Electron energy filtering by a nonplanar potential to enhance the thermoelectric power factor in bulk materials

Je-Hyeong Bahk,[1,*] Zhixi Bian,[2] and Ali Shakouri[1,2]

[1]*Birck Nanotechnology Center, Purdue University, West Lafayette, Indiana 47907, USA*

[2]*Department of Electrical Engineering, University of California, Santa Cruz, California 95064, USA*

(Received 29 August 2012; revised manuscript received 15 November 2012; published 14 February 2013)

We present a detailed theory on electron energy filtering by the nonplanar potential introduced by dispersed nanoparticles or impurities in bulk materials for enhancement of the thermoelectric power factor. When electrons with energies below a certain cut-off energy are prevented from participating in conduction through the material, the Seebeck coefficient and thus the thermoelectric power factor can be drastically enhanced. Instead of using planar heterostructures which require elaborate epitaxial techniques, we study embedded nanoparticles or impurities so that the conservation of lateral momentum limiting electron transport at heterointerfaces is no longer a limiting factor. Based on the Boltzmann transport equations under the relaxation time approximation, the optimal cut-off energy level that maximizes the power factor is calculated to be a few k_BT above the Fermi level, and is a function of the scattering parameter, Fermi level, and temperature. The maximized power factor enhancement is quantified as a function of those parameters. The electronic thermal conductivity and Lorenz number are also shown to be suppressed by the electron filtering to further enhance the thermoelectric figure of merit. We find that the power factor of PbTe at 300 K could be enhanced by more than 120% when the cut-off energy level is 0.2 eV or higher and the carrier density higher than 5×10^{19} cm^{-3}. Finally we propose the use of distributed resonant scatterings to partially realize the nonplanar electron filtering in bulk materials.

DOI: 10.1103/PhysRevB.87.075204 PACS number(s): 72.15.Jf, 72.10.Bg

I. INTRODUCTION

Thermoelectric energy conversion has been drawing great attention as a viable solution to waste heat recovery and microchip hotspot cooling.[1,2] The efficiency of thermoelectric devices depends directly on the dimensionless figure of merit $ZT = S^2\sigma T/(\kappa_l + \kappa_e)$, of the materials used, where S is the Seebeck coefficient, σ is the electrical conductivity, T is the absolute temperature, κ_l is the lattice thermal conductivity, and κ_e is the electronic thermal conductivity. Significant enhancements of the thermoelectric figure of merit have been reported in recent years for nanostructured materials such as superlattices and nanocomposites, and attributed mainly to the lattice thermal conductivity reduction via additional interface phonon scattering in these materials.[3–6]

Completely different approaches are required, however, for enhancing the so-called thermoelectric power factor $S^2\sigma$ in the numerator of ZT since this property is related to charge carrier transport, and there is a trade-off relation between the Seebeck coefficient and the electrical conductivity. Hicks and Dresselhaus theoretically predicted drastic power factor enhancement by the modified density of states in low-dimensional materials such as quantum wells and wires in 1993.[7] Since then, a rebirth of interest has been made in thermoelectric material research, and a considerable number of research activities have been conducted on various low-dimensional and nanostructured materials for thermoelectric energy conversion.[8,9] Recently, Heremans *et al.* demonstrated a large power factor enhancement in Tl-doped PbTe at high temperatures, and attributed this enhancement to the distorted density of states by the Tl resonant level inside the valence band of PbTe.[10] The concept of modulation doping has been experimentally demonstrated in SiGe nanocomposites to improve electrical conductivity over bulk values, thus enhancing the power factor.[11] Bahk *et al.* proposed that the electrical conductivity and power factor can be enhanced over a wide temperature range when embedded nanoparticles of a few nanometers in diameter donate charge carriers instead of the conventional impurity dopants, thereby replacing the stronger impurity scattering with weaker nanoparticle scattering.[12] Also, it has been theoretically predicted that energetically sharp resonant electron scattering by embedded core-shell nanoparticles that form quasibound states inside the band can significantly enhance the power factor at low temperatures.[13]

The concept of electron energy filtering was originally proposed using planar barriers in thermionic energy conversion devices. In 1994, vacuum thermionic coolers that utilize a potential barrier between electrode and vacuum were proposed by Mahan.[14] The requirement for efficient cooling at room temperature was the work function of the electrode metal to be about 0.3–0.4 eV, which was not feasible to achieve. Shakouri and Bowers proposed in 1997 to use semiconductor heterostructures for the selective emission of hot electrons over a barrier layer to enhance the cooling performance.[15] By controlling the conduction band offsets in heterostructures, high cooling power density at room temperature was predicted in the nonlinear thermionic transport regime. Later, Shakouri *et al.* proposed that tall barrier, highly degenerate multilayers could achieve thermoelectric power factors an order of magnitude higher than bulk values.[16] They considered thermionic electron transport perpendicular to the planar barriers and the barrier layers which were thick enough to prevent tunneling.

In the case of electron transport perpendicular to planar barriers, the lateral momentum component is conserved during transport over the barriers analogous to the light refraction in multilayers. This is called lateral momentum conservation.[17] The transverse momentum component needs to be larger than

a certain value determined by conduction band offset for the electron to go over the potential barrier. Vashaee and Shakouri pointed out that the key requirement for a large thermoelectric power factor enhancement in planar barrier materials is that the lateral momentum conservation is relaxed, such that the momentum criterion for the transport at the interfaces is replaced by the energy criterion depicting that most of the electrons having energies greater than the cut-off energy are allowed to participate in transport over the barrier.[17] The idea was that this lateral momentum nonconservation allows a much larger number of hot electrons to participate in the emission process, so that the suppression of electrical conductivity by barriers is significantly alleviated while a large enhancement of the Seebeck coefficient by hot electron filtering is still achieved. However, Kim *et al.* recently found that the enhancement in emission current due to the lateral momentum nonconservation could be modest because the smallest number of modes in the well and barrier layers limits the emission current.[18]

Due to these complications in the planar carrier filtering scheme, it is necessary to realize a nonplanar carrier energy filtering without extended planar potential barriers, but instead with embedded discrete nanoparticles or impurities. In the latter case, the lateral momentum is not conserved since there is no translational invariance, and thus the power factor can be enhanced in bulk materials. In an ideal case of nonplanar energy filtering, the modeling of carrier transport becomes simply having a cut-off energy level in the bulk transport calculations, below which all the charge carriers are filtered out of the transport.

In this paper we present a general theory on electron energy filtering, as similarly shown in Ref. 19, with a cut-off energy in the Boltzmann transport theory under the relaxation time approximation for bulk materials. We extend this theory to find the optimal cut-off energy level and corresponding maximum power factor enhancement as a function of the scattering parameter, Fermi level, and temperature in Sec. III. An approximate expression for the optimal cut-off energy level is provided. Variations of the electronic thermal conductivity, Lorenz number, and the Hall factor by the filtering effect are also quantified as functions of the aforementioned parameters in Sec. IV. Then the proposed theory is applied to a real material PbTe with a nonparabolic band and realistic energy-dependent scattering time taken into account in Sec. V. In Sec. VI, after a brief investigation of various nonplanar potential scatterers, such as step-potential nanoparticles, resonant core-shell nanoparticles, and resonant impurities, we propose the use of distributed resonant scatterings in Lorentzian shapes to partially realize the nonplanar electron energy filtering in bulk materials. Finally, Sec. VII concludes the paper.

II. ELECTRON TRANSPORT IN BULK MATERIALS WITH A CUT-OFF ENERGY

In equilibrium, the distribution of electrons is given by the Fermi-Dirac distribution

$$f_0(E) = \frac{1}{1+\exp[(E-E_F)/k_BT]}, \tag{1}$$

where E is the electron energy, and E_F is the Fermi level, k_B is the Boltzmann constant, and T is the absolute temperature. In this paper we assume that all energies are referenced to the conduction band minimum. In near equilibrium with a small external field and a temperature gradient, the transport properties can be derived from the Boltzmann transport equation under the relaxation time approximation. They are all integral functions of the differential conductivity given by

$$\sigma_d(E) = e^2\tau(E)v^2(E)\rho_{\mathrm{DOS}}(E)\left(-\frac{\partial f_0}{\partial E}\right), \tag{2}$$

where e is the electron charge, $\tau(E)$ is the total electron momentum relaxation time, $v(E)$ is the electron velocity in one direction, i.e., $v^2(E) = 2E/(3m^*)$ for a parabolic band, m^* is the effective mass, and $\rho_{\mathrm{DOS}}(E)$ is the density of states. The Fermi window factor $(-\partial f_0/\partial E)$ is a bell-shaped function centered at the Fermi level with a few k_BT width, and represents the distribution of electrons that contributes to conduction.

If there is a cut-off energy level E_C below which all the electrons are completely blocked from participating in conduction, the calculations of the transport properties are reduced to the integrals over energy from E_C to infinity such that

$$\sigma = \int_{E_C}^{\infty}\sigma_d(E)dE, \tag{3}$$

$$S = \frac{1}{eT}\frac{\int_{E_C}^{\infty}\sigma_d(E)(E-E_F)dE}{\int_{E_C}^{\infty}\sigma_d(E)dE}. \tag{4}$$

In this paper we ignore the sign of Seebeck coefficient as we deal with one-type carrier transport.

For a parabolic band and $\tau(E) = \tau_0E^r$, where τ_0 is a constant, and r is called the scattering parameter, (3) and (4) can be simplified as

$$\sigma = C_1F_0(E_C), \tag{5}$$

$$S = C_2\left[\frac{F_1(E_C)}{F_0(E_C)} - E_F\right], \tag{6}$$

and thus the power factor becomes

$$S^2\sigma = C_1C_2^2\frac{[F_1(E_C) - E_FF_0(E_C)]^2}{F_0(E_C)}, \tag{7}$$

where $C_1 = 2^{3/2}e^2(m^*)^{1/2}\tau_0/(3\pi^2\hbar^3)$ and $C_2 = 1/(eT)$ are all constants, and

$$\begin{aligned} F_s(E_C) &= \int_{E_C}^{\infty}E^{\frac{3}{2}+r+s}\left(-\frac{\partial f_0}{\partial E}\right)dE \\ &= \left(\frac{3}{2}+r+s\right)\int_{E_C}^{\infty}E^{\frac{1}{2}+r+s}f_0(E)dE \\ &\quad + E_C^{\frac{3}{2}+r+s}f_0(E_C), \end{aligned} \tag{8}$$

with $s = 0$, 1, or 2. Note that the first term in the second line of (8) is an incomplete Fermi-Dirac integral of the order of $(\frac{1}{2}+r+s)$.

The scattering parameter r manifests the dominant scattering mechanism in a material. For example, r is -0.5 for

the acoustic phonon deformation potential scattering, $r = 1.5$ for the ionized impurity scattering, and $r \approx 0.5$ for the polar optical phonon scattering. In general, however, more than a single scattering mechanism are combined together to make the total scattering time. For most semiconductors, r typically falls in between -0.5 and 0.5.

Here we define

$$g(E_C) = \frac{F_1(E_C)}{F_0(E_C)}, \tag{9}$$

such that the Seebeck coefficient can be rewritten from (6) as

$$S = C_2 \left[g(E_C) - E_F \right]. \tag{10}$$

The $g(E_C)$ is a monotonically increasing function with E_C, and is always larger than both E_C and E_F. More details about the derivation and proof for this are given in Appendix A. Therefore, according to (10), it is evident that the Seebeck coefficient increases with increasing E_C, which is the consequence of the electron filtering. On the other hand, $F_s(E_C)$ is a monotonically decreasing function with E_C (see Appendix A for the proof), which implies from (5) that the electrical conductivity decreases with increasing E_C. However, the power factor can still be enhanced due to much larger enhancement of the Seebeck coefficient than the suppression of the electrical conductivity.

The electronic thermal conductivity with a cut-off energy is given by

$$\kappa_e = \frac{C_1}{e^2 T} \left\{ F_2(E_C) - \frac{[F_1(E_C)]^2}{F_0(E_C)} \right\}. \tag{11}$$

The electronic thermal conductivity is related to the electrical conductivity via the Lorenz number L, which is defined as

$$L = \frac{\kappa_e}{\sigma T}. \tag{12}$$

Thus, from (5) and (11), (12) becomes

$$L = \frac{1}{e^2 T^2} \left\{ \frac{F_2(E_C)}{F_0(E_C)} - \left[\frac{F_1(E_C)}{F_0(E_C)} \right]^2 \right\}. \tag{13}$$

The Lorenz number is a function of E_F, although it becomes a constant to be $L_0 = (\pi k_B)^2/(3e^2) \approx 2.44 \times 10^{-8}$ W Ω K^{-2} for metals or in the degenerate limit ($E_F \gg 0$) in the bulk ($E_C = 0$), as described by the Wiedemann-Franz law. L deviates from L_0, monotonically decreasing, as E_F goes down, or the carrier density decreases in semiconductors. Also the Lorenz number monotonically decreases with increasing E_C, and saturates as E_C goes very high for a given E_F. More details about this are discussed in Sec. IV.

The Hall effect measurements can be used to obtain the effective carrier density n_e by measuring the Hall coefficient R_H in the case of one-type carrier transport from

$$R_H = \alpha \frac{1}{n_e q}, \tag{14}$$

where α is the Hall factor, and q is $-e$ for electrons, and $+e$ for holes. The Hall factor is given by[20]

$$\alpha = \frac{\langle \tau^2 \rangle}{\langle \tau \rangle^2} = \frac{F_0^{r=2r}(E_C) F_0^{r=0}(E_C)}{[F_0(E_C)]^2}, \tag{15}$$

where $F_0^{r=2r}(E_C)$ and $F_0^{r=0}(E_C)$ are $F_0(E_C)$ defined by (8) with $s = 0$, but with r replaced by $2r$, and 0, respectively. The Hall factor is close to unity for the materials with $-0.5 \leqslant r \leqslant 0.5$. The effective carrier density measured by the Hall effect measurements does not include the carriers below the cut-off energy because their overall displacement becomes zero by the cut off. Thus,

$$n_e = \int_{E_C}^{\infty} \rho_{\mathrm{DOS}}(E) f_0(E) dE, \tag{16}$$

which decreases with increasing E_C.

In nonideal or practical cases of the energy filtering, the modeling can be more complicated than just having a cut-off energy in the transport calculations. For example, nonplanar potentials created by embedded nanoparticles or impurities of various kinds in a bulk material can be used to modify the transport of charge carriers via energy-dependent scattering times to realize the energy filtering. If these additional energy-dependent scatterings by nanoparticles or impurities are exclusively targeted at the carriers in a specific energy region only, those carriers can be effectively prevented from participating in conduction by the extensive scatterings while others are not affected and allowed to transport, so the energy filtering is realized. An accurate calculation of the energy-dependent scattering time by any arbitrary spherically symmetric potential is possible by the partial wave method.[21] Assuming that all the scatterings are independent, the energy-dependent scattering time by the nanoparticles $\tau_{\mathrm{NP}}(E)$ is combined with other existing scattering times such as the phonon scattering $\tau_{\mathrm{PH}}(E)$ and the ionized impurity scattering $\tau_{\mathrm{II}}(E)$ in energy space as

$$\frac{1}{\tau(E)} = \frac{1}{\tau_{\mathrm{PH}}(E)} + \frac{1}{\tau_{\mathrm{II}}(E)} + \frac{1}{\tau_{\mathrm{NP}}(E)} + \cdots. \tag{17}$$

Then the total scattering time is plugged in (5), (6), and (11), with $E_C = 0$ as in bulk, to calculate the transport properties of the bulk material embedded with the nanoparticles. We will go into more details about the use of various nanoparticles and impurities in Sec. VI.

III. OPTIMAL CUT-OFF ENERGY AND POWER FACTOR

The optimal cut-off energy $E_{C,\mathrm{opt}}$ that maximizes the power factor for a given Fermi level, scattering parameter, and temperature can be found by differentiating (7) with respect to E_C and matching the derivative to zero, which reduces simply to

$$g(E_{C,\mathrm{opt}}) = 2E_{C,\mathrm{opt}} - E_F. \tag{18}$$

The derivation of (18) is provided in Appendix B. Figure 1(a) shows the graphical method of finding the optimal cut-off energy based on (18), in which the x coordinate of the intersecting point between the two curves, $y = g(E_C)$ and $y = 2E_C - E_F$, represents the optimal cut-off energy. As an example, $E_{C,\mathrm{opt}}$ is found to be 0.14 eV for $E_F = 0.1$ eV and $r = 0.5$ at 300 K from Fig. 1(a). Figure 1(b) shows the ratios of the Seebeck coefficient, the electrical conductivity, and the power factor to their bulk values as functions of the cut-off energy under the same condition as in Fig. 1(a). As anticipated

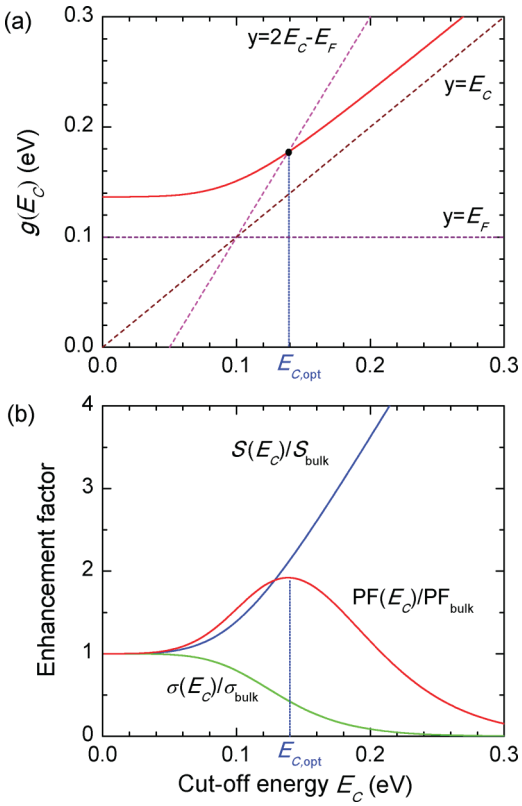

FIG. 1. (Color online) (a) $g(E_C)$ (solid curve) defined by (9) as a function of cut-off energy E_C for $E_F = 0.1$ eV, $r = 0.5$, and $T = 300$ K. The x coordinate (0.14 eV) of the intersecting point (filled dot) between y $= g(E_C)$ and $y = 2E_C - E_F$ is the optimal cut-off energy ($E_{C,\mathrm{opt}}$) that maximizes the power factor. (b) The ratios of the Seebeck coefficient (S), electrical conductivity (σ), and the power factor (PF) to their bulk counterparts as a function of the cut-off energy under the same conditions as in (a). The power factor is maximized at $E_{C,\mathrm{opt}} = 0.14$ eV, which confirms the value obtained from (a).

in the previous section, the Seebeck coefficient increases with the cut-off energy while the electrical conductivity decreases. The power factor has a maximum, which is about 90% larger than the bulk value, when the cut off is at 0.14 eV, which confirms the optimal cut-off energy value obtained from the graphical method shown in Fig. 1(a).

Figure 2 shows the difference between the optimal cut-off energy and the Fermi level as a function of the Fermi level for a varying scattering parameter ($-3/2 \leqslant r \leqslant 3/2$) at two different temperatures, 300 and 600 K. The optimal cut-off energy is always higher than the Fermi level, but the difference becomes smaller and saturated as the Fermi level goes higher. The trend is almost an exponential decay. For a larger scattering parameter, the optimal cut-off energy is higher for a fixed Fermi level and temperature. The variation of the optimal cut-off energy with the change of the scattering parameter is more significant at a lower Fermi level. At a very high Fermi level, the variation is within $0.35k_BT$ when the scattering parameter is varied from $-3/2$ to $+3/2$. The optimal cut-off energy increases almost linearly with temperature for a fixed Fermi level and scattering parameter.

Although one can obtain these optimal cut-off energies by numerically solving (18) under various conditions, an

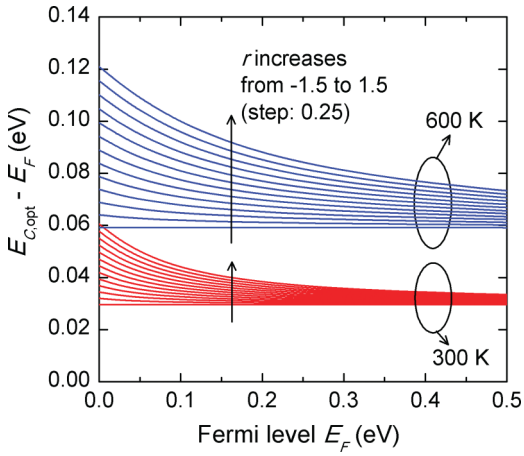

FIG. 2. (Color online) Difference between the optimal cut-off energy and the Fermi level ($E_{C,\mathrm{opt}} - E_F$) as a function of the Fermi level (E_F) at 300 and 600 K. For each temperature, the scattering parameter (r) is varied from -1.5 to 1.5 with a step of 0.25.

approximate expression for the optimal cut-off energy as a function of Fermi level, scattering parameter, and temperature can be found as

$$E_{C,\mathrm{opt}} = E_F + [A\exp(-\beta E_F) + B]k_BT, \tag{19}$$

where

$$\begin{aligned} A &= \{(-1.48\times10^{-4})T + 0.39\}(r+1.5),\\ B &= \{(1.47\times10^{-4})T + 6.66\times10^{-3}\}r\\ &\quad + \{(2.2\times10^{-4})T + 1.15\},\\ \beta &= 1/\{(2.0\times10^{-4})T + 0.04\}. \end{aligned}$$

This expression is accurate within 5% error for Fermi levels less than 0.4 eV over the temperature range from 200 to 800 K. Within 10% error, it is good for Fermi levels less than 0.5 eV from 100 to 800 K. In particular, (19) can overestimate the optimal cut-off energy at very high Fermi level and very high temperatures, or at very low Fermi level and very low temperatures.

The optimal cut-off energy is saturated at about $1.14k_BT \sim 1.30k_BT$ above the Fermi level at 300 K, and about $1.14k_BT \sim 1.45k_BT$ above the Fermi level at 600 K when the Fermi level is very high (>0.4 eV). If the Fermi level is low, close to the conduction band minimum, then the upper bound of the optimal cut-off energy is about $2.33k_BT$ above the Fermi level for $r = 3/2$, and the lower bound remains the same, $1.14k_BT$ above the Fermi level, for $r = -3/2$ at all temperatures.

Figure 3 shows the effects of temperature and scattering time on $g(E_C)$ and the Seebeck coefficient. Here, for convenience, we define Δ and Δ' as

$$\Delta = g(0) - E_F, \tag{20}$$

$$\Delta' = g(E_{C,\mathrm{opt}}) - E_F. \tag{21}$$

Figure 3(a) shows Δ and Δ' at $T = 300$ K as an example. From (10), the Seebeck coefficient at the optimal cut-off energy is $S_{\mathrm{opt}} = C_2\Delta$', and the bulk Seebeck coefficient is $S_{\mathrm{bulk}} = C_2\Delta$.

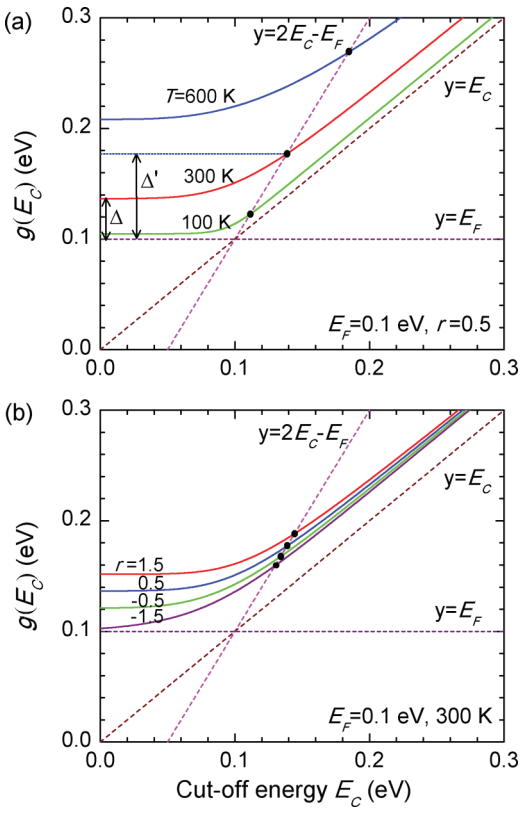

FIG. 3. (Color online) $g(E_C)$ (solid curves) defined by (9) as a function of cut-off energy E_C, (a) for varying temperature, $T = 100$, 300, and 600 K, with fixed $E_F = 0.1$ eV, $r = 0.5$, and (b) for varying scattering time $r = -1.5, -0.5, 0.5$, and 1.5, with fixed $E_F = 0.1$ eV, and $T = 300$ K. In (a), Δ and Δ' are defined.

Thus, the Seebeck enhancement factor is

$$\frac{S_{\text{opt}}}{S_{\text{bulk}}} = \frac{\Delta'}{\Delta}. \tag{22}$$

In degenerate limit, the bulk Seebeck coefficient can be expressed by[22]

$$S_{\text{bulk}} = \frac{\pi^2}{3}\frac{k_B^2 T}{e}\frac{(r+3/2)}{E_F}, \tag{23}$$

which implies that the Seebeck coefficient is proportional to temperature for a fixed Fermi level. Since $S_{\text{bulk}} = C_2\Delta$ and $C_2 = 1/eT$, we find Δ is approximately proportional to T^2 in degenerate limit. In practice, typically Δ is slightly slower than T^2. From Fig. 3(a), Δ' is almost proportional to T. Therefore, the Seebeck enhancement factor is approximately proportional to $1/T$ for a fixed Fermi level. This indicates that the effect of the electron filtering on power factor becomes weaker at higher temperatures, following $\sim 1/T$.

A similar discussion is possible for the scattering parameter. Since S_{bulk} is proportional to $(r + 3/2)$ in degenerate limit according to (23), Δ is also proportional to $(r + 3/2)$ in degenerate limit. On the other hand, Δ' is only slightly changed with r as shown in Fig. 3(b). As the Fermi level further increases, this change becomes smaller and approaches zero. Therefore, the Seebeck enhancement factor is proportional to $1/(r + 3/2)$ in degenerate limit. This indicates that a very large Seebeck enhancement, and thus very large power factor enhancement, is possible for r close to $-3/2$ when the Fermi level is high or the doping density is high. However, this is due to very small S_{bulk}. In fact, S_{opt} is higher for a larger r with a fixed Fermi level, and becomes almost independent of r when the Fermi level is very high, but S_{bulk} is significantly lowered as r decreases and approaches $-3/2$.

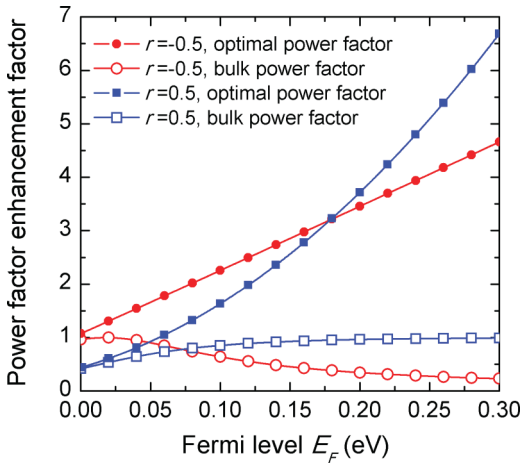

FIG. 4. (Color online) Enhancement of power factor (curves with filled symbols) by optimal electron filtering with the optimal cut-off energy shown in Fig. 2 as a function of Fermi level for $r = -0.5$ (circles) and 0.5 (squares) at 300 K. The bulk power factors (curves with open symbols) with no electron filtering are also shown for comparison. All the power factor values are normalized to the maximum bulk power factor for each r.

Figure 4 shows the enhancement of power factor by the optimal electron filtering as a function of Fermi level in comparison with the bulk power factor for two scattering parameters, $r = -0.5$ and 0.5. For the calculation of the optimal power factors, the optimal cut-off energy obtained for each Fermi level shown in Fig. 2 was applied, while no cut off was assumed for the bulk power factor, i.e., $E_C = 0$. For each r, the power factor values were normalized to the maximum bulk power factor in Fig. 4. For a lower r, the peak of the bulk power factor occurs at a lower Fermi level because the Seebeck coefficient drops more rapidly with Fermi level for a lower r. The Seebeck coefficient decreases more slowly with Fermi level for a higher r, so that the peak bulk power factor lies at a relatively higher Fermi level. The optimal power factor increases almost linearly with Fermi level for $r = -0.5$, and achieves a factor of 3.5 enhancement over the maximum bulk value at the Fermi level $\sim$0.2 eV, and more than a factor of 4.5 enhancement at the Fermi level $\sim$0.3 eV. The optimal power factor increases with Fermi level for $r = 0.5$ as well, but at a slower rate than for $r = -0.5$ at the low Fermi level region due to the smaller bulk values. However, as the Fermi level goes further up, the increasing rate for $r = 0.5$ becomes faster than that for $r = -0.5$, and achieves more than a factor of 6.5 enhancement over the maximum bulk value at the Fermi level $\sim$0.3 eV. In general, a material with a lower scattering parameter can achieve a larger power factor enhancement in the low Fermi level region by the optimal energy filtering condition, but if the Fermi level is high enough, a larger power factor enhancement is possible for a higher scattering parameter material.

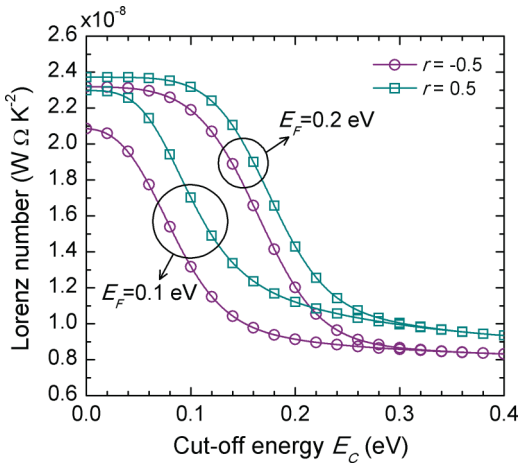

FIG. 5. (Color online) Variation of the Lorenz number by energy filtering as a function of cut-off energy for two different Fermi level $E_F = 0.1$ and 0.2 eV, and two different scattering parameter $r = -0.5$ and 0.5 at 300 K.

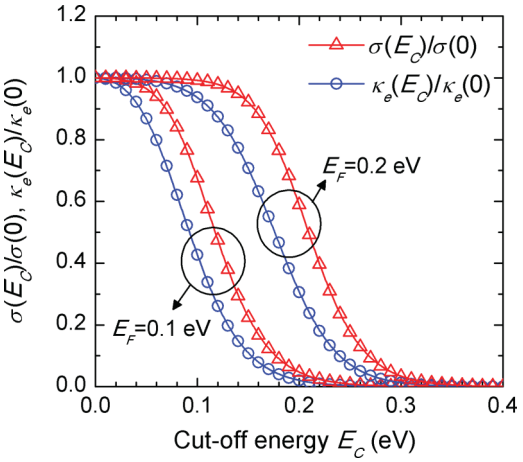

FIG. 6. (Color online) Reduction of the electrical conductivity (σ) and the electronic thermal conductivity (κ_e) from their bulk values by energy filtering as a function of cut-off energy for two different Fermi level $E_F = 0.1$ and 0.2 eV, and $r = -0.5$ at 300 K.

IV. LORENZ NUMBER, ELECTRONIC THERMAL CONDUCTIVITY, FIGURE OF MERIT, AND HALL FACTOR

Figure 5 shows the Lorenz number given by (12) as a function of E_C for two different scattering parameter $r = -0.5$ and 0.5, and two different Fermi levels $E_F = 0.1$ and 0.2 eV, at 300 K. Lorenz number rapidly decreases when E_C falls into the energy range near the Fermi level where most of the electrons participating in conduction is distributed, i.e., in the Fermi window. When E_C goes much higher than the Fermi level, the variation of Lorenz number with E_C becomes small, and L saturates to a certain value that depends on the scattering parameter. For example, Lorenz number converges to 0.82×10^{-8} W Ω K^{-2} for $r = -0.5$, and to 0.90×10^{-8} W Ω K^{-2} for $r = 0.5$ at 300 K regardless of the Fermi level when the cut-off energy level is very high. At higher temperatures, the high cut-off energy limit of Lorenz number slightly increases. For example, at 600 K, the Lorenz number saturates to 0.88×10^{-8} W Ω K^{-2} for $r = -0.5$, and to 1.04×10^{-8} W Ω K^{-2} for $r = 0.5$.

At the optimal cut off, Lorenz number already becomes significantly lower than the bulk value. As shown in Fig. 5, $L = 1.0 \times 10^{-8}$ W Ω K^{-2} when the optimal cut-off energy level $E_{C,\mathrm{opt}} = 0.23$ eV is used for $E_F = 0.2$ eV and $r = -0.5$ at 300 K, which is about 57% lower than the bulk value 2.3×10^{-8} W Ω K^{-2}. At 600 K, this Lorenz number at optimal cut off slightly increases to 1.06×10^{-8} W Ω K^{-2} with $E_{C,\mathrm{opt}} = 0.27$ eV, which is about 49% lower than the bulk value 2.09×10^{-8} W Ω K^{-2}.

Electronic thermal conductivity is lowered from its bulk value by electron energy filtering since the carrier density contributing to the thermal conduction is reduced by the filtering, just like the carrier density contributing to the electrical conduction is reduced. At the same time, the reduction of Lorenz number by filtering causes the electronic thermal conductivity to further decrease. Figure 6 shows the electrical conductivity and the electronic thermal conductivity normalized to their bulk values as a function of E_C for $r = -0.5$ at 300 K. It is clearly seen that the electronic thermal conductivity decreases more rapidly than the electrical conductivity does as E_C goes up. At the optimal cut off ($E_{C,\mathrm{opt}} = 0.23$ eV) for $E_F = 0.2$ eV, κ_e is reduced to merely a 13% of the bulk value, while σ becomes a 30% of the bulk value. The Lorenz number has been reduced to a 43% of the bulk value in this case, which accounts for the difference between the reduction factors of κ_e and σ such that (13%) = (30%) × (43%).

Since the electronic thermal conductivity is suppressed significantly by the optimal filtering to about a 10%–20% of the bulk value, the lattice thermal conductivity can dominate the total thermal conductivity in most cases when the optimal cut off is applied. In this case, the total thermal conductivity in the denominator of ZT becomes almost a constant, and thus the optimal cut-off energy that maximizes the power factor also becomes optimal for maximizing ZT. However, it is still possible that the electronic thermal conductivity dominates over the lattice one in those materials in which the lattice thermal conductivity is enormously suppressed by effective phonon scatterings by nanostructuring. In this case, since the electronic thermal conductivity decreases more rapidly than the electrical conductivity with increasing cut-off energy level, ZT will keep increasing as the cut-off energy further increases. Then, the upper limit of ZT will be determined when the lattice thermal conductivity limits the decreasing rate of the total thermal conductivity with E_C to be slower than that of the power factor.

Figure 7 shows the Hall factor as a function of cut-off energy for varying scattering parameter and $E_F = 0.1$ eV at 300 K. Typically the Hall factor is very close to unity for degenerate semiconductors having the scattering parameters between −0.5 and 0.5. For the materials having the absolute value of scattering parameter larger than 0.5, the Hall factor can deviate much from unity as shown in Fig. 7. As E_C increases, the Hall factor decreases from the bulk values, and converges to unity regardless of the scattering parameter.

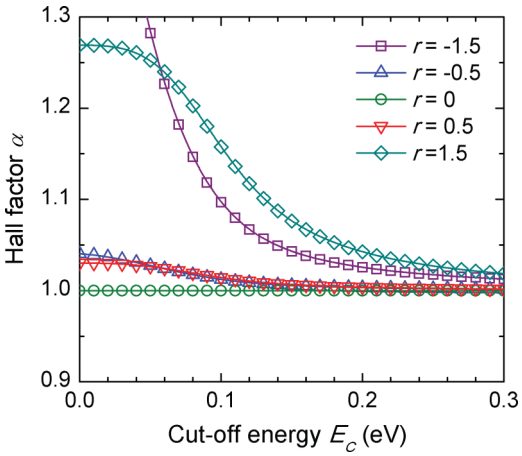

FIG. 7. (Color online) Hall factor as a function of cut-off energy for varying scattering parameter with $E_F = 0.1$ eV at 300 K.

V. POWER FACTOR ENHANCEMENT IN PbTe

So far we have assumed a parabolic band and a simple energy dependency of the scattering time with a constant scattering parameter for the calculations of optimal cut-off energy, power factor, and other transport properties. In real materials such as PbTe, however, the band structure can be highly nonparabolic, and the scattering time is a complicated function of energy since several different scattering mechanisms are involved in electron scattering, such as acoustic and optical phonon scatterings, and impurity scattering, all of which have different energy dependencies.

PbTe and PbTe-based alloys have long been known as important thermoelectric materials for power generation applications.[22] Recently, nanostructured PbTe-based materials have shown improved figures of merit.[4,23–26] PbTe has the conduction band minima at the L valleys in the Brillouin zone, and the band structure is strongly nonparabolic and anisotropic. We used a single-band model for the n-type PbTe with 4 degeneracy of the L valleys based on the modified Kane model accounting for the nonparabolic band and the anisotropic effective masses.[27,28] Temperature-dependent effective masses from Ref. 28 are used in our calculations. Major scattering mechanisms for PbTe such as acoustic/optical phonon deformation potential scattering, and polar optical phonon scattering have been included for the energy-dependent scattering time. We have also included the screened Coulomb impurity scattering and the short range vacancy deformation potential scattering, which can be important at very high doping densities.[27] These band parameters and the energy-dependent scattering time were plugged in (3) and (4) to calculate the power factors with cut offs in n-type PbTe.

Figure 8 shows the Seebeck coefficient, electrical conductivity, electronic thermal conductivity, Lorenz number, and the maximized power factor of n-type PbTe by the optimal electron energy filtering as a function of carrier density at 300 K. The optimal cut-off energy level used at each carrier density is plotted in Fig. 8(c) along with the bulk Fermi level. The bulk properties of PbTe are also plotted together for comparison. As shown in Fig. 8(a), the Seebeck coefficient is greatly enhanced by the optimal filtering, while the reduction

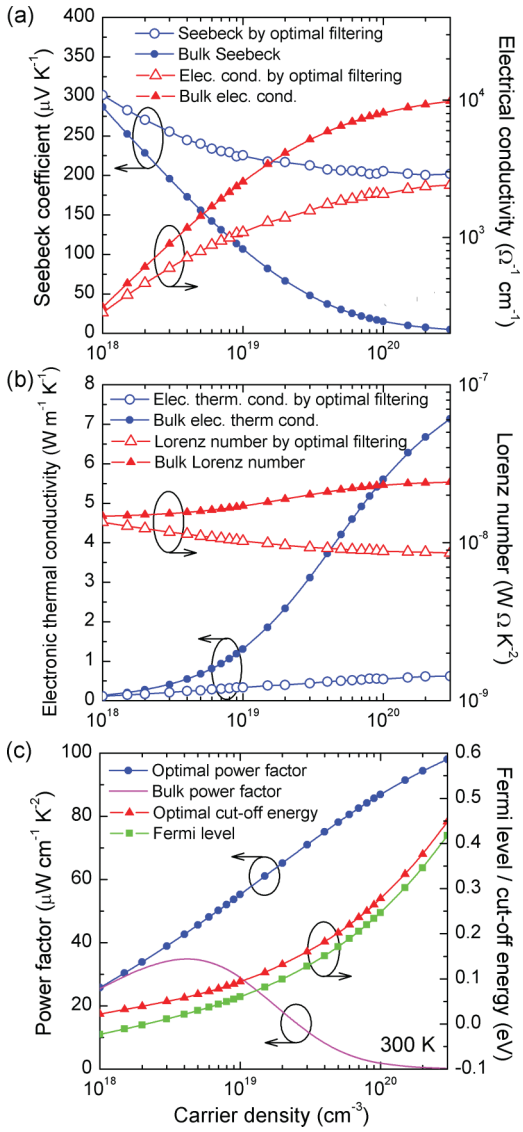

FIG. 8. (Color online) (a) Seebeck coefficient and electrical conductivity, (b) electronic thermal conductivity and Lorenz number, (c) optimized power factor of n-type PbTe by the optimal filtering as a function of carrier density at 300 K, in comparison with the bulk counterparts. The optimal cut-off energy used and the Fermi level as a function of carrier density are also shown on the right y axis in (c).

of the electrical conductivity is relatively small. The Seebeck coefficient is maintained very high to be $\sim$200 μV/K even at very high carrier densities by the optimal filtering. Moreover, the electronic thermal conductivity is also maintained very low $\sim$0.5 W/m K at the high carrier densities due to both the reduced electrical conductivity and Lorenz number ($\sim$0.8 W Ω K^{-2}) by the energy filtering. This low electronic thermal conductivity allows one to push the carrier density to a very high level in this material to achieve a large ZT enhancement by the energy filtering without worrying about significant electronic thermal conductivity rising.

Bulk n-type PbTe has the maximum power factor of 35 μW/cm K^2 at around 4×10^{18} cm^{-3} carrier density at 300 K. If the optimal cut-off energy is applied for electron filtering at high carrier densities, the power factor can be

enhanced significantly to achieve higher than 80 μW/cm K^2 at 1×10^{20} cm^{-3} carrier density or higher. The optimized power factor will keep increasing with carrier density as long as the cut-off energy is optimized at each carrier density.

Since the cut-off energy cannot be easily altered once it is fixed in a real material, one can use Fig. 8(c) to find the optimal carrier density for a given cut-off energy and the resulting maximum power factor in n-PbTe at 300 K. Since the region where the power factor is largely enhanced is quite localized in both cut-off energy and carrier density spaces, optimizing carrier density for a given cut-off energy is equivalent to optimizing cut-off energy for a given carrier density. They are in one-to-one correspondence. For example, the resulting optimal power factor for a 0.2 eV cut-off energy is found from Fig. 8(c) to be 78 μW/cm K^2 at 5×10^{19} cm^{-3} carrier density, which is about a 120% enhancement of power factor over the bulk maximum.

As predicted in the previous section, the optimal cut-off energy for PbTe is found to be a few $k_B T$ above the Fermi level as shown in Fig. 8(c). At low carrier density where the Fermi level is close to the conduction band minimum, the difference between the optimal cut-off energy and the Fermi level is about 40 meV in PbTe. At higher carrier densities, the difference is reduced slightly to 32 meV. These values correspond to the case of $r = -0.5$ at 300 K in Fig. 2, implying that the acoustic/optical phonon deformation potential scattering is a dominant scattering mechanism in PbTe at 300 K. However, the power factor enhancement in PbTe shown in Fig. 8(c) is lower than the prediction shown in Fig. 4 for $r = -0.5$. For example, about a 210% enhancement was predicted for 0.17 eV Fermi level (with 0.2 eV cut-off energy) in Fig. 4, while a 120% enhancement from the maximum bulk power factor of 35 to 78 μW/cm K^2 was obtained for the same cut-off energy level in n-PbTe as shown in Fig. 8(c). This reduction in power factor enhancement is due to the large nonparabolicity in the PbTe band structure.

Figure 9 shows the similar variation of the transport properties with carrier density and the power factor enhancement results by the optimal filtering in n-PbTe at 600 K. The Seebeck coefficient remains high ~200 μV/K, and the electronic thermal conductivity remains low ~0.5 W/m K by the optimal filtering at high carrier density region at 600 K. These values did not change much from the values obtained by the optimal filtering at 300 K in the high carrier density limit, which implies that the condition of optimal filtering reduces the dependency of these two properties on scattering and temperature significantly at very high carrier densities. However, the electrical conductivity by the optimal filtering at 600 K is about 2.5 times lower than that at 300 K at the carrier densities higher than 1×10^{20} cm^{-3} mainly due to the reduced mobility by stronger phonon scatterings at higher temperatures. As a result, the optimal power factor at 600 K is about 2.5 times lower than that at 300 K at the high carrier densities. Still, the optimal power factor keeps increasing with carrier density. A power factor of 35.6 μW/cm K^2 can be achieved at 7×10^{19} cm^{-3} by the optimal filtering with a 0.2 eV cut-off energy, and a power factor of 43.2 μW/cm K^2 at 3×10^{20} cm^{-3} with a 0.4 eV cut-off energy in PbTe at 600 K, which correspond to, respectively, 55% and 88% enhancements over the maximum bulk power factor of 23 μW/cm K^2 at this temperature.

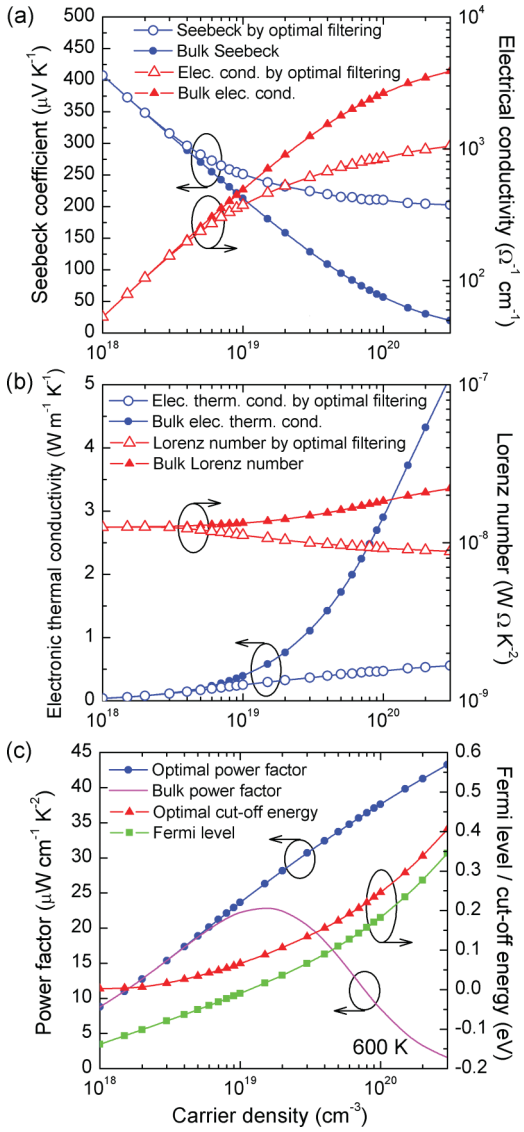

FIG. 9. (Color online) (a) Seebeck coefficient and electrical conductivity, (b) electronic thermal conductivity and Lorenz number, (c) optimized power factor of n-type PbTe by the optimal filtering as a function of carrier density at 600 K, in comparison with their bulk counterparts. The optimal cut-off energy used and the Fermi level as a function of carrier density are also shown on the right y axis in (c).

Since the electronic thermal conductivity remains low (~0.5 W/m K) and the power factor keeps increasing with increasing carrier density by the adjusted cut-off energy, it is possible to achieve very high figure of merits at high carrier densities by energy filtering. Assuming the lattice thermal conductivity to be 1.0 W/m K at 600 K for PbTe, the total thermal conductivity becomes 1.45 W/m K with the electronic thermal conductivity of 0.45 W/m K, and the power factor becomes 35.6 μW/cm K^2 at carrier density 7×10^{19} cm^{-3} when a cut-off energy level of 0.2 eV is used. These altogether make $ZT \sim 1.48$ at 600 K. If a 0.4 eV cut-off energy level is used at carrier density 3×10^{20} cm^{-3}, then the resulting total thermal conductivity of 1.55 W/m K and power factor of 43.2 μW/cm K^2 give $ZT \sim 1.66$ at 600 K. For comparison, the bulk n-type PbTe has the maximum $ZT \sim 0.97$ at

carrier density 8×10^{18} cm^{-3} at 600 K. This optimal carrier density for maximum ZT, 8×10^{18} cm^{-3}, is lower than the optimal carrier density for maximum power factor, which is 1.5×10^{19} cm^{-3}, because the electronic thermal conductivity increased quite significantly from 0.3 to 0.6 W/m K when the carrier density increases from 8×10^{18} to 1.5×10^{19} cm^{-3}.

VI. DISTRIBUTED RESONANT SCATTERINGS

One question we now have is, how can we realize the electron energy filtering in real materials. One possible way to suppress electron transport in a certain energy range is to put an extensive amount of scatterings in that particular energy range, so that electrons experience very slow mobility within the energy region and thus are predominantly prevented from contributing to conduction. Scatterings cannot completely block the electron transport in reality because they cannot be infinitely strong. But a sufficient selectivity in energy can mimic the ideal filtering. Another important requirement for electron filtering is a sharp cut off in energy selection. Since most of the electrons that participate in conduction are near the Fermi level, only a sharp edge near the Fermi level in transport can effectively achieve strong energy selection and thus significant Seebeck enhancement. The often-used Mott formula given below by (24) also backs this requirement,[2]

$$S = \frac{\pi^2}{3}\frac{k_B^2 T}{e}\left.\frac{d\ln[\sigma_d(E)]}{dE}\right|_{E=E_F}, \tag{24}$$

which states that the Seebeck coefficient can be significantly enhanced if the slope of the differential conductivity with respect to energy at the Fermi level is very large. The slope of the differential conductivity can be increased when there is a sharp edge in the scattering time, the group velocity, or the density of states. It is noted that (24) is derived from the general formula given by (4) in some limited cases such as at low temperatures where the Fermi window is very narrow in energy.

It has been known that steplike potential offsets within small nanoparticles embedded in bulk materials cannot achieve such a scattering time with a sharp positive slope with respect to energy.[12] These step potential barriers or wells instead create scattering times that have a slow negative slope with respect to energy because the electron wavelengths are too large to see the potential step in such a small spatial region, so the scattering is Rayleigh-like. When nanoparticles are ionized, however, a slowly varying screened Coulomb potential is formed around the nanoparticles, so the electrons in the low energy region are effectively scattered by the spatially wide potential profile. Therefore, the scattering time by ionized nanoparticles can have a positive slope in energy, but the slope is still not sharp enough to be used for electron filtering. Instead, they can replace the ionized impurity scattering to enhance the electrical conductivity.[12]

One can find energetically sharp scattering times from resonances. Recently it has been reported that core-shell structured nanoparticles can have sharp resonant scattering times at the quasibound energy states formed inside the core region.[13] The width and position of the resonant scattering dip can be controlled by the well depth in the core region. However, the dips in scattering time were not very deep due to the existence of nonresonant scatterings by potential offsets at the heterointerfaces.

A very deep and narrow resonant scattering time can be obtained by resonant impurities. In the 1950s, Friedel found that transition metal impurities can induce a sharp increase of resistivity in a small energy region by their resonant electron scattering.[29] According to the paper, the Coulomb potential due to the ionization of d orbitals in transition metal atoms is strongly screened by the covalent electrons, which can shift up the bound states into the conduction band, and make them resonate with the free electron states in the band. Since the phase shift of electrons by the resonant scattering changes more rapidly from 0 to π with energy for the higher azimuthal quantum number l, the broadening decreases as l increases. This may provide another controllability of the resonance width. Group III elements such as Tl, and group IV elements such as Sn have also been found to create resonant levels in the band structures of IV-VI semiconductors,[30] and Bi_2Te_3,[31] respectively. Since the impurities have infinitesimally small sizes, there is no additional nonresonant scattering by potential offsets in contrast to the core-shell nanoparticles, and thus the scattering time can be approximated by the bell-shaped Lorentzian function given by[32]

$$\tau_{\rm res}(E) = \tau_{0,\rm res}\left[1 + \left(\frac{E - E_{\rm res}}{\gamma/2}\right)^2\right], \tag{25}$$

where $E_{\rm res}$ is the resonant energy level of the impurity, γ is the full width at half maximum (FWHM) of the resonance, and $\tau_{0,\rm res}$ is the minimum scattering time reached at $E = E_{\rm res}$. The parameter $\tau_{0,\rm res}$ is inversely proportional to the concentration of the resonant impurities. This sharp and deep resonant impurity scattering is a good candidate for creating a sharp cut-off edge near the Fermi level. Also, if different kinds of resonant impurities with different resonant energy levels are put together with appropriate depths and widths and distributed in energy at appropriate positions in a bulk material, a suppression of electron transport over a wide energy range along with a sharp cut off is possible. A single sharp resonant dip may not be effective in modifying the transport at room temperature or higher, because its width in energy is typically much narrower than that of the Fermi window at high temperatures, so that the major portion of the distribution of conducting electrons is not affected by the resonance. By using multiple resonant dips distributed in energy that cover the half width of the Fermi window, the most of the conducting electrons distributed in the lower-energy half of the Fermi window below the Fermi level could be filtered out to enhance the power factor at high temperatures.

In Fig. 10 we show two examples of using distributed resonant scatterings as a means to partially realize the electron filtering. With five 5 meV wide resonances evenly spaced between 0.1 and 0.2 eV as shown in Fig. 10(a), most of the low-energy half of the Fermi window at 300 K can be effectively covered to suppress electron transport in that energy range. There are local peaks of scattering time between the resonances where the scattering time is much higher than the minimum, such that electron filtering is weak there. One can make the resonances wider to suppress the local peaks in scattering time, except for the far right resonance at the highest

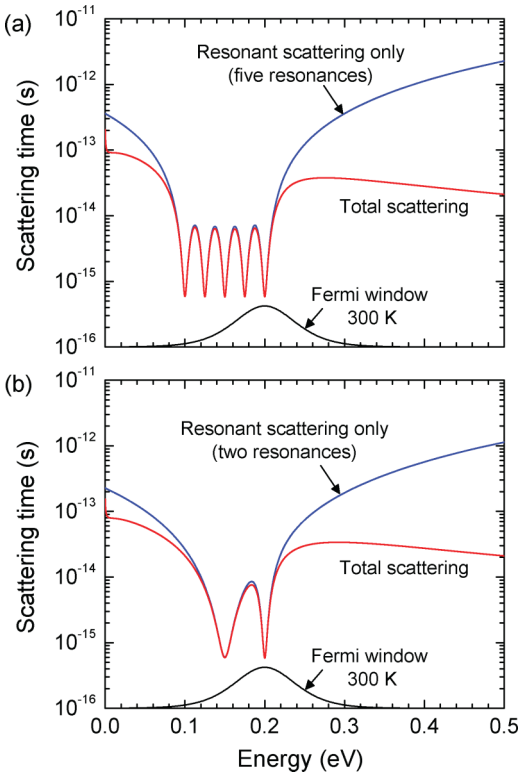

FIG. 10. (Color online) Proposed use of distributed resonant scatterings as a means to realize the electron energy filtering by suppressing electron transport in the left half of the Fermi window at 300 K. (a) Five resonances with resonant centers evenly spaced between 0.1 and 0.2 eV, widths of 5 meV, and minimum scattering times at 6×10^{-16} s and (b) two resonances with resonant centers at 0.15 and 0.2 eV, widths of 15 and 5 meV, respectively, and minimum scattering times at 6×10^{-16} s. The total scattering includes all other relevant scattering times such as phonon scatterings in PbTe.

energy which defines the sharp cut off of energy filtering near the Fermi level. However, if the resonances are too wide, the high-energy shoulder of the sharp edge can be affected by the long tails of the broad resonances, such that the electron filtering effect is weakened.

In addition, using only two resonances, partial electron filtering is possible as shown in Fig. 10(b). In this case, the first resonance at the lower energy region can be wider (γ = 15 meV) than the second resonance at the higher energy region (γ = 5 meV). The spacing between the two is selected to be larger (50 meV) than that in the case of five resonances (25 meV) to cover as much space as possible with a fewer number of resonances. Depths of the resonant scatterings are selected carefully to achieve the largest power factor enhancements. Typically the depths of scattering dips need to be sufficiently deep, about two orders of magnitude lower than the background scattering time. They should not be too deep, however, since the long tail of the scattering time can lower the high-energy shoulder of the sharp edge at cut off, such that the slope of the scattering time with respect to energy becomes smaller near the cut-off energy.

Figure 11 shows the resulting Seebeck coefficient, electrical conductivity, and power factor in n-type PbTe at 300 K by the two cases of distributed resonant scatterings described in

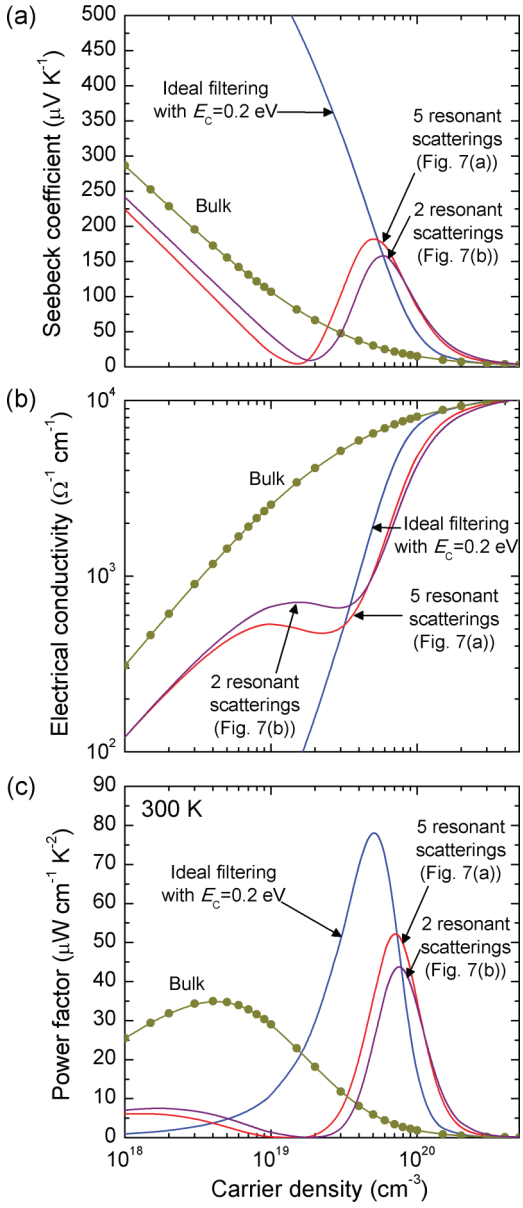

FIG. 11. (Color online) (a) Seebeck coefficient, (b) electrical conductivity, and (c) power factor of n-type PbTe at 300 K as a function of carrier density for the two cases of distributed resonant scatterings described in Fig. 10. For comparison, the properties by an ideal electron filtering with a 0.2 eV cut off, and the bulk values (curve with filled circles) are also presented.

Fig. 10 in comparison with the case of an ideal electron filtering with a fixed cut off at 0.2 eV, and with the bulk values. Unlike the ideal filtering case, the distributed resonant scattering cases may have lower Seebeck coefficients than the bulk at low carrier density region as shown in Fig. 11(a). When the Fermi level is positioned below the lowest-energy resonance, the high-energy electrons above the Fermi level will be scattered strongly by the resonances. As a result, the average transport energy is lowered than that of the bulk, and the Seebeck coefficient is reduced. As the carrier density changes from low to high, the Fermi level increases accordingly; more of the low-energy electrons below the Fermi level fall into the energy window affected by the resonances, experiencing significant

mobility suppression, while more of the high-energy electrons above the Fermi level get unaffected by the resonances staying out of the resonance-affected energy region. As a result, the average transport energy rises back up and the Seebeck coefficient increases, which is the consequence of the energy filtering. As the Fermi level goes further up beyond the optimal cut-off energy, then the Seebeck coefficient decreases again and finally converges to the bulk value.

The mobility suppression is largest when the Fermi level falls in the middle of the resonant energy window, so that the largest portion of conducting electrons experience the resonant scatterings. This leads to a local dip in the electrical conductivity curve for the two distributed resonant scattering cases plotted in Fig. 11(b). At high carrier density region, the electrical conductivity for the ideal filtering case is slightly higher than those for the distributed resonant scattering cases, since the right-hand side of the highest-energy resonance in the latter cases extends beyond 0.2 eV, while electron energy is cut off exactly at 0.2 eV for the ideal filtering case.

The maximum power factor of 52 μW/cm K^2 can be obtained for the five resonant scatterings, and 43 μW/cm K^2 for the two resonant scatterings both at 7.6×10^{19} cm^{-3} carrier density, which correspond to 50% and 25% enhancements, respectively, over the maximum bulk value of 35 μW/cm K^2 at 300 K. These enhancements are lower than the 120% enhancement obtained by ideal electron filtering with a similar cut-off energy at 0.2 eV. This is due to the incomplete suppression of electron transport by the resonant scatterings, in particular, at the local peaks of scattering time between the resonances. These uncovered energy regions are wider when a smaller number of resonant scatterings are used. As a result, the power factor enhancement can be lower for a smaller number of resonant scatterings. The Fermi window is wider at higher temperature, proportional to T. Hence, a larger number of wider resonant scatterings might be needed to sufficiently cover the lower-energy half of the Fermi window at higher temperatures.

One important note is, if the scattering spectrum is significantly modified to form a band pass or stop, the density of states and dispersion curves can be modified as well. For example, Heremans *et al.*[10] used a 2% atomic fraction of Tl impurities in PbTe to enhance the power factor by the modification of the density of states. In order to utilize the resonant scatterings without a significant modification of the band structure, we may need to use a lower amount of impurities than 2%. In the future, the amount of impurities that must be used to achieve sufficiently strong electron filtering should be determined experimentally. When the density of states is modified, the local effective mass at the energy region is also increased, which in turn, may, simultaneously, reduce the mobility and electrical conductivity. Thus special care might be necessary to achieve a large power factor enhancement at high impurity concentrations.

VII. CONCLUSIONS

We have investigated the electron energy filtering effect in bulk materials where the lateral momentum conservation is no longer an issue. For a parabolic band and a simple energy-dependent electron momentum scattering time represented by a single constant scattering parameter, we have quantified the optimal cut-off energy and the maximum power factor enhancement factor as functions of the Fermi level, scattering parameter, and temperature. The variations of the electronic thermal conductivity, Lorenz number, and Hall factor were also investigated under the energy filtering. This theory can also be applied to real materials with nonparabolic bands and complicated energy-dependent scattering time such as PbTe. It is found that the nonparabolicity can reduce the power factor enhancement factor, but the optimal cut-off energy is quite similar to that obtained for a parabolic band. We have also proposed the use of distributed resonant scatterings of Lorentzian shapes to partially realize the electron filtering effect in real materials. A 50% power factor enhancement is achieved by five narrow resonant scatterings properly distributed in energy space in PbTe at 300 K. With a lower number of resonant scatterings, the power factor enhancement can be smaller.

APPENDIX A: CHARACTERISTICS OF $F_S(E_C)$ AND $g(E_C)$

The $F_S(E_C)$ and $g(E_C)$ functions are defined by (8) and (9), respectively. In particular $F_S(0)$ with $E_C = 0$ can be rewritten using the complete Fermi-Dirac integral defined by

$$\mathcal{F}_j(\eta) = \frac{1}{\Gamma(j+1)} \int_0^\infty \frac{x^j}{\exp(x-\eta)+1} dx, \tag{A1}$$

where Γ is the gamma function, so that

$$F_s(0) = \left(\frac{3}{2}+r+s\right)(k_B T)^{\frac{3}{2}+r+s}\Gamma \times \left(\frac{3}{2}+r+s\right)\mathcal{F}_{\frac{1}{2}+r+s}\left(\frac{E_F}{k_B T}\right). \tag{A2}$$

From (8) and (9),

$$g(E_C)-E_C = \frac{\left(\frac{5}{2}+r\right)\int_{E_C}^\infty E^{\frac{3}{2}+r} f_0(E)dE + E_C^{\frac{5}{2}+r} f_0(E_C)}{\left(\frac{3}{2}+r\right)\int_{E_C}^\infty E^{\frac{1}{2}+r} f_0(E)dE + E_C^{\frac{3}{2}+r} f_0(E_C)} - E_C = \frac{\int_{E_C}^\infty E^{\frac{1}{2}+r}\left[\left(\frac{5}{2}+r\right)E - \left(\frac{3}{2}+r\right)E_C\right] f_0(E)dE}{\left(\frac{3}{2}+r\right)\int_{E_C}^\infty E^{\frac{1}{2}+r} f_0(E)dE + E_C^{\frac{3}{2}+r} f_0(E_C)}. \tag{A3}$$

Since all the integrands are positive, and $(\frac{5}{2}+r)E - (\frac{3}{2}+r)E_C > 0$ for $E \geqslant E_C$ and $r \geqslant -3/2$, therefore,

$$g(E_C) > E_C \quad \text{for} \quad \forall E_C \geqslant 0. \tag{A4}$$

Differentiating (8),

$$\begin{aligned} F_s'(E_C) &= \frac{d}{dE_C}\left[\left(\frac{3}{2}+r+s\right)\int_{E_C}^\infty E^{\frac{1}{2}+r+s} f_0(E)dE\right] + \frac{d}{dE_C}\left(E_C^{\frac{3}{2}+r+s} f_0(E_C)\right) \\ &= -\left(\frac{3}{2}+r+s\right)E_C^{\frac{1}{2}+r+s} f_0(E_C) + \left(\frac{3}{2}+r+s\right)E_C^{\frac{1}{2}+r+s} f_0(E_C) + E_C^{\frac{3}{2}+r+s} f_0'(E_C) \\ &= E_C^{\frac{3}{2}+r+s} f_0'(E_C). \end{aligned}$$

Thus,

$$F_s'(E_C) = E_C^{\frac{3}{2}+r+s} f_0'(E_C). \tag{A5}$$

Since $f_0' < 0$ and $E_C \geqslant 0$, from (A5),

$$F_s' \leqslant 0, \tag{A6}$$

where the equality holds only when $E_C = 0$.

Also, from (A5),

$$\frac{F_1'(E_C)}{F_0'(E_C)} = \frac{E_C^{\frac{5}{2}+r} f_0'(E_C)}{E_C^{\frac{3}{2}+r} f_0'(E_C)} = E_C. \tag{A7}$$

The derivative of $g(E_C)$ is given from (A7) by

$$\begin{aligned} g' &= \frac{F_1'F_0 - F_1F_0'}{F_0^2} = \frac{F_0'}{F_0}\left(\frac{F_1'}{F_0'} - \frac{F_1}{F_0}\right) \\ &= \frac{F_0'}{F_0}(E_C - g). \end{aligned} \tag{A8}$$

From (A6), $F_0' \leqslant 0$ (equality when $E_C = 0$), $F_0 > 0$, and $(E_C - g) < 0$ from (A4). Therefore,

$$g'(E_C) \geqslant 0, \tag{A9}$$

where the equality holds only when $E_C = 0$.

Now, check $g(0)$: From (A2),

$$g(0) = \frac{\left(\frac{5}{2}+r\right)(k_BT)\Gamma\left(\frac{5}{2}+r\right)\mathcal{F}_{\frac{3}{2}+r}\left(\frac{E_F}{k_BT}\right)}{\left(\frac{3}{2}+r\right)\Gamma\left(\frac{3}{2}+r\right)\mathcal{F}_{\frac{1}{2}+r}\left(\frac{E_F}{k_BT}\right)}.$$

Since $\Gamma(j+1) = j\Gamma(j)$,

$$g(0) = \frac{\left(\frac{5}{2}+r\right)(k_BT)\mathcal{F}_{\frac{3}{2}+r}\left(\frac{E_F}{k_BT}\right)}{\mathcal{F}_{\frac{1}{2}+r}\left(\frac{E_F}{k_BT}\right)}. \tag{A10}$$

It is known for the complete Fermi-Dirac Integral $\mathcal{F}_j(\eta)$ that

$$\mathcal{F}_j(\eta) \to \frac{\eta^{j+1}}{\Gamma(j+2)} \quad \text{as} \quad \eta \to \infty.$$

Thus, from (A10),

$$\begin{aligned} \lim_{E_F\to\infty}[g(0) - E_F] &= \lim_{E_F\to\infty} \frac{\left(\frac{5}{2}+r\right)(k_BT)\mathcal{F}_{\frac{3}{2}+r}\left(\frac{E_F}{k_BT}\right)}{\mathcal{F}_{\frac{1}{2}+r}\left(\frac{E_F}{k_BT}\right)} - E_F \\ &= \frac{\left(\frac{5}{2}+r\right)(k_BT)\frac{\left(\frac{E_F}{k_BT}\right)^{\frac{5}{2}+r}}{\Gamma\left(\frac{7}{2}+r\right)}}{\frac{\left(\frac{E_F}{k_BT}\right)^{\frac{3}{2}+r}}{\Gamma\left(\frac{5}{2}+r\right)}} - E_F \\ &= E_F - E_F = 0. \end{aligned}$$

Therefore,

$$\lim_{E_F\to\infty}[g(0) - E_F] = 0. \tag{A11}$$

Since $g(0)$ asymptotically approaches E_F as E_F increases to infinity, and $g(0) - E_F$ decreases monotonically with increasing E_F,

$$g(0) > E_F \quad \text{for} \quad \forall E_F. \tag{A12}$$

Therefore, by (A9) and (A12),

$$g(E_C) > E_F \quad \text{for} \quad \forall E_F \text{ and } E_C \geqslant 0. \tag{A13}$$

APPENDIX B: OPTIMAL CUT-OFF ENERGY

The Seebeck coefficient with a cut-off energy E_C is given by (6), $S(E_C) = C_2[g(E_C) - E_F]$. For convenience, we removed a negative sign in (6), so that the Seebeck is positive for n-type semiconductors. The bulk Seebeck coefficient is $S(0) = C_2[g(0) - E_F]$ with $E_C = 0$. By (A12), $S(0)$ is always positive.

Since $g(E_C)$ increases monotonically with E_C according to (A9),

$$S(E_C) = C_2[g(E_C) - E_F] \geqslant S(0) = C_2[g(0) - E_F]. \tag{B1}$$

Thus, the Seebeck coefficient is enhanced by a positive cut-off energy E_C, and keeps increasing with E_C as $g(E_C)$ increases with E_C.

The electrical conductivity is given by (5), $\sigma(E_C) = C_1 \cdot F_0(E_C)$. According to (A6), F_0 monotonically decreases with increasing E_C. Therefore, the electrical conductivity decreases with increasing E_C.

The power factor $S^2\sigma$ is given by (7) to be

$$S^2\sigma = C_1C_2\frac{[F_1(E_C) - E_FF_0(E_C)]^2}{F_0(E_C)}.$$

The power factor is maximized when its derivative is equal to zero, such that

$$\begin{aligned} \frac{dS^2\sigma}{dE_C} &= \frac{C}{e^2T^2}\frac{d}{dE_C}\frac{\left(F_1^2 - 2E_FF_1F_0 + E_F^2F_0^2\right)}{F_0} \\ &= \frac{C}{e^2T^2}\frac{2F_1F_1'F_0 - 2E_FF_1'F_0^2 + E_F^2F_0^2F_0' - F_1^2F_0'}{F_0^2} \\ &= 0. \end{aligned}$$

Thus we get

$$2F_1F_1'F_0 - 2E_FF_1'F_0^2 + E_F^2F_0^2F_0' - F_1^2F_0' = 0. \tag{B2}$$

Here we apply (A7) $F_1'/F_0' = E_C$ to (B2) to find

$$\left(\frac{F_1}{F_0} - E_C\right)^2 = (E_F - E_C)^2. \tag{B3}$$

The solution to (B3) is

$$\frac{F_1}{F_0} = E_F \quad \text{or} \quad \frac{F_1}{F_0} = 2E_C - E_F.$$

However, by (A13), the first solution cannot be accepted. Therefore,

$$\frac{F_1}{F_0} = g(E_C) = 2E_C - E_F. \tag{B4}$$

The optimal cut-off energy $E_{C,\text{opt}}$ that maximizes the power factor is the one that satisfies (B4).

*jbahk@purdue.edu

[1] L. E. Bell, Science **321**, 1457 (2008).
[2] A. Shakouri, Annu. Rev. Mater. Res. **41**, 399 (2011).
[3] R. Venkatasubramanian, E. Siivola, T. Colpitts, and B. O'Quinn, Nature (London) **413**, 597 (2001).
[4] K. F. Hsu, S. Loo, F. Guo, W. Chen, J. S. Dyck, C. Uher, T. Hogan, E. K. Polychroniadis, and M. G. Kanatzidis, Science **303**, 818 (2004).
[5] B. Poudel, Q. Hao, Y. Ma, Y. C. Lan, A. Minnich, B. Yu, X. Yan, D. Z. Wang, A. Muto, D. Vashaee *et al.*, Science **320**, 634 (2008).
[6] J. M. O. Zide, J.-H. Bahk, R. Singh, M. Zebarjadi, G. Zeng, H. Lu, J. P. Feser, D. Xu, S. L. Singer, Z. X. Bian *et al.*, J. Appl. Phys. **108**, 123702 (2010); **110**, 059902(E) (2011).
[7] L. D. Hicks and M. S. Dresselhaus, Phys. Rev. B **47**, 12727 (1993).
[8] J. R. Sootsman, D. Y. Chung, and M. G. Kanatzidis, Angew. Chem., Int. Ed. **48**, 8616 (2009).
[9] C. J. Vineis, A. Shakouri, A. Majumdar, and M. G. Kanatzidis, Adv. Mater. **22**, 3970 (2010).
[10] J. P. Heremans, V. Jovovic, E. S. Toberer, A. Saramat, K. Kurosaki, K. Charoenphakdee, S. Yamanaka, and J. F. Snyder, Science **321**, 554 (2008).
[11] M. Zebarjadi, G. Joshi, G. Zhu, B. Yu, A. Minnich, Y. Lan, X. Wang, M. Dresselhaus, Z. Ren, and G. Chen, Nano Lett. **11**, 2225 (2011).
[12] J.-H. Bahk, Z. Bian, M. Zebarjadi, P. Santhanam, R. Ram, and A. Shakouri, Appl. Phys. Lett. **99**, 072118 (2011).
[13] J.-H. Bahk, P. Santhanam, Z. Bian, R. Ram, and A. Shakouri, Appl. Phys. Lett. **100**, 012102 (2012).
[14] G. D. Mahan, J. Appl. Phys. **76**, 4362 (1994).
[15] A. Shakouri and J. E. Bowers, Appl. Phys. Lett. **71**, 1234 (1997).
[16] A. Shakouri, C. Labounty, P. Abraham, J. Piprek, and J. E. Bowers, Mater. Res. Soc. Proc. **545**, 449 (1999).
[17] D. Vashaee and A. Shakouri, Phys. Rev. Lett. **92**, 106103 (2004).
[18] R. Kim, C. Jeong, and M. S. Lundstrom, J. Appl. Phys. **107**, 054502 (2010).
[19] E. Flage-Larsen and O. M. Løvvik, in *Thermoelectrics and Its Energy Harvesting: Materials, Preparation, and Characterization in Thermoelectrics*, edited by D. M. Rowe (CRC, Boca Raton, FL, 2012), Chap. 10.
[20] R. A. Smith, *Semiconductors*, 2nd ed. (Cambridge University Press, London, 1979).
[21] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1949).
[22] G. S. Nolas, J. Sharp, and H. J. Goldsmid, *Thermoelectrics: Basic Principles and New Materials Developments* (Springer, New York, 2001).
[23] T. C. Harman, P. J. Taylor, M. P. Walsh, and B. E. LaForge, Science **297**, 2229 (2002).
[24] P. F. P. Poudeu, J. D'Angelo, A. D. Downey, J. L. Short, T. P. Hogan, and M. G. Kanatzidis, Angew. Chem., Int. Ed. **45**, 3835 (2006).
[25] J. Androulakis, C.-H. Lin, H.-J. Kong, C. Uher, C.-I. Wu, T. Hogan, B. A. Cook, T. Caillat, K. M. Paraskevopoulos, and M. G. Kanatzidis, J. Am. Chem. Soc. **129**, 9780 (2007).
[26] K. Biswas, J. He, I. D. Blum, C.-I. Wu, T. P. Hogan, D. N. Seidman, V. P. Dravid, and M. G. Kanatzidis, Nature (London) **489**, 414 (2012).
[27] D. I. Bilc, S. D. Mahanti, and M. G. Kanatzidis, Phys. Rev. B **74**, 125202 (2006).
[28] C. J. Vineis, T. C. Harman, S. D. Calawa, M. P. Walsh, R. E. Reeder, R. Singh, and A. Shakouri, Phys. Rev. B **77**, 235202 (2008).
[29] J. Friedel, Can. J. Phys. **34**, 1190 (1956).
[30] S. Ahmad, K. Hoang, and S. D. Mahanti, Phys. Rev. Lett. **96**, 056403 (2006).
[31] M. K. Zhitinskaya, S. A. Nemov, and T. E. Svechnikova, Phys. Solid State **40**, 1297 (1998).
[32] Y. I. Ravich, in *CRC Handbook of Thermoelectrics*, edited by D. M. Rowe (CRC, Boca Raton, FL, 1995), Chap. 7.