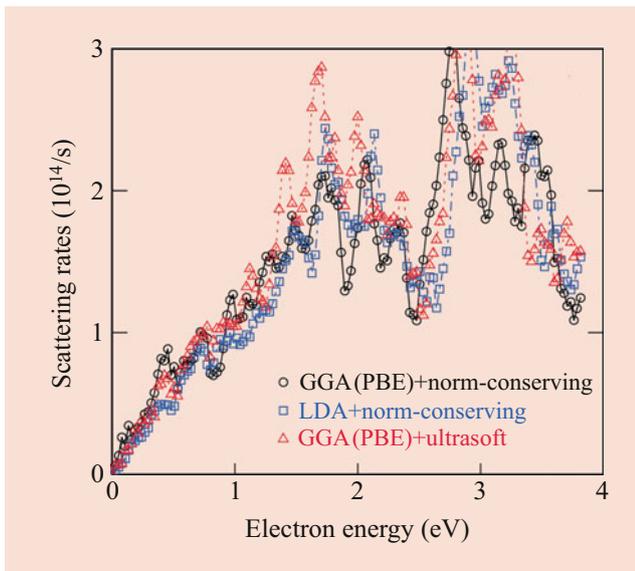
Fig. 42.4 As in the previous figure, but now the scattering rates shown have been calculated using different pseudopotentials and exchange-correlation functionals, as indicated. As in Fig. 42.3, the data represent an average over all "initial" wave vectors (distributed on a uniform mesh in the Brillouin zone) and bands at a given energy. The lines connecting the symbols are only a guide to the eye. (After M. V. Fischetti et al. [51], with permission from the American Institute of Physics ©2019 American Institute of Physics)
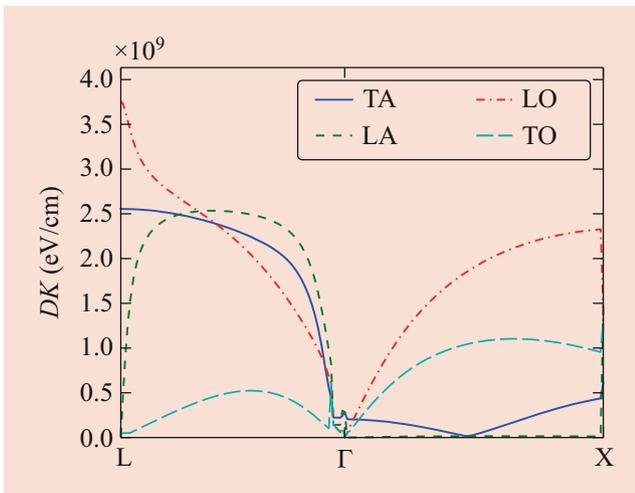
Fig. 42.5 Deformation potentials in bulk silicon for an electron starting at the valence band at Γ, transition to the conduction band along one of the Γ-L or Γ-X symmetry lines. For BTBT, the values at 0.85 in the Γ-X direction are the relevant quantity, and we find $D_{TA} = 3.5 \times 10^8$ eV/cm, $D_{TO} = 1.0 \times 10^9$ eV/cm, and $D_{LO} = 2.3 \times 10^9$ eV/cm. (After W. G. Vandenberghe et al. [17], with permission from the American Institute of Physics ©2015 American Institute of Physics)

As an another important example of deformation potentials obtained from first principles, we consider the deformation potentials that determine the phonon-assisted inter-band tunneling exploited in tunnel field-effect transistors (tFETs). In this case, we have used the Vienna *ab initio* simulation

package (VASP) [13] to solve the Kohn-Sham equation. We have used the finite-displacement method to find the phonon frequencies, polarization vectors, and electron-phonon *inter-band* matrix elements. The full details of the calculation can be found in Ref. [17].

In Fig. 42.5 we exhibit the calculated deformation potential for an electron transitioning from the top of the valence band to the first conduction band. We compute the interaction with all phonon branches along the Γ-L and the Γ-X high symmetry lines. Silicon has six phonon branches, but the two transverse branches (TA and TO) are twofold degenerate. Inspecting Fig. 42.5, we see that only the interaction with the TA vanishes. Values of the deformation potentials are quite large and measure up to 4×10^9 eV/cm.

The deformation potentials relevant for BTBT are found at 0.85 in the $\Gamma-X$ direction, and their calculated values are $D_{TA} = 3.5 \times 10^8$ eV/cm, $D_{TO} = 1.0 \times 10^9$ eV/cm, and $D_{LO} = 2.3 \times 10^9$ eV/cm. These values are much larger than previous theoretical estimates [98], but this is in line with the recent experimental estimate $D_{TO} \approx 10$–13×10^8 eV/cm from Ref. [99].

These results are very relevant to many practical applications since inter-band tunneling determines the probability of band-to-band transitions in silicon. In indirect-gap semiconductors, such as Si, band-to-band transitions (BTBT) are assisted by phonons and proceed from the valence band maximum (Γ) to the conduction band minimum, located near the X symmetry point.

42.4.2 Group-III Nitrides

Since wide band gap semiconductors are very important for power electronics, it is important to study their transport properties not just near thermal equilibrium, as it is assumed when calculating the low-field carrier mobility. Studies need to account for the high kinetic energy that electrons can reach at the high electric fields that are present in high-voltage devices. Group-III nitrides in the wurtzite phase constitute a key family of semiconductors that are employed in power devices. They have been investigated extensively theoretically using full-band Monte Carlo simulations based on empirical pseudopotentials and with electron-phonon scattering rates calculated using constant deformation potentials calibrated to experimental data, similarly to what had been done for Si before the "advent" of DFT, as discussed in Sect. 42.2. As in all empirical approaches, these studies – most notably those described in Refs. [100, 101] – must rely on experimental data that are often hard to obtain and interpret.

Most notably, the value of the all-important energy of the satellite valleys at the symmetry point L in GaN plays a crucial role in determining its high-field transport properties. Experimental data [102] and theoretical calculations using calibrated empirical pseudopotentials and DFT (VASP) [103]

have resulted in vastly different values for this quantity (0.8–1 eV experimentally, from Refs. [104, 105], 2.1–2.3 eV theoretically, from Refs. [106, 107]). This has resulted, in turn, in diverging opinions regarding the role of Auger recombination in these devices, a process that plays a major role at high injection levels in power devices. Therefore, it is of utmost interest to investigate this issue theoretically, coupling DFT and Monte Carlo studies to experimental data and establish the correctness of the theory compared to experiments. Here we shall not go into this controversial issue, but we shall see how *ab initio* methods can be used in two group-III nitrides of high interest in this context, GaN (with a band gap of ∼3.4 eV at 300 K [108]) and AlN (with a band gap of ∼6.026 eV at 300 K [109]).

The study has been performed obtaining the electronic band structure and the phonon dispersion using Quantum ESPRESSO [15, 16] with norm-conserving pseudopotentials [110] and the local density approximation (LDA) for the exchange-correlation functional. The dynamical matrix determined using DFPT was used to obtain phonon spectra, and the electron-phonon scattering rates have been calculated using the "electron-phonon coupling using Wannier functions" (EPW) program [76, 111, 112]. The maximally localized Wannier functions are obtained from the Wannier90 software package [113].

GaN

The band structure of wurtzite GaN along selected high symmetry lines is shown in Fig. 42.6a. We also list in Table 42.1 the energies of the satellite valleys compared with previous results [100, 101, 104–107, 114, 115], (one should recall the controversy we have mentioned above regarding the energy of the satellite L-valleys.) Rather than relaxing the structure,

we have used the experimentally measured lattice constants $a = 3.215$ Å and $c = 5.241$ Å [116].

Table 42.2 shows the ability of DFT to give reliable values for the effective masses. These are extracted from the curvature of the electron dispersion at the bottom of the first-conduction band valleys. The effective mass at the Γ point of the first conduction band computed from DFT is 0.2 m_0 (where m_0 is the free electron mass). This value is in good agreement with the experimental values reported by Drechsler [117] (0.2 m_0) and Witowski [118] (0.222 m_0). The effective mass at the Γ point in the second conduction band is only slightly larger, 0.28 m_0.

Figure 42.6b shows the phonon dispersion of the wurtzite GaN. The longitudinal and transverse sound velocities are 7930 m/s and 3994 m/s, respectively [125]. Note the rather large energy of the highest-energy optical phonon, about 91.2 meV at Γ. The calculated polar and nonpolar scattering with optical and acoustic phonons are shown in Fig. 42.7a, b, respectively. Note that the polar scattering rate with acoustic phonons (*i.e.*, piezoelectric scattering) has a strength that is very similar to strength of the polar scattering with optical phonons at low energies. This shows that DFT correctly confirms the well-known fact that piezoelectric scattering plays an important role in determining the low-field electron mobility even at room temperature, unlike other III–V compound semiconductors like GaAs or InAs. In these figures we show also the scattering rates calculated with the analytical formulas presented in Ref. [90] to fit the first-principles scattering rates. These "analytic" scattering rates have been calculated using an effective mass of 0.2 m_0, a nonparabolicity factor of 0.58 eV^{-1}, and an intravalley deformation potential in the Γ-valley of ∼4.0 eV. For the nonpolar scattering rates with optical phonons, we have used an optical deformation
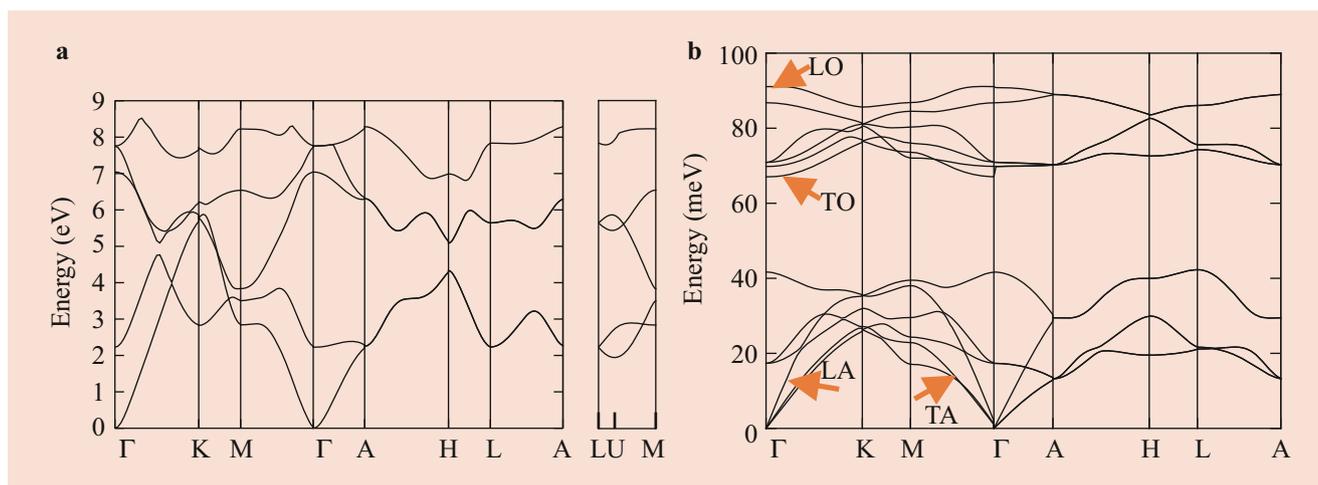
Fig. 42.6 (**a**) Calculated energy dispersion of the first five conduction bands of wurtzite GaN along several high symmetry lines; (**b**) Full phonon dispersion along several high symmetry lines of wurtzite GaN. The longitudinal and transverse acoustic phonon modes and the long-range transverse and longitudinal optical phonons are indicated by arrows. (After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society)

Table 42.1 The conduction band energies at different satellite valleys (in units of eV) in wurtzite GaN and AlN

		K	U	L	A	Γ^3	M
GaN	This work	2.82	1.95	2.23	2.25	2.23	2.84
	LDA+G_0W_0 (Ref. [107])	3.06	2.32	2.60	2.43	2.50	3.27
	GW (Ref. [106])	3.1		2.5	2.6	2.4	3.0
	EPM(Ref. [114])	1.59	1.34			2.14	1.87
	EPM (Ref. [100])	3.08		2.43	2.37	2.35	3.05
	Expt. (Ref. [104])			0.90 ± 0.08			
	Expt. (Ref. [105])			0.97 ± 0.02			
	Expt. (Ref. [115])		1.34				
AlN	This work	0.64	0.61	0.92	2.11	2.27	1.33
	LDA+G_0W_0 (Ref. [107])	0.84	0.67	0.99	2.28	2.65	1.57
	GW (Ref. [106])	0.9		1.1	2.5	2.5	1.6
	EPM (Ref. [114])	0.9	1.05			2.49	1.68
	EPM (Ref. [100])	1.26		1.32	2.33	2.46	1.95

After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society The conduction band energy at the Γ-valley of the first conduction band is set as the zero reference energy

Table 42.2 Electron effective mass (in units of the free electron mass m_0) along different crystalline directions at the first conduction band of different valleys and acoustic phonons′ sound velocities (in units of m/s) of GaN

	This work	Other work
$m_\Gamma^\star(\Gamma \to K)$	0.20	0.18[a], 0.36[b]
$m_K^\star(K \to \Gamma)$	0.64	0.47[c]
$m_K^\star(K \to M)$	0.65	
$m_M^\star(M \to K)$	0.21	0.565[c]
$m_M^\star(M \to \Gamma)$	4.18	
$m_\Gamma^\star(\Gamma \to M)$	0.20	0.18[a], 0.33[b], 0.2[c], 0.2[d], 0.21[e], 0.2283[f], 0.2[g], 0.23[h]
$m_\Gamma^\star(\Gamma \to A)$	0.18	0.20[a], 0.27[b], 0.1846[f], 0.19[h]
$m_A^\star(A \to \Gamma)$	−0.54	
$m_A^\star(A \to H)$	0.37	
$m_U^\star(U \to L/M)$	0.33	0.442[a], 0.285[b], 0.25[c]
$m_L^\star(L \to H)$	0.32	
$m_L^\star(L \to A)$	1.02	
$m_A^\star(A \to L)$	0.37	
$c_s^{LA} (\Gamma \to A)$	7930	7641[c], 6600[i]
$c_s^{TA} (\Gamma \to A)$	3994	4110[c], 2700[i]

After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society
[a]Ref. [119], LDA
[b]Ref. [120], LCAO
[c]Ref. [114], EPM
[d]Ref. [121]
[e]Ref. [122]
[f]Ref. [100], EPM
[g]Ref. [117], Expt
[h]Ref. [123], LMTO
[i]Ref. [124]
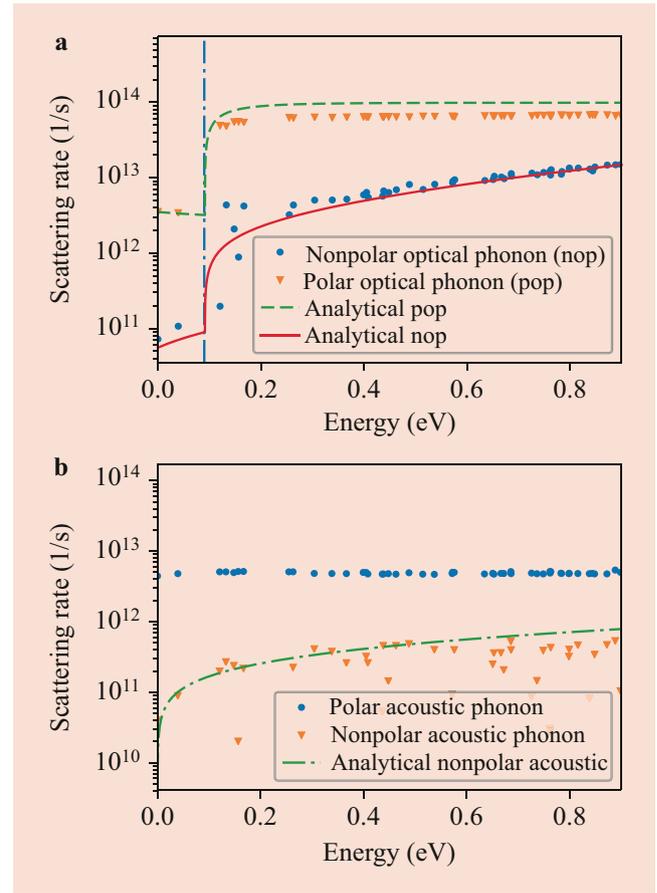
Fig. 42.7 Room-temperature (300 K) polar and nonpolar scattering rates with optical (**a**) and acoustic phonons (**b**) in wurtzite GaN calculated from first principles. A fine q-mesh of 80×80×80 and a Gaussian smearing of 0.1 eV are used in the calculations. The scattering rates calculated using analytical formula are also shown in (**a**) and (**b**). The fitting parameters are given in the text. (After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society)

potential of 3.8×10^8 eV/cm. Such analytic expressions are a convenient way to simplify transport calculations when one wished to avoid the use of full-band Monte Carlo simulations. A similar analytic formula – that approximates quite well the DFT results – can be found to approximate the polar scattering rate with longitudinal optical (LO) phonons using the conventional Fröhlich expression [126] assuming a static relative dielectric constant of 8.9 and a high-frequency relative dielectric constant of 5.35 [127].

Figure 42.8 shows the low-field and high-field transport characteristics along the Γ-K symmetry line in the basal plane. These have been obtained by solving the Boltzmann transport equation with the full-band Monte Carlo method outlined in Sect. 42.3.4 with first-principles band structure and electron-phonon scattering rates. Our simulation results show a very high peak of the drift velocity of about 2.8×10^7 cm/s. Figure 42.8 also shows the average electron en-

ergies as a function of the homogeneous electric field. For electric fields smaller than 150 kV/cm, the average electron energies are close to the thermal energy. Above 150 kV/cm, the average electron energy increases sharply. It increases slowly again when the field strength exceeds 500 kV/cm. The evolution of the particle valley occupancy is the underlying reason for the trend. The trend shown here is similar to the results from the multi-valley Monte Carlo simulation results [124, 128].

The distribution of MC particles occupying the first and the second conduction bands in the first Brillouin zone is shown in Fig. 42.9 for a given electric field strength of 800 kV/cm. For particles (representing electrons) occupying the first conduction band, most of the particles are located around the Γ-valley with momenta exhibiting a noticeable drift along the reverse direction of the applied electric field, with a few particles located around the 12 satellite U-valleys. For particles occupying the second conduction band, most particles are distributed along the Γ-A symmetry line. Some particles occupy also the L-valleys. The particle occupation of the valleys is consistent with their energy splitting from the conduction band minimum. A two-dimensional plot in the $(k_x - k_y)$ basal plane is given for the convenience.

AlN

As we have done in the case of GaN, we show in Fig. 42.10a the band structure of the wurtzite AlN. Similarly, we list in Table 42.1 the conduction band energies of several satellite valleys. The experimental lattice constants of wurtzite AlN used in the calculations are $a = 3.110$ Å and $c = 4.980$ Å [127]. Table 42.3 lists the electron effective masses in different valleys of the lowest conduction band, compared with

previous results [100, 119, 120, 123, 129]. Also in this case, the results obtained from DFT are in satisfactory agreement with the available experimental information. Note that for the first conduction band at the Γ-point, the effective mass is close to an isotropic value of 0.31 m_0, the same value used by Albrecht et al. [129]. From experimental measurements, it has been estimated that the electron effective mass ranges from 0.233 m_0 to 0.336 m_0 [130]. The effective mass for the second conduction band at Γ is 0.36 m_0.

Figure 42.10b illustrates the phonon dispersion. The LO energy at Γ-valley is ~110 meV. The longitudinal and the transverse acoustic-phonon sound velocities are 10,877 m/s and 5880 m/s, respectively, listed in Table 42.3. Figure 42.11a

Fig. 42.8 Characteristics of the electric field versus average electron kinetic energy (left axis) and versus drift velocity (right axis) for wurtzite GaN and AlN along the Γ-K high symmetry line in the basal plane. The data for GaN are plotted with solid black lines and the data for AlN are plotted with dot-dashed red lines. (After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society)

Fig. 42.9 Simulated distribution of MC particles occupying the first conduction band (**a**) and the second conduction band (**b**) of wurtzite GaN shown in reciprocal space for an electric field strength of 800 kV/cm. (After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society)
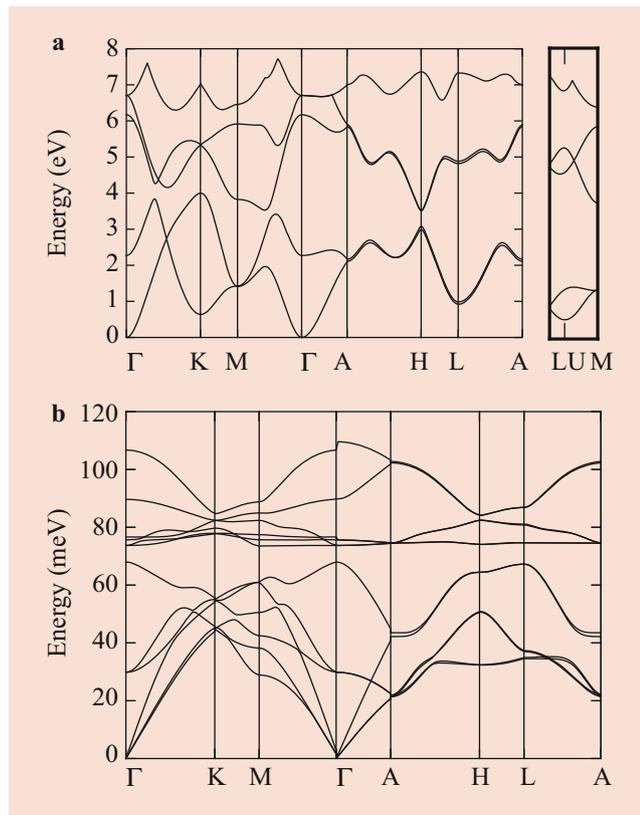
Fig. 42.10 (**a**) The energy dispersion of the first five conduction bands along several high symmetry lines of wurtzite AlN. (**b**) Full phonon dispersion along several high symmetry lines of wurtzite AlN. (After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society)

Table 42.3 Electron effective mass (in units of the free electron mass m_0) along different crystalline directions at the first conduction band of different valleys and acoustic phonons' sound velocities (in units of m/s) of AlN

	This work	Other work
$m_\Gamma^\star (\Gamma \rightarrow K)$	0.31	0.42[a]
$m_K^\star (K \rightarrow \Gamma)$	0.64	0.54[b], 0.488[c]
$m_K^\star (K \rightarrow M)$	0.65	0.54[b]
$m_M^\star (M \rightarrow K)$	0.25	0.629[c]
$m_M^\star (M \rightarrow \Gamma)$	1.20	
$m_\Gamma^\star (\Gamma \rightarrow M)$	0.31	0.40[a], 0.31[b], 0.26[c], 0.35[d], 0.25[e], 0.3433[f]
$m_\Gamma^\star (\Gamma \rightarrow A)$	0.33	0.33[a], 0.35[b], 0.33[e], 0.2938[f]
$m_A^\star (A \rightarrow \Gamma)$	−0.64	
$m_A^\star (A \rightarrow H)$	0.55	
$m_U^\star (U \rightarrow L/M)$	0.43	0.39[a], 0.495[c]
$m_L^\star (L \rightarrow H)$	0.40	
$m_L^\star (L \rightarrow A)$	0.89	
$m_A^\star (A \rightarrow L)$	0.55	
$c_s^{LA} (\Gamma \rightarrow A)$	10,877	
$c_s^{TA} (\Gamma \rightarrow A)$	5880	

After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society
[a]Ref. [120], LCAO
[b]Ref. [129]
[c]Ref. [114], EPM
[d]Ref. [123], LMTO
[e]Ref. [119], LDA
[f]Ref. [100], EPM

presents the first-principles scattering rates for polar and nonpolar scattering with optical phonons. Figure 42.11a presents both polar and nonpolar scattering with acoustic phonons. We have used a Gaussian smearing of 0.1 eV to deal with the energy-conserving Dirac-delta functions. Interestingly, the rate for polar scattering with acoustic phonons is of the same order of magnitude as the rate for polar scattering with optical phonons for electrons with energies below 110 meV. Above 110 meV, the polar scattering rate with optical phonons is about ten times larger than the piezoelectric scattering rate. The scattering rates calculated from the analytical formula are also plotted in Fig. 42.11a, b. In order to fit these rates to the DFT results, we have used a relative static dielectric constant of 9.14 and a high-frequency relative dielectric constant of 4.84 for polar scattering with optical phonons [131]. Using a nonparabolicity factor of 0.35 eV^{-1}, we have obtained an "effective" nonpolar acoustic deformation potential of 7.0 eV and a nonpolar optical deformation potential of 1.32×10^9 eV/cm.

The velocity-field and average carrier energy-field characteristics are compared in Fig. 42.8 with those of GaN. An electron mobility of ~450 cm^2/(V · s) at room temperature

is extracted from the low electric field region. Taniyasu et al. have measured a room-temperature electron mobility of 426 cm^2/(V · s) for n-type AlN with Si doping concentration of 3×10^{17} cm^{-3} [132]. The peak drift velocity of AlN is smaller than that of GaN, and the corresponding critical field is larger. This is due to the heavier Γ-valley effective mass of AlN and also the smaller U-valley minima. We show in Fig. 42.12 the occupation of electrons at the first and the second conduction bands in reciprocal space at the field strength of 800 kV/cm. Particles populating the Γ-valley exhibit the expected momentum shift along the reverse direction of the electric field. The occupation of the satellite U- and K-valleys is larger in AlN than in GaN, a result of the smaller energy splitting from the conduction band minimum. The L-valleys and the M-valleys in the second conduction band are also occupied at large electric fields.

As "dry" as they may be, the approach and result that we have presented in this section – discussed in more detail in Ref. [133] – hopefully will shed some light on the still open issue of the energy of the satellite L-valleys in GaN, on the role of Auger recombination and its inverse process, impact ionization, in power devices based on group-III/nitrides semiconductors. The uncertainty (and controversy) we have
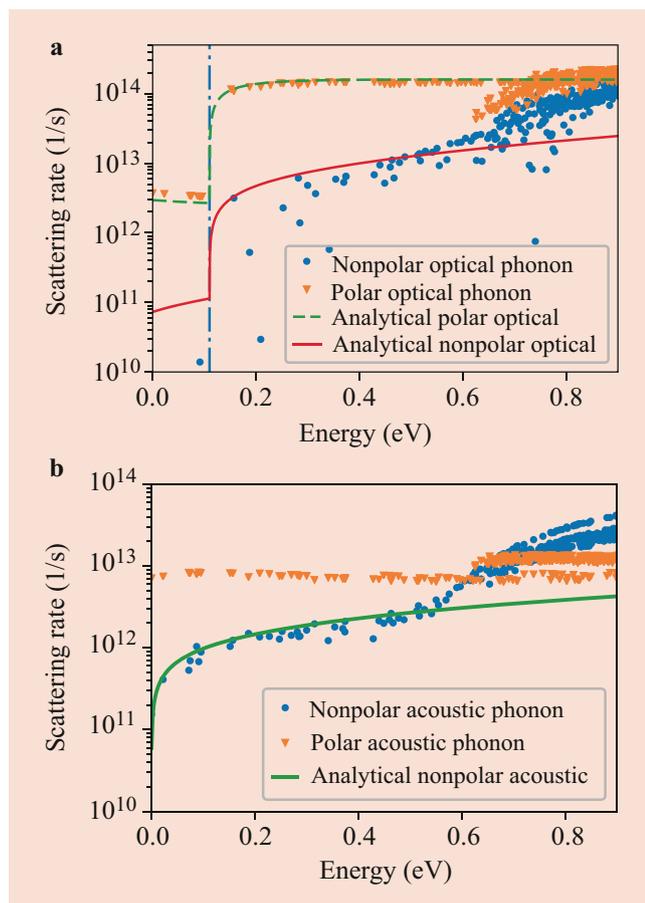
Fig. 42.11 Room-temperature (300 K) polar and nonpolar scattering rates with optical (**a**) and acoustic phonons (**b**) in wurtzite AlN calculated from first principles. A fine q-mesh of $40 \times 40 \times 40$ and a Gaussian smearing of 0.1 eV are used in the calculations. (After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society)

Fig. 42.12 For AlN, the distribution of MC particles on (**a**) the first conduction band and (**b**) the second conduction band of AlN in reciprocal space at an electric field strength of 800 kV/cm. (After J. Fang et al. [133], with permission from the American Physical Society ©2019 American Physical Society)

mentioned at the beginning of this section hints at a more general issue that we shall discuss in Sect. 42.7, namely, the accuracy and reliability of *ab initio* methods in guiding us when dealing with electronic transport.

42.4.3 2D Materials

Since the first isolation of graphene [5], two-dimensional materials have attracted considerable interests not only for the unusual physical properties that some of them exhibit (topological properties, such as quantum spin Hall effect, or superconductivity, for example) but also since extremely thin channels are required in order to continue the never-ending scaling of semiconductor devices. Therefore, in this section we consider some of these 2D materials and see what information *ab initio* methods can give us about their charge-transport properties and assess which of them – if

any – could be used as the active channel material in ultra-scaled transistors or which of their unique properties could be exploited. We limit our attention mostly to homogeneous electronic transport (low-field mobility and high-field velocity and energy characteristics) in some of the popular 2D materials such as phosphorene, silicene, and germanene, calculated by full-band synchronous ensemble Monte Carlo method using the *ab initio* methods described above. We shall consider devices only cursorily, looking at a 10 nm gate-length phosphorene-based field-effect transistor (FET), to show the final connection that it is now possible to make between DFT and device characteristics.
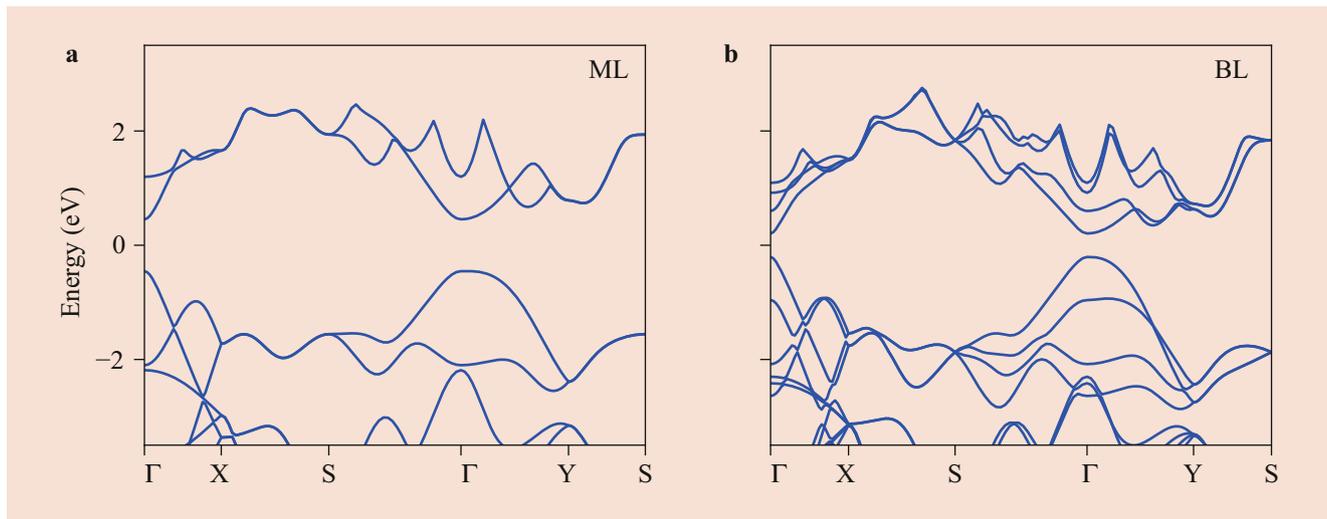
Fig. 42.13 The band structure of (**a**) monolayer phosphorene and (**b**) bilayer phosphorene calculated using VASP

Phosphorene

Phosphorene, mono- or few-layer black phosphorus, has gained wide interest due to high electron mobility observed in its bulk counterpart [134,135]. In this section, we present the results of low- and high-field electron/hole transport studies in monolayer and bilayer phosphorene. Reference [97] provides a more detailed description on transport calculations for phosphorene.

Single layers of black phosphorus, phosphorene, have a puckered honeycomb structure with a rectangular Brillouin zone. We show in Fig. 42.13 the electronic band structures for monolayer and bilayer phosphorene, calculated using the Vienna *ab initio* simulation package (VASP) [11–14], plotted along symmetry lines. VASP uses using projector-augmented wave (PAW) [136] pseudopotentials, and we have employed the Perdew-Burke-Ernzerhof generalized gradient approximations (PBE-GGA) [94] for the exchange-correlation functional. The relaxed lattice constants obtained are 4.62 Å and 3.30 Å for monolayers and 4.51 Å and 3.30 Å for bilayers. A direct band gap of 0.91 eV and 0.51 eV at the Γ symmetry point is obtained for mono- and bilayer. These values are lower than experimental values [137–139]. This is to be expected, since DFT is known to underestimate the band gap. However, this is not going to affect significantly our calculations since here we do consider inter-band transitions between the valence and conduction bands. The band structure of phosphorene exhibits also two additional valleys, the so-called Q and Y satellite valleys, along the $\Gamma - Y$ symmetry direction. At high fields and large doping densities, electrons tend to populate these valleys. The calculated effective masses for both electrons and holes in monolayers and bilayers at Γ are shown in Table 42.4. Both the valence and the conduction bands are anisotropic with a smaller effective mass along the armchair direction. The electron energy and

Table 42.4 Calculated effective masses, m_Γ^e and m_Γ^h, in monolayer and bilayer phosphorene

	m_Γ^e		m_Γ^h	
	Armchair	Zigzag	Armchair	Zigzag
ML	0.14	1.24	0.14	N/A
BL	0.10	1.33	0.09	3.08

After G. Gaddenane et al. [97], with permission from the American Physical Society ©2018 American Physical Society

velocity are calculated on a fine mesh around the Γ symmetry point (145×205 **k** points in the first quadrant) in order to account for the anisotropy and nonparabolicity of the electron dispersion.

We have calculated the phonon dispersion curves for monolayer and bilayer phosphorene calculated using the PHONOPY computer program [140]. In Fig. 42.14 we show the resulting dispersion along symmetry lines. One feature that is clearly evident in that figure – and that is common to all 2D materials – is the parabolic dispersion of the out-of-plane acoustic phonons, the flexural modes usually called ZA phonons. The effect of these phonons on carrier transport is detrimental in some of the 2D materials, such as silicene and germanene, that lack horizontal mirror (σ_h) symmetry [141]. However, in the case of phosphorene, the electron/hole-ZA phonon coupling is forbidden at the first order. Comparing the phonon dispersion of bilayers (Fig. 42.14b) with that of monolayers (Fig. 42.14a), we note the presence of low energy optical modes in bilayers. These are interlayer mode which are caused by weak interlayer coupling of entire unit cells in different layers that oscillate out-of-phase along either the in-plane (LO, TO) or out-of-plane (ZO) direction. Obviously, their low energy is due to the heavy mass of the entire unit cells that oscillate.
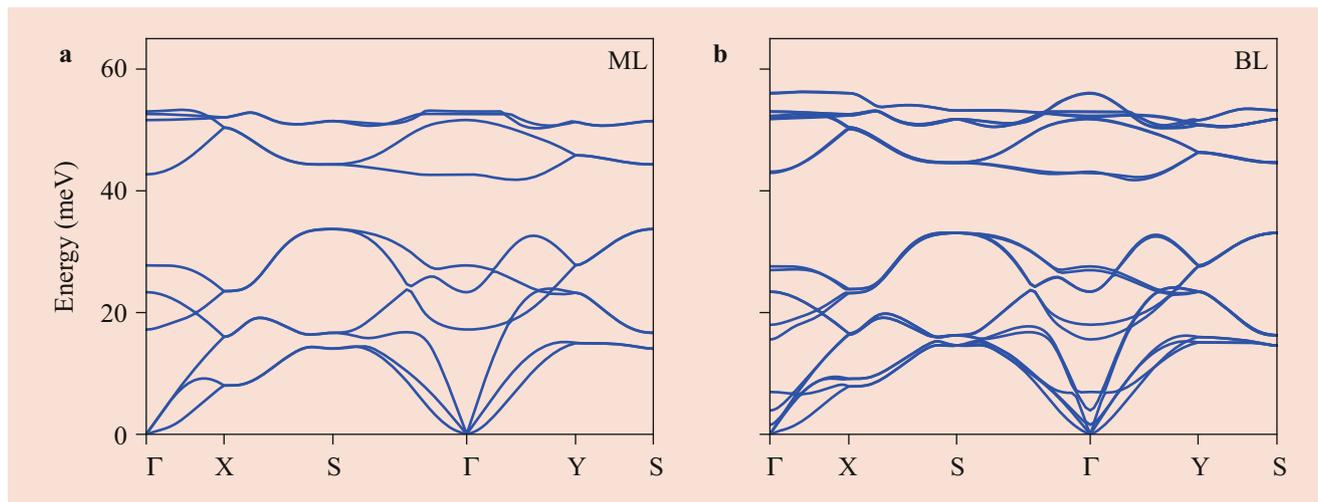
Fig. 42.14 The phonon dispersion for (**a**) monolayer phosphorene and (**b**) bilayer phosphorene calculated using PHONOPY

Figure 42.15 shows the angular average of the electron/hole-phonon scattering rate as a function of carrier kinetic energy for monolayer and bilayer phosphorene. The matrix elements are evaluated following the finite-displacement method described above. These are calculated on a very fine mesh to account for the angular dependence and wavefunction overlap effects. This is a crucial consideration, since failure to do so has shown to give inaccurate scattering rates and to overestimate the carrier mobility. Indeed, this is one of the main causes of the large discrepancy ($\mu_e = 20$–$26{,}000\,\mathrm{cm^2V^{-1}s^{-1}}$) among the mobility values present in the literature [97, 137, 142–145]. Such a strong angular dependence is shown in Fig. 42.16. As we have just mentioned, in 2D materials with horizontal mirror symmetry (σ_h-symmetric crystal), scattering of electrons/holes with the ZA phonons is forbidden at first order [141]. Therefore, this process can be ignored in the case of phosphorene, a σ_h-symmetric crystal. In monolayers, in-plane acoustic modes dominate the intravalley scattering processes with a strong backward scattering for electrons (Fig. 42.15a), and optical modes for holes (Fig. 42.15b). For electrons, intervalley scattering between Γ and Q is dominated by the 32 meV optical phonon which has a rather large deformation potential of 1.7×10^9 eV/cm. On the contrary, in bilayers in-plane acoustic modes and low-energy interlayer optical phonons dominate the intravalley scattering processes for both electrons and holes (Fig. 42.15c, d), and the 32 meV optical phonon dominates intervalley scattering for electrons. We should remark that, due to the low energy of these interlayer modes and to the numerous "band crossings" among themselves and with acoustic modes, it is difficult to separate clearly their contribution from the contribution of the acoustic modes.

Table 42.5 lists the calculated low-field carrier mobility for mono- and bilayer phosphorene in the armchair and zigzag direction. When the electron/hole-phonon matrix elements calculations are treated correctly by accounting the angular dependence, including wavefunction overlap effects, including scattering by all phonon modes and intra- and intervalley processes, the mobility obtained is on the lower end of the range of values present in the literature. In both mono- and bilayers, the carrier mobility is larger along the armchair direction compared to zigzag direction due to lower effective mass along the armchair direction. The carrier mobility does not change significantly going from monolayers to bilayers. In fact, the hole mobility decreases slightly.

Figures 42.17 and 42.18 show the velocity-field and energy-field characteristics for electrons and holes in monolayer and bilayer phosphorene. The carrier mobility for electrons and holes obtained from the velocity-field characteristics along the armchair and zigzag directions is in good agreement with the value obtained from the diffusion constant. The saturated velocity for electrons in both mono- and bilayers is relatively low (4×10^6 cm/s in the armchair direction and 1×10^6 cm/s in the zigzag direction) because of strong intervalley transfer of electrons from the Γ to the Q and Y valleys (Fig. 42.19). For holes, the velocity does not saturate in the zigzag direction even at high fields due to extremely low hole mobility along that direction (Figs. 42.17b and 42.17d). The average carrier energy for electrons remains close to the thermal energy (≈ 25 meV) up to a field of 10^5 V/cm for both monolayers and bilayers (42.18). However, the Ohmic regime extends to higher fields, especially in the case of holes accelerated along the zigzag direction, because of extremely low mobility along that direction (Fig. 42.18b, d).

Finally, we show in Fig. 42.20 the structure and current-voltage characteristics of a double-gate FET with a phosphorene channel, a 10 nm gate length, assuming Al_2O_3 as top and bottom oxide with an equivalent oxide thickness
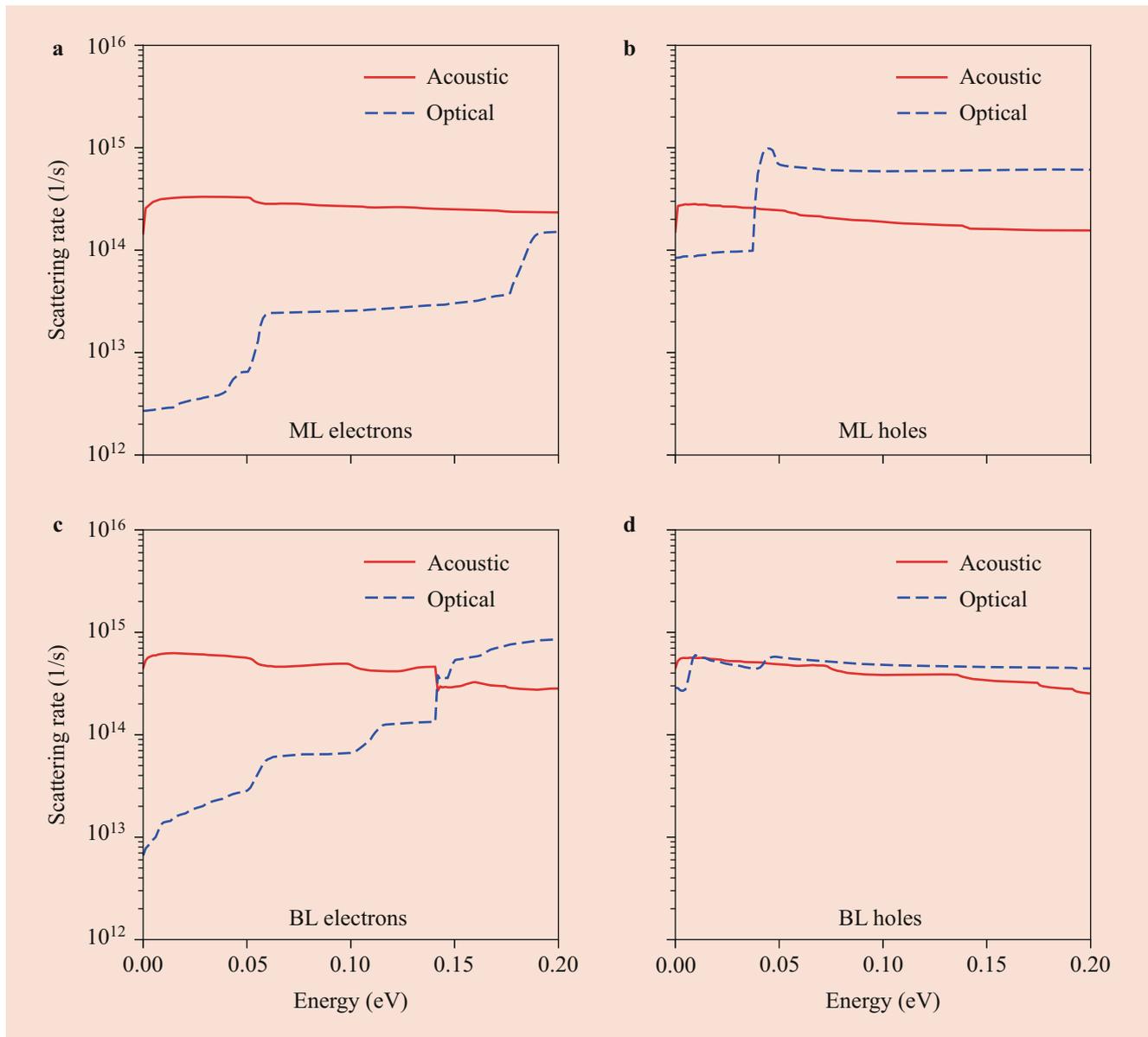
Fig. 42.15 Electron-phonon (left frames, (**a**) and (**c**)) and hole-phonon (right frames, (**b**) and (**d**)) scattering rates for monolayer phosphorene in (**a**) and (**b**), for bilayer phosphorene in (**c**) and (**d**), at 300 K, where the matrix element is obtained from VASP

(EOT) of 0.7 nm. Despite the poor mobility we have calculated for free-standing intrinsic monolayers, the performance of the device is satisfactory: an excellent I_{on}/I_{off} ratio of about 10^4, a transconductance g_m exceeding $1600\,\mu S/\mu m$ at $V_{DD} \approx 0.5$ V, a subthreshold slope of 60–70 mV/decade, and a drain-induced barrier lowering (DIBL) of 10 mV/V. Only the on current fails to meet the latest target of the International Technology Roadmap for Semiconductor (ITRS) [146]. Indeed, the mobility is only one of the many factors that control the performance of a device, especially since at the nanoscale transport approaches the ballistic limit. Moreover, the mobility may increase at high carrier density, (thanks to dielectric screening by the free carriers and Pauli blocking),

thus minimizing the negative effects of a high resistance in the source and drain regions.

Silicene and Germanene

Monolayer silicon (silicene) and germanium (germanene) have gained wide interest due to the tremendous role played by these semiconductors in their bulk form in the semiconductor industry in the last several decades. In this section we present the results of electron transport in these materials in monolayer form, following the procedure described above. Reference [147] provides for a more detailed description on how transport calculations have been performed to study these materials using first-principles methods.

Fig. 42.16 Acoustic electron-phonon matrix element for electrons plotted for an initial and final electron kinetic energies ≈40 meV as functions of the angle ϕ of the final wave vector $\mathbf{k}'$ with respect to the armchair direction. The initial wave vector $\mathbf{k}$ is taken to be along the armchair direction

Table 42.5 Calculated electron and hole mobility, μ_e and μ_h, in monolayer and bilayer phosphorene

	μ_e cm^{2}V^{-1}s^{-1}		μ_h cm^{2}V^{-1}s^{-1}	
	Armchair	Zigzag	Armchair	Zigzag
ML	20	10	19	2.4
BL	14	7	12	2

Adapted from G. Gaddenane et al. [97], with permission from the American Physical Society ©2018 American Physical Society

the calculations. In Fig. 42.22 the various phonon branches are labeled only for convenience; such an identification of the various branches (acoustic or optical, longitudinal or transverse) should be interpreted with a grain of salt, since these branches are often somewhat mixed, and such a distinction is physically unjustified for wave vectors at the edge of the Brillouin zone.

Figure 42.23 shows the electron-phonon scattering rates for silicene and germanene calculated at 300 K. The lack of horizontal mirror (σ_h) symmetry in these materials results in a very strong coupling, in principle diverging, with the long wavelength flexural (ZA) phonons [141]. This effect is further enhanced by the Dirac-like dispersion which causes strong back scattering, effect that is due to the degeneracy of the valence and conduction band at the K symmetry point. In the absence of a process which suppresses the ZA phonons, the resulting mobility in these materials is extremely low (10^{-3} to 10^{-2} cm^2/V·s). In the calculations presented here, we have assumed a long wavelength cutoff of 1 nm for the ZA phonons to avoid the divergence. The cutoff is chosen such that intravalley scattering is suppressed and intervalley scattering is included. In both silicene and germanene, scattering with ZA and LA phonons dominates the scattering, in particular the K-K′ intervalley process.

Without delving too deeply into the extremely interesting problem of assessing which mechanism may suppress, or at least damp, the first-order interaction (one-phonon processes) with the ZA modes (issue discussed in detail in Ref. [141]), we should mention that, besides obvious effects like the finite grain-size and the rippling or crumpling of the layer, their anharmonic coupling to in-plane acoustic modes [151–154] has been shown to renormalize their dispersion from a square (quadratic) law to a power-law of the form q^r, with r being approximately 3/2. Moreover, in supported and/or gated layers, Amorim and Guinea [155] have shown that at long wavelengths the renormalized dispersion of the ZA phonons, $\tilde{\omega}^{(ZA)}(q)$, approaches a nonzero constant, taking the form $\tilde{\omega}^{(ZA)}(q) \sim \sqrt{\omega^{(ZA)}(q)^2 + \omega_0^2}$, where $\omega^{(ZA)}(q)$ is the dispersion in free-standing layers and $\omega_0 \approx \sqrt{g/\rho_{2D}}$, ρ_{2D} being the mass density of the 2D layer and g is a measure of the strength of the coupling between the layer and the substrate (or gate insulator). Typically, $\hbar\omega_0$ is of the order of a few meV, small but nonzero, which is what matters. These considerations, all derived in the continuum limit, solve the

Figures 42.21 and 42.22 show the electronic band structure and phonon dispersion for ideal free-standing silicene and germanene, calculated by DFT using the VASP and PHONOPY packages. Similar to graphene, both these materials exhibit a Dirac-like electronic dispersion. However, the Fermi velocities we have obtained for silicene ($\approx 5.3 \times 10^7$ cm/s) and germanene ($\approx 5.3 \times 10^7$ cm/s) are lower than in graphene ($\approx 9.5 \times 10^7$ cm/s), and, unlike graphene, their structures are buckled. The lattice and buckling constants we have obtained are 3.86 Å and 0.48 Å for silicene and 4.05 Å and 0.71 Å for germanene, respectively. Furthermore, our calculations were performed by neglecting the contributions of spin-orbit (SO) coupling for computational efficiency. This should not be a source of any major concern, since it has been shown that the SO interaction opens a rather small band gap in these materials (about 1.5 meV for silicene [148–150] and 25 meV for germanene [148, 149]). Given these small values, the effect of the SO interaction on the transport properties should be negligible for silicene and would result only in a slight overestimation of the mobility and drift velocity in germanene. We have chosen a supercell of size $5 \times 5 \times 1$ unit-cells when calculating the phonon spectra, in order to avoid nonphysical "negative frequencies" (in reality, imaginary frequencies) of the low-energy acoustic phonons. Often these negative (square)frequencies are interpreted as a sign of the thermodynamic instability of the crystal, but in many cases they are simply the result of numerical artifacts when cells of excessively small dimensions are used to speed up

42

Fig. 42.17 Drift velocity *vs* field at 300 K for electrons (left frames, (**a**) and (**c**)) and holes (right frames, (**b**) and (**d**)), in monolayer phosphorene (**a**) and (**b**), and bilayer phosphorene (**c**) and (**d**)

Mermin-Wagner puzzle [156] but do not necessarily result in a negligible role of the ZA phonons in depressing the carrier mobility. In particular, even accounting for the emergence of a gap in supported and/or gated layers, the coupling with the Rayleigh waves of the substrate and/or gate [155, 157] may still play a huge role. Recently, Rudenko et al. [158] have concluded that even in nonsymmetric materials with a parabolic dispersion, the effect is small, amounting to no more than a 35% reduction of the mobility due to second-order interactions (two-phonon processes), arguing that even

in this case the first-order coupling of the carriers to the acoustic flexural modes is forbidden by the rotational symmetry of the crystal. However, these conclusions are reached by approximating the 2D layer as a membrane of zero thickness, thus ignoring the "internal structure" of the membrane, an effect that may result in the breaking of the rotational and inversion symmetry of a non-σ_h-symmetric membrane. The situation is even less clear in the case of non-σ_h-symmetric materials with a Dirac-like electronic dispersion, like silicene and germanene.
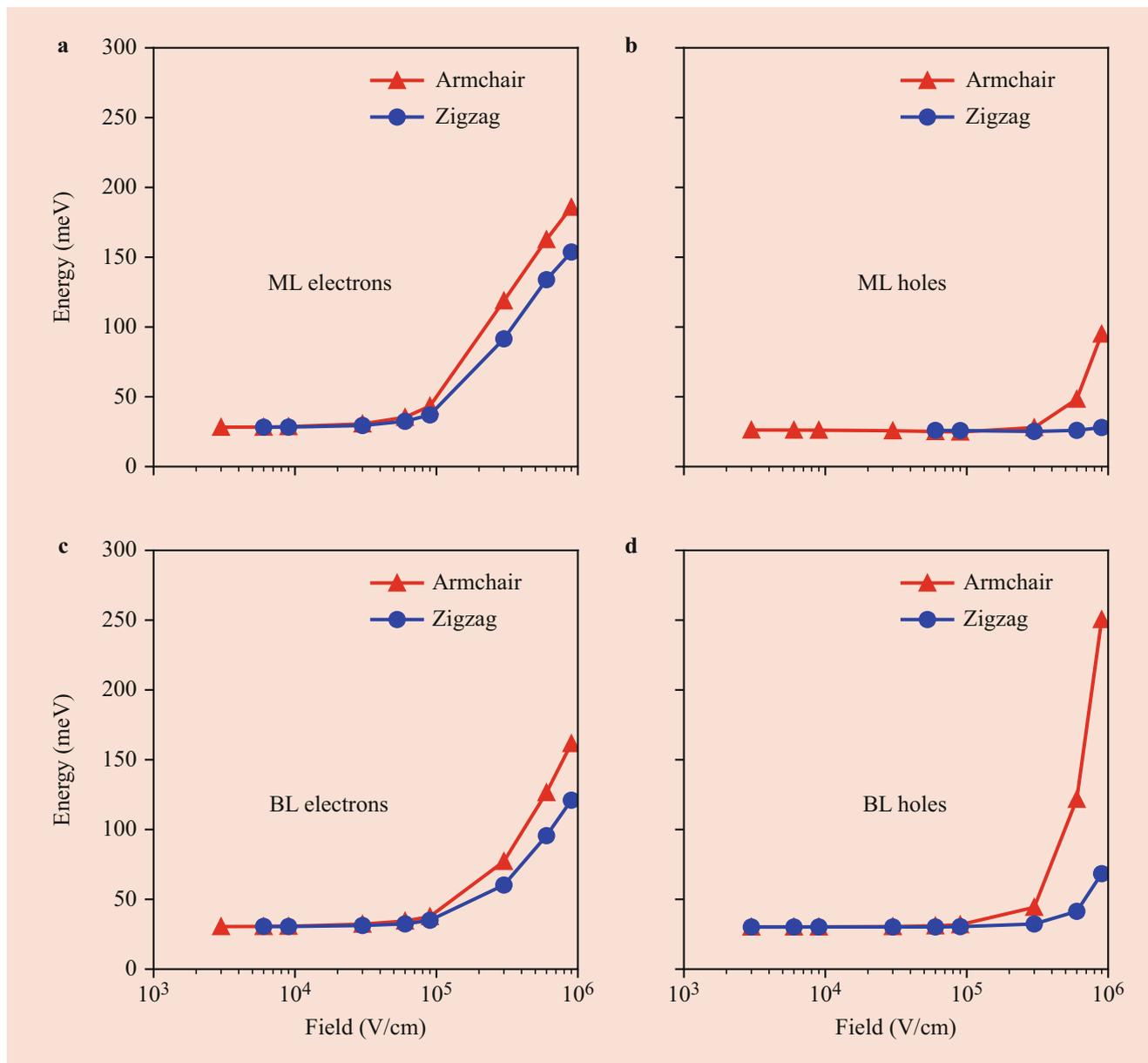
Fig. 42.18 Energy *vs*. field at 300 K for electrons (left frames, (**a**) and (**c**)) and holes (right frames, (**c**) and (**d**)), in monolayer phosphorene (**a**) and (**b**), and bilayer phosphorene (**c**) and (**d**)

Table 42.6 lists the electron mobility calculated along the [01] direction (Γ-K) for various scattering processes from the diffusion constant. With a 1 nm ZA phonon cutoff, the calculated electron mobility is 701 cm²/V·s in silicene and 2327 cm²/V·s in germanene. When we neglect scattering by ZA phonons completely, the mobility rises to 1296 cm²/V·s and 4549 cm²/V·s, respectively. When intervalley scattering processes assisted by both LA and ZA phonons are ignored, the mobility increases by an order of magnitude.

Figure 42.24 shows the velocity-field and energy-field characteristics for silicene and germanene for an electric field applied along the [01] direction. The electron mobility ex-tracted from the velocity-field characteristics is in agreement with the value obtained from the evaluation of the diffusion constant. At low fields, electrons remain thermal and exhibit Ohmic behavior in both materials. However, at higher fields, germanene exhibits negative differential mobility due to Γ-K intervalley scattering. Lower scattering rates and a smaller energy difference between the Dirac point K and the bottom of the Γ valley in germanene result in a significant transfer of electrons to the Γ valley, a lower velocity region. However, in silicene, we observe no significant population of the Γ valley due to large energy difference between the two valley minima, and we obtain a saturation velocity of $\approx 8 \times 10^6$ cm/s.
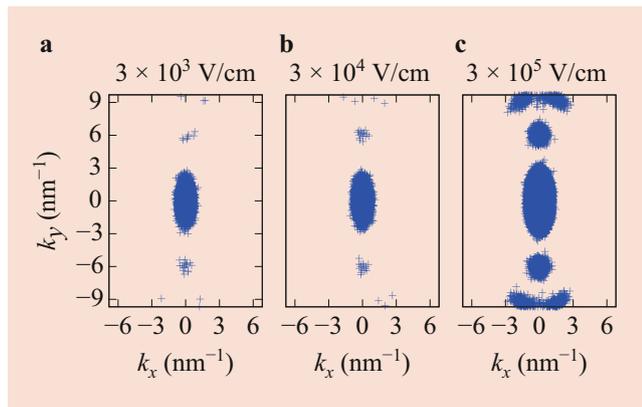
Fig. 42.19 Distribution in reciprocal space for electrons in monolayer phosphorene for a field of 3×10^3 (**a**), 3×10^4 (**b**), and 3×10^5 V/cm (**c**), along the armchair direction. Note the $\Gamma - Q$ intervalley transfer at center, the $\Gamma - Y$, and $Q - Y$ intervalley transfer at the highest field

This difference can be seen in Fig. 42.25, which compares the momentum distribution of electrons in silicene and germanene at high fields.

In summary, at least judging from the DFT results we have just reviewed, and at least in principle, silicene and germanene may exhibit satisfactory properties. However, understanding the role played by the acoustic flexural modes in determining their charge-transport behavior remains an open issue. Moreover, and probably more important, processing these materials (growth/deposition of the layers, their passivation, doping, and gating, formation of Ohmic contact, etc.) could constitute a formidable challenge.

42.5 Dielectric Response of Low-dimensional Materials

In Sect. 42.1, when outlining the content of this chapter, we mentioned the need to use a position-dependent dielectric constant (or function) when simulating nanoscale devices using atomistic models. Indeed, looking at the top frame of Fig. 42.20, a solution of Poisson equation requires the knowledge of the location of the interface between the 2D layers (for phosphorene a thin layer but still of a finite thickness) and the top and bottom insulators. When using conventional continuum models, approximating the location of this interface, the straight line drawn in the diagram will suffice. Not so when using atomistic models. In this case, the smooth transition between the dielectric and the 2D layer has to be resolved accurately. Moreover, a more intriguing question is how to define the dielectric constants (in plane and out of plane) of the 2D layer itself.

The case illustrated in Fig. 42.20 is only one example. Indeed, devices based on the many low-dimensionality nanomaterials of present interest, such as graphene nanoribbons
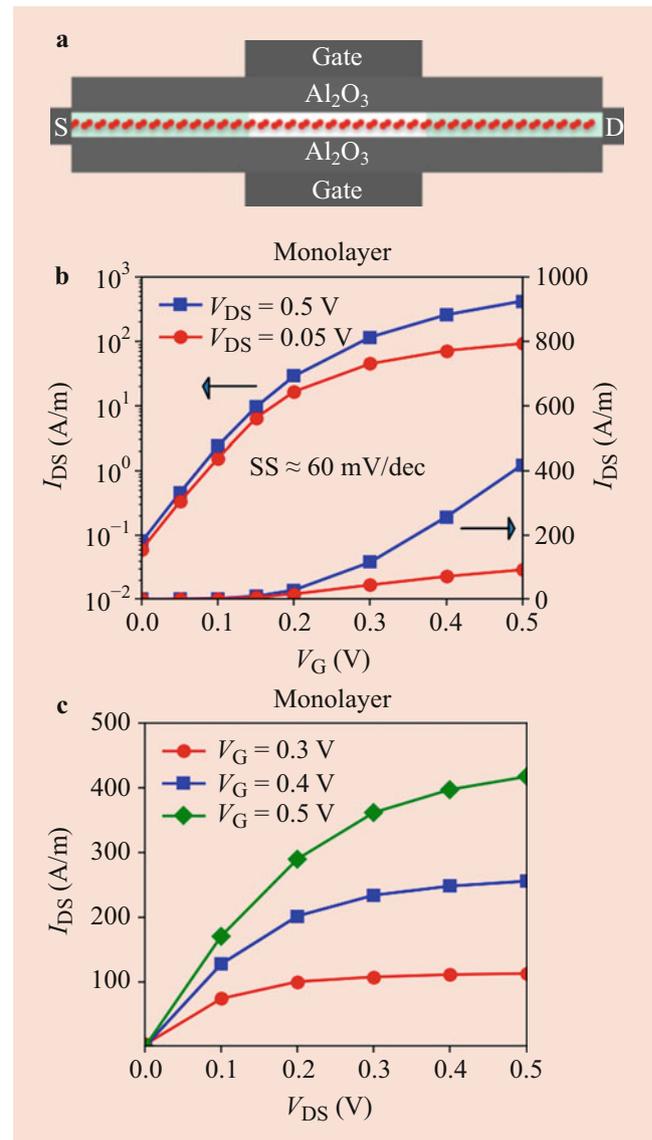
Fig. 42.20 (**a**) Schematic of the 10 nm gate-length phosphorene-based FET we have simulated. (**b**) Transfer characteristics of the device. (**c**) output characteristics of the device

[159], carbon nanotubes [160], silicon nanowires [161], graphene [5], silicene [162], phosphorene [163], and transition metal dichalcogenide (TMD) monolayers [164], all require this knowledge. Indeed, their dielectric properties determine macroscopic quantities, such as the gate capacitance and transconductance, and transport properties, as affected by dielectric screening and by the coupling of collective modes (plasmons) in the layers and insulators. However, the dielectric properties of these nanomaterials are not well studied, either experimentally or theoretically, compared with the bulk form of these materials. Different from the bulk materials which can usually be characterized by a single isotropic macroscopic dielectric constant if it has a
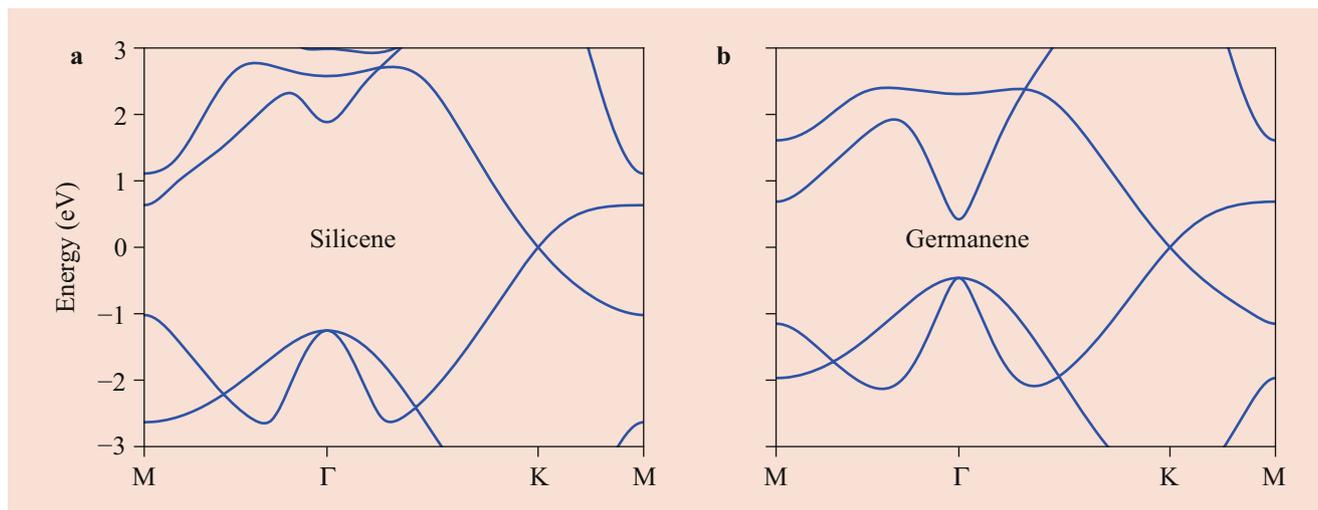
Fig. 42.21 The band structure of (**a**) silicene and (**b**) germanene calculated using VASP
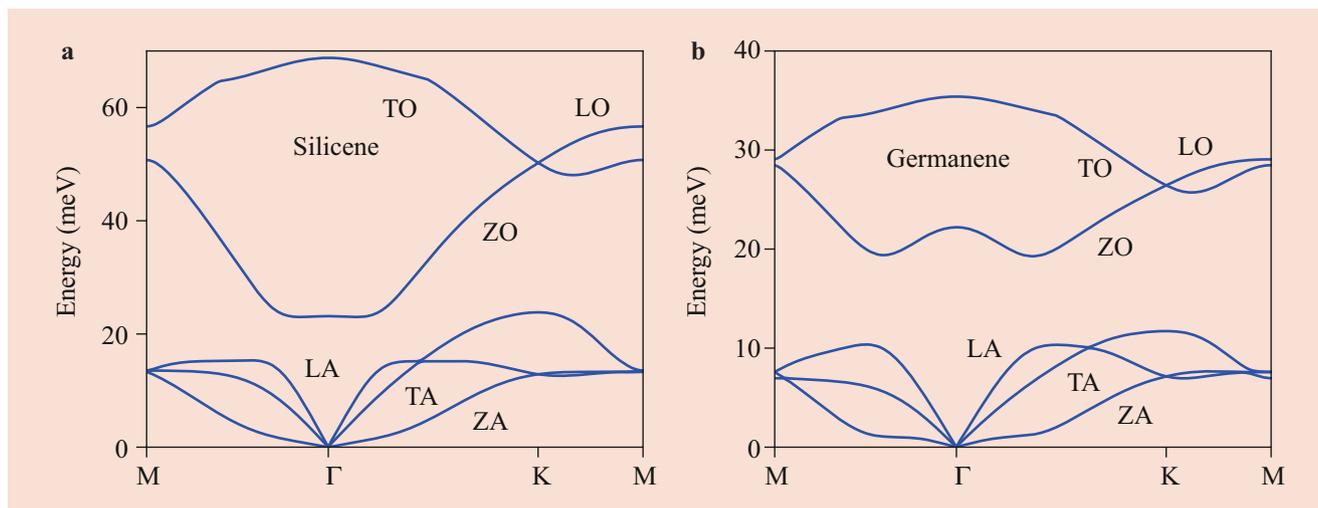
Fig. 42.22 The phonon dispersion of (**a**) silicene and (**b**) germanene calculated using PHONOPY

cubic crystal structure, low-dimensional nanomaterials are anisotropic by definition, and the study of their dielectric properties is more complex. Anisotropic tensorial dielectric properties are expected.

Here we shall outline how *ab initio* methods can provide the missing information. To keep the treatment simple and general, we shall consider only empirical pseudopotentials. Whereas this may be viewed as not a "first-principles" approach, the formalism can be extended to self-consistent pseudopotentials (DFT) in a straightforward way.

Several theoretical methods have been developed to study either the macroscopic or the microscopic dielectric response of nanosystems [165–176]. However, not much attention has been paid to the microscopic dielectric tensor of low-dimensionality nanomaterials. Therefore, here we present

the derivation of a microscopic Poisson equation based on the density-density response function [177, 178] and its application to the study of these systems. We consider the particular case of the static dielectric properties of armchair-edge graphene nanoribbons (aGNRs) and graphene, but the method can be extended to other nanomaterials, such as Si nanowires (NWs, as illustrated in Fig. 42.30) and TMD monolayers (as recently reported by Laturia et al. using DFT [179]).

42.5.1 Density-Density Response Function

In the static linear response theory [177, 178, 180], the density-density response function, $\chi(\mathbf{r}, \mathbf{r}')$, describes the
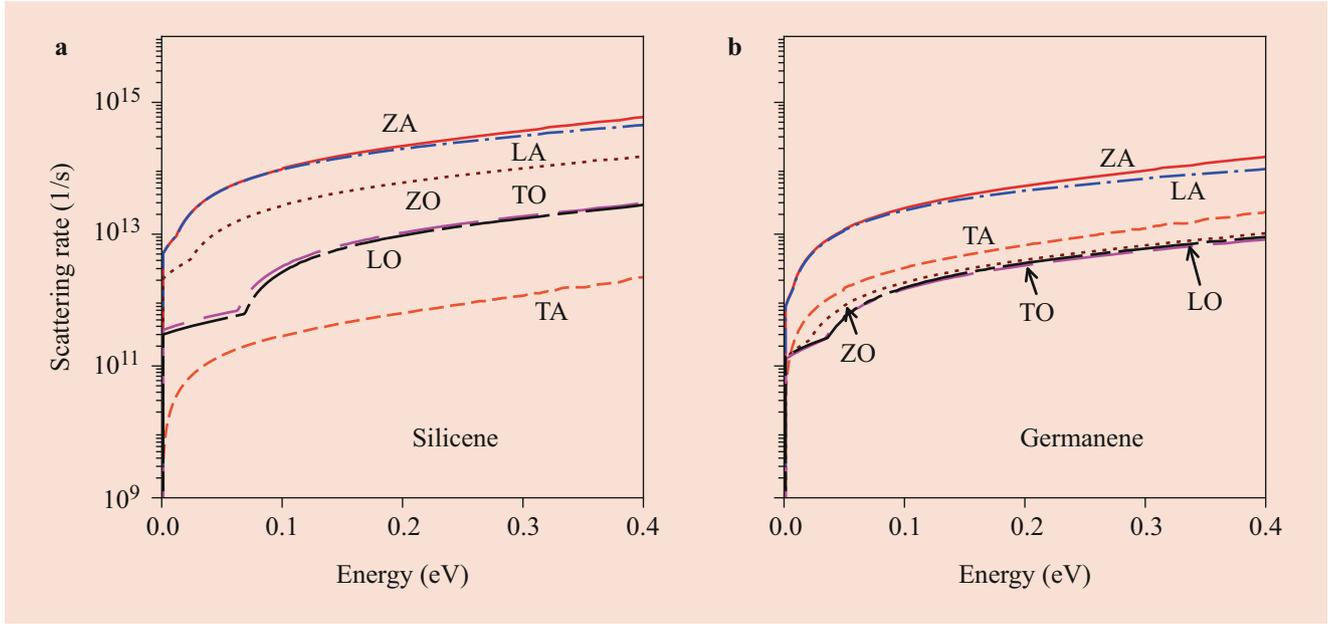
Fig. 42.23 Electron-phonon scattering rates for (**a**) silicene and (**b**) germanene per mode at 300 K, where the matrix elements are obtained from VASP

Table 42.6 Electron mobility in silicene and germanene in the [01] direction

	Silicene (germanene) ($cm^2V^{-1}s^{-1}$)
1 nm ZA phonon cutoff	$7.01 \times 10^2 (2.32 \times 10^3)$
Ignoring ZA phonons (intra- and intervalley)	$1.29 \times 10^3 (4.55 \times 10^3)$
Only intravalley with in-plane acoustic phonons	$4.3 \times 10^4 (2.6 \times 10^4)$
Only intervalley with acoustic phonons	$8.17 \times 10^2 (3.5 \times 10^3)$
Only optical phonons (intra- and intervalley)	$8.2 \times 10^3 (4.2 \times 10^4)$

After G. Gaddenane et al. [147], with permission from the American Institute of Physics ©2018 American Institute of Physics)

change of the charge density $e\delta\rho(\mathbf{r})$ when the external potential undergoes a small change, $\delta v_{ext}(\mathbf{r}')$, so that the electric polarizability (or "susceptibility") is $\chi(\mathbf{r},\mathbf{r}') = e\delta\rho(\mathbf{r})/\delta v_{ext}(\mathbf{r}')$. As an example, we first illustrate the independent-particle density-density response function $\chi^0(\mathbf{r},\mathbf{r}')$ of an armchair-edge graphene nanoribbon with a width given by seven atomic lines, here called a "7-aGNR." As mentioned above, in order to reduce the computational effort, we employ the local empirical pseudopotentials proposed by Kurokawa [181]. These pseudopotentials surprisingly yield the correct band structure of both sp^3-coordinated C in the face-center cubic phase(diamond) as well as sp^2-coordinated C, such as graphene [60]. The band

structure of the ribbon has been calculated using a supercell size of 1.72 × 0.85 nm in the nonperiodic directions and an energy cutoff of 10 Rydberg, resulting in 2237 plane waves in the basis set. The static independent-particle polarizability matrix is then calculated as [180, 182, 183]:

$$\chi^0_{\mathbf{GG}'}(\mathbf{q}) = \frac{4}{\Omega}$$
$$\sum_{c,v,\mathbf{k}} \frac{\langle v, \mathbf{k}|e^{-i(\mathbf{q+G})\cdot\mathbf{r}}|c, \mathbf{k+q}\rangle \langle c, \mathbf{k+q}|e^{i(\mathbf{q+G}')\cdot\mathbf{r}'}|v, \mathbf{k}\rangle}{E^v_{\mathbf{k}} - E^c_{\mathbf{k+q}}},$$

(42.35)

using the eigenvalues $E^n_{\mathbf{k}}$ ($n = c, v$ for the conduction and valence bands, respectively) and eigenfunctions $|n, \mathbf{k}\rangle$ obtained from the calculation of the band structure. In Eq. (42.35), $\mathbf{G}$ and $\mathbf{G}'$ are reciprocal lattice vectors, $\mathbf{q}$ and $\mathbf{k}$ are the 2D vectors within the first Brillouin zone, and Ω denotes the crystal volume. In Eq. (42.35), the sum over the indices c and v denote the sum over all conduction and valence bands, respectively. The real-space susceptibility $\chi^0(\mathbf{r},\mathbf{r}')$ can be obtained by transforming $\chi^0_{\mathbf{GG}'}(q_z \to 0)$:

$$\chi^0(\mathbf{r},\mathbf{r}') = \frac{1}{\Omega_{cell}} \sum_{\mathbf{GG}'} e^{i\mathbf{G}\cdot\mathbf{r}} \chi^0_{\mathbf{GG}'}(q_z \to 0) e^{-i\mathbf{G}'\cdot\mathbf{r}'}, \quad (42.36)$$

where Ω_{cell} is the volume of the supercell.

The screened response function, $P(\mathbf{r}, \mathbf{r}')$, relates the induced charge, $e\rho_{ind}(\mathbf{r})$, to the total potential, $\varphi_{tot}(\mathbf{r})$, through:
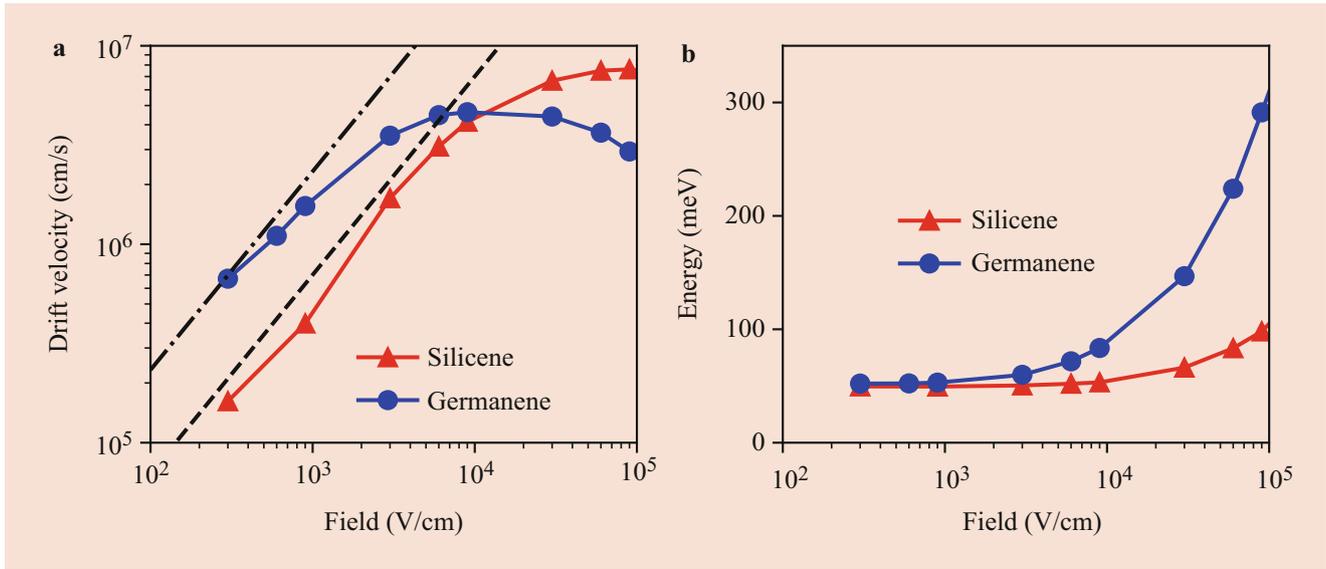
Fig. 42.24 (**a**) Drift velocity *vs* field and (**b**) average-energy *vs* field at 300 K for electrons in silicene and germanene
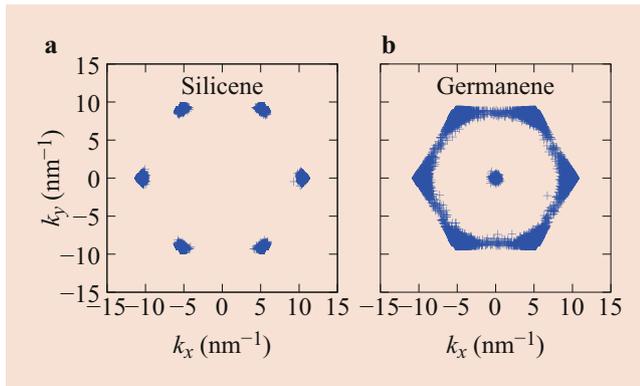
Fig. 42.25 Distribution of electrons in reciprocal space in (**a**) silicene and (**b**) germanene at a high field ($\approx 10^5$ V/cm) along the [01] direction

$$e\rho_{\text{ind}}(\mathbf{r}) = e^2 \int P(\mathbf{r},\mathbf{r}')\varphi_{\text{tot}}(\mathbf{r}')\mathrm{d}\mathbf{r}', \qquad (42.37)$$

Within the random-phase approximation (RPA), $\chi^0(\mathbf{r},\mathbf{r}')$, obtainable as the Fourier-transform of the dielectric matrix, is $P(\mathbf{r}, \mathbf{r}')$. The function $\chi^0(\mathbf{r},\mathbf{r}')$, averaged along z, is shown in Fig. 42.26 for a perturbation at $\mathbf{r}'$, for $\mathbf{r}'$ is at the ribbon center (a) and at the bottom-left vacuum region (b), respectively.

Two important features of the density-density response function are its exponential spatial decay, like the electron density, and its scalar nature. The first feature enables us to employ a small supercell – as long as it is sufficiently large to guarantee the independence of the calculated band structure on its size – to study the dielectric response. The second feature allows us to avoid unwanted complexity in constructing the linear system of equations when solving the 'ordinary'

Poisson equation with tensorial dielectric constants, as shown in the following.

42.5.2 Microscopic Poisson Equation

Of particular interest from the perspective of atomistic device simulations is the possibility of deriving a microscopic Poisson equation that accounts "exactly" for the position-dependent polarization of the system without having to employ any position-dependent "dielectric constant." This can be done by combining Gauss's law, $\nabla \cdot \mathcal{E}_{\text{tot}}(\mathbf{r}) = e\rho_{\text{tot}}(\mathbf{r})/\epsilon_0 = (e\rho_{\text{ext}}(\mathbf{r}) + e\rho_{\text{ind}}(\mathbf{r}))/\epsilon_0$, with the relation between the electric field and electrostatic potential, $\mathbf{E}_{\text{tot}}(\mathbf{r}) = -\nabla\varphi_{\text{tot}}$, in the Coulomb gauge and using Eq. (42.37), to express the induced charge as a function of the total potential:

$$\epsilon_0 \nabla^2 \varphi_{\text{tot}}(\mathbf{r}) + e^2 \int P(\mathbf{r},\mathbf{r}')\varphi_{\text{tot}}(\mathbf{r}')\mathrm{d}\mathbf{r}' = -e\rho_{\text{ext}}(\mathbf{r}), \quad (42.38)$$

where ϵ_0 is the vacuum permittivity, $\varphi_{\text{tot}}(\mathbf{r})$ is the total potential, and $e\rho_{\text{ext}}(\mathbf{r})$ is the external charge. Equation (42.38) can be straightforwardly discretized and solved when Dirichlet/Neumann/periodic boundary conditions are applied. Assuming linear response and embracing the RPA, the solution of this microscopic Poisson equation with the polarizability $\chi^0(\mathbf{r},\mathbf{r}')$ – calculated using the procedure outlined above and either DFT of empirical pseudopotentials – provides the exact total microscopic potential for any arbitrary distribution of free charge.

In so doing, the total potential is obtained immediately, but separating the distinct contributions of the external potential
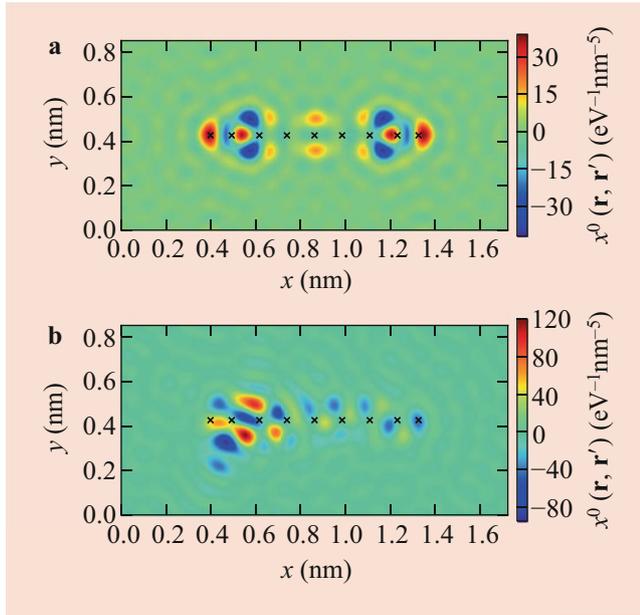
Fig. 42.26 The independent-particle density-density response function, $\chi^0(\mathbf{r}, \mathbf{r}')$, of a 7-aGNR in response to a long-wavelength perturbing electric field $q_z \rightarrow 0$, (**a**) when the perturbation $\mathbf{r}'$ is at the ribbon center (*i.e.*, $\mathbf{r}' = (0.86 \, \text{nm}, 0.43 \, \text{nm})$), and (**b**) when the perturbation $\mathbf{r}'$ is at the bottom-left vacuum (*i.e.*, $\mathbf{r}' = (0.43 \, \text{nm}, 0.21 \, \text{nm})$). The microscopic distribution of the polarization charge can be observed. The seven carbon atoms in the center of the ribbon and the two terminating hydrogen atoms on the edges are indicated with the black cross symbols. (After J. Fang et al. [245], with permission from the American Physical Society ©2016 American Physical Society)

and of the induced (polarization) potential requires a brief discussion. Since we have assumed that the aGNR (or any other nanostructure) is embedded in vacuum, the polarization charge lies exclusively within the aGNR. This can be calculated using Eq. (42.37). Once this is known, the induced polarization potential can be obtained from:

$$\varphi_{\text{ind}}(\mathbf{r}) = e \int \frac{\rho_{\text{ind}}(\mathbf{r}')}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \quad (42.39)$$

In turn, having obtained the polarization potential, the applied external potential (*i.e.*, the bare unscreened perturbation) can be obtained from the difference between the total potential and the polarization potential. Finally, the external, polarization, and total electric fields can be calculated from the negative of the gradients of the corresponding potentials.

In two dimensions, averaging the charge distribution along the out-of-plane direction z, one should use the two-dimensional Green's function of the Poisson equation. Thus, the polarization potential is given by:

$$\varphi_{\text{ind}}(\mathbf{r}) = e \int \frac{\rho_{\text{ind}}(\mathbf{r}') \ln(|\mathbf{r} - \mathbf{r}'|/L)}{2\pi\epsilon_0} d\mathbf{r}', \quad (42.40)$$
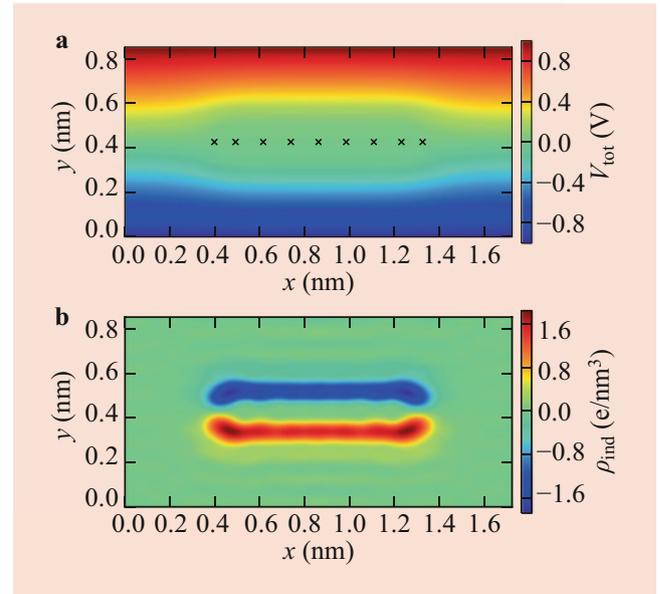
Fig. 42.27 The total potential (**a**) and the induced charge density (**b**) for a 'horizontal' parallel-plate capacitor when 1 V is applied to the top plate and -1 V to the bottom plate. In response to the external field, a dipole is formed. This counteracts the external electric field with a polarization field. (After J. Fang et al. [245], with permission from the American Physical Society ©2016 American Physical Society)

where L some arbitrary scaling length. In practice, to the integration over the singularity of the Green's function $\ln(|\mathbf{r} - \mathbf{r}'|/L)$, we can instead calculate the polarization potential by simulating the system using a larger region and Fourier-transforming the polarization charge to obtain $e\rho_{\text{ind}}(\mathbf{r}) \rightarrow e\rho_{\text{ind}}(\mathbf{q})$. Then, the polarization potential is finally obtained from $\varphi_{\text{ind}}(\mathbf{q}) = e\rho_{\text{ind}}(\mathbf{q})/(\epsilon_0|\mathbf{q}|^2) \rightarrow \varphi_{\text{ind}}(\mathbf{r})$.

42.5.3 Microscopic Dielectric Tensor of 1D Materials

To compute the microscopic dielectric tensor along the non-periodic directions of 1D materials, Eq. (42.38) can be used in the case of two nanocapacitors, a "horizontal" one with plates parallel to the aGNR surface and a "vertical" one with plates perpendicular to the aGNR surface. Since at $q_z \rightarrow 0$, the net polarization cannot be built to screen the external potential or electric field applied along the axial direction of the 1D materials, the axial dielectric constant is unity. For finite values of q_z, Eq. (42.38) can be applied to study the axial component of the dielectric tensor by assuming a periodic external (perturbing) potential or electric field and considering a nanocapacitor with length equal to the (spatial) period of the external perturbation. Here we restrict our attention to the simpler case of $q_z \rightarrow 0$.
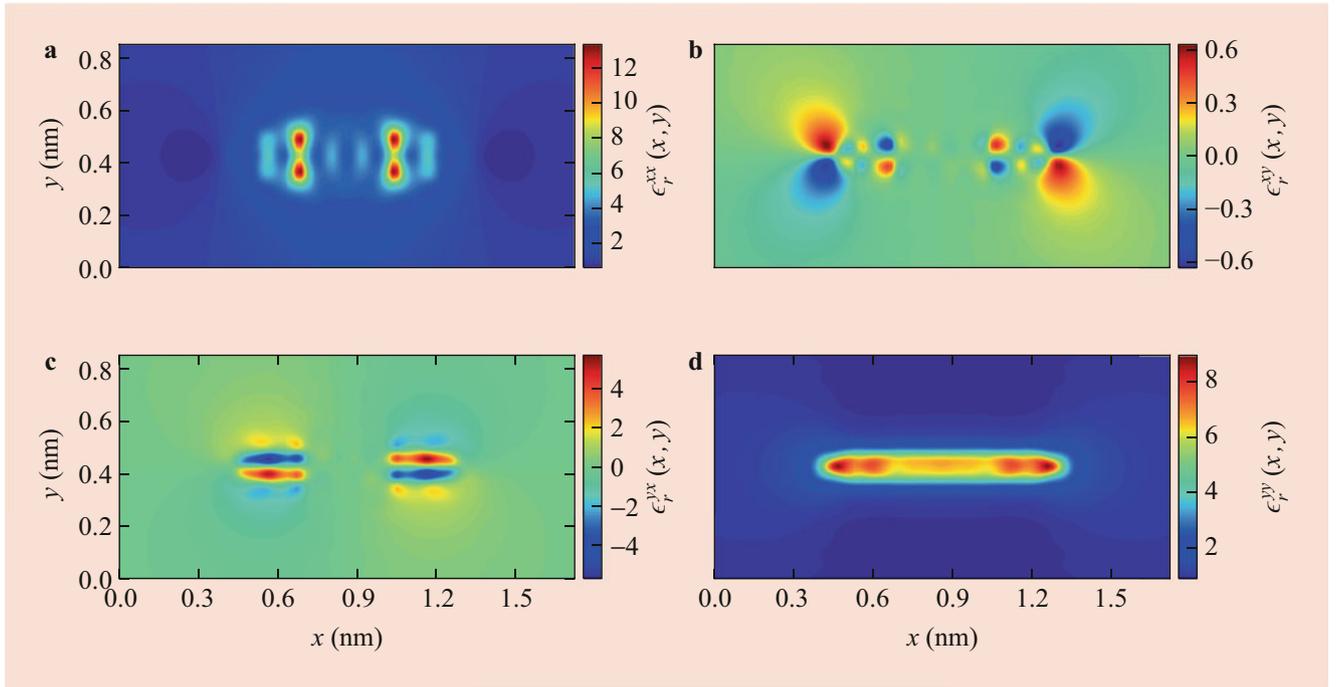
Fig. 42.28 The microscopic dielectric tensor of a 7-aGNR in vacuum. (**a**) and (**d**) show the diagonal components, and (**b**) and (**c**) show the off diagonal components. (After J. Fang et al. [245], with permission from the American Physical Society ©2016 American Physical Society)

42

The perturbation is defined by applying 1 V on the top plate and −1 V on the bottom plate of the horizontal nanocapacitor. Note that periodic boundary conditions are applied on the left and right boundaries and there are no free charges inside the capacitor. The spatial variation of the total potential and the induced charge, obtained by solving Eq. (42.38) in the two-dimensional case, are shown in Fig. 42.27. In the top frame, note the curvature of the potential distribution around the nanoribbon; in the bottom frame, one can clearly see the induced dipole. In the top frame, we can observe the curved potential distribution around the nanoribbon region; in the bottom frame, we can see the induced dipole. Although not shown, it is easy to obtain the potential and induced charge for the vertical nanocapacitor.

The microscopic external electric field and total electric field for the horizontal ($E_{ext/tot}^{hor}$) and the vertical ($E_{ext/tot}^{ver}$) capacitor can be easily derived having obtained $\varphi_{tot}(\mathbf{r})$ and $e\rho_{ind}(\mathbf{r})$. The microscopic dielectric tensor of a 1D nanostructure in the confined (x, y) plane (*i.e.*, the transverse dielectric tensor) is then evaluated as:

$$\begin{bmatrix} \epsilon_r^{xx}(\mathbf{r}) & \epsilon_r^{xy}(\mathbf{r}) \\ \epsilon_r^{yx}(\mathbf{r}) & \epsilon_r^{yy}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} \mathcal{E}_{ext,x}^{hor}(\mathbf{r}) & \mathcal{E}_{ext,x}^{ver}(\mathbf{r}) \\ \mathcal{E}_{ext,y}^{hor}(\mathbf{r}) & \mathcal{E}_{ext,y}^{ver}(\mathbf{r}) \end{bmatrix} \begin{bmatrix} \mathcal{E}_{tot,x}^{hor}(\mathbf{r}) & \mathcal{E}_{tot,x}^{ver}(\mathbf{r}) \\ \mathcal{E}_{tot,y}^{hor}(\mathbf{r}) & \mathcal{E}_{tot,y}^{ver}(\mathbf{r}) \end{bmatrix}^{-1},$$

(42.41)

where $\mathcal{E}_{ext/tot,x/y}^{hor/ver}(\mathbf{r})$ is the x or y component of the external/total electric field for the horizontal/vertical nanocapacitor and $\epsilon_r^{ij}(\mathbf{r})$ ($i, j = x, y$) is a component of the dielectric tensor

which indicates the dielectric response along the direction i when the perturbing electric field is applied along the direction j.

Microscopic Dielectric Tensor of a 7-aGNR

Using the empirical pseudopotentials of Ref. [181] for carbon and the terminating H atoms, we have used the method just outlined to calculate the microscopic dielectric tensor of a 7-aGNR. This is shown in Fig. 42.28. The anisotropy of the microscopic dielectric properties can be observed. Observing the distribution of both $\epsilon_r^{xx}(\mathbf{r})$ and $\epsilon_r^{yy}(\mathbf{r})$, the ribbon region exhibits a large microscopic dielectric constant whereas it approaches ϵ_0 in the vacuum region. Nevertheless, inside the ribbon, $\epsilon_r^{xx}(\mathbf{r})$ exhibits more pronounced microscopic oscillations compared to $\epsilon_r^{yy}(\mathbf{r})$. The off-diagonal elements of the microscopic dielectric tensor, $\epsilon_r^{xy}(\mathbf{r})$ and $\epsilon_r^{yx}(\mathbf{r})$, exhibit the inversion symmetry of the aGNR. Macroscopically, the off-diagonal elements of the microscopic dielectric tensor vanish. This is because x and y are two of the three principal axes of the nanosystem.

Ribbon-Width Dependence of the Dielectric Permittivity

We have simulated also a 3-aGNR, a 5-aGNR and a 6-aGNR. Using empirical pseudopotentials, we find for each of these ribbons a band gap of 1.19 eV, 0.43 eV, and 1.03 eV, respectively, in contrast to the band gap of 1.70 eV for the 7-aGNR. The variation of $\epsilon_r^{yy}(\mathbf{r})$ along the horizontal
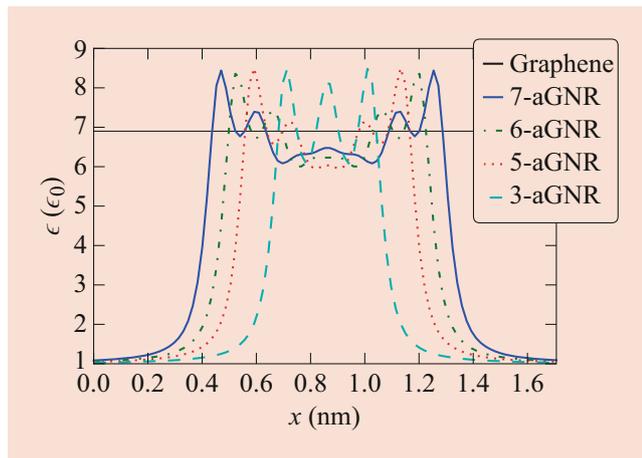
Fig. 42.29 Out-of-plane microscopic dielectric constant, $\epsilon_r^{yy}(\mathbf{r})$, in the horizontal plane for different ribbon widths. The local permittivity in the interior of all ribbons except for the 3-aGNR shows values close to $6\,\epsilon_0$. The local permittivity at the edges of the ribbon shows similar oscillating peaks. The black horizontal straight line, $\epsilon = 6.9\epsilon_0$, indicates the macroscopic dielectric constant of graphene at the carbon-atom plane, as shown in Fig. 42.31. (After J. Fang et al. [245], with permission from the American Physical Society ©2016 American Physical Society)

plane of a 3-aGNR, a 5-aGNR, a 6-aGNR, and a 7-aGNR is plotted in Fig. 42.29. The permittivity in the center of all ribbons, except the 3-aGNR, is approximately $6\,\epsilon_0$, and the permittivity at the edges of the ribbon shows similar peaks and valleys oscillating around $7\epsilon_0$. These features show that ribbon size barely affects the out-of-plane microscopic dielectric constant.

Microscopic Dielectric Tensor of a 3 × 3 Silicon Nanowire

To touch upon silicon, although only very briefly, in Fig. 42.30 we show results obtained for the *xx*-component of the dielectric tensor of a Si nanowire, $\epsilon_r^{xx}(\mathbf{r})$, following the same procedure and using the empirical local pseudopotentials of Ref. [184]. Given the small cross section of the nanowire (only 3 × 3 atoms) and the fact that it is free-standing (surrounded by vacuum, it is not surprising to see that the dielectric function inside the nanowire reaches only a small value, less than 4. This is the result of extremely small nanostructures inheriting the polarization (and, so, the dielectric constant) of the surrounding vacuum or bulk material.

42.5.4 Microscopic Dielectric Permittivity of 2D Materials

Similar to the study of aGNRs, we construct a nanocapacitor with graphene as dielectric and the capacitor plates are placed parallel to the graphene surface. The potential perturbation is applied by imposing a voltage on the plates and then
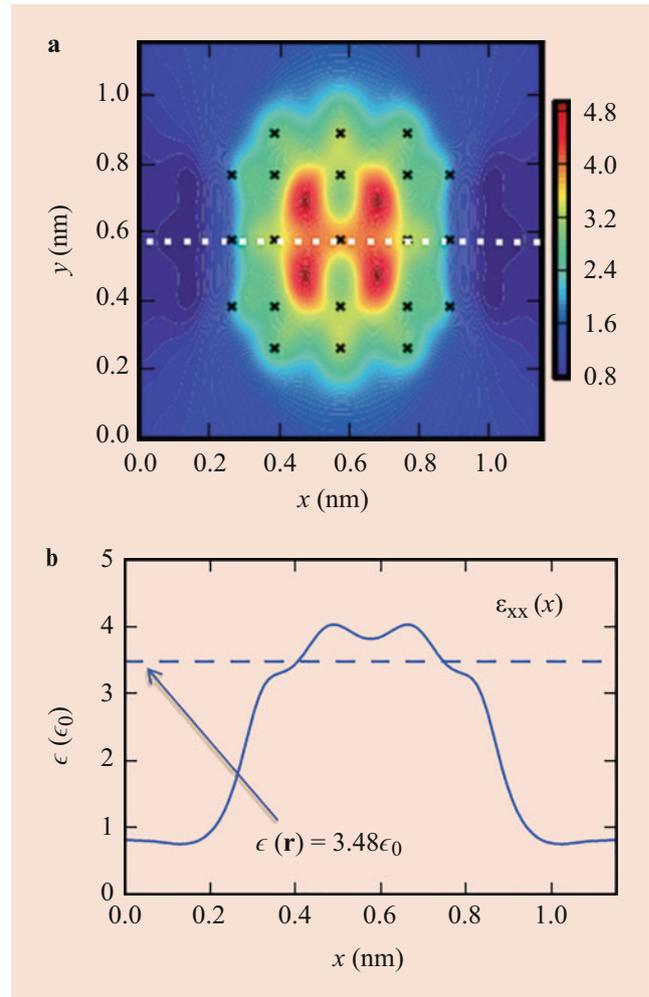
Fig. 42.30 (**a**) variation on the cross-sectional plane of the calculated diagonal *xx* component of the dielectric tensor of a small (3 × 3 atom in cross section) Si nanowire. (**b**) the same, but plotted along the dashed line in the left figure, showing how the maximum value of the dielectric constant in this very small nanostructure is affected by the surrounding medium, vacuum in this case. Note also the smooth transition between the semiconductor and the surrounding medium

we solve the microscopic Poisson equation. By solving the microscopic Poisson equation, Eq. (42.38), we obtain the induced charge density.

The induced charge density averaged on the (x, z) plane is shown in Fig. 42.31. The long-range oscillations of the induced charge are caused by local-field effects [171]. The polarization profile along *y* can be calculated by applying Gauss's law in 1D [185]. This is also shown in Fig. 42.31. The microscopic dielectric permittivity decays from $6.9\,\epsilon_0$ in the carbon-atom plane to the vacuum permittivity within approximately 0.1 nm. This shows that the out-of-plane dielectric permittivity of the aGNRs resembles that of graphene. We also evaluate the capacitance per unit area of the graphene nanocapacitor as D/V, where D is the displacement field and V is the applied voltage. We find a capacitance of
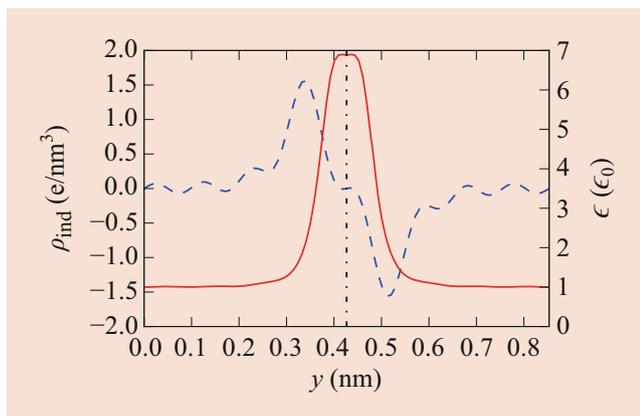
Fig. 42.31 Induced charge density (left *y*-axis, dashed blue line) when a voltage, $V = 2$ V, is applied to the graphene nanocapacitor and the microscopic dielectric constant of graphene along the direction perpendicular to the graphene plane (right *y*-axis, solid red line). In response to the potential perturbation, a dipole is formed across the graphene plane, as shown by the induced charge density. The microscopic dielectric permittivity exhibits a peak value of $6.9\epsilon_0$ on the carbon-atom plane and decays to the vacuum permittivity about 0.1 nm away from the carbon-atom plane, indicated by the dashed vertical line. (After J. Fang et al. [245], with permission from the American Physical Society ©2016 American Physical Society)

13.28 F/m^2. Equivalently, attributing an effective dielectric thickness, d, and a macroscopic dielectric constant, ϵ_{mac}, to graphene, the capacitance can also be calculated as $V/D = d/\epsilon_{\mathrm{mac}} + (d_0 - d)/\epsilon_0$ (d_0 is the distance between the capacitor plates) since graphene and vacuum are connected in series. Assuming a macroscopic dielectric constant of $6.9\epsilon_0$, the effective dielectric thickness of graphene is about 0.22 nm. This corresponds roughly to the distance of the center of the respective induced charge layer comprising the dipole shown in Fig. 42.31.

42.5.5 Discussion

We have demonstrated that a direct solution of the microscopic Poisson equation incorporates all quantum effects on the electrostatics. This offers a physical-sound avenue for atomistic self-consistent quantum transport simulations in nanotransistors [186] without knowledge of the dielectric tensor of the nanosystems. This is different from the traditional treatment of the Poisson equation in atomistic device simulations reported by Guo et al. in Refs. [187, 188] and Fiori et al. in Ref. [189]. In these simulations the Poisson equation is solved with macroscopic empirical approaches, either using an arbitrarily assumed uniform dielectric permittivity throughout the entire devices or by defining interfaces as "straight ϵ-discontinuity lines," lines arbitrarily drawn among atoms. Thus, physical accuracy is not ensured.

42.6 Quantum Transport

In this final section, we deal with transport beyond the semiclassical limit. Other chapters in this volume discuss this issue in great detail from the perspective of the Wigner function formalism (Chap. 43 by P. Dollfus) and of the nonequilibrium Green's function formalism (Chap. 44 by S. Datta), as implemented using a tight-binding model for the electronic dispersion (Chap. 45 by G. Klimeck). Therefore, we give here only a brief introduction to the subject.

The discussion presented so far has been based implicitly on the idea that electronic transport can be described correctly using semiclassical models, the BTE in particular. However, as early as in 1950, doubts were raised regarding the suitability of such a model to deal with transport at electric fields and energies so large that the associated short electron scattering times render the electronic Bloch states an unacceptably poor approximation of the correct eigenstates of the total Hamiltonian (*i.e.*, that includes also the perturbation). Bardeen and Shockley considered this problem for the case of Si [190], concluding that a semiclassical approach is satisfactory in most cases of practical technological interest. However, the first full-band Monte Carlo study [50] was performed to study impact ionization in GaAs, a process that results in an electron lifetime, τ, as short as 1 fs. Capasso et al. [191] argued that in such cases collisional broadening renders the semiclassical BTE invalid, that is, the energy broadening of the electronic state, $\Delta E \sim \hbar/\tau$, becomes comparable to – or even exceeds – the energy itself of the state. This implies a failure of first-order perturbation theory and Fermi's golden rule. This problem had been already investigated in the context of high-field transport in polar insulators using Feynman path integrals [192]. However, the daunting nature of this formulation prompted, instead, the use of Monte Carlo simulations that account, somehow, for these "broadening" effects [193], vaguely in the spirit of the Barker and Ferry formulation [194]. However, Monte Carlo simulations that employ scattering rates computed beyond first-order perturbation theory soon proved to be affected by serious artifacts due to violations energy conservation. High-field transport was later studied by Jacoboni et al. [195, 196] using the concept of Wigner paths. However, the basic physical nature of this problem is so difficult that the use of *ab initio* methods in this context remains completely unexplored.

A second, and completely different, limitation of semiclassical transport stems, instead, from the ultra-fast dynamics of electrons coupled to elementary excitations of a crystal when stimulated, for example, by ultrashort laser pulses in pump/probe experiments. The eigenstates of the total Hamiltonian are now coupled modes, such as electron-phonon modes, and the semiconductor Bloch equations (SBE), essentially a density-matrix formulation handled by solving the

Heisenberg equations of motion, have been successfully used in the past (see, *e.g.*, Ref. [197] for a comprehensive review).

A third, and final, limitation of the BTE is the limitation that at present is raising the strongest concerns in devices scaled at the nanometer length. This limitation is caused by the small size of the structures that are used by the VLSI technology. Quantum-size (or confinement) effects, which historically originated from studies of the two-dimensional electron gas (2DEG) that is formed in Si inversion layers [198], does not present a particular problem: indeed, transport must be studied in a reduced dimensionality (much as we have seen here when dealing with 2D materials) but can still be treated semiclassically, as long as the length of the active layer (channel) of the device is much longer than the electron wavelength or phase-coherence length. However, in devices whose channels are as short as 10 nm or less, so dimensions comparable to the wavelength of the charge carriers and of their decoherence and scattering length, electrons (or holes) may traverse the active region almost coherently (ballistic transport), so that their full quantum mechanical behavior – neglected by the BTE and in what we wrote so far – must be included in any formulation of transport that we must use to study these devices. Tunneling is another genuine and important quantum mechanical effect that must be captured, since it affects small devices via leakage through thin gate oxides and source-to-drain tunneling and that may also be exploited not only in electrical-erasable programmable random-access memories (E^2PROM and flash-EPROM) but also in tFETs and interlayer tunneling between 2D materials [199]. While tunneling has been modeled using semiclassical expressions, this approach can fail for general nonuniform fields [200–202].

A varied set of methods have been developed to tackle the quantum mechanical description of electronic devices. This variety is caused by the many choices and approximation made:

1. The physics that is included in the (effective) Hamiltonian: Exchange and correlation, spin-orbit coupling, magnetism, etc. Similarly to the historical evolution of methods used to deal with semiclassical transport (as discussed in Sect. 42.2, empirical or semiempirical methods have been used at first, notably empirical tight-binding in the first computer program based on the nonequilibrium Green's function method, NEMO [203–205], empirical pseudopotentials [39, 186, 206, 207], or linear combination of atomic bands [208]. However, also in this case the progress made on numerical methods and hardware have begun to make it possible to consider first-principles approaches. Indeed empirical tight-binding has been replaced by its DFT counterpart [209, 210], and we have already mentioned approaches based on all-electron calculations [64, 65] and even earlier DFT itself [29].

2. The basis that is used to discretize the Hamiltonian: Linear combination of localized orbitals (LCAO) – as in many tight-binding studies [211], plane waves, Bloch waves, and Wannier functions. This choice affect also first-principles formulations, as we have seen before.

3. The representation of the nonequilibrium state, that is, the formulation of the quantum transport equations, such as single particle wavefunctions, Landauer-Büttiker formulation [212–214], density matrix [197, 215], Wigner functions [196, 216, 217], or nonequilibrium Green's functions (NEGF [218–220]). Even in the ballistic case, the quantum-transmitting boundary method (QTBM) [221] can be chosen in place of a (somewhat improperly termed but highly popular) "ballistic NEGF" approach.

One additional crucial issue is constituted by inelastic scattering, that is, collisions with scatterers with internal degrees of freedom, so that memory of the phase of the carriers is lost (phase decoherence). When using the Wigner formulation, inelastic processes are often included in an approximated way via a Boltzmann-like collisional integral, since the "correct" formulation [222] is numerically intractable. The NEGF formulation is equally daunting, unless the scattering processes are "diagonalized" in some *ad hoc* fashion (so, rendered local) or unless small systems are considered, such as molecules [33, 34]. Nevertheless, this last method is arguably the method of choice, since, *in principle*, it is valid over the entire range of device dimensions, from long "semiclassical" channels to ultrashort "ballistic" devices. In particular, numerical difficulties notwithstanding, it can describe correctly the all-important intermediate regime, in which the device length is comparable to the other characteristic length-scales, namely, the electron wavelength, and the scattering/decoherence length associated with the various scattering processes, mainly, electron-phonon scattering, scattering with impurities or surface roughness.

In this section, we limit our discussion to length-scales short enough to render scattering processes a weak perturbation, starting from the ballistic limit, and extending in a natural way our methodology to include weak scattering processes. We employ plane waves since we consider them more suitable than localized orbitals to treat extended conduction states and we employ the simpler QTBM, rather than NEGF, because of its better numerical efficiency in the ballistic limit, a limit in which QTBM and NEGF are virtually identical from a physical perspective.

When considering *ab initio* transport within the density functional theory (DFT) framework, two options are available to approximate the Hamiltonian. The first option is to use the full self-consistent Kohn-Sham Hamiltonian, including exchange and correlation based on the nonequilibrium charge. However, since devices are large systems, self-consistently solving the Kohn-Sham Hamiltonian is com-

putationally expensive within the DFT framework. Since transport properties are only determined by electron states near the Fermi level, a second option is to use an effective low-energy Hamiltonian that is computationally less expensive. Such an effective Hamiltonian can be constructed from *ab initio* simulations of representative systems. A further benefit is that empirical data can supplement *ab initio* data in the construction of the effective Hamiltonian to correct for the deficiencies of DFT, such as underestimated band gaps and incorrect effective masses. Therefore, here we describe quantum electron transport within the pseudopotential approximation, using a plane-wave basis. However, the formulation can be easily generalized to first-principles methods. We first describe the ballistic transport formalism based on the extended states. We then detail the numerical approach based on envelope functions. Next, we introduce the addition of scattering using the Pauli Master Equation. Finally, we demonstrate our method by simulating highly scaled transistors based on graphene nanoribbons and silicon nanowires.

42.6.1 Ballistic Transport

In the ballistic limit, electron dynamics is described by the extended wavefunctions in the device and its infinitely long contacts. The extended wavefunctions are the solutions of the Schrödinger equation:

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + V_{\text{tot}}(\mathbf{r}) \right] \psi_{\kappa \nu, E}(\mathbf{r}) = E \psi_{\kappa \nu, E}(\mathbf{r}), \qquad (42.42)$$

where E is the energy of the electron injected from contact κ and ν is a mode index. The potential energy $V_{\text{tot}}(\mathbf{r})$ captures all of the potential energy felt by an individual particle; this includes the ionic potential, the mean-field Hartree potential generated by the free charges, doping and electrostatic gates, as well as exchange and correlation contributions. In the contacts, the total potential is assumed to be lattice periodic, yielding Bloch wave solutions with a well-defined wave vector $\mathbf{k}_{\kappa \nu}$.

We assume that the device's contacts are connected to large equilibrium reservoirs, each at a fixed chemical potential μ_κ. When a bias is applied to the device (between contacts κ and κ'), it is driven out of equilibrium by a relative change in the Fermi level of the contacts, *i.e.*, $\mu_\kappa - \mu_{\kappa'} = V_{\text{bias}}$. To describe the statistics, we introduce the ballistic density matrix, which is diagonal in the basis of the extended states:

$$\rho^{(\text{bal})}_{\kappa \nu E, \kappa' \nu' E'} = \delta_{\kappa \kappa'} \delta_{\nu \nu'} \delta_{E E'} f_{\text{FD}}(E - \mu_\kappa), \qquad (42.43)$$

where the Fermi-Dirac distribution is given by

$$f_{\text{FD}}(E - \mu_\kappa) = \left[1 - \exp\left(\frac{E - \mu_\kappa}{k_{\text{B}} T} \right) \right]^{-1}. \qquad (42.44)$$

Using this density matrix, the expressions for electron density, as well as current density are readily obtained:

$$\rho^{(\text{bal})}(\mathbf{r}) = \sum_{\kappa \nu} \int dE \left(\frac{dE_{\kappa \nu}}{dk} \right)^{-1}$$
$$\times \left| \psi_{\kappa \nu, E}(\mathbf{r}) \right|^2 f_{\text{FD}}(E - \mu_\kappa), \qquad (42.45)$$

$$\mathbf{J}^{(\text{bal})}(\mathbf{r}) = \sum_{\kappa \nu} \frac{e\hbar}{m} \int dE \left(\frac{dE_{\kappa \nu}}{dk} \right)^{-1}$$
$$\times \text{Im}[\psi_{\kappa \nu, E}(\mathbf{r}) \nabla \psi_{\kappa \nu, E}(\mathbf{r})] f_{\text{FD}}(E - \mu_\kappa), \quad (42.46)$$

where, since we have assumed that the injecting wavefunctions have been δ-normalized in their respective contacts, $(dE_{\kappa \nu}/dk)^{-1}$ is the density of states of mode ν contact κ.

42.6.2 Numerical Approach

In a numerical implementation, Eq. (42.42) needs to be discretized and the infinitely long contacts need to be truncated. In a periodic system, Eq. (42.42) is readily discretized on plane waves, $\psi_{\kappa \nu, E} = \sum_{\mathbf{G}} u_{\kappa \nu, E, G} e^{i\mathbf{G} \cdot \mathbf{r}} e^{i\mathbf{k}_{\kappa \nu, E} \cdot \mathbf{r}}$ using Bloch's theorem. Plane waves offer a mathematically complete basis that can be tuned to very high accuracy through the kinetic energy cutoff. By using the fast Fourier transform algorithm, plane waves have a relatively low computational cost and almost optimal scalability with system size. However, the application of bias and gate potentials breaks the periodic nature of the total potential in Eq. (42.42). This prevents the use of Bloch's theorem to express the electronic wavefunctions in the device. To circumvent this problem, the envelope function approach was developed [39, 186, 206]. More recently we developed a significantly more efficient method based on Bloch waves [223]. Once discretized, the truncation of the numerical domain is achieved by replacing the infinite contacts with a self-energy contribution, regardless of the chosen discretization.

While the Bloch wave method demonstrates significant progress in numerical efficiency, the envelope function has the benefit of conceptual clarity. For this reason, we highlight here the envelope function approach and its associated self-energies as an introduction to quantum transport simulations using pseudopotentials. We advise the interested reader to research advanced techniques, such as the Bloch wave-based method [223].

Envelope Functions

In the envelope-wavefunction approximation, we express the full electronic wavefunction at a given energy E as:

$$\psi_{\kappa v,E}(\mathbf{r}) = \sum_{\mathbf{G}} \phi_{\kappa v,E,\mathbf{G}}(\mathbf{r}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{42.47}$$

where the envelope function in the periodic case $e^{i\mathbf{k}_{\kappa v,E}\cdot\mathbf{r}}$ is replaced by the more general envelope $\phi_{\kappa v,E,\mathbf{G}}$. The dependence of the envelope ϕ on $\mathbf{r}$ results indeed from the breaking of the periodicity caused by the applied bias and doping.

Using the envelope expansion in Eq. (42.47) in the Schrödinger equation (Eq. 42.42) yields an equation for the envelope functions

$$\sum_{\mathbf{G}'} \left\{ \left[\frac{\hbar^2}{2m}(-i\nabla + G)^2 + \bar{V}_{\text{tot}}(\mathbf{r}) \right] \delta_{\mathbf{G},\mathbf{G}'} + V_{\mathbf{G}-\mathbf{G}'} \right\}$$
$$\times \phi_{\kappa v,E,\mathbf{G}}(\mathbf{r}) = E\phi_{\kappa v,E,\mathbf{G}}(\mathbf{r}) \tag{42.48}$$

where we have separated the total potential $V_{\text{tot}}(\mathbf{r}) = \bar{V}_{\text{tot}}(\mathbf{r}) + \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}$ into a slowly varying part $\bar{V}_{\text{tot}}(\mathbf{r})$ and the periodic lattice potential, captured by $V_{\mathbf{G}}$.

After discretization of the $\mathbf{r}$ coordinate, and using a finite difference scheme for the gradient, we obtain a sparse matrix equation:

$$\mathbf{H}\boldsymbol{\phi} = E\boldsymbol{\phi} \Leftrightarrow \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \mathbf{T}^\dagger & \mathbf{D}_{i-1} & \mathbf{T} & 0 & 0 & \cdot \\ \cdot & 0 & \mathbf{T}^\dagger & \mathbf{D}_i & \mathbf{T} & 0 & \cdot \\ \cdot & 0 & 0 & \mathbf{T}^\dagger & \mathbf{D}_{i+1} & \mathbf{T} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot \\ \boldsymbol{\phi}(\mathbf{r}_{i-1}) \\ \boldsymbol{\phi}(\mathbf{r}_i) \\ \boldsymbol{\phi}(\mathbf{r}_{i+1}) \\ \cdot g \end{bmatrix}$$
$$= E \begin{bmatrix} \cdot \\ \boldsymbol{\phi}(\mathbf{r}_{i-1}) \\ \boldsymbol{\phi}(\mathbf{r}_i) \\ \boldsymbol{\phi}(\mathbf{r}_{i+1}) \\ \cdot \end{bmatrix}. \tag{42.49}$$

where the explicit matrix structure is given for the one-dimensional case, for which $\mathbf{D}_i$ and $\mathbf{T}$ are block matrices of size $(N_{\mathbf{G}} \times N_{\mathbf{G}})$ [186]. The vector containing the discretized envelope function contains sub-vectors of size $N_{\mathbf{G}}$:

$$\boldsymbol{\phi}(\mathbf{r}_i) = [\phi_{\mathbf{G}_1}(\mathbf{r}_i), \phi_{\mathbf{G}_2}(\mathbf{r}_i), \dots, \phi_{\mathbf{G}_{N_{\mathbf{G}}}}(\mathbf{r}_i)]^\mathsf{T}. \tag{42.50}$$

Self Energies

In order to solve Eq. 42.49 numerically, we truncate it to the device region, i.e.:

$$\boldsymbol{\phi} = [\boldsymbol{\phi}(\mathbf{r}_1), \boldsymbol{\phi}(\mathbf{r}_2), \dots, \boldsymbol{\phi}(\mathbf{r}_N)]^\mathsf{T}. \tag{42.51}$$

This necessitates the introduction of boundary conditions (the so-called quantum-transmitting boundary conditions) at the contacts that can inject and absorb waves. Since we have forced the potential in each contact to be periodic, the solutions in the contacts are Bloch waves. Specifically, for a given energy E, the envelope functions of the eigenmodes $\psi_\kappa^{(c)}$ of a contact κ located at $\mathbf{r}_{i_\kappa}$ are found, using Bloch's theorem, as:

$$\left[\mathbf{T}^\dagger e^{-ik_\kappa \Delta} + \mathbf{D}_{i_\kappa} + \mathbf{T} e^{ik_\kappa \Delta} \right] \boldsymbol{\phi}_\kappa^{(c)}(\mathbf{r}_{i_\kappa}) = E\boldsymbol{\phi}_\kappa^{(c)}(\mathbf{r}_{i_\kappa}), \tag{42.52}$$

where the Δ is distance between r_{i_κ} and $r_{i_\kappa - 1}$.

For each energy E, Eq. (42.52) admits $2N_{\mathbf{G}}$ solutions $(k_\kappa, \boldsymbol{\phi}_\kappa^{(c)})$. Classifying the contact solutions into incoming and outgoing modes, we assemble the "Bloch matrices," $\left[\boldsymbol{\phi}_\kappa^{(c)} \right]_{\text{in}}$ and $\left[\boldsymbol{\phi}_\kappa^{(c)} \right]_{\text{out}}$, containing the respective solutions as their columns. Using these Bloch matrices, we calculate the contact self-energies directly as:

$$[\boldsymbol{\Sigma}]_\kappa = \mathbf{T}' \left[\boldsymbol{\phi}_\kappa^{(c)} \right]_{\text{out}} \text{diag} \left[e^{ik_\kappa \Delta} \right]_{\text{out}} \left[\boldsymbol{\phi}_\kappa^{(c)} \right]_{\text{out}}^{-1}, \tag{42.53}$$

where $\text{diag}\left[e^{ik_\kappa \Delta} \right]_{\text{out}}$ is a diagonal matrix containing the phase factors of their respective solutions in the Bloch matrices and $\mathbf{T}'$ is either $\mathbf{T}$ or $\mathbf{T}^\dagger$ depending on the position of the contact. From right to left, the effect of the matrices in Eq. (42.53) is to (1) project the envelope at the edge onto the basis of contact solution, (2) propagate each solution one step (Δ) inside the contact, (3) reassemble the envelope function, and (4) apply the truncated part of the Hamiltonian. For a detailed derivation of the various parts, we refer the reader to Ref. [186].

With the definition of the self-energies $[\boldsymbol{\Sigma}]_\kappa$, the envelope functions ψ obey the following linear system of differential equations [186]:

$$(\mathbf{H} - E\mathbf{I} + [\boldsymbol{\Sigma}]_\kappa)\boldsymbol{\phi} = [\mathbf{r.h.s.}]_\kappa^{\text{inj}}, \tag{42.54}$$

where $\mathbf{I}$ is the identity matrix and $[\mathbf{r.h.s.}]_\kappa^{\text{inj}}$ is the contact injection from contact κ that decides which mode gets injected.

Self-Consistency

At a given bias, we solve Eq. (42.54) for all electron injection energies with a rough initial guess on the external potential and obtain their respective envelope functions and the spatial electronic wavefunctions. We calculate the carrier density using Eq. (42.45), summing over all the energy states with their corresponding thermal occupation set by the Fermi level of each contact. We then solve the (macroscopic or microscopic, see Sect. 42.5.2) Poisson equation, to update the Hartree potential and, if accounted for, update the exchange and correlation contributions. This procedure is repeated using Newton-Rhapson iteration until the root-mean-square potential-difference between two successive it-

erations is smaller than a preset convergence criterion. To accelerate convergence, we succesfully employed the direct inversion of the iterative subspace (DIIS) method, also known as Pulay mixing in the chemistry community [223]. After convergence, we calculate the current density from the final envelope functions.

42.6.3 Scattering—The Pauli Master Equation

In realistic small devices, the ballistic limit is rarely, if ever, achieved. To account for scattering in short devices where scattering is still relatively weak, we can use the Pauli Master equation approach [39, 224–226]. In this approach, we replace the noninteracting ballistic density matrix $\rho^{(\mathrm{bal})}$ with the density matrix ρ for the interacting system.

By invoking the Van Hove limit, Fischetti [225] showed that the time-dependence of the elements of the density matrix is given by:

$$
\begin{aligned}
\frac{\partial}{\partial t}\rho_{\kappa \nu E, \kappa' \nu' E'}(t) = &-\frac{i}{\hbar}(E - E')\rho_{\kappa \nu E, \kappa' \nu' E'}(t) \\
&- \sum_{\kappa'' \nu'' E''} \left[W_{\kappa'' \nu'' E'', \kappa \nu E} + W_{\kappa'' \nu'' E'', \kappa' \nu' E'} \right] \rho_{\kappa \nu E, \kappa' \nu' E'}(t) \\
&+ \sum_{\kappa'' \nu'' E''} \delta_{\kappa \nu E, \kappa' \nu' E'} W_{\kappa \nu E, \kappa'' \nu'' E''} \rho_{\kappa'' \nu'' E'', \kappa'' \nu'' E''}(t).
\end{aligned}
\tag{42.55}
$$

Thanks to the Van Hove limit, where we assume the scattering process completes much faster than the dynamics of the electrons themselves, we can ignore memory effects and phase information. The scattering rates $W(\kappa \nu E, \kappa' \nu' E')$ for a perturbing potential V' are then obtained using Fermi's golden rule, in much the same way as for the semiclassical Monte-Carlo approach:

$$
W(\kappa \nu E, \kappa' \nu' E') = \frac{\left| \langle \kappa \nu E | V' | \kappa' \nu' E' \rangle \right|^2}{|E - E'|}.
\tag{42.56}
$$

When considering small devices, where electrons scatter only rarely before being absorbed by a contact, the off-diagonal parts of the density matrix are negligible and the Pauli Master equation reduces to:

$$
\begin{aligned}
\frac{\partial}{\partial t}\rho_{\kappa \nu E, \kappa \nu E}(t) \\
= \sum_{\kappa' \nu' E'} \left[W_{\kappa \nu E, \kappa' \nu' E'} \rho_{\kappa' \nu' E', \kappa' \nu' E'}(t) \right. \\
\left. - W_{\kappa' \nu' E', \kappa \nu E} \rho_{\kappa \nu E, \kappa \nu E}(t) \right].
\end{aligned}
\tag{42.57}
$$

To account for the exchange of particles through the contact reservoirs, Fischetti [224] introduced an empirical term:

$$
\begin{aligned}
\left(\frac{\partial}{\partial t}\rho_{\kappa \nu E, \kappa \nu E}(t) \right)_{\mathrm{res}} \\
= |A_{\kappa \nu E}|^2 \, v_{\kappa \nu E, \perp} \left[f_{\mathrm{FD}}(E - \mu_\kappa) - \rho_{\kappa \nu E, \kappa \nu E}(t) \right],
\end{aligned}
\tag{42.58}
$$

where particles (waves) are injected from a contact κ that is in equilibrium with chemical potential μ_κ with a group velocity $v_{\kappa \nu E, \perp}$ normal to the contact. $A_{\kappa \nu E}$ is the normalization factor applied to the injected wavefunction such that $\int_\Omega \mathrm{d}^3 r |\psi_{\kappa \nu E}(\mathbf{r})|^2 = 1$, where Ω is the volume of the device.

For device transport we are mostly interested in the steady state. In this case, the injection from the contacts needs to balance the scattering, *i.e.*, the right-hand sides of Eqs. (42.57) and (42.58) must be equal. The resulting linear matrix equation can be solved very efficiently using existing linear algebra packages.

42.6.4 Example 1: Graphene Nanoribbon Transistors

To study transport through graphene nanoribbons, we use the empirical pseudopotentials from Ref. [181] with an energy cutoff of 15 Ry. The atoms have been positioned in a perfect sp^2 lattice, without relaxation of the hydrogen atom positions. We calculate the band structure of several armchair graphene nanoribbons (aGNR), shown in Fig. 42.32a. In the given energy range, 7-, 10-, 13-aGNRs show two conducting bands, whereas the 6-aGNR shows only one conducting band. Fischetti et al. have shown that the electron effective mass of aGNRs at the conduction band minimum oscillates, so that, for example, $m_6^* < m_7^* < m_{13}^* < m_{10}^*$, where m_n^* represents the effective mass of the n-aGNR [60]. Figure 42.32b shows the dispersion of the lowest conduction band for 3-, 6-, 9-, and 12-aGNRs, and it indicates that $m_{12}^* < m_9^* < m_6^* < m_3^*$ [60]. The difference of the effective mass can result in different transport behavior of each ribbon, especially in the subthreshold region of device operation, as we will discuss later.

Ribbons with different body widths exhibit different band gaps, as shown in Fig. 42.32c [199]. The 3n-aGNRFETs and (3n+1)-aGNRFETs with different channel lengths and different ribbon widths are simulated to investigate their device performance. The (3n+2)-aGNRs are not of interest here, since their small band gap results in a very poor subthreshold leakage as a consequence of a high tunneling current.

Transport Characteristics of aGNRFETs

The I_{DS}-V_{GS} characteristics are presented in Fig. 42.33a,b. Clearly, transistors with different widths show different transfer characteristics. We show the ribbon width dependence of the current at a gate overdrive, V_{GS}-$V_{\mathrm{th}} = 0.25$ V, and subthreshold swing in Fig. 42.35.
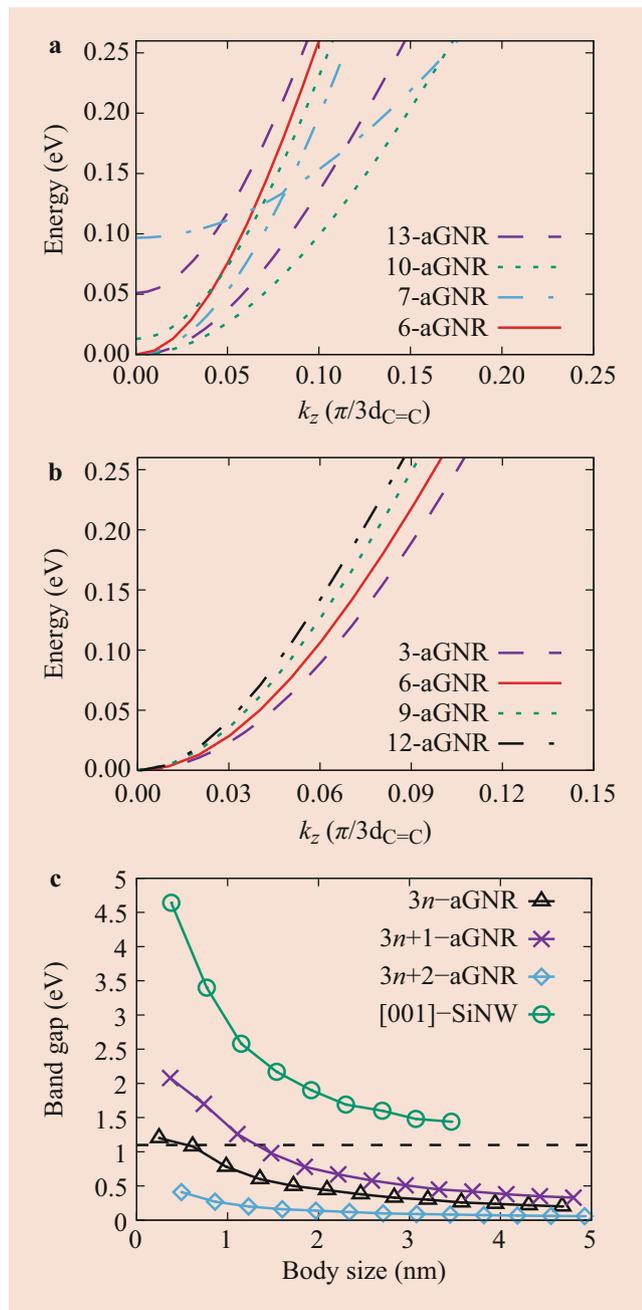
Fig. 42.32 (**a**, **b**) Conduction band dispersion of several aGNRs in a given energy range from E_{CBM}, the conduction band minimum, to $E_{CBM} + 10k_BT$ ($T = 300$ K). The energy of E_{CBM} is set as zero. (**c**) Body-size dependence of the band gap of aGNRs and [001]-oriented silicon nanowires. The horizontal dashed line indicates the band gap of bulk silicon, which is 1.1 eV. The index n is an integer which characterizes the ribbon width. The lines connecting the symbols are a guide to the eye. (After Ref. [206], with permission from IEEE)

Figure 42.32c illustrates the band gap of aGNRs with different body size (*i.e.*, ribbon width). For a ribbon wider than 5 nm, the band gap approaches the vanishing band gap of graphene. The band gap increases as the ribbon width de-
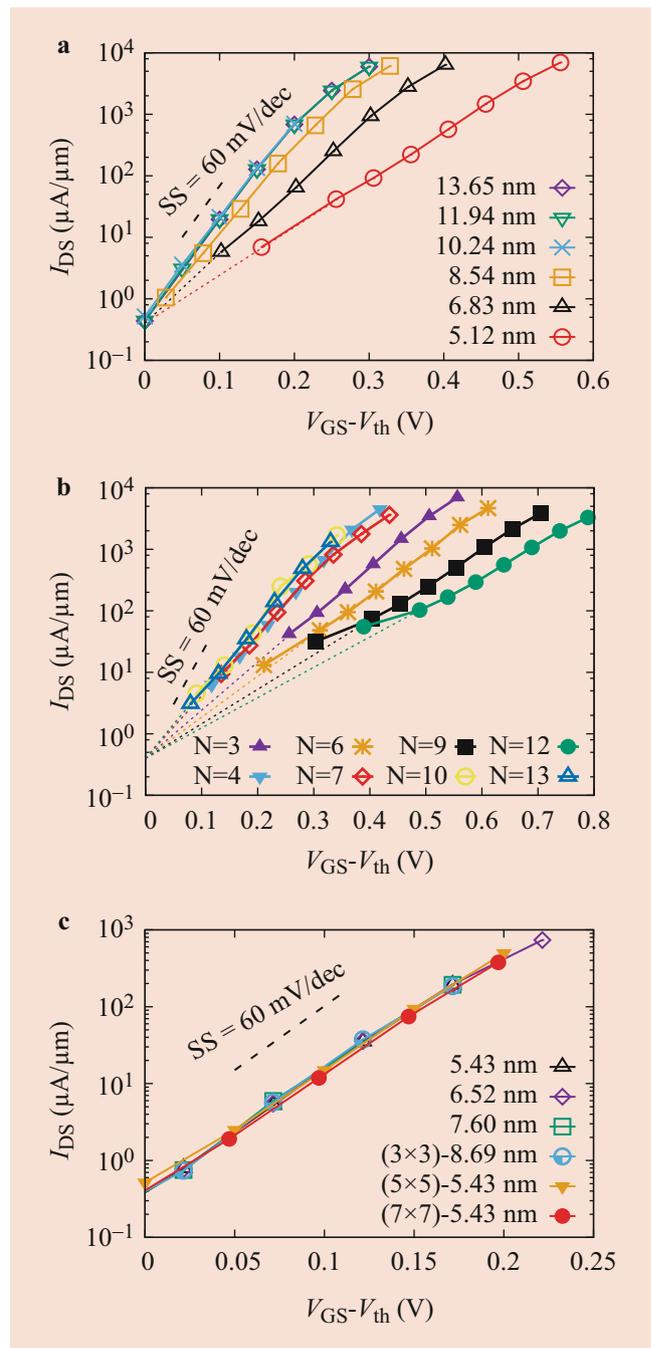
Fig. 42.33 (**a**) Gate-length dependence of transfer characteristics for 3-aGNRFETs at $V_{DS} = 0.1$ V (taken as representative of $V_{DD} = 0.4$ V, thanks to the saturation of the $I_{DS} - V_{DS}$ curves for $V_{DS} > 0.1$ V). The device gate length, L_G, ranges from 5.12 nm to 13.65 nm. (**b**) I_{DS}-V_{GS} transfer characteristics at $V_{DS} = 0.1$ V for aGNRFETs with the same gate length $L_G = 5.12$ nm and with different widths. (**c**) I_{DS}-V_{GS} transfer characteristics of four (3×3)-SiNWFETs with different gate lengths (from 5.43 nm to 8.69 nm) and of a (5×5)-SiNWFET and a (7×7)-SiNWFET both with a gate length of 5.43 nm, at $V_{DS} = 0.1$ V. The dotted lines extending to the threshold voltage have been obtained by extrapolating the current calculated at higher gate bias using the calculated subthreshold slope. The 60 mV/dec slope and the lines connecting the symbols are a guide to the eye. (After Ref. [206], with permission from IEEE)

creases as a consequence of quantum confinement [227]. The band gap for a ribbon with a width about 4 nm is about 0.4 eV, which matches the experimental result reported in Ref. [228]. The gap of a 7-aGNR is also experimentally observed to be ~1.8 eV [229]. These facts highlight the accuracy of band structure calculation based on empirical pseudopotentials, while maintaining computational efficiency. The 3-aGNRFETs with a body width $W_{body} = 0.25$ nm and a gate length L_G ranging from 5.12 nm (12 cells along z) to 13.65 nm (32 cells along z) are simulated. The source and the drain regions are both set as ~6 nm, and we assume no gate-channel overlap. Both the source and the drain regions are doped as n-type with a density of 10^8 m^{-1}, and the channel region is doped as p-type with a density of 10^8 m^{-1}. Abrupt junctions are assumed. We use "vacuum" as a virtual gate insulator in the transport model. The size of "vacuum," t_{vac}, and its dielectric constant, ϵ, are related to the EOT as $t_{vac}/\epsilon = \text{EOT}/\epsilon_{SiO_2}$, where EOT is set as 0.43 nm and ϵ_{SiO_2} is the dielectric constant of silicon dioxide.

The current-voltage characteristics of these 3-aGNRFETs are shown in Fig. 42.33a. We define here the threshold voltage, V_{th}, as the gate bias at which the drain current reaches a preset value $I_{off} = 0.4\,\mu\text{A}/\mu\text{m}$. This definition allows us to compare various devices at the same I_{off} and the same gate overdrive $V_{GS} - V_{th} = 0.25$ V. Also we assume a power supply voltage $V_{DD} = 0.4$ V [1]. Since I_{DS} saturates for $V_{DS} > 0.1$ V, we shall consider the on-state at $V_{DS} = 0.1$ V as representative of the performance at $V_{DS} = V_{DD}$. When the gate length is larger than 10 nm, the current at a given bias is approximately independent of gate length. This is due to the assumption of ballistic transport. This is in contrast with "the linear model" or "the quadratic model" in traditional metal-oxide-semiconductor field-effect transistors (MOSFETs) according to which carriers can scatter within the device when traveling along the channel. Thus, increasing the device length will increase the channel resistance proportionally and reduce the current (*i.e.*, $I_{DS} \propto 1/L$) [230]. For devices with a gate length below 10 nm, the current decreases with a decrease of the gate length at the same given gate overdrive due to the significant quantum-mechanical source-to-drain tunneling effect that causes a very poor turn-off behavior and thus threshold voltage shift for devices with a ultrashort gate length [231].

This is shown as well in Fig. 42.34a which compares current for devices with different channel lengths at a given gate overdrive, $V_{GS} - V_{th} = 0.25$ V. The current does not show a strong dependence on the channel length when it is above 10 nm, whereas it decreases by two orders of magnitude when the channel length is below 10 nm.

The subthreshold behavior is shown in Fig. 42.34b. The subthreshold swing reaches the ideal performance of 60 mV/dec for traditional MOSFETs [232] when $L_G > 10$ nm. However, a shorter-channel device presents a degraded subthreshold swing due to the source-to-drain tunneling. Thus, for a given power supply V_{DD}, the I_{on}/I_{off} current ratio decreases for a shorter device.

42.6.5 Example 2: Silicon Nanowire FETs

Gate-all-around SiNWFETs with square and smaller cross sections are simulated as well. The SiNWs are with axis along the [100] crystallographic (transport) direction and with {100} surfaces since this yields a high ballistic conductance (at least in large cross-section NWs, Ref. [40], Fig. 7.34) while retaining a transport mass that for small cross-section NWs is large enough [233] to minimize source-to-drain tunneling. The band gap of [001]-oriented SiNWs with a square cross section and different body sizes (*i.e.*, NW side lengths) is also shown in Fig. 42.32c. When the NW side length is larger than 3.5 nm, the band gap approaches the band gap of bulk silicon. The electron effective mass at the conduction band minimum is $1.94m_0$, $0.88m_0$, and $0.73m_0$, respectively, for (3×3) SiNW, (5×5) SiNW, and (7×7) SiNW. The (3×3) SiNWFETs with gate lengths ranging from 5.43 nm (10 cells along z) to 8.69 nm (16 cells along z) are simulated. All of these nanowires have three atomic layers along each side of the cross section and thus have a side length of 0.39 nm (Ref. [186], Fig. 3). The surfaces of the NWs are hydrogen-terminated. The current-voltage transfer characteristics of these devices are shown in Fig. 42.33c. The current is the same at a given V_{DS} and V_{GS} for all gate lengths, and thus they have the same threshold voltage. Thus, there is no appreciable source-to-drain tunneling for a (3×3) SiNWFET with a ~5 nm gate length. Normalizing the current at $V_{GS} - V_{th} \approx 0.25$ V and $V_{DS} = 0.1$ V to the NW side length (footprint), the current density reaches around $1000\,\mu\text{A}/\mu\text{m}$, which satisfies the performance requirement for continued device scaling suggested by ITRS [146]. The subthreshold swing is 66 mV/decade and I_{on}/I_{off} ratio is $\sim2 \times 10^3$.

In Fig. 42.34a we compare their performance with aGNRFETs in a similar range of gate lengths. It is obvious that SiNWFETs have a much larger current than aGNRFETs when the gate length is below 10 nm. Especially, at the 5 nm gate length, (3×3) SiNWFET shows a current ~25 times larger than 3-aGNRFET.

Both a (5×5) SiNWFET and a (7×7) SiNWFET are simulated to investigate the body-size effect of NWs on the device performance. They both have a gate length of 5.43 nm and a side length of 0.78 nm and 1.17 nm. These large cross-section SiNWFETs show very similar current density with the (3×3) SiNWFETs at the same gate overdrive.

The current at a gate overdrive of 0.25 V and the subthreshold swing are illustrated in Fig. 42.35. In Fig. 42.34 we compare the performance of SiNWFETs with that of aGNRFETs at different body sizes. The SiNWFETs show a
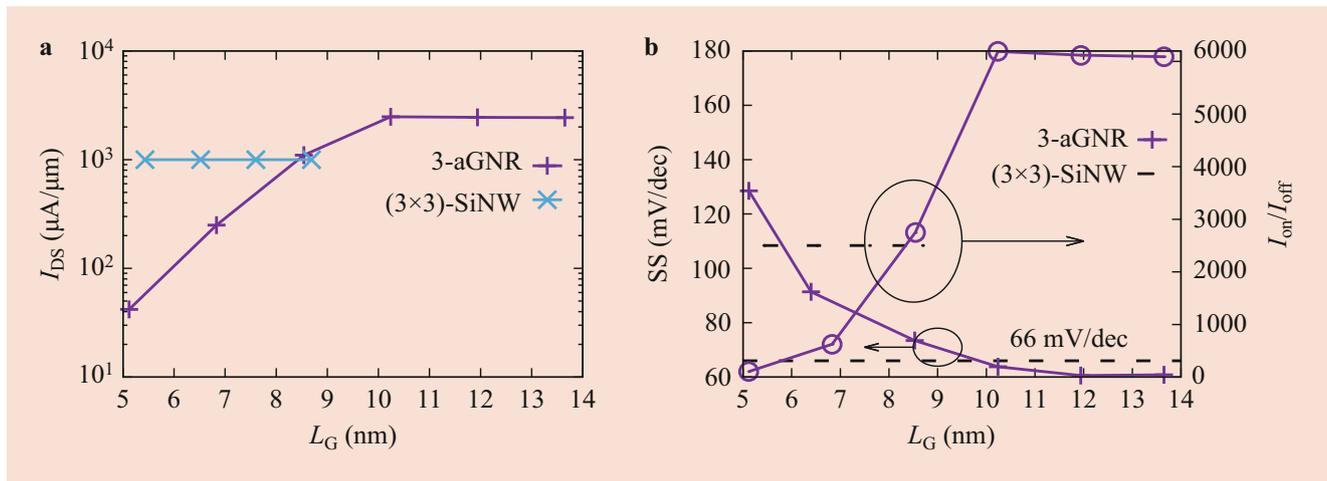
Fig. 42.34 (**a**) Drain-source current I_{DS} for 3-aGNRFETs and (3×3)-SiNWFETs with different L_G at $V_{DS} = 0.1$ V and $V_{GS} - V_{th} = 0.25$ V. (**b**) Subthreshold swing (SS) and I_{on}/I_{off} ratio as a function of L_G. The horizontal dashed line indicates the subthreshold swing, 66 mV/dec, for (3×3)-SiNWFETs in the given channel lengths. Lines connecting the symbols are a guide to the eye. (After Ref. [206], with permission from IEEE)
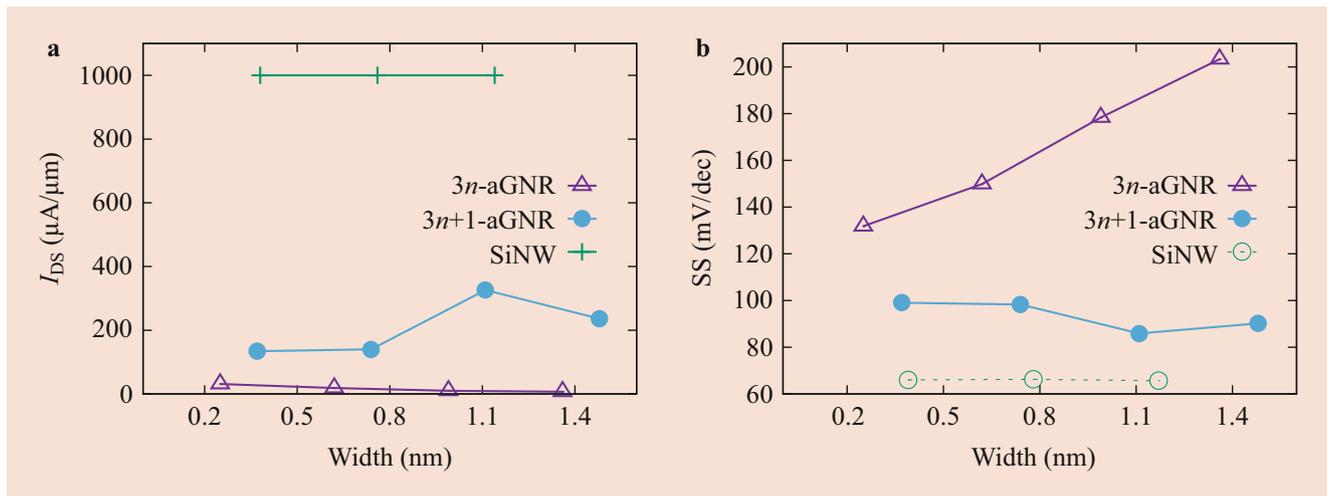
Fig. 42.35 (**a**) The current I_{DS} at $V_{DS} = 0.1$ V and $V_{GS} - V_{th} = 0.25$ V for aGNRFETs and SiNWFETs with different width and a similar gate length (~5 nm). (**b**) The SS as a function of width. The lines connecting the symbols are a guide to the eye. (After Ref. [206], with permission from IEEE)

current more than three times larger than aGNRFETs for the same gate length of 5 nm. To sum the studies shown above, SiNWFETs offer the avenue to scale transistors down to 5 nm channel lengths.

42.7 Conclusions and Outlook

So far, we have shown how DFT can be used to calculate electron-phonon scattering rates and transport properties of several bulk and two-dimensional materials. Comparing the power of these *ab initio* methods to what was possible as recently as 20 years ago (as discussed in Sect. 42.2), we can appreciate the enormous "cultural" value of the progress

made. We must also express optimism about the possibility of investigating the transport properties of new materials without having to wait for expensive and time-consuming experiments that provide the information needed to assess the performance of devices based on these materials.

However, in a few occasions, we have also hinted at some uncertainty that still affects the accuracy and reliability of results obtained using *ab initio* methods. Broadly speaking, these difficulties originate from three causes: *(i)* the use of unsatisfactory approximations to calculate the carrier mobility and other transport characteristics, *(ii)* numerical "shortcuts" that reduce the computational efforts but introduce unwanted artifacts and errors, and *(iii)* intrinsic limitations of our knowledge of the basic physics.

(i) Regarding the accuracy and correctness of the formulations employed to compute transport characteristics, we have discussed in Ref. [97] how several results, labeled *ab initio*, rely on DFT calculations of the changes of the band structure under strain (the "proper" deformation potentials) to estimate the carrier mobility using oversimplified assumptions. To summarize what we wrote in Ref. [97] (and as also discussed in Refs. [234, 235]), these assumptions consists in the use of Takagi formula [236] to calculate the carrier mobility for a two-dimensional electron gas (2DEG):

$$\mu = \frac{e\hbar^3 C_{2D}}{k_B T m^* m_d E_1^2} , \qquad (42.59)$$

where m^* and m_d are the conductivity and density-of-states effective masses, respectively, C_{2D} is the longitudinal or transverse elastic constant of the 2D material, and E_1 is the "proper" deformation potential mentioned above and calculated from the energy-shift, $\Delta E_{c,v}$, of the relevant band-edge (conduction for electron transport, valence for holes), under a relative change, $\Delta a/a_0$, of the lattice constant a_0:

$$E_1 = a_0 \frac{\Delta E_{c,v}}{\Delta a} . \qquad (42.60)$$

This formula has been used to calculate extremely high values for the carrier mobility in silicene ($\approx 2 \times 10^5$ cm^2 V^{-1} s^{-1} for both electrons and holes [237]) and germanene ($\approx 6 \times 10^5$ cm^2 V^{-1} s^{-1} for both electrons and holes [238]), ignoring coupling to flexural acoustic modes [141]. Qiao et al. [137] have also used Eq. (42.59) to obtain an equally impressive electron and hole mobility in phosphorene. However, while approximately correct in the context of Si inversion layers, as originally intended by Takagi and coworkers, its validity should be considered carefully when extending it to other materials. Indeed, Eq. (42.59) accounts only for longitudinal phonons (unless the shear deformation potential is also accurately considered, as done in Ref. [144] for phosphorene), it ignores scattering with optical phonons, intervalley processes and ignores the anisotropy (*i.e.*, the dependence on the scattering angle) of the electron-phonon matrix element. This anisotropy is extremely important when estimating the carrier mobility, since it affects strongly the momentum relaxation rate. This has been shown even in bulk Si [36], strained bulk Si [239], and Si inversion layers [240]. Liao et al. [142] and Gaddemane et al. [97] have shown how this is also of paramount importance in monolayer phosphorene. Finally, Eq. (42.59) ignores effects that may result from the overlap integral of the initial and final wavefunctions, effects that are instead captured by the finite-displacement method or by DFPT. Therefore, the impressive values for the carrier mobility reported in Refs. [137, 237], among others, should

Table 42.7 Theoretical calculations of the 300 K electron and hole mobility, μ_e and μ_h, in monolayer and bilayer phosphorene

Reference	μ_e (cm^{2}V^{-1}s^{-1})		μ_h (cm^{2}V^{-1}s^{-1})	
	Armchair	Zigzag	Armchair	Zigzag
Monolayers				
Qiao et al.[a]	1100	80	640–700	10,000–26,000
Jin et al.[b]	210	40	460	90
Rudenko et al.[c]	738	114	292	157
Rudenko et al.[d]	∼700		∼250	
Trushkov et al.[e]	625	82		
Liao et al.[f]	170	50	170	35
This work[g]	20	10	19	2.4
This work[h]	21	10	19	3
This work[i]	25	5		
Bilayers				
Qiao et al.[a]	600	140–160	2600–2800	1300–2200
Jin et al.[b]	1020	360	1610	760
This work[g]	14	7	12	2

After G. Gaddenane et al. [97], with permission from the American Physical Society ©2018 American Physical Society
[a]Ref. [137]
[b]Ref. [143], Monte Carlo and DFT (DFPT)
[c]Ref. [144], LA and TA, one-phonon processes
[d]Ref. [144], LA and TA, one- and two-phonon processes
[e]Ref. [145], LA and TA, at a density of 10^{13} electrons/cm^2
[f]Ref. [142], DFT (DFPT)
[g]Monte Carlo and DFT (finite differences), acoustic and optical phonons
[h]Monte Carlo and DFT (DFPT)
[i]Kubo-Greenwood, acoustic phonons only, elastic and equipartition approximation

be regarded as no more than extremely optimistic upper bounds.

We show in Table 42.7 (reproduced from Ref. [97]) the large discrepancies among the values of the electron mobility in phosphorene calculated using several different approximations or methods. As serious as the issue may be, paradoxically we should consider it a "minor problem," since reliable methods to calculate the electron mobility are available, such as the EPW computer package [77] and the by-now "conventional" full-band Monte Carlo simulations we have employed here.

(ii) Regarding numerical issues, we have also discussed in them in Ref. [97], and we have mentioned above in Sect. 42.4.1. These may be considered mundane issues, since the use of a sufficiently fine discretization of the Brillouin zone does not constitute an insurmountable conceptual obstacle. Even if today's hardware renders the task painful and (perhaps) impossible when dealing with high-energy transport, we are growing accustomed to ever-increasing computer resources, in terms of both speed and memory.

(iii) Much more fundamental are the reasons causing the dependence of the results on the choice of exchange-and-correlation functionals and even (very surprising to us) of the pseudopotentials used when performing *ab initio* calcula-

Table 42.8 Calculated electron mobility in monolayer WS_2

	SSSP-PBE without SO	ONCV-PBE without SO	ONCV-PBE with SO	ONCV-LDA without SO	ONCV-LDA with SO
Bandgap (eV)	1.81	1.835	1.56	1.99	1.66 (indirect)
ΔE_{KQ} (meV)	210	182	97	22	73
μ (cm^2/(Vs))	767	380	215	17	127

Table 42.9 Calculated electron mobility in monolayer TMDs

μ (cm^2/Vs)	SSSP-PBE without SO	ONCV-PBE without SO	ONCV-PBE with SO	ONCV-LDA without SO	ONCV-LDA with SO
MoS_2	127	147	145	127	116
$MoSe_2$	78	92	101	18	19
$MoTe_2$	49	43	50	17	11
WS_2	767	380	215	37	127
WSe_2	275	180	46	23	88
WTe_2	64	166	181	60	55

tions. In Sect. 42.4.3 we have already observed the surprising wide range of values for the calculated electron mobility in monolayer phosphorene that have been reported in the literature. Unfortunately, a similar puzzling observation can be made for monolayer TMDs. A typical example is provided by WS_2. As shown in Tables 42.8, the electron mobility calculated using DFT varies from a promising 767 cm^2/(Vs) to a disappointing 17 cm^2/(Vs). Indeed, different values are obtained when using different exchange and correlation functionals (PBE [94] or LDA [95]), different pseudopotentials ("standard solid-state pseudopotentials," SSSP [241] or "optimized norm-conserving," ONCV [242]) and depending on whether or not one accounts for the spin-orbit (SO) interaction.

The cause of this problem is twofold: first, the band structure of monolayer WS_2 shows conduction band minima at the six K symmetry points (the doubly degenerate K valleys) and local minima at a point along the Γ-K line (the sixfold degenerate Q valleys). The energy separation, ΔE_{KQ}, between these two sets of valleys obtained choosing different "flavors" of DFT using Quantum ESPRESSO [15, 16] varies between 22 and 210 meV. Since the K-Q intervalley scattering rates calculated using EPW [76] are quite strong, clearly different values of ΔE_{KQ} result in vastly different values of the mobility. In this case, we may blame the well-known inability of DFT – a ground-state theory thanks to the Hohenberg-Kohn theorems [243] – to reproduce the correct excitation spectrum with the accuracy required when performing transport calculations. However, in some sense, experimental information can guide us in choosing the proper combination of exchange-correlation functionals and pseudopotentials. In so doing, the problem is addressed very pragmatically, so that we have to accept some compromises in how we define the term "*ab initio*."

A similar problem affects our understanding of high-field electron transport in Si and GaN: regarding Si, Fig. 42.3 shows that DFT/EPW calculations yield peaks of the density of states at energies that are slightly smaller than those measured experimentally (and reproduced, by definition, by empirical pseudopotentials). This "compression" of the conduction bands results in a larger density of states that, in turn, results in scattering rates that are slightly larger than those predicted by rigid-(pseudo)ion calculations based on empirical pseudopotentials. Similarly, in GaN the energy of the satellite L-valleys computed from DFT seems to be inconsistent with some experimental data, as we have mentioned in Sect. 42.4.2. Ultimately and not surprisingly, the problem originates from the approximations that we must necessarily make in dealing with the exchange and correlation functionals that enter the Kohn-Sham equation. We should add, parenthetically, that accounting for the spin-orbit interaction still presents some fundamental issues.

Much more serious is the second cause of the broad range of calculated results. As shown in Table 42.9, even in materials in which intervalley scattering does not control transport, some differences are seen. If one may expect some dependence of the results on the choice made for the exchange-correlation functionals, more worrisome is the observation that even the choice made for the pseudopotentials matters. The situation is acceptable only in some particular cases, such as for MoS_2, but it remains generally unsatisfactory.

As a final important remark, we should note that results like those reported in Tables 42.8 and 42.9 have been obtained assuming free-standing 2D layers. Experiments are often performed on layers supported by insulating substrates, often also gated (occasionally supported *and* gated), and at a nonzero carrier density. Under these circumstances, dielectric screening by the dielectric environment (as reported for impurity scattering in Ref. [244]) and free-carrier screening may improve the mobility, whereas scattering with the hybrid plasmon/optical-phonon interface modes (often called "remote-phonon" scattering, somewhat improperly) may depress it. Therefore, comparison with experimental data should be made keeping these effects in mind.

If the choice of functionals and pseudopotential is an issue, using an all-electron method may solve the latter problem and employing the GW approximation may address the former issue. At present, such calculations are possible only in some particular cases (also employing the simpler G_0W_0 approximation). However, having already remarked how accustomed we have grown in witnessing the progress of numerical algorithms and computer hardware, we may speculate that in a not-too-distant future the term *ab initio*

will be considered synonymous of all-electron/GW, rather than DFT.

References

1. Hänsch, W.: The Drift Diffusion Equation and Its Applications in MOSFET Modeling. Springer, Wien (1991)
2. Baccarani, G.: Hfields: Proceedings of the NASECODE IV Conference, pp. 3–12. Boole Press (1985)
3. Rudan, M., Odeh, F.: COMPEL **5**, 149–183 (1987)
4. Fischetti, M.V., Laux, S.E.: Phys. Rev. B **38**, 9721–9745 (1988)
5. Geim, A.K., Novoselov, K.S.: Nat. Mater. **6**(3), 183–191 (2007)
6. Zhang, W., et al.: Nano Res. **7**, 1731–1731 (2014)
7. Houssa, M., Dimoulas, A., Molle, A. (eds.): 2D Materials for Nanoelectronics. CRC Press-Taylor&Francis (2016)
8. Molle, A., et al.: Nat. Mater. **16**, 163–163 (2017)
9. Jones, R.O.: Rev. Mod. Phys. **87**, 897–897 (2015)
10. Hasnip, P.J., et al.: Philos. Trans. Series A Math. Phys. Eng. Sci. **372**, 20130270 (2014)
11. Kresse, G., Hafner, J.: Phys. Rev. B **47**, RC558 (1993)
12. Kresse, G.: *Ab initio* molekular dynamik für flüssige metalle. Ph.D. Thesis (1993)
13. Kresse, G., Furthmüller, J.: Comput. Mater. Sci. **6**, 15 (1996)
14. Kresse, G., Furthmüller, J.: Phys. Rev. B **54**, 11169 (1996)
15. Giannozzi, et al.: J. Phys. Condens. Matter **21**, 395502 (2009)
16. Giannozzi, P., et al.: J. Phys. Condens. Matter **29**(46), 465901 (2017). https://doi.org/10.1088%2F1361-648x%2Faa8f79
17. Vandenberghe, W.G., Fischetti, M.V.: Appl. Phys. Lett. **106**, 013505 (2015)
18. Baroni, S., et al.: Rev. Mod. Phys. **73**, 515 (2001)
19. Giustino, F.: Rev. Mod. Phys. **89**, 015003 (2017)
20. Ziman, J.M.: Electrons and Phonons. Oxford University Press, Oxford (1958)
21. Zollner, S., Gopalan, S., Cardona, M.: Appl. Phys. Lett. **68**, 614 (1989)
22. Zollner, S., Gopalan, S., Cardona, M.: J. Appl. Phys. **68**, 1682 (1990)
23. Zollner, S., Gopalan, S., Cardona, M.: Phys. Rev. B **44**, 13446–13451 (1991)
24. Zollner, S., Gopalan, S., Cardona, M.: Semicond. Sci. Technol. **7**, B137 (1992)
25. Fischetti, M.V., Higman, J.: Monte Carlo Device Simulation: Full Band and Beyond, pp. 123–160. Kluwer Academic (1991)
26. Yoder, P.D., et al.: Semicond. Sci. Technol. **7**, B357 (1992)
27. Yoder, P.D., Natoli, V.D., Martin, R.M.: J. Appl. Phys. **73**, 4378 (1993)
28. Yoder, P.D., Natoli, V.D., Martin, R.M.: Semicond. Sci. Technol. **9**, 852 (1994)
29. Choi, H.J., Ihm, J.: Phys. Rev. B **59**, 2267 (1999)
30. Gunst, T., et al.: Phys. Rev. B **93**, 035414 (2016)
31. Hadjisavvas, G., Tsetseris, L., Pantelides, S.T.: IEEE Electron Device Lett. **28**, 1018–1020 (2007)
32. Restrepo, O.D., Varga, K., Pantelides, S.T.: Appl. Phys. Lett. **94**, 212103 (2009)
33. Di Ventra, M., Pantelides, S.T., Lang, N.D.: Phys. Rev. Lett. **84**, 979–982 (2000)
34. Pecchia, A., et al.: Phys. E Low-Dimension. Syst. Nanostruct. **19**, 139 (2003)
35. Bardeen, J., Shockley, W.: Phys. Rev. **80**, 72–80 (1950)
36. Herring, C., Vogt, E.: Phys. Rev. **101**, 944–961 (1956)
37. Cohen, M.L., Bergstresser, T.K.: Phys. Rev. **141**, 779 (1966)
38. Chelikowsky, J.R., Cohen, M.L.: Phys. Rev. B **14**, 556 (1976)
39. Fischetti, M., et al.: Nano-Electronic Devices: Semiclassical and Quantum Transport Modeling, pp. 183–247. Springer, New York (2011)
40. Fischetti, M.V., Vandenberghe, W.G.: Advanced Physics of Electron Transport in Semiconductors and Nanostructures. Springer, Berlin (2016)
41. Allen, P.B., Heine, V.: J. Phys. C: Solid State Phys. **9**, 2305–2312 (1976)
42. Allen, P.B., Cardona, M.: Phys. Rev. B **27**, 4760 (1983)
43. Lautenschlager, P., Allen, P.B., Cardona, M.: Phys. Rev. B **31**, 2163 (1985)
44. Zollner, S., Gopalan, S., Cardona, M.: Solid State Commun. **77**, 485–488 (1991)
45. Varshni, Y.P.: Physica **34**, 149–154 (1967)
46. Poncé, S., et al.: J. Chem. Phys. **143**, 102813 (2015)
47. Maradudin, A.A., et al.: Solid State Physics, Suppl. 3, 2nd edn. Academic Press, New York (1971)
48. Kunc, K., Nielsen, O.H.: Comput. Phys. Commun. **17**, 413 (1976)
49. Shah, J., et al.: Phys. Rev. Lett. **59**, 2222 (1987)
50. Shichijo, H., Hess, K.: Phys. Rev. B **23**, 4197–4207 (1981)
51. Fischetti, M.V., et al.: Appl. Phys. Lett. **114**, 222104 (2019)
52. Tang, J.Y., Hess, K.: J. Appl. Phys. **54**, 5139 (1983)
53. Tang, J.Y., Hess, K.: J. Appl. Phys. **54**, 5145 (1983)
54. Brennan, K., Hess, K.: Solid-State Electron. **27**(4), 347–357 (1984)
55. Kunikiyo, T., et al.: J. Appl. Phys. **75**, 297–312 (1994)
56. Abramo, A., et al.: IEEE Trans. Electron Devices **ED-41**, 1646 (1994)
57. Fischetti, M.V., et al.: J. Technol. Comput. Aided Design **10**, 1 (1996)
58. Cartier, E.A., et al.: Appl. Phys. Lett. **62**, 3339 (1993)
59. Fang, J., et al.: IEEE Trans. Nucl. Sci. **66**, 444 (2019)
60. Fischetti, M.V., et al.: J. Phys. Condens. Matter **25**(47), 473202 (2013)
61. Wunsch, B., et al.: New J. Phys. **8**, 318 (2006)
62. Weinert, M., Wimmer, E., Freeman, A.J.: Phys. Rev. B **26**, 4571 (1982)
63. Schwarz, K., Blaha, P., Madsen, G.K.H.: Comput. Phys. Commun. **147**, 71–76 (2002)
64. Chen, Z., Polizzi, E.: Phys. Rev. B **82**, 205410 (2010)
65. Levin, A.R., Zhang, D., Polizzi, E.: Comput. Phys. Commun. **183**, 2370–2375 (2012)
66. Ke, S.H.: Phys. Rev. B **84**, 205415 (2011)
67. Mortensen, J.J., Hansen, L.B., Jacobsen, K.W.: Phys. Rev. B **71**, 035109 (2005)
68. Enkovaara, J., et al.: J. Phys. Condens. Matter **22**, 253202 (2010)
69. Clark, S.J., et al.: Z. Kristallogr. **220**, 567–570 (2005)
70. Gonze, X., et al.: Comput. Mater. Sci. **25**, 478 (2002)
71. Izquierdo, J., et al.: Phys. Rev. B **61**, 13639 (2000)
72. Soler, J.M., et al.: J. Phys. Condens. Matter **14**, 2745–2779 (2002)
73. Smidstrup, S., et al.: Phys. Rev. B **96**, 195309 (2017)
74. Smidstrup, S., et al.: J. Phys. Condens. Matter (2019)
75. Stokbro, K., et al.: Ann. N. Y. Acad. Sci. **1006**, 212–226 (2002)
76. Poncé, S., et al.: Comput. Phys. Commun. **209**, 116–133 (2016)
77. Poncé, S., Margine, E.R., Giustino, F.: Phys. Rev. B **97**, 121201(R) (2018)
78. Laasonen, K., et al.: Phys. Rev. B **47**(16), 10142 (1993)
79. Pickett, W.E.: Comput. Phys. Rep. **9**(3), 115–197 (1989)
80. Blöchl, P.E.: Phys. Rev. B **50**(24), 17953 (1994)
81. Kresse, G., Joubert, D.: Phys. Rev. B **59**, 1758 (1999)
82. Feynman, R.P.: Phys. Rev. **56**, 340 (1939)
83. Fletcher, R.: Practical Methods of Optimization. Wiley, London (2013)

42

84. Baroni, S., et al.: Rev. Mod. Phys. **73**(2), 515 (2001)
85. Chaput, L., et al.: Phys. Rev. **84**, 094302 (2011)
86. Togo, A., Chaput, L., Tanaka, I.: Phys. Rev. **91**, 094306 (2015)
87. Togo, A., Tanaka, I.: Scr. Mater. **108**, 1–5 (2015)
88. Kurosawa, T.: J. Phys. Soc. Jpn. **21**, 66 (1966)
89. Price, P.J.: Lasers, Junctions, Transport, vol. 14, pp. 249–308. Academic Press, New York, San Francisco, London (1979)
90. Jacoboni, C., Reggiani, L.: Rev. Mod. Phys. **55**, 645 (1983)
91. Jacoboni, C., Lugli, P.: The Monte Carlo Method for Semiconductor Device Simulation. Springer, Wien New York (1991)
92. Gilat, G., Raubenheimer, L.J.: Phys. Rev. B **144**, 390 (1966)
93. Gilat, G., Raubenheimer, L.J.: Phys. Rev. B **147**, 670 (1966)
94. Perdew, J.P., Burke, K., Ernzerhof, M.: Phys. Rev. Lett. **77**, 3865–3868 (1996)
95. Perdew, J.P., Zunger, A.: Phys. Rev. Lett. **77**, 3865 (1996)
96. Vanderbilt, D.: Phys. Rev. B **41**, 7892 (1990)
97. Gaddemane, G., et al.: Phys. Rev. B **98**(11), 115416 (2018)
98. Rivas, C., et al.: Appl. Phys. Lett. **78**(6), 814–816 (2001)
99. Kao, K.H., et al.: J. Appl. Phys. **116**(21), 214506 (2014)
100. Goano, M., et al.: J. Appl. Phys. **88**(11), 6467–6475 (2000)
101. Farahmand, M., et al.: IEEE Trans. Electron Devices **48**(3), 535–542 (2001)
102. Iveland, J., et al.: Phys. Rev. Lett. **110**, 177406 (2013)
103. Bertazzi, F., et al.: Appl. Phys. Lett. **106**(3), 061112 (2015)
104. Piccardo, M., et al.: Phys. Rev. B **89**(48), 235124 (2014)
105. Marcinkevicius, S., et al.: Phys. Rev. B **94**(48), 235205 (2016)
106. Rubio, A., et al.: Phys. Rev. B **94**(48), 11810 (1993)
107. Delaney, K.T., Rinke, P., Van de Walle, C.G.: Appl. Phys. Lett. **94**(19), 191109 (2009)
108. Chow, T.P., Ghezzo, M.: MRS Proc. **423**, 69–73 (1996)
109. Guo, Q., Yoshida, A.: Jpn. J. Appl. Phys. **33**(5R), 2453 (1994)
110. Schlipf, M., Gygi, F.: Comput. Phys. Commun. **196**, 36–44 (2015)
111. Verdi, C., Giustino, F.: Phys. Rev. Lett. **115**, 176401 (2015)
112. Giustino, F., Cohen, M.L., Louie, S.G.: Phys. Rev. B **76**, 165108 (2007)
113. Mostofi, A.A., et al.: Comput. Phys. Commun. **178**(9), 685–699 (2008)
114. Bulutay, C.: Semicond. Sci. Technol. **17**(10), L59 (2002)
115. Sun, C.K., et al.: Phys. Rev. B **59**, 13535–13538 (1999)
116. Lagerstedt, O., Monemar, B.: Phys. Rev. B **19**, 3064–3070 (1979)
117. Drechsler, M., et al.: Jpn. J. Appl. Phys. **34**(9B), L1178 (1995)
118. Witowski, A., et al.: Appl. Phys. Lett. **75**(26), 4154–4155 (1999)
119. Suzuki, M., Uenoyama, T.: J. Appl. Phys. **80**(12), 6868–6874 (1996)
120. Xu, Y.N., Ching, W.Y.: Phys. Rev. B **48**, 4335–4351 (1993)
121. Wang, S., et al.: J. Electron. Mater. **47**, 1560–1568 (2018)
122. Vitanov, S., Nedjalkov, M., Palankovski, V.: Numerical Methods and Applications, pp. 197–204. Springer, Berlin Heidelberg, Berlin, Heidelberg (2007)
123. Kim, K., et al.: Phys. Rev. B **56**, 7363–7375 (1997)
124. Bhapkar, U.V., Shur, M.S.: J. Appl. Phys. **82**(4), 1649–1655 (1997)
125. Sarasamak, K., Limpijumnong, S., Lambrecht, W.R.L.: Phys. Rev. B **82**, 035201 (2010)
126. Fröhlich, H.: Adv. Phys. **3**, 325 (1954)
127. Levinshtein, M.E., Rumyantsev, S.L., Shur, M.S.: Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe. Wiley, London (2001)
128. Chen, S., Wang, G.: J. Appl. Phys. **103**(2), 023703 (2008)
129. Albrecht, J., et al.: J. Appl. Phys. **83**(3), 1446–1449 (1998)
130. Silveira, E., et al.: Phys. Rev. B **71**, 041201 (2005)
131. Collins, A.T., Lightowlers, E.C., Dean, P.J.: Phys. Rev. **158**, 833–838 (1967)
132. Taniyasu, Y., Kasu, M., Makimoto, T.: Appl. Phys. Lett. **89**(18), 182112 (2006)
133. Fang, J., et al.: Phys. Rev. Appl. **11**, 044045 (2019)
134. Akahama, Y., Endo, S., Narita, S.I.: J. Phys. Soc. Jpn. **52**(6), 2148–2155 (1983)
135. Morita, A.: Appl. Phys. A **39**(4), 227–242 (1986)
136. Blöchl, P.E.: Phys. Rev. B **50**, 17953–17979 (1994)
137. Qiao, J., et al.: Nat. Commun. **5**, 4475 (2014)
138. Tran, V., et al.: Phys. Rev. B **89**, 235319 (2014)
139. Zhang, S., et al.: ACS Nano **8**(9), 9590–9596 (2014)
140. Togo, A., Tanaka, I.: Scr. Mater. **108**, 1–5 (2015)
141. Fischetti, M.V., Vandenberghe, W.G.: Phys. Rev. B **93**, 155413 (2016)
142. Liao, B., et al.: Phys. Rev. B **91**(23), 235419 (2015)
143. Jin, Z., et al.: Appl. Phys. Lett. **109**(5), 053108 (2016)
144. Rudenko, A., Brener, S., Katsnelson, M.: Phys. Rev. Lett. **116**(24), 246401 (2016)
145. Trushkov, Y., Perebeinos, V.: Phys. Rev. B **95**(7), 075436 (2017)
146. International Technology Roadmap for Semiconductor: Process integration, devices, and structures
147. Gaddemane, G., et al.: J. Appl. Phys. **124**(4), 044306 (2018)
148. Liu, C.C., Jiang, H., Yao, Y.: Phys. Rev. B **84**(19), 195430 (2011)
149. Liu, C.C., Feng, W., Yao, Y.: Phys. Rev. Lett. **107**(7), 076802 (2011)
150. Yakovkin, I.: Surf. Sci. **662**, 1–5 (2017)
151. Mariani, E., von Oppen, F.: Phys. Rev. Lett. **100**, 076801 (2008)
152. Mariani, E., von Oppen, F.: Phys. Rev. B **82**, 195403 (2008)
153. Gornyi, I.V., Kachorovskii, V.Y., Mirlin, A.D.: Phys. Rev. B **86**, 165413 (2012)
154. Gornyi, I.V., Kachorovskii, V.Y., Mirlin, A.D.: Phys. Rev. B **92**, 155428 (2015)
155. Amorim, B., Guinea, F.: Phys. Rev. B **88**, 115418 (2013)
156. Mermin, N.D., Wagner, H.: Phys. Rev. Lett. **17**, 1133 (1966)
157. Ong, Z.Y., Pop, E.: Phys. Rev. B **84**, 075471 (2011)
158. Rudenko, A.N., et al.: Phys. Rev. B **100**, 075417 (2019)
159. Han, M.Y., Kim, P.: Nano Convergence **1**(1), 1–11 (2014)
160. Dürkop, T., et al.: Nano Lett. **4**(1), 35–39 (2004)
161. Cui, Y., et al.: Nano Lett. **3**(2), 149–152 (2003)
162. Tao, L., et al.: Nat. Nanotechnol. **10**(3), 227 (2015)
163. Li, L., et al.: Nat. Nanotechnol. **9**(5), 372 (2014)
164. Desai, S.B., et al.: Science **354**(6308), 99–102 (2016)
165. Delerue, C., Lannoo, M., Allan, G.: Phys. Rev. B **68**(11), 115411 (2003)
166. Giustino, F., Pasquarello, A.: Phys. Rev. B **71**(14), 144104 (2005)
167. Nakamura, J., et al.: J. Appl. Phys. **99**(5), 054309 (2006)
168. Zhang, G., et al.: IEEE Electron Device Lett. **29**(12), 1302–1305 (2008)
169. Cartoixà, X., Wang, L.W.: Phys. Rev. Lett. **94**, 236804 (2005)
170. Kunc, K., Resta, R.: Phys. Rev. Lett. **51**, 686–689 (1983)
171. Resta, R., Kunc, K.: Phys. Rev. B **34**, 7146–7157 (1986)
172. Benedict, L.X., Louie, S.G., Cohen, M.L.: Phys. Rev. B **52**, 8541–8549 (1995)
173. Kozinsky, B., Marzari, N.: Phys. Rev. Lett. **96**, 166801 (2006)
174. Vasiliev, I., Öğüt, S., Chelikowsky, J.R.: Phys. Rev. Lett. **78**, 4805–4808 (1997)
175. Wang, L.W., Zunger, A.: Phys. Rev. Lett. **73**(7), 1039–1042 (1994)
176. Meyer, B., Vanderbilt, D.: Phys. Rev. B **63**, 205426 (2001)
177. Del Sole, R., Fiorino, E.: Phys. Rev. B **29**(8), 4631–4645 (1984)
178. Harl, J.: The linear response function in density functional theory. Ph.D. Thesis, Vienna, Austria (2008)
179. Laturia, A., de Put, M.L.V., Vandenberghe, W.G.: NPJ 2D Mater. Appl. **2**, 6 (2018)
180. Hybertsen, M.S., Louie, S.G.: Phys. Rev. B **35**(11), 5585 (1987)
181. Kurokawa, Y., et al.: Phys. Rev. B **61**, 12616 (2000)
182. Adler, S.L.: Phys. Rev. **126**(2), 413 (1962)
183. Wiser, N.: Phys. Rev. **129**(1), 62 (1963)
184. Zhng, S.B., Yeh, C., Zunger, A.: Phys. Rev. B **48**, 11204 (1993)
185. Shi, N., Ramprasad, R.: Phys. Rev. B **74**(4), 045318 (2006)

186. Fang, J., et al.: J. Appl. Phys. **119**(3), 035701 (2016)
187. Guo, J.: Nanoscale **4**, 5538–5548 (2012)
188. Ouyang, Y., Yoon, Y., Guo, J.: IEEE Trans. Electron Dev. **54**(9), 2223–2231 (2007)
189. Fiori, G., Iannaccone, G.: IEEE Electron Device Lett. **28**(8), 760–762 (2007)
190. Bardeen, J., Shockley, W.: Phys. Rev. **80**, 69 (1950)
191. Capasso, F., et al.: J. Appl. Phys. **53**, 3324 (1982)
192. Thornber, K.K., Feynman, R.P.: Phys. Rev. B **1**, 4099 (1979)
193. Kim, K., Mason, B.A., Hess, K.: Phys. Rev. B **36**, 6547 (1987)
194. Barker, J.R., Ferry, D.K.: Phys. Rev. Lett. **42**, 1779 (1979)
195. Pascoli, M., et al.: Phys. Rev. B **58**, 3503 (1998)
196. Jacoboni, C., Bordone, P.: Rep. Progr. Phys. **67**, 1033 (2004)
197. Rossi, F., Kuhn, T.: Rev. Mod. Phys. **74**, 895 (2002)
198. Ando, T., Fowler, A.B., Stern, F.: Rev. Mod. Phys. **54**, 437 (1982)
199. Van de Put, M.L., et al.: J. Appl. Phys. **119**(21), 214306 (2016). https://doi.org/10.1063/1.4953148
200. Vandenberghe, W.G., et al.: Appl. Phys. Lett. **98**(14), 143503 (2011). https://doi.org/10.1063/1.3573812
201. Van de Put, M.: Eurocon 2013, pp. 2134–2139. IEEE (2013). https://doi.org/10.1109/EUROCON.2013.6731011
202. Verreck, D., et al.: J. Appl. Phys. **115**(5), 053706 (2014). https://doi.org/10.1063/1.4864128
203. Lake, R., Datta, S.: Phys. Rev. B **45**, 6670 (1992)
204. Lake, R., et al.: J. Appl. Phys. **81**, 7845 (1997)
205. Klimeck, G., et al.: Comput. Model. Eng. Sci. **3**, 601 (2002)
206. Fang, J., et al.: IEEE Trans. Electron Devices **64**(6), 2758–2764 (2017)
207. Pala, M.G., Palestri, P.: J. Appl. Phys. **126**, 055703 (2019)
208. Esseni, D., Palestri, P.: Phys. Rev. B **72**, 165342 (2005)
209. Tan, Y.P., et al.: Phys. Rev. B **92**, 085301 (2015)
210. Szabó, À., et al.: IEEE Trans. Electron Devices **65**, 4180 (2008)
211. Luisier, M., Schenk, A., Fichtner, W.: J. Appl. Phys. **100**, 043713 (2006)
212. Landauer, R.: IBM J. Res. Dev. **1**, 233 (1957)
213. Büttiker, M.: Phys. Rev. B **38**, 9735 (1988)
214. Büttiker, M., Prêtre, A., Thomas, H.: Phys. Rev. Lett. **70**, 4114 (1992)
215. Iotti, R.C., Ciancio, E., Rossi, F.: Phys. Rev. B **72**, 125347 (2005)
216. Weinbub, J., Ferry, D.K.: Appl. Phys. Rev. **5**, 041104 (2018)
217. de Put, M.L.V., Sorée, B., Magnus, W.: J. Comput. Phys. **350**, 314 – 325 (2017). https://doi.org/10.1016/j.jcp.2017.08.059
218. Kadanoff, L.P., Baym, G.: Quantum Statistical Mechanics. Benjamin, New York (1962)
219. Keldysh, L.V.: Sov. Phys. J. Exp. Theor. Phys. **20**, 1018 (1965)
220. Haug, H., Jauho, A.P.: Quantum Kinetics in Transport and Optics of Semiconductors. Springer, Berlin (1996)
221. Lent, C.S., Kirkner, D.J.: J. Appl. Phys. **67**, 6353 (1998)
222. Jonasson, O., Knezevic, I.: J. Comput. Electron. **14**, 041104 (2018)
223. Van de Put, M.L., Fischetti, M.V., Vandenberghe, W.G.: Comput. Phys. Commun. **244**, 156–169 (2019). https://doi.org/10.1016/j.cpc.2019.06.009
224. Fischetti, M.V.: J. Appl. Phys. **83**(1), 270–291 (1998). https://doi.org/10.1063/1.367149
225. Fischetti, M.V.: Phys. Rev. B **59**(7), 4901–4917 (1999). https://doi.org/10.1103/PhysRevB.59.4901
226. Knezevic, I., Novakovic, B.: J. Comput. Electron. **12**(3), 363–374 (2013). https://doi.org/10.1007/s10825-013-0474-7
227. Han, M.Y., et al.: Phys. Rev. Lett. **98**(20), 206805 (2007)
228. Wang, X., Dai, H.: Nat. Chem. **2**(8), 661–665 (2010)
229. Huang, H., et al.: Sci. Rep. **2**, 983 (2012)
230. Taur, Y., Ning, T.H.: Fundamentals of modern VLSI devices. Cambridge Univ. Press, Cambridge, (2013)
231. Kawaura, H., Sakamoto, T., Baba, T.: Appl. Phys. Lett. **76**(25), 3810–3812 (2000)
232. Cheung, K.P.: Proceedings of 2010 International Symposium on VLSI Technology, System and Application, pp. 72–73. IEEE (2010)
233. Neophytou, N., et al.: IEEE Trans. Electron Devices **55**(6), 1286–1297 (2008)
234. Nakamura, Y., et al.: Adv. Electron. Mater. **3**, 1700143 (2017)
235. Poncé, S., et al.: Rep. Prog. Phys. **83**, 036501 (2020)
236. Takagi, S.I., et al.: IEEE Trans. Electron Devices **41**, 2357 (1994)
237. Shao, Z.G., et al.: J. Appl. Phys. **114**, 093712 (2013)
238. Ye, X.S., et al.: R. Soc. Chem. Adv. **4**, 21216 (2014)
239. Fischetti, M.V., Laux, S.E.: J. Appl. Phys. **80**, 2234 (1996)
240. Fischetti, M.V., Laux, S.E.: Phys. Rev. B **48**, 2244 (1993)
241. Prandini, G., et al.: https://doi.org/10.24435/materialscloud:2018.0001/v1
242. Hamann, D.R.: Phys. Rev. B **88**, 085117 (2013)
243. Hohenberg, P., Kohn, W.: Phys. Rev. **136**(3B), B864 (1994)
244. Ong, Z.Y., Fischetti, M.V.: Phys. Rev. B **88**, 165316 (2013)
245. Fang, J., et al.: Phys. Rev. B **94**, 045318 (2016)

42

Massimo Fischetti (PhD from the University of California – Santa Barbara) has been at the IBM T. J. Watson Research Center (1983–2005) and at the University of Massachusetts – Amherst (2005–2010). He is now Texas Instruments Distinguished professor in Nanoelectronics at the University of Texas at Dallas. His interests focus on the theory of electron transport in solids.

William Vandenberghe received the M.Sc. degree (2007) and the Ph.D. degree (2012) from KU Leuven, Belgium. From 2012–2016, he was a research associate/scientist at the University of Texas at Dallas in the Department of Materials Science and Engineering. Dr. Vandenberghe was appointed assistant professor at UT Dallas in 2016 and promoted to Associate Professor in 2021. He studies electronic transport at the nanoscale using theoretical methods.

Maarten L. Van de Put received his PhD from the University of Antwerp in 2016, working at imec in Belgium. He joined the University of Texas at Dallas as a post-doctoral researcher. He currently works as a Research Scientist, focussing on the development of advanced modeling and simulation capabilities to study quantum transport in nanostructures and devices.

Jingtian Fang received his PhD from the University of Texas at Dallas in 2016. He worked at Vanderbilt University Department of Physics and Astronomy as postdoctoral scholar and Boston University Department of Electrical and Computer Engineering as postdoctoral research associate from 2016 to 2018. Currently he works in Synopsys, Inc. as senior application engineer.

Gautam Gaddemane received his PhD from the University of Texas at Dallas in 2018. He has worked as a postdoctoral researcher at KU Leuven/imec, Belgium from 2019–2021. He is currently working as R & D Engineer in the TCAD group at imec, Belgium. His research interests include materials and device modeling using semi-classical and quantum approaches.