ISSN 1674-1056
http://cpb.iphy.ac.cn
http://iopscience.iop.org/cpb
中国物理 B
Chinese Physics
中国物理
Chinese Physics B

Group velocity matters for accurate prediction of phonon-limited carrier mobility

Qiao-Lin Yang(杨巧林), Hui-Xiong Deng(邓惠雄), Su-Huai Wei(魏苏淮), and Jun-Wei Luo(骆军委)
Citation: Chin. Phys. B, 2021, 30 (**8**): 087201. DOI: 10.1088/1674-1056/ac0133
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Different noncollinear magnetizations on two edges of zigzag graphene nanoribbons

Yang Xiao(肖杨), Qiaoli Ye(叶巧利), Jintao Liang(梁锦涛), Xiaohong Yan(颜晓红), and Ying Zhang(张影)
Chin. Phys. B, 2020, 29 (**12**): 127201. DOI: 10.1088/1674-1056/abc0e4

Effect of weak disorder in multi-Weyl semimetals

Zhen Ning(宁震), Bo Fu(付博), Qinwei Shi(石勤伟), Xiaoping Wang(王晓平)
Chin. Phys. B, 2020, 29 (**7**): 077202. DOI: 10.1088/1674-1056/ab9612

Semi-analytic study on the conductance of a lengthy armchair honeycomb nanoribbon including vacancies, defects, or impurities

Fateme Nadri, Mohammad Mardaani, Hassan Rabani
Chin. Phys. B, 2019, 28 (**1**): 017202. DOI: 10.1088/1674-1056/28/1/017202

Temperature dependence on the electrical and physical performance of InAs/AlSb heterojunction and high electron mobility transistors

Jing Zhang(张静), Hongliang Lv(吕红亮), Haiqiao Ni(倪海桥), Zhichuan Niu(牛智川), Yuming Zhang(张玉明)
Chin. Phys. B, 2018, 27 (**9**): 097201. DOI: 10.1088/1674-1056/27/9/097201

Investigation of the surface orientation influence on 10-nm double gate GaSb nMOSFETs

Shaoyan Di(邸绍岩), Lei Shen(沈磊), Zhiyuan Lun(伦志远), Pengying Chang(常鹏鹰), Kai Zhao(赵凯), Tiao Lu(卢朓), Gang Du(杜刚), Xiaoyan Liu(刘晓彦)
Chin. Phys. B, 2017, 26 (**4**): 047201. DOI: 10.1088/1674-1056/26/4/047201

Group velocity matters for accurate prediction of phonon-limited carrier mobility*

Qiao-Lin Yang(杨巧林)[1,2], Hui-Xiong Deng(邓惠雄)[1,2], Su-Huai Wei(魏苏淮)[3], and Jun-Wei Luo(骆军委)[1,2,4,†]

[1] State Key Laboratory of Superlattices and Microstructures, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China
[2] Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
[3] Beijing Computational Science Research Center, Beijing 100193, China
[4] Beijing Academy of Quantum Information Sciences, Beijing 100193, China

(Received 23 April 2021; revised manuscript received 12 May 2021; accepted manuscript online 14 May 2021)

First-principles approaches have recently been developed to replace the phenomenological modeling approaches with adjustable parameters for calculating carrier mobilities in semiconductors. However, in addition to the high computational cost, it is still a challenge to obtain accurate mobility for carriers with a complex band structure, e.g., hole mobility in common semiconductors. Here, we present a computationally efficient approach using isotropic and parabolic bands to approximate the anisotropy valence bands for evaluating group velocities in the first-principles calculations. This treatment greatly reduces the computational cost in two ways: relieves the requirement of an extremely dense $\boldsymbol{k}$ mesh to obtain a smooth change in group velocity, and reduces the 5-dimensional integral to 3-dimensional integral. Taking Si and SiC as two examples, we find that this simplified approach reproduces the full first-principles calculation for mobility. If we use experimental effective masses to evaluate the group velocity, we can obtain hole mobility in excellent agreement with experimental data over a wide temperature range. These findings shed light on how to improve the first-principles calculations towards predictive carrier mobility in high accuracy.

Keywords: electron–phonon interaction, phonon-limited hole mobility, Boltzmann transport equation

PACS: 72.10.–d, 72.10.Bg, 72.20.Fr **DOI:** 10.1088/1674-1056/ac0133

1. Introduction

Carrier mobility is one of the fundamental parameters of semiconducting materials as it characterizes how fast a carrier can move across the device when pulled by an electric field.[1,2] In general, carriers in a semiconductor can be scattered by phonons, impurities, lattice defects such as dislocations, neutral defects, surfaces and interfaces, and other carriers (e.g., scattering between electrons and holes). All these scattering processes limit the carrier mobility by reducing the mean scattering time τ, which is a phenomenological scattering time introduced to account for the scattering of the carriers by impurities, phonons, and so on. The scattering time usually depends on the crystal wave vector $\boldsymbol{k}$ since, in the absence of an external field, the carriers at thermal equilibrium are distributed in bands according to the Fermi–Dirac distribution function $f^0_{i\boldsymbol{k}} = \{\exp[(E_{i\boldsymbol{k}} - E_{\rm F})/k_{\rm B}T] + 1\}^{-1}$, where $E_{i\boldsymbol{k}}$ is the carrier energy, $E_{\rm F}$ Fermi energy, $k_{\rm B}$ Boltzmann constant, and T carrier temperature. The application of an external electric field $\boldsymbol{F}$ modifies the distribution of carriers to be described by a perturbed distribution function $f_{i\boldsymbol{k}}(\boldsymbol{r})$ with its variation following the semiclassical Boltzmann transport equation (BTE), in which the carriers are treated as particles regarding their mean free path should be much larger than the de Broglie wavelength and the correlation between the carriers is ignored.[3] The Boltzmann transport equation can be simplified by using the relaxation time approximation[3]

$$\left(\frac{\partial f_{i\boldsymbol{k}}}{\partial t}\right)\Big|_{\rm scatt} \approx \frac{f_{i\boldsymbol{k}} - f^0_{i\boldsymbol{k}}}{\tau_{i\boldsymbol{k}}}, \tag{1}$$

which is based on the assumption that for small changes in $f_{i\boldsymbol{k}}$ the net effect of the scattering processes is to cause carriers to return to equilibrium in a characteristic time $\tau_{i\boldsymbol{k}}$. From the linearization of the BTE, we can now obtain the current density due to carriers in band i,

$$\boldsymbol{j}_i = -\frac{e}{\Omega}\int \boldsymbol{v}_{i\boldsymbol{k}} f_{i\boldsymbol{k}} \frac{{\rm d}\boldsymbol{k}}{\Omega_{\rm BZ}}, \tag{2}$$

$$\boldsymbol{j}_i = \frac{e^2}{\Omega}\int \tau_{i\boldsymbol{k}}\boldsymbol{v}_{i\boldsymbol{k}}\left(\frac{\partial f^0_{i\boldsymbol{k}}}{\partial E_{i\boldsymbol{k}}}\right)(\boldsymbol{v}_{i\boldsymbol{k}}\cdot\boldsymbol{F})\frac{{\rm d}\boldsymbol{k}}{\Omega_{\rm BZ}}, \tag{3}$$

where Ω and $\Omega_{\rm BZ}$ are the volume of the crystalline unit cell and the first Brillouin zone, respectively, and $\boldsymbol{v}_{i\boldsymbol{k}} = \hbar^{-1}\partial E_{i\boldsymbol{k}}/\partial\boldsymbol{k}$ is the velocity of carriers with wave vector $\boldsymbol{k}$. From the microscopic form of Ohm's law, the current density is also related to the carrier mobility, $\boldsymbol{j} = e(n_{\rm e}\mu_{\rm e} + n_{\rm h}\mu_{\rm h})\boldsymbol{F}$, where $n_{\rm e}$ and $n_{\rm h}$ are the total densities of electrons and holes, and $\mu_{\rm e}$ and $\mu_{\rm h}$ are the averaged electron and hole mobilities, respectively. We can then define the mobility tensor for carriers in band i,

$$\mu_{i,\alpha\beta} = -\frac{e}{n_i\Omega}\int \frac{{\rm d}\boldsymbol{k}}{\Omega_{\rm BZ}}\left(\frac{\partial f^0_{i\boldsymbol{k}}}{\partial E_{i\boldsymbol{k}}}\right)\boldsymbol{v}_{i\boldsymbol{k},\alpha}\boldsymbol{v}_{i\boldsymbol{k},\beta}\tau_{i\boldsymbol{k}}, \tag{4}$$

*Project supported by the National Natural Science Foundation of China (Grant Nos. 11925407 and 61927901) and the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (Grant No. ZDBS-LY-JSC019).

†Corresponding author. E-mail: jwluo@semi.ac.cn

where n_i is the steady-state concentration of free carriers in band i. Finally, the overall electron mobility will be given by the sum over all conduction bands (CBs) or valleys, weighted by n_i, that is,

$$\mu_{\mathrm{e},\alpha\beta} = \frac{1}{n_{\mathrm{e}}} \sum_{i\in\mathrm{CB}} n_i \mu_{i,\alpha\beta}, \tag{5}$$

and the overall hole mobility is

$$\mu_{\mathrm{h},\alpha\beta} = \frac{1}{n_{\mathrm{h}}} \sum_{i\in\mathrm{VB}} n_i \mu_{i,\alpha\beta}, \tag{6}$$

In order to calculate carrier mobility we have to obtain the scattering time $\tau_{i\boldsymbol{k}}$, which is defined as the inverse of the total scattering rate at which carriers with a specific wave vector $\boldsymbol{k}$ in band i scatter to any other state,

$$\frac{1}{\tau_{i\boldsymbol{k}}} = \sum_{j\boldsymbol{k}'} S(i\boldsymbol{k}, j\boldsymbol{k}')[1 - f^0_{j\boldsymbol{k}'}]. \tag{7}$$

Here $S(i\boldsymbol{k}, j\boldsymbol{k}')$ is the probability per unit time that a carrier with wave vector $\boldsymbol{k}$ in band i will be scattered into another state $\boldsymbol{k}'$ in band j and can be calculated using Fermi's golden rule

$$S(i\boldsymbol{k}, j\boldsymbol{k}') = \frac{2\pi}{\hbar} \left|H_{ij}(\boldsymbol{k}, \boldsymbol{k}')\right|^2 \delta(E_{j\boldsymbol{k}'} - E_{i\boldsymbol{k}} - \Delta E), \tag{8}$$

where $H_{ij}(\boldsymbol{k}, \boldsymbol{k}') = \langle i\boldsymbol{k}|H_{\mathrm{scatt}}|j\boldsymbol{k}'\rangle$, H_{scatt} is the Hamiltonian for the scattering processes which conserve both energy and wave vector, and ΔE is the corresponding change in energy of carrier. Carriers in a semiconductor are scattered by their interaction with the phonons (both acoustic and optical phonons), ionized impurities, neutral defects, surfaces and interfaces, and other carriers (e.g., scattering between electrons and holes).[4] It was realized that that in common semiconductors, carriers are predominantly scattered by phonons in a large temperature range, apart from its other roles in the hot-electron thermalization process.[3] Therefore, in this work, we focus on phonon-limited carrier mobilities in semiconductors. We will assume that a carrier with initial energy $E_{i\boldsymbol{k}}$ and wave vector $\boldsymbol{k}$ in band i is scattered to another state with energy $E_{j\boldsymbol{k}+\boldsymbol{q}}$ and wave vector $\boldsymbol{k}+\boldsymbol{q}$ in band j via emission or absorption of a phonon with energy $\hbar\omega_{\boldsymbol{q}\nu}$ and wave vector $\boldsymbol{q}$ in phonon mode ν, the phonon-limited $\tau_{i\boldsymbol{k}}$ is then given by[5,6]

$$\begin{aligned}\frac{1}{\tau_{i\boldsymbol{k}}} = {} & \frac{2\pi}{\hbar} \sum_{j\nu} \int \frac{\mathrm{d}\boldsymbol{q}}{\Omega_{\mathrm{BZ}}} \left|H_{ij\nu}(\boldsymbol{k}, \boldsymbol{q})\right|^2 \\ & \times \Big[\left(1 + n_{\boldsymbol{q}\nu} - f^0_{j\boldsymbol{k}}\right) \delta\left(E_{i\boldsymbol{k}} - E_{j\boldsymbol{k}+\boldsymbol{q}} + \hbar\omega_{\boldsymbol{q}\nu}\right) \\ & + \left(n_{\boldsymbol{q}\nu} + f^0_{j\boldsymbol{k}}\right) \delta\left(E_{i\boldsymbol{k}} - E_{j\boldsymbol{k}+\boldsymbol{q}} - \hbar\omega_{\boldsymbol{q}\nu}\right)\Big],\end{aligned} \tag{9}$$

where $f^0_{i\boldsymbol{k}}$ and $n_{\nu\boldsymbol{q}}$ are the Fermi–Dirac and Bose–Einstein distribution functions for carriers and phonons, respectively, and the matrix elements $H_{ij\nu}(\boldsymbol{k}, \boldsymbol{q})$ are defined by[6]

$$\begin{aligned}\hat{H}_{\text{e-ph}} &= \sum_{j\boldsymbol{k}', i\boldsymbol{k}} \langle \Psi_{j\boldsymbol{k}'} \left|\hat{H}_{\text{e-ph}}\right| \Psi_{i\boldsymbol{k}}\rangle \hat{c}^\dagger_{j\boldsymbol{k}'} \hat{c}_{i\boldsymbol{k}} \\ &= N_p^{-1/2} \sum_{ij\nu} H_{ij\nu}(\boldsymbol{k}, \boldsymbol{q}) \hat{c}^\dagger_{j\boldsymbol{k}+\boldsymbol{q}} \hat{c}_{i\boldsymbol{k}} (\hat{a}^\dagger_{-\boldsymbol{q}\nu} + \hat{a}_{\boldsymbol{q}\nu}),\end{aligned} \tag{10}$$

where $\hat{H}_{\text{e-ph}}$ is the electron–phonon coupling Hamiltonian to the first order in the atomic displacement, $\hat{a}^\dagger_{-\boldsymbol{q}\nu}$ and $\hat{a}_{\boldsymbol{q}\nu}$ represent the creation and destruction operators of phonons with energy $\hbar\omega_{\boldsymbol{q}\nu}$, and $\hat{c}^\dagger_{j\boldsymbol{k}+\boldsymbol{q}}$ and $\hat{c}_{i\boldsymbol{k}}$ for the creation and destruction operators of carriers in states $\Psi_{j\boldsymbol{k}+\boldsymbol{q}}$ and $\Psi_{i\boldsymbol{k}}$, respectively. The detailed information about the electron–phonon matrix element $H_{ij\nu}(\boldsymbol{k}, \boldsymbol{q})$ is given in Ref. [6].

If we learn the electron–phonon matrix elements, we are ready to calculate the carrier mobility by inserting Eq. (9) into Eq. (5). However, the electron–phonon matrix elements were historically treated semi-empirically considering the fact that the potential seen by carriers is disturbed by the displacement of the atoms due to thermal vibrations. Such perturbation causes exchange of energy and momentum between carrier and phonons, which then impacts the carrier scattering time. Specifically, carriers can interact with phonons via deformation potentials since the energy variation of the band edges is related to the lattice strain caused by atomic displacement due to acoustic phonons[7] and nonpolar optical phonons.[8] The former is called acoustic (AC) deformation potential scattering with the interaction Hamiltonian[9,10]

$$H_{\mathrm{AC}} = \sum_{i,j} \frac{1}{2} \left(\frac{\partial u_j}{\partial x_i} \mathcal{D}^{ji} + \mathcal{D}^{ji} \frac{\partial u_j}{\partial x_i}\right), \tag{11}$$

where $\mathcal{D}^{ji}$ is the deformation potential of valence band (for holes) and of conduction band (for electrons) and $\partial u_j/\partial x_i$ is the strain[11] induced by the the lattice displacement, which is $\boldsymbol{u}(\boldsymbol{r}) = \sum_{\boldsymbol{q}} \sqrt{\frac{\hbar}{2MN\omega_{\boldsymbol{q}}}} e_{\boldsymbol{q}} [a_{\boldsymbol{q}} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} + a^\dagger_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}]$ if we consider an elastic wave. The latter is nonpolar optical (NPO) deformation potential scattering with the interaction Hamiltonian[10,12]

$$H_{\mathrm{NPO}} = \frac{1}{2} \sum_{xx'} u_{xx'} \mathcal{D}^{xx'}_{\mathrm{opt}}, \tag{12}$$

where $\mathcal{D}^{xx'}_{\mathrm{opt}}$ is the optical deformation potential. In polar semiconductors, carriers can also couple with optical phonons via the induced longitudinal electrical field[13] as the macroscopic polarization field generated by the out-of-phase vibration [longitudinal-optical (LO) phonon with phonon energy $\hbar\omega_{\mathrm{LO}}$] in electric dipoles tends to strongly couple with carriers, termed polar optical phonon (POP) scattering with the interaction Hamiltonian given by[11]

$$H_{\mathrm{POP}} = \mathrm{i} \frac{1}{\varepsilon_0} \left(\frac{e^2\hbar}{2V\gamma\omega_{\mathrm{LO}}}\right)^{1/2} \sum_{\boldsymbol{q}} \left(a_{\boldsymbol{q}} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} + a^+_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}\right). \tag{13}$$

An analogy of POP scattering to longitudinal-acoustic (LA) phonons is the piezoelectric (PZ) scattering[14] with interaction Hamiltonian given by[11]

$$H_{\mathrm{PZ}} = \sum_{\boldsymbol{q}} \frac{e e^*_{\mathrm{PZ}}}{\varepsilon^*} \sqrt{\frac{\hbar}{2MN\omega_{\boldsymbol{q}}}} \left(a_{\boldsymbol{q}} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} + a^+_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}\right). \tag{14}$$

Here e^*_{PZ} and ε^* are the effective piezoelectric constant and effective dielectric constant, respectively, M is the total mass in a unit cell, and N is the number of unit cells per unit volume.

The strength of carrier–phonon coupling of these four types of scattering mechanisms is characterized by a variety of deformation potentials responsible for strains induced by lattice vibrations in acoustic and optical phonons.[15,16] However, some of these deformation potentials are usually difficult to measure accurately and thus taken as adjustable parameters. It may yield a fortuitous agreement between the experimental values and simple model calculations as evidenced by the fact that experimentally measured hole mobility data can be fitted well both with and without polar-mode scattering for semiconductors.[16,17]

Such uncertainties seem to be avoided by a recently developed first-principles approach[6,18,19] to evaluate the electron–phonon matrix elements $H_{ijv}(\boldsymbol{k},\boldsymbol{q})$. However, it costs massive computational resources because extremely fine grids of $\boldsymbol{k}$ and $\boldsymbol{q}$ are required to predict reliable mobilities. Specifically, Ma *et al.*[20] found that the convergence of mobility values to within 0.5% requires $300\times300\times300$ $\boldsymbol{k}$ points and $150\times150\times150$ $\boldsymbol{q}$ points for the prediction of hole mobility in GaAs. Despite electron–phonon matrix elements on such dense Brillouin-zone grids are obtained by means of the Wannier interpolation scheme, which allows the determination of the e–ph coupling matrix fully from first-principles calculations with a much sparse Brillouin-zone grid, say, $6\times6\times6$ $\boldsymbol{k}$ and $6\times6\times6$ $\boldsymbol{q}$ uniform meshes,[20] the 3-dimensional integration of Eq. (5) is still very expensive. This high-dimensional integration can be reduced, if not avoided, by adopting a variety of approximations, see Ref. [21] for a recently brief review. Particularly, in order to derive analytical formulas for carrier mobility based on the phenomenological electron–phonon interaction models, one usually considers the simple case of a nondegenerate electron (or hole) gas in a parabolic band with an isotropic effective mass m^* in a cubic crystal.[4] The Fermi–Dirac distribution function for a non-degenerate electron gas can also be approximated by the Boltzmann distribution. The introduced error in this approximation is less than 5% for the high anisotropic valence bands,[15] which is much less than the errors arising from uncertainty in band structures. For instance, Poncé *et al.*,[5] showed that the predicted Si hole mobility can also vary from 502 $cm^2/V{\cdot}s$ to 820 $cm^2/V{\cdot}s$ depending on the levels of band structure theories used, illustrating that carrier mobility is very sensitive to the band structure.

In this work, taking hole mobility as an example, we propose a simplified method for the evaluation of mobilities (SMM) by combining first-principles density functional perturbation theory for the prediction of $H_{ijv}(\boldsymbol{k},\boldsymbol{q})$ and an assumption of a parabolic band with an isotropic effective mass m^* for each of three topmost valence bands. The latter leads the $\boldsymbol{k}$-dependent $\boldsymbol{v}_{i\boldsymbol{k}}$ to be a constant group velocity characterized by m^* in Eq. (5). This method overcomes the difficulties caused by both the aforementioned full first-principles approach and the semi-empirical approach. We indeed find that this SMM can still give a reasonable result for the phonon-limited scattering the same as first-principles method. However, different from the full first-principles method, once we have obtained the imaginary part of the carrier self-energy or relaxation time via the full first-principle method, we can exact the band-resolved and phonon-limited hole mobility via SMM. Therefore, the method cannot only reduce computation resources but also conduce to analyzing band-resolved phonon-limited hole mobility with a clear physical image, which will be further discussed in our future works.

2. Methodology

To simplify the evaluation of hole mobilities but keeping the accurate description of electron–phonon interaction the same as first-principles approach, the assumption of an isotropic and parabolic band with an isotropic effective mass is implemented to estimate the group velocity of holes.[4] In this approximation, the multiplication of group velocities in α and β directions, respectively, in Eq. (5) can be simplified to

$$v_{i\boldsymbol{k},\alpha}v_{i\boldsymbol{k},\beta}\approxeq\frac{2\varepsilon_{i\boldsymbol{k}}}{m_i^*}, \tag{15}$$

where $1/m_i^*=(1/m_i^{[100]}+1/m_i^{[110]}+1/m_i^{[111]})/3$ is the averaged effective mass for an anisotropic valence band i ($i=$ heavy-hole, light-hole, and spin–orbit split bands). Inserting Eq. (15) into Eq. (7), the mobility of holes in the valence band i can be rewritten as

$$\mu_i=-\frac{1}{n_i\Omega}\frac{2e}{m_i^*}\int\varepsilon_{i\boldsymbol{k}}\cdot\frac{\mathrm{d}\boldsymbol{k}}{\Omega_{\mathrm{BZ}}}\cdot\frac{\partial f^0_{i\boldsymbol{k}}}{\partial\varepsilon_{i\boldsymbol{k}}}\cdot\tau_{i\boldsymbol{k}}. \tag{16}$$

The advantage using the approximation of an isotropic group velocity to replace $\boldsymbol{v}_{i\boldsymbol{k}}$ is that we can reduce significantly the density of the $\boldsymbol{k}$-mesh to save a large fraction of computational cost. Because the group velocity $\boldsymbol{v}_{i\boldsymbol{k}}$ is extremely sensitive to the variation of $\boldsymbol{k}$, a very dense $\boldsymbol{k}$-mesh is usually required in order to have a smooth change in $\boldsymbol{v}_{i\boldsymbol{k}}$.[6,18,19] Thus, the integration over $\boldsymbol{k}$-space in Eq. (16) is extremely computational cost because of the required dense $\boldsymbol{k}$-mesh. Here, we only need a moderate dense $\boldsymbol{k}$-mesh for $\tau_{i\boldsymbol{k}}$ (details see below).

The integration over $\boldsymbol{k}$-space in the above equation [Eq. (16)] can now be replaced by an integration over the energy ε using the density of states (DOS) $\rho(\varepsilon)$,

$$\begin{aligned}\frac{1}{\Omega}\int\frac{\mathrm{d}\boldsymbol{k}}{\Omega_{\mathrm{BZ}}}&=\frac{1}{\Omega\Omega_{\mathrm{BZ}}}\iiint\mathrm{d}k_x\mathrm{d}k_y\mathrm{d}k_z\\&=\frac{1}{(2\pi)^3}\int\rho(\varepsilon)\mathrm{d}\varepsilon.\end{aligned} \tag{17}$$

We then rewrite Eq. (16) as

$$\mu_i = -\frac{1}{n_i(2\pi)^3}\frac{2e}{m_i^*}\int \frac{\partial f_i^0(\varepsilon_i)}{\partial \varepsilon_i}\tau_i(\varepsilon_i)\rho(\varepsilon_i)\varepsilon_i \mathrm{d}\varepsilon_i. \quad (18)$$

For a parabolic band in three dimensions, the DOS $\rho(\varepsilon)$ is given by (including spin degeneracy)[4]

$$\rho(\varepsilon) = \frac{1}{2\pi^2}\left(\frac{2m_i^*}{\hbar^2}\right)^{3/2}\varepsilon_i^{1/2}. \quad (19)$$

The above treatment allows us to keep the sum over band index j and phonon mode ν in Eq. (9) and we can further rewrite Eq.(16) as

$$\begin{aligned}\mu_i &= \sum_{j,\nu} -\frac{1}{n_i(2\pi)^3}\frac{2e}{m_i^*}\int \frac{\partial f_i^0(\varepsilon_i)}{\partial \varepsilon_i}\rho(\varepsilon_i)\tau_{i,j,\nu}(\varepsilon_i)\varepsilon_i \mathrm{d}\varepsilon_i \\ &= \frac{e}{m_i^*}\cdot\sum_{j,\nu}\frac{\langle\tau_{i,j,\nu}\varepsilon_i\rangle}{\langle\varepsilon_i\rangle}. \end{aligned} \quad (20)$$

Here, $\tau_{i,j,\nu}$ is the scattering time from band i to band j induced by phonon mode ν, which satisfies the Matthiessen's rule $\tau_i^{-1} = \sum_{j,\nu}(\tau_{i,j,\nu})^{-1}$; and $\langle\tau_{i,j,\nu}\varepsilon_i\rangle$ and $\langle\varepsilon_i\rangle$ stand for

$$\langle\varepsilon_i\tau_{i,j,\nu}\rangle = \int \frac{\partial f_i^0}{\partial \varepsilon_i}\rho(\varepsilon_i)\tau_{i,j,\nu}\varepsilon_i \mathrm{d}\varepsilon_i, \quad (21)$$

$$\langle\varepsilon_i\rangle = \int \frac{\partial f_i^0}{\partial \varepsilon_i}\rho(\varepsilon_i)\varepsilon_i \mathrm{d}\varepsilon_i. \quad (22)$$

This treatment enables us to assess the intraband and interband scatterings for holes. Inserting Eq. (20) into Eq. (7), we finally obtain the total hole mobility μ_h.

If having the valence band structure, electron–phonon coupling matrix elements, and phonon dispersion for a specific semiconductor, we are ready to predict its hole mobility. We obtain these quantities following Giustino and his co-workers[6,22] by carrying out first-principles calculations within density-functional theory (DFT) as implemented in the EPW code of the QUANTUM ESPRESSO[23] distribution, in conjunction with a generalized Wannier–Fourier interpolation technique.[24] To obtain the total energy and band structure for Si and SiC, we perform calculations using a 2-atom primary cell with a $12\times12\times12$ Monkhorst–Pack $\boldsymbol{k}$-point mesh for integration over the Brillouin zone, and a 400-eV cutoff for the plane-wave basis set. In order to include spin–orbit coupling (SOC), we adopt the optimized norm-conserving Vanderbilt (ONCV) fully relativistic pseudopotentials[25–27] generated by the ONCVPSP code. In the case of phonon-related calculations, we initially choose a $6\times6\times6$ coarse $\boldsymbol{q}$ and $\boldsymbol{k}$ meshes, which are interpolated to a dense grid with 4720 $\boldsymbol{k}$ points and 43288 $\boldsymbol{q}$ points[28] at the near region of Γ point via the Wannier interpolation. The Fermi energy level is set to 0.3 eV above the VBM to account for non-degenerate hole gas. SOC is always included in our calculations for both electronic structure or phonon dispersion since it has a remarkable impact on the prediction of the hole mobility as reported in different materials.[5,20,29] For example, Ponce *et al.*[5] found that the inclusion of SOC will increase hole mobility by 10%.

In our calculations, we take an experimental lattice constant of 5.43 Å for Si the same as the full first-principles calculations for mobility.[5] The DFT estimated fundamental bandgap of Si is 0.4351 eV, which is significantly smaller than the experimental value of 1.12 eV[30] due to the well known DFT error in bandgap. This underestimation in bandgap will artifically introduce scattering channels from valence bands to conduction bands for holes at high temperature. Therefore, we have to rigidly shift the whole conduction band up to give a 1.12 eV bandgap in order to get a reasonable scattering channel of holes. For SiC, we take a zinc-blend structure (3C-SiC) with a lattice constant of 5.645 Å.[31] The DFT estimated bandgap of 3C-SiC is 1.384 eV smaller than the experimental value of 2.417 eV. Because the DFT bandgap is large enough to suppress the artificial interband scattering between conduction and valence bands, we will not shift the conduction band in calculations of carrier mobility. It is also worth to note that the current EPW code[22] will not correctly describe the LO–TO splitting at $q=0$ for all polar materials because it does not include the non-analytic term. Therefore, the effect of phonon at $q=0$ is ignored in calculations considering its contribution to carrier–phonon scattering can be significantly reduced by increasing the number of $\boldsymbol{q}$ points in the $\boldsymbol{q}$-mesh.

3. Results

Here, we take Si and SiC as two examples to validate our proposed simplified approach for the full first-principles calculations of carrier mobility. Regarding Si and SiC are highly anisotropy in the valence bands, our simplified approach based on the assumption of isotropic and parabolic bands should give the worst results in these two semiconductors. Figure 1 shows our predicted Si hole mobility as a function of temperature in comparison with the best result obtained from the full first-principles calculations.[5] After considering effects of Brilouin-zone sampling and iterative solution of Boltzmann transport equation, DFT exchange and correlation, spin–orbit coupling, many-body quasiparticle corrections to the band structure, corrections to the DFT screening in the electron–phonon coupling matrix elements, and lattice thermal expansion, Ponce *et al.*[5] suggested that combination of several methods (named PBE EXP SOC+SCR+IBTE) should be the best computational setup for the prediction of carrier mobility. We see from Fig. 1 that our predicted hole mobility is in good agreement with that from the full first-principles approach above room temperature. Divergence becomes more visible as temperature is going down. To make such comparison, we intentionally take the effective masses of Si valence bands for calculating the group velocity from the corre-

sponding first-principles calculation rather than from the extensively cited experimental data.[4,32] The electron–phonon coupling matrix elements and band structure for energy- and momentum-conversation rules in Eq. (9) are obtained in an exact same way as in the full first-principle calculation. Therefore, the only difference between these two compared methods occurs in group velocities. The valence bands of semiconductors are complex because of the shapes of their constant energy surfaces referred to as warped spheres. The warping occurs along the [100] and [111] directions because of the cubic symmetry of the zinc-blende crystal. Hole transport is commonly believed to be substantially affected by the warping and anisotropy of the valence bands. Here, we demonstrate that we reproduce reasonably well the full first-principles hole mobility of Si by using $\boldsymbol{k}$-independent group velocities with an assumption of isotropic and parabolic valence bands although Si has the highest degree of anisotropy among common semiconductors as it has the largest ratio of C/B (Dresselhaus–Kip–Kittel (DKK)[33] valence band parameters).

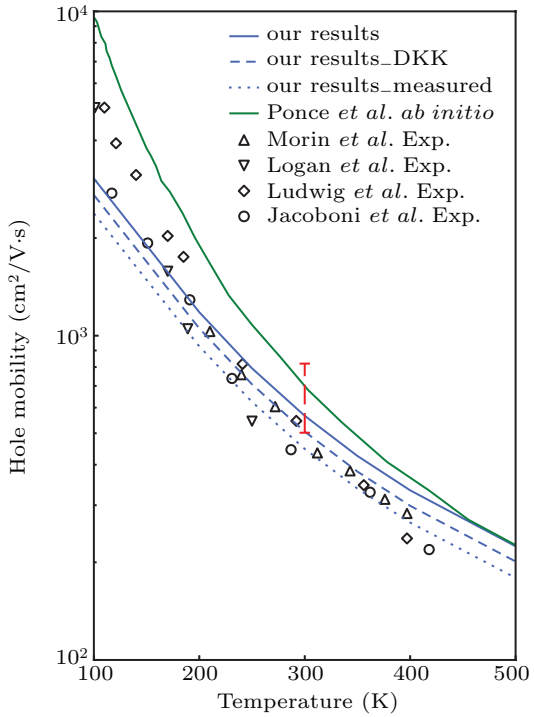

Fig. 1. Hole mobilities of Si as a function of temperature, calculated from our simplified approach with the effective mass from the first-principle calculation (blue solid line),[5] and from the DKK band parameters (blue dashed line)[34] and from the experimental measurement (blue dotted line),[4,32] compared with theoretical data from the full first-principle calculation (green solid line),[5] and the experimental data (open symbols). Experiments are from Ref. [35] (△), Ref. [36] (▽), Ref. [37] (◇), and Ref. [38] (◯).

Polar scattering arising from optical phonons (POP) and acoustic phonons (PZ) is absent in the nonpolar diamond structure Si. In order to examine the effect of these two scattering mechanisms on the robust of our simplified approach, we do a comparison for polar SiC as shown in Fig. 2. It exhibits that our simplified approach gives a very good agreement in an even wider temperature range from 200 K to 700 K, which has available experimental[39–41] and full first-principles data.[42]

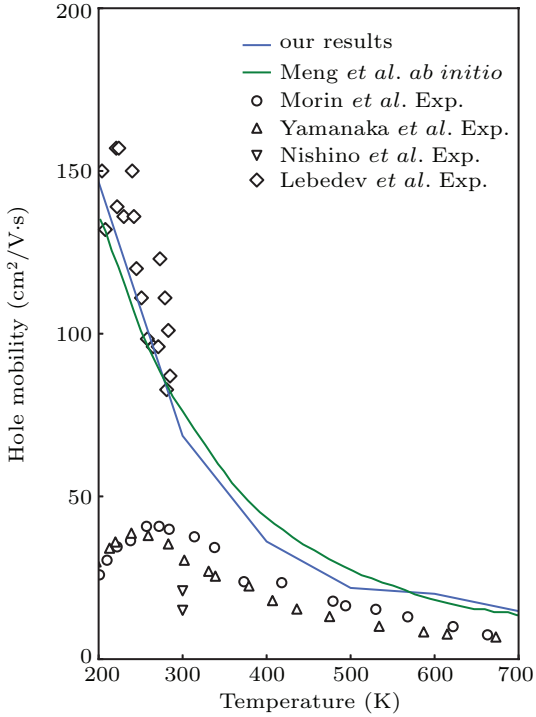

Fig. 2. Hole mobilities of 3C-SiC as a function of temperature, calculated from the the full first-principle Boltzmann formalism (blue solid line), compared with theoretical data from conventional relaxation time approximation[42] (green solid), and the experimental values from Refs. [39–41].

From these two comparisons for both nonpolar and polar semiconductors with the high degree of anisotropic valence bands among semiconductors, we have demonstrated that our simplified approach is reasonable and efficient alternative to the full first-principles calculations. The accuracy of the simplified approach should be even better in semiconductors with less anisotropy in bands.

4. Discussion

Figure 1 shows that the Si hole mobility predicted by the best tuned full first-principles calculations is still away from the experimental measured values over the large temperature range. Specifically, Ponce *et al.*[5] found that the Si hole mobility at room temperature can vary from 658 cm^2/V·s to 820 cm^2/V·s predicted by the full first-principles calculations considering the effects of exchange and correlation, spin–orbit coupling, many-body quasiparticle corrections, corrections to the DFT screening in the electron–phonon coupling matrix elements, and lattice thermal expansion. The best computational setup (considering the effect of many-body correlations on the band extreme within the GW quasiparticle approximation and correcting the DFT screening via the most accurate description of the screening) predicted a hole mobility of 658 cm^2/V·s, which is at least 30% higher than the experimental data (450–510 cm^2/V·s[30,34]). This overestimation on the hole mobility

was identified due to the overestimation of the band dispersion, which is characterized by a smaller effective mass, particularly for the heavy hole band along the [100] direction (see Table 1). The experimental heavy hole mass along [100] direction is $0.46m_e$,[4] whereas the first-principle predictions are usually smaller than $0.243m_e$.[5,20] Ponce *et al.*[5] found that, if they replace the valence band structure by a set of parabolic bands fitting to the experimental effective masses in their best first-principles calculations, they can further improve the hole mobility to $\mu_h = 502$ cm^2/V·s which is much close to the experimental values.

Carrier mobility depends on the band structure in two aspects, one is the density of final states in the scattering process, which is the described by energy- and momentum-conservation in Eq. (9), the other one is the group velocity as given in Eq. (5). In full first-principles calculations, it is difficult to separate the influences of these two aspects to hole mobility. Our simplified approach offers the chance to separate them since we can only consider the group velocity from experimental effective masses and keep the calculation of Eq. (9) using full first-principles (here, the band structure for energy- and momentum-conservation is taken from GW). In doing so, we adopt the experimental effective masses based on DKK valence band parameters,[34] and obtain a Si hole mobility $\mu_h = 505$ cm^2/V·s, which is very close to the experimental values. Figure 1 further shows that our improved hole mobility is actually in great agreement with experimental data over a wide range of temperatures from 100 K to 500 K. Since our improved data is so close to Ponce *et al.* obtained value of $\mu_h = 502$ cm^2/V·s, we can safely conclude that the change of the density of states of the final states in Eq. (9) induced by the overestimation of the band dispersion by the first-principles calculations has a negligible effect on the hole mobility.

In recent years, extensive works[20,22,29,42] have reported that very dense $\boldsymbol{k}$ and $\boldsymbol{q}$ meshes are required in order to calculate accurately carrier mobility. The purpose is to have the band structure via Wannier interpolation in an extremely fine $\boldsymbol{k}$-mesh, from which we can obtain the carrier group velocity smoothly. We have demonstrated that we can use a parabolic band to obtain the velocity group and sparse $\boldsymbol{k}$ and $\boldsymbol{q}$ meshes (with one-order of magnitude smaller in numbers of $\boldsymbol{k}$ and $\boldsymbol{q}$) to reproduce well the full first-principles results. We can learn that, in order to obtain an accurate carrier mobility, it is crucial to have correct group velocities reflecting the real band structure. It is not necessary to describe these group velocities on a very dense $\boldsymbol{k}$ mesh. On the other hand, it is well known that the first principles make it difficult to obtain the exact effective mass and group velocity, even with very complex GW many-body perturbation theory and modified Becke and Johnson (mBJ) method.[43] Kim[43] also argued that it is almost impossible to simultaneously obtain both band-gap and effective mass close to the experimental value, no matter how to adjust the ratio of functional. Therefore, it is impossible to obtain an accurate hole mobility by the full first-principles calculations unless one can correct the valence bands to have correct effective masses. The key to reduce the computational resources is to obtain the exact electronic structure on coarse $\boldsymbol{k}$ and $\boldsymbol{q}$ meshes. At present, in addition to fitting band structures from the measured effective masses, the atomistic semipseudopotential method (SEPM)[44] is also a good choice. Hence, if perfectly describing the effective mass of the hole band using the two approaches combined with the scattering rate Eq. (9), we can deduce that the approximate group velocity does not cause non-negligible uncertainty and the hole mobility can be evaluated accurately.

Table 1. Comparison between effective masses of silicon calculated using GW quasiparticle approximation with SOC[5] and experimental data from Refs. [4,32,34], in units of the electron mass. Expt. 1 means the effective mass calculated from the DKK band parameters,[34] and Expt. 2 represents the directly measured effective mass.[4,32] m_d^* and m_c^* denote the density-of-states and conductivity effective masses, respectively.

Band		GW_SOC	Expt. 1	Expt. 2
Heavy hole	[100]	0.2430	0.2747	0.4600
	[110]	0.5120	0.5795	0.5300
	[111]	0.6430	0.7381	0.5600
	m_d^*	0.4309	0.4898	0.5149
	m_c^*	0.3935	0.4464	0.5131
Light hole	[100]	0.2020	0.2041	0.1700
	[110]	0.1400	0.1468	0.1600
	[111]	0.1320	0.1392	0.1600
	m_d^*	0.1551	0.1609	0.1633
	m_c^*	0.1525	0.1587	0.1632
Split-off hole	[100]	0.2260	0.2342	0.2300
	[110]	0.2250	0.2342	0.2300
	[111]	0.2270	0.2342	0.2300
	m_d^*	0.2260	0.2342	0.2300
	m_c^*	0.2260	0.2342	0.2300
(A, B, $\|C\|$)		(–4.53, –0.418, 5.16)	(–4.27, –0.630, 4.93)	(–4.03, –1.854, 2.15)

5. Conclusion

In this article, we presented a simplified hole mobility computational approach based on the EPW software in the field of electron–phonon calculations. In comparison with the full first-principle and experimental values, two physically relevant examples showcasing the traits of the simplified approach were presented in detail.

Acknowledgment

We acknowledge fruitful discussions with Wu Li and Fanchen Meng.

References

[1] Di L J, Dai X Y, *et al.* 2018 *Acta Phys. Sin.* **67** 027101 (in Chinese)
[2] Deng H X, Luo J W and Wei S H 2018 *Chin. Phys. B* **27** 117104
[3] Lundstrom M 2000 *Fundamentals of carrier transport*, 2nd ed edn. (Cambridge: Cambridge University Press)
[4] Yu P Y and Cardona M 2010 *Fundamentals of semiconductors: physics and materials properties* 4th edn. (Berlin: Springer)
[5] Ponce S, Margine E R and Giustino F 2018 *Phys. Rev. B* **97** 121201
[6] Giustino F 2017 *Rev. Mod. Phys.* **89** 015003
[7] Shockley W and Bardeen J 1950 *Phys. Rev.* **77** 407
[8] Herring C and Vogt E 1956 *Phys. Rev.* **101** 944
[9] Tiersten M 1961 *IBM J. Res. Dev.* **5** 122
[10] Lawaetz P 1968 *Phys. Rev.* **174** 867
[11] Hamaguchi C 2010 *Basic semiconductor physics* 2nd ed edn. (Berlin: Springer-Verlag)
[12] Bir G L and Pikus G E 1974 *Symmetry and strain-induced effects in semiconductors* (New York: Wiley)
[13] Frohlich H 1954 *Adv. Phys.* **3** 325
[14] Meijer H and Polder D 1953 *Physica* **19** 255
[15] Kranzer D 1974 *Phys. Status Solidi A* **26** 11
[16] Wiley J D 1971 *Phys. Rev. B* **4** 2485
[17] Wiley J D and DiDomenico M 1970 *Phys. Rev. B* **2** 427
[18] Li W 2015 *Phys. Rev. B* **92** 075405
[19] Liu T H, Zhou J, Liao B, Singh D J and Chen G 2017 *Phys. Rev. B* **95** 075206
[20] Ma J, Nissimagoudar A and Li W 2018 *Phys. Rev. B* **97** 045201
[21] Deng T, *et al.* 2020 *npj Comput. Mater.* **6** 1
[22] Ponce S, Margine E R, Verdi C and Giustino F 2016 *Comput. Phys. Communs.* **209** 116
[23] Giannozzi P, *et al.* 2017 *J. Phys.: Condens. Matter* **29** 465901
[24] Mostofi A A, *et al.* 2014 *Comput. Phys. Commun.* **185** 2309
[25] Hamann D R 2013 *Phys. Rev. B* **88** 085117
[26] Schlipf M and Gygi F 2015 *Comput. Phys. Commun.* **196** 36
[27] Scherpelz P, Govoni M, Hamada I and Galli G 2016 *J. Chem. Theory Comput.* **12** 3523
[28] Ponce S, Jena D and Giustino F 2019 *Phys. Rev. B* **100** 085204
[29] Liu T H, *et al.* 2018 *Phys. Rev. B* **98** 081203
[30] https://www.azom.com/article.aspx?ArticleID=8346
[31] http://www.ioffe.ru/SVA/NSM/Semicond/SiC/hall.html
[32] Sze S M and Ng K K 2007 *Physics of semiconductor devices* 3rd edn. (Hoboken: Wiley-Interscience)
[33] Dresselhaus G, Kip A F and Kittel C 1955 *Phys. Rev.* **98** 368
[34] Bimberg D, *et al.* 1982 *Physics of Group IV Elements and III-V Compounds / Physik der Elemente der IV. Gruppe und der III-V Verbindungen, Condensed Matter* (Berlin: Springer-Verlag)
[35] Morin F J and Maita J P 1954 *Phys. Rev.* **96** 28
[36] Logan R A and Peters A J 1960 *J. Appl. Phys.* **31** 122
[37] Ludwig G W and Watters R L 1956 *Phys. Rev.* **101** 1699
[38] Jacoboni C, Canali C, Ottaviani G and Alberigi Quaranta A 1977 *Solid-State Electron.* **20** 77
[39] Yamanaka M, Daimon H, Sakuma E, Misawa S and Yoshida S 1987*J. Appl. Phys.* **61** 599
[40] Nishino S, Powell J A and Will H A 1983 *Appl. Phys. Lett.* **42** 460
[41] Lebedev A A, *et al.* 2008 *Semicond. Sci. Technol.***23** 075004
[42] Meng F, Ma J, He J and Li W 2019 *Phys. Rev. B* **99** 045201
[43] Kim Y S, Marsman M, Kresse G, Tran F and Blaha P 2010 *Phys. Rev. B* **82** 205212
[44] Wang L and Zunger A 1994 *J. Chem. Phys.* **100** 2394