SAND2014-17861J

Density Functional Theory Calculations of Activation Energies for Non-radiative Carrier Capture by Deep Defect Levels in Semiconductors

N. A. Modine, A. F. Wright, and S. R. Lee

Sandia National Laboratories, Albuquerque, New Mexico 87185-1315, USA

Corresponding Author Email: namodin@sandia.gov

Abstract

Carrier recombination due to defects can have a major impact on device performance. The rate of defect-induced carrier recombination is determined by both defect levels and carrier capture cross-sections. Kohn-Sham density functional theory (DFT) has been widely and successfully used to predict defect levels in semiconductors and insulators, but only recently has work begun to focus on using DFT to determine carrier capture cross-sections. Lang and Henry worked out the fundamental theory of carrier-capture cross-sections in the 1970s and showed that, in most cases, room temperature carrier-capture cross-sections differ between defects primarily due to differences in the carrier capture activation energies. We present an approach to using DFT to calculate carrier capture activation energies that does not depend on perturbation theory or an assumed configuration coordinate, and we demonstrate this approach for the -3/-2 level of the Ga vacancy in wurtzite GaN.

1. Introduction

Point defects in semiconductors and insulators may exist in more than one charge state. Defects can interact with the band edges by capturing and emitting carriers in order

to change between these charge states. These processes are technologically important since they induce carrier recombination, which can have a dramatic influence on device performance. [1,2] The rate of defect-induced carrier recombination is determined by both the thermodynamic levels and carrier capture cross-sections associated with the particular defect. Both of these types of parameters are needed in models that determine the rate of recombination induced by defects.[1,2,3,4] Both measurements and calculations of defect levels are a mainstay of modern materials science. However, measurements and calculations of carrier capture cross-sections have been performed much less frequently.

Furthermore, experiments that measure defect levels do not directly identify their atomistic origin (e.g., vacancy, interstitial, antisite, etc.). Theoretical calculations of defect levels for specific defects can be compared with experimental measurement, but it often remains difficult to unambiguously assign measured defect levels to a particular defect species. Calculations of additional defect properties, such as carrier capture cross-sections, would allow the comparison of more than one property and greatly aid in this identification.

The most widely used approach to calculating defect levels relies on Kohn-Sham density functional theory (DFT).[5] In cases where comparisons have been made between theory and experiment, defect levels calculated by DFT (using periodic finite-size supercells and then corrected *a posterior* for interactions between supercells and the alignment of the delocalized band edge states to experiment[6,7,8,9,10]) often agree with experimental levels to within a few tenths of an electron volt (eV). [11,12,13,14,15,16] Given the success of DFT in calculating defect levels and other properties of defects, it makes

sense to develop an approach by which DFT can be used to determine carrier capture cross-sections.

In the 1970s, based on a combination of experiment and theory, Lang and Henry developed the fundamental theory of non-radiative carrier capture by deep defect levels in semiconductors. [17,18] Building on previous work by Hwang and Rhys[19], Lax[20], and Kubo and Toyazawa[21], they established that multiphonon emission (MPE) is the predominant mechanism by which carriers are captured by deep defect levels. This mechanism will be reviewed in detail in Section 2 below, but a key result is that, at moderate to high temperatures, typical carrier capture cross-sections σ have the form

$$\sigma = \sigma_\infty \exp(-E_A/k_B T), \qquad (1)$$

where σ_∞ is the high temperature limit of σ, E_A is a carrier capture activation energy, k_B is the Boltzmann constant, and T is the temperature. Lang and Henry found that σ_∞ varies in a relatively small range with $\sigma_\infty = (0.5 - 4.0) \times 10^{-15}$ cm^2 for neutral defects, an enhancement of 16-24 for an attractive Coulomb interaction between the defect and the carrier, and a similar decrease for a repulsive interaction. However, room temperature carrier capture cross-sections differ by many orders of magnitude due to variations in E_A between different defect and levels. Therefore, the focus of this article will be on using DFT to determine carrier capture activation energies. At temperatures where Eq. 1 is valid, the ions essentially move in a classical fashion as they climb the barrier E_A to the point where carrier capture can occur. Thus, one advantage of our focus on carrier capture activation energies is that we will be able to treat the ions as classical particles.

In contrast, other recent work on using DFT to determine carrier capture cross-sections has focused on a fully quantum mechanical, perturbative approach to calculating carrier capture cross-sections directly. [22,23] These sophisticated treatments allow effects that require a quantum mechanical treatment of the ions, such as carrier capture by tunneling at low temperature, to be correctly included, and they are likely to give accurate answers in many cases. They should be particularly accurate when the structures of the initial and final defect states are similar, and the defect does not have to move very far from its relaxed configuration in order to capture the carrier. However, in some cases, the defect might need to move quite far from its relaxed configuration in order to capture a carrier. In such cases, it is not clear that quantities (e.g. electron-phonon coupling constants, phonon modes and frequencies, etc.) calculated at the relaxed structure will remain valid in the region where the electronic transition occurs. Likewise, when the initial and final defect states are quite different (e.g., a split interstitial that changes into a tetrahedral interstitial) it is unclear how to choose a configuration coordinate for the capture process. Since our classical treatment of the ions allows us to explore configuration space far from the relaxed defect structures, we can address these issues in a straightforward manner.

The purpose of this paper is to seek a method to calculate carrier capture activation energies using DFT. We develop this method as follows: In Section 2, we review the fundamental theory of carrier capture activation energies as developed by Lang and Henry. Then, in Section 3, we translate the key aspects of this fundamental theory into quantities that can be determined using DFT. We proceed in Section 4 by presenting an

algorithm to construct an optimized configuration coordinate in multidimensional space, before giving some concluding remarks in Section 5.

2. Fundamental Theory of Carrier Capture Activation Energies

As mentioned above, Lang and Henry established MPE as the most common mechanism for non-radiative carrier capture at a deep defect level in semiconductors.[17,18] Figure 1 describes the basic physical processes associated with MPE in schematic form. In the basic theory, it is assumed that there exists a one-dimensional path in the configuration space of the system, called the "configuration coordinate", along which the atoms move during the carrier capture or emission processes. In Fig. 1, motion along this coordinate is parameterized by the x-coordinate, which applies to both the upper and lower panels of the figure.

As shown in Fig. 1a (the upper panel of Fig. 1), the energy of the system changes as the system moves along the configuration coordinate. Consider carrier capture or emission by the $q-1/q$ level of a defect. Three energies, corresponding to different charge states of the defect and different numbers of carriers in the band edges, are important to the process. These energies are indicated by the three lines in Fig. 1a. The blue line gives the energy of a defect in charge state q with no carriers in the band edges, the black line gives the energy of the defect in charge state $q-1$ with a hole in the valence band, and the red line gives the energy of the defect in charge state q with a hole in the valence band and an electron in the conduction band.

It is assumed that the configuration coordinate passes through the minimum energy configuration for both charge states. Thus, the minimum of the blue curve corresponds to the fully relaxed structure and energy of the defect in charge state q. Likewise, the minimum of the black line corresponds to the fully relaxed structure and energy of the defect in charge state $q-1$, but the energy is increased by the (coordinate-independent) energy of a hole in the valence band. The red line in the same as the blue line but shifted upward by the energy of both a hole in the valence band and an electron in the conduction band, which is equal to the band gap of the material. Since the process depends only on differences in energies, the zero of the energy scale has been taken arbitrarily to be the energy of the fully relaxed defect in charge state q.

The lower panel in Fig. 1 (Fig. 1b) shows the defect level (in black) as the system moves along the configuration coordinate. We have chosen to reference the defect level to the valence band edge (in blue), and thus, the defect level corresponds to the difference between the black and blue curves in Fig. 1a. Likewise, the defect level crosses the conduction band (in red), when the black curve crosses the red curve in Fig. 1a. This defect level differs from the thermodynamic level that is usually discussed in defect theory. In this case, both charge states are constrained to be in the same configuration, which is determined by the configuration coordinate, while the thermodynamic level is given by the difference in energy between the two charge states when each charge state is separately relaxed to its minimum energy configuration. In Fig. 1, the $q-1/q$ thermodynamic level of the defect is given by the energy difference between the minimum of the black line and the minimum of the blue line.

Each of the energies in Fig. 1a corresponds to a system with total charge q, and thus these energies can be compared directly without consideration of the Fermi level. In order for the system to make a non-radiative transition between two states, energy must be conserved. In a simple classical picture of the MPE process, this can occur when the black line crosses either the blue line or the red line in Fig. 1a. These points correspond to the configurations where the defect level crosses into either the valence band or the conduction band in Fig. 1b.

If the system consists of a defect in charge state $q-1$ along with a hole in the valence band, it will thermally fluctuate along the configuration coordinate with its potential energy given by the black line in Fig. 1a. Occasionally, it can reach the crossing point where its energy is the same as a defect in charge state q, and at that point, the hole can be captured with a commensurate change in defect charge state from $q-1$ to q. In order for this to happen, the motion along the configuration coordinate must have thermal energy equal to the energy difference between the crossing point of the black and blue curves and the minimum of the black curve. This energy difference corresponds to the activation energy for the defect in charge state $q-1$ to capture a hole.[17,18]

Once the hole has been captured, the defect will begin to move with a new potential energy function given by the blue line in Fig. 1a. As the system moves back down the blue curve, an additional amount of kinetic energy equal to the energy difference between the minima of the black and blue curves will be deposited into vibrations along the configuration coordinate. Since this additional energy, which is equal to the depth of the thermodynamic defect level relative to the valence band edge, is typically large compared to the energy of the phonons in the system, this process results in the emission of many

phonons and leaves the system in a highly excited vibrational state. This behavior is the essence of the MPE process and explains how a large number of photons can be emitted in order to absorb the large electronic energy released during carrier capture by a deep level.[17,18]

The reverse process occurs when a defect in charge state q thermally fluctuates along the configuration coordinate with its potential energy given by the blue line in Fig. 1a. If it can overcome a barrier given by the energy difference between the crossing point of the black and blue curves and the minimum of the blue curve, it can reach the crossing point and non-radiatively emit a hole with a commensurate change in defect charge state from q to $q-1$. The barrier for hole emission is larger than the barrier for hole capture by the energy difference between the minima of the black and blue curves in Fig. 1a, which again is equal to the depth of the thermodynamic defect level relative to the valence band edge. If we follow the conventional procedure, we factor out this additional barrier for carrier emission, so that capture and emission of a carrier have the same cross-section, with carrier emission being suppressed by an additional factor that is exponential in the thermodynamic defect level depth divided by the thermal energy.

Electron capture and emission occur by analogous processes. Note that the systems used to define the black and red curves in Fig. 1a both contain a hole in the valence band, which does not participate in these processes. The presence of this hole is important in the above discussion of hole capture, but it shifts both the black and red curves in Fig. 1a by the same constant factor and has no effect on the results for electron capture and emission. A system containing a defect in charge state q along with an electron in the conduction band can thermally fluctuate along the configuration coordinate with its

potential energy given by the red curve in Fig. 1a. If it reaches the crossing point between the black and red curves, the defect can non-radiatively capture the electron with the defect charge state changing to $q-1$. In the process, it must overcome a barrier given by the energy difference between the crossing point of the red and black curves and the minimum of the red curve. This is the barrier for a defect in charge state q to capture an electron. Likewise, a defect in charge state $q-1$ can thermally fluctuate along the configuration coordinate with its potential energy given by the black curve in Fig. 1a. If it reaches the crossing point between the black and red curves, it can emit an electron with its charge state changing to q. The additional energy kinetic energy deposited during electron capture and the additional barrier that must be overcome during electron emission is given by the energy difference between the minima of the red and black curves, which is equal to the depth of the thermodynamic defect level relative to the conduction band edge.

3. DFT Calculations of Carrier Capture Activation Energies

We wish to translate the above fundamental theory into the language of DFT calculations in the hopes of using DFT to obtain carrier capture activation energies for specific defect levels. There are two significant obstacles impeding this translation: (1) DFT is a theory of electronic ground states, while, all but the lowest energy values in Fig. 1a formally correspond to excited states of the defect containing system, and (2) The configuration space in a DFT calculation consists of three coordinates for each atom in the system while the above theory assumes that a one-dimensional configuration

coordinate has somehow been identified. Confronting the second obstacle will be the subject of the next section, while this section will deal with the first concern.

In DFT calculations, it is not possible to create a system with a defect in charge state $q-1$ and a hole in the valence band (or a defect in charge state $q+1$ and an electron in the conduction band) independently from a system with a defect in charge state q. One merely chooses the total charge of the system, and DFT will find the electronic configuration that has the lowest energy. This prevents a direct calculation of the curves in Fig. 1 using DFT. A partial solution to this problem lies in observing that a carrier in a band edge of an infinite crystal is delocalized throughout the crystal and should have a vanishing interaction with a localized defect. Thus, the energy of a system containing a defect and a carrier should simply be the energy of a defect plus a constant energy for adding the carrier. If a suitable energy for adding each carrier can be determined, this should allow us to evaluate portions of the curves in Fig. 1a using DFT calculations.

However, the above procedure provides only a partial solution because a defect in charge state q can equally well turn into a defect in charge state $q-1$ and a hole (or a defect in charge state $q+1$ and an electron) if the energy of the latter system is lower. Thus, DFT calculations cannot be used to evaluate the blue and red curves in Fig. 1a to the right of the crossing point between the black and blue curves. Likewise, DFT calculations cannot be used to evaluate the black curve in Fig. 1a to the left of the crossing point between the black and red curves. Fortunately, the carrier capture activation energies given by the above theory depend on the crossing point energies, but not the behavior of the curves outside of the crossing points. Thus, there is hope that the carrier capture activation energies can be obtained from DFT calculations. In fact, for

DFT calculations in finite supercells, we will see that there is also a region inside the crossing points where the calculated energies become undependable. However, when this region is sufficiently small, we can extrapolate the curves obtained outside this region to make reasonable estimates of the curves in Fig. 1 throughout the crossing point regions, the crossing point energies, and the carrier capture activation energies.

We will now return to the question of determining suitable energies for adding a hole in the valence band or an electron in the conduction band. One might think that the Kohn-Sham eigenvalues at the valence band maximum and conduction band minimum, or perhaps, the experimentally determined ionization potential and electron affinity, would be appropriate. However, outside the crossing points in Fig. 1, our approach gives us two different ways to calculate the energy of the defect plus a carrier. For example, to the right of the crossing point between the black and blue curves, we can obtain the energy of the defect in charge state $q-1$ plus a hole either by performing a DFT calculation with a total excess charge q, or by performing a DFT calculation with a total excess charge of $q-1$ and adding the (yet to be determined) energy for adding a hole. For consistency and in order to help in interpreting our DFT results, we would like these two approaches to, as nearly as possible, give the same energy.

In recent work, we have explored the case where charge added to a DFT calculation for a defect-containing supercell (which might already be charged) goes into band-edge-like states rather than localizing on the defect. [24] In this case, the energy for adding the additional charge depends only slightly on the defect identity, with a repulsive interaction between the defect and the added carrier giving a slightly higher energy that decreases with increasing supercell size, and a long-range attractive interaction giving a small

reduction in energy corresponding to formation of a shallow defect state. However, this energy is a rather strong function of the supercell size and the Brillouin sampling approach - An entire unit of charge is added to the band edge states, and for typical supercell sizes used in DFT calculations, this corresponds to filling electron states substantially above the conduction band minimum or hole states substantially below the valence band maximum. This consequential dependence on the supercell size and the other DFT technical parameters can be captured by approximating the energy for adding the additional charge to the defect containing supercell by the energy for adding the same charge to the corresponding defect-free bulk cell.[24] This approximation is independent of the particular defect under consideration or its specific configuration, and this is the approach that we will use to define the energies for adding a hole in the valence band or an electron in the conduction band when analyzing our DFT calculations for carrier capture activation energies.

To be specific, consider a DFT calculation for a defect using a particular supercell and set of technical parameters (Brillouin-zone sampling, basis set, pseudopotential set, etc.). Let $E^D(q,\vec{R})$ be the total energy of defect D in charge state q when the atomic positions are given by the generalized coordinate $\vec{R}$. We can perform DFT calculations for a defect-free bulk supercell of the same size using the same set of technical parameters. Suppose $E^B(0)$ is the total energy of the neutral bulk supercell, $E^B(-1)$ is the total energy of the bulk supercell with one electron added, and $E^B(+1)$ is the total energy of the bulk supercell with one electron removed. We can then define the energy to add an electron to the bulk supercell

$$\Delta^B(-1) = E^B(-1) - E^B(0), \qquad (2)$$

and the negative of the energy to add a hole to the bulk supercell

$$\Delta^B(+1) = E^B(0) - E^B(+1). \qquad (3)$$

We have chosen the sign in the definition of $\Delta^B(+1)$ to be consistent with our previous work, where we have shown that $\Delta^B(-1)$ and $\Delta^B(+1)$ provide (approximate) upper and lower bounds on the range of defect levels that can be obtained from DFT calculations using a particular supercell and set of technical parameters.[24]

We can then approximate the energy of a system with a defect in charge state $q-1$ and a hole in the valence band by

$$E_h^D\left(q-1,\vec{R}\right) \equiv E^D\left(q-1,\vec{R}\right) - \Delta^B(+1), \quad (4)$$

Likewise, we can approximate the energy of a system with a defect in charge state q, an electron in the conduction band, and a hole in the valence band by

$$E_{eh}^D\left(q,\vec{R}\right) \equiv E^D\left(q,\vec{R}\right) + \Delta^B(-1) - \Delta^B(+1) \qquad (5)$$

The quantity

$$E_{GAP}^B \equiv \Delta^B(-1) - \Delta^B(+1) = E^B(-1) + E^B(+1) - 2E^B(0) \qquad (6)$$

is equivalent in form to the *fundamental gap* defined by Mori-Sanchez, Cohen, and Yang[25] for an infinite bulk system, but here applied to a finite supercell. In the limit of an infinite supercell, E_{GAP}^B becomes the difference in the valence band maximum and conduction band minimum eigenvalues (the Kohn-Sham gap) for Kohn-Sham calculations using semilocal exchange-correlation functionals (e.g., LDA or GGA) or for generalized Kohn-Sham calculations with hybrid functionals (e.g., PBE0, B3LYP, or HSE). However, for typical supercells used in DFT calculations, this quantity is substantially larger than the Kohn-Sham gap and can approach the experimental gap even with semilocal exchange-correlation functionals.[24]

We now have approximations for each of the energies appearing in Fig. 1a, and we wish to define the defect level appearing in Fig. 1b. The defect level $\Delta^D(q-1/q)$is often defined to be the value of the Fermi level where the defect in charge state q has the same formation energy as the defect in charge state $q-1$. This definition can be expressed in terms of supercell energies as [24,26,27]

$$\Delta^D\left(q-1/q,\vec{R}\right)=E^D\left(q-1,\vec{R}\right)-E^D\left(q,\vec{R}\right)-\mu_e, \quad (7)$$

where the defect levels are defined relative to the reference energy μ_e. When calculating defect levels using DFT, it is fairly common to take $\mu_e = \varepsilon_{VBE}$, where ε_{VBE} is the energy of the bulk valence band maximum. Instead, we will take $\mu_e = \Delta^B(+1)$, which gives the expression

$$\Delta^D\left(q-1/q,\vec{R}\right)=E^D\left(q-1,\vec{R}\right)-E^D\left(q,\vec{R}\right)-\Delta^B(+1). \quad (8)$$

This choice simply shifts all band edges and levels by a constant value, and it ensures that our defect level corresponds to the difference between $E_h^D\left(q-1,\vec{R}\right)$ and $E^D\left(q,\vec{R}\right)$. With this reference energy, our defect levels should be (approximately) bound between 0 and E_{GAP}^B.[24]

We now have DFT-based expressions for all of the quantities appearing in Fig. 1. In order to demonstrate DFT calculations of these quantities and our subsequent analysis to determine carrier capture activation energies, we have chosen an example system: the -3/-2 level of the Ga vacancy (V_{Ga}) in wurtzite GaN. This example is interesting because it is one of the levels that possibly contributes to the broad yellow luminescence often seen in GaN.[28,29] Although other defects such as a carbon impurity on a nitrogen site are likely to contribute also to luminescence in this frequency range,[30,31,32] radiative capture of an electron by V_{Ga} in the -2 charge state, possibly in a complex with another defect species,

has been considered as an important source of yellow luminescence. If radiative capture is the dominant electron capture mechanism, the cross-section for non-radiative capture must be small at the relevant temperature. Photoluminescence lifetime measurements indicate a radiative electron capture cross-section of $(2.7 \pm 0.5) \times 10^{-21}$ cm^{-2} for the defect responsible for yellow luminescence. [33] Using Eq. 1 along with typical values of σ_∞,[17,18] we obtain $E_A > 14\, k_B T \sim 0.6$ eV for the radiative process to remain the dominant process at 500 K. A calculated electron capture barrier that is at least this large would be consistent with -3/-2 level of V_{Ga} being a possible contributor to yellow luminescence.

Our DFT calculations were performed with the open-source Socorro code, [34] using a plane-wave basis with a 30 Rydberg cutoff to represent the Kohn-Sham orbitals, the Projector Augmented Wave (PAW) method [35], and the LDA[5,36] for exchange and correlation. Thirteen electrons were treated as valence ($3d^{10}$, $4s^2$, and $4p^1$) for Ga, and five electrons were treated as valence ($2s^2$ and $2p^3$) for N. In order to model the V_{Ga} defect, we removed one Ga atom from a 72 atom supercell consisting of 3×3×2 primitive wurtzite cells. A 2×2×2 Monkhorst-Pack [37] mesh was used to sample the Brillouin zone, and occupations of the Kohn-Sham orbitals were calculated using a Fermi distribution with kT = 6.8×10^{-2} eV. Structures were relaxed until the maximum force component was less than 13 meV/Å. A spatially uniform compensating background charge density was used to neutralize the cell for charge state calculations. [38] No attempt to realign or post-process energies or levels in order to remove interactions between the periodically repeated supercells or align the calculations to experimental band structure was made. Detailed results for the gallium vacancy in wurtzite GaN using a similar set of

computational parameters with the VASP code [39] and Vanderbilt ultrasoft pseudopotentials [40] can be found in Ref. 41.

Our example system is intended to provide a computationally undemanding demonstration of our procedure for calculating carrier capture activation energies, and our computational parameters are not intended to provide fully converged, quantitative predictions of defect properties. In particular, the size of the supercell and the Brillouin zone sampling would likely need to be increased in order to get-well converged results. Furthermore, a better approximation for the exchange-correlation functional, such as one of the hybrid functionals, might be needed in order to get details of the defect behavior correct. [23] Drawing dependable physical conclusions from our calculations will require future work testing whether these issues significantly affect our results.

In order to demonstrate our procedure for using DFT to calculate carrier capture activation energies independently of the multidimensional search for an appropriate configuration coordinate, which is the subject of the next section, we have chosen an assumed configuration coordinate. Let $\vec{R}_{-2}$ and $\vec{R}_{-3}$ be generalized coordinates describing the positions of all of the atoms in the relaxed V_{Ga} structures for the -2 and -3 charge states respectively. Then, our configuration coordinate λ parameterizes the set of configurations given by

$$\vec{R}(\lambda) = \vec{R}_{-2} + \lambda\left(\vec{R}_{-3} - \vec{R}_{-2}\right)/\left\|\vec{R}_{-3} - \vec{R}_{-2}\right\|, \qquad (9)$$

which describes linear extrapolation along the line through $\vec{R}_{-2}$ and $\vec{R}_{-3}$. This is the same configuration coordinate that is assumed in the perturbative approach described in Ref. 23. The expression $\left\|\vec{R}_{-3} - \vec{R}_{-2}\right\|$ indicates the Euclidean norm of the vector

$\vec{R}_{-3} - \vec{R}_{-2}$, and therefore, λ measures the generalized distance between the configurations $\vec{R}(\lambda)$ and $\vec{R}_{-2}$.

Our calculations indicate that the relaxed configurations of V_{Ga} in both the -2 and -3 charge states have C_{3v} symmetry with three equivalent neighboring N atoms and one inequivalent neighboring N atom, which is located along the *c* axis from the vacancy site. For purposes of defining directions, orient the V_{Ga} structure so that the vacancy site is the center, and the neighboring atom along the *c* axis is above the vacancy. Relative to their bulk positions, all of the neighboring atoms relax away from the vacancy site for both charge states, with the *c* axis atom moving up and the three equivalent atoms moving down and outward. We find that the structures of the -2 and -3 charge states differ slightly. In moving from -2 charge state to the -3 charge state, the *c* axis neighbor moves up by 0.028 Å while the three equivalent neighbors move up by 0.003 Å and inward by 0.019 Å. As our configuration coordinate increases, these atoms will continue to move along these directions, while negative values of the configuration coordinate correspond to motion in the opposite directions.

Figure 2 shows the results of our DFT calculations for V_{Ga} with this assumed configuration coordinate. The upper panel (Fig. 2a) shows the three DFT energies that we believe correspond to the energies that appear in the fundamental theory of non-radiative carrier capture. The blue line plots $E^D\left(-2, \vec{R}(\lambda)\right)$, the black line plots $E_h^D\left(-3, \vec{R}(\lambda)\right)$, and the red line plots $E_{eh}^D\left(-2, \vec{R}(\lambda)\right)$. In the lower panel (Fig. 2b), the black line plots the defect level $\Delta^D\left(-3/-2, \vec{R}(\lambda)\right)$, while the blue and black lines give the lower and upper (approximate) bounds on our defect levels, which are zero and E_{GAP}^B respectively. Our DFT calculations give $E_{GAP}^B = 3.35$ eV, while the experimental band

gap of wurtzite GaN is 3.47 eV at 0 K and 3.39 eV at 300K. [42] We obtain a 2.18 eV Kohn-Sham band gap at the Γ point and a 2.67 eV Kohn-Sham band gap using our Brillouin zone sampling, so the agreement between E_{GAP}^{B} and experiment is a result of both our Brillouin zone sampling and band filling effects.[24]

We can compare Fig. 2 to Fig. 1. One of the most notable differences is that the curves in Fig. 2 do not cross. As we discussed above, this reflects the fact that the -2 charge state becomes a -3 charge state plus a hole if we move far enough to the right, and the -3 charge state becomes a -2 charge state plus an electron if we move far enough to the left. The relative vertical positioning of the curves is independently determined from calculations for the corresponding bulk system, so it is testament to the accuracy of our approximations that the curves in Fig. 2 merge cleanly.

From Fig. 2a, it is difficult to determine where the curves would cross if this merging did not occur. However, the defect level in Fig. 2b shows a clearly linear region that bends over to follow the lower bound. This non-linear region near the bound indicates that the -2 charge state is beginning to convert to a -3 charge state with the extra electron coming from the valence band and that the calculated energy for the -2 charge state can no longer be trusted in this region. In contrast, the defect level deviates from linear behavior well before it reaches the upper bound. Reference to the Fig. 2a shows that when this happens the defect has already gained about 5 eV of energy due to the distortion as it moves along the configuration coordinate, and the deviation from linear behavior probably results from a significant reorganization of the electronic structure in response to this large distortion.

The dotted line in Fig. 2b indicates a linear extrapolation of the defect level in the linear region to the bounds. This linear extrapolation represents an attempt to determine the behavior of the defect level if the defects in our calculations could not change charge state, and the points where the linear extrapolation crosses the bounds approximately indicate where the curves in Fig. 2a would cross if this were the case. One could equally well extrapolate to experimentally determined band edges if the DFT results could be aligned with the experimental band structure by, for example, aligning a carefully converged defect level calculation with well-established experimental results. Given the close agreement between E_{GAP}^{B} and the experimental band gap in this case and remembering that our calculations are intended to illustrate our approach rather than provide accurate predictions, we believe that it is reasonable to extrapolate to the bounds.

The vertical dotted lines in Fig. 2a indicate the values of the configuration coordinate where our defect level extrapolation indicates that the curves should cross if the defect charge state could be fixed. For the crossing point on the right, evaluating the difference between the black curve (which cleanly represents a -3 charge state in this region) and its minimum (the relaxed energy for the -3 charge state) gives a hole capture activation energy of 0.89 eV. Similarly, for the crossing point on the left, evaluating the difference between the blue curve (which cleanly represents a -2 charge state in this region) and its minimum (the relaxed energy for the -2 charge state) gives an electron capture activation energy of 10.1 eV. These are very large carrier capture activation energies, which in part reflects the small difference in structure between the -2 and -3 charge states and the large band gap of GaN. The enormous activation energy for non-radiative electron capture is consistent with radiative capture being the dominant process. However, the very large

activation energy for hole capture suggests that so few holes would be captured (at least be a non-radiative process) that very little luminesce would be observed. Temperature dependent photoluminescence measurements indicate a hole capture cross-section of $(2.7 \pm 1.3) \times 10^{-14}$ cm^{-2} for the main process responsible for yellow luminescence in GaN near room-temperature.[33] Even considering that the Coulomb attraction to a -3 defect would enhance the rate of hole capture, this suggests that the barrier for hole capture could not be larger than a few $k_B T$ if the -3/-2 level of V_{Ga} is this defect. In addition, the activation energy for electron capture, and the associated distortion of the defect structure, is so large that it is very difficult to believe that our assumed configuration coordinate remains valid in this region. We will now present an approach to optimizing the configuration coordinate used in our calculations and see whether it significantly changes these results.

4. Multidimensional Optimization of Carrier Capture Processes

In this section, we wish to construct an algorithm for finding activation energies for carrier capture similar to the various transition state finding algorithms that are widely used to find activation energies for fixed charge state processes. [43,44] If we were not concerned with the limitations of DFT, we could define the transition state for hole capture by the $q - 1/q$ level of a defect as the lowest energy point in configuration space where the energy of the defect in charge state q is the same as the energy of the defect in charge state $q - 1$ plus a hole in the valence band. Likewise, we could define the transition state for electron capture as the lowest energy point in configuration space where the energy of the defect in charge state $q - 1$ is the same as the energy of the

defect in charge state q plus an electron in the conduction band. For the same reason that we expect thermally activated, fixed charge state processes to be dominated by trajectories whose energy exceeds the transition state energy by no more that the thermal energy $k_B T$, we could expect carrier capture processes to be dominated by trajectories with energies within $k_B T$ of the relevant carrier capture transition state energy. Therefore, the difference in energy between the carrier capture transition state and the relaxed defect would give the activation energy for carrier capture. However, these definitions cannot be used directly when DFT calculations are used to calculate our defect energies.

We showed above that DFT calculations using finite supercells become unreliable as the energy of one charge state approaches the energy of another charge state plus a carrier, which is exactly the condition defining these transitions states. In fact, due to anticrossing between adiabatic states, the above definitions might be problematical even if we had access to a method that would allow us to efficiently perform the required excited state calculations. In order to overcome this problem, we will generalize the concept of a carrier capture transition state to a carrier capture transition pathway consisting of a series of configurations that allow us to extrapolate to the carrier capture transition state while remaining in a region of configuration space where DFT calculations for both charge states are reliable. Like transition pathways for thermally activated changes in configuration, we do not expect typical trajectories of systems undergoing the transition, which involve non-negligible momentum, to coincide with the transition pathway. In addition, there may be more than one useful definition of a carrier

capture transition pathway. However, we will provide a definition that we believe is a proper generalization of the configuration coordinate concept to multidimensional space.

What characterizes our "in theory" definitions of the carrier capture transition states is that they are the lowest energy configurations consistent with a fixed energy difference between the two charge states. By Eq. 8, this is equivalent to a constraint on the defect level. The hole capture transition state is defined by a minimization over configurations with the defect level at the valence band edge, while the electron capture transition state is defined by a minimization over configurations with the defect level at the conduction band edge. A logical extension of these definitions to a pathway is to consider the defect level as a parameter that defines the configurations in the pathway. For each defect level, we can perform a constrained energy minimization over all configurations consistent with that defect level. As the defect level approaches the valence band edge, this pathway should approach the hole capture transition state, while as the defect level approaches the conduction band edge, the pathway should approach the electron capture transition state.

If we calculate this pathway using DFT energies, the results may become unreliable close to the carrier capture transition states for the same reason that our DFT calculations with an assumed configuration coordinate became unreliable close to the crossing points. However, for defect levels intermediate between the upper and lower bounds, the energies of both charge states should be accurately determined, and the carrier capture transition pathway should be well defined. This well-behaved portion of the pathway can then be used to extrapolate to the carrier capture transition states.

There are a couple of important properties of the carrier capture transition pathway defined in this manner that make it suitable for use as a configuration coordinate in

carrier capture calculations. First, if we consider only the constrained subspace of configurations consistent with a specific defect level, the energies of the two charge states differ only by a constant factor, which is determined by the defect level. Therefore, within this subspace, minimization with respect to the energy of charge state q is exactly equivalent to minimization with respect to the energy of charge state $q-1$. This ensures that the calculated pathways for hole capture, electron emission, electron capture, and hole emission are all the same. Also, when the defect level is constrained to the value

$$E^D\left(q-1,\vec{R}_q\right)-E^D\left(q,\vec{R}_q\right)-\Delta^B(+1) \qquad (10)$$

where $\vec{R}_q$ is the relaxed configuration of charge state q, the constrained subspace must include the configuration $\vec{R}_q$. Since the configuration $\vec{R}_q$ is the global minimum of $E^D\left(q,\vec{R}\right)$, it must also be the minimum over the constrained subspace. Therefore, the pathway must include the relaxed configuration of charge state q. Likewise, when the defect level is constrained to the value

$$E^D\left(q-1,\vec{R}_{q-1}\right)-E^D\left(q,\vec{R}_{q-1}\right)-\Delta^B(+1) \qquad (11)$$

where $\vec{R}_{q-1}$ is the relaxed configuration of charge state q, the minimum over the constrained subspace must be $\vec{R}_{q-1}$. Therefore, the pathway must include the relaxed configuration of charge state $q-1$. Together, these results satisfy the condition that the configuration coordinate should pass through the minimum energy configurations of both charge states.

Just as there may be local minima in the energy as a function defined on the space of all configurations, there may be local minima in the energy as a function defined on the constrained subspace. As the defect level defining the constraint varies, the global

energy minimum may occasionally change from one local minimum to another leading to a discontinuity in the pathway. This process is analogous to a first order phase transition in thermodynamics where the global free energy minimum switches from one local free energy minimum to another leading to a discontinuity in the order parameter. In this sense, the carrier capture transition pathway considered in this work differs from a conventional configuration coordinate. However, given the smoothness of the energy function, we should expect few such discontinuities in our pathway, and unless such a discontinuity occurs close to one of the carrier capture transition states, they should not seriously impede our extrapolation to the transition state.

If energies and forces for the defect charge states are calculated using DFT, the constrained minimization defining the configurations in the carrier capture transition pathway can be easily accomplished by a combination of a relaxation step perpendicular to the force difference and a Newton step parallel to the force difference. Let $\vec{F}^D(q,\vec{R})$ be the force on charge state q and $\vec{F}^D(q-1,\vec{R})$ be the force on charge state $q-1$. Consider a minimization with the defect level constrained to δ. Define the sum and difference forces $\vec{\Omega} = \vec{F}^D(q,\vec{R}) + \vec{F}^D(q-1,\vec{R})$ and $\vec{\Theta} = \vec{F}^D(q,\vec{R}) - \vec{F}^D(q-1,\vec{R})$. $\vec{\Theta}$ is the gradient of $\Delta^D(q-1/q,\vec{R}_i)$, so moves perpendicular to $\vec{\Theta}$ do not change $\Delta^D(q-1/q,\vec{R}_i)$ to linear order. Let α be a damping parameter, which must be chosen small enough to avoid oscillations in the relaxing system. Then, the iteration

$$\vec{R}_{i+1} = \vec{R}_i + \alpha\left(\vec{\Omega} - \vec{\Theta}\left(\vec{\Theta}\cdot\vec{\Omega}\right)/\left(\vec{\Theta}\cdot\vec{\Theta}\right)\right) + \vec{\Theta}\left(\delta - \Delta^D(q-1/q,\vec{R}_i)\right)/\left(\vec{\Theta}\cdot\vec{\Theta}\right), \qquad (12)$$

where $\vec{\Omega}$ and $\vec{\Theta}$ are recalculated at each $\vec{R}_i$, will accomplish the constrained minimization. The first term relaxes the system in directions perpendicular to $\vec{\Theta}$ in order to minimize the

energy within the constrained subspace, while the second term moves along $\vec{\Theta}$ in order to maintain the constraint. The steepest descent relaxation may be replaced by more efficient methods such as quenched dynamics or conjugate gradient as long as all terms are kept orthogonal to $\vec{\Theta}$. As is usual with multidimensional minimization, the algorithm should be started from a number of low symmetry configurations in order to make sure that the global minimum has been located.

The carrier capture transition path is then calculated by performing this constrained minimization for a series of defect levels δ. If we wish to construct the entire path from electron capture / emission to hole capture / emission, these values of δ should range between the lower and upper bounds (0 and E_{GAP}^{B} using our reference for the defect levels). Given a series of configurations in the path, we can then construct a configuration coordinate that measures the distance along the path by summing the distances between sequential configurations in the path.

We have applied this multidimensional optimization procedure to our V_{Ga} example. The results are shown in Fig. 3. The general behavior of the results in Fig. 2 and Fig. 3 are similar, but there are some notable differences. Moving from the minimum of the blue curve to the right (toward hole capture / emission), our results do not change significantly from those we obtained with an assumed coordinate. In fact, if the data in Figs. 2a and 3a were plotted together, the curves in Fig. 3a in this region would be only slightly below the curves in Fig. 2a. This indicates that the assumed configuration coordinate that we used in constructing Fig. 2 is actually a good approximation. Moving to the left from the minimum of the blue curve, we initially find a similar picture. The configurations retain C_{3v} symmetry, and the energies of the two charge states are only

slightly lower when following the optimized pathway. However, at a configuration coordinate of about -0.4 Å, the picture suddenly changes. The configurations in the pathway begin a complex distortion that fully breaks the symmetry of the structure to C_1. Comparison with Fig. 2 shows that this distortion results in a major reduction in the energy required to reach a given defect level, which we can expect to reduce the barrier for electron capture.

In order to reveal the character of the electron capture pathway once the symmetry breaking occurs, Fig. 4 shows the optimized configuration corresponding to a 3.22 eV defect level. This value is close to the upper bound, but before the deviation from linear behavior due to the loss of an electron from the -3 charge state to the conduction band becomes significant. In Fig. 4, one of the N atom first neighbors of the missing Ga atom has moved past its three Ga neighbors to form a dimer with another N atom. This second N atom, which was originally a third neighbor of the missing Ga atom, has broken a N-Ga back-bond in order to form the dimer. Since breaking this N-Ga bond leaves a Ga dangling bond, the net effect of this structural reorganization is to convert a N dangling bond into a Ga dangling bond while keeping the total number of bonds unchanged. In the -3 charge state of V_{Ga}, the initial N dangling bond is filled, while the final Ga dangling bond prefers to be empty. Thus, when carried to completion, this process results in electron emission, while the reverse process results in electron capture. Although the energies associated with this process are sufficiently high that it would almost never occur in practice, these results demonstrate the power of our multidimensional optimization procedure to predict complex, but physically sensible pathways that might be very difficult to identify otherwise.

Fig. 3b shows that as the C_1 distortion progresses, the defect level eventually resumes a nearly linear climb toward the upper bound. In this case, it does not make sense to fit the data to a single linear expression. Instead, there are linear regions near each bound, and we separately extrapolate the defect level to the bounds. These extrapolations are indicated by the dotted lines in Fig. 3b. There is some ambiguity in exactly which points should be included in these linear extrapolations, but it is easy to see that different procedures do not lead to very big changes in the final results. The vertical dotted lines in Fig. 3a indicate the values of the configuration coordinate where our extrapolations indicates that the defect level should cross the bounds if the linear behavior was maintained. For the crossing point on the right, evaluating the difference between the black curve and its minimum gives a hole capture activation energy of 0.83 eV, which is only a slight reduction from the value we obtained with our assumed configuration coordinate. Similarly, for the crossing point on the left, evaluating the difference between the blue curve and its minimum gives an electron capture activation energy of 3.45 eV. Thus, our optimization of the carrier capture pathway reduces the electron capture activation energy by almost a factor of 3. However, the barrier still easily exceeds the 0.6 eV value needed to be consistent with a radiative mechanism for this process.

Finally, we would like to make a few comments on the behavior of the optimized pathway when the defect level is very close to or beyond the bounds. The algorithm described above does not limit the changes in configuration to the atoms near the defect. As the bounds are reached and defect charge is lost to delocalized band edge states, the bulk-like region surrounding the defect can begin to distort in an attempt to raise the

energy of these delocalized states. In our example system, this is observed most dramatically as the defect level reaches the lower bound. Trying to extend the carrier capture pathway to values of the defect level below the last point shown in Fig. 3 leads to a large, discontinuous, delocalized distortion in structure. The effect is less dramatic as the defect level reaches the upper bound, and the calculated pathway remains continuous. However, a delocalized distortion is observed, the defect level moves above the bound, and correspondingly the black line moves above the red line in Fig. 3a. These anomalous results are an artifact of the DFT calculations and the fact that the defect can lose charge to band edge states. Thus, these results emphasize the importance of avoiding the crossing regions in DFT calculations of carrier capture activation energies.

5. Conclusions

In this article, we have developed an approach to calculating carrier capture activation energies using DFT that does not depend on a perturbation expansion around the relaxed defect configurations. Therefore, it allows us to fully incorporate non-linear effects. In the process, we addressed difficulties arising from the ground state nature of DFT and the multidimensional nature of configuration space. Our approach allows us to calculate rather than assume a configuration coordinate for the carrier capture process, and this configuration coordinate can represent an arbitrary nonlinear pathway in configuration space.

We demonstrated our approach with calculations for the-3/-2 level of the Ga vacancy (V_{Ga}) in wurtzite GaN. In this case, our results suggest that approaches based on perturbation theory or an assumed coordinate would work well for hole capture but fail

dramatically for electron capture. We obtain an extremely large electron capture activation energy, which is consistent with the suggestion that a radiative process is responsible for electron capture by this defect level. We also obtain a very large hole capture activation energy, which does not seem to be consistent with the measured hole capture cross-section associated with yellow luminescence. It is possible that better convergence of our calculations with respect to supercell size and Brillouin zone sampling could reduce the activation energy for hole capture. A more accurate exchange-correlation functional, such as one of the hybrid functionals, could also change our results. Also, it has been suggested that yellow luminescence arises not from V_{Ga} itself, but from a complex of V_{Ga} with another defect. It is easy to imagine how a nearby defect could increase the displacement between the -2 and -3 charge states of the defect and reduce the activation energy for hole capture. Finally, it is possible V_{Ga} does not contribute significantly to yellow luminescence, and another defect such as a carbon impurity on a nitrogen site is exclusively responsible.[30,31,32]

Finally, we should note that our approach takes a purely classical view of ionic motion, which contrasts with the very sophisticated, fully quantum mechanical calculations involved in modern perturbative approaches to carrier capture. We believe that our work is complementary to these perturbative treatments of the problem and that some combination of the approaches is likely to provide the ultimate solution.

Acknowledgements

This work was supported by Sandia's Solid-State Lighting Science Energy Frontier Research Center, sponsored by the Department of Energy Office of Basic Energy Science. Sandia National Laboratories is a multi-program laboratory managed and

operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

---

[1] W. Shockley and W.T. Read, Jr., Phys. Rev. **87**, 853 (1953).

[2] R.N. Hall, Phys. Rev. **83**, 288 (1951); **87**, 387 (1952).

[3] C.-T. Sah and W. Shockley, Phys. Rev. **109**, 1103 (1958).

[4] N. A. Modine, A. M. Armstrong, M. H. Crawford, and W. W. Chow, J. Appl. Phys. **114**, 144502 (2013).

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

[6] Y. Bar-Yam and J.D. Joannopoulos, Phys. Rev. B **30**, 1844 (1984).

[7] G. Makov and M.C. Payne, Phys. Rev. B **51**, 4014 (1995).

[8] S. Lany and A. Zunger, Phys. Rev. B **78**, 235104 (2008)..

[9] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, Rev. Mod. Phys. **86**, 253 (2014).

[10] A.F. Wright and N.A. Modine, Phys. Rev. B **74**, 235209 (2006).

[11] H.-P. Komsa and A. Pasquarello, Phys. Rev. B **84**, 075207 (2011).

[12] A. Alkauskas, P. Broqvist, and A. Pasquarello, Phys. Rev. Lett. **101**, 046405 (2008).

[13] P.A. Schultz, Phys. Rev. Lett. 96, 246401 (2006).

[14] P.A. Schultz and O.A. von Lilienfeld, Modelling Simul. Mater. Sci. Eng. **17**, 084007 (2009).

[15] A. F. Wright, Phys. Rev. B **74**, 165116 (2006).

[16] R.R. Wixom and A.F. Wright, Phys. Rev. B **74**, 205208 (2006).

[17] D.V. Lang and C.H. Henry. Phys. Rev. Lett. **35**, 1525 (1975).

[18] C.H. Henry and D.V. Lang, Phys. Rev. B **15**, 989 (1977).

[19] K. Huang and A. Rhys, Proc. R. Soc. A **204**, 406 (1950).

[20] M. Lax, J. Chem. Phys. **20**, 1752 (1952).

[21] R. Kubo and I. Toyazawa, Prog. Theor. Phys. **13**, 160 (1955).

[22] L. Shi and L.-W. Wang, Phys. Rev. Lett. **109**, 245501 (2012).

[23]A. Alkauskas, Q. Yan, and C. G. Van de Walle, Phys. Rev. B **90**, 075202 (2014).

[24] N. A. Modine, A. F. Wright, and S. R. Lee, Comput. Mater. Sci. **92**, 431 (2014).

[25] P. Mori-Sanchez, A.J. Cohen, and W. Yang, Phys. Rev. Lett. 100, 146401 (2008).

[26] A. Alkauskas, P. Broqvist, and A. Pasquarello, Phys. Stat. Sol. B **248**, 775 (2011).

[27] C. D. Latham, M. Alatalo, R. M. Nieminen, R. Jones, S. Oberg, and P. R. Bridden, Phys. Rev. B **72**, 235205 (2005).

[28] T. Matilla and R. M. Nieminen, Phys. Rev. B **55**, 9571 (1997) and references therein.

[29] J. Neugebauer and C. G. Van de Walle, Appl. Phys. Lett. **69**, 503 (1996) and references therein.

[30] T. Ogino and M. Aoki, Jpn. J. Appl. Phys. **19**, 2395 (1980).

[31] C. H. Seager, D. R. Tallant, J. Yu, and W. Gotz, J. Lumin. **106**, 115 (2004).

[32] J. L. Lyons, A. Janotti, and C. G. Van de Walle, Appl. Phys. Lett. **97**, 152108 (2010).

[33] M. Reschikov and H. Morkoc, J. Appl. Phys **97**, 061301 (2005).

[34] See http://dft.sandia.gov for a detailed description.

[35] P. E. Blochl, Phys. Rev. B **30**, 17953 (1994).

[36] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980); J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[37] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[38] Y. Bar-Yam and J.D. Joannopoulos, Phys. Rev. B **30**, 1844 (1984).

[39] See http://www.vasp.at.

[40]D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[41]A. F. Wright, J. Appl. Phys. **90**, 1164 (2001).

[42] See <http://www.ioffe.ru/SVA/NSM/Semicond/GaN/bandstr.html>.

[43] H. Jonsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker, (World Scientific, 1998), p. 385.

[44] G. Henkelman and H. Jonsson, J. Chem. Phys. **111**, 7010 (1999).

Figure Captions

Figure 1: Schematic showing the basic physical mechanism of non-radiative carrier capture at a deep defect level in a semiconductor by multiphonon emission. The x-coordinate in both panels is a one-dimensional configuration coordinate that parameterizes the motion of the atoms in the system along some path in configuration space. The upper panel (Fig. 1a) shows the energy of the defect for various charge states of the defect and occupations of the band edges, which are designated as follows: (blue) defect charge state q with no carriers in the band edges, (black) defect charge state $q-1$ with a hole, and (red) defect charge state q with a hole and an electron. The lower panel (Fig. 1b) shows the defect level (black) along with the valence band edge (blue) and the conduction band edge (red).

Figure 2: Results of our DFT calculations for non-radiative carrier capture by the -3/-2 level of the Ga vacancy in wurtzite GaN with an assumed configuration coordinate. The x-coordinate in both panels parameterizes the motion of the atoms along a straight line in configuration space that passes through the relaxed structures of the -2 and -3 charge states. The upper panel (Fig. 2a) shows the energy of the defect for various charge states of the defect and occupations of the band edges, which are designated as follows: (blue) defect charge state -2 with no carriers in the band edges, (black) defect charge state -3 with a hole, and (red) defect charge state -2 with a hole and an electron. The lower panel (Fig. 2b) shows the -3/-2 defect level (black) along with the lower (blue) and upper (red) bounds on the defect level. The dotted line in Fig. 2b shows a linear extrapolation of the

defect level, and the vertical dotted lines in Fig. 2a indicate where this extrapolation crosses the bounds.

Figure 3: Results of our DFT calculations for non-radiative carrier capture by the -3/-2 level of the Ga vacancy in wurtzite GaN with an optimized configuration coordinate. The x-coordinate in both panels parameterizes the motion of the atoms along path consisting of configurations that minimize the total energy of the defects for a particular defect level. The upper panel (Fig. 3a) shows the energy of the defect for various charge states of the defect and occupations of the band edges, which are designated as follows: (blue) defect charge state -2 with no carriers in the band edges, (black) defect charge state -3 with a hole, and (red) defect charge state -2 with a hole and an electron. The lower panel (Fig. 3b) shows the -3/-2 defect level (black) along with the lower (blue) and upper (red) bounds on the defect level. The dotted lines in Fig. 3b show a linear extrapolation of the defect level near to each bound, and the vertical dotted lines in Fig. 3a indicate where these extrapolations cross the bounds.

Figure 4: A ball-and-stick model showing a configuration from the optimized electron capture pathway for the -3/-2 level of the Ga vacancy in wurtzite GaN. The configuration corresponds to a defect level close to the upper bound. The green balls represent Ga atoms, while the brown balls represent N atoms. The 4 N atoms and 12 Ga atoms on the left are the first and second neighbors, respectively, of the missing Ga atom, while the N dimer that forms during the electron capture / emission process is apparent on the right.

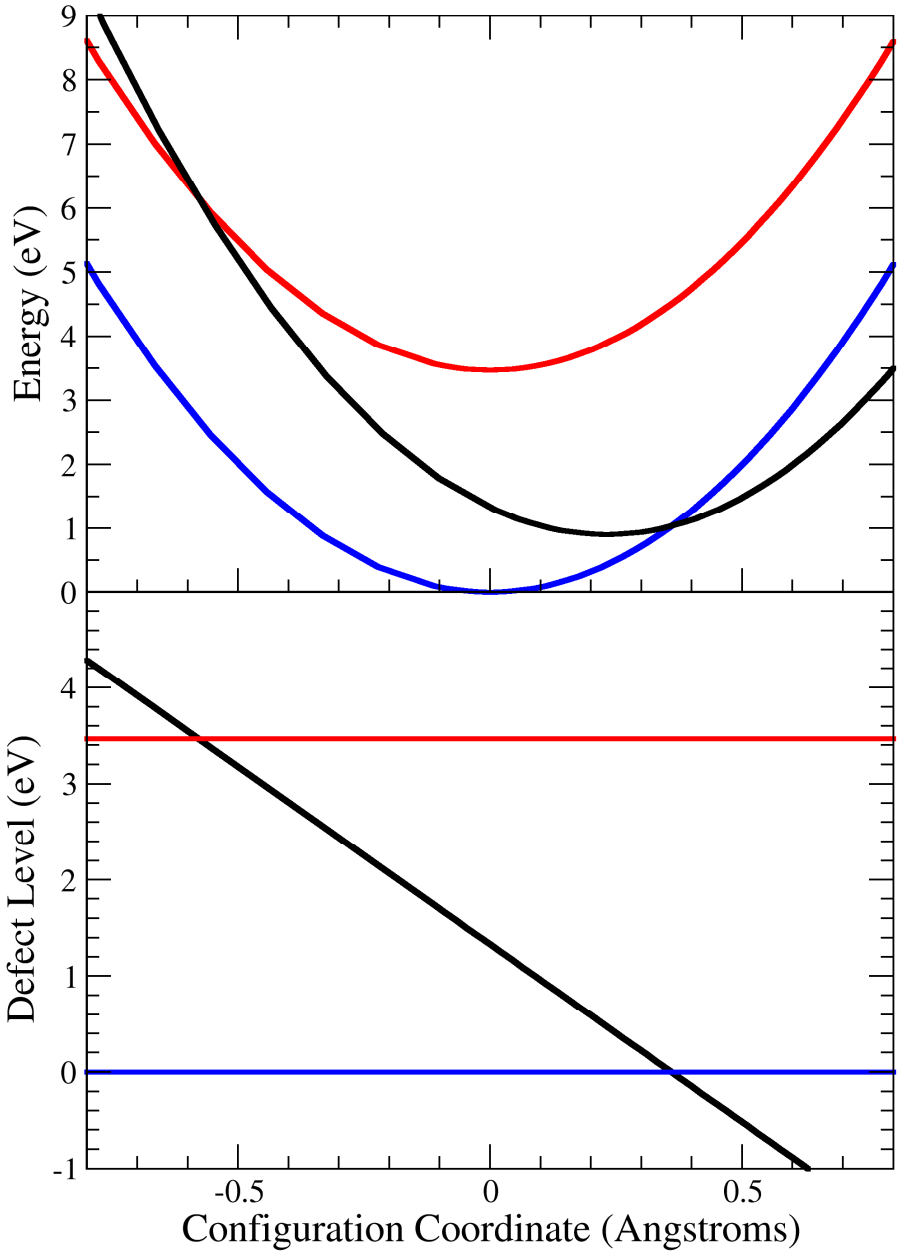
Energy (eV)
9
8
7
6
5
4
3
2
1
0
Defect Level (eV)
4
3
2
1
0
-1
-0.5
0
0.5
Configuration Coordinate (Angstroms)

Figure 2:

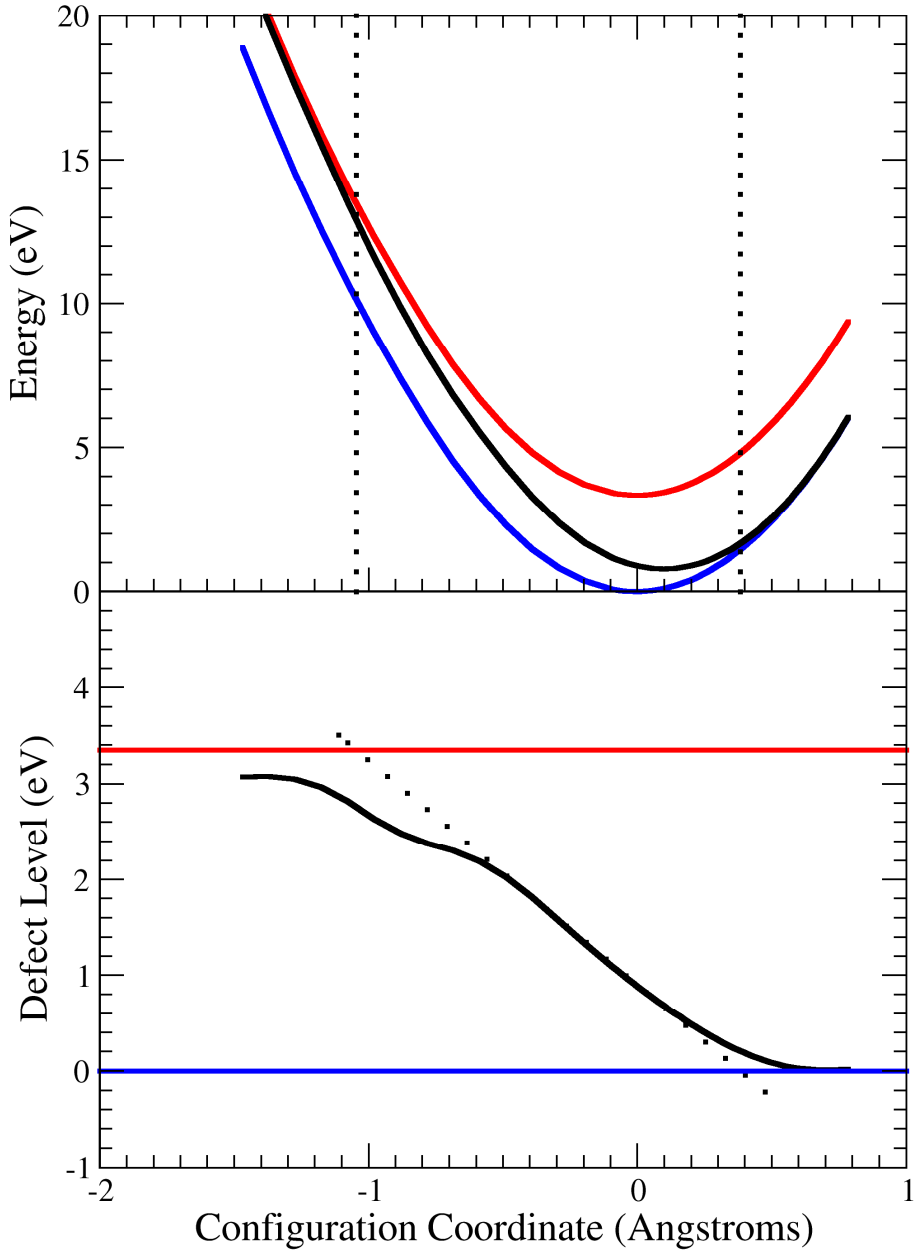
Energy (eV)
20
15
10
5
0
Defect Level (eV)
4
3
2
1
0
-1
-2
-1
0
1
Configuration Coordinate (Angstroms)

Figure 3:

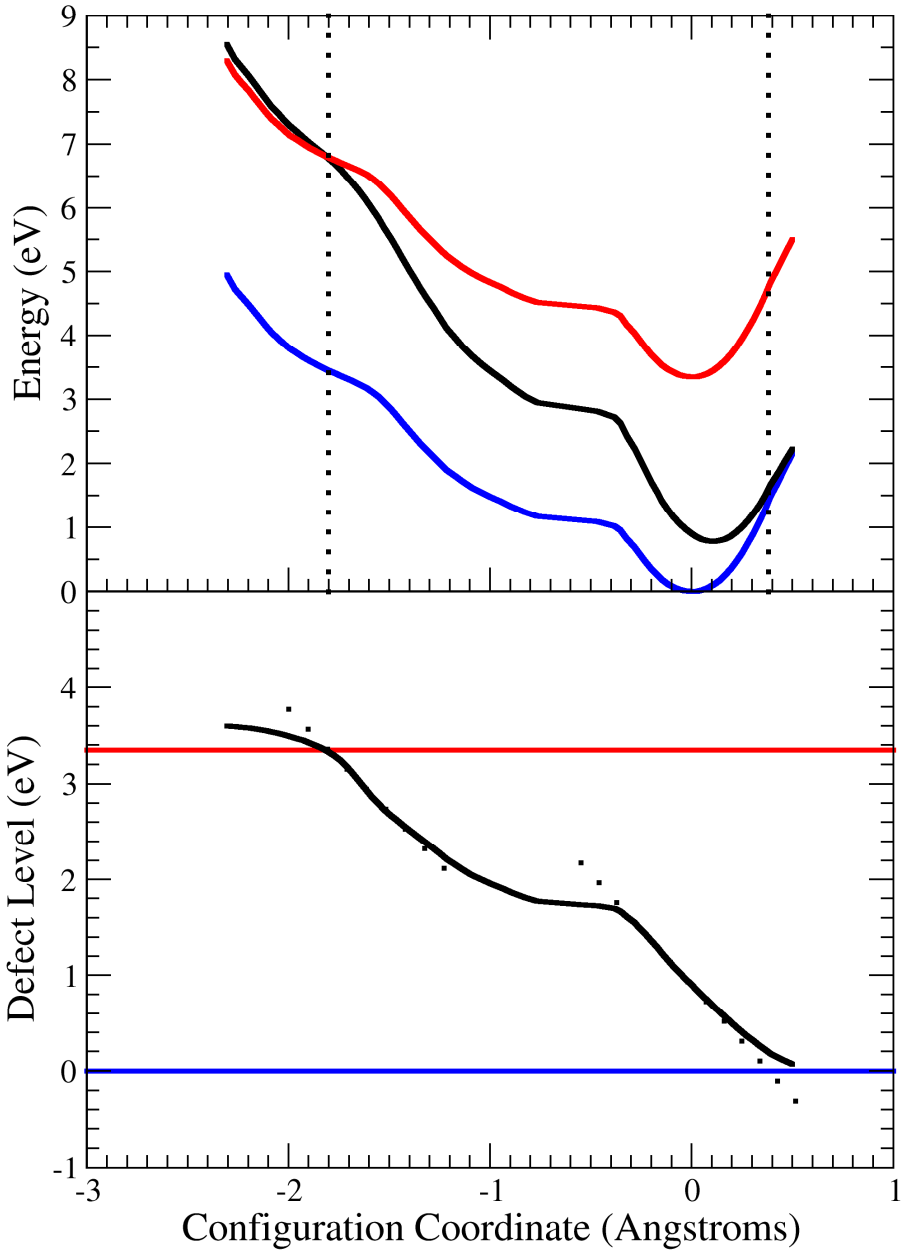
Energy (eV)
Defect Level (eV)
Configuration Coordinate (Angstroms)
9
8
7
6
5
4
3
2
1
0
-1
-3
-2
-1
0
1

Figure 4: