

Thermopower of correlated semiconductors: Application to FeAs₂ and FeSb₂Jan M. Tomczak,¹ K. Haule,¹ T. Miyake,^{2,3} A. Georges,⁴ and G. Kotliar¹¹*Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854, USA*²*Nanosystem Research Institute, AIST, Tsukuba 305-8568, Japan*³*Japan Science and Technology Agency, CREST, Kawaguchi 332-0012, Japan*⁴*Centre de Physique Théorique, Ecole Polytechnique, CNRS, 91128 Palaiseau Cedex, France*

(Received 3 June 2010; revised manuscript received 16 July 2010; published 9 August 2010)

We investigate the effect of electronic correlations onto the thermoelectricity of semiconductors and insulators. Appealing to model considerations, we study various many-body renormalizations that enter the thermoelectric response. We find that, contrary to the case of correlated metals, correlation effects do not *per se* enhance the Seebeck coefficient or the figure of merit, for the former of which we give an upper bound in the limit of vanishing vertex corrections. For two materials of current interest, FeAs₂ and FeSb₂, we compute the electronic structure and thermopower. We find FeAs₂ to be well described within density-functional theory and the therefrom deduced Seebeck coefficient to be in quantitative agreement with experiment. The capturing of the insulating ground state of FeSb₂, however, requires the inclusion of many-body effects, in which we succeed by applying the *GW* approximation. Yet, while we get qualitative agreement for the thermopower of FeSb₂ at intermediate temperatures, the tremendously large Seebeck coefficient at low temperatures is found to violate our upper bound, suggesting the presence of decisive (e.g., phonon mediated) vertex corrections.

DOI: [10.1103/PhysRevB.82.085104](https://doi.org/10.1103/PhysRevB.82.085104)

PACS number(s): 72.20.Pa, 71.27.+a, 31.15.A-

I. INTRODUCTION

While recent efforts to design materials with enhanced thermoelectric properties were mainly focused on reducing the lattice contributions to the thermal conductivity by superstructures or nanostructures,¹ interest in the potential merits of electronic correlation effects was revived by the discovery of large Seebeck coefficients in transition-metal compounds, such as FeSi,² Na_xCoO₂,³ and FeSb₂,⁴ with the latter displaying an astonishing response of up to $S = -45$ mV/K at 12 K.⁴ Indeed, on a model level,⁵⁻⁸ as well as for realistic compounds,^{9,10} correlation effects were shown to enhance the Seebeck coefficient in metals and transition-metal oxides.^{11,12}

In this work we address the thermoelectric response of correlated semiconductors and insulators. Our aim is to set up the general formalism for discussing the thermoelectric response with the ultimate goal to understand the origin of the large thermoelectricity observed in correlated semiconductors and to search for high performance thermoelectrics in this class of materials. This work parallels the analysis done for correlated metals in Refs. 6 and 10.

In particular, we investigate whether the electronic structure and correlation effects alone can account for the very different magnitudes in the Seebeck coefficient of the two isostructural and isoelectronic compounds FeSb₂ and FeAs₂. In the case of FeSb₂, it has indeed been conjectured that electronic correlations are at the origin of the huge thermoelectric response.^{4,13-15}

The setup of the paper is the following. First we describe the FeX₂ materials. Then we shall first extent general textbook considerations for the Seebeck coefficient to include the important aspect of carrier selective electronic renormalizations, as well as properties beyond the picture of coherent band structures. We use this framework to make general arguments on how to obtain high values for the figure of merit

ZT . Given the sizes of the charge gaps, effective masses, and other parameters, these considerations allow us to put constraints on the possible regimes of materials of interest. In Sec. IV we apply realistic electronic-structure tools and show that for iron arsenide, FeAs₂, both the electronic structure, as well as the thermoelectric response can be understood as that of a conventional semiconductor and described quantitatively by *ab initio* band-structure methods. For the iron antimonide on the other hand, standard density-functional theory (DFT) (Ref. 16) based methods are known to be insufficient to account for the electronic structure.^{17,18} To show that correlation effects play an important role, we employ a hybrid functional approach,¹⁹ as well as Hedin's *GW* approximation,²⁰ with the latter yielding results in good agreement with the experimental charge gap. As to the thermoelectric response, however, we find qualitative agreement with experiment only at intermediate temperatures (35–70 K). Our analysis shows that the low-temperature Seebeck coefficient of FeSb₂ is incompatible with a local electronic picture, suggesting the importance of vertex corrections and nonlocal self-energy effects (that we neglect), or the presence of a substantial phonon-drag effect,²¹ as is, e.g., found in the classical example of *p*-type germanium.²²

II. MATERIALS

Despite the structural similarity of FeSb₂ and FeAs₂, experimental findings point to markedly different properties, heralding a varying importance of correlation, and, potentially, electron-phonon effects.

FeAs₂ is an insulator with a gap of 0.2–0.22 eV,^{14,23} as obtained from the activation behavior in the resistivity at temperatures of 200K and higher. Below 200 K, the influence of impurities is pivotal:^{14,23} the resistivity has a metallic slope before resuming, below 30 K and down to 10 K, an activation law with an energy of 0.01 eV. Further, below 10

K, the resistivity exhibits activation with 6 K (0.5 meV).¹⁴ The Hall coefficient is negative for all temperatures.²³ Congruently, the Seebeck coefficient is negative as well. From its room-temperature value $-200 \mu\text{V}/\text{K}$,²³ it grows in magnitude upon cooling, to reach $-7 \text{ mV}/\text{K}$ at 12 K,¹⁴ before it vanishes towards zero temperature.¹⁴

In the case of FeSb₂, optical spectroscopy finds a small gap of 432 K (37 meV) at low temperatures²⁴ but witnesses the development of a Drude-type peak at 70 K and above. The concomitant transfer of spectral weight is found to extend over an energy range as high as 1 eV,²⁴ i.e., a scale that is much larger than the initial gap, a common harbinger of correlation effects.²⁵ The resistivity of FeSb₂, on the other hand, has three distinct temperature regimes that exhibit activated behavior: in the range of 50–100 K the activation energy corresponds to a gap of 300 K (26 meV).^{4,26} From 20 K down to 10 K, the resistivity shows a shoulderlike behavior with an activation energy of $\Delta/2=3 \text{ meV}$ while below 5 K, extrinsic impurities are believed to be at the origin of a weakly temperature dependent resistivity following an activation behavior with 0.04–0.09 meV.⁴ The resistivity is anisotropic^{4,23,26} and some experiments find metallic transport behavior ($dp/dT > 0$) for selected directions above 40 K.^{26,27} This anisotropy is also seen in Hall measurements.²³ As a matter of fact, the Hall coefficient even changes sign for some polarizations [at 100 K (Ref. 23) or 40 K (Ref. 28)], with predominant electron character (n type) below these temperatures. The Seebeck coefficient at 300 K is found to be 15–40 $\mu\text{V}/\text{K}$, with the sign depending on the polarization.^{4,29} Upon lowering the temperature, the Seebeck coefficient passes a local maximum ($\partial^2 S/\partial T^2 > 0$) at around 40 K, before turning towards very large negative values, reaching, depending on the polarization and the sample, a global extrema of up to $-45 \text{ mV}/\text{K}$ at slightly above 10 K.⁴ Below this temperature, the coefficient drops sharply in magnitude and practically vanishes at 5 K and below. Interestingly, the largest thermopower is thus found in the temperature range where the resistivity has the shoulderlike behavior. Noteworthy, this regime is concomitant with the appearance of a prominent feature in the “electronic” specific heat.¹⁴ No such feature is found for the arsenide.³⁰ However, the unlocking of spins in FeSb₂ becomes appreciable only beyond this regime at around 150 K, where the entropy reaches $R \log 2$, owing to a second and larger hump in the specific heat and in congruity with the susceptibility.^{26,31} Indeed FeSb₂ becomes paramagnetic above 100 K,²⁶ and a Curie-type downturn appears at temperatures above 350 K (Ref. 31) whereas the susceptibility of FeAs₂ is flat up to 350 K.¹⁴ That the low-temperature feature in the specific heat of FeSb₂ has no spin signature might indicate that its contribution to the entropy is associated with either the charge degrees of freedom or an electron-phonon effect. The importance of electron-electron effects in FeSb₂ is further highlighted by the fact that various properties are very sensitive with respect to changes in the carrier density. Doping the system with electrons, e.g. FeSb_{2-x}Sn_x,²⁹ or holes, e.g., FeSb_{2-x}Te_x,^{32,33} instantly metalizes the compound, sometimes generates a Curie law at low temperatures³² and reduces the Seebeck coefficient.^{29,33} Thermoelectric properties of FeSb_{2-x}Te_x are indeed that of a

correlated metal, i.e., the low-temperature Seebeck coefficient is linear in T with an enhancement factor of 15 via the effective mass.¹⁵

Despite these indications for correlation effects, some experimental findings for the antimonide are quantitatively reproducible by conventional band-structure methods. Volume and bulk-modulus³⁴ are very well captured within the generalized gradient approximation (GGA) of DFT.³⁵ Also, the finding of small electron and hole pockets in FeSb₂ (Refs. 17 and 18) (see also below) within band-structure methods could simply be attributed to the well documented underestimation of charge gaps within DFT. Moreover, a calculation within the local-density approximation (LDA) with a Hartree-type Coulomb interaction (LDA+ U) suggested that FeSb₂ while being paramagnetic, could be close to a ferromagnetic instability.¹⁷ Weak ferromagnetism was then indeed found in Fe_{1-x}Co_xSb₂.³⁶

III. TRANSPORT FORMULAE, GENERAL, AND MODEL CONSIDERATIONS

Within the Kubo formalism the Seebeck coefficient—that relates the gradients of temperature and electrical field—is given by (see, e.g., Refs. 9 and 10)

$$S = -\frac{k_B A_1}{|e| A_0}, \quad (1)$$

where the current-current (current-heat current) correlation function A_0 and (A_1) is given by

$$A_n = \int d\omega \beta^n (\omega - \mu)^n \left(-\frac{\partial f_\mu}{\partial \omega} \right) \Xi(\omega). \quad (2)$$

Here, f_μ is the Fermi function, μ is the Fermi level, and Ξ is the transport kernel. If vertex corrections are neglected, the transport kernel can be expressed (in matrix notation) as

$$\Xi(\omega) = \sum_{\mathbf{k}} \text{Tr}[v(\mathbf{k})A(\mathbf{k}, \omega)v(\mathbf{k})A(\mathbf{k}, \omega)] \quad (3)$$

with the Fermi velocity $v_{ij}(\mathbf{k}) = -\frac{ie}{m} \langle \Psi_{\mathbf{k}i} | \nabla | \psi_{\mathbf{k}j} \rangle$ and the spectral function $A_{ij}(\mathbf{k}, \omega)$, where $\psi_{\mathbf{k}i}$ is a complete set of one-electron basis functions, such as Kohn-Sham (KS) orbitals. Using transport coefficients Eq. (2), we can further express (see, e.g., Ref. 6) the dc conductivity, the thermal conductivity and the figure of merit as

$$\sigma = \frac{2\pi e^2}{\hbar V} A_0, \quad (4)$$

$$\kappa = \kappa_L + \frac{2\pi k_B^2}{\hbar V} T \left(A_2 - \frac{A_1^2}{A_0} \right), \quad (5)$$

$$ZT = \frac{S^2 \sigma T}{\kappa}, \quad (6)$$

where V is the unit-cell volume and κ_L the thermal lattice conductivity.

The chemical potential, μ , is obtained by the requirement of the charge neutrality,

$$n - p - n_{D^+} = 0, \quad (7)$$

where

$$\left. \begin{array}{l} n \\ p \end{array} \right\} = \sum_{\mathbf{k}} \int_{-\infty}^{\infty} d\omega \begin{cases} A^v(\mathbf{k}, \omega) f_{\mu}(\omega) \\ A^c(\mathbf{k}, \omega) [1 - f_{\mu}(\omega)] \end{cases} \quad (8)$$

is the number of electrons (holes). Here, $A^{c,v}(k, \omega)$ is the valence/conduction spectral function. We also allowed for the presence of ionized donor impurities $n_{D^+} = n_D [1 + 2e^{-\beta(E_D - \mu)}]^{-1}$ of concentration n_D at an energy E_D .

Model considerations

Here we find it instructive to extend on the usual textbook considerations (see, e.g., Ref. 37) and generalize to include carrier dependent masses, renormalizations, as well as finite (yet energy independent) scattering amplitudes. This makes it possible to investigate the important effects of particle-hole asymmetry, carrier coherence, and allows us later to discuss the consistency of a purely diffusive thermopower for a given material.

Assuming a Lorentzian line shape of the conduction (c) and valence (v) spectral functions

$$A_{coh}^{c,v}(k, \omega) = \frac{Z_k^{c,v}}{\pi} \frac{\Gamma_k^{c,v}}{(\omega - \xi_k^{c,v})^2 + (\Gamma_k^{c,v})^2}, \quad (9)$$

i.e., we limit the discussion to the coherent part, A_{coh} , (of weight Z_k) of the full spectrum, $A = A_{coh} + A_{incoh}$, with a quasiparticle dispersion ξ_k , and an elastic scattering of amplitude Γ_k . Within this approximation, one finds for the number of electrons³⁸

$$n = \sum_k \left\{ \frac{1}{2} - \frac{1}{\pi} \text{Im} \psi \left[\frac{1}{2} + \frac{\beta}{2\pi} [\Gamma_k + i(\xi_k - \mu)] \right] \right\}, \quad (10)$$

where $\psi(z)$ is the digamma function and $\beta = 1/(k_B T)$ the inverse temperature. The above expression reduces to the usual $\sum_k f(\mu - \xi_k)$ in the coherent limit ($\Gamma = 0$). The response functions, Eq. (2), can be also be expressed analytically (here, we restrict ourselves to $n=0, 1$) as

$$A_n = \frac{\beta^{n+1}}{4\pi^3} \sum_k \frac{Z_k^2}{\Gamma_k} v_k^2 \left\{ (\xi_k - \mu)^n \text{Re} \psi'(z) - \frac{\beta}{2\pi} [(\xi_k - \mu)^n \Gamma_k \text{Re} \psi''(z) + n \Gamma_k^2 \text{Im} \psi''(z)] \right\}, \quad (11)$$

where the arguments of the derivatives of the digamma function are

$$z = \left\{ \frac{1}{2} + \frac{\beta}{2\pi} [\Gamma_k - i(\xi_k - \mu)] \right\}.$$

The first contribution in the curly brackets of Eq. (11) is the leading term in the coherent limit ($\Gamma \rightarrow 0$). Indeed with $\beta/(2\pi^2) \text{Re} \psi'[1/2 + i\beta x/(2\pi)] = -f'(x)$, one recovers the Boltzmann expressions for transport coefficients,³⁷ in which

appears an *ad hoc* lifetime $\tau = 1/(2\Gamma)$. In the above expression, however, the influence of *finite* scatterings is not restricted to the prefactor but since the spectrum broadens, a wider energy range becomes activated for supplying charge carriers and, as a result, the Fermi statistics assumes the digamma form, and, also, higher order terms appear. Therewith, contrary to the Boltzmann description at small Γ , coherence effects do not in general cancel in the ratio A_1/A_0 of the Seebeck coefficient.

For illustrative purposes, we now analyze Eq. (11) in terms of special cases for simple quadratic dispersions: We consider bare bands $\epsilon_{c,v}^k = \frac{\hbar^2 k^2}{2m_{c,v}^0}$ and an interacting dispersion

$$\xi_{c,v}^k = \pm \Delta/2 \pm \frac{\hbar^2 k^2}{2m_{c,v}^*}, \quad (12)$$

where Δ is the charge gap, m^* the effective mass of the carriers, and the origin of the chemical potential is chosen at the midgap point. Further, we assume the Fermi velocities to be given by the group velocity³⁹

$$v_k = \frac{1}{\hbar} \partial_k \epsilon_k \quad (13)$$

and weights, Z , and scattering rates, Γ , that are independent of momentum.

Large gap coherent semiconductor

In the limit of a coherent system ($\Gamma \ll 1$) with a large gap ($\beta|\Delta/2 - \mu| \gg 1$), Eq. (11) can be simplified to

$$A_n^{c/v} = (\pm 1)^n \frac{3\lambda}{\sqrt{2\pi^5} \beta^3} e^{-\beta\Delta/2} \left\{ [\beta(\pm\mu - \Delta/2)]^n - \frac{5}{2} n \right\}, \quad (14)$$

where all carrier specific parameters have been gathered in

$$\lambda_{c,v} = \frac{Z^2 m_*^{5/2}}{\Gamma m_0^2} e^{\pm\beta\mu}. \quad (15)$$

Therewith the Seebeck coefficient becomes

$$S = - \frac{k_B A_1^c + A_1^v}{|e| A_0^c + A_0^v}, \quad (16)$$

$$= \frac{1}{|e|T} \left(\mu - \frac{\Delta}{2} \delta\lambda \right) - \frac{5}{2} \frac{k_B}{|e|} \delta\lambda, \quad (17)$$

where the asymmetry parameter $\delta\lambda$ (that depends on μ and T) is given by

$$\delta\lambda = \frac{\lambda^c - \lambda^v}{\lambda^c + \lambda^v}. \quad (18)$$

Hence, a large Seebeck coefficient can be achieved by an interplay of the gap, Δ , the anisotropy or asymmetry $\delta\lambda$ in the transport function—stemming from either the densities of states ($m_{c,v}^0$), different bandwidth narrowings ($m_{c,v}^*$), scattering amplitudes ($\Gamma_{c,v}$), or quasiparticle weights ($Z_{c,v}$).

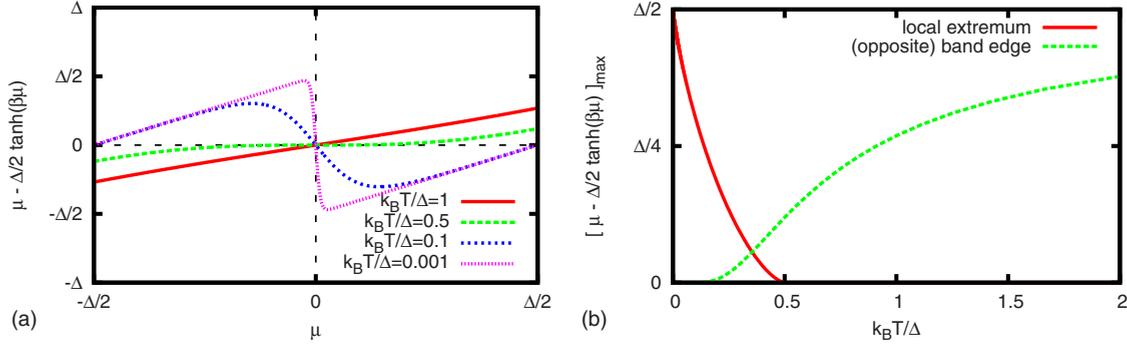


FIG. 1. (Color online) The symmetric, coherent, and large gap semiconductor. (a) coefficient of the $1/T$ low-temperature behavior of the Seebeck coefficient as a function of the chemical potential and for different temperatures. (b) temperature evolution of the extremal value of the $1/T$ coefficient of the thermopower.

Upper limit for thermopower in a semiconductor

The position of the chemical potential plays an important role in maximizing the thermopower of a semiconductor. Let us take the example of two equivalent bands. In this case, the term in brackets in Eq. (17) becomes

$$\mu - \Delta/2 \tanh(\beta\mu). \quad (19)$$

This coefficient of the $1/T$ behavior is displayed in Fig. 1(a) as a function of the chemical potential for a few different temperatures. At high temperatures the optimal chemical potential, which maximizes thermopower, is near the gap edges. At low temperature $k_B T < \Delta/2$, the $1/T$ coefficient shows a local extremum. The Seebeck coefficient vanishes at the point of particle-hole symmetry ($\mu=0$ in the symmetric case considered here⁴⁰) and the optimal location of the chemical potential at low temperature is in the direct vicinity of that point, hence very close to the center of the gap.

Fig. 1(b) displays the value of the $1/T$ coefficient in this extremum and the value at the gap edge as a function of temperature. For $k_B T \gtrsim 0.3\Delta$ the maximum value of the thermopower is achieved when the chemical potential is at the gap edge, and for lower temperature, it is achieved close to the middle of the gap, where the thermopower can reach the maximum value of $S = \Delta/(2eT)$.

In an asymmetric case, the thermopower can be larger than this maximum value, however, for a given charge gap Δ , there is always an upper bound for the Seebeck coefficient, namely,

$$|S(T)e| \leq \Delta/T + 5/2k_B. \quad (20)$$

This is because the asymmetry is bounded to an absolute value of one $|\delta\lambda| \leq 1$. This extremal value corresponds to the fictitious system in which only one type of the two carriers contributes to the thermoelectricity, e.g., the conduction electrons and with the chemical potential being, in that case, at the edge of the valence band.

Thus the correlation effects, such as small Z in the conduction band and large Z in the valence band, can enhance the thermopower of a semiconductor. However, this effect is limited by the form of Eq. (17) allowing maximum S bounded by Eq. (20). The possible merits of electron-hole asymmetry for the case of metals is discussed in Ref. 10.

Model semiconductor in the presence of donor impurities

With the goal of understanding the thermopower and the figure of merit in a renormalized semiconductor in a very general setup, including the presence of impurities, we now numerically study the model based on the response functions Eq. (11). As before, we assume parabolic dispersions Eq. (12), with the band structure depicted in the inset of Fig. 2(b): excitations of different effective masses are separated in energy by a gap Δ and we allow for the presence of donor impurities, situated at an energy E_D , as measured from the middle of the gap. We again assume transition matrix elements to be given by the group velocity, Eq. (13).

We choose the parameters compatible with the band structure of FeAs₂: we consider a gap $\Delta=0.2$ eV and, unless stated otherwise, an impurity level at $E_D=95$ meV, as inferred from the low-temperature activation behavior of the resistivity.¹⁴ In our current treatment, we assume that the impurity carriers have vanishing Fermi velocities and thus their only effect is to shift the chemical potential.

To fix the particle-hole asymmetry, we note that at high temperatures, the number of ionized impurities are irrelevant with respect to the number of conduction and valence carriers, and the chemical potential follows the intrinsic behavior. In the coherent limit of the large gap semiconductor (see above) one finds that

$$\mu = 3k_B T/4 \ln(\eta_v/\eta_c), \quad \text{with } \eta = m^*/m_0. \quad (21)$$

In this regime the resistivity shows an activation law with the activation gap $\Delta/2$. Of course, the situation in a real material can be much more complicated (several types of impurities, temperature dependence of the gap, etc.). Using our *ab initio* data for FeAs₂ (that is presented below in Secs. IV and V A), and assuming the approximate validity of Eq. (21) for non-parabolic dispersions, one finds the ratio of the valence and conduction effective mass $\eta_v/\eta_c=2.5$ for FeAs₂, which we will use for all the following model calculations.

Further, we note that a uniform weight-factor Z cancels in the Seebeck coefficient whereas in the figure of merit it can be seen as a scaling factor of the thermal lattice conductivity, for which we assume $\kappa_L/Z^2=250$ W(Km)⁻¹.⁴¹ Moreover, we use a unit-cell volume of 80 Å³.

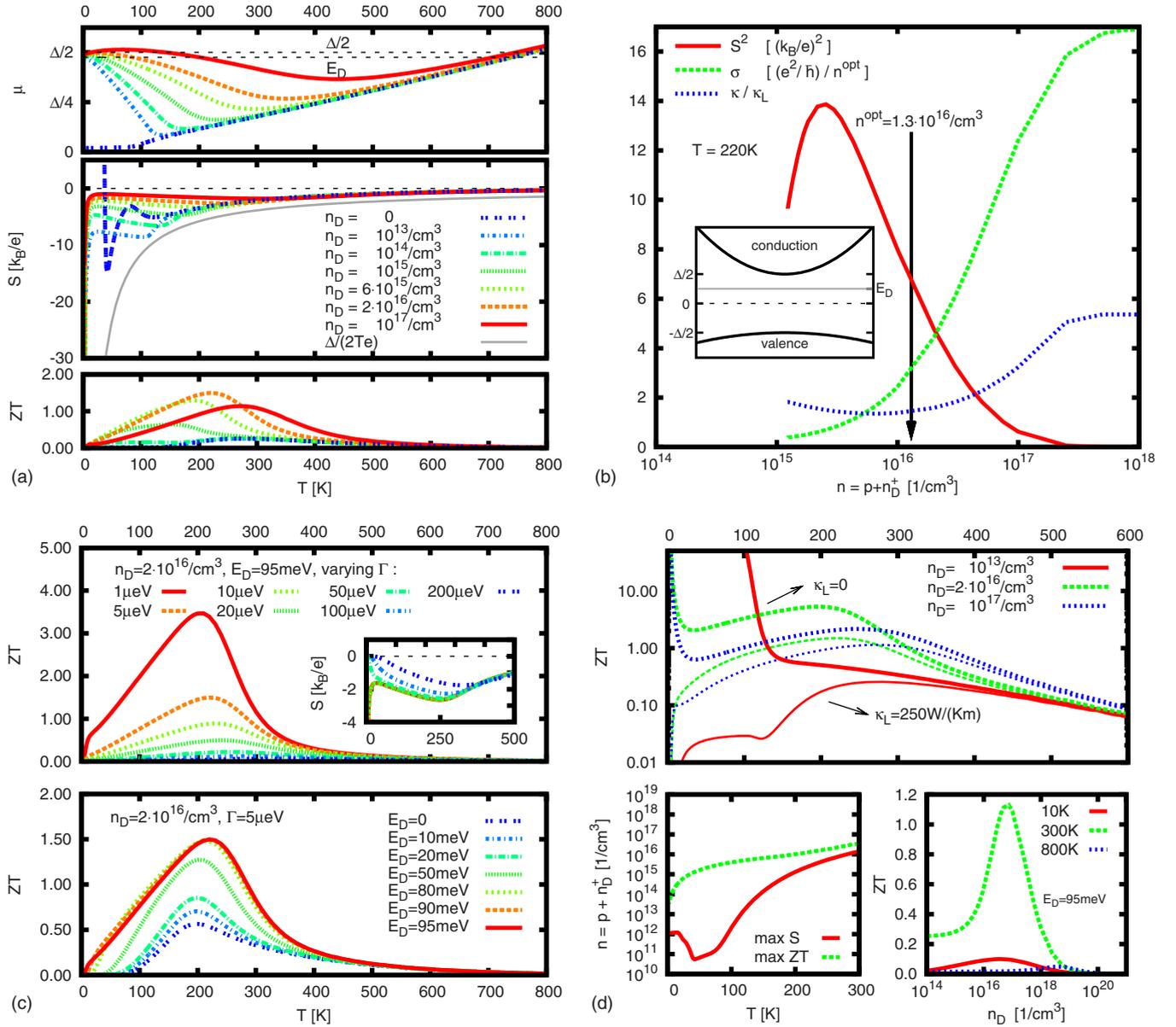


FIG. 2. (Color online) Thermoelectric properties of the model semiconductor with parameters inspired from FeAs₂. (a) from top to bottom: chemical potential, Seebeck coefficient, and figure of merit ZT as a function of temperature for different impurity concentrations. (b) ingredients to the figure of merit $ZT = S^2 \sigma T / \kappa$ at $T = 220$ K as a function of carrier density. The inset shows the general setup of asymmetric valence and conduction dispersions, with an impurity level E_D in the gap Δ . (c) effect of the scattering amplitude Γ (top) and the impurity level position E_D (bottom) onto the figure of merit ZT , and the thermopower (inset). (d) top: comparison of the figure of merit from (a) with the purely electronic figure of merit ($\kappa_L = 0$), bottom left: carrier density that maximizes $|S|$ and ZT (for the chemical potential in the upper half of the gap), bottom right: ZT as a function of the donor concentration for several temperatures, and a fixed impurity level $E_D = 95$ meV. We assume a unit-cell volume $V = 80 \text{ \AA}^3$ and a thermal lattice conductivity (scaled with the quasiparticle weight Z , see text) of $\kappa_L / Z^2 = 250 \text{ W}/(\text{Km})$.

Having thus fixed the size of the gap and the asymmetry, the principle parameters to vary in this setup are the concentration of impurities n_D and the scattering rate Γ (that we assume to be orbital independent). We will also study the dependence on the position of the impurity band E_D away from the value motivated by the experimental resistivity.

(a) *Impurity concentration.* With these parameters, we display in Fig. 2(a) the temperature dependence of the chemical potential, the Seebeck coefficient, and the figure of

merit for various impurity concentrations n_D , and for a constant scattering rate $\Gamma = 5 \text{ \mu eV}$.

In the intrinsic case, $n_D = 0$, the chemical potential is indeed linear above a certain temperature that is related to the scattering rate Γ . Below this regime, the chemical potential is almost temperature independent. Since Γ is small, the point of charge neutrality at zero temperature is very close to the midgap point (where it is in the coherent case). The point of particle-hole compensation for the thermopower, however, is

higher in energy than the midgap point, as rationalized, above, for the symmetric semiconductor. As a consequence, the chemical potential is actually moving through the symmetry point of the Seebeck coefficient and we expect it to change sign as a function of temperature. This is clearly seen in Fig. 2(a).

At finite doping and at low temperatures, the donor level plays the role of the valence band, and the chemical potential (top panel) is between E_D and the conduction band. In a coherent semiconductor the chemical potential is $\mu \approx (\Delta/2 - \delta/2) - k_B T \ln(n/n_D^+)$, where $\delta = \Delta/2 - E_D$ is the impurity activation energy, seen in the resistivity of this regime ($\rho \sim e^{-\beta\delta/2}$). At very low temperatures, μ is pinned to E_D . At intermediate temperatures the chemical potential smoothly connects with the intrinsic high-temperature slope, as shown in Fig. 2(a). For large enough concentration of impurities ($n_D > 10^{16}/\text{cm}^3$), the chemical potential can even go inside the conduction band at some intermediate temperature, which can result in a shoulder, or even a metallic slope in the resistivity [cf. the mentioned transport measurements on FeSb_2 and FeAs_2 (Ref. 14)].

For finite impurity concentration ($n_D > 0$), the Seebeck coefficient displayed in Fig. 2(a) may or may not be enhanced at a given temperature, depending on whether or not the additional carriers bring the chemical potential closer to its optimal value. As explained above for the case of two-particle-hole symmetric bands, at fixed temperature there exists a value of the chemical potential, which maximizes the Seebeck coefficient. In our asymmetric setup with $\eta_v/\eta_c > 1$, the optimum chemical potential is located above the midgap point. In the limit of vanishing impurity concentration, the midgap remains the point of charge neutrality at zero temperature while the optimal chemical potential is very near that point, hence the Seebeck coefficient is a very strong function of temperature in this limit, and can even change sign, as seen in the middle panel of Fig. 2(a). We note, however, that the fundamental extremum, established by the size of the gap, is always respected as is evident in Fig. 2(a).

The efficiency of the thermoelectrical material is determined by its figure of merit ZT , which we also plot in the lower panel of Fig. 2(a). ZT can be greatly enhanced by the presence of impurities and its maximum is not necessarily in close vicinity of the thermopower maximum. Indeed the largest ZT for the current parameters is achieved at about $T=220$ K for a concentration $n_D=2 \times 10^{16}/\text{cm}^3$. The Seebeck coefficient at this point is actually smaller than in the intrinsic limit. The position of the impurity with respect to the gap edge has a large effect on the optimal impurity density. While for our specific choice of parameters, $\Delta/2 - E_D = 5$ meV, the optimal density is $n_D = 2 \times 10^{16}/\text{cm}^3$, we notice that for larger gaps and/or larger separation of the impurity level from the gap edge, the optimal impurity density can reach a value as large as $\sim 10^{20}/\text{cm}^3$.⁴² To elucidate the origin of an optimal density further, we show in Fig. 2(b) the dependence of the quantities entering the expression of ZT as a function of the particle density ($n = p + n_D^+$) at fixed $T=220$ K, the temperature which maximizes ZT in Fig. 2(a). We use here the total particle density because in this case the description becomes independent of the bare impurity concentration and the level position and, in particular, the addi-

tional carriers can find their origin from multiple impurity sources.

For the given gap, and $T=220$ K no smaller densities than $10^{15}/\text{cm}^3$ can be accessed. The optimum n_D found in Fig. 2(a) translates into $n = 1.3 \times 10^{16}/\text{cm}^3$, mainly as a trade-off between decreasing S^2 (less entropy per carrier) and increasing σ (larger conductivity with more carriers). In particular, we note that S achieves its maximum for smaller concentration of carriers than the figure of merit.

The thermal conductivity in this range varies very slowly with concentration, which is not surprising since we fixed the lattice contribution to thermal conductivity to a fixed value of $\kappa_L/Z^2 = 250 \text{ W(Km)}^{-1}$.

Having chosen parameters to represent FeAs_2 , we note that the experimentally measured carrier concentration in this compound, as inferred from Hall measurements, is $5 \times 10^{17}/\text{cm}^3$ in the range of 60–170 K,²³ which is higher than the density that optimizes ZT in our model. Thus, it seems conceivable that by a deliberate change in the impurity concentration or position, an increase in the figure of merit of the specimen can be achieved.

(b) *Lifetime effects.* Next we pick the impurity concentration which maximized ZT , and we investigate the role of the scattering rate for the figure of merit and the Seebeck coefficient in Fig. 2(c). The lifetime has two effects: through the change in chemical potential and directly through the dependence of the response functions Eq. (11) on scattering rate Γ . It is this latter effect that causes the Seebeck coefficient to vanish at low temperatures for a sufficiently large scattering amplitude, as can be seen in the inset of the top panel of Fig. 2(c). The increase in scattering rate reduces the absolute value of both the Seebeck coefficient and the figure of merit, hence long lifetimes are preferred in thermoelectric materials.

The upper limit of the Seebeck coefficient has been discussed above. As a function of the scattering rate Γ , figure of merit ZT is limited as well. If lifetimes are long, the dependence of the response functions thereof is linear and thus cancels in the dimensionless ratio ZT if there are no lattice contributions to the thermal conductivity. Therefore, with decreasing Γ , the figure of merit converges towards the purely electronic limit in which $\kappa_L = 0$.

(c) *Position of the impurity level.* In Fig. 2(c) we show the dependence of the figure of merit on the position of the impurity level E_D . We fix the impurity concentration to $n_D = 2 \times 10^{16}/\text{cm}^3$ and the scattering rate to $\Gamma = 5 \text{ } \mu\text{eV}$. It is clear from Fig. 2(c) that a maximum ZT is achieved when the donor level is very close to the conduction band, which is located at $\Delta/2 = 100 \text{ meV}$.

This can again be understood from the optimal number of carriers: Indeed, in order to reach the ideal electron density of $1.3 \times 10^{16}/\text{cm}^3$, the chemical potential must be rather close to the conduction band. Since at $T=200$ K and a gap of $\Delta=0.2$ eV, the compensating holes cannot come from the valence electrons, they have to be supplied from the impurity band, hence, for an impurity concentration $n_D = 2 \times 10^{16}/\text{cm}^3$ that is larger than the needed number (if completely ionized), also E_D must be very close to the conduction band.

Next we study the sensitivity of the figure of merit to the lattice thermal conductivity. With thick line in Fig. 2(d) we

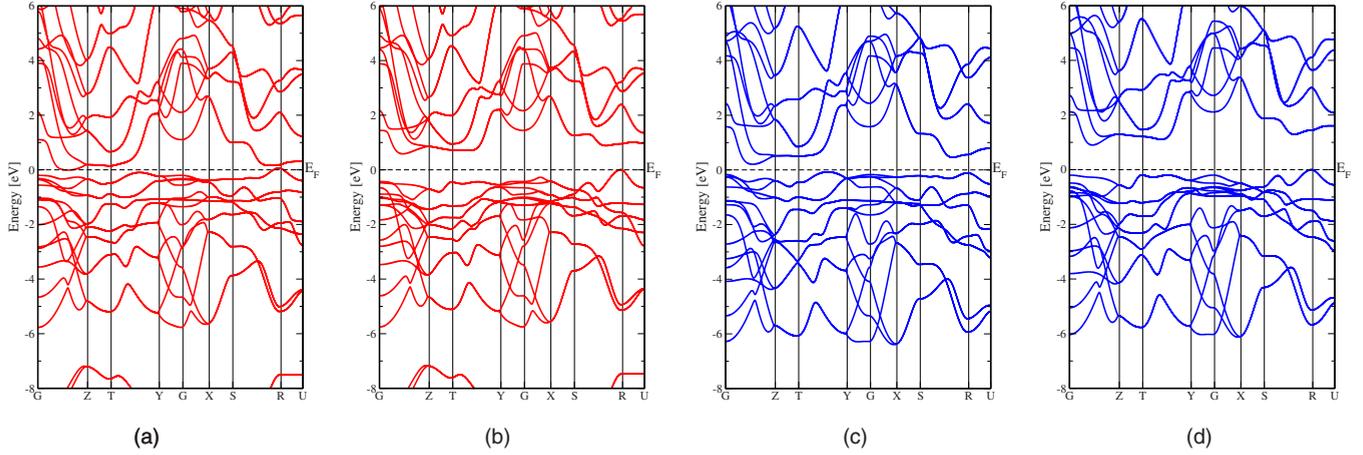


FIG. 3. (Color online) Band structures of FeSb₂ (a) GGA, (b) Hybrid functional, and FeAs₂ (c) GGA, and (d) Hybrid functional.

show the ZT in the absence of lattice thermal conductivity κ_L and with a thin line is shown ZT for a constant value of $\kappa_L/Z^2=250$ W/(Km). The figure of merit is clearly enhanced when the lattice conductivity is reduced, hence the desire for a “phonon glass” (see, e.g., the review¹), i.e., a solid which has a low phonon mean-free path such as to prevent substantial heat conduction by lattice vibration modes. The effect of lattice thermal conductivity is most enhanced at low temperature and for small impurity concentration.

In Fig. 2(d) lower panel we plot the optimal carrier concentration, which maximizes S or ZT , as a function of temperature. The ZT curve is monotonically increasing function of temperature, hence for best performance at higher temperature we need more impurity carriers. To maximize ZT we need larger impurity concentration than we need to maximize S . Finally, we also display the value of ZT as a function of impurity concentration for a few representative temperatures. For our setup of parameters, the figure of merit is very sharply peaked at room temperature around a carrier concentration ($n_D \approx 10^{17}/\text{cm}^3$).

In conclusion, these model considerations give guidance as to where to look for promising thermoelectric materials. In particular, we showed that unless vertex corrections or strongly frequency dependent lifetimes are of pivotal importance—electronic correlation effects are not in the position to enhance the thermopower of a gapped system, they can only shift the asymmetry of the contributions for electrons and holes. Indeed we found the thermopower (of purely electronic origin) to have an upper bound that is given by the size of the gap.

IV. ELECTRONIC STRUCTURE OF FeSb₂ AND FeAs₂

FeSb₂ and FeAs₂ both crystallize in the (regular) marcasite structure, have the orthorhombic space group $Pnmm$ and there are two formula units per unit cell (see Ref. 43). The iron ions are surrounded by distorted pnictogen octahedra that share corners along the c axis (see, e.g., Ref. 34). In the ligand-field picture, the compounds have a Fe $3d^4$ configuration and the $3d$ orbitals are split into e_g and lower lying t_{2g}

orbitals. The inequivalence of Fe-pnictogen distances causes the t_{2g} to split further into two degenerate lower and one higher orbital. In this picture, the compounds are in an insulating low spin state with the two degenerate t_{2g} orbitals filled.⁴³ From this perspective, metalization of FeSb₂ is driven by a temperature induced population of the third t_{2g} orbital.^{34,43} Previous band-structure calculations, however, suggest a more covalent picture,¹⁸ in the sense that stabilization occurs for d orbitals that point towards the ligands, i.e., in particular, lowering the e_g orbitals. In FeSb₂ this happens to the extent that LDA calculation actually yield a metallic ground state.^{17,18,29,44}

From the perspective of electronic-structure methods, at zero temperature, the challenge is hence to obtain an insulating ground state for FeSb₂. We therefore compare how three different approaches, GGA, Hybrid functionals and the GW approximation perform in this problem. In preparation for future work which should include correlations beyond GW and to clarify what would be the starting Hamiltonian to describe these materials, we obtain transfer-matrix elements and estimate the values of the interaction using the constrained random-phase-approximation (cRPA) method.

(d) *Band structures.* Our results for the band structure of FeSb₂ given by the GGA (Ref. 45) of DFT as implemented in the WIEN2K package⁴⁶ are displayed in Fig. 3(a). We used the atomic positions at room temperature ($a=5.83$ Å, $b=6.54$ Å, and $c=3.20$ Å).^{47,48} Congruent with previous works,^{17,18,29} the GGA ground state is metallic with small electron pockets halfway between the Γ and Z symmetry points, and corresponding hole pockets at all corners, R , of the orthorhombic Brillouin zone. Crucial to the understanding of the gap mechanism within more sophisticated techniques (see the GW discussion below) is to note the different orbital characters of the pockets.^{17,18} To quantify this, we transform the local coordinate system of the d orbitals into a basis, in which the local projection of the d block of the (GGA) Hamiltonian is as diagonal as possible. In this coordinate system, the x and z axes point (almost) towards the antimonide atoms,⁴⁹ and the e_g orbitals exhibit the expected bonding/antibonding splitting. In this basis, the electron pocket is mainly of d_{xy} character, and the hole pocket is formed by the now degenerate d_{xz} and d_{yz} (t_{2g}) orbitals. The

TABLE I. Occupations of the d orbitals (in the muffin spheres) within the transformed local coordinate system (see text for details). The t_{2g} orbitals d_{xz} and d_{yz} are degenerate and mainly account for the hole pocket. The electron pocket is of d_{xy} character.

FeSb ₂	d_{z^2}	$d_{x^2-y^2}$	d_{xz}	d_{yz}	d_{xy}
n	1.27	1.42	1.42	1.42	1.06

respective occupations (within the muffin spheres) are shown in Table I.

The GGA band structure of FeAs₂ [we use $a=5.3$ Å, $b=5.98$ Å, and $c=2.88$ Å (Ref. 23)] is shown in Fig. 3(c). With respect to FeSb₂, the chemical pressure of the larger As atoms is almost isotropic, and the c/a ratio remains virtually constant (as a function of external pressure, the ratio slightly decreases³⁴). In consequence, and as is apparent from the graph, the bands of FeAs₂ are much akin to those of FeSb₂ and could have roughly been obtained from a rigid-band shift.⁵⁰ While within GGA the gap of FeSb₂ was underestimated (no gap at all), the value 0.28 eV for FeAs₂ is just slightly too large with respect to the experimental 0.2–0.22 eV.^{14,23} On a qualitative level, one could thus say that a DFT calculation seems to work rather well for FeAs₂. Below we will explain why we believe this to be a mere coincidence.

(e) *Maximally localized Wannier function and cRPA.* From a conceptual point of view, we find it insightful to note, and compare, the hierarchy of transfer-matrix elements and the magnitude of local Hubbard interactions. Starting from a full-potential (FP) linear muffin-tin orbital (LMTO) LDA computation,⁵¹ we construct maximally localized Wannier functions for the subsystem consisting of the Fe 3d and Sb 5p orbitals, as described, e.g., in Ref. 52–54, and find that the (largest) nearest-neighbor hopping amplitudes in FeSb₂ are $t_{dd} \sim 0.2\text{--}0.3$ eV, $t_{pd} \sim 0.95$ eV, and $t_{pp} \sim 0.7$ eV. The significance of both the Sb 5p dispersion and the large hybridization was heralded already in the band structure, Fig. 3(a) and Refs. 17 and 18, as well as in the strongly mixed orbitals characters,^{17,18} see also the recent work Ref. 55. The two transfers t_{pd} , t_{pp} being of comparable magnitude, the system thus lies between the canonical Anderson model ($t_{pp} \gg t_{pd}$) and cupratelike compounds ($t_{pp} \ll t_{pd}$). The centers of gravity of the d and p bands are separated by $\Delta_{ct}=4.2$ eV. In FeAs₂ the pd transfers are larger, as expected from the chemical pressure and the covalent/hybridizationlike character of the gap. We find $t_{dd} \sim 0.25\text{--}0.3$ eV, $t_{pd} \sim 1.1$ eV, and $t_{pp} \sim 0.8$ eV, and $\Delta_{ct}=5.0$ eV.

To investigate the strength of local interactions, we compute the Hubbard U using the cRPA (Refs. 54 and 56) technique. Since in FeSb₂ and FeAs₂ the eigenvalues of the pd subspace are entangled with higher energy bands, we employ the scheme presented in Ref. 57.

In relation to the rather small energetic shift needed to deplete the (GGA) pockets in FeSb₂, we find, see Table II, that the orbital dependence of the Hubbard U within the d orbitals is of notable 5%—an effect hitherto mostly neglected in methods for correlated materials that start from a

TABLE II. cRPA values for the Hubbard U (in eV) of FeSb₂ for the Fe 3d orbitals in the pd setup of maximally localized Wannier functions in the local coordinate system.

FeSb ₂	d_{z^2}	$d_{x^2-y^2}$	d_{xz}	d_{yz}	d_{xy}
d_{z^2}	8.5	7.0	7.3	7.3	7.0
$d_{x^2-y^2}$	7.0	8.8	7.2	7.2	7.5
d_{xz}	7.3	7.2	8.6	7.0	7.3
d_{yz}	7.3	7.2	7.0	8.6	7.3
d_{xy}	7.0	7.5	7.3	7.3	8.8

parametrized Hubbard-type Hamiltonian. Before, we alluded to the different orbital characters of the pockets. Here, we indeed find for the d_{xz} and d_{yz} orbital that mainly account for the hole pocket a value of $U=8.6$ eV while for the orbital corresponding to the character of the electron pocket the interaction is slightly larger with $U_{d_{xy}}=8.8$ eV. This differentiation, in principle, favors a charge transfer towards a gap opening. Further matrix elements of the interaction are $U_{pd} \sim 2.5$ eV and $U_{pp} \sim 4$ eV. The hierarchy of the interaction strength within the d shell is already seen in the bare, i.e., unscreened, Coulomb interaction and is thus linked with the construction of the localized orbitals. On general grounds, larger matrix elements of the unscreened interaction are indeed expected for orbitals that have stronger hybridizations with other orbitals, thus are more spatially delocalized.⁵⁸ In analogy with the pressure dependence of Coulombic interactions in a localized basis,^{58,59} the bare interaction of FeAs₂ is larger than that of FeSb₂, in the cubic reference frame we find, e.g., $V_{d_{z^2}}=22.7$ eV and $V_{d_{x^2-y^2}}=22.6$ eV (the corresponding values for FeSb₂ are $V_{d_{z^2}}=22.0$ eV, $V_{d_{x^2-y^2}}=21.5$ eV). Moreover, with respect to the antimonide, the Kohn-Sham eigenvalues of the arsenide move towards higher energies, therewith reducing screening strengths and causing significantly larger values also for the Hubbard U : $U_{d_{z^2}}=11.0$ eV and $U_{d_{x^2-y^2}}=10.7$ eV.

(f) *Hybrid functional approach.* Previous attempts to produce an insulating band structure for FeSb₂ were made within the LDA+ U scheme,⁶⁰ where the paramagnetic state (LDA) was found to be stable below a critical $U=2.6$ eV with respect to a ferromagnetically ordered phase (LDA+ U).¹⁷ Here, we use a hybrid functional approach¹⁹ and Fig. 3(b) displays the resulting band structure, where the B3PW91 functional was used for the d orbitals of the iron atoms.⁶¹ We also note that while given the freedom, the system does not develop any magnetic moment within this setup (in LDA+ U it necessarily does). The band-structure features an indirect gap of about 0.6 eV, i.e., it is by far larger than in experiment. This points towards, both, a serious underestimation of static correlations within the previously used GGA, and the lacking of dynamical effects in the hybrid functional approach that will work to reduce the size of the gap.

The hybrid functional band structure of FeAs₂ is displayed in Fig. 3(d). As for the antimonide, the admixing of exact exchange shifts valence and conduction bands further apart, and the resulting gap of ~ 0.9 eV is much too large with respect to experiment.

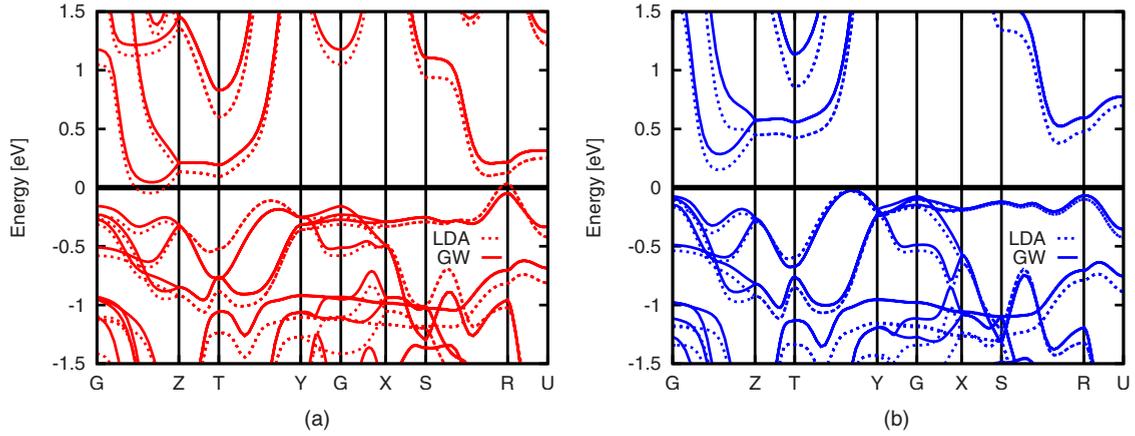


FIG. 4. (Color online) (a) FeSb_2 , (b) FeAs_2 . Band structure in the GW approximation (full lines), in comparison with LDA (dashed).

(g) *GW approximation.* To investigate the dynamical effects of electronic correlations in our compounds, we applied Hedin’s (nonselfconsistent) GW approximation,²⁰ which has proven to be quite successful for semiconductors,^{62,63} in its FP-LMTO realization⁶⁴ to both, FeSb_2 and FeAs_2 . In Figs. 4(a) and 4(b) we display, besides the FP-LMTO (LDA) Kohn-Sham energies ϵ_{KS} , the band structure obtained by taking into account (perturbatively) the energy shifts as provided by the GW self-energy

$$\epsilon_{\text{GW}} \approx Z[\epsilon_{\text{KS}} + \text{Re} \Sigma(\epsilon_{\text{KS}})] \quad (22)$$

with $Z^{-1} = 1 - \partial_{\omega} \text{Re} \Sigma|_{\omega=\epsilon_{\text{KS}}}$. In the case of FeSb_2 this indeed opens a charge gap in agreement with experiment. We note that, consistent with the above hybrid functional calculation, as well as with the discussed orbital-dependent interaction strength, the static part of the GW self-energy, i.e., setting $Z=1$ in Eq. (22), yields a too large gap of ~ 0.2 eV. Thus, it is the dynamics of the self-energy, therewith a true correlation effect, that scales down the gap size with respect to Hartree-type approaches a situation quite akin to that of correlated band insulators.^{65,66} Indeed, the real part of the diagonal matrix elements of the self-energy are linear in frequency over an extended energy range of up to 10 eV. While the derivative of these elements are basically orbitally independent within the Fe $3d$, and Sb $5p$ orbital subsets, respectively, the different hybridizations and also the different off-diagonal elements yield for the antimonide, in the Kohn-Sham basis, a minimal value (eigenvalue of the self-energy derivative matrix) of $Z \approx 0.52$ for “bands” near the Fermi level, and $Z \approx 0.6-0.7$ for higher lying bands.⁶⁷ Concomitant with the linear slope of the real part, the imaginary part of the self-energy is basically quadratic, but notably asymmetric with respect to the Fermi level, see Table III.

Despite the larger values of the Hubbard U , the correlation dynamics is less pronounced in FeAs_2 and values of Z reach a minimum of $Z \approx 0.6$ for excitations closest to the Fermi level. Also, lifetime effects are both smaller in magnitude, and less asymmetric, see Table III. While for FeSb_2 the GW approach has the correct trend with respect to experiments, the gap of FeAs_2 also slightly increases from its LDA value, thus departing a bit further from the experimental value of $\sim 0.2-0.22$ eV. We note that while the gap size

within LDA/GGA and GW are comparable, the physics is not: In the GW , the size of the gap is a result of an almost compensation between static (exchangelike) contributions (see hybrids) and the dynamical correlations. The induced bandwidth narrowing within the GW distinguishes its excitations from the KS spectrum.

V. REALISTIC SEEBECK COEFFICIENTS FOR FeAs_2 AND FeSb_2

In the light of the above considerations for the thermopower of semiconductors, the principle puzzle now is why FeSb_2 while having a gap that is about seven times smaller than that of FeAs_2 , has a Seebeck coefficient that is (up to) five times larger.

For the calculation of the realistic Seebeck coefficient given in Eq. (1), we employ the Fermi velocity matrix elements of the optics implementation⁶⁸ of WIEN2K, and compute the correlation functions according to Eq. (2). We choose a dense momentum mesh with 13,320 irreducible k points.

A. FeAs_2

The band structure underlying the theoretical Seebeck coefficient of FeAs_2 is the GGA result shown in Fig. 3(c). Since the size of the gap is important for the magnitude of the thermopower, we scale down its size from its GGA value of 0.28 eV to the experimental 0.2 eV. In the model calculation above, the frequency integration is performed analytically and the momentum summation is turned into an integral over the density of states, allowing for a sufficiently

TABLE III. Comparison of the asymmetry of the scattering amplitude within the GW approximation. Extraction by fitting the average d orbital self-energy (in the Kohn-Sham basis) by $\Im \Sigma(|\omega| < 5 \text{ eV}) = -\Gamma \omega^2$.

Γ (eV^{-1})	FeSb_2	FeAs_2
$\omega < 0$	0.15	0.08
$\omega > 0$	0.02–0.05	0.02–0.03

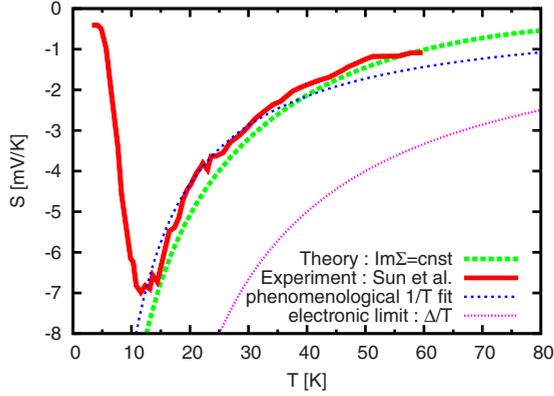


FIG. 5. (Color online) Thermopower of FeAs₂. Shown are the theoretical Seebeck coefficient for x polarization, using a constant self-energy, and the GGA band structure, with the gap scissored to 0.2 eV. Experimental results are of Sun *et al.* (Ref. 14) (x polarization). Also displayed is a simple $1/T$ fit corresponding to the large gap semiconductor, Eq. (17), yielding $\Delta/2\delta\lambda - \mu \approx 86$ meV. Further indicated is the largest possible purely electronic Seebeck coefficient for FeAs₂.

precise evaluation of the response functions to account for effects at very low temperature. The discrete momentum mesh in the *ab initio* case requires a finite broadening of the excitations. Here, we use a frequency and momentum independent scattering rate/self-energy $\Gamma \sim 20$ meV. Since this broadening is larger than the relevant temperature scale of the impurity dominated regime in FeAs₂, this does not allow one to capture the lifetime induced decrease in the thermopower or a precise enough determination of the chemical potential at very low temperatures. Therefore we do not consider the effect of impurities in the realistic part of the work for FeAs₂. Instead, we extract from our *ab initio* data the linear dependence of the chemical potential at high temperatures, and use this asymmetry via Eq. (21) of the semiconductor model (along with a reduced broadening of $\Gamma = 5$ μ eV) to access the chemical potential at low temperatures and we further eliminate in the transport kernel residual weight of valence and conduction bands inside the gap.⁶⁹ These limitations will exclude the realistic probing of temperatures in which impurity and coherence effects are crucial, which for FeAs₂ is experimentally found to be the case below ~ 12 K.

In Fig. 5 we show our theoretical Seebeck coefficient of FeAs₂ as a function of temperature (thick dashed curve), and compare it to experimental results¹⁴ for the same polarization. The agreement is excellent for all temperatures above 12 K, i.e., in the intrinsic, not impurity dominated, temperature regime.

Also shown is a simple fit, using the formula Eq. (17) for the large gap semiconductor. The individual determination of the parameters $\delta\lambda$ and μ (Δ given by experiment) is ambiguous, given the scale of the low-temperature thermopower (millivolt per kelvin) with respect to the high-temperature Heikes limit (which is on the order of $k_B/e = 86$ μ V/K). In Fig. 5 we show results for the large gap model for $\Delta/2\delta\lambda - \mu = 85$ meV (blue dotted curve), which is compatible with the constraints $|\delta\lambda| \leq 1$, and $\mu \leq \Delta/2$.

The decrease in the Seebeck coefficient at low temperature (that is beyond our numerical precision) can be understood from our model considerations. Indeed both scattering effects, as well as the presence of impurities can be at the origin of the upturn of the thermopower, as seen in Figs. 2(c) and 2(a), respectively.

B. FeSb₂

The situation is entirely different for FeSb₂. The maximal measured Seebeck coefficient $S(T=10$ K) is -45 mV/K.⁴ If one takes the maximum possible asymmetry parameter, $\delta\lambda = 1$, and one assumes the chemical potential to be at the most favorable position, i.e., $\mu = -\Delta/2$, the charge gap must be larger than $\Delta = 0.45$ eV to explain the value of the measured thermopower in terms of our purely electronic model. The experimental charge gap, however, is only $\Delta \approx 30$ meV. We are thus led to suspect that the large thermopower of FeSb₂ at low temperature is not purely of electronic origin.

A possible scenario, mentioned in the literature, is that the very large Seebeck coefficient is mainly caused by a substantial phonon drag, i.e., by an electron drift induced by a scattering with phonons. While there is no conclusive evidence that this effect is operational in FeSb₂, there are several reasons why it is more likely to be present in this material than in FeAs₂. Since the thermopower is a measure for the entropy per carrier, the phonon contribution to the Seebeck coefficient will be proportional to the lattice specific-heat times the electron-phonon coupling constant divided by the electron density. Given Debye temperatures of 348 K for FeSb₂,^{14,29} and 510 K for FeAs₂,¹⁴ the specific heat of FeSb₂ will be larger than that of FeAs₂. The charge-carrier concentration at temperatures where the thermopower is maximal, on the other hand, is larger for FeSb₂:¹⁵ for the best sample $n \sim 8 \times 10^{14}/\text{cm}^3$ whereas for FeAs₂ $n \sim 5 \times 10^{14}/\text{cm}^3$.¹⁵

The electron-phonon coupling is the least accessible ingredient from the theoretical point of view. Experimentally there are some insinuations: First of all, the low-temperature feature seen in the specific heat of FeSb₂, that has no analogue in the spin response and is absent in the arsenide,³⁰ could originate from a substantial electron-phonon coupling, charge ordering, excitonic or polaronic effects from an enhanced coupling to the lattice.

Also the nuclear-spin-lattice relaxation rate increases below 40 K,⁷⁰ i.e., in the regime where the thermopower starts its huge magnification whereas an activation law decrease ($\Delta = 473$ K) is found above 50 K,⁷⁰ in rough accordance with the intrinsic gap.

Moreover, optical spectroscopy witnesses a large change in phonon lifetimes across the metal-insulator transition, suggesting an important electron-lattice coupling.²⁴ Recently, also polarized Raman-scattering experiments gave indications for a notable electron-phonon coupling, that is strongly temperature dependent below 40 K.⁷¹ Further, we note that, as expected for substantial phonon-drag contributions to the Seebeck coefficient, the magnetothermopower of FeSb₂ is very low for those samples that exhibit the largest response without magnetic field.⁴ A decrease in the phonon mean-free

path by nonelectronic scattering (i.e., in particular, by imperfections) is expected to lower the respective effect in the thermopower and indeed the Seebeck coefficient of polycrystalline samples²⁹ and thin films³³ was found to be significantly smaller than for single crystals while having the same high-temperature behavior.⁷⁰ Recently, also substituted $\text{FeSb}_{2-x}\text{As}_x$ was investigated.¹⁵ Interestingly, it was found that the above mentioned increase in the susceptibility, starting at around 50 K, is stable with respect to the substitution whereas the shoulder in the resistivity at 10–20 K is flattened out, and the Seebeck coefficient decreases. In the phonon-drag picture this would, again, be owing to a decrease in the phonon mean-free path for nonelectronic scattering due to the presence of the As “impurities.”

Comparing the thermopower of the antimonide and the arsenide (see Fig. 1 in Ref. 14) one notes that the Seebeck coefficient of FeAs_2 is larger than that of FeSb_2 at 35 K and higher. This might indicate—if the phonon-drag picture holds—that the effective electron phonon coupling in FeSb_2 has sufficiently decreased (by umklapp and phonon-phonon scattering) so that the thermopower is now dominated by the electronic degrees of freedom, i.e., the larger gap in FeAs_2 causes a larger response. Yet, we note that optical spectroscopy²⁴ and some transport measurements^{26,27} see metallic behavior above 70 K, or already above 40 K, respectively, an effect not captured by our one-particle approach. Hence we will focus on the temperature range from 35 K upwards to, at best, 70 K.

Since the GGA Kohn-Sham spectrum is metallic, we opted for using the hybrid functional calculation [see Fig. 3(b)], albeit with a gap scissored to the experimental value $\Delta=0.03$ eV, to compute the theoretical Seebeck coefficient. Moreover, we assume the presence of donor impurities at $E_D=9$ meV, corresponding to an activation energy $\delta=\Delta/2-E_D=6$ meV as is seen in the resistivity in the range of 5–15 K,⁴ and we use an impurity concentration $n_D=10^{17}/\text{cm}^3$. This concentration yields $n_D^+ + p \approx 7 \times 10^{16}/\text{cm}^3$ at 20 K, in rough accordance with the respective hole concentration of $4 \times 10^{17}/\text{cm}^3$ found in Hall measurements.²⁸

As in the model case, we limit the influence of impurities to their effect on the chemical potential according to Eq. (8).⁷² At high temperature, the latter is linear in T as expected, and, using Eq. (21), we find an effective-mass ratio $\eta_v/\eta_c = m_v^*/m_c^* = 0.23$ when using a constant scattering rate, and the very similar $\eta_v/\eta_c = 0.25$ when using the imaginary parts of the self-energy from the GW calculation, i.e., the anisotropy is mainly propelled by the spectral weight and the Fermi velocities, and the GW scattering actually slightly reduces the particle-hole asymmetry in the current case. We further note that the asymmetry is opposite to that of FeAs_2 , where we found $\eta_v/\eta_c = 2.5 > 1$.

Thus obtained Seebeck coefficient is displayed in Fig. 6, along with experimental results on single crystals⁴ for the three polarizations along the crystallographic axes. In the limited range (discussed above), starting at 35 K and extending towards 70 K (indicated by the gray gradient in Fig. 6), we find good agreement with experiments: Both, the order of magnitudes, as well as the hierarchy of polarizations is captured within our approach. Below 35 K, the single-crystal

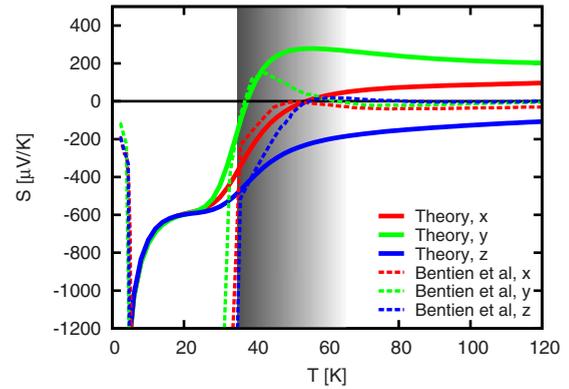


FIG. 6. (Color online) Thermopower of FeSb_2 . Shown are our theoretical together with experimental results from Ref. 4 for measurements along the crystallographic orientations x , y , z , as indicated. We can expect reasonable agreement in the temperature range indicated by the gray area. See text for details.

experiment reaches stellar magnitudes of up to -45 mV/K,⁴ that we argued to be beyond our approach which neglects vertex corrections. Measurements (not shown) using a polycrystalline sample,²⁹ and on films with preponderant $\langle 101 \rangle$ orientation³³ display Seebeck coefficients that at low temperatures never surpass -500 $\mu\text{V}/\text{K}$ and -200 $\mu\text{V}/\text{K}$, respectively, while having the exact same high-temperature behavior, advocating a disorder or decoherence induced lowering of the electron drift. At intermediate temperatures, those experiments agree qualitatively with both the single-crystal measurements, and our theoretical results.

VI. CONCLUSIONS

In conclusion we have considered the problem of thermoelectricity in correlated insulators and semiconductors. We developed a simple toy model to study how the various many-body renormalizations enter the thermoelectric response. We used LDA, hybrid density-functional theory and GW methods to carry out a comparative study of two systems of current experimental and theoretical interest FeAs_2 and FeSb_2 . The ratio between strength of the Hubbard U and the bandwidth of FeAs_2 and FeSb_2 are comparable and so is the correlation strength. In FeAs_2 DFT is qualitatively correct while in FeSb_2 correlation effects beyond DFT are essential for obtaining an insulating ground state and the one shot GW approximation succeeds in that respect. Indeed, using this method, we obtained good agreement with the experimental values of the gap for both materials.

The tools developed in this work were sufficient to describe the thermoelectric response of FeAs_2 quantitatively. This framework is not as successful for the FeSb_2 compound, and in particular it fails to explain the remarkably high low-temperature thermopower discovered by Bentien *et al.*⁴ Our work implies that the latter cannot be understood in the context of local correlations and one should focus either on vertex corrections to the transport coefficients or on nonlocal self-energy effects characteristic to the proximity to a quantum critical point. In this context we notice that within

LDA+ U this material is close to a ferromagnetic instability.¹⁷

An important form of vertex corrections describe the phonon-drag effect. A framework to estimate quantitatively these effects in conjunction with *ab initio* methods, are currently not available. Above, we mentioned several experimental findings that suggest the presence of this mechanism in FeSb₂, providing a strong incentive to further development in this vein.

Future work should include explicit calculations on correlated insulators using LDA+dynamical mean field theory to compare with the results of the toy model calculations. Furthermore, the investigation of vertex corrections on the thermoelectricity together with the effects of nonlocal self-

energies that go beyond the quasiparticle approximation should be considered.

ACKNOWLEDGMENTS

We thank K. Behnia, Kyoo Kim, and J. Snyder for interesting discussions. This international collaboration is supported by the NSF-materials world network under Grant No. NSF DMR 0806937 and by the PUF program. We are grateful to the KITP who hosted the program "Towards Materials Design Using Strongly Correlated Electron Systems" where part of this research was carried out. J.M.T. further thanks AIST, Japan, for hospitality.

-
- ¹G. J. Snyder and E. S. Toberer, *Nature Mater.* **7**, 105 (2008).
- ²A. Sakai, F. Ishii, Y. Onose, Y. Tomioka, S. Yotsuhashi, H. Adachi, N. Nagaosa, and Y. Tokura, *J. Phys. Soc. Jpn.* **76**, 093601 (2007).
- ³I. Terasaki, Y. Sasago, and K. Uchinokura, *Phys. Rev. B* **56**, R12685 (1997).
- ⁴A. Bentiën, S. Johnsen, G. K. H. Madsen, B. B. Iversen, and F. Steglich, *EPL* **80**, 17008 (2007).
- ⁵H. Schweitzer and G. Czycholl, *Phys. Rev. Lett.* **67**, 3724 (1991).
- ⁶G. Pálsson and G. Kotliar, *Phys. Rev. Lett.* **80**, 4775 (1998).
- ⁷V. S. Oudovenko and G. Kotliar, *Phys. Rev. B* **65**, 075102 (2002).
- ⁸C. Grenzebach, F. B. Anders, and G. Czycholl, *Physica B* **378-380**, 690 (2006).
- ⁹V. S. Oudovenko, G. Pálsson, K. Haule, G. Kotliar, and S. Y. Savrasov, *Phys. Rev. B* **73**, 035120 (2006).
- ¹⁰K. Haule and G. Kotliar, *Properties and Applications of Thermoelectric Materials, Proceedings of the NATO Advanced Research Workshop on Properties and Application of Thermoelectric Materials, Hvar, Croatia, 21-26 September 2008*, chapter Thermoelectrics Near the Mott Localization–Delocalization Transition, pages 119–131. NATO Science for Peace and Security Series B: Physics and Biophysics. Springer Netherlands, 2009.
- ¹¹K. Held, R. Arita, V. I. Anisimov, and K. Kuroki, *Properties and Applications of Thermoelectric Materials*, NATO Science for Peace and Security, Series B: Physics and Biophysics (Springer, Netherlands, 2009), pp. 141–157.
- ¹²R. Arita, K. Kuroki, K. Held, A. V. Lukoyanov, S. Skornyakov, and V. I. Anisimov, *Phys. Rev. B* **78**, 115121 (2008).
- ¹³P. Sun, N. Oeschler, S. Johnsen, B. B. Iversen, and F. Steglich, *Phys. Rev. B* **79**, 153308 (2009).
- ¹⁴P. Sun, N. Oeschler, S. Johnsen, B. B. Iversen, and F. Steglich, *Appl. Phys. Express* **2**, 091102 (2009).
- ¹⁵P. Sun, N. Oeschler, S. Johnsen, B. B. Iversen, and F. Steglich, *Dalton Trans.* **39**, 1012 (2010).
- ¹⁶W. Kohn, *Rev. Mod. Phys.* **71**, 1253 (1999).
- ¹⁷A. V. Lukoyanov, V. V. Mazurenko, V. I. Anisimov, M. Sigrist, and T. M. Rice, *Eur. Phys. J. B* **53**, 205 (2006).
- ¹⁸G. K. H. Madsen, A. Bentiën, S. Johnsen, and B. B. Iversen, *25th International Conference on Thermoelectrics: ICT '06*, Vienna, 2006 (IEEE, New York, 2006), pp. 579–581.
- ¹⁹A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- ²⁰L. Hedin, *Phys. Rev.* **139**, A796 (1965).
- ²¹C. Herring, *Phys. Rev.* **96**, 1163 (1954).
- ²²T. H. Geballe and G. W. Hull, *Phys. Rev.* **94**, 1134 (1954).
- ²³A. K. L. Fan, G. H. Rosenthal, H. L. McKinzie, and A. Wold, *J. Solid State Chem.* **5**, 136 (1972).
- ²⁴A. Perucchi, L. Degiorgi, R. Hu, C. Petrovic, and V. F. Mitrovic, *Eur. Phys. J. B* **54**, 175 (2006).
- ²⁵M. J. Rozenberg, G. Kotliar, and H. Kajueter, *Phys. Rev. B* **54**, 8452 (1996).
- ²⁶C. Petrovic, J. W. Kim, S. L. Budko, A. I. Goldman, P. C. Canfield, W. Choe, and G. J. Miller, *Phys. Rev. B* **67**, 155205 (2003).
- ²⁷R. Hu, V. F. Mitrović, and C. Petrovic, *Phys. Rev. B* **74**, 195130 (2006).
- ²⁸R. Hu, V. F. Mitrović, and C. Petrovic, *Appl. Phys. Lett.* **92**, 182108 (2008).
- ²⁹A. Bentiën, G. K. H. Madsen, S. Johnsen, and B. B. Iversen, *Phys. Rev. B* **74**, 205105 (2006).
- ³⁰F. Steglich and P. Sun (private communication).
- ³¹T. Koyama, Y. Fukui, Y. Muro, T. Nagao, H. Nakamura, and T. Kohara, *Phys. Rev. B* **76**, 073203 (2007).
- ³²R. Hu, V. F. Mitrović, and C. Petrovic, *Phys. Rev. B* **79**, 064510 (2009).
- ³³Y. Sun, S. Johnsen, P. Eklund, M. Sillassen, J. Böttiger, N. Oeschler, P. Sun, F. Steglich, and B. B. Iversen, *J. Appl. Phys.* **106**, 033710 (2009).
- ³⁴C. Petrovic, Y. Lee, T. Vogt, N. Dj. Lazarov, S. L. Budko, and P. C. Canfield, *Phys. Rev. B* **72**, 045103 (2005).
- ³⁵X. Wu, G. Steinle-Neumann, S. Qin, M. Kanzaki, and L. Dubrovinsky, *J. Phys.: Condens. Matter* **21**, 185403 (2009).
- ³⁶R. Hu, R. P. Hermann, F. Grandjean, Y. Lee, J. B. Warren, V. F. Mitrović, and C. Petrovic, *Phys. Rev. B* **76**, 224422 (2007).
- ³⁷J. M. Ziman, *Electrons and Phonons: The Theory of Transport Phenomena in Solids*, Oxford Classic Texts in the Physical Sciences (Oxford University Press, Amen House, London 1960).
- ³⁸As is the usual custom when not knowing the incoherent part of the spectrum (as, e.g., in slave boson techniques), we omit the Z factors in the determination of the particle numbers.

- ³⁹For a discussion on this kind of Peierls approximation see, e.g., Ref. 73.
- ⁴⁰We note that (even in the absence of impurities) the point of charge neutrality and the particle-hole symmetric point for the thermoelectric response are distinct from each other in the asymmetric case, in particular, the latter is largely influenced by the Fermi velocities.
- ⁴¹This is a typical order of magnitude for our compounds of interest, see, e.g., Ref. 14. For simplicity we are neglecting the sizable temperature dependence of the thermal lattice conductivity.
- ⁴²G. Mahan, B. Sales, and J. Sharp, *Phys. Today* **50**(3), 42 (1997).
- ⁴³J. B. Goodenough, *J. Solid State Chem.* **5**, 144 (1972).
- ⁴⁴In Ref. 17 a gap of 0.3eV was extracted from the density of states. Yet, the displayed bandstructure is metallic.
- ⁴⁵J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- ⁴⁶P. Blaha, K. Schwarz, G.-K.-H. Madsen, D. Kvasnicka, and J. Luitz, *Wien2k, an Augmented Plane Wave Plus Local Orbitals Program for Calculating Crystal Properties* (Vienna University of Technology, Austria, 2001).
- ⁴⁷H. Holseth and A. Kjekshus, *Acta Chem. Scand.* **23**, 3043 (1969).
- ⁴⁸Results (not shown) for low-temperature positions (Ref. 34) do not qualitatively differ as far as the following discussion is concerned.
- ⁴⁹The new unit vectors are $x=(-0.627, 0.326, 0.707)$, $y=(0.627, -0.326, 0.707)$, and $z=(0.462, 0.887, 0)$.
- ⁵⁰We note that the former hole pockets on the Brillouin-zone corners are slightly surpassed as the uppermost valence band by momentum regions on the way from the midpoint of the xz face (Y point), up z -wards towards the edge point T .
- ⁵¹M. Methfessel, M. van Schilfhaarde, and R. I. Casali, in *Electronic Structure and Physical Properties of Solids: The Uses of the LMTO Method*, Lecture Notes in Physics Vol. 535, edited by H. Dreyse (Springer, Berlin/Heidelberg, 2000), pp. 114–147.
- ⁵²N. Marzari and D. Vanderbilt, *Phys. Rev. B* **56**, 12847 (1997).
- ⁵³I. Souza, N. Marzari, and D. Vanderbilt, *Phys. Rev. B* **65**, 035109 (2001).
- ⁵⁴T. Miyake and F. Aryasetiawan, *Phys. Rev. B* **77**, 085122 (2008).
- ⁵⁵J. Kunes, R. Arita, P. Wissgott, A. Toschi, H. Ikeda, and K. Held, [arXiv:1004.3934](https://arxiv.org/abs/1004.3934), *Comput. Phys. Commun.* (to be published).
- ⁵⁶F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, *Phys. Rev. B* **70**, 195104 (2004).
- ⁵⁷T. Miyake, F. Aryasetiawan, and M. Imada, *Phys. Rev. B* **80**, 155134 (2009).
- ⁵⁸J. M. Tomczak, T. Miyake, R. Sakuma, and F. Aryasetiawan, *Phys. Rev. B* **79**, 235133 (2009).
- ⁵⁹J. M. Tomczak, T. Miyake, and F. Aryasetiawan, *Phys. Rev. B* **81**, 115116 (2010).
- ⁶⁰V. I. Anisimov, J. Zaanen, and O. K. Andersen, *Phys. Rev. B* **44**, 943 (1991).
- ⁶¹In WIEN2K, the hybrid functional approach is implemented only within the atomic spheres. Therewith one can choose to apply corrections to specific atomic characters.
- ⁶²F. Aryasetiawan and O. Gunnarsson, *Rep. Prog. Phys.* **61**, 237 (1998).
- ⁶³G. Onida, L. Reining, and A. Rubio, *Rev. Mod. Phys.* **74**, 601 (2002).
- ⁶⁴T. Kotani, M. van Schilfhaarde, and S. V. Faleev, *Phys. Rev. B* **76**, 165106 (2007).
- ⁶⁵J. Kuneš and V. I. Anisimov, *Phys. Rev. B* **78**, 033109 (2008).
- ⁶⁶M. Sentef, J. Kuneš, P. Werner, and A. P. Kampf, *Phys. Rev. B* **80**, 155116 (2009).
- ⁶⁷We stress that the parameter Z occurs here only formally as parameter in the frequency expansion of the self-energy and should not be confounded with the quasiparticle weight in a Fermi liquid.
- ⁶⁸C. Ambrosch-Draxl and J. O. Sofo, *Comput. Phys. Commun.* **175**, 1 (2006).
- ⁶⁹While for the sake of the frequency integration and momentum sum the broadening could be reduced when using even more k points (we use a mesh of 13,320 irreducible points), it is also needed in the matrix inversion that yields the spectral functions. In the pure band-structure case this can evidently be circumvented, e.g., by employing the tetrahedron method. The latter we do not use since our implementation allows for a frequency dependent self-energy, in which case the tetrahedron methods becomes more cumbersome to handle, see, e.g., Ref. 74.
- ⁷⁰A. A. Gippius, M. Baenitz, A. K. Rajarajan, E. M. Bruening, K. Okh otnikov, R. Walstedt, A. Strydom, J. Mydosh, and F. Steglich, *J. Phys.: Conf. Ser.* **150**, 042040 2009.
- ⁷¹N. Lazarević, Z. V. Popović, R. Hu, and C. Petrovic, *Phys. Rev. B* **81**, 144302 (2010).
- ⁷²The accounting for impurity effects in the *ab initio* case is easier to handle for FeSb₂ than for FeAs₂ since, both, the charge gap is much smaller and the temperatures that we look at larger.
- ⁷³J. M. Tomczak and S. Biermann, *Phys. Rev. B* **80**, 085117 (2009).
- ⁷⁴K. Haule, C.-H. Yee, and K. Kim, *Phys. Rev. B* **81**, 195107 (2010).