# Prototypical many-body signatures in transport properties of semiconductors

Matthias Pickem ◉, Emanuele Maggio ◉, and Jan M. Tomczak ◉
*Institute of Solid State Physics, TU Wien, A-1040 Vienna, Austria*

We devise a methodology for charge, heat, and entropy transport driven by carriers with finite lifetimes. Combining numerical simulations with analytical expressions for low temperatures, we establish a comprehensive and thermodynamically consistent phenomenology for transport properties in semiconductors. We demonstrate that the scattering rate (inverse lifetime) is a relevant energy scale: It causes the emergence of several characteristic features in each transport observable. The theory is capable to reproduce, with only a minimal input electronic structure, the full temperature profiles measured in correlated narrow-gap semiconductors. In particular, we account for the previously elusive low-$T$ saturation of the resistivity and the Hall coefficient, as well as the (linear) vanishing of the Seebeck and Nernst coefficients in systems such as $FeSb_2$, $FeAs_2$, $RuSb_2$, and $FeGa_3$.

## I. INTRODUCTION

Transport properties, such as resistivity ($\rho$), magnetoresistance (MR), thermal conductance ($\kappa$), and the coefficients of Hall, Seebeck, and Nernst ($R_H$, $S$, $\nu$), are among the most widely investigated quantities in materials science. They provide essential information for characterizing new materials and for elucidating physical phenomena. To extract microscopic information from measurements requires a fundamental understanding of how carriers transport charge, heat, and entropy. When simulating transport properties, an adequate inclusion of scattering processes is particularly crucial. These limit the lifetime of carriers, lead to a decay of currents, and can have various origins, such as electron-electron or electron-phonon interactions as well as defects or impurities.

Here, we establish the prototypical signatures of finite electronic lifetimes in transport properties of (narrow-gap) semiconductors [1,2]. To this end, we build on the Kubo methodology of Ref. [3], which captures (in)coherence effects beyond the reach of semiclassical Boltzmann approaches [4–6], while incurring a comparable numerical cost. Our theory reveals that the scattering is a *relevant energy scale*: Through its interplay with the gap energy, a rich temperature profile emerges in all observables. In this scenario, *extrinsic* effects only enter indirectly, by limiting the lifetimes of *intrinsic* carriers. Previous semiclassical techniques, where scattering merely scales the amplitude of conduction, often require explicit in-gap states to provide additional carriers so as to mimic the experimental temperature dependence.

We highlight this insight in Fig. 1 for a simple two-band modeling of the colossal-thermopower material $FeSb_2$ [7–10]. The large *magnitude* of its $S$ and $\nu$ originates from the phonon-drag effect [11–16].[1] Here, we focus on the *characteristic temperatures* that mark features across various

observables [9]: for instance, inflection points in the resistivity $\rho$ and the Seebeck coefficient $S$ correlate with maxima in the Hall and Nernst coefficients $R_H$, $\nu$. This intriguing, but by no means uncommon [1,17,18], temperature profile, was previously advocated to derive from *extrinsic* in-gap states [14–16,19–21].

Here, instead, we reproduce all qualitative temperature features in $FeSb_2$ in the absence of explicit in-gap levels, exclusively by endowing the *intrinsic* valence and conduction carriers with a *finite scattering rate* $\Gamma(T) = \Gamma_0 + \gamma T^2$. Our findings establish a phenomenology for transport properties in semiconductors: Below a temperature $T_\rho^*$, $\rho$ saturates [3] instead of growing exponentially [see the Boltzmann result (dashed line) in Fig. 1]. $R_H$ also saturates (below $T_{R_H}^* < T_\rho^*$), indicating that residual scattering leads to a finite density of states even at absolute zero. Finite lifetimes also mend the violation of the third law of thermodynamics of Boltzmann's relaxation-time approximation: instead of diverging, the Seebeck coefficient $S$ vanishes linearly for $T \to 0$. Combined to the power factor $S^2\sigma$ and the figure of merit $zT$, our findings have practical relevance for thermoelectric applications. In narrow-gap semiconductors, these quantities exhibit large values at intermediate to low temperatures when scattering processes are properly accounted for. Material surveys based on Boltzmann approaches for coherent band structure instead fail to even qualitatively describe $S^2\sigma$ and $zT$. Finally, in congruence with experiment, a sharp low-$T$ feature emerges in the Nernst coefficient that, again, has no analog in the Boltzmann treatment.

Looking at the available experimental literature, we find a number of narrow-gap semiconductors [1] that exhibit qualitatively similar temperature profiles as the ones displayed in Fig. 1. For instance, other marcasite

---

[1]The phonon enhancement of the electron diffusion is expected to be smooth in temperature, i.e., it does not introduce characteristic features. Further, a suppression of the phonon drag in polycrystalline

samples yields Seebeck coefficients [12,13,105,106] comparable to our modeling.
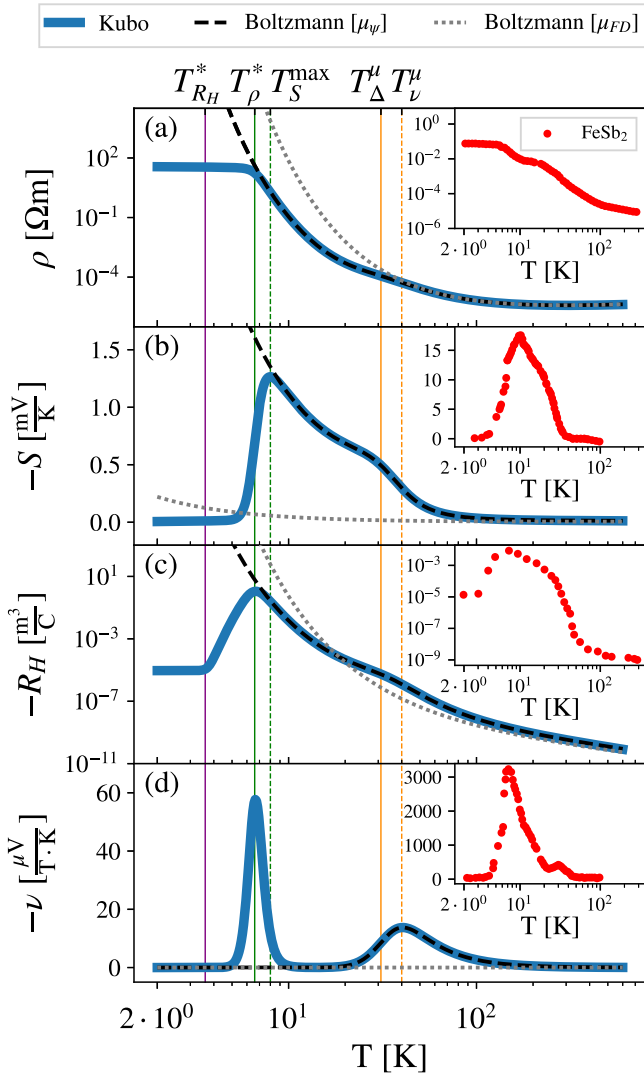
FIG. 1. Prototypical transport in semiconductors. Comparing a two-band model $\epsilon^0_{\mathbf{k}n} = -\sum_{i=x,y,z} 2t_n \cos(k_i) + (-1)^n(6t_n + \Delta_0/2)$, with $t_1 = 250$ meV, $t_2 = -265$ meV, band gap $\Delta_0 = 60$ meV; effective mass $Z^{-1} = 2$, scattering rate $\Gamma(T) = (5 \times 10^{-5} + 10^{-7}/\mathrm{K}^2\, T^2)$ eV to experiments on FeSb$_2$ (Insets) from Ref. [9]. (a) Resistivity, coefficients of (b) Seebeck, (c) Hall, and (d) Nernst *without* impurity in-gap states. Vertical lines mark characteristic temperatures (from left to right): saturation onset of the Hall coefficient $T^*_{R_H}$ and the resistivity $T^*_\rho$, the maximal Seebeck coefficient $T^{\max}_S$, onset of second activated regime $T^\mu_\Delta$, and the high-temperature Nernst peak $T^\mu_\nu$. Dashed black (gray) lines show Boltzmann results using a chemical potential $\mu_\psi$ ($\mu_{FD}$), that accounts for lifetime and thermal (only thermal) broadening.

compounds (FeAs$_2$, RuSb$_2$ [17,22–24], CrSb$_2$ [25]), silicides (FeSi [26–28], RuSi [29–31]), Heusler systems (e.g., Fe$_2$VAl [32–36]), other intermetallic compounds [37] (e.g., FeGa$_3$, RuGa$_3$ [18,38,39]), as well as Kondo insulators (e.g., Ce$_3$Bi$_4$Pt$_3$ [3,40–45]), strongly suggesting that our scenario based on carriers with finite lifetimes is prototypical for a wide array of different systems.

The paper is organized as follows: After introducing the formal background in Sec. II, we present the methodological advances of our approach in Sec. III. Then, in Sec. IV, we

set out to establish a phenomenology of transport properties of (narrow-gap) semiconductors. In Sec. IV A we provide analytical results for the low-temperature asymptotic behavior of all considered transport functions. In Sec. IV B we survey how the size of the charge gap $\Delta$, the particle-hole asymmetry $\alpha$, and the scattering rate $\Gamma$ control the temperature profile of observables. Crucially, we elucidate how finite lifetimes conspire with $\Delta$ and $\alpha$ to cause the emergence of several characteristic energy scales that appear in all transport observables. In Sec. V, we discuss implications for materials simulations on the basis of a simple modeling of selected intermetallic narrow-gap semiconductors [1]. Finally, we end with a summary in Sec. VI.

## II. METHODOLOGICAL CONTEXT

In this section, we lay out the setting in which we consider transport properties. The purpose is mainly to introduce the considered transport quantities and some necessary notation. For more detailed derivations, the reader is referred to specialized literature [46–48], as well as Refs. [49,50]. The conceptual advances beyond this setting will be presented in Sec. III, with more details in Appendix A.

### A. Linear response

In linear response, transport quantities are based on correlation functions that specify measurable observables of a system in the presence of external perturbations (electric field, magnetic field, temperature gradient, etc.). In our case these processes are described, on the imaginary-time ($\tau$) axis, by

$$\chi^{\alpha\beta}_{\mathbf{j}^a\mathbf{j}^b}(\mathbf{q}, \tau) = \frac{1}{V}\langle \mathcal{T}_\tau j^a_\alpha(\mathbf{q}, \tau) j^b_\beta(-\mathbf{q}, 0)\rangle \tag{1}$$

with the time-ordering operator $\mathcal{T}_\tau$, the charge ($a, b = 1$) and heat ($a, b = 2$) current operator $\mathbf{j}^a_\alpha$ in the Cartesian direction $\alpha, \beta \in \{x, y, z\}$, and $V$ indicating the unit-cell volume. From them, the usual (retarded) Onsager coefficients $\mathcal{L}$ for dipolar transitions ($\mathbf{q} = 0$) are obtained by first Fourier transforming Eq. (1) into bosonic Matsubara frequencies

$$\chi^{\alpha\beta}_{\mathbf{j}^a\mathbf{j}^b}(\mathbf{q}, i\omega_n) = \int_0^\beta d\tau\, e^{i\omega_n\tau} \chi^{\alpha\beta}_{\mathbf{j}^a\mathbf{j}^b}(\mathbf{q}, \tau), \tag{2}$$

analytical continuation to real frequencies $i\omega_n \rightarrow \omega + i\delta$ and then taking the dynamic limit

$$\mathcal{L}^{\alpha\beta}_{ab} = \lim_{\omega \rightarrow 0^+} \frac{1}{\omega} \mathrm{Im}[\chi^{\alpha\beta}_{\mathbf{j}^a\mathbf{j}^b}(\mathbf{q} = 0, \omega)]. \tag{3}$$

In the presence of an external magnetic field $B$ in direction $\gamma \in \{x, y, z\}$, one needs to instead evaluate the expectation value

$$\chi^{B,\alpha\beta\gamma}_{\mathbf{j}^a\mathbf{j}^b}(\mathbf{q}, \tau) = \frac{1}{V}\langle T_\tau j^a_\alpha(\mathbf{q}, \tau) j^b_\beta(\mathbf{0}, 0)\rangle_{B_\gamma} \tag{4}$$

for the Hamiltonian that includes the field [49,51,52], and the resulting Onsager coefficients will be denoted

$$\mathcal{L}^{B,\alpha\beta\gamma}_{ab} = \lim_{\omega \rightarrow 0^+} \frac{1}{\omega} \mathrm{Im}[\chi^{B,\alpha\beta\gamma}_{\mathbf{j}^a\mathbf{j}^b}(\mathbf{q} = 0, \omega)]. \tag{5}$$

From the above, the observable transport tensors can be derived. Namely, the conductivity $\sigma$, the resistivity $\rho$, the thermopower (Seebeck coefficient) $S$, the electronic thermal conductivity $\kappa$, the Hall conductivity $\sigma^B$, the Hall coefficient $R_H$, and the Nernst coefficient $\nu$:

$$\sigma_{\alpha\beta} = \mathcal{L}_{11}^{\alpha\beta}, \tag{6}$$

$$\rho_{\alpha\beta} = \left(\mathcal{L}_{11}^{-1}\right)^{\alpha\beta}, \tag{7}$$

$$S_{\alpha\beta} = -\frac{1}{T}\left(\mathcal{L}_{11}^{-1}\right)^{\alpha i}\mathcal{L}_{12}^{i\beta}, \tag{8}$$

$$\kappa_{\alpha\beta} = \frac{1}{T}\left[\mathcal{L}_{22}^{\alpha\beta} - \mathcal{L}_{12}^{\alpha i}\left(\mathcal{L}_{11}^{-1}\right)^{ij}\mathcal{L}_{12}^{j\beta}\right], \tag{9}$$

$$\sigma_{\alpha\beta\gamma}^B = \mathcal{L}_{11}^{B,\alpha\beta\gamma}, \tag{10}$$

$$R_{H,\alpha\beta\gamma} = \left(\mathcal{L}_{11}^{-1}\right)^{\alpha i}\mathcal{L}_{11}^{B,ij\gamma}\left(\mathcal{L}_{11}^{-1}\right)^{j\beta}, \tag{11}$$

$$\nu_{\alpha\beta\gamma} = -\frac{1}{T}\left(\mathcal{L}_{11}^{-1}\right)^{\alpha i}\left[\mathcal{L}_{11}^{B,ij\gamma}\mathcal{L}_{12}^{jk} - \mathcal{L}_{12}^{B,ij\gamma}\mathcal{L}_{11}^{jk}\right]\left(\mathcal{L}_{11}^{-1}\right)^{k\beta}. \tag{12}$$

For models and materials whose unit cell's axes are orthogonal, as we are considering here, the Einstein summations over Cartesian directions simplify to a single expression, e.g., for an electric current in the $x$ direction and a magnetic field in the $z$ direction, the resulting Hall coefficient in the $y$ direction is $R_{H,xyz} = (\mathcal{L}_{11}^{-1})^{xx}\mathcal{L}_{11}^{B,xyz}(\mathcal{L}_{11}^{-1})^{yy}$. Later, we will also address the (empirical) Wiedemann-Franz law

$$L = \frac{\kappa}{\sigma T} \tag{13}$$

as well as the thermoelectric power factor

$$PF = S^2\sigma \tag{14}$$

and figure of merit

$$zT = \frac{S^2\sigma T}{\kappa}. \tag{15}$$

### B. Integral equations

#### 1. One vs multiparticle correlations

The transport observables, even when featuring one-particle currents $\mathbf{j}_\alpha^a$,[2] in Eqs. (1) and (4) probe multiparticle excitations. Diagrammatically, Eq. (1) can be described as the sum of all possible two-particle processes, with the leading term corresponding to the independent propagation of a particle-hole pair (bubble diagram). Magnetotransport quantities on the other hand stemming from Eq. (4) can be shown [49–52] to involve all possible three-particle processes. Diagrams that (unlike the bubble) contain interconnected propagation lines are commonly referred to as vertex corrections [53–56]. These can lead to collective phenomena such as excitons, $\pi$-tons [57–59], and other polaritons. In this work, following the spirit of the dynamical mean-field theory [60], we are neglecting vertex corrections. In this approximation,

which amounts to the infinite-dimensional limit, vertex corrections vanish for all considered observables [60–63].[3]

Assuming that the one-particle Green's function $G_\mathbf{k}(\omega)$ is diagonal in the chosen band or orbital basis, the Onsager coefficients (3)–(5), can be written as

$$\mathcal{L}_{ab}^{\alpha\beta} = \frac{\pi\hbar e^{(4-a-b)}}{V}\sum_{\substack{n,m \\ \mathbf{k},\sigma}}\mathcal{K}_{ab}(\mathbf{k},n,m)M^{\alpha\beta}(\mathbf{k},n,m), \tag{16}$$

$$\mathcal{L}_{ab}^{B,\alpha\beta\gamma} = \frac{4\pi^2\hbar e^{(5-a-b)}}{3V}\sum_{\substack{n,m \\ \mathbf{k},\sigma}}\mathcal{K}_{ab}^B(\mathbf{k},n,m)M^{B,\alpha\beta\gamma}(\mathbf{k},n,m) \tag{17}$$

with the electron charge $e$, and the sums running over band indices $n$, $m$, Brillouin zone momentum $\mathbf{k}$, and spin $\sigma$. Here, the $M^{(B)}$ collect the dipolar transition matrix elements that depend on the Cartesian directions $\alpha$, $\beta$ (and $\gamma$) and are given in the next paragraph. The kernel functions $\mathcal{K}^{(B)}$, instead, contain the two- (three-) particle expectation value of the fermionic operators that make up the currents $\mathbf{j}^a$. Neglecting vertex corrections (see above), they can be expressed as

$$\mathcal{K}_{ab}(\mathbf{k},n,m) = \int_{-\infty}^{\infty}d\omega\,\omega^{(a+b-2)}\left(-\frac{\partial f}{\partial\omega}\right)A_{\mathbf{k}n}(\omega)A_{\mathbf{k}m}(\omega), \tag{18}$$

$$\mathcal{K}_{ab}^B(\mathbf{k},n,m) = \int_{-\infty}^{\infty}d\omega\,\omega^{(a+b-2)}\left(-\frac{\partial f}{\partial\omega}\right)A_{\mathbf{k}n}^2(\omega)A_{\mathbf{k}m}(\omega), \tag{19}$$

where $A_{\mathbf{k}n}(\omega) = -1/\pi\,\mathrm{Im}G_{\mathbf{k}n}(\omega)$ is the spectral function associated with the retarded one-particle Green's function. Energies $\omega$ are measured with respect to the Fermi level $\mu$. Thus, within our approximations, many-body (scattering) effects enter the transport properties only through the renormalization of the one-particle propagators.

#### 2. Transition-matrix elements

Concerning the transition-matrix elements, we make use of the Peierls approximation [62,64–66]. Therein, Fermi velocities are the derivative of the bare dispersion[4] $v_{\mathbf{k}n}^\alpha = 1/\hbar\partial_{k_\alpha}\epsilon_{\mathbf{k}n}^0$, and the matrix elements $M^{(B)}$ in Eqs. (16) and (17) can be expressed as [49]

$$M^{\alpha\beta}(\mathbf{k},n,n) = v_{\mathbf{k}n}^\alpha v_{\mathbf{k}n}^\beta, \tag{20}$$

$$M^{B,\alpha\beta\gamma}(\mathbf{k},n,n) = \varepsilon_{\gamma ij}v_{\mathbf{k}n}^\alpha c_{\mathbf{k}n}^{\beta i}v_{\mathbf{k}n}^j, \tag{21}$$

where $\varepsilon_{\gamma ij}$ is the Levi-Civita symbol and the curvature of the bare dispersion is encoded in $c_{\mathbf{k}n}^{\alpha\beta} = 1/\hbar\partial_{k_\alpha}\partial_{k_\beta}\epsilon_{\mathbf{k}n}^0$. Standard Fermi velocities obtained in the band basis only account for intraband transitions. In a more general framework [51,62],

---

[2]The heat-current $\mathbf{j}^2$ is only of one-particle nature when, as we assume here, interactions are local [62].

[3]at least in the absence of multi-band effects [66]. For the vanishing of vertex corrections in infinite dimensions for massless fermions, see Ref. [109].

[4]When the Hamiltonian of the system is expressed by (continuum) field operators, the charge density commutes with the interaction terms.

also interband transitions can be included in a Peierls-type fashion.

## III. METHODOLOGICAL ADVANCEMENT

### A. Approximation: Linearized self-energy

As seen in Eqs. (18) and (19), the derivative of the Fermi function ensures that transport properties are dominated by energies close to the Fermi level.[5] Then, also in the quantity that encodes many-body renormalizations, the electron self-energy $\Sigma$, only the low-energy behavior is relevant. Hence, for the purpose of transport properties and in the absence of polelike structures within several $k_B T$ of the Fermi level, the self-energy can be linearized:

$$\Sigma_{\mathbf{k}n}(\omega) \approx \mathrm{Re}\,\Sigma_{\mathbf{k}n}(0) + \left(1 - Z_{\mathbf{k}n}^{-1}\right)\omega - i\Gamma_{\mathbf{k}n}^0. \quad (22)$$

In other words, the central assumption is that for transport properties the temperature dependence of renormalizations is more important than that on frequency. An implicit higher-frequency dependence can, however, be included by linearizing the self-energy around the band energies $\epsilon_{\mathbf{k}n}^0$. For the scattering rate, for instance, instead of evaluating $\Gamma_{\mathbf{k}n}^0 = -\mathrm{Im}\,\Sigma_{\mathbf{k}n}(\omega = 0)$ at the Fermi level, one can use $\Gamma_{\mathbf{k}n}^0 = -\mathrm{Im}\,\Sigma_{\mathbf{k}n}(\omega = \epsilon_{\mathbf{k}n}^0)$.

With Eq. (22), the coherent part of the spectrum (of weight $Z$) is of Lorentzian form

$$A_{\mathbf{k}n}(\omega) = \frac{Z_{\mathbf{k}n}}{\pi} \frac{\Gamma_{\mathbf{k}n}}{(\omega + \mu - \epsilon_{\mathbf{k}n})^2 + \Gamma_{\mathbf{k}n}^2} \quad (23)$$

with $\Gamma = Z\Gamma^0$ and $\epsilon = Z[\epsilon^0 + \mathrm{Re}\,\Sigma(0)]$ the renormalized scattering rate and dispersion, respectively.

### B. Linear response transport quantities for finite scattering

The central innovation of this paper is the observation that, in the current setting, the integrals in Eqs. (18) and (19) can be performed analytically, circumventing costly and (for small $\Gamma$) unstable numerical integrations. Indeed, also the evaluation of the particle number simplifies, one finds [11]

$$N = \sum_{\mathbf{k},n,\sigma} \int_{-\infty}^{\infty} d\omega\, f(\omega) A_{\mathbf{k}n}(\omega) = \sum_{\mathbf{k},n,\sigma} \left( \frac{1}{2} - \frac{1}{\pi} \mathrm{Im}\,\psi(z_{\mathbf{k}n}) \right) \quad (24)$$

with the digamma function $\psi$ evaluated at $z_{\mathbf{k}n} = \frac{1}{2} + \frac{\beta}{2\pi}[\Gamma_{\mathbf{k}n} + i(\epsilon_{\mathbf{k}n} - \mu)]$, where $\beta = 1/(k_B T)$ is the inverse temperature.[6] Finite lifetimes (inverse scattering rate) explicitly enter through the digamma function, describing the thermal and lifetime smearing of excitations on an equal footing. Consequently, the energy states now obey a $\Gamma$-modified

[6]In this expression, the quasiparticle weight $Z$ has been set to one, as, e.g., customary in slave-boson approaches. This procedure implicitly assumes the transfer of spectral weights $1 - Z$ to be symmetrical in the sense that it does not alter the chemical potential $\mu$. If a many-body electronic structure and, thus, $\mu$ is provided, the self-energy is only linearized in the transport kernels.
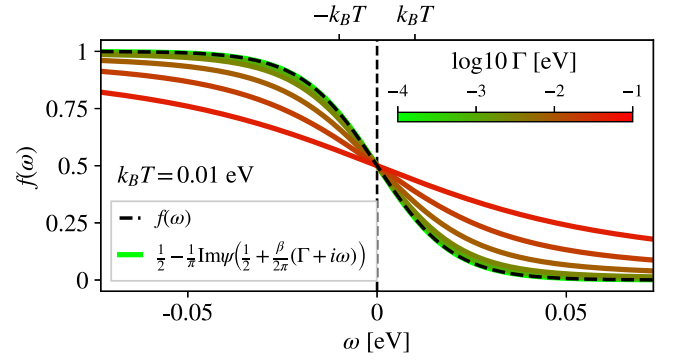


FIG. 2. Lifetime enhanced broadening. Comparison between purely thermal broadening [dashed black line, Fermi function $f(\omega)$] and lifetime-enhanced broadening described by Eq. (24) for various scattering rates (solid colored lines).

Fermi-Dirac statistic, displayed in Fig. 2. Crucially, even for $T = 0$ this distribution is not steplike, provided that $\Gamma > 0$. In Sec. IV B, we explore the impact of the carrier density behavior on the chemical potential and all derived transport properties.

Treating the smearing of excitations due to thermal ($\beta$) and lifetime ($\Gamma$) effects on an equal footing also affects the transport kernels. For the intraband transport kernels $\mathcal{K}_{ab}^{(B)}(\mathbf{k}, n) \equiv \mathcal{K}_{ab}^{(B)}(\mathbf{k}, n, n)$ of Eqs. (18) and (19) we derive the LINRETRACE expressions (see Appendix A)

$$\mathcal{K}_{11}(\mathbf{k}, n) = \frac{Z^2 \beta}{4\pi^3 \Gamma} \left[ \mathrm{Re}\,\psi_1(z) - \frac{\beta\Gamma}{2\pi} \mathrm{Re}\,\psi_2(z) \right], \quad (25)$$

$$\mathcal{K}_{12}(\mathbf{k}, n) = \frac{Z^2 \beta}{4\pi^3 \Gamma} \left[ a\,\mathrm{Re}\,\psi_1(z) - \frac{a\Gamma\beta}{2\pi} \mathrm{Re}\,\psi_2(z) - \frac{\Gamma^2 \beta}{2\pi} \mathrm{Im}\,\psi_2(z) \right], \quad (26)$$

$$\mathcal{K}_{22}(\mathbf{k}, n) = \frac{Z^2 \beta}{4\pi^3 \Gamma} \left[ (a^2 + \Gamma^2)\mathrm{Re}\,\psi_1(z) + \frac{\beta}{2\pi}\Gamma(\Gamma^2 - a^2)\mathrm{Re}\,\psi_2(z) - \frac{\beta}{\pi}a\Gamma^2\mathrm{Im}\,\psi_2(z) \right], \quad (27)$$

$$\mathcal{K}_{11}^B(\mathbf{k}, n) = \frac{Z^3 \beta}{16\pi^4 \Gamma^2} \left[ 3\,\mathrm{Re}\,\psi_1(z) - \frac{3\Gamma\beta}{2\pi}\mathrm{Re}\,\psi_2(z) + \frac{\Gamma^2\beta^2}{4\pi^2}\mathrm{Re}\,\psi_3(z) \right], \quad (28)$$

$$\mathcal{K}_{12}^B(\mathbf{k}, n) = \frac{Z^3 \beta}{16\pi^4 \Gamma^2} \left[ 3a\,\mathrm{Re}\,\psi_1(z) - \frac{3a\Gamma\beta}{2\pi}\mathrm{Re}\,\psi_2(z) - \frac{\Gamma^2\beta}{2\pi}\mathrm{Im}\,\psi_2(z) + \frac{a\Gamma^2\beta^2}{4\pi^2}\mathrm{Re}\,\psi_3(z) + \frac{\Gamma^3\beta^2}{4\pi^2}\mathrm{Im}\,\psi_3(z) \right], \quad (29)$$

$\mathcal{K}_{22}^B(\mathbf{k}, n)$

$$
= \frac{Z^3 \beta}{16\pi^4 \Gamma^2} \bigg[ (\Gamma^2 + 3a^2) \text{Re}\psi_1(z) - \frac{\beta \Gamma(\Gamma^2 + 3a^2)}{2\pi} \text{Re}\psi_2(z)
$$

$$
- \frac{\beta a \Gamma^2}{\pi} \text{Im}\psi_2(z) - \frac{\beta^2 \Gamma^2 (\Gamma^2 - a^2)}{4\pi^2} \text{Re}\psi_3(z)
$$

$$
+ \frac{\beta^2 a \Gamma^3}{2\pi^2} \text{Im}\psi_3(z) \bigg], \tag{30}
$$

where $\psi_i$ is the $i$th derivative of the digamma function $\psi$ evaluated at $z = \frac{1}{2} + \frac{\beta}{2\pi}(\Gamma + ia)$, with $a = \epsilon - \mu$. Momentum $\mathbf{k}$ and band $n$ indices of $Z$, $\Gamma$, and $a$ have been omitted for "brevity."

The above equations allow for a simple symmetry analysis. For instance, we see that $\mathcal{K}_{12}$ is odd with respect to $a$. As a consequence, electron and hole contributions to the Seebeck coefficient (8) have opposite signs. $\mathcal{K}_{11}^B$, instead, is even in $a$. Therefore, the Hall coefficient $R_H$ [Eq. (11)] actually does not distinguish electron and hole contributions through their energies (the sign of $a$), but thanks to the sign of the dispersion's curvature entering the matrix element $M^B$ in Eq. (20).

Manifestly, the above kernel functions are far more complicated than the familiar expressions of the semiclassical Boltzmann approach in the constant relaxation-time approximation (cf., e.g., Refs. [2,4–6,67]). However, the latter are recovered from the above formulas as the leading terms in the limit of infinite lifetimes, i.e., $\Gamma \to 0^+$. This is most readily seen for the kernel underlying the conductivity: Noting that $\beta/(2\pi^2)\text{Re}\psi_1[1/2 + i\beta a/(2\pi)] = -f'(a)$, one finds (see Appendix C), to leading order in the scattering rate, the familiar expression

$$
\sigma_{\alpha\beta} = \frac{e^2 Z^2}{V} \sum_{\mathbf{k}\sigma} \tau v_{\mathbf{k}}^\alpha v_{\mathbf{k}}^\beta (-\partial f / \partial \omega)_{\omega = \epsilon_{\mathbf{k}} - \mu}, \tag{31}
$$

with the lifetime $\tau = \hbar/(2\Gamma)$, and a renormalization factor $Z^2$ commonly not included. From this point of view, the inclusion of finite lifetimes in Eqs. (25)–(30) leads to the following:

(1) An effectively different statistic: the Fermi function is replaced with a digamma function in which thermal broadening is supplemented by an energy smearing $\Gamma$ corresponding to finite lifetimes (cf. Fig. 2).

(2) All transport kernels have, beyond the explicitly leading terms (e.g., $1/\Gamma \text{Re}\Psi_1$ in $\mathcal{K}_{11}$), contributions from higher-order polygamma functions $\Psi_i$ ($i > 1$).

We will discuss the effects of both, later on.

Computational results in this paper have been obtained using the transport properties code LINRETRACE [68]. While the Eqs. (25)–(30) are quite involved, they are, in fact, on par with Boltzmann approaches employing the relaxation-time approximation in terms of numerical complexity and evaluation speed. In fact, many Boltzmann codes [4–6,69–72] could readily upgrade their electronic transport capabilities by switching to the kernels presented here. Conversely, LINRETRACE [68] could benefit from being interfaced with code packages [69,70,72,73] that provide electronic lifetimes from electron-phonon scattering.

## IV. A PHENOMENOLOGY OF TRANSPORT IN SEMICONDUCTORS

### A. Low-temperature expansion

Signatures of finite electronic lifetimes are largest at low temperatures, where qualitative deviations from Boltzmann behavior are most pronounced. Our first goal therefore is to provide *simple* phenomenological formulas for transport observables at low temperatures. To this end, we first expand the polygamma functions $\psi_i(1/2 + \bar{z})$ in the kernel equations (25)–(30) around $\bar{z} = \infty$, i.e., $T = 0$ (the resulting expressions can be found in Appendix B). Second, we assume a simplistic electronic structure; indeed, we note that in metals and semimetals only states in the direct vicinity of the Fermi level contribute sizably to conduction of charge and heat. In gapped systems, instead, transport will be dominated by the conduction (valence) band minimum (maximum). Both constraints effectively limit the parts of the Brillouin zone relevant to transport. To gain a qualitative insight, we therefore radically forgo the momentum integration in Eqs. (16) and (17): We consider a single nondispersive level at an energy $\epsilon$ and constant transition matrix elements $M^{(B)}$. Assuming further a scattering rate $\Gamma$ independent of temperature, we obtain the following prototypical dependencies:

$$
\sigma \propto e^2 \frac{Z^2}{\pi^2} \frac{\Gamma^2}{(a^2 + \Gamma^2)^2} \bigg[ 1 + \frac{2\pi^2}{3} \frac{5a^2 - \Gamma^2}{(a^2 + \Gamma^2)^2} (k_B T)^2 + O(T^4) \bigg], \tag{32}
$$

$$
\sigma^B \propto e^3 \frac{4Z^3}{3\pi^3} \frac{\Gamma^3}{(a^2 + \Gamma^2)^3} \bigg[ 1 + \pi^2 \frac{7a^2 - \Gamma^2}{(a^2 + \Gamma^2)^2} (k_B T)^2 + O(T^4) \bigg], \tag{33}
$$

$$
S \propto -\frac{k_B}{e} \frac{4\pi^2}{3} \frac{a}{a^2 + \Gamma^2}
$$
$$
\times \bigg[ k_B T + \frac{\pi^2}{15} \frac{55a^2 - 53\Gamma^2}{(a^2 + \Gamma^2)^2} (k_B T)^3 + O(T^5) \bigg], \tag{34}
$$

$$
R_H \propto \frac{4\pi}{3e} \frac{1}{Z\Gamma} \bigg[ a^2 + \Gamma^2 + \frac{\pi^2}{3} (k_B T)^2 + O(T^4) \bigg], \tag{35}
$$

$$
\kappa \propto \frac{Z^2}{3} \frac{\Gamma^2}{(a^2 + \Gamma^2)^2} \bigg[ k_B T - \frac{16\pi^2 a^2}{3(a^2 + \Gamma^2)^2} (k_B T)^3 + O(T^5) \bigg], \tag{36}
$$

$$
\nu \propto -\frac{8\pi^2}{9} k_B \frac{aZ\Gamma}{(a^2 + \Gamma^2)^2}
$$
$$
\times \bigg[ k_B T - \frac{4\pi^2}{3} \frac{8a^2 - \Gamma^2}{(a^2 + \Gamma^2)^2} (k_B T)^3 + O(T^5) \bigg], \tag{37}
$$

where $a = \epsilon - \mu$ indicates the position of the renormalized level $\epsilon = Z\epsilon^0$ with respect to the chemical potential $\mu$. We now discuss the above asymptotic behavior and compare to Boltzmann approaches in the relaxation-time approximation (see Table I for a summary of the $T \to 0$ limit). Note that the above equations describe the low-$T$ response for a single level. If several states are contributing, terms can be simply added up for the conductivities. For composite quantities, such

TABLE I. Zero-temperature limits of transport properties in stoichiometric gapped systems for a finite scattering rate $\Gamma$. Equations (32) and (35) lead to saturation in the resistivity and the Hall coefficient, while the Boltzmann signal diverges. Entropy transport complies with the laws of thermodynamics (Seebeck $S \to 0$ for $T \to 0$), while $S$ unphysically diverges in the relaxation-time approximation. $\nu_{1/2}$ denote the two contributions to the Nernst coefficient, Eq. (12). All limiting behaviors of LINRETRACE are congruent with experiments (see Figs. 1 and 11).

| $\lim T \to 0$ | LINRETRACE | Boltzmann |
|---|---|---|
| $\rho$ | $\rho_{\mathrm{sat}}$ | $\infty$ |
| $S$ | 0 | $\infty$ |
| $\kappa$ | 0 | 0 |
| $R_H$ | $R_{H,\mathrm{sat}}$ | $\infty$ |
| $\nu = \nu_1 - \nu_2$ | 0 | 0 |
| $\nu_{1/2}$ | 0 | $\infty$ |

as the Seebeck or the Hall coefficient, however, contributions to the Onsager coefficients [Eq. (3)] have to be summed before they are combined into the observable quantities.

*a. Charge transport.* One of the main observations is that in the zero-temperature limit $T \to 0$ the electrical ($\sigma$) [3] and Hall conductivity ($\sigma^B$) remain finite in the presence of residual scattering ($\Gamma > 0$). In Eqs. (18) and (19) the derivative of the Fermi function becomes increasingly narrow with decreasing temperature. Alone, this temperature-dependent energy cutoff would lead to the typical activated behavior and is well described in Boltzmann theory. However, for $\Gamma > 0$, the Lorentzian shape of the spectral function (23) allows states away from the Fermi level to still contribute to conduction even at $T = 0$, as incoherent spectral weight spills into the gap [see Fig. 3(b) and cf. the effective particle distribution function in Fig. 2]. This residual conductivity is at the heart of the resistivity saturation in (nontopological) Kondo insulators and has been discussed in detail in Ref. [3]. Similarly, the Hall conductivity in Eq. (10) and, hence, the Hall coefficient in Eq. (11) saturate for $\Gamma > 0$. The Boltzmann approximation [see Eq. (31)], on the other hand, relies solely on the Fermi function to select states with sharply defined energies $\epsilon$. Then, conductivities must strictly approach zero in gapped systems

for $T \to 0$. Since the electrical and the Hall conductivity have the identical temperature scaling, Eq. (11) implies a diverging Hall coefficient in Boltzmann's relaxation-time approximation [see Fig. 1(c)].[7]

*b. Thermoelectric transport.* In essence, the Seebeck and Nernst effects can be understood as entropy carried by charged currents [47,48,74]. The third law of thermodynamics states that at zero temperature the entropy $S_0$ of the system must be minimal. In a perfect crystal lattice without ground-state degeneracy this minimal value must be 0 since there is only one possible microstate [$S_0 = k_B \ln(\Omega); \Omega = 1$], requiring the Seebeck as well as the Nernst coefficient to vanish for $T \to 0$. This is respected in our framework: Similar to the case of metals [75], we find $S \sim T$ ($T \to 0$) in a semiconductor with finite lifetimes. As discussed in more detail in Sec. IV B 3 b, it is residual conduction from incoherent states that leads to a weakly metal-like Seebeck coefficient. In the Boltzmann limit, instead, $S$ unphysically diverges in a semiconductor: $S(T) \propto \frac{1}{T}$.

The Nernst coefficient vanishes in both formalisms. In the Boltzmann case, this is hidden in the two terms making up Eq. (12) ($\nu = \nu_1 - \nu_2$): while both $\nu_i$ ($i = 1, 2$) diverge, they cancel exactly when combined. In the Kubo formalism, both terms $\nu_i$ separately approach 0. Further, for finite $\Gamma$, $\nu \propto T$ at lowest temperatures, again akin to the behavior of metals [76] and (see Sec. IV B 5 b) connected to conduction from intrinsic, but incoherent, in-gap states.

In all, Eqs. (32)–(37) establish a low-temperature phenomenology of transport in semiconductors. The derived asymptotic behavior overcomes limitations of semiclassical descriptions and is congruent with experimental observations (see Fig. 1 above and Fig. 11 below).

### B. Prototypical transport properties of narrow-gap systems

In this section we leave the low-temperature limit and study the full temperature dependence of the transport observables

---

[7]The higher-order kernels $\mathcal{K}_{12}^{(B)}$ and $\mathcal{K}_{22}^{(B)}$ are accompanied by an additional $\omega$ and $\omega^2$ factor in Eqs. (18) and (19), respectively. Therefore, the active energy window is additionally suppressed, causing these kernels to vanish for $T \to 0$.
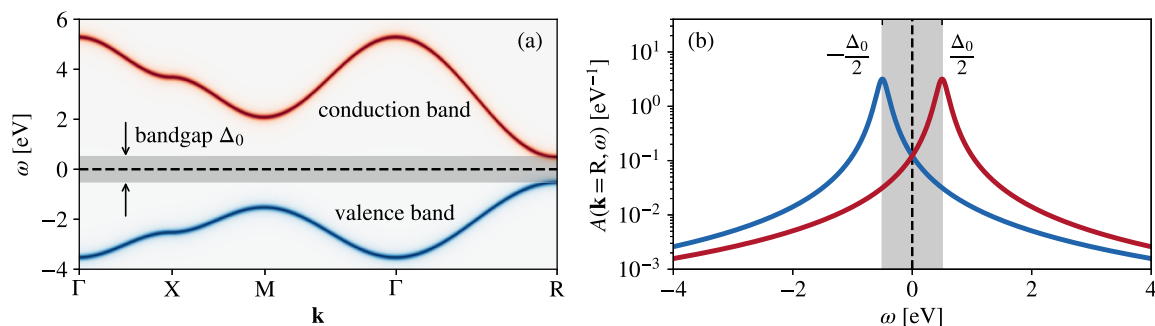


FIG. 3. Model electronic structure. (a) Broadened band structure $\epsilon_{\mathbf{k}n}^0 = -\sum_{i=x,y,z} 2t_n \cos(k_i) + (-1)^n (6t_n + \Delta_0/2)$ with $t_1 = 0.25$ eV, $t_2 = -0.40$ eV, and $\Delta_0 = 1.0$ eV. The broadening $\Gamma = 0.1$ eV is exaggerated in order to illustrate the effects. (b) Spectral functions at $\mathbf{k} = \mathrm{R} = (\pi, \pi, \pi)$ where the direct gap is exactly $\Delta_0$ (we here set the quasiparticle weight to unity $Z = 1$). The Lorentzian form (23) of the spectral functions causes weight to spill over into the band gap. As a consequence, for $T \to 0$, transport properties exhibit metal-like signatures, such as residual conduction (=resistivity saturation) and a linear-in-$T$ Seebeck coefficient.

from Eqs. (6)–(12). Motivated by experimental transport measurements of intermetallic hybridization-gap semiconductors, such as $FeSb_2$, $FeAs_2$, FeSi, and $FeGa_3$, their Ru analogs and others, we consider a simple, asymmetric two-band electronic structure

$$\epsilon_{\mathbf{k}n}^0 = - \sum_{i=x,y,z} 2t_n \cos(k_i) + (-1)^n (6t_n + \Delta_0/2) \quad (38)$$

with $n = 1, 2$ for the valence and conduction band, respectively, and fix the filling to $N = 2$ (half-filling). We use a generic lattice constant $a_{\text{lattice}} = 1$ Å (cf. Appendix D) and $60 \times 60 \times 60$ ($200 \times 200 \times 200$) $k$ points for the Kubo (Boltzmann) calculations to achieve $k$-grid convergence. Particle-hole asymmetry is introduced by hopping parameters $t_n$ that are different for the valence band (VB, $n = 1$) and the conduction band (CB, $n = 2$). We measure the degree of asymmetry via

$$\alpha = \left| \frac{t_2}{t_1} \right|. \quad (39)$$

The two bands of the narrow-gap semiconductor are then additionally endowed with the same, finite and, for the time being, temperature-independent scattering rate $\Gamma$, while we set the quasiparticle weight to unity, $Z = 1$. This setup generalizes the symmetric ($\alpha = 1$) two-band model for which we studied the resistivity in Ref. [3]. An example electronic structure is displayed in Fig. 3. Even though motivated by said materials, we keep the electronic structure deliberately simple, so as to isolate qualitative trends and the prototypical temperature dependencies of transport properties. A more realistic setup with a temperature-dependent scattering rate will be introduced in Sec. IV C.

The three parameters that describe the transport equations of our model are the band gap $\Delta_0$, the band asymmetry $\alpha$, and the scattering rate $\Gamma$. As a first step we simply scan through each parameter individually while keeping the other two fixed. The results for these parameter sweeps are shown in Fig. 4. The clear protagonist of transport properties beyond Boltzmann results based on band structures is the scattering rate $\Gamma$. It influences transport in two ways: First, $\Gamma > 0$ leads, in a particle-hole asymmetric system, to a nontrivial temperature dependence of the chemical potential, which, in turn, influences charge and energy transport. Boltzmann approaches that use band structures as input fully miss this ingredient as only thermal (not lifetime) broadening is included in the chemical potential search. Second, contributions to the transport kernels (25)–(30) are, contrary to Boltzmann approaches in the relaxation-time approximation, not simply proportional to the carrier lifetime $1/\Gamma$, but exhibit an intricate $\Gamma$ dependence that influences the temperature profile of transport properties. In the following, we will disentangle these two ingredients.

### 1. Chemical potential and activated carriers

For sharply defined valence and conduction states $\epsilon_{\mathbf{k}}$, i.e., $\Gamma = 0$, the position of the chemical potential $\mu$ is driven through the thermal broadening of the Fermi function. In particular, one can show (see, e.g., Ref. [11]) that $\mu$ approaches the middle of the gap for $T \to 0$, with a temperature slope

that depends on the particle-hole asymmetry. If the lifetime of valence and conduction states is finite, this description is no longer valid. Indeed, determined via Eq. (24), the chemical potential is intrinsically dependent on the scattering rate. Figure 3(b), that displays a spectral function at a selected $k$ point, illustrates why this is the case: the Lorentzian width of the spectral function results in a small but finite weight of incoherent in-gap states that the chemical potential has to account for. As seen in the bottom panel of Fig. 4, $\mu$ follows the result of the Fermi function at high temperatures, where thermal broadening dominates over the Lorentzian in-gap weight. Below a temperature $T_\nu^\mu$, however, $\mu$ starts to strongly deviate. In particular, it no longer extrapolates to the midgap point (here set to zero), but to a finite value that increases with growing particle-hole asymmetry $\alpha$ and band gap $\Delta_0$, while being only weakly dependent on the residual scattering $\Gamma$. This behavior can be rationalized through the low-temperature expansion of the occupation in Eq. (24), which, for a state at $a = \epsilon - \mu$, yields

$$N = \frac{1}{2} - \frac{1}{\pi} \text{Im} \ln(\Gamma + ia) + \frac{\pi}{3} \frac{a\Gamma}{(a^2 + \Gamma^2)^2} (k_B T)^2 + O(T^4). \quad (40)$$

Given that, in a semiconductor, the scattering rate $\Gamma$ is orders of magnitude smaller than the band energies, changes in the band structure ($a = \epsilon - \mu$) will dominate the chemical potential at low temperatures. Noteworthy, the evolution of the chemical potential $\mu$ shown in the lower panels of Fig. 4 [its deviation from the intrinsic Fermi-Dirac result (gray dashed in bottom right panel) at $T_\nu^\mu$, its inflection point $T_\Delta^\mu$, and the eventual saturation at a finite position] is reminiscent of the exhaustion and extrinsic regime in semiconductors with impurity-derived in-gap states [77]. There, changes in the chemical potential are driven by donated electrons or holes, i.e., a change in the total number of charge carriers. In our scenario, instead, the total number of electrons stays constant, but the finite lifetimes of intrinsic carriers causes excitations to widen, spilling incoherent spectral weight into the gap, so that the chemical potential has to adapt. Consequently, even at lowest temperatures, the number of activated carriers

$$n_e = \sum_{\mathbf{k}, n \geqslant \text{CB}, \sigma} \left( \frac{1}{2} - \frac{1}{\pi} \text{Im} \psi(z_{\mathbf{k}n}) \right), \quad (41)$$

$$n_h = \sum_{\mathbf{k}, n \leqslant \text{VB}, \sigma} \left( \frac{1}{2} + \frac{1}{\pi} \text{Im} \psi(z_{\mathbf{k}n}) \right) \quad (42)$$

*must* remain finite for $\Gamma > 0$, as shown in Fig. 4 (second panel row from the bottom). In more detail, at any temperature, the number of activated electrons $n_e$ and holes $n_h$ (per unit cell) necessarily balance each other, $n_e = n_h$, in the stoichiometric (half-filled) case considered here. At high temperatures $n_e$ follows the result for the coherent ($\Gamma = 0$) band structure (indicated in dashed gray). In this regime, dominated by thermal activation across the gap $\Delta_0$, $n_e$ is exponentially suppressed upon cooling; for $k_B T \ll \Delta_0$: $n_e(\Gamma = 0; T) \sim e^{-\Delta_0/k_B T}$. For finite $\Gamma$, the discussed deviations in the chemical potential reflect in the carriers available
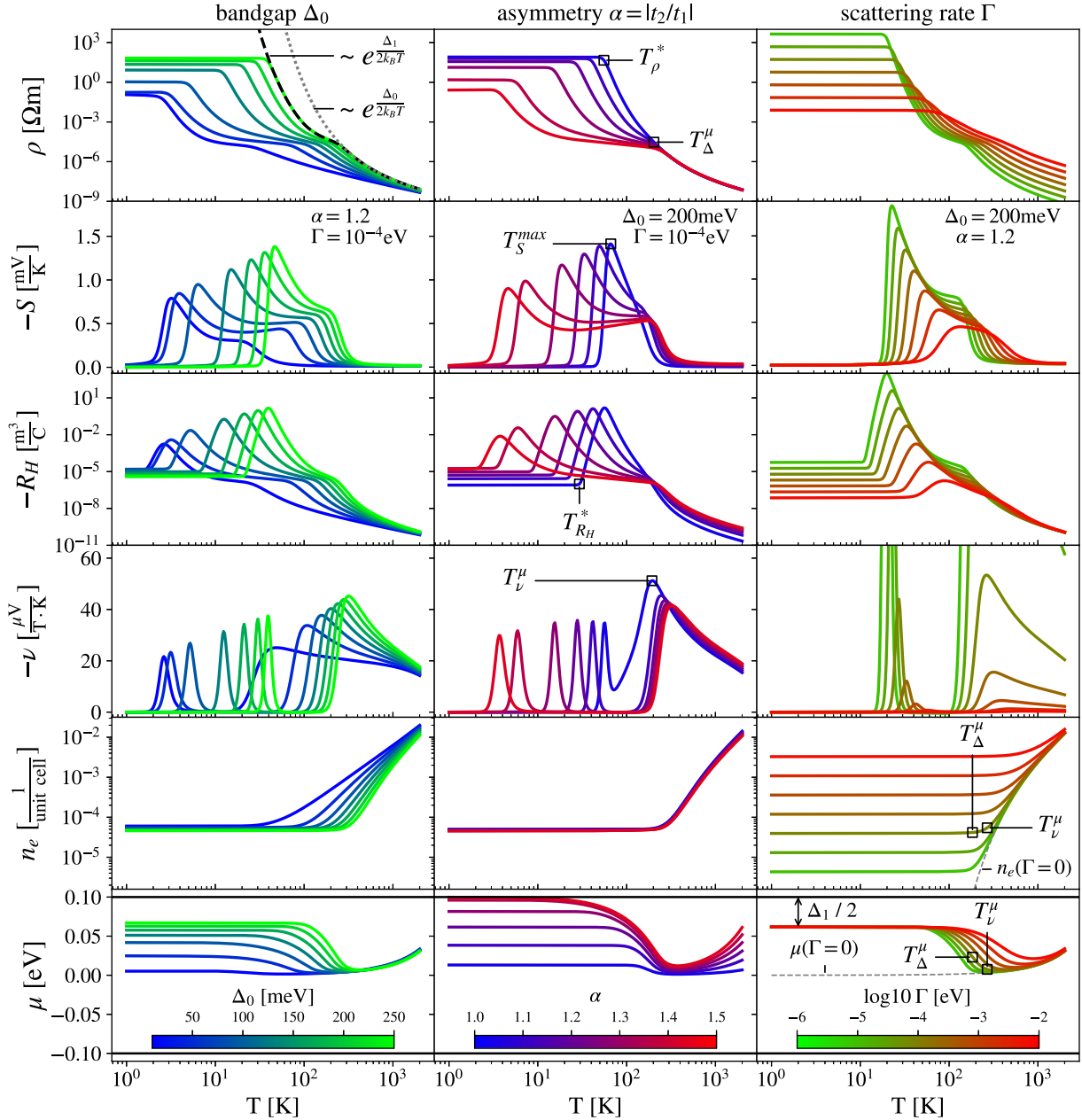
FIG. 4. Transport in semiconductors: a parameter scan. Left column: varying band gaps $\Delta_0$ for fixed hoppings $t_1 = 0.25$ eV and $t_2 = -0.30$ eV (asymmetry $\alpha = 1.2$) and scattering rate $\Gamma = 10^{-4}$ eV; middle column: varying asymmetries $\alpha$ for fixed band gap $\Delta_0 = 200$ meV and scattering rate $\Gamma = 10^{-4}$ eV; right column: varying scattering rate $\Gamma$ for fixed $\Delta_0 = 200$ meV and $\alpha = 1.2$. Throughout, a quasiparticle weight $Z = 1$ is used. From top to bottom we show the resistivity $\rho$, the coefficients of Seebeck $S$, Hall $R_H$, Nernst $\nu$, the activated number of electrons $n_e$ (holes $n_h = n_e$), and the chemical potential $\mu$. Notable characteristics: For finite $\Gamma$, the asymmetry $\alpha$ leads to a chemical potential that does not converge to the gap midpoint for $T \to 0$. This $\Gamma$-induced deviation in $\mu$ causes the high-temperature features at $T_\Delta^\mu$ and $T_\nu^\mu$. There, the resistivity transitions between an activated regime corresponding to the fundamental gap $\Delta_0$ to one with a reduced energy $\Delta_1$, given by twice the distance between $\mu(T \to 0)$ and the nearest band edge. The characteristic $T_S^{\max}$, $T_\rho^*$, and $T_{R_H}^*$ instead are driven by the quantum kernels and the importance of higher-order polygamma functions in them: the resistivity and the Hall coefficient saturate at $T_\rho^*$ and $T_{R_H}^*$, respectively. The latter signals a finite density of carriers at the Fermi level even at absolute zero, congruent with $n_e > 0$. The Seebeck coefficient is suppressed and vanishes for $T \to 0$ in accordance with the laws of thermodynamics. The Nernst coefficient peaks one more time before also trending to zero for $T \to 0$.

for conduction: Below the temperature $T_\nu^\mu$, $n_e$ no longer shrinks exponentially and, at $T_\Delta^\mu$, transitions into a regime in which the number of available carriers is virtually independent of temperature. In this low-temperature regime,

the dominant control parameter for the number of carriers is the scattering rate $\Gamma$, whereas asymmetry and the size of the gap hardly affect $n_e(T \to 0)$ on the shown exponential scale.

### 2. Electric resistivity

Signatures of the described evolution of the number of carriers available for conduction are readily seen in the resistivity in the top row of Fig. 4. The activated behavior above $T_\nu^\mu$, purely determined by the band gap $\Delta_0$, transitions into a second activated regime realized below $T_\Delta^\mu$, with an effectively reduced band gap $\Delta_1 < \Delta_0$. As indicated in Fig. 4 (bottom right panel), $\Delta_1$ measures (twice) the distance between the saturated $\mu(T \lesssim T_\Delta^\mu)$ and the nearest band edge (for $\alpha > 1$, the conduction band). Again, this behavior is reminiscent of an impurity-driven extrinsic regime. There, $\Delta_1$ would measure the difference between the extrinsic impurity level on the one hand and the conduction or valence band on the other. In both scenarios, extrinsic in-gap states vs intrinsic states with finite lifetimes, changes in conduction reflect a modification in the chemical potential. Here, our theory provides a complementary microscopic origin for the appearance of the chemical potential-driven characteristic temperature scales $T_\nu^\mu$ and $T_\Delta^\mu$. Note that for particle-hole symmetric systems, where the chemical potential is temperature independent, no such crossover exists and there is only one activationlike regime [3].

Cooling further, also the second activated regime is bounded from below: at a temperature $T_\rho^*$, the resistivity enters a saturation regime. Contrary to the features at higher temperatures, $T_\rho^*$ has no signature in the chemical potential, but derives entirely from the physics encoded in the kernel function (25). As discussed in detail in Ref. [3] the crossover temperature $T_\rho^*$ and the saturation value $\rho(T \to 0)$ strongly depend on the scattering rate and the band gap. Noteworthy, the influence of $\Gamma$ inverts as a function of temperature: at high $T$, a larger scattering rate increases the resistivity. This is the conventional behavior, also realized in metals. At low $T$, however, where conduction is driven by incoherent spectral weight inside the gap, the resistivity understandably *decreases* with a growing scattering rate (see top right panel in Fig. 4). Here, we extend the previous analysis [3] and demonstrate that also the band asymmetry $\alpha$ has a strong effect on the conduction. In the asymmetric case, the chemical potential must be positioned closer to the conduction ($\alpha > 1$) or valence ($\alpha < 1$) band so that the correct number of electrons in the system is occupied. Therewith, the majority of carriers, those that reside in the centers of the Lorentz-broadened peaks in the spectral function, conduct more and freeze-out at a lower temperature, i.e., $T_\rho^*$ and the corresponding saturation value $\rho(T \to 0)$ decreases with $\alpha$.

To summarize, the resistivity of an *intrinsic* narrow-gap semiconductor with a finite scattering rate has four regimes: (1) $T > T_\nu^\mu$, the activated high-temperature region that is well described in Boltzmann theory; (2) $T_\Delta^\mu < T < T_\nu^\mu$, a narrow regime in which the chemical potential starts to sense the incoherent spectral weight inside the gap and adjusts accordingly; (3) $T_\rho^* < T < T_\Delta^\mu$, a regime at intermediate temperatures in which the ($\Gamma$-imposed) chemical potential shift has led to a reduced activation energy for valence ($\alpha < 1$) or conduction ($\alpha > 1$) carriers; (4) $T < T_\rho^*$, a regime of resistivity saturation in which thermal activation is frozen out but a residual conductivity, driven by incoherent in-gap weight, remains finite.

### 3. Seebeck coefficient

The three temperatures $T_\rho^* < T_\Delta^\mu < T_\nu^\mu$ that separate the four regimes in the electrical resistivity also account for features in the Seebeck coefficient (second row in Fig. 4): the increase of $S$ starting from high temperatures is interrupted by the crossover of the chemical potential at $T_\nu^\mu$. Depending on the parameters, the transition to the maximum amplitude at lower temperatures can then either be smooth (large scattering rate) and monotonous (large gap) or be accompanied by a significant shoulder (large band gap, small scattering rates). In extreme cases this shoulder transforms into a local peak (small gap, strong asymmetry), i.e., the temperature dependence can be nonmonotonous. The temperature at which the Seebeck coefficient has its global maximum amplitude $S^{\max}$ is linked to $T_\rho^*$. In fact, $S$ peaks at $T_S^{\max}$, consistently slightly above the onset of the resistivity saturation regime ($T_S^{\max} \gtrsim T_\rho^*$). Below this global peak temperature, the Seebeck coefficient drops rather abruptly. In the zero-temperature limit, it follows the metal-like linear behavior $S(T) \sim T$, anticipated in Sec. IV A.

This rich structure is *absent* when the Boltzmann approach is applied to the band structure $\epsilon_{\mathbf{k}n}^0$ of Eq. (38). The features associated with $T_\Delta^\mu$ and $T_\nu^\mu$ are missed if finite lifetimes are unaccounted for in the search of the chemical potential; the characteristic features further below are absent owing to the simple structure of the Boltzmann transport kernels. Indeed, for a momentum- and state-independent scattering rate $\Gamma$, the kernels $\mathcal{L}_{11}$ and $\mathcal{L}_{12}$ in the Boltzmann approximation are both merely proportional to $\Gamma^{-1}$. Then, given by their ratio, Eq. (8), Boltzmann's Seebeck coefficient is independent of the scattering rate. Manifestly, this approximation is a severe oversimplification even for extremely small $\Gamma$ (see Fig. 4, right column, second panel from the top).

Next, we will comment on two approximate tools that are popular for the analysis of thermoelectric measurements or simulations.

*a. The Goldsmid-Sharp gap.* Goldsmid and Sharp [78] motivated that the size of a semiconductor's gap could be gauged from the peak amplitude of the Seebeck coefficient:

$$\Delta \approx 2e|S^{\max}| \cdot T_S^{\max}. \tag{43}$$

This estimate works decently for both *n*- and *p*-type semiconductors [78], although deviations of the order of a factor of 2 are not uncommon. The relation is used as a coarse analysis tool [79–81] in simulations and experiments and even as a descriptor in high-throughput materials discovery studies [82]; Eq. (43) was deduced for a coherent, large-gap, particle-hole symmetric semiconductor in which impurity states move the chemical potential so as to optimize the Seebeck coefficient [11,78]. Allowing for particle-hole asymmetry, the Seebeck coefficient can, however, be larger, while it is always bounded by $|S| \leqslant |\Delta/(eT) + S(\infty)|$, where $S(\infty)$ is the comparatively small high-temperature limit [$|S(\infty)| = O(k_B/e)$] [11]. Since the original argument [78] relies on replacing the Fermi-Dirac with the classical Maxwell-Boltzmann statistics, further deviations occur if $k_B T_S^{\max} \not\ll \Delta$ [83].
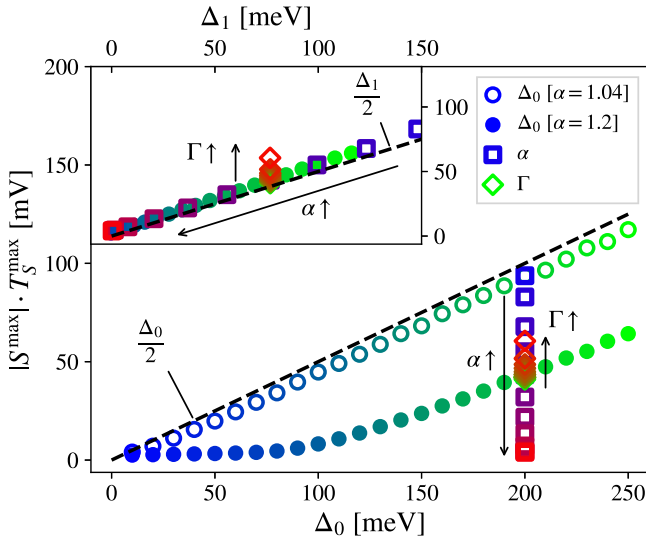
FIG. 5. Goldsmid-Sharp gap. We display the gap estimate $|S^{max}| \cdot T_S^{max}$ for scans of the band gap $\Delta_0$ for asymmetries $\alpha = 1.04$ (open circles) and $\alpha = 1.2$ (closed circles). Also shown is a scan of the asymmetry (open squares) and of the scattering rate (open diamonds), using the same parameters as in Fig. 4. The expression $|S^{max}| \cdot T_S^{max}$ deviates strongly from $\Delta_0/2$ (dashed line) when the system's asymmetry is at least moderate. Only in the vicinity of particle-hole symmetry ($\alpha = 1.04$, open circles) do we find good agreement for Eq. (43). An increase in asymmetry leads to a suppression of $|S^{max}| \cdot T_{max}$ while different scattering rates have minimal effects. Instead, plotting $|S^{max}| \cdot T_{max}$ against the effective gap $\Delta_1$ [cf. $\rho(T)$ in Fig. 4], the various scans collapse onto the $\Delta_1/2$ line (see inset): the Goldsmid-Sharp gap expression reliably estimates the effective gap $\Delta_1$.

Here, we scrutinize the Goldsmid-Sharp relation (43) for our two-band model: In Fig. 5 we report $|S^{max}| \cdot T_S^{max}$ extracted from the parameter scans of Fig. 4 as well as for an additional scan for an almost particle-hole symmetric system $\alpha = 1.04$ (open circles in the main panel). For the latter, the Goldsmid-Sharp gap indeed provides a very accurate estimation of the fundamental gap $\Delta_0$. For systems with more asymmetrical electronic structures, however, Eq. (43) yields poor results: $|S^{max}| \cdot T_S^{max}$ largely underestimates the charge gap (filled circles). Looking at the corresponding resistivities, Seebeck coefficients, and the chemical potential in Fig. 4 reveals the reason: The thermopower is largest at the lower end of the second activated regime of the resistivity $T_\rho^* \lesssim T_S^{max} < T_\Delta^\mu$. This regime emerges when a finite scattering rate pushes the chemical potential towards the lighter band (in our case the conduction band). As a consequence, the crucial gap, as apparent in the resistivity, is the effective $\Delta_1$, not $\Delta_0$. This observation suggests that the Goldsmid-Sharp gap does not monitor the band gap $\Delta_0$ but the effective gap $\Delta_1$. This hypothesis is confirmed in the inset of Fig. 5: plotted as a function of the emergent $\Delta_1$, the scatter plot of $|S^{max}| \cdot T_S^{max}$ collapses onto a single line indicating $\Delta_1/2$. This statement is largely independent of the scattering rate: $|S^{max}| \cdot T_S^{max}$ only slightly increases with $\Gamma$, leading to a mild overestimation of $\Delta_1$.

In all, in the realistic presence of a scattering rate, the Goldsmid-Sharp expression quite accurately gauges the effective gap $\Delta_1$ that controls transport at intermediate temperatures, but which can be significantly smaller than the true band gap $\Delta_0$.

*b. The Mott formula.* In metals, conduction is largely dominated by states in the vicinity of the chemical potential. Then, performing a low-temperature Sommerfeld expansion of the conductivity is justified. Doing so for the Boltzmann relaxation-time approximation (31) yields a convenient expression for the Seebeck coefficient

$$S \approx -\frac{\pi^2}{3e} k_B^2 T \frac{\partial \ln \sigma(\mu)}{\partial \mu}, \tag{44}$$

which is a simplified version of the so-called *Mott formula* of the thermopower [84,85]. Here, $\sigma(\mu)$ is the electrical conductivity for varying chemical potential $\mu$. Clearly, the above approximation is inaccurate for *coherent* semiconductors, where conduction is driven by conduction and/or valence states that are far (more than several $k_B T$) from the chemical potential. Manifestly, our general kernel functions therefore do not verify Eq. (44). However, as we demonstrated [3], finite lifetimes may drive residual conduction in semiconductors, leading to resistivity saturation. The incoherent in-gap states associated with this phenomenon might provide the metallicity required to justify expanding the derivative of the Fermi function around the chemical potential (Sommerfeld expansion). Therefore, Eq. (44) is expected to hold in the saturation regime, where conduction is dominated by said incoherent in-gap weight. Evaluating Eq. (44) for the residual term of the conductivity given in Eq. (32) indeed yields the lowest-order expression of the low-$T$ Seebeck coefficient (34). This validity of the Mott formula for metals in the resistivity saturation regime of a semiconductor provides a direct link between residual charge conduction [$\sigma(T \to 0) > 0$] and a metal-like linear-in-$T$ thermoelectric Seebeck signal [$S(T \to 0) \propto T$].

### 4. Hall coefficient

From high temperatures down to its maximum, the Hall coefficient $R_H$ (third panel in Fig. 4) exhibits a qualitatively similar dependency on $\Delta_0$, $\alpha$, and $\Gamma$ as the Seebeck coefficient. However, $R_H$ peaks at a slightly smaller temperature that matches $T_\rho^*$ from the resistivity. Also, instead of vanishing, $R_H$ saturates below a temperature $T_{R_H}^*$ ($< T_\rho^*$), in agreement with the low-$T$ expansion (35) and experiment (see Figs. 1 and 11). This temperature arrangement is expected since the spectral function enters Eq. (19) to higher order compared to Eq. (18). Therefore, the $\mathcal{K}_{11}^B$ kernel effectively senses a smaller amount of incoherent in-gap weight than $\mathcal{K}_{11}$. Consequently, the temperature, below which the thermal selection of valence and conduction carriers via $(-\frac{\partial f}{\partial \omega})$ can be neglected in comparison to the incoherent in-gap weight that drives the residual conduction, is effectively reduced. At large temperatures ($T > T_\nu^\mu$, Boltzmann regime), $R_H$ is dominantly controlled by the band gap and the particle-hole asymmetry, while the dependence on the scattering rate is weak. Instead, at low $T$, it is mostly the scattering rate that controls both the peak value and the saturation limit $R_H(T \to 0)$.
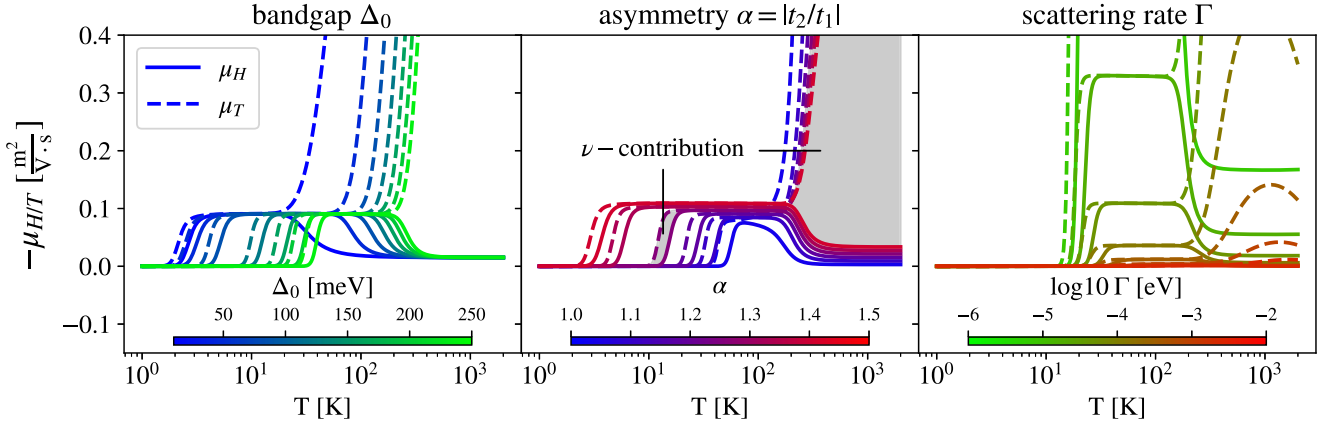
FIG. 6. Hall and thermal mobilities. Both mobilities vanish in the zero-temperature limit and coincide with each other at intermediate temperatures, marking the ranges where the Nernst coefficient vanishes in Fig. 4. This transition, however, takes place at slightly different temperature between $\mu_H$ and $\mu_T$ giving rise to the low-temperature $\nu$ contribution at roughly $T_\rho*$ (first shaded area). At large temperature the two mobilities diverge from each other. $\mu_T$ increases drastically while $\mu_H$ gets suppressed, marking the second $\nu$ contribution which is peaked at $T_\nu^\mu$ (second shaded area).

### 5. Nernst coefficient

The Nernst coefficient exhibits the most dramatic features.[8] Starting from high temperature, $\nu$ increases and reaches a first peak at $T_\nu^\mu$ where the chemical potential starts to transition towards its saturation regime. The Nernst then quickly drops to zero (unless the asymmetry is very small) and remains suppressed in the intermediate regime ($T_\rho^* < T < T_\Delta^\mu$; the second activation regime of $\rho$). The transition into the $\rho$-saturation regime at $T_\rho^*$ is then accompanied by a second large and sharp peak in $\nu$ before it finally converges linearly to zero for $T \rightarrow 0$. Similar to the Hall coefficient, the biggest changes in the temperature profile of $\nu$ are achieved by varying the band gap and the asymmetry, while absolute values are mostly controlled by the scattering rate.

Focusing on (a) the large low-temperature peak and (b) the $T \rightarrow 0$ limit, we analyze the fabric of the Nernst coefficient through different representations.

*a. Hall and thermal mobilities.* Using the expression (8) of the Seebeck coefficient, we can rewrite the Nernst coefficient (12) as

$$\nu = S(\mu_H - \mu_T), \tag{45}$$

where $\mu_H = \mathcal{L}_{11}^B/\mathcal{L}_{11}$ is the Hall mobility, and $\mu_T = \mathcal{L}_{12}^B/\mathcal{L}_{12}$ its thermal analog, introduced by Sun *et al.* [9]. From this point of view, a finite Nernst signal can only appear when there is a mismatch in the mobilities. The two mobilities $\mu_H$ and $\mu_T$ are shown in Fig. 6 for varying (a) band gap, (b) asymmetry, and (c) scattering rate. At large $T$, $-\mu_T \gg -\mu_H$, resulting in a sizable Nernst coefficient. At intermediate temperatures, both mobilities exhibit a temperature- and gap-independent plateau of *equal* magnitude: the Nernst signal vanishes. At low $T$ there is again a mobility mismatch $-\mu_T > -\mu_H$, which is responsible for the pronounced

low-temperature peak. Qualitatively, this behavior mirrors the analysis of the experimental mobilities of FeSb$_2$ from Ref. [9] that we reproduce in the inset of Fig. 11(h). Also in experiment, the mobility mismatch is notable only at low and high temperatures.

*b. Mott formula for the Nernst coefficient.* Analogous to the Mott formula of the Seebeck coefficient (44), a Sommerfeld expansion can be used to obtain an approximate formula for the Nernst coefficient. Using $\mathcal{L}_{12}^B \approx \frac{\pi^2 k_B^2}{3e} T^2 \frac{\partial \mathcal{L}_{11}^B(\mu)}{\partial \mu}$, valid at low $T$ for Boltzmann-derived Onsager coefficients of metals [86], one finds [87]

$$\nu \approx -\frac{\pi^2}{3e} k_B^2 T \frac{\partial \mu_H}{\partial \mu}, \tag{46}$$

where $\mu_H = \mathcal{L}_{11}^B/\mathcal{L}_{11}$ is again the Hall mobility.[9,10] As for the Seebeck coefficient, we find the link Eq. (46) between transport of charge ($\mu_H$) and entropy ($\nu$) to hold in the low-$T$ saturation regime (in this case $T < T_{R_H}^*$). In other words, the lowest-order terms in the low-temperature expansions (32) and (33) fulfill Eq. (46). In this sense, the saturation of both $\sigma_{xx}$ and $\sigma_{xy}^B$ dictates the Nernst coefficient to vanish linearly for $T \rightarrow 0$. This behavior, otherwise typical for metals [76], is indeed experimentally observed in correlated narrow-gap semiconductors [see Fig. 11(h) for the example of FeAs$_2$]. In metals, however, the variation of the charge and Hall conductivities with the chemical potential is usually small. Then, Eq. (46) means that also the Nernst coefficient will be very small, a statement referred to as Sondheimer cancellation [76,86,88]. Instead, as we have seen here, a changing chemical potential can notably manipulate the residual conductivities of an incoherent semiconductor.

---

[8]Our survey of the Nernst coefficient can be compared to the semiclassical relaxation-time approximation pioneered recently in Ref. [67] that includes effects to leading order in $1/\Gamma$.

---

[9]This "Mott formula" for the Nernst coefficient is often written using the Hall angle $\tan(\theta_H)/B = \mu_H$ with the magnetic field $B$.

[10]Alternatively, the temperature derivative of the Hall mobility can be seen as a source for a finite Nernst signal $\propto T \partial \mu_H/\partial T$ [110].
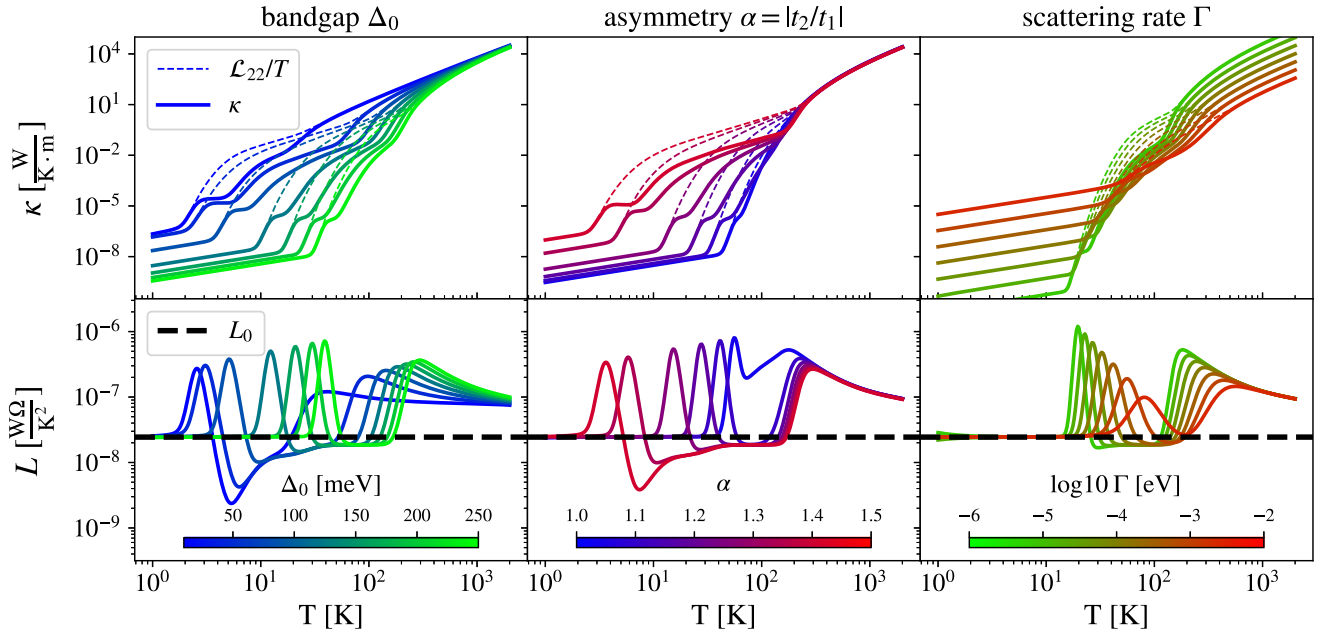
FIG. 7. Thermal conductivity and Wiedemann-Franz law. The same parameter sets as in Fig. 4 are used. In addition to the total thermal conductivity $\kappa$ (top row) we plot also the contribution $\mathcal{L}_{22}/T$ (dashed) individually. The Lorenz ratio (bottom row) converges to $L(T \to 0) = L_0 = \frac{\pi^2 k_B^2}{3e^2}$ (see text for details).

### 6. Thermal conductivity and Lorenz ratio

Next, we discuss the electronic contribution to the thermal conductivity $\kappa$. In the low-temperature $\rho$-saturation regime we find the linear behavior from Eq. (36). Increasing the temperature results in various kinks and shoulders. Again, we can separate the influence of a changing chemical potential from the inner structure of the transport kernel functions: While the pure $\mathcal{L}_{22}$ contribution [first term in Eq. (9); dashed lines in Fig. 7] only experiences the transition stemming from the chemical potential, the shoulders in the intermediate regime derive from the $\mathcal{L}_{12}$ contribution [second term in Eq. (9)]. Combined to the Lorenz ratio $L$ [see Eq. (13)], we find a complex temperature dependence: In the zero-temperature limit $L(T)$ clearly converges to the Sommerfeld value of the Lorenz number $L_0 = \frac{\pi^2 k_B^2}{3e^2}$. This can also be confirmed via the low-$T$ expansion

$$L = L_0 + \left(\frac{k_B}{e}\right)^2 \frac{16\pi^4}{45} \frac{5a^2 - 2\Gamma^2}{(a^2 + \Gamma^2)^2} k_B^2 T^2 + O(T^4) \quad (47)$$

(see Appendix B for more details). This result is expected since in our theory both the electrical and heat currents are transported by the same carriers, subject to the same elastic scattering mechanism. If inelastic scattering effects, e.g., via an electron-phonon coupling, were to be included, this unison will be jeopardized [89]. Then, the Wiedemann-Franz law can be strongly violated at low (but finite) temperatures, with $L$ being notably suppressed [90].

In the opposite limit $T \to \infty$, we find $L(T \to \infty) = 0$. Note that this result assumes a temperature-independent scattering rate. For specific conditions on $\Gamma(T)$, the Lorenz ratio converges to $L_0$ at high temperatures, as will be discussed in Sec. IV C.

### 7. Power factor and figure of merit

Finally, we consider the thermoelectric power factor $S^2\sigma$ and the figure of merit $zT$, given by Eqs. (14) and (15) and shown in Fig. 8. In order to achieve somewhat realistic $zT$ values we add a (dominant) phonon contribution to the thermal conductivity $\kappa = \kappa_e + \kappa_{ph}$ using an optimistic $\kappa_{ph} = 10\frac{W}{K^2 m}$. First, we note that $S^2\sigma$ is seemingly large at the upper end of the shown temperature window. However, this behavior originates from the temperature independence of the scattering rate, that we assume here for illustrative purposes. Indeed, $\Gamma = \text{const}$ causes (for large $T$) a very small resistivity (see top row of Fig. 4) that overcompensates the drop in the Seebeck coefficient. In practice, the scattering rate itself is typically temperature dependent. As explained below (see Sec. IV C 2), a reasonable $\Gamma(T)$ causes both the power factor and $zT$ to vanish quite rapidly at large temperatures [see Fig. 9 (lowest two panels)]. We therefore focus on the lower-temperature structure in $S^2\sigma$ and $zT$ in Fig. 8, that is equally present when a realistic $\Gamma(T)$ is used. The displayed peak in the power factor and $zT$ is the result of the usual compromise [47,48] between large $S$ and small $\rho$.[11] We find the optimal power factor to occur in the vicinity of $T_\nu^\mu$, the onset of the high-temperature crossover from the first into the second activated-$\rho$ regime. Peak temperatures move up (down) for a growing band gap (scattering rate), while peak amplitudes benefit from larger gaps, larger asymmetry, but a smaller scattering rate.

As the ratio of power factor and thermal conductivity, $zT$ inherits its structure from the former, while the latter modulates the overall magnitude. Indeed, band-gap variations keep

---

[11]Interestingly, this conventional tradeoff was recently found to be broken in an ultrathin oxide film near its Mott transition [111].
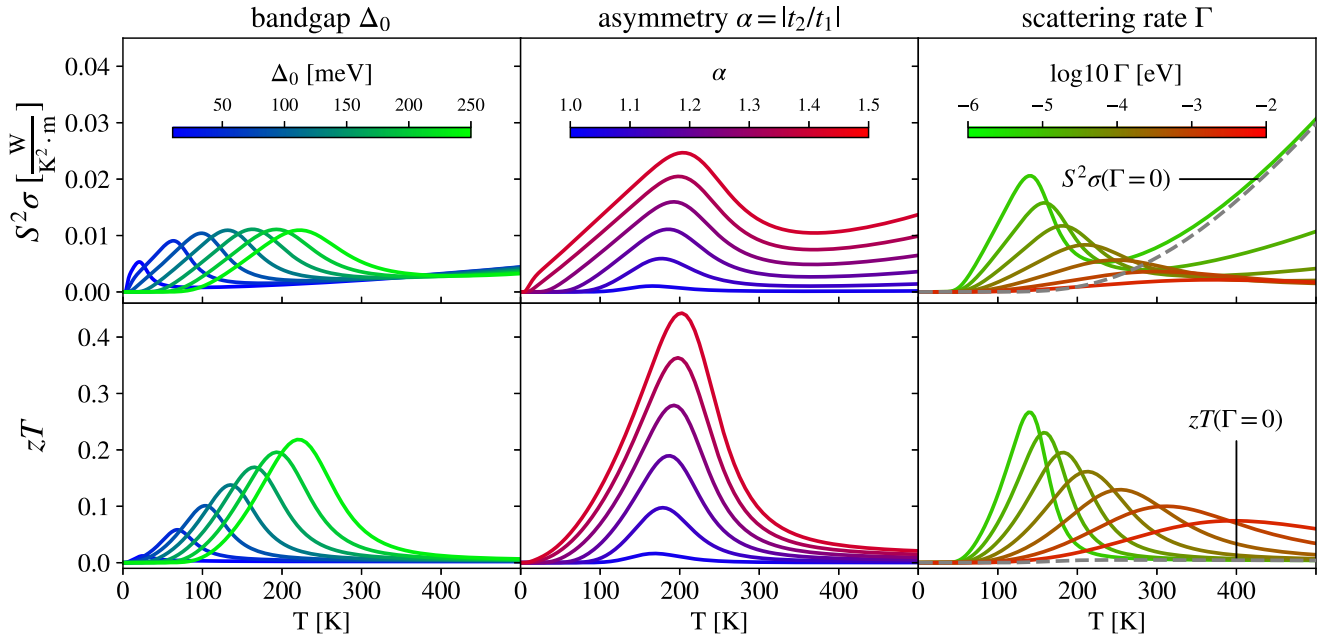
FIG. 8. Power factor and figure of merit. The power factor $S^2\sigma$ (top row) is peaked at the transition of the first into the second activated regime of $\rho = 1/\sigma$. For $T \to 0$, $S^2\sigma$ is suppressed due to the low-temperature limit of $S(T) \sim T$. At intermediate temperatures, $S^2\sigma$ develops an important peak that is undetectable in Boltzmann theory (gray, dashed line). The high-temperature increase in the power factor is nonphysical and disappears when a $T$-dependent scattering is included (see Sec. IV C 2). The figure of merit $zT$ (bottom row) shows similar features but, due to the thermal-conductivity weighing, peak sizes are affected differently. Note: In $zT$ we included a constant phonon contribution of $\kappa_{ph} = 10 \frac{W}{K^2\,m}$. Again, the low-temperature peak $zT$, indicative of potential merit in thermoelectric devices, is completely missing in the Boltzmann regime (gray, dashed line).

the peak amplitude of $S^2\sigma$ essentially unchanged but move the peak position. The maximal $zT$ then increases for larger band gaps, as the (here: electrical) thermal conductivity dwindles. Variations of the band asymmetry, instead, only change the size of the $zT$ maximum, while increasing scattering rates lower the peak amplitude and shift it to higher temperatures.

Importantly, the just described peak in both $S^2\sigma$ and $zT$ is *absent* when the Boltzmann approach is applied to the band structure (dashed gray lines in the right panels of Fig. 8). Indeed, we find that maximal thermoelectric performance is realized in the temperature range delimited by $T_\Delta^\mu$ and $T_\nu^\mu$. These characteristic scales are driven (see above) by changes in the chemical potential $\mu$, caused by the finite lifetimes of conduction and valence states, an effect beyond mere thermal activation. As a consequence, assessing the potential of narrow-gap semiconductors for thermoelectric applications on the basis of Boltzmann theory applied to coherent electronic band structures is virtually meaningless. A (high-throughput) screening of materials [91–93] that neglects finite electronic lifetimes of intrinsic carriers may miss potentially favorable compounds.

### C. Temperature-dependent scattering rate

In the previous section, we held the scattering rate $\Gamma$ constant to unravel the prototypical variations of transport observables with respect to gap, particle-hole asymmetry, and the scattering rate itself. Setting $\Gamma(T) = \Gamma_0$ led to some effects not observed in experiments. In particular, the resistivity became vanishingly small in the intermediate- to

high-temperature regime ($\frac{\rho(T<T_\rho^*)}{\rho(T>T_\nu^\mu)} \sim 10^5$). Indeed, only when temperature reaches a value greater than the system's band-width, the resistivity starts again to increase (not shown). Experiments probing narrow-gap semiconductors, however, witness an insulator-to-metal crossover above a temperature that is still small with respect to the charge gap [1]. In FeSi ($\Delta \sim 50$ meV $= k_B \times 580$ K), for example, the slope of the resistivity turns positive above 300 K [94], while in opti-cal spectroscopy for FeSb$_2$ ($\Delta \sim 30$ meV $= k_B \times 350$ K) a Drude-type peak starts developing at around 100 K [95,96]. Clearly, this metallization is beyond mere thermal activation of carriers across the charge gap. Theoretically, this phe-nomenon has been attributed to incoherent spectral weight spilling into the gap and was advocated to derive from elec-tronic correlation effects [28] or thermal disorder [97,98]. In the correlations' picture, the Hund's rule coupling drives a scattering rate that grows quadratically with tempera-ture [28,31]. Therefore, we will restrict ourselves in the following to scattering rates with a polynomial temperature dependence.

#### 1. General considerations

Without a growing scattering rate the Lorenz ratio $L$ approaches zero in the high-temperature limit and an unrea-sonably large power factor $S^2\sigma$ appears in the intermediate-temperature regimes (see the high-$T$ upturn in Fig. 8). If, instead, we consider a residual scattering rate plus a term with a polynomial temperature dependence $\Gamma(T) = \Gamma_0 + \gamma T^\eta$, where $\gamma > 0$, $\eta > 0$, the argument $z$ of the polygamma
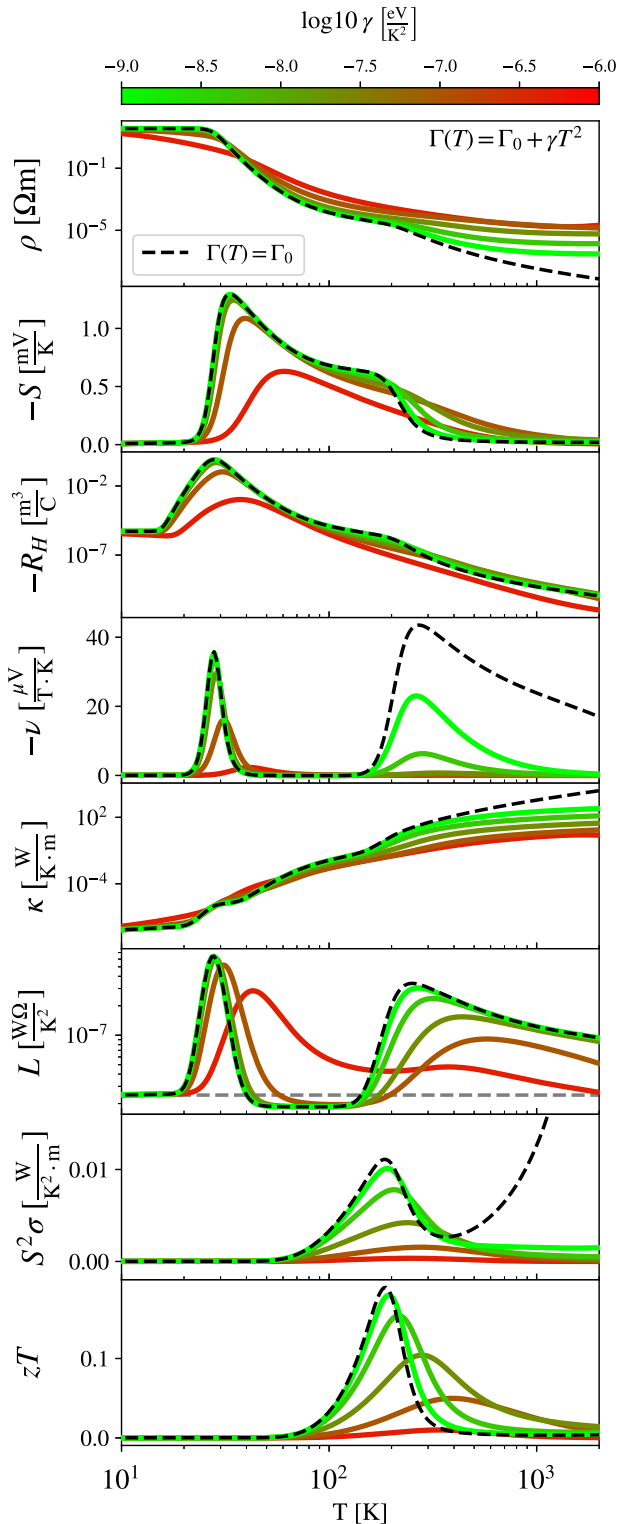
FIG. 9. Temperature-dependent scattering rate. Effects of $\Gamma(T) = \Gamma_0 + \gamma T^2$, with residual scattering $\Gamma_0 = 10^{-4}$ eV and varying $\gamma$, for fixed band gap $\Delta_0 = 200$ meV and asymmetry $\alpha = 1.2$. The dashed black lines are results for $\gamma = 0$. The additional temperature dependence causes a resistivity upturn at high temperatures which also results in a smoothed Seebeck and Hall coefficient. This metallic trend directly removes the nonphysical upturn in the power factor while simultaneously causing the Lorenz factor $L$ to converge towards $L_0$ (horizontal, dashed, gray line).

functions $\psi_i(z)$ becomes

$$z(T) = \frac{1}{2} + \left[ \frac{\Gamma_0 + ia}{2\pi k_B T} + \frac{\gamma T^\eta}{2\pi k_B T} \right]. \quad (48)$$

Scattering rates that increase slower than linearly ($\eta < 1$) lead to arguments that converge to $z(T \to \infty) = \frac{1}{2}$; exact linear behavior leads to $z(T \to \infty) = \frac{1}{2} + \frac{\gamma}{2\pi k_B}$ while $\eta > 1$ leads to a diverging $z(T)$.

In the first two cases, the Lorenz ratio simplifies in leading order to $L(T) \sim O(T^{2\eta-2})$. $\eta < 1$ therefore implies a vanishing Lorenz ratio while $\eta = 1$ implies some saturation value $L(\gamma)$, which, numerically, is generally orders of magnitude smaller than $L_0$. If the scattering rate increases faster ($\eta > 1$), the same Taylor series of the polygamma functions that was applied in the zero-temperature limit (see Appendix B) can be employed. Consequently, the high-temperature limit is identical to the low-temperature limit and thus $L(T \to \infty) = L_0$.

### 2. Fermi-liquid-like scattering

Dynamical mean-field theory calculations suggest that the scattering rate evolves quadratically with temperature for both Kondo insulators [3] and $d$-electron-based narrow-gap semiconductors [28,31]. From here on, we therefore assume a Fermi-liquid-like

$$\Gamma(T) = \Gamma_0 + \gamma T^2. \quad (49)$$

Figure 9 illustrates how the additional scattering term changes the transport for a range of $\gamma$ values (shades green to red) compared to $\gamma = 0$ (dashed black line), for a fixed band gap $\Delta_0 = 200$ meV, residual scattering $\Gamma_0 = 10^{-4}$ eV, and asymmetry $\alpha = 1.2$.

Overall, the increased scattering smoothes all considered quantities. A minimal $\gamma$ is sufficient to suppress the high-temperature shoulder in $\rho$ at $T_\nu^\mu$ and causes a slight upturn at high temperatures. The saturation regime is instead stable up to rather large $\gamma$. Naturally, the added scattering term only notably influences transport above temperatures for which $\gamma T^2 \sim \Gamma_0$. In this vein the high-temperature shoulder of the Seebeck coefficient is smoothed out and parts of the signal are pushed up in temperature. Quite generally, the increased scattering leads to less pronounced peaks which are shifted to higher temperatures. Since the shoulder in $S$ at $T_\Delta^\mu$ was responsible for the strong signal in the power factor and the figure of merit (see above), significant qualitative changes are expected for $\gamma > 0$: Aside from the suppression of the nonphysical increase of the power factor at high temperatures, both $S^2\sigma$ and $zT$ are equally attenuated for $\gamma > 0$ and their peaks shift up in temperature, as expected. Again, we included in $zT$ a phonon contribution $\kappa_{\rm ph} = 10 \frac{\rm W}{\rm K^2\,m}$ to the thermal conductivity. As already seen in the $\Gamma_0$ scan in Fig. 4, a noticeable suppression is observed for the high-temperature peak of the Nernst coefficient. The second, low-$T$ peak in the Nernst coefficient (that is absent in Boltzmann approaches) is instead relatively stable with $\gamma$ as it occurs at low enough temperatures $T \sim T_\rho^*$. As discussed in Sec. IV C 1 the Fermi-liquid-like scattering rate enforces that the high-temperature limit of the Lorenz ratio converges to $L_0$. This is evident in Fig. 9: for the largest scattering rate, $L$ reaches $L_0$ within the shown temperature window.

## V. MODELING MATERIALS

Having established an understanding of how relevant parameters drive changes in transport observables, we now turn to material-specific simulations. While still focusing on a minimal description, we attempt to reproduce the temperature profiles of transport properties in selected narrow-gap semiconductors, as well as trends among them.

### A. FeSb$_2$: Characteristic temperature scales without impurity states

First, we discuss the result for FeSb$_2$ shown in the front Fig. 1. FeSb$_2$ is a correlated semiconductor [7–9,95] with a narrow gap of $\Delta \gtrsim 30$ meV, as extracted from activation-law fits of the resistivity [7,9] or the magnetic susceptibility [99,100]. Consistent with $GW$ and $GW$+DMFT simulations [10,11], we model FeSb$_2$ with a noninteracting gap $\Delta_0 = 60$ meV and an effective mass enhancement $Z^{-1} = 2$. We find that a small asymmetry $\alpha = 1.06$ mimics the material well. Finally, we assume a scattering rate of the form (49), where the parameters of the residual scattering $\Gamma_0$ and the prefactor $\gamma$ of the quadratic term are adjusted by hand. We find that best agreement with experiment is reached for $\Gamma(T) = 5 \times 10^{-5}$ eV $+ 10^{-7} \frac{\text{eV}}{\text{K}^2} T^2$ (see Fig. 1). This scattering rate is quite realistic: dynamical mean-field calculations for the related correlated narrow-gap semiconductor FeSi [28,31] yield comparable values. Having optimized the electronic structure parameters so that the simulated resistivity qualitatively follows the experiment, the temperature profiles of all other transport observables for FeSb$_2$ automatically fall into place (see Fig. 1). The approach therewith in particular verifies the experimentally observed correlation between features in different response functions.

At low temperatures, the onset of resistivity saturation at $T_\rho^*$ is accompanied by a peak in both the Hall and the Nernst coefficients at the identical temperature and a peak in the Seebeck coefficient at a slightly higher temperature, all of which is congruent with experiment. This low-$T$ behavior of transport properties is encoded in the linear response kernel functions. Agreement with experiment confirms that our approximations for the kernels, linearized self-energy, omission of vertex corrections, conserves the essential physics. Instead, in previous modelings of FeSb$_2$, based on semiclassical approaches [14–16], resistivities and the Hall coefficient either diverged at low $T$ or had to be suppressed by impurity states, e.g., by forcing the chemical potential into the conduction band. An alternative scenario for residual conduction in FeSb$_2$ could be provided by the recent observation of metallic surface states [10,101]. Whether these weakly dispersive states can account for the typical low-$T$ characteristics across all transport observables remains to be seen. We also note that for the topological insulator SmB$_6$ conduction by surface states and residual bulk conduction from finite lifetimes coexist [3].

Moving to higher temperatures, the resistivity exhibits two distinct activation regimes. We find the shoulder in-between $T_\Delta^\mu < T < T_\nu^\mu$ to be driven by changes in the chemical potential. Therefore, if the chemical potential only accounts for the thermal broadening of excitations, as commonly done in the context of Boltzmann approaches for band theory methods, all structure at intermediate temperatures is lost (gray, dotted lines in Fig. 1). If, instead, the chemical potential, e.g., via Eq. (24), accounts for the scattering rate, Boltzmann simulations do capture the intermediate-$T$ features (see black, dashed lines in Fig. 1). Alternatively, the temperature dependence of the chemical potential can be engineered by assuming in-gap impurity states [14–16]. Given that transport observables exhibit three to four distinct regimes, phenomenological modelings actually used up to three impurity levels to properly guide the chemical potential [14]. In our description, no impurity states are required: According to the presented phenomenology for transport in narrow-gap semiconductors, the intriguing temperature dependence in FeSb$_2$ exclusively originates from its *intrinsic* electronic structure. Instead of being set by *explicit* energy levels inside the gap, characteristic temperatures emerge through the interplay of relevant energy scales: the gap, the hopping, temperature, and, crucially, the scattering rate.

In the following section, we investigate the influence of potential impurity states in more detail for FeSb$_2$. Thereafter, we will see that *explicit* impurity states are not fully out of the picture for other materials, but likely account for at least some aspects of conduction at intermediate temperatures in semiconductors with gaps $\Delta > 50$ meV.

Before, however, a few comments regarding thermoelectricity in FeSb$_2$ are in order. While our approach neatly captures the *temperature profile* of transport observables, we do not reproduce the large *amplitude* of the Seebeck and Nernst coefficients. In fact, this is expected, as FeSb$_2$ violates the upper bound $|S| \leqslant \Delta/T$ for a thermopower exclusively driven by electron diffusion [11]. As alluded to in the Introduction, this riddle was successfully solved [11–16] by attributing the colossal amplitude to the phonon-drag effect. Simply speaking, the thermal gradient also leads to a nonequilibrium phonon distribution. Working to equilibrate the thermal gradient, phonons then scatter with electrons dominantly towards the cold end of the sample, which is also the direction of the net electronic diffusion. Thereby momentum is constructively injected into the electronic subsystem, significantly boosting thermoelectric effects. This well-known phenomenon [102] continues to receive renewed interest, in the context of electronic correlations [1,14,16] (the effect is large when coupling to heavy electrons), the phonon-engineering pathway to efficient thermoelectrics [103], or both [104]. Crucially for our argument here, this phonon enhancement of the electric response is expected to be smooth in temperature, so as to not produce additional features in transport observables. Indeed, while experimental peak amplitudes cover almost an order of magnitude $S^{\text{max}} \sim 5$–$45 \frac{\text{mV}}{\text{K}}$ across different samples [7,9,17,21], the corresponding characteristic temperature profiles are almost identical. Crucially, Pokharel *et al.* [12] and Takahashi *et al.* [13] demonstrated that the phonon drag in FeSb$_2$ can be consistently suppressed by geometric constraints. With their severely limited phonon mean-free path, polycrystalline samples are then expected to yield Seebeck amplitudes compatible with the purely electronic diffusion simulated here. Indeed, experimental peak amplitudes for polycrystalline samples $S^{\text{max}} \sim O(0.1$–$1) \frac{\text{mV}}{\text{K}}$ [12,13,105,106] are comparable to our

modeling (see Fig. 1). With the phonon drag thus mainly scaling the amplitude of the thermoelectric response, previous modelings including this effect had to explicitly introduce in-gap impurity levels [14–16] to generate the experimentally evidenced characteristic temperature scales. Here, we showed that the electron diffusion part of the Seebeck and Nernst coefficients has the correct temperature profile without the need for *ad hoc* in-gap levels, provided that finite lifetimes of intrinsic carriers are accounted for consistently.

### B. FeSb$_2$: Explicit impurity states?

To strengthen the argument that in-gap impurity states are not crucial for an understanding of transport properties of FeSb$_2$, we study the explicit inclusion of such states. In Fig. 10 (bottom panel) we compare several ways to obtain the needed chemical potential: (i) $\mu_\psi$ (solid blue) indicates an occupation determined via Eq. (24), that accounts for both thermal broadening and the finite lifetimes of valence and conduction states (data reproduced from Fig. 1); (ii) $\mu_{FD}$ (black) that only includes thermal broadening via the Fermi-Dirac distribution; (iii) $\mu_{FD}$ + imp (pink dashed) in which an in-gap impurity level has been designed to mimic $\mu_\psi$; and (iv) $\mu_\psi$ + imp (green dashed) in which the same impurity level is added in the presence of finite lifetimes of intrinsic states.

Clearly, the Fermi-Dirac description of the chemical potential (black), in which the chemical potential converges towards the midgap point (origin of energy) for $T \to 0$, yields very different transport functions (upper panels). Their temperature profiles do not agree with experiment [cf. Fig. 1 (insets)]. Still, the resistivity and the Hall coefficient saturate (albeit at values different than in the "$\mu_\psi$" scenario) and the Nernst and Seebeck coefficients vanish for $T \to 0$ since these low-temperature properties stem from the employed LINRETRACE kernels [Eqs. (25)–(30)]. As a consequence, if the chemical potential $\mu_\psi$, that drives both the features at intermediate temperatures and influences the saturation values, could be mimicked by other means, transport properties will be very similar to the intrinsic "$\mu_\psi$" picture. And, indeed, imitating the temperature dependence of $\mu_\psi$ through the inclusion of a single explicit donor level (at an energy $E_D = 26$ meV, degeneracy $g = 1$, and density $\rho_D = 5 \times 10^{-6}$ per unit cell), these "$\mu_{FD}$ + imp" results (pink dashed) are very close to the $\mu_\psi$ results. In turn, if we include the same impurity level on top of the lifetime effects that drive $\mu_\psi$ [a combination labeled "$\mu_\psi$ + imp" (green dashed) in Fig. 10], nothing much happens. In other words, if finite lifetimes of intrinsic valence and conduction states are properly accounted for, extrinsic in-gap states have little influence on transport properties in FeSb$_2$. This finding strengthens our alternative scenario in which the driver of the characteristic temperature profile in transport properties is the scattering rate.

### C. Related materials: FeAs$_2$, FeGa$_3$, RuSb$_2$

We now extend our transport study to other materials. In the right column of Fig. 11 we reproduce experimental data of various intermetallic semiconductors. In order of increasing
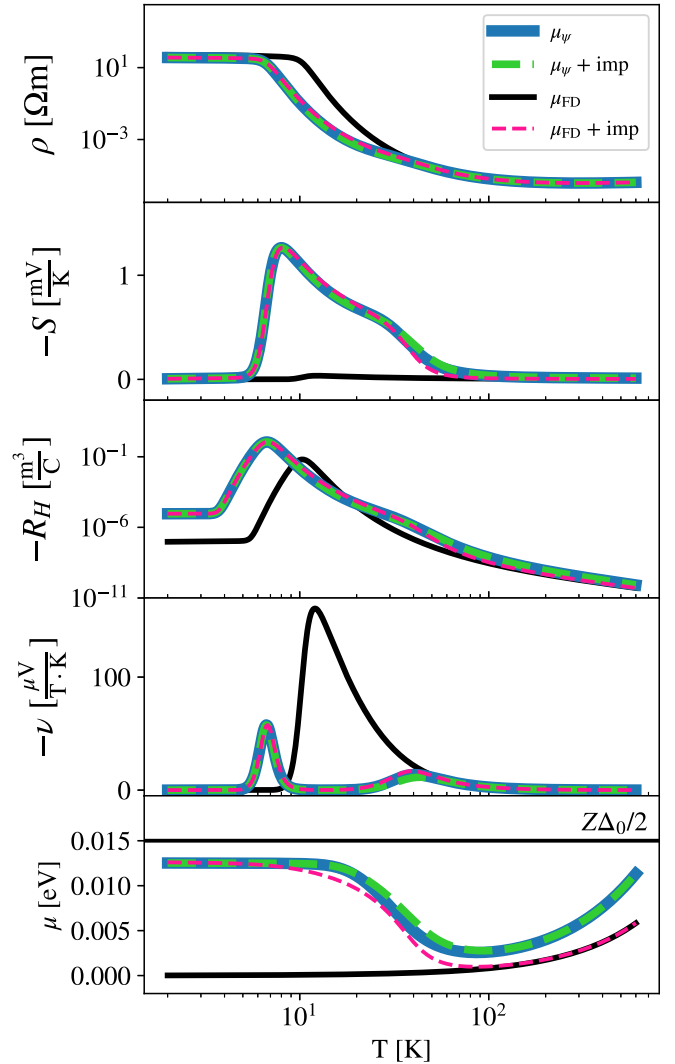


FIG. 10. FeSb$_2$ and impurity in-gap states. We imitate the chemical potential determined via Eq. (24) ($\mu_\psi$) with the Fermi-Dirac distribution through the presence of an explicit impurity level ($\mu_{FD}$ + imp) in the vicinity of the conduction band. We employ the same parameters as in Fig. 1. Additionally, an impurity donor level with density $\rho_D = 5 \times 10^{-6} \frac{1}{\text{unit cell}}$, degeneracy $g = 1$ is positioned at $E_D = 26$ meV below the conduction band (before renormalization via $Z$). The chemical potential is then adjusted according to $N \equiv n_{occ}(\mu) - \frac{\rho_D}{1 + gf(\mu - E_D)}$. For comparison we also show the (very different) chemical potential determined by the Fermi-Dirac distribution without the donor level ($\mu_{FD}$) and the digamma-computed chemical potential with the additional impurity ($\mu_\psi$ + imp) which is indistinguishable from $\mu_\psi$. All shown transport quantities employ the full LINRETRACE kernels (25)–(30).

gap, FeSb$_2$ [9], FeAs$_2$ [9], RuSb$_2$ [17], and FeGa$_3$ [18,38]. For all considered compounds, the charge gap can be extracted directly from the high-temperature behavior of the resistivity (or optical data). With the exception of FeSb$_2$ ($Z^{-1} = 2$), we do not apply a quasiparticle renormalization ($Z = 1$). Indeed, larger hybridization-gap semiconductors are expected to exhibit less correlation signatures [1] and also the substitution of a 3$d$ transition metal with its 4$d$ homolog
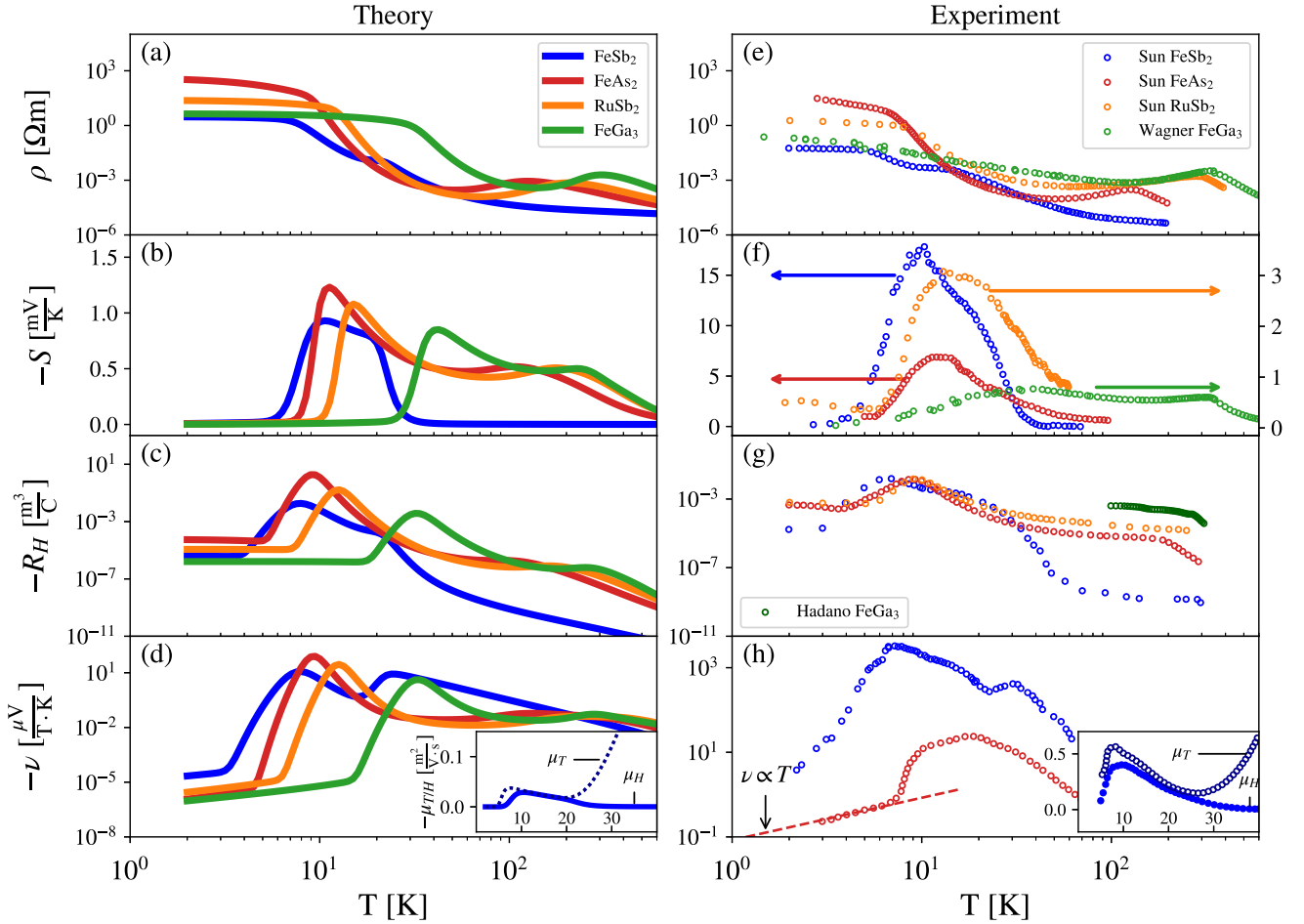
FIG. 11. Modeling of intermetallic narrow-gap semiconductors. The experimental temperature profile of transport observables of FeSb$_2$ [9], FeAs$_2$ [9], RuSb$_2$ [17], and FeGa$_3$ ([100] orientation) [18,38] is simulated with finite lifetimes and a single donor level at a distance $E_D$ below the conduction band. Lifetimes dominantly determine transport at intermediate to low temperatures. Instead, the donor level virtually only affects higher temperatures near $T_\Delta^\mu$ and $T_\nu^\mu$. In particular, in congruence with experiment, the level accounts for the metallic slope in the resistivity seen in all materials other than FeSb$_2$. To limit the number of adjustable parameters, band gaps were fixed to experimental values: using $Z = 0.5$ for $\Delta_{\mathrm{FeSb_2}} = Z \times \Delta_{0,\mathrm{FeSb_2}} = 30$ meV and $Z = 1$ in $\Delta_{\mathrm{FeAs_2}} = 200$ meV; $\Delta_{\mathrm{RuSb_2}} = 290$ meV; $\Delta_{\mathrm{FeGa_3}} = 500$ meV. All modeling parameters are listed in Table II. Jointly, finite lifetimes and the impurity level yield an accurate description without the need for other electronic structure details.

will reduce correlation effects [31], as explicitly shown for Fe$_{1-x}$Ru$_x$Sb$_2$ [24,107].

While the resistivity of FeSb$_2$ only displays a shoulder at $T \sim T_\Delta^\mu$ [successfully modeled with $\Gamma(T)$], a distinct peak can be observed in the three other materials. The metallic slope $\partial\rho/\partial T > 0$ at temperatures below said peak *cannot* be replicated with a chemical potential that is driven by finite lifetimes through Eq. (24) alone. Indeed, the transition of the chemical potential must occur more abruptly in temperature, making explicit impurity states a necessity to achieve agreement with experiment. We find that deploying a single donor level near the conduction band allows us to reproduce the qualitative behavior of all considered materials across all transport observables.

*a. Resistivity.* For FeSb$_2$ this leads to minor improvements in $T_\Delta^\mu$ and $T_\nu^\mu$ of the initial fit (Fig. 1). For Fe/RuAs$_2$ and FeGa$_3$ the engineered chemical potential causes the resistivity to decrease when cooling below their $T_\nu^\mu$ (metallic slope), before it rises again to enter the second activated region and,

eventually, the saturation regime below $T_\rho^*$. The overall agreement is astounding: with a single impurity level characteristic temperatures, qualitative features, and even relative amplitudes between various materials can be accurately modeled.[12]

*b. Seebeck coefficient.* Unsurprising from the previous analysis and the optimal parameters listed in Table II the magnitude of the theoretical Seebeck coefficients [Fig. 11(b)] does not differ significantly in the modeling. For all materials considered we observe a peak amplitude in the range of $|S^{\max}| = 0.8$–$1.3 \frac{\mathrm{mV}}{\mathrm{K}}$ positioned at a respective $T_S^{\max}$ slightly above the saturation temperatures $T_\rho^*$ of Fig. 11(a). While lacking the phonon-drag boost to the Seebeck magnitude, our treatment still captures quite well both the dominant peak's

---

[12]The only noticeable deviation constitutes FeGa$_3$ where the chosen experiment exhibits a less clear transition between an activated and a saturated region. See, however, the *c*-axis resistivity in Ref. [38].

TABLE II. Electronic structure parameters for simulated materials. Quasiparticle renormalization $Z$, band gap $\Delta_0$, band asymmetry $\alpha = |t_2/t_1|$, scattering rate coefficients in $\Gamma(T) = \Gamma_0 + \gamma T^2$, and a single donor level of concentration $\rho_D$ positioned at a distance $E_D$ below the conduction band. Gap sizes $\Delta = Z\Delta_0$ are kept fixed to experimental values during the (manual) optimization procedure.

| | $Z$ | $\Delta_0$ (meV) | $\alpha$ | $\Gamma_0$ (eV) | $\gamma$ ($\frac{\mathrm{eV}}{\mathrm{K}^2}$) | $E_D$ (meV) | $\rho_D$ ($\frac{1}{\text{unit cell}}$) |
|---|---|---|---|---|---|---|---|
| FeSb$_2$ | 0.5 | 60 | 1.02 | $1.5 \times 10^{-4}$ | $8 \times 10^{-7}$ | 20 | $6 \times 10^{-8}$ |
| FeAs$_2$ | 1 | 200 | 1.12 | $1.5 \times 10^{-5}$ | $3 \times 10^{-7}$ | 15 | $1.5 \times 10^{-4}$ |
| RuSb$_2$ | 1 | 290 | 1.12 | $7 \times 10^{-5}$ | $2 \times 10^{-7}$ | 18 | $3.5 \times 10^{-4}$ |
| FeGa$_3$ | 1 | 500 | 1.04 | $3 \times 10^{-4}$ | $2 \times 10^{-7}$ | 40 | $1.5 \times 10^{-3}$ |

position at $T_S^{\max}$ and the high-temperature shoulder (peak) at $T_\Delta^\mu$ for FeSb$_2$ (FeGa$_3$). Unfortunately, no data for FeAs$_2$ and RuSb$_2$ are available for higher temperatures: For them, we anticipate an additional Seebeck feature where the respective resistivities are peaked.

*c. Hall coefficient.* The agreement to experiment for the Hall coefficient, Fig. 11(g), is comparable to that of the resistivity: as in the experiments, the position of the peak in $R_H$ virtually coincides with the $T_\rho^*$ crossover in the resistivity. Equally, the hierarchy across materials is captured for the saturation value $R_H(T \to 0)$, and also the decay at higher temperatures mirrors the experiment. Clearly, the high-temperature shoulder is connected to the resistivity peak at $T_\Delta^\mu$. We therefore expect $R_H$ of RuSb$_2$ to similarly drop if temperatures slightly beyond the shown experimental range were probed.

*d. Nernst coefficient.* For the Nernst coefficient less experimental data are available [see Fig. 11(h)]. For the cases of FeSb$_2$ and FeAs$_2$ the qualitative agreement between simulation and measurements is reasonably good. Of course, what has been said about the Seebeck coefficient of FeSb$_2$ also applies to its Nernst signal: it is substantially boosted by the phonon drag [14,16], limiting us to discussing the qualitative temperature profile.[13] Not suffering from this intricacy, clearer agreement is seen for FeAs$_2$: there, the experimental low-temperature signal neatly follows the linear behavior $\nu \propto T$ (dashed line) as derived in Eq. (37) and illustrated in Fig. 11(d).

For FeSb$_2$, we also indicate the Hall and thermal mobilities of Eq. (45) in the insets of Figs. 11(d) and 11(h). The experimental data [9] qualitatively match the theoretical prediction: at high temperatures a divergence between $\mu_T$ and $\mu_H$ is observed, giving rise to FeSb$_2$'s smaller Nernst peak at $T \approx 40$ K. Below, at intermediate temperatures $T \approx 20$ K, the two mobilities almost coincide (i.e., $\nu$ is suppressed). At low temperatures $T \approx 10$ K, again a slight mismatch occurs, giving rise to the prominent low-temperature peak.

---

[13]Contrary to $S$, $\nu$ depends on the lattice constant and scales according to $\nu \propto a_{\text{lattice}}^2$ (see Appendix D). Using instead of our generic $a = 1$ Å, a realistic lattice constant, FeSb$_2$'s Nernst amplitude is, in principle, not out of reach of the electron diffusion picture.

### D. Perspective

The previous section made clear that with reasonable scattering rates and (for larger gap systems) an explicit impurity in-gap level, all experimental transport coefficients can be qualitatively matched with an essentially featureless band structure. This emphasizes the notion that most, if not all, of the relevant transport physics in narrow-gap semiconductors are purely determined by the interplay of the gap, the chemical potential profile (shaped by temperature, carrier lifetimes and, potentially, impurity states), and the scattering rate. Electronic structure intricacies, such as details of the band structure beyond the gap value and optical transition elements, all seem to play only a secondary role. Further, we evidenced that, at low temperatures ($T_{R_H}^* < T_\rho^* < T_S^{\max}$), features are controlled by the scattering rate through the (quantum = beyond-Boltzmann) transport kernels, whereas the higher-temperature features ($T_\Delta^\mu < T_\nu^\mu$) are determined through the behavior of the chemical potential, which can be driven by the scattering rate of intrinsic carriers as well as by extrinsic impurity in-gap states.

Future extensions of the presented formalism (and software package [68]) could include the ability to describe phonon-drag contributions to thermoelectric observables. This advance could remedy our current inability to quantitatively match the amplitude of the Seebeck coefficient, in particular of single crystalline FeSb$_2$. Furthermore, an inclusion of (topological) surface conduction and in-depth comparisons of their importance vis-à-vis the discussed bulk conduction is desirable, also in view of FeSb$_2$ [10,101]. Finally, also anomalous bulk contributions could be included, following, e.g., the recent Ref. [108].

### VI. SUMMARY AND CONCLUSIONS

We conceptualized an efficient linear response transport formalism: a low-energy expansion of quasiparticle renormalizations enabled performing frequency integrations in dominant Feynman diagrams *analytically*. This algorithmic innovation, implemented in the LINRETRACE [68] package, allows accurate simulations down to temperatures where full Kubo calculations are cumbersome and Boltzmann techniques insufficient. We applied the methodology to narrow-gap semiconductors and established a comprehensive phenomenology of their transport properties.

First, we analytically extracted low-temperature characteristics of various transport observables: in congruence with experiments, both the resistivity and the Hall coefficient saturate at finite values for $T \to 0$. The Seebeck and Nernst coefficients, instead, vanish linearly in the zero-temperature limit, consistent with thermodynamic expectations. These behaviors are beyond the reach of semiclassical approaches like Boltzmann theory in the relaxation-time approximation, highlighting the importance of a fully quantum mechanical description.

Next, we simulated transport properties as a function of temperature for varying band gap, particle-hole asymmetry, and scattering rate. This survey establishes the prototypical temperature dependence of transport in narrow-gap semiconductors to be structured by five emergent characteristic

temperatures: $T^*_{R_H} < T^*_\rho < T^{\max}_S < T^\mu_\Delta < T^{\max}_\nu$: At high temperatures, $T > T^\mu_\nu$, the shape of all observables is controlled by the band gap $\Delta$ and our equations yield results identical to Boltzmann approaches. Upon cooling, $T^\mu_\Delta < T < T^\mu_\nu$, the system experiences a crossover from the activated, Boltzmann-type regime to a renormalized activation region $T^*_\rho < T < T^\mu_\Delta$, with an associated energy $\Delta_1 < \Delta$. We find this crossover to be driven by the chemical potential: finite lifetimes of valence and conduction states cause incoherent spectral weight to spill into the gap. Below a characteristic temperature, these incoherent carriers can no longer be neglected with respect to the charges that are thermally activated across the gap $\Delta_0$. In particle-hole asymmetrical systems, the chemical potential then has to adapt to preserve charge neutrality by moving to a position separated by only $\Delta_1$ from the top (or bottom) of the valence (or conduction) band. This *intrinsic* mechanism provides an explanation alternative to the common *extrinsic* scenario in which the moving of the chemical potential (and the associated shoulder in the resistivity) is attributed to the presence of impurity in-gap states. Finally, at low enough temperature $T < T^*_\rho$ ($T < T^*_{R_H}$), the system enters lifetime-dominated regimes in which the resistivity (the Hall coefficient) saturates and thermoelectric observables vanish linearly.

In semiclassical approaches, the lifetime of excitations appears as a mere prefactor of, say, the conductivity. In the quantum formulation derived here, instead, the scattering rate is a relevant control parameter that can compete with other energy scales of the problem. Indeed, the emergence of all characteristic transport features is a direct consequence of the interplay of the charge gap and the scattering rate. In other words, in our transport equations, thermal and lifetime broadening are described on an equal footing.

Crucial for potential applications, we find the LINRE-TRACE methodology to be essential to describe the temperature region where the thermoelectric power factor and the figure of merit are optimal. Materials discovery efforts based on conventional Boltzmann approaches, instead, are likely to overlook promising compounds. To mend this shortcoming, we note that our methodology allows for an easy integration into already established code bases, heralding future high-throughput material scans using, e.g., phenomenological scattering rates.

The established transport phenomenology further allowed us to fit experimental measurements and reverse engineer scattering rates (a task previously highlighted for Kondo insulators [3]), activated carrier densities, and (if needed) impurity in-gap states. For the examples FeSb$_2$, RuSb$_2$, FeAs$_2$, and FeGa$_3$, all characteristic features across transport observables were well captured, including trends between the different compounds. We therefore believe our phenomenology to be prototypical in the wider context of narrow-gap semiconductors [1].

(a) $\mathcal{I}_{11}(\mathbf{k}, n; \Omega; i\nu_m > 0)$

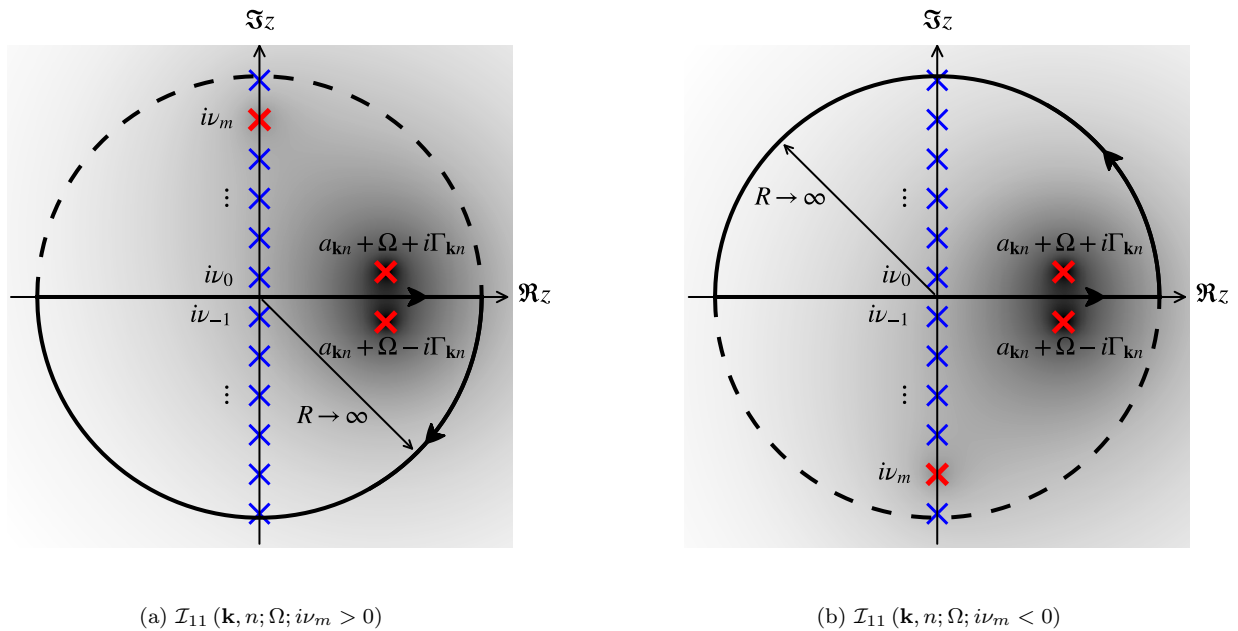(b) $\mathcal{I}_{11}(\mathbf{k}, n; \Omega; i\nu_m < 0)$

FIG. 12. Contour integration. (a) Extending the desired integral along $\omega = \mathrm{Re}z$ ($\mathrm{Im}z = 0$), the contour is closed via the *lower* half-plane for $\nu_m > 0$ or (b) closed via the *upper* half-plane for $\nu_m < 0$. The poles of the Fermi function are located on the imaginary axis $z = i\nu_m = i\frac{\pi}{\beta}(2m + 1)$ while the poles of the spectral function are located at $z = a_{\mathbf{k}n} + \Omega \pm i\Gamma_{\mathbf{k}n}$. Due to functional decay $f(z) = O(z^{a+b-7})$ (see text) in the limit of $R \to \infty$ the half-circles do not contribute. A straightforward residue evaluation (inside the closed contour) is thus sufficient to calculate the initial integral.

Calculations were performed on the Vienna Scientific Cluster (VSC).

## APPENDIX A: TRANSPORT KERNELS

In this Appendix we illustrate the contour integration and Matsubara summations necessary to evaluate the transport kernels. Again, we will restrict ourselves to the intraband case of Eqs. (18) and (19). For brevity we will abridge the notation and drop the full momentum and band dependence of the spectral function $A(\omega)$, quasiparticle weight $Z$, scattering rate $\Gamma$, and energy $a = \epsilon - \mu$.

### 1. Contour integration

Starting from the generalized transport kernels ($a \in \{1, 2\}$, $b \in \{1, 2\}$) from Eq. (18)

$$\mathcal{K}_{ab}(\mathbf{k}, n) = \int_{-\infty}^{\infty} d\omega\, \omega^{(a+b-2)} \left(-\frac{\partial f}{\partial \omega}\right) A^2(\omega) \qquad (A1)$$

we insert the Matsubara representation of the derivative of Fermi function

$$-\frac{\partial f}{\partial \omega} = \lim_{\Omega \to 0^+} \frac{f(\omega) - f(\omega + \Omega)}{\Omega}$$

$$= \lim_{\Omega \to 0^+} \frac{1}{\beta} \sum_m \frac{1}{\Omega} \left(\frac{1}{i\nu_m - \omega} - \frac{1}{i\nu_m - \omega - \Omega}\right) \quad (A2)$$

as well as the (coherent part of the) spectral function

$$A(\omega) = \frac{Z\Gamma}{\pi} \frac{1}{(\omega - a)^2 + \Gamma^2}, \qquad (A3)$$

where the fermionic Matsubara frequencies are $\nu_m = (2m + 1)\frac{\pi}{\beta}$ with $m \in \mathbb{Z}$. The resulting expression

$$\mathcal{K}_{ab}(\mathbf{k}, n) = \int_{-\infty}^{\infty} d\omega\, \frac{Z^2\Gamma^2}{\pi^2} \frac{\omega^{(a+b-2)}}{[(\omega - a)^2 + \Gamma^2]^2}$$

$$\times \frac{1}{\beta} \sum_m \lim_{\Omega \to 0^+} \frac{1}{\Omega} \left(\frac{1}{i\nu_m - \omega} - \frac{1}{i\nu_m - \omega - \Omega}\right) \qquad (A4)$$

can be abbreviated with

$$\mathcal{I}_{ab}(\mathbf{k}, n; \Omega; i\nu_m) = \int_{-\infty}^{\infty} d\omega\, \frac{(\omega - \Omega)^{(a+b-2)}}{[(\omega - a - \Omega)^2 + \Gamma^2]^2} \frac{1}{i\nu_m - \omega} \qquad (A5)$$

as

$$\mathcal{K}_{ab}(\mathbf{k}, n) = \frac{Z^2\Gamma^2}{\pi^2} \frac{1}{\beta} \lim_{\Omega \to 0^+} \left[\frac{1}{\Omega} \sum_m [\mathcal{I}_{ab}(\mathbf{k}, n; 0; i\nu_m)\right.$$

$$\left. - \mathcal{I}_{ab}(\mathbf{k}, n; \Omega; i\nu_m)]\right]. \qquad (A6)$$

For finite (positive) scattering rates $\Gamma > 0$, Eq. (A5) is an integral over a function with three distinct poles in the complex plane: $z = a + \Omega + i\Gamma$, $z = a + \Omega - i\Gamma$, and $z = i\nu_m$. This function decays with $z \to \infty$ as $O(z^{a+b-7})$ which for all considered values of $a$ and $b$ is stronger than $O(z^{-2})$, ensuring that any infinitely large arc in the complex plane will not contribute. Our desired integral, located on the real axis, can therefore be extended to a closed loop and thus be expressed as a sum of residues (see Fig. 12). By always choosing the half-circle opposite to the pole of the Matsubara frequency we can restrict the evaluation to exactly one (higher-order) pole

$$\mathcal{I}_{ab}(\mathbf{k}, n; \Omega; i\nu_m) = \oint_{\mathcal{C}} dz\, \frac{(z - \Omega)^{(a+b-2)}}{[(z - a - \Omega)^2 + \Gamma^2]^2} \frac{1}{i\nu_m - z}$$

$$= -\text{sign}(\nu_m)\, 2\pi i\, \text{Res}_{z=(a+\Omega-i\,\text{sign}(\nu_m)\Gamma)} \frac{(z - \Omega)^{(a+b-2)}}{[(z - a - \Omega)^2 + \Gamma^2]^2} \frac{1}{i\nu_m - z}. \qquad (A7)$$

Due to the different mathematical integration directions, positive and negative fermionic Matsubara frequencies result in differing signs. Evaluating the residue at $z = [a + \Omega - i\,\text{sign}(\nu_m)\Gamma]$ results in the following expressions:

$$\mathcal{I}_{11}(\mathbf{k}, n; \Omega; i\nu_m) = \frac{\pi}{2\Gamma^3} \left[\frac{i\Gamma\,\text{sign}(\nu_m)}{[i\nu_m - a - \Omega + i\Gamma\text{sign}(\nu_m)]^2} + \frac{1}{[i\nu_m - a - \Omega + i\Gamma\,\text{sign}(\nu_m)]}\right], \qquad (A8)$$

$$\mathcal{I}_{12}(\mathbf{k}, n; \Omega; i\nu_m) = \frac{\pi}{2\Gamma^3} \left[\frac{\Gamma^2 + i(a + \Omega)\Gamma\,\text{sign}(\nu_m)}{[i\nu_m - a - \Omega + i\Gamma\,\text{sign}(\nu_m)]^2} + \frac{(a + \Omega)}{[i\nu_m - a - \Omega + i\Gamma\,\text{sign}(\nu_m)]}\right] - \Omega\mathcal{I}_{11}(\mathbf{k}, n; \Omega; i\nu_m), \qquad (A9)$$

$$\mathcal{I}_{22}(\mathbf{k}, n; \Omega; i\nu_m) = \frac{\pi}{2\Gamma^3} \left[\frac{i(a + \Omega)^2\Gamma\,\text{sign}(\nu_m) + 2(a + \Omega)\Gamma^2 - i\Gamma^3\text{sign}(\nu_m)}{[i\nu_m - a - \Omega + i\Gamma\,\text{sign}(\nu_m)]^2} + \frac{\Gamma^2 + (a + \Omega)^2}{[i\nu_m - a - \Omega + i\Gamma\,\text{sign}(\nu_m)]}\right]$$

$$- 2\Omega\mathcal{I}_{12}(\mathbf{k}, n; \Omega; i\nu_m) + \Omega^2\mathcal{I}_{11}(\mathbf{k}, n; \Omega; i\nu_m). \qquad (A10)$$

### 2. Matsubara sums

The second step is to perform the Matsubara sums

$$\mathcal{I}_{ab}(\mathbf{k}, n; \Omega) = \frac{1}{\beta} \sum_{m=-\infty}^{\infty} \mathcal{I}_{ab}(\mathbf{k}, n; \Omega; i\nu_m). \qquad (A11)$$

Using the series representation of the digamma and polygamma functions

$$\psi(z) = -\gamma + \sum_{n=1}^{\infty}\left(\frac{1}{n} - \frac{1}{n+z}\right), \qquad (A12)$$

$$\psi_{m>0}(z) = (-1)^{m+1}m!\sum_{k=0}^{\infty}\frac{1}{(z+k)^{m+1}}, \qquad (A13)$$

the summations appearing in Eqs. (A8)–(A10) result in

$$\frac{1}{\beta}\sum_{m=-\infty}^{\infty}\frac{1}{iv_m - a + i\Gamma\,\text{sign}(v_m)} = -\frac{1}{\pi}\text{Im}\psi(z), \qquad (A14)$$

$$\frac{1}{\beta}\sum_{m=-\infty}^{\infty}\frac{1}{[iv_m - a + i\Gamma\,\text{sign}(v_m)]^2} = -\frac{\beta}{2\pi^2}\text{Re}\psi_1(z), \qquad (A15)$$

$$\frac{1}{\beta}\sum_{m=-\infty}^{\infty}\frac{\text{sign}(v_m)}{[iv_m - a + i\Gamma\,\text{sign}(v_m)]^2} = -\frac{i\beta}{2\pi^2}\text{Im}\psi_1(z), \qquad (A16)$$

$$\frac{1}{\beta}\sum_{m=-\infty}^{\infty}\frac{1}{[iv_m - a + i\Gamma\,\text{sign}(v_m)]^3} = \frac{\beta^2}{8\pi^2}\text{Im}\psi_2(z), \qquad (A17)$$

$$\frac{1}{\beta}\sum_{m=-\infty}^{\infty}\frac{\text{sign}(v_m)}{[iv_m - a + i\Gamma\,\text{sign}(v_m)]^3} = -\frac{i\beta^2}{8\pi^3}\text{Re}\psi_2(z) \qquad (A18)$$

with $z = \frac{1}{2} + \frac{\beta}{2\pi}(\Gamma + ia)$. Then, the transport integrals simplify to

$$\mathcal{I}_{11}(\mathbf{k}, n; \Omega) = \frac{\pi}{2\Gamma^3}\left[\frac{\beta\Gamma}{2\pi^2}\text{Im}\psi_1\left(z + \frac{i\beta\Omega}{2\pi}\right)\right.$$
$$\left. -\frac{1}{\pi}\text{Im}\psi\left(z + \frac{i\beta\Omega}{2\pi}\right)\right], \qquad (A19)$$

$$\mathcal{I}_{12}(\mathbf{k}, n; \Omega) = \frac{\pi}{2\Gamma^3}\left[-\frac{\Gamma^2\beta}{2\pi^2}\text{Re}\psi_1\left(z + \frac{i\beta\Omega}{2\pi}\right)\right.$$
$$+\frac{(a+\Omega)\Gamma\beta}{2\pi^2}\text{Im}\psi_1\left(z + \frac{i\beta\Omega}{2\pi}\right)$$
$$\left. -\frac{(a+\Omega)}{\pi}\text{Im}\psi\left(z + \frac{i\beta\Omega}{2\pi}\right)\right]$$
$$-\Omega\mathcal{I}_{11}(\mathbf{k}, n; \Omega), \qquad (A20)$$

$$\mathcal{I}_{22}(\mathbf{k}, n; \Omega) = \frac{\pi}{2\Gamma^3}\left[\frac{(a+\Omega)^2\Gamma\beta - \Gamma^3\beta}{2\pi^2}\text{Im}\psi_1\left(z + \frac{i\beta\Omega}{2\pi}\right)\right.$$
$$-\frac{2(a+\Omega)\Gamma^2\beta}{2\pi^2}\text{Re}\psi_1\left(z + \frac{i\beta\Omega}{2\pi}\right)$$

$$\left. -\frac{\Gamma^2 + (a+\Omega)^2}{\pi}\text{Im}\psi\left(z + \frac{i\beta\Omega}{2\pi}\right)\right]$$
$$-2\Omega\mathcal{I}_{12}(\mathbf{k}, n; \Omega) + \Omega^2\mathcal{I}_{11}(\mathbf{k}, n; \Omega). \qquad (A21)$$

### 3. Dynamic limit

Taylor expanding the frequency-dependent digamma and polygamma ($m > 0$) functions around $z$,

$$\psi\left(z + \frac{i\beta\Omega}{2\pi}\right) = \psi(z) + \frac{i\beta\Omega}{2\pi}\psi_1(z) + O(\Omega^2), \qquad (A22)$$

$$\psi_m\left(z + \frac{i\beta\Omega}{2\pi}\right) = \psi_m(z) + \frac{i\beta\Omega}{2\pi}\psi_{m+1}(z) + O(\Omega^2) \qquad (A23)$$

and evaluating the limit

$$\mathcal{K}_{ab}(\mathbf{k}, n) = \frac{Z^2\Gamma^2}{\pi^2}\lim_{\Omega\to 0^+}\frac{1}{\Omega}[\mathcal{I}_{ab}(\mathbf{k}, n; 0) - \mathcal{I}_{ab}(\mathbf{k}, n; \Omega)], \qquad (A24)$$

we finally arrive at the intraband equations listed in the main text:

$$\mathcal{K}_{11}(\mathbf{k}, n) = \frac{Z^2\beta}{4\pi^3\Gamma}\left[\text{Re}\psi_1(z) - \frac{\beta\Gamma}{2\pi}\text{Re}\psi_2(z)\right], \qquad (A25)$$

$$\mathcal{K}_{12}(\mathbf{k}, n) = \frac{Z^2\beta}{4\pi^3\Gamma}\left[a\,\text{Re}\psi_1(z) - \frac{a\Gamma\beta}{2\pi}\text{Re}\psi_2(z)\right.$$
$$\left. -\frac{\Gamma^2\beta}{2\pi}\text{Im}\psi_2(z)\right], \qquad (A26)$$

$$\mathcal{K}_{22}(\mathbf{k}, n) = \frac{Z^2\beta}{4\pi^3\Gamma}\left[(a^2+\Gamma^2)\text{Re}\psi_1(z)\right.$$
$$\left. +\frac{\beta}{2\pi}\Gamma(\Gamma^2 - a^2)\text{Re}\psi_2(z) - \frac{\beta}{\pi}a\Gamma^2\text{Im}\psi_2(z)\right]. \qquad (A27)$$

The equivalent calculation is easily performed for the magnetic kernel functions

$$\mathcal{K}_{ab}^B(\mathbf{k}, n) = \int_{-\infty}^{\infty}d\omega\,\omega^{(a+b-2)}\left(-\frac{\partial f}{\partial\omega}\right)A_{\mathbf{k}n}^3(\omega) \qquad (A28)$$

which only differ from their nonmagnetic counterparts by an additional spectral function. The integrals

$$\mathcal{I}_{ab}^B(\mathbf{k}, n; \Omega; iv_m) = \int_{-\infty}^{\infty}d\omega\,\frac{(\omega - \Omega)^{(a+b-2)}}{[(\omega - a - \Omega)^2 + \Gamma^2]^3}\frac{1}{iv_m - \omega} \qquad (A29)$$

then evaluate to

$$\mathcal{I}_{11}^B(\mathbf{k}, n; \Omega; iv_m) = \frac{i\pi}{8\Gamma^5}\left[\frac{2i\Gamma^2}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]^3} + \frac{3\Gamma\,\text{sign}(v_m)}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]^2} + \frac{-3i}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]}\right], \qquad (A30)$$

$$\mathcal{I}_{12}^B(\mathbf{k}, n; \Omega; iv_m) = \frac{i\pi}{8\Gamma^5}\left[\frac{2i(a+\Omega)\Gamma^2 + 2\Gamma^3\text{sign}(v_m)}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]^3} + \frac{3(a+\Omega)\Gamma\,\text{sign}(v_m) - i\Gamma^2}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]^2} + \frac{-3i(a+\Omega)}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]}\right]$$
$$-\Omega\mathcal{I}_{11}^B(\mathbf{k}, n; \Omega; iv_m), \qquad (A31)$$

$$\mathcal{I}_{22}^{B}(\mathbf{k}, n; \Omega; iv_m) = \frac{i\pi}{8\Gamma^5} \left[ \frac{2i(a + \Omega)^2\Gamma^2 + 4(a + \Omega)\Gamma^3\text{sign}(v_m) - 2i\Gamma^4}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]^3} + \frac{-2i(a + \Omega)\Gamma^2 + \Gamma^3\text{sign}(v_m) + 3(a + \Omega)^2\Gamma\,\text{sign}(v_m)}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]^2} \right.$$

$$\left. + \frac{-i\Gamma^2 - 3i(a + \Omega)^2}{[iv_m - a - \Omega + i\Gamma\,\text{sign}(v_m)]} \right] - 2\Omega\mathcal{I}_{12}^{B}(\mathbf{k}, n; \Omega; iv_m) + \Omega^2\mathcal{I}_{11}^{B}(\mathbf{k}, n; \Omega; iv_m). \tag{A32}$$

Performing the Matsubara sums and taking the dynamical limit, the transport kernels results in the magnetic transport kernels

$$\mathcal{K}_{11}^{B}(\mathbf{k}, n) = \frac{Z^3\beta}{16\pi^4\Gamma^2} \left[ 3\,\text{Re}\,\psi_1(z) - \frac{3\Gamma\beta}{2\pi}\text{Re}\,\psi_2(z) + \frac{\Gamma^2\beta^2}{4\pi^2}\text{Re}\,\psi_3(z) \right], \tag{A33}$$

$$\mathcal{K}_{12}^{B}(\mathbf{k}, n) = \frac{Z^3\beta}{16\pi^4\Gamma^2} \left[ 3a\,\text{Re}\,\psi_1(z) - \frac{3a\Gamma\beta}{2\pi}\text{Re}\,\psi_2(z) - \frac{\Gamma^2\beta}{2\pi}\text{Im}\,\psi_2(z) + \frac{a\Gamma^2\beta^2}{4\pi^2}\text{Re}\,\psi_3(z) + \frac{\Gamma^3\beta^2}{4\pi^2}\text{Im}\,\psi_3(z) \right], \tag{A34}$$

$$\mathcal{K}_{22}^{B}(\mathbf{k}, n) = \frac{Z^3\beta}{16\pi^4\Gamma^2} \left[ (\Gamma^2 + 3a^2)\text{Re}\,\psi_1(z) - \frac{\beta\Gamma(\Gamma^2 + 3a^2)}{2\pi}\text{Re}\,\psi_2(z) - \frac{\beta a\Gamma^2}{\pi}\text{Im}\,\psi_2(z) \right.$$

$$\left. - \frac{\beta^2\Gamma^2(\Gamma^2 - a^2)}{4\pi^2}\text{Re}\,\psi_3(z) + \frac{\beta^2 a\Gamma^3}{2\pi^2}\text{Im}\,\psi_3(z) \right]. \tag{A35}$$

## APPENDIX B: LOW-TEMPERATURE EXPANSION

In order to evaluate the low-temperature behavior we exploit the Taylor expansion around $\bar{z} = \infty$ of the digamma function

$$\Psi\left(\frac{1}{2} + \bar{z}\right) = \ln(\bar{z}) + \frac{1}{24\bar{z}^2} - \frac{7}{960\bar{z}^4} + O(\bar{z}^{-6}), \tag{B1}$$

and all higher-order polygamma functions

$$\Psi_1\left(\frac{1}{2} + \bar{z}\right) = \frac{1}{\bar{z}} - \frac{1}{12\bar{z}^3} + \frac{7}{240\bar{z}^5} + O(\bar{z}^{-7}), \tag{B2}$$

$$\Psi_2\left(\frac{1}{2} + \bar{z}\right) = -\frac{1}{\bar{z}^2} + \frac{1}{4\bar{z}^4} - \frac{7}{48\bar{z}^6} + O(\bar{z}^{-8}), \tag{B3}$$

$$\Psi_3\left(\frac{1}{2} + \bar{z}\right) = \frac{2}{\bar{z}^3} - \frac{1}{\bar{z}^5} + \frac{7}{8\bar{z}^7} + O(\bar{z}^{-9}). \tag{B4}$$

Equations (25)–(30) then become

$$\mathcal{K}_{11} = \frac{Z^2}{\pi^2}\frac{\Gamma^2}{(a^2 + \Gamma^2)^2}\left[1 + \frac{2\pi^2}{3}\frac{5a^2 - \Gamma^2}{(a^2 + \Gamma^2)^2}k_B^2T^2 + O(T^4)\right], \tag{B5}$$

$$\mathcal{K}_{12} = \frac{4Z^2}{3}\frac{a\Gamma^2}{(a^2 + \Gamma^2)^3} \times \left[k_B^2T^2 + \frac{7\pi^2}{5}\frac{5a^2 - 3\Gamma^2}{(a^2 + \Gamma^2)^2}k_B^4T^4 + O(T^6)\right], \tag{B6}$$

$$\mathcal{K}_{22} = \frac{Z^2}{3}\frac{\Gamma^2}{(a^2 + \Gamma^2)^2} \times \left[k_B^2T^2 + \frac{14\pi^2}{5}\frac{5a^2 - \Gamma^2}{(a^2 + \Gamma^2)^2}k_B^4T^4 + O(T^6)\right], \tag{B7}$$

$$\mathcal{K}_{11}^{B} = \frac{Z^3}{\pi^3}\frac{\Gamma^3}{(a^2 + \Gamma^2)^3}\left[1 + \pi^2\frac{7a^2 - \Gamma^2}{(a^2 + \Gamma^2)^2}k_B^2T^2 + O(T^4)\right], \tag{B8}$$

$$\mathcal{K}_{12}^{B} = \frac{2Z^3}{\pi}\frac{a\Gamma^3}{(a^2 + \Gamma^2)^4}$$

$$\times \left[k_B^2T^2 + \frac{28\pi^2}{15}\frac{7a^2 - 3\Gamma^2}{(a^2 + \Gamma^2)^2}k_B^4T^4 + O(T^6)\right], \tag{B9}$$

$$\mathcal{K}_{22}^{B} = \frac{Z^3}{3\pi}\frac{\Gamma^3}{(a^2 + \Gamma^3)^3}$$

$$\times \left[k_B^2T^2 + \frac{21\pi^2}{5}\frac{7a^2 - \Gamma^2}{(a^2 + \Gamma^2)^2}k_B^4T^4 + O(T^6)\right]. \tag{B10}$$

Therefore, for any $\lim_{T\to 0}\Gamma(T) > 0$ the resistivity as well as the Hall coefficient will saturate. Furthermore, the Seebeck coefficient, the Nernst coefficient, as well as the thermal conductivity will tend to zero in a linear fashion.

## APPENDIX C: KERNEL APPROXIMATIONS

In order to better understand the transport kernels, we consider two types of approximations to the full kernel expressions: the $\psi_1$ approximation, as the name suggests, consists of simply using the lowest *explicit* order in $\Gamma$, resulting in expressions that functionally depend only on the first-order polygamma function, e.g.,

$$\mathcal{K}_{11}^{\psi_1}(\mathbf{k}, n) = \frac{Z^2\beta}{4\pi^3\Gamma}\left[\text{Re}\,\psi_1(z) \underbrace{- \frac{\beta\Gamma}{2\pi}\text{Re}\,\psi_2(z)}_{\stackrel{!}{=}0}\right]. \tag{C1}$$

Due to the additional, *implicit* $\Gamma$ dependence through the polygamma argument $z = \frac{1}{2} + \frac{\beta}{2\pi}(\Gamma + ia)$, this approximation is not reasonable in the zero-temperature limit where the $\psi_2$ term becomes equally important (see Appendix B). A better approximation is to go one step further and take the full leading $\Gamma$ dependence into consideration: Expanding for small
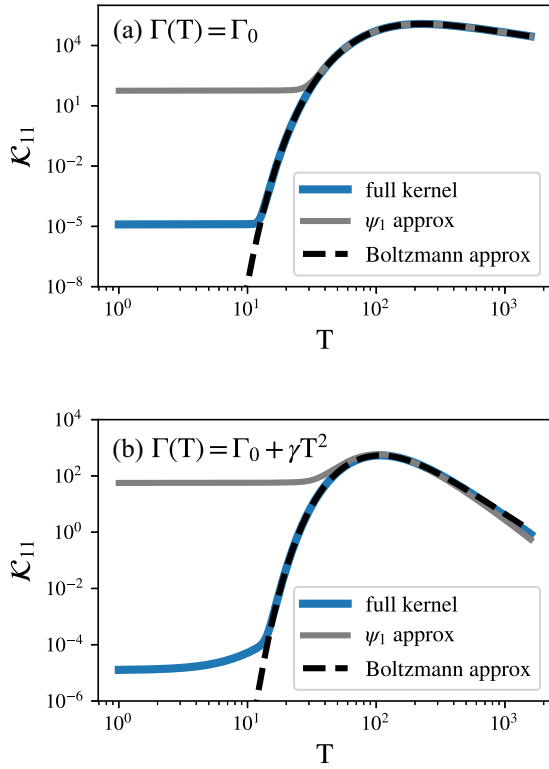
FIG. 13. $\psi_1$ approximation and Boltzmann approximation of the full kernel expression $\mathcal{K}_{11}$ for $Z = 1$, $a = 0.03$, $\Gamma_0 = 10^{-5}$, and $\gamma = 10^{-7}$. (a) Kernel behavior with temperature-independent scattering rate $\Gamma(T) = \Gamma_0$. (b) Kernel behavior with Fermi-liquid-like scattering rate $\Gamma(T) = \Gamma_0 + \gamma T^2$. Using the lowest-order polygamma function is a good approximation at high temperature even for relatively large scattering rates. At low enough temperatures, we find a crossover where both terms of the kernel contribute equally. Surprisingly, the Boltzmann approximation follows the full kernel expression longer, however, does not saturate.

$\beta\Gamma$, i.e., around $z = \frac{1}{2} + \frac{i\beta a}{2\pi}$,

$$\mathrm{Re}\,\Psi_n\left(\frac{1}{2} + \frac{\beta}{2\pi}(\Gamma + ia)\right)$$
$$= \mathrm{Re}\,\psi_n\left(\frac{1}{2} + \frac{i\beta a}{2\pi}\right) + \frac{\beta\Gamma}{2\pi}\mathrm{Re}\,\psi_{n+1}\left(\frac{1}{2} + \frac{i\beta a}{2\pi}\right)$$
$$+ O(\beta^2\Gamma^2), \tag{C2}$$

and inserting Eq. (C2) into the initial kernel expression results in an exact cancellation of all terms linear in $\Gamma$, recovering the Boltzmann expression

$$\mathcal{K}_{11}^{\mathrm{Boltzmann}}(\mathbf{k}, n) = \frac{Z^2\beta}{4\pi^3\Gamma}\left[\underbrace{\mathrm{Re}\,\psi_1\left(\frac{1}{2} + \frac{i\beta a}{2\pi}\right)}_{-\frac{2\pi^2}{\beta}f'(a)} + \underbrace{O(\beta^2\Gamma^2)}_{\stackrel{!}{=}0}\right] \tag{C3}$$

which ignores terms of the order $(\beta\Gamma)^2$ and higher.

The full kernel and the corresponding approximations, illustrated in Fig. 13, show that the $\psi_1$ approximation produces a finite saturation value in the zero-temperature limit, that, however, emerges at too large $T$. Taking the implicit $\Gamma$ dependence into account, the resulting Boltzmann approximation (dashed) leads to a better agreement at intermediate temperatures. Yet, this comes with a tradeoff: the Boltzmann kernel is able to trace the full kernel down to lower temperature, but does not saturate, $\mathcal{K}_{11}^{\mathrm{Boltzmann}}(T \to 0) = 0$. Identical behaviors can be observed for the higher-order transport kernels. Combined to observable transport tensors this deficiency in the limiting behavior causes the usual problem of the relaxation-time approximation with its nonphysical entropy transport in gapped systems. Employing the full kernels on the other hand does not suffer from this problem and entropy-transport quantities become thermodynamically consistent.

## APPENDIX D: SCALING WITH LATTICE CONSTANTS AND BAND STRUCTURE

Due to the reductionist approach in the main text (only nearest-neighbor hopping, cubic lattice constant $a_{\mathrm{lattice}} = 1$ Å) the absolute values of the transport properties were for the most part ignored. A first step towards more realism is using a proper lattice spacing $a_{\mathrm{lattice}}$: the (cubic) unit-cell volume is $V = a_{\mathrm{lattice}}^3$, while the optical elements within the Peierls approximation (Sec. II B 2) scale with $M \propto a_{\mathrm{lattice}}^2$ and $M^B \propto a_{\mathrm{lattice}}^4$. Combined, the Onsager coefficients scale like

$$\mathcal{L}_{ab} \propto a_{\mathrm{lattice}}^{-1}, \tag{D1}$$

$$\mathcal{L}_{ab}^B \propto a_{\mathrm{lattice}}. \tag{D2}$$

Equations (6)–(12) then naturally lead to

$$\rho \propto a_{\mathrm{lattice}}, \tag{D3}$$

$$S \equiv \mathrm{const}, \tag{D4}$$

$$R_H \propto a_{\mathrm{lattice}}^3, \tag{D5}$$

$$\nu \propto a_{\mathrm{lattice}}^2. \tag{D6}$$

The second step is a more realistic band structure and transition matrix elements. An input from density functional calculations and including beyond-Peierls matrix elements is straightforward within LINRETRACE [68]. Here, however, we wanted to restrict ourselves to a tight-binding model with hoppings $t$ to extract the most essential physics. Then, matrix elements transform like $M \propto t^2$ and $M^B \propto t^3$. In a multiband situation, an overall scaling of results for varying hopping parameters is not possible.

[1] J. M. Tomczak, Thermoelectricity in correlated narrow-gap semiconductors, J. Phys.: Condens. Matter (Topical Review) **30**, 183001 (2018).

[2] S. Poncé, W. Li, S. Reichardt, and F. Giustino, First-principles calculations of charge carrier mobility and conductivity in bulk semiconductors and two-dimensional materials, Rep. Prog. Phys. **83**, 036501 (2020).

[3] M. Pickem, E. Maggio, and J. M. Tomczak, Resistivity saturation in Kondo insulators, Commun. Phys. **4**, 226 (2021).

[4] G. K. H. Madsen and D. J. Singh, BoltzTraP. A code for calculating band-structure dependent quantities, Comput. Phys. Commun. **175**, 67 (2006).

[5] G. K. H. Madsen, J. Carrete, and M. J. Verstraete, BoltzTraP2, a program for interpolating band structures and calculating semi-classical transport coefficients, Comput. Phys. Commun. **231**, 140 (2018).

[6] G. Pizzi, D. Volja, B. Kozinsky, M. Fornari, and N. Marzari, BoltzWann: A code for the evaluation of thermoelectric and electronic transport properties with a maximally-localized Wannier functions basis, Comput. Phys. Commun. **185**, 422 (2014).

[7] A. Bentien, S. Johnsen, G. K. H. Madsen, B. B. Iversen, and F. Steglich, Colossal Seebeck coefficient in strongly correlated semiconductor $FeSb_2$, Europhys. Lett. **80**, 39901 (2007).

[8] C. Petrovic, Y. Lee, T. Vogt, N. D. Lazarov, S. L. Bud'ko, and P. C. Canfield, Kondo insulator description of spin state transition in $FeSb_2$, Phys. Rev. B **72**, 045103 (2005).

[9] P. Sun, W. Xu, J. M. Tomczak, G. Kotliar, M. Søndergaard, B. B. Iversen, and F. Steglich, Highly dispersive electron relaxation and colossal thermoelectricity in the correlated semiconductor $FeSb_2$, Phys. Rev. B **88**, 245203 (2013).

[10] A. Chikina, J.-Z. Ma, W. H. Brito, S. Choi, P. Sémon, A. Kutepov, Q. Du, J. Jandke, H. Liu, N. C. Plumb, M. Shi, C. Petrovic, M. Radovic, and G. Kotliar, Correlated electronic structure of colossal thermopower $FeSb_2$: An ARPES and ab initio study, Phys. Rev. Res. **2**, 023190 (2020).

[11] J. M. Tomczak, K. Haule, T. Miyake, A. Georges, and G. Kotliar, Thermopower of correlated semiconductors: Application to $FeAs_2$ and $FeSb_2$, Phys. Rev. B **82**, 085104 (2010).

[12] M. Pokharel, H. Zhao, K. Lukas, Z. Ren, C. Opeil, and B. Mihaila, Phonon drag effect in nanocomposite $FeSb_2$, MRS Commun. **3**, 31 (2013).

[13] H. Takahashi, R. Okazaki, S. Ishiwata, H. Taniguchi, A. Okutani, M. Hagiwara, and I. Terasaki, Colossal Seebeck effect enhanced by quasi-ballistic phonons dragging massive electrons in $FeSb_2$, Nat. Commun. **7**, 12732 (2016).

[14] M. Battiato, J. M. Tomczak, Z. Zhong, and K. Held, Unified Picture for the Colossal Thermopower Compound $FeSb_2$, Phys. Rev. Lett. **114**, 236603 (2015).

[15] H. Matsuura, H. Maebashi, M. Ogata, and H. Fukuyama, Effect of phonon drag on Seebeck coefficient based on linear response theory: Application to FeSb2, J. Phys. Soc. Jpn. **88**, 074601 (2019).

[16] R. Masuki, T. Nomoto, and R. Arita, Origin of anomalous temperature dependence of the Nernst effect in narrow-gap semiconductors, Phys. Rev. B **103**, L041202 (2021).

[17] P. Sun, N. Oeschler, S. Johnsen, B. B. Iversen, and F. Steglich, Narrow band gap and enhanced thermoelectricity in $FeSb_2$, Dalton Trans. **39**, 1012 (2010).

[18] M. Wagner-Reetz, D. Kasinathan, W. Schnelle, R. Cardoso-Gil, H. Rosner, Y. Grin, and P. Gille, Phonon-drag effect in $FeGa_3$, Phys. Rev. B **90**, 195206 (2014).

[19] H. Takahashi, R. Okazaki, I. Terasaki, and Y. Yasui, Origin of the energy gap in the narrow-gap semiconductor $FeSb_2$ revealed by high-pressure magnetotransport measurements, Phys. Rev. B **88**, 165205 (2013).

[20] Q. Du and C. Petrovic, Optimal carrier concentration for $FeSb_2$ colossal thermopower, Appl. Phys. Lett. **118**, 233901 (2021).

[21] Q. Du, L. Wu, H. Cao, C.-J. Kang, C. Nelson, G. L. Pascut, T. Besara, T. Siegrist, K. Haule, G. Kotliar, I. Zaliznyak, Y. Zhu, and C. Petrovic, Vacancy defect control of colossal thermopower in $FeSb_2$, npj Quantum Mater. **6**, 13 (2021).

[22] P. Sun, N. Oeschler, S. Johnsen, B. B. Iversen, and F. Steglich, Huge thermoelectric power factor: $FeSb_2$ versus $FeAs_2$ and $RuSb_2$, Appl. Phys. Express **2**, 091102 (2009).

[23] P. Sun, N. Oeschler, S. Johnsen, B. B. Iversen, and F. Steglich, Thermoelectric properties of the narrow-gap semiconductors $FeSb_2$ and $RuSb_2$ : A comparative study, J. Phys.: Conf. Ser. **150**, 012049 (2009).

[24] M. K. Fuccillo, Q. D. Gibson, M. N. Ali, L. M. Schoop, and R. J. Cava, Correlated evolution of colossal thermoelectric effect and Kondo insulating behavior, APL Mater. **1**, 062102 (2013).

[25] B. C. Sales, A. F. May, M. A. McGuire, M. B. Stone, D. J. Singh, and D. Mandrus, Transport, thermal, and magnetic properties of the narrow-gap semiconductor $CrSb_2$, Phys. Rev. B **86**, 235136 (2012).

[26] R. Wolfe, J. H. Wernick, and S. E. Haszko, Thermoelectric properties of FeSi, Phys. Lett. **19**, 449 (1965).

[27] B. C. Sales, O. Delaire, M. A. McGuire, and A. F. May, Thermoelectric properties of Co-, Ir-, and Os-doped FeSi alloys: Evidence for strong electron-phonon coupling, Phys. Rev. B **83**, 125209 (2011).

[28] J. M. Tomczak, K. Haule, and G. Kotliar, Signatures of electronic correlations in iron silicide, Proc. Natl. Acad. Sci. USA **109**, 3243 (2012).

[29] H. Hohl, A. P Ramirez, C. Goldmann, G. Ernst, and E. Bucher, Transport properties of RuSi, RuGe, OsSi, and quasi-binary alloys of these compounds, J. Alloys Compd. **278**, 39 (1998).

[30] B. Buschinger, C. Geibel, J. Diehl, M. Weiden, W. Guth, A. Wildbrett, S. Horn, and F. Steglich, Preparation and low temperature properties of FeSi-type RuSi, J. Alloys Compd. **256**, 57 (1997).

[31] J. M. Tomczak, K. Haule, and G. Kotliar, Thermopower of the correlated narrow gap semiconductor FeSi and comparison to RuSi, in *New Materials for Thermoelectric Applications: Theory and Experiment*, NATO Science for Peace and Security Series B: Physics and Biophysics, edited by V. Zlatić and A. Hewson (Springer, Dordrecht, 2013), pp. 45–57, doi: 10.1007/978-94-007-4984-9_4.

[32] Y. Nishino, M. Kato, S. Asano, K. Soda, M. Hayasaki, and U. Mizutani, Semiconductorlike Behavior of Electrical Resistivity in Heusler-Type $Fe_2VAl$ Compound, Phys. Rev. Lett. **79**, 1909 (1997).

[33] M. Kato, Y. Nishino, U. Mizutani, and S. Asano, Electronic, magnetic and transport properties of $(Fe_{1-x}V_x)_3Al$ alloys, J. Phys.: Condens. Matter **12**, 1769 (2000).

[34] T. Naka, A. M. Nikitin, Y. Pan, A. de Visser, T. Nakane, F. Ishikawa, Y. Yamada, M. Imai, and A. Matsushita, Composition induced metal-insulator quantum phase transition in the Heusler type $Fe_2VAl$, J. Phys.: Condens. Matter **28**, 285601 (2016).

[35] T. Nakama, Y. Takaesu, K. Yagasaki, T. Naka, A. Matsushita, K. Fukuda, and Y. Yamada, Transport properties of Heusler compounds $Fe_{3-x}V_xAl$, J. Phys. Soc. Jpn. **74**, 1378 (2005).

[36] B. Hinterleitner, I. Knapp, M. Poneder, Y. Shi, H. Müller, G. Eguchi, C. Eisenmenger-Sittner, M. Stöger-Pollach, Y.

Kakefuda, N. Kawamoto, Q. Guo, T. Baba, T. Mori, S. Ullah, X.-Q. Chen, and E. Bauer, Thermoelectric performance of a metastable thin-film Heusler alloy, Nature (London) **576**, 85 (2019).

[37] M. S. Likhanov and A. V. Shevelkov, Intermetallic compounds with non-metallic properties, Russian Chem. Bull. **69**, 2231 (2020).

[38] Y. Hadano, S. Narazu, M. A. Avila, T. Onimaru, and T. Takabatake, Thermoelectric and magnetic properties of a narrow-gap semiconductor $FeGa_3$, J. Phys. Soc. Jpn. **78**, 013702 (2009).

[39] M. B. Gamża, J. M. Tomczak, C. Brown, A. Puri, G. Kotliar, and M. C. Aronson, Electronic correlations in $FeGa_3$ and the effect of hole doping on its magnetic properties, Phys. Rev. B **89**, 195102 (2014).

[40] M. F. Hundley, P. C. Canfield, J. D. Thompson, Z. Fisk, and J. M. Lawrence, Hybridization gap in $Ce_3Bi_4Pt_3$, Phys. Rev. B **42**, 6842 (1990).

[41] J. C. Cooley, M. C. Aronson, and P. C. Canfield, High pressures and the Kondo gap in $Ce_3Bi_4Pt_3$, Phys. Rev. B **55**, 7533 (1997).

[42] K. Katoh and T. Takabatake, Crystal growth and semiconducting properties of $Ce_3Bi_4Pt_3$, J. Alloys Compd. **268**, 22 (1998).

[43] N. Wakeham, P. F. S. Rosa, Y. Q. Wang, M. Kang, Z. Fisk, F. Ronning, and J. D. Thompson, Low-temperature conducting state in two candidate topological Kondo insulators: $SmB_6$ and $Ce_3Bi_4Pt_3$, Phys. Rev. B **94**, 035127 (2016).

[44] D. J. Campbell, Z. E. Brubaker, C. Roncaioli, P. Saraf, Y. Xiao, P. Chow, C. Kenney-Benson, D. Popov, R. J. Zieve, J. R. Jeffries, and J. Paglione, Pressure-driven valence increase and metallization in the Kondo insulator $Ce_3Bi_4Pt_3$, Phys. Rev. B **100**, 235133 (2019).

[45] J. M. Tomczak, Realistic many-body theory of Kondo insulators: Renormalizations and fluctuations in $Ce_3Bi_4Pt_3$, arXiv:1904.01346.

[46] G. D. Mahan, *Many-particle Physics* (Plenum, New York, 1990).

[47] V. Zlatic and R. Monnier, *Modern Theory of Thermoelectricity* (Oxford University Press, Oxford, 2014).

[48] K. Behnia, *Fundamentals of Thermoelectricity* (Oxford University Press, Oxford, 2015).

[49] W. Xu, *Transport and Magnetic Properties of Correlated Electron Systems* (Rutgers State University Press, Rutgers, NJ, 2014).

[50] M. Pickem, Electronic correlations and transport phenomena in Mott and Kondo Materials, Ph.D. thesis, TU Wien, 2022.

[51] P. Voruganti, A. Golubentsev, and S. John, Conductivity and Hall effect in the two-dimensional Hubbard model, Phys. Rev. B **45**, 13945 (1992).

[52] H. Kohno and K. Yamada, A General Expression for Hall Coefficient Based on Fermi Liquid Theory, Prog. Theor. Phys. **80**, 623 (1988).

[53] H. Kontani, *Transport Phenomena in Strongly Correlated Fermi Liquids*, Springer Tracts in Modern Physics (Springer, Berlin, 2013).

[54] H. Kontani, General formula for the thermoelectric transport phenomena based on Fermi liquid theory: Thermoelectric power, Nernst coefficient, and thermal conductivity, Phys. Rev. B **67**, 014408 (2003).

[55] R. Nourafkan and A.-M. S. Tremblay, Hall and Faraday effects in interacting multiband systems with arbitrary band topology and spin-orbit coupling, Phys. Rev. B **98**, 165130 (2018).

[56] J. Vučičević, J. Kokalj, R. Žitko, N. Wentzell, D. Tanasković, and J. Mravlje, Conductivity in the Square Lattice Hubbard Model at High Temperatures: Importance of Vertex Corrections, Phys. Rev. Lett. **123**, 036601 (2019).

[57] A. Kauch, P. Pudleiner, K. Astleithner, P. Thunström, T. Ribic, and K. Held, Generic Optical Excitations of Correlated Systems: $\pi$-tons, Phys. Rev. Lett. **124**, 047401 (2020).

[58] O. Simard, S. Takayoshi, and P. Werner, Diagrammatic study of optical excitations in correlated systems, Phys. Rev. B **103**, 104415 (2021).

[59] P. Worm, C. Watzenböck, M. Pickem, A. Kauch, and K. Held, Broadening and sharpening of the Drude peak through antiferromagnetic fluctuations, Phys. Rev. B **104**, 115153 (2021).

[60] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Rev. Mod. Phys. **68**, 13 (1996).

[61] A. Khurana, Electrical Conductivity in the infinite-dimensional Hubbard Model, Phys. Rev. Lett. **64**, 1990 (1990).

[62] I. Paul and G. Kotliar, Thermal transport for many-body tight-binding models, Phys. Rev. B **67**, 115131 (2003).

[63] A. A. Markov, G. Rohringer, and A. N. Rubtsov, Robustness of the topological quantization of the Hall conductivity for correlated lattice electrons at finite temperatures, Phys. Rev. B **100**, 115102 (2019).

[64] A. J. Millis, Optical conductivity and correlated electron physics, in *Strong Interactions in Low Dimensions*, Physics and Chemistry of Materials with Low-Dimensional Structures, edited by L. Degiorgi and D. Baeriswyl (Kluwer Academic, Dordrecht, 2004), Vol. 25, p. 195ff.

[65] J. M. Tomczak and S. Biermann, Optical properties of correlated materials: Generalized Peierls approach and its application to $VO_2$, Phys. Rev. B **80**, 085117 (2009).

[66] J. M. Tomczak, Spectral and Optical Properties of Correlated Materials, Ph.D. thesis, Ecole Polytechnique, France, 2007.

[67] M. Zebarjadi, S. E. Rezaei, M. S. Akhanda, and K. Esfarjani, Nernst coefficient within relaxation time approximation, Phys. Rev. B **103**, 144404 (2021).

[68] M. Pickem, E. Maggio, and J. M. Tomczak, The Linear Response Transport Centre (LinReTraCe) https://github.com/linretrace/linretrace.

[69] S. Poncé, E. R. Margine, C. Verdi, and F. Giustino, EPW: Electronphonon coupling, transport and superconducting properties using maximally localized Wannier functions, Comput. Phys. Commun. **209**, 116 (2016).

[70] J.-J. Zhou, J. Park, I.-Te Lu, I. Maliyov, X. Tong, and M. Bernardi, Perturbo: A software package for *ab initio* electron-phonon interactions, charge transport and ultrafast dynamics, Comput. Phys. Commun. **264**, 107970 (2021).

[71] X. Li, Z. Zhang, J. Xi, D. J. Singh, Y. Sheng, J. Yang, and W. Zhang, TransOpt. A code to solve electrical transport properties of semiconductors in constant electronphonon coupling approximation, Comput. Mater. Sci. **186**, 110074 (2021).

[72] N. H. Protik, C. Li, M. Pruneda, D. Broido, and P. Ordejón, The elphbolt *ab initio* solver for the coupled electron-phonon Boltzmann transport equations, arXiv:2109.08547.

[73] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, ShengBTE: A solver of the Boltzmann transport equation for phonons, Comput. Phys. Commun. **185**, 1747 (2014).

[74] K. Behnia and H. Aubin, Nernst effect in metals and superconductors: a review of concepts and experiments, Rep. Prog. Phys. **79**, 046502 (2016).

[75] K. Behnia, D. Jaccard, and J. Flouquet, On the thermoelectricity of correlated electrons in the zero-temperature limit, J. Phys.: Condens. Matter **16**, 5187 (2004).

[76] K. Behnia, The Nernst effect and the boundaries of the Fermi liquid picture, J. Phys.: Condens. Matter **21**, 113101 (2009).

[77] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1972).

[78] H. Goldsmid and J. Sharp, Estimation of the thermal band gap of a semiconductor from Seebeck measurements, J. Electron. Mater. **28**, 869 (1999).

[79] C. Fu, M. Yao, X. Chen, L. Z. Maulana, X. Li, J. Yang, K. Imasato, F. Zhu, G. Li, G. Auffermann, U. Burkhardt, W. Schnelle, J. Zhou, T. Zhu, X. Zhao, M. Shi, M. Dressel, A. V. Pronin, G. J. Snyder, and C. Felser, Revealing the intrinsic electronic structure of 3d half-Heusler thermoelectric materials by angle-resolved photoemission spectroscopy, Adv. Sci. **7**, 1902409 (2020).

[80] A. L. Kutepov and A. Ruth, Electronic structure and thermoelectric properties of CoAsSb with post-DFT approaches, Appl. Phys. A **126**, 137 (2020).

[81] F. Garmroudi, A. Riss, M. Parzer, N. Reumann, H. Müller, E. Bauer, S. Khmelevskyi, R. Podloucky, T. Mori, K. Tobita, Y. Katsura, and K. Kimura, Boosting the thermoelectric performance of $Fe_2VAl$-type Heusler compounds by band engineering, Phys. Rev. B **103**, 085202 (2021).

[82] T. Deng, J. Recatala-Gomez, M. Ohnishi, D. V. M. Repaka, P. Kumar, A. Suwardi, A. Abutaha, I. Nandhakumar, K. Biswas, M. B. Sullivan, G. Wu, J. Shiomi, S.-W. Yang, and K. Hippalgaonkar, Electronic transport descriptors for the rapid screening of thermoelectric materials, Mater. Horiz. **8**, 2463 (2021).

[83] Z. M. Gibbs, H.-S. Kim, H. Wang, and G. J. Snyder, Band gap estimation from temperature dependent Seebeck measurementdeviations from the 2e|Smax|Tmax relation, Appl. Phys. Lett. **106**, 022112 (2015).

[84] M. Cutler and N. F. Mott, Observation of Anderson localization in an electron gas, Phys. Rev. **181**, 1336 (1969).

[85] M. Jonson and G. D. Mahan, Mott's formula for the thermopower and the Wiedemann-Franz law, Phys. Rev. B **21**, 4223 (1980).

[86] Y. Wang, Z. A. Xu, T. Kakeshita, S. Uchida, S. Ono, Y. Ando, and N. P. Ong, Onset of the vortexlike Nernst signal above $T_c$ in $La_{2-x}Sr_xCuO_4$ and $Bi_2Sr_{2-y}La_yCuO_6$, Phys. Rev. B **64**, 224519 (2001).

[87] V. Oganesyan and I. Ussishkin, Nernst effect, quasiparticles, and *d*-density waves in cuprates, Phys. Rev. B **70**, 054503 (2004).

[88] E. H. Sondheimer and A. H. Wilson, The theory of the galvanomagnetic and thermomagnetic effects in metals, Proc. R. Soc. London A **193**, 484 (1948).

[89] W.-R. Lee, A. M. Finkel'stein, K. Michaeli, and G. Schwiete, Role of electron-electron collisions for charge and heat transport at intermediate temperatures, Phys. Rev. Res. **2**, 013148 (2020).

[90] A. Lavasani, D. Bulmash, and S. Das Sarma, Wiedemann-Franz law and Fermi liquids, Phys. Rev. B **99**, 085104 (2019).

[91] I. Opahle, G. K. H. Madsen, and R. Drautz, High throughput density functional investigations of the stability, electronic structure and thermoelectric properties of binary silicides, Phys. Chem. Chem. Phys. **14**, 16197 (2012).

[92] S. Wang, Z. Wang, W. Setyawan, N. Mingo, and S. Curtarolo, Assessing the Thermoelectric Properties of Sintered Compounds via High-Throughput *Ab-Initio* Calculations, Phys. Rev. X **1**, 021012 (2011).

[93] W. Chen, J.-H. Pohls, G. Hautier, D. Broberg, S. Bajaj, U. Aydemir, Z. M. Gibbs, H. Zhu, M. Asta, G. J. Snyder, B. Meredig, M. A. White, K. Persson, and A. Jain, Understanding thermoelectric properties from high-throughput calculations: trends, insights, and comparisons with experiment, J. Mater. Chem. C **4**, 4414 (2016).

[94] S. Bocelli, F. Marabelli, R. Spolenak, and E. Bauer, Evolution of the optical response from a very narrow gap semiconductor to a metallic material in $(Fe_xMn_{1-x})Si$, Mater. Res. Soc. Symp. Proc. **402**, 361 (1996).

[95] A. Perucchi, L. Degiorgi, R. Hu, C. Petrovic, and V. F. Mitrovic, Optical investigation of the metal-insulator transition in $FeSb_2$, Eur. Phys. J. B **54**, 175 (2006).

[96] C. C. Homes, Q. Du, C. Petrovic, W. H. Brito, S. Choi, and G. Kotliar, Unusual electronic and vibrational properties in the colossal thermopower material $FeSb_2$, Sci. Rep. **8**, 11692 (2018).

[97] T. Jarlborg, Temperature-dependent electronic structure: from heavy fermion behaviour to phase stability, Rep. Prog. Phys. **60**, 1305 (1997).

[98] O. Delaire, K. Marty, M. B. Stone, P. R. C. Kent, M. S. Lucas, D. L. Abernathy, D. Mandrus, and B. C. Sales, Phonon softening and metallization of a narrow-gap semiconductor by thermal disorder, Proc. Natl. Acad. Sci. USA **108**, 4725 (2011).

[99] R. Hu, V. F. Mitrović, and C. Petrovic, Anisotropy in the magnetic and transport properties of $Fe_{1-x}Co_xSb_2$, Phys. Rev. B **74**, 195130 (2006).

[100] T. Deguchi, K. Matsubayashi, Y. Uwatoko, T. Koyama, T. Kohara, H. Nakamura, Y. Mitsui, and K. Koyama, Magnetic measurements of narrow-gap semiconductor $FeSb_2$ under high pressure, Mater. Trans. **61**, 1476 (2020).

[101] K.-J. Xu, S.-D. Chen, Yu He, J. He, S. Tang, C. Jia, E. Yue Ma, S.-K. Mo, D. Lu, M. Hashimoto, T. P. Devereaux, and Z.-X. Shen, Metallic surface states in a correlated d-electron topological Kondo insulator candidate $FeSb_2$, Proc. Natl. Acad. Sci. USA **117**, 15409 (2020).

[102] C. Herring, Theory of the thermoelectric power of semiconductors, Phys. Rev. **96**, 1163 (1954).

[103] J. Zhou, B. Liao, B. Qiu, S. Huberman, K. Esfarjani, M. S. Dresselhaus, and G. Chen, Ab initio optimization of phonon drag effect for lower-temperature thermoelectric energy conversion, Proc. Natl. Acad. Sci. USA **112**, 14777 (2015).

[104] M. Kimura, X. He, T. Katase, T. Tadano, J. M. Tomczak, M. Minohara, R. Aso, H. Yoshida, K. Ide, S. Ueda, H. Hiramatsu, H. Kumigashira, H. Hosono, and T. Kamiya, Large phonon drag thermopower boosted by massive electrons and phonon leaking in LaAlO3/LaNiO3/LaAlO3 heterostructure, Nano Lett. **21**, 9240 (2021).

[105] A. Bentien, G. K. H. Madsen, S. Johnsen, and B. B. Iversen, Experimental and theoretical investigations of strongly correlated $FeSb_{2-x}Sn_x$, Phys. Rev. B **74**, 205105 (2006).

[106] A. V. Sanchela, A. D. Thakur, and C. V. Tomy, Enhancement in thermoelectric properties of $FeSb_2$ by Sb site deficiency, J. Materiomics **1**, 205 (2015).

[107] Q. Du, X. Tong, Yu Liu, and C. Petrovic, Suppression of thermal conductivity and electronic correlations in $Fe_{1-x}Ru_xSb_2$ ($0 \leqslant x \leqslant 0.6$), Appl. Phys. Lett. **118**, 171904 (2021).

[108] J. Mitscherling, Longitudinal and anomalous Hall conductivity of a general two-band model, Phys. Rev. B **102**, 165151 (2020).

[109] N. Wagner, S. Ciuchi, A. Toschi, B. Trauzettel, and G. Sangiovanni, Resistivity Exponents in 3D Dirac Semimetals from Electron-Electron Interaction, Phys. Rev. Lett. **126**, 206601 (2021).

[110] P. Sun, B. Wei, J. Zhang, J. M. Tomczak, A. M. Strydom, M. Søndergaard, B. B. Iversen, and F. Steglich, Large Seebeck effect by charge-mobility engineering, Nat. Commun. **6**, 7475 (2015).

[111] T. Katase, X. He, T. Tadano, J. M. Tomczak, T. Onozato, K. Ide, B. Feng, T. Tohei, H. Hiramatsu, H. Ohta, Y. Ikuhara, H. Hosono, and T. Kamiya, Breaking of thermopower-conductivity trade-off in $LaTiO_3$ film around Mott insulator to metal transition, Adv. Sci. **8**, 2102097 (2021).