Multiphonon carrier capture by defects in semiconductors

Laura R. Nichols [1,*] Guanzhi Li [2,*] Yue Yu [2] Jun Jiang [2] Andrew O'Hara,[1,†]
Georgios D. Barmparis [3] Sokrates T. Pantelides [1,4,‡] and X.-G. Zhang [2,§]

[1]*Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235, USA*
[2]*Department of Physics and the Quantum Theory Project, University of Florida, Gainesville, Florida 32611, USA*
[3]*Institute of Theoretical and Computational Physics, Department of Physics, University of Crete, Heraklion 71003, Greece*
[4]*Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, Tennessee 37235, USA*

(Received 12 March 2024; revised 5 December 2024; accepted 23 December 2024; published 7 January 2025)

We report developments and corresponding calculations in the first-principles theory of multiphonon processes in solids, enabling direct calculations of nonradiative capture cross sections of both thermalized and hot carriers by defects in semiconductors. We present an improved derivation of the theory, implement an efficient time-domain integration scheme that enables summing over all possible multiphonon configurations for dissipating the energy released by the carrier capture, and extend the time-domain formulation to allow changes in phonon frequencies before and after capture. We examine zeroth- and first-order nonequilibrium contributions to the electron-capture cross section of a triply hydrogenated vacancy in Si and the first-order equilibrium hole-capture coefficient of a C_N substitutional impurity in GaN. In Si, we find that the zeroth-order term is two orders of magnitude larger than the first-order term. In GaN, the present results are within range of the sparse available experimental data at high temperatures but below the single experimental point at low temperatures. We see an improvement over previous calculations and a decreased temperature dependence that we attribute to the inclusion of all phonon modes.

DOI: 10.1103/PhysRevB.111.045201

I. INTRODUCTION

Electron wave functions in solids are functions of not only the electron coordinates but of the nuclear coordinates as well. This dependence is the fundamental reason for electron-phonon coupling. The Born-Oppenheimer approximation (BOA), also known as the adiabatic approximation, is usually adopted in first-principles calculations. In the BOA, the nuclear coordinates are "frozen," and only the electron coordinates are treated explicitly. The solutions are electronic quantum states that only depend on the nuclear coordinates parametrically. The sum of the total energy of each electronic configuration and the corresponding energy of nucleus-nucleus interactions as a function of the nuclear coordinates form a separate potential energy surface that, together with the kinetic-energy operator, define the Hamiltonian of the nuclear system. The dynamics along such surfaces is considered "adiabatic." When this total energy is expanded to second order in nuclear displacements, the solutions of the corresponding Hamiltonian problem are the vibrational quanta that we call phonons.

Considering the change in the electronic wave functions with changes in nuclear positions amounts to the inclusion of terms beyond the BOA, known as nonadiabatic terms. Such terms can cause electrons to transition from one potential energy surface to another. When mediated by the creation or annihilation of a single phonon quantum, these transitions amount to the well-known electron-phonon scattering that limits the conductivity of metals and the mobilities of electrons and holes in semiconductors.

When a defect is present in a semiconductor, an electron or hole can transition from a band state to a localized defect state in the energy gap, with the transition energy dissipated to the lattice by multiphonon processes. Typically, such "carrier capture" and the inverse process of "carrier emission" occur routinely under equilibrium conditions. Experiments such as deep-level transient spectroscopy (DLTS) [1] measure the corresponding capture and emission coefficients. In devices, it is common to have so-called hot electrons or holes with high energies outside the Fermi-Dirac distribution that can also be captured. A hot carrier can also undergo inelastic scattering into a lower-energy band state while the transition energy is dissipated primarily to the defect, which undergoes a transformation (e.g., a bound hydrogen atom is released or the defect transits to a metastable state [2]).

Theories for carrier capture by defects were developed and debated beginning in 1950, continuing into the 1980s [3–6]. These theories were based on the assumption that multiphonon-mediated carrier capture can be caused only by nonadiabatic terms (i.e., terms in the total Hamiltonian that are dropped in the adiabatic approximation). In 1950, Huang and Rhys [3] carried out analytical calculations by using a single phonon mode to approximate multiphonon processes.

*These authors contributed equally to this work.
†Present address: Department of Physics, Western Michigan University, Kalamazoo, MI 49008, USA.
‡Contact author: pantelides@vanderbilt.edu
§Contact author: xgz@ufl.edu

In 2014, Alkauskas *et al.* [7] reported the first density-functional-theory (DFT)-based calculations of capture coefficients at a defect in a semiconductor. They adopted the conventional assumption that the first-order nonadiabatic term is responsible for the carrier capture and used a single, specially selected effective phonon mode to carry out the multiphonon sum.

The following year, 2015, we reported the discovery that, though the nonadiabatic term is needed to cause carrier capture by a defect under equilibrium conditions, it is possible to obtain a Franck-Condon-like carrier capture within the adiabatic limit under nonequilibrium conditions [8] (i.e., in the presence of an electrical current or hot carriers, which is typical in both electronic and optoelectronic devices under operating conditions). Within the harmonic approximation, the adiabatic transitions only occur if the defect relaxes after the carrier-capture event. Instead of considering a single phonon mode, we introduced a Monte Carlo sampling method to sum the contributions from multiphonon configurations with an increasing number of distinct phonon modes until convergence. We reported results for the adiabatic term for a model defect in Si: a vacancy in which all but one of the four Si dangling bonds are hydrogenated. Converged results were obtained using 12 distinct phonon modes. For the nonadiabatic term, however, we only provided an estimate for the ratio between the magnitudes of the adiabatic term and the nonadiabatic term, without actually calculating the nonadiabatic term.

We have since found that the Monte Carlo scheme that was used for the calculation of the adiabatic term [8] would overwhelm today's supercomputers if directly applied to computing the first-order term. However, we now note that in a 1981 paper Huang [6] showed that the so-called first-order nonadiabatic transition [8] is in fact contained within the adiabatic approximation. Thus, what we referred to as zeroth-order (adiabatic) and first-order (nonadiabatic) contributions [8] are in fact both adiabatic. In the remainder of the paper, we will simply refer to the two terms as zeroth order and first order. Their definitions will become clear in the mathematical exposition. We should also point out that there is no zeroth-order contribution to equilibrium capture.

The goal of the present paper is as follows. First, we report an improved derivation of the theory developed in our previous work [8] and replace the Monte Carlo sampling method with a faster and more accurate time-domain-integration method [9] that allows participation of all modes in the dissipation via multiphonon processes of the energy released during capture. We also generalize the time-domain formalism to allow the phonon frequencies to change after capture. We then calculate the zeroth-order and first-order nonequilibrium contributions to the electron-capture cross section of a triply hydrogenated vacancy in Si, studied in our previous paper [8], and the first-order equilibrium hole-capture coefficient of a C_N substitutional impurity in GaN, which is one of the defects investigated by Alkauskas *et al.* [7].

For electron capture by a triply hydrogenated vacancy in Si, we find that the first-order term is roughly two orders of magnitude smaller than the zeroth-order term, compared to the roughly four orders of magnitude we estimated in Ref. [8]. For hole capture by C_N impurities in GaN, we find that, by including all phonon modes in the calculation, the first-order term of this theory has a weaker temperature dependence than the previous calculation [7], bringing it closer to available experimental data, which show nearly no dependence on temperature. More extensive and more accurate experimental data would be highly desirable to fully validate the theory.

II. THEORY AND METHOD

A. Theoretical framework

The theoretical framework largely follows that of Ref. [8]. Here, we provide a more streamlined and complete account of the theory that underlies the present calculations, fully enabling readers to derive all of the equations that are used for computation. Changes will be noted during the exposition.

We begin with the total many-electron Hamiltonian for a crystal containing the defect,

$$H = H_{\mathrm{e}} + H_{\mathrm{I}}. \tag{1}$$

The electronic part H_{e} is a function of the set of all electronic positions $\{\mathbf{r}_\alpha\}$ and the set of all nuclear positions $\{\mathbf{R}_k\}$ due to the Coulomb interaction between the electrons and the ions:

$$H_{\mathrm{e}}(\{\mathbf{r}_\alpha\},\{\mathbf{R}_k\}) = K_{\mathrm{e}}(\{\mathbf{r}_\alpha\}) + V_{\mathrm{ee}}(\{\mathbf{r}_\alpha\}) + V_{\mathrm{eI}}(\{\mathbf{r}_\alpha\},\{\mathbf{R}_k\}). \tag{2}$$

The ionic part H_{I} is only a function of $\{\mathbf{R}_k\}$,

$$H_{\mathrm{I}}(\{\mathbf{R}_k\}) = K_{\mathrm{I}}(\{\mathbf{R}_k\}) + V_{\mathrm{II}}(\{\mathbf{R}_k\}), \tag{3}$$

where K_{e} and K_{I} are the electronic and ionic kinetic energies and V_{ee}, V_{II}, and V_{eI} are the electron-electron, ion-ion, and electron-ion interaction potentials, respectively. The eigenvalue equation for the full Hamiltonian is

$$H(\{\mathbf{r}_\alpha\},\{\mathbf{R}_k\})|\Psi\rangle = E^{\mathrm{T}}|\Psi\rangle, \tag{4}$$

where the wave function $|\Psi\rangle$ describes all electrons and nuclei in the solid and E^{T} is the total energy of the system. The corresponding Hamiltonian and electronic and ionic positions for the pristine crystal (without the defect) is indicated by a superscript 0:

$$H^0(\{\mathbf{r}^0_\alpha\},\{\mathbf{R}^0_k\}) = H^0_{\mathrm{e}}(\{\mathbf{r}^0_\alpha\},\{\mathbf{R}^0_k\}) + H^0_{\mathrm{I}}(\{\mathbf{R}^0_k\}). \tag{5}$$

The Born-Oppenheimer approximation (BOA) allows us to separate the total wave function into an electronic and an ionic wave function. The electronic eigenvalue problems for the perfect and defect-containing crystals are then given by

$$H^0_{\mathrm{e}}(\{\mathbf{r}^0_\alpha\},\{\mathbf{R}^0_k\})|\Phi^0_l\rangle = E^0_l(\{\mathbf{R}^0_k\})|\Phi^0_l\rangle \tag{6}$$

and

$$H_{\mathrm{e}}(\{\mathbf{r}_\alpha\},\{\mathbf{R}_k\})|\Phi_l\rangle = E_l(\{\mathbf{R}_k\})|\Phi_l\rangle, \tag{7}$$

respectively. Here, the electronic wave functions depend only parametrically on the nuclear coordinates (i.e., the derivatives of the electr onic wave functions with respect to the nuclear coordinates are neglected). We adopt density-functional theory (DFT) and the Kohn-Sham formalism to simplify the above many-electron problems into systems of noninteracting particles. Within the Kohn-Sham formalism, the electronic wave function is given by a Slater determinant of noninteracting, single-particle orbitals. The defect electronic wave

function can then be written

$$|\Phi_l\rangle = \frac{1}{\sqrt{N!}}|\det\{\phi_1 \cdots \phi_N\}\rangle, \tag{8}$$

where N is the total number of electrons.

The corresponding eigenvalue problem for the ions is

$$[H_{\rm I}(\{\mathbf{R}_k\}) + E_l(\{\mathbf{R}_k\})]|X_{ln}\rangle = E^{\rm T}_{ln}|X_{ln}\rangle. \tag{9}$$

Here, the eigenvalue $E^{\rm T}_{ln}$ is the total energy of the system, with l labeling the electronic states and n labeling ionic (phonon) states. The product of the electronic and ionic wave functions, $|\Psi_{ln}\rangle = |\Phi_l\rangle|X_{ln}\rangle$, is viewed as the BOA solution to Eq. (4), and provides a basis for the calculation of the transition rate for multiphonon carrier capture by a defect.

We now regroup the terms in the ionic eigenvalue problem to get

$$[K_{\rm I}(\{\mathbf{R}_k\}) + W_l(\{\mathbf{R}_k\})]|X_{ln}\rangle = E^{\rm T}_{ln}|X_{ln}\rangle, \tag{10}$$

where the total potential energy for the ions is defined as

$$W_l(\{\mathbf{R}_k\}) \equiv V_{\rm II}(\{\mathbf{R}_k\}) + E_l(\{\mathbf{R}_k\}). \tag{11}$$

For a given electronic state l, we can expand this potential around the equilibrium positions $\{\mathbf{R}^{(l)}_k\}$ and subtract out the equilibrium potential $W_l(\{\mathbf{R}^{(l)}_k\})$ from both sides in Eq. (10). If we retain only the second-order term in the potential expansion (harmonic approximation), Eq. (10) describes the vibrating nuclei as a set of coupled harmonic oscillators. We adopt the usual approach of treating the ions as classical particles, which allows us to construct and diagonalize the dynamical matrix into normal modes, each of which is represented by a Hamiltonian,

$$H_j = \tfrac{1}{2}\dot{q}_j^2 + \tfrac{1}{2}\omega_j q_j^2, \tag{12}$$

defined in terms of the canonical displacement q_j. The eigenvalue problems of these single-phonon-mode Hamiltonians are

$$H_j|\chi_{n_j}\rangle = \left(n_j + \tfrac{1}{2}\right)\hbar\omega_j|\chi_{n_j}\rangle, \tag{13}$$

where n_j is the occupation number of mode j. We denote the initial phonon configuration as $n \equiv \{n_1, n_2, \ldots, n_N\}$. Since the phonon modes are independent harmonic oscillators, the initial-state phonon wave function $|X_{in}\rangle$ is the product of the single-oscillator wave functions $|\chi_{n_j}\rangle$:

$$|X_{in}\rangle = \prod_{j=1}^{N}|\chi_{n_j}\rangle. \tag{14}$$

In general, after capture there is a change in the phonon frequencies to ω'_j and a shift in the equilibrium positions, Δq_j, so that the final, displaced Hamiltonian of phonon mode j is

$$H'_j = \tfrac{1}{2}\dot{q}_j^2 + \tfrac{1}{2}\omega_j'^2(q_j + \Delta q_j)^2, \tag{15}$$

and the corresponding eigenvalue problem is

$$H'_j|\chi'_{n'_j}\rangle = \left(n'_j + \tfrac{1}{2}\right)\hbar\omega'_j|\chi'_{n'_j}\rangle. \tag{16}$$

The Duschinsky matrix is approximated by a permutation matrix. $n' \equiv \{n'_1, n'_2, \ldots, n'_N\}$ defines a single possible final-state phonon configuration with wave function

$$|X_{fn'}\rangle = \prod_{j=1}^{N}\left|\chi'_{n'_j}\right\rangle. \tag{17}$$

The total phonon energies Θ_n and $\Theta'_{n'}$ of the initial and final phonon configurations n and n' are written as

$$\Theta_n = \sum_{j=1}^{N}\left(n_j + \frac{1}{2}\right)\hbar\omega_j \tag{18}$$

and

$$\Theta'_{n'} = \sum_{j=1}^{N}\left(n'_j + \frac{1}{2}\right)\hbar\omega'_j. \tag{19}$$

We assume that the initial phonon frequencies ω_j are the same for the different initial band states, but we allow for a change in the frequencies to ω'_j after capture. Because we consider a change in the equilibrium positions after capture, the initial and final phonons must be defined by expansions around the different equilibrium positions to account for a complete conservation of energy, including the relaxation energy. Doing so leads to

$$E^{\rm T}_{in} = \Theta_n + W_i\left(\left\{\mathbf{R}^{(i)}_k\right\}\right) \tag{20}$$

and

$$E^{\rm T}_{fn'} = \Theta'_{n'} + W_f\left(\left\{\mathbf{R}^{(f)}_k\right\}\right). \tag{21}$$

These total energies must be the same. This energy conservation will be reflected in the δ function in Fermi's golden-rule formula for the transition rate as the energy difference

$$E^{\rm T}_{fn'} - E^{\rm T}_{in} = \Theta'_{n'} - \Theta_n + W_{if}, \tag{22}$$

where the notation

$$W_{if} \equiv W_f\left(\left\{\mathbf{R}^{(f)}_k\right\}\right) - W_i\left(\left\{\mathbf{R}^{(i)}_k\right\}\right) \tag{23}$$

is used for brevity.

B. Total transition rate

We consider a carrier that approaches an isolated defect from far away and is suddenly perturbed by the defect. As in scattering theory, this incoming carrier is described by a single band-state wave function in a pristine crystal. On the other hand, the defect state that the electron finally transitions into and all other single-electron states that do not participate in the scattering/capture process belong to states of the defected solid. To facilitate derivations involving states belonging to different solids, we introduce an auxiliary set of complete basis, $|\tilde{\Phi}_l\rangle$, using

$$\tilde{H}_e(\{\mathbf{r}_\alpha\}, \{\mathbf{R}_k\})|\tilde{\Phi}_l\rangle = \tilde{E}_l(\{\mathbf{R}_k\})|\tilde{\Phi}_l\rangle, \tag{24}$$

where an "unperturbed Hamiltonian" $\tilde{H}_e(\{\mathbf{r}_\alpha\}, \{\mathbf{R}_k\})$ will be defined below. This basis must have the following properties. First, one of the states, $|\tilde{\Phi}_i\rangle$, corresponds to the initial electronic state and contains the band state occupied by the incoming carrier, which is approximately $|\phi^0_i\rangle$:

$$|\tilde{\Phi}_i\rangle \approx \frac{1}{\sqrt{N!}}\left|\det\left\{\phi_1 \cdots \phi^0_i \cdots \phi_N\right\}\right\rangle. \tag{25}$$

Second, another state, $|\tilde{\Phi}_f\rangle$, corresponds to the final electronic state,

$$|\tilde{\Phi}_f\rangle \approx |\Phi_f\rangle = \frac{1}{\sqrt{N!}}|\det\{\phi'_1 \cdots \phi'_f \cdots \phi'_N\}\rangle, \tag{26}$$

where $|\phi'_f\rangle$ is the single-electron wave function of the defect level that captures the carrier, and the superscript prime is used to denote all the single-electron wave functions in the final electronic state.

To ensure that the sudden turn-on of the defect potential, $\Delta H \equiv H_e - H_e^0$, affects only the incoming-carrier wave function $|\phi_i^0\rangle$ but not any other single-electron wave functions, the perturbing Hamiltonian H_1^{BO} is defined as the coupling matrix elements of the defect potential ΔH between the initial (i) and final (f) states [8]:

$$H_1^{\mathrm{BO}} = |\tilde{\Phi}_i\rangle\langle\tilde{\Phi}_i|\Delta H|\tilde{\Phi}_f\rangle\langle\tilde{\Phi}_f| + |\tilde{\Phi}_f\rangle\langle\tilde{\Phi}_f|\Delta H|\tilde{\Phi}_i\rangle\langle\tilde{\Phi}_i|. \tag{27}$$

The full basis set $|\tilde{\Phi}_l\rangle$ in Eq. (24) is then defined as the eigenstates of the "unperturbed" Hamiltonian,

$$\tilde{H}_e \equiv H_e - H_1^{\mathrm{BO}}. \tag{28}$$

In this definition, H_1^{BO} is treated as a Förster-like [10] perturbation. Note that the basis $|\tilde{\Phi}_l\rangle$ is only needed to help with the derivation and is never computed.

Electron-phonon coupling is usually defined through a series expansion of the defect Hamiltonian H_e around the equilibrium nuclear positions of the initial electronic configuration (dropping the dependence on the electronic coordinates here for brevity),

$$H_e(\{\mathbf{R}_k\}) = H_e\big(\{\mathbf{R}_k^{(i)}\}\big) + \sum_{k'} \big(\mathbf{R}_{k'} - \mathbf{R}_{k'}^{(i)}\big) \cdot \nabla_{\mathbf{R}_{k'}} H_e\big|_{\{\mathbf{R}_k^{(i)}\}} + \cdots, \tag{29}$$

with

$$H_{\mathrm{ep}} \equiv H_e(\{\mathbf{R}_k\}) - H_e\big(\{\mathbf{R}_k^{(i)}\}\big). \tag{30}$$

Note that, despite the naming convention, this electron-phonon-coupling term is still within the Born-Oppenheimer approximation because it just represents part of the series expansion of the electronic Hamiltonian. However, Huang [6] showed that, to first order, the perturbation from linear electron-phonon coupling on the static, frozen-phonon wave functions is equivalent to the perturbation due to the nonadiabatic derivatives left out of the Born-Oppenheimer approximation when the wave functions are expanded as

$$|\Phi_l(\{\mathbf{R}_k\})\rangle = \big|\Phi_l\big(\{\mathbf{R}_k^{(i)}\}\big)\big\rangle + \sum_{l'} \kappa_{l'l}\big|\Phi_{l'}\big(\{\mathbf{R}_k^{(i)}\}\big)\big\rangle, \tag{31}$$

where

$$\kappa_{l'l} \equiv \frac{\big\langle\Phi_{l'}\big(\{\mathbf{R}_k^{(i)}\}\big)\big|H_{\mathrm{ep}}\big|\Phi_l\big(\{\mathbf{R}_k^{(i)}\}\big)\big\rangle}{E'_{l'}\big(\{\mathbf{R}_k^{(i)}\}\big) - E'_l\big(\{\mathbf{R}_k^{(i)}\}\big)}, \tag{32}$$

with the energies $E'_l(\{\mathbf{R}_k^{(i)}\})$ shifted from $E_l(\{\mathbf{R}_k^{(i)}\})$ by the diagonal contribution of H_{ep}.

In order to consider the two perturbations on the same unperturbed states $|\tilde{\Phi}_l\rangle$, we must define $\tilde{H}_{\mathrm{ep}}$ in terms of $\tilde{H}_e$ that is logically equivalent to H_{ep} in terms of H_e (Appendix B shows that the resulting linear electronic matrix element is equal to that of the usual system with states $|\Phi_l\rangle$ and Hamiltonian H_e). To first order, the total transition rate out of an initial state $|\tilde{\Phi}_i\rangle|X_{in}\rangle$ into some final state $|\tilde{\Phi}_f\rangle|X_{fn'}\rangle$ with the carrier in the defect and an arbitrary phonon configuration is given by Fermi's golden rule:

$$\Gamma_{in} = \frac{2\pi}{\hbar}\sum_{fn'}\big|\langle X_{fn'}|\langle\tilde{\Phi}_f|\big(H_1^{\mathrm{BO}} + \tilde{H}_{\mathrm{ep}}\big)|\tilde{\Phi}_i\rangle|X_{in}\rangle\big|^2 \times \delta(\Theta'_{n'} - \Theta_n + W_{if}). \tag{33}$$

For the first term, we plug in the equivalent version of Eq. (31) in terms of $|\tilde{\Phi}_l\rangle$ and $\tilde{H}_e$ and expand the electronic matrix element to first order in all perturbations to get

$$\big\langle\tilde{\Phi}_f(\{\mathbf{R}_k\})\big|H_1^{\mathrm{BO}}(\{\mathbf{R}_k\})\big|\tilde{\Phi}_i(\{\mathbf{R}_k\})\big\rangle \approx \big\langle\tilde{\Phi}_f\big(\{\mathbf{R}_k^{(i)}\}\big)\big|H_1^{\mathrm{BO}} \times \big(\{\mathbf{R}_k^{(i)}\}\big)\big|\tilde{\Phi}_i\big(\{\mathbf{R}_k^{(i)}\}\big)\big\rangle. \tag{34}$$

Combining this with a linear approximation of $\tilde{H}_{\mathrm{ep}}$ for the second term gives the total result

$$\Gamma_{in} = \frac{2\pi}{\hbar}\sum_{fn'}\big|M_e^{\mathrm{BO}}\big(\{\mathbf{R}_k^{(i)}\}\big)\big|^2|\langle X_{fn'}|X_{in}\rangle|^2\delta(\Theta'_{n'} - \Theta_n + W_{if}) + \frac{2\pi}{\hbar}\sum_{fn'}\Bigg|\sum_{k'}\langle\tilde{\Phi}_f\big(\{\mathbf{R}_k^{(i)}\}\big)\big|\big[\nabla_{\mathbf{R}_{k'}}\tilde{H}_e\big]_{\{\mathbf{R}_k^{(i)}\}}\big|\tilde{\Phi}_i\big(\{\mathbf{R}_k^{(i)}\}\big)\big\rangle \times \langle X_{fn'}|\big(\mathbf{R}_{k'} - \mathbf{R}_{k'}^{(i)}\big)|X_{in}\rangle\Bigg|^2\delta(\Theta'_{n'} - \Theta_n + W_{if}), \tag{35}$$

where $M_e^{\mathrm{BO}}(\{\mathbf{R}_k^{(i)}\}) = \langle\tilde{\Phi}_f(\{\mathbf{R}_k^{(i)}\})|H_1^{\mathrm{BO}}(\{\mathbf{R}_k^{(i)}\})|\tilde{\Phi}_i(\{\mathbf{R}_k^{(i)}\})\rangle$. The first term in Eq. (35) is a zeroth-order contribution and the second term is a first-order contribution, where order refers to the power of $\mathbf{R}_{k'} - \mathbf{R}_{k'}^{(i)}$ in the phonon matrix element [8]. The present derivation ensures internal consistency in the two terms, which removes the extra term shown in Eq. (19) in Ref. [8]. In Ref. [8], the zeroth-order and first-order terms were designated as adiabatic and nonadiabatic, respectively. The above derivation by Huang [6], however, reveals that both terms are fully within the adiabatic, Born-Oppenheimer approximation because the derivatives of the electronic wave functions with respect to nuclear coordinates are converted to derivatives of the electronic Hamiltonian.

The theory of capture and how the electronic transition interacts with the phonon vibrations is fundamentally many body. However, as shown in the Appendixes and in the sections below, the electronic matrix elements are all evaluated with single-electron wave functions. It is conventional, therefore, to refer to the single-electron wave functions as electronic states. From now on, we adopt this convention and refer to the many-electron states as electronic configurations, as will be defined later.

C. Zeroth-order transition rate from time-domain integration

The zeroth-order transition rate is

$$\Gamma_{in}^{(0)} = \frac{2\pi}{\hbar}\big|M_e^{\mathrm{BO}}\big(\{\mathbf{R}_k^{(i)}\}\big)\big|^2\sum_{fn'}|\langle X_{fn'}|X_{in}\rangle|^2\delta(\Theta'_{n'} - \Theta_n + W_{if}). \tag{36}$$

This transition rate can be expressed in terms of the line-shape function $F_i^{(0)}$,

$$\Gamma_i^{(0)} = \frac{2\pi}{\hbar} g \left| M_{\mathrm{e}}^{\mathrm{BO}}\left(\{\mathbf{R}_k^{(i)}\}\right)\right|^2 F_i^{(0)}, \tag{37}$$

where $F_i^{(0)}$ is derived in Appendix C. The initial-phonon-state index n is dropped because it is assumed to be the thermal equilibrium distribution by default. The degeneracy factor g results from summing over (nearly) degenerate final electronic states [7]. For the two examples considered in this work, $g = 4$ for GaN:C_N [7] and $g = 1$ for the model Si dangling bond.

To reach faster convergence, the transition rates are converted to an integration over time, as detailed in Appendix C. The kernel of the time integration is a generalization of the so-called "generating function" [11–14], to the case of different initial and final phonon frequencies,

$$G^{(0)}(t) = \exp\left[i \sum_j D_j^{(0)}(t) \right], \tag{38}$$

where

$$D_j^{(0)}(t) = -2\left[i S_j^{-1} A_j(t) - {S'_j}^{-1} \cot(\omega'_j t/2) \right]^{-1}, \tag{39}$$

$$A_j(t) = \frac{e^{i\omega_j t}(\bar{n}_j + 1) + \bar{n}_j}{e^{i\omega_j t}(\bar{n}_j + 1) - \bar{n}_j}, \tag{40}$$

$$S_j = \frac{\omega_j}{2\hbar}(\Delta q_j)^2 \tag{41}$$

is the Huang-Rhys factor, and

$$\bar{n}_j = \frac{1}{e^{\beta\hbar\omega_j} - 1} \tag{42}$$

is the Bose-Einstein distribution. This more general form simplifies to the previous time-domain formulation [9,11] when $\omega_j = \omega'_j$, as shown in Appendix C. A Lorentzian smearing, corresponding to the lifetime broadening, is added to the electronic energy levels to ensure convergence. In the time domain, the smearing appears as a factor $e^{-\gamma t}$ in the integrand. The final expression for the zeroth-order transition rate is

$$\Gamma_i^{(0)} = \frac{2}{\hbar^2} g \left| M_{\mathrm{e}}^{\mathrm{BO}}\left(\{\mathbf{R}_k^{(i)}\}\right)\right|^2 \mathrm{Re} \int_0^{+\infty} G^{(0)}(t) e^{\frac{iW_{if}t}{\hbar} - \gamma t} dt. \tag{43}$$

To evaluate the electronic matrix element, Appendix A shows that

$$\begin{aligned} M_{\mathrm{e}}^{\mathrm{BO}}\left(\{\mathbf{R}_k^{(i)}\}\right) &\equiv \left\langle \tilde{\Phi}_f\left(\{\mathbf{R}_k^{(i)}\}\right) \middle| H_1^{\mathrm{BO}}\left(\{\mathbf{R}_k^{(i)}\}\right) \middle| \tilde{\Phi}_i\left(\{\mathbf{R}_k^{(i)}\}\right) \right\rangle \\ &\approx \left\langle \phi'_f \middle| \phi_i^0 \right\rangle \left[E_f\left(\{\mathbf{R}_k^{(i)}\}\right) - E_i\left(\{\mathbf{R}_k^{(i)}\}\right) \right]. \end{aligned} \tag{44}$$

The last expression involves an overlap integral of two known Kohn-Sham orbitals and an electronic energy difference, which are both readily available from DFT calculations. The evaluation of the overlap integral is described in detail in Appendix C of Ref. [8].

D. First-order transition rate

The first-order transition rate can be written in terms of the phonon eigenmodes q_j:

$$\Gamma_{in}^{(1)} = \frac{2\pi}{\hbar} \sum_{fn'} \left| \sum_j M_j \langle X_{fn'} | q_j | X_{in} \rangle \right|^2 \delta(\Theta'_{n'} - \Theta_n + W_{if}). \tag{45}$$

In this expression, the first-order electronic matrix elements M_j are defined in terms of the many-electron wave functions $|\Phi_i\rangle$ and $|\Phi_f\rangle$ in the relaxed initial positions and are derived in Appendix B to be

$$M_j = \langle \Phi_f | \partial_{q_j} H_{\mathrm{e}} | \Phi_i \rangle. \tag{46}$$

The single-electron expression for M_j is more complex than that for the zeroth order, whereby we refer the reader to Appendix B.

We can expand the norm square of the sum in Eq. (45) as

$$\left| \sum_j M_j \langle X_{fn'} | q_j | X_{in} \rangle \right|^2 = \sum_{jj'} M_j \langle X_{fn'} | q_j | X_{in} \rangle M_{j'}^* \langle X_{in} | q_{j'}^* | X_{fn'} \rangle. \tag{47}$$

The canonical coordinates q_j have the form $c_j e^{-i\omega_j t}$, where $c_j = |c_j| e^{i\varphi_j}$ are time-independent complex constants, whose phases φ_j are randomly distributed. The phases cancel when $j = j'$, but the cross terms contain a phase difference $e^{i(\varphi_j - \varphi_{j'})}$. Over many random configurations (or, equivalently, a sufficiently long time during which many phonons are created and annihilated), these cross terms average to zero, so we can write the transition rate as

$$\Gamma_{in}^{(1)} = \frac{2\pi}{\hbar} \sum_{fn'} \sum_j \left| M_j \langle X_{fn'} | q_j | X_{in} \rangle \right|^2 \delta(\Theta'_{n'} - \Theta_n + W_{if}). \tag{48}$$

Following the same procedure as for the computation of the zeroth-order transition rate, the first-order term is converted into an integration over time,

$$\Gamma_i^{(1)} = \frac{2}{\hbar^2} g \,\mathrm{Re} \int_0^{+\infty} \sum_{j'} \frac{G_{j'}^{(1)}(t)}{Z_{j'}} \left[\prod_{j \neq j'} \frac{G_j^{(0)}(t)}{Z_j} \right] e^{i\frac{W_{if}}{\hbar}t} \, dt, \tag{49}$$

where

$$\begin{aligned} G_{j'}^{(1)}(t) = |M_{j'}|^2 \sum_{n'_{j'}} \sum_{n_{j'}} \left| \left\langle \chi'_{n'_{j'}} \middle| q_{j'} \middle| \chi_{n_{j'}} \right\rangle \right|^2 \\ \times e^{i(n'_{j'}\omega'_{j'} - n_{j'}\omega_{j'})t} e^{-\beta\left(n_{j'} + \frac{1}{2}\right)\hbar\omega_{j'}}. \end{aligned} \tag{50}$$

The transition rate in Eq. (49) can alternatively be written as

$$\Gamma_i^{(1)} = \frac{2}{\hbar^2} g \,\mathrm{Re} \int_0^{+\infty} \left[\frac{\partial}{\partial\lambda} \prod_{j=1}^{N} \frac{G_j^{(0)}(t) + \lambda G_j^{(1)}(t)}{Z_j} \bigg|_{\lambda \to 0} \right] \times e^{i\frac{W_{if}}{\hbar}t} \, dt. \tag{51}$$

We evaluate $G_j^{(1)}(t)$ in a similar manner as $G_j^{(0)}(t)$ and find

$$G_j^{(1)}(t) = |M_j|^2 G_j^{(0)}(t) D_j^{(1)}(t), \tag{52}$$

where

$$D_j^{(1)}(t) = -\left[\frac{\Delta q_j}{2S_j}\right]^2 D_j^{(0)}(t)\left\{D_j^{(0)}(t)[A_j(t)]^2 - \frac{1}{2}\frac{\omega_j}{\omega_j'} \times \left[A_j(t)\cot\frac{\omega_j' t}{2} - \frac{\omega_j \cot(\omega_j' t/2) - i\omega_j' A_j(t)}{\omega_j A_j(t) + i\omega_j' \cot(\omega_j' t/2)}\right]\right\}. \tag{53}$$

Plugging back into Eq. (51), simplifying, and inserting the same Lorentzian smearing as in the zeroth-order term, we finally obtain

$$\Gamma_i^{(1)} = \frac{2}{\hbar^2} g \,\mathrm{Re}\int_0^{+\infty}\left[\sum_j |M_j|^2 D_j^{(1)}(t)\right] G^{(0)}(t) e^{\frac{iW_{if}t}{\hbar} - \gamma t} dt. \tag{54}$$

III. COMPUTATIONAL CONSIDERATIONS

All electronic-structure and phonon calculations are performed in identical supercells that correspond to the initial and final atomic and electronic configurations of the capture process. In particular, for the cases we study in this paper, the supercells are neutral: In the initial configuration, a carrier is in a band state while the point defect is in a nominally charged state so that, when the carrier transitions into a defect energy level in the gap, the point defect is in the neutral state in the resulting final configuration. The final configuration is the ground state of the system, whereby the appropriate electronic structure calculations are based on the ground-state DFT. In contrast, the initial configuration is an excited state of the same system, whereby constrained-DFT calculations are appropriate. The total DFT energies of the initial and final configurations are precisely the energies W_i and W_f, respectively, in Eq. (23) that appear in the δ function of the Fermi golden rule, as in Eq. (33) and other subsequent expressions for transition rates.

More specifically, in the case of Si in the initial configuration, an electron is in a conduction-band state while the dangling bond is in a nominally positively charged state, i.e., both the spin-up and spin-down defect energy levels in the gap are empty. During the capture process, the electron transitions into one of these two empty levels, so that, in the final configuration, one of the two energy levels in the gap is occupied and the dangling bond is in its neutral state. Similarly, in the case of GaN in the initial configuration, a hole is in a valence-band state while the C_N defect is in a nominally negatively charged state, i.e., all of the six p-like defect energy levels in the gap are occupied. During the capture process, the hole transitions into one of these occupied energy levels, so that, in the final configuration, one of the six defect levels in the gap is empty and the C_N defect is in its neutral state.

All calculations are based on DFT with projector augmented-wave (PAW) pseudopotentials [15,16] and are performed using the computer code VASP [16–19]. In the Si calculations, the Perdew-Burke-Ernzerhof (PBE) form [20] of the generalized-gradient approximation is used to get the wave functions and conduction-band eigenvalues across multiple k points. These eigenvalues are used only for the energies of conduction-band states relative to the band minimum. We also use the PBE functional for relaxations and phonon calculations. On the other hand, we use the hybrid functional of Heyd, Scuseria, and Ernzerhof (HSE) [21] to get accurate total energies, from which the capture transition energy is obtained, and the bound-state defect energy level. The functional is adapted by tuning to $\alpha = 0.23$ to reproduce the experimental band gap. For GaN, we use HSE for all relaxation, electronic structure, and phonon calculations with $\alpha = 0.31$, which reproduces the experimental band gap and produces the same lattice relaxation around the C impurity as in Ref. [7].

Neither pure Si nor pure GaN is magnetic. However, in any carrier-capture process, the defect has an unpaired spin in either the initial or final state. Therefore, all defect calculations are spin polarized. The free carrier in the initial state also has a spin; however, its interaction with the defect spin is negligible. In a finite-size supercell, a spin-polarized free carrier would artificially interact with the defect spin because of the confinement of the supercell. To remove this artifact, free carriers are treated as non-spin-polarized (i.e., with equal occupation of 0.5 in both spin channels). In all calculations, we completely turn off symmetry.

We get the total potential energies $W_i(\{\mathbf{R}_k^{(i)}\})$ and $W_f(\{\mathbf{R}_k^{(f)}\})$ directly from the VASP total-energy output in the initial and final relaxed configurations, for which we use constrained-DFT and ground-state-DFT calculations, respectively, in neutral supercells. One should ideally do separate total-energy calculations with the carrier in each initial band state. Such an undertaking, however, is not practical. Instead, we approximate the potential energy difference W_{if} as

$$W_{if} \approx W_f\left(\{\mathbf{R}_k^{(f)}\}\right) - \left[W_{\mathrm{ref}}\left(\{\mathbf{R}_k^{(i)}\}\right) + \varepsilon_i - \varepsilon_{\mathrm{ref}}\right], \tag{55}$$

where the total energy is calculated only for a single reference initial state for each system, as described below. The assumption here is that the ion-ion interaction potential does not change significantly between band states so that the additional energy is only the eigenvalue energy difference between each initial band state and the reference band state.

In the case of Si in the initial configuration, the electronic-structure calculation for the reference configuration is carried out in the constrained-DFT scheme with the electron at the bottom of the conduction band. The resulting total energy is then W_{ref}. Since in the case of Si we calculate nonequilibrium capture coefficients for electrons that approach the defect in a perfect-crystal Bloch state, the eigenvalue difference $\varepsilon_i - \varepsilon_{\mathrm{ref}}$ is obtained from a perfect-crystal calculation, with $\varepsilon_{\mathrm{ref}}$ being the energy of the conduction-band minimum. In the final-state configuration, the band electron occupies the defect level in the energy gap, i.e., the defect is neutral and a standard ground-state-DFT calculation is performed.

In the case of GaN in the initial configuration, the reference configuration is defined as a hole in the third highest valence state at the Γ point. The choice of the third highest valence band is made because it allows easier convergence than the top two bands that are closer in energy. Rather than converging the Γ point self-consistently, however, it is more

accurate to converge the $2 \times 2 \times 2$ Monkhorst-Pack (MP) k-point mesh, then calculate the Γ-point eigenvalues and wave functions from that charge density. The electronic-structure calculations are once more performed using the constrained-DFT scheme, with the hole evenly split among the third highest valence states of the $2 \times 2 \times 2$ MP k points, then the total energy is adjusted by the energy difference between the third highest band of those k points versus that of the Γ point to get the total energy of the Γ-point reference configuration. The energies of the three hole states are then calculated from the reference configuration by adding the relevant eigenvalue differences. The final results are the Boltzmann-weighted average over the three hole states,

$$\Gamma^{(1)}(T) = \frac{\sum_i \Gamma_i^{(1)}(T) e^{W_{if}/k_B T}}{\sum_i e^{W_{if}/k_B T}}, \tag{56}$$

where W_{if} is negative, representing energy transfer into the phonon vibrations. In the final configuration, the excess electron on the defect is removed and the valence band states are fully occupied, i.e., the defect is neutral and a standard ground-state-DFT calculation is performed. Since in GaN we are calculating the hole-capture coefficient under equilibrium conditions in order to compare with available experiments and prior calculations, the eigenvalue differences are evaluated in the presence of the defect. The band eigenvalues from the ground-state system are most accurate, so we use those for all calculations where only band eigenvalues are needed.

Due to hybridization with the valence and conduction bands in both Si and GaN, the bound state in the gap has a nonzero dispersion in the Brillouin zone. The dispersion in Si is of order 0.1 eV in the PBE calculation and should be smaller with HSE, whereas in the case of GaN, because of the smaller supercell but larger band gap, the dispersion is comparable, of order 0.2 eV. This fictitious dispersion is a finite-cell artifact that goes to zero with larger supercells. Here, the defect-band state at the Γ point is adopted as the defect energy level in the gap and is used for all total energies and eigenvalue differences that involve the defect level.

To calculate the zeroth-order electronic matrix elements given in Eq. (44), we need to isolate the electronic-only energy difference $E_f(\{\mathbf{R}_k^{(i)}\}) - E_i(\{\mathbf{R}_k^{(i)}\})$. We utilize the total potential energy difference

$$\begin{aligned} W_f\big(\{\mathbf{R}_k^{(i)}\}\big) - W_i\big(\{\mathbf{R}_k^{(i)}\}\big) &= \big[V_{\rm II}\big(\{\mathbf{R}_k^{(i)}\}\big) + E_f\big(\{\mathbf{R}_k^{(i)}\}\big)\big] \\ &\quad - \big[V_{\rm II}\big(\{\mathbf{R}_k^{(i)}\}\big) + E_i\big(\{\mathbf{R}_k^{(i)}\}\big)\big] \qquad (57) \\ &= E_f\big(\{\mathbf{R}_k^{(i)}\}\big) - E_i\big(\{\mathbf{R}_k^{(i)}\}\big) \qquad (58) \end{aligned}$$

taken directly from the total energy output by VASP. We utilize an approximation similar to that used for the energy argument in the δ function to consider many initial band states:

$$\begin{aligned} W_f\big(\{\mathbf{R}_k^{(i)}\}\big) - W_i\big(\{\mathbf{R}_k^{(i)}\}\big) \approx W_f\big(\{\mathbf{R}_k^{(i)}\}\big) - \big[W_{\rm ref}\big(\{\mathbf{R}_k^{(i)}\}\big) \\ + \varepsilon_i - \varepsilon_{\rm ref}\big]. \end{aligned} \tag{59}$$

The calculation of the first-order electronic matrix elements in terms of one-electron wave functions and their derivatives with respect to q entails two parts: the calculation of the pertinent phonon states to perform the derivatives and the calculation of the electronic matrix elements themselves. For the phonons, we use PHONOPY [22] with VASP as the force calculator. For the calculation of the electronic matrix elements, the reduction from many-electron to one-electron wave functions is carried out in Appendix B.

The process of expressing the first-order many-electron matrix elements in Eq. (46) in terms of one-electron wave functions imposes a requirement that the one-electron wave functions must be eigenfunctions of the same Hamiltonian, namely the same electronic configuration. We show in Appendix B that using the single-electron states from the initial configuration yield simpler equations than using those of the final configuration. We then transform the matrix element into a single-electron overlap multiplied by an eigenvalue difference $\varepsilon_i - \varepsilon_f$. We show in Appendix B that the two possible choices for the single-electron overlap—initial-configuration and final-configuration wave functions—are in fact equivalent to first order. We opted for the final-configuration wave functions, which are obtained by a ground-state calculation of the neutral defect. DFT is a ground-state variational theory, so the ground-state electronic configuration is generally more accurate than constrained DFT calculations. This choice also has added bonuses: A Sommerfeld factor [7] to correct for small-supercell effects is not warranted because the defect is neutral, and only a single set of calculations needs to be performed for the single final configuration rather than all three of the initial configurations that we consider. On the other hand, no such simplification exists for the eigenvalue difference. Therefore, its evaluation remains in the relaxed initial configuration. As in Eqs. (55) and (59), we calculate these initial-configuration eigenvalue differences using a single reference state and adjust for the splitting between the different initial states using the more accurate final-state eigenvalue differences:

$$\varepsilon_i - \varepsilon_f = (\varepsilon_{\rm ref} - \varepsilon_f)^{\rm initial} + (\varepsilon_i - \varepsilon_{\rm ref})^{\rm final}. \tag{60}$$

The value of the Lorentzian smearing, $\hbar\gamma \approx 1$ meV, for the Si system is taken from the theoretical estimate of the lifetime broadening of electrons in the Si conduction band due to Coulomb interactions [23]. For the GaN system, we use $\hbar\gamma \approx 3.7$ meV, based on an experimental measurement of the half width at half maximum of the band edge peak in the GaN photoluminescence spectrum [24]. The convergence criterion for the time-domain integral is set to $e^{-\gamma t} \leqslant 10^{-4}$ for Si and $e^{-\gamma t} \leqslant 10^{-5}$ for GaN. The time step was decreased to convergence, resulting in $dt = 2.1 \times 10^{-9}$ ps for Si and $dt = 5 \times 10^{-9}$ ps for GaN. The oscillatory nature of the integrand is limited by the highest possible phonon frequency, which is independent of temperature.

IV. RESULTS

A. Electron capture by triply hydrogenated vacancy in Si

In the previously published work [8] the zeroth-order electron-capture cross section by a nominally positively charged defect in this system at 300 K was calculated using a Monte Carlo sampling method for phonon configurations. Here, we use the more accurate time-domain integration to calculate both the zeroth- and first-order terms.

In all calculations, we use a $4 \times 4 \times 4$ supercell of the 8-atom conventional fcc unit cell, comprising 512 atomic sites,

and a plane-wave energy cutoff of 400 eV. For relaxations and the first-order matrix elements, we use only the Γ k point. For the zeroth-order electronic matrix element $M_e^{\mathrm{BO}}(\{\mathbf{R}_k^{(i)}\})$, we use a $3 \times 3 \times 3$ Γ-centered k-point mesh to obtain matrix elements for more energy levels.

We calculate the electron-capture cross section for a range of initial states i using the relationship

$$\sigma_i = \frac{\Gamma_i \Omega}{v_{g,i}}, \tag{61}$$

where Ω is the volume over which the wave functions are normalized, i.e., the supercell volume, and $v_{g,i} = 1/\hbar|\nabla_{\mathbf{k}}\varepsilon_i(\mathbf{k})|$ is the group velocity of the incident electron in the perfect crystal. To estimate the group velocity at a given k point, we calculate the energy eigenvalues at four nearby k points, with $\delta k = -0.02$, -0.01, 0.00, and 0.01 $2\pi/a$, where $\delta k = 0.00$ is the original k point. Four k points are needed to properly line up the energy bands as they cross and can be linear or quadratic. When band crossing occurs, the bands can be matched correctly using symmetry considerations. The derivative of energy with respect to k in each direction k_x, k_y, and k_z was computed using a least-squares fit. As a check, numerical derivatives produce group velocities that are consistent with the point-group symmetries.

The final cross section as a function of the initial state is jagged due to having discrete band levels in a finite supercell. To smooth the curve, we apply Gaussian broadening, where each point at a calculated eigenenergy is smeared according to a Gaussian function with a width η based on the group velocity $v_{g,i}$ at that energy,

$$\eta = \hbar v_{g,i} \Delta k, \tag{62}$$

where Δk is the distance between points in the k mesh.

The capture cross section as a function of the electronic transition energy $E_f(\{\mathbf{R}_k^{(i)}\}) - E_i(\{\mathbf{R}_k^{(i)}\})$ is shown in Fig. 1. The zeroth-order calculation incorporates several improvements over the previous work [8]. In the present work, we include more accurate energies using a hybrid exchange functional, spin polarization, and the same neutral supercell for the initial and final atomic and electronic configurations of the capture process. We also implement an improved phonon line-shape function using the time-integration method, which sums over all possible phonon configurations, and we use independent, rather than averaged, matrix elements for each band state. The electron group velocity is also more accurate as it is calculated directly from the band structure.

With these improvements, the calculated cross section at the conduction-band minimum is roughly an order of magnitude smaller than the previous value [8], but it has the same qualitative features, including the exponential decay of the cross section as a function of the incident electron energy [8]. The results also show that the cross section is dominated by the zeroth-order contribution, as previously estimated in Ref. [8].

B. Hole capture by substitutional carbon in GaN

The yellow luminescence band in high-purity, n-type GaN is attributed to hole capture by negatively charged isolated C_N defects [7,25]. More specifically, photogenerated holes in the valence bands are captured by the defects (the hole capture of interest here), making the defects neutral. Recombination of the trapped holes with conduction-band electrons emits the yellow luminescence. The hole-capture coefficient can be experimentally measured from the yellow luminescence spectrum [26–29].

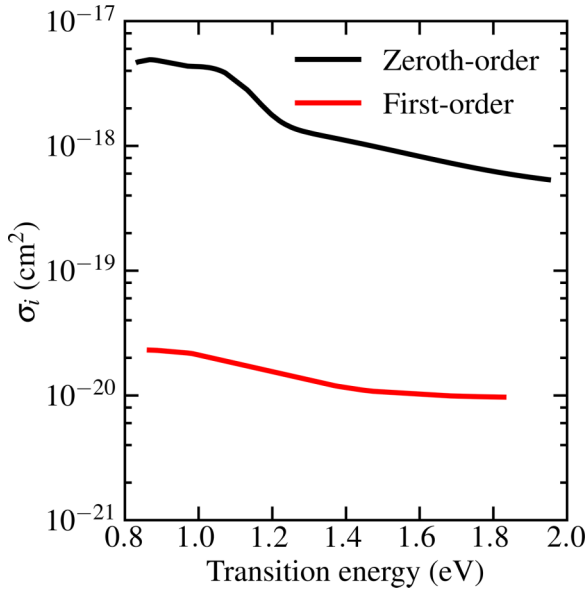

FIG. 1. Zeroth- and first-order contribution to electron-capture cross section (σ_i) of a triply hydrogenated, positively charged vacancy in Si (a model dangling bond) as a function of the transition energy, namely the total electronic energy difference $E_f(\{\mathbf{R}_k^{(i)}\}) - E_i(\{\mathbf{R}_k^{(i)}\})$ as defined in Eq. (58).

In order to compare directly with the results from Alkauskas *et al.* [7], we use a 96-atom $3 \times 2 \times 2$ supercell of the wurtzite conventional unit cell (see the example of a supercell structure in Zhao *et al.* [30]) with their cell parameters $a = 3.20$ Å, $c = 5.19$ Å, and $u = 0.377$. The relaxation and phonon calculations are performed using a $2 \times 2 \times 2$ Monkhorst-Pack (MP) k-point mesh with an energy cutoff of 400 eV for the plane waves. These choices reproduce the relaxation around the C impurity found by Alkauskas *et al.* [7]. For the calculation of the total-energy differences and the band-energy splittings at the Γ point, we carried out a Γ-point calculation using the self-consistent charge density obtained in the more accurate $2 \times 2 \times 2$ MP mesh, instead of calculating the Γ point alone self-consistently. However, for computational feasibility, the first-order matrix elements are calculated using a Γ-only k-point mesh and a plane-wave cutoff of 600 eV for increased accuracy in the wave-function derivatives.

The top of the GaN valence bands comprises a doublet and a singlet. The presence of a point defect with large lattice relaxation splits the doublet into two singlets (increasing the supercell size, which is not currently feasible, would reduce this splitting). The calculated W_{if} for the capture of a hole at the highest, second highest, and third highest valence band states are 0.92, 0.94, and 1.17 eV, where the splitting of 0.02 eV between the top two levels is an artifact of the finite-size supercell. We only calculated hole capture from these

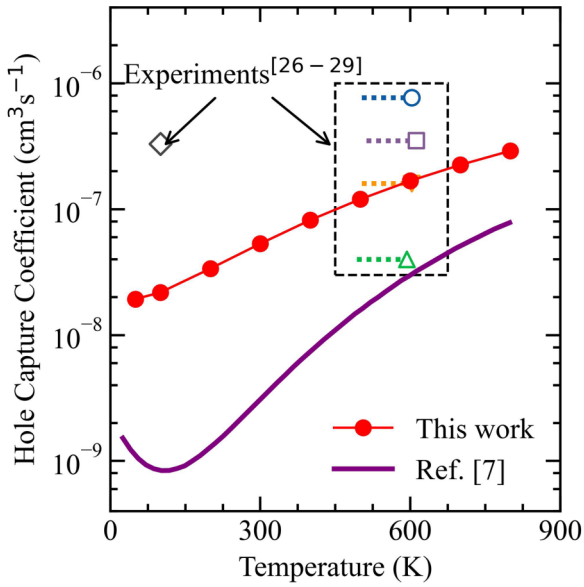

FIG. 2. The equilibrium hole-capture coefficient for C impurities at N sites in GaN. The hole-capture coefficient reported by Alkauskas *et al.* [7] at $W_{if} = 1.02$ eV (ΔE in their notation) is also displayed in purple. Reported experimental values of capture coefficients are displayed as open symbols, and dotted lines indicate the range of temperatures in which they are reliable [26–29]. The four different symbols in the dashed-line rectangle are data extracted from four different samples.

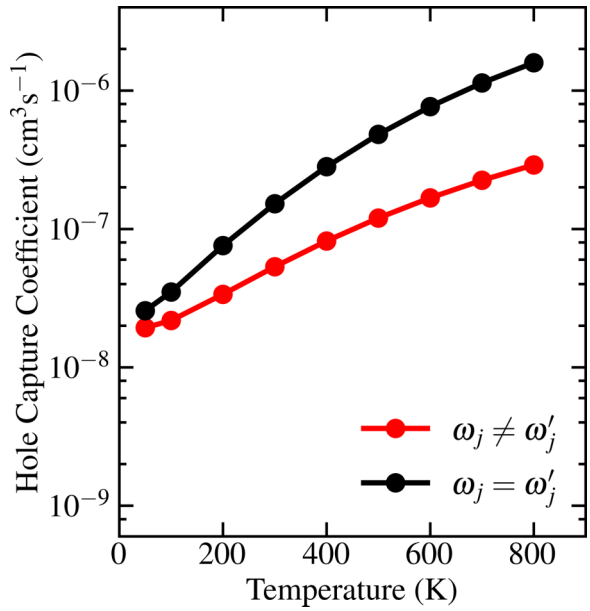

FIG. 3. Comparison plot of the hole-capture coefficient calculated using the phonon frequencies of the initial and final configurations ($\omega_j \neq \omega'_j$) and invoking the approximation of using the phonon frequencies of the final configuration for both configurations ($\omega_j = \omega'_j$).

three split levels in order to compare our results directly with those of Alkauskas *et al.* [7].

To compare our results with experimental data [26–29], we calculate the hole-capture coefficient C_p, which is equal to the transition rate divided by the C_N-defect concentration, or, equivalently,

$$C_p = \Gamma\Omega, \tag{63}$$

where Ω is the volume over which the wave functions are normalized (i.e., the supercell volume). Assuming that convergence with respect to supercell size is reached, in the limit of a single hole captured in a single defect, C_p in Eq. (63) represents the theoretical value to be compared with experimental data, even though the experimental defect concentrations are much smaller than the theoretical concentration in the supercell calculations.

The experimental values [26–29] of the hole-capture coefficient range from 0.4×10^{-7} to 7.7×10^{-7} cm^3/s with an average of 3.3×10^{-7} cm^3/s in four samples in the temperature range of 500–600 K, and 3.3×10^{-7} cm^3/s at 100 K. The hole-capture coefficients estimated from experiments show a negligible temperature dependence. The experimental values of the hole-capture coefficient are estimated from an analysis of photoluminescence intensities. In photoluminescence, hole capture is considered to be an equilibrium process, in which the zeroth-order term is zero, so only the first-order equilibrium term is included in Fig. 2. The results show that the hole-capture coefficient features an increase of around an order of magnitude from 50 to 800 K, in good agreement with the high-temperature experimental data (500–800 K), but a little over an order of magnitude lower than the single experimental point available at lower temperatures. It is worth noting that the approximation of using the same set of phonon frequencies in the initial and final state does impact the magnitude and temperature dependence of the capture coefficient. The difference is very small at low temperatures and approaches an order of magnitude at higher temperatures, as shown in Fig. 3.

We compare the present equilibrium hole-capture coefficient to that calculated by Alkauskas *et al.* [7], as shown in Fig. 2. The present capture coefficient is larger by more than an order of magnitude at low temperatures but within a factor 4 at high temperatures, being closer overall to the sparse experimental data. It also shows a much weaker temperature dependence. In order to understand the differences between the two results, we consider the most significant differences between the two treatments.

The first major difference between the two calculations is the inclusion of all phonons in the multiphonon configurations by which the transition energy is dissipated instead of a single effective phonon mode. The existence of significantly more multiphonon configurations to dissipate the transition energy is the major contributor to the overall larger capture-coefficient values in the present calculations. Moreover, the weaker temperature dependence of the present results at low temperatures is caused by the fact that there always exist many phonons at temperatures below $k_B T$.

A second major difference between the two calculations is that, in the present formulation, inclusion of the Sommerfeld correction is not necessary. As shown in Appendix B, the calculation of the overlaps using either the final- or initial-configuration wave functions is equivalent to first order. Thus, our results shown in Fig. 2 do not contain a Sommerfeld correction.

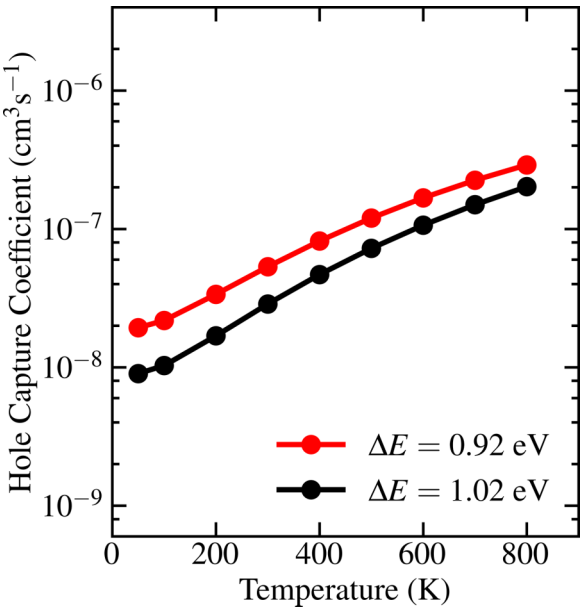

FIG. 4. Comparison between the calculated $\Delta E = 0.92$ eV and a hypothetical $\Delta E = 1.02$ eV, showing minimal sensitivity to uncertainties in the defect energy level.

A final difference between the current multiphonon formulation and the single-phonon-mode scheme of Ref. [7] is the lack of sensitivity to the accuracy of the calculated transition energy ΔE. Whereas in that scheme a 10% change in ΔE induces a one-to-two orders of magnitude change in the capture coefficient, we find that in the present multiphonon scheme such a change induces a change of only a factor $\sim$2 (Fig. 4).

V. SUMMARY

In summary, we have developed a first-principles methodology to study nonradiative carrier capture by defects in semiconductors using a comprehensive theory of inelastic multiphonon carrier-capture processes, identifying both zeroth-order and first-order contributions, with both terms (to first order, as shown by Huang [6]) being within the adiabatic, Born-Oppenheimer approximation. An accurate, fast-converging time-integration method is implemented to calculate the phonon line-shape function. We present results for a triply hydrogenated vacancy in Si and a C_N substitutional impurity in GaN.

For Si, we calculate both the zeroth-order and first-order nonequilibrium contributions to the carrier-capture process and compare the relative magnitude of each. We see good agreement with our previous results [8]. For the C_N substitutional impurity in GaN, we calculate the equilibrium capture coefficient, for which only the first-order contribution is nonzero, to compare with available sparse experimental data [26–29] and previous first-principles calculations [7]. We find a larger capture coefficient, closer to the experimental data, and a decrease in the temperature dependence, both of which can be attributed to the inclusion of all phonons in constructing multiphonon configurations to dissipate the transition energy. Discrepancies between the theoretical and experimental results still remain, especially at low temperatures. Additional experiments are needed to test the theories.

From our investigation, we believe that increasing the accuracy of the electronic energies through the use of larger supercells would have the largest impact on improving the calculations. However, scaling to larger supercells while still using the more accurate hybrid functionals is currently limited by computational feasibility. The primary limiting factors for supercell size are the phonon calculation and, for the first-order term, derivatives of electronic wave functions with respect to each phonon mode, especially using high-level exchange functionals. Each of these steps require $3N - 3$ number of DFT calculations, where N is the number of atoms in the supercell. Given that a typical DFT calculation scales as $O(N^3)$, the total computational time scales as $O(N^4)$. Consequently, extending the model to more complex defect structures faces the fundamental challenge common to all electronic structure calculations: accelerating the $O(N^3)$ DFT step.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation under Grant No. 1508898. It was also supported as part of the Center for Molecular Magnetic Quantum Materials, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award No. DE-SC0019330, by the Air Force Office of Scientific Research Center of Excellence on Radiation Effects via Grant No. FA9550-22-1-0012, and by the McMinn Endowment at Vanderbilt University. L.R.N. was supported by a Department of Energy Computational Science Graduate Fellowship under Award No. DE-SC0021110. Computing at Vanderbilt was performed at the Department of Defense High Performance Computing Modernization Program (HPCMP) and at the Department of Energy National Energy Research Scientific Computing Center (NERSC).

APPENDIX A: ZEROTH-ORDER ELECTRONIC MATRIX ELEMENTS

The zeroth-order matrix element is defined as

$$M_e^{BO} = \langle \tilde{\Phi}_f | H_1^{BO} | \tilde{\Phi}_i \rangle. \tag{A1}$$

The wave functions $|\tilde{\Phi}_l\rangle$ are not directly available for computation, so we must transform the matrix element to use only the defect wave functions $|\Phi_l\rangle$ and the perfect-crystal wave functions $|\Phi_l^0\rangle$. The defect wave functions $|\Phi_l\rangle$ can be written as a first-order perturbative expansion in terms of $|\tilde{\Phi}_l\rangle$,

$$|\Phi_l\rangle \approx |\tilde{\Phi}_l\rangle + \sum_{l' \neq l} \frac{\langle \tilde{\Phi}_{l'} | H_1^{BO} | \tilde{\Phi}_l \rangle}{\tilde{E}_l - \tilde{E}_{l'}} |\tilde{\Phi}_{l'}\rangle \tag{A2}$$

or

$$|\Phi_l\rangle \approx |\tilde{\Phi}_l\rangle + \frac{\Delta H_{if}}{\tilde{E}_l - \tilde{E}_i} \delta_{fl} |\tilde{\Phi}_i\rangle - \frac{\Delta H_{fi}}{\tilde{E}_f - \tilde{E}_l} \delta_{il} |\tilde{\Phi}_f\rangle, \tag{A3}$$

where

$$\Delta H_{fi} = \langle \tilde{\Phi}_f | \Delta H | \tilde{\Phi}_i \rangle = \langle \tilde{\Phi}_f | H_1^{BO} | \tilde{\Phi}_i \rangle \tag{A4}$$

is the zeroth-order electronic matrix element. Using this relationship and Eq. (A3), we find that

$$\langle\tilde{\Phi}_f|H_1^{\mathrm{BO}}|\tilde{\Phi}_i\rangle = \langle\Phi_f|\tilde{\Phi}_i\rangle(\tilde{E}_f - \tilde{E}_i). \tag{A5}$$

Making use of Eqs. (25) and (26), we obtain

$$\langle\Phi_f|\tilde{\Phi}_i\rangle \approx \frac{1}{N!}\langle\det\{\phi_1'\cdots\phi_f'\cdots\phi_N'\}|\det\{\phi_1\cdots\phi_i^0\cdots\phi_N\}\rangle. \tag{A6}$$

Because the perturbation between the Förster wave functions and the defect wave functions is only the coupling between the initial and final many-electron states, we can make the approximation that all of the other states are unchanged under the perturbation and remain orthonormal to ϕ_i^0. Because of the orthonormality, for any single permutation on the left, only the identical permutation on the right survives and the only remaining inner product is $\langle\phi_f'|\phi_i^0\rangle$. We can then cancel the factor of $1/N!$ out front, giving

$$\langle\tilde{\Phi}_f|H_1^{\mathrm{BO}}|\tilde{\Phi}_i\rangle = \langle\phi_f'|\phi_i^0\rangle(\tilde{E}_f - \tilde{E}_i). \tag{A7}$$

Finally, the total electronic energy difference $\tilde{E}_f - \tilde{E}_i$ can be approximated as the total electronic energy difference from the defect system $E_f - E_i$:

$$\langle\tilde{\Phi}_f|H_1^{\mathrm{BO}}|\tilde{\Phi}_i\rangle = \langle\phi_f'|\phi_i^0\rangle(E_f - E_i). \tag{A8}$$

APPENDIX B: FIRST-ORDER ELECTRONIC MATRIX ELEMENT

In order to perform practical calculations of the first-order matrix element, we must show that it can be transformed from being in terms of the Förster wave functions and Hamiltonian, $|\tilde{\Phi}_l\rangle$ and $\tilde{H}_{\mathrm{e}}$, to being in terms of the defect wave functions and Hamiltonian, $|\Phi_l\rangle$ and H_{e}, of the initial ion positions:

$$\begin{aligned}&\langle\Phi_f(\{\mathbf{R}_k^{(i)}\})|\nabla_{\mathbf{R}_{k'}}H_{\mathrm{e}}|\Phi_i(\{\mathbf{R}_k^{(i)}\})\rangle\\ &\quad= \langle\tilde{\Phi}_f(\{\mathbf{R}_k^{(i)}\})|\nabla_{\mathbf{R}_{k'}}\tilde{H}_{\mathrm{e}}|\tilde{\Phi}_i(\{\mathbf{R}_k^{(i)}\})\rangle.\end{aligned} \tag{B1}$$

Moving forward, we drop notating the dependence on $\mathbf{R}_k^{(i)}$ for brevity. From the expansion of $|\Phi_l\rangle$ in terms of $|\tilde{\Phi}_l\rangle$ from Appendix A, we can write

$$\begin{aligned}\langle\Phi_f|\nabla_{\mathbf{R}_{k'}}H_{\mathrm{e}}|\Phi_i\rangle = &\left(\langle\tilde{\Phi}_f| + \frac{\Delta H_{fi}}{\tilde{E}_f - \tilde{E}_i}\langle\tilde{\Phi}_i|\right)\\ &\times\ \nabla_{\mathbf{R}_{k'}}H_{\mathrm{e}}\left(|\tilde{\Phi}_i\rangle - \frac{\Delta H_{fi}}{\tilde{E}_f - \tilde{E}_i}|\tilde{\Phi}_f\rangle\right).\end{aligned} \tag{B2}$$

Expanding and keeping terms that are linear in the matrix elements ΔH_{fi} and $\langle\tilde{\Phi}_l|\nabla_{\mathbf{R}_{k'}}H_{\mathrm{e}}|\tilde{\Phi}_{l'}\rangle$ gives

$$\langle\Phi_f|\nabla_{\mathbf{R}_{k'}}H_{\mathrm{e}}|\Phi_i\rangle \approx \langle\tilde{\Phi}_f|\nabla_{\mathbf{R}_{k'}}H_{\mathrm{e}}|\tilde{\Phi}_i\rangle. \tag{B3}$$

We know that $H_{\mathrm{e}} = \tilde{H}_{\mathrm{e}} + H_1^{\mathrm{BO}}$, so

$$\langle\Phi_f|\nabla_{\mathbf{R}_{k'}}H_e|\Phi_i\rangle = \langle\tilde{\Phi}_f|\nabla_{\mathbf{R}_{k'}}\tilde{H}_e|\tilde{\Phi}_i\rangle + \langle\tilde{\Phi}_f|\nabla_{\mathbf{R}_{k'}}H_1^{\mathrm{BO}}|\tilde{\Phi}_i\rangle. \tag{B4}$$

Expanding $\langle\tilde{\Phi}_f|\nabla_{\mathbf{R}_{k'}}H_1^{\mathrm{BO}}|\tilde{\Phi}_i\rangle$ gives only terms that are second order in perturbations ΔH and $\tilde{H}_{\mathrm{ep}}$, so we can neglect that term and say that

$$\langle\Phi_f|\nabla_{\mathbf{R}_{k'}}H_{\mathrm{e}}|\Phi_i\rangle \approx \langle\tilde{\Phi}_f|\nabla_{\mathbf{R}_{k'}}\tilde{H}_{\mathrm{e}}|\tilde{\Phi}_i\rangle. \tag{B5}$$

Going back to the full first-order matrix element, we can transition directly from the electron-phonon coupling operator being in terms of the coordinates $\mathbf{R}_k$,

$$H_{\mathrm{ep}} = H_{\mathrm{e}}(\{\mathbf{R}_k\}) - H_{\mathrm{e}}\left(\{\mathbf{R}_k^{(i)}\}\right), \tag{B6}$$

to being in terms of the generalized coordinates q_j,

$$H_{\mathrm{ep}} = H_{\mathrm{e}}(\{q_j\}) - H_{\mathrm{e}}(\{q_j = 0\}). \tag{B7}$$

Plugging the linear term of this expansion into the first-order term would then give

$$\Gamma_{in}^{(1)} = \frac{2\pi}{\hbar}\sum_{fn'}\left|\sum_j M_j\langle X_{fn'}|q_j|X_{in}\rangle\right|^2\delta(\Theta_{n'}' - \Theta_n + W_{if}), \tag{B8}$$

where

$$M_j \equiv \langle\Phi_f|\partial_{q_j}H_{\mathrm{e}}|\Phi_i\rangle. \tag{B9}$$

Following the typical density-functional-theory formalism, the wave functions can be written as Slater determinants and the Hamiltonian can be written as a sum of fictitious, noninteracting, single-particle Hamiltonians, giving

$$\langle\Phi_f|\partial_{q_j}H_{\mathrm{e}}|\Phi_i\rangle = \langle\phi_f|\partial_{q_j}h_{\mathrm{e}}|\phi_i\rangle + \sum_{l=1}^{N-1}\langle\phi_l|\partial_{q_j}h_{\mathrm{e}}|\phi_l\rangle, \tag{B10}$$

where $|\phi_l\rangle$ are the single-particle orbitals from the same Hamiltonian h_{e}. To eliminate the last term in Eq. (B10), we choose h_{e} to be that of the initial configuration. This choice allows us to use the Helmann-Feynman theorem, whereby the second term represents the total force on all of the nontransitioning particles, which is zero. Consequently, we can equivalently write M_j in the single-particle form

$$M_j = \langle\phi_f|\partial_{q_j}h_{\mathrm{e}}|\phi_i\rangle, \tag{B11}$$

where $|\phi_i\rangle$ and $|\phi_f\rangle$ are single-electron wave functions from the relaxed initial configuration.

In order to evaluate M_j, we first look at a perturbative expansion of the final wave functions and manipulate that to get to an alternate form of M_j. We can approximate the final wave functions after some infinitesimal displacement δq_j of the atoms from the initial positions along the phonon eigenvector of mode j as

$$|\phi_f^j\rangle \approx |\phi_f\rangle + \sum_{l\neq f}\frac{\langle\phi_l|\left(\partial_{q_j}h_{\mathrm{e}}\right)\delta q_j|\phi_f\rangle}{\varepsilon_l - \varepsilon_f}|\phi_l\rangle. \tag{B12}$$

We can then evaluate

$$\begin{aligned}\left\langle\phi_f^j|\phi_i\right\rangle &= \langle\phi_f|\phi_i\rangle + \sum_{l\neq f}\frac{\langle\phi_f|\left(\partial_{q_j}h_{\mathrm{e}}\right)\delta q_j|\phi_l\rangle}{\varepsilon_l - \varepsilon_f}\langle\phi_l|\phi_i\rangle,\\ \left\langle\phi_f^j|\phi_i\right\rangle &= \langle\phi_f|\phi_i\rangle + \frac{\langle\phi_f|\left(\partial_{q_j}h_{\mathrm{e}}\right)\delta q_j|\phi_i\rangle}{\varepsilon_i - \varepsilon_f}.\end{aligned} \tag{B13}$$

Numerically, the overlap $\langle\phi_f|\phi_i\rangle$ is not identically zero, so we preserve that term for increased numerical accuracy. The overlaps in the sum are higher order, so the terms with $\langle\phi_l|\phi_i\rangle$ and $l \neq i$ can be safely neglected. Equation (B13) can be

rearranged to get

$$M_j = \langle\phi_f|\partial_{q_j} h_e|\phi_i\rangle = \frac{\varepsilon_i - \varepsilon_f}{\delta q_j}\left(\langle\phi_f^j|\phi_i\rangle - \langle\phi_f|\phi_i\rangle\right). \quad \text{(B14)}$$

The constrained DFT calculations needed for the overlaps from the initial configuration are difficult to converge over all of the phonon modes, and a separate set of calculations would be needed for each possible initial state. The final, ground-state configuration, however, is easier to converge and benefits from increased numerical accuracy as DFT is a ground-state variational theory. Below, we show that to first order the overlaps from either configuration are equivalent.

We relate the initial-configuration single-electron Hamiltonian h_e and the final-configuration single-electron Hamiltonian h'_e through

$$h'_e = h_e + \Delta h_e, \quad \text{(B15)}$$

where $|\phi_l\rangle$ are eigenfunctions of h_e and $|\phi'_l\rangle$ are those of h'_e. We relate the wave functions for the states i and f of the two configurations using perturbation theory:

$$|\phi'_i\rangle = |\phi_i\rangle + \sum_{l\neq i} \frac{\langle\phi_l|\Delta h_e|\phi_i\rangle}{\varepsilon_l - \varepsilon_i}|\phi_l\rangle. \quad \text{(B16)}$$

The Hamiltonian for $|\phi_f'^j\rangle$ is $h'_e + (\partial_j h'_e)\delta q_j = h_e + \Delta h_e + (\partial_j h_e)\delta q_j + (\partial_j \Delta h_e)\delta q_j$. We treat the term $(\partial_j \Delta h_e)\delta q_j$ as higher order and drop it. The wave function is then

$$\begin{aligned}|\phi_f'^j\rangle &= |\phi_f\rangle + \sum_{l\neq f} \frac{\langle\phi_l|\Delta h_e + (\partial_j h_e)\delta q_j|\phi_f\rangle}{\varepsilon_l - \varepsilon_f}|\phi_l\rangle \\ &= |\phi_f^j\rangle + \sum_{l\neq f} \frac{\langle\phi_l|\Delta h_e|\phi_f\rangle}{\varepsilon_l - \varepsilon_f}|\phi_l\rangle. \quad \text{(B17)}\end{aligned}$$

Using these to perform the inner product and keeping only terms to first order, we find

$$\langle\phi_f'^j|\phi'_i\rangle \approx \langle\phi_f^j|\phi_i\rangle. \quad \text{(B18)}$$

Thus, either the initial- or final-configuration wave functions can be used to calculate the matrix elements. Replacing the overlaps in Eq. (B14) with the final-configuration, ground-state overlaps gives

$$M_j = \frac{\varepsilon_i - \varepsilon_f}{\delta q_j}\left(\langle\phi_f'^j|\phi'_i\rangle - \langle\phi'_f|\phi'_i\rangle\right). \quad \text{(B19)}$$

Note that the eigenvalue difference should still come from the relaxed initial configuration and the initial relaxed positions should still be the starting point for the displacement, as the the perturbation expansion was carried out in the relaxed initial configuration. The physically but not numerically zero overlap $\langle\phi_f|\phi_i\rangle$ is replaced with the equivalent overlap from the final configuration for consistency.

APPENDIX C: INTEGRATION METHOD FOR THE LINE-SHAPE FUNCTIONS

The zeroth-order line-shape function for a single initial phonon configuration is defined as

$$F_{in}^{(0)} = \sum_{n'} |\langle X_{fn'}|X_{in}\rangle|^2 \delta(\Theta'_{n'} - \Theta_n + W_{if}). \quad \text{(C1)}$$

Expanding the wave functions and accounting for the thermal distribution of initial phonons, we have the total zeroth-order line-shape function,

$$\begin{aligned}F_i^{(0)} = \sum_{\{n'_j\}}\sum_{\{n_j\}} &\left[\prod_{j=1}^{N} \frac{e^{-\beta(n_j+\frac{1}{2})\hbar\omega_j}}{Z_j}\left|\langle\chi'_{n'_j}|\chi_{n_j}\rangle\right|^2\right] \\ &\times \delta\left(\Theta'_{\{n'_j\}} - \Theta_{\{n_j\}} + W_{if}\right), \quad \text{(C2)}\end{aligned}$$

where

$$Z_j = \sum_{n_j=0}^{\infty} e^{-\beta(n_j+\frac{1}{2})\hbar\omega_j} \quad \text{(C3)}$$

is the partition function of phonon mode j. Here, we change notation from n, n' to $\{n_j\}, \{n'_j\}$ to show explicitly that the sum over the initial and final phonon states is over all possible initial and final configurations. Following Ref. [9], the δ function can be rewritten as a time integration using a Fourier transform:

$$\delta(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\omega t}dt = \frac{1}{\pi}\,\mathrm{Re}\int_0^{+\infty} e^{i\omega t}dt. \quad \text{(C4)}$$

The zeroth-order line-shape function then becomes

$$\begin{aligned}F_i^{(0)} = \frac{1}{\pi\hbar}\,\mathrm{Re}\int_0^{+\infty}\sum_{\{n'_j\}}\sum_{\{n_j\}}&\left[\prod_{j=1}^{N}\frac{e^{-\beta(n_j+\frac{1}{2})\hbar\omega_j}}{Z_y j}\left|\langle\chi'_{n'_j}|\chi_{n_j}\rangle\right|^2\right] \\ &\times e^{i\frac{\Theta'_{\{n'_j\}}-\Theta_{\{n_j\}}}{\hbar}t} e^{i\frac{W_{if}}{\hbar}t}\,dt. \quad \text{(C5)}\end{aligned}$$

Using Eqs. (18) and (19), we can bring the phonon-energy exponential and the summations over the phonon configurations inside the product,

$$\begin{aligned}F_i^{(0)} = \frac{1}{\pi\hbar}\,\mathrm{Re}\int_0^{+\infty}\prod_{j=1}^{N}&\left[\frac{1}{Z_j}\sum_{n'_j}\sum_{n_j}\left|\langle\chi'_{n'_j}|\chi_{n_j}\rangle\right|^2\right. \\ &\left.\times e^{i(n'_j\omega'_j - n_j\omega_j)t} e^{-\beta(n_j+\frac{1}{2})\hbar\omega_j}\right] e^{i\frac{W_{if}}{\hbar}t}\,dt, \quad \text{(C6)}\end{aligned}$$

where the energy of the entire configuration becomes the energy of a single mode with a given number of phonons and we allow the phonon frequencies to be different between the initial (ω_j) and final (ω'_j) configurations. Note that because the initial sums were over all possible configurations, no new cross terms are introduced.

Define the product part of the integrand as

$$G^{(0)}(t) \equiv \prod_{j=1}^{N} G_j^{(0)}(t), \quad \text{(C7)}$$

where

$$G_j^{(0)}(t) = \frac{1}{Z_j}\sum_{n'_j}\sum_{n_j}\left|\langle\chi'_{n'_j}|\chi_{n_j}\rangle\right|^2 e^{i(n'_j\omega'_j - n_j\omega_j)t} e^{-\beta(n_j+\frac{1}{2})\hbar\omega_j}. \quad \text{(C8)}$$

Using the definition of the single-mode Hamiltonians, H_j and H'_j, given in Eqs. (12)–(16), Eq. (C8) can be expressed as a trace:

$$G_j^{(0)}(t) = \frac{1}{Z_j}\,\mathrm{Tr}\left(e^{\frac{i\hat{H}'_j t}{\hbar}} e^{-\frac{i\hat{H}_j t}{\hbar} - \beta\hat{H}_j}\right). \tag{C9}$$

Given the fact that the trace is invariant under a change of basis, Eq. (C9) can be rewritten in the normal-mode q-coordinate basis in the final-state configuration of the defect after carrier capture:

$$G_j^{(0)}(t) = \frac{1}{Z_j}\int_{-\infty}^{+\infty} dq \langle q|e^{\frac{i\hat{H}'_j t}{\hbar}} e^{-\frac{i\hat{H}_j t}{\hbar} - \beta\hat{H}_j}|q\rangle = \frac{1}{Z_j}\int_{-\infty}^{+\infty} dq \int_{-\infty}^{+\infty} dq' \langle q|e^{\frac{i\hat{H}'_j t}{\hbar}}|q'\rangle\langle q'|e^{-\frac{i\hat{H}_j t}{\hbar} - \beta\hat{H}_j}|q\rangle. \tag{C10}$$

Each term in the integral represents a quantum propagator $K(q, q'; t) = \langle q|\hat{U}(t)|q'\rangle$, where $\hat{U}(t)$ in the Schrödinger picture is $e^{-i\hat{H}t/\hbar}$. The propagator of a one-dimensional quantum harmonic oscillator is the Mehler kernel (without m because we use q instead of x):

$$\langle q|e^{-i\hat{H}t/\hbar}|q'\rangle = \left[\frac{\omega}{2\pi i\hbar\sin(\omega t)}\right]^{\frac{1}{2}} \exp\left[\frac{i\omega}{2\hbar}\frac{(q^2+q'^2)\cos(\omega t) - 2qq'}{\sin(\omega t)}\right]. \tag{C11}$$

A similar expression was used by Kubo and Toyozawa [11]. Equation (C10) then leads to

$$G_j^{(0)}(t) = \frac{1}{Z_j}\int_{-\infty}^{+\infty} dq \int_{-\infty}^{+\infty} dq' \left[\frac{\omega_j}{2\pi i\hbar\sin(\omega_j T_1)}\right]^{\frac{1}{2}} \left[\frac{\omega'_j}{2\pi i\hbar\sin(\omega'_j T_2)}\right]^{\frac{1}{2}} \exp\left[\frac{i\omega_j}{2\hbar}\frac{(q^2+q'^2)\cos(\omega_j T_1) - 2qq'}{\sin(\omega_j T_1)}\right]$$
$$\times \exp\left\{\frac{i\omega'_j}{2\hbar}\frac{\left[(q+\Delta q_j)^2 + (q'+\Delta q_j)^2\right]\cos(\omega'_j T_2) - 2(q+\Delta q_j)(q'+\Delta q_j)}{\sin(\omega'_j T_2)}\right\}, \tag{C12}$$

where

$$T_1 = t - i\hbar\beta, \quad T_2 = -t. \tag{C13}$$

Using the Gaussian integral

$$\int_{-\infty}^{+\infty} \exp\left(\frac{1}{2}iax^2 + iJx\right) dx = \left(\frac{2\pi i}{a}\right)^{\frac{1}{2}} \exp\left(\frac{-iJ^2}{2a}\right), \tag{C14}$$

the integration is evaluated to yield

$$G_j^{(0)}(t) = \exp\left[iD_j^{(0)}(t)\right], \tag{C15}$$

where

$$D_j^{(0)}(t) = -2\left[iS_j^{-1}A_j(t) - S_j'^{-1}\cot(\omega'_j t/2)\right]^{-1}, \tag{C16}$$

$$A_j(t) = \frac{e^{i\omega_j t}(\bar{n}_j+1) + \bar{n}_j}{e^{i\omega_j t}(\bar{n}_j+1) - \bar{n}_j}, \tag{C17}$$

the average phonon distribution functions are

$$\bar{n}_j = \frac{1}{e^{\beta\hbar\omega_j} - 1}, \tag{C18}$$

and the Huang-Rhys factors S_j and S'_j are defined as

$$S_j = \frac{\omega_j}{2\hbar}(\Delta q_j)^2 \tag{C19}$$

and

$$S'_j = \frac{\omega'_j}{2\hbar}(\Delta q_j)^2, \tag{C20}$$

with Δq_j being a projection of the change between the initial and final relaxed atomic positions on the phonon eigenvectors. The prefactor $1/Z_j$ in $G_j^{(0)}(t)$ is canceled due to the requirement that when $\Delta q_j = 0$, $G_j^{(0)}(t) = 1$. When $\omega_j = \omega'_j$, $G_j^{(0)}(t)$ is simplified to

$$G_j^{(0)}(t) = \exp\left\{S_j[(\bar{n}_j+1)e^{i\omega_j t} + \bar{n}_j e^{-i\omega_j t} - (2\bar{n}_j+1)]\right\}. \tag{C21}$$

Plugging Eq. (C21) back into Eq. (C7) gives Eq. (38).

Likewise, the first-order term is derived from

$$G_{j'}^{(1)}(t) = |M_{j'}|^2 \sum_{n'_{j'}} \sum_{n_{j'}} \left|\left\langle \chi'_{n'_{j'}} \middle| q_{j'} \middle| \chi_{n_{j'}} \right\rangle\right|^2 e^{i(n'_{j'}\omega'_{j'} - n_{j'}\omega_{j'})t} e^{-\beta(n_{j'}+\frac{1}{2})\hbar\omega_{j'}}. \tag{C22}$$

Following a similar procedure as for the zeroth-order term, this is evaluated as

$$G_j^{(1)}(t) = |M_j|^2 G_j^{(0)}(t) D_j^{(1)}(t), \tag{C23}$$

where

$$D_j^{(1)}(t) = -\left[\frac{\Delta q_j}{2S_j}\right]^2 D_j^{(0)}(t)\left\{D_j^{(0)}(t)[A_j(t)]^2 - \frac{1}{2}\frac{\omega_j}{\omega'_j}\left[A_j(t)\cot\frac{\omega'_j t}{2} - \frac{\omega_j \cot(\omega'_j t/2) - i\omega'_j A_j(t)}{\omega_j A_j(t) + i\omega'_j \cot(\omega'_j t/2)}\right]\right\}. \tag{C24}$$

For $\omega'_j = \omega_j$, $D_j^{(1)}(t)$ reduces to

$$D_j^{(1)}(t) = \frac{\hbar}{2\omega_j}\{\bar{n}_j e^{-i\omega_j t} + (\bar{n}_j + 1)e^{i\omega_j t} + S_j[1 + \bar{n}_j e^{-i\omega_j t} - (\bar{n}_j + 1)e^{i\omega_j t}]^2\}. \tag{C25}$$

[1] D. V. Lang, Deep-level transient spectroscopy: A new method to characterize traps in semiconductors, J. Appl. Phys. **45**, 3023 (1974).

[2] S. T. Pantelides, Y. Puzyrev, X. Shen, T. Roy, S. DasGupta, B. R. Tuttle, D. M. Fleetwood, and R. D. Schrimpf, Reliability of III–V devices—The defects that cause the trouble, Microelectron. Eng. **90**, 3 (2012).

[3] K. Huang and A. Rhys, Theory of light absorption and non-radiative transitions in *F*-centres, Proc. R. Soc. Lond. Ser. A **204**, 406 (1950).

[4] R. Kubo, Thermal ionization of trapped electrons, Phys. Rev. **86**, 929 (1952).

[5] B. K. Ridley, Multiphonon, non-radiative transition rate for electrons in semiconductors and insulators, J. Phys. C: Solid State Phys. **11**, 2323 (1978).

[6] K. Huang, Adiabatic approximation theory and static coupling theory of nonradiative transition, Sci. Sin. **24**, 27 (1981).

[7] A. Alkauskas, Q. Yan, and C. G. Van de Walle, First-principles theory of nonradiative carrier capture via multiphonon emission, Phys. Rev. B **90**, 075202 (2014).

[8] G. D. Barmparis, Y. S. Puzyrev, X.-G. Zhang, and S. T. Pantelides, Theory of inelastic multiphonon scattering and carrier capture by defects in semiconductors: Application to capture cross sections, Phys. Rev. B **92**, 214111 (2015).

[9] L. Shi, K. Xu, and L.-W. Wang, Comparative study of *ab initio* nonradiative recombination rate calculations under different formalisms, Phys. Rev. B **91**, 205315 (2015).

[10] V. May and O. Kühn, *Charge and Energy Transfer Dynamics in Molecular Systems* (Wiley, Hoboken, NJ, 2008).

[11] R. Kubo and Y. Toyozawa, Application of the method of generating function to radiative and non-radiative transitions of a trapped electron in a crystal, Prog. Theor. Phys. **13**, 160 (1955).

[12] A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. V. de Walle, First-principles theory of the luminescence lineshape for the triplet transition in diamond NV centres, New J. Phys. **16**, 073026 (2014).

[13] R. Borrelli, A. Capobianco, and A. Peluso, Generating function approach to the calculation of spectral band shapes of free-base chlorin including Duschinsky and Herzberg–Teller effects, J. Phys. Chem. A **116**, 9934 (2012).

[14] A. Baiardi, J. Bloino, and V. Barone, General time dependent approach to vibronic spectroscopy including Franck–Condon, Herzberg–Teller, and Duschinsky effects, J. Chem. Theory Comput. **9**, 4097 (2013).

[15] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[16] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59**, 1758 (1999).

[17] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, Phys. Rev. B **47**, 558 (1993).

[18] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[19] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. **6**, 15 (1996).

[20] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**, 3865 (1996).

[21] J. Heyd, G. E. Scuseria, and M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential, J. Chem. Phys. **118**, 8207 (2003).

[22] A. Togo and I. Tanaka, First principles phonon calculations in materials science, Scr. Mater. **108**, 1 (2015).

[23] P. T. Landsberg and D. J. Robbins, Lifetime broadening of a parabolic band edge of a pure semiconductor at various temperatures, Solid-State Electron. **28**, 137 (1985).

[24] R. Zhang and T. F. Kuech, Photoluminescence of carbon *in situ* doped GaN grown by halide vapor phase epitaxy, Appl. Phys. Lett. **72**, 1611 (1998).

[25] M. A. Reshchikov, D. O. Demchenko, A. Usikov, H. Helava, and Y. Makarov, Carbon defects as sources of the green and yellow luminescence bands in undoped GaN, Phys. Rev. B **90**, 235203 (2014).

[26] M. A. Reshchikov and R. Y. Korotkov, Analysis of the temperature and excitation intensity dependencies of photoluminescence in undoped GaN films, Phys. Rev. B **64**, 115205 (2001).

[27] M. A. Reshchikov and H. Morkoç, Luminescence properties of defects in GaN, J. Appl. Phys. **97**, 061301 (2005).

[28] M. A. Reshchikov, A. Usikov, H. Helava, Y. Makarov, V. Prozheeva, I. Makkonen, F. Tuomisto, J. H. Leach, and K. Udwary, Evaluation of the concentration of point defects in GaN, Sci. Rep. **7**, 9297 (2017).

[29] M. A. Reshchikov, Measurement and analysis of photoluminescence in GaN, J. Appl. Phys. **129**, 121101 (2021).

[30] F. Zhao, M. E Turiansky, A. Alkauskas, and C. G. Van de Walle, Trap-assisted Auger-Meitner recombination from first principles, Phys. Rev. Lett. **131**, 056402 (2023).