2012

On the best bandstructure for thermoelectric performance: A Landauer Perspective

Changwook Jeong
Birck Nanotechnology Center, Purdue University, jeong.changwook@gmail.com

Raseong Kim
Birck Nanotechnology Center, Purdue University, kim369@purdue.edu

Mark S. Lundstrom
Purdue University, lundstro@purdue.edu

On the best bandstructure for thermoelectric performance: A Landauer perspective

Changwook Jeong, Raseong Kim,[a] and Mark S. Lundstrom
Network for Computational Nanotechnology, Birck Nanotechnology Center, Purdue University, West Lafayette, Indiana 47907, USA

(Received 26 March 2012; accepted 2 May 2012; published online 7 June 2012)

The question of what bandstructure produces the best thermoelectric device performance is revisited from a Landauer perspective. We find that a delta-function transport distribution function (TDF) results in operation at the Mahan-Sofo upper limit for the thermoelectric figure-of-merit, ZT. We show, however, the Mahan-Sofo upper limit itself depends on the bandwidth (BW) of the dispersion, and therefore, a finite BW dispersion produces a higher ZT when the lattice thermal conductivity is finite. Including a realistic model for scattering profoundly changes the results. Instead of a narrow band, we find that a broad BW is best. The prospects of increasing ZT through high valley degeneracy or by distorting the density-of-states are discussed from a Landauer perspective. We conclude that while there is no simple answer to the question of what bandstructure produces the best thermoelectric performance, the important considerations can be expressed in terms of three parameters derived from the bandstructure—the density-of-states, $D(E)$, the number of channels, $M(E)$, and the mean-free-path, $\lambda(E)$. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4727855]

I. INTRODUCTION

The performance of thermoelectric (TE) devices is related to a dimensionless figure of merit, ZT,

$$ZT = \frac{S^2\sigma T}{\kappa_{ph} + \kappa_{el}}, \tag{1}$$

where S is the Seebeck coefficient, σ the electrical conductivity, κ_{ph} the lattice thermal conductivity, and κ_{el} the electronic thermal conductivity. Early work developed TE technology with a figure of merit of about one,[1] but subsequent progress was stalled for several decades. Recent progress has, however, been significant, and there are now several reports of ZTs above one,[2] which have been largely achieved by reducing the lattice thermal conductivity, which dominates the denominator of Eq. (1). Figure 1 is a plot of ZT vs. $\kappa_{tot}(=\kappa_{ph}+\kappa_{el})$ for several different TE materials including recent materials with $ZT > 1$.[3–22] Also shown (dashed line) is the result that would be obtained if the power factor (PF) ($S^2\sigma$) of each material was the same as that of silicon. The conclusion is that the performance of a thermoelectric material is largely determined by its thermal conductivity. The power factors of good TE materials are all similar. This raises the question of what controls the magnitude of the power factor and provides an opportunity to further increase ZT by power factor engineering. This paper addresses the question: "How is the electronic structure of a material related to its power factor?"

For conventional TE materials with approximately parabolic energy bands, the power factor is well understood.[23,24] High power factors require high mobility to increase σ, and ionized impurity scattering should dominate to enhance S.[24] In a seminal paper, Mahan and Sofo asked the question: "What shape of a bandstructure would produce the highest thermoelectric performance?" They concluded that materials with a δ-function "transport distribution function," (TDF) would be best.[25] Subsequently, Nishio and Hirano[26] showed that in the absence of thermal conduction by the lattice, a single energy channel leads to "electronic efficiencies" at the Carnot limit. Similar conclusions were reached by Humphrey and Linke.[27] In a recent paper, Nakpathomkun *et al.* argued that the power delivered to a load is the important measure of performance and that for such purposes, ZT is not the best figure of merit.[28] Nakpathomkun concluded that the TDF should have a finite bandwidth (BW $\sim 2.25\ k_BT$) for maximum power output, although the maximum efficiency (for $\kappa_{ph} = 0$) still occurs for a δ-function TDF. The "best bandstructure question" has also been explored recently by Fan *et al.*[29] who concluded that for a normalized TDF (i.e., the area under the TDF vs. energy curve is bounded), the δ-function TDF is best, but for a bounded TDF (the maximum value is limited), a narrow but finite width of the TDF is best. Very recently, Zhou *et al.* considered the optimal BW question and concluded that the existence of an optimal BW depends strongly on the scattering model used.[30] Finally, we note the recent experiments reporting improved TE performance in materials with a resonant level[8] and in PbTeSe alloys that display a high degree of valley degeneracy,[31] which increases the density-of-states, $D(E)$, near the Fermi level.

To continue to increase performance of thermoelectric materials, the electronic performance must be enhanced.[2,32] Previous study has clarified several of issues, but a number of questions remain:

(1) What physical constraints should be placed on the TDF? Before we explore the best bandstructure question, the physical constraints to be placed on the TDF must be clarified.

[a]Current address: Components Research, Intel Corporation, Hillsboro, Oregon 97124, USA.

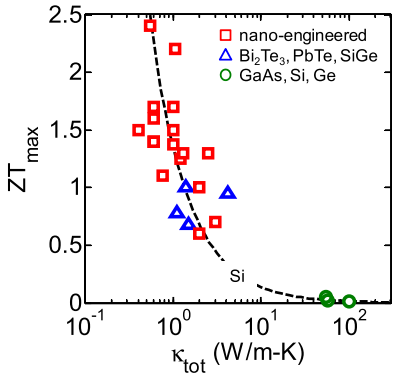

FIG. 1. The maximum ZT vs. total κ. Conventional semiconductors used in IC industry (circle): Si 300 K,[3,4] Ge 300 K,[5] GaAs 300 K;[21] bulk TE materials used in TE devices (triangles): Bi_2Te_3 300 K,[6] $Bi_xSb_{2-x}Te_3$ 300 K,[7] $Si_{80}Ge_{20}$ 1275 K,[22] PbTe 300 K;[8] nano-engineered materials (square): Si nanowire (NW) 300 K,[9] Si NW 200 K,[10] Bi_2Te_3/Sb_2Te_3 superlattice (SL) 300 K,[11] PbTe/PbSeTe quantum dot (QD) SL 300 K,[12] Tl-PbTe 773 K,[8] $Na_{1-x}Pb_mSb_yTe_{2+m}$ 650 K,[13] $Pb_xSn_{1-x}Te$-PbS 642 K,[14] $AgPb_mSbTe_{m+2}$ 800 K,[15] PbSbTe 700 K,[16] p-PbTe-SrTe 800 K,[17] Bi_2Te_3 300 K,[18] $Bi_xSb_{2-x}Te_3$ 475 K,[7] n-$Si_{80}Ge_{20}$ 1275 K,[22] Si 1275 K,[3] n-LaTe 1273 K,[19] n-$Yb_xCo_4Sb_{12+y}$ 800 K.[20] The dashed line is the ZT that would be obtained if the power factor of each material were the same as that of silicon.

(2) How does the BW of the TDF affect TE performance? Under what conditions is a δ-function TDF the best and under what conditions is a narrow band the best? When a narrow band is best, what determines the optimum BW? Finally, how do these results relate to the original arguments of Mahan and Sofo?
(3) How does scattering affect optimum bandstructure?
(4) How should the improved performance of materials with a high valley degeneracy[31] or with a resonant energy levels[8] that distort the density-of-states be understood?
(5) Is there a best bandstructure for TE performance?

Our goal in this paper is to answer these questions. We use a Landauer approach, which is equivalent to the Boltzmann transport equation for crystalline semiconductors in the diffusive limit but has advantages of mathematical simplicity and physical transparency. This is most apparent with regard to the so-called transport distribution, a central quantity in thermoelectric theory[25] whose physical interpretation is unclear. In the Landauer approach, the transport distribution acquires a clear physical interpretation—it is proportional to the number of channels available for conduction times the mean-free-path (MFP) for backscattering, which makes it easy to identify the appropriate physical constraints to place on the TDF.

Following Nakpathomkun *et al.*,[28] we shall assess thermoelectric performance using two different metrics: (1) the maximum thermoelectric efficiency and (2) the maximum power that a thermoelectric generator delivers to a load. The first is of theoretical interest and the second of practical interest. As discussed in the Appendix, as ZT approaches infinity, the maximum thermoelectric efficiency approaches the Carnot efficiency, but the efficiency when the maximum power is delivered to a load approaches one-half of the Carnot limit, the so-called Curzon-Ahlborn limit.[33] Both operating conditions will be considered.

The paper is organized as follows: Section II summarizes the approach. The expressions presented are those of standard thermoelectric theory with only one difference—the transport distribution is expressed in Landauer form. Section III A is a short discussion of the single energy case (a δ-function TDF). This section sets the stage for understanding the subsequent results and relates this paper to some previous studies. Section III B is a short discussion of one-dimensional (1D) thermoelectrics. The simplicity of the 1D problem provides a clear illustration of how the number of channels for conduction, $M(E)$, is related to the density-of-states, $D(E)$, and allows us to address question (1) above. Question (2) is discussed in Secs. III C and III D. Section III C examines how the BW of the dispersion affects TE performance, extending the analysis of Sec. III B to three-dimensional (3D) thermoelectrics. In Sec III D, it is shown that the conclusion of Mahan and Sofo is correct, if properly understood. Section III E is a discussion of scattering and addresses question (3), and Secs. III F and III G address question (4). The paper's conclusions are summarized in Sec. IV, where our perspective on question (5) is presented.

II. APPROACH

A. Thermoelectric coefficients

We begin with a brief review of the Landauer approach[34] to TE transport. The TE transport parameters are[5,25]

$$\sigma = \int_{-\infty}^{+\infty} dE\sigma'(E), \tag{2a}$$

$$S = \left(\frac{k_B}{q}\right)\int_{-\infty}^{+\infty} dE(E-E_F)\frac{\sigma'(E)}{\sigma}, \tag{2b}$$

$$\kappa_0 = T\left(\frac{k_B}{q}\right)^2\int_{-\infty}^{+\infty} dE(E-E_F)^2\sigma'(E), \tag{2c}$$

$$\kappa_{el} = \kappa_0 - S^2\sigma T, \tag{2d}$$

where E_F is the Fermi level and $\sigma'(E)$ is the so-called differential conductivity. For 3D bulk diffusive materials,

$$\sigma'(E) = \frac{2q^2}{h}\left(\frac{M(E)}{A}\right)\lambda(E)\left(-\frac{\partial f_0}{\partial E}\right) = q^2\Sigma(E)\left(-\frac{\partial f_0}{\partial E}\right), \tag{3}$$

where $2q^2/h$ is the quantum of conductance, $M(E)$ is the number of conducting channels at a given energy, E, A is the cross-sectional area, $\lambda(E)$ is the mean-free-path for backscattering, and f_0 is the Fermi-Dirac distribution. In Eq. (3), $\Sigma(E)$ is the so-called TDF, which arises from a solution to the Boltzmann transport equation.[25] The TDF depends on both bandstructure and scattering. In the Landauer approach, $\Sigma(E)$ is proportional to the product of $M(E)$, which depends only on bandstructure and $\lambda(E)$, which depends on bandstructure and the scattering physics. For semiclassical transport in the diffusive limit, the Boltzmann and Landauer approaches are mathematically identical; we use the

Landauer approach in this paper because it provides a simple and clear physical interpretation of the TE transport distribution function.

B. Bandstructure model

Given a bandstructure, $E(k)$, $M(E)$ is easily obtained by simply counting the bands that cross the energy of interest.[5] For the purposes of this paper, we seek a simple but realistic bandstructure model that gives $E(k)$ across the entire Brillouin zone (BZ) and for which the BW of the dispersion can be varied from broad to narrow in order to explore the effects of bandwidth on the TE coefficients. A simple, nearest neighbor tight-binding (TB) model,

$$E(k) = 2t_0(1-\cos k_x a) + 2t_0(1-\cos k_y a) + 2t_0(1-\cos k_z a), \tag{4}$$

will be used. In Eq. (4), $t_0 = \hbar^2/2m_e a^2$ with a and m_e being the lattice constant and the effective electron mass, respectively. The BW of the electron dispersion is $12t_0$. We change the BW while assuming $a = 5\times10^{-10}$ m, which keeps the total number of states fixed. It is important to note that the TB model gives $E(k)$ across the entire BZ and that we do not assume parabolic energy bands (i.e., $E(k) \neq \hbar^2k^2/2m_e$). In the case of a large BW, however, only states near the bottom of the band, which are nearly parabolic, are occupied, and we recover the expected results for parabolic energy bands. For the small BW case, however, the TDF approaches a δ-function, and much different results are obtained. Because the TDF is derived from a physically sensible dispersion, no artificial constraints are placed on the TDF.

III. RESULTS AND DISCUSSIONS

A. Single energy case

We begin with a short discussion of the single energy case (a δ-function TDF), which has received a good deal of attention and forms one end of the spectrum of BWs that we will explore. When all the channels are at $E=E_0$, $M(E)=M_0{}^*\delta(E-E_0)$, and the differential conductivity becomes $\sigma'(E)=\sigma_0\delta(E-E_0)$. In this case, the thermoelectric coefficients become

$$\sigma = \sigma_0, \tag{5a}$$

$$S = \left(-\frac{k_B}{q}\right)\left(\frac{E_0-E_F}{k_BT}\right), \tag{5b}$$

$$\kappa_0 = T\left(\frac{k_B}{q}\right)^2\left(\frac{E_0-E_F}{k_BT}\right)^2\sigma_0, \tag{5c}$$

$$\kappa_{el} = \kappa_0 - S^2\sigma_0 T = 0. \tag{5d}$$

Here, with a constant MFP (λ_0) being assumed, σ_0 is found to be

$$\sigma_0 = \frac{2q^2}{h}\left(\frac{M_0\lambda_0}{k_BT}\right)\frac{e^{\eta_F}}{(1+e^{\eta_F})^2}, \tag{5e}$$

where $\eta_F = (E_F - E_0)/k_BT$ and $\eta_F \approx \pm 2.4$ for the maximum power factor. The result, Eq. (5e), agrees with Mahan and Sofo[25] but not with Zhou *et al.*[30] who found $\sigma_0 = 0$.

For the single energy case, the electronic heat conductivity, κ_{el}, is zero. This occurs because κ_{el} defines the heat flow under open-circuit conditions. If all the current flows at $E=E_0$, then zero current means that no electrons are flowing, so there can be no heat current. For the single energy case,

$$\frac{\kappa_{el}}{\sigma} = LT = 0, \tag{5f}$$

which shows that the Lorenz number, L, is zero. For a parabolic energy band under strongly degenerate conditions, $L = (\pi^2/3)(k_B/q)^2$, but we shall see that as the BW of the dispersion decreases, L decreases and approaches Eq. (5f) in the limit of zero bandwidth.

B. One-dimensional analysis

Here, we illustrate how the number of conducting channels, $M(E)$, is related to the density-of-states, $D(E)$, by using a simple 1D example that illustrates the physical constraint that should be imposed on the TDF. Figures 2(a) and 2(b) show a plot of the 1D dispersion, $E(k) = 2t_0(1-\cos k_x a)$, and the corresponding $D(E)$ for BWs of ~0.1 and ~0.6 eV. At a given energy, E, the number of states that participate in transport is the number of conducting channels, $M(E)$, which is often referred to as the number of (transverse) modes in analogy with the modes of an electromagnetic waveguide. Given an accurate dispersion, $M(E)$ can be readily computed by counting the bands that cross the energy of interest[5] and is shown in Fig. 2(c). It can be seen that although the $D(E)$ goes to infinity, $M(E)$ remains bounded, independent of bandwidth. (Zhou *et al.*[30] also pointed out that while $D(E)\to\infty$, $\Sigma(E)$ remains finite.) Note that $\int D(E)\,dE$ is independent of bandwidth because we fix the total number of states, but it is the peak value of $M(E)$ that is independent of bandwidth, not $\int M(E)\,dE$. This 1D example demonstrates that the answer to the first question posed in Sec. I is that for a given dispersion, the maximum of $M(E)$ is fixed. Fixing

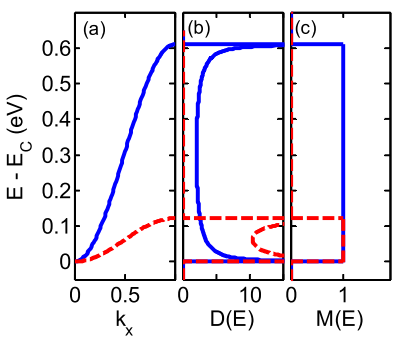

FIG. 2. (a) The 1D dispersion, $E(k) = 2t_0(1-\cos k_x a)$ with two different BWs of a dispersion, ~0.1 (dashed line) and ~0.6 eV (solid line). (b) The corresponding the density-of-states, $D(E)$, and (c) the number of conducting channels, $M(E)$, which represent the number of states that participate in transport at a given energy E. Note that although the $D(E)$ goes to infinity, $M(E)$ remains bounded independent of bandwidth.

the electrical conductivity when varying the BW (Ref. 27) or fixing the area under TDF vs. E (Ref. 29) lead to non-physical results.

The counting bands method can be extended to the two-dimensional (2D) and the 3D cases. The 1D procedure is repeated for each transverse wave vector so that the entire BZ of the material is spanned. The resulting number of conducting channels is integrated over transverse momentum at a given energy to find the $M(E)$. This method is used next in 3D. Finally, we also note that while not obvious, $M(E)$ is related to the density-of-states according to[5]

$$M(E) = \frac{h}{2}\langle v_x^+(E)\rangle D(E), \tag{6}$$

where $\langle v_x^+(E)\rangle$ is the average velocity in the direction of transport at energy, E, and $D(E)$ is the density of states per spin. The density of states is per unit length in 1D, per unit area in 2D, and per unit volume in 3D. The number of channels is a number in 1D, a number per unit width in 2D, and a number per unit area in 3D, where the width and cross sectional area are normal to the direction of current flow.

C. Three-dimensional analysis: Constant mean-free-path

In this section, we extend our analysis to 3D and evaluate the TE coefficients at $T = 300$ K to address question (2). The BW of the dispersion is varied from very narrow to very wide while assuming a constant MFP. Figure 3 shows the computed density-of-states, $D(E)$, and number of channels, $M(E)$, for small and large BW dispersions. The $M(E)$ characteristics display a peak value of ~0.6 times the number of atoms in the cross section—independent of bandwidth. As was observed for 1D, the total number of states (area under the $D(E)$ curve) is independent of BW but the area under the $M(E)$ curve depends on BW. Finally, Fig. 3 also shows that the parabolic band assumption (dashed line) matches the full-band TB results (solid line) only near the bottom of the band where $D(E) \propto E^{1/2}$ and $M(E) \propto E$.[5]

Next, TE performance for 3D bulk is assessed for two different conditions: (1) the maximum TE efficiency and (2) the maximum power that a thermoelectric generator delivers to a load.[25] The load resistance and the location of the Fermi level are co-optimized in order to extract the maximum efficiency or the maximum power output. For each of the two different operating conditions, the efficiency and power output are calculated as a function of the BW.

We first evaluate TE performance for zero lattice thermal conductivity, $\kappa_{ph} = 0$, and the results are shown in Fig. 4(a), where the TE efficiency is normalized by the Carnot efficiency, η_C. For this case, the maximum power output is obtained for a moderate BW band, but the maximum efficiency occurs for a δ-function like narrow band. In agreement with Nakpathomkun,[28] we find that as the TDF (or $M(E)$) approaches a δ-function, the maximum efficiency approaches the Carnot efficiency, but no useful power can be delivered to a load. However, the δ-function TDF does produce a finite power under maximum power conditions with an efficiency of one-half the Carnot efficiency, the Curzon-Ahlborn limit.[33]

Figure 4(b) shows the results with the more realistic case, i.e., a finite $\kappa_{ph} = 0.5$ W/m K, which is about 2–3 times

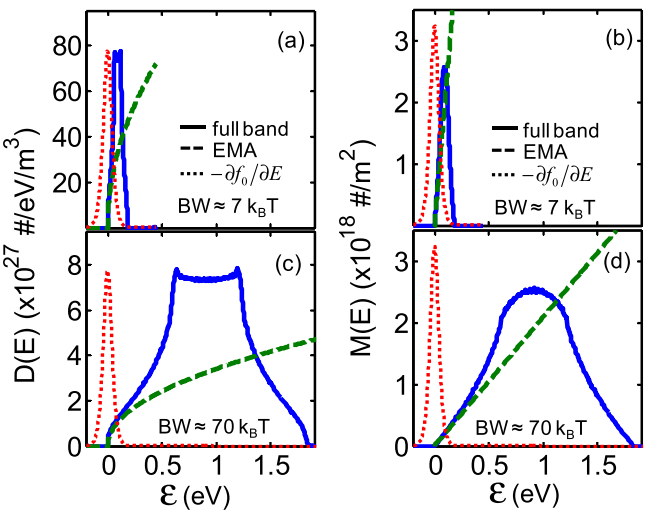

FIG. 3. (a) and (c) The 3D density-of-states, $D(E)$, for the narrow and the broad BW bands. (b) and (d) The number of conducting channels, $M(E)$, for the narrow and the broad BW bands. Full-band calculations (solid line) are compared to effective mass approximation (EMA, dashed line). Based on the EMA, $D(E) = m_e\sqrt{2m_e(E-E_C)}/2\pi^2\hbar^3$ and $M(E) = m_e(E-E_C)/2\pi\hbar^2$, where E_C is the band edge. Fitted effective masses (m_e) at the bottom of band are $m_e = m_0$ for the broad band and $m_e = 10\,m_0$ for the narrow band, where m_0 is the electron rest mass. It is seen that the $D(E)$ and the $M(E)$ obtained from parabolic band assumption (dashed line) match well the full-band TB results (solid line) only at the bottom of the band. Dotted line is the arbitrarily normalized "window function," $W = (-\partial f_0/\partial E)$, where f_0 is Fermi-Dirac distributions. For horizontal axis, $\varepsilon = E - E_C$ for $D(E)$ and $M(E)$ and $\varepsilon = E - E_F$ for $W = (-\partial f_0/\partial E)$, where E_F is the Fermi level. In Fig. 3, we assume $E_C = E_F$ which is a typical condition for optimum performance.

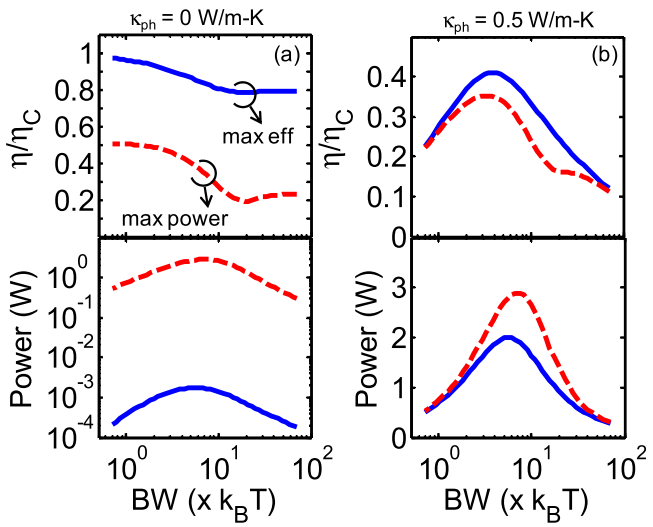

FIG. 4. Efficiency normalized by Carnot efficiency η_C (upper panel) and power for (a) zero lattice thermal conductivity, $\kappa_{ph} = 0$, and (b) a finite lattice thermal conductivity, $\kappa_{ph} = 0.5$ W/m K are plotted as a function of the BW. Efficiency and power are evaluated from two different perspectives. Solid line: condition for the maximum thermoelectric efficiency. Dashed line: condition for the maximum power that a thermoelectric generator delivers to a load. The load resistance and the location of the Fermi level are co-optimized in order to extract the maximum efficiency or the maximum power output.

smaller than the lattice thermal conductivity of Bi_2Te_3. In contrast to the case of $\kappa_{ph} = 0$, the maximum efficiency now occurs for a moderate BW, instead of for the narrowest BW. Note that the maximum power occurs for a moderate BW for both zero and finite κ_{ph}. If we repeat the calculations using a smaller (larger) value of κ_{ph}, we find only a slight decrease (increase) in the optimum BW. Next, we discuss how the BW affects the four TE transport parameters.

The TE coefficients assuming zero lattice thermal conductivity, $\kappa_{ph} = 0$, are shown in Fig. 5. For each value of the BW, we found the optimal location of the Fermi level to maximize *ZT*. As the BW decreases, *ZT* diverges. Therefore, the highest *ZT* and efficiency is obtained for a δ-function like band. As seen in Figs. 5(b) and 5(c), this occurs mainly because as the BW approaches zero, κ_{el} approaches zero while σ approaches a finite value, Eq. (5e), and therefore the Lorenz number, *L*, approaches zero as discussed in Sec. III A. For the large BW case, Fig. 5(b) shows that we obtain the expected results for a parabolic energy band, i.e., for large BWs, *L* saturates at a value slightly above 2. Recall that for a parabolic energy band under strongly degenerate conditions, $L/(k_B/q)^2 = \pi^2/3$ and for non-degenerate conditions with a constant MFP, $L/(k_B/q)^2 = 2$. The Wiedemann-Franz "law" states that there is a relation between the electrical conductivity and the electronic component of the thermal conductivity, but the specific value of the *L* depends on bandstructure, scattering, and the location of the Fermi level. As noted by Mahan and Bartkowiak,[35] it should be regarded as a "rule of thumb" rather than a law.

In contrast to the case of $\kappa_{ph} = 0$, Fig. 6(a) shows that the highest *ZT* occurs for a moderate BW when $\kappa_{ph} = 0.5$ W/m K.

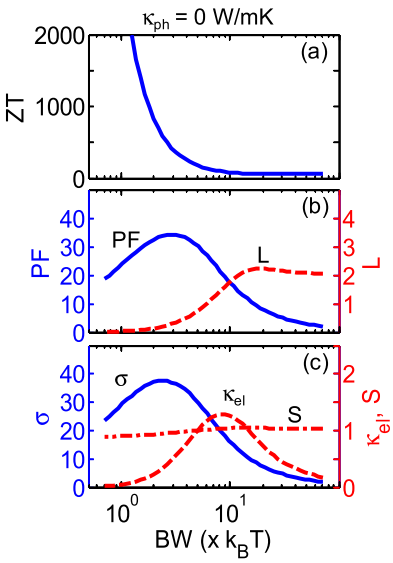

FIG. 5. For zero lattice thermal conductivity ($\kappa_{ph} = 0$), (a) *ZT*, (b) the PF and the Lorenz number (*L*), and (c) the Seebeck coefficient (*S*), the electrical conductivity (σ), and the electronic thermal conductivity (κ_{el}) are plotted. The units of PF, *L*, *S*, σ, and κ_{el} in the plots are 10^{-6} W/m-K^2, $(k_B/q)^2$, 10^{-3}V/K, 1/Ω-m, and 5×10^{-5} W/m-K, respectively. For each value of the BW, we found the optimal location of the Fermi level to maximize *ZT*. As the BW decreases, the highest *ZT* (i.e., efficiency) is obtained for a delta-function like narrow band. This result occurs mainly because as the BW approaches zero, κ_{el} approaches zero and therefore the *L* approaches zero. Note that the maximum PF still appears at a moderate BW.

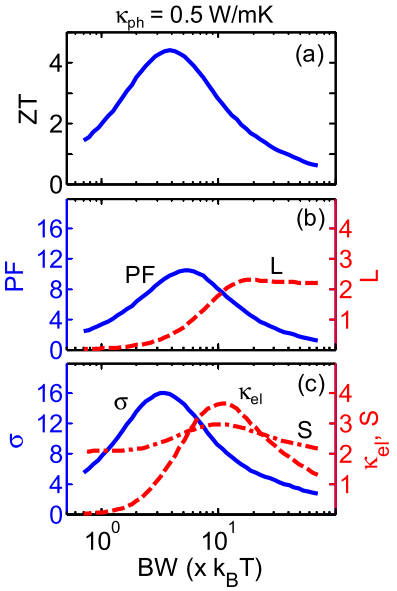

FIG. 6. For a finite lattice thermal conductivity ($\kappa_{ph} = 0.5$ W/m K), (a) the *ZT*, (b) the PF and the Lorenz number (*L*), and (c) the Seebeck coefficient (*S*), the electrical conductivity (σ), and the electronic thermal conductivity (κ_{el}) are plotted. The units of PF, *L*, *S*, σ, and κ_{el} in the plots are 10^{-3} W/m-K^2, $(k_B/q)^2$, 10^{-4}V/K, 10^4/Ω-m, and 10^{-1} W/m-K, respectively. For each value of the BW, we found the optimal location of the Fermi level to maximize *ZT*. In contrast with the case of $\kappa_{ph} = 0$, the highest *ZT* occurs for the moderate BW mainly because of the BW dependence of the power factor. Since the BW has a rather small effect on *S*, the strong variation of σ vs. BW explains the shape of the power factor vs. BW. The optimum BW for σ occurs when the width of the Fermi window matches the width of $M(E)$.

As seen in Fig. 6(b), for this case, the optimum BW for highest *ZT* is mainly determined by the BW dependence of the power factor rather than that of the *L*. Figure 6(c) shows that the BW has a strong effect on σ, but it has a rather small effect on *S*. The stronger variation of σ vs. BW than that of *S* vs. BW explains the shape of the power factor vs. BW characteristic in Fig. 6(b).

The results shown in Fig. 6 can be understood in terms of the width of the Fermi "window function," $(-\partial f_0/\partial E)$, and the distribution of conducting channels, as plotted in Fig. 3(a). The width of the Fermi window function is a few k_BT, so when the bandwidth of the dispersion is less than this value, σ decreases. The optimum BW for σ occurs when the width of the Fermi window matches the BW of $M(E)$. As the BW of the dispersion increases, the channels are more spread out, so given the finite width of the Fermi window function, a decreasing fraction of the channels can participate in electrical conduction, and σ decreases (compare Figs. 3(b) and 3(d)). Note that the peak of κ_{el} occurs for a somewhat larger BW than that of σ because of the $(E - E_F)^2$ factor in Eqs. (2c) and (2d).

We have discussed why a δ-function like TDF maximizes the efficiency for $\kappa_{ph} = 0$ and why a TDF with a moderate BW of a few k_BT maximizes the efficiency for a finite κ_{ph}. For the maximum power output, however, Fig. 4 shows that a moderate BW is best in either case. This occurs because the power output is proportional to the power factor (as discussed in Appendix), and the power factor displays its maximum at a moderate BW regardless of the value of κ_{ph}.

Achieving a moderate BW band by coherent transport in a superlattice, however, is not an effective approach because in that case, most of the channels are filtered out. Molecular thermoelectrics is another possibility.[36] This might lead to high efficiency but not to high power, because although molecular levels can be sharp (possibly too sharp), one still needs a large number of channels in a small energy range. Packing molecules closely may broaden the levels and degrade performance.

D. The Mahan and Sofo upper limit

In previous sections, we have shown that for a constant MFP, a moderate BW (a few k_BT) is best for the practical case of a finite κ_{ph}. This conclusion holds for both the efficiency and the power output when we consider a constant MFP. This fact has been pointed out in previous studies;[25–30] we have provided a simple, physical explanation in terms of the need to match the width of the Fermi window to the width of the transport distribution or $M(E)$ and also explained the appropriate physical constraints on the TDF. In this section, we address the question of how these theoretical studies relate to the original arguments.

The Mahan-Sofo upper limit to ZT can be readily obtained by using Eqs. (1) and (2d), from which ZT can be written as

$$ZT = \frac{\kappa_0}{\kappa_{ph}} \times \left(\frac{1 - \kappa_{el}/\kappa_0}{1 + \kappa_{el}/\kappa_{ph}} \right). \tag{7a}$$

Since the term in the brackets is always less than 1, it can be seen that

$$ZT \leq \frac{\kappa_0}{\kappa_{ph}}, \tag{7b}$$

which is the Mahan-Sofo upper limit.[25] Mahan and Sofo[25] also showed that a bandstructure that produces a δ-function TDF (a single energy channel) gives the upper limit. This can be readily understood from the fact that $\kappa_{el} = 0$ for the single energy case, so Eq. (7a) shows that the thermoelectric figure of merit reaches its upper limit, $ZT = \kappa_0/\kappa_{ph}$. Figure 7 shows the computed κ_0/κ_{ph} (upper limit of ZT) vs. BW of the dispersion (dashed line) along with the computed ZT vs. BW (solid line). It can be seen that $ZT \leq \kappa_0/\kappa_{ph}$ is always true, and that in agreement with the prediction of Mahan and Sofo,[25] ZT approaches its upper limit for the narrowest BW. Although we assumed a constant MFP and a finite κ_{ph}, we find that the conclusion that $ZT \leq \kappa_0/\kappa_{ph}$ and $ZT = \kappa_0/\kappa_{ph}$ for δ-function TDF are independent of the specific scattering model and value of κ_{ph}. The important point, however, is that κ_0 depends on the BW, so the upper limit itself depends on BW and shows peak value at a BW of a few k_BT where the maximum ZT occurs. The highest ZT, therefore, occurs for a BW that results in operation well below the Mahan-Sofo upper limit.

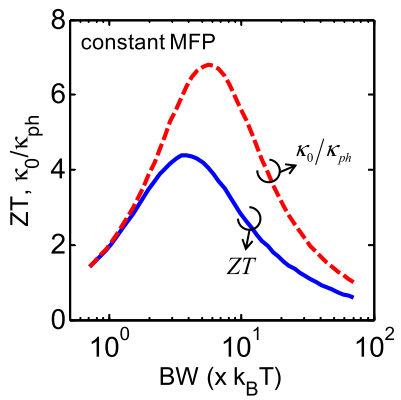

FIG. 7. For a finite lattice thermal conductivity ($\kappa_{ph} = 0.5$ W/m-K), the ZT (solid line) and its upper limit, κ_0/κ_{ph} (dashed line), are plotted as a function of BW. For each assumed bandwidth, the optimal location of the Fermi level is determined to maximize ZT. Here, a constant mean-free-path is assumed. It can be seen that $ZT \leq \kappa_0/\kappa_{ph}$ is always true and the ZT approaches its upper limit for the narrowest BW.

E. Role of scattering

In previous sections, we showed that a narrow TDF with a BW of a few k_BT gives the best TE performance. The only exception is that when $\kappa_{ph} = 0$, the maximum efficiency (but not the maximum power delivered to a load) occurs for a δ-function TDF. We also revisited the Mahan-Sofo limit and showed that while the upper limit is obtained for a δ-function TDF, better efficiency can be obtained by operating below the BW dependent upper limit using a TDF with a BW of a few k_BT. These results answer question (2), but before we conclude that a narrow band is best, however, we should realize that our use of the same MFP for all bandwidths is physically unreasonable. One advantage of the Landauer approach is that it separates the TDF into a part that depends only on bandstructure, $M(E)$, and a part that depends both on bandstructure and scattering physics, the MFP. We turn now to the question of how scattering affects TE performance and we shall see that although narrow TDFs have been much discussed,[25–30] they are probably not the best for TE performance.

Recent work by Zhou *et al.*[30] and Jeong *et al.*[37] has examined three models for scattering: (1) the constant mean-free-path discussed here in previous sections, (2) a constant scattering time, and (3) a scattering rate proportional to the density-of-states, $\tau^{-1}(E) = C_{el}D(E)$. The constant MFP can be justified for parabolic energy bands, but it is hard to justify over a wide range of BWs. The constant scattering time is commonly used, but hard to justify under any circumstances. A scattering rate that is proportional to the density-of-states follows directly from Fermi's Golden Rule and should describe acoustic phonon scattering, which typically dominates for good thermoelectrics.

Extensive calculations for the three scattering models have been presented recently,[30,37] so we only review the conclusions here. As discussed in previous sections, for a constant mean-free-path, ZT is mostly determined by the BW dependence of σ, which is maximized when the BW of the TDF matches that of the Fermi window function. Similar results are obtained for the constant scattering time case. For the most realistic scattering model, however, it is found that there is no optimum BW.[30,37] Instead, ZT continuously decreases as the BW decreases. The reason is clear in a Landauer picture. According to Eq. (3), σ is proportional to the

TDF, which is the Landauer picture proportional to the product of the number of channels, $M(E)$, and the MFP, $\lambda(E)$. As the BW decreases, the number of channels in the Fermi window increases, which should increase the conductivity, but the density of states near the Fermi level also increases, which increases the scattering rate and decreases the MFP. In our isotropic bandstructure model, the smaller BW corresponds to a larger effective mass and smaller velocity. Since the MFP is the product of velocity and scattering time, it decreases faster than $M(E)$ increases so ZT decreases as the BW decreases.

In contrast to several previous studies and to the discussion in earlier sections of this paper, which used overly simplified treatments of scattering, we conclude that for best TE performance, wide (dispersive) bands are the best. This point can also be seen from the expression for the conductivity,

$$\sigma = \int \frac{M(E)}{A}\lambda(E)\left(-\frac{\partial f_0}{\partial E}\right)dE. \tag{8a}$$

Recall that $M(E) \propto |\upsilon_x^+|D(E)$, where $|\upsilon_x^+|$ is the average velocity in the direction of transport at energy, E. Recall also that $\lambda(E) \propto |\upsilon_x^+|\tau(E)$ and that $1/\tau(E) \propto D(E)$, so Eq. (8a) becomes

$$\sigma \propto \int |\upsilon_x^+(E)|^2\left(-\frac{\partial f_0}{\partial E}\right)dE. \tag{8b}$$

Equation (8b) shows that the conductivity is proportional to the square of the average velocity in the Fermi window. High velocities occur for light effective masses (large BWs), so for a realistic model of scattering, we conclude that a wide band, not a narrow band, is best.

F. High valley degeneracy

The analysis in the previous section showed that it is hard to increase the power factor in a single band by increasing the density-of-states near the Fermi level because of the tradeoff between the number of channels and the mean-free-path. It is generally understood, however, that a high degree of valley degeneracy is beneficial for thermoelectric performance,[1] and recently, this approach has produced significant increases in performance.[31] This leads to the question of how high valley degeneracy affects the power factor.

The benefits of valley degeneracy can be understood with a very simple model. As shown in previous sections, the power factor is mainly controlled by the behavior of the conductivity, σ, and when a realistic model for scattering is assumed (proportional to the density of states), large bandwidths, for which the parabolic band assumption holds, are best. Accordingly, we assume two spherical, parabolic band semiconductors, the first with an effective mass of m_1^* and the second with m_2^*. The first semiconductor has a valley degeneracy of N_{V1}, and the second has only a single valley, i.e., $N_{V2} = 1$. We compare these two semiconductors at the same density of states ($D(E)$) and ask "How do the power factors of these two semiconductors with the same densities of states compare?"

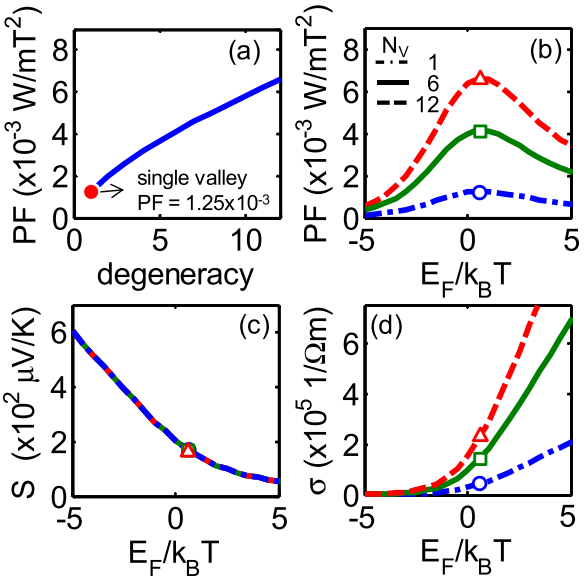

FIG. 8. (a) The computed power factor vs. valley degeneracy is plotted. (b) The power factor (PF), (c) S, and (d) σ are plotted as a function of Fermi level for three cases of $N_V = 1$, 6, and 12. Symbols represent values at optimal Fermi level. It can be seen that PF of multi-valley semiconductor is improved and that the enhancement is attributed to the increase of σ.

The computed power factor vs. valley degeneracy is plotted in Fig. 8(a). The calculations assume $m_2^* = m_0$ and that $\tau(E) = C_{el}/D(E)$ with C_{el} selected to produce average MFP ($\langle\langle\lambda\rangle\rangle$) of 10 nm for a single valley. The calculations confirm the expectation that valley degeneracy produces higher performance. For $N_V = 6$, about a factor of 3 increase of the power factor can be achieved in this model multi-valley structure. Figures 8(b)–8(d) show the power factor, S, and σ as a function of Fermi level for an isotropic single valley ($N_{V2} = 1$) and multi-valley ($N_{V1} = 6$ and $N_{V1} = 12$). It is found that the S vs. E_F characteristics are the same for the three cases, but Fig. 8(d) shows that the conductivity increases with valley degeneracy.

Additional insight into the benefits of valley degeneracy can be gained from Fig. 9, which compares $D(E)$, $M(E)$, $\lambda(E)$, and the transport distribution, $M(E)\lambda(E)$ for the three cases. For this calculation, we forced $D(E)$ to be the same in the three cases (Fig. 9(a)), so the scattering times, $\tau(E)$, are also the same. In the multi-valley cases, we combine the contributions of several light mass bands. In an isotropic single valley, the same density-of-states is achieved by increasing the effective mass, which lowers the velocity. As shown in Fig. 9(b), $M(E)$ is higher for the multiple valley case because $M(E) \propto \upsilon(E)D(E)$. Fig. 9(c) shows that $\lambda(E)$ is also higher for the multiple valley case because $\lambda(E) \propto \upsilon(E)\,\tau(E)$. Note that with a parabolic band, $\lambda(E)$ is energy-independent. Since the transport distribution is proportional to $M(E)\lambda(E)$, it is considerably higher for the multiple valley case, as shown in Fig. 9(d). The $|\upsilon_x^+(E)|^2$ term in Eq. (8b) is larger in multi-valley case. Stated another way, $M(E) \propto \upsilon(E)D(E)$ and $\lambda(E) \propto \upsilon(E)\,\tau(E)$. While $D(E)$ and $\tau(E)$ are the same in the three cases, the velocity is higher in the multi-valley cases, so both $M(E)$ and $\lambda(E)$ are larger when a high density-of-states is obtained by combining light mass valleys.

On the best bandstructure for thermoelectric performance: A Landauer perspective

Changwook Jeong, Raseong Kim, and Mark S. Lundstrom

Citation: Journal of Applied Physics **111**, 113707 (2012); doi: 10.1063/1.4727855
View online: http://dx.doi.org/10.1063/1.4727855
View Table of Contents: http://scitation.aip.org/content/aip/journal/jap/111/11?ver=pdfcov
Published by the AIP Publishing

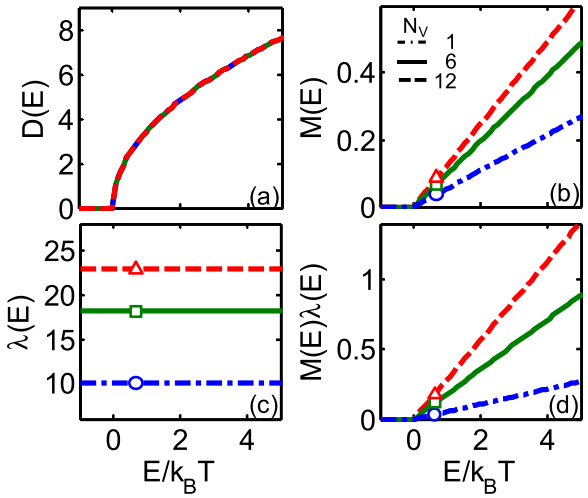

FIG. 9. (a) Density-of-states ($D(E)$), (b) number of conduction channels ($M(E)$), (c) MFP for backscattering ($\lambda(E)$), and (d) $M(E)\lambda(E)$ are plotted for three cases of $N_V = 1$, 6, and 12. Symbols represent values at optimal Fermi level. $M(E)$ and $\lambda(E)$ are higher for the multiple valley case because $M(E) \propto \upsilon(E)D(E)$ and $\lambda(E) \propto \upsilon(E)\,\tau(E)$. Since the transport distribution is proportional to $M(E)\lambda(E)$, it is considerably higher for the multiple valley case.

The improved PF for multiple valleys is due to the high conductivity, $\sigma = (2q^2/h)\langle M\rangle\langle\langle\lambda\rangle\rangle$. The improved power factor is attributed to increases in both the average MFP, $\langle\langle\lambda\rangle\rangle$, and in the number of channels in the Fermi window, $\langle M\rangle$. For example, $\langle M\rangle = 9.5 \times 10^{16}\,\mathrm{m}^{-2}$ and $\langle\langle\lambda\rangle\rangle = 18\,\mathrm{nm}$ for the multi-valley case of $N_V = 6$ and $\langle M\rangle = 5.2 \times 10^{16}\,\mathrm{m}^{-2}$ and $\langle\langle\lambda\rangle\rangle = 10\,\mathrm{nm}$ for the single valley case.

G. Distorted density of states

Next, we examine the possibility of improving TE performance with a distorted density-of-states, $D(E)$.[8,38,39] To illustrate the effect of a distorted $D(E)$, we consider a model semiconductor for which the lower band is an isotropic single valley with an effective mass of m_0 and the upper band has an effective mass of $10\,m_0$. The $10\times$ larger effective mass induces sharp increase of $D(E)$, which is similar to the effect of a resonant level. We compare the power factor of this semiconductor to that of an isotropic single valley with an effective mass of m_0. Two different scenarios for scattering are considered; the first assumes a constant MFP, $\langle\langle\lambda\rangle\rangle$, with a value of 10 nm. The second scenario assumes that $\tau(E) = C_{el}/D(E)$ with C_{el} selected to produce $\langle\langle\lambda\rangle\rangle$ of 10 nm for an isotropic single valley. In practice, we expect the results to lie between these two limits. We compare TE performance at the optimal location of the Fermi level while varying the band-offset, ΔE_C, between the lower and upper bands.

Figure 10(a), the computed power factor vs. ΔE_C for the constant MFP case, shows that the best performance is obtained when $\Delta E_C = 0$. The maximum performance is much better than that of the single, small mass valley, and slightly better than that of a single, large mass valley. Figure 10(b) shows that the maximum PF occurs when the Fermi level is located near the bottom of the large mass valley. Fig. 10(c) shows that a non-monotonic behavior of $S(E_F)$ when $\Delta E_C > 0$ maintains a relatively large Seebeck

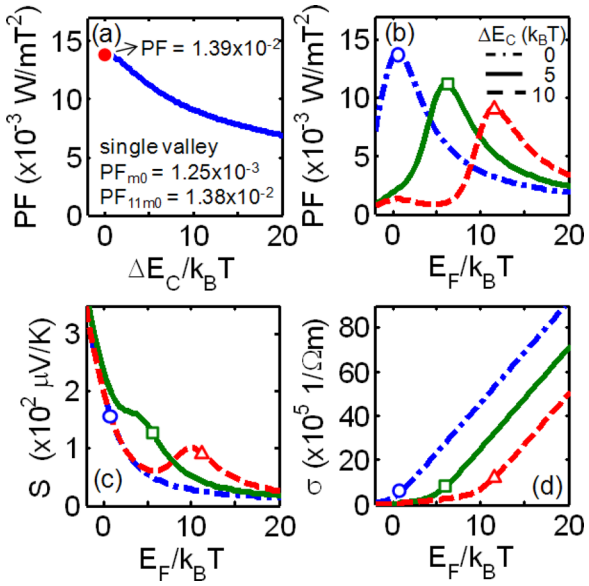

FIG. 10. (a) For constant MFP, the PF vs. ΔE_C for a material with lower band and upper band, where ΔE_C is the band-offset. Lower band is isotropic single valley with an effective mass of m_0 and upper band has an effective mass of $10m_0$. Note that PFs for single valley with an effective mass of m_0 and $11m_0$ are 1.25×10^{-3} W/mT^2 and 1.38×10^{-2} W/mT^2, respectively. Upper band with heavy effective mass produces higher performance regardless of ΔE_C in comparison to the power factor of a single light mass valley and best performance is obtained when $\Delta E_C = 0$. (b)–(d) The PF, Seebeck coefficient (S), and electrical conductivity (σ) vs. Fermi level for three cases of $\Delta E_C = 0$, 5, and $10k_BT$. Comparing to the case of single valley, it is found that a significant increase in σ lead to improved power factor. Non-monotonic Seebeck coefficient behavior (Fig. 10(c)) maintains large S at the degenerate limit. Symbols represent values at optimal Fermi level.

coefficient under degenerate conditions. These results can be understood from Fig. 11. The density of states for three different valley offsets are shown in Fig. 11(a), and Fig. 11(b) shows the corresponding $M(E)$. The case of $\Delta E_C = 0$ produces the largest M at any energy. Because the MFP is constant (Fig. 11(c)), the transport distribution

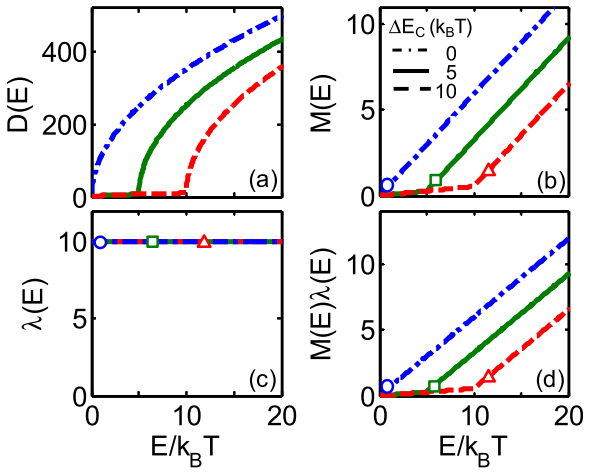

FIG. 11. (a) Density-of-states ($D(E)$), number of conduction channels ($M(E)$), MFP for backscattering ($\lambda(E)$), and $M(E)\lambda(E)$ vs. Fermi level are plotted for three cases of $\Delta E_C = 0$, 5, and $10k_BT$ for constant MFP. Since MFP is constant, the sharp increase of $D(E)$ leads to a sharp increase of $M(E)$ and TDFs (i.e., $M(E)\lambda(E)$). The resulting strong energy dependence of $M(E)\lambda(E)$ produces non-monotonic Seebeck coefficient behavior. Symbols represent values at optimal Fermi level.

(which is proportional to $M(E)\lambda(E)$) is largest for all energies for $\Delta E_C = 0$, which leads to the higher power factor. As discussed in Sec. III E, however, the assumption of a constant mean free path for this composite band is unrealistic. Next, we discuss the more realistic case where the scattering rate is proportional to the total density-of-states, $\tau(E) = C_{el}/D(E)$. This scattering rate produces a significantly different TDF, which leads to significantly different results.

Figure 12(a) shows the computed power factor vs. ΔE_C for the case of $\tau(E) = C_{el}/D(E)$. Also shown is the power factor for a single light mass valley (red circle). As discussed in Sec. III E, a single heavy mass valley produces much lower performance. Figure 12(a) shows an enhanced power factor when ΔE_C is larger than about $10\,k_BT$. For example, about a factor of two increase can be achieved for $\Delta E_C = 15\,k_BT$. This is in stark contrast to the case of constant MFP for which best performance is obtained with $\Delta E_C = 0$. This behavior occurs because for the case of $\tau(E) = C_{el}/D(E)$, the increase in $D(E)$ leads to large scattering rates if the upper band with its large $D(E)$ is located within the Fermi window. For $\Delta E_C = 0$, σ is reduced by a factor of $\sim$10 compared to that for the single light mass valley.

Figures 12(b)–12(d) plot the power factor, S and σ vs. Fermi level for $\Delta E_C = 5$, 10, and $15\,k_BT$ and show that the improved power factor for $\Delta E_C >\sim 10\,k_BT$ is mainly attributed to a large σ while maintaining a large S of $\sim$120 μV/K in the degenerate limit (which is due to the non-monotonic $S(E_F)$ characteristic). For $\Delta E_C = 10k_BT$, $\sigma = 1.88 \times 10^5\,\Omega^{-1}\mathrm{m}^{-1}$ which is about 4 $\times$ larger than σ for a single light mass valley. It is found that at the optimal Fermi level, σ keeps increasing

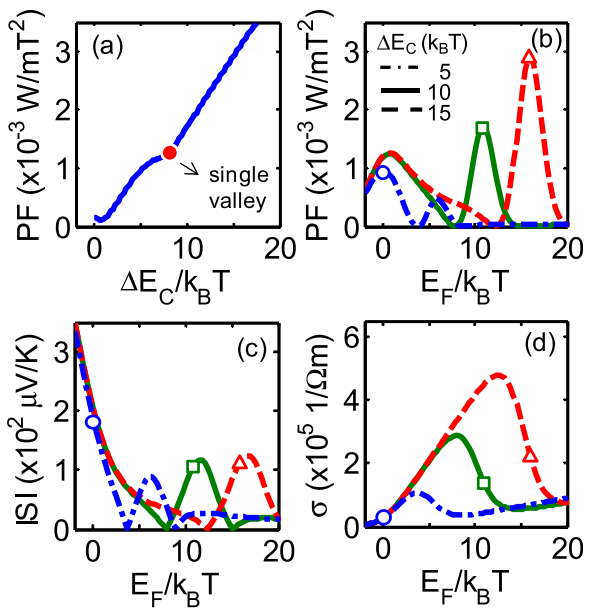

FIG. 12. (a) The PF vs. ΔE_C for a material with lower band and upper band, where ΔE_C is the band-offset. Lower band is isotropic single valley with an effective mass of m_0 and upper band has an effective mass of $10m_0$. Red circles represent the PF for a single valley. The improved PF is obtained for $\Delta E_C >\sim 10k_BT$. (b)–(d) The PF, Seebeck coefficient (S), and electrical conductivity (σ) vs. Fermi level for three cases of $\Delta E_C = 5$, 10, and $15k_BT$. It can be seen that non-monotonic Seebeck coefficient behavior (Fig. 9(c)) improves S at the degenerate limit. As ΔE_C becomes large, σ at optimal Fermi level continues to increase. Therefore, the enhanced power factor is obtained when ΔE_C is larger than about $10k_BT$.

with ΔE_C because the effective number of conduction channels contributed by the lower band increases for large ΔE_C. In practice, achieving this performance would depend on the ability to dope the semiconductor so that E_F is near ΔE_C.

The non-monotonic behavior of $S(E_F)$ is observed for both constant MFP and for $\tau(E) = C_{el}/D(E)$. In fact, for the second case, $S(E_F)$ actually changes sign for $E_F > \Delta E_C$. This can be understood from the $D(E)$, $M(E)$, $\lambda(E)$, and $M(E)\lambda(E)$ characteristics as plotted in Figs. 11 and 13. For constant MFP, the sharp increase of $D(E)$ leads to a sharp *increase* of $M(E)$ and TDF (i.e., $M(E)\lambda(E)$) when the upper bands are available. In contrast, for the second case, the sharp increase of $D(E)$ leads to sharp *decrease* in $\lambda(E)$ and TDF due to large scattering rates when we assume $\tau(E) = C_{el}/D(E)$. From Eq. (2b), the Seebeck coefficient in the degenerate limit is given by

$$S = \frac{\pi^2 k_B{}^2 T}{-3q} \frac{d\left(\ln\lambda(E)M(E)\right)}{dE}\bigg|_{E=E_F}, \tag{9}$$

which is the so-called Mott formula. Consequently, the strong energy dependence of the $M(E)\lambda(E)$ near the edge of the upper band leads to the non-monotonic $S(E_F)$. The drop in $M(E)\lambda(E)$ for the second case causes a change in sign of S.

The simple models considered in this discussion show that we should expect improved TE performance with increasing valley degeneracy. In practice, the valleys may be anisotropic, which provides additional opportunities to increase the number of channels without decreasing the MFP. (Some example calculations are discussed in the Appendix.) We also showed that a semiconductor with a locally distorted $D(E)$ near the Fermi level can display an enhanced power factor through increase of σ and non-monotonic Seebeck coefficient characteristics. However, the increase of power factor only happens when lower and upper bands are

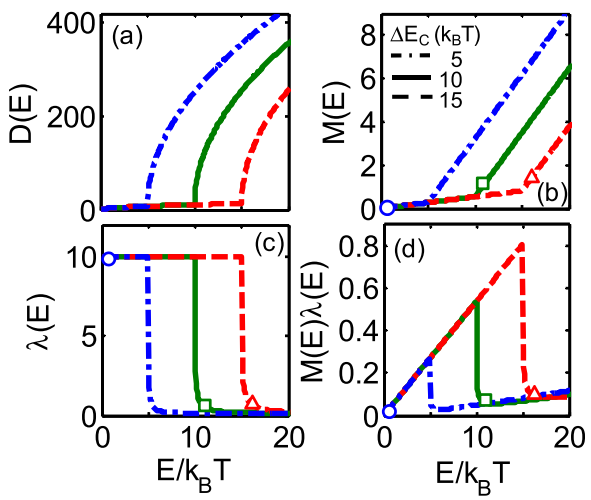

FIG. 13. Density-of-states ($D(E)$), number of conduction channels ($M(E)$), MFP for backscattering ($\lambda(E)$), and $M(E)\lambda(E)$ vs. Fermi level are plotted for three cases of $\Delta E_C = 5$, 10, and $15k_BT$. The MFP is decreased when upper bands are available. As a result, $M(E)\lambda(E)$ is reduced. The resulting sharp decrease of $M(E)\lambda(E)$ produces non-monotonic Seebeck coefficient behavior.

engineered in an appropriate way, so the benefits of locally distorted density-of-states should be carefully considered on a case-by-case basis. Finally, these model calculations show how profoundly scattering influences the results underscoring a point recently made by Zhou *et al.*[30] and illustrating how these effects can be understood by interpreting the TDF from a Landauer perspective.

IV. SUMMARY

In this paper, we set out to answer several questions and summarize the answers as follows.

(1) What physical constraints should be placed on the TDF?
The TDF(E) can be written as a product of the number of channel, $M(E)$, and the mean-free-path, $\lambda(E)$. Each of these two quantities is well-defined and directly relatable to the underlying electronic bandstructure. For a given dispersion, the maximum of $M(E)$ is fixed not the area under $M(E)$ vs. E.

(2) How does the BW of the TDF affect TE performance?
When the lattice thermal conductivity is zero, a δ-function TDF produces an electronic efficiency at the Carnot limit, but no power can be delivered to a load. For a constant MFP (independent of BW), a narrow TDF maximizes the power delivered to a load—for both zero and finite lattice thermal conductivity. For a finite lattice thermal conductivity, it also maximizes the efficiency. The BW should match the width of the Fermi window. A δ-function TDF produces an ZT at the Mahan-Sofo limit, $ZT = \kappa_0/\kappa_{ph}$, but this upper limit itself depends on the BW, so higher ZTs result for higher BWs where operation is below the Mahan-Sofo limit.

(3) How does scattering affect optimum bandstructure?
Scattering profoundly changes these conclusions. If, instead of a constant MFP, we assume that the scattering rate is proportional to the density-of-states, we conclude that a very broad band is better than a narrow band.

(4) How should the improved performance of materials with a high valley degeneracy[31] or with a resonant energy levels[8] that distort the density-of-states be understood?
It is best to achieve a high density of states through valley degeneracy with a number of light mass valleys, as opposed to a single heavy mass valley because the higher velocity of the light mass valley increases both $M(E)$ and $\lambda(E)$. Offsetting the valleys in energy can enhance the Seebeck coefficient, but it degrades the conductivity. With a higher upper valley effective mass and the appropriate energy offset, TE performance can be enhanced, but the results are sensitive to the specifics of scattering.

(5) Is there a best bandstructure for TE performance?
Although there is no simple answer, the general considerations are clear. Assuming that scattering rate is proportional to density-of-states, high average velocities in the Fermi window produce the best results, so materials with a small density-of-states are best. The small $D(E)$ helps to increase scattering times. The large velocity times this $D(E)$ gives a significant number of channels for conduction, and the large velocity also increases the MFP.

For more complex thermoelectric performance, materials can be compared in terms of three well-defined physical parameters: (1) the average velocity in the direction of transport, $|\upsilon_x^+(E)|$, (2) the density-of-states, $D(E)$, and (3) the distribution of channels in energy, $M(E)$. Each of these three parameters can be easily extracted from a bandstructure, $E(k)$. Alternatively, we could express the three parameters as $\upsilon(E)$, $D(E)$, and $\upsilon(E)D(E)$. High velocities lead to long MFPs, low densities-of-states produce long scattering times and long MFPs, and large numbers of channels increase the conductivity. A super linear increase of $M(E)$ is also beneficial for the Seebeck coefficient. As illustrated in the discussion of distorted bandstructures, specific results depend very much on specifics of scattering (e.g., electron-phonon coupling constants), but these three parameters should provide useful guidance in assessing the performance of materials.

ACKNOWLEDGMENTS

This work was supported by MARCO Materials Structures and Devices (MSD) Focus Center and computational services were provided by the Network for Computational Nanotechnology (NCN). The authors also acknowledge illuminating discussions with S. Datta, H. Linke, and A. Shakouri.

APPENDIX A: THE MAXIMUM POWER CONDITION AND THE MAXIMUM EFFICIENCY CONDITION

For the maximum power condition and the maximum efficiency condition, we will find the power delivered to the load resistance and the efficiency. The power delivered to the load,

$$P_{Load} = I^2 R_{Load}, \tag{A1}$$

where I is the current flow due to temperature gradient and is expressed as

$$I = \frac{S\Delta T}{R + R_{Load}}, \tag{A2}$$

where R is the resistance of thermoelectric devices. So Eq. (A1) becomes

$$P_{Load} = \left(\frac{S\Delta T}{R + R_{Load}}\right)^2 R_{Load}. \tag{A3}$$

Equation (A3) can be expressed as

$$P_{Load} = \frac{S^2\Delta T^2}{R}\frac{M}{(1+M)^2}, \tag{A4}$$

where M is the ratio of the resistance of R_{Load} to R, $M \equiv R_{Load}/R$. The efficiency is given by

$$\eta = \frac{P_{Load}}{W_{in}} = \frac{I^2 R_{Load}}{ST_h I + K_{tot}(T_h - T_c)}, \tag{A5}$$

where W_{in} is the total heat flow from the hot side.

1. The maximum power condition, $R_{Load}=R$

First, it can be seen from Eq. (A3) or (A4) that the power is maximized for $R_{Load} = R$ and is given by

$$P^{\dagger}_{Load} = \frac{S^2}{4R}\Delta T^2 = PF\frac{\Delta T^2}{4}, \tag{A6}$$

with $PF = S^2/R$ being the power factor. The power delivered to the load for this condition is a function of PF, not ZT. Under this condition, the current is obtained to be $I = S\Delta T/2R$ from Eq. (A2). Inserting this current into Eq. (A5), the efficiency under the maximum power condition is given by

$$\begin{aligned}\eta^{\dagger} &= \frac{(S\Delta T/2R)^2 R}{ST_h(S\Delta T/2R) + K_{tot}\Delta T} \\ &= \frac{Z\Delta T}{2ZT_{hot}+4} \\ &= \frac{Z}{(2Z+4/T_{hot})}\frac{\Delta T}{T_{hot}},\end{aligned} \tag{A7}$$

where $Z = S^2/(RK_{tot})$. The efficiency is a function of Z.

2. The maximum efficiency condition, $R_{Load}=R\sqrt{1+ZT}$

We can maximize the efficiency in Eq. (A5) when $M \equiv R_{Load}/R$ is given by

$$M = \sqrt{1+ZT} \tag{A8}$$

Under this condition, the maximum efficiency is obtained from Eq. (A5) to be

$$\eta^{*} = \left\{\frac{\sqrt{1+ZT}-1}{\sqrt{1+ZT}+T_{cold}/T_{hot}}\right\}\frac{\Delta T}{T_{hot}}. \tag{A9}$$

This efficiency is also a function of Z. Under the maximum efficiency condition (i.e., $R_{Load} = R\sqrt{1+ZT}$), the power delivered to the load P^{*}_{Load} is obtained from Eq. (A3),

$$\begin{aligned}P^{*}_{Load} &= \left(\frac{S\Delta T}{R+R\sqrt{1+ZT}}\right)^2 R\sqrt{1+ZT} \\ &= \left(\frac{S^2\Delta T^2}{R}\right)\frac{\sqrt{1+ZT}}{(1+\sqrt{1+ZT})^2} \\ &= P^{\dagger}_{Load}\frac{4\sqrt{1+ZT}}{(1+\sqrt{1+ZT})^2},\end{aligned} \tag{A10}$$

where we also used Eq. (A6). P^{*}_{Load} is a function of Z as well as PF since $P^{\dagger}_{Load}$ is a function of PF. In addition, we can see that $P^{*}_{Load} \le P^{\dagger}_{Load}$ always holds because $4\sqrt{1+ZT} \le (1+\sqrt{1+ZT})^2$. Now we have the efficiency (Eqs. (A7) and (A9)) and the power (Eqs. (A6) and (A10)) under both conditions.

APPENDIX B: THE EFFECT OF ANISOTROPIC BANDS

To illustrate the effect of anisotropic valleys in a multi-valley materials, we consider multi-valley materials with a valley degeneracy of $N_V = 6$. For ellipsoidal bands, the degree of anisotropy (k) is represented by the ratio of the longitudinal effective mass (m^{*}_{l}) to transverse effective mass (m^{*}_{t}), i.e., $k = m^{*}_{l}/m^{*}_{t}$. We compare an anisotropic multi-valley material to a single valley material with an effective mass of m^{*}_{2} at the same density of states. It is found that anisotropic multi-valleys produces higher performance than isotropic multi-valleys. This occurs mainly because a larger degree of anisotropy (i.e., larger k) leads to increases in $\langle\langle\lambda\rangle\rangle$ and $\langle M\rangle$ (i.e., enhanced power factor).

The results can be understood in a following way. With the same density-of-states,

$$N_V\left((m^{*}_{l}m^{*2}_{t})^{1/3}\right)^{3/2} = (m^{*}_{2})^{3/2} \tag{B1}$$

or

$$m^{*}_{2} = N_V^{2/3}(m^{*}_{l}m^{*2}_{t})^{1/3} \tag{B2}$$

For each equivalent ellipsoidal band, the number of conduction channels is the density-of-states in the 2D plane transverse to the transport direction[1] (i.e., $\sqrt{m^{*}_{l}m^{*}_{t}}$). The MFP for backscattering is proportional to velocity times scattering time. The velocity is proportional to $1/\sqrt{m^{*}_{t}}$ but the scattering times are the same if we make the physically reasonable assumption that the scattering rate is proportional to the density-of-states. From Eq. (B1) or (B2), the ratio of σ for anisotropic multi-valley (σ_1) to that for isotropic single valley (σ_2) is

$$\frac{\sigma_1}{\sigma_2} \approx \frac{M_1}{M_2}\times\frac{\lambda_1}{\lambda_2} = \frac{N_V\sqrt{m^{*}_{t}m^{*}_{l}}}{m^{*}_{2}}\times\left(\frac{m^{*}_{2}}{m^{*}_{t}}\right)^{1/2} = N_V^{2/3}k^{1/3}. \tag{B3}$$

Equation (B3) reduces to an isotropic multi-valley case at $k=1$. The simple models considered here show that we should expect improved power factor with increasing valley degeneracy and with the degree of anisotropy. Also note that the conductivity of the multiband semiconductor depends more strongly on valley degeneracy than on the degree of anisotropy.

[1]A. F. Ioffe, *Semiconductor Thermoelements and Thermoelectric Cooling* (Infosearch Limited, 1957).
[2]J. R. Sootsman, D. Y. Chung, and M. G. Kanatzidis, Angew. Chem., Int. Ed. **48**, 8616–8639 (2009).
[3]S. K. Bux, R. G. Blair, P. K. Gogna, H. Lee, G. Chen, M. S. Dresselhaus, R. B. Kaner, and J.-P. Fleurial, Adv. Funct. Mater. **19**, 2445–2452 (2009).
[4]L. Weber and E. Gmelin, Appl. Phys. A: Mater. Sci. Process. **53**, 136–140 (1991).
[5]C. Jeong, R. Kim, M. Luisier, S. Datta, and M. Lundstrom, J. Appl. Phys. **107**, 023707 (2010).
[6]H. J. Goldsmid, *Thermoelectric Refrigeration* (Plenum, 1964).
[7]Y. Ma, Q. Hao, B. Poudel, Y. Lan, B. Yu, D. Wang, G. Chen, and Z. Ren, Nano Lett. **8**, 2580–2584 (2008).
[8]J. P. Heremans, V. Jovovic, E. S. Toberer, A. Saramat, K. Kurosaki, A. Charoenphakdee, S. Yamanaka, and G. J. Snyder, Science **321**, 554–557 (2008).
[9]A. I. Hochbaum, R. Chen, R. Diaz Delgado, W. Liang, E. C. Garnett, M. Najarian, A. Majumdar, and P. Yang, Nature **451**, 163–167 (2008).
[10]A. I. Boukai, Y. Bunimovich, J. Tahir-Kheli, J.-K. Yu, W. A. Goddard III, and J. R. Heath, Nature **451**, 168–171 (2008).
[11]R. Venkatasubramanian, E. Siivola, T. Colpitts, and B. O'Quinn, Nature **413**, 597–602 (2001).
[12]T. C. Harman, P. J. Taylor, M. P. Walsh, and B. E. LaForge, Science **297**, 2229–2232 (2002).

[13]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**, 3835–3839 (2006).
[14]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**, 9780–9788 (2007).
[15]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**, 818–821 (2004).
[16]J. R. Sootsman, H. Kong, C. Uher, J. J. D'Angelo, C.-I. Wu, T. P. Hogan, T. Caillat, and M. G. Kanatzidis, Angew. Chem., Int. Ed. **47**, 8618–8622 (2008).
[17]K. Biswas, J. He, Q. Zhang, G. Wang, C. Uher, V. P. Dravid, and M. G. Kanatzidis, Nat. Chem. **3**, 160–166 (2011).
[18]M. G. Kanatzidis, Chem. Mater. **22**, 648–659 (2010).
[19]A. F. May, J.-P. Fleurial, and G. J. Snyder, Phys. Rev. B **78**, 125205 (2008).
[20]H. Li, X. Tang, Q. Zhang, and C. Uher, Appl. Phys. Lett. **93**, 252109 (2008).
[21]G. Homm, P. J. Klar, J. Teubert, and W. Heimbrodt, Appl. Phys. Lett. **93**, 042107 (2008).
[22]X. W. Wang, H. Lee, Y. C. Lan, G. H. Zhu, G. Joshi, D. Z. Wang, J. Yang, A. J. Muto, M. Y. Tang, J. Klatsky, S. Song, M. S. Dresselhaus, G. Chen, and Z. F. Ren, Appl. Phys. Lett. **93**, 193121 (2008).
[23]H. J. Goldsmid, Br. J. Appl. Phys. **11**, 209 (1960).
[24]G. D. Mahan, *Solid State Physics* (Academic, 1997), pp. 81–157.
[25]G. D. Mahan and J. O. Sofo, Proc. Natl. Acad. Sci. U.S.A. **93**, 7436–7439 (1996).
[26]Y. Nishio and T. Hirano, Jpn. J. Appl. Phys. **36**, 5181–5182 (n.d.).
[27]T. E. Humphrey and H. Linke, Phys. Rev. Lett. **94**, 096601 (2005).
[28]N. Nakpathomkun, H. Q. Xu, and H. Linke, Phys. Rev. B **82**, 235428 (2010).
[29]Z. Fan, H.-Q. Wang, and J.-C. Zheng, J. Appl. Phys. **109**, 073713 (2011).
[30]J. Zhou, R. Yang, G. Chen, and M. S. Dresselhaus, Phys. Rev. Lett. **107**, 226601 (2011).
[31]Y. Pei, X. Shi, A. LaLonde, H. Wang, L. Chen, and G. J. Snyder, Nature **473**, 66–69 (2011).
[32]B. C. Sales, Int. J. Appl. Ceram. Technol. **4**, 291–296 (2007).
[33]F. L. Curzon, Am. J. Phys. **43**, 22 (1975).
[34]R. Landauer, IBM J. Res. Dev. **1**, 223 (1957).
[35]G. D. Mahan and M. Bartkowiak, Appl. Phys. Lett. **74**, 953–954 (1999).
[36]C. M. Finch, V. M. Garcia-Suarez, and C. J. Lambert, Phys. Rev. B **79**, 033405 (2009).
[37]C. Jeong, R. Kim, and M. Lundstrom, e-print arXiv:1103.1274v1 (2011).
[38]J.-H. Lee, J. Wu, and J. C. Grossman, Phys. Rev. Lett. **104**, 016602 (2010).
[39]C. Jeong and M. Lundstrom, J. Electron. Mater. **40**, 738–743 (2011).