Microscopic origins of surface states on nitride surfaces*

Chris G. Van de Walle a) and David Segev
Materials Department, University of California, Santa Barbara, California 93106

(Received 31 July 2006; accepted 15 September 2006; published online 27 April 2007)

We report a systematic and comprehensive computational study of the electronic structure of GaN and InN surfaces in various orientations, including the polar c plane, as well as the nonpolar a and m planes. Surface band structures and density-of-states plots show the energetic position of surface states, and by correlating the electronic structure with atomistic information we are able to identify the microscopic origins of each of these states. Fermi-level pinning positions are identified, depending on surface stoichiometry and surface polarity. For polar InN we find that all the surface states are located above the conduction-band minimum, and explain the source of the intrinsic electron accumulation that has been universally observed on InN surfaces. © 2007 American Institute of Physics. [DOI: 10.1063/1.2722731]

I. INTRODUCTION

Group III nitrides have experienced a remarkable ascent over the past 15 years, from being mere subjects of academic interest to materials of choice for green, blue, and ultraviolet (UV) optoelectronics, as well as high-power radio-frequency electronics. Considerable efforts have been invested in improving the performance of these devices. In many cases, the surface plays a key role. During growth, the surface structure determines morphology and host atom, as well as impurity incorporation and therefore ultimately the crystal quality. The position of the Fermi level at the surface often affects the incorporation of defects and impurities during growth.1 At the same time, as demonstrated by recent x-ray photoemission experiments on GaN (0001),2 the Fermi-level position at the surface depends strongly on growth conditions.

It has also been shown that due to its proximity to the surface, the electron channel in nitride-based high-electron mobility transistors (HEMTs) is very sensitive to the surface conditions; indeed, the surface has been proposed to be the main source of electrons for the two-dimensional electron gas (2DEG).3 Furthermore, the typical growth process along the polar (0001) direction results in spontaneous and piezoelectric polarization fields,4 which are beneficial for electronic devices such as HEMTs, but detrimental to the radiative efficiency of quantum wells for light emitters.5 There has therefore been a strong recent focus on growth of nitrides along nonpolar [m(1100) and a(1120) plane] orientations, which eliminate the polarization fields along the quantum-well direction.5–7 In the case of InN, additional motivation for our study is provided by the observation of a very high, and so far unexplained, electron accumulation on InN surfaces.8,9 This high surface sheet charge density compiles measurements of the electrical properties of the underlying bulk InN layer,8 and may interfere with the performance of InN-based devices. Efforts to eliminate this electron accumulation have not been successful so far.

In Sec. II we report a survey of experimental results for the surface band structure of GaN and InN. This survey shows that although valuable information has been obtained, a consistent interpretation has not yet been achieved. In particular, a detailed understanding of the connection between electronic properties and microstructure of nitride surfaces is still lacking. The dependence on crystal orientation and growth conditions is also far from understood. Past studies for a variety of semiconductors have established that first-principles calculations based on density functional theory (DFT) offer the most accurate description of structural and electronic properties.10 Nevertheless, most studies to date have suffered from the underestimation of the band gap inherent to DFT, which severely complicates the study of surface band structures.11–14 The case of InN is even more acute since DFT calculations in the local density approximation (LDA) produce a negative band gap, prohibiting any interpretation of surface electronic structure.

In the present study we have overcome this problem by applying an extensively tested modification to the pseudopotentials within the pseudopotential-DFT method, allowing us to evaluate the surface states within a band gap close to the experimental value, but still based on a self-consistent scheme and a full relaxation of the atomic structures. Using this approach, we have carried out a comparative study of the electronic properties of the stable reconstructions on polar and nonpolar surfaces of GaN and InN, as a function of the surface stoichiometry. Our results are presented in terms of two stoichiometry regimes: one corresponding to moderate cation/anion (Ga/N or In/N) ratios, the other to high ratios. Low cation/anion ratios typically result in poor-quality growth,15 and will not be addressed here.

In the next section we review experimental results that are available in the literature. Section III describes our methodology, and Sec. IV contains the main results.

---

*This paper is based on a talk presented by the authors at the 28th International Conference on the Physics of Semiconductors, which was held 24–28 July 2006, in Vienna, Austria. Contributed papers for that conference may be found in “Physics of Semiconductors: 28th International Conference on the Physics of Semiconductors,” AIP Conference Proceedings No. 893 (AIP, Melville, NY, 2007); see http://proceedings.aip.org/proceedings/confproceed/893.jsp

a)Electronic mail: vandewalle@mrl.ucsb.edu
II. SURVEY OF EXPERIMENTAL RESULTS

Several experimental studies have addressed the electronic properties of GaN surfaces through the determination of band bending and surface Fermi-level pinning.\textsuperscript{2,16–18} Band dispersion has also been investigated using photoemission.\textsuperscript{19–25} Long et al.\textsuperscript{16} used photoelectron (PES) and x-ray photoemission spectroscopy (XPS) on the clean GaN(0001) surface. In the limit of low temperature and low doping concentration, $E_F-E_v$ was measured to be about 2.7 eV on the surface of n-type GaN, and about 1.3 eV on the surface of p-type GaN. $E_F$ is the Fermi level and $E_v$ the position of the valence-band maximum (VBM). Similar results were obtained by Wu et al.,\textsuperscript{17} who measured an $E_F-E_v$ value at room temperature of about 2.6 eV, with a corresponding band bending of 0.75±0.1 eV for n-type GaN, while for p-type GaN they found $E_F-E_v=1.0$ eV and a band bending of 0.75±0.1 eV. Cho et al.\textsuperscript{18} showed that UV illumination strongly reduces the band bending on the (0001) surface of as-grown n-type GaN, from an initial value of about 1 eV to a final value of 0.75 eV. This effect was attributed to the presence of electrons that occupy surface states, and whose concentration can be reduced by optical excitation.

All these experimental data were obtained by preparing the GaN surface in such a way as to remove adsorbates that could affect the surface electronic properties. However, as shown by Kočan et al.\textsuperscript{2} using in situ XPS, in a growth environment the growth conditions themselves, such as the Ga/N ratio, can strongly affect the Fermi level at the GaN surface. $E_F$ was found to vary from $E_F-E_v=2.89$ eV under nearly stoichiometric growth conditions to $E_F-E_v=1.65$ eV under Ga-rich conditions where metallic Ga forms on the surface.\textsuperscript{3}

Band-dispersion results were obtained through angle-resolved photoemission spectroscopy (ARPES) on the GaN(0001) (Refs. 19–23) and GaN(000\bar{1}) (Refs. 24 and 25) surfaces. The reconstructions were deduced from low-energy electron diffraction measurements on clean surfaces, although the presence of coexisting Ga- and N-terminated surfaces could not be excluded.\textsuperscript{19,25} Due to uncertainties in the surface structure and stoichiometry, and the potential presence of adsorbates, it is in general very difficult to unambiguously explain the origin of the surface states based solely on ARPES. Nevertheless, the studies indicate that on the GaN(0001) surface a nondispersive, occupied state is present close to the VBM.\textsuperscript{19–21} The experimentalists related this state to Ga dangling bonds; theoretically, Wang et al.\textsuperscript{26} attributed it to the presence of Ga adatoms on $T_2$ sites (on top of sublayer N atoms and bonded to three Ga surface atoms).

On the GaN(000\bar{1}) surface, Kowalski et al.\textsuperscript{24} observed nondispersive occupied surface states in the vicinity of the VBM and at about 1.2 eV above it. After subsequent Ga deposition a nondispersive surface state was induced close to the VBM, and another, more dispersive, state at about 1.8 eV above the VBM. In contrast, Ryan et al.\textsuperscript{25} observed both dispersive and nondispersive surface states on “clean” surfaces, which were attributed by Wang et al.\textsuperscript{11} to the coexistence of clean and Ga-covered regions on the GaN(000\bar{1}) surface.

Studies of the electronic properties of InN surfaces are scarcer, mostly due to difficulties in obtaining high-quality material. The available studies focused on measurements of band bending and Fermi-level pinning on the InN(0001) surface.\textsuperscript{9,27,28} A downward band bending of ~0.6–0.7 eV was observed, associated with donor-type surface states with a density of about $2.5(\pm0.2) \times 10^{13}$ cm$^{-2}$, and a surface Fermi level of about 1.6±0.1 eV above the VBM.\textsuperscript{9,28}

Our survey indicates that even though a variety of experimental results are available, a number of conflicting results have been reported and a consistent interpretation is lacking. First-principles calculations of the type reported in this article can be very helpful in sorting out these issues.

III. METHODOLOGY

We use DFT within either the local density approximation (LDA) (Refs. 29 and 30) or the generalized gradient approximation (GGA).\textsuperscript{31} With LDA the band gap of GaN is 2.20 eV, while with GGA it is as small as 1.67 eV. Ultrasoft\textsuperscript{32} as well as norm-conserving\textsuperscript{33,34} pseudopotentials were tested. Calculations performed with LDA and GGA result in different positions of surface states within the respective band gap, indicating the difficulty of interpreting the results and the need for a systematic improvement in the band structure. After testing a number of approaches, we found we could obtain very reliable results with a method based on modified pseudopotentials. The modification is performed by including an atom-centered repulsive potential of Gaussian shape, in the spirit of an approach proposed by Christensen in the context of linearized muffin-tin orbital (LMTO) calculations.\textsuperscript{35} The potential is applied at the all-electron stage of the pseudopotential generation, within LDA and the norm-conserving scheme. Our modified pseudopotentials are generated to yield direct experimental band gaps of 3.40 eV for GaN and 0.75 eV for InN, using a plane-wave cutoff of 60 Ry. They yield theoretical lattice parameters of $a=3.22$ Å, $c/a=1.623$, and $u=0.377$ for GaN, and $a=3.58$ Å, $c/a=1.617$, and $u=0.379$ for InN, within 1% of the experimental values.

We have carried out extensive tests to ensure that, on the one hand, calculated atomic structures are still reliable and very close to LDA or GGA results, and on the other hand the band structure closely matches experiment. These tests\textsuperscript{36} included a systematic comparison and analysis of atomic structures and band structures generated with LDA, with GGA, and with the modified pseudopotentials; comparisons with the LDA+$U$ method,\textsuperscript{37,38} and comparisons of the electronic structure of neutral point defects with results obtained with the self-interaction and relaxation-corrected (SIRC) pseudopotential method.\textsuperscript{39} Our tests for both GaN and InN showed that our modified pseudopotentials result in very similar bulk, as well as surface structures, and also very similar energetics. This allows us to systematically and consistently use these pseudopotentials for all aspects of the calculations, structural as well as electronic.

The surfaces are simulated using slabs of up to ten double layers of GaN or InN, with a vacuum region of up to 25 Å. The bottom surface of the slab is passivated with frac-
As a check, we calculated surface reconstructions on GaN(0001) and GaN(0001) surfaces and found structures, as well as energies, to be in agreement with published results.\textsuperscript{12–14,40} $\text{N}_2$ molecules and Ga bulk act as boundaries for the chemical potentials; as the Ga chemical potential increases, the (001) surface undergoes a transition from a 2\times2 N-adatom to a 2\times2 Ga-adatom reconstruction [2\times2 Ga\textsubscript{74} (0001)] for moderate Ga/N ratios, and then to a laterally contracted Ga-bilayer structure at highly Ga-rich conditions. The (0001) surface evolves from a 2\times2 Ga-adatom reconstruction [2\times2 Ga\textsubscript{73} (0001)] to a 1\times1 Ga-adlayer reconstruction [1\times1 In\textsubscript{ato} (0001)] with increasing Ga/N ratio.

For InN(0001) we find a 2\times2 In\textsubscript{74} structure at moderate In/N ratio and a contracted In-bilayer structure at highly In-rich conditions. On the (0001) surface of InN we find the 2\times2 In\textsubscript{73} structure to be unstable compared to the 1\times1 In\textsubscript{ato} (0001) structure.

### A. Polar surfaces—Moderate Ga(In)/N ratios

In the 2\times2 Ga\textsubscript{74} (0001) (2\times2 Ga\textsubscript{73} (0001)) structure, the Ga adatom binds to three surface Ga (N) atoms, and one surface Ga (N) remains threefold coordinated. Band structures for the 2\times2 Ga\textsubscript{74} (0001) and 2\times2 Ga\textsubscript{73} (0001) reconstructions, which are stable at moderate Ga/N ratios, are shown in Fig. 1(a) and Fig. 1(b), respectively. The 2\times2 Ga\textsubscript{74} (0001) structure results in a semiconducting surface, with two sets of surface states appearing within the band gap. By inspecting the electronic charge densities associated with the surface states we have been able to identify their microscopic origins. The upper, empty states \textit{at} $\sim$0.6 eV below the conduction-band minimum (CBM) arise from the dangling bond on the Ga adatom, with some charge density also on nearby sublayer N atoms, and from the dangling bond on the threefold-coordinated Ga surface atom. The lower states at $\sim$1.7 eV above the VBM are fully occupied and are well localized on bonds between the Ga adatom and three Ga surface atoms. The degeneracy occurring at the $\Gamma$ point can be attributed to a mirror symmetry with respect to a plane perpendicular to the surface. The upper, empty state in the 2\times2 Ga\textsubscript{73} (0001) structure is also mainly derived from the Ga-adatom dangling bond, while the lower occupied state corresponds to a nitrogen dangling bond and is close in energy to the VBM.

The 2\times2 Ga\textsubscript{74} (0001) and 2\times2 Ga\textsubscript{73} (0001) reconstructions give rise to states with rather small dispersion, as can be seen in Fig. 1 and from the fairly sharp peaks observed in the density of states (DOS) plotted in Fig. 2(a). These unoccupied and occupied states have an areal density of $\sim6\times10^{14}$ cm$^{-2}$ each. The unoccupied surface states can therefore readily accept electrons taken from a near-surface depletion layer, effectively acting to pin the Fermi level at the surface of $n$-type GaN at $\sim$0.6 eV below the CBM for the (0001) plane, and $\sim$1.2 eV above the VBM for the (0001) plane. Our value for Fermi-level pinning on $n$-type GaN(0001) ($E_F-E_v=2.8$ eV) compares very well with the experimental data of Long \textit{et al.}\textsuperscript{16} ($E_F-E_v=2.7$ eV, in the limit of low doping concentration), as well as with the results of Koçan \textit{et al.}\textsuperscript{2} (2.89 eV) and of Wu \textit{et al.}\textsuperscript{17} (2.6 eV).

Our value of $\sim$1.2 eV for GaN(0001) seems to be in disagreement with the data of Kowalski \textit{et al.},\textsuperscript{24} who found...
essentially no Fermi-level pinning on the surface of n-type GaN(0001). However, it should be noted that Kowalski et al. observed nondispersive occupied states in the vicinity of the VBM and at about 1.2 eV above the VBM on the surface of a 1 × 1 reconstructed GaN(0001) surface. These energies are very similar to those of the two surface states observed for our 2×2 GaN(0001) reconstructed surface shown in Fig. 1(b). The seeming disagreement between the periodicities of the surface reconstructions may arise from disorder, attributed to the adatom occurring in different positions within the 2×2 unit cells and thus obscuring the 2×2 reconstruction.

Conversely, the occupied surface states induced by the 2×2 GaN(0001) structure can accept holes derived from a depletion layer on p-type GaN, pinning the Fermi level at ~1.7 eV above the VBM. The large difference between the pinning level of n-type GaN and p-type GaN has also been observed experimentally, albeit with somewhat smaller values of 1.3 eV (Ref. 16) and 1.0 eV. In contrast, little or no band bending is expected for the 2×2 GaN(0001) structure on p-type GaN since the N-derived occupied states are almost resonant with the valence band. This result differs from the measured Fermi-level pinning of 1.6 eV above the VBM for p-type GaN(0001) found by Ryan et al. However, the results can be reconciled if we assume a relatively high Ga coverage was present on the sample surface in the experiments of Ref. 25, as explained in Sec. IV C.

The electronic band structures of the 2×2 InN(0001) and 1×1 InN(0001) reconstructed surfaces of InN found at moderate In/N ratios are shown in Figs. 1(c) and 1(d). Two sets of surface states appear for the 2×2 InN(0001) structure, with similar character as in the case of GaN. However, InN has a very small band gap and large electron affinity, resulting in a low energetic position of the CBM at the Γ point. As a consequence, both sets of surface states occur at energies above the CBM, as shown in the DOS plot in Fig. 2(b). The presence of the occupied In-In bond states above the CBM provides an immediate explanation for the observed electron accumulation on InN polar surfaces. Because the number of surface states is much larger than the number of available bulk states in the near-surface accumulation layer, the surface Fermi-level position is approximately determined by the position of the upper portion of the occupied surface state, which is ~0.6 eV above the CBM. This result compares well with the experimental value of ~0.8 eV. The 1×1 InN(0001) reconstruction, finally, results in relatively dispersive states due to the interaction between neighboring In atoms. The surface is metallic, with a Fermi level at about 0.3 eV above the CBM, as shown in Fig. 1(d).

B. Nonpolar surfaces—Moderate Ga(In)/N ratios

For the nonpolar m and a surfaces, we consistently find that at moderate Ga/N ratios the GaN dimer structure is most stable. The N atom relaxes outward (toward an sp^3 configuration) while the Ga atom relaxes inward (toward sp^2), accompanied by a charge transfer from the Ga dangling bond to the N dangling bond. Consequently, we obtain an unoccupied Ga dangling bond state at ~0.7 eV below the CBM for the m plane [Fig. 3(a)], and two unoccupied Ga dangling bond states at ~0.5 eV below the CBM for the a plane [Fig. 3(b)] (the unit cell of the a plane is twice as large as that of the m plane). The energy level of the Ga-derived dangling bond is very nearly the same as found for the polar (0001) surface, resulting in similar Fermi-level pinning on n-type GaN [see Fig. 4(a)]. In contrast, the occupied surface states behave quite differently on nonpolar versus polar (0001) surfaces: on the m and a planes; these states are associated with dangling bonds on the N atom and overlap with the valence band, i.e., they do not create levels within...
the band gap, as shown in Figs. 3(a), 3(b), and 4(a).

Similar results apply in the case of the InN m and a planes, i.e., the occupied surface states are close to or below the VBM, as shown in Figs. 3(c), 3(d), and 4(b). Therefore, for moderate In/N ratios we predict an absence of electron accumulation on the nonpolar InN(1100) and (1120) surfaces, in contrast with the polar surfaces. The unoccupied surface states associated with In dangling bonds are well above the CBM, with no discernible effect on the electronic properties of the surface.

C. Polar surfaces–High Ga(In)/N ratios

Under highly Ga-rich conditions, the (0001) polar surface is covered by a laterally contracted double layer of Ga. At this high coverage, the Ga-Ga bonding and dangling-bond states that were distinct at moderate Ga/N ratios now strongly interact, leading to a large energy dispersion within the band gap. This is reflected in the DOS shown in Fig. 5(a). The cations in the bilayer form relatively strong metallic bonds within the plane and between the two adlayers, and the Fermi level is located at ~1.8 eV above the VBM, close to the experimental value of 1.7 eV reported in Ref. 2. Inspection of the charge density distribution for states around the Fermi level shows them to be localized on Ga-Ga bonds within the underlying Ga adlayer, with some admixture of dangling bonds on the upper adlayer. The 1×1 Ga_{ad} (0001) structure involves a Ga adlayer that binds to the uppermost N layer of the GaN surface, with a Fermi level at ~1.6 eV above the VBM. This compares favorably with the measured Fermi-level pinning position of 1.6 eV above the VBM for p-type GaN(0001) found by Ryan et al., if we assume a relatively high Ga coverage on the surface of their samples.

Similar physics applies in the case of the In bilayer on InN (0001) [Fig. 5(b)], where we find the Fermi level at ~0.7 eV above the CBM, very close to the 0.6 eV found for moderate In/N ratios. For the 1×1 In_{ad} (0001) structure the Fermi level is ~0.3 eV above the CBM, again similar to the result for moderate In/N ratios.

D. Nonpolar surfaces–High Ga(In)/N ratios

For highly Ga-rich conditions we identified a number of reconstructions that have not previously been reported; details will be published elsewhere. For the purposes of the present article, what matters is that these surface reconstructions involve metallic adlayers, and therefore the electronic structure is again characterized by highly dispersive bands crossing the Fermi level at ~1.8 eV above the VBM.

In the case of InN, we find that the Fermi level is located about 0.6 eV above the CBM, as for the In bilayer on the (0001) polar surface. Nonpolar surfaces of InN therefore exhibit electron accumulation when In adlayers are present on the surface. However, it may be possible to remove such adlayers in postgrowth processing, resulting in a surface consisting purely of In-N dimers, for which we found an absence of electron accumulation.

V. SUMMARY

We have performed a systematic and comprehensive computational study of reconstructed GaN and InN surfaces in various orientations, including (1010) (m plane), (1120) (a plane), as well as the polar (0001) (+c) and (0001) (−c) planes. The calculations were based on density-functional theory and an extensively tested approach for correcting the band-gap error through use of modified pseudopotentials, enabling us to provide a realistic prediction of the energetic position of surface states based on self-consistent and fully relaxed atomic structures. Our calculations for GaN allowed us to identify the microscopic origins of Fermi-level pinning that lead to depletion layers on the surface of n-type and p-type material. For InN we found that on polar surfaces all the surface states are located above the CBM. Fermi-level pinning occurs due to occupied surface states above the CBM, for all In/N ratios, thus explaining the observed electron accumulation. Our studies predict the absence of electron accumulation on nonpolar surfaces of InN exhibiting dimer reconstructions. Detailed comparisons with experiment, where available, were provided.

ACKNOWLEDGMENTS

This work was supported in part by ONR under Contract No. N00014–02-C-0433, monitored by C. Baatar, through a subcontract from the Palo Alto Research Center Inc., and by the NSF MRSEC Program under Award No. DMR05–20415. It made use of the CNSI Computing Facility under NSF Grant No. CHE-0321368. We are grateful to A. Janotti, J. Speck, G. Kohlmueller, and C. Gallinat for useful discussions.