Numerical study of the thermoelectric power factor in ultra-thin Si nanowires

Neophytos Neophytou · Hans Kosina

Abstract Low dimensional structures have demonstrated improved thermoelectric (TE) performance because of a drastic reduction in their thermal conductivity, κ_l. This has been observed for a variety of materials, even for traditionally poor thermoelectrics such as silicon. Other than the reduction in κ_l, further improvements in the TE figure of merit ZT could potentially originate from the thermoelectric power factor. In this work, we couple the ballistic (Landauer) and diffusive linearized Boltzmann electron transport theory to the atomistic $sp^3d^5s^*$-spin-orbit-coupled tight-binding (TB) electronic structure model. We calculate the room temperature electrical conductivity, Seebeck coefficient, and power factor of narrow 1D Si nanowires (NWs). We describe the numerical formulation of coupling TB to those transport formalisms, the approximations involved, and explain the differences in the conclusions obtained from each model. We investigate the effects of cross section size, transport orientation and confinement orientation, and the influence of the different scattering mechanisms. We show that such methodology can provide robust results for structures including thousands of atoms in the simulation domain and extending to length scales beyond 10 nm, and point towards insightful design directions using the length scale and geometry as a design degree of freedom. We find that the effect of low dimensionality on the thermoelectric power factor of Si NWs can be observed at diameters below $\sim$7 nm, and that quantum confinement and different transport orientations offer the possibility for power factor optimization.

N. Neophytou (✉) · H. Kosina
Institute for Microelectronics, TU Wien, Gußhausstraße
27-29/E360, 1040 Wien, Austria
e-mail: neophytou@iue.tuwien.ac.at

H. Kosina
e-mail: kosina@iue.tuwien.ac.at

Keywords Thermoelectrics · Tight-binding · Atomistic · $sp^3d^5s^*$ · Boltzmann transport · Seebeck coefficient · Thermoelectric power factor · Silicon · Nanowire · ZT
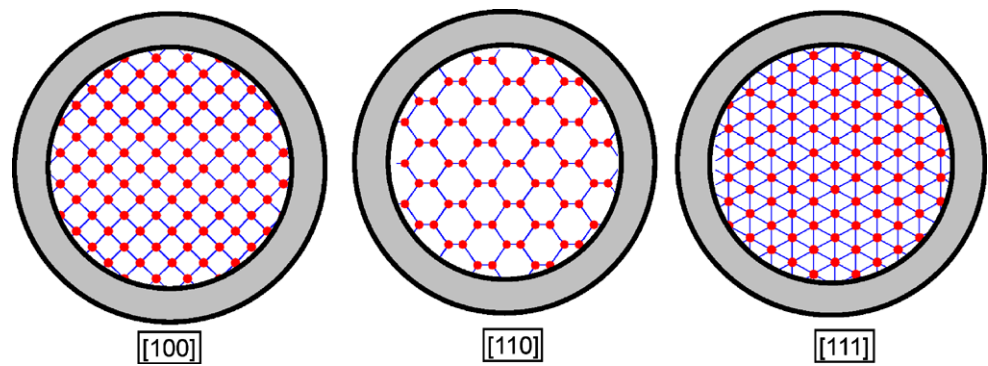
1 Introduction

The ability of a material to convert heat into electricity is measured by the dimensionless figure of merit $ZT = \sigma S^2 T/(\kappa_e + \kappa_l)$, where σ is the electrical conductivity, S is the Seebeck coefficient, and κ_e and κ_l are the electronic and lattice part of the thermal conductivity, respectively. The interrelation between σ, S, and κ_e in bulk materials keeps ZT low [1]. Some of the best thermoelectric materials are compounds of Bi, Te, Pb, Sb, Ag, and exhibit $ZT \sim 1$ [1, 2]. Recently, however, using low-dimensional structures, it was demonstrated that ZT could be greatly increased compared to their bulk counterparts, setting the stage for highly efficient TE energy conversion.

It was initially suggested that thermoelectric efficiency could be improved at the nanoscale because of two reasons: (i) Low-dimensionality and quantum size effects could improve the Seebeck coefficient [3], and (ii) Small feature sizes enhance phonon scattering on nanoscale interfaces and reduce thermal conductivity [4]. Indeed, large improvements of ZT in low-dimensional structures such as 0D quantum dots, 1D nanowires (NWs), 2D superlattices and bulk nanocomposites have recently been achieved [4–14]. This was even achieved for common materials, and importantly Si based systems such as Si, SiGe, and SiC [12–15]. Silicon, the most common semiconductor with the most advanced industrial processes, is a poor TE material with $ZT_{bulk} \sim 0.01$. Si NWs, on the other hand, have demonstrated $ZT \sim 1$, a $100\times$ increase [12–15], and they are now considered as

Fig. 1 Zincblende lattice of cylindrical nanowires in the [100], [110], and [111] orientations

emerging candidates for high efficiency and large volume production TE applications [16].

Most of the benefit to the measured *ZT* values of NWs originates from a dramatic reduction in the lattice thermal conductivity κ_l [15, 17–20]. It has very recently become evident, however, that benefits from κ_l reduction are reaching their limits, and further increases of *ZT* can only be achieved through improvements in the power factor σS^2 [16, 21]. By nanostructuring, the electronic structure could be engineered to tune the Seebeck coefficient [3, 7, 22] and the electrical conductivity [23] independently, which could maximize σS^2. For example, Hicks and Dresselhaus suggested that the sharp features in the low-dimensional density of states function *DOS(E)* can improve the Seebeck coefficient [3, 7]. Mahan and Sofo have further shown that thermoelectric energy conversion through a single energy level (0D channel) can reach the Carnot efficiency when κ_l is zero [24]. Because of the strong interconnection between σ and S, and their dependence on the geometrical features, involved simulation capabilities that account for the atomistic nature over large length scales are necessary in order to guide the design of such devices.

In this work the atomistic $sp^3d^5s^*$-spin-orbit-coupled ($sp^3d^5s^*$-SO) tight-binding model [25–29] is used to calculate the electronic structure of thin silicon NWs. Two transport formalisms are employed to calculate the thermoelectric coefficients σ, S, and the power factor σS^2: (i) The Landauer formalism [30–34], and linearized Boltzmann theory [23, 24, 35]. We describe the numerical methodologies and the approximations used, and demonstrate why such methodology is appropriate and efficient for this purpose. We consider different NW diameters, different transport orientations ([100], [110], [111]), different cross section geometries and various relevant scattering mechanisms. Using experimental values for κ_l in Si NWs, we estimate the *ZT* figure of merit. Our results explore effects of bandstructure features resulting from scaling the channel cross sections on the TE coefficients. Design optimization directions based on bandstructure engineering in low-dimensional channels are identified.

The paper is organized as follows: In Sect. 2 we describe the Landauer approach which is used to investigate the effect of the geometrical features on the electronic structures and the thermoelectric coefficients of ultra-scaled Si NWs. In Sect. 3 we describe the numerical approach to couple the TB model and Boltzmann transport theory, and the approximations used. In Sect. 4 we investigate the effects of NW cross section size, orientation, and scattering mechanisms on the thermoelectric coefficients. Finally, in Sect. 5 we conclude.

2 Ballistic Landauer approach for TE coefficients

The NW bandstructure is calculated using the 20 orbital atomistic tight-binding $sp^3d^5s^*$-SO model [25, 28], which is sufficiently accurate and inherently includes the effects of different transport and quantization orientations. We consider infinitely long, uniform, silicon NWs in the [100], [110] and [111] transport orientations as shown in Fig. 1, with also different cross section shapes. We assume passivated surfaces. The passivation technique details are provided in Appendix B [36]. These geometrical features have an impact on the electronic structure and the transport properties. Figure 2 shows examples of n-type NW electronic structures. The lowest subbands are shifted to the same origin $E = 0$ eV for comparison purposes. For brevity, only half of the k-space is shown. Figures 2a and b show the dispersions of [100] NW with diameters of 3 nm and 12 nm, respectively. As the diameter is reduced, the number of subbands is reduced, and the relative shift between the Γ and off-Γ valleys also changes. The degeneracies (η) of the Γ and off-Γ valleys are $\eta = 4$ and $\eta = 1$, respectively. Figures 2c and d show the corresponding dispersions for the [111] NW. The shape of the dispersion is different for different orientations. The degeneracy of this valley is $\eta = 6$. For [110] NWs, Fig. 2e shows the dispersion of the 3 nm wide and 12 nm tall rectangular NW (strong (1–10) surface quantization), whereas Fig. 2f the dispersion of the 12 nm wide and 3 nm tall NW (strong (001) surface quantization),
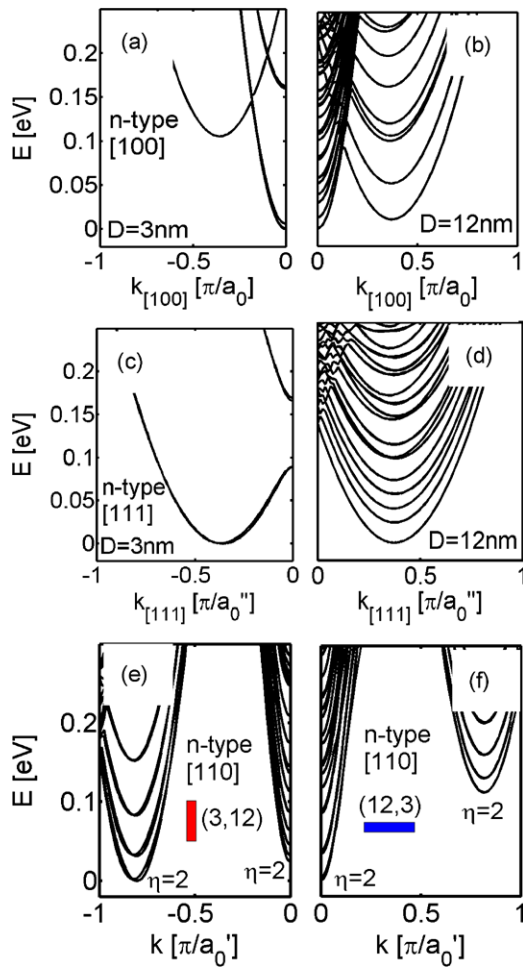
Fig. 2 Dispersions of n-type NWs in various orientations and diameters/side lengths. (**a**) [100], $D = 3$ nm. (**b**) [100], $D = 12$ nm. (**c**) [111], $D = 3$ nm. (**d**) [111], $D = 12$ nm. (**e–f**) Dispersions of rectangular NWs with widths (W) and heights (H): (**e**) [110] NW, $W = 3$ nm, $H = 12$ nm. (**f**) [110] NW, $W = 12$ nm, $H = 3$ nm. a_0, a_0' and a_0'' are the unit cell lengths for the wires in the [100], [110], and [111] orientations, respectively. The *filled rectangles* indicate the NW cross section

which produce different electronic structures. The degeneracy of both, the Γ and off-Γ valleys is $\eta = 2$. This dependence of the dispersions on geometry will result in different electronic and thermoelectric characteristics.

In this section, the ballistic Landauer formalism [30] is used to extract the TE coefficients. Although ballistic transport cannot be achieved in a realistic thermoelectric device, the results in this section indicate the upper performance limit, and it is a "fast" way to identify whether geometry could have an effect on TE properties through bandstructure engineering. The results from this method are compared to the results from the diffusive Boltzmann transport method in Sect. 3.

In the Landauer formalism the current is given by:

$$J = -\frac{q_0}{L} \sum_{k>0} v_k f_1 - \frac{q_0}{L} \sum_{k<0} v_k f_2 \qquad (1a)$$

$$= -\frac{q_0}{L} \sum_{k>0} v_k (f_1 - f_2), \qquad (1b)$$

where v_k is the bandstructure velocity, and f_1, f_2 are the Fermi functions of the left and right contacts, respectively. Auxiliary functions $R^{(\alpha)}(f_1, f_2, T)$ can be defined as:

$$R^{(\alpha)} = \frac{-\frac{q_0}{L} \sum_{k>0} v_k (f_1 - f_2) (E_k - \mu_1)^{\alpha}}{(\mu_1 - \mu_2)}, \qquad (2)$$

where μ_1, μ_2 are the contact Fermi levels, and E_k is the subband dispersion relation. This formula is the same as the one described in references [31, 37], where for small driving fields ΔV, the linearization $f_1 - f_2 = -q_0 \Delta V \frac{\partial f_1}{\partial E}$ is applied. Here, however, the computation is explicitly performed in k-space rather than energy-space. From these functions, the conductance G, the Seebeck coefficient S, and the electronic part of the thermal conductivity κ_e, can be derived as

$$G = R^{(0)}, \qquad (3a)$$

$$S = \frac{1}{T} \frac{R^{(1)}}{R^{(0)}}, \qquad (3b)$$

$$\kappa_e = \frac{1}{T} \left[R^{(2)} - \frac{[R^{(1)}]^2}{R^{(0)}} \right]. \qquad (3c)$$

Using this approach, the power factor (defined as $\sigma S^2 = G/\text{Area} \cdot S^2$) has been calculated. It is shown in Fig. 3 as a function of the one-dimensional carrier concentration, for cylindrical n-type NWs in the three transport orientations [100], [110] and [111] for two different diameters $D = 3$ nm and $D = 12$ nm. Comparing the magnitude of the power factor for $D = 3$ nm, the [111] NW with a 6-fold degenerate band has a higher power factor than the other NWs. The [100] NW, with a 4-fold degenerate Γ-valley follows, whereas the [110] NW with a 2-fold degenerate Γ-valley has the lowest power factor. Subbands with higher degeneracies, or subbands with edges very close in energy, improve the Seebeck coefficient which can be beneficial to the power factor. We show in Sect. 3, however, that once scattering is included in the calculation, the conductivity is degraded, which turns out to be a more dominant effect than the increase in Seebeck coefficient. For $D = 12$ nm in Fig. 3, the NW bandstructure becomes bulk-like, and any orientation effects that existed because of the bandstructure differences in lower diameters, are now smeared out. The interesting observation, however, is that under ballistic conditions, it seems that it is possible to improve the thermoelectric power factor by feature size scaling, in agreement with other theoretical ballistic transport studies, [31–34]. The magnitude of these benefits, however, is only within a factor of two.

To also emphasize the effect of the different confinement orientations, Fig. 4a shows the power factor for n-type [100] NWs, as a function of the carrier concentration for different

confinement conditions for *rectangular* NWs. Starting from the 12 nm × 12 nm "bulk-like" NW, we examine two cases: (i) we reduce the size of one of the sides, i.e. the height H is scaled to $H = 3$ nm, while the width W is kept at $W = 12$ nm in decrements of 1 nm (red lines). This rep-
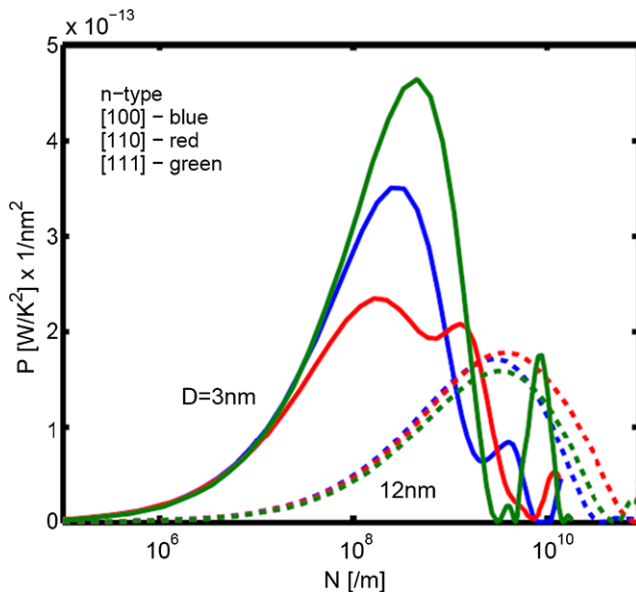

Fig. 3 (Color online) The thermoelectric power factor versus the 1D carrier concentration under ballistic transport conditions for n-type NWs of $D = 3$ nm and $D = 12$ nm in the [100] (*blue*), [110] (*red*), and [111] (*green*) orientations

resents the case of scaling from bulk towards "thin-body" devices. (ii) We scale both the width W and height H simultaneously down to $W = H = 3$ nm in decrements of 1 nm (blue lines). In both cases, decreasing the feature size of either side increases the peak of the power factor. The increase is larger when both sides are scaled, noted (3, 3). In this case, cross section scaling is beneficial for the power factor. Those benefits are ∼50%, and appear for side sizes below ∼7 nm (for sizes above that the power factor saturates).

Figure 4b, shows the same features for the n-type [110] NWs. The sides are [1–10] in the width and [001] in the height directions. Two device families are shown: (i) Devices with constant width along [1–10] at $W = 3$ nm, while the height along [001] varies from $H = 3$ nm to $H = 12$ nm (thin and tall NWs—red lines). (ii) Devices with the reverse aspect ratio, for which W varies from $W = 3$ nm to 12 nm, while H is fixed at $H = 3$ nm (wide and thin NWs—blue lines). The peaks of the power factors of the first device series (red lines) are higher than those of the second device series (blue lines). Interestingly, they are even higher than the peak of the fully scaled 3 nm × 3 nm NW, indicating that cross section scaling is not always beneficial, even for ballistic channels. The relative performance in these channels, as in the case of the ones described in Fig. 3, originate from the higher Seebeck coefficient, which is a consequence of the larger number of subbands/degeneracies in the electronic structure of this nanowire near the conduction band edge. The 3 nm × 12 nm NW has a higher performance than

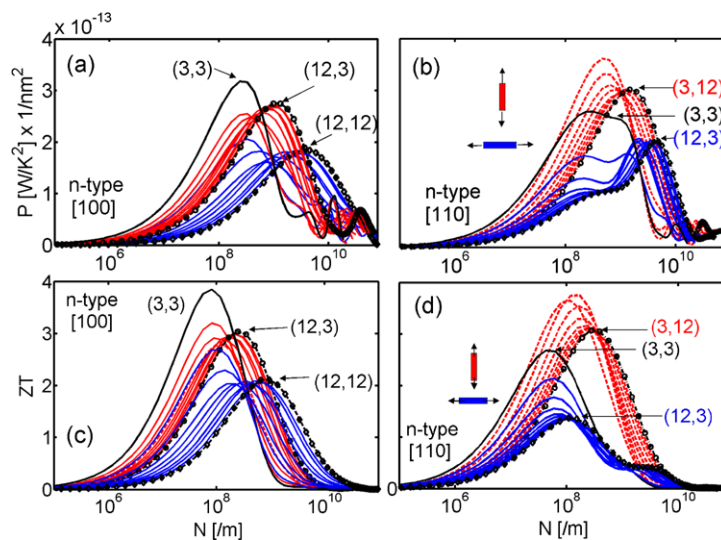**Fig. 4** (Color online) Thermoelectric features for n-type [100] (*left column*) and [110] (*right column*) NWs versus the 1D carrier concentration under ballistic transport conditions. (**a, c**) [100] NWs. Red lines: NWs with cross section sizes $W = 3$ nm to 12 nm, while $H = 3$ nm fixed (wide and thin NWs, approaching a thin-body). *Blue lines*: square NWs with cross section sizes $W = H = 3$ nm to $W = H = 12$ nm. Increments in sizes are of 1 nm. (**a**) Power factor σS^2. (**c**) ZT figure of merit. (**b, d**) [110] NWs. *Red lines*: NWs with cross section sizes

$W = 3$ nm fixed, and $H = 3$ nm to 12 nm (thin and tall NWs, approaching a thin-body device). *Blue lines*: NWs with cross section sizes $W = 3$ nm to 12 nm and $H = 3$ nm fixed (thin and wide NWs, approaching a thin-body). Increments in sizes are of 1 nm. (**b**) Power factor σS^2. (**d**) The ZT figure of merit. The *filled rectangles* indicate the NW cross section. $\kappa_l = 2$ W/mK is used for the thermal conductivity
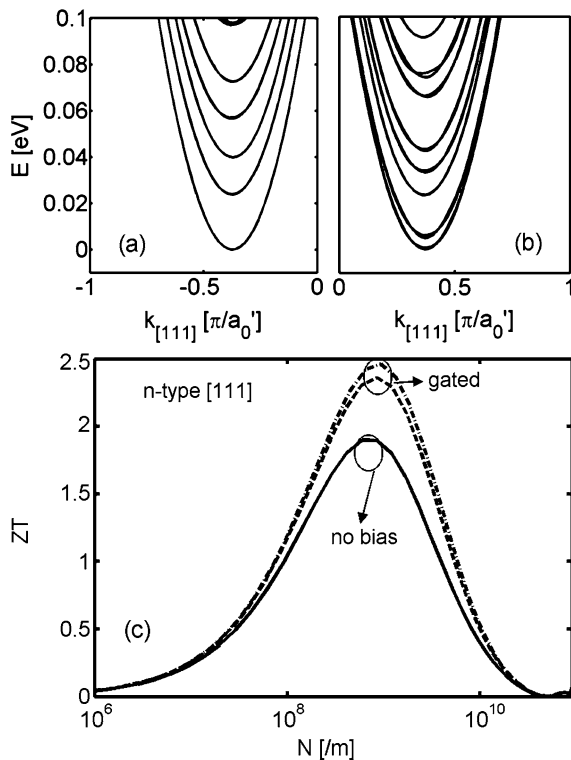
Fig. 5 The effect of gate field electrostatic potential on the electronic structure and the ballistic ZT figure of merit of the n-type [111] $D = 12$ nm NW. (**a**) The bandstructure for a flat potential profile. (**b**) The bandstructure under strong inversion, for $V_G - V_T = 1.0$ V. (**c**) The ZT figure of merit for the NW versus carrier concentration under no bias, and under large gate bias cases

the 12 nm × 3 nm because, as previously shown in Fig. 2e, the band edges of the Γ and off-Γ valleys are nearby in energy. For the 12 nm × 3 nm NW in Fig. 2f, the off-Γ valleys are higher in energy and do not participate in transport.

Figures 4c and d show the figure of merit ZT values of the devices in Figs. 4a and b, respectively, using a single value $\kappa_l = 2$ W/mK for the lattice part of the thermal conductivity, which was experimentally demonstrated for NWs [13, 15, 18]. ZT follows the shape and trends of the power factor. Interestingly, under ballistic assumptions, very high ZT values up to 4 can be achieved. We emphasize that such a low value for the thermal conductivity has experimentally only been achieved in rough or distorted NWs. We still use it, however, although for electrons we consider ballistic transport. Our intention here is to provide an idealized upper value for the ZT in Si NWs. As we will describe later on in Sect. 4, such values cannot be obtained once surface roughness scattering is incorporated. On the other hand, other methods for achieving very low thermal conductivity values have been theoretically proposed, which do not rely on surface roughness. Markussen et al., has proposed that Si nanowires, having surfaces decorated with molecules could also significantly reduce thermal conductivity, for which case our results are more relevant [38].

Another possibility to further improve thermoelectric performance is by adjusting the band positioning through gating. The gate electric field, similar to transistor devices, could shift the bands and change the thermoelectric properties. Figure 5 demonstrates this effect. Figure 5a shows the electronic structure of the $D = 12$ nm [111] n-type nanowire under flat potential in the cross section, whereas Fig. 5b under high gate inversion conditions. The separation of the bands has changed, and this results in an improvement of the thermoelectric ballistic ZT value by ~40%, which is a significant improvement. Careful design of the subband placement is, therefore, needed for improved performance. The nanostructure geometry enters the design through subband engineering. The tight-binding (TB) model is particularly suited for this, because the computational domain can be extended beyond 10 nm, and the effect of length scale can be properly investigated.

3 Linearized Boltzmann approach for TE coefficients

The ballistic Landauer approach emphasizes the effect of the Seebeck coefficient through subband positioning, whereas the conductivity of the channel is not affected by the otherwise enhanced scattering in ultra-narrow channels. In this section, we describe an approach to couple the TB model to linearized Boltzmann transport theory in order to investigate thermoelectric (TE) properties in 1D Si NWs in the diffusive transport regime. Several approximations are made in order to make the computation more robust, without affecting the essence of the conclusions. The entire procedure is described in detail in our previous works [23, 39]. Here, we only present the basic formalism, but we focus on the numerical and computational details of the method.

In Linearized Boltzmann theory, the TE coefficients are defined as:

$$\sigma = q_0^2 \int_{E_0}^{\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) \Xi(E), \tag{4a}$$

$$S = \frac{q_0 k_B}{\sigma} \int_{E_0}^{\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) \Xi(E) \left(\frac{E - E_F}{k_B T} \right), \tag{4b}$$

$$\kappa_0 = k_B^2 T \int_{E_0}^{\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) \Xi(E) \left(\frac{E - E_F}{k_B T} \right)^2, \tag{4c}$$

$$\kappa_e = \kappa_0 - T\sigma S^2. \tag{4d}$$

The transport distribution function $\Xi(E)$ is defined as [24, 35]:

$$\Xi(E) = \frac{1}{A} \sum_{k_x, n} v_n^2(k_x) \tau_n(k_x) \delta(E - E_n(k_x))$$

$$= \frac{1}{A} \sum_n v_n^2(E) \tau_n(E) g_{1D}^n(E). \tag{5}$$

where $v_n(E) = \frac{1}{\hbar}\frac{\partial E_n}{\partial k_x}$ is the bandstructure velocity, $\tau_n(k_x)$ is the momentum relaxation time for a carrier with wavenumber k_x in subband n, and

$$g_{1D}^n(E_n) = \frac{1}{2\pi\hbar}\frac{1}{|v_n(E)|} \quad (6)$$

is the density of states for 1D subbands (per spin). The transition rate $S_{n,m}(k_x, k_x')$ for a carrier in an initial state k_x in subband n to a final state k_x' in subband m is extracted from the atomistic dispersions using Fermi's Golden Rule [40]:

$$S_{n,m}(k_x, k_x') = \frac{2\pi}{\hbar}|H_{k_x',k_x}^{m,n}|^2\delta(E_m(k_x') - E_n(k_x) - \Delta E). \quad (7)$$

Usually, the momentum relaxation times are calculated by:

$$\frac{1}{\tau_n(k_x)} = \sum_{m,k_x'} S_{n,m}(k_x, k_x')\left(1 - \frac{|p_m(k_x')|}{|p_n(k_x)|}\cos\vartheta\right) \quad (8)$$

where in 1D the angle ϑ can take only two values $\vartheta = 0$ and $\vartheta = \pi$ [40, 41].

In this work, we calculate the relaxation times by:

$$\frac{1}{\tau_n(k_x)} = \sum_{m,k_x'} S_{n,m}(k_x, k_x')\left(1 - \frac{v_m(k_x')}{v_n(k_x)}\right). \quad (9)$$

Both are simplifications of the actual expression that involves an integral equation for τ_n [41–44]:

$$\frac{1}{\tau_n(k_x)} = \sum_{m,k_x'} S_{n,m}(k_x, k_x')$$
$$\times \left(1 - \frac{v_m(k_x')\tau_m(k_x')f_m(k_x')}{v_n(k_x)\tau_n(k_x)f_n(k_x)}\right). \quad (10)$$

While self-consistent solutions of this equation may be found, this is computationally very expensive, especially for atomistic calculations. Therefore, it is common practice in the literature to simplify the problem [45–48], and often sufficiently accurate results are obtained using the above approximations [42, 43]. For a parabolic dispersion, the use of (8) and (9) is equivalent. For a generalized dispersion, however, where the effective mass of the subbands is not well defined and the valleys appear in various places in the Brillouin zone, and the use of (9) is advantageous.

The matrix element between a carrier in an initial state k_x in subband n and a carrier in a final state k_x' in subband m is defined as:

$$H_{k_x',k_x}^{m,n} = \frac{1}{\Omega}\int_{-\infty}^{\infty}\int_R F_m(\vec{R})^* e^{-ik_x'x}U_S(\vec{r})F_n(\vec{R})$$
$$\times e^{ik_xx}d^2Rdx, \quad (11)$$

where the total wavefunction is decomposed into a plane wave e^{ik_xx} in the x-direction, and a bound state $F_v(\vec{R})$ in

the transverse, in-plane, with $\vec{R}$ being the in-plane vector. $U_S(\vec{r})$ is the scattering potential and Ω is the normalization volume. We note here that (11), and later on (14a) and (14b) and (20) involve integrals of the function $F_{m/n}$ over the NW in-plane R. However, $F_{m/n}$ is only sampled on the atomic sites. In the actual calculation the integrals over R are converted to summations over the atomic sites. The procedure is described in detail in Appendix A.

Elastic and inelastic scattering processes are taken into account. We consider bulk phonons and following the same rules when selecting the final states for scattering as in bulk Si. For n-type nanowires (NWs), the elastic processes due to elastic acoustic phonons, surface roughness (SRS), and impurity scattering are only treated as intra-valley processes, whereas inelastic processes due to inelastic phonons are only treated as inter-valley (IVS). An example of such transitions is shown in Fig. 6 for the $D = 3$ nm [110] NW. Although all valleys from the bulk Si electronic structure collapse from 3D to 1D k-space in our calculations, we carefully chose the final scattering states for each event by taking into account the degeneracies of the projected valleys for each orientation differently, as also indicated in Fig. 6. For inelastic transitions all six f- and g-type processes are included [40, 49]. For p-type NWs we consider ADP (acoustic deformation potential) and ODP (optical deformation potential) processes which can be intra-band and inter-band as well as intra-valley and inter-valley.

For the scattering rate calculation, we extend the usual approach for 3D and 2D thin-layer scattering commonly described in the literature [40], to 1D electronic structures. For phonon scattering, the relaxation rate of a carrier in a specific subband n as a function of energy is given by [23, 39]:

$$\frac{1}{\tau_{ph}^n(E)}$$
$$= \frac{\pi}{\hbar}\frac{(N_\omega + \frac{1}{2} \mp \frac{1}{2})}{\rho\hbar\omega_{ph}}\left(\frac{1}{L_x}\sum_{m,k_x'}\frac{|K_{\vec{q}}|^2}{A_{nm}^{k_xk_x'}}\delta_{k_x',k_x\pm q_x}\right.$$
$$\left.\times \delta(E_m(k_x') - E_n(k_x) \pm \hbar\omega_{ph})\left(1 - \frac{v_m(k_x')}{v_n(k_x)}\right)\right), \quad (12)$$

where $\hbar\omega_{ph}$ is the phonon energy, and we have used $\Omega = AL_x$. For optical deformation potential scattering (ODP for holes, IVS for electrons) it holds $|K_{\vec{q}}|^2 = D_O^2$, whereas for acoustic deformation potential scattering (ADP or IVS) it holds $|K_{\vec{q}}|^2 = q^2D_{ADP}^2$, where D_0 and D_{ADP} are the scattering deformation potential amplitudes. Specifically for elastic acoustic deformation potential scattering (ADP), after applying the equipartition approximation, the
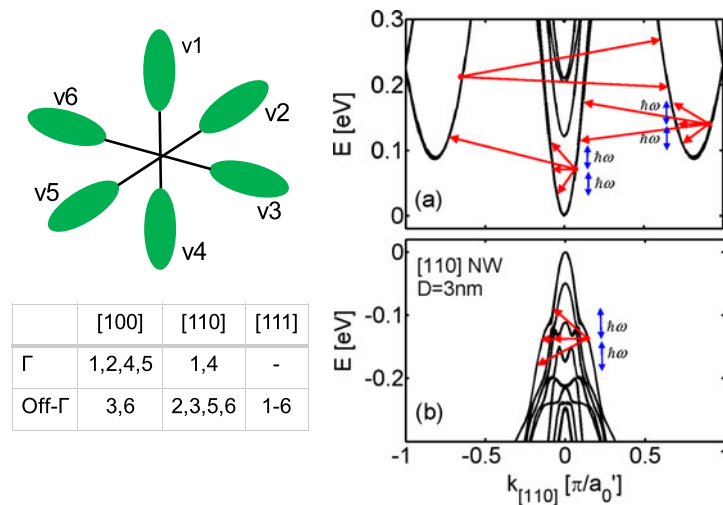
Fig. 6 Dispersions of the [110] NW of $D = 3$ nm with the scattering mechanisms indicated. (**a**) n-type NW. Intra-valley elastic and inter-valley inelastic (IVS) processes are considered (between the three valleys), following the bulk silicon scattering selection rules. For NWs in different orientations the Γ and off-Γ valley degeneracies, and the bulk relaxation rate becomes:

valleys from where they originate are shown in the table. Following the bulk scattering selection rules, however, each of the valleys is considered independently. (**b**) p-type NW. Elastic and inelastic processes are considered within the entire bandstructure. Intra- and inter-valley scattering is considered

relaxation rate becomes:

$$\frac{1}{\tau_{ADP}^n(E)} = \frac{2\pi}{\hbar} \frac{D_{ADP}^2 k_B T}{\rho v_s^2} \left(\frac{1}{L_x} \sum_{m,k_x'} \frac{1}{A_{k_x k_x'}^{nm}} \delta_{k_x', k_x \pm q_x} \right.$$

$$\left. \times \delta\left(E_m(k_x') - E_n(k_x)\right)\left(1 - \frac{v_m(k_x')}{v_n(k_x)}\right) \right), \quad (13)$$

where v_s is the sound velocity in Si.

In the expressions above, the quantities in the right-hand-side are all k-resolved when computed from the electronic structure $E(k)$, whereas the scattering rate in the left-hand-side is a function of energy. The δ-function in (12) and (13) states energy conservation. Numerically, the $E(k)$ relation needs to be discretized in energy. All states are sorted in energy and at a particular energy, arrays with all relevant k-states from all subbands are constructed.

One of the computationally most demanding steps in terms of memory requirements is the calculation of $A_{k_x k_x'}^{nm}$, the wavefunction overlap between the final and initial states. The calculation of this quantity involves an integral of the form:

$$\int_R \rho_{k_x', k_x}^{m,n}(\vec{R})^2 d^2 R, \quad (14a)$$

$$\rho_{k_x', k_x}^{m,n}(\vec{R}) = F_{m,k_x'}(\vec{R})^* F_{n,k_x}(\vec{R}). \quad (14b)$$

For the larger NWs, this calculation of the matrix elements imposes a huge computational burden. All wavefunctions of every k-state for every subband need to be stored because it is not known *a priori* for each initial state which are the corresponding final scattering k-states and at which

subbands when calculating the electronic structure, as indicated in the scattering examples of Fig. 6. The $D = 12$ nm structures that could include 5500 atoms each described by 20 orbitals, and a typical k-space grid of 200 points and considering 100 subbands, require several tens of Gbytes for the storage of the wavefunctions alone. For computational efficiency, therefore, we use the following scheme: on each atom we add the probability density of the components of each multi-orbital wavefunction, and afterwards perform the final/initial state overlap multiplication. In such way, we approximate the form factor components of a lattice atom at a specific location R_0 by:

$$\left| \rho_{k_x', k_x}^{m,n} \right|^2 = \sum_\alpha F_{n,k_x}^\alpha F_{m,k_x'}^{\alpha}{}^* \sum_\beta F_{n,k_x}^\beta F_{m,k_x'}^{\beta}{}^*$$

$$\approx \sum_{\alpha,\beta} F_{n,k_x}^\alpha F_{n,k_x}^{\beta}{}^* \delta_{\alpha,\beta} \sum_{\alpha,\beta} F_{m,k_x'}^\alpha F_{m,k_x'}^{\beta}{}^* \delta_{\alpha,\beta}$$

$$= \sum_\alpha \left| F_{n,k_x}^\alpha \right|^2 \sum_\beta \left| F_{m,k_x'}^\beta \right|^2$$

$$\equiv \left\langle F_{n,k_x} \right\rangle^2 \left\langle F_{m,k_x'} \right\rangle^2 \quad (15)$$

where α, β run over the tight-binding orbitals of a specific atom. With this the overlaps are computed using the probability density of each state, as in a single orbital (i.e. effective mass) model, although we still keep the k_x-dependence of the wavefunctions. The approximation in (15) is important because it reduces the memory needed in the computation by $20\times$, allowing simulations of large NW cross sections with only minimal reduction in accuracy. Indeed, our numerical
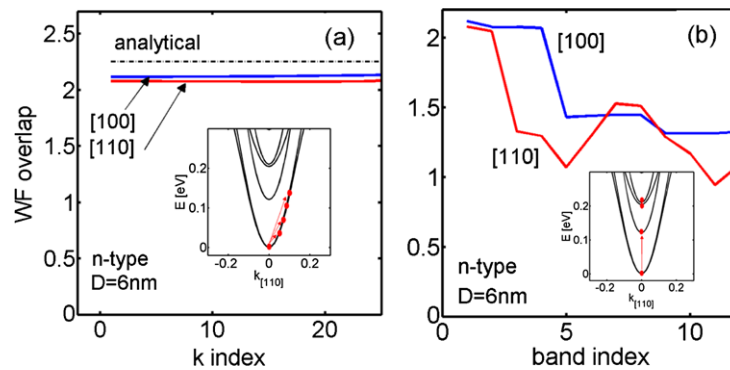
Fig. 7 The wavefunction overlap integral between a state at $k = 0$ in the first subband with (**a**) states of different k in the first subband, and (**b**) with states at $k = 0$ but different subbands in units of $1/A$, where A is the area of the NW. Results for n-type [100] and [110] NWs of $D = 6$ nm are shown. The analytical value for the integral is

9/4 for intra-band transitions, and 1 for inter-band transitions. *Insets*: Schematics indicating the initial and final $E(k)$ states (the dispersions are of the $D = 3$ nm, [110] n-type NW, for which the transitions can be more easily visualized)

overlaps agree with the analytical expressions for the wave-function overlaps if one assumes sine or cosine wavefunctions and parabolic bands, which can be derived to be $9/4A$ for intra-band and $1/A$ for inter-band transitions, where A is the cross section area of the NW [31, 40]. This is clearly indicated in Fig. 7, where we show the wavefunction overlap for the n-type [100] and [110] NWs with $D = 6$ nm between the state $k = 0$, in subband $n = 1$, and several final states in units of $1/A$. In Fig. 7a, the intra-band transitions are shown with final states in subband $m = 1$, and varying k-values. The wavefunction overlaps are indeed very close to the analytical value of $9/4$. In Fig. 7b, the inter-band transitions are shown with final states in subbands $m = 1, 2, \ldots, 12$ and $k = 0$. The first point, for $m = 1$ is the intraband transition which gives $\sim 9/4$, whereas for higher bands the overlaps reduce to lower values around ~ 1. The values are very close to the analytical ones, do not have significant k-dependence, and should not affect the qualitative nature of the results significantly. The price to pay, however, is that with this simplification the phase information for the wavefunctions is lost, and the selection rules are incorporated into the scattering rate calculation "by hand". However, this treatment is consistent with that for scattering in bulk and ultra-thin-layer structures reported in the literature. Still, even after this simplification, the storage of the probability density for the larger diameter NWs still requires several Giga bytes of memory.

For surface roughness (SR), we assume a 1D exponential autocorrelation function [50] for the roughness given by:

$$\langle \delta(\rho) \delta(\rho' - \rho) \rangle = \Delta_{rms}^2 e^{-\sqrt{2}|\rho|/L_C} \qquad (16)$$

with $\Delta_{rms} = 0.48$ nm and $L_C = 1.3$ nm [48]. We derive the surface roughness matrix element assuming that SR only causes a band edge shift. The scattering strength is given by the shift in the subband edges with diameter scaling

$\Delta E_{C,V}/\Delta D$ [51, 52]. The transition rate is derived as:

$$S_{n,m}^{SRS}(k_x, k_x') = \frac{2\pi}{\hbar} \left(\frac{q_0 \Delta E_{C,V}}{\Delta D} \right)^2 \left(\frac{2\sqrt{2} \Delta_{rms}^2 L_C}{2 + q_x^2 L_C^2} \right)$$
$$\times \delta \left(E_m(k_x') - E_n(k_x) \right), \qquad (17)$$

where $q_x = k_x - k_x'$. As described by various authors, the band edge variation is the cause of the major impact of SRS in ultra-scaled channels [48, 51–54]. In Refs. [48, 52] it was shown that the SRS limited low-field mobility in ultra-thin nanostructures follows a L^6 behavior, where L is the confinement length scale, originating from this subband shift due to the variation of L. This SRS model is a simplified one, compared to the ones described in Refs. [48, 55–57] that account for additional Coulomb effects, the wavefunction deformation at the interface, and the position of electrons in the channel. These effects are ignored here since they only cause quantitative changes in our results, whereas our focus is on qualitative trends that originate from geometry-induced electronic structure variations.

Figures 8a and b show the shift in the band edges $\Delta E / \Delta D$ as a function of diameter for the conduction and valence subbands, respectively. Indeed, the trends follow a D^{-3} power law both for electrons and holes as expected, with some minor deviations. For the n-type, the lowest valleys have slightly lower band edge shifts compared to the upper valleys. In the calculation of the SRS, the Γ and off-Γ valleys are taken separately into account when calculating the scattering rate. The orientation dependence is more evident in the case of p-type NWs. The band edge shifts are larger for the [100] NWs, whereas the band edges of the [111] NW are affected the least by diameter variations. The sensitivity of the band edges can be directly correlated with a confinement effective mass m_C^*. Using the simple notion of a particle in a box where the ground state energy is
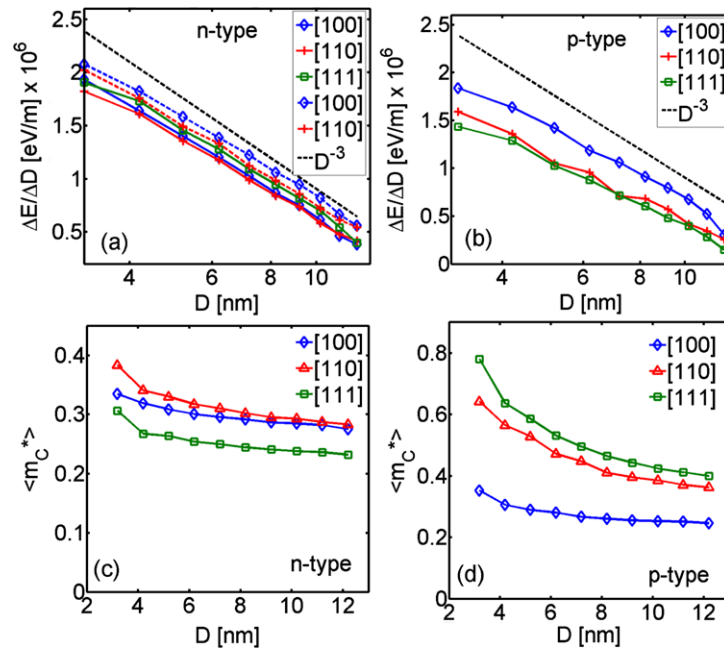
Fig. 8 (Color online) Change in band edges as a function of diameter. Results for NWs in the [100] (*diamond—blue*), [110] (*cross—red*), and [111] (*square—green*) transport orientations are shown. (**a**) Conduction band. Results for the lower valleys (*solid*), and upper valleys (*dashed*) are shown. (**b**) Valence band. The *dashed-black line* indicates the D^{-3} law. (**c**–**d**) The average confinement effective mass for NWs in different orientations versus the diameter. This is calculated from the change in the subband edges with confinement using the particle in a box quantization picture. Results for [100] (*diamond—blue*), [110] (*triangle—red*) and [111] (*square—green*) transport orientated NWs are shown. (**c**) n-type NWs. (**d**) p-type NWs

$E = \pi^2 \hbar^2 / 2 m_C^* D^2$, approximate values for m_C^* can be extracted. These are shown in Fig. 8c and d. For n-type NWs, the [110] orientation shows the largest m_C^*, whereas for p-type NWs the [110] and [111] NWs have the largest m_C^*. The slight deviation in the band edges from the D^{-3} law at smaller diameters, which reduces the rate of increase in the scattering matrix element, are also reflected as an increase in the confinement effective mass. The value of m_C^* of the n-type NWs lies between the longitudinal and transverse bulk Si masses of $m_l = 0.9 m_0$ and $m_t = 0.19 m_0$. For p-type NWs, the m_C^* values for the larger NW diameters are close to the bulk Si heavy-hole mass $m_{hh} = 0.4 m_0$. For the [100] orientation they remain in that region for the smaller diameters as well. For the [110] and [111] orientations on the other hand, m_C^* increases as the diameter is reduced. This is an important observation that indicates that the p-type [111] and [110] NWs will be less sensitive to surface roughness scattering (SRS). For thermoelectric materials this can be especially important since SRS is needed for the reduction in thermal conductivity κ_l. The fact that an intrinsic band-structure mechanism makes the conductivity more tolerant to SRS could help in power factor optimization in such channels in which rough boundaries are favored.

For ionized impurity scattering the scattering potential is approximated by:

$$U_S(\vec{r}) = \frac{q_0^2}{4\pi \kappa_s \varepsilon_0} \frac{e^{-\sqrt{(\vec{R}-\vec{R}')^2+x^2}/L_D}}{\sqrt{(\vec{R}-\vec{R}')^2+x^2}}, \tag{18}$$

where $\vec{R}$ is the position of an electron in the 2D cross section at $x = 0$, influenced by an impurity at $(x, \vec{R}')$, and the x direction is assumed to extend to infinity. The 3D screening length L_D is given by:

$$L_D = \sqrt{\frac{\kappa_s \varepsilon_0 k_B T}{q_0^2 n} \frac{\Im_{-1/2}(\eta_F)}{\Im_{-3/2}(\eta_F)}}. \tag{19}$$

where $\Im_\alpha(\eta_F)$ is the Fermi-Dirac integral of order α, and n is the carrier concentration. The matrix element for electron-impurity scattering then becomes:

$$H_{k_x',k_x}^{m,n}(\vec{R}') = \int_R \frac{\langle F_{n,k_x}(\vec{R})\rangle}{\sqrt{A}}$$
$$\times \left(\frac{1}{L_x} \int_{-\infty}^{\infty} \frac{q_0^2 e^{iq_x x}}{4\pi \kappa_s \varepsilon_0} \frac{e^{-\sqrt{(\vec{R}-\vec{R}')^2+x^2}/L_D}}{\sqrt{(\vec{R}-\vec{R}')^2+x^2}} dx \right)$$
$$\times \frac{\langle F_{m,k_x'}(\vec{R})\rangle}{\sqrt{A}} d^2 R \tag{20}$$

Fig. 9 (Color online)
Low-field, phonon-limited
transport characteristics for
NWs in the [100] (*blue*), [110]
(*red*) and [111] (*green*)
orientations. (**a**) The transport
distribution function (TD) of the
$D = 3$ nm n-type NWs. (**b**) The
mobility of $D = 3$ nm n-type
NWs. (**c**) The transport
distribution function (TD) of
$D = 12$ nm p-type NWs.
(**d**) The mobility of $D = 12$ nm
p-type NWs
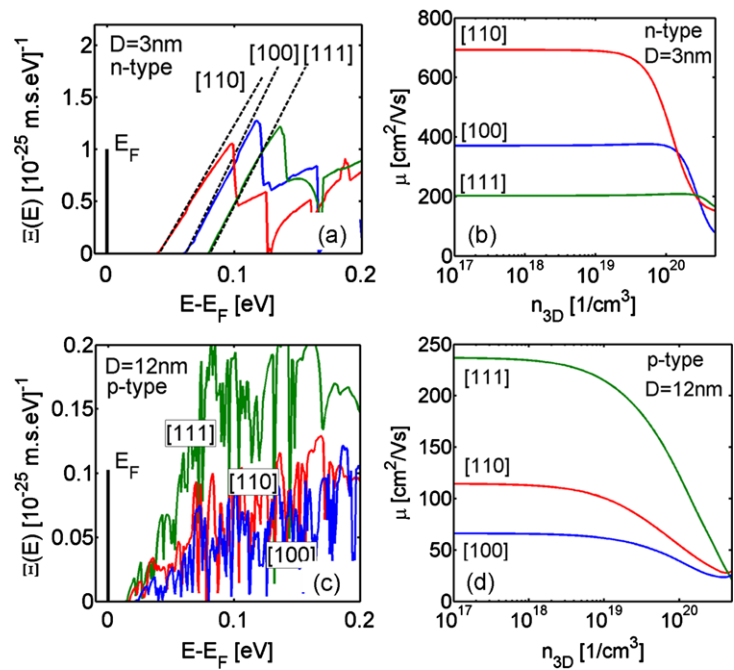
where the expression in the brackets is the Green's function of the infinitely long channel device. For a cylindrical channel, the expression in the parenthesis is the modified Bessel function of second kind of order zero, $K_0(q, \vec{R}')$ [48, 58–60].

The total transition rate due to impurity scattering is computed after taking the square of the matrix element, multiplying by $N_I L_x$, the number of impurities in the normalized cross sectional area of the NW in the length of the unit cell, and integrating over the distribution of impurities in the cross sectional area (over $\vec{R}'$). The impurities are assumed to be distributed uniformly in the volume considered.

The transport distribution function (TD) in (5) turns out to be a very convenient means to understand the effect of the electronic structure on the thermoelectric (TE) coefficients. Figure 9a shows the phonon-limited TDs for n-type NWs of $D = 3$ nm. The TDs for the three different orientations [100], [110] and [111] are shown. There are two observations that determine the performance of the NWs. (i) The low energy linear region, where only one subband participates in transport, with slope proportional to η/m^*, where η is the degeneracy of the subband [39]. (ii) The separation of the TD from the Fermi level, η_F. The closer the TD is to the Fermi level for a particular carrier concentration and the higher its slope, the larger the conductivity and mobility will be. This is shown in Fig. 9b, where the orientation dependence of the mobility correlates with the order the TDs appear with respect to the Fermi level. It is important to note that since the transport and density-of-states (DOS) effective masses (m^*) are the same for NWs, a reduction in m^* will not only reduce η_F in order to keep the carrier concentra-

tion fixed, but it will also increase the TD slope, finally having a doubly positive impact on the conductivity [39, 61]. Figure 9c shows a different situation, regarding the TDs for p-type NWs of $D = 12$ nm in the three orientations. As the NW diameter increases, the DOS of the NWs approaches the bulk DOS, and η_F is the same for all NWs. Their slope, however, is different, which is reflected in the large anisotropy in the mobility in Fig. 9d. Note that there are more subbands which result in more peaks in the TDs of the larger NWs compared to the narrower ones.

As mentioned previously, one of the approximations used is that of *bulk* phonons. Bulk phonons provide an ease of modeling and allow the understanding of the bandstructure effects on the TE coefficients, still with good qualitative accuracy in the results. Confined phonons in NWs can have very different dispersions and properties than bulk. However, the effect of phonon confinement for the thinnest NWs examined in this work is not that strong; it can be of the order of 10–20% (reduction in conductivity), and declines fast as the diameter increases [45–47, 62]. In the literature it is common to employ higher than bulk deformation potential values to account for phonon confinement [48, 63–65]. Here we use deformation potential parameters $D_{ODP}^{holes} = 13.24 \times 10^{10}$ eV/m, $D_{ADP}^{holes} = 5.34$ eV, and $D_{ADP}^{electrons} = 9.5$ eV from Refs. [45, 46, 61] which are more suitable for NWs. All other electron-phonon coupling parameters are the bulk values taken from [40]. The qualitative behavior of our results mostly depends on the shape of the bandstructure and not on the strength of the phonon scattering mechanisms. For a more quantitative description of the results, phonon confinement has to be accounted for.
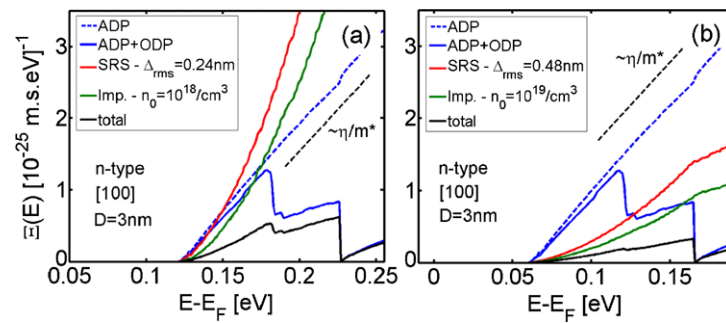
Fig. 10 (Color online) The transport distribution function (TD) $\Xi(E)$ for n-type [100] NW of $D = 3$ nm under different scattering conditions (i) ADP phonon-limited (*blue dashed*), (ii) ADP-IVS phonon-limited (*blue*), (iii) SRS limited (*red*), and (iv) impurity limited (*green*).

(**a**) Carrier concentration n $= 10^{18}/cm^3$, $\Delta_{rms} = 0.24$ nm. (**b**) Carrier concentration n $= 10^{19}/cm^3$, $\Delta_{rms} = 0.48$ nm. The impurity concentration n_0 is equal to the carrier concentration in each case

However, even in that case, phonon scattering is not the major scattering mechanism in NW devices that are suited for TE applications. This is clearly demonstrated in Fig. 10, again using the TD features for ADP-limited (blue-dotted), ADP-ODP-limited (blue-solid), SRS-limited (red), impurity scattering-limited (green), and the TD including all scattering mechanisms (black). In Fig. 10a moderate values for surface roughness ($\Delta_{rms} = 0.24$ nm) and impurity concentration ($n_0 = 10^{18}/cm^3$) are used. Strong scattering will lower the TD value and degrade conductivity. From the important low energy region, we observe that both SRS and impurity scattering mechanisms are stronger than phonon scattering. Figure 10b shows the same features, but with $\Delta_{rms} = 0.48$ nm and $n_0 = 10^{19}/cm^3$, which are more relevant for high performance thermoelectric devices (around the peak of the power factor as it will be shown in Sect. 4). SRS and impurity scattering are much stronger than phonon scattering. A calculation of the phonon contribution to the total scattering rate shows that it is only 12% and 6% in the situations of Fig. 10a and b, respectively, even with the larger than bulk deformation potentials values [23]. The strongest mechanism is impurity scattering, which dominates the scattering processes at such high concentrations. Indeed, this is in agreement with impurity scattering in bulk Si which reduces the mobility by almost an order of magnitude from the phonon-limited value at such high concentrations [66]. This shows that the details of phonon scattering strength for NW devices might not be of great importance to the total channel conductivity. This also demonstrates the importance of modulation doping in achieving high thermoelectric performance.

4 Si nanowire thermoelectric coefficients

Geometrical features such as diameter and orientation will affect the electronic structure, and influence the electrical conductivity and the Seebeck coefficient. If one considers
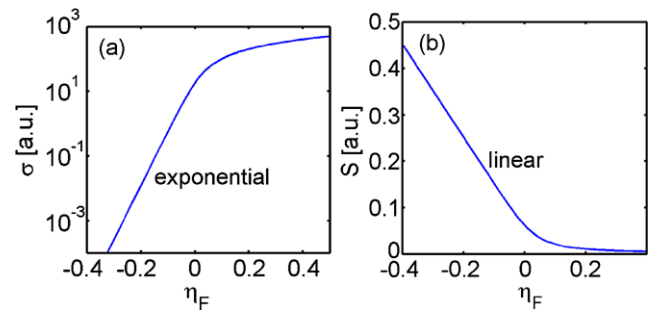

Fig. 11 The electrical conductivity (**a**) and Seebeck coefficient (**b**) versus the distance of the conduction band from the Fermi level, $\eta_F = E_F - E_C$. A simple parabolic band and scattering rates proportional to the density of final states are assumed

a specific carrier concentration, the influence of geometry shows up is two ways: (i) The band edges (or transport distribution functions TD) shift with respect to the Fermi level as the geometry changes. (ii) The effective masses (or carrier velocities) change. The changes will be different for different NW cases. A change in $\eta_F = E_F - E_C$ will affect both the conductivity and the Seebeck coefficient. This effect is shown in Figs. 11a and b, respectively, using a simple 1D subband and effective mass approximation. Changes in η_F affect the conductivity exponentially, but affect the Seebeck coefficient only linearly, (and in an inverse way). The conductivity, therefore, is affected much more than the Seebeck coefficient. At a specific carrier concentration, changes in η_F can happen as follows: (i) In a NW channel with only a few subbands, once the diameter is reduced, η_F increases in the negative direction in order to keep the carrier concentration constant as explained in detail in Refs. [23, 39]. This reduces the conductivity exponentially. (ii) The DOS changes through electronic structure modifications and η_F will adjust to keep the carrier concentration constant.

As a consequence, since the electronic structures of the NWs in different orientations are different, η_F will differ as well, resulting in orientation and geometry dependence of

TE performance. Figure 12 shows the power factor for the n-type (solid) and p-type (dashed) NWs with $D = 10$ nm, in the [100] (blue), [110] (red), and [111] (green) transport orientations. The Boltzmann transport formalism was used. Some orientation dependence can be observed. Especially in p-type NWs the [111] orientation gives almost ∼2× higher power factor than the other two p-type NW orientations. Note that p-type NWs perform lower than the n-type NWs for this NW diameter, but this difference is less severe for smaller diameters [23].

The conductivity usually degrades with diameter reduction because of the enhancement of scattering mechanisms
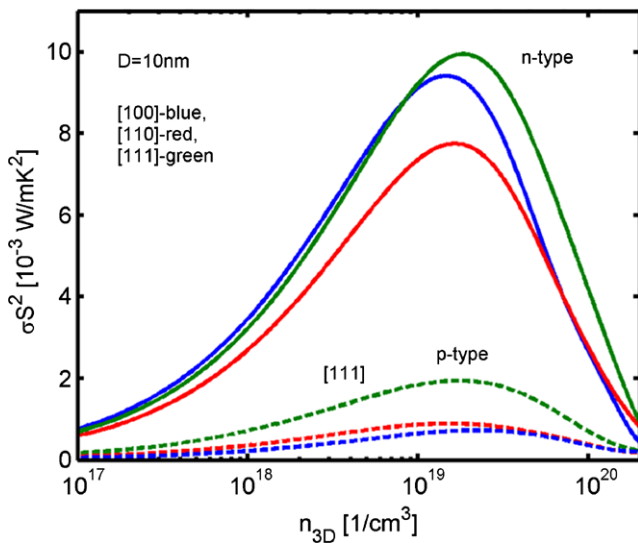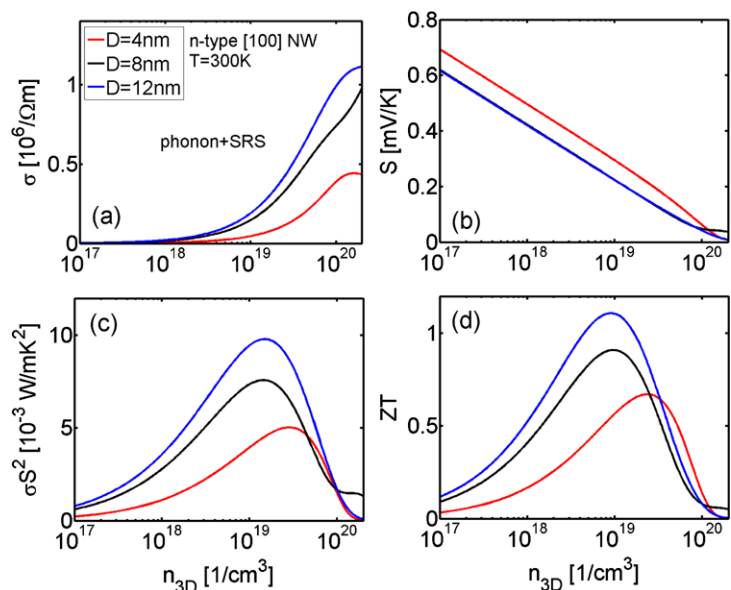

Fig. 12 (Color online) The phonon-limited thermoelectric power factor for $D = 10$ nm, n-type (*solid*) and p-type (*dashed*) NWs of different transport orientations versus the carrier concentration. Orientations are [100] (*blue*), [110] (*red*), and [111] (*green*)

such as phonon and surface roughness scattering (SRS) at smaller feature sizes. Figure 13 shows the effect of the diameter reduction on the TE coefficients for the [100] n-type NW at room temperature. Phonon scattering and SRS are considered. Figure 13a shows that the electrical conductivity decreases as the diameter of the NW is reduced. On the other hand, the Seebeck coefficient in Fig. 13b increases for the smaller diameters due to an η_F increase. Overall, the power factor in Fig. 13c decreases with diameter, because the conductivity is degraded much more than the Seebeck coefficient is improved. Using an experimentally measured value for the thermal conductivity $\kappa_l = 2$ W/mK [15, 18], we compute the *ZT* figure of merit in Fig. 13d. *ZT* is reduced with diameter reduction, following the trend in the power factor. Two important observations can be made at this point: (i) The conclusions are different from what previously described in Figs. 3 and 4 for ballistic transport. The increase in the power factor and *ZT* at reduced feature sizes is not observed when scattering is incorporated. On the contrary, the performance is degraded, because of a reduction in the conductivity. (ii) $ZT \sim 0.5$–1 can be achieved in Si NWs, in agreement with recent experimental measurements [12, 13] (reduced from $ZT \sim 4$ under ballistic considerations in Fig. 4). On the other hand, the value $\kappa_l = 2$ W/mK used for the calculation of *ZT* is measured for Si NWs of diameters $D = 15$ nm [15, 18]. This might be even smaller for smaller NW diameters or even orientation dependent [67, 68]. The power factor and *ZT* could potentially change and higher performance could be achieved. Nevertheless, the magnitude of these results is in agreement with other reports, both theoretical [37, 69] and experimental [12–14, 65].

The results in Fig. 13 only consider phonon scattering and SRS. The peak of the power factor, however, appears at carrier concentrations of 10^{19}/cm^3. In order to reach such

Fig. 13 (Color online) Thermoelectric coefficients versus carrier concentration for n-type [100] NWs of $D = 4$ nm (*red*), 8 nm (*black*) and 12 nm (*blue*), at 300 K. Phonon scattering plus SRS are included. (**a**) The electrical conductivity. (**b**) The Seebeck coefficient. (**c**) The power factor. (**d**) The *ZT* figure of merit

Fig. 14 (Color online) Thermoelectric coefficients versus carrier concentration for a $D = 5$ nm, n-type NW, in the [100] transport orientation at 300 K. Different scattering mechanisms are considered, phonons (*blue*), phonons plus SRS (*black*), and phonon plus SRS plus impurity scattering (*red*). For impurity scattering, $n_0 = n_{3D}$ is assumed. (**a**) The electrical conductivity. (**b**) The Seebeck coefficient. (**c**) The power factor. (**d**) The ZT figure of merit
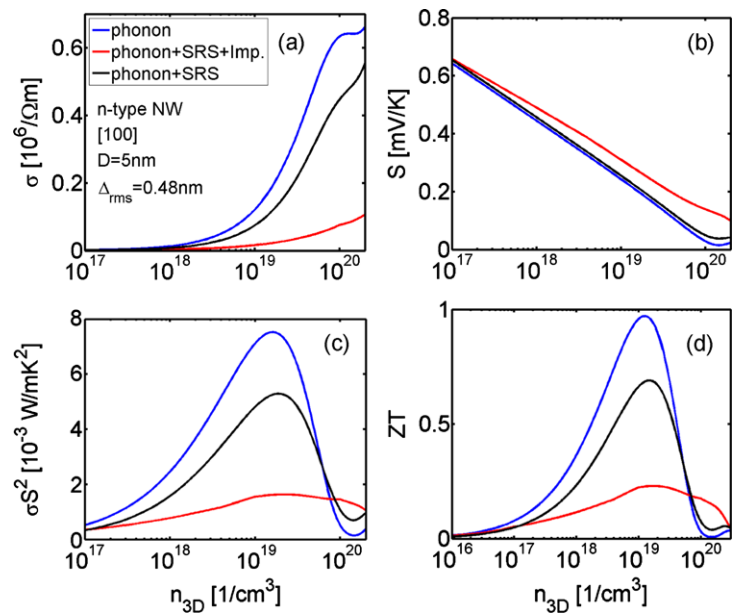
concentration high doping levels are required and the effect of impurity scattering thus cannot be excluded. In Fig. 14, we demonstrate the effect of different scattering mechanisms for the n-type [100] NW of diameters $D = 5$ nm. The conductivity in Fig. 14a is strongly degraded from the phonon-limited values (blue) once surface roughness scattering-SRS (black) and most importantly impurity scattering (red) are included in the calculation. The impurity concentration used at each instance is equal to the carrier concentration. The Seebeck coefficient in Fig. 14b does not change significantly with the introduction of additional scattering mechanisms because it is independent of scattering at first order [31]. The conductivity dominates the power factor, which is drastically reduced due to SRS and mostly impurity scattering (Fig. 14c). This can also reduce the ZT as shown in Fig. 14d from $ZT \sim 1$ down to $ZT \sim 0.2$. Since impurity scattering is such a strong mechanism, for high performance NW TEs alternative doping schemes need to be employed such as modulation doping or charge transfer techniques [65, 70–72].

5 Conclusions

We presented a methodology that couples the atomistic $sp^3d^5s^*$-SO tight-binding model to two different transport formalisms: (i) Landauer ballistic and (ii) Linearized Boltzmann theory for calculating the thermoelectric power factor in ultra-thin Si nanowires. We introduced some approximations needed to make such methodology robust and efficient, and explained the differences in the conclusions obtained from these two different transport methods. Using this formalism the computational domain can be extended

to "large" feature sizes (>10 nm) still accounting for all atomistic effects, so that the length scale degree of freedom can be properly used as a design parameter. We show that geometrical features such as cross section and orientations could potentially provide optimization directions for the thermoelectric power factor in NWs. In the Si NWs investigated, low-dimensionality and geometrical features affect the electrical conductivity much more than the Seebeck coefficient. The conductivity is, therefore, the quantity that controls the behavior of the power factor and the figure of merit ZT, in contrast to the current view that the low-dimensional features could provide benefits through improvements in the Seebeck coefficient. We finally show that impurity scattering is the strongest scattering mechanism in nanowire thermoelectric channels, and ways that allow high carrier concentration without direct doping could largely improve the performance.

Appendix A: Wavefunction overlap integral/numerical calculation of the sum

The wavefunctions in tight-binding are sampled on the atomic sites. Equations (11) and (20) involve integrations over the in-plane R perpendicular to the nanowire axis. In the calculation the integrals over R are performed by transforming the integrals to summations over the atomic sites N. Below we demonstrate how this is performed for the calculation of the wavefunction overlaps in the case of phonon scattering. The matrix element needs to be squared in the

calculation of the scattering rates. What is required is integration of the type:

$$\frac{1}{A^{m,n}_{k'_x,k_x}} = \frac{1}{A^2}\int_R \rho^{m,n}_{k'_x,k_x}(\vec{R})^2 d^2R, \quad \text{where} \tag{A.1a}$$

$$\rho^{m,n}_{k'_x,k_x}(\vec{R}) = F_{m,k'_x}(\vec{R})^* F_{n,k_x}(\vec{R}) \tag{A.1b}$$

and $1/A^2$ originates from wavefunction normalization.

We convert the integral to a sum by

$$\frac{1}{A}\int_R (\bullet)d^2R \approx \frac{1}{A}\sum_R (\bullet)\Delta A = \frac{1}{N}\sum_R (\bullet), \tag{A.2a}$$

where $\Delta A = A/N$, and N is the number of atomic sites in the unit cell of the NW.

Therefore,

$$\frac{1}{A^{m,n}_{k'_x,k_x}} = \frac{1}{A^2}\int_R \rho^{m,n}_{k'_x,k_x}(\vec{R})^2 d^2R$$

$$= \frac{1}{AN}\sum_R \rho^{m,n}_{k'_x,k_x}(\vec{R})^2$$

$$= \frac{1}{A}\left(\frac{1}{N}\sum_R \rho^{m,n}_{k'_x,k_x}(\vec{R})^2\right). \tag{A.3}$$

In the wavefunction normalization, the usual expression in integral or summation form is:

$$\frac{1}{A}\int_R F_{n,k_x}(\vec{R})F_{n,k_x}(\vec{R})^* dA = 1, \quad \text{or} \tag{A.4a}$$

$$\frac{1}{N}\sum_R F_{n,k_x}(\vec{R})F_{n,k_x}(\vec{R})^* = 1. \tag{A.4b}$$

Numerically, however, the wavefunctions provided by the eigenvalue solvers are already normalized and give:

$$\sum_R \tilde{F}_{n,k_x}(\vec{R})\tilde{F}_{n,k_x}(\vec{R})^* = 1. \tag{A.5}$$

where $F_{n,k_x} = \sqrt{N}\tilde{F}_{n,k_x}$.

The expression in (A.3) then becomes:

$$\frac{1}{A^{m,n}_{k'_x,k_x}} = \frac{1}{A}\left(N\sum_R \tilde{\rho}^{m,n}_{k'_x,k_x}(\vec{R})^2\right),$$

where $\tilde{\rho}^{m,n}_{k'_x,k_x}(\vec{R})$ is calculated using the actual F expressions given by the eigenvalue solver, and: $\rho^{m,n}_{k'_x,k_x}(\vec{R}) = N^2\tilde{\rho}^{m,n}_{k'_x,k_x}(\vec{R})$.

Appendix B: Bond passivation (sp³-hybridization) in tight-binding, following from Ref. [36]

The passivation of the bonds that reside outside the domain of the NW, is done using a sp³ hybridized scheme. The construction of the Hamiltonian assumes sums of couplings between atomic basis orbitals (orbital-space). This means that each on-site element represents a specific orbital and has contributions from four bonds (couplings). In order to passivate a specific bond, a transformation to the hybridized sp³ space is performed. This means that the transformed matrix reside in the hybridized "bond-space", in which all hybridized orbitals are aligned along the four bond directions. The on-side element of the hybridized orbital along the dangling bond direction that is to be passivated is then raised to a large value (30 eV), in order to be placed away from the energies of interest and not to affect the bandstructure calculation. The bonds from an anion to the four cations and vice versa, are formed primarily by sp³-hybridization as a linear combination of only the s and p orbitals. The sp³ hybridized orbitals from an anion to the cations are:

$$|sp^3\rangle^{a\to c}_{[111]} = \tfrac{1}{2}(|s\rangle + |p_x\rangle + |p_y\rangle + |p_z\rangle)$$

$$|sp^3\rangle^{a\to c}_{[\bar{1}\bar{1}1]} = \tfrac{1}{2}(|s\rangle - |p_x\rangle - |p_y\rangle + |p_z\rangle)$$

$$|sp^3\rangle^{a\to c}_{[1\bar{1}\bar{1}]} = \tfrac{1}{2}(|s\rangle + |p_x\rangle - |p_y\rangle - |p_z\rangle)$$

$$|sp^3\rangle^{a\to c}_{[\bar{1}1\bar{1}]} = \tfrac{1}{2}(|s\rangle - |p_x\rangle + |p_y\rangle - |p_z\rangle)$$

$$\Rightarrow \begin{bmatrix} |sp^3\rangle^{a\to c}_{[111]} \\ |sp^3\rangle^{a\to c}_{[\bar{1}\bar{1}1]} \\ |sp^3\rangle^{a\to c}_{[1\bar{1}\bar{1}]} \\ |sp^3\rangle^{a\to c}_{[\bar{1}1\bar{1}]} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} |s\rangle \\ |p_x\rangle \\ |p_y\rangle \\ |p_z\rangle \end{bmatrix}$$

$$= V^{a\to c}_{sp^3} \tag{B.1}$$

whereas the sp³ hybridized orbitals from a cation to the four anions are:

$$|sp^3\rangle^{c\to a}_{[\bar{1}\bar{1}\bar{1}]} = \tfrac{1}{2}(|s\rangle - |p_x\rangle - |p_y\rangle - |p_z\rangle)$$

$$|sp^3\rangle^{c\to a}_{[11\bar{1}]} = \tfrac{1}{2}(|s\rangle + |p_x\rangle + |p_y\rangle - |p_z\rangle)$$

$$|sp^3\rangle^{c\to a}_{[\bar{1}11]} = \tfrac{1}{2}(|s\rangle - |p_x\rangle + |p_y\rangle + |p_z\rangle)$$

$$|sp^3\rangle^{c\to a}_{[1\bar{1}1]} = \tfrac{1}{2}(|s\rangle + |p_x\rangle - |p_y\rangle + |p_z\rangle)$$

$$\Rightarrow \begin{bmatrix} |sp^3\rangle^{c\to a}_{[\bar{1}\bar{1}\bar{1}]} \\ |sp^3\rangle^{c\to a}_{[11\bar{1}]} \\ |sp^3\rangle^{c\to a}_{[\bar{1}11]} \\ |sp^3\rangle^{c\to a}_{[1\bar{1}1]} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} |s\rangle \\ |p_x\rangle \\ |p_y\rangle \\ |p_z\rangle \end{bmatrix}$$

$$= V^{c\to a}_{sp^3} \tag{B.2}$$

The passivation is then achieved by a transformation as follows:

$$[H]_{Hybrid} = V_{sp^3}[H_{E(sp)}]V_{sp^3}^\dagger, \tag{B.3}$$

where:

$$H_{E(sp)} = \begin{bmatrix} E_s & & & \\ & E_{px} & & \\ & & E_{py} & \\ & & & E_{pz} \end{bmatrix},$$

is the on-site matrix consisting only of the s and p orbitals. Once the transformation takes place, the on-site elements of the hybridized space matrix along the bonds to be passivated are raised by $(h_{sp3})_{i,i} = 30$ eV. Finally, a back transformation into the orbital space will give the passivated matrix elements:

$$[H]_{Passiv.} = V_{sp^3}^\dagger[H_{Hybrid} + h_{sp^3}]V_{sp^3},$$

$$\text{where } h_{sp^3} = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & a_3 & \\ & & & a_4 \end{bmatrix}, \tag{B.4}$$

with a_i been 30 eV or zero, depending on whether the bond i is passivated or not.

References

1. Snyder, G.J., Toberer, E.S.: Nat. Mater. **7**, 105–114 (2008)
2. Majumdar, A.: Sci. Mater. **303**, 777–778 (2004)
3. Hicks, L.D., Dresselhaus, M.S.: Phys. Rev. B **47**(24), 16631 (1993)
4. Venkatasubramanian, R., Siivola, E., Colpitts, T., O' Quinn, B.: Nature **413**, 597–602 (2001)
5. Broido, D.A., Reinecke, T.L.: Phys. Rev. B **51**, 13797 (1994)
6. Harman, T.C., Taylor, P.J., Walsh, M.P., LaForge, B.E.: Science **297**, 2229–2232 (2002)
7. Dresselhaus, M., Chen, G., Tang, M.Y., Yang, R., Lee, H., Wang, D., Ren, Z., Fleurial, J.-P., Gagna, P.: Adv. Mater. **19**, 1043–1053 (2007)
8. Shakouri, A.: In: IEEE International Conference on Thermoelectrics, pp. 495–500 (2005)
9. Sofo, J.O., Mahan, G.D.: Appl. Phys. Lett. **65**, 2690 (1994) (3pp)
10. Kim, W., Singer, S.L., Majumdar, A., Vashaee, D., Bian, Z., Shakouri, A., Zeng, G., Bowers, J.E., Zide, J.M.O., Gossard, A.C.: Appl. Phys. Lett. **88**, 242107 (2006)
11. Wu, Y., Fan, R., Yang, P.: Nano Lett. **2**(2), 83 (2002)
12. Boukai, A.I., Bunimovich, Y., Tahir-Kheli, J., Yu, J.-K., Goddard, W.A. III, Heath, J.R.: Nature **451**, 168–171 (2008)
13. Hochbaum, A.I., Chen, R., Delgado, R.D., Liang, W., Garnett, E.C., Najarian, M., Majumdar, A., Yang, P.: Nature **451**, 163–168 (2008)
14. Tang, J., Wang, H.-T., Lee, D.H., Fardy, M., Huo, Z., Russell, T.P., Yang, P.: Nano Lett. **10**(10), 4279–4283 (2010)
15. Li, D., Wu, Y., Fang, R., Yang, P., Majumdar, A.: Appl. Phys. Lett. **83**, 3186–3188 (2003)
16. Nielsch, K., Bachmann, J., Kimling, J., Böttner, H.: Adv. Energy Mater. **1**, 713 (2011)
17. Chen, G.: Semicond. Semimet. **71**, 203–259 (2001)
18. Chen, R., Hochbaum, A.I., Murphy, P., Moore, J., Yang, P., Majumdar, A.: Phys. Rev. Lett. **101**, 105501 (2008)
19. Li, D., Huxtable, S.T., Abramsin, A.R., Majumdar, A.: Trans. ASME **127**, 108–114 (2005)
20. Martin, P., Aksamija, Z., Pop, E., Ravaioli, U.: Phys. Rev. Lett. **102**, 125503 (2009)
21. Vineis, C.J., Shakouri, A., Majumdar, A., Kanatzidis, M.C.: Adv. Mater. **22**, 3970–3980 (2010)
22. Jaworski, C.M., Kulbachinskii, V., Heremans, J.P.: Phys. Rev. B **80**, 125208 (2009)
23. Neophytou, N., Kosina, H.: Phys. Rev. B **83**, 245305 (2011)
24. Mahan, G.D., Sofo, J.O.: Proc. Natl. Acad. Sci. USA **93**, 7436–7439 (1996)
25. Boykin, T.B., Klimeck, G., Oyafuso, F.: Phys. Rev. B **69**(11), 115201 (2004)
26. Klimeck, G., Ahmed, S., Hansang, B., Kharche, N., Clark, S., Haley, B., Lee, S., Naumov, M., Ryu, H., Saied, F., Prada, M., Korkusinski, M., Boykin, T.B., Rahman, R.: IEEE Trans. Electron Devices **54**(9), 2079–2089 (2007)
27. Klimeck, G., Ahmed, S., Kharche, N., Korkusinski, M., Usman, M., Prada, M., Boykin, T.B.: IEEE Trans. Electron Devices **54**(9), 2090–2099 (2007)
28. Neophytou, N., Paul, A., Lundstrom, M., Klimeck, G.: IEEE Trans. Electron Devices **55**(6), 1286–1297 (2008)
29. Neophytou, N., Paul, A., Klimeck, G.: IEEE Trans. Nanotechnol. **7**(6), 710–719 (2008)
30. Landauer, R.: IBM J. Res. Dev. **1**, 223 (1957)
31. Kim, R., Datta, S., Lundstrom, M.S.: J. Appl. Phys. **105**, 034506 (2009)
32. Jeong, C., Kim, R., Luisier, M., Datta, S., Lundstrom, M.: J. Appl. Phys. **107**, 023707 (2010)
33. Liang, G., Huang, W., Koong, C.S., Wang, J.-S., Lan, J.: J. Appl. Phys. **107**, 014317 (2010)
34. Neophytou, N., Wagner, M., Kosina, H., Selberherr, S.: J. Electron. Mater. **39**(9), 1902–1908 (2010)
35. Scheidemantel, T.J., Ambrosch-Draxl, C., Thonhauser, T., Badding, J.V., Sofo, J.O.: Phys. Rev. B **68**, 125210 (2003)
36. Lee, S., Oyafuso, F., Von Allmen, P., Klimeck, G.: Phys. Rev. B **69**, 045316 (2004)
37. Vo, T.T.M., Williamson, A.J., Lordi, V., Galli, G.: Nano Lett. **8**(4), 1111–1114 (2008)
38. Markussen, T., Jauho, A.-P., Brandbyge, M.: Phys. Rev. Lett. **103**, 055502 (2009)
39. Neophytou, N., Kosina, H.: Phys. Rev. B **84**, 085313 (2011)
40. Lundstrom, M.: Fundamentals of Carrier Transport. Cambridge University Press, Cambridge (2000)
41. Ferry, D.K., Goodnick, S.M.: Transport in Nanostructures. Cambridge University Press, Cambridge (1997)
42. Fischetti, M.V.: J. Appl. Phys. **89**, 1232 (2001)
43. Fischetti, M.V., Ren, Z., Solomon, P.M., Yang, M., Rim, K.: J. Appl. Phys. **94**, 1079 (2003)
44. Esseni, D., Abramo, A.: IEEE Trans. Electron Devices **50**(7), 1665 (2003)
45. Buin, A.K., Verma, A., Svizhenko, A., Anantram, M.P.: Nano Lett. **8**(2), 760–765 (2008)
46. Buin, A.K., Verma, A., Anantram, M.P.: J. Appl. Phys. **104**, 053716 (2008)
47. Ramayya, E.B., Vasileska, D., Goodnick, S.M., Knezevic, I.: J. Appl. Phys. **104**, 063711 (2008)
48. Jin, S., Fischetti, M.V., Tang, T.: J. Appl. Phys. **102**, 83715 (2007)
49. Jacoboni, C., Reggiani, L.: Rev. Mod. Phys. **55**, 645 (1983)
50. Goodnick, S.M., Ferry, D.K., Wilmsen, C.W., Lilienthal, Z., Fathy, D., Krivanek, O.L.: Phys. Rev. B **32**, 8171 (1985)

51. Sakaki, H., Noda, T., Hirakawa, K., Tanaka, M., Matsusue, T.: Appl. Phys. Lett. **51**(23), 1934 (1987)
52. Uchida, K., Takagi, S.: Appl. Phys. Lett. **82**(17), 2916–2918 (2003)
53. Fang, T., Konar, A., Xing, H., Jena, D.: Phys. Rev. B **78**, 205403 (2008)
54. Wang, J., Polizzi, E., Ghosh, A., Datta, S., Lundstrom, M.: Appl. Phys. Lett. **87**, 043101 (2005)
55. Prange, R.E., Nee, T.-W.: Phys. Rev. **168**, 779 (1968)
56. Ando, T., Fowler, A., Stern, F.: Rev. Mod. Phys. **54**, 437 (1982)
57. Esseni, D.: IEEE Trans. Electron Devices **51**(3), 394 (2004)
58. Lee, J., Spector, H.N.: J. Appl. Phys. **57**(2), 366 (1985)
59. Sarma, S.D., Lai, W.: Phys. Rev. B **32**(2), 1401 (1985)
60. Yuh, P., Wang, K.L.: Appl. Phys. Lett. **49**(25), 1738 (1986)
61. Neophytou, N., Kosina, H.: Nano Lett. **10**(12), 4913–4919 (2010)
62. Donetti, L., Gamiz, F., Roldan, J.B., Godoy, A.: J. Appl. Phys. **100**, 013701 (2006)
63. Fischetti, M.V., Laux, S.E.: J. Appl. Phys. **80**, 2234 (1996)
64. Yamada, T., Ferry, D.K.: Solid-State Electron. **38**, 881 (1995)
65. Ryu, H.J., Aksamija, Z., Paskiewicz, D.M., Scott, S.A., Lagally, M.G., Knezevic, I., Eriksson, M.A.: Phys. Rev. Lett. **105**, 256601 (2010)
66. Kosina, H., Kaiblinger-Grujin, G.: Solid-State Electron. **42**(3), 331–338 (1998)
67. Markussen, T., Jauho, A.-P., Brandbyge, M.: Nano Lett. **8**(11), 3771–3775 (2008)
68. Aksamija, Z., Knezevic, I.: Phys. Rev. B **82**, 045319 (2010)
69. Markussen, T., Jauho, A.-P., Brandbyge, M.: Phys. Rev. B **79**, 035415 (2009)
70. Goswami, S., Siegert, C., Shamim, S., Pepper, M., Farrer, I., Ritchie, D.A., Ghosh, A.: Appl. Phys. Lett. **97**, 132104 (2010)
71. Zhang, P., Tewaarwerk, E., Park, B.-N., Savage, D.E., Celler, G.K., Knezevic, I., Evans, P.G., Eriksson, M.A., Lagally, M.G.: Nat. Lett. **439**, 703 (2006)
72. He, T., He, J., Lu, M., Chen, B., Pan, H., Reus, W.F., Nolte, W.M., Nackashi, D.P., Franzon, P.D., Tour, J.M.: J. Am. Chem. Soc. **128**, 14537 (2006)