Journal of Applied Physics

Tutorial: Defects in semiconductors—Combining experiment and theory

Cite as: J. Appl. Phys. **119**, 181101 (2016); https://doi.org/10.1063/1.4948245
Submitted: 17 November 2015 . Accepted: 17 December 2015 . Published Online: 12 May 2016

Audrius Alkauskas, Matthew D. McCluskey, and Chris G. Van de Walle

Tutorial: Defects in semiconductors—Combining experiment and theory

Audrius Alkauskas,[1] Matthew D. McCluskey,[2] and Chris G. Van de Walle[3,a)]
[1]Center for Physical Sciences and Technology, Vilnius, LT-01108, Lithuania
[2]Department of Physics and Astronomy, Washington State University, Pullman, Washington 99164-2814, USA
[3]Materials Department, University of California, Santa Barbara, California 93106-5050, USA

(Received 17 November 2015; accepted 17 December 2015; published online 12 May 2016)

Point defects affect or even completely determine physical and chemical properties of semiconductors. Characterization of point defects based on experimental techniques alone is often inconclusive. In such cases, the combination of experiment and theory is crucial to gain understanding of the system studied. In this tutorial, we explain how and when such comparison provides new understanding of the defect physics. More specifically, we focus on processes that can be analyzed or understood in terms of configuration coordinate diagrams of defects in their different charge states. These processes include light absorption, luminescence, and nonradiative capture of charge carriers. Recent theoretical developments to describe these processes are reviewed. *Published by AIP Publishing.*
[http://dx.doi.org/10.1063/1.4948245]

I. INTRODUCTION

Every material contains defects; perfect materials simply do not exist. While it may cost energy to create a defect, configurational entropy renders it favorable to incorporate a certain concentration of defects, since this lowers the free energy of the system.[1] Therefore, even in equilibrium, we can expect defects to be present; kinetic limitations sometimes lead to formation of additional defects. Note that all of these considerations also apply to *impurities* that are unintentionally present in the growth or processing environment. Of course, impurities are often intentionally introduced to tailor the properties of materials. Doping of semiconductors with acceptors and donors is essential for electronic and optoelectronic applications. In the following, we will use the word "defect" as a generic term to cover both intrinsic defects (vacancies, self-interstitials, and antisites) and impurities.

Since defects are unavoidable, we must consider the effects they have on the properties of materials. These effects can be considerable, to the point of determining the functionality of the material, as in *p*- or *n*-type doping. Point defects play a key role in diffusion: virtually all diffusion processes are assisted by point defects. Defects are often responsible for degradation of a device. Even in the absence of degradation, defects can limit the performance of a device. Compensation by native point defects can decrease the level of doping that can be achieved. Defects with energy levels within the band gap can act as recombination centers, impeding carrier collection in a solar cell or light emission from a light-emitting diode. Sometimes, these effects can be used to advantage: luminescence centers in wide-band-gap materials can be used to emit light at specified wavelengths; or single-spin centers (such as the nitrogen–vacancy (NV) center in diamond) can act as an artificial atom and serve as a qubit in a quantum information system.[2,3] Finally, sometimes, one deliberately wants to grow materials with *many* defects. Examples are materials for ultrafast optoelectronic switches or semiconductors used to optically generate THz pulses, where defect densities should be large enough so that carrier lifetimes are as short as a few picoseconds.[4]

All of these scenarios require *control* over the defect population. Such control, in turn, requires accurate knowledge of the mechanisms by which defects incorporate and affect materials properties. Over the past 60 years, tremendous progress has been made in the experimental and theoretical characterization of defects in semiconductors—much of it documented in the proceedings of the biennial International Conference on Defects in Semiconductors (ICDS). The first such conference was held in 1959, Gatlinburg, TN, USA, chaired by Cleland.[5] Starting in the 1970s, computational studies, in combination with experimental efforts, have yielded new perspectives and quantitative detail about the impact of defects.

The present tutorial is not intended as a comprehensive review of the theoretical foundations or computational approaches; these have been documented in a number of review papers,[6–8] a series of articles,[9] and books.[10,11] Books that cover experimental aspects include Lannoo and Bourgoin,[12] Pantelides,[13] Stavola,[14] Spaeth and Overhof,[15] and McCluskey and Haller.[16] In the present tutorial, we aim to focus on specific areas where theory/computation and experiment have converged to provide unprecedented insight into the physical mechanisms that govern defect behavior. The ability to directly compare theoretical results with experimental measurements is essential for validation of the computational approach. In addition, it empowers theory to help interpret and explain experimental observations and ultimately to reliably predict structures and properties that can subsequently be experimentally implemented and observed.

The paper is structured as follows. Section II gives a general background on defect levels, first-principles calculations,

Note: This Invited Tutorial is part of the Special Topic section "Defects in Semiconductors" guest-edited by Filip Tuomisto and Ilja Makkonen (Department of Applied Physics, Aalto University, Finland), and published in Journal of Applied Physics **119**, 18 (2016). See http://scitation.aip.org/content/aip/journal/jap/119/18

a)vandewalle@mrl.ucsb.edu

and experimental methods. Section III discusses configuration coordinate diagrams, which are important for describing deep-level defects. Section IV presents specific case studies. Conclusions and general observations are presented in Sec. V. Finally, an Appendix discusses the validity of the one-dimensional configuration coordinate diagram.

II. BACKGROUND

A. Defect levels and wavefunctions

This tutorial focuses on electrically active defects. These are defects that can exist in different charge states. For example, single acceptors can exist in either a negatively charged or neutral state. Single donors exist in either the positively charged or neutral state. Other defects can exhibit more complexity. Substitutional gold in silicon, for example, can be negative, neutral, or positive. Defects that can act both as donors and acceptors are called *amphoteric*. Electrically active defects have at least one defect level (or transition level) in the band gap, somewhere between the valence-band maximum (VBM) and conduction-band minimum (CBM). This level has an associated defect wavefunction, a state to which the electron is added or removed when the charge state of the defect changes.

If a transition level is positioned such that the defect is likely to be thermally ionized at room temperature, it is conventionally referred to as a shallow level. Otherwise, it is called a deep level. An alternative definition is based on the degree of localization of the wave function of an electron (or hole) associated with the defect. A delocalized wave function (on the order of many lattice constants) is the characteristic of a shallow level, while a wave function localized on a length scale of an atomic bond indicates a deep level. A more detailed discussion of shallow levels can be found in Sec. II D 2 of Ref. 6.

B. First-principles calculations

As in so many areas of solid-state physics, density functional theory (DFT)[17,18] has emerged as the most powerful approach for assessing the properties of defects. DFT calculations, typically carried out in a supercell geometry,[6] yield reliable information about atomic structure, including all relaxations of the host atoms. Electronic structure proved to be a greater challenge. Traditional functionals, such as the still widely used local density approximation (LDA) and generalized gradient approximation (GGA), severely underestimate band gaps of semiconductors and insulators, and therefore the position of defect levels suffers from large uncertainties. This can be addressed by going beyond DFT, for instance using many-body perturbation theory (typically in the *GW* approach),[19] but this tends to be computationally expensive and difficult to execute in a self-consistent fashion. It is usually applied to address only the electronic structure, based on atomic structures obtained from DFT.[20,21]

In recent years, hybrid functionals[22] have offered a powerful way to overcome these limitations: not only do they produce band structures in much better agreement with experiment, but they also provide a much more reliable description of charge localization, which is essential for accurate modeling of low-symmetry defects or structures that give rise to polaron formation.[23,24] In particular, the screened hybrid functional of Heyd, Scuseria, and Ernzerhof (HSE)[25] has proven reliable to predict formation energies and transition levels of native defects and impurities.[26–30]

Another important advance has been the ability to correct for errors that arise from the use of supercells to describe charged defects. While these supercells can typically be made large enough to minimize interactions between a neutral defect and its mirror images, the long range of the Coulomb interaction renders this essentially impossible in the case of charged defects. Explicit correction schemes are therefore essential, and a supercell-size correction scheme based on the rigorous treatment of electrostatics was described in Refs. 31 and 32.

C. Experimental techniques

In this section, we give a brief overview of experimental techniques that are most common and most useful for studying defects.

1. Secondary ion mass spectrometry (SIMS)

SIMS allows accurate measurement of impurity concentrations as a function of depth. The detection limit can be as low as $10^{13}\,\mathrm{cm}^{-3}$,[33] but it depends on the type of impurity. Impurities that are commonly found in the environment, such as nitrogen or carbon, are most difficult to detect in low concentrations.[34] SIMS is not suitable for determining intrinsic-defect concentrations: finding the needle in the haystack is hard enough—finding any missing or additional hay is essentially impossible. A notable exception is when isotopes are used, a very powerful approach for studying self-diffusion.[35]

2. Positron annihilation spectroscopy (PAS)

PAS identifies point defects by measuring the annihilation of positrons with electrons.[36] It is most powerful for detection of neutral or negative vacancies, which can trap positrons efficiently. The increased lifetime of the trapped positron provides a signature for the defect. In addition, the momentum distribution of the emitted gamma rays provides information about the local environment.

3. X-ray measurements

X-ray diffraction can measure lattice parameters with high accuracy and can sometimes be used to assess impurity-induced changes of the atomic structure.[37,38] Extended X-ray absorption fine structure (EXAFS)[39] produces details about atomic structure and bond lengths around an impurity, but only in the case of impurities with relatively heavy mass.

4. Electron paramagnetic resonance (EPR)

EPR is an excellent tool that can provide detailed information about concentrations, chemical identity, symmetry, and lattice environment of defects or impurities.[40] Measured

hyperfine parameters can be directly compared with values based on wave functions obtained from first-principles calculations, thus allowing an explicit identification of the microscopic structure.[41] Since EPR relies on the presence of unpaired electrons, sometimes optical excitation is necessary to generate a metastable charge state with a nonzero spin density. Optically detected magnetic resonance (ODMR) is a variant of the technique that offers additional information about the defect-induced levels.[42]

5. Electrical measurements

Hall-effect measurements provide details about carrier concentrations, concentrations of compensating defects, and mobility. The identity of the electrically active and scattering centers must be inferred from other measurements. Deep level transient spectroscopy (DLTS)[43] is a powerful technique to observe defect levels in the band gap, under conditions where the final charge state can fully relax to its equilibrium configuration after the transition. These measurements should therefore be compared with thermodynamic transition levels, as discussed in Sec. III.

6. Vibrational spectroscopy

The presence of a defect modifies the vibrational spectrum of the solid, often leading to local vibrational modes (LVMs) that are detected with Raman spectroscopy or Fourier-transform infrared spectroscopy.[44] These modes can be calculated with first-principles calculations,[45,46] greatly aiding in identifying the nature and local structure of the defect.

7. Optical measurements

In a photoluminescence (PL) experiment, above-gap light excites electrons from the valence band to the conduction band (CB).[47] In addition, electrons can be excited from or to defect levels. When the excited electrons transition to lower energy levels, they can emit light to conserve energy, resulting in a peak in the PL spectrum. In a photoluminescence excitation (PLE) experiment, the PL intensity is measured as a function of excitation photon energy. This gives an absorption profile for the defect. While these techniques yield optical signatures of defects with high resolution,[48] they cannot by themselves identify their nature or character. Comparison with first-principles calculations can be particularly powerful here.

III. CONFIGURATION COORDINATE DIAGRAMS

Defect physics is almost as complex as solid-state physics itself. In studying defects, one is faced with the same variety of different phenomena as for solids in general: electronic and nuclear magnetism, thermodynamics, optical properties, mechanical properties, transport, etc. In this tutorial, we focus on one specific but very important aspect of defect physics: the coupling of electronic and atomic structure. This is best described in terms of *configuration coordinate diagrams*, defined below. On the one hand, analysis of experimental results in terms of these diagrams provides a lot of insight about defects. On the other hand, first-principles calculations comprise a tool that can directly provide such information. Thus, in this regard, the comparison between experiment and theory is a very powerful combination to study defects.

A. Definition

As discussed in Sec. II A, the wavefunction of a defect can be either delocalized or localized. If it is delocalized, then the electron density is very low at the defect and does not exert much influence on the atomic positions. This is the case for shallow defects. If the wavefunction is localized (on the defect itself and/or a few neighboring atoms), then this has a strong effect on how the atoms rearrange themselves locally when the charge of the defect changes or, more generally, when the electronic state of the defect changes. We will focus on the properties of deep, localized defects.

The change of the defect geometry associated with the change of its electronic state is best described by a configuration coordinate diagram.[16,49] For simplicity, consider a diatomic model, where an impurity atom attaches to a host atom. We will look at the case where the defect is an acceptor. When the Fermi level is at the CBM, the defect's electronic ground state is negative (−) and its excited state is neutral (0).

The defect is described by a configuration coordinate (cc) Q, which in this simple picture is proportional to the bond length. At equilibrium, the energy E is minimized at a specific value of Q. To first order, the bond behaves like a harmonic spring, so $E \sim Q^2$. Consider the defect in its ground state (i.e., the negatively charged state; see Fig. 1). During absorption, an incoming photon can excite the electron into the CB, thus neutralizing the acceptor. In the classical Franck-Condon approximation,[49–52] this optical transition occurs instantaneously: the ions are slow, and so their position does not change when the electronic state changes. This gives us the classical absorption energy E_{abs}, shown in Fig. 1. Because the defect has lost an electron, the bond strength is different and the bond length has a new equilibrium Q value. It relaxes to this new equilibrium, losing energy in the process. The amount of energy lost is the *Franck-Condon shift*, d^e_{FC} (Fig. 1). If the energy separation between the equilibrium configurations in the ground and the excited state is E_{therm}, then

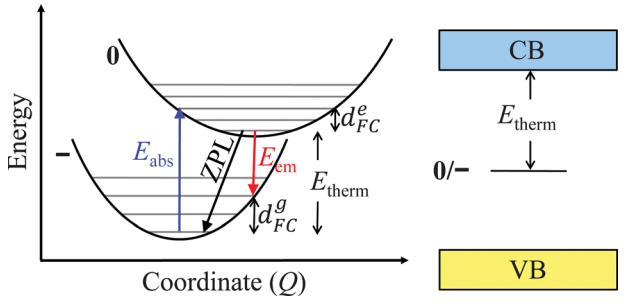

FIG. 1. Configuration coordinate diagram, showing important energies and optical transitions. For this example, E_{therm} gives the acceptor level relative to the CBM.

$$E_{\text{abs}} = E_{\text{therm}} + d^e_{FC}, \tag{1}$$

where E_{therm} is the thermal energy of the transition: it is the energy difference between the ground state and the excited state in their equilibrium configurations. An alternative term often used is the zero phonon line (ZPL) energy, E_{ZPL}.

Once in the excited state, the acceptor can recapture the electron, often emitting a photon. This is the luminescence, or light emission, process. Within the same classical Franck-Condon approximation, such a transition occurs for a fixed nuclear configuration, and its energy is E_{em} (Fig. 1). Finally, the system relaxes to its ground-state equilibrium, losing more energy. The energy lost in the ground state is d^g_{FC}.

Because of relaxations in the excited state and the ground state, the emission energy is lower than the absorption energy. The total energy difference between absorption and emission is called the *Stokes shift*. We see that the Stokes shift is given by the sum of Franck-Condon shifts in the ground and excited state

$$E_{\text{abs}} - E_{\text{em}} = d^e_{FC} + d^g_{FC}. \tag{2}$$

The classical Franck-Condon approximation is often sufficient to interpret optical experiments. For example, it can be used to understand the distinction between electrical and optical properties of the same defect.[6,16] However, one can go beyond this approximation by including the vibrational broadening of the transitions,[50–52] both in absorption and emission. This concept is best illustrated using the same cc diagram (Fig. 1), taking into account the vibrational sub-levels.

It is possible to define an effective vibrational frequency that is representative of exactly the same atomic motion that is "encoded" in our configuration coordinate Q. In the simple case of the diatomic molecule discussed above, this will simply be linear vibrations of the molecule. Let these frequencies be ω_g and ω_e for the ground and the excited state. If the curvatures of the two parabolas are identical, we will then use the symbol ω_0 to represent this frequency. Going beyond the classical Franck-Condon approximation, we now have to consider transitions between vibronic levels of defects, as explained in Sec. III B. A very important parameter in discussing the vibrational broadening is the so-called Huang-Rhys factor S.[53] In the case of equal vibrational frequencies in the ground and the excited state, it is given by

$$S = \frac{d_{FC}}{\hbar\omega_0}. \tag{3}$$

The Huang-Rhys factor S essentially quantifies the number of phonons emitted during optical transition, either absorption or emission.

Often the real motion of the atoms around the defect associated with the change of the electronic state is more complex than the model discussed above. For example, when the charge of the Mg acceptor in GaN changes from 0 to −1 or vice versa, this mostly involves a change of the geometry of five atoms:[54,55] the Mg atom itself, one out of four nitrogen atoms that are nearest neighbors to Mg, as well as three Ga atoms that are bonded to this N atom. The configuration coordinate then describes the *collective* motion of atoms that captures the essential physics of the process. It might be surprising that this one-dimensional approximation to what is essentially a multi-dimensional problem (where the dimensionality is $3N$, N being the number of atoms in the system) is sufficient. The beauty of 1D cc diagrams is that often they *are* sufficient.[51] As discussed in the Appendix, this is particularly the case for defects with strong electron–phonon coupling. In certain cases, the validity of this approximation can be demonstrated rigorously; we discuss one such example in the Appendix.

B. Emission and absorption

In this section, we discuss absorption and emission processes more quantitatively, focusing in particular on the energy dependence of absorption cross-sections and luminescence lineshapes. As above, let us consider optical processes that involve a single acceptor. Optical absorption is the process that can be described via

$$A^- + h\nu \rightarrow A^0 + e^-, \tag{4}$$

while luminescence is given by

$$A^0 + e^- \rightarrow A^- + h\nu. \tag{5}$$

After optical absorption, the electron ends up in a continuum of CB states, even at low temperatures. In the case of emission at low temperatures, both the initial and the final electronic state of the electron can be described by single electronic states. Because of this, we start our discussion with luminescence, because it is conceptually a simpler process.

Emission: To describe the vibrational broadening of emission, we have to sum up all possible transitions between the vibrational levels in the excited state and those in the ground state. Within the 1D model, the normalized luminescence lineshape is given by[49]

$$L(\hbar\omega, T) = \sum_{n,m} w_m(T) |\langle \chi_{em} | \chi_{gn} \rangle|^2 \delta(E_{ZPL} + m\hbar\omega_e - n\hbar\omega_g - \hbar\omega). \tag{6}$$

The sum runs over all vibrational levels in the excited state (m) and the ground state (n). $w_m(T)$ is the thermal occupation factor. In principle, at low temperatures ($m = 0$), Eq. (6) describes a series of equidistant peaks: the emission of a photon without any phonons ($n = 0$) is the ZPL, while the $n = 1$, 2, 3… peaks are referred to as *phonon replicas*.

In practice, especially for defects with large Huang-Rhys factors ($S \gg 1$), these phonon replicas are not observed, for reasons described in the Appendix. In calculations, the Dirac delta functions $\delta(\hbar\omega)$ (which reflect energy conservation) can be replaced by functions of finite width (e.g., Gaussian functions). Equation (6) corresponds to the quantum variant of the Franck-Condon approximation, which assumes that the transition dipole moment that describes the coupling between the excited state and the ground state does not depend on Q. $\langle \chi_{em} | \chi_{gn} \rangle$ are Franck-

Condon overlap integrals between vibrational levels in the ground and the excited state. When $\omega_e = \omega_g$ and $m = 0$, these integrals have a simple analytical expression[48]

$$|\langle\chi_{e0}|\chi_{gn}\rangle|^2 = e^{-S}\frac{S^n}{n!}, \tag{7}$$

where S is the Huang-Rhys factor. At temperatures much higher than $\hbar\omega_e/k_B$, the lineshape $L(\hbar\omega, T)$ becomes Gaussian-shaped with a full width at half maximum W

$$W(T) = W_0 \coth\left(\frac{\hbar\omega_e}{k_B T}\right), \tag{8}$$

where k_B is the Boltzmann constant. A general expression for W_0, valid when $\omega_e \neq \omega_g$, is given in Refs. 56 and 57. When $\omega_e = \omega_g = \omega_0$,

$$W_0 = \hbar\omega_0\sqrt{8\ln 2\, S} \approx 2.355\hbar\omega_0 S^{1/2}. \tag{9}$$

Thus, measuring the temperature-dependent broadening of defect luminescence bands, one can directly measure both the vibrational frequencies [Eq. (8)] and the Huang-Rhys factors [Eq. (9)]. While the general theory that we discussed above was known for many years (e.g., Refs. 50 and 51), an important point is that it is now possible to calculate all the relevant parameters that enter into the description completely from first principles.[20,55,58–60] Examples of this are given in Sec. IV.

Absorption: The theory behind the lineshape of optical absorption follows a similar logic as for emission. However, one has to include the sum over final electronic states that represent the electron in the conduction band. Various models have been developed to simulate the shape of the absorption onset.[61] Following the treatment by Kopylov and Pikhtin,[62] the absorption cross section σ is given by

$$\sigma(\hbar\omega) \propto \int_0^\infty \frac{e^{-(E+E_{\rm abs}-\hbar\omega)^2/\Gamma^2}\sqrt{E}\,dE}{\hbar\omega(E+E_{\rm abs})^2}, \tag{10}$$

where $\hbar\omega$ is the photon energy and Γ is a broadening parameter. In the limit of no vibrational broadening ($\Gamma = 0$), Eq. (10) reduces to

$$\sigma(\hbar\omega, \Gamma = 0) \propto \frac{\sqrt{\hbar\omega - E_{\rm abs}}}{(\hbar\omega)^3}. \tag{11}$$

This absorption profile has a sharp onset at $\hbar\omega = E_{\rm abs}$. As Γ increases, the onset spreads out (Fig. 2).

The broadening parameter depends on the Franck-Condon shift ($d_{\rm FC}$), ground state vibrational frequency ($\omega_{\rm g}$), and excited state vibrational frequency ($\omega_{\rm e}$). At absolute zero temperature,

$$\Gamma_0 = \frac{\omega_g}{\omega_{\rm e}}\sqrt{2d_{FC}\hbar\omega_g}. \tag{12}$$

When $\omega_e = \omega_g = \omega_0$, this simplifies to

$$\Gamma_0 = \sqrt{2d_{FC}\hbar\omega_0} \approx 1.414\hbar\omega_0 S^{1/2}. \tag{13}$$

Note that this equation is essentially the same as Eq. (9), just written for the Gaussian broadening parameter that appears in Eq. (10) rather than the width of the Gaussian function. As the temperature T is raised, the higher vibrational states of the electronic ground state become populated, leading to more broadening. Similar to Eq. (8), the temperature-dependent broadening is given by

$$\Gamma = \Gamma_0\sqrt{\coth(\hbar\omega_0/k_B T)}. \tag{14}$$

This temperature dependence is valuable because it allows one to determine ω_0 experimentally.

C. Nonradiative capture of charge carriers

In addition to radiative transitions discussed in Sec. III B, cc diagrams are also important to understand nonradiative capture of carriers. Here, we will focus only on transitions that occur via the mechanism of multiphonon emission (MPE).

The theory of MPE has a rich history. In the context of defects, it was first suggested by Huang and Rhys in the 1950s[53] and was subsequently developed by Gummel and Lax,[63] Kubo and Toyozawa,[64] Henry and Lang,[65] and others. The nonradiative transition rate is usually calculated within first-order perturbation theory. First-principles calculations of nonradiative capture coefficients have appeared only in recent years.[66–69] References 67 and 69 in particular, provide a historical overview of the developments in the theory.

A 1D cc diagram pertaining to a nonradiative capture process is shown in Fig. 3. The example illustrates a capture of a hole by a negatively charged acceptor. In contrast to absorption or luminescence discussed previously (Fig. 1), now the excited state corresponds to the negative charge state, while the electronic ground state corresponds to the neutral charge state. This happens because for this particular process, we study the exchange of charge with the valence band rather than the conduction band. The energy difference between the two states is ΔE, and the two potential energy curves are offset horizontally by ΔQ. This results in a

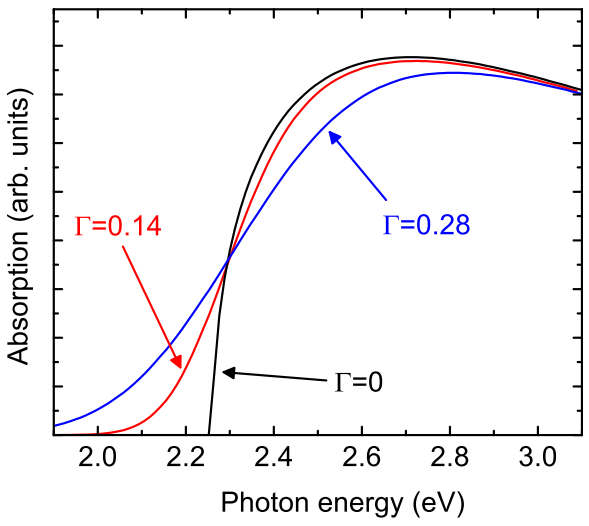

FIG. 2. Absorption profile for a deep-level acceptor with $E_{\rm abs} = 2.26$ eV, according to Eq. (10). Broadening parameters (Γ) are in units of eV.

Franck-Condon shift d_{FC}^{g}, equivalent to the Franck-Condon shift for optical transitions. ΔE_b is the classical barrier for a transition from the upper curve to the lower potential energy curve, determined as the difference between the intersection point of the two curves and the minimum energy of the excited state. If the curvatures of the two parabolas are equal ($d_{FC}^{g} = d_{FC}^{e} = d_{FC}$), the classical barrier is given by

$$\Delta E_b = \frac{(\Delta E - d_{FC})^2}{4 d_{FC}}. \tag{15}$$

The process of MPE is possible only when the potential energy curves of the ground and the excited are offset, i.e., when there is a nonzero Franck-Condon shift d_{FC}. If d_{FC} were 0, the phonon selection rules $\Delta m = \pm 1$ (m is phonon quantum number) would allow only transitions where one phonon is emitted or absorbed. Since maximum phonon energies in wide-band-gap semiconductors are of the order of 0.1 eV, such mechanisms would not be important for deep centers. It is exactly this offset of the potential energy surfaces, and thus the breaking of the strict phonon selection rule, that enables nonradiative transitions, whereby energies much larger than phonon energies can be dissipated (hence the term "multiphonon emission").

Nonradiative carrier capture via MPE occurs in two steps. The first step represents the capture of a carrier from a delocalized bulk-like state to a localized defect state, whereby the electronic energy is conserved, but the system ends up in a vibrationally excited state. In the cc picture, this could be represented by a horizontal transition between the two potential energy curves (Fig. 3). In the second step, the vibrationally excited state returns to local equilibrium via phonon relaxation. The second process is much faster than the first one. Therefore, the rate-limiting step for MPE is the electronic transition.

Taking holes as our example, the nonradiative capture rate R_p per unit volume per unit time is given by

$$R_p = C_p N^{-} p. \tag{16}$$

Here, C_p is the capture coefficient, N^{-} is the concentration of defects in the negative charge state, and p is the hole density. The quantum-mechanical expression for C_p can be found, for example, in Ref. 65 [Eq. (13)] or Ref. 67 [Eq. (22)].

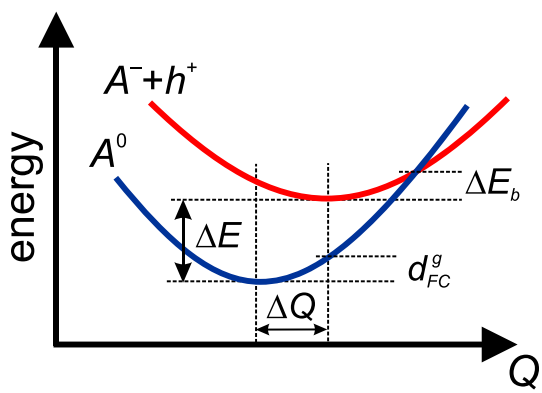

FIG. 3. Configuration coordinate diagram illustrating nonradiative capture of a hole by a deep acceptor (A).

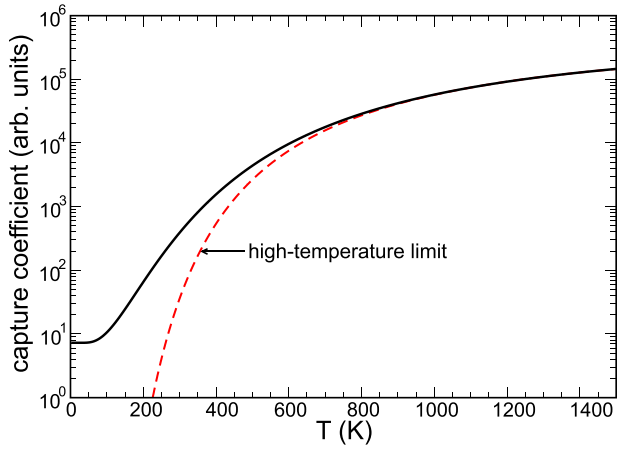

FIG. 4. Nonradiative capture coefficient at a model defect system. Solid line: quantum-mechanical calculation based on 1D cc diagram. Dashed line: high-temperature limit showing temperature-activated nature of the process.

To see how the parameters of the 1D cc diagram affect nonradiative transitions (or how these parameters can be determined from experimental measurements), let us consider a model defect with these parameters: $\Delta E = 0.80$ eV, $d_{FC} = 0.243$ eV, $\hbar\omega_0 = 30$ meV, $S = 8.1$ (cf. Fig. 3). From Eq. (15), this yields $\Delta E_b = 0.32$ eV. In Fig. 4, we show a calculated nonradiative capture coefficient (solid line) as a function of temperature. The magnitude of the capture coefficient is proportional to the square of the electron-phonon coupling matrix element.

At very low temperatures, the capture coefficient reaches a constant value. This represents quantum-mechanical tunneling between the two potential energy curves. At higher temperatures, the capture coefficient increases, revealing that there is an activation energy for the transition. If we fit the high-temperature limit to a function $C_1 \exp(-\Delta E_b'/k_B T)$, we obtain $E_b' = 0.24$ eV. An even better approach [cf. Eq. (28) of Ref. 65] is to fit to a function $C_2 \exp(-\Delta E_b''/k_B T)/\sqrt{T}$, which presents a correct analytical limit to the capture coefficient for $k_B T \gg \hbar\omega_0$. In this case, we obtain $\Delta E_b'' = 0.30$ eV, much closer to the classical barrier $\Delta E_b = 0.32$ eV. Henry and Lang[65] showed that in the limit of very large Huang-Rhys factors S and high temperatures, the latter fit yields exactly the classical barrier ΔE_b. Thus, we come to the conclusion that measurements of temperature-dependent capture coefficients provide a means to measure ΔE_b. If, at the same time, the energy difference ΔE (which is the acceptor ionization energy in our case) is known or is measured separately, one can determine d_{FC} as per Eq. (15). Therefore, this allows us to construct the 1D cc diagram for this process, as in the case of radiative transitions.

We note that the discussion based on the cc picture is analogous to the Marcus theory of electron transfer, often used to describe nonradiative processes in liquids, molecular crystals, etc.[70] The cc picture is similarly useful to study nonradiative intra-defect transitions, i.e., there is a transition between different forms of the defect in the same charge state.

IV. EXAMPLES

A. Nitrogen in ZnO

While ZnO has potential for a range of optoelectronic application, the problem of *p*-type doping remains a major challenge. Substitutional nitrogen (N_O) was once thought to be a shallow acceptor that could lead to *p*-type ZnO.[71] DFT calculations using hybrid[72] and so-called Koopmans-corrected functionals[73] challenged that view, showing the acceptor level to be 1.3–1.6 eV above the VBM. Motivated by the theoretical work, experiments were performed on bulk ZnO:N crystals.[74] These crystals were *n*-type, so the nitrogen acceptors were negatively charged (N_O^-) at thermal equilibrium. By constructing a cc diagram, theory and experiment joined forces to show that N_O is, in fact, a deep acceptor.

Room-temperature PL measurements on ZnO:N showed a "red" luminescence band,[74] centered at 1.7 eV, in good agreement with the prediction of Ref. 72. To obtain an absorption profile, PLE experiments were performed, where the intensity of red luminescence was recorded versus excitation photon energy (Fig. 5). When the photon energy is above the threshold, it excites an electron from N_O^- to the CB. The electron is then recaptured by N_O^0, resulting in red emission.

The PLE spectrum was fit using Eq. (10). At 10 K, the model parameters are $E_{abs} = 2.26$ eV and $\Gamma_0 = 0.14$ eV. At 300 K, the parameters are $E_{abs} = 2.24$ eV and $\Gamma = 0.18$ eV. The increase in vibrational broadening allows us to estimate the ground-state vibrational frequency. From Eq. (14), we obtain $\hbar\omega_0 = 35$ meV. This value is in good agreement with the calculations of Ref. 55, namely, $\hbar\omega_g = 40$ meV.

To simplify the modeling, we assume the ground and excited vibrational frequencies are the same, $\omega_e = \omega_g$. With this approximation, we can use Eq. (12) to solve for the Franck-Condon shift

$$d_{FC} = \Gamma_0^2/(2\hbar\omega_0). \quad (17)$$

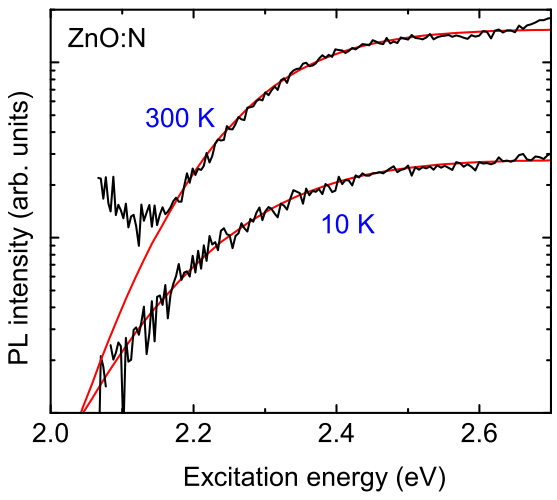

FIG. 5. PLE spectra for ZnO:N at 10 K and 300 K. The smooth red lines are fits according to Eq. (10).

Equation (17) yields $d_{FC} = 0.28$ eV. From Eq. (1), the thermodynamic energy at 10 K is $E_{therm} = 2.26–0.28 = 1.98$ eV. Therefore, the (0/–) acceptor level lies 1.98 eV below the CBM. Since the low-temperature band gap of ZnO is 3.44 eV, the acceptor level is 1.46 eV above the VBM.

It is instructive to compare this experimentally derived acceptor level, 1.46 eV, with theories that produce the correct ZnO band gap. Theories using the HSE functional give an acceptor level of 1.3 eV,[72,75] which increases to 1.46 eV when a correction for the finite supercell is applied.[75] Quantum Monte Carlo simulations give an energy of 1.6(3) eV.[76] On the experimental side, photo-EPR measurements on ZnO:N indicate the acceptor level to be ~1.6 eV above the VBM.[77,78] In summary, there is strong consensus that the N_O acceptor level is 1.3–1.6 eV above the VBM. Such a deep level cannot contribute a significant density of holes at room temperature, and hence it can be firmly concluded that nitrogen is not a shallow acceptor in ZnO.

B. DX centers

A *DX* center is a donor impurity that relaxes away from its substitutional site, becoming a deep-level defect.[61,79] The most-studied system is Si in $Al_xGa_{1-x}As$. For $Al_xGa_{1-x}As$ with $x > 0.22$, or GaAs under pressures > 2 GPa, Si becomes a *DX* center (Fig. 6). In this configuration, it accepts an electron and is negatively charged. When exposed to light, a photon of energy E_{opt} can excite the electron into the CB, neutralizing the defect. The Si atom relaxes to its substitutional site and acts as shallow donor. In order to revert to the *DX* ground state, the defect must capture an electron and surmount a barrier (~0.2 eV). At low temperatures (<180 K), this barrier is large enough to cause the Si atom to remain in its metastable donor state for hours or days. This leads to persistent photoconductivity (PPC), an increase in free-carrier density that persists even after the light source is turned off.[80]

DFT calculations by Chadi and Chang[81] indicated that, in its *DX* configuration, the Si atom occupies an interstitial site neighboring three As atoms (Fig. 6). This model has been tested by a range of experiments. Electrical characterization techniques such as Hall effect and DLTS have determined the thermodynamic energy and capture barrier as a function of *x*.[79] Positron annihilation spectroscopy showed that the *DX* center is vacancy-like, consistent with the open

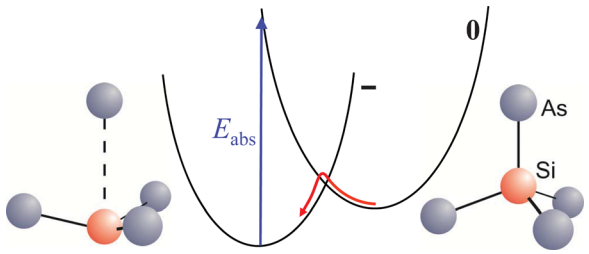

FIG. 6. Configuration coordinate diagram for the *DX* center in AlGaAs. In the *DX* configuration (left), the Si atom is displaced from its substitutional site. In the shallow-donor configuration (right), the Si atom occupies the substitutional site. The red arrow depicts a barrier for transforming from the shallow-donor to the *DX* configuration.

volume created by the displacement of the Si atom.[82] IR spectroscopy on GaAs:Si under pressure showed that the Si local vibrational mode of the *DX* state is 2% lower than the shallow-donor state, in agreement with calculations.[83]

Beyond $Al_xGa_{1-x}As$, *DX* centers have been discovered in various III–V and II–VI semiconductors.[79,84] The unique phenomenon of PPC led researchers to propose that *DX* centers could be used for holographic memory, where data are optically written and read throughout the bulk of a crystal.[85] A major practical problem is that PPC only occurs at low temperatures. Recently, room-temperature PPC was discovered in $SrTiO_3$ crystals and attributed to a *DX*-like defect.[86] When exposed to photons >2.9 eV, the free-electron density increases by over two orders of magnitude, with negligible decay at room temperature.

C. Carbon in GaN

Bulk and epitaxially grown GaN layers often exhibit yellow luminescence (YL), peaking at 2.1–2.2 eV. The origin of the YL has been discussed for a long time. In the 1980s, Ogino and Aoki[87] suggested that YL is related to the presence of carbon impurities in the GaN. However, at the time, it was believed that simple substitutional carbon is a shallow acceptor with a defect level ∼0.2 eV from the valence band. This would give rise to photoluminescence peaking at about 3.3 eV. Thus, it appeared that substitutional carbon cannot cause YL.

As in the case of ZnO:N_O, the problem was solved by accurate first principles calculations based on hybrid density functionals.[88] Calculations show the (0/–) level occurs at 0.9–1.0 eV above the VBM.[88,89] Combined with a Franck-Condon shift $d_{FC} = 0.46$ eV,[88] this yields luminescence peaking at ∼2.1 eV, explaining the experimental result. The actual 1D cc diagram for the luminescence process is shown in Fig. 7. In Fig. 8, we show the comparison of the calculated[55] and measured[90] luminescence lineshapes. For a more meaningful comparison, the ZPL of the theoretical curve was shifted upward by 0.08 eV. It is clear that the calculations agree very well with experiments. Moreover, the theoretical Huang-Rhys factors ($S_g = 11$) and effective phonon frequencies ($\hbar\omega_0 = 42$ meV) also agree well with experimental results reported in Ref. 87 ($S_g = 12.8 \pm 1.8$, $\hbar\omega_0 = 41 \pm 5$ meV) and Ref. 91 ($S_g = 8.4$, $\hbar\omega_0 = 55$ meV).

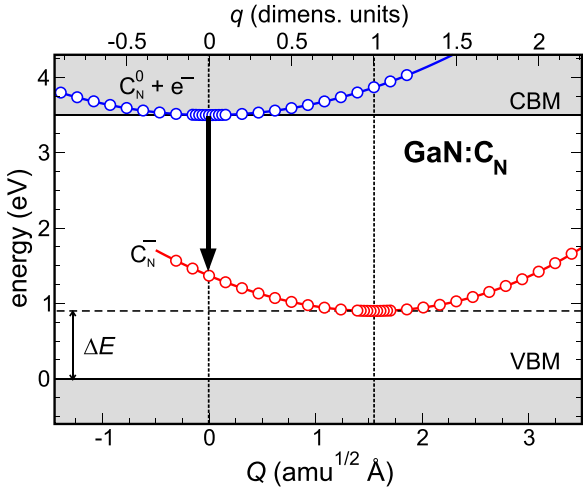

FIG. 7. Calculated configuration coordinate diagram of substitutional carbon in GaN. The vertical solid arrow illustrates the optical transition that gives rise to yellow luminescence. $\Delta E = 0.9$ eV is the ionization energy of the carbon acceptor.[88] For the quantitative definition of the configuration coordinate Q, see Ref. 55.

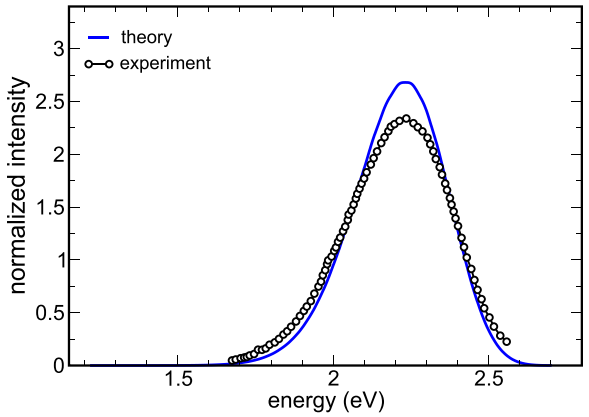

FIG. 8. Calculated (Ref. 55) and measured (Ref. 90) luminescence lineshapes for GaN:C_N.

In Ref. 67, calculations of nonradiative capture of a hole by a negatively charged substitutional carbon were reported. The results were compared to capture coefficients determined from thermal quenching of YL, measured in the temperature range 500–700 K. When the decrease of the band gap of GaN at these elevated temperatures is taken into account, the theoretical capture coefficient $C_p = 3.1 \times 10^{-7}$ cm^3 s^{-1} agrees very well with experimental results[91] $C_p = (3-6) \times 10^{-7}$ cm^3 s^{-1}. Note that the calculated capture coefficient depends very sensitively on the energy difference ΔE (cf. Fig. 3). Getting the order-of-magnitude right is already a challenge for such calculations. Keeping this in mind, the agreement between experiment and theory reported in Ref. 67 is indeed excellent; it confirms the accuracy of the 1D cc diagram pertaining to substitutional carbon in GaN calculated using hybrid functionals.

V. CONCLUSIONS

In conclusion, the combination of theory and experiment yields detailed insights into the properties of defects in semiconductors that could not be obtained from computations or from measurements alone. With the advent of more accurate density functional methods, such as hybrid functionals, and other advances, such as supercell-size corrections, theoretical and computational work has become increasingly important to interpret experimental results. Configuration coordinate diagrams, derived from theory and experiment, provide crucial information about deep-level defects. The examples in Sec. IV have illustrated that important conclusions can be drawn from such a combined theory/computation/experiment approach.

In spite of (or maybe because of) these successes, this remains an active area of new developments. On the side of theory, more accurate but still computationally tractable ways to incorporate finite-temperature effects need to be developed. More efficient methods also need to be developed for treating electron-phonon interactions in cases where one

needs to go beyond the 1D configuration coordinate model. The expanding interest in two-dimensional materials has focused attention on the development and implementation of functionals that correctly treat van der Waals interactions.[92] On the experimental side, scanning-probe and electron microscopy provide images which, when combined with first-principles calculations, yield information at the atomic scale.

Given the crucial impact of defects and impurities on the properties and technological applications of materials, the fruitful interaction of theory and experiment will continue to gain in importance and produce essential results.

ACKNOWLEDGMENTS

We acknowledge A. Janotti, Q. Yan, C. E. Dreyer, and J. L. Lyons for fruitful collaborations and N. A. Modine for discussions on the classical barrier during nonradiative capture. Work at UCSB and WSU was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES) under Award Nos. DE-SC0010689 (UCSB) and DE-FG02-07ER46386 (WSU). A.A. acknowledges support by the Marie Skłodowska-Curie Action of the European Union (project Nitride-SRH, Grant No. 657054).

APPENDIX: VALIDITY OF THE 1D CONFIGURATION COORDINATE DIAGRAM

In Sections III and IV, we used one dimensional cc diagrams to analyze absorption, luminescence, and nonradiative capture at defects. The question might arise: is this not overly simplistic? The 1D cc diagram corresponds to one phonon mode of a particular frequency. Defects significantly perturb the underlying crystal lattice, and therefore one expects a complicated spectrum of quasi-local modes. How can we reconcile this expectation with the success of 1D cc diagrams?

1D cc diagrams are certainly *not* universally valid.[49,51] However, the validity of the 1D approximation can be demonstrated for certain classes of problems. Let us again consider luminescence at a defect. When we go beyond the cc representation, we have to include the coupling to all the possible modes. The easiest way to achieve this is by introducing the so-called spectral density of electron-phonon coupling $S(\varepsilon)$[64]

$$S(\varepsilon) = \sum_k S_k \delta(\varepsilon - \varepsilon_k). \tag{A1}$$

Here, the sum runs over all phonon modes k with frequencies ω_k, and $\varepsilon_k = \hbar\omega_k$. S_k is a partial Huang-Rhys factor, which, in analogy with the total Huang-Rhys factor S introduced in Eq. (3), shows the average number of phonons of type k emitted during the optical transition. The total Huang-Rhys factor is

$$S = \int_0^{\hbar\omega_{max}} S(\varepsilon) d\varepsilon, \tag{A2}$$

ω_{max} being the largest phonon frequency in the system. Let us consider a model defect, for which we choose the spectral density $S(\varepsilon)$ shown in Fig. 9. $S(\varepsilon)$ has the following features. (i) Coupling to lower-energy modes in the energy range 0–55 meV; there is broad asymmetric maximum at 35 meV. (ii) Coupling to higher-energy modes in the energy range 65–100 meV with a well-pronounced symmetric peak at 80 meV. (iii) $S(\varepsilon) = 0$ in the energy range 55–65 meV. In our hypothetical example, this happens because there are no phonons with these energies, representing a gap between acoustical and optical modes, and the defect itself does not introduce additional modes in this gap. The total Huang-Rhys factor is $S = 15$, a typical value for deep acceptors in wide-band-gap semiconductors such as ZnO and GaN.[55,91] From the knowledge of the spectral density $S(\varepsilon)$, one can define an average phonon energy

$$\varepsilon_0 = \frac{1}{S} \int_0^{\hbar\omega_{max}} S(\varepsilon)\varepsilon d\varepsilon, \tag{A3}$$

which in our examples turns out to be 60 meV (Fig. 9). Note that, per our assumption, there are no actual phonons with this particular energy.

The luminescence lineshape $L(\varepsilon)$ is uniquely determined by the spectral density $S(\varepsilon)$ and the zero-phonon-line energy $E_{ZPL} = E_{\text{therm}}$. One possible way to determine $L(\varepsilon)$ from the knowledge of $S(\varepsilon)$ and E_{ZPL} is provided by means of the so-called generating function.[64] We do not provide explicit formulas here but refer the reader to the literature, e.g., Eqs. (8)–(10) in Ref. 59. The only other parameter that enters into the calculation is the full width at half-maximum of the zero-phonon line, γ. In all subsequent examples, we choose $\gamma = 2$ meV. γ represents the inhomogeneous broadening that is present also for T = 0 K.

To calculate the luminescence lineshape within a 1D model, one essentially needs only three parameters: E_{ZPL}, S, and the average phonon energy ε_0, defined in Eq. (A3). The analytical form of the lineshape is given by [cf. Eqs. (6) and (7)]

$$L(\varepsilon) = \sum_n e^{-S} \frac{S^n}{n!} g_\sigma(E_{ZPL} - n\varepsilon_0 - \varepsilon), \tag{A4}$$

where $g_\sigma(x)$ is a Gaussian function with a smearing parameter σ, which represents the replacement of all phonon modes with just one effective mode.

Before providing the results for our defect (let us call it "defect C") with $S = 15$, it is instructive to study two other cases, where the form of the spectral density $S(\varepsilon)$ is the same as in Fig. 9, but the function $S(\varepsilon)$ is scaled, so that for "defect A" $S = 0.3$ (weak electron-phonon coupling), whereas for "defect B" $S = 5$ (intermediate electron-phonon coupling). The calculated luminescence lineshape for "defect A" is shown in Fig. 10(a). We chose $E_{ZPL} = 2.0$ eV. In this case, most of the luminescence is in the ZPL (red arrow). The fraction of light in the ZPL is given by $e^{-S} \approx 0.75$. Thus, only 0.25 of luminescence intensity is in the phonon sideband. By analyzing this sideband [inset in Fig. 10(a)], we can clearly identify the phonon replicas that correspond to the peaks in the spectral density at 35 meV and 80 meV. First-order peaks are pronounced, and second-

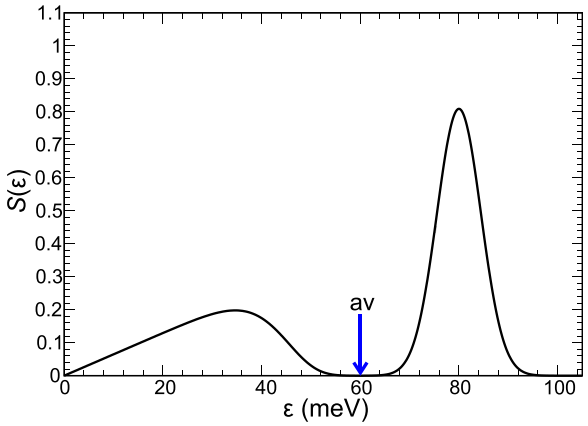

FIG. 9. The spectral density of electron-phonon coupling $S(\varepsilon)$ pertaining to a model defect. The average ("av") phonon energy is indicated by a vertical arrow.

order peaks can be identified, but higher-order peaks can no longer be distinguished in the spectrum. In the inset of Fig. 10(a), we also show a fit (dashed line) of the luminescence lineshape to a 1D formula given in Eq. (A4). The free parameter we have in the fit is the Gaussian smearing parameter σ, and we choose $\sigma = 5$ meV to get the best overall fit. Clearly, 1D model is not very good. Most importantly, it does not reproduce the position of phonon replicas. Also, it cannot reproduce the width of these replicas and the ZPL at the same time no matter what σ one chooses. For example, choosing a smaller σ would yield a better agreement for the ZPL, but a worse one for the phonon replicas.

The lineshape for "defect B" is shown in Fig. 10(b). In this case, only $e^{-5} = 0.007$ of luminescence intensity is in the ZPL, while the majority is in the phonon sideband. While close to the ZPL, we can still identify features that originate from phonons of various frequencies, these features become broader, and the lineshape smoother, further from the ZPL (smaller energies). In this case, the 1D model performs better. While it is still not able to reproduce the fine structure of the lineshape, it provides a fair approximation to the overall lineshape.

When the electron-phonon coupling increases even further, as is the case for our "defect C" with $S = 15$, the luminescence lineshape becomes a smooth function, where one can no longer identify contributions from different phonon modes, as shown in Fig. 10(c). Furthermore, the luminescence intensity of the ZPL is practically 0. It turns out that in this case, i.e., when $S \gg 1$, one does not really need to know the fine structure of the spectral density $S(\varepsilon)$, and a 1D model is a very good approximation. The resulting lineshape is shown in Fig. 10(c). We observe that the centers-of-mass of the two curves are identical, both equal to $E_{ZPL} - S\varepsilon_0 = 1.10$ eV, which is exactly the classical Franck-Condon transition energy E_{em} corresponding to emission. However, the lineshape pertaining to the full calculation is slightly more asymmetric. We conclude that while the 1D model does not *exactly* reproduce the full calculation, it clearly provides a highly accurate description of the luminescence lineshape, and thus the phonon physics, for defects with strong electron-phonon coupling ($S \gg 1$). In our example, we assumed that phonon frequencies in the ground and the excited states were equal. This is not a restriction; the usefulness of the 1D approximation can be similarly shown with explicit calculations[55] in case the frequencies are different.

Our example has demonstrated why the 1D cc diagram picture is so useful, since many defects exhibit strong electron-phonon coupling. Even for cases with weak or intermediate electron-phonon coupling, the 1D picture can still be valuable as an approximation to the overall lineshape, even though the fine structure cannot be captured.

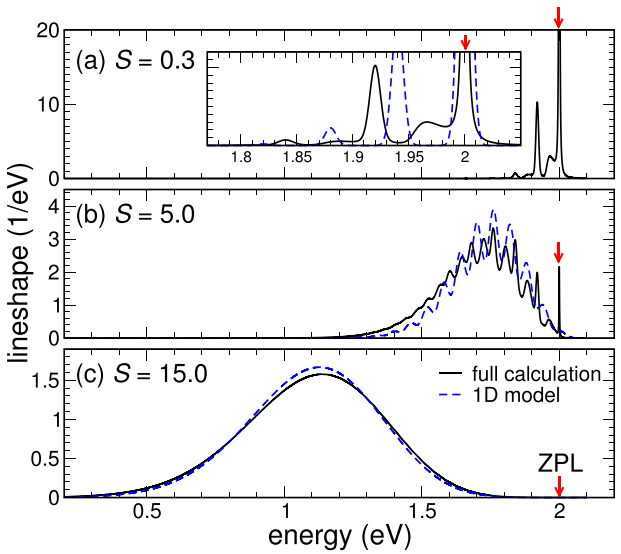

FIG. 10. Luminescence lineshapes of three model defects. (a) Weak electron-phonon coupling (Huang-Rhys factor $S = 0.3$). (b) Intermediate electron-phonon coupling ($S = 5.0$). (c) Strong electron-phonon coupling ($S = 15.0$). The ZPL is indicated with a (red) arrow. Solid line shows the actual luminescence lineshape, and the dashed line shows a calculation based on a one-dimensional configuration coordinate diagram ("1D model").

[1]N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (W. B. Saunders Co., 1976).
[2]L. Gordon, J. R. Weber, J. B. Varley, A. Janotti, D. D. Awschalom, and C. G. Van de Walle, MRS Bull. **38**, 802 (2013).
[3]L. Childress, R. Walsworth, and M. D. Lukin, Phys. Today **67**(10), 38 (2014).
[4]A. Krotkus, J. Phys. D: Appl. Phys. **43**, 273001 (2010).
[5]J. Appl. Phys. **30**, 1117 (1959).
[6]C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, Rev. Mod. Phys. **86**, 253 (2014).
[7]S. K. Estreicher, Phys. Status Solidi B **217**, 513 (2000).
[8]C. G. Van de Walle and J. Neugebauer, J. Appl. Phys. **95**, 3851 (2004).
[9]Phys. Status Solidi B **248**, 17 (2011).
[10]*Theory of Defects in Semiconductors*, edited by D. A. Drabold and S. K. Estreicher (Springer-Verlag, Berlin, 2007).
[11]*Advanced Calculations for Defects in Materials*, edited by A. Alkauskas, P. Deák, J. Neugebauer, A. Pasquarello, and C. G. Van de Walle (Wiley, New York, 2011).
[12]M. Lannoo and J. Bourgoin, *Point Defects in Semiconductors I: Theoretical Aspects* (Springer-Verlag, Berlin, 1981); *Point Defects in Semiconductors II: Experimental Aspects* (Springer-Verlag, Berlin, 1983).
[13]*Deep Centers in Semiconductors: A State-of-the-Art Approach*, 2nd ed., edited by S. T. Pantelides (Gordon and Breach Science, Yverdon, 1992).
[14]*Identification of Defects in Semiconductors*, in Semiconductors and Semimetals Vols. 51A and 51B, edited by M. Stavola (Academic Press, 1998).

[15]M. Spaeth and H. Overhof, *Point Defects in Semiconductors and Insulators: Determination of Atomic and Electronic Structure from Paramagnetic Hyperfine Interactions* (Springer, 2003).
[16]M. D. McCluskey and E. E. Haller, *Dopants and Defects in Semiconductors* (CRC Press, 2012).
[17]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965).
[18]W. Kohn, Rev. Mod. Phys. **71**, S59 (1999).
[19]M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[20]P. Rinke, A. Schleife, E. Kioupakis, A. Janotti, C. Rödl, F. Bechstedt, M. Scheffer, and C. G. Van de Walle, Phys. Rev. Lett. **108**, 126404 (2012).
[21]W. Chen and A. Pasquarello, J. Phys.: Condens. Matt. **27**, 133202 (2015).
[22]J. P. Perdew, K. Burke, and M. Ernzerhof, J. Chem. Phys. **105**, 9982 (1996).
[23]G. Pacchioni, F. Frigoli, D. Ricci, and J. A. Weil, Phys. Rev. B **63**, 054102 (2000).
[24]J. L. Lyons, A. Janotti, and C. G. Van de Walle, J. Appl. Phys. **115**, 012014 (2014).
[25]J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003); Erratum: J. Chem. Phys. **124**, 219906 (2006).
[26]J. L. Gavartin, D. Muñoz Ramo, A. L. Shluger, G. Bersuker, and B. H. Lee, Appl. Phys. Lett. **89**, 082908 (2006).
[27]A. Alkauskas and A. Pasquarello, Physica B **401–402**, 670 (2007).
[28]F. Oba, A. Togo, I. Tanaka, L. Paier, and G. Kresse, Phys. Rev. B **77**, 245202 (2008).
[29]J. L. Lyons, A. Janotti, and C. G. Van de Walle, Phys. Rev. B **89**, 035204 (2014).
[30]P. Deák, B. Aradi, T. Frauenheim, E. Janzén, and A. Gali, Phys. Rev. B **81**, 153203 (2010).
[31]C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Phys. Rev. Lett. **102**, 016402 (2009).
[32]C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Phys. Status Solidi B **248**, 1067 (2011).
[33]See http://www.eag.com/mc/sims-sensitivity-detection-limits.html for sensitivity limits for various impurities.
[34]C. G. Van de Walle and J. Neugebauer, Annu. Rev. Mater. Res. **36**, 179 (2006).
[35]H. Bracht and E. E. Haller, Phys. Rev. Lett. **81**, 393 (1998).
[36]F. Tuomisto and I. Makkonen, Rev. Mod. Phys. **85**, 1583 (2013).
[37]C. G. Van de Walle, Phys. Rev. B **68**, 165209 (2003).
[38]A. Janotti, B. Jalan, S. Stemmer, and C. G. Van de Walle, Appl. Phys. Lett. **100**, 262104 (2012).
[39]P. A. Lee, P. H. Citrin, P. Eisenberger, and B. M. Kincaid, Rev. Mod. Phys. **53**, 769 (1981).
[40]G. D. Watkins, "Identification of defects in semiconductors," in *Semiconductors and Semimetals*, edited by M. Stavola (Academic Press, San Diego, 1999), Vol. 51A, Chap. 1, p. 1.
[41]C. G. Van de Walle and P. E. Blöchl, Phys. Rev. B **47**, 4244 (1993).
[42]T. A. Kennedy and E. R. Glaser, "Identification of defects in semiconductors," in *Semiconductors and Semimetals*, edited by M. Stavola (Academic Press, San Diego, 1999), Vol. 51A, Chap. 3, p. 93.
[43]P. M. Mooney, "Identification of defects in semiconductors," in *Semiconductors and Semimetals*, edited by M. Stavola (Academic Press, San Diego, 1999), Vol. 51B, Chap. 2, p. 93.
[44]M. D. McCluskey, J. Appl. Phys. **87**, 3593 (2000).
[45]S. K. Estreicher, D. Backlund, T. M. Gibbons, and A. Doçaj, Modell. Simul. Mater. Sci. Eng. **17**, 084006 (2009).
[46]S. Limpijumnong, J. E. Northrup, and C. G. Van de Walle, Phys. Rev. B **68**, 075206 (2003).
[47]J. I. Pankove, *Optical Processes in Semiconductors* (Dover, New York, 1971).
[48]G. Davies, "Identification of defects in semiconductors," in *Semiconductors and Semimetals*, edited by M. Stavola (Academic Press, San Diego, 1999), Vol. 51B, Chap. 1, p. 1.
[49]A. M. Stoneham, *Theory of Defects in Solids: Electronic Structure of Defects in Insulators and Semiconductors* (Oxford University Press, Oxford, 1975).
[50]M. Lax, J. Chem. Phys. **20**, 1752 (1952).
[51]J. J. Markham, Rev. Mod. Phys. **31**, 956 (1959).
[52]I. S. Osad'ko, Usp. Fiz. Nauk. **128**, 21 (1979).
[53]K. Huang and A. Rhys, Proc. R. Soc. A **204**, 406 (1950).
[54]J. L. Lyons, A. Janotti, and C. G. Van de Walle, Phys. Rev. Lett. **108**, 156403 (2012).
[55]A. Alkauskas, J. L. Lyons, D. Steiauf, and C. G. Van de Walle, Phys. Rev. Lett. **109**, 267401 (2012).
[56]S. Shinoya, T. Koda, K. Era, and H. Fujirawa, J. Phys. Soc. Jpn. **19**, 1157 (1964).
[57]M. A. Reshchikov and H. Morkoç, J. Appl. Phys. **97**, 061301 (2005).
[58]M. K. Kretov, I. M. Iskandarova, B. V. Potapkin, A. V. Scherbinin, A. M. Srivastava, and N. F. Stepanov, J. Lumin. **132**, 2143 (2012).
[59]A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. Van de Walle, New J. Phys. **16**, 073026 (2014).
[60]J. L. Lyons, A. Alkauskas, A. Janotti, and C. G. Van de Walle, Phys. Status Solidi B **252**, 900 (2015).
[61]D. V. Lang, R. A. Logan, and M. Jaros, Phys. Rev. B **19**, 1015 (1979).
[62]A. A. Kopylov and A. N. Pikhtin, Sov. Phys. Solid State **16**, 1200 (1975).
[63]H. Gummel and M. Lax, Ann. Phys. **2**, 28 (1957).
[64]R. Kubo and Y. Toyozawa, Prog. Theor. Phys. **13**, 160 (1955).
[65]C. H. Henry and D. V. Lang, Phys. Rev. B **15**, 989 (1977).
[66]L. Shi and L.-W. Wang, Phys. Rev. Lett. **109**, 245501 (2012).
[67]A. Alkauskas, Q. Yan, and C. G. Van de Walle, Phys. Rev. B **90**, 075202 (2014).
[68]L. Shi, K. Xu, and L.-W. Wang, Phys. Rev. B **91**, 205315 (2015).
[69]G. D. Barmparis, Y. S. Puzyrev, X.-G. Zhang, and S. T. Pantelides, Phys. Rev. B **92**, 214111 (2015).
[70]R. A. Marcus, Ann. Rev. Phys. Chem. **15**, 155 (1964).
[71]M. D. McCluskey and S. J. Jokela, J. Appl. Phys. **106**, 071101 (2009).
[72]J. L. Lyons, A. Janotti, and C. G. Van de Walle, Appl. Phys. Lett. **95**, 252105 (2009).
[73]S. Lany and A. Zunger, Phys. Rev. B **81**, 205209 (2010).
[74]M. C. Tarun, M. Zafar Iqbal, and M. D. McCluskey, AIP Adv. **1**, 022105 (2011).
[75]S. Sakong, J. Gutjahr, and P. Kratzer, J. Chem. Phys. **138**, 234702 (2013).
[76]E. Ertekin, "Quantum Monte Carlo calculations for nitrogen in ZnO," unpublished.
[77]J. E. Stehr, D. M. Hofmann, and B. K. Meyer, J. Appl. Phys. **112**, 103511 (2012).
[78]M. D. McCluskey, C. D. Corolewski, J. Lv, M. C. Tarun, S. T. Teklemichael, E. D. Walter, M. G. Norton, K. W. Harrison, and S. Ha, J. Appl. Phys. **117**, 112802 (2015).
[79]P. M. Mooney, J. Appl. Phys. **67**, R1 (1990).
[80]R. J. Nelson, Appl. Phys. Lett. **31**, 351 (1977).
[81]D. J. Chadi and K. J. Chang, Phys. Rev. Lett. **61**, 873 (1988).
[82]J. Mäkinen, T. Laine, K. Saarinen, P. Hautojärvi, C. Corbel, V. M. Airaksinen, and P. Gibart, Phys. Rev. Lett. **71**, 3154 (1993).
[83]J. A. Wolk, M. B. Kruger, J. N. Heyman, W. Walukiewicz, R. Jeanloz, and E. E. Haller, Phys. Rev. Lett. **66**, 774 (1991).
[84]S. H. Wei and S. B. Zhang, Phys. Rev. B **66**, 155211 (2002).
[85]L. Hesselink, S. S. Orlov, and M. C. Bashaw, Proc. IEEE **92**, 1231 (2004).
[86]M. C. Tarun, F. A. Selim, and M. D. McCluskey, Phys. Rev. Lett. **111**, 187403 (2013).
[87]T. Ogino and M. Aoki, Jpn. J. Appl. Phys. **19**, 2395 (1980).
[88]J. L. Lyons, A. Janotti, and C. G. Van de Walle, Appl. Phys. Lett. **97**, 152108 (2010).
[89]S. G. Christenson, W. Xie, Y. Y. Sun, and S. B. Zhang, J. Appl. Phys. **118**, 135708 (2015).
[90]C. H. Seager, D. R. Tallant, J. Yu, and W. Götz, J. Lumin. **106**, 115 (2004).
[91]M. A. Reshchikov, AIP Conf. Proc. **1583**, 127 (2014).
[92]H. Peelaers and C. G. Van de Walle, J. Phys. Condens. Matter **26**, 305502 (2014).