Theoretical study of the band-gap anomaly of InN

Pierre Carrier and Su-Huai Wei[a)]
National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, Colorado 80401

(Received 21 July 2004; accepted 22 November 2004; published online 14 January 2005)

Using a band-structure method that includes the correction to the band-gap error in the local-density approximation (LDA), we find that the band gap for InN is 0.8±0.1 eV, in good agreement with recent experimental data, but is much smaller than previous experimental value of ~1.9 eV. The unusually small band gap for InN is explained in terms of the high electronegativity of nitrogen and, consequently, the small band-gap deformation potential of InN. The possible origin of the measured large band gaps is discussed in terms of the nonparabolicity of the bands, the Moss–Burstein shift, and the effect of oxygen. Based on the error analysis of our LDA-corrected calculations we have compiled the band-structure parameters for wurtzite AlN, GaN, and InN. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1849425]

I. INTRODUCTION

III-nitrides are usually considered as wide-band-gap materials that have applications in devices such as ultraviolet/blue/green light-emitting diodes and lasers.[1] However, recent measurements suggest that the band gap of wurtzite (WZ) InN is below 1.0 eV,[2–6] much smaller than the 1.89-eV band gap[7] widely accepted in the past to interpret experimental data[1] and to fit empirical pseudopotentials for modeling InN and related alloy properties.[8,9] If InN indeed has a less than 1.0-eV band gap, which is even smaller than that for InP (1.4 eV),[10] then InN and its III-nitride alloys could also be suitable for low-band-gap device applications such as future-generation solar cells because the nitride alloys can cover the whole solar spectrum range.

The possible low band gap of InN also provides a challenge to understand the general chemical trends of semiconductor band gaps.[11,12] Conventional wisdom holds that for common-anion (or cation) III-V semiconductors, the direct band gap at Γ increases as the cation (or anion) atomic number decreases. This observation is strongly supported by experimental data[10] shown in Fig. 1. This trend would hold for common-cation InX (X=N, P, As, and Sb) compounds if E_g(InN)=1.9 eV, as previously reported.[7] However, the trend will be broken if E_g(InN)~0.8 eV, as reported in recent measurements.[2–6]

To understand the origin of the InN band-gap anomaly, we have performed band-structure calculations within the local-density approximation[13] (LDA) in which the LDA band-gap error are corrected using a semiempirical method.[12] We find that the band gap of WZ InN is 0.8±0.1 eV, in good agreement with recent experimental measurements. We show that the reason that InN has a smaller band gap than InP is due to the much larger electronegativity and the much smaller band-gap deformation potential for InN. The possible origin of the measured apparent large band gaps is discussed in terms of the nonparabolicity of the bands, the carrier-induced band-gap shift, and the effect of oxygen. The semiempirical approach is also applied to analyze the LDA error in the calculated band-structure parameters. Based on this analysis, more realistic band-structure parameters for the III-nitrides are proposed.

II. METHOD OF CALCULATIONS

The band-structure calculations in this study are performed using the fully relativistic (including spin-orbit coupling), general potential, linearized augmented plane-wave (LAPW) method.[14] The band structures are calculated at experimental lattice constants.[10] The absorption coefficients for the nitrides are calculated using the optical package in WIEN2K.[15] Details of the calculation method is described in Ref. 12.

Although LDA is accurate in predicting the ground-state properties such as the lattice parameters (Table I), it is well known that it severely underestimates the semiconductor band gap.[16] Direct calculations beyond the LDA is currently difficult.[12,17–20] This is partly because the presence of the In 4d orbitals in the valence bands and partly because the LDA-calculated band gap for InN is negative.[21] To correct the LDA band-gap error, we use a well-established approach[12] by adding to the LDA potential δ-functionlike external potentials[22,23] inside the muffin-tin (MT) spheres centered at each atomic site α

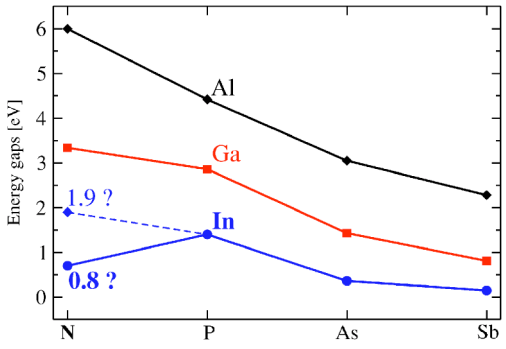

FIG. 1. Band gaps as a function of anions for III-V semiconductors. The common- cation trend is broken for InN if its band gap is 0.8 eV.

[a)]Electronic mail: suhuai_wei@nrel.gov

TABLE I. Calculated LDA structural parameters and bulk moduli of AlN, GaN, and InN. Results are compared with the available experimental data (in parentheses). ΔE_{ZB-WZ} is the calculated total energy difference between the ZB and WZ phases. Positive number indicates that the WZ structure is more stable, in agreement with the experiment.

Properties	AlN	GaN	InN
a(Å)	3.098(3.112)	3.170(3.189)	3.546(3.544)
c/a(Å)	1.601(1.601)	1.625(1.626)	1.612(1.613)
u	0.3819	0.3768	0.3790
B_{WZ}(kbar)	2145	2046	1484
a_{ZB}(Å)	4.364(4.36)	4.488(4.50)	4.975(4.98)
B_{ZB}(kbar)	2158	2063	1498
ΔE_{ZB-WZ},(meV/2-atom)	45	11	21

$$V_{\text{ext}}^{\alpha}(r) = \overline{V}^{\alpha} + V_0^{\alpha}\left(\frac{r_0^{\alpha}}{r}\right)e^{-(r/r_0^{\alpha})^2}, \tag{1}$$

and performed the calculation self-consistently. Specifically, the parameters in Eq. (1) are fitted first only to the available experimental energy levels and to the quasiparticle energies[16] at *high-symmetry k* points for AlP, GaP, InP, and GaN.[23] The *same parameters* are then used to predict the band gaps of arsenides, antimonides, and nitrides. To find the band gap for the wurtzite structure, we add the LDA-calculated band-gap differences between the WZ and zincblende (ZB) compounds to the calculated band gaps for the ZB compound. Compared to our directly calculated results for the wurtzite structure, we find that this procedure is reliable; the overall uncertainty of the predicted band gap associated with this fitting procedure is estimated to be less than 0.1 eV.

The effective mass is calculated using two definitions. First, for the transport effective mass we assume the derivative $dE(\mathbf{k})/d\mathbf{k}$, which defines the electron group velocity, to be the same as a parabolic band with effective mass m_T, i.e.,

$$m_T[E(\mathbf{k})] = \frac{\hbar^2 k}{dE(\mathbf{k})/d\mathbf{k}}. \tag{2}$$

Second, for the density-of-states (DOS) effective mass, which defines the Moss–Burstein shift,[24] we assume that the DOS of the true band to be the same as a parabolic band with effective mass m_D, i.e.,

$$m_D(E) = \left\{\frac{\int_0^E [m_T(\epsilon)]^{3/2}\sqrt{\epsilon}d\epsilon}{\int_0^E \sqrt{\epsilon}d\epsilon}\right\}^{2/3}, \tag{3}$$

where $m_T(E)$ is the transport effective mass defined in (2) and the energy zero is set at the conduction-band minimum (CBM). For wurtzite compounds with anisotropic band we define $m_D^* = [(m_D^{\perp})^2 m_D^{\parallel}]^{1/3}$, where $m_D^{\perp}$ and $m_D^{\parallel}$ are the effective masses perpendicular and parallel to the c axis, respectively.

The two definitions are identical at Γ, or if the band is parabolic, but could be significantly different at large $\mathbf{k}$ when the band is nonparabolic. For most semiconductors,[10] $E(\mathbf{k})$ contains additional terms of the order k^4; therefore, near the CBM the electron effective mass increases linearly as a function of the band-edge energy E and the slope of m_T vs E is larger than the slope of m_D vs E. It is well known that LDA also underestimates the calculated effective masses. Similar to the treatment for the band gap, we have fitted our results for GaN and GaAs, and applied the same procedures to calculate the effective masses for AlN and InN.

TABLE II. Calculated band gaps at Γ for ZB and WZ III-V compounds at experimental (exp) lattice constants using the LDA-plus-correction (LDA +C) methods. The $E_g^{\text{LDA+C}}$ values with an (*) are fitted values, whereas all the others are predicted values. Our calculated results are compared with the available experimental data (Ref. 10). The last column shows the error between the predicted values and the experimental data.

Compound	a_{exp} (Å)	$E_g^{\text{LDA+C}}$ (eV)	E_g^{exp} (eV)	$\lvert\delta E_g\rvert$ (eV)
AlSb	6.133	2.28	2.32	0.04
GaSb	6.096	0.81	0.81	0.00
InSb	6.479	0.15	0.24	0.09
AlAs	5.660	3.05	3.13	0.08
GaAs	5.653	1.43	1.52	0.09
InAs	6.058	0.36	0.42	0.06
AlP	5.467	4.42*	4.38	0.04
GaP	5.451	2.86*	2.86	0.00
InP	5.869	1.40*	1.46	0.06
AlN	4.360	6.00	···	···
GaN	4.500	3.34*	3.32	0.02
InN	4.980	0.70	···	···
AlN(WZ)	a=3.112 c=4.982 u=0.3819	5.95	6.1	0.15
GaN(WZ)	a=3.189 c=5.185 u=0.3768	3.49	3.5	0.01
InN(WZ)	a=3.544 c=5.718 u=0.3790	0.85	···	···

III. RESULTS AND DISCUSSIONS

A. Band-gap anomaly of InN

The predicted direct band gaps at the Γ point for the III-V semiconductors using the approach described in Ref. 12 are shown in Table II. These values are compared with the available experimental data.[10] We find that for nearly *all* the III-V semiconductors, the differences between the predicted and the experimental band gaps are less than 0.1 eV. For InN, however, our predicted value of 0.85 eV is much smaller than the previous experimental value[7] of 1.9 eV, but it is in very good agreement with recent experimental measurements.[2–6] For AlN, our predicted band gap of 6.0 eV is also close to recent photoluminescence measurement of 6.1 eV,[25] which is smaller than the previously accepted value around 6.3 eV.[1,10,26]

To understand the origin of the small band gap of InN and the general chemical trends of the semiconductor band gaps, we studied the chemical and size contributions of the band gap in III-V semiconductors.[12] First, we find that at the fixed phosphide volume, the band gaps of the common-cation system decrease from *M*Sb to *M*P to *M*As to *M*N (*M*=Al, Ga, and In), following the same trend of the anion atomic valence *s* orbital energies.[12] This is because the CBM

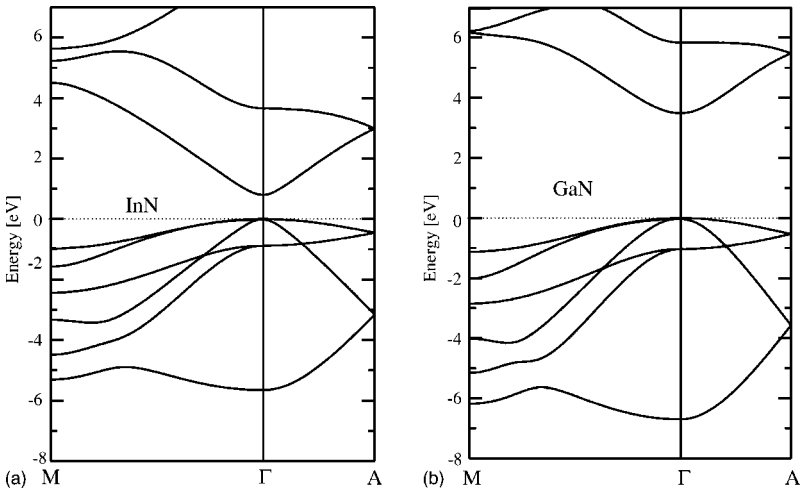

FIG. 2. Calculated band structure of (a) wurtzite InN and (b) wurtzite GaN. The energy zero is set at VBM.

at the Γ point is an anion *s* plus cation *s* state. Because the N 2*s* orbital energy is much lower in energy than the Sb 5*s*, As 4*s*, and P 3*s* orbital energies, respectively, the band gap of the nitrides is also much lower than the corresponding antimonides, arsenides, and phosphides at fixed volume. Next, we investigate the size or volume deformation contribution to the band gap. We find that all the compounds have negative volume deformation potentials at Γ, i.e., when the volume decreases, the band gap increases.[27] Therefore, it is clear that the observed chemical trends, that the smaller compound has larger band gap, are mainly due to the large deformation potential. However, for InN, although its volume is about 49% smaller than InP, its band-gap deformation potential $a_g = dE_g/dlnV$ is small due to the large ionicity and relatively large bond length of InN $[a_g(\mathrm{InN}) = -4.2\ \mathrm{eV}]$.[27] Because of this small $|a_g|$, the contribution due to the size or deformation potential is not sufficient to reverse the order of the band gap due to the contribution of the chemical effect. This explains why InN has a band gap of about 0.6 eV smaller than that of InP. We find that a similar situation also exists in II-VI semiconductors. Indeed, the experimental data[10] show that the ZnO band gap of 3.4 eV is smaller than the ZnS band gap of 3.8 eV. Our calculations also show that CdO and HgO would have band gaps that are about 0.5 eV smaller than the band gaps of CdS and HgS, respectively, if they could all exist in the ZB phase.

B. Possible origin of the measured large band gap for InN

Our calculation and analysis above show that the fundamental band gap of InN is indeed small, around 0.8 eV. To understand the origin of some of the experiments which show a large band gap of InN,[7,28,29] we have performed a detailed study of the band structure of InN. Experimentally, one often assumes that the band-edge states near Γ is parabolic and the dipole transition matrix element is nearly independent of **k**; therefore, the absorption coefficient squared, α^2, is a linear function of the absorption energy *E*. The fundamental band gap, thus, can be obtained from the interception with the energy axis by drawing a straight line in the α^2 vs *E* plot.[7,28,29] To test the validity of this assumption, we show in Fig. 2 the calculated band structure of wurtzite InN. Figure 3 shows the calculated electron effective masses, and Fig. 4 shows our calculated absorption coefficients of InN. For comparison, we also calculate the band structure and absorption coefficients of GaN. We find that the conduction band of InN is strongly nonparabolic, due to the small band gap, which leads to strong coupling between valence and conduction bands (Fig. 2). This is confirmed from the calculated electron effective masses (Fig. 3), which increase significantly with the band-edge energy or electron concentration. If the conduction band was parabolic the electron effective mass would be a constant. Because of the large

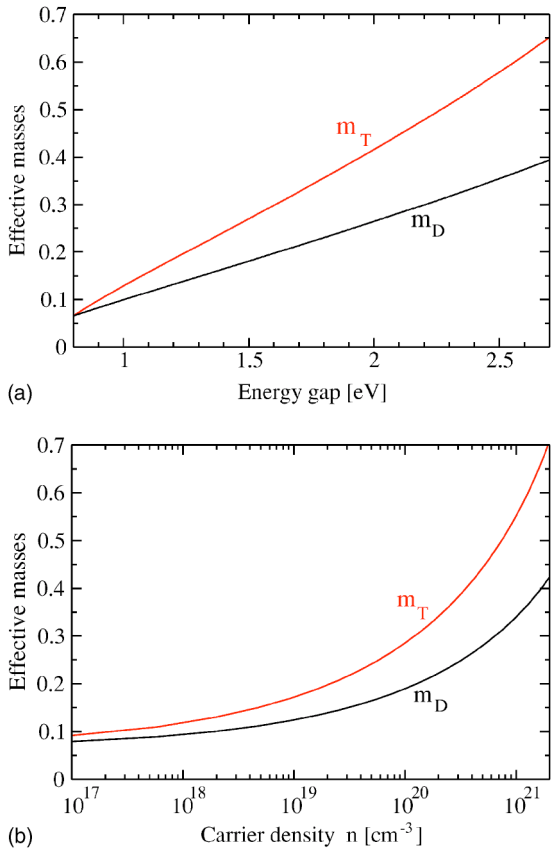

FIG. 3. Calculated electron effective masses as a function of (a) the absorption edge energy and (b) the carrier density. The energy zero is at VBM. The definition of m_T and m_D are given in Eqs. (2) and (3), respectively.

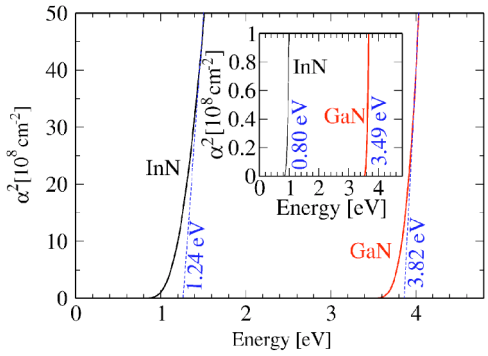

FIG. 4. Calculated absorption coefficients of InN and GaN. We show that using the linear extrapolation technique the apparent measured band gaps depend on the scale used in the extrapolation.

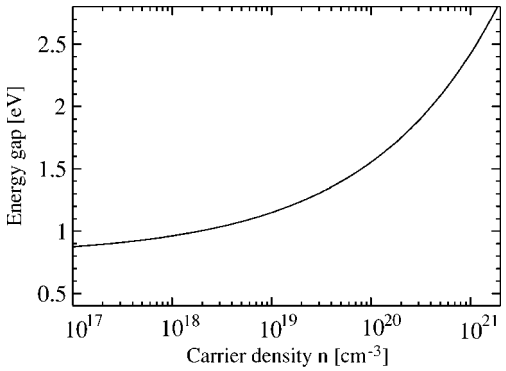

FIG. 5. Calculated Moss–Burstein shift of the absorption edge energy as a function of the carrier density.

deviation from the parabolic band, the calculated α^2 is not a linear function of E (Fig. 4). Therefore, if one uses the linear extrapolation technique to determine the band gap, the derived apparent band gaps depend on where the straight lines are drawn. For example, as shown in Fig. 4 for InN, using large values of the absorption coefficient to draw the straight line, one can obtain an apparent band gap that is about 0.4 eV larger than the fundamental band gap. The dependence is relatively smaller for GaN which has a larger effective mass than that of InN, thus a larger density of states near the CBM and a sharper increase of the absorption coefficient. We notice that the samples used to obtain the large InN band gaps[7,28,29] often have poor sample quality and the band gaps are usually estimated from the absorption spectra with large absorption coefficients.

In addition, we have investigated the effect of oxygen on the band gap of InN.[28] First, we assume that O substitutes N, forming InN_xO_{1-x} alloy. To balance the charge, In vacancy (V_{In}) is also introduced (with a ratio of 3 to 1 between O_N and V_{In}). Our calculations show that V_{In} strongly binds with O_N. However, the band gap of InN_xO_{1-x} alloy decreases compared to bulk InN. This is consistent with the fact that O_N induces a donor level below the InN CBM and V_{In} creates an acceptor level above the InN valence-band maximum (VBM). Second, we studied the $(InN)_{2n}(In_2O_3)_m(001)$ superlattices. We notice that in In_2O_3, In atoms form a slightly distorted fcc sublattice that matches nicely with InN in the ZB structure. Therefore, In_2O_3 could be a useful substrate for growing ZB InN. The calculated valence-band offset between InN and In_2O_3 is 0.6 ± 0.2 eV, with a type-I band alignment. However, the change of the band gap of the superlattice is small. In a superlattice $(InN)_4(In_2O_3)_2(001)$, the band-gap increase is only about 0.3 eV. Therefore, oxygen may not be an important fact in determining the measured InN band gap.[30]

Finally, we notice that the InN samples that show large band gaps are often heavily n-type doped.[7,28,29] This suggests that the measured absorption edge can be shifted by the Moss–Burstein effect.[24] Figure 5 shows our calculated absorption edge energy as a function of the carrier density. We see that the absorption edge increases with the carrier density from 0.8 eV for intrinsic InN to ~2.5 eV for the sample with electron concentration of $\sim10^{21}$ cm^3. These results are consistent with recent experimental observation.[31]

IV. BAND-STRUCTURE PARAMETERS OF NITRIDES

In the last few years several excellent review articles[32,33] and books[1,26] have been published to describe the band-structure parameters of InN and other III-nitrides. The recommended band-structure parameters are based on the collections of available experimental data and theoretical calculations. The semiempirical approach used in this study enabled us to analyze the degree of the systematic LDA errors on the band-structure parameters of the III-nitrides. For example, we find that the LDA did not only underestimate the band gap, but also overestimated the coupling between the conduction band and the valence light hole band. This introduces error in the LDA-calculated crystal-field splittings Δ_{CF} and the spin-orbit splittings Δ_0 at the top of the valence band, as well as the electron and light hole effective masses. Based on the LDA error analysis and comparison with the available experimental data we propose here a set of band-structure parameters for the III-nitrides shown in Table III. We suggest that these parameters should be compared directly with the experimental data and be used to fit empirical pseudopotentials[8,9] to study the nitride systems.

V. SUMMARY

In summary, using a semiempirical LDA-based band-structure method with band-gap correction we have systematically studied the chemical trends of the band-gap variation in III-V semiconductors. We find that InN has a band gap of 0.8 ± 0.1 eV, in good agreement with recent experimental measurements. The possible origin of the measured large band gaps is analyzed. We find that the Moss–Burstein shift should be the dominant factor in determining the measured InN band gap. Based on the error analysis of our LDA-calculated results using the semiempirical approach, we have proposed a set of the band-structure parameters for wurtzite AlN, GaN, and InN, which can be compared directly with the experimental data.

TABLE III. Recommended band-structure parameters at Γ for unstrained AlN, GaN, and InN. The properties shown in this table are the band gap E_g, the spin-orbit splitting Δ_0, the crystal-field splitting Δ_{CF}, the valence-band splittings ΔE_{12} and ΔE_{13}, and the effective masses m parallel and perpendicular to the c axis. The averaged effective mass can be obtained using $m^* = [(m^{\perp})^2 m^{\parallel}]^{1/3}$. For AlN, the indices A, B, and C for the effective masses corresponds to $\Gamma_{7\mathrm{VBM}}$, Γ_{9v}, and Γ_{7v} states, respectively (Ref. 25). For GaN and InN, the orders of A, B, and C correspond to $\Gamma_{9\mathrm{VBM}}$, Γ_{7v}, and Γ_{7v} states, respectively (Ref. 34).

Properties	AlN		GaN		InN	
E_g(WZ)(eV)	6.10		3.51		0.78	
E_g(ZB)(eV)	6.15		3.35		0.70	
Δ_0(meV)	19		16		5	
Δ_{CF}(meV)	−224		25		19	
ΔE_{12}(meV)	218		8		3	
ΔE_{13}(meV)	237		33		21	
	Effective masses					
	$\perp$	$\parallel$	$\perp$	$\parallel$	$\perp$	$\parallel$
$m_A(m_0)$	4.35	0.28	0.39	2.04	0.14	2.09
$m_B(m_0)$	0.67	3.50	0.43	0.85	0.13	0.50
$m_C(m_0)$	0.68	3.43	1.05	0.19	0.81	0.07
$m_e(m_0)$	0.33	0.32	0.22	0.20	0.07	0.06

ACKNOWLEDGMENTS

We would like to thank I. G. Batyrev, X. Nie, S. B. Zhang, and Clas Persson for many valuable discussions. This work is supported by the U.S. Department of Energy, Contract No. DE-AC36-99GO10337.

[1] H. Morkoç, *Nitride Semiconductors and Devices* (Springer, New York, 1999).
[2] T. Inushima, V. V. Mamutin, V. A. Vekshin, S. V. Ivanov, T. Sakon, M. Motokawa, and S. Ohoya, J. Cryst. Growth **227/228**, 481 (2001).
[3] V. Yu. Davydov *et al.*, Phys. Status Solidi B **230**, R4 (2002); **229**, R1 (2002); **234**, 787 (2002).
[4] J. Wu *et al.*, Appl. Phys. Lett. **80**, 3967 (2002).
[5] J. Wu, W. Walukiewicz, K. M. Yu, J. W. Ager III, E. E. Haller, H. Lu, and W. J. Schaff, Appl. Phys. Lett. **80**, 4741 (2002).
[6] A. Yamamoto, K. Sugita, H. Takatsuka, A. Hashimoto, and V. Yu. Davydov, J. Cryst. Growth **261**, 275 (2004).
[7] T. L. Tansley and C. P. Foley, J. Appl. Phys. **59**, 3241 (1986).
[8] K. Kim and A. Zunger, Phys. Rev. Lett. **86**, 2609 (2001).
[9] P. R. C. Kent and A. Zunger, Appl. Phys. Lett. **79**, 1977 (2001).
[10] *Semiconductors: Data Handbook*, edited by O. Madelung (Springer, Berlin, 2004).
[11] B. R. Nag, Phys. Status Solidi B **233**, R8 (2002).
[12] S.-H. Wei, X. Nie, I. G. Batyrev, and S. B. Zhang, Phys. Rev. B **67**, 165209 (2003).
[13] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965).
[14] S.-H. Wei and H. Krakauer, Phys. Rev. Lett. **55**, 1200 (1985); D. J. Singh, *Planewaves, Pseudopotentials, and the LAPW Method* (Kluwer, Norwell, 1994).
[15] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, and J. Luitz, WIEN2K code, Vienna University of Technology, November 2001; P. Blaha, K. Schwarz, P. Sorantin, and S. B. Trickey, Comput. Phys. Commun. **59**, 399 (1990).
[16] X. Zhu and S. G. Louie, Phys. Rev. B **43**, 14142 (1991).
[17] M. van Schilfgaarde, A. Sher, and A.-B. Chen, J. Cryst. Growth **178**, 8 (1997); A. Sher, M. van Schilfgaarde, M. A. Berding, S. Krishnamurthy, and A.-B. Chen, MRS Internet J. Nitride Semicond. Res. **4S1**, G5.1 (1999).
[18] T. Kotani and M. van Schilfgaarde, Solid State Commun. **121**, 461 (2002).
[19] D. Vogel, P. Krüger, and J. Pollmann, Phys. Rev. B **55**, 12836 (1997); C. Stampfl, C. G. Van de Walle, D. Vogel, P. Krüger, and J. Pollmann, *ibid.* **61**, R7846 (2000).
[20] K. A. Johnson and N. W. Ashcroft, Phys. Rev. B **58**, 15548 (1998).
[21] F. Bechstedt and J. Furthmüller, J. Cryst. Growth **246**, 315 (2002).
[22] N. E. Christensen, Phys. Rev. B **30**, 5753 (1984).
[23] S.-H. Wei and A. Zunger, Phys. Rev. B **57**, 8983 (1998).
[24] T. S. Moss, Proc. Phys. Soc. London, Sect. B **67**, 775 (1954); E. Burstein, Phys. Rev. **93**, 632 (1954).
[25] J. Li, K. B. Nam, M. L. Nakarmi, J. Y. Lin, H. X. Jiang, P. Carrier, and S.-H. Wei, Appl. Phys. Lett. **83**, 5163 (2003).
[26] *Properties of Group III Nitrides*, edited by J. H. Edgar (INSPEC, London, 1994).
[27] S.-H. Wei and A. Zunger, Phys. Rev. B **60**, 5404 (1999).
[28] Motlan, E. M. Goldys, and T. L. Tansley, J. Cryst. Growth **241**, 165 (2002).
[29] K. Osamura, K. Nakajima, Y. Murakami, P. H. Shingu, and A. Ohtsuki, Solid State Commun. **11**, 617 (1972); K. Osamura, S. Naka, and Y. Murakami, J. Appl. Phys. **46**, 3432 (1975).
[30] M. Yoshimoto, H. Yamamoto, W. Huang, H. Harima, J. Saraie, A. Chayahara, and Y. Horino, Appl. Phys. Lett. **83**, 3480 (2003).
[31] J. Wu, W. Walukiewicz, W. Shan, K. M. Yu, J. W. Ager III, E. E. Haller, H. Lu, and W. J. Schaff, Phys. Rev. B **66**, 201403 (2002).
[32] I. Vurgaftman and J. R. Meyer, J. Appl. Phys. **94**, 3675 (2003).
[33] A. G. Bhuiyan, A. Hashimoto, and A. Yamamoto, J. Appl. Phys. **94**, 2779 (2003).
[34] G. D. Chen, M. Smith, J. Y. Lin, H. X. Jiang, S.-H. Wei, M. Asif Khan, and C. J. Sun, Appl. Phys. Lett. **68**, 2784 (1996).