Impact of phonons and spin-orbit coupling on Auger recombination in InAs

Jimmy-Xuan Shen,[1] Daniel Steiauf,[2] Andrew McAllister,[3] Guangsha Shi,[3] Emmanouil Kioupakis,[3] Anderson Janotti,[2,*] and Chris G. Van de Walle[2,†]

[1]*Department of Physics, University of California, Santa Barbara, California 93106-9530, USA*

[2]*Materials Department, University of California, Santa Barbara, California 93106-5050, USA*

[3]*Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA*

(Received 20 December 2018; revised manuscript received 22 August 2019; published 7 October 2019)

We extend the methodology for calculating Auger recombination from first principles to include spin-orbit interaction, which allows us to study materials in which the spin-orbit splitting and the band gap are of similar magnitude. We use this methodology to compute the direct and indirect Auger coefficients in InAs and related alloys. The direct process involves only Coulomb interaction, while the indirect process is mediated by phonons. We show how the spin-orbit split-off states allow for efficient direct Auger recombination, greatly enhancing the excitation rate of Auger holes due to a near resonance between the spin-orbit splitting and the band gap of the material. We find that the direct *hole-hole-electron* Auger recombination coefficient decays exponentially with increasing band gap, while the *electron-electron-hole* Auger recombination coefficient decays faster than exponentially. Additionally, we modeled the phonon-assisted process in InAs and show that the direct Auger process is dominant.

DOI: 10.1103/PhysRevB.100.155202

I. INTRODUCTION

Auger recombination is an important nonradiative carrier recombination mechanism in semiconductors. It reduces the efficiency of light-emitting devices at high power and has been shown to play a significant role in the efficiency-droop and green-gap problems of nitride light-emitting diodes [1–3]. Auger processes also contribute to efficiency loss in photovoltaics [4] and limit the direct current performance of bipolar transistors [5]. Auger recombination is particularly critical in narrow-gap semiconductors such as InAs-based heterostructures. These materials find applications in mid-wavelength infrared (3–5 μm) diode lasers [6] as well as in interband cascade lasers with InAs/GaInSb quantum wells [7]. The reduction in quantum efficiency due to Auger recombination tends to govern the lasing threshold, especially at high carrier concentrations and high temperatures [8]. Understanding the microscopic details of the Auger recombination mechanism in these materials will aid in identifying engineering solutions to mitigate its impact.

General trends in the Auger coefficients of common semiconductors were examined by Bulashevich and Karpov [9]. They noted that the Auger coefficients in direct-band-gap materials fall off exponentially with increasing band gap until they reach a value around 10^{-30} cm^6 s^{-1}. A microscopic understanding of the Auger recombination process in low-band-gap materials such as InAs will help us better understand these trends.

In the Auger process (Fig. 1), the energy resulting from an electron-hole pair recombination is not emitted as a photon but transferred to a third carrier, either an Auger electron (in the *e-e-h* process) or an Auger hole (in the *h-h-e* process). In the direct *e-e-h* process illustrated in Fig. 1(b), an initial electron-hole pair (states 1 and 3) recombines across the band gap and transfers its energy to a third electron (state 2) that gets excited to a high-energy state in the conduction band (state 4). Such direct Auger recombination is mediated purely by Coulomb interactions. However, additional second-order interactions involving the emission or absorption of a phonon [Fig. 1(c)] can also contribute significantly to the total Auger rate [2,10].

First-principles calculations based on density functional theory (DFT) are well suited for the predictive modeling of optoelectronic materials [11]. Auger carriers are excited to states with energy approximately equal to the gap away from the band edges. DFT calculations can reliably predict the energies and wave functions throughout the Brillouin zone (BZ) and at higher energies. Based on such calculations, it was revealed that in wide-band-gap semiconductors such as GaN Auger recombination is dominated by the indirect, phonon-assisted, process [2,12]. It is still an open question to what extent phonon-assisted processes contribute to Auger recombination in smaller-gap materials.

Calculations of the Auger rate in narrow-gap materials require an accurate treatment of spin-orbit coupling (SOC), which modifies the band structure near the band extrema. In InAs, the magnitude of the spin-orbit (SO) splitting (0.39 eV [13]) is comparable to the band gap (0.35 at 300 K [14] and 0.43 eV at 0 K [13]) and enables additional direct Auger recombination processes that would otherwise be forbidden by energy and momentum conservation. In addition, the lack of inversion symmetry in InAs breaks the spin degeneracy of the bands. Hence, careful consideration of the SOC is needed.

*Present address: Department of Materials Science and Engineering, University of Delaware, Newark, Delaware 19716, USA.

†Corresponding author: vandewalle@mrl.ucsb.edu

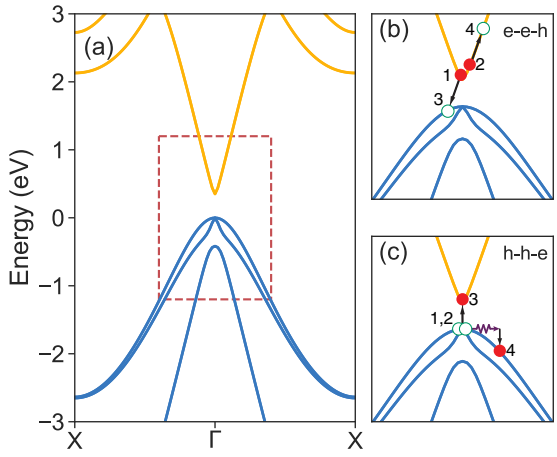

FIG. 1. (a) Calculated band structure of InAs. The indicated region near the center of the Brillouin zone is enlarged in panels (b) and (c) to illustrate Auger processes: (b) the direct *e-e-h* process and (c) the phonon-assisted *h-h-e* process.

In this paper, we report a fully first-principles prediction of SOC effects in direct and phonon-assisted Auger recombination, which requires significant modifications to previously developed methodology [12]. When applied to InAs, the results show that SOC significantly enhances direct Auger recombination. The resulting rate for the direct Auger process in InAs is a factor of 50 larger than for the phonon-assisted process. We also find that, as the band gap of the material increases, the direct Auger process becomes comparatively less relevant. The Auger recombination coefficients decay approximately exponentially as a function of band gap, and this strong dependence means that the direct Auger process can be significantly reduced by alloying.

II. METHODOLOGY

A. Computational details

Our first-principles calculations are based on DFT [15,16]. The electronic structure is computed with projector augmented wave potentials as implemented in the Vienna *Ab-initio* Simulation Package (VASP) [17,18]. The *d* electrons in In and As were treated as core electrons, and a plane-wave energy cutoff of 500 eV was used. The Auger rate is highly sensitive to the eigenvalues near the band extrema, requiring an accurate description of the band structure. We use the modified Becke-Johnson functional within the local-density approximation (MBJLDA) [19,20]; a fit to a highly accurate band structure calculated with the Heyd-Scuseria-Ernzerhof functional [21] yielded a c parameter of 1.15 for the MBJLDA functional. The resulting band structure is shown in Fig. 1(a). The calculated LDA lattice parameter is $a = 6.06$ Å, in very good agreement with experiment [22]. The phonon spectra and electron-phonon coupling matrix elements are obtained with density functional perturbation theory [23]. Since electron-phonon interactions are not implemented in VASP at this time, we have used the Quantum ESPRESSO code [24]. Perturbations to the wave functions due to SOC are implemented using the vector part of the *ab initio* pseudopotentials [25].

The reciprocal space is sampled with a $60 \times 60 \times 60$ Γ-centered $\boldsymbol{k}$-point mesh for the electron and phonon states. The screened Coulomb interaction is evaluated using a static model dielectric function [12] and the experimental dielectric constant of 12.3 for InAs [26]. We account for the screening by free carriers with the Debye model. For the evaluation of the phonon-assisted Auger rate we assume that the initial states of the three carriers are all at Γ—the same approximation as employed in Refs. [3,10] to render the problem computationally tractable.

B. Inclusion of SOC for direct Auger process

The methodology in Ref. [12] must be modified to treat systems with strong SOC. We will first discuss how the spin quantum number is accounted for in non-SOC systems; then we will show how spin-orbit coupling is implemented in our calculations.

Our computational approach is based on time-dependent perturbation theory using Fermi's "golden rule." We represent the states in the band structure using composite index $\boldsymbol{i} = [\boldsymbol{k}_i n_i]$, where $\boldsymbol{k}$ is the momentum quantum number and n is the band index. In the spin-degenerate case, for each state $\boldsymbol{i}$ there are two possible spin quantum numbers $\sigma_i = \pm\frac{1}{2}$. In the absence of SOC, the eigenvalues do not depend on σ_i, so the total direct Auger recombination rate, $R^{\text{direct}}_{\text{Auger}}$, is given by

$$R^{\text{direct}}_{\text{Auger}} = \frac{2\pi}{\hbar} \sum_{\mathbf{1234}} \sum_{\sigma_1\sigma_2\sigma_3\sigma_4} f_1 f_2 (1-f_3)(1-f_4) \times |M_{\mathbf{1234}}|^2 \delta(\varepsilon_1 + \varepsilon_2 - \varepsilon_3 - \varepsilon_4), \tag{1}$$

where f_i denotes the Fermi occupation number of the electronic state indexed by $\boldsymbol{i}$ with energy ε_i. The quantity inside the absolute value is the screened Coulomb matrix element:

$$M_{\mathbf{1234}} \equiv \underbrace{\langle \mathbf{12} \,|\, \hat{W} \,|\, \mathbf{34} \rangle}_{M^d_{\mathbf{1234}}} \delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4} - \underbrace{\langle \mathbf{12} \,|\, \hat{W} \,|\, \mathbf{43} \rangle}_{M^x_{\mathbf{1234}}} \delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}, \tag{2}$$

where the first term is a *direct* term and the second is an *exchange* term, accounting for the antisymmetry under fermion exchange. The *direct* and *exchange* terms can be obtained by evaluating the two-body Coulomb interaction $\hat{W}(\boldsymbol{r}_\alpha, \boldsymbol{r}_\beta)$ between the states represented in real space as $\psi_{\boldsymbol{i}}(\boldsymbol{r})$:

$$\langle \mathbf{12} \,|\, \hat{W} \,|\, \mathbf{34} \rangle \equiv \iint d\boldsymbol{r}_\alpha d\boldsymbol{r}_\beta \psi^*_{\mathbf{1}}(\boldsymbol{r}_\alpha)\psi^*_{\mathbf{2}}(\boldsymbol{r}_\beta)\hat{W}(\boldsymbol{r}_\alpha, \boldsymbol{r}_\beta)\psi^*_{\mathbf{3}}(\boldsymbol{r}_\alpha)\psi^*_{\mathbf{4}}(\boldsymbol{r}_\beta). \tag{3}$$

After the spin summation over the $2^4 = 16$ combined spin configurations, we recover the expression for the Auger rate from Ref. [12]:

$$R^{\text{direct}}_{\text{Auger}} = 2\frac{2\pi}{\hbar} \sum_{\mathbf{1234}} f_1 f_2 (1-f_3)(1-f_4) \times \left(\left| M^d_{\mathbf{1234}} - M^x_{\mathbf{1234}} \right|^2 + \left| M^d_{\mathbf{1234}} \right|^2 + \left| M^x_{\mathbf{1234}} \right|^2 \right) \times \delta(\varepsilon_1 + \varepsilon_2 - \varepsilon_3 - \varepsilon_4). \tag{4}$$

SOC is included via a perturbing Hamiltonian $\Delta\hat{H}^{\mathrm{SOC}}$ expressed in terms of the spectrum of Kohn-Sham states $|\boldsymbol{k}n\sigma\rangle$ at each k point [25]:

$$\langle \boldsymbol{k}n\sigma \mid \Delta\hat{H}^{\mathrm{SOC}} \mid \boldsymbol{k}n'\sigma'\rangle. \tag{5}$$

This Hamiltonian is then diagonalized to obtain the energy eigenvalues and spinor wave functions at that k point. In non-spin-degenerate materials like InAs, the wave function must be represented as a spinor wave function where the two spin components are independent:

$$|\boldsymbol{i}\rangle \equiv |\psi_{\boldsymbol{i}}^{\uparrow}\rangle \otimes |\uparrow\rangle + |\psi_{\boldsymbol{i}}^{\downarrow}\rangle \otimes |\downarrow\rangle\,. \tag{6}$$

The Auger rate in the non-spin-degenerate case is given by an expression similar to Eq. (1) where the summation over the index $\boldsymbol{i}$ in the spinor basis [Eq. (6)] now replaces the separate summations over the band and spin indices:

$$\begin{aligned} R_{\mathrm{Auger}}^{\mathrm{direct}} &= \frac{2\pi}{\hbar}\sum_{\mathbf{1234}} f_{\mathbf{1}}f_{\mathbf{2}}(1-f_{\mathbf{3}})(1-f_{\mathbf{4}})\left|M^{d}_{\mathbf{1234}} - M^{x}_{\mathbf{1234}}\right|^2 \\ &\quad\times\ \delta(\varepsilon_{\mathbf{1}}+\varepsilon_{\mathbf{2}}-\varepsilon_{\mathbf{3}}-\varepsilon_{\mathbf{4}}). \end{aligned} \tag{7}$$

The matrix element in Eq. (7) corresponds to Eq. (2) but with nonorthonormal spin components. To evaluate this matrix element, we must sum over each spin component of the wave functions ($\psi_{\boldsymbol{i}}^{\uparrow}$ and $\psi_{\boldsymbol{i}}^{\downarrow}$) for the initial states:

$$\begin{aligned} M^{d}_{\mathbf{1234}} &\equiv \langle \mathbf{12}|\hat{W}|\mathbf{34}\rangle \\ &= \langle\psi_{\mathbf{1}}^{\uparrow}\psi_{\mathbf{2}}^{\uparrow}|\hat{W}|\psi_{\mathbf{3}}^{\uparrow}\psi_{\mathbf{4}}^{\uparrow}\rangle + \langle\psi_{\mathbf{1}}^{\uparrow}\psi_{\mathbf{2}}^{\downarrow}|\hat{W}|\psi_{\mathbf{3}}^{\uparrow}\psi_{\mathbf{4}}^{\downarrow}\rangle \\ &\quad\times\langle\psi_{\mathbf{1}}^{\downarrow}\psi_{\mathbf{2}}^{\uparrow}|\hat{W}|\psi_{\mathbf{3}}^{\downarrow}\psi_{\mathbf{4}}^{\uparrow}\rangle + \langle\psi_{\mathbf{1}}^{\downarrow}\psi_{\mathbf{2}}^{\downarrow}|\hat{W}|\psi_{\mathbf{3}}^{\downarrow}\psi_{\mathbf{4}}^{\downarrow}\rangle, \end{aligned} \tag{8}$$

$$\begin{aligned} M^{x}_{\mathbf{1234}} &\equiv \langle \mathbf{12}|\hat{W}|\mathbf{43}\rangle \\ &= \langle\psi_{\mathbf{1}}^{\uparrow}\psi_{\mathbf{2}}^{\uparrow}|\hat{W}|\psi_{\mathbf{4}}^{\uparrow}\psi_{\mathbf{3}}^{\uparrow}\rangle + \langle\psi_{\mathbf{1}}^{\uparrow}\psi_{\mathbf{2}}^{\downarrow}|\hat{W}|\psi_{\mathbf{4}}^{\uparrow}\psi_{\mathbf{3}}^{\downarrow}\rangle \\ &\quad\times\langle\psi_{\mathbf{1}}^{\downarrow}\psi_{\mathbf{2}}^{\uparrow}|\hat{W}|\psi_{\mathbf{4}}^{\downarrow}\psi_{\mathbf{3}}^{\uparrow}\rangle + \langle\psi_{\mathbf{1}}^{\downarrow}\psi_{\mathbf{2}}^{\downarrow}|\hat{W}|\psi_{\mathbf{4}}^{\downarrow}\psi_{\mathbf{3}}^{\downarrow}\rangle. \end{aligned} \tag{9}$$

The two terms ($|M^{d}_{\mathbf{1234}}|^2$ and $|M^{x}_{\mathbf{1234}}|^2$) in Eq. (4) are absent in Eq. (7), reflecting the fact that Eq. (7) does not include a spin summation over the $\sigma_{\boldsymbol{i}}$'s. Instead, the number of allowed band indices for each state $|\boldsymbol{i}\rangle$ is doubled for the non-spin-degenerate case.

C. Inclusion of SOC for phonon-assisted Auger recombination

Following the same argumentation as for the direct process in the previous section, the expression for the indirect Auger rate is

$$\begin{aligned} R_{\mathrm{Auger}}^{\mathrm{indirect}} &= \frac{2\pi}{\hbar}\sum_{\mathbf{1234},\nu\boldsymbol{q}} f_{\mathbf{1}}f_{\mathbf{2}}(1-f_{\mathbf{3}})(1-f_{\mathbf{4}})\left(n_{\nu\boldsymbol{q}}+\frac{1}{2}\pm\frac{1}{2}\right) \\ &\quad\times|\tilde{M}_{\mathbf{1234};\nu\boldsymbol{q}}|^2\delta(\epsilon_{\mathbf{1}}+\epsilon_{\mathbf{2}}-\epsilon_{\mathbf{3}}-\epsilon_{\mathbf{4}}\mp\hbar\omega_{\nu\boldsymbol{q}}). \end{aligned} \tag{10}$$

Here, the $n_{\nu\boldsymbol{q}}$ are the phonon-occupation numbers from Bose-Einstein statistics:

$$n_{\nu\boldsymbol{q}} = \frac{1}{e^{\hbar\omega_{\nu\boldsymbol{q}}/k_{\mathrm{B}}T}-1}, \tag{11}$$

and the phonon energy ($\hbar\omega_{\nu\boldsymbol{q}}$) is accounted for in the energy conservation delta function. The upper (lower) sign represents the phonon-emission (absorption) process. The generalized matrix element ($\tilde{M}$, with the subscript omitted) is given by

$$|\tilde{M}|^2 = |\tilde{M}^1+\tilde{M}^2+\tilde{M}^3+\tilde{M}^4-\tilde{M}^5-\tilde{M}^6-\tilde{M}^7-\tilde{M}^8|^2, \tag{12}$$

where each of the terms ($\tilde{M}^1$ to $\tilde{M}^8$) corresponds to one of the phonon-assisted processes depicted in Fig. 3 of Ref. [12]:

$$\tilde{M}^{1}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{g_{\mathbf{1}\boldsymbol{m};\nu}M^{d}_{\boldsymbol{m}\mathbf{234}}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{1}}\pm\hbar\omega_{\nu\boldsymbol{q}}+i\eta}, \tag{13}$$

$$\tilde{M}^{2}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{g_{\mathbf{2}\boldsymbol{m};\nu}M^{d}_{\mathbf{1}\boldsymbol{m}\mathbf{34}}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{2}}\pm\hbar\omega_{\nu\boldsymbol{q}}+i\eta}, \tag{14}$$

$$\tilde{M}^{3}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{M^{d}_{\mathbf{12}\boldsymbol{m}\mathbf{4}}g_{\mathbf{3}\boldsymbol{m};\nu}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{3}}\mp\hbar\omega_{\nu\boldsymbol{q}}+i\eta}, \tag{15}$$

$$\tilde{M}^{4}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{M^{d}_{\mathbf{123}\boldsymbol{m}}g_{\mathbf{4}\boldsymbol{m};\nu}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{4}}\mp\hbar\omega_{\nu\boldsymbol{q}}+i\eta}, \tag{16}$$

$$\tilde{M}^{5}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{g_{\mathbf{1}\boldsymbol{m};\nu}M^{x}_{\boldsymbol{m}\mathbf{234}}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{1}}\pm\hbar\omega_{\nu\boldsymbol{q}}+i\eta}, \tag{17}$$

$$\tilde{M}^{6}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{g_{\mathbf{2}\boldsymbol{m};\nu}M^{x}_{\mathbf{1}\boldsymbol{m}\mathbf{34}}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{2}}\pm\hbar\omega_{\nu\boldsymbol{q}}+i\eta}, \tag{18}$$

$$\tilde{M}^{7}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{M^{x}_{\mathbf{12}\boldsymbol{m}\mathbf{4}}g_{\mathbf{3}\boldsymbol{m};\nu}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{3}}\mp\hbar\omega_{\nu\boldsymbol{q}}+i\eta}, \tag{19}$$

$$\tilde{M}^{8}_{\mathbf{1234};\nu\boldsymbol{q}} = \sum_{\boldsymbol{m}}\frac{M^{x}_{\mathbf{123}\boldsymbol{m}}g_{\mathbf{4}\boldsymbol{m};\nu}}{\epsilon_{\boldsymbol{m}}-\epsilon_{\mathbf{4}}\mp\hbar\omega_{\nu\boldsymbol{q}}+i\eta}. \tag{20}$$

Following the notation in Ref. [12], $\boldsymbol{m}$ is the index of an intermediate state and η is the inverse lifetime of that intermediate state. The direct (M^d) and indirect (M^x) screened Coulomb matrix elements are defined in Eqs. (8) and (9). The electron-phonon matrix elements (g) are evaluated using density functional perturbation theory for the spin-degenerate system, and expressed in the spinor wave-function basis of the form in Eq. (6) via a transformation

$$U_{i,j\sigma_j} = (|\psi_{\boldsymbol{i}}^{\uparrow}\rangle\otimes|\uparrow\rangle + |\psi_{\boldsymbol{i}}^{\downarrow}\rangle\otimes|\downarrow\rangle)(\langle\psi_j|\otimes\langle\sigma_j|), \tag{21}$$

from the product-state wave functions to spinor wave functions.

III. RESULTS AND DISCUSSIONS

We assume equal densities of free electrons and holes ($n=p$), which is typical in optoelectronic devices with high carrier injection. The Auger coefficients can be defined by $C_{\mathrm{Auger}} = R_{\mathrm{Auger}}/n^3$. The resulting direct and phonon-assisted coefficients at n=10^{18} cm^{-3} are tabulated in Table I. At the experimental band gap of 0.35 eV we computed a direct Auger coefficient (*e-e-h* plus *h-h-e*) of 8.60×10^{-28}cm^6 s^{-1}. The computed phonon-assisted coefficient, 1.5×10^{-29} cm^6 s^{-1}, is more than an order of magnitude smaller. A survey of experimental literature [27–32] reveals significant variations among the reported Auger coefficients, ranging

TABLE I. Computed Auger coefficients C (in units of 10^{-30} cm^6 s^{-1}) for InAs at 300 K with a carrier concentration of 10^{18} cm^{-3}.

	Direct	Direct (no SOC)	Phonon	Phonon (no SOC)
e-e-h	496	387	9	10
h-h-e	364	84	6	9
Total	860	471	15	19

from 1–3 $\times 10^{-28}$ cm^6 s^{-1} in InAs based quantum wells [32] to bulk values of 6.0×10^{-27} in Ref. [29] and 6.5×10^{-26} cm^6 s^{-1} in Ref. [27]. The model theory of Takeshima [33] predicted an Auger coefficient of 1.8×10^{-27} cm^6 s^{-1} (at $T = 300$ K and $n = 10^{18}$ cm^{-3}).

While the experimental measurements of the Auger recombination rates are conducted at room temperature, the effective temperature of the carriers can be much higher than the lattice temperature. To evaluate the Auger coefficients at higher carrier temperatures, we performed a series of calculations for the direct Auger process between 150 and 450 K and fit the results to an Arrhenius law [34] [$C(T) = C_\infty \exp(-E_A/k_BT)$], as shown in Fig. 2. The resulting activation energies are $E_A^{e\text{-}e\text{-}h} = 0.021$ eV and $E_A^{h\text{-}h\text{-}e} = 0.011$ eV.

We now describe an approximate approach that allows us to qualitatively comment on recombination processes in materials related to InAs. InAs-based alloys can be formed by alloying with GaAs (band gap 1.43 eV) or InSb (band gap 0.17 eV), both of which have band structures similar to InAs [19]. Explicit alloy calculations require large supercells and are computationally prohibitive with the k-point densities needed for the Auger calculations. We therefore approximate the alloy electronic structure by modifying the band gap of InAs while keeping all of the Coulomb and electron-phonon matrix elements fixed—commonly referred to as a *scissors shift*. In doing so we are neglecting additional scattering mechanisms due to alloy disorder and focusing purely on the band-gap dependence. Figure 3 shows the calculated coefficients as a function of the shifted band gap. While the Auger coefficients predicted at different band gaps are not intended to quantitatively represent the actual alloys, shifting the band gap still allows us to probe the dependence of the Auger process on different features in the band structure.

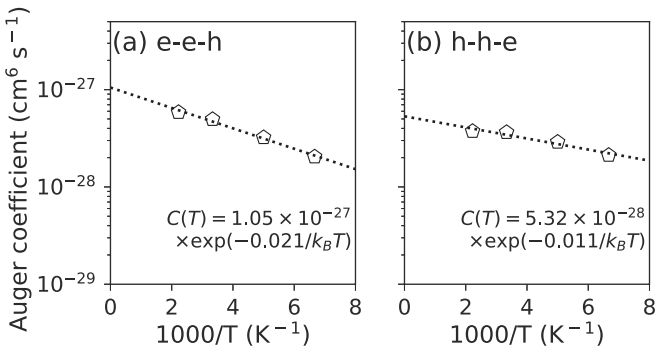

FIG. 2. Computed direct Auger coefficients at the experimental band gap ($E_g = 0.35$ eV) with a carrier concentration of 10^{18} cm^{-3}, for temperatures between 150 and 450 K, fitted to an Arrhenius relationship.

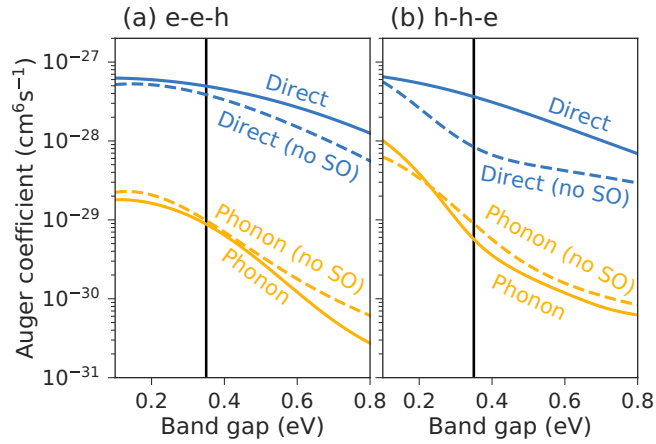

FIG. 3. Auger coefficient vs scissors-shift-adjusted band gap for direct (solid blue) and phonon-assisted (solid orange) Auger recombination due to (a) *e-e-h* and (b) *h-h-e* processes. The calculated coefficients without SO interactions for the direct (dashed blue) and phonon-assisted (dashed orange) processes are included for comparison. The gray vertical line indicates the experimental band gap of InAs. The calculations are carried out at 300 K with a carrier concentration of 10^{18} cm^{-3}.

A. Direct Auger process

Figure 3 shows that the direct Auger process is stronger than the indirect Auger process across the entire range of band-gap values for both *e-e-h* and *h-h-e*. The decrease of the direct *e-e-h* Auger coefficient as a function of increasing band gap can be attributed to the lack of higher-energy states (roughly one band gap higher in energy than the conduction-band minimum) near the center of the BZ. This renders energy and momentum conservation difficult to satisfy when three of the states must come from the same highly dispersive conduction band [states 1, 2, and 4 in Fig. 1(b)]. The decrease in the direct *h-h-e* coefficient depends strongly on the near resonance of the SOC and follows an exponential trend, which reflects the tail of the Fermi-Dirac distributions. The same trends can be observed in previous studies of GaN [12] and GaAs [10].

Inclusion of SOC effects modifies the band structure near the valence-band maximum and significantly enhances the *h-h-e* direct Auger coefficient. This is most evident for band-gap values close to the SO-splitting energy of InAs (see Fig. 3), where a near resonance occurs. At the band gap of InAs, the *h-h-e* Auger coefficient is 3.64×10^{-28} cm^6 s^{-1} when SOC is included (see Table I), more than a factor of 4 greater than the coefficient computed without SOC (8.4×10^{-29} cm^6 s^{-1}). The inclusion of SOC has a less dramatic effect on the *e-e-h* process (4.96×10^{-28} cm^6 s^{-1} with SOC and 3.87×10^{-28} cm^6 s^{-1} without). The near coincidence of the band gap with the SO splitting makes InAs close to a *worst case scenario* in terms of impact of SO coupling on Auger recombination. For this reason, it is often assumed that the rapid Auger recombination in InAs is primarily due to the *h-h-e* process. However, our calculations show that the *e-e-h*

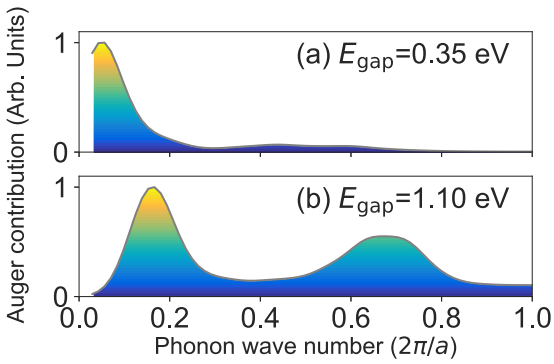

FIG. 4. Normalized contributions to the phonon-assisted Auger rate plotted as a function of phonon wave number q for (a) a band gap of 0.35 eV and (b) a band gap of 1.1 eV.

recombination is just as important and, in fact, accounts for a majority of the direct Auger recombination in InAs. The high *e-e-h* rate can be attributed to the large difference in dispersion between the conduction and valence bands. In an *e-e-h* process, the final Auger electron state occurs at roughly one band-gap energy above the initial state, but since the conduction band is very dispersive the associated momentum is quite small. This momentum can easily be provided by the hole state, since the valence-band dispersion is low and hence occupation of higher-momentum states is significant.

B. Phonon-assisted Auger process

In contrast to the direct process, in the indirect process a phonon supplies the additional momentum (the phonons do not contribute much to energy conservation), in principle allowing the Auger electron and hole to be scattered anywhere in the BZ. At the carrier concentrations and temperatures considered here, the initial electron and hole states represent a very small portion of the BZ—0.03% for electrons and 2% for holes. Any small amount of momentum difference in the initial states can easily be compensated by the additional phonon, which justifies our approximation that the initial states are all at Γ. The small momentum differences will modify the matrix elements; however, any corrections to the phonon-assisted rate would not change our conclusion that the direct process dominates across the entire range of band gaps considered in Fig. 3.

Figure 4 shows contributions to the total Auger rate from phonons of different absolute wave numbers. At the 0.35-eV band gap of InAs [Fig. 4(a)], the vast majority of phonon-assisted Auger contributions comes from the center of the BZ, near Γ. In this regime, the energy and momentum conservation for the phonon-assisted Auger processes is very similar to the direct process, and the phonon-assisted Auger process behaves as a higher-order correction to the direct Auger recombination, with the phonon-assisted rate more than an order of magnitude smaller than the direct rate (Table I).

In contrast, at larger band gaps [Fig. 4(b)], states from other regions of the BZ can be involved via a phonon. These states account for a much greater portion of the total phonon-assisted Auger rate, with barely any contributions coming from the zone center (since there are no states in the vicinity of Γ that would allow energy conservation to be satisfied with a low-wave-number phonon). For even higher band gaps, the majority of states involved in the recombination process are no longer near the zone center (where the direct process occurs), and the indirect Auger process becomes stronger relative to the direct process.

In wide-band-gap materials (e.g., GaN [12] with a band gap of 3.5 eV) the phonon-assisted process becomes the dominant process. This is because the direct Auger coefficient falls off exponentially at large band gaps, while the phonon-assisted coefficient decays much slower, roughly as an inverse power law above 0.5 eV. The fact that the direct Auger coefficient is decreasing much faster than the phonon-assisted Auger coefficient at higher band gaps explains the trends observed in Ref. [9]. A crossover between the direct and phonon-assisted rates will occur at a band gap above 1.5 eV. We acknowledge that this is well beyond the limit where our approximate treatment of alloys is expected to be accurate. However, we believe the trends to be correct, and the observation that direct Auger recombination dominates in direct-gap materials at lower band gaps while the phonon-assisted Auger process becomes dominant at higher band gaps should be general.

Since the Auger coefficients generally decrease with increasing band gap, suppression of Auger recombination can in principle be achieved with band-gap engineering. This can be accomplished by alloying or straining the material, but is not always feasible if specific band-gap values are required. However, the strong dependence of the Auger coefficient on SOC opens an additional dimension in the design space, allowing for suppression of Auger recombination by eliminating the resonance between the SO splitting and the band gap [35].

IV. CONCLUSION

In conclusion, we have developed methodology to include spin-orbit coupling in first-principles simulations [12] of direct and indirect Auger recombination. Our calculations for InAs show that inclusion of SOC enhances the *h-h-e* recombination coefficient by more than a factor of 4. The direct Auger process is dominant in InAs, accounting for more than 98% of the total Auger rate. The direct process decreases exponentially with increasing band gap, making the phonon-assisted process relatively more important in larger-gap materials.

ACKNOWLEDGMENTS

Discussions with J. Bowers and G. Kresse are gratefully acknowledged. J.-X.S. was supported by the U.S. Department of Energy (DOE) Office of Science, Basic Energy Sciences under Grant No. DE-SC0010689. A.M. acknowledges support from the National Science Foundation (NSF) Graduate Research Fellowship Program through Grant No. DGE 1256260. G.S. and E.K. acknowledge support by the NSF CAREER award through Grant No. DMR-1254314. A.J. was supported by NSF Grant No. DMR-1652994. Computational resources were provided by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

[1] Y. C. Shen, G. Mueller, S. Watanabe, N. Gardner, A. Munkholm, and M. Krames, Appl. Phys. Lett. **91**, 141101 (2007).

[2] E. Kioupakis, P. Rinke, K. T. Delaney, and C. G. Van de Walle, Appl. Phys. Lett. **98**, 161107 (2011).

[3] E. Kioupakis, Q. Yan, D. Steiauf, and C. G. Van de Walle, New J. Phys. **15**, 125006 (2013).

[4] M. Green, IEEE Trans. Electron Devices **31**, 671 (1984).

[5] H.-W. Chiang, J. C. Rode, P. Choudhary, and M. J. W. Rodwell, J. Appl. Phys. **116**, 164509 (2014).

[6] D. H. Chow, R. H. Miles, T. C. Hasenberg, A. R. Kost, Y.-H. Zhang, H. L. Dunlap, and L. West, Appl. Phys. Lett. **67**, 3700 (1995).

[7] W. W. Bewley, C. L. Canedy, C. S. Kim, M. Kim, C. D. Merritt, J. Abell, I. Vurgaftman, and J. R. Meyer, Opt. Express **20**, 3235 (2012).

[8] J. R. Meyer, C. L. Felix, W. W. Bewley, I. Vurgaftman, E. H. Aifer, L. J. Olafsen, J. R. Lindle, C. A. Hoffman, M.-J. Yang, B. R. Bennett, B. V. Shanabrook, H. Lee, C.-H. Lin, S. S. Pei, and R. H. Miles, Appl. Phys. Lett. **73**, 2857 (1998).

[9] K. A. Bulashevich and S. Yu. Karpov, Phys. Status Solidi C **5**, 2066 (2008).

[10] D. Steiauf, E. Kioupakis, and C. G. Van de Walle, ACS Photonics **1**, 643 (2014).

[11] J. M. Rondinelli and E. Kioupakis, Annu. Rev. Mater. Res. **45**, 491 (2015).

[12] E. Kioupakis, D. Steiauf, P. Rinke, K. T. Delaney, and C. G. Van de Walle, Phys. Rev. B **92**, 035207 (2015).

[13] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001).

[14] Z. M. Fang, K. Y. Ma, D. H. Jaw, R. M. Cohen, and G. B. Stringfellow, J. Appl. Phys. **67**, 7034 (1990).

[15] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[16] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[17] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[18] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).

[19] Y.-S. Kim, M. Marsman, G. Kresse, F. Tran, and P. Blaha, Phys. Rev. B **82**, 205212 (2010).

[20] F. Tran and P. Blaha, Phys. Rev. Lett. **102**, 226401 (2009).

[21] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **124**, 219906 (2006).

[22] O. Madelung, U. Rössler, and M. Schulz, Indium arsenide (InAs) lattice parameters, thermal expansion, in *Group IV Elements, IV-IV and III-V Compounds. Part A - Lattice Properties* (Springer, Berlin, Heidelberg, 2001), pp. 1–6.

[23] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).

[24] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys. Condens. Matter **21**, 395502 (2009).

[25] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 2920 (1986).

[26] S. Adachi, J. Appl. Phys. **53**, 8775 (1982).

[27] V. L. Dalal, W. A. Hicinbothem, and H. Kressel, Appl. Phys. Lett. **24**, 184 (1974).

[28] K. L. Vodopyanov, H. Graener, C. C. Phillips, and T. J. Tate, Phys. Rev. B **46**, 13194 (1992).

[29] J. R. Lindle, J. R. Meyer, C. A. Hoffman, F. J. Bartoli, G. W. Turner, and H. K. Choi, Appl. Phys. Lett. **67**, 3153 (1995).

[30] J. E. L. Hollis, Proc. Phys. Soc. **91**, 151 (1967).

[31] S. Marchetti, M. Martinelli, and R. Simili, J. Phys.: Condens. Matter **14**, 3653 (2002).

[32] W. W. Bewley, J. R. Lindle, C. S. Kim, M. Kim, C. L. Canedy, I. Vurgaftman, and J. R. Meyer, Appl. Phys. Lett. **93**, 041118 (2008).

[33] M. Takeshima, Phys. Rev. B **25**, 5390 (1982).

[34] A. McAllister, D. Åberg, A. Schleife, and E. Kioupakis, Appl. Phys. Lett. **106**, 141901 (2015).

[35] K. J. Cheetham, A. Krier, I. P. Marko, A. Aldukhayel, and S. J. Sweeney, Appl. Phys. Lett. **99**, 141110 (2011).