First-principles calculations of indirect Auger recombination in nitride semiconductors

Emmanouil Kioupakis, Daniel Steiauf, Patrick Rinke, Kris T. Delaney, and Chris G. Van de Walle

First-principles calculations of indirect Auger recombination in nitride semiconductors

Emmanouil Kioupakis,[1, 2] Daniel Steiauf,[1] Patrick Rinke,[1, 3] Kris T. Delaney,[1] and Chris G. Van de Walle[1]

[1]*Materials Department, University of California, Santa Barbara, California 93106-5050, U.S.A.*
[2]*Department of Materials Science and Engineering,*
University of Michigan, Ann Arbor, Michigan 48109, U.S.A.
[3]*COMP/Department of Applied Physics, Aalto University, P.O. Box 11100, Aalto FI-00076, Finland*
(Dated: July 9, 2015)

Auger recombination is an important non-radiative carrier recombination mechanism in many classes of optoelectronic devices. The microscopic Auger processes can be either direct or indirect, mediated by an additional scattering mechanism such as the electron-phonon interaction and alloy disorder scattering. Indirect Auger recombination is particularly strong in nitride materials and affects the efficiency of nitride optoelectronic devices at high powers. Here we present a first-principles computational formalism for the study of direct and indirect Auger recombination in direct-band-gap semiconductors and apply it to the case of nitride materials. We show that direct Auger recombination is weak in the nitrides and cannot account for experimental measurements. On the other hand, carrier scattering by phonons and alloy disorder enables indirect Auger processes that can explain the observed loss in devices. We analyze the dominant phonon contributions to the Auger recombination rate and the influence of temperature and strain on the values of the Auger coefficients. Auger processes assisted by charged-defect scattering are much weaker than the phonon-assisted ones for realistic defect densities and not important for the device performance. The computational formalism is general and can be applied to the calculation of the Auger coefficient in other classes of optoelectronic materials.

PACS numbers: 72.20.Jv, 63.20.kd, 85.60.Bt, 85.60.Jb

I. INTRODUCTION

Auger recombination is an important non-radiative carrier recombination mechanism that is well recognized as a loss mechanism in optoelectronic devices such as light-emitting (LEDs) and laser diodes.[1] In the Auger process, the energy released through an electron-hole recombination event, approximately equal to the band gap of the material, is transferred via Coulomb scattering to a third free carrier that is excited to a higher-energy state [Fig. 1(a)]. The third carrier can be either an electron [in which case the process is called an electron-electron-hole (e-e-h) process] or a hole [hole-hole-electron (h-h-e)]. The overall energy and momentum of the carriers is conserved, and for non-degenerate carriers the recombination rate scales with the third power of the carrier density. As a result, it becomes an important carrier recombination mechanism at high carrier densities and reduces the efficiency of solid-state light emitters at high currents. In particular, Auger recombination has been shown to be involved in the efficiency reduction of nitride LEDs and lasers at high power.[2–13] The direct Auger recombination process [Fig. 1(a)] has been studied both analytically (e.g., Refs. 14–17) as well as with first-principles calculations.[18–22]

An additional type of Auger recombination mechanisms are those assisted by scattering due to phonons, defects, alloying, etc. [Fig. 1(b)]. The coupling of charge carriers to lattice vibrations in semiconductors gives rise to important quantum processes such as optical absorption in indirect-gap materials[23,24] and free-carrier absorption in semiconductors[25,26] and transparent conducting
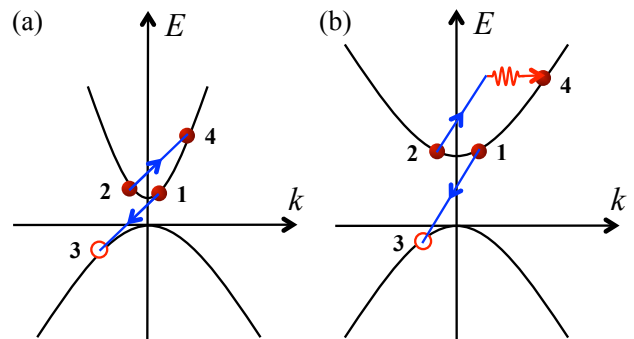

FIG. 1. (Color online) Schematic diagram of (a) direct and (b) indirect electron-electron-hole Auger recombination processes. (a) In the direct case, an electron in the conduction band (1) recombines with a hole in the valence band (3) while the excess energy and momentum is transferred to a second electron (2) that gets excited to a higher conduction-band state (4). (b) The indirect process is assisted by a carrier-scattering mechanism, such as electron-phonon coupling, alloy disorder, or defect scattering, which provides additional momentum and enables Auger transitions to a wider range of final conduction-band states throughout the first Brillouin zone.

oxides.[27] Phonon-assisted Auger processes have until recently only been treated with model calculations for the band structure and/or the electron-phonon coupling matrix elements, and/or only a fraction of all the relevant microscopic processes have been accounted for.[28–34] We have developed a set of first-principles computational approaches to study direct and indirect Auger recom-

bination in nitride materials. Some initial results of our work have been reported elsewhere.[35–37] Our calculations allowed us to conclude that Auger recombination, and in particular the indirect process assisted by electron-phonon coupling or alloy-disorder scattering, is especially strong in nitride materials. The calculated coefficient is sufficient to explain the efficiency droop in nitride light emitters. Moreover, we have found that phonon-mediated processes dominate Auger recombination in GaAs[37] and NaI.[38]

In the present paper, we present the full formalism and computational details for the first-principles calculation of direct and indirect Auger recombination in direct-band-gap semiconductors. Our computational formalism makes use of an array of first-principles tools, such as the maximally localized Wannier function method for the efficient interpolation of energy eigenvalues[39–42] and density functional perturbation theory[43,44] for the calculation of phonon frequencies and electron-phonon coupling matrix elements. We discuss the results we obtained for wurtzite nitride materials and we show that phonon- and alloy-scattering-assisted Auger recombination dominates in this technologically important class of wide-band-gap semiconductors. The approximations we developed to reduce the computational cost are also described.

The article is organized as follows. In Section II we present the relevant equations for the calculation of the direct and indirect Auger coefficients in semiconductors. In Section III we describe the computational formalism we implemented and the approximations we used for the calculations. In Section IV we present our results for nitride materials: In Subsection IV A we discuss our results for direct Auger recombination in GaN. In Subsection IV B we present our results for the alloy-scattering-assisted Auger recombination in $In_{0.25}Ga_{0.75}N$. In Subsection IV C we describe the results for the phonon-assisted Auger recombination in GaN, analyze the contributions by the various phonons, and discuss the effects of temperature and strain. In Subsection IV D we discuss charge-defect-assisted Auger recombination in GaN, and finally in Subsection IV E we present the cumulative values of the Auger recombination coefficients.

II. AUGER RECOMBINATION FORMALISM

A. Direct Auger recombination

The formalism for the calculation of the direct Auger recombination rate has been reported previously;[18,19,22,45] for completeness we provide a brief review here. The Auger recombination rate is calculated starting from Fermi's golden rule, which gives the transition probability per unit time in terms of matrix elements of the perturbation and the quasiparticle energies. For the case of electrons and holes in a solid, the Auger re-
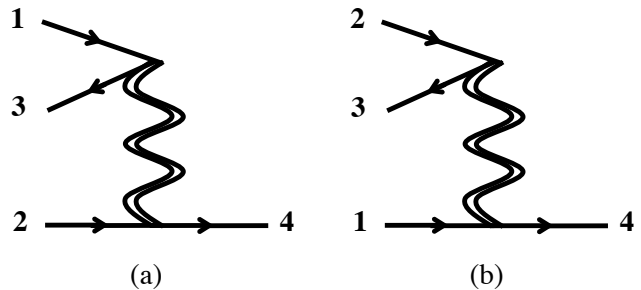

FIG. 2. Feyman diagrams for the (a) direct and (b) exchange terms of the Auger recombination process of Eq. (2). Diagram (a) corresponds to the direct process shown in Fig. 1(a).

combination rate is given by:

$$R = 2\frac{2\pi}{\hbar} \sum_{\mathbf{1234}} P|M_{\mathbf{1234}}|^2 \delta(\epsilon_{\mathbf{1}} + \epsilon_{\mathbf{2}} - \epsilon_{\mathbf{3}} - \epsilon_{\mathbf{4}}), \quad (1)$$

where the bold indices are composite band and $\mathbf{k}$-point indices $[\mathbf{1} \equiv (n_1, \mathbf{k}_1)]$, the factor of 2 accounts for spin, P is a statistics factor that accounts for the occupation numbers and ensures that transitions occur only from occupied to empty fermion states,

$$P = f_{\mathbf{1}}f_{\mathbf{2}}(1 - f_{\mathbf{3}})(1 - f_{\mathbf{4}}),$$

and f are free-carrier occupation numbers according to Fermi-Dirac statistics. The δ-function ensures energy conservation, while momentum conservation is imposed by the matrix elements.

The perturbing Hamiltonian is the screened Coulomb interaction between the carriers and involves two terms, the direct and the exchange one, that account for the antisymmetry of the many-body wave function under fermion exchange:[45]

$$|M_{\mathbf{1234}}|^2 \equiv |M_{\mathbf{1234}}^d - M_{\mathbf{1234}}^x|^2 + |M_{\mathbf{1234}}^d|^2 + |M_{\mathbf{1234}}^x|^2, \quad (2)$$

where the direct (M^d) and exchange (M^x) terms are given by matrix elements of the screened Coulomb interaction (W) between electron and hole wave functions (ψ):

$$M_{\mathbf{1234}}^d \equiv \langle \psi_{\mathbf{1}}\psi_{\mathbf{2}}|W|\psi_{\mathbf{3}}\psi_{\mathbf{4}}\rangle, \quad (3)$$
$$M_{\mathbf{1234}}^x \equiv \langle \psi_{\mathbf{1}}\psi_{\mathbf{2}}|W|\psi_{\mathbf{4}}\psi_{\mathbf{3}}\rangle, \quad (4)$$

and are shown in Figs. 2(a) and (b), respectively. The matrix elements of the screened Coulomb interaction are given by:

$$\langle \psi_{\mathbf{1}}\psi_{\mathbf{2}}|W|\psi_{\mathbf{3}}\psi_{\mathbf{4}}\rangle =$$
$$= \iint d\mathbf{r}_1 d\mathbf{r}_2 \psi_{\mathbf{1}}^*(\mathbf{r}_1)\psi_{\mathbf{2}}^*(\mathbf{r}_2)W(\mathbf{r}_1, \mathbf{r}_2)\psi_{\mathbf{3}}(\mathbf{r}_1)\psi_{\mathbf{4}}(\mathbf{r}_2)$$
$$= \frac{1}{V} \sum_{\mathbf{G}} \delta_{\mathbf{k}_1+\mathbf{k}_2, \mathbf{k}_3+\mathbf{k}_4+\mathbf{G}'} \tilde{W}(\mathbf{k}_1 - \mathbf{k}_3 + \mathbf{G})$$
$$\times I_{\mathbf{1,3}}(\mathbf{G})I_{\mathbf{2,4}}(\mathbf{G}' - \mathbf{G}), \quad (5)$$

where $\boldsymbol{G}'$ is the Umklapp vector that brings $\boldsymbol{k}_4 = \boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3 + \boldsymbol{G}'$ back into the first Brillouin zone. Ignoring local-field effects,[46] the Fourier-transformed Coulomb interaction $(\tilde{W})$ is given by:

$$W(\boldsymbol{r}_1, \boldsymbol{r}_2) = W(\boldsymbol{r}_1 - \boldsymbol{r}_2) \equiv \frac{1}{V} \sum_{\boldsymbol{q}} \tilde{W}(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot(\boldsymbol{r}_1 - \boldsymbol{r}_2)}$$

$$= \frac{1}{V} \sum_{\boldsymbol{q}} \frac{1}{\varepsilon(\boldsymbol{q})} \frac{4\pi e^2}{q^2 + \lambda^2} e^{i\boldsymbol{q}\cdot(\boldsymbol{r}_1 - \boldsymbol{r}_2)} \quad (6)$$

where $\varepsilon(\boldsymbol{q})$ is the static dielectric function of the material and λ is the inverse screening length due to the free carriers. The overlap integrals, I, are given by:

$$I_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{G}) \equiv \sum_{\boldsymbol{G}_1} c_{\boldsymbol{\alpha}}^*(\boldsymbol{G}_1) c_{\boldsymbol{\beta}}(\boldsymbol{G}_1 - \boldsymbol{G})$$

$$= \int_{\text{cell}} u_{\boldsymbol{\alpha}}^*(\boldsymbol{r}) u_{\boldsymbol{\beta}}(\boldsymbol{r}) e^{i\boldsymbol{G}\cdot\boldsymbol{r}} d\boldsymbol{r}, \quad (7)$$

where c are the plane-wave components of the lattice-periodic part of the Bloch functions u.

The Auger rate is then calculated using:

$$\frac{dn}{dt} = \frac{R}{V} = \frac{4\pi}{\hbar} \frac{1}{V_{\text{cell}}^3} \frac{1}{N_{\boldsymbol{k}}^3} \sum_{\boldsymbol{123}n_4} f_{\boldsymbol{1}} f_{\boldsymbol{2}} (1 - f_{\boldsymbol{3}})(1 - f_{\boldsymbol{4}})$$

$$\times |V M_{\boldsymbol{1234}}|^2 \delta(\epsilon_{\boldsymbol{1}} + \epsilon_{\boldsymbol{2}} - \epsilon_{\boldsymbol{3}} - \epsilon_{\boldsymbol{4}}), \quad (8)$$

where $\boldsymbol{4} \equiv (n_4, \boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3 + \boldsymbol{G}')$. In the remainder of this paper, we assume an equal density of free electrons and holes $(n = p)$, which is usually the case in optoelectronic devices for high injected carrier densities and under steady-state conditions. For non-degenerate carriers, which can be described with Boltzmann statistics, the Auger recombination rate scales with the third power of the carrier density.[45] At high carrier concentrations, however, the carriers become degenerate and need to be described by Fermi-Dirac statistics. In this case, the exponent of the power law can take values lower than 3, an effect known as phase-space filling.[47] In the following, we will calculate density-dependent Auger coefficients $C = C(n)$ defined by:

$$C(n) \equiv \frac{R(n)}{n^3 V}. $$

B. Indirect Auger recombination

In this Subsection, we present the complete formalism for the calculation of the indirect recombination rate. Various aspects of this formalism have been presented before in Refs. 28–33. The transition probability per unit time $(\mathscr{P})$ for indirect Auger processes from an initial state (I) to a final state (F) via an intermediate state (M) is determined based on second-order Fermi's golden rule:[48]

$$\mathscr{P}_{\text{I}\to\text{F}} = \frac{2\pi}{\hbar} \left| \sum_{\text{M}} \frac{\mathscr{H}_{\text{IM}} \mathscr{H}_{\text{MF}}}{E_{\text{M}} - E_{\text{I}}} \right|^2 \delta(E_{\text{F}} - E_{\text{I}})$$

where the perturbation Hamiltonian $(\mathscr{H})$ is the combination of the screened Coulomb interaction plus the scattering mechanism (e.g., the electron-phonon interaction), and E_{I}, E_{F}, and E_{M} are the energies of the initial, final, and intermediate states.

By summing over all initial and final states, we obtain an expression for the recombination rate similar to the expression for direct Auger recombination [Eq. (1)]. The equation for phonon-assisted Auger recombination is:

$$R = 2\frac{2\pi}{\hbar} \sum_{\boldsymbol{1234},\nu\boldsymbol{q}} \tilde{P} |\tilde{M}_{\boldsymbol{1234};\nu\boldsymbol{q}}|^2 \delta(\epsilon_{\boldsymbol{1}} + \epsilon_{\boldsymbol{2}} - \epsilon_{\boldsymbol{3}} - \epsilon_{\boldsymbol{4}} \mp \hbar\omega_{\nu\boldsymbol{q}}),$$
$$(9)$$

where ν is the phonon mode and $\boldsymbol{q}$ is the phonon wave vector. The upper (lower) sign corresponds to the phonon-emission (absorption) process, respectively. The difference with direct Auger recombination is that the phonon energy $(\hbar\omega_{\nu\boldsymbol{q}})$ enters the energy-conserving δ-function (the phonon momentum $\hbar\boldsymbol{q}$ appears in the momentum-conserving matrix elements later). The statistics factor P is modified to account for the additional phonon absorption or emission process,

$$\tilde{P} = f_{\boldsymbol{1}} f_{\boldsymbol{2}} (1 - f_{\boldsymbol{3}})(1 - f_{\boldsymbol{4}}) \left(n_{\nu\boldsymbol{q}} + \frac{1}{2} \pm \frac{1}{2} \right),$$

where $n_{\nu\boldsymbol{q}}$ are the phonon occupation numbers given by Bose-Einstein statistics:

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

and the matrix element has been generalized to include the scattering to the virtual intermediate state.

There are eight diagrams that enter the calculation of the generalized matrix element that correspond to all possible combinations of the electron-phonon and screened-Coulomb interaction Hamiltonians (direct and exchange) for both parallel and antiparallel initial carrier spin configurations. These diagrams are summed in groups and then squared, depending on whether they correspond to distinguishable or indistinguishable quantum processes. The generalized matrix element $(\tilde{M})$ is given by:

$$|\tilde{M}|^2 = \left| \tilde{M}^1 + \tilde{M}^2 + \tilde{M}^3 + \tilde{M}^4 \right.$$
$$\left. - \tilde{M}^5 - \tilde{M}^6 - \tilde{M}^7 - \tilde{M}^8 \right|^2 +$$
$$+ \left| \tilde{M}^1 + \tilde{M}^2 + \tilde{M}^3 + \tilde{M}^4 \right|^2 +$$
$$+ \left| \tilde{M}^5 + \tilde{M}^6 + \tilde{M}^7 + \tilde{M}^8 \right|^2, \quad (10)$$

where each of the $\tilde{M}^1$ to $\tilde{M}^8$ terms corresponds to each

of the phonon-assisted process depicted in Figs. 3(a)–(h):

$$\tilde{M}^1_{\mathbf{1234};\nu q} = \sum_m \frac{g_{\mathbf{1}m;\nu} M^d_{m\mathbf{234}}}{\epsilon_m - \epsilon_{\mathbf{1}} \pm \hbar\omega_{\nu q} + i\eta} \tag{11}$$

$$\tilde{M}^2_{\mathbf{1234};\nu q} = \sum_m \frac{g_{\mathbf{2}m;\nu} M^d_{\mathbf{1}m\mathbf{34}}}{\epsilon_m - \epsilon_{\mathbf{2}} \pm \hbar\omega_{\nu q} + i\eta} \tag{12}$$

$$\tilde{M}^3_{\mathbf{1234};\nu q} = \sum_m \frac{M^d_{\mathbf{12}m\mathbf{4}} g_{m\mathbf{3};\nu}}{\epsilon_m - \epsilon_{\mathbf{3}} \mp \hbar\omega_{\nu q} + i\eta} \tag{13}$$

$$\tilde{M}^4_{\mathbf{1234};\nu q} = \sum_m \frac{M^d_{\mathbf{123}m} g_{m\mathbf{4};\nu}}{\epsilon_m - \epsilon_{\mathbf{4}} \mp \hbar\omega_{\nu q} + i\eta} \tag{14}$$

$$\tilde{M}^5_{\mathbf{1234};\nu q} = \sum_m \frac{g_{\mathbf{1}m;\nu} M^x_{m\mathbf{234}}}{\epsilon_m - \epsilon_{\mathbf{1}} \pm \hbar\omega_{\nu q} + i\eta} \tag{15}$$

$$\tilde{M}^6_{\mathbf{1234};\nu q} = \sum_m \frac{g_{\mathbf{2}m;\nu} M^x_{\mathbf{1}m\mathbf{34}}}{\epsilon_m - \epsilon_{\mathbf{2}} \pm \hbar\omega_{\nu q} + i\eta} \tag{16}$$

$$\tilde{M}^7_{\mathbf{1234};\nu q} = \sum_m \frac{M^x_{\mathbf{12}m\mathbf{4}} g_{m\mathbf{3};\nu}}{\epsilon_m - \epsilon_{\mathbf{3}} \mp \hbar\omega_{\nu q} + i\eta} \tag{17}$$

$$\tilde{M}^8_{\mathbf{1234};\nu q} = \sum_m \frac{M^x_{\mathbf{123}m} g_{m\mathbf{4};\nu}}{\epsilon_m - \epsilon_{\mathbf{4}} \mp \hbar\omega_{\nu q} + i\eta}, \tag{18}$$

and η is the inverse of the lifetime of the intermediate state. For example, the process in Fig. 1(b) is given by the $\tilde{M}^4$ term [Fig. 3(d)]. We made sure to evaluate all terms using the same set of calculated wave functions in order to preserve the correct phase information and ensure that the cross terms among the various paths are properly taken into account.

The electron-phonon coupling matrix elements (g) are given in terms of the electron and hole wave functions and the derivative $\partial_{\nu q}V$ of the self-consistent potential due to a collective ionic displacement by phonon mode νq:[49]

$$g_{n\mathbf{k},m\mathbf{k+q};\nu} = \left(\frac{\hbar}{2M_0\omega_{\nu q}}\right)^{1/2} \langle \psi_{n\mathbf{k}} | (\partial_{\nu q}V)^* | \psi_{m\mathbf{k+q}} \rangle,$$

where M_0 is the total mass of the atoms in the unit cell.

III. COMPUTATIONAL FORMALISM

Calculations for the electronic energies and wave functions were performed using density functional theory (DFT)[50,51] within the local-density approximation for the exchange-correlation potential[52,53] and the plane-wave pseudopotential method.[44,54] The semicore $3d$ electrons of Ga and $4d$ electrons of In are treated as valence electrons to get accurate electronic and structural properties.[55] Effects due to the spin-orbit interaction, which are small in nitride semiconductors, were not included in the calculations. After a self-consistent calculation was performed to obtain the charge density, we determined the maximally-localized Wannier functions[39–41] which we subsequently used to interpolate the band
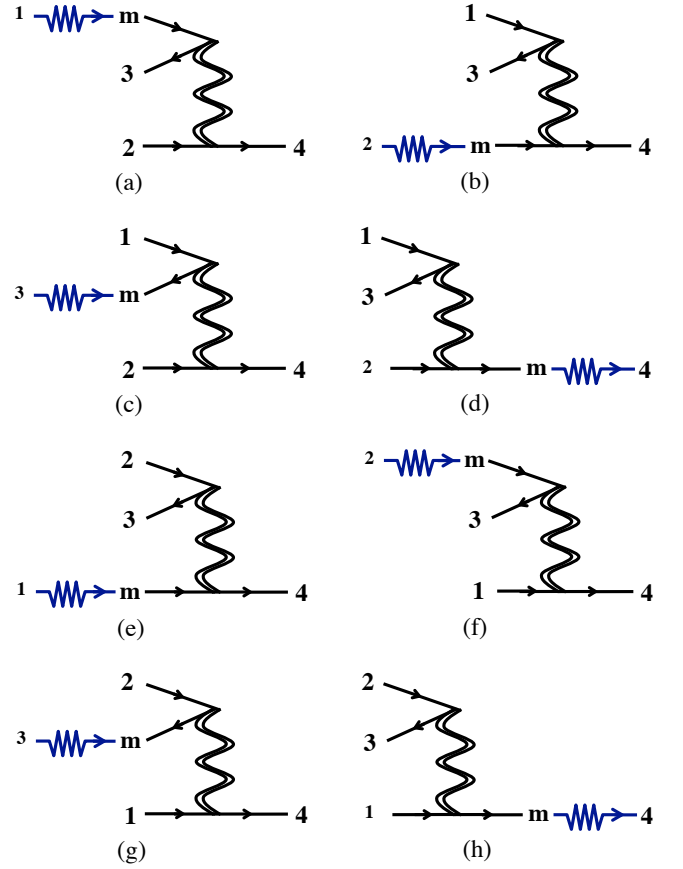

FIG. 3. (Color online) (a)–(h) Schematic diagrams corresponding to each of the microscopic indirect Auger recombination processes of Eq. (10) ($\tilde{M}^1$–$\tilde{M}^8$, respectively).

structure from coarse $\mathbf{k}$-grids ($8\times8\times8$ for the GaN calculations in Sec. IV A and $4\times4\times4$ for the $In_{0.25}Ga_{0.75}N$ calculations described in Sec. IV B) to arbitrarily fine $\mathbf{k}$-grids (as fine as $64\times64\times32$ for GaN and $24\times24\times12$ for $In_{0.25}Ga_{0.75}N$) in the first Brillouin zone. This enables the efficient determination of the band energies near the band extrema, which we subsequently use to determine the free-carrier quasi-Fermi levels, without the need to perform costly non-self-consistent DFT calculations for these fine $\mathbf{k}$-grids. The band structure is not significantly affected by the addition of the free carriers for carrier densities that are relevant for most devices (up to 10^{20} cm^{-3}). The delta functions in the rate equations are approximated by finite-width Gaussians. All $\mathbf{k}$ grids mentioned in this paper are unshifted (i.e., they include the Γ point). We expect that shifted $\mathbf{k}$ grids yield similar results provided they are sufficiently converged.

To describe the distribution of the free carriers over states near the band extrema, we interpolated the bands to fine $\mathbf{k}$-point meshes in the Brillouin zone and determined the quasi-Fermi levels as a function of the free-carrier density and temperature. We then applied a cutoff criterion to limit the number of $\mathbf{k}$-points used in subsequent Auger rate calculations. The criterion was to

keep only $\boldsymbol{k}$-points which correspond to band energies within an energy cutoff from the corresponding band edge, $E_{\text{cutoff}} = \epsilon_{\text{F}} + MkT$, where ϵ_{F} is the free-carrier quasi-Fermi level and M is an integer. The integer cutoff parameter M is determined such that the free-carrier density obtained after applying the cutoff criterion differs by less than 1% from its starting value. This ensures that most of the free charge carriers are accounted for when calculating the Auger rates. We repeat this procedure to converge the Auger coefficients with respect to the $\boldsymbol{k}$-grid spacing.

To further reduce the number of $\boldsymbol{k}$-points used in Auger rate calculations, the sums over hole states are sampled with coarser grids that are half as dense along each linear dimension as the corresponding electron sums. This choice relies on the fact that holes in the nitrides have a larger effective mass than electrons and thus occupy a larger $\boldsymbol{k}$-space volume around Γ that can be effectively sampled with coarser grids. Since we eventually converge our calculations with respect to the grid spacing, the choice of sparser hole grids speeds up the calculations without loss of accuracy.

For the calculation of the matrix elements of the screened Coulomb interaction [Eq. (5)] we used a model dielectric function:[56]

$$\varepsilon(q) = 1 + \{[\varepsilon_\infty - 1]^{-1} + \alpha(q/q_{\text{TF}})^2 + \hbar^2 q^4/(4m^2\omega_p^2)\}^{-1},$$

where ε_∞ is the dielectric constant due to electronic screening, q_{TF} is the Thomas-Fermi wave vector, ω_p is the plasma frequency, and $\alpha = 1.563$ is an empirical parameter. This model has been shown to accurately reproduce the first-principles dielectric function of zinc-blende GaN calculated in the random-phase-approximation.[57] For the wave functions used to calculate the matrix elements we employed a lower plane-wave cut-off energy (50 Ry) than the self-consistent field calculations (90 Ry) to speed up the calculation. The choice of this lower plane-wave cut-off energy affects the calculated Auger rates by at most 10%. For the dielectric constant of GaN needed by the model, we used the directionally averaged experimental dielectric constant, $\varepsilon_\infty = (\varepsilon_{\infty\parallel}\varepsilon_{\infty\perp}^2)^{1/3} = 5.5$,[58,59] where $\varepsilon_\parallel$ and $\varepsilon_\perp$ are the dielectric constants for polarization parallel and perpendicular to the c-axis of the wurtzite structure, respectively. The dielectric constant of $In_{0.25}Ga_{0.75}N$ ($\varepsilon_\infty = 7.0$) was estimated with a model extrapolation[60] to a band gap of 2.4 eV, the value of the 25% InGaN alloy as predicted from hybrid-density-functional calculations.[61]

For the free-carrier screening wave vector of electrons λ_{e} we used the Debye-Hückel equation for non-degenerate carriers ($\lambda_{\text{e}}^2 = 4\pi n e^2/\epsilon_\infty k_{\text{B}}T$) if the quasi-Fermi energy of electrons referenced to the energy of the conduction-band minimum (E_{CBM}) is $\epsilon_{\text{F}} - E_{\text{CBM}} < 0$ or $0 \leq \epsilon_{\text{F}} - E_{\text{CBM}} < \frac{3}{2}k_{\text{B}}T$, and the Thomas-Fermi equation for degenerate carriers [$\lambda_{\text{e}}^2 = 6\pi n e^2/\epsilon_\infty(\epsilon_{\text{F}} - E_{\text{CBM}})$] if $\epsilon_{\text{F}} - E_{\text{CBM}} \geq \frac{3}{2}k_{\text{B}}T$. We use similar equations for the screening wave vector of holes λ_{h} and sum the contributions by electrons and holes to obtain the total free-carrier screening wave vector $\lambda^2 = \lambda_{\text{e}}^2 + \lambda_{\text{h}}^2$.

For the calculation of the phonon-assisted Auger coefficients, we need to make approximations in the way we calculate the indirect Auger rate. The technical challenge is that the full calculation of Eq. (9) that accounts for all possible initial and final electronic states is computationally very expensive. The high computational cost arises because of the fine grids needed to sample the free-carrier distributions in reciprocal space and the large number of associated phonon calculations. The approximation we make is based on the fact that the free carriers are confined in reciprocal space to the vicinity of the band extrema near the Γ point of the Brillouin zone. As a result, the wave functions of the initial electron and hole states can be effectively approximated by the conduction- and valence-band wave functions at the Brillouin-zone center. We will later show that this is a valid approximation for the case of wide-band-gap nitrides.

Within the approximation that the initial-state wave functions are those at the Γ point, the sums over indices **1** to **3** yield $N/2$, where N is the total number of free electrons or holes and the factor of 2 arises because we already accounted for the carrier spin in the matrix elements. State **4** is initially empty since it is located at an energy on the order of the band gap (i.e., several eV) away from the corresponding band edge. In the case of the e–e–h process, for example, the expression we need to calculate reduces to the form:

$$R = 2\frac{2\pi}{\hbar}\frac{N^3}{8}\sum_{n\nu\boldsymbol{q}}\left(n_{\nu\boldsymbol{q}} + \frac{1}{2} \pm \frac{1}{2}\right)|\tilde{M}_{c\Gamma,c\Gamma,v\Gamma,n\boldsymbol{q};\nu\boldsymbol{q}}|^2$$
$$\times\,\delta(\epsilon_{c\Gamma} + \epsilon_{c\Gamma} - \epsilon_{v\Gamma} - \epsilon_{n\boldsymbol{q}} \mp \hbar\omega_{\nu\boldsymbol{q}}),$$
(19)

which involves only a single sum over the phonon wave vectors and is computationally feasible. A similar expression can be derived for the h–h–e case:

$$R = 2\frac{2\pi}{\hbar}\frac{N^3}{8}\sum_{n\nu\boldsymbol{q}}\left(n_{\nu\boldsymbol{q}} + \frac{1}{2} \pm \frac{1}{2}\right)|\tilde{M}_{v_1\Gamma,v_2\Gamma,c\Gamma,n\boldsymbol{q};\nu\boldsymbol{q}}|^2$$
$$\times\,\delta(-\epsilon_{v_1\Gamma} - \epsilon_{v_2\Gamma} + \epsilon_{c\Gamma} + \epsilon_{n\boldsymbol{q}} \mp \hbar\omega_{\nu\boldsymbol{q}}),$$
(20)

where v_1 and v_2 are hole band indices. The eight components of the generalized matrix element in Eq. (19) are given by:

$$\tilde{M}^1_{c\Gamma,c\Gamma,v\Gamma,n\boldsymbol{q};\nu\boldsymbol{q}} = \sum_m \frac{g_{c\Gamma,m\boldsymbol{q};\nu}M^d_{m\boldsymbol{q},c\Gamma;v\Gamma,n\boldsymbol{q}}}{\epsilon_{m\boldsymbol{q}} - \epsilon_{c\Gamma} \pm \hbar\omega_{\nu\boldsymbol{q}} + i\eta}. \quad (21)$$

Similar expressions hold for $\tilde{M}^2$ to $\tilde{M}^8$. Since we are neglecting the spin-orbit interaction, the heavy- and light-hole bands are degenerate at Γ, and we average their contribution in the evaluation of Eqs. (19) and (20). The phonon frequencies and electron-phonon coupling matrix elements were calculated using density functional perturbation theory[43,44] for $\boldsymbol{k}$-points on a $24\times24\times12$ grid that fall in the irreducible part of the first Brillouin zone of GaN.

IV. RESULTS AND DISCUSSION

A. Direct Auger recombination in GaN

We performed calculations for the direct Auger coefficient in bulk GaN and show that it cannot account for the efficiency reduction of nitride LEDs. In our calculations of Auger rates, we treat the band gap of GaN as an adjustable parameter and rigidly shift the conduction bands with a scissors operator. This serves to correct the band-gap problem of density functional theory and to model the effect of alloying GaN with InN to form $In_xGa_{1-x}N$ alloys. By varying the band-gap value, the third carrier that participates in the Auger process is excited to a different range of higher-energy states for different gap values and this leads to a variation of the Auger rate with respect to the adjusted band gap. The screened-Coulomb-interaction matrix elements are calculated using the wave functions of GaN. The Auger coefficients due to direct e–e–h and h–h–e Auger recombination processes are plotted as a function of the adjusted band gap in Figs. 4(a) and 4(b), respectively. The values of the direct e–e–h and h–h–e Auger coefficients for pure GaN are simply given by the value of the Auger coefficients for the experimental band-gap value of GaN (3.5 eV[62]). The broadening of the delta function in Eq. (1) is set to 0.1 eV and the free-carrier densities and temperatures to $n = p = 10^{19}$ cm^{-3} and 300 K, respectively. The different curves correspond to different grids that we employed to sample the part of the Brillouin zone occupied by free carriers, illustrating the convergence with respect to the grid spacing.

The data indicate that both e–e–h and h–h–e direct Auger processes are very weak for GaN (less than 10^{-35} cm^{6}s^{-1} for $E_{gap} = 3.5$ eV) and the corresponding coefficients cannot account for the experimentally measured values in nitride devices (10^{-31}–10^{-30} cm^{6}s^{-1}).[3–11] The coefficients increase drastically for decreasing band-gap values in the range of 2.4–3.5 eV, which is the typical band-gap range for nitride devices. However, even the larger Auger coefficients at lower band-gap values cannot explain the experimental measurements. The coefficient of the h–h–e process increases exponentially for decreasing adjusted band gap and reaches a value of 3.5×10^{-33} cm^{6}s^{-1} for band-gap values in the green ($E_{gap} = 2.4$ eV). The e–e–h process increases for decreasing band gap and displays a peak for a band-gap value of 2.3 eV due to interband Auger transitions to the second conduction band. The magnitude of the Auger coefficient at this peak (5.7×10^{-33} cm^{6}s^{-1}), however, is too small to account for the experimental observations. This result is different from earlier calculations[20] that used a different approach to model the Auger process. The discrepancy has been traced back to the inadvertent omission of a normalization factor and the treatment of the long-range part of the Coulomb interaction in Ref. 20. Our present results for the direct Auger process are in qualitative agreement with the work of Bertazzi et al.[63] who
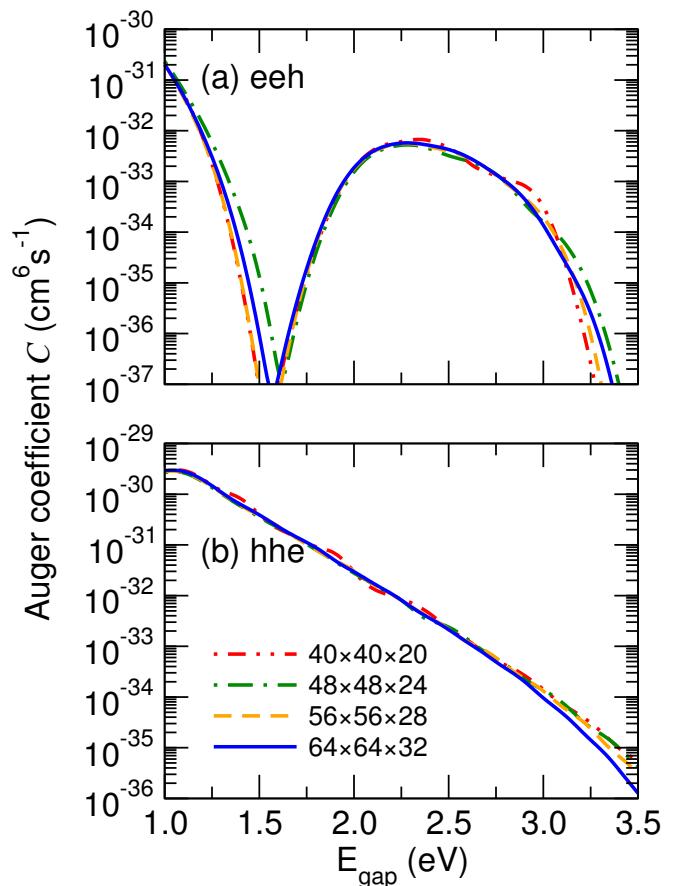

FIG. 4. (Color online) Auger coefficients of GaN for a carrier density of 10^{19} cm^{-3} and a temperature of 300 K due to (a) electron-electron-hole (e-e-h) and (b) hole-hole-electron (h-h-e) Auger processes plotted as a function of the scissors-shift-adjusted band gap for various k-grid sampling densities.

also observed a peak of the e–e–h Auger coefficient, but at a different energy ($E_{gap} = 2.9$ eV). We attribute the quantitative differences between our work and that of Bertazzi et al.[63] to the different band structures (fully first-principles in pure GaN versus fitted pseudopotentials for virtual-crystal InGaN) and dielectric-function models used in the respective calculations. Finally, in the 1.0–1.5 eV energy range the e–e–h coefficient increases exponentially for decreasing band-gap values, as expected for direct Auger recombination,[64] because only direct intraband Auger transitions are possible for this band-gap range.

We also examined the dependence of the direct Auger coefficients on the free-carrier density. Figures 5(a) and 5(b) show the dependence of the e–e–h and h–h–e direct Auger processes on the free-carrier density. The dependence of the Auger coefficient on the carrier density is not monotonic as a function of the adjusted band gap. For small adjusted band gaps (less than 2.0 eV) the h–h–e Auger coefficient is reduced for increasing carrier density. This is due to phase-space filling effects.[47] Free carriers at low densities are nondegenerate and this results in a cu-

FIG. 6. (Color online) Structure of the quasirandom alloy cell used for the calculation of the Auger coefficient in $In_{0.25}Ga_{0.75}N$ (coordinates from Ref. 67).
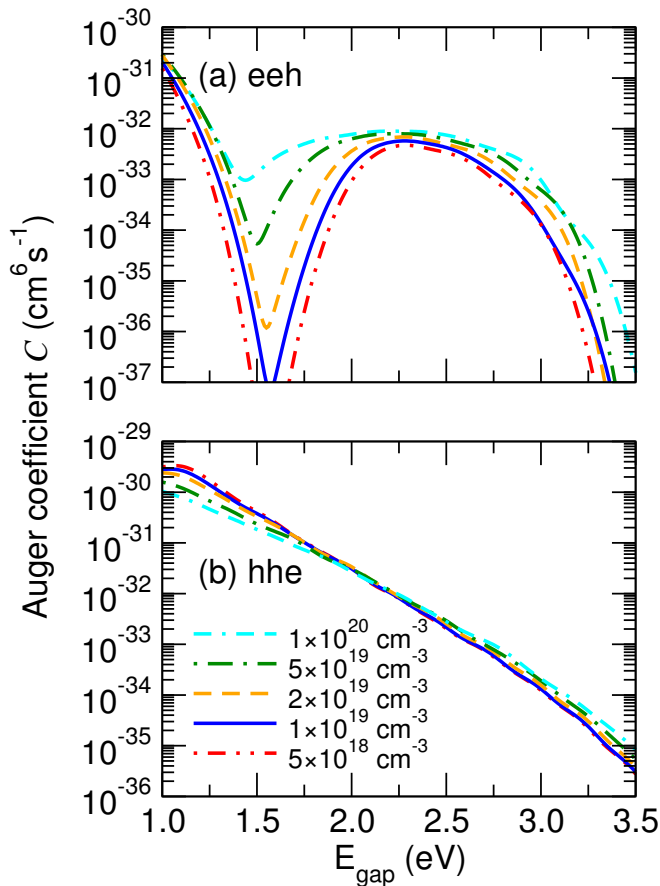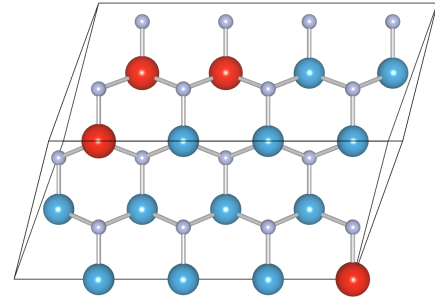
FIG. 5. (Color online) Auger coefficients of GaN at 300 K due to (a) electron-electron-hole (e-e-h) and (b) hole-hole-electron (h-h-e) Auger recombination processes plotted as a function of the scissors-shift-adjusted band gap for various free-carrier densities.

B. Alloy-scattering-assisted Auger recombination in $In_{0.25}Ga_{0.75}N$

Alloying GaN with InN forms $In_xGa_{1-x}N$ alloys with band gaps desirable for optoelectronic applications. In Sec. IV A we modeled the resulting variation in Auger rates by changing the band gap. However, alloying also introduces an additional carrier scattering mechanism due to the breaking of the translational periodicity by the substitution of Ga atoms with In. As a result, the momentum-conservation requirement of Eq. (5) imposed by the crystal periodicity is relaxed for the alloy. Consequently, Auger transitions that are forbidden in the perfect crystal because of the translational symmetry are allowed in the alloy as a result of the substitutional disorder. The corresponding matrix elements acquire nonzero values. By performing calculations for an InGaN alloy supercell, which explicitly contains the effect of alloy disorder, we show that alloying enables additional Auger recombination processes that enhance the Auger coefficient.

Previous calculations of Auger recombination in alloys used the virtual-crystal approximation to model the effect of alloying. However, the virtual crystal is a perfectly periodic arrangement of atoms that misses important aspects of the actual alloy structure, such as substitutional disorder and atomic relaxations. The importance of these two factors has been highlighted, for instance, in the case of $TiO_{2(a-x)}S_{2x}$, where they break the symmetry of optical matrix elements and enhance interband optical absorption,[65] and also in the case of free-carrier absorption in $In_xGa_{1-x}N$.[25,26] In order to capture these effects in our simulations, we have to explicitly study Auger recombination in an alloy supercell that includes both kinds of alloy atoms. A computationally feasible way to do this is with the 32-atoms special quasirandom alloy structure[66] for the 25% alloy composition (Fig. 6).[67] This particular supercell has the special property that it is the optimal 32-atom supercell that reproduces the short-range correlation function of the random alloy.

The scattering of free carriers by the alloy disorder is found to be important for Auger recombination and sig-
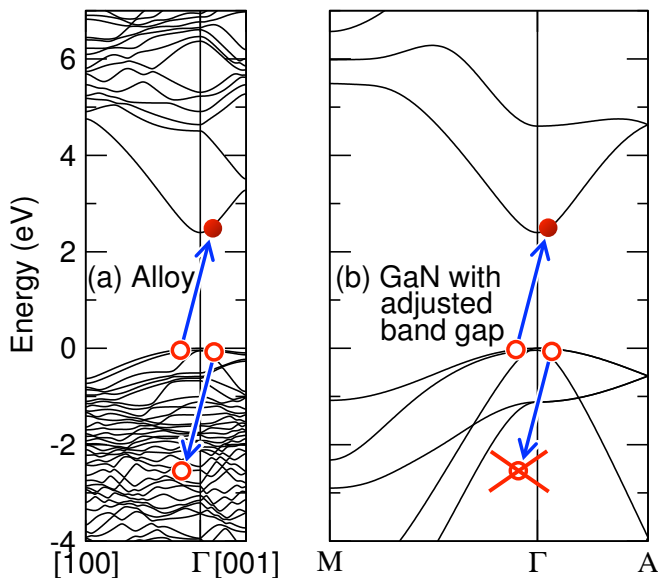
bic carrier-dependence of the Auger rate. For degenerate carrier densities, however, the Boltzmann approximation breaks down and carriers need to be described by the Fermi-Dirac distribution, which reduces the cubic density dependence of the Auger rate. In other words, the Auger coefficient becomes density-dependent at high carrier densities. The Auger coefficient increases as a function of carrier density for larger vales of the adjusted band gap for h–h–e and for all adjusted band-gap values for e–e–h. This is because direct Auger recombination for large values of the adjusted band gap is inhibited by the limited density of states available for momentum-conserving Auger transitions. However, as the carrier density increases the free carriers occupy a more extended region of the Brillouin zone and more phase-space becomes available for Auger transitions.

FIG. 7. (Color online) Scissors-shift-adjusted band structure of $In_{0.25}Ga_{0.75}N$ calculated using (a) the quasirandom alloy structure and (b) bulk GaN with an adjusted band gap to match that of the alloy. Alloying enables additional Auger transitions, such as the illustrated h–h–e process, that are not possible without the zone folding and symmetry breaking introduced by alloying.

nificantly increases the Auger coefficients compared to pure GaN. Figure 7(a) shows the scissors-shift-adjusted band structure of $In_{0.25}Ga_{0.75}N$ calculated for the 32-atoms alloy supercell shown in Fig. 6, while Figure 7(b) shows the band structure of GaN with a band gap adjusted to the gap of $In_{0.25}Ga_{0.75}N$ ($E_{gap} = 2.4$ eV). The two figures show that the disorder introduced by alloying breaks the symmetry and folds the band structure, which in turn enables additional Auger transitions and enhances the Auger coefficient. These transitions cannot be modeled within the virtual-crystal approximation because the latter cannot describe the folded bands and the associated symmetry breaking. Our first-principles calculations take into account the short-range alloy disorder and incorporate the effect of alloy scattering on the Auger coefficients.

Once the additional scattering processes are taken into account, the dependence of the Auger rates on the band gap can again be examined by plotting results as a function of a scissors-shift-adjusted band gap, keeping the matrix elements fixed to those of the $In_{0.25}Ga_{0.75}N$ alloy. Figures 8(a) and (b) illustrate the dependence of the e–e–h and h–h–e alloy-scattering-assisted Auger coefficients on the scissors-shift-adjusted band gap, as well as the convergence with respect to the Brillouin-zone-sampling grid density. The broadening of the delta function in Eq. (1) has been set to 0.3 eV, the free-carrier densities are $n = p = 10^{19}$ cm^{-3}, and the temperature is 300 K. The coefficients are on the order of 3×10^{-31} cm^{6}s^{-1} for the adjusted gap value that matches the 25%-alloy band
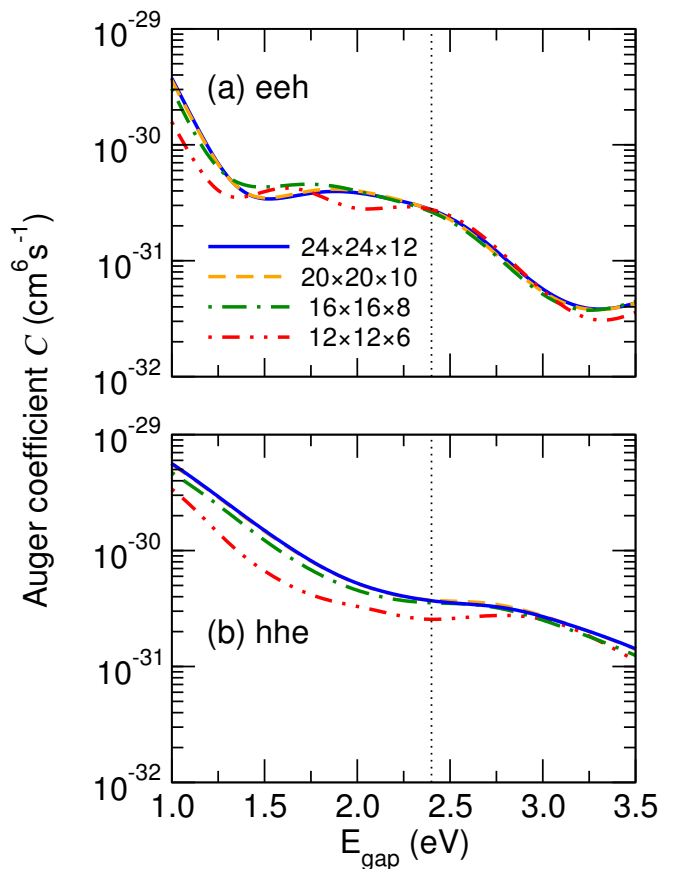

FIG. 8. (Color online) Auger coefficients of $In_{0.25}Ga_{0.75}N$ for a carrier density of 10^{19} cm^{-3} and a temperature of 300 K due to (a) electron-electron-hole (e-e-h) and (b) hole-hole-electron (h-h-e) Auger processes, plotted as a function of the band gap for various Brillouin-zone sampling densities. The band gap is treated as an adjustable variable in order to model a varying alloy composition. The variation of the alloy scattering potential with alloy composition on these Auger coefficients has not been included in this Figure, but it is examined in Subsection IV E.

gap ($E_{gap} = 2.4$ eV), much larger than the corresponding values for GaN.

The dependence of the alloy-assisted Auger coefficients on the free-carrier density is shown in Figs. 9(a) and (b). Both the e–e–h and the h–h–e coefficients decrease for increasing densities for every value of the adjusted band gap. This implies that for the alloy-assisted Auger case, the effects of phase-space filling are more important than the enabling of additional Auger transitions for higher carrier densities. This is expected, since bands farther from the zone center have already been made accessible for direct Auger transitions by the zone folding due to alloying.

The density dependence of the alloy-scattering-assisted Auger coefficients of $In_{0.25}Ga_{0.75}N$ at 300 K for an adjusted band-gap value of 2.4 eV is plotted as a function of carrier concentration in Fig. 10. The values of the Auger coefficients decrease linearly for increasing carrier
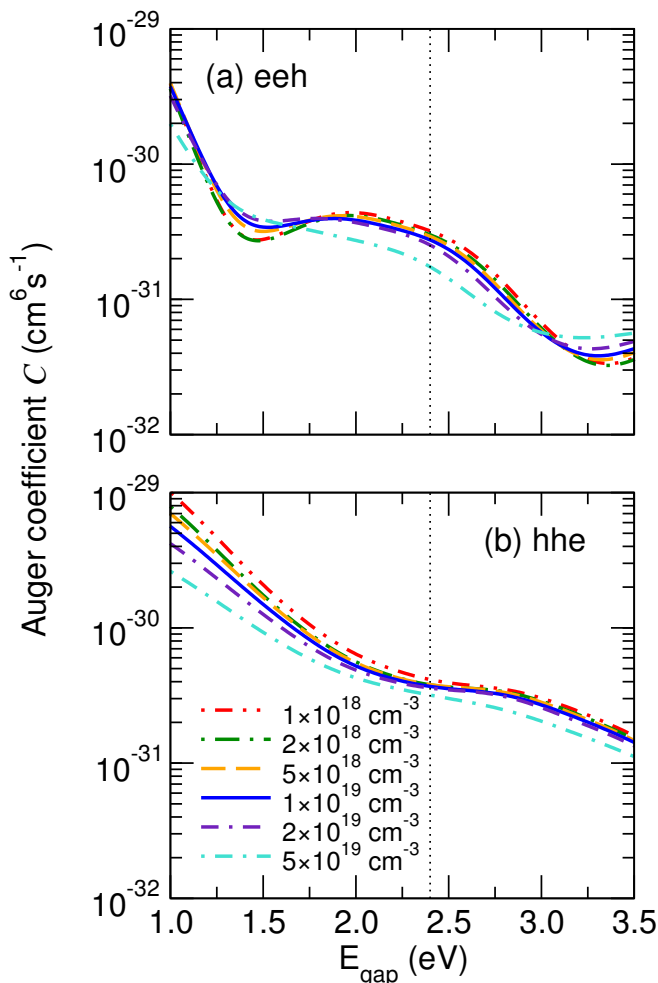
FIG. 9. (Color online) Alloy-assisted Auger coefficients of $In_{0.25}Ga_{0.75}N$ at 300 K due to (a) electron-electron-hole (e-e-h) and (b) hole-hole-electron (h-h-e) Auger recombination processes, plotted as a function of the scissors-shift-adjusted band gap for various free-carrier densities. The dotted line corresponds to the band gap of $In_{0.25}Ga_{0.75}N$ (2.4 eV) as predicted from hybrid functional calculations.[61]

FIG. 10. (Color online) Carrier-density dependence of the alloy-scattering-assisted Auger coefficients of $In_{0.25}Ga_{0.75}N$ at 300 K. The coefficients decrease for increasing carrier densities due to the transition from nondegenerate to degenerate carrier statistics (phase-space filling).

densities in the $2 \times 10^{18} - 5 \times 10^{19}$ cm^{-3} range, which is the typical range of free-carrier concentrations under LED operating conditions. The decrease of the Auger coefficients with increasing density is due to the transition from nondegenerate to degenerate carrier statistics (phase-space filling) in this carrier-density regime and to the increasing screening of the Coulomb interaction matrix elements by free carriers. We fit the first-principles data in the $2 \times 10^{18} - 5 \times 10^{19}$ cm^{-3} carrier-concentration range with a linear equation of the form

$$C(n)/10^{-31}\text{cm}^6\text{s}^{-1} = a + bn/10^{19}\text{cm}^{-3}, \qquad (22)$$

where a and b are dimensionless fitting parameters with values $a = 3.1$, $b = -0.26$ for e–e–h and $a = 3.9$, $b = -0.13$ for h–h–e.

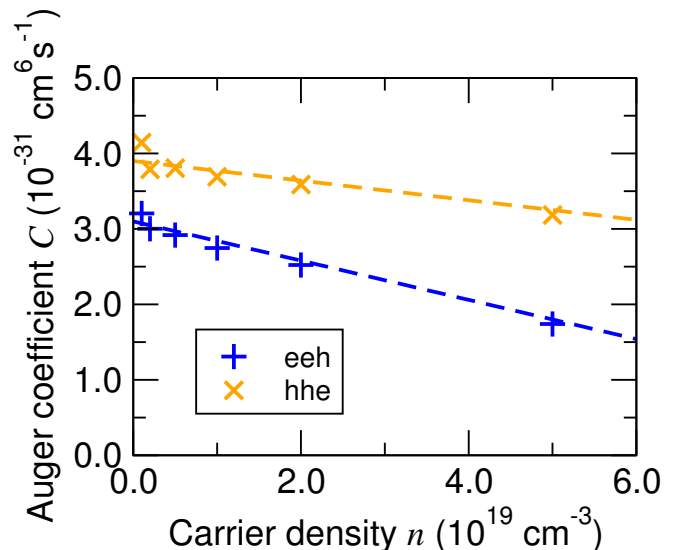We note that the calculated density dependence of the alloy-scattering-assisted Auger coefficients differs from the expression $C(n) = C_0/(1 + n/n_0)$ which has been used in the literature to describe the effect of phase-space filling.[8,47] Our present results show that the rate of decrease of the Auger coefficient with density in the $2 \times 10^{18} - 5 \times 10^{19}$ cm^{-3} range (which includes the peak of the internal quantum efficiency of LEDs[9]) is smaller than the rate of decrease of the radiative coefficient (B).[36] The more rapid decrease of the B coefficient in comparison to the C coefficient with increasing free-carrier density may explain the asymmetry of the internal quantum efficiency curve reported in the literature,[68] without invoking an additional fourth-order or higher-power carrier-loss mechanism.

C. Phonon-assisted Auger recombination in GaN

1. Phonon-assisted Auger coefficients

Figures 11(a) and 11(b) display our calculated results of the phonon-assisted Auger coefficients at 300 K for the e–e–h and h–h–e processes, respectively. To model the various alloy compositions we again used the calculated parameters for GaN and adjusted the band gap with a rigid scissors shift. By varying the band gap, we can model transitions to different sets of final states and this enables us to model phonon-assisted Auger recombination for a range of $In_xGa_{1-x}N$ compositions. The broadening of the δ function in Eqs. (19) and (20), as well as the imaginary component of the energy denominator in Eq. (21) were set to 0.3 eV. The phonon-assisted Auger coefficients are much larger than the values of the direct
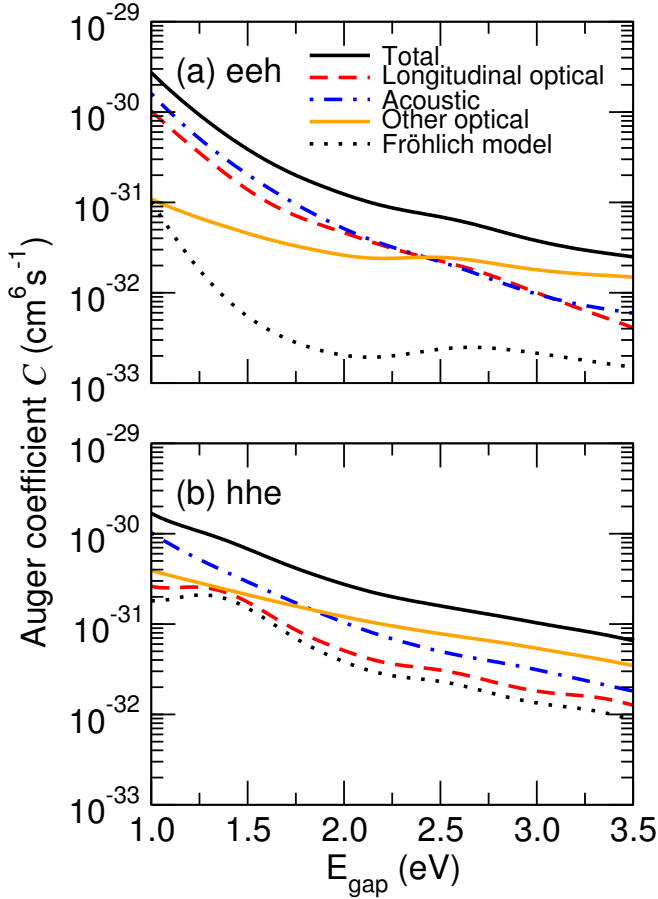
FIG. 11. (Color online) Contributions by the various phonon modes [longitudinal optical (LO), acoustic, and other optical modes] to the total phonon-assisted Auger coefficient of GaN for the (a) e–e–h and (b) h–h–e processes at 300 K. The phonon-assisted Auger coefficient using the Fröhlich model for the electron-phonon coupling is also shown for comparison.

followed by scattering by the acoustic phonon modes and the polar optical phonons. Figures 11(a) and 11(b) also include a comparison of the first-principles results to the phonon-assisted Auger coefficients calculated using the Fröhlich expression for the electron-phonon interaction potential, which models the scattering of carriers by polar interactions with the longitudinal-optical phonons[45] and has been shown to agree with the first-principles results at long wavelengths:[25]

$$\Delta V_{\text{el-ph}} = -\frac{i}{q} \left[\frac{2\pi e^2 \hbar \omega_{\text{LO}}}{V} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right) \right]^{1/2}.$$

For the model parameters, we used $\varepsilon_0 = (\varepsilon_{0\parallel} \varepsilon_{0\perp}^2)^{1/3} = 9.79$,[59] $\varepsilon_\infty = (\varepsilon_{\infty\parallel} \varepsilon_{\infty\perp}^2)^{1/3} = 5.5$,[59] and $\hbar \omega_{\text{LO}} = 92$ meV.[69] For band-gap values in the 2.4–3.0 eV energy range, the Auger coefficients calculated with the Fröhlich model are smaller than the first-principles ones by approximately one order of magnitude. For the h–h–e process the Fröhlich model agrees well with the first-principles data for the LO-phonon contribution to the scattering, but for the e–e–h case the Fröhlich model cannot account even for that contribution. As we will show later, the reason for this discrepancy is because the phonons involved in the scattering process for band-gap values in the 2.4–3.0 eV range are short-ranged ones (i.e., they involve large momentum transfer) and cannot be captured by the Fröhlich model, which is derived for the long-wavelength limit.

3. Limitations of approximate band-structure and phonon models

The limitations of approximate phonon models are further illustrated by an analysis of the wave vectors of the phonons that provide the dominant contribution to phonon-assisted Auger recombination. Figure 12 shows the contributions to the phonon-assisted Auger coefficients from phonons with different momenta $\boldsymbol{q}$ as a function of the adjusted band gap (to model the effect of alloying, but using parameters for GaN) and the norm of the phonon momentum $|\boldsymbol{q}|$. The plots indicate that for band gap values in the 2.4–3.0 eV range, the dominant contributions to the Auger coefficient originate from scattering by phonons with momenta comparable to the Brillouin-zone dimensions [i.e., comparable to the distance from the Γ to the A point at the boundary of the first Brillouin zone of wurtzite (Fig. 7)]. This indicates that the phonon-scattering events involved in the indirect Auger process in nitride devices with wavelengths from violet to green are short ranged. This is dictated by the band structure and by the fact that the higher conduction- (valence-)band states, which accommodate the electrons (holes) excited by the e–e–h (h–h–e) processes, are located at wave vectors comparable to the size of the first Brillouin zone.

Our analysis indicates that since phonon-assisted

Auger coefficient in GaN. The coefficients increase for decreasing band-gap values, indicating that Auger recombination is enhanced for higher-In-composition alloys.

2. Phonon mode and wave vector contributions to phonon-assisted Auger coefficients

One interesting question to examine is which phonons contribute to the Auger process and how well their contributions to the Auger processes are captured with more approximate theoretical treatments, as opposed to the first-principles density functional theory treatment used in the present work. The contributions from each type of vibrational mode to the Auger coefficient at 300 K are also plotted in Figs. 11(a) and 11(b) for the e–e–h and h–h–e processes, respectively. For both kinds of Auger processes, and for band-gap values in the 2.4–3.5 eV range, the dominant electron-phonon processes are due to optical-phonon deformation-potential scattering,
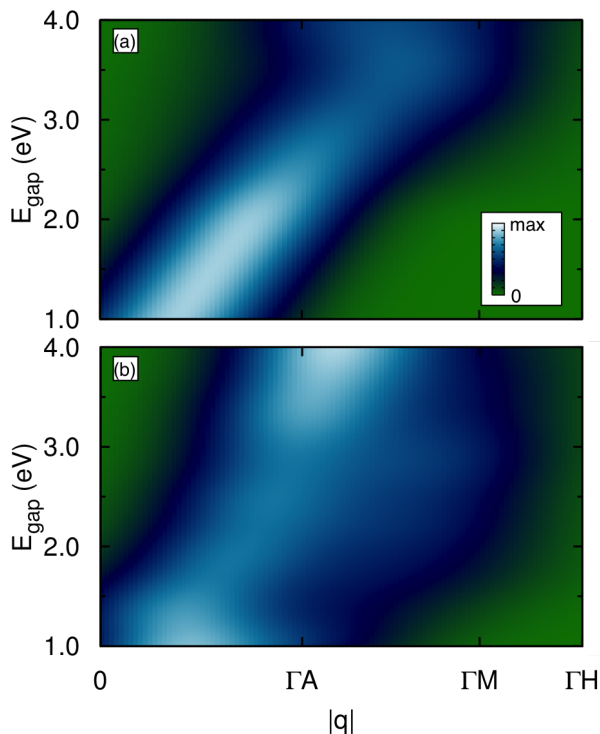
FIG. 12. (Color online) Relative magnitude of phonon-assisted Auger coefficients of GaN at 300 K due to (a) electron-electron-hole (e-e-h) and (b) hole-hole-electron (h-h-e) Auger recombination processes, plotted as a function of the scissors-shift-adjusted band gap and analyzed in terms of the norm of the q-points that contribute to the recombination process. The norms of the vectors from the Γ $(0, 0, 0)$ point to the A $(0, 0, \frac{1}{2})$, M $(\frac{1}{2}, 0, 0)$, and H $(\frac{2}{3}, \frac{1}{3}, \frac{1}{2})$ points are denoted along the horizontal axis. The coordinates of the points are given in relative coordinates with respect to the reciprocal lattice vectors.

Auger recombination is a short-range effect, it will not be significantly affected by quantum confinement. The values of the phonon-assisted Auger coefficients for quantum wells with a width of 2.5–3 nm (i.e., 5–6 times the lattice constant along the c axis), which is typical for commercial nitride LEDs, are expected to be approximately equal to their bulk values. Our results also show that the values of the phonon-assisted Auger coefficients cannot significantly be suppressed by engineering approaches such as nanopatterning, because these methods do not modify materials properties on the scale of the unit cell.

These results also validate the approximation of using Γ-point-only wave functions for the initial electron and hole states. It is known from $\boldsymbol{k} \cdot \boldsymbol{p}$ theory that the difference between the wave functions at Γ and at a $\boldsymbol{k}$-point in the vicinity of Γ is proportional to $|\boldsymbol{k}|$.[70] Since matrix elements are proportional to the square of the wave function, the error we make by approximating the initial-state wave functions of the carriers with those at Γ is on the order of the square of the ratio of the free-carrier Fermi

wave vector versus the dominant-phonon wave vectors, i.e., the Brillouin-zone dimensions. For $n = 10^{19}$ cm^{-3}, the free-carrier Fermi wave vector at $T = 0$ K is on the order of 0.03 a_B^{-1}, while the Brillouin-zone dimension (distance from Γ to A point) is 0.64 a_B^{-1}. Therefore, the error introduced by this approximation is on the order of 0.2%, which is sufficient for our purposes.

Another conclusion that can be drawn from the results of Figs. 11 and 12 is that the calculations using $\boldsymbol{k} \cdot \boldsymbol{p}$ theory for the band structure and the Fröhlich model for the electron-phonon coupling employed in Ref. 33 cannot properly capture Auger recombination in wide-band-gap nitride materials. This is because $\boldsymbol{k} \cdot \boldsymbol{p}$ theory is inaccurate for the bands and wave functions of states in the energy range at 2.4–3.0 eV away from the band edges, which is relevant for Auger recombination in the nitrides, and therefore cannot accurately include the contribution of these states to the overall recombination rate. Moreover, Figs. 11(a) and (b) illustrate that the dominant contributions to the phonon-assisted Auger recombination rate arises from deformation-potential optical phonon scattering and from acoustic phonons, while the LO phonons contribute only a small fraction to the overall rate. In addition, the results in Figs. 11(a) and (b) show that the Fröhlich model cannot accurately capture the contribution of the LO phonons to the Auger rate and that a microscopic description of electron–LO-phonon coupling is necessary. The proper description of the energies and wave functions of states away from the band edges as well as the incorporation of all phonon modes in electron-phonon scattering provided by the first-principles calculations of this work is essential to obtain accurate results for phonon-assisted Auger recombination in the nitrides.

4. Appropriateness of second-order perturbation theory

Our results also show that another previous study of phonon-assisted Auger recombination in the nitrides,[71] which is based on the formalism of Bardyzewski and Yevick,[72] does not incorporate the dominant phonon-assisted processes and significantly underestimates the magnitudes of the phonon-assisted Auger coefficients in the nitrides. The method used in Ref. 71 incorporates the effect of electron-phonon coupling in the broadening of the carrier spectral functions, but it does not consider the additional momentum provided to the carriers by phonon emission or absorption [Eq. (3) in Ref. 71]. In other words, the indirect Auger processes facilitated by the additional momentum provided by the phonons that excite carriers to higher electronic states near the edges of the Brillouin zone, which we found to dominate phonon-assisted Auger recombination, were not considered in Ref. 71. We therefore attribute the difference between our results and the ones reported in Ref. 71 to the omission of the dominant phonon terms in the latter work.

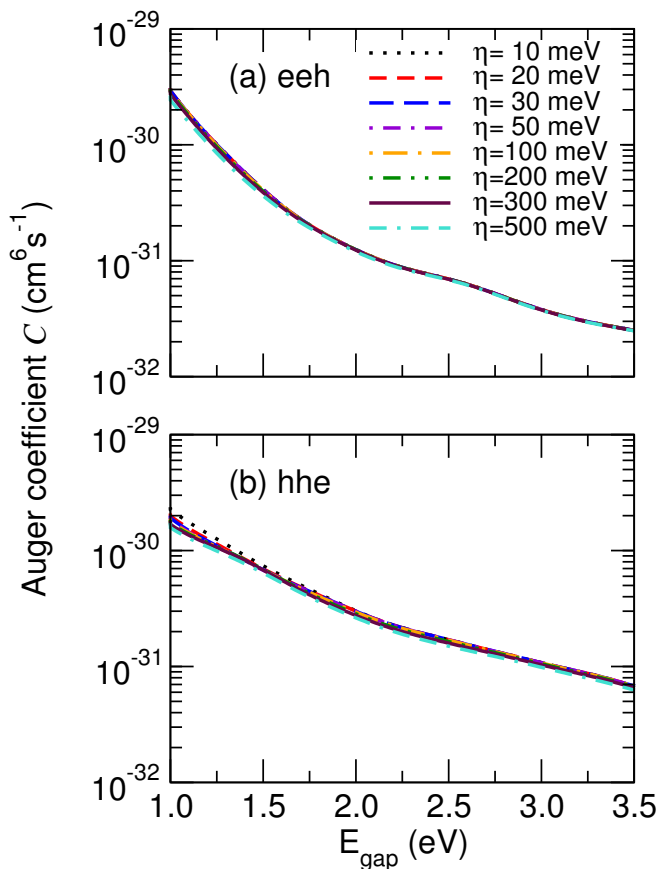It was also argued in Ref. 71 that a treatment of

FIG. 13. (Color online) Phonon-assisted Auger coefficient as a function of the scissors-shift-adjusted band gap of GaN for various values of the energy broadening parameter η in the energy denominator of Eq. (21).

phonon-assisted Auger recombination with second-order perturbation theory as performed in the present work may be inadequate because of the need to use the imaginary broadening parameter (η) in the denominator of the matrix elements in Eq. (21). We have explicitly examined the dependence of the calculated phonon-assisted e–e–h and h–h–e Auger coefficients on the value of the broadening parameter η used in Eq. (21). Figure 13 shows that the calculated coefficients remain approximately constant as the broadening parameter varies over almost two orders of magnitude. For adjusted band-gap values in the 2.4–3.5 eV range, the Auger coefficients change only by 2% for the e–e–h and by 14% for the h–h–e processes as the broadening parameter varies by a factor of 50. The dependence on the broadening parameter is slightly greater for band-gap values in the 1.0–2.0 eV range, but still the e–e–h coefficient changes by at most 21% and the h–h–e coefficient by 43% at most.

The imaginary part of the energy denominator may in general affect the transition rates calculated with second-order perturbation theory if, as pointed out in Ref. 71, resonant direct transitions to the intermediate states are possible. However, this is not the case for phonon-assisted Auger recombination in the nitrides. The intermediate states involved are far from being resonant and the real part of the energy denominator of the matrix elements in Eq. (21) dominates. Therefore, the theoretical treatment of phonon-assisted Auger recombination in terms of perturbation theory yields results that do not significantly depend on the chosen value of the imaginary broadening parameter. We also note that for materials for which resonant intermediate states do matter, the imaginary part of the energy denominator induced by electron-phonon coupling can be calculated with first-principles methods (e.g., Ref. 73) and incorporated in our formalism.

5. Temperature dependence of phonon-asissted Auger coefficients

Figure 14 shows the temperature dependence of the phonon-assisted Auger coefficients of GaN as a function of the scissors-shift-adjusted band gap. The Auger coefficients increase at higher temperatures because the phonon occupation numbers in Eq. (9) are a monotonically increasing function of temperature. We note that phonon-assisted Auger recombination is possible even at absolute zero temperature because the phonon-emission-assisted process does not require a finite thermal phonon population.

Our calculated temperature dependence of the Auger coefficient is in agreement with experimental measurements. Galler et al.[74] measured the radiative and Auger coefficients for a single-quantum-well InGaN device. After accounting for the reduced overlap due to polarization fields, they determined the values of the Auger coefficient as a function of temperature. Their data indicates that the Auger coefficient in InGaN is a monotonically increasing function of temperature, in agreement with our findings for phonon-assisted Auger recombination. Moreover, the internal-quantum-efficiency data reported by Laubsch et al.[6] show that Auger recombination in InGaN occurs even at cryogenic temperatures (4 K). This is in contrast with the expected exponential temperature dependence of direct Auger recombination[45] but in agreement with our finding that phonon-emission-assisted Auger processes are possible even at absolute zero temperature.

6. Influence of strain on phonon-asissted Auger coefficients

We also explored the influence of strain on the value of the phonon-assisted Auger coefficients, but we found that the effect is very small. Strain modifies the relative energy of the valence bands near the maximum at Γ[75] and influences the electron-hole recombination. It is therefore conceivable that certain strain conditions could yield a favorable valence-band ordering that minimizes the Auger coefficients. To test this hypothesis, we calcu-
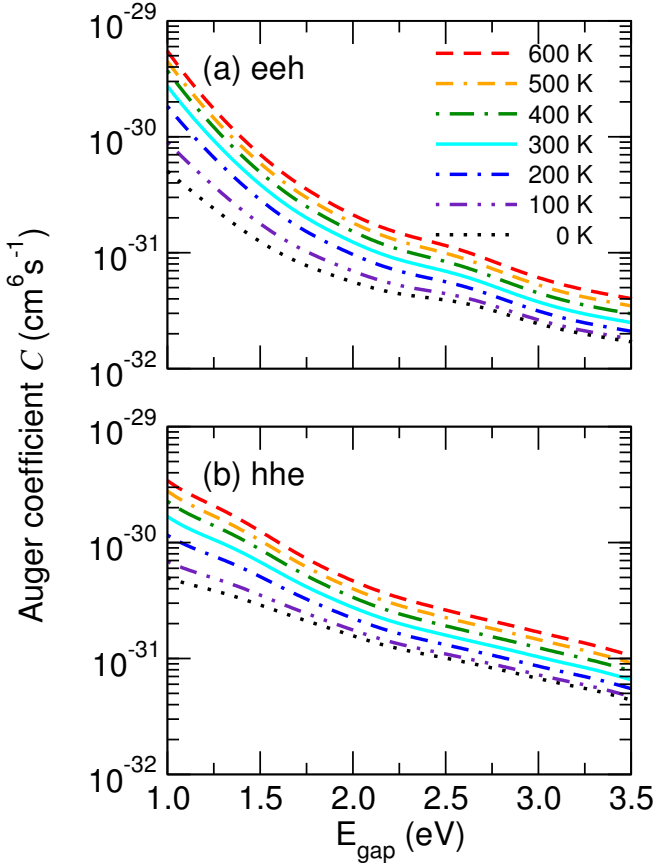
FIG. 14. (Color online) Temperature dependence of the (a) e–e–h and (b) h–h–e phonon-assisted Auger coefficient of GaN as a function of the scissors-shift adjusted band gap.

lated the phonon-assisted Auger coefficients in strained GaN. We applied a 0.5% tensile strain along one of the primitive vectors in the c-plane of the wurtzite structure, which breaks the valence-band degeneracy and allows us to distinguish transitions involving various hole states. The value of the Auger coefficients for various InGaN compositions was again calculated by applying a rigid scissors shift to the band gap. The calculated results are shown in Fig. 15. The e–e–h phonon-assisted Auger coefficient depends weakly on the choice of hole state [Fig. 15(a)]. This is expected by symmetry since electrons occupy Ga s states, while the top of the valence band is made of N p orbitals oriented along different directions. On the other hand, we expect variations for the h–h–e coefficient depending on whether the two holes occupy the same or different valence bands. Our results, however, show that this dependence is weak [Fig. 15(b)]. Therefore, the aplication of strain in order to reorder the valence bands and alter the hole state occupations does not significantly reduce the value of the phonon-assisted Auger recombination rate, which is the primary Auger recombination mechanism in the nitrides.
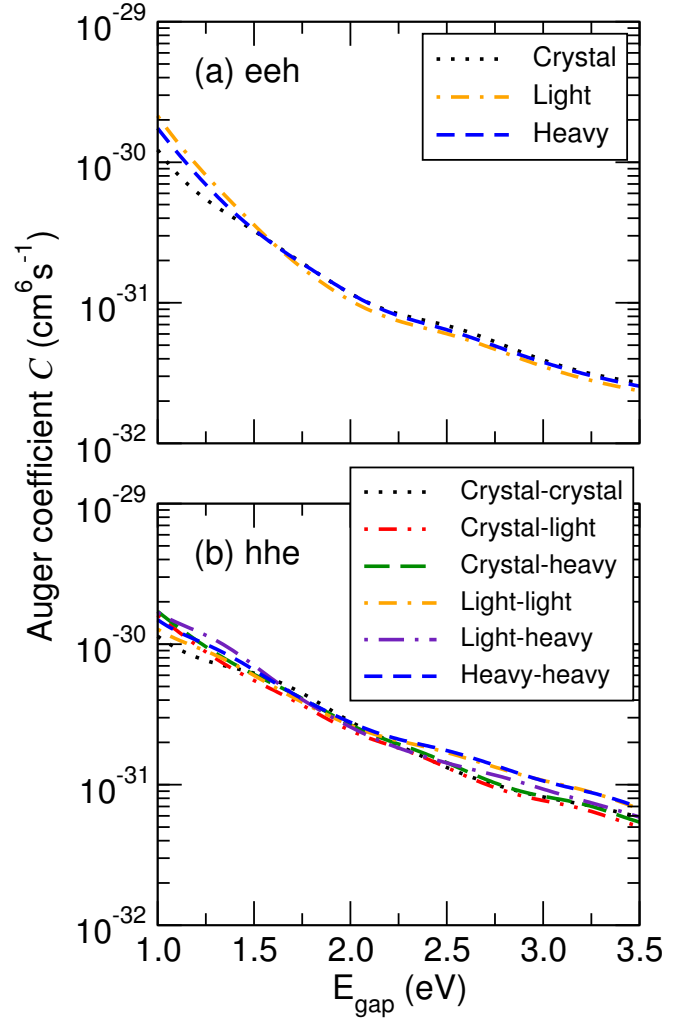

FIG. 15. (Color online) Phonon-assisted Auger recombination in strained GaN as a function of the scissors-shift-adjusted band gap for the (a) electron-electron-hole (eeh) and (b) hole-hole-electron (hhe) processes. The various curves correspond to the various hole states (heavy hole, light hole, or crystal-field-split hole) that participate in the Auger transitions.

D. Charged-defect-assisted Auger recombination in GaN

Another scattering mechanism that can in principle contribute to indirect Auger recombination is scattering due to charged defects. The formalism for the calculation of the charged-defect-assisted Auger coefficient is similar to the phonon-assisted case, except the phonon frequencies in Eqs. (9) are set to zero, and the electron-phonon scattering matrix elements (g) to the corresponding charged-defect-scattering ones, $g_{\text{defect}} = \langle \boldsymbol{k} | \Delta V_{\text{defect}} | \boldsymbol{k} + \boldsymbol{q} \rangle$, where ΔV_{defect} is the screened Coulomb potential of a defect of charge Z:

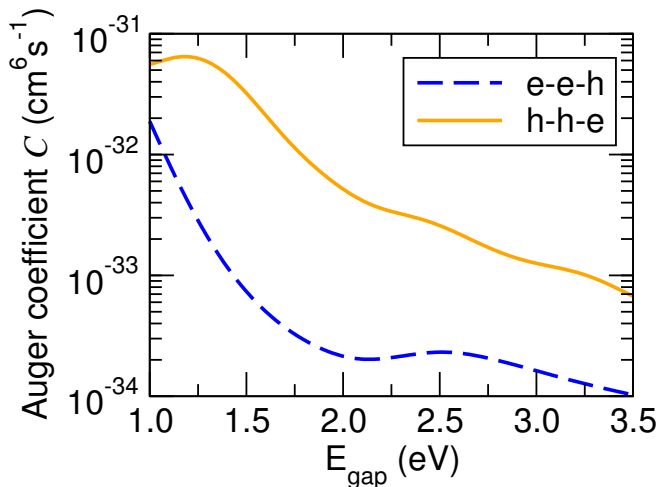$$\Delta V_{\text{defect}} = \frac{4\pi Z e^2}{\varepsilon(q)(q^2 + \lambda^2)}.$$

FIG. 16. (Color online) Charged-defect-assisted Auger recombination in GaN for a charged-defect concentration of 10^{19} cm^{-3}.

The calculations were performed using band-structure and screening parameters for GaN and rigidly shifting the band gap to model the effect of alloying. We investigated the scattering induced by negatively charged defects on Ga sites. The results for the e–e–h and h–h–e processes for a free-carrier concentration of 10^{19} cm^{-3} are shown in Fig. 16. We assumed a singly charged defect ($Z = 1$) with a density of 10^{19} cm^{-3}. This density is actually much higher than observed defect densities, for instance for the Ga vacancies that are the most common point defects in n-type GaN.[76,77] Figure 16 illustrates that charged-defect-assisted Auger recombination is weak for realistic densities of charged defects. Even if we assume a concentration of 10^{19} cm^{-3} of triply charged defects ($Z = 3$, the charge state expected for gallium vacancies in n-type GaN), the calculated charged-defect-assisted Auger coefficients increase only by a factor of 9 and are still smaller by one order of magnitude compared to the phonon-assisted results. In order to get an effect comparable to phonon-assisted Auger, the defect density would need to be on the order of 10^{21} cm^{-3} of singly charged defects or 10^{20} cm^{-3} of triply charged ones, which are unrealistically high and not encountered in actual devices. Therefore, charged-defect-assisted Auger recombination is an unlikely candidate to explain the efficiency droop in nitride light emitters. This is in agreement with experimental findings that the Auger coefficient does not change significantly for samples of varying quality.[3] We note that our results only included the effect of long-range Coulomb scattering by charged impurities, and do not include effects such as central-cell corrections,[78] short-range scattering by neutral defects, or Auger transitions to localized defect states.[79] Further work that considers these effects is needed to fully evaluate the interplay of defects and Auger recombination in nitride materials.

E. Cumulative Auger recombination coefficients

The contributions to the total Auger coefficients, as well as the cumulative result, for a carrier concentration of 10^{19} cm^{-3} as a function of the band gap of InGaN are plotted in Fig. 17. The band-gap dependence of the direct, phonon-assisted, and charged-defect-assisted Auger coefficients is obtained by applying a rigid scissors shift to the band gap of GaN. The alloy-scattering-assisted Auger coefficients calculated for In$_{0.25}$Ga$_{0.75}$N have been extrapolated to model other alloy compositions by applying a rigid scissors shift to the band gap (which was the only effect included in Fig. 8) and multiplying with $x(1-x)/(0.25 \cdot 0.75)$. This is because the alloy-scattering-assisted Auger recombination rate is proportional to the square of the alloy-disorder-scattering matrix elements, which are proportional to $x(1-x)$.[45] The band-gap bowing equation of In$_x$Ga$_{1-x}$N from Ref. 61 was inverted to yield the alloy composition as a function of the band gap. Our results reveal that the primary contribution to Auger recombination in In$_x$Ga$_{1-x}$N alloys stems from the scattering of charge carriers by the alloy disorder, while phonon-mediated processes provide a secondary contribution to Auger recombination. Our results also show that the recombination rate due to direct Auger processes is negligible compared to the phonon- and alloy-scattering assisted terms, and that charge-defect-assisted Auger recombination is negligible for realistic defect densities in actual devices. Our present results for the phonon-assisted and alloy-scattering-assisted coefficients slightly differ from values we reported previously,[35] mainly due to the different model used in the present work to describe the screening by free carriers.

The calculated values for the Auger coefficients are in agreement with experimentally reported values for nitride LEDs and lasers, and thus they highlight the important role of these non-radiative recombination mechanisms during the operation of nitride optoelectronic devices at high power.[35] The total value of the Auger coefficient for the e–e–h process ranges from 7×10^{-32} cm^{6}s^{-1} for a gap of 3.0 eV to 4×10^{-31} cm^{6}s^{-1} when the gap is 2.4 eV, while the h–h–e auger coefficient ranges from 3×10^{-31} cm^{6}s^{-1} for a gap value of 3.0 eV to 6×10^{-31} cm^{6}s^{-1} for a 2.4 eV gap. The overall Auger coefficient accounting for both e–e–h and h–h–e contributions amounts to $0.3 - 1 \times 10^{-30}$ cm^{6}s^{-1} in the range of In compositions corresponding to LEDs in the violet to green part of the visible spectrum.

V. SUMMARY

In this work, we presented a first-principles methodology and the associated computational details for the determination of the direct and indirect Auger recombination rate in direct-band-gap semiconductors. With minor modifications the methodology can also be applied to indirect-gap materials. The formalism has been applied
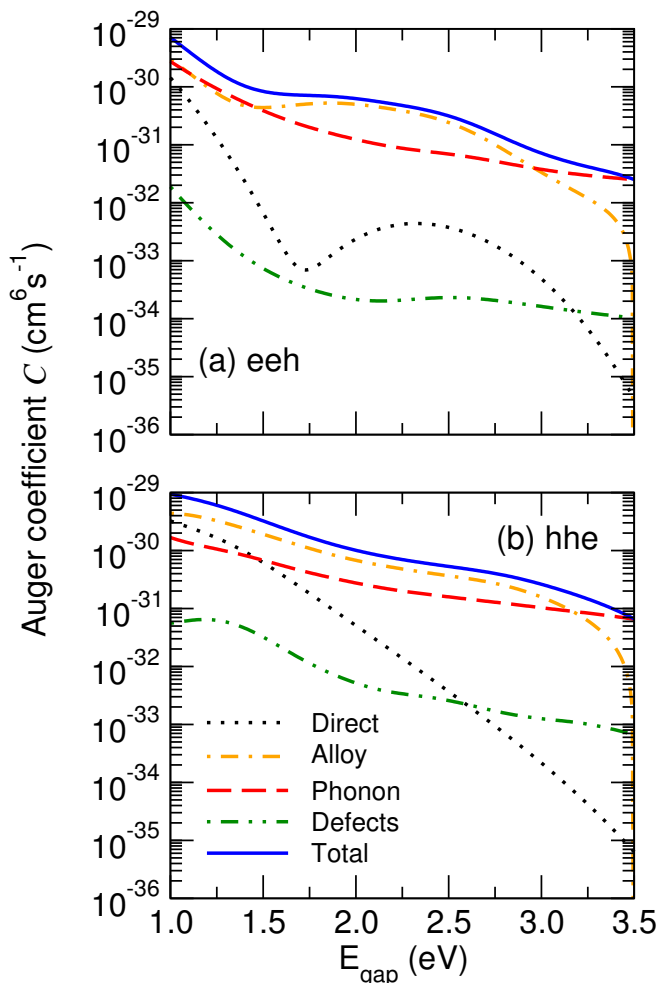
FIG. 17. (Color online) Summary of the calculated results for the investigated direct and indirect Auger recombination mechanisms in InGaN for free electron and hole concentrations of 10^{19} cm^{-3} and a temperature of 300 K. The contributions by the individual processes as well as the cumulative Auger coefficients are shown for (a) electron-electron-hole and (b) hole-hole-electron processes. In both cases, the results show that phonon-assisted Auger recombination is the dominant Auger recombination mechanism in InGaN.

to study Auger recombination in the group-III nitride materials used in optoelectronic devices. The sensitivity to various parameters and approximations used in the calculations has been examined. The results indicate that direct Auger recombination is not important in this class of wide-band-gap materials. Indirect Auger processes, mediated by electron-phonon and alloy scattering, are much stronger and the corresponding coefficients are of the magnitude that has been shown to be sufficient to account for the efficiency droop of nitride LEDs.[35] The formalism is not limited to nitrides and can be applied to study Auger recombination in other classes of materials both from a fundamental materials physics point of view and to understand their performance in high-power optoelectronic applications. Future improvements to the methodology could include the first-principles calculation of the dielectric function within the random-phase approximation[46] and the inclusion of alloy effects on the band structure and matrix elements for phonon-assisted Auger recombination.

[1] E. F. Schubert, *Light-emitting diodes* (Cambridge University Press, Cambridge, 2006).

[2] N. F. Gardner, G. O. Müller, Y. C. Shen, G. Chen, S. Watanabe, W. Götz, and M. R. Krames, Appl. Phys. Lett. **91**, 243506 (2007).

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

[4] M. Meneghini, N. Trivellin, G. Meneghesso, E. Zanoni, U. Zehnder, and B. Hahn, J. Appl. Phys. **106**, 114508 (2009).

[5] M. Zhang, P. Bhattacharya, J. Singh, and J. Hinckley, Appl. Phys. Lett. **95**, 201108 (2009).

[6] A. Laubsch, M. Sabathil, W. Bergbauer, M. Strassburg, H. Lugauer, M. Peter, S. Lutgen, N. Linder, K. Streubel, J. Hader, J. V. Moloney, B. Pasenow, and S. W. Koch, phys. stat. sol. (c) **6**, S913 (2009).

[7] A. Laubsch, M. Sabathil, J. Baur, M. Peter, and B. Hahn, IEEE T. Electron Dev. **57**, 79 (2010).

[8] A. David and M. J. Grundmann, Appl. Phys. Lett. **96**, 103504 (2010).

[9] A. David and M. J. Grundmann, Appl. Phys. Lett. **97**, 033501 (2010).

[10] W. G. Scheibenzuber, U. T. Schwarz, L. Sulmoni, J. Dorsaz, J.-F. Carlin, and N. Grandjean, J. Appl. Phys. **109**, 093106 (2011).

[11] M. Brendel, A. Kruse, H. Jönen, L. Hoffmann, H. Bremers, U. Rossow, and A. Hangleiter, Appl. Phys. Lett. **99**, 031106 (2011).

[12] E. Kioupakis, Q. Yan, and C. G. Van de Walle, Appl. Phys. Lett. **101**, 231107 (2012).

[13] J. Iveland, L. Martinelli, J. Peretti, J. S. Speck, and C. Weisbuch, Phys. Rev. Lett. **110**, 177406 (2013).

[14] A. R. Beattie and P. T. Landsberg, Proc. R. Soc. Lon. A **249**, 16 (1959).

[15] A. Sugimura, J. Appl. Phys. **51**, 4405 (1980).

[16] N. K. Dutta, J. Appl. Phys. **54**, 1236 (1983).

[17] A. Haug, D. Kerkhoff, and W. Lochmann, phys. stat. sol. (b) **89**, 357 (1978).

[18] D. B. Laks, G. F. Neumark, and S. T. Pantelides, Phys. Rev. B **42**, 5176 (1990).

[19] S. Picozzi, R. Asahi, C. B. Geller, and A. J. Freeman, Phys. Rev. Lett. **89**, 197601 (2002).

[20] K. T. Delaney, P. Rinke, and C. G. Van de Walle, Appl. Phys. Lett. **94**, 191109 (2009).

[21] M. Heinemann and C. Heiliger, Journal of Applied Physics **110**, 083103 (2011).

[22] M. Govoni, I. Marri, and S. Ossicini, Phys. Rev. B **84**, 075215 (2011).

[23] J. Noffsinger, E. Kioupakis, C. G. Van de Walle, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **108**, 167402 (2012).

[24] C. E. Patrick and F. Giustino, J. Phys.: Condens. Matter **26**, 365503 (2014).

[25] E. Kioupakis, P. Rinke, A. Schleife, F. Bechstedt, and C. G. Van de Walle, Phys. Rev. B **81**, 241201 (2010).

[26] E. Kioupakis, P. Rinke, and C. G. Van de Walle, Appl. Phys. Express **3**, 082101 (2010).

[27] H. Peelaers, E. Kioupakis, and C. G. Van de Walle, Appl. Phys. Lett. **100**, 011914 (2012).

[28] D. M. Eagles, P. Phys. Soc. **78**, 204 (1961).

[29] L. Huldt, phys. stat. sol. (a) **33**, 607 (1976).

[30] W. Lochmann, phys. stat. sol. (a) **40**, 285 (1977).

[31] W. Lochmann, phys. stat. sol. (a) **45**, 423 (1978).

[32] A. Haug, Solid State Communications **22**, 537 (1977).

[33] B. Pasenow, S. W. Koch, J. Hader, J. V. Moloney, M. Sabathil, N. Linder, and S. Lutgen, phys. stat. sol. (c) **6**, S864 (2009).

[34] F. Bertazzi, M. Goano, and E. Bellotti, in *Proc. of IWCE* (2010) pp. 179–182.

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

[36] E. Kioupakis, Q. Yan, D. Steiauf, and C. G. Van de Walle, New Journal of Physics **15**, 125006 (2013).

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

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

[39] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[40] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001).

[41] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Comput. Phys. Commun. **178**, 685 (2008).

[42] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Rev. Mod. Phys. **84**, 1419 (2012).

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

[44] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. 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).

[45] B. K. Ridley, *Quantum Processes in Semiconductors* (Clarendon Press, Oxford, 1982).

[46] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[47] J. Hader, J. V. Moloney, and S. W. Koch, Appl. Phys. Lett. **87**, 201112 (2005).

[48] F. Bassani and G. Pastori Parravicini, *Electronic States and Optical Transitions in Solids* (Pergamon Press, 1975).

[49] F. Giustino, M. L. Cohen, and S. G. Louie, Phys. Rev. B **76**, 165108 (2007).

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

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

[52] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[53] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[54] J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C: Solid State Phys. **12**, 4409 (1979).

[55] C. Stampfl and C. G. Van de Walle, Physical Review B **59**, 5521 (1999).

[56] G. Cappellini, R. Del Sole, L. Reining, and F. Bechstedt, Phys. Rev. B **47**, 9892 (1993).

[57] M. Palummo, R. D. Sole, L. Reining, F. Bechstedt, and G. Cappellini, Solid State Commun. **95**, 393 (1995).

[58] D. D. Manchon, A. S. Barker, P. J. Dean, and R. B. Zetterstrom, Solid State Commun. **8**, 1227 (1970).

[59] A. Barker Jr and M. Ilegems, Phys. Rev. B **7**, 743 (1973).

[60] M. J. Bergmann and J. H. C. Casey, J. Appl. Phys. **84**, 1196 (1998).

[61] P. G. Moses and C. G. Van de Walle, Appl. Phys. Lett. **96**, 021908 (2010).

[62] I. Vurgaftman and J. R. Meyer, J. Appl. Phys. **94**, 3675 (2003).

[63] F. Bertazzi, M. Goano, and E. Bellotti, Appl. Phys. Lett. **97**, 231118 (2010).

[64] K. A. Bulashevich and S. Y. Karpov, phys. stat. sol. (c) **5**, 2066 (2008).

[65] A. Schleife, P. Rinke, F. Bechstedt, and C. G. Van de Walle, J. Phys. Chem. C **117**, 4189 (2013).

[66] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Phys. Rev. Lett. **65**, 353 (1990).

[67] D. Shin, R. Arróyave, Z.-K. Liu, and A. Van de Walle, Phys. Rev. B **74**, 024204 (2006).

[68] Q. Dai, Q. Shan, J. Cho, E. F. Schubert, M. H. Crawford, D. D. Koleske, M.-H. Kim, and Y. Park, Appl. Phys. Lett. **98**, 033506 (2011).

[69] M. Giehler, M. Ramsteiner, O. Brandt, H. Yang, and K. H. Ploog, Appl. Phys. Lett. **67**, 733 (1995).

[70] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 2001).

[71] F. Bertazzi, M. Goano, and E. Bellotti, Appl. Phys. Lett. **101**, 011111 (2012).

[72] W. Bardyszewski and D. Yevick, J. Appl. Phys. **58**, 2713 (1985).

[73] C.-H. Park, F. Giustino, M. L. Cohen, and S. G. Louie, Phys. Rev. Lett. **99**, 086804 (2007).

[74] B. Galler, P. Drechsel, R. Monnard, P. Rode, P. Stauss, S. Froehlich, W. Bergbauer, M. Binder, M. Sabathil, B. Hahn, and J. Wagner, Appl. Phys. Lett. **101**, 131111 (2012).

[75] Q. Yan, P. Rinke, M. Scheffler, and C. G. Van de Walle, Appl. Phys. Lett. **95**, 121111 (2009).

[76] C. G. Van de Walle and J. Neugebauer, J. Appl. Phys. **95**, 3851 (2004).

[77] F. Tuomisto, J.-M. Mäki, C. Rauch, and I. Makkonen, J. Cryst. Growth **350**, 93 (2012).

[78] H. I. Ralph, G. Simpson, and R. J. Elliott, Physical Review B **11**, 2948 (1975).

[79] A. Hangleiter, Physical Review B **35**, 9149 (1987).

[80] K. Momma and F. Izumi, J. Appl. Crystallogr. **44**, 1272 (2011).