

Electronic structure of intrinsic defects in *c*-gallium nitride: Density functional theory study without the jellium approximation

Arthur H. Edwards ¹, Peter A. Schultz ², and Richard M. Dobzynski³

¹*Air Force Research Laboratory, Space Vehicles Directorate, 3550 Aberdeen Ave. SE, Kirtland Air Force Base, New Mexico 87117-5776, USA*

²*Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*

³*RMD Materials Science, LLC, 125 Currituck Sound Dr., Currituck, North Carolina 27929, USA*

 (Received 24 February 2021; revised 15 April 2022; accepted 25 April 2022; published 10 June 2022)

We report the first *nonjellium*, systematic, density functional theory (DFT) study of intrinsic and extrinsic defects and defect levels in zinc-blende (cubic) gallium nitride. We use the local moment counter charge (LMCC) method, the standard Perdew-Becke-Ernzerhoff (PBE) exchange-correlation potential, and two pseudopotentials, where the Ga $3d$ orbitals are either in the core (d^0) or explicitly in the valence set (d^{10}). We studied 64, 216, 512, and 1000 atom supercells, and demonstrated convergence to the infinite limit, crucial for delineating deep from shallow states near band edges, and for demonstrating the elimination of finite cell-size errors. Contrary to common claims, we find that exact exchange is *not* required to obtain defect levels across the experimental band gap. As was true in silicon, silicon carbide, and gallium arsenide, the extremal LMCC defect levels of the aggregate of defects yield an effective LMCC defect band gap that is within 10% of the experimental gap (3.3 eV) for both pseudopotentials. We demonstrate that the gallium vacancy is more complicated than previously reported. There is dramatic metastability—a nearest-neighbor nitrogen atom shifts into the gallium site, forming an antisite, nitrogen vacancy pair, which is more stable than the simple vacancy for positive charge states. Our assessment of the d^0 and d^{10} pseudopotentials yields minimal differences in defect structures and defect levels. The better agreement of the d^0 lattice constant with experiment suggests that the more computationally economical d^0 pseudopotentials are sufficient to achieve the fidelity possible within the physical accuracy of DFT, and thereby enable calculations in larger supercells necessary to demonstrate convergence with respect to finite size supercell errors.

DOI: [10.1103/PhysRevB.105.235110](https://doi.org/10.1103/PhysRevB.105.235110)

I. INTRODUCTION

Gallium nitride is a crucial component for several electronic and optical applications. It forms the active layer of AlGaIn/GaN high electron mobility transistors used for RF communication and for power electronics. It is also used for blue LED's and, hence, is crucial to the recent diode lighting advances. There is an emerging, significant interest in the metastable, cubic phase of gallium nitride because of the absence of spontaneous polarization fields, and because the band gap, 3.3 eV, compared to 3.45 eV for the hexagonal wurtzite structure, is amenable to green light emission [1–3]. Finally, cubic gallium nitride is a bright single photon emitter [4–8]. While there are many experimental and theoretical studies of wurtzite gallium nitride, including important recent surveys [9–12], there are few experimental studies on cubic gallium nitride [13,14], and no systematic computational studies of defects in cubic gallium nitride for the last fifteen years. The earlier density functional theory (DFT) studies that included defects in the cubic phase [15–18] were on very small supercells, and were subject to the persistent band gap problem, wherein the standard local density (LDA) and generalized gradient (GGA) approximations predict one-electron band gaps that are roughly 40% smaller than experiment. Also, these studies paid little or no attention to the fundamental electro-

static problems within the *jellium* approximation, where the net charge on a defect, which would lead to divergent total energy in a supercell model, is compensated by a uniform background charge density. As a result of these deficiencies their predictions for defect levels are now considered unreliable. With increased technological interest, and with recent developments in theory—methods that circumvent the Kohn-Sham (K-S) band gap problem, and that impose rigorous and accurate control of boundary conditions [19,20]—a new focus specifically on the defects in cubic (zinc-blende) gallium nitride is merited.¹

More recent defect calculations in gallium nitride have focused on the wurtzite structure using hybrid exchange-DFT (HE-DFT) methods, where the functional includes some form of exact exchange [21–23] to deliberately fit the K-S band gap to experiment, and where more intentional approaches for treating electrostatic boundary conditions [19,20,24–28] have been exploited. Over the past decade, these methods have become commonplace, driven by the widely held belief that HE-DFT is required to obtain the full spectrum of defects across the band gap. However, use of HE-DFT typically

¹We should note the recent rediscovery of some of the principles articulated in Refs. [19,20,71].

TABLE I. Lattice constant, K-S band gap, elastic constants, and the bulk modulus B for zinc-blende gallium nitride. LO=local orbital, PP=pseudopotential, PW=plane wave, MB=mixed basis, FLAPW=fully linearized augmented plane wave, FP-LMTO=full potential linearized muffin tin orbital, d^{10} = Ga $3d$ orbitals explicitly included, d^0 = Ga $3d$ orbitals subsumed into the core, GGA-generalized gradient approximation. Numbers inside (outside) parentheses for B are from a Birch Murnaghan fit $[(C_{11} + 2C_{12})/3]$.

	Approximation	a (Å)	K-S E_{gap} (eV)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	B (GPa)
Current Work	PBE d^{10} LO	4.588	1.50	242	122	127	162 (163)
	PBE d^0 LO	4.482	2.40	261	124	151	170 (167)
Ref. [42]	GGA d^{10} PP PW	4.55	1.69	255	133	177	173
Ref. [43]	GGA FLAPW	4.55	1.51				
Ref. [44]	GGA FLAPW	4.55	1.53				
Ref. [45]	GGA d^{10} PP PW	4.59	1.28				
Current Work	LDA d^{10} LO	4.496	1.99	274	145	150	188 (192)
	LDA d^0 LO	4.390	2.99	291	146	169	194 (193)
Ref. [46]	LDA FLAPW	4.41	1.521	274	166	199	212 (202)
Ref. [47]	LDA FP-LMTO	4.46	2.0	296	154	206	201
Ref. [48]	LDA d^0 PP PW	4.54		285	161	149	202
Expt.		4.52 [49]	3.3 [50]				

comes at a cost of using smaller supercells, as exact exchange is computationally expensive. Moreover, a more systematic assessment of the relative performance of the HE-DFT and DFT + U methods for transition metal defects in gallium nitride suggests that modification of the functional to obtain a better band gap does not lead to an accurate description of the electronic structure of defects [29,30]. Finally, all of these calculations use a *jellium* approximation.

We use the local moment counter charge (LMCC) method, where a localized charge distribution compensates a charged defect state, to solve the Poisson equation with the correct boundary conditions in a supercell model [19,20]. This provides a more rigorous treatment of charged defects and enables accurate calculations of defect energy levels—defect levels obtained from total energy differences. In the LMCC method, the K-S band gap is used solely to determine if a given defect eigenstate is within the band gap, and, thus, that the charge state is localized. An effective defect band gap is determined from the extremal values of computed defect levels derived from transitions between localized charge states in a large ensemble of defect calculations. In Si, Schultz found an effective defect gap that spanned the full experimental gap (1.2 eV), despite a K-S gap of only 0.6 eV, using semi-local functionals [31,32]. For GaAs, four different chemical contexts (two functionals, with $3d$ -core and $3d$ -valence Ga pseudopotentials) yielded K-S band gaps ranging between 0.13 and 0.84 eV [33], while the defect level results in Fig. 4 of Ref. [33] revealed minimal impact of the K-S band gap upon the computed defect levels.

A significant computational savings could be gained if the Ga semicore d electrons were subsumed into the pseudopotential rather than treated as valence electrons. Although gallium pseudopotentials have been systematically studied relative to bulk properties in Ga-group V semiconductors [34], their effect on the fidelity of related defects has not been assessed in gallium nitride. In this paper, we perform a systematic study of defects in cubic gallium nitride with both d -core (“ d^0 ”) and d -valence (“ d^{10} ”) pseudopotentials to inform future simulations in more complex systems of the benefits and shortcomings of each pseudopotential.

The balance of the paper is organized as follows. In Sec. II, we discuss details of the calculational technique and of our results on bulk gallium nitride. We present our estimate of the band gap based on the ensemble of defect calculations in Sec. III A, and detailed discussions of individual defects in Sec. III B, where we compare our results to the literature. We conclude in Sec. IV. We include Tables I–VI, and Figs. 1–4 in Ref. [35].

II. METHOD

We performed the calculations reported here using SE-QUEST [36], a local orbital, DFT code. The basis functions are double- ζ quality plus polarization functions. For vacancies, we add a set of floating Gaussian functions to eliminate the basis undercompleteness problem. We used the Perdew-Becke-Ernzerhoff (PBE) version of the generalized gradient approximation [37], and Troullier-Martins pseudopotentials [38]. Specifically, we have used a pseudopotential where the gallium $3d$ orbitals are treated explicitly (d^{10}), and a pseudopotential where they are subsumed into the core orbitals (d^0). The construction of these gallium pseudopotentials is discussed at length in Ref. [34]. We also used a spin-polarized version of PBE. Except where noted, all charge states of the defects we studied followed Hund’s rule; high-spin states were preferred.

As mentioned in Sec. I, for charged defects, we used the LMCC method [19,20], wherein the net charge on the defect is compensated by a localized charge density, rather than a constant charge density (*jellium*) over the entire supercell. This enables the use of correct boundary conditions for an isolated, charged defect. For all charged defects, we use the Jost approximation [39], shown in Eq. (1), to approximate the long-range polarization energy of the bulk dielectric outside of the supercell volume,

$$E_{\text{pol}} = -\frac{Q^2}{2R_{\text{Jost}}} \left[1 - \frac{1}{\epsilon_0} \right], \quad (1)$$

where Q is the net charge on the defect, R_{Jost} is the radius of the sphere representing the supercell volume, and ϵ_0 is

TABLE II. Summary of approximations in previous results to which we refer below. (a) Methods used for vacancies, antisites, and the nitrogen interstitial. (b) Methods used for the divacancy. PP=pseudopotential.

Ref.	VdW-2004 [16]	GR-2013 [12]	Diallo-2016 [11]	Lyons-2017 [10]	Gao-2019 [9]
PP/PAW	PP: Ga-3d ⁰	PP: Ga-3d ¹⁰	PAW: Ga-3d ¹⁰	PAW: Ga-3d ⁰	PAW Ga-3d (?)
Cell size (atoms)	96	96	128	96	128
Finite cell correction	none	Refs. [26,55]	Ref. [25]	Ref. [26]	Ref. [26]
Functional	LDA-SP	sX-LDA [21,22]	HSE06-SP	HSE06-SP	HSE06-SP
<i>k</i> mesh	2 × 2 × 2	2 × 2 × 2	2 × 2 × 2	(unknown)	4 × 4 × 2
	Ref.	Gohda-2010 [56]	Ganch-2006 [15]		
	PP/PAW	PAW: Ga-3d ¹⁰	PAW: Ga-3d(?)		
Cell size (atoms)		96	300		
Finite cell correction		none	none		
Functional		PBE	LDA		
<i>k</i> -mesh		4 × 4 × 4	2 × 2 × 2		

the static dielectric constant for the bulk crystal. In these calculations we use the experimental value for ϵ_0 . For a thorough discussion of this approximation, see Ref. [40]. Finally, the potential of a charged defect and of the perfect crystal are asymptotically aligned infinitely far from the defect. In this way, all defect total energies are referenced to a single, unknown chemical potential defined by this alignment.

We determine which defect charge states are localized, and hence states in the gap, from the computed one-electron K-S eigenstates. We calculate the projected density of states (PDOS) on various atoms of interest using a generalized Mulliken population analysis, and by using a Gaussian broadening function around each K-S eigenenergy. We use the position of the one-electron K-S state within the K-S band gap as a qualitative guide to localization. If the highest occupied level of a charge state is within the K-S gap, as evidenced in the PDOS, it is a candidate for localization. The amplitude of these peaks is a direct measurement of localization. If the highest occupied PDOS peak is near a band edge, we probe its behavior as a function of supercell size [41]. The localization of a truly deep level should be either constant or increase as a function of cell size. In this study, we used 64, 216, 512, and 1000 atom cells to verify that we eliminate the finite cell artifacts, including hybridization of defect wave functions with either valence or conduction band states (which would be clear in the PDOS), by studying the convergence of both one-electron properties and defect levels arising from total energy calculations.

As discussed in Sec. III B, we predict localized charge states that were unreported in previous wurtzite calculations. In those cases, we used a 300 atom wurtzite supercell to determine whether these new charge states persist in that phase.

Having determined the truly localized charge states for each defect, we calculate the defect levels directly as ionization potentials between these with respect to the single chemical potential, rather than indirectly from energies of formation. While this chemical potential is unknown, the complete defect level spectrum is known relative to this one unknown, so that the defect level spectra of the different defects are correctly aligned with each other. At this point in the method, we have all of the defect levels that arise from localized states for different defects in correct registry with one another on an energy scale determined by the chemical potential. We still need to locate the valence and conduction band edges relative to these defect spectra. We determine the upper and lower bounds of the effective band gap, the implicit valence and conduction band edges, from the total breadth in energy of all defect levels that arise from localized states. The smallest ionization potential bounds the CBE from below, and the largest ionization potential bounds the VBE from above.

To validate our local orbital approximation, we calculated band structures, equilibrium geometries, formation energies, and elastic properties for a two atom primitive cell of zincblende gallium nitride. In Table II, we give our results for lattice constant, (K-S) band gap, and selected elastic constants, and compare with previous calculations for both LDA and GGA results. For all of the current work reported in Table II, we used an $8 \times 8 \times 8$ Monkhorst-Pack *k* grid [51] and a $48 \times 48 \times 48$ *r* grid. These are well-converged results with respect to both *r*- and *k*-grid spacings. Both FLAPW and LMTO are all-electron calculations that use mixed basis sets. Even though all of our defect calculations used the PBE approximation, we include bulk LDA properties because there are very few published PBE calculations of zinc-blende gallium nitride, and only one reporting elastic constants [42]. It

TABLE III. The predicted spatial symmetry type for each of the lowest energy (high-spin state) configuration of each stable charge state for V_{Ga} .

Charge State	+1	0	-1	-2	-3
Slater determinant	$ a_1 \uparrow x \uparrow y \uparrow z \uparrow \rangle$	$ x \uparrow y \uparrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow y \downarrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow y \downarrow z \uparrow z \downarrow \rangle$
w-f symmetry	A ₂	A ₂	T ₂	T ₂	A ₁
defect symmetry	T _d	T _d	T _d	T _d	T _d

TABLE IV. The predicted spatial symmetry type for each of the lowest energy (high-spin state) configurations of each stable charge state for V_N , and the calculated defect symmetry. Symmetries in parentheses are for d^0 .

Charge State	+3	+2	+1	0	-1
Slater Determinant	$ >$	$ a_1 \uparrow >$	$ a_1 \uparrow a_1 \downarrow >$	$ a_1 \uparrow a_1 \downarrow x \uparrow >$	$ x \uparrow y \uparrow >$
w-f symmetry	A_1	A_1	A_1	T_2	T_1
defect symmetry	T_d	T_d	T_d	T_d (pD_{2d})	$pD_{2d}?$ (pD_{2d})

is clear that our local orbital technique, either LDA or PBE, agrees satisfactorily with other plane wave, pseudopotential calculations, as well as with all-electron calculations, at least as well as they agree with one another. Including gallium $3d$ orbitals in the basis leads to consistently larger lattice constants. Our d^{10} lattice constant matches Ref. [45] and is within 1% of the value in Ref. [42] a plane wave pseudopotential calculation and is within 1% of FLAPW-GGA calculations, where $3d$ orbitals are explicitly part of the set of valence orbitals. Also, our calculated band gap is in excellent agreement with Refs. [43,44], and within 12% of Ref. [42], although in worse agreement (greater than 15% difference) with Ref. [45]. Our PBE d^{10} elastic constants are between 5% and 30% different from the most similar calculation in Ref. [42]. The largest difference is in the shear constant C_{44} . Aside from the comparison with previous calculations, our d^0 calculations give a reduced theoretical lattice constant (2.3% smaller), and an increased band gap (60% larger) relative to the d^{10} results. These two quantities are directly related. When we calculate the d^{10} band gap at the d^0 lattice parameter we obtain 1.98 eV.

III. RESULTS

A. The defect band gap

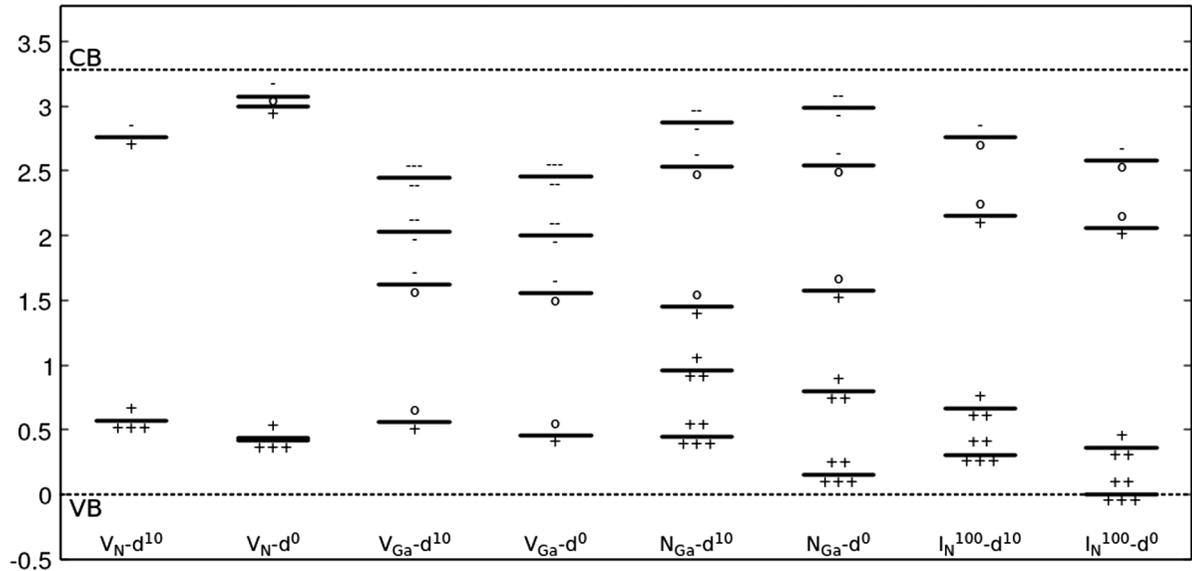
Before we discuss individual defects, we show in Fig. 1 the level spectra for all defects we studied to illustrate visually our estimate of the defect band gap. We have interlaced the d^{10} and d^0 results, facilitating defect-by-defect comparison. As discussed in Sec. II, the VBE is bounded from above by the largest (deepest) ionization energy, which is the $+2/+3$ transition of the split-100 nitrogen interstitial, $I_N^{<100>}$. The conduction band edge is bounded from below by the smallest (highest) computed defect ionization energy, which for this collection is the $-2/-1$ level of the zinc substituting on the nitrogen site, Zn_N . These bounds define the extent of the predicted effective defect band gap. These band edges, and the inferred defect band gap, are bounds, not direct predictions of the VBE and CBE. A survey of additional defects could discover levels that extend these bounds outwards. The calculations for this selection of defects represents a lower bound on the effective defect gap. The CBE marked in Figs. 1(a) and 1(b), 3.30 eV above the valence band edge, indicates the experimental gap [50].

For cubic gallium nitride, the two pseudopotentials predict different estimates for the effective defect band gap. The d^0 prediction is 3.58 eV for the 216 atom cell (3.47 eV for the 1000 atom cell), while the d^{10} prediction is 2.91 eV, assuming that the band edges are at the extremal defect levels. If we assume a margin of 0.05 eV between the last defect level and the actual band edge, to account for effective mass-like shallow

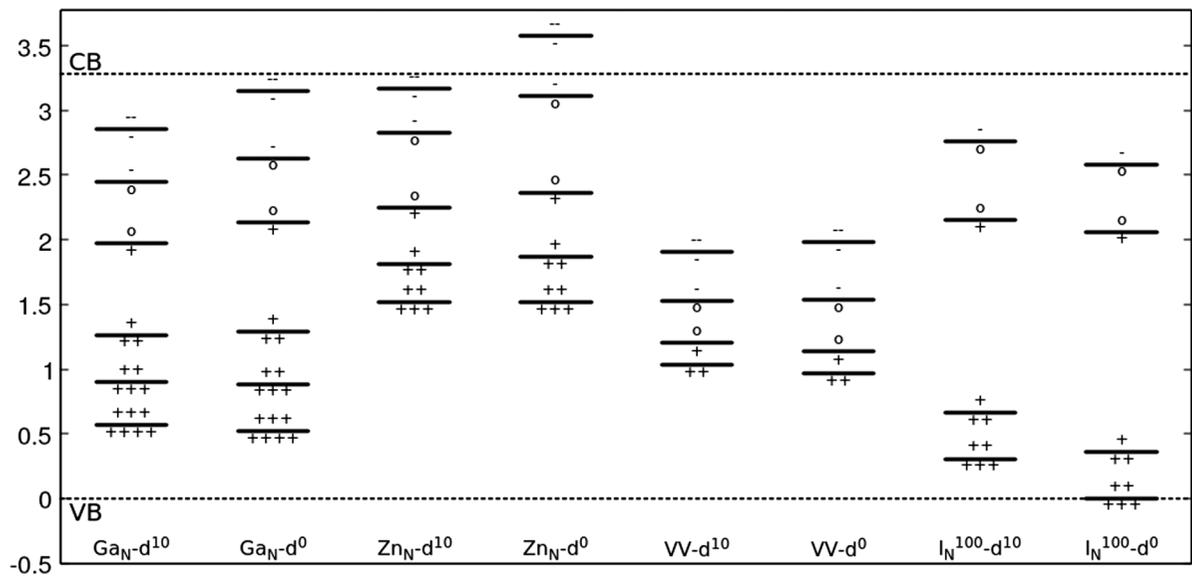
levels, then the band gaps are 3.57 and 3.01 eV—each 0.3 eV, or 9%, away from the experimental value. While the shapes of the individual defect spectra are generally the same for both pseudopotentials, the d^0 spectra are uniformly, though not dramatically, wider. Considering that the band gap is not fitted to the experimental result, either by adjusting the fraction of exact exchange in the exchange-correlation potential, or by using some variant of a “scissors operator,” these results lend further support to the prior claim by Schultz and others that the LMCC method gives an excellent overall representation of the band gaps of a variety of materials, now including silicon [31], gallium arsenide [33,52], cesium iodide [53], silicon carbide [54], and gallium nitride, without the need for exact exchange.

In Fig. 2, we show the degree of convergence of our defect levels with respect to supercell size for the d^0 pseudopotential. Results are qualitatively similar for the d^{10} calculations. A $3 \times 3 \times 3 k$ grid, shifted away from the Γ point, was required for the 64 atom cells. For all others, a $2 \times 2 \times 2 k$ grid, also shifted away from the Γ point, sufficed. In all cases, we used the Monkhorst-Pack method for k -grid generation [51]. While there are clear changes between the 64 atom and the 216 atom supercells, beyond that, with the exception of the $-2/-1$ level of the zinc substitutional, the changes are minor. We note that, even at 1000 atoms, the $-2/-1$ defect level in Fig. 2(c) is not fully converged. By itself, level convergence does not determine whether this is a gap defect level. If it were an effective mass state, the level would move closer to the band edge, as the wave function is allowed to expand with cell size. For this defect, the localization obtained from PDOS is essentially constant in the highest occupied, one-electron defect state in the gap. In fact, the sensitivity to cell size, combined with the stable localization, makes this a good lower bound for the conduction band edge. Note also that in Fig. 2(b), for the $<100>$ split nitrogen interstitial, the $2+/3+$ level is very stable, even though it is extremely close to our predicted valence band edge. This state is accidentally shallow—it arises from two strongly localized charge states, each insensitive to the proximity of the supercell boundaries. Taken together, the spectra of these three defects demonstrate the insensitivity of the electrostatics to the cell size.

In Fig. 3, we show the energies of formation for the neutral charge states of the vacancy and antisite defects as a function of supercell size. While the d^0 formation energies are consistently larger than the d^{10} , the more important feature is that both of these sets of results show that the rate of convergence with respect to cell-size is mixed. For some defects there is significant variation between the 64 atom and 216 atom supercells, with little difference between the 216 atom and the 1000 atom supercells. This result is important because all of the most recent studies in wurtzite, Refs. [9–12], use 96 or 128 atom cells, and claim that their results are well converged. The



(a)



(b)

FIG. 1. Defect spectra for intrinsic and extrinsic defects in zinc-blende gallium nitride for the d^{10} and d^0 pseudopotentials. VB and CB are separated by the experimental band gap [50]. All levels here are for the 216 atom supercell.

neutral results are independent of the approximations used to account for finite cell size electrostatics, so that the variations point to either or both of defect banding in small cells, and poorly converged relaxations.

Based on the results shown in Figs. 2 and 3 (and on the rest of our results), with few exceptions, our calculated defect levels are converged for the 216-atom cell. Unless otherwise stated, the results in the following sections on individual defects are from the 216-atom cell.

B. Individual defects

We focus on the on-site defects, vacancies, including the divacancy, antisites, and substitutional defects, leaving a de-

tailed discussion of interstitials to a future publication. Unless noted otherwise, descriptions of the electronic structure of individual charge states apply qualitatively to both the d^{10} and the d^0 calculations. We compare our predictions to several previous sets of results for defects in the wurtzite structure. We compare to Van de Walle and Neugebauer [16], as this was a careful, general review of defect physics, and included a discussion of the relative merits of d^0 and d^{10} pseudopotentials for bulk properties. Those results stand as a benchmark of LDA without the inclusion of finite size effects. We also compare to other recent HE-DFT results that do include electrostatic, finite cell size corrections, albeit based on the *jellium* approximation. In Table II, we summarize the

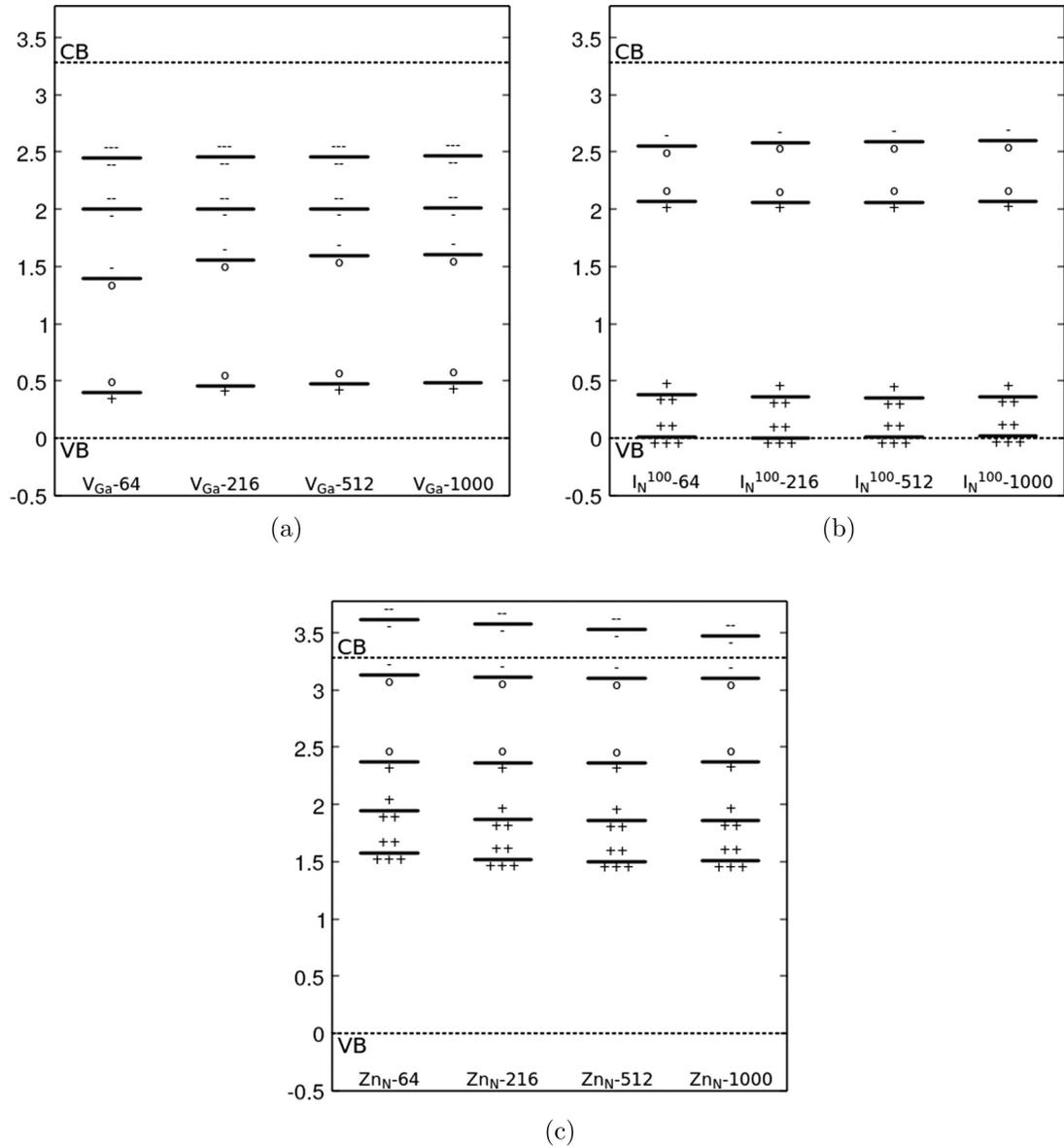


FIG. 2. Defect levels for the (a) the gallium vacancy, (b) the nitrogen split interstitial, and (c) zinc substitution on the nitrogen lattice for the d^0 calculations as a function of supercell size.

approximations used in these calculations. Gillen and Robertson [12] used an LDA screened exchange method, while Lyons and Van de Walle [10], Diallo and Demchenko [11], and Gao and coworkers [9] used the Heyd-Scuseria-Ernzerhoff (HSE-06) [23] exchange correlation potential modifying the PBE method. All of the referenced calculations are for the wurtzite structure. However, we expect the results to be similar to zinc-blende because the local chemistry is very similar. Neugebauer and Van de Walle also claimed that their unpublished zinc-blende results were similar to wurtzite [16].

In our discussion of all approximately tetrahedral defects, we will appeal to a molecular orbital treatment similar to Schultz [41], and inspired by Watkins [57]. The orbitals are shown in Fig. 4 for a typical antisite. For vacancies, the absence of the central atom makes little qualitative difference—the bond orbitals are replaced by sp^3 orbitals facing the vacancy. The ideal, unrelaxed antisites, single

vacancies, and substitutional defects have tetrahedral (T_d) symmetry. We will use the states in Fig. 4 to build up many-electron, Slater determinants, from which we can derive the symmetry of various spin states, and predict the charge states of each defect that are unstable to Jahn-Teller distortions.² A concrete example is given in Sec. III B 1. In Fig. 5, we show the principal distortions we applied. In most cases, these were sufficient to capture the equilibria obtained from unconstrained relaxations starting from less symmetric initial

²Even though in standard DFT, there is no need for Slater determinants, as there are only one-electron operators, we know that the true wave function has to be antisymmetric with respect to particle interchange. Furthermore, the use of Slater determinants leads to the correct assignment of symmetry for the many-electron states.

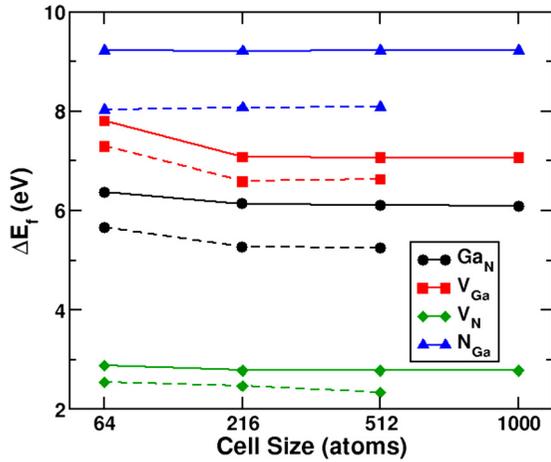


FIG. 3. Formation energies for neutral vacancy and antisite defects as a function of supercell size. Solid lines are d^0 , dashed lines are d^{10} . All values are for gallium-rich conditions.

geometries. They also give a systematic process for finding other metastable, symmetry-constrained states.

1. Gallium vacancy

The gallium vacancy was initially thought to be the simplest of the intrinsic defects. However, aside from the simple vacancy, there is a site-shift conformation, discussed in Sec. III B 2. For the simple vacancy, we predict five charge states in the gap, +1 to -3 [see Fig. 1(a)]. To discuss individual charge states, we refer to Fig. 6, where we show schematically the filling scheme for the a_1 and t_2 states. In (a), we show the standard occupation in defect-free gallium nitride of 8 electrons (5 from nitrogen and 3 from gallium). Removing a gallium atom removes 3 electrons, leaving two possibilities: a high-spin quartet, and a low-spin doublet, represented in Figs. 6(b) and 6(c), respectively. Ignoring the purely symmetric, doubly occupied a_1 manifold, the Slater determinant for the high-spin, neutral charge state is $|x \uparrow y \uparrow z \uparrow \rangle$. The purely antisymmetric space part transforms as A_2 , a one-dimensional irreducible representation (irrep.), so

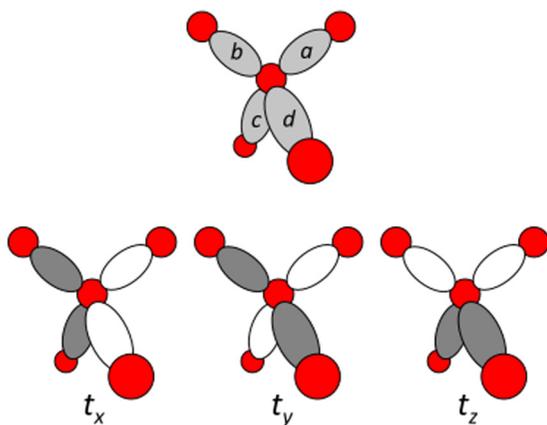


FIG. 4. Schematic of molecular orbitals used to describe defects. From Ref. [41], with permission.

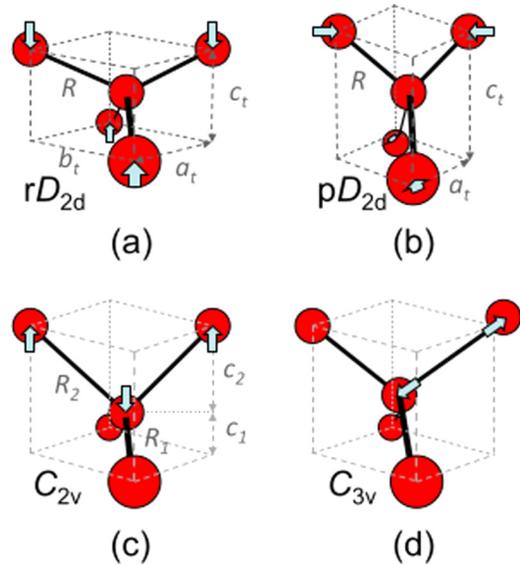


FIG. 5. Symmetric distortions applied to intrinsic defects. From Ref. [41], with permission. pD_{2d} and rD_{2d} refer to paired and resonant structures, where pairs of nearest-neighbor atoms either form bond pairs structures, or move apart.

this spin state has no Jahn-Teller instability, and we expect a purely tetrahedral equilibrium geometry. In Table III, we show the Slater determinant, the wave function irrep. and the optimized symmetry for the lowest energy spin state of each charge state. Even though the -1 and -2 charge states are predicted to be Jahn-Teller distorted, the geometric distortions are very small—changes in the fourth decimal place for distances, and in the third decimal place for angles. The most interesting result is that the ground state of the $+1$ charge state is localized, and that it is a spin quintet. No other spin state is localized.

In Fig. 7, we show our defect level spectra compared with selected previous calculations. Older studies in wurtzite gallium nitride (see, for example, Refs. [15,16,56,58]), as well as the more recent work of Gillen and Robertson [12],

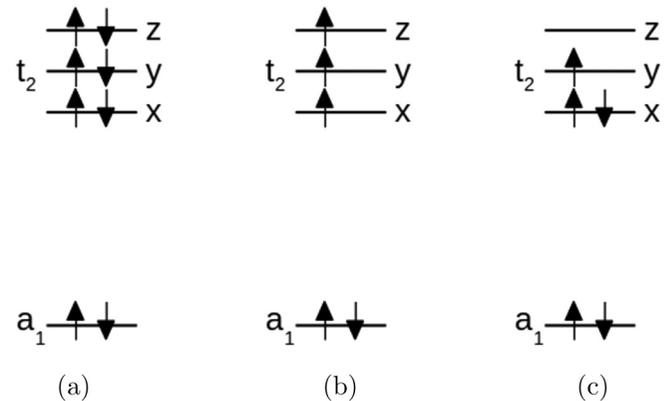


FIG. 6. Schematic diagram for T_d states arising from bond orbitals in Fig. 4. We use the standard labels for t_2 basis functions. (a) Normal bulk occupation. (b) Neutral V_{Ga} high spin. (c) Neutral V_{Ga} low spin.

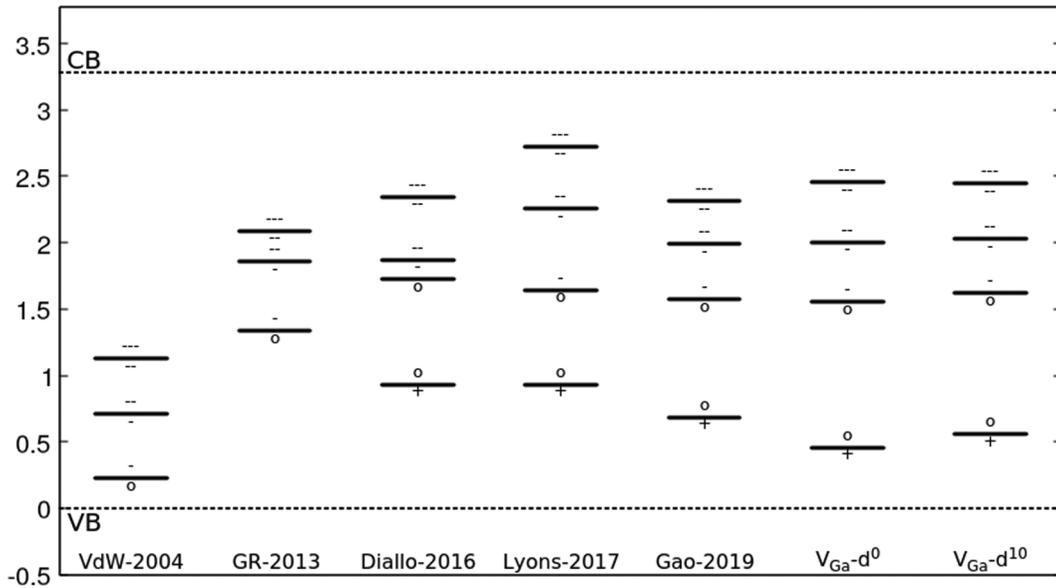


FIG. 7. Defect levels for gallium vacancy from previous calculations (wurtzite) and the current results (zinc-blende). VdW-2004 from Ref. [16], GR-2013 from Ref. [12], Diallo-2016 from Ref. [11], Lyons-2017 from Ref. [10], and Gao-2019 from Ref. [9].

predicted that the only charge states in the gap were 0, -1 , -2 , and -3 . However, in agreement with the current results, the recent HSE calculations of Diallo and Demchenko [11], Lyons and van de Walle [10], and Gao *et al.* [9] predicted that the $+1$ state is in the gap. That this was absent from the Gillen-Robertson result [12] may simply reflect a neglect of the possibility of a quintet. The current results, and the recent HSE calculations, place the $-/0$ defect level close to, or above, mid gap, compared to the prior prediction of a near-valence band state. In fact, all of the recent predictions for this level are within 0.3 eV of each other. Both current calculations, with and without explicit presence of the gallium $3d$ orbitals, predict that the $0/+$ defect level is ~ 0.4 eV closer to the VB edge than both the Diallo-Demchenko and the Lyons-Van de Walle results.

2. Site-shifted gallium vacancy

If one of the nearest-neighbor nitrogen atoms hops into the gallium vacancy site, it creates an intimate $V_N:N_{Ga}$ pair, which we label as a site-shift complex. This class of defects has been studied earlier in gallium arsenide [59–61] and silicon carbide [40]. Figures 8(a) and 8(b) show the canonical geometry for the positive charge states, and for the neutral and -1 charge states, respectively. Here, N^* denotes the antisite nitrogen atom. All the positive charge states have C_{3v} symmetry, while the neutral and -1 charge states undergo a strong distortion, wherein the antisite nitrogen atom becomes twofold-coordinated. We note that the twofold conformation for the $+1$ charge state is only favored by 2 meV for the d^{10} pseudopotential.

In Fig. 9, we show the calculated defect levels from the current calculations. The current calculations predict that this defect has five stable charge states in the gap, from $+3$ to -1 . While all of the charge states in Fig. 9 refer to local geometrical minima, several of these are metastable. The simple gallium vacancy is preferred in the -1 and neutral charge

states, while the site-shifted $V_N:N_{Ga}$ complex is favored for the $+1$ charge state, and is the only stable structure for the $+2$ and $+3$.

3. Nitrogen vacancy

Because of its low formation energy, as seen in Fig. 3, the nitrogen vacancy is believed to be the most abundant intrinsic defect. It was initially suggested to account for the intrinsic n -type character of the material. However, for positive charge states, the formation energy increases linearly with the Fermi level (see, for example, Eq. (1) and Fig. 3 in Ref. [10]), so that, for n -type gallium nitride, the formation energy would prohibit large densities. Hence, the observed intrinsic n -character must arise from accidental incorporation of extrinsic species.

As seen in Fig. 1, we predict five charge states in the gap, $+3$ to -1 . The d^{10} pseudopotential predicts a $-U_{\text{eff}}$ to exclude the $+2$ and the neutral charge states. The d^0 pseudopotential gives nearly the same qualitative physics. It predicts all charge states will be present, but that both the $+2$ and neutral charge states exist within very small intervals of the Fermi energy. We summarize the electronic structure of the individual charge states in the current calculations in Table IV. Several subtle features bear comment. First, we predict from group theory that all of the positive charge states retain T_d symmetry for both pseudopotentials. Second, we predict that both the neutral and -1 charge states have degenerate ground states, belonging to either a T_1 or T_2 manifold, so that they should be unstable to a Jahn-Teller distortion. However, this would only be true if the one electron t_2 states were all in the gap. The two pseudopotentials do not give the same results, as summarized in Table V. For all supercell sizes, the d^0 pseudopotential predicts that the neutral and -1 charge states are localized, and that the ground state is pD_{2d} . In the -1 charge state there are two pairs of bonded gallium atoms. In the neutral charge state, the gallium-gallium bonds are essentially broken ($R_{Ga-Ga} = 2.96$ and 3.17 Å), but the pD_{2d} symmetry

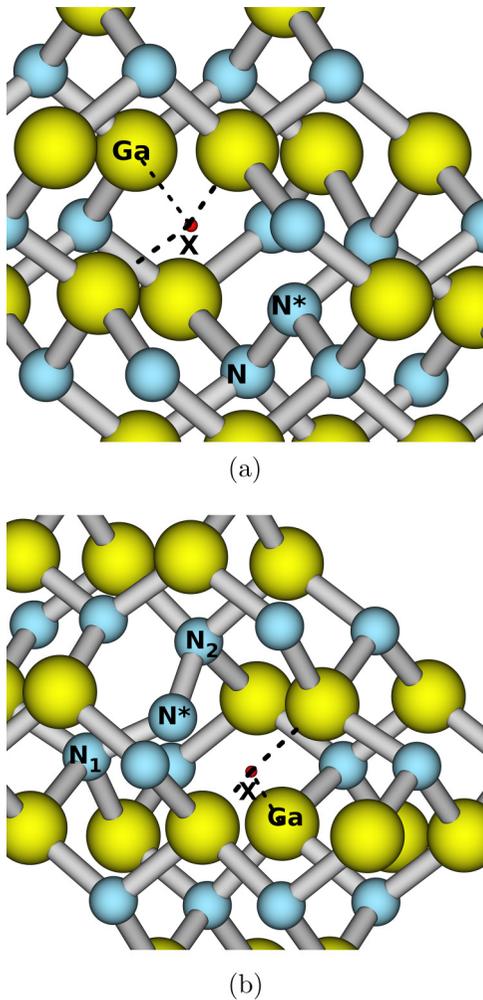


FIG. 8. Canonical geometries for all charge states of the $V_N:N_{Ga}$ site-shift complex. (a) is lowest energy for all positive charge states, while (b) is lowest for 0 and -1 charge states. The red sphere (X) indicates the tetrahedral center of the nitrogen vacancy. N^* indicates the antisite nitrogen atom.

persists. There is a metastable T_d geometry that is 0.49 eV (0.28 eV) higher in energy for the 216 atom (512 atom) supercell, but this is deceptive, because the t_2 states are above the conduction band, so that the highest energy electrons are actually at the bottom of the conduction band, and, hence, delocalized.

The predictions of the d^{10} pseudopotential are qualitatively different. For the 216 atom supercell, it also predicts an equilibrium pD_{2d} geometry, although the stability relative

TABLE V. Geometries and relative stabilities of the -1 charge state of the nitrogen vacancy for the d^0 and d^{10} pseudopotentials as a function of supercell size. Numbers outside (inside) parentheses are for 216 atom (512 atom) supercells.

	$\Delta E(pD_{2d}-T_d)$ (eV)	R_{Ga-Ga} (\AA)
d^0	-0.49 (-0.28)	2.68 (2.69)
d^{10}	-0.01 (-)	2.73 (-)

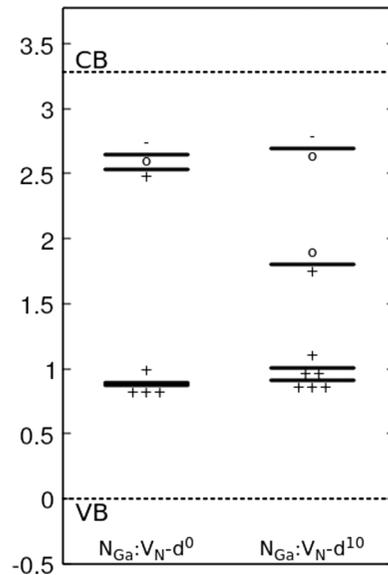


FIG. 9. Current results for defect levels for the $V_N:N_{Ga}$ site-shift complex.

to the T_d geometry is very small (0.01 eV). For the 512 atom supercell, there is no pD_{2d} geometry. Instead, the K-S states accommodating these additional electrons rise above the K-S conduction band edge, so a local charge state is not stabilized. The equilibrium geometry is approximately tetrahedral. Thus, for the larger cell, we do not predict a $-1/+1$ defect level in the gap. The d^{10} pseudopotential also does not predict any distortion in the neutral charge state, for the same reason; the t_2 states are above the K-S CB edge for both the 216 and the 512 atom supercells. The significantly larger d^0 K-S band gap accounts for this qualitative difference between the d^0 and the d^{10} level structures. As discussed in Sec II, we contend that the d^0 results are more reliable, because they reproduce more faithfully the bulk properties.

In Fig. 10, we compare our defect level spectra with results from the same representative studies we used for the gallium vacancy [9–12,16]. There is general qualitative agreement between all of the calculations in the lower half of the band gap—all predict a $+1/+3$ defect level near the valence band edge. All predict a $-U_{\text{eff}}$. The current calculations are in essential agreement. The d^{10} pseudopotential predicts a $-U_{\text{eff}}$, while the d^0 pseudopotential predicts a very small, positive U_{eff} (0.02 eV). All $+1/+3$ levels are within 0.2 eV of each other. The levels in the upper half of the bandgap are less consistent, even across recent, HE-DFT studies. Diallo and Demchenko [11] predict no localized states above the $+1$ charge state, while Lyons and Van de Walle [10] and Gao *et al.* [9] do not predict a localized -1 charge state. Gillen and Robertson [12], predict a $-U_{\text{eff}}$ between the $+1$ and -1 states, in qualitative agreement with the current d^0 result. These differences have important implications. Diallo and Demchenko predict that the nitrogen vacancy is a deep donor and hence could be an abundant source of excess electrons, as was originally proposed by Van de Walle and Neugebauer, while Lyons and Van de Walle, and Gao *et al.* do not. In agreement with Gillen and Robertson, we predict that the

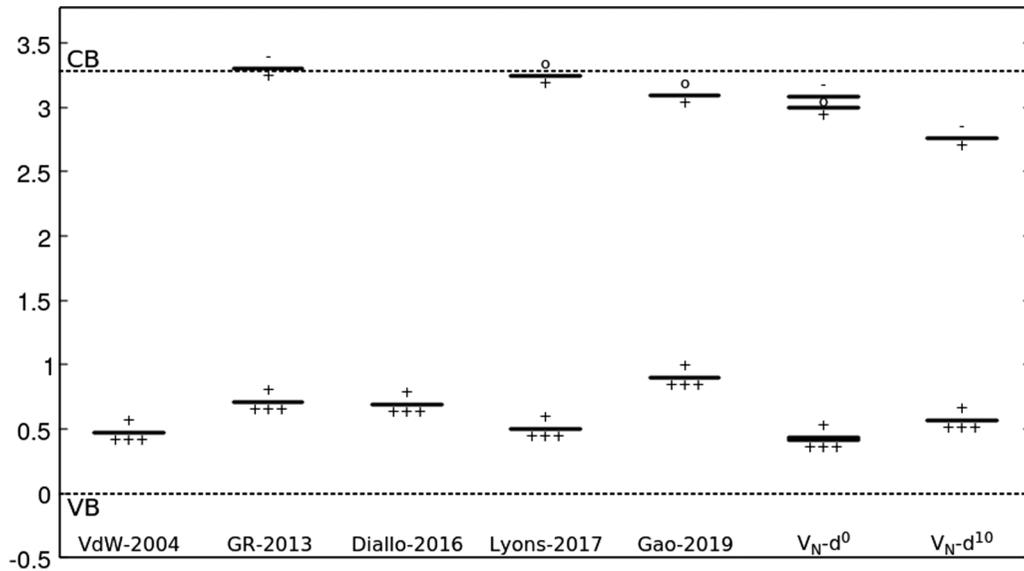


FIG. 10. Defect levels for the nitrogen vacancy. See Fig. 7 for references. All previous results were for wurtzite-gallium nitride.

nitrogen vacancy is an acceptor that traps conduction band electrons.

4. Gallium antisite

Because the gallium atom is so much larger than nitrogen, we expect steric limitations on distortions around the gallium antisite. Nevertheless, we find significant distortions away from the ideal tetrahedral conformation in several charge states. We predict seven charge states within the band gap, from +4 to -2 . In Table VI, we summarize the electronic structure and geometry of each stable charge state. The d^0 and d^{10} pseudopotentials predict the same qualitative geometries for all charge states with the exception of the +3 state, where the d^0 pseudopotential predicts virtual degeneracy between a short-bond (sb) C_{3v} and C_{2v} , while the d^{10} pseudopotential gives predicts that the C_{2v} symmetry is favored by 65 meV.

In Fig. 11, we compare with wurtzite results of Refs. [9–11,16]. All predict at least +3 through -1 levels in the gap, and all but Gao *et al.* [9] predict the +4 state in the gap. The positions of the 0/+ and $-/0$ levels are in good agreement across calculations. In agreement with Diallo and Demchenko [11], the current calculations predict the +3/+4 level to be significantly lower in the gap—within ~ 0.5 eV of the VB edge—than any other previous calculations. We also predict that the +2/+3 is also much lower than all the previous results. Finally, in agreement with Diallo and Demchenko [11], and with Gao *et al.* [9], we predict that the

$-2/-1$ level is in the gap. However, in our calculations, the -2 charge state does not persist in the gap in the wurtzite structure.

5. Nitrogen antisite

Because the nitrogen atom is so much smaller than the gallium, and because the nitrogen-nitrogen bond is so much smaller than the equilibrium nitrogen-gallium bond, the nitrogen antisite equilibrium geometries have very large distortions from the ideal tetrahedral coordination. There are several metastable states, and some of these are so close in total energy that in equilibrium we would expect significant fractional populations, implying that experimental determination could be difficult. We predict six charge states within the band gap, from +3 to -2 .

We start our discussion of the individual charge states with the -2 charge state, as this is only predicted to be localized in the current calculations. The geometry, shown in Fig. 12, is the qualitative archetype for all nonpositive charge states. The N_1-N^* distance varies in analogy with the diatomic nitrogen ion. As electrons are removed from the -2 charge state, $R_{N^*-N_2}$ decreases by 0.14 Å between the -2 and the neutral charge state. The added charge also strengthens the bonds between N_1 and its gallium neighbors.

The +3 charge state is predicted to be stable in all of the more recent calculations. However, the +3 equilibrium geometry, as shown in Fig. 13, is unreported elsewhere.

TABLE VI. Slater determinant, wave function symmetry in tetrahedral configuration, and the calculated defect symmetry for each charge state for the gallium antisite. For C_{3v} , sb (lb) indicates that the bond between the antisite atom and its neighbor along the symmetry direction is shorter (longer) than the other three.

Charge state	+4	+3	+2	+1	0	-1	-2
Slater determinant	$ a_1 \uparrow a_1 \downarrow \rangle$	$ x \uparrow \rangle$	$ x \uparrow y \uparrow \rangle$	$ x \uparrow y \uparrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow y \downarrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow y \downarrow z \uparrow z \downarrow \rangle$
w-f symmetry	A_1	T_2	T_1	A_2	T_1	T_2	A_1
defect symmetry	T_d	$C_{3v}(\text{sb})/C_{2v}$	$C_{3v}(\text{lb})$	T_d	pD_{2d}	rD_{2d}	T_d

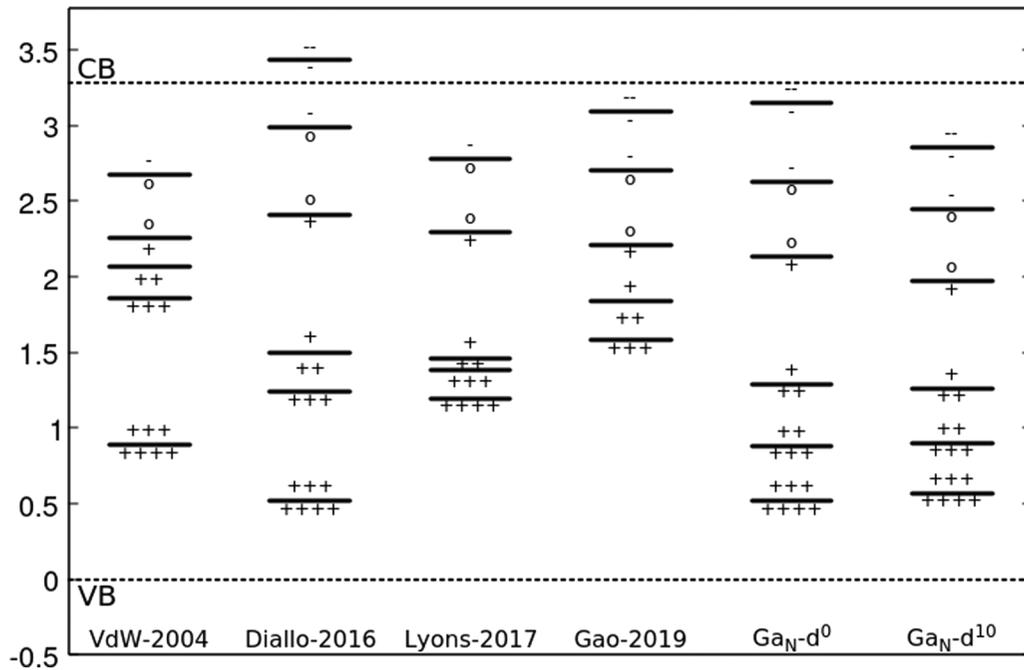


FIG. 11. Defect levels for the gallium antisite defect. See Fig. 7 for references. All previous results were for wurtzite-gallium nitride

Qualitatively, this is the equilibrium geometry for all three positive charge states. This could be described as a long-bond C_{3v} structure, but that would understate the extent of the distortion. Here, the antisite atom, N^* , is bonded strongly to three neighbors in sp^2 hybridization, with no bonding to the fourth neighbor. In the +2 and +1 charge states N_1 and N^* are more pyramidal. The $\angle N_1-N^*-N_2$ and $\angle N^*-N_1-Ga_2$ angle increase by 10° and 5° , respectively. In all cases, the distance between N_1 and N^* is large enough to preclude bonding.

All charge states display metastability. For the +3 charge state, there is a purely tetrahedral metastable state 0.54 eV higher in total energy than the geometry shown in Fig. 13. The +2 charge state has two metastable states. The lower energy has C_{2v} symmetry, and the higher energy is purely tetrahedral. Interestingly, both of these states are within 10 meV of the C_{3v} ground state, even though there is significant distortion, and

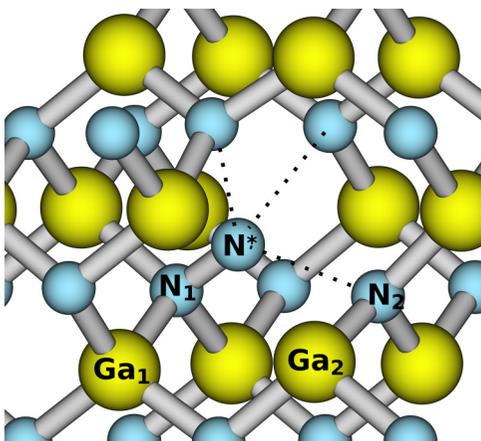


FIG. 12. Equilibrium geometry of the -2 charge state of the nitrogen antisite.

even rebonding. The +1 charge state also has two metastable states. The closest in energy is a C_{2v} structure, where the antisite atom bonds to only two of its neighbors. Above that, there is a short bond C_{3v} structure, similar to that in Fig. 12.

In Fig. 14, we compare our predicted defect levels with those in Refs. [9–11,16]. There are strong similarities and differences between these defect spectra. All three of the most recent calculations predict a $2+/3+$ level in the gap, which is absent from the results of Van de Walle and Neugebauer [16]. The current calculations capture an additional level ($2-/-$), only predicted by Gao *et al.* [9]. In our wurtzite calculation, we predict that a localized -2 charge state persists, so we expect a $-2/-1$ level there as well.

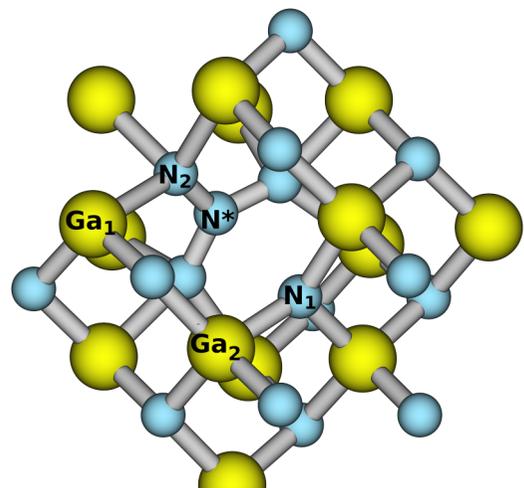


FIG. 13. Equilibrium geometry for the +3 charge state of the nitrogen antisite. N^* is the antisite nitrogen.

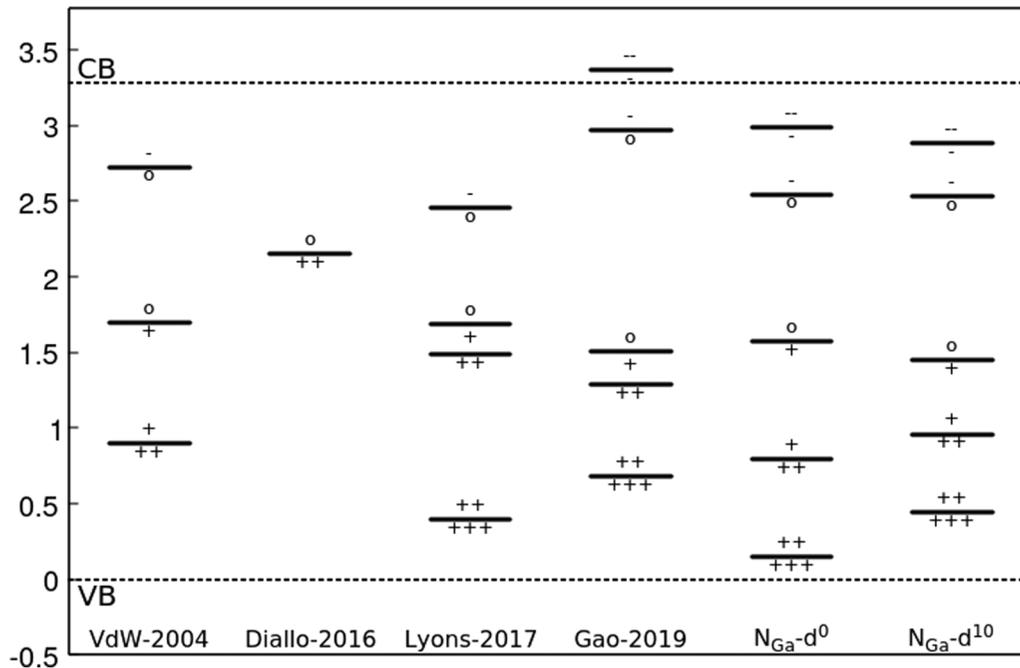


FIG. 14. Defect levels for the nitrogen antisite defect. See Fig. 7 for references. All previous results were for wurtzite-gallium nitride.

6. Divacancy

Positron annihilation experiments on irradiated wurtzite gallium nitride indicate that the principal nonradiative recombination center is the divacancy, VV [62]. We find that the neutral formation energy for VV, 6.35 eV (6.02 eV) for the d^0 (d^{10}) pseudopotential, is actually less than that of V_{Ga} in a gallium-rich environment. We predict five charge states in the gap, +2 to -2 . As noted by Gohda and Oshiyama [56], we find that there are rather small variations in geometries for the nonpositive charge states, while the positive charge states exhibit dramatic differences for the positions of the gallium atoms. For the nonpositive charge states the gallium atoms move toward the vacancy center by $\sim 5\%$, while in the positive charge states, they move away by $\sim 19\%$. The relaxations of the positive charge states are easily rationalized: the orbitals arising from nitrogen lone-pair orbitals are lower in energy, and hence have greater electron affinity, than the gallium dangling orbitals. In all of the negative charge states, at least one of the gallium orbitals will be occupied, leading to pyramidal local geometry. In the neutral singlet state, the gallium orbitals are completely unoccupied. They rehybridize toward sp^2 , and relax into the plane of the three nitrogen neighbors. However, in the neutral triplet ground state, one of the gallium orbitals retains occupation, frustrating the rehybridization. In all positive charge states, the gallium orbitals are empty so that the hybridization is, again, approximately sp^2 . Finally, while the two current calculations agree well for all four defect levels, there is a subtle difference in spin states. The d^{10} calculation +2 state is a doublet, while the d^0 prefers the quartet spin state.

In Fig. 15, we show the predicted defect levels for the divacancy for the current calculations compared with those from Refs. [11,15,56]. As shown in Table II, Gohda and Oshiyama [56] used the PBE variant of GGA, and PAW poten-

tials, with explicit inclusion of the $3d$ orbitals in the valence set, with no mention of the treatment of finite cell corrections. Diallo and Demchenko [11] used the HSE hybrid functional excluding the gallium $3d$ states from the valence set, and used two terms of the Makov and Payne correction for finite cell sizes. Ganchenkova and Nieminen [15] used LDA with no finite cell-size correction. The results from these studies differ significantly. While Diallo and Demchenko [11] and the current calculations place the top-most $2-/-1-$ defect level within ~ 0.07 eV of each other, Gohda and Oshiyama [56] predicted that level is absent. Rather, the $2-/-1+$ level exhibited $-U_{\text{eff}}$ behavior and the highest energy defect level

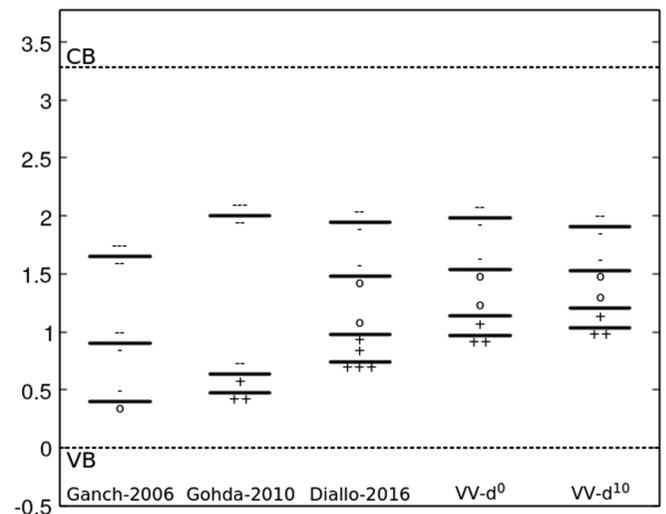


FIG. 15. Defect levels for the divacancy from Refs. [15] (Ganch-2006), [56] (Gohda-2010), [11] (Diallo-2016), and from the current calculations. All previous results were for wurtzite-gallium nitride.

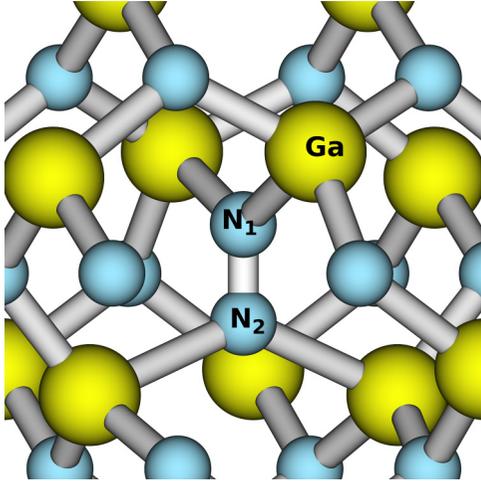


FIG. 16. Equilibrium geometry of the nitrogen interstitial in all non-negative charge states.

was a $3 - /2-$ transition. We find no evidence in our calculations for a localized -3 charge state, although Ganchenkova and Nieminen [15] also predicted a -3 charge state. While Gohda and Oshiyama [56] and the current calculations predict the lowest energy localized charge state is the $+2$, Diallo and Demchenko [11] predicted the $+3$ charge state is stable over a fairly wide range (~ 0.81 eV) of the Fermi energy. We find no evidence of a $+3$ charge state in our calculations, even in the 1000 atom supercell.

7. Nitrogen interstitial

We defer detailed discussion of the interstitials to a future publication. However, the nitrogen interstitial has the lowest energy defect level for either intrinsic or extrinsic defects, and, thus, is the basis for our assignment of the valence band edge. We give a brief discussion of our results here. We predict five charge states in the gap, $+3$ to -1 . In Fig. 16, we show the canonical geometry for the $\langle 100 \rangle$ split nitrogen interstitial. This is the equilibrium geometry for the $+3$, $+2$, $+1$, and neutral charge states. The N_1 - N_2 bond length ranges from 1.17 Å ($+3$) to 1.29 Å (neutral) for the d^{10} pseudopotential. The -1 charge state adopts a split- $(110)_N$ geometry, shown in Fig. 17. Here the N_1 - N_2 bond length is 1.54 Å. Not only is this the lowest energy geometry we could find, but it is also the only geometry with the Fermi level in the K-S band gap, indicating true localization. For this charge state, the PDOS reveals that the one-electron states in the gap have very large nitrogen content. Neither Ga_1 nor Ga_2 in Fig. 17 contribute 10% as strongly as N_1 and N_2 . Again, our results are qualitatively identical for d^0 and d^{10} .

In Fig. 18, we show the defect levels from the current calculations, and from prior studies [9–11,16]. Several features bear notice. With the exception of Lyons and Van de Walle [10], all calculations predict the lowest energy charge state in the gap is the $+3$ state, while all calculations predict that the -1 is the highest energy charge state. All of the most recent calculations, starting with Diallo and Demchenko [11], predict that the $-1/0$ state is within a volt of the conduction band edge, and that the $0/+1$ level is above mid-gap.

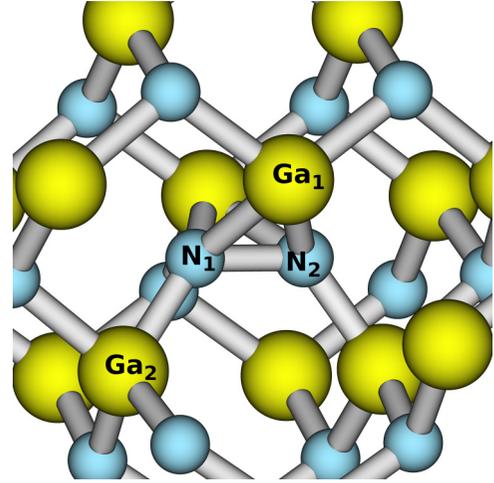


FIG. 17. Equilibrium geometry for -1 charge state of the nitrogen interstitial.

8. $Zn:N$

The last defect we consider is zinc substituting on a nitrogen site. As mentioned in Sec. III A, we included this defect because it gives an extremal value to bound the conduction band edge. While zinc is expected to substitute for gallium, Watkins and co-workers have hypothesized that this defect, minus one nearest-neighbor gallium atom, is the possible origin of a strong optically detected magnetic resonance signal [63]. We find, however, that the $Zn_N:V_{Ga}$ complex is unstable for all charge states. The Zn shifts to the Ga site to become a zinc decorated nitrogen vacancy—still with three equivalent gallium atoms. In Table VII, we give the symmetry and electronic structure for each stable charge state. The equilibrium geometries are mostly D_{2d} , either paired or resonant. In all cases, both pseudopotentials have the same equilibrium symmetry. Besides the subtle changes in defect symmetry, R_{Zn-Ga} decreases monotonically as electrons are added to the $+3$ charge state, illustrating the bonding nature of the t_2 orbitals. The total change in bond length from $+3$ to -2 is 0.24 (0.26) Å for the d^0 (d^{10}) pseudopotential.

The level structure for for the Zn_N defect for both d^0 and d^{10} is shown in Fig. 1 for the 216 atom supercell, and for d^0 as a function of supercell size in Fig. 2(c). While, ideally six electrons could be accommodated, the -3 state is not localized. Finally, we find that the -2 charge state is not localized in the wurtzite structure.

IV. DISCUSSION AND CONCLUSION

We have presented the first nonjellium calculations of defect structures and defect levels for mostly intrinsic defects in zinc-blende gallium nitride. As demonstrated in Sec. III A, the effective defect band gap, calculated from the extremal defect levels predicted by the LMCC method, is in good agreement with experiment without including exact exchange. This agreement is in line with previous LMCC results in silicon [31,33], gallium arsenide [33,52], and, most recently, in silicon carbide [40]. As in the previous studies, we have demonstrated that our results are converged, and that we

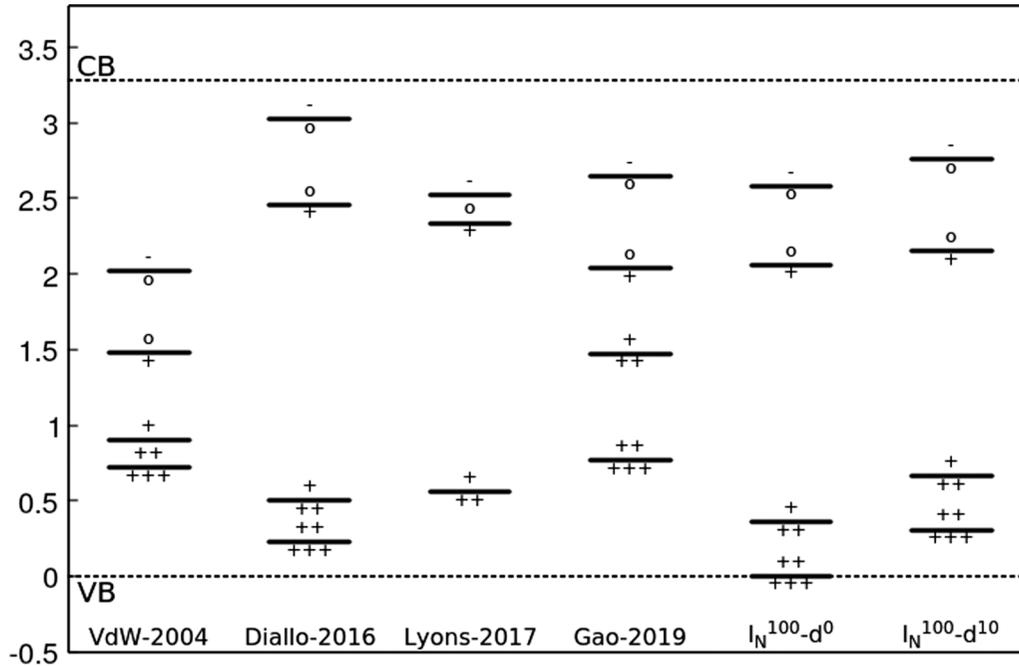


FIG. 18. Defect levels for the nitrogen interstitial defect. See Fig. 7 for references. All previous results were for wurtzite-gallium nitride.

have eliminated electrostatic, finite-cell effects by studying the same defects in 64-, 216-, 512-, and 1000-atom supercells. This convergence study has enabled identification of defect levels arising from well-localized states very near the valence and conduction band edges. Several specific topics warrant discussion.

A. Comparison with wurtzite results

In Sec. III, we compared prior results on defects in wurtzite GaN to the current results for the zinc-blende phase. Across Figs. 7, 10, 11 and 14, we remarked on two salient features. First, the defect levels from prior work, including the HE-DFT results, are rarely in agreement with each other. There are significant differences in positions within the band gap for the same defect levels, and there are disagreements about the total number of localized charge states. This is especially noteworthy for the HE-DFT calculations, all of which, by design, produce nearly the same value for the K-S band gap [64]. Second, our LMCC results resemble more closely the HE-DFT results than the prior PBE results reported in Refs. [15,16,56]. It is fair to say that the LMCC results are in as good agreement with the other HE-DFT results as HE-DFT results are with each other. Of course, neither LMCC nor HE-DFT can prove accuracy. That requires demonstrating agreement with unambiguous experiment, and there is little

direct experiment that ties defect levels in the gap to specific physical models. Nevertheless, it is reassuring that two such different approaches lead to some consensus. This has both practical and fundamental implications. Practically, the agreement implies that LMCC is a useful tool—that its predictions are as valid as those from HE-DFT jellium, and that we can investigate larger systems more routinely. Fundamentally, it implies that prior claims that smaller splittings between defect levels are evidence of deficient localization in PBE [65], need revision. The level splittings in, for example, Ref. [16] are the result of the *jellium* approximation and the attendant incorrect electrostatic boundary conditions, rather than the functional. Third, the predicted LMCC defect conformations are in very good qualitative agreement with prior results, with a few notable exceptions (see Sects. III B 2 and III B 5.) This is expected, as local geometries are driven by symmetry and simple molecular orbital principles.

B. Cell-size convergence

We note our results are demonstrated to be converged using supercells containing up to 1000 atoms, where electrostatic cell-size errors and defect banding should be much smaller. This has given us confidence in our assignments of localized charge states. We have also demonstrated that for some defects, energies of formation are not converged at 64 atoms, and

TABLE VII. Slater determinant, wave-function symmetry, and defect symmetry in tetrahedral configuration for each charge state of the zinc substitutional on the nitrogen site.

Charge state	+3	+2	+1	0	-1	-2
Slater determinant	$ a_1 \uparrow a_1 \downarrow \rangle$	$ x \uparrow \rangle$	$ x \uparrow y \uparrow \rangle$	$ x \uparrow y \uparrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow z \uparrow \rangle$	$ x \uparrow x \downarrow y \uparrow y \downarrow z \uparrow \rangle$
wave function irrep.	A_1	T_2	T_1	A_2	T_1	T_2
defect symmetry	T_d	pD_{2d}	pD_{2d}	T_d	pD_{2d}	rD_{2d}

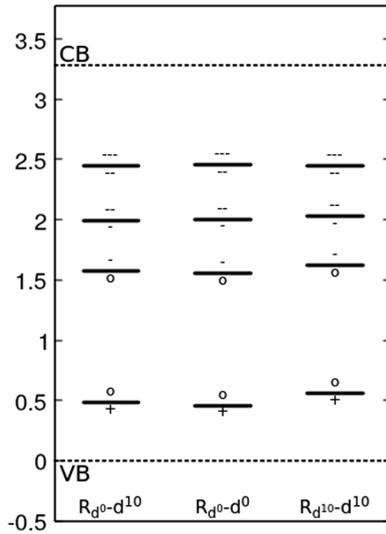


FIG. 19. Defect levels for the gallium vacancy for the d^{10} pseudopotential calculated at the d^0 (far left) and the d^{10} (far right) lattice constant, compared to the same levels for the d^0 pseudopotential calculated at the d^0 lattice constant.

that at least 216 atom cells are required. We should note that Burr and Cooper [66] have shown supercell size can be crucial even for neutral defects, where electrostatic effects should be very small (dipole and higher order). They showed that 96 atom supercells can give incorrect ordering of energies for different internal structures of neutral Schottky defects, and that this misordering is a direct result of poor representation of elastic properties in smaller cells. The misordering was independent of the level of theory: force field calculations gave the same misordering as DFT in 96 atom cells. In larger cells containing 324 atoms, and where elastic properties are represented better, the energetic ordering changes for all levels of theory.

C. Role of d functions

This is the first defect study systematically to compare results for pseudopotentials that include and exclude the gallium $3d$ orbitals in the valence set. Several previous reports have suggested their inclusion is crucial to obtain correct defect structures and levels [67–69], while Van de Walle and Neugebauer have claimed that a nonlinear core correction to the d^0 gallium pseudopotential suffices [16] to give results equivalent to the d^{10} . Thus, there lacks clear consensus about either the necessity or the utility of their inclusion. With few exceptions, our results predict the same qualitative geometries as stable and metastable, independent of the pseudopotential. The computed defect levels for d^0 and d^{10} are qualitatively and quantitatively similar (with noted exceptions). The only consistent difference is a compression of splittings for the d^{10} relative to the d^0 , which leads to a 15% reduction in the predicted effective defect band gap, which is consistent with the compression of the K-S band gap, as discussed in Sect. II. In Fig. 19, we show a comparison between the d^{10} results at the d^{10} and d^0 lattice constants, and the d^0 results at the d^0 lattice constant, for the gallium vacancy. There are similar

results for the other defects showing that the compression of levels is primarily the consequence of a change in lattice constant. In fact when the extremal defect levels ($-2/ -1$ for Zn_N , and $+2/ +3$ for N_I) for the d^{10} pseudopotential are calculated at the d^0 lattice constant, the predicted defect band gap increases by 0.5 eV, putting the d^0 and d^{10} predictions within 0.1 eV of each other. So, when lattice constants are equal, the defect band gap is approximately independent of the pseudopotential, as it was in Ref. [33]. Careful reinspection of the results in Fig. 4 of Ref. [33] shows identical trends in gallium arsenide that we see in gallium nitride. The d^{10} pseudopotentials give larger lattice constants, smaller K-S band gaps, and compressed defect level spectra compared to d^0 pseudopotentials. It was not remarked upon there because all of the effects were much smaller—a 0.5% change of lattice constant in gallium arsenide, compared to a 2.3% change in gallium nitride, and a 0.06 eV change in effective defect band gap in gallium arsenide, compared to a 0.6 eV change in gallium nitride.

The neutral formation energies for d^0 are consistently larger than for d^{10} , although the differences range from a few tenths of an eV for the nitrogen vacancy, to over 1 eV for the nitrogen antisite. More importantly, we find that there can be substantial variation in formation energies as a function of supercell size, although these appear to converge for cell sizes at and above 216 atoms. These calculations show clearly there is no meaningful utility in including gallium $3d$ orbitals in gallium nitride calculations, while the d^0 results for some bulk properties, such as lattice constant and elastic constants, are in better agreement with experiment.

D. Accuracy of defect band gap

The fidelity of the predicted effective defect band gap compared to experiment rests on the existence of defect levels very near both experimental band edges. It is fortuitous that we and others have been able to find defects with near-edge defect levels in many of the previously studied materials, including silicon [31], gallium arsenide [33,41,52], and silicon carbide [40,54], and that we have been able to do so in cubic gallium nitride. In cesium iodide, for example, the LMCC method covered 5 eV of a measured gap between 6.1 and 6.4 eV [53]. While a full band gap is not guaranteed, the relative positions of the defect levels are rigorously converged with respect to supercell size.

ACKNOWLEDGMENTS

AFRL acknowledges generous support from the Air Force Office of Scientific Research (AFOSR) through Contract No. FA9550-17RVCOR505. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy, or the US. Government. This work (Sandia) was supported by a Laboratory Directed Research and Development (LDRD) project, and was funded, in part, at the Center

for Integrated Nanotechnologies, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science. We gratefully acknowledge the AFRL High Performance Computing centers for their generous al-

lotment of computational resources. Finally, we thank Andrew Pineda, Renee Van Ginhoven, Andrew Mounce, and Normand Modine for valuable discussions during the course of this study.

-
- [1] L. Y. Lee, *Mater. Sci. Technol.* **33**, 1570 (2017).
- [2] C. Bayram and R. Liu, *Proc. SPIE Int. Soc. Opt. Eng.* **10111**, 101110Y (2017).
- [3] R. Liu, R. Schaller, C. Q. Chen, and C. Bayram, *ACS Photonics* **5**, 955 (2018).
- [4] I. Aharonovich, D. Englund, and M. Toth, *Nat. Photonics* **10**, 631 (2016).
- [5] A. M. Berhane, K.-Y. Jeong, Z. Bodrog, S. Fiedler, T. Schröder, N. V. Triviño, T. Palacios, A. Gali, M. Toth, D. Englund *et al.*, *Adv. Mater.* **29**, 1605092 (2017).
- [6] A. M. Berhane, K.-Y. Jeong, C. Bradac, M. Walsh, D. Englund, M. Toth, and I. Aharonovich, *Phys. Rev. B* **97**, 165202 (2018).
- [7] Y. Zhou, Z. Wang, A. Rasmita, S. Kim, A. Berhane, Z. Bodrog, G. Adamo, A. Gali, I. Aharonovich, and W.-B. Gao, *Sci. Adv.* **4**, eear3580 (2018).
- [8] M. Nguyen, T. Zhu, M. Kianinia, F. Massabuau, I. Aharonovich, M. Toth, R. Oliver, and C. Bradac, *Appl. Phys. Lett. Mat.* **7**, 081106 (2019).
- [9] Y. Gao, D. Sun, X. Jiang, and J. Zhao, *J. Appl. Phys.* **125**, 215705 (2019).
- [10] J. L. Lyons and C. G. Van de Walle, *npj Comput. Mater.* **3**, 12 (2017).
- [11] I. C. Diallo and D. O. Demchenko, *Phys. Rev. Appl.* **6**, 064002 (2016).
- [12] R. Gillen and J. Robertson, *J. Phys.: Condens. Matter* **25**, 405501 (2013).
- [13] S. A. Church, S. Hammersly, P. W. Mitchell *et al.*, *J. Appl. Phys.* **123**, 185705 (2018).
- [14] S. V. Novikov, A. J. Kent, and C. T. Foxon, *Prog. Cryst. Growth Charact. Mater.* **63**, 25 (2017).
- [15] M. G. Ganchenkova and R. M. Nieminen, *Phys. Rev. Lett.* **96**, 196402 (2006).
- [16] C. G. Van de Walle, J. Neugebauer *et al.*, *J. Appl. Phys.* **95**, 3851 (2004).
- [17] T. Mattila, A. P. Seitsonen, and R. M. Nieminen, *Phys. Rev. B* **54**, 1474 (1996).
- [18] I. Gorczyca, A. Svane, and N. E. Christensen, *Phys. Rev. B* **60**, 8147 (1999).
- [19] P. A. Schultz, *Phys. Rev. B* **60**, 1551 (1999).
- [20] P. A. Schultz, *Phys. Rev. Lett.* **84**, 1942 (2000).
- [21] D. M. Bylander and L. Kleinman, *Phys. Rev. B* **41**, 7868 (1990).
- [22] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy, *Phys. Rev. B* **53**, 3764 (1996).
- [23] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207 (2003).
- [24] G. Makov and M. C. Payne, *Phys. Rev. B* **51**, 4014 (1995).
- [25] S. Lany and A. Zunger, *Phys. Rev. B* **78**, 235104 (2008).
- [26] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, *Phys. Rev. Lett.* **102**, 016402 (2009).
- [27] H.-P. Komsa and A. Pasquarello, *Phys. Rev. Lett.* **110**, 095505 (2013).
- [28] Y. Kumagai and F. Oba, *Phys. Rev. B* **89**, 195205 (2014).
- [29] T. Zakrzewski and P. Boguslawski, *J. Alloys Compd.* **664**, 565 (2016).
- [30] O. Volnianska, T. Zakrzewski, and P. Boguslawski, *J. Chem. Phys.* **141**, 114703 (2014).
- [31] P. A. Schultz, *Phys. Rev. Lett.* **96**, 246401 (2006).
- [32] This insensitivity is demonstrated implicitly for silicon in Fig. 3 of Ref. [31], and Fig. 2 of Ref. [33]. These figures show the defect levels, and the defect band gap for LDA and PBE, respectively. Both the defect band gap and the individual defect levels are insensitive to changes in the exchange-correlation potential. The variation of the K-S band gap is mild (0.05 eV), but in both cases the defect gap is within 0.05 eV of experiment, while the K-S gap is small by roughly 45%.
- [33] P. A. Schultz and A. H. Edwards, *Nucl. Instrum. Methods Phys. Res., Sect. B* **327**, 2 (2014).
- [34] O. A. von Lilienfeld and P. A. Schultz, *Phys. Rev. B* **77**, 115202 (2008).
- [35] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.105.235110> for details of calculations presented here.
- [36] P. A. Schultz, seqQuest code (unpublished) see <http://dft.sandia.gov/Quest/>.
- [37] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [38] N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).
- [39] W. Jost, *J. Chem. Phys.* **1**, 466 (1933).
- [40] P. A. Schultz, R. M. Van Ginhoven, and A. H. Edwards, *Phys. Rev. B* **103**, 195202 (2021).
- [41] P. A. Schultz, *Phys. Rev. B* **93**, 125201 (2016).
- [42] Z. Usman, C. Cao, G. Nabi, W. S. Khan, T. Mehmood, and S. Hussain, *J. Phys. Chem. A* **115**, 6622 (2011).
- [43] R. Riane, S. F. Matar, and L. Hammerelaine, *Z. Naturforsch.* **63**, 1231 (2008).
- [44] L. E. Ramos, L. K. Teles, L. M. R. Scolfaro, J. L. P. Castineira, A. L. Rosa, and J. R. Leite, *Phys. Rev. B* **63**, 165210 (2001).
- [45] C. Stampfl and C. G. Van de Walle, *Phys. Rev. B* **59**, 5521 (1999).
- [46] M. B. Kanoun, Ph.D. thesis, Abou-Bakr Belkaid University, 2004.
- [47] K. Kim, W. R. L. Lambrecht, and B. Segall, *Phys. Rev. B* **53**, 16310 (1996).
- [48] K. Shimada, T. Sota, and K. Suzuki, *J. Appl. Phys.* **84**, 4951 (1998).
- [49] V. D. Compán García, I. E. Orozco Hinostrroza, A. Escobosa Echavarría, E. López Luna, A. G. Rodríguez, and M. A. Vidal, *J. Cryst. Growth* **418**, 120 (2015).
- [50] G. Ramírez-Flores, H. Navarro-Contreras, A. Lastras-Martínez, R. C. Powell, and J. E. Greene, *Phys. Rev. B* **50**, 8433 (1994).
- [51] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [52] P. A. Schultz and O. A. von Lilienfeld, *Modelling Simul. Mater. Sci. Eng.* **17**, 084007 (2009).

- [53] R. M. Van Ginhoven and P. A. Schultz, *J. Phys.: Condens. Matter* **25**, 495504 (2013).
- [54] M. A. Anders, P. M. Lenahan, A. H. Edwards, P. A. Schultz, and R. M. Van Ginhoven, *J. Appl. Phys.* **124**, 184501 (2018).
- [55] G. Fisherauer, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 1179 (1997).
- [56] Y. Gohda and A. Oshiyama, *J. Phys. Soc. Jpn.* **79**, 083705 (2010).
- [57] G. D. Watkins, *Deep Centers in Semiconductors* (Gordon and Breach, 1986), Chap. 3, p. 177.
- [58] J. Neugebauer and C. G. Van de Walle, *Phys. Rev. B* **50**, 8067 (1994).
- [59] H. Overhof and J.-M. Spaeth, *Phys. Rev. B* **72**, 115205 (2005).
- [60] C. Delerue, *Phys. Rev. B* **44**, 10525 (1991).
- [61] G. A. Baraff and M. Schluter, *Phys. Rev. Lett.* **55**, 2340 (1985).
- [62] S. F. Chichibu, A. Uedono, K. Kojima, H. Ikeda, K. Fujito, S. Takashima, M. Edo, K. Ueno, and S. Ishibashi, *J. Appl. Phys.* **123**, 161413 (2018).
- [63] G. D. Watkins, L. S. Vlasenko, and C. Bozdog, *Phys. B: Condens. Matter* **308**, 62 (2001).
- [64] Recently, Demchenko and Reshchikov [70] have argued that it is crucial that the HSE potential be parametrized to satisfy the generalized Koopman's relation. That is, along with constrain-
- ing the K-S band gap to agree with experiment, that the first unoccupied K-S eigenvalue be equal to the electron addition energy. We argue that, as a practical matter, this is unimportant. The authors point out that the usual parametrization leads to a very small error in the Koopman's relation, yet the previous calculations we compare to, all of which use roughly the same parametrization, vary dramatically. Furthermore, our total energy results are independent of the K-S eigenvalues.
- [65] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, *Rev. Mod. Phys.* **86**, 253 (2014).
- [66] P. A. Burr and M. W. D. Cooper, *Phys. Rev. B* **96**, 094107 (2017).
- [67] V. Fiorentini, M. Methfessel, and M. Scheffler, *Phys. Rev. B* **47**, 13353 (1993).
- [68] J. Neugebauer and C. G. Van de Walle, *MRS Proc.* **339**, 687 (1994).
- [69] K. Karch, F. Bechstedt, and T. Pletl, *Phys. Rev. B* **56**, 3560 (1997).
- [70] D. O. Demchenko and M. A. Reshchikov, *J. Appl. Phys.* **127**, 155701 (2020).
- [71] M. Chagas da Silva, M. Lorke, B. Aradi, M. Farzalipour Tabriz, T. Frauenheim, A. Rubio, D. Rocca, and P. Deak, *Phys. Rev. Lett.* **126**, 076401 (2021).