Thermoelectric transport effects beyond single parabolic band and acoustic phonon scattering

Heng Wang, Ramya Gurunathan, Chenguang Fu, Runzi Cui, Tiejun Zhu and G. Jeffrey Snyder

Thermoelectric materials have been extensively studied for applications in solid-state power generation and cooling. Progress has been made over the past decade in multiple materials systems, hence, it becomes increasingly valuable to be able to analytically model the transport behavior to optimize materials and compare different systems. The well-known effective mass modeling approach is often used to fit the data to the form expected for a single, parabolic band, with charge carrier scattering dominated by acoustic phonons, i.e., deformation potential scattering. However, many high-performance thermoelectric materials benefit from having multiple bands (multi-valley) and many have non-parabolic bands or complex scattering. Understanding how these effects alter properties from that given by the effective mass model provides rational strategies for new materials. In this review, we discuss how this can be done in three scenarios. The first is how to evaluate the influence of point defects on charge carrier mobilities, as well as thermal conductivity. Established methods are available for considering additional scattering mechanisms for phonons and electrons. We focus on how to determine the parameters used in modeling that require the least amount of fitting. We discuss the thermoelectric transport in two different types of materials: lead chalcogenides and half-Heuslers. The second scenario involves systems with multiple sets of conduction or valence bands, which are not necessarily aligned. We discuss different conditions in hypothetical materials systems by considering quality factors for each set of bands. We then demonstrate how the lessons learned are reflected in real thermoelectric materials systems. The third scenario has resonant dopants, and lead chalcogenides have become model systems. These dopants create a distortion in the density of states; inherently, the parabolic dispersion assumption can no longer be used. It is possible nonetheless, to quantitatively understand thermoelectric transport properties, providing insights on how to best utilize resonant dopants. Finally, we provide an outlook, identifying limitations and challenges to solve in order to model, and better yet, predict the thermoelectric performance of different materials.

1. Introduction

Thermoelectric effects enable the direct conversion of energy between heat and electricity. This is made possible entirely by solid-state materials, with the help of mobile charge carriers. Their applications start with thermocouples, which have widespread use as temperature sensors in many applications. What motivates researchers today are power generation and cooling devices based on high-performance thermoelectric materials. These devices have no moving parts, need no maintenance, are compact and quiet, and are suitable for localized/distributed applications.

As of 2017, commercial thermoelectric devices have a sizable market close to a billion US dollars. A decisive majority of this is made up of thermoelectric coolers (Peltier devices) and existing applications are numerous. For example, consumer appliances and medical devices, such as PCR (polymerase chain reaction) equipment used for virus testing, as well as vaccine storage cabinets play important roles in fighting against COVID-19. TECs also provide cooling or temperature regulation for microelectronic and telecommunication devices such as laser diodes, photodetectors, and CCD cameras. In addition, thermoelectric generators have been powering multiple space missions.

Regardless of power generation or cooling, the conversion process can be discussed via the Seebeck effect: the generation of a voltage across an electric conductor that is proportional to
the temperature difference, and the proportionality is the Seebeck coefficient. A large Seebeck coefficient alone is not necessarily the indicator of better performance, regardless of the application. Instead, the measure of the efficiency of the energy conversion process is the figure of merit. The thermoelectric material figure of merit is given by \( zT = S^2T/\rho k \), where \( S \) is a material's Seebeck coefficient, \( \rho \) the resistivity, and \( k \) the thermal conductivity.  

Exceptional thermoelectric performance was first found in semiconductors like Bi$_2$Te$_3$, IV–VI semiconductors such as PbTe, and Si$_1-x$Ge$_x$. A few breakthroughs were made in the 1990s with the concept of PGEC (phonon glass electron crystal) proposed by Slack. Compounds like skutterudites, clathrates, and Zintl phases were identified as promising thermoelectrics, where the structure features a covalently-bonded backbone and loosely-bonded ions. These systems have become the choice for next-generation thermoelectric generators, replacing PbTe and Si$_1-x$Ge$_x$. More examples of thermoelectric compounds include, to name a few: Bi$_{1-x}$Sb$_x$, alloy, Mg$_2$Si alloys, half-Heusler compounds, layered oxytellurides, BiCuSeO, superionic Cu$_x$Se, Zintl phase Mg$_4$Sb$_2$, ductile Ag$_2$S, lead-free chalcogenides GeTe, and SnSe. In addition, organic/polymeric materials are also being investigated, which hold great promise in printable and flexible thermoelectric devices.
flexible devices. Many great reviews are available that cover aspects of materials,37–40 strategy,41–45 as well as devices.44–46

All of these good thermoelectric materials have their electronic properties, $S$ and $\sigma$ (electrical conductivity), trending similarly,37 as expected from a single parabolic band semiconductor. This is because the $S-\sigma$ relationship is mostly determined by the presence of a band edge, requiring only an adequate band gap. Thus the presence of multiple bands, nonparabolicity, unconventional scattering in many materials does not fundamentally change the trends in transport properties that make the effective mass model48 so useful to compare and optimize materials. It is the deviations of the effective mass model, changes in $m^*$, and weighted mobility49 $\mu_w$ with doping, alloying, and temperature that indicate multi-band or nonparabolic effects.41,50,51

For optimizing thermoelectric materials, the quality factor $B$ is best used because it best removes most of the effects of the variation in charge carrier concentration $n$. Thus, comparing the quality factor, and its components, the weighted mobility that determines the electronic quality factor and the lattice thermal conductivity is the best way to compare materials for the maximum possible $\varepsilon T$.

For a single band system, the best achievable $\varepsilon T$, when carrier density is optimized, is determined by the quality factor $B$. Originally referred to by Chasmar and Stratton52 as the $\beta$ parameter, similar concepts have been used by many researchers under different names and expressions.8,53–59 In general,

$$B \propto \mu_w m^* T / \kappa_L$$

(1)

where $m^*$ is the effective mass (in $m_e$), $\mu_w$ is the characteristic mobility at the nondegenerate limit, and $\kappa_L$ is the lattice thermal conductivity. We suggest the use of the following form:59

$$B = \frac{2k_\text{B}h C_1 N_V}{3\pi m^*_L \varepsilon^3 \kappa_L} T$$

(2)

$N_V$ is the valley degeneracy, $C_1$ is the averaged longitudinal elastic constant, $\varepsilon$ is the deformation potential coefficient, and $m^*_L$ is the inelastic effective mass. The two expressions (and other similar ones) have no fundamental difference; using eqn (2) can help avoid confusion, for example, given everything else is the same, smaller (inertial) effective mass $m^*_L$ is beneficial for thermoelectrics. $\mu_w$ is the carrier-energy-independent part of the mobility observed in a system at carrier densities relevant to thermoelectrics (heavily doped). It is not equivalent to the mobility measured on a specific specimen; it is often not the mobility measured in the most-intrinsic, highest-quality, lightly-doped samples. The latter are often documented values:60 1500 cm$^2$ V$^{-1}$ s$^{-1}$ for n-Si, 500 cm$^2$ V$^{-1}$ s$^{-1}$ for p-Si, 8000 cm$^2$ V$^{-1}$ s$^{-1}$ for n-GaAs, etc. Replacing $\mu_w$ with $C_1$ and $\varepsilon$ suggests that deformation potential scattering is the dominant mechanism, which is in many cases a good approximation. Lastly, eqn (2) makes the definition of $B$ easy for each band in a multi-band system.

Understanding measured properties with transport physics and modeling is an important component of rational research toward better materials. The effective mass model gives the material parameters ($m^*, \mu_w, \kappa_L, \Gamma$) from measured properties ($S, \sigma, \kappa, \Gamma$) that would result if the material could be described by a single parabolic band with only acoustic phonon scattering of electrons (deformation potential scattering), often called the SPB model (for an introduction, see Chapter 3 in ref. 61 or ref. 62). Because real systems are not SPB, for example, the non-parabolic Kane band dispersion relation63 is often better suited for narrow-gap semiconductors, the SPB parameters ($m^*, \mu_w, \kappa_L$) will show some variation with temperature and doping.64 In the following sections, we will go over three different strategies for altering the electronic properties: introducing alloy scattering, multiple band systems, and resonant impurities, giving quantitative modeling methods in each case. Some implications on materials design are also discussed.

2. Alloys with isovalent substitution

Isovalent substitution is perhaps one of the first used strategies to make better thermoelectrics. The success of Bi$_2$Sb$_3$Te$_3$ and Si$_{1-x}$Ge$_x$ solid solutions is among the first lessons learned by every new researcher in the field.

2.1 Thermal conductivity

Solid solutions have lower thermal conductivities. If the goal is to understand and model thermal conductivities in solid solutions with proper experimental inputs, the method suggested by Callaway and Klemens can be used even for complex alloys with vacancies and interstitials.72 This method uses experimentally determined lattice thermal conductivity $\kappa_L$ of pristine compounds, together with a few physical properties as input, to predict thermal conductivities in solid solutions.

The applicability of this method relies on the assumption that the phonon scattering mechanism is dominated by the Umklapp process (this could include contributions from the normal process) in the pristine compound, and in solid solutions point defects are, predominantly, the only added phonon scattering source. The relation between lattice thermal conductivity of the pristine compound $\kappa_L$ and the solid solution $\kappa_L$ is given as follows:

$$\frac{\kappa_L}{\kappa_L} = \frac{\arctan(u)}{\kappa_L}$$

(3)

where $u$ is given by

$$u^2 = \frac{\pi \theta_D \Omega}{2 \hbar \kappa_L}$$

(4)

$\theta_D$ is the Debye temperature, $\Omega$ is the volume per atom in the structure, $\nu$ is the average speed of sound, and $\Gamma$ is the scattering parameter. Klemens’ analysis took into account influences from mass difference, binding force difference, and strain field induced by a point defect (Fig. 1). The full expression of $\Gamma$ contains 3 terms, capturing the change in mass, site radius, and force constants at the defect site (see eqn (12) in ref. 74). Abeles simplified it into two terms, for binary ($A_1B_x$ type) or pseudo-binary ($A_1B_x(A_2C_y)$ type) systems:
Fig. 1 Illustration of a substitutional point defect in a crystalline lattice. The defect brings a different atomic mass, different bonding strengths with its neighboring atoms, and strains to its surrounding lattice; all have an impact on thermal conductivity in solid solutions where large numbers of substitutions are present.

\[
\Gamma = x(1 - x) \left[ \left( \frac{\Delta M}{M} \right)^2 + e \left( \frac{\Delta a}{a} \right)^2 \right] \tag{5}
\]

\(\Delta M\) and \(\Delta a\) are the differences in mass and lattice constants between two constituents, \(M\) and \(a\) are the molar mass and lattice constant of the alloy, respectively. The parameter \(e\) is determined by the Grüneisen parameter \(\gamma\) and elastic properties like the Poisson ratio \(\nu\). Abeles wrote \(e\) as follows:75

\[
e = \frac{2}{9} (G + 6.4\gamma \frac{1 + \nu}{1 - \nu} ) \tag{6}
\]

The scaling factor (6.4) of the strain effect comes from the assumption of a cubic lattice with octahedral coordination at each lattice site. The coefficient represents the number of residual bonds perturbed by a strain field decaying as \((1/d^3)\), where \(d\) is the distance from the point defect.68 A factor of 6.4 neglects the scattering of the nearest neighbor bond and is recommended for vacancies since bonds to neighboring sites are absent. For substitutional defects, a factor of 8.4 is recommended by Klemens.68 Note that Abeles used 6.4 instead of 8.4 as suggested in Klemens' analysis,68 which is inherited by most works afterwards.

\(G\) in eqn (6) is a ratio between the contrast in bulk modulus (\(\Delta K/K\)) and that in the bonding length (\(\Delta R/R\)). The relationship between bulk modulus and the volume of the unit cell has been studied for a lot of compounds. For instance, see Anderson’s study76,77 on minerals of different types. A general trend exists among similar compounds: \(KV^2\) = constant (Fig. 2). From an In \(K\) vs. In \(V\) plot, it was found that when \(AK\) and \(AR\) are small, \(G\) is approximately three times the slope that \(G = 4\) for covalent IV and III–V structures, and more ionic II–VI and I–VII structures \(G = 3\). The value of \(G\) should be based on material systems. For Pb chalcogenides, we suggest the use of \(G = 3\).

With all parameters in eqn (6) determined, the value of \(\varepsilon\) could be calculated for any binary compounds. For PbTe, PbSe and PbS, \(\varepsilon\) was calculated to be 100, 110, and 150, respectively.

For PbTe-based systems, \(\varepsilon\) was previously suggested to be 65 by Alekseeva,83 which was obtained from fitting experimental results. This value has been used in lots of studies. We suggested the use of 100 as it has a clear physical basis and the modeling results remain reasonably consistent with experiments. On applying eqn (6) to different systems, we found that for Si1–Gei and Ga1–IniAs, the calculated \(\varepsilon\) was very close to those from fitting,75,84 whereas for Mg2Si1–SniSn, a significant discrepancy was seen (23 from fitting85 vs. 67 from calculation).

In compounds with mixed ionic and covalent bonding character, uncertainty around the choice of atomic or ionic radii when defining the strain scattering term can further confound predictions.68 Bearing some uncertainty, the thermal conductivity of binary (pseudo-binary) solid solutions can be predicted with Klemens’ method, with no fitting parameter needed.

Abeles has suggested25 that when there is more than one type of substitutional defect, eqn (5) should be generalized to eqn (7), where \(i\) indexes the species that can occupy the lattice, including the host atom and any impurity defects:

\[
\Gamma = \sum_i \Gamma_i = \sum_i \xi_i \left[ \left( \frac{\Delta M_i}{M} \right)^2 + e \left( \frac{\Delta a_i}{a} \right)^2 \right] \tag{7}
\]
This formulation is only well-defined for a monatomic lattice, in which all host and impurity atoms occupy the same sublattice.

Yang et al. introduced another method, which is widely adopted and successful when applied to complex structures like skutterudites. In this method, \( \Gamma \) has two parts \( \Gamma_m \) and \( \Gamma_s \), representing the mass fluctuation and strain fluctuation (for a case with \( n \) sublattices, and each sublattice has two different atoms ‘1’ and ‘2’, see ref. 73 and 87 for detailed descriptions):

\[
\Gamma_m = \frac{\sum_{i=1}^{n} c_i \left( \frac{M_i}{M} \right)^2 f_i f_i' \left( M_i' - M_i \right)^2}{\sum_{i=1}^{n} c_i}
\]

(8)

\[
\Gamma_s = \frac{\sum_{i=1}^{n} c_i \left( \frac{M_i}{M} \right)^2 f_i f_i' \left( r_i' - r_i \right)^2}{\sum_{i=1}^{n} c_i}
\]

(9)

Yang’s method was derived from Slack’s work discussing the influence of isotopes on thermal conductivity. Note that the mass term \( \frac{M_i}{M} \) is used in eqn (9) on strain fluctuation. A probable rationale is that the vibrations of heavier sublattices have larger impacts on the overall vibrational modes of the structure, and thus perturbations on each sublattice should be weighed in accordance with this term, although no clear reasoning was explicitly provided.

The two treatments appear more different than they actually are. The most general formulation for the point defect scattering parameter of a polyatomic lattice (i.e. multiple sublattices) with an arbitrary number of defects was first presented by Berman, Foster, and Ziman for isotope scattering. Here, the scattering parameter \( \Gamma \) has the basic form of the averaged atomic mass variance in the lattice divided by the squared averaged atomic mass. The form below, extended to include strain fluctuations, is discussed in detail in eqn (10):73

\[
\Gamma = \Gamma_m + \Gamma_s = \frac{\langle \Delta M^2 \rangle}{\langle M \rangle^2} + \frac{\langle \Delta R^2 \rangle}{\langle R \rangle^2}
\]

(10)

There are two levels of averaging required in order to treat polyatomic lattices. Here, the bar in \( M \) or \( V \) signifies a configurational average over all arrangements of the alloy. The angular brackets \( \langle \cdot \rangle \) signify a stoichiometric average over all the elements in the formula unit. More explicitly, \( \langle \Delta M^2 \rangle \) is given as follows:

\[
\langle \Delta M^2 \rangle = \frac{\sum_{i} c_i \sum_{n} x_i (M_i - \bar{M}_n)^2}{\sum_{i} c_i}
\]

(11)

and where \( n \) indexes the sublattices in the formula unit, \( c_i \) is the degeneracy or stoichiometric ratio of each sublattice, and \( i \) again indexes all species that can occupy a given sublattice, including the host atom and any substitutional defect.

Most implementations of \( \Gamma \), as detailed above, are based on “per atom” quantities, with the perturbation defined by atomic mass or radius variances and the defect volume \( \Omega \) defined as the volume per atom. However, several texts, including the original work of Klemens, define \( \Gamma \) instead of “per unit cell” quantities, which quantify the perturbation in terms of the mass and volume change of the primitive unit cell when it is occupied by a specific population of substitutional defects. This treatment is in keeping with a monatomic lattice approximation, in which all atoms in the primitive unit cell are effectively summed together into a single, vibrating mass. It is shown that either formulation will provide equivalent results, as long as the user is consistent in using “per atom” or “per unit cell” definitions. Several discrepancies exist in model results reported in the literature when parameter definitions have a mixed basis. Often eqn (10) is defined in terms of atomic masses and radii, but then \( \Omega \) in eqn (4) is defined as the volume of the unit cell, leading to an overestimation of the thermal conductivity reduction due to alloy scattering.

In several cases, this literature discrepancy has obscured the unique scattering characteristics of off-stoichiometric defects, like vacancies and interstitial atoms. In the vacancy and interstitial cases, the removal or addition of bonds leads to a large potential energy perturbation and additional scattering. A simple Klemens model based on the virial theorem relates the potential energy perturbation to the average atomic mass in the solid (\( \langle M \rangle \)). Therefore, in this model, the scattering parameter can be determined solely using the \( \Gamma_m \) term when \( \Delta M = M_d + 2\langle M \rangle \), where \( M_d \) is the missing/added mass of the vacancy/interstitial defect. In practice, vacancy and interstitial defects can have about 10 times the scattering strength (\( \Gamma \)) of a simple, substitutional defect.

2.2 Charge carrier mobility

Regardless of how effective point defects are in the scattering of phonons, they cause the scattering of charge carriers as well. For isovalent substitutional defects, the atomic site is not charged. Scattering is caused by perturbations to the periodic lattice potential and is called alloy scattering.

To briefly explain how the relaxation time is determined, consider a lattice containing only two types of atoms A and B, and define the average lattice potential in a disordered system:

\[
\bar{U} = xU_A + (1 - x)U_B
\]

(12)

\( x \) is the concentration of atom A in the alloy. For a lattice site occupied by atom A, the local potential perturbation Hamiltonian is as follows:

\[
H_{\text{alloy},A} = U_A - \bar{U} = (1 - x)\Delta U
\]

(13)

where \( \Delta U = U_A - U_B \). The scattering probability by A atoms with concentration \( x \) is

\[
|M_{k,A}|^2 = n_A \sum_{k'} |\langle k | H_{\text{alloy},A} | k' \rangle|^2 = \frac{x(1 - x)^2 \Delta U^2}{\Omega}
\]

(14)

Similarly, the scattering probability by B atoms is

\[
|M_{k,B}|^2 = n_B \sum_{k'} |\langle k | H_{\text{alloy},B} | k' \rangle|^2 = \frac{(1 - x)x^2 \Delta U^2}{\Omega}
\]

(15)
So the total relaxation time\(^{96}\) (parabolic band) is

\[
\tau_{\text{alloy}} = \frac{\hbar}{2\pi \left| M_{A,A} \right|^2 \left| M_{B,B} \right|^2 \sqrt{\xi(x)^{-1}}}
\]

or

\[
\tau_{\text{alloy}} = \frac{\hbar}{\pi^2 \hbar^3} \frac{1}{2 \pi^2 \Omega n^2 (k_B T)^{3/2} \xi \sqrt{(1 - x) \Delta U^2}}
\]

(16)

This simple derivation, based on defining the defect perturbation with respect to a virtual crystal lattice potential \(\langle \hat{U} \rangle\), is the same approach used to derive the classical Nordheim’s Rule often used to model the excess resistivity in metal alloys.\(^{97}\) Nordheim’s rule and the simple scattering potential above for binary alloys can also be derived from the coherent potential approximation (CPA), in which a self-energy term, determined self-consistently, is added to each site of the disordered lattice, generating an effective mean field that scatters charge carriers.\(^{98}\) The CPA is discussed as the preferred method for alloy scattering and has been used to derive higher-order extensions of the alloy scattering model, in which there are two or more alloying elements in the system.\(^{99,100}\) In these multicomponent alloys, the scattering probability is\(^{100}\)

\[
\left| M \right|^2 = \frac{1}{\Omega} \sum_i \sum_j x_i x_j (U_i - U_j)^2
\]

(17)

Alloy scattering has the same (reduced) carrier energy dependence \(\epsilon^{-1/2}\) as deformation potential scattering, which means the \(S\) vs. \(n\) (Pisarenko) relation would be exactly the same before and after alloying, giving no change in effective mass. Correspondingly, the \(\mu\) vs. \(n\) relation for solid solutions would also have the same trend as the pristine compound, just as if the deformation potential had increased. The difference, however, is clear in the temperature dependence, with \(T^{1/2}\) for alloy scattering, and \(T^{-3/2}\) for deformation potential scattering.

The first explicit expression of relaxation time \(\tau_{\text{alloy}}\) was developed by Harrison and Hauser\(^{101}\) for nondegenerate III–V semiconductors. We used this expression in previous works; \(\tau_{\text{alloy}}\) for a parabolic band is given by

\[
\tau_{\text{alloy}} = \frac{8 \hbar^4}{3 \sqrt{2} \pi \Omega \xi (1 - x) \Delta U^2 m_0^2 (k_B T)^{3/2} \xi} \epsilon^{1/2}
\]

(18)

Compared with eqn (16) from a much-simplified derivation, there is only a small difference in the numerical constants. In fact, after a short literature survey, we found that the constant for \(\tau_{\text{alloy}}\) used by Makowski and Glicksman\(^{97}\) was roughly 2 times that in eqn (18). The one used by Chattopadhyay\(^{103}\) is half of that in eqn (18). Mahrotra used\(^{104}\) eqn (16). Caution is needed when comparing results from different researchers.

The alloy scattering potential \(\Delta U\), following the simple derivation, is the offset of lattice potentials between substitutional atoms and the atoms being substituted. Unfortunately, based on lead chalcogenides, we found that none of the potential offsets obtained directly from another method can explain the experimental mobility reduction in solid solutions. Historically, Brooks\(^{109}\) suggested using the band gap difference between the 0% and 100% substituted compounds. This was later found to provide inaccurate values, whereas, the difference in electron affinity (hence the position of band edges relative to vacuum level) is a better measure.\(^{96,110}\) This is clearly not supported by results from the PbTe–PbSe or PbSe–PbS system: in each case, the conduction band offset is about 0.1 eV, whereas \(\Delta U\) from the mobility data is around 1 eV (Fig. 3a and b). Band alignments were calculated \(via\) DFT for coherently embedded nanoscale clusters;\(^{108,111–113}\) these cannot be applied in the case of point defects. In cases where near-perfect band alignment was found for nano-clusters,\(^{108}\) experimental results with solid solutions\(^{107}\) have shown significant alloy scattering (Fig. 3c).

So far, \(\Delta U\) can only be found as a fitting parameter, and predicting mobility changes in solid solutions is still a goal to reach. Nonetheless, a qualitative link between the band position offset and \(\Delta U\) might be justifiable: \(\Delta U\) is around 1 eV in n-type PbTe–PbSe and PbSe–PbS, and 3 eV in PbSe–SrSe. The band position offset is around 0.1 eV for the former two cases and 0.8 eV for the latter. Further study in p-type PbSe–PbS would be helpful; since the valence band offset in this case is very small (0.03 eV), \(\Delta U\) much less than 1 eV, and a minimum mobility reduction are expected. We note here that a similar discrepancy was seen in the coefficient \(\Xi\) for deformation potential scattering: from fitting mobility data \(\Xi\) in lead chalcogenides were found\(^{114}\) to be over 20 eV, whereas those from calculation\(^{115}\) were no more than a few eV.

Often, disorders from point defects in lead chalcogenides are not ‘beneficial’. However, studies in other material systems, for example, many half-Heusler compounds, have demonstrated that point defects (iso-valent substitution or doping) are capable of reducing thermal conductivity without a heavy penalty on carrier scattering.\(^{117–120}\) There is no guideline to predict what substitutions, or which systems can benefit from disorder. To shed some light on this, we observed that in lead chalcogenides, the strain contrast has a large contribution to the thermal conductivity reduction \(\epsilon\) from eqn (6) is large in these cases), whereas in half-Heuslers the strain contrast is almost negligible\(^{117,118}\) (especially, when between a period-5 and a period-6 element). The more fundamental reason could be that the bands in lead chalcogenides are from more-overlapped, less-hybridized \(p\)-orbitals of cations and anions (hence, small effective masses, high mobilities), which are sensitive to interatomic distance changes. In contrast, bands in half-Heuslers are from less-overlapped, highly-hybridized \(d\)-orbitals (hence, large effective masses, low mobilities), which are insensitive to changes in the local atoms.

Therefore, selecting an alloying element that has a large mass difference and a small atomic size difference with the host atom should be more likely to generate a beneficial disorder. Such candidates are more likely to be found when both alloying and host atoms are transition elements. The lanthanide contraction leads to a small covalent radius difference between the period-6 elements (post-lanthanides) and their isoelectronic period-5 counterparts. Following this, we have found a strategy (Fig. 4) particularly suitable for half-Heusler compounds containing transition metals ranging from periods 4 to 6: n-type Zr\(_{1-x}\)Hf\(_x\)NiSn, p-type...
could be other factors such as the change in band structures. One of the reasons why Bi$_2$Te$_{1-x}$Se$_x$ and Bi$_2$Se$_{3-x}$Te$_x$ have better performance than Bi$_2$Te$_3$ is that band gaps in these alloys are larger, thus thermal excitation of minority carriers is inhibited. The change in defect chemistry$^{133}$ and carrier concentration is another factor such that no additional doping is necessary. Band convergence happens in systems (Fig. 5) such as$^{17,24,134–139}$ Bi$_2$–Sb–Te$_3$, Bi$_2$–Se–Te$_3$, Si$_1$–Gex$_3$, PbTe$_1$–Se$_3$, Mg$_x$Si$_1$–Sn$_x$, or Mg$_x$Bi$_2$–Sb$_2$ (different explanations should be noted$^{140,144}$), so much so that in most cases (except for Si$_1$–Gex$_3$), the optimum solid solution composition is consistent with what is required for band convergence.

3. Systems with multiple bands

Thermoelectrics often demonstrate better performances when they take advantage of complex band structures. Symmetry-imposed degeneracy of multiple carrier pockets is seen in most high-performance systems. In other cases, materials have benefited from the presence of additional bands with small energy offsets from the primary conducting bands. Enabling band convergence$^{107,137,142}$ where the edge of different bands came to the same energy, has been a dependable strategy to achieve higher $zT$s.

The maximum possible thermoelectric properties of a two-band system have been discussed in several studies.$^{143}$ Some are for electron–hole systems$^{144,145}$ as many thermoelectrics are narrow bandgap semiconductors. Consider a system with two bands having a reduced energy offset for band convergence.

$$zT = \frac{S^2\sigma}{k}T = \frac{(S_1\sigma_1 + S_2\sigma_2)^2}{(\sigma_1 + \sigma_2)}$$

Here, the subscripts denote the properties of each band, with “1” being the primary band, and “2” the secondary band. All terms in eqn (19) can be expressed in terms of the quality factor $B$, and Fermi integrals$^{146}$:

$$\sigma = 3\left(\frac{k_B}{e}\right)^{-2}kL B F_{-2}^{-1}$$

$$S\sigma = 3\left(\frac{k_B}{e}\right)^{-1}kL B^{(1)} F_{-2}^{-1} - \eta^{0} F_{-2}^{-1}$$

$$L + S^2 = \left(\frac{k_B}{e}\right)^{2} \frac{2}{\eta} F_{-2}^{-1} - \eta^{1} F_{-2}^{-1} + \eta^{2}$$

$\eta = E_F/k_B T$ is the reduced chemical potential of carriers. $F$ is the generalized Fermi integral for non-parabolic Kane bands$^{63,146}$.
(take $\alpha = k_B T/E_g = 0$ for parabolic bands):

$$nF_{m}^{\alpha}(\eta, \alpha) = \int_{m}^{\infty} \left( -\frac{\partial f}{\partial e} \right) \varepsilon^2 (\varepsilon + e)^2 \left[ (1 + 2\varepsilon\alpha)^2 + 2 \right]^2 \, d\varepsilon \quad (23)$$

Thus, the overall $zT$ of a two-band system can be expressed analytically. The detailed equation is lengthy yet simply algebra. $zT$ is a function of $B_1, B_2, A$, and $\eta$ only, which also means the optimized $zT$ is only a function of $B_1, B_2, A$. Fig. 6 and 7 present the $zT$ of a hypothetical two-band system (both are parabolic). In Fig. 6 we take $B_1 = 0.35$, which is about the quality factor of the L valence band\(^{146}\) in PbSe, and in Fig. 7 $B_1 = 0.7$, which is about that of the L conduction band\(^{146}\) in PbSe. The value of $B_2$ is allowed to change within a reasonable range based on known quality factors. We consider 0.17 to be the lower bound for meaningful contribution to $zT$, whereas 1.4 is almost unrealistically high based on known examples.

Note that we have neglected a potential adverse influence on carrier mobility when multiple bands come closer: the effect of inter-band/valley scattering. It is reasonable to expect that the optimum $zT$ in a two-band system would be lower than expected from the simple analysis above, which assumes no inter-valley scattering. In simple terms, the question is the strength of inter-valley scattering as compared to intra-valley scattering.\(^{59}\) If there is a symmetry requirement forbidding inter-valley scattering, there is a clear benefit. However, even when the scattering is allowed, there is little reason to doubt that a higher $B$ and $zT$ would result.

Consider inter-band scattering when the initial and final states belong to different pockets with their extrema at the same $k$ point. The inter-band scattering is very similar to intra-band scattering because of the small change in the wave vector. One such example is the multi-band system of Mg$_2$X (X: Si, Ge, Sn). An expression of relaxation time was suggested\(^{147,148}\) and an inter-band deformation potential $D$ was introduced. In principle, $D$ can be determined if both functional forms of the two band states are known; this is by no means easily available information. Experimentally, it is impossible to tell the inter-band processes from the normally discussed intra-band processes since they have the same energy and temperature dependence. First-principles calculation of the electron–phonon interaction would be the only method to study this. In general, this type of multi-band system tends to see increased inter-band scattering as two bands move close. It is now known that in certain systems,\(^{149}\) the convergence of bands is not beneficial due to increased carrier scattering. Nonetheless, in other systems like Mg$_2$X (X: Si, Ge, Sn), the...
convergence is beneficial and the inter-band scattering was reported as insignificant.\textsuperscript{147}

The inter-valley scattering is a little different when the two pockets have extrema at different \( k \) points. Since a large change in the \( k \) vector is needed, this process resembles that of the optical phonon scattering and is inelastic in nature,\textsuperscript{150–153} which is distinguishable from intra-valley deformation potential scattering processes. The best-known example of inter-valley scattering is perhaps in n-type Si where inter-valley scattering was found to be twice as significant as intra-valley scattering.\textsuperscript{151,154–156} However, several more recent results\textsuperscript{157–161} have indicated that the inter-valley scattering is important but not greater than intra-valley scattering. Overall, there have been only very limited examples (ref. 162 which is a theoretical study) where the convergence of multiple valleys has failed to improve the thermoelectric performance.

As long as inter-band/valley scattering would not completely offset the benefit from converging multiple bands, the following principles can be concluded (the dominant scattering mechanism is still deformation potential scattering):

1. The optimized \( zT \) is determined only by quality factors of both bands and their offset, not the individual band characteristics, such as effective mass.

2. The quality factor of the first band sets the baseline for \( zT \). The second band adds its contribution depending on the offset between them; the smaller the offset, the larger the contribution. If a system has a first band with a low-quality factor, its \( zT \) will not be high unless a better second band takes over.

3. Depending on the ratio \( B_2/B_1 \), the second band starts to contribute when \( \Delta \) is below \( \sim 2 \) small \( B_2/B_1 \) to \( \sim 4 \) large \( B_2/B_1 \).

---

**Fig. 5** Band convergence in thermoelectric solid solutions. (a) Hole pockets of Bi\textsubscript{2}Sb\textsubscript{3}Te\textsubscript{3} and (b) schematic valence band positions with different compositions, from ref. 134. (c) Hole pockets in PbTe and (d) schematic valence band positions at different temperatures in PbTe\textsubscript{1-x}Se\textsubscript{x}, from ref. 137. (e) Schematic band positions of Mg\textsubscript{2}Si\textsubscript{1-x}Sn\textsubscript{x}, from ref. 138. Used with permission, Copyright the American Physical Society. (f) Schematic band positions of Mg\textsubscript{2}Sb\textsubscript{2-x}Bi\textsubscript{x}, from ref. 24.

**Fig. 6** Modeled \( zT \) of a two-band system with the two bands separated by reduced energy, \( \Delta \). The quality factor for the first band \( B_1 = 0.35 \). (a)–(d) Cases with different quality factors \( B_2 \) for the second band.
On the other hand, two bands with $\Delta \leq 1$ are almost as good as converged (though which one comes first matters). More studies are needed to understand whether the inter-band/valley scattering intensity has notable changes within such a range of $\Delta$.

4. In most cases, the maximum $zT$ is achieved only at one $\eta$, which is close to the edge of the first band (slightly below, for a conduction band), regardless of $B_2$. A good strategy should always be reducing $\Delta$, instead of excessive doping to let $\eta$ ‘reach a better band’. This is consistent with our conclusion regarding resonant doping in the next section, that the resonant level should be close to the band edge.

5. Only for extreme $B_2/B_1$ ratios, the optimum $\eta$ moves away towards the edge of the second band, and the maximum $zT$ is found when $\eta$ is near the edge of the second band. $B_1$ is very small in this case (Fig. 8 and 9). This is likely the case when the first band is an in-gap defect band. Even though the optimum $\eta$ followed the second band, the first band still leaves its negative impact on $zT$.

6. When the two bands are close ($\Delta < 1$) the optimum $\eta$ slightly shifts toward the second band. This shift is less than $0.5k_BT$, even when the second band has a $B$ factor twice as high. The optimum $\eta$ for a two-band system is approximately the same as that for the first band alone. The carrier density needs to be properly increased to keep the same $\eta$.

This discussion so far is simply based on numerical simulation. Nonetheless, we can consider a few real examples where experimental findings reflect such principles:

### 3.1 La$_{3-x}$Te$_4$

La$_{3-x}$Te$_4$ is an excellent n-type material at high temperatures ($T > 1000$ K). Band structure calculation indicated two sets of conduction bands separated by about 0.3 eV (Fig. 10a), which is equivalent to $\Delta = 2.8$ at 1273 K. Fermi distribution can create finite occupancy of the second band even at lower temperatures (400 K, Fig. 10b) with composition control (doping). However, as suggested by our discussion above, at this value of $\Delta$, the second band would have limited contribution to the overall performance. La$_{3-x}$Te$_4$ has decent $zT$ at lower temperatures, as well with optimized $zT > 0.6$ at 700 K, suggesting a decent quality factor $B_1$ for its first conduction band. Thus, if we assume the first conduction band has a $B_1$ that scales with $T$, of conduction bands separated by reduced energy, $\Delta$. The quality factor for the first band $B_1 = 0.7$. (a)–(c) Cases with different quality factors $B_2$ for the second band.

Fig. 7 Modeled $zT$ of a two-band system with the two bands separated by reduced energy, $\Delta$. The quality factor for the first band $B_1 = 0.7$. (a)–(c) Cases with different quality factors $B_2$ for the second band.

Fig. 8 Modeled $zT$ of an extreme two-band (hypothetical) system with the two bands separated by reduced energy, $\Delta$. The first band is set to be a very poor one with $B_1 = 0.05$ ($zT$ below 0.2), while the second band has a decent $B_2 = 0.5$. 

### 3.1 La$_{3-x}$Te$_4$

La$_{3-x}$Te$_4$ is an excellent n-type material at high temperatures ($T > 1000$ K). Band structure calculation indicated two sets of conduction bands separated by about 0.3 eV (Fig. 10a), which is equivalent to $\Delta = 2.8$ at 1273 K. Fermi distribution can create finite occupancy of the second band even at lower temperatures (400 K, Fig. 10b) with composition control (doping). However, as suggested by our discussion above, at this value of $\Delta$, the second band would have limited contribution to the overall performance. La$_{3-x}$Te$_4$ has decent $zT$ at lower temperatures, as well with optimized $zT > 0.6$ at 700 K, suggesting a decent quality factor $B_1$ for its first conduction band. Thus, if we assume the first conduction band has a $B_1$ that scales with $T$,
at 1273 K the observed maximum $zT$ of 1.1 matches our expectation without the contribution from a second band. The optimum carrier density was found around 2 to 4 x $10^{20}$ cm$^{-3}$, consistent with the expected doping level for a DoS effective mass of the first band (1.4 $m_e$). This further suggested that no significant number of carriers populated the second conduction band. Thus, even though La$_{3-x}$Te$_4$ is a system with two conduction bands, it is likely that the second band has limited contribution to the material's high $zT$.

### 3.2 CoSb$_3$

CoSb$_3$ has two conduction bands. Historically, Sofo and Mahan$^{166}$ found from the calculated band structure that CoSb$_3$ has a non-parabolic Kane band and a triply degenerate parabolic band at the $\Gamma$ point. A previous study$^{20}$ using optical absorption (300 K) identified a second conduction band off $\Gamma$ at 0.08 eV above the primary band edge (corresponding to $\Delta = 3.1$ at 300 K, and 1.2 at 800 K, as shown in Fig. 11b and d). This second conduction band has a complex Fermi surface (Fig. 11c), a very large DoS effective mass, a high equivalent degeneracy $N_v$, and a quality factor $B_2$ that is likely comparable or even better than the first one (we estimated $B_1$ to be around 0.4 at 800 K). Different from La$_{3-x}$Te$_4$, the small $\Delta$ makes the second band influence perceptible even at room temperature: the observed effective mass from the Seebeck coefficient steadily increases with carrier density.$^9$ At a high doping level of $5 \times 10^{20}$ cm$^{-3}$, the Seebeck coefficient was still$^9$ around $-100$ $\mu$V K$^{-1}$. Assuming the same $\Delta$, we could estimate the optimum $zT$ at 800 K to reach 1.2 (see Fig. 6b). Better yet, temperature-dependent absorption$^{20}$ measurements confirmed that $\Delta$ decreases with temperature and the two sets of bands converge at around 800 K. The convergence could potentially bring maximum $zT$ to 1.5 at 800 K.

Experimentally, single-filled CoSb$_3$ achieved $zT = 1.2$ at 800 K. This value is lower due to the narrow bandgap of CoSb$_3$ (at 0.23 eV at 300 K). In fact, based on Fig. 9c, the optimum $\eta$ with converged bands is between $-0.5$ to $-1.0$, suggesting that the Seebeck coefficient should be between $-240$ to $-270$ $\mu$V K$^{-1}$ when $zT$ is maximized. Experimental values from different studies are around $-180$ $\mu$V K$^{-1}$, which is an indicator of the valence band influence.

### 3.3 Bi$_3$Te$_3$

Bi$_3$Te$_3$ is an outstanding system due to its inherently low lattice thermal conductivity and complex conduction and valence band Fermi surfaces. Both its conduction band side and valence band side Fermi surface have two (three) sets of valleys emerging at different energies$^{167,168}$ (Fig. 12). The primary valence band is six-fold degenerate. The edge of the second set of six valleys has an offset of 0.03 eV ($\Delta = 1.2$) from the band edge. At even higher energies, a total of 12 pockets merge around Z point forming a ring structure. On the conduction band side, the primary conduction band is also six-fold, the second set of two pockets along the Z–Z line emerge at 0.05 eV.
processes, we can find\textsuperscript{167} that the conduction bands have a noticeable influence from the additional pockets. Nonetheless, if we assume that the primary bands dominate transport, we expect in p-type alloys that the maximum $zT$ is not sensitive to small changes in alloy composition, given that carrier densities are properly tuned; this was observed experimentally.\textsuperscript{178,179}

On the other hand, for n-type alloys, we know that (a) smaller $\Delta$ are required for the second band to affect overall $zT$, (b) $zT$ of a two-band system is primarily determined by the first band. Thus, we expect the maximum $zT$ to be more sensitive to composition changes, especially on the Se-rich side (when the low-quality band becomes primary). Again, this was observed experimentally.\textsuperscript{180,181} If a pseudo-ternary system can be found, which also converges the additional six electron pockets, notably higher $zT$s can be expected for n-type alloys.

4. Resonant doping

In thermoelectrics, dopants are point defects that provide free charge carriers thus shifting the position of the Fermi level. Often, this is considered as (approximately) their only impact on the material’s transport properties (with exemptions for very heavy doping). The resonant dopants are a special type of dopant, which, in addition to providing free carriers, interact with conducting bands, leading to unique transport behavior.\textsuperscript{182–187}

The success of resonant doping is well-known\textsuperscript{188,189} in Tl-doped PbTe and In-doped SnTe, and the enhancement of effective mass $m^*$ as seen from the higher Seebeck coefficient in the Pisarenko ($S$ vs. $n$) relation. More recent studies suggest that resonant doping exists in more systems such as Bi$_2$Te$_3$,\textsuperscript{187} GeTe,\textsuperscript{190} half-Heuslers,\textsuperscript{191,192} As$_2$Te$_3$,\textsuperscript{193} and Zn$_2$Sb.\textsuperscript{194}

In classic semiconductor physics, the dopants that provide free carriers are depicted by discrete defect levels below the edge of the density-of-states of the conduction band (taking donors as an example, Fig. 13a). The defect levels are so-called “hydrogen-like” levels, meaning they are not from a specific atomic orbit from the donor, but effective ones to account for the coulombic interaction between the positively charged ions. However, when the donor is a resonant one, its atomic orbit is used (a $\Delta$ level). As such, a resonant donor can be considered as a resonant acceptor (with the same atomic orbit), which is a broad concept for resonant doping. A donor-D$_0$ acceptor-acceptor resonance scheme is presented in Fig. 13b. Thus, the resonant donors, such as As$_2$Te$_3$,\textsuperscript{193} and half-Heuslers,\textsuperscript{191,192} are considered as resonant donors.

In their alloys, the two sets converged around the composition of Bi$_0.25$Sb$_1.75$Te$_3$ (a larger effective mass\textsuperscript{171,172} at this composition is also an indicator). In Bi$_2$Se$_3$, the primary electron pocket is\textsuperscript{173} the one along $Z$--$Z$; alloying with Bi$_2$Te$_3$ allows these two pockets to converge with the six pockets in Bi$_2$Te$_3$ around the composition\textsuperscript{172} of Bi$_2$Te$_2.7$Se$_0.3$.

Assuming it is comparable to its counterpart\textsuperscript{167} in Sb$_2$Te$_3$, the second set of six hole pockets in Bi$_2$Te$_3$ should have a quality factor (0.20, take $\kappa_L$ same as Bi$_2$Te$_3$) that is close to the first set ($B_C$ in Bi$_2$Te$_3$). On the other hand, assuming the two pockets along $Z$--$Z$ in Bi$_2$Te$_2.7$Se$_0.3$ have characteristics from Bi$_2$Se$_3$, we can see that they have a much lower quality factor (0.09, take $\kappa_L$ same as Bi$_2$Te$_3$). While convergence happened in both alloys, different valleys were converged. The higher quality factor of the second band made the p-type Bi$_0.25$Sb$_1.75$Te$_3$ the one\textsuperscript{174–177} with higher $zT$, while in binary Bi$_2$Te$_3$ it was the opposite.

Additionally, from the modeling above, when the two bands have similar $B_C$ $zT$s will be close to the maximum when the two are very close, regardless of which one is the primary. Thus, we expect to observe $zT$s close to the maximum without significant deviation.

\textbf{Fig. 10} (a) The calculated band structure of La$_{3-x}$Te$_x$ close to the conduction band minimum, showing secondary bands around 0.3 eV above the primary band. (b) The Pisarenko relation at 400 K (inset 1000 K) shows an increase in the effective mass as carrier density increases. The red curve is from first-principles calculations based on the calculated band structure. The blue curve is from a multi-parabolic band model with effective masses and band offsets from first-principles calculations. The black curve is from a semi-empirical multi-parabolic band model, where the effective mass of the first band was adjustable and set to 0.844. From ref. 165, used with permission, Copyright the American Physical Society.
cores and the free electrons. These large polaron levels are below the band minimum so carriers are free only when the orbits from neighboring defects overlap or activation energy (thermal or electromagnetic) is provided to dissociate them.
from cores. In heavily doped semiconductors, the discrete defect levels interact with each other, become “band-like”, and overlap with the conduction band. This results in metallic transport behavior and donors are always mobile.

In the case of resonant doping, the “hydrogen-like” levels from dopants are found above the minimum of the conduction band density of states. With energy relaxation, electrons will dissociate from these states forming free carriers in the conduction band. Different from regular dopants, resonant defects continue to interact with carriers (beyond normal Coulombic interactions that give rise to ionized scattering), which depends heavily on carrier energy and is strongest when the defect level aligns with the Fermi level (Fig. 13b). This interaction causes carriers to ‘resonate’ between continuous band states and the ‘discrete, defect states’ (keep in mind they are not actual atomic states, and the carriers do not actually transit between different states). It is worth noting that all known examples of resonant dopants are elements that have multiple valences. Although there is no definitive evidence yet, that multiple valences (such as Tl$^{1+}$ and Tl$^{3+}$) coexist in these cases (some recent study did suggest existence of mixed valence, in contrast with previous studies such as ref. 186,196).

Identification of resonant behavior is in most cases based on an increased effective $m^*$ either by density-of-states DFT calculations or increased Seebeck coefficient over the regular carrier density dependence (Pisarenko relation). In principle, strong energy-dependent resonance leads to Fermi level pinning, which is seen as the saturation of carrier concentration measured by the Hall effect when the dopant concentration changes. Additionally, upon sufficient co-doping or counter-doping with a regular dopant, the carrier concentration could increase or decrease again (un-pinning). The pinning and un-pinning behavior, combined with Seebeck coefficients, serve as a more definitive indicator of resonant doping experimentally. As seen in Fig. 14, the measured carrier densities showed clear pinning and un-pinning in PbTe with resonant dopants of Tl (ref. 200, 14b) and Ti (ref. 197, 14c), and PbSe with resonant dopants of Tl (ref. 195, 14a); the same is likely although not as clear in SnTe with resonant dopants of In (ref. 201, 14d).

Resonant doping could be beneficial, however, optimization is necessary to reach the best possible thermoelectric performance. Using Na and Ti co-doped PbSe as a platform, a descriptive model was proposed so carrier density and temperature-dependent properties can be modeled.

In this model, only the density-of-states distortion from resonant states was taken into account. Additional resonant scattering mechanisms were not considered, for the reasons discussed in ref. 186 and 195. The change in the density-of-states affects transport properties either through the change in carrier group velocity, or the change in the scattering probability, so each transport property can be expressed.

The distortion of DoS was described by an additional Lorentzian term to the total DoS:

$$g_{res}(E) = \frac{N_i H}{2\pi} \cdot \frac{1}{(E - E_{res})^2 + \left(\frac{\Gamma}{2}\right)^2}$$

$E_{res}$ is the central energy of the resonant levels/band. $\Gamma$ is the resonant bandwidth that marks the full-width-at-half-maximum of the Lorentzian distribution in eqn (24). Integrating eqn (24) over all energies gives $2N_i H \cdot (N_i$ is the total density of resonant dopants). $H$ is an adjustable, unitless parameter between 0 and 1, it can be seen as a measure of how strongly the localized defect level is ‘hybridized’ with band states.

To express the transport properties, the carrier group velocity has to be determined. ‘Resonating’ between the band states and localized states would reduce the group velocity of

![Fig. 13](https://www.materials-advances.com/fig13.png)

Fig. 13 An illustration of hydrogen-like levels from (a) regular dopants and (b) resonant dopants (ref. 195). In dilute cases, these levels are treated as non-interacting, discrete levels, whereas in degenerate cases, these levels form bands. In the case of regular dopants, the defect levels do not cause changes to the total density-of-states. In the case of resonant dopants, the additional density-of-states from defects is Lorentzian and interacts with the continuous band states, distorting the total density-of-states.
carriers; to capture this, the following relations were considered:

\[ v_g = \frac{1}{h} \left( \frac{dk}{dE} \right)^{-1} \]  

(25)

and

\[ g(E) = \frac{N_g k}{\pi^2} \frac{dk}{dE} \]

(26)

such that (for a non-parabolic, Kane band)

\[ v_g = \frac{1}{\hbar} N_g g(E)^{-1} k_{\text{band}}^2, \]

\[ k_{\text{band}} = \frac{1}{\hbar} \left[ 2m^*E \left( 1 + \frac{E}{E_g} \right) \right]^{1/2} \]

(27)

These are the key treatments of this model; complete equations can be found in ref. 195.

Note that three parameters were introduced in this model: \( E_{\text{res}}, \Gamma, \) and \( H. E_{\text{res}} \) can be determined by solving for \( E_{\text{res}} = E_g \) using the ‘pinned’ carrier concentration. For \( \Gamma \) and \( H, \) it is not clear if they can be determined from experiments. In the previous work they were taken as adjustable parameters. All three parameters are set as constants for a given resonant system. Fig. 15 illustrates how each of them affects the Seebeck coefficient and mobility at 300 K.

With reasonable accuracy, the modeled carrier density dependence of the Seebeck coefficients and mobility were consistent with experimental results in the applicable temperature range. Although similar calculations\(^{186}\) (same scattering mechanism, DoS from KKR-CPA) had successfully described the Seebeck coefficient in PbTe doped with Tl, the work on PbSe provided a complete picture that describes both Seebeck coefficients and carrier mobility, as well as their temperature dependence.

Based on this model, some general guidelines regarding resonant doping can be summarized:

1. Resonant doping leads to stronger mobility reduction and, therefore, is not always beneficial for thermoelectrics.\(^{202}\) However, a moderate increase (10–20%) in the power factor under optimum \( n \) is possible in some cases. Tl doping in PbTe is the best-known example of resonant doping but the \( zT \) was not higher than PbTe doped with regular dopants. Tl doping in PbSe is not beneficial for \( zT \). So far, the only case\(^{189}\) to achieve higher \( zT \) than regular cases with a resonant dopant is In in SnTe.

2. The impact of ‘resonance’ can be seen at different temperatures (assuming the nature of the defects does not...
change, only the result of DoS distortion). The unique ‘hump’ of the Seebeck coefficient within a certain carrier density range may not always be seen. In many cases, the impact on the Seebeck coefficient is similar to an increased effective mass.\textsuperscript{203}

(3) Further tuning of the carrier density with a regular dopant is likely needed. Optimization should involve co-doping or counter-doping.

(4) The optimum position for a resonant level/band is close to the continuous band edge, with a small offset into the band. This conclusion is consistent with the study using tight-binding plus transport property modeling.\textsuperscript{204}

(5) The DoS distortion causes the $zT$ to peak at higher Seebeck coefficients,\textsuperscript{205} whereas for systems approximated by a single parabolic band and deformation potential scattering, $zT$ tends to peak\textsuperscript{48,206} when the Seebeck coefficient is around 250 $\mu$V K$^{-1}$. This could be favorable for certain device applications.\textsuperscript{207}

This is an example of how transport property modeling can help to guide material optimization, which would otherwise be tedious to study on a trial-and-error basis. This is true even when two parameters in the model are obtained from fitting, and the modeling still needs a good amount of experimental input.

5. Outlook

Thermoelectrics often have better performance when they take advantage of complex physical phenomena. The challenge is to characterize complex phenomena in a manner that can be engineered using the simple tuning parameters, composition, and processing conditions, available to synthetic chemists. The analytic models discussed here represent simplifications that describe the essential physics of the complex phenomena using a few scalar parameters that can be changed with synthesis conditions. Although not an exact theory, the formulations are based on the correct physics so that they should characterize a large fraction of the effect in a way that can be interpolated and extrapolated, perhaps more accurately than a machine learning algorithm, for example.
Many of the parameters used in these models are obtained from fitting experimental measurements. Often they cannot be determined in advance using other properties; for instance, the alloy scattering potential $\Delta U$, which we only know qualitatively where it comes from and can only guess its order of magnitude. For some of the parameters used to characterize resonant scattering, even such a qualitative relation is not clear. In the case of point defects on thermal conductivity, we argued that a simple, analytical approach is possible where all parameters can be determined from observable properties. This is the case mostly for mass contrast or simple structures like Si, Ge, and rock-salt lead chalcogenides. More studies on diverse and complex materials are needed for such models to be predictive. Although easily interpolated to similar systems with small compositional variations, it is difficult to extrapolate by predicting the parameters in complex materials without full experimental characterization.

While we currently only use the fitting of experimental results to obtain these parameters, the rapid advances of computational methods may revolutionize the process. It is tedious experimental work to prepare and exactly measure the same samples with only one synthetic parameter changed. However, such a computational, virtual experiment, is now relatively easy, even if not exactly applicable (e.g. approximations to the physics or missing microstructural effects). This provides an immense opportunity for theory and experiment to collaborate toward an engineering goal.

One must keep in mind that the analytical models are essentially isotropic, spectral averages of various related physical phenomena. It should not be surprising when parameters that fit experiments differ from values determined computationally or even other experiments. The deformation coefficient $\Xi$ of lead chalcogenides can be very different from calculated values (for example, ref. 115). This is at least partly because of other carrier scattering mechanisms, such as deformation potential scattering from optical phonons, inter-valley scattering, and potential scattering from ionized impurities, which are not distinguished from the often-called ‘acoustic phonon’ deformation potential. The realization, for example, that optical rather than acoustic phonons are primarily responsible for temperature-dependent electron scattering in complex crystals leads us to think of the deformation coefficient $\Xi$ as an effective parameter for phonon scattering (rather than just acoustic phonons). Similarly, the alloy scattering potential $\Delta U$ is more complicated than just electronegativity or work function differences. Additionally, different crystal symmetry and atomic coordination could make the analytical equations slightly different. Overall, the basic modeling approach is a compromise between simplicity, generality, and accuracy.

Compared to inorganic semiconductors, thermoelectric transport properties in organic or composite semiconductors are very poorly understood. Taking organic semiconductors as an example, historically, researchers used theories such as variable range hopping based on amorphous inorganic semiconductors to explain general trends. The results were not satisfactory. Recently, this has been revisited and it was suggested that organic semiconductors can be understood with a transport theory similar to that used for crystalline semiconductors containing homogeneous itinerant states and a mobility edge due to their disordered nature. This is important progress towards successful modeling that can be used to optimize materials. The models are even more reliant on fitting to determine key parameters since their exact physical mechanisms are poorly understood. Given the large variations in reported properties (presumably rooted in complex polymeric chemistry and microstructure), there is much left to understand and study within the physics and modeling of transport properties in organic semiconductors.

Keep in mind that the selection of thermoelectric materials is not all about higher $zT$s. It is true that for both power generation and cooling performance, $zT$ is the only metric, given no limitations on other device design factors such as geometry and interfaces: higher (average) $zT$ means better performance. This statement is, however, for a single leg made of a uniform material only. A better $zT$ for a leg does not always translate to better device performance. Material design is a multi-faceted problem when device-building is the ultimate goal.

**Conflicts of interest**

There are no conflicts to declare.

**Acknowledgements**

H. W. acknowledges the start-up support from Illinois Institute of Technology, and the support from the Educational and Research Initiative Fund provided by Illinois Institute of Technology. R. L. G. and G. J. S. acknowledge the support of award 70NANB19H005 from U.S. Department of Commerce, National Institute of Standards and Technology as part of the Center for Hierarchical Materials Design (CHiMaD). C. F. and T. Z. acknowledge the support by the Fundamental Research Funds for the Central Universities.

**References**

1. L. E. Bell, *Science*, 2008, 321, 1457–1461.
2. G. J. Snyder, S. LeBlanc, D. Crane, H. Pangborn, C. E. Forest, A. Rattner, L. Borgsmiller and S. Priya, *Joule*, 2021, 5, 748–751.
3. Global Thermoelectric Modules Market Size And Forecast To 2025, https://www.marketwatch.com/press-release/global-thermoelectric-modules-market-size-and-forecast-to-2025-2020-04-09, Accessed May, 2021.
4. G. J. Snyder and A. H. Snyder, *Energy Environ. Sci.*, 2017, 10, 2280–2283.
5. A. F. Ioffe, *Semiconductor thermoelements, and Thermoelectric cooling*, Infosearch, Ltd, 1957.
6. F. D. Rosi, E. F. Hockings and N. E. Lindenblad, *RCA Rev.*, 1961, 22, 82–121.
7. D. A. Wright, *Metall. Rev.*, 1970, 15, 147–160.
8 G. A. Slack, in CRC Handbook of Thermoelectrics, ed. D. M. Rowe, CRC Press, Boca Raton, 1995, pp. 407–440.

9 T. Caillat, A. Borshchevsky and J. P. Fleural, J. Appl. Phys., 1996, 80, 4442–4449.

10 G. S. Nolas, D. T. Morelli and T. M. Tritt, Annu. Rev. Mater. Sci., 1999, 29, 89–116.

11 V. L. Kuznetsov, L. A. Kuznetsova, A. E. Kaluzhin and D. M. Rowe, J. Appl. Phys., 2000, 87, 7871–7875.

12 G. J. Snyder, M. Christensen, E. Nishibori, T. Caillat and B. B. Iversen, Nat. Mater., 2004, 3, 458–463.

13 E. S. Toberer, A. F. May and G. J. Snyder, Chem. Mater., 2010, 22, 624–634.

14 T. C. Holgate, R. Bennett, T. Hammel, T. Caillat, S. Keyser and B. Sievers, J. Electron. Mater., 2014, 44, 1814–1821.

15 A. M. Ibrahim and D. A. Thompson, Mater. Chem. Phys., 1985, 12, 29–36.

16 B. Lenoir, H. Scherrer and T. Caillat, in Semiconductors and Semimetals, ed. T. M. Tritt, Elsevier, 2001, vol. 69, pp. 101–137.

17 V. K. Zaitsev, M. I. Fedorov, E. A. Gurieva, I. S. Eremin, P. P. Konstantinov, A. Y. Samunin and M. V. Vedernikov, Phys. Rev. B: Condens. Matter Mater. Phys., 2006, 74, 045207.

18 S. K. Bux, M. T. Yeung, E. S. Toberer, G. J. Snyder, R. B. Kaner and J.-P. Fleurial, J. Mater. Chem., 2011, 21, 12259.

19 Q. Shen, L. Chen, T. Goto, T. Hirai, J. Yang, G. P. Meisner and C. Uher, Appl. Phys. Lett., 2001, 79, 4165–4167.

20 C. Yu, T.-J. Zhu, R.-Z. Shi, Y. Zhang, X.-B. Zhao and J. He, Acta Mater., 2009, 57, 2757–2764.

21 L. D. Zhao, D. Berardan, Y. L. Pei, C. Byl, L. Pinsard-Gaudart and N. Dragoè, Appl. Phys. Lett., 2010, 97, 092118.

22 J. Sui, J. Li, J. He, Y.-L. Pei, D. Berardan, H. Wu, N. Dragoè, W. Cai and L.-D. Zhao, Energy Environ. Sci., 2013, 6, 2916.

23 H. L. Liu, X. Shi, F. F. Xu, L. L. Zhang, W. Q. Zhang, L. D. Chen, Q. Li, C. Uher, T. Day and G. J. Snyder, Nat. Mater., 2012, 11, 422–425.

24 K. Imasato, S. D. Kang, S. Ohno and G. J. Snyder, Mater. Horiz., 2018, 5, 59–64.

25 J. Mao, H. Zhu, Z. Ding, Z. Liu, G. A. Gamage, G. Chen and Z. Ren, Science, 2019, 365, 495–498.

26 X. Shi, H. Chen, F. Hao, R. Liu, T. Wang, P. Qiu, U. Burkhartd, Y. Grin and L. Chen, Nat. Mater., 2018, 17, 421–426.

27 J. Dong, F.-H. Sun, H. Tang, J. Pei, H.-L. Zhuang, H.-H. Hu, B.-P. Zhang, Y. Pan and J.-F. Li, Energy Environ. Sci., 2019, 12, 1396–1403.

28 L.-D. Zhao, G. Tan, S. Hao, J. He, Y. Pei, H. Chi, H. Wang, S. Gong, H. Xu, V. P. Dravid, C. Uher, G. J. Snyder, C. Wolverton and M. G. Kanatzidis, Science, 2016, 351, 141–144.

29 Z. Fan and J. Ouyang, Adv. Electron. Mater., 2019, 5, 1800769.

30 J. Tang, Y. Chen, S. R. McCuskey, L. Chen, G. C. Bazan and Z. Liang, Adv. Electron. Mater., 2019, 5, 1800943.

31 A. K. Menon, R. M. W. Wolfe, S. Kommandur and S. K. Yee, Adv. Electron. Mater., 2019, 5, 1800884.
57. D. Tuomi, J. Electrochem. Soc., 1984, 131, C91–C91.
58. D. Tuomi, J. Electrochem. Soc., 1984, 131, 2319–2325.
59. H. Wang, Y. Pei, A. D. LalLonde and G. J. Snyder, in Thermoelectric Nanomaterials, ed. K. Koumoto and M. Takao, Springer, Berlin, Heidelberg, 2013, vol. 182.

Electronic Archive on Physical Properties of Semiconductors at Ioffe Physical Technical Institute, http://www.ioffe.ru/SVA/NSM/Semicond/index.html, accessed Dec, 2021.
61. H. J. Goldsmid, Introduction to Thermoelectricity, Springer, Berlin, Heidelberg, 2010.
62. A. F. May and G. Snyder, in Materials, Preparation, and Characterization in Thermoelectrics: Thermoelectrics and Its Energy Harvesting, ed. D. M. Rowe, CRC Press, Boca Raton, 2012.
63. Y. I. Ravich, B. A. Efimova and I. A. Smirnov, Semiconducting lead chalcogenides, Plenum Press, New York, 1970.
64. W. Zawadzki, Adv. Phys., 1974, 23, 435–522.
65. W. M. Yin and F. D. Rosi, Solid-State Electron., 1972, 15, 1121–1140.
66. J. P. Dismukes, L. Ekstrom, E. F. Steigmeier, I. Kudman and D. B. Rabs, J. Appl. Phys., 1964, 35, 2899–2907.
67. H. J. Goldsmid, J. Appl. Phys., 1961, 32, 2198–2202.
68. P. G. Klemens, Proc. Phys. Soc., London, Sect. A, 1955, 68, 1113–1128.
69. J. Callaway, Phys. Rev., 1959, 113, 1046–1051.
70. P. G. Klemens, Phys. Rev., 1960, 119, 507–509.
71. R. Gurunathan, R. Hanus and G. J. Snyder, Mater. Horiz., 2020, 7, 1452–1456.
72. G. Tan, S. Hao, R. C. Hanus, X. Zhang, S. Anand, T. P. Bailey, A. J. E. Rettie, X. Su, C. Uber, V. P. Dravid, G. J. Snyder, C. Wolverton and M. G. Kanatzidis, ACS Energy Lett., 2018, 3, 705–712.
73. R. Gurunathan, R. Hanus, M. Dylla, A. Katre and G. J. Snyder, Phys. Rev. Appl., 2020, 13, 034011.
74. H. Wang, A. D. LaLonde, Y. Z. Pei and G. J. Snyder, Adv. Funct. Mater., 2013, 23, 1586–1596.
75. B. Abeles, Phys. Rev., 1963, 131, 1906–1911.
76. O. L. Anderson and J. E. Nafe, J. Geophys. Res., 1965, 70, 3951–3963.
77. D. L. Anderson and O. L. Anderson, J. Geophys. Res., 1970, 75, 3494–3500.
78. Y. I. Ravich, in Lead Chalcogenides Physics and Applications, ed. D. Khokhlov, Taylor & Francis, New York, 2003, vol. 18.
79. G. A. Slack and V. G. Tsoukala, J. Appl. Phys., 1994, 76, 1665–1671.
80. K. Koga, K. Akai, K. Oshiro and M. Matsura, Phys. Rev. B: Condens. Matter Mater. Phys., 2005, 71, 155119.
81. S. Ganeshan, S. L. Shang, Y. Wang and Z. K. Liu, J. Alloys Compd., 2010, 498, 191–198.
82. D. Rached, M. Rabah, N. Benkhettou, B. Soudini and H. Abid, Phys. Status Solidi B, 2004, 241, 2529–2537.
83. G. T. Alekseeva, B. A. Efimova, L. M. Ostrovskaya, O. S. Serebryanikova and M. I. Tsypin, Sov. Phys. Semiconductors, 1971, 4, 1122–1125.
84. E. F. Steigmeier and B. Abeles, Phys. Rev., 1964, 136, A1149–A1155.
85. V. K. Zaitsev, E. N. Tkaleenko and E. N. Nikitin, Sov. Phys. Solid State, 1969, 11, 221–224.
86. B. R. Ortiz, H. Peng, A. Lopez, P. A. Parilla, S. Lany and E. S. Toberer, Phys. Chem. Chem. Phys., 2015, 17, 19410–19423.
87. J. Yang, G. P. Meisner and L. Chen, Appl. Phys. Lett., 2004, 85, 1140–1142.
88. G. A. Slack, Phys. Rev., 1962, 126, 427–441.
89. R. Berman, E. L. Foster, J. M. Ziman and F. E. Simon, Proc. R. Soc. London, Ser. A, 1956, 237, 344–354.
90. G. P. Meisner, D. T. Morelli, S. Hu, J. Yang and C. Uher, Phys. Rev. Lett., 1998, 80, 3551–3554.
91. Y. Wang, F. Li, L. Xu, Y. Sui, X. Wang, W. Su and X. Liu, Inorg. Chem., 2011, 50, 4412–4416.
92. Y. Pei and D. T. Morelli, Appl. Phys. Lett., 2009, 94, 122112.
93. G. Tan, W. G. Zeier, F. Shi, P. Wang, G. J. Snyder, V. P. Dravid and M. G. Kanatzidis, Chem. Mater., 2015, 27, 7801–7811.
94. P. G. Klemens, Phys. B, 1999, 263–264, 102–104.
95. C. A. Ratsifaritana and P. G. Klemens, in Phonon Scattering in Condensed Matter, ed. H. J. Maris, Springer US, Boston, MA, 1980, pp. 259–262.
96. J. W. Harrison and J. R. Hauser, Phys. Rev. B: Solid State, 1976, 13, 5347–5350.
97. L. Makowski and M. Glikman, J. Phys. Chem. Solids, 1973, 34, 487–492.
98. L. M. Scarfone, Phys. Rev. B: Solid State, 1973, 7, 4435–4449.
99. K. I. Wysokinski, Phys. Status Solidi B, 1976, 77, K139–K142.
100. F. Goedsche and G. Vojta, Phys. Status Solidi B, 1978, 85, K81–K84.
101. J. W. Harrison and J. R. Hauser, J. Appl. Phys., 1976, 47, 292–300.
102. J. R. Hauser, M. A. Littlejohn and T. H. Glisson, Appl. Phys. Lett., 1976, 28, 458–461.
103. D. Chattopadhyay, Solid State Commun., 1994, 91, 149–151.
104. S. R. Mehrotra, A. Paul and G. Klimeck, Appl. Phys. Lett., 2011, 98, 173503.
105. H. Wang, J. Wang, X. Cao and G. J. Snyder, J. Mater. Chem. A, 2014, 2, 3169.
106. T. S. Stavitskaya, Investigation of Electronic Properties of Thermoelectric Lead Telluride and Some of Its Alloys, DSc thesis, A. F. Ioffe Physical-Technical Institute, Leningrad, 1968.
107. H. Wang, Z. M. Gibbs, Y. Takagiwa and G. J. Snyder, Energy Environ. Sci., 2014, 7, 804.
108. Y. Lee, S.-H. Lo, J. Androulakis, C.-I. Wu, L.-D. Zhao, D.-Y. Chung, T. P. Hogan, V. P. Dravid and M. G. Kanatzidis, J. Am. Chem. Soc., 2013, 135, 5152–5160.
109. H. Brooks, Theory of the Electrical Properties of Germanium Silicon, Academic Press Inc, NY, 1955.
110. M. A. Littlejohn, J. R. Hauser, T. H. Glisson, D. K. Ferry and J. W. Harrison, Solid-State Electron., 1978, 21, 107–114.
111. K. Biswas, J. Q. He, Q. C. Zhang, G. Y. Wang, C. Uher, V. P. Dravid and M. G. Kanatzidis, Nat. Chem., 2011, 3, 160–166.
C. Felser, D. J. Singh, T. Zhu, J. Yang and J. Ma, Nat. Commun., 2020, 11, 3142.

209 J. D. Wiley and M. Didomeni, Phys. Rev. B: Solid State, 1970, 2, 427–433.

210 M. V. Fischetti and S. E. Laux, J. Appl. Phys., 1996, 80, 2234–2252.

211 R. D’Souza, J. Cao, J. D. Querales-Flores, S. Fahy and I. Savić, Phys. Rev. B: Condens. Matter Mater. Phys., 2020, 102, 115204.

212 H. Wang, X. Cao, Y. Takagiwa and G. J. Snyder, Mater. Horiz., 2015, 2, 323–329.

213 S. Ihnatsenka, X. Crispin and I. V. Zozoulenko, Phys. Rev. B: Condens. Matter Mater. Phys., 2015, 92, 035201.

214 C. Juhasz and S. O. Kasap, J. Phys. D: Appl. Phys., 1985, 18, 721–729.

215 S. D. Kang and G. J. Snyder, Nat. Mater., 2017, 16, 252–257.

216 S. A. Gregory, R. Hanus, A. Atassi, J. M. Rinehart, J. P. Wooding, A. K. Menon, M. D. Losego, G. J. Snyder and S. K. Yee, Nat. Mater., 2021, 20, 1414–1421.

217 M. T. Agne, F. R. L. Lange, J. P. Male, K. S. Siegert, H. Volker, C. Pistorak, A. Poitz, T. Siegrist, S. Maier, G. J. Snyder and M. Wuttig, Matter, 2021, 4, 2970–2984.

218 P. Walmsley, C. Liu, A. D. Palczewski, P. Giraldo-Gallo, C. G. Olson, I. R. Fisher and A. Kaminski, Phys. Rev. B, 2018, 98, 184506.

219 T. Keiber, F. Bridges, B. C. Sales and H. Wang, Phys. Rev. B, 2013, 87, 144104.

220 S. Waddington, P. Weightman and A. D. C. Grassie, J. Phys. C: Solid State Phys., 1988, 21, 2695–2700.