TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY
Julkaisu 772 • Publication 772

Hannu-Pekka Komsa

First-Principles Study of Nitrogen in GaAsN – Defects and Interfaces

Tampere 2008

Tampereen teknillinen yliopisto. Julkaisu 772
Tampere University of Technology. Publication 772

Hannu-Pekka Komsa

First-Principles Study of Nitrogen in GaAsN – Defects and Interfaces

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Sähkötalo Building, Auditorium S3, at Tampere University of Technology, on the 28th of November 2008, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2008

ISBN 978-952-15-2068-6 (printed)
ISBN 978-952-15-2111-9 (PDF)
ISSN 1459-2045

Abstract

Semiconductors are an essential part of the modern society. Transistors, solid-state lasers, solar cells, and several other semiconductor devices have shaped the world drastically. The most familiar material is perhaps silicon found in all the computer chips. Optoelectronic components, on the other hand, are usually based on materials such as GaAs, which already has a long history. On the contrary, the potential of dilute nitride GaAsN material, which is the subject of this thesis, was only recently discovered.

The rapid progress of semiconductor technology would have been impossible without basic materials research, which in the recent years have ever-increasingly taken the form of computational research (fittingly, executed on a piece of silicon). For new materials, the *first-principles* methods which do not rely on any empirical parameters are important on predicting the unknown material properties reliably. More precisely, we have employed the density-functional theory, which is a good compromise between speed and accuracy.

Various GaAsN material properties show unconventionally strong non-linear dependence as a function of the nitrogen concentration. These properties originate from the notable concentration of nitrogen up to alloy concentrations, although nitrogen is essentially a defect forming a deep resonance state. The purpose of this thesis is to understand nitrogen in various environments typically encountered in practical devices. However, the limitations of the methods need to be understood and carefully controlled in order to obtain meaningful results. Moreover, the interpretation of the computational results, and matching with the experimental results, is not always straightforward. These considerations are then an inseparable part of the discussion.

First, the isolated nitrogen defects in GaAs are discussed. This includes our studies of interstitial nitrogen, where we have found the most likely locations for interstitial nitrogen as a function of Fermi-energy and growth conditions. Nitrogen concentration is then increased up to alloy concentrations, and the properties are investigated relying mostly on the extensive studies presented in the literature. After obtaining understanding of dilute nitrides, we can apply it to more complex systems.

The nitrogen interaction with other material components is studied in the last part of the thesis and comprises a large part of our research: The study of the beryllium doping effect on nitrogen alloying and nitrogen effect on beryllium defects allows us to find the causes for the experimentally observed p-type carrier compensation and observe cancellation of nitrogen-induced band-gap reduction. Our computational studies are used in the interpretation of the Raman spectra of nitrogen-related vibrational modes in InGaAsN. Finally, the GaAsN/GaAs and InGaAsN/GaAs interface calculations show a long-range interaction due to

nitrogen and a strong valence-band offset dependence on the local geometry of nitrogen and indium.

In many instances, this thesis displays how seemingly surprising results of more complex systems can be explained by relying only on a few fundamental physical concepts of isolated defects or simple alloys. The role of computational physics is then to find (more) quantitative values and the winner among the competitive situations.

Preface

The introduction part of this thesis might seem a bit lengthy for being just an introduction to the articles, and it is not a comprehensive coverage of any of the subjects considered in this thesis. It mostly tries to explain the concepts that I struggled with during the research. These might be obvious to some, but hopefully also provide a bit of insight to others. In my case, it took about 2–3 years of graduate studies to build a concise picture of the current state of research, the physics in question and the possibilities for my own research. This is not all due to reading papers, but also due to teaching, discussions with supervisors and coworkers, and also supervising others, all of which should not be overlooked. The genuine interest to the topic only woke at about this time, which then feeded the work motivation for the last 1–2 years. I was lucky enough that some of the research that was started just to calculate some data, revealed more interesting results. Thus, I have to thank all the people involved for the enlightened choice of the research topic.

The first words of gratitude are naturally directed towards Professor Tapio Rantala, who were willing to provide me with research topic, motivation, time, and financial support. Second words go to my second supervisor Docent Eero Arola for the additional insight and always digging through all the fine details.

I am grateful for the external examinees Professors Jukka Tulkki and Peter Kratzer, as well as my opponent Prof. Kurt Schroeder, for taking the time to review this thesis on a short notice.

I also have to thank all my colleagues at the g-wing and especially, in a somewhat chronological order: Markku Leino, Ville Arpiainen, Jussi Ojanen, Ilkka Kylänpää, Matti Viitala, Mikael Kuisma, Tommi Kortelainen and Eero Kokkonen. Naturally, I also owe a create deal (of intellectual debt) to my research colleagues in Helsinki University of Technology (Katri Laaksonen), Optoelectronic Reseach Centre (Emil-Mihai Pavelescu, Janne Pakarinen), and University of Turku (Pekka Laukkanen, Marko P. J. Punkkinen).

Finally, there is the constant background support of my family, which is never forgotten. They create a balancing force against the undesirable perturbations in life. In addition, I thank Kyoko for creating many desirable perturbations.

I am grateful for the financial support I have received during these years from several sources: TULE/QUEST project, the Department of Physics, Jenny and Antti Wihuri foundation and Finnish Foundation for Technology Promotion (TES). The computational resources were provided by the IT center for science (CSC) and Akaatti (as a part of M-Grid).

All figures are reprinted with permission from the authors and publisher. All articles are reprinted with permission from the publisher.

List of publications

The thesis consists of the introductory part and the following publications published in refereed journals:

Paper I: E. Arola, J. Ojanen, H.-P. Komsa and T. T. Rantala, *Atomic and electronic structures of N interstitials in GaAs*, Physical Review B 72, 045222 (2005)

Paper II: K. Laaksonen, H.-P. Komsa, T. T. Rantala and R. M. Nieminen, *Nitrogen interstitial defects in GaAs*, Journal of Physics: Condensed Matter 20, 235231 (2008)

Paper III: K. Laaksonen, H.-P. Komsa, E. Arola, T. T. Rantala and R. M. Nieminen, *Computational study of $GaAs_{1-x}N_x$ and $GaN_{1-y}As_y$ alloys and arsenic impurities in GaN*, Journal of Physics: Condensed Matter 18, 10097 (2006)

Paper IV: H.-P. Komsa, E. Arola, E. Larkins and T. T. Rantala, *Band offset determination of the GaAs/GaAsN interface using the DFT method*, Journal of Physics: Condensed Matter 20, 315004 (2008)

Paper V: H.-P. Komsa, E. Arola and T. T. Rantala, *Band offset of InGaAs(N)/GaAs superlattice interfaces from first principles*, Applied Physics Letters 92, 262101 (2008)

Paper VI: E.-M. Pavelescu, J. Wagner, H.-P. Komsa, T. T. Rantala, M. Dumitrescu and M. Pessa, *Nitrogen incorporation into GaInNAs lattice-matched to GaAs: the effects of growth temperature and thermal annealing*, Journal of Applied Physics 98, 083524 (2005)

Paper VII: H.-P. Komsa, E. Arola, J. Pakarinen, C. S. Peng and T. T. Rantala, *Beryllium doping of GaAs and GaAsN from first principles*, Submitted to Physical Review B

The author also participated in the work in the following publications not included in this thesis:

M. P. J. Punkkinen, P. Laukkanen, K. Kokko, M. Ropo, M. Ahola-Tuomi, I. J. Väyrynen, H.-P. Komsa, T. T. Rantala, M. Pessa, M. Kuzmin, L. Vitos, J. Kollár and B. Johansson, *Surface core-level shifts of GaAs(100)(2×4) from first principles*, Physical Review B 76, 115334 (2008)

P. Laukkanen, M. P. J. Punkkinen, H.-P. Komsa, M. Ahola-Tuomi, K. Kokko, M. Kuzmin, J. Adell, J. Sadowski, R. E. Perälä, M. Ropo, T. T. Rantala, I. J. Väyrynen, M. Pessa, L. Vitos, J. Kollár, S. Mirbt and B. Johansson, *Anomalous Bismuth-Stabilized (2×1) Reconstructions on GaAs(100) and InP(100) Surfaces*, Physical Review Letters 100, 086101 (2008)

M. P. J. Punkkinen, P. Laukkanen, H.-P. Komsa, M. Ahola-Tuomi, N. Räsänen, K. Kokko, M. Kuzmin, J. Adell, J. Sadowski, R. E. Perälä, M. Ropo, T. T. Rantala, I. J. Väyrynen, M. Pessa, L. Vitos, J. Kollár S. Mirbt and and B. Johansson, *Bismuth-stabilized (2×1) and (2×4) reconstructions on GaAs(100) surfaces: A combined first-principles, photoemission and scanning tunneling microscopy study*, Physical Review B 78, 195304 (2008)

The author's contributions to the papers included in the thesis:

	Research work	Writing
I	VASP part	minor
II	half	half
III	GaAsN part	GaAsN part
IV	major	major
V	major	major
VI	Calculations	Calculations
VII	major	major

Contents

List of Abbreviations

BAC	band-anticrossing
BEEM	ballistic electron emission microscopy
CB	conduction band
CBE	conduction band edge
CBM	conduction band maximum
CBO	conduction band offset
DFT	density-functional theory
DOS	density of states
ER	electroreflectance
GGA	generalized gradient approximation
GW	not an abbreviation, but the product of Green's function G and screened interaction W
KS	Kohn-Sham
LDA	local density approximation
LDOS	local density of states
MBE	molecular beam epitaxy
MOCVD	metalorganic chemical vapour deposition
PAW	projector-augmented wave
PL	photoluminescence
QW	quantum well
SO	spin-Orbit
SOC	spin-Orbit Coupling
VB	valence band
VBE	valence band edge
VBM	valence band maximum
VBO	valence band offset
XAS	x-ray absorption spectroscopy
XC	exchange-correlation

Chapter 1

Introduction

At the end of the 20th century, the apparently simple semiconductor material GaAsN grabbed the attention of many semiconductor physicists. Nitrogen substitution into GaAs was previously thought to be possible only at the doping concentrations. Mostly owing to the advances in the crystal growth techniques, it became possible to incorporate up to few per cent of nitrogen into GaAs. This new material system provided many surprises: especially, the large non-linear reduction of the band gap, and changes in the effective mass, in the mobility, and carrier lifetimes.

New materials always find their application. The small band gap lends itself well to the long-wavelength optoelectronic applications such as lasers, photodetectors, semiconductor saturable absorber mirrors, and solar cells, and also the base material in heterojunction bipolar transistors. At the same, other properties of GaAsN might prove beneficial to some of these applications and harmful to others.

More interesting for the basic research, even after the gush of research on the material system, GaAsN still provides surprises at the microscopic level. That is, the topic of my research on a more abstract level is the application of the knowledge of the fundamental properties of the material on the various properties of the more complex systems. This, in its part, also bridges the gap between the basic research and the applications.

The topic is mostly treated from the perspective of computational (and theoretical) physics. Therefore, the methods and concepts related to the density functional theory and its application to the systems in question are first discussed in chapter 2. The elemental properties of GaAsN, from the defects to alloys are discussed in chapter 3. Finally, chapter 4 contains the application to "complex" systems, and in my case the InGaAsN properties, (In)GaAsN/GaAs interfaces and beryllium defects in GaAsN.

Chapter 2

Computational methods and models

The methods and models needed for the computational study of the systems of interest are covered here briefly concerning the electronic structure methods (for more details, see e.g. Kohanoff [1], Thijssen [2], Payne et al. [3]). The defect and the band offset calculation methodology is covered more thoroughly.

2.1 Electronic Structure Calculations

In this thesis, we will be mostly interested in the total energies of the systems consisting of frustratingly many particles. This can be done by solving the Schrödinger equation (the time-independent form within the Born-Oppenheimer approximation):

$$\hat{H}\Psi(\mathbf{r};\mathbf{R}) = E\Psi(\mathbf{r};\mathbf{R}). \tag{2.1}$$

The total energy E of a system $\hat{H}$ of N electrons at positions $\{\mathbf{r}\}$ and P nuclei at $\{\mathbf{R}\}$, consists generally of the kinetic energies and the mutual interactions of the particles leading to a Hamiltonian of the form (in Hartree atomic units)

$$\begin{aligned} \hat{H} &= -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{I=1}^{P}\sum_{i=1}^{N}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &= \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}}, \end{aligned} \tag{2.2}$$

$$\tag{2.3}$$

where the three energy terms are for the electron kinetic, attractive electron-nucleus and repulsive electron-electron energies, respectively. The electron-nucleus interactions can be considered as forming an external potential, since the nuclei are fixed in the Born-Oppenheimer approximation.

For the lowest energy solution, the ground-state solution, instead of solving the above eigenvalue problem, it is also possible to try to find the wavefunction Ψ giving the lowest energy E from the equation

$$E = \langle\Psi|\hat{H}|\Psi\rangle = E[\Psi] \tag{2.4}$$

The energy functional $E[\Psi]$ in the last form simply gives the total energy for a given wavefunction. Since this functional is known, the optimization process becomes straightforward, in principle.

2.1.1 Density Functional Theory

The total wavefunction Ψ contains all the information about the system, whether it be in the ground state or in the excited state. In the heart of the density functional theory is a theorem, which states that the ground state electron density also contains all the information about the system. Instead of having a functional that gives the total energy for a given wavefunction, we now need a functional that gives the total (ground state) energy and is minimized for the ground state density.

More precisely, the Hohenberg-Kohn theorems [4] state that:

Theorem 1 *The external potential is univocally determined by the electronic density, besides a trivial additive constant. Corollary: There* exists *a one-to-one mapping between the ground state electron density and the ground state wavefunction (of a many-particle system).*

Theorem 2 *The ground state density minimizes the total electronic energy of the system.*

Hohenberg-Kohn theorems only state that the total energy functional exists, but nothing about its form. Writing the energy functional for interacting electrons has proven to be difficult. Instead of considering electrons experiencing the interactions with other electrons along with some external potential, one could alternatively consider non-interacting electrons experiencing an effective potential, which is constructed in the spirit of mean-field theories. The difficult many-body effects are now included in this one-particle potential, which can be approximated successfully with a relative ease.

Kohn and Sham [5] were the first to tackle the problem this way. More precisely, the total energy functional is

$$E[\rho] = T_{\mathrm{s}}[\psi] + E_{\mathrm{ext}}[\rho] + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho] \tag{2.5}$$

where

$$T_{\mathrm{s}}[\psi] = \sum_{i=1}^{N} \langle \psi_i | -\frac{1}{2}\nabla_i^2 | \psi_i \rangle \tag{2.6}$$

$$E_{\mathrm{ext}}[\rho] = \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r}. \tag{2.7}$$

$$E_{\mathrm{H}}[\rho] = \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.8}$$

The first term is the kinetic energy of the *noninteracting* electron gas, the second term is the interaction with the external potential, usually the nuclei, the third term is the mean-field like Coulomb interaction of the electrons (called the Hartree term), and the last exchange-correlation term contains everything that is not already included in the other terms and its exact form is generally unknown.

Taking the variation with respect to $\langle\psi_i|$ and defining the electron density as

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2, \tag{2.9}$$

we get the Kohn-Sham (KS) equations for the single-particle wavefunctions

$$\left\{-\frac{1}{2}\nabla_1^2 + v(\mathbf{r}_1) - \int \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\rho(\mathbf{r}_2)d\mathbf{r}_2 + V_{\mathrm{xc}}(\mathbf{r}_1)\right\}\psi_i(\mathbf{r}_1) = \epsilon_i\psi_i(\mathbf{r}_1) \tag{2.10}$$

where

$$V_{\mathrm{xc}}(\mathbf{r}) = \frac{\partial E_{\mathrm{xc}}[\rho]}{\partial \rho}. \tag{2.11}$$

Notice that $\{\psi_i\}$ are the wavefunctions of the non-interacting electrons in an effective potential (also called the Kohn-Sham states or orbitals) and have no strict physical meaning. In practice however, observables calculated from the KS states are close to the experimental values. The total energy can also be written with the aid of the one-electron energies ϵ_i. This construction requires the self-consistent field (SCF) solution, where the potential is solved from an initial guess for the electron density (or, in practice, effective potential), which is consequently used to solve the KS-equations and eventually the new electron density (effective potential).

We are left with one more task: finding the XC-energy term, or good approximation for it. As all the unknown many-body interactions are dumped here, it can also depend on everything. Especially, it is non-local. Formulating non-local XC-functionals and calculating with them is painful, and therefore (semi-)local approximations are almost always used. The simplest, the most important, and the one used all throughout this work is the local density approximation (LDA) [6], where the XC-energy locally is estimated to be the XC-energy of a homogeneous interacting electron gas with the same density as in that point:

$$E_{\mathrm{XC}}[\rho] = \int \rho(\mathbf{r})\epsilon_{\mathrm{xc}}(\rho(\mathbf{r}))d\mathbf{r} \tag{2.12}$$

As a next step, some of the non-locality can be included by making the ϵ_{xc} depend also on the gradient(s) of the electron density. These generalized gradient approximations (GGA) give better results in some systems and worse in others.

The choice of LDA has some important implications (or restrictions). GGA is generally better for molecules as the electron density is far from homogeneous. On the other hand, LDA seems to work better for metals and semiconductors and has the added advantage that there is only one LDA. The DFT-LDA method gives the band gap roughly half of the experimental band gap. Usually, even in the case of the *exact* DFT for finite systems, only the highest occupied eigenvalue has a rigorous physical meaning: it is the first ionization energy [7, 8]. However, according to Perdew and Levy [9], for non-metallic *crystals* exact DFT gives correctly the middle of the band gap $-\frac{1}{2}(I+A)$ (where I and A are the ionization energy and electron affinity, respectively). Although LDA is an approximation, this still gives a point of reference.

Later in this thesis, a term "electrostatic potential" is often used, which refers to $V_{\mathrm{ext}}[\rho] + V_{\mathrm{H}}[\rho]$ with the corresponding energy term $E_{\mathrm{ext}}[\rho] + E_{\mathrm{H}}[\rho]$. It is worth pointing out that in Eq. 2.10, adding a constant to the electrostatic

potential shifts the one-electron energies by the same amount, although the corresponding change in the total energy depends on the electron density in Eq. 2.5 and especially Eq. 2.7. If a constant is added to the electrostatic potential, this constant is also transferred to the single-particle energies.

The DFT is a ground-state calculation meaning that it will find the 0 K solution (without the nuclei zero-point motion). In this case, the Fermi-Dirac distribution becomes a step-function with the electron chemical potential usually defined as the limit $\mu_e(0) = \lim_{T\to 0} \mu_e(T)$. Fermi-energy is often defined as the energy of the highest occupied state at 0 K, which would correspond to the valence-band maximum (VBM) in intrinsic semiconductors. Unfortunately, in semiconductor physics, the concepts of Fermi-energy and the electron chemical potential are used interchangeably, and this (minor) sin is also committed in this thesis.

2.1.2 Basis sets and core electron description

The numerical solution of the KS-equations requires attention to some practical details: everything has to be discretized for a numerical solution, and further approximations are often needed.

Representation of all the wavefunctions, potentials, and densities in the real-space grids is certainly possible, but in the case of a periodic crystal the plane-wave basis set is often a more efficient choice. Even a rather small set is often sufficient for a good description of the delocalized electrons such as found in the valence states of crystals. Moreover, for the parts where a real-space presentation is needed (e.g. XC-energy evaluation), the fast Fourier transform provides efficient transformation. On the other hand, the plane-waves are not efficient at describing the localized structure of the wavefunctions near the atomic cores.

The electrons close to the nuclei are usually so tightly confined that changes in the surroundings have negligible effect on the corresponding one-electron wave-functions. Therefore, we could "freeze" the core electrons and include them to the external potential along with the nuclei. Calculations can now be performed for a smaller number of electrons. Moreover, this eliminates the localized structure of the wavefunctions which was problematic in the plane-wave basis. Earlier, this was mostly done with the "pseudopotential" approach, in which a smooth atomic potential is searched for, giving for the valence electrons the same scattering properties, wavefunction outside the core, and/or eigenenergies as the true potential.

In my work, I have used the projector-augmented wave (PAW) scheme, which has been gaining popularity lately as it should give the all-electron calculation accuracy with the computational cost of a pseudo-potential calculation. In PAW formalism, there is a linear transformation that maps the true wavefunctions with the complete nodal structure onto smooth auxiliary wavefunctions. Basically, the transformation, which is defined only inside a sphere near the core, removes the orthogonality with the core electrons. In this way, the nodal structure can also be evaluated at a much finer grid at the core region, even if the auxiliary wavefunctions are smooth. Again, this is advantageous for the plane-wave basis. Only valence electrons are explicitly included in the calculations, but whenever observables need to be computed, the full nodal structure can be obtained by applying the (inverse) transformation. In practice, one has to be careful in deciding which electrons are frozen and which are explicitly included

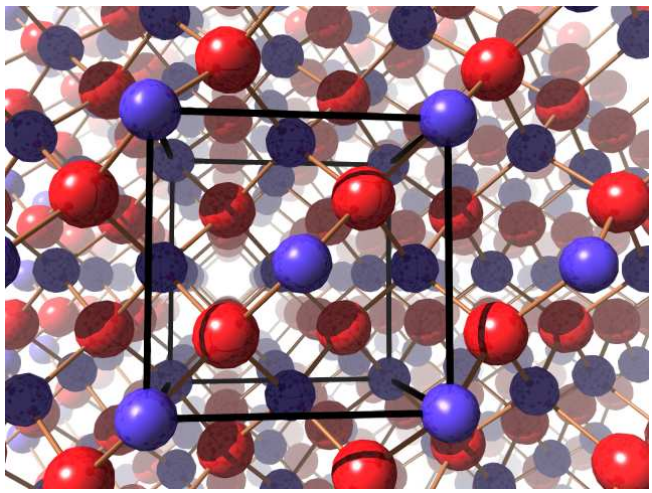

Figure 2.1: A zinc-blende crystal and a unit cell.

in the calculations as this can affect the calculated properties.

All calculations in this thesis are performed using the VASP code [10, 11, 12].

2.2 Models

2.2.1 Atomic Structure

The GaAs, like most of the III-V compound semiconductors, has the zinc-blende crystal structure, which is shown in Figure 2.1. In this thesis, I will not concentrate on the bulk properties of pure materials, but the properties of dilute alloys, defects, and some nanostructures of dilute alloys. Thus, large computational supercells are needed. Compromise between the accuracy of the method and the number of atoms is always inevitable. With the first-principles methods employed in this thesis, the maximum number of atoms at the time was around 200–300 atoms. This equals to a supercell of only about $3 \times 3 \times 3$ unit cells or a cube of about 15 Å per side for GaAs. In the alloy calculations, this corresponds to minimum composition of about 1%.

For a quick overview of some of the most important defect types, a two-dimensional illustration is shown in figure 2.2. With the above-mentioned supercell sizes, this corresponds to a minimum defect concentration of about 2×10^{20} cm^{-3}; a very high concentration.

The same limitations apply to calculations of interfaces. A long supercell is created with one side of the cell being material A and the other side being material B. Within the supercell there are two A/B interfaces (they can be equivalent or inequivalent) and again, in total there will be an infinite amount of interfaces periodically. Using 256 atoms, gives us supercells consisting of $2 \times 2 \times 8$ or $1 \times 1 \times 32$ unit cells, The total supercell lengths with GaAs lattice constant are then about 40 Å and 160 Å, respectively, and half of them for each material. The latter is of the same size as the thinnest quantum wells (8 nm). Naturally, also the widths of these supercells present limitations. Luckily, some

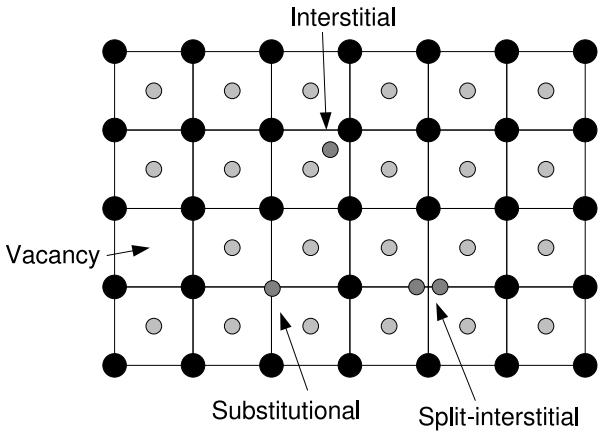

Figure 2.2: A two-dimensional crystal illustrating some of the most important defect types.

of the interface properties are so local, that they can be correctly extracted even from a very short supercell. In general, only the interfaces (not quantum wells) are studied in this thesis, in which case it is sufficient to have the interfaces far enough to be "uncoupled".

2.2.2 Global energy reference and k-points

The periodicity affects the computational framework in two major ways. First, the solutions of the Schrödinger equation in the periodic potential can be shown to be of the form:

$$\psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) \exp(-i\mathbf{k} \cdot \mathbf{r}) \tag{2.13}$$

where $u_{\mathbf{k}}(\mathbf{r})$ is the wavefunction in the unit cell, possessing the lattice translational periodicity, and $\mathbf{k}$ is the wave-vector (or the "crystal momentum"). Although the number of $\mathbf{k}$-vectors is still basically infinite (or quasi-continuous), now that they are confined to a finite region in $\mathbf{k}$-space, we can also discretize it.

Second, there is no global energy reference. As mentioned, the total energy is often all we need, which leads us to trouble. Usually, when estimating the Hartree-energy E_{H} in Eq. 2.8, the Coulomb interaction should be summed over all the infinite number of unit cells. For an efficient implementation, Ewald summation scheme is often employed. However, if the unit cell has net dipole, the sum is conditionally convergent i.e., depending on the order of terms in the sum, the result can be anything. Physically, this can be related to the fact that infinite crystal has no surface. For a polar crystal, surface has a charge density and subsequently the shape of the surface governs the absolute value of Hartree-potential inside the crystal. When the summation scheme is fixed, the "shape of the surface" is fixed, and the energies can be compared, even if there is no physically meaningful reference for the energies. That is, the energy differences are reliable.

2.2.3 Chemical Potentials

Energy differences are not reliable if the number of particles differ, because in this case the energy difference would contain absolute energy of the particles. Also, in our calculations, matter does not disappear or spawn from emptiness. Still, there are often different number of particles in the supercells and, for the comparison of energies, we need to obtain the energy for the excess/missing atoms. One could calculate the energy of an atom in vacuum, energy per atom in a molecule or in bulk or something else. In order to model the particle exchange in a more general framework, we have to get acquainted with the concept of chemical potential [13]. Assuming constant temperature T, pressure P, and number of other particle species $N_j (j \neq i)$, the chemical potential is defined as

$$\mu_i \equiv \mu_i(T, P, N_1, N_2, \ldots, N_s) = \left(\frac{\partial G}{\partial N_i} \right)_{T,P,N_j(j \neq i)}, \tag{2.14}$$

where G is the Gibbs free energy. In short: with a change in the number of particles in a system, the chemical potential is the corresponding change in the Gibbs free energy (per particle). All the mechanisms and processes that may be related to the change in the particle number are incorporated into the chemical potential; e.g. changes in volume, entropy, or electric field.

Alternative formulations in different conditions are:

$$\mu_i = \left(\frac{\partial U}{\partial N_i} \right)_{S,V,N_j(j \neq i)} = \left(\frac{\partial A}{\partial N_i} \right)_{T,V,N_j(j \neq i)}, \tag{2.15}$$

where U is the internal energy and $A = U - TS$ is the Helmholtz free energy.

The free energy obtains its minimum value in thermodynamic equilibrium. Equilibrium means that the reactions in which the particle numbers change occur at the same rate in both directions, so that the particle numbers stay constant. From here it follows that $\sum_i n_i \mu_i = 0$, and for example $\mu_{N_2} = \mu_N + \mu_N$ or $\mu_{\mathrm{GaAs}} = \mu_{\mathrm{Ga}} + \mu_{\mathrm{As}}$. In essence, the total system can be divided into smaller and smaller subsystems, but the chemical potential of particle specie should always be the same. Even though the particles are mostly atoms and molecules in this chapter, they can be also electrons, in which case the term "electron chemical potential" is also used.

Out of equilibrium, if the number of particles of one type is very low it is likely that the reactions would lead to generation of these particles. This corresponds to a low chemical potential for the said particles. On the other hand, the reactions on the opposite direction are very unlikely simply because there are not enough particles. The chemical potential pressure dependence for ideal gas (in constant temperature) can be derived from the basic thermodynamics [13], giving

$$\mu = \mu_0 + kT \ln \left(\frac{p}{p_0} \right) \tag{2.16}$$

where μ_0 and p_0 refer to the standard state, i.e. 1 bar. Hence, $\mu = -\infty$ would correspond to complete lack of particles.

For visualization purposes and to keep the chemical potential constant, the particle exchange is often considered with a reservoir. Often, there are many processes built in to the chemical potential and modeling this reservoir is impossible or at least very difficult. For example, for a nitrogen in GaAsN, are the

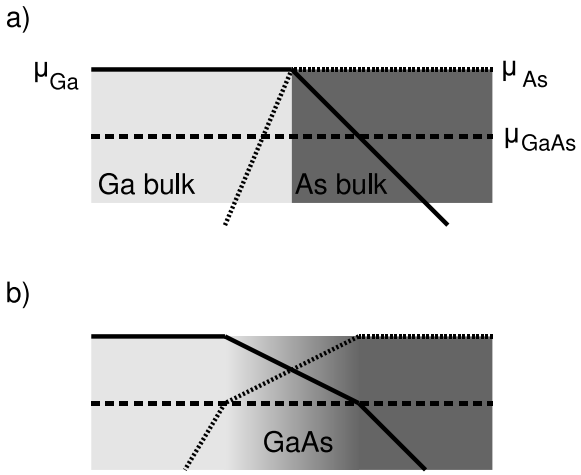

Figure 2.3: The chemical potentials of gallium, arsenic and GaAs for a "pure" interface of gallium and arsenic (a) and after formation of GaAs on the interface (b).

nitrogen atoms exchanged with GaN precipitates, N_2 gas outside the sample or simply with the nitrogen at some location inside the bulk? Most likely a little bit of all of them and modified with the processes needed to move the nitrogen in the lattice and so on. In order to find the correct value for the energy of an atom, the complete system with its environment would need to be modeled; in which case, we wouldn't need the concept of chemical potential in the first place. The inherent problems in tying together the statistical thermodynamical variables with QM calculations might seem pointless, but the advantage is that the chemical potential can be taken as a variable having values within some limits that correspond to well-defined cases.

Let's consider an (artificial) example of formation of GaAs by joining together bulk pieces of gallium and arsenic. The lowest energy phases for the gallium and arsenic are the metallic bulk phases. The maxima for the μ_{Ga} and μ_{As} are then given by $\mu_{Ga}^{max} = \mu_{Ga[bulk]}$ and $\mu_{As}^{max} = \mu_{As[bulk]}$. The gallium or arsenic atom in a phase with higher energy should eventually be driven into the phase of lowest energy; that is why we choose it, although, in MBE, arsenic could come to the surface as As_4 molecule, giving perhaps a better choice for the maximum. The GaAs formation from the bulk constituents is exothermic.

$$E_{tot}[GaAs] = \mu_{Ga[bulk]} + \mu_{As[bulk]} + \Delta H_f[GaAs]. \tag{2.17}$$

The $\Delta H_f[GaAs]$ is the "heat of formation", which is a useful concept also because it can be easily both measured and calculated. The Figure 2.3 shows the situation soon after joining the bulks. Some gallium is diffusing into the arsenic

side which results in a finite chemical potential, but smaller than that of gallium bulk. The same is happening for the arsenic. The sum of the gallium and arsenic chemical potentials near the interface is higher than that of GaAs. GaAs starts forming and keeps forming until the situation of Figure 2.3b is reached. Outside the GaAs region, there is not enough gallium or arsenic to create any more GaAs: $\mu_{\mathrm{Ga}} + \mu_{\mathrm{As}} < \mu_{\mathrm{GaAs}}$. The GaAs region is in equilibrium, but it is not pure GaAs. Closer to gallium bulk, the GaAs has more and more gallium antisites and arsenic vacancies until the material becomes gallium bulk. From this example, we see that it is possible to also find the lower limits for μ_{Ga} and μ_{As}, in equilibrium with GaAs, as

$$\begin{aligned} \mu_{Ga}^{min} &= E_{tot}[GaAs] - \mu_{As[bulk]} \\ \mu_{As}^{min} &= E_{tot}[GaAs] - \mu_{Ga[bulk]}. \end{aligned} \tag{2.18}$$

This equation can also be understood in the following way: if $\mu_{Ga} + \mu_{As[bulk]}$ is less than $E_{tot}[GaAs]$, then GaAs will not form, as it is not the lowest energy phase.

The above example is unrealistic in itself, but for example when annealing GaAs sample in the presence of high arsenic pressure, the pressure can be related to arsenic chemical potential. In calculations, it might be better to consider bulk, surface and vacuum as separate phases. When considering the growth of semiconductors, it is often the surface processes during growth which determine the final atomic structure.

For nitrogen incorporation to GaAs, if we take the equilibrium with GaN phase, we get a nitrogen chemical potential so low that the nitrogen solubility to GaAsN alloy would be less than 10^{14} cm^{-3}, about 8 orders of magnitude smaller than experimentally achieved. Zhang and Wei [14] showed that substituting nitrogen in to the surface sites is energetically favorable with respect to bulk sites, which would explain the higher nitrogen concentration. In this case, the equilibrium is with N_2 gas and the defect formation energy is also smaller. Moreover, not only atoms, but also nitrogen molecules start get substituted into the lattice.

One more complication to this is that MBE growth is really not in equilibrium. MBE growth proceeds at a rather low temperature ($\approx$ 500–800 °C) at which the surface defect mobility is not that high. However, often the thermal annealing after the growth brings the sample quite close to equilibrium. In MOCVD growth equilibrium is better satisfied as the temperatures are above 1000 °C.

2.3 Defects

The defects, whether intentional or unintentional, have a dominant role in most electrical and also optical properties of semiconductors. The rich varierty of defects along with their technological importance still guarantee interest to their computational research.

First, it is important to make the distinction between the “shallow” and the “deep” defects. A good overview can be found in Ref. 15, but in short: Shallow defects can be described well with the effective mass theory in which the microscopic details can be forgotten so that both the potential and the

wavefunction envelope change relatively little at the unit-cell size scale. Deep defects have strong potential within the unit-cell size scale and subsequently the wavefunction is localized. Often, shallow defect levels are close to band edges (within few tens of meV) and deep levels closer to the midgap, but not always, and this should not be used as a definition. Also, defining them through the localization seems less ambiguous especially since many defect properties depend more on the localization of the states than on the energetic position. The large potential of deep defects can be caused for example by broken bonds, strain or difference in electronegativity. We have concentrated on the deep defects, since they are less studied and more suited for the DFT within periodic supercells. It is also worth mentioning that usually deep defects have minor contribution on the electrical properties, but they are important as recombination centers (with mostly non-radiative mechanism).

2.3.1 Formation energy

The methods presented here are also comprehensively reviewed by Van de Walle and Neugebauer [16].

The defect concentration in a thermodynamic equilibrium at a temperature T is

$$\begin{aligned} c &= N_{\text{sites}} \exp(-G^f/kT) \\ &= N_{\text{sites}} N_{\text{config}} \exp(-E^f/kT) \end{aligned} \tag{2.19}$$

where E^f is the formation energy (or enthalpy), N_{sites} is the number of possible sites for the defect, and N_{config} is the number of atomic configurations per site (e.g. orientation). The latter form simply follows from $G = H - TS$ and $S = k\ \ln\Omega$.

The formation energy of a neutral defect X is

$$E^f[X] = E_{\text{tot}}[X] - E_{\text{tot}}[\text{host, bulk}] - \sum_i n_i \mu_i \tag{2.20}$$

where $E_{\text{tot}}[\text{host, bulk}]$ is the energy of the supercell of the host bulk, $E_{\text{tot}}[X^q]$ is the energy of the supercell with the defect included, n_i gives the number of atoms added to the supercell ($n_i > 0$) or removed from the supercell ($n_i < 0$) when creating the defect, and μ_i are the corresponding chemical potentials. The energy comparison is now stoichiometric, as needed.

The formation energy of a defect in a charge state q, on the other hand is

$$\begin{aligned} E^f[X^q] = E_{\text{tot}}[X^q] - E_{\text{tot}}[\text{host, bulk}] \\ - \sum_i n_i \mu_i + q[E_F + E_v + \Delta V] \end{aligned} \tag{2.21}$$

In the last term, electron exchange with the electron reservoir, which is (on average) at the Fermi-energy E_F, is taken into account. Fermi-energy is given with respect to the VBM. It could also be written with the help of electron chemical potential $\mu_e = [E_F + E_v + \Delta V]$.

In the above, the VBM energy in the energy reference of the defect calculation is required. This is handled by the so-called "potential correction" term

ΔV in Eq. 2.21. It is calculated as the difference of the electrostatic potential in the host ($\overline{V}_H$)) and defect ($\overline{V}_D$) supercells far from the defect

$$E_v = \epsilon_{\mathrm{VBM,D}} = \epsilon_{\mathrm{VBM,H}} + (\overline{V}_D - \overline{V}_H) = \epsilon_{\mathrm{VBM,H}} + \Delta V \tag{2.22}$$

where $\epsilon_{\mathrm{VBM,D}}$ and $\epsilon_{\mathrm{VBM,H}}$ are the valence-band maxima of the defect and host supercells, respectively. For illustration see Figure 2 in Ref. 17. The advantages and disadvantages related to this approach are further discussed in section 2.3.3.

2.3.2 Ionization levels

Let's consider in more detail the effects of the last term in Eq. 2.21. The higher the Fermi-energy is, the more energy is gained (on average) by moving the electron from the reservoir in to the defect. That is, the formation energy of a negatively charged defect gets lower at higher values of E_F, and vice versa for the positively charged defects. This is clear also from equation 2.21 when q is negative or positive. If the defect has the lowest formation energy in the negative charge state, then a neutral defect would like to receive an electron from the electron reservoir i.e., it behaves like an acceptor. In a similar fashion, if a neutral defect likes to donate an electron to the electron reservoir, it becomes positively charged and this can be related with a donor.

When the formation energies of all the possible charge states are calculated, it often occurs that the charge state giving the lowest formation energy depends on the Fermi-energy. Consequently, there will be a value of E_F where the formation energy is the same for two charge states, which will be denoted as the ionization level $\epsilon(q_1/q_2)$. This level is the same as the energy required to promote an electron from the valence band into the defect for acceptor levels or demoting an electron from the defect to the valence band for donor levels, as will be shown later.

The atomic structures of the defect in charge states q_1 and q_2 are probably different. If the atomic relaxation occurs during the ionization process, it is called a thermodynamic transition. If not, it is called an optical transition. For example, former transitions are observed in deep-level transient spectroscopy (DLTS) and latter in the photoluminescence (PL) spectroscopy.

In order to see how the ionization level is related to ionization energy, we start by writing the formation energies for the neutral and a singly charged defects

$$\begin{aligned} E^f[X^0] &= E_{\mathrm{tot}}[X^0] - E_{\mathrm{tot}}[\text{host, bulk}] - \sum_i n_i\mu_i + 0[E_F + E_v + \Delta V] \\ E^f[X^-] &= E_{\mathrm{tot}}[X^-] - E_{\mathrm{tot}}[\text{host, bulk}] - \sum_i n_i\mu_i - 1[E_F + E_v + \Delta V] \end{aligned} \tag{2.23}$$

Mark them equal $E^f[X^0] = E^f[X^-]$, a lot of terms cancel out, and we are left with

$$E_{\mathrm{tot}}[X^0] = E_{\mathrm{tot}}[X^-] - [E_F + E_v + \Delta V] \tag{2.24}$$

The Fermi-energy at which this equation holds is called the acceptor level E_A (w.r.t. VBM) and we can write it as

$$E_A = E_{\mathrm{tot}}[X^-] - E_{\mathrm{tot}}[X^0] - [E_v + \Delta V] = \epsilon(0/-) \tag{2.25}$$

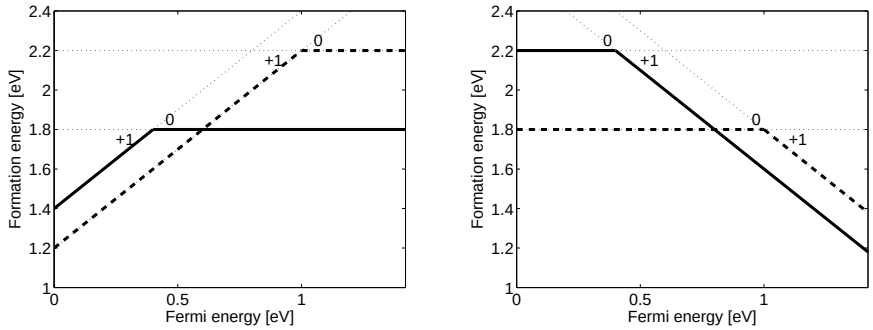

Figure 2.4: Schematic of two examples of deep donor levels (left) and deep acceptor levels (right) in GaAs with $E_g = 1.42$ eV. Dotted lines are showing the formation energies of each charge state and thicker lines the charge state with the lowest formation energy.

This equation describes the energy difference between a system with a negatively charged defect and a system with a neutral defect and an electron at the VBM. Exactly as is the usual definition for the acceptor level.

We can do the same thing for donors by equalling the formation energies of neutral and positively charged defects $E^f[X^0] = E^f[X^+]$, from where we get

$$E_D = E_{\text{tot}}[X^0] - E_{\text{tot}}[X^+] - [E_v + \Delta V] = \epsilon(+/0) \tag{2.26}$$

Here we compare a system with a neutral defect to a system with a positive defect and an electron in the valence band. The standard definition of the donor level is with respect to the CBM

$$E'_D = E_g - E_D = E_{\text{tot}}[X^+] - E_{\text{tot}}[X^0] + [E_c + \Delta V] \tag{2.27}$$

That is, E'_D gives the energy required to move an electron from the defect in to the conduction band.

Typical formation energy diagrams are shown in figure 2.4 for two acceptors and two donors. As calculated from Eq. 2.19, if the Fermi-energy is at the middle of the gap, the donor marked with solid line is dominantly on the neutral charge state and the donor marked with dashed line is on the positive charge state. If, for example, the Fermi-energy were higher, near the $\epsilon(+/0)$ of the dashed donor, then the formation energy for both charge states is roughly equal and they are present at roughly equal concentrations. Naturally, similar reasoning applies to acceptor levels.

There is no reason to be limited only to the neutral and singly occupied charge states. Some defects can accept or donate more electrons or in rare cases do both (amphoteric). An example is shown in figure 2.5 for the GaN intrinsic defects, where various different types can be observed. Gallium vacancy can occur from neutral to up to +3 charge state. The meaning of the $\epsilon(-/2-)$ ionization level is: how much energy is needed to bring an electron from valence band to the defect when it is already singly occupied. It is often higher than the $\epsilon(0/-)$ level, because of the repulsion of the electrons in the defect. If the $\epsilon(-/2-)$ level is below $\epsilon(0/-)$ level, then it would be energetically favored to

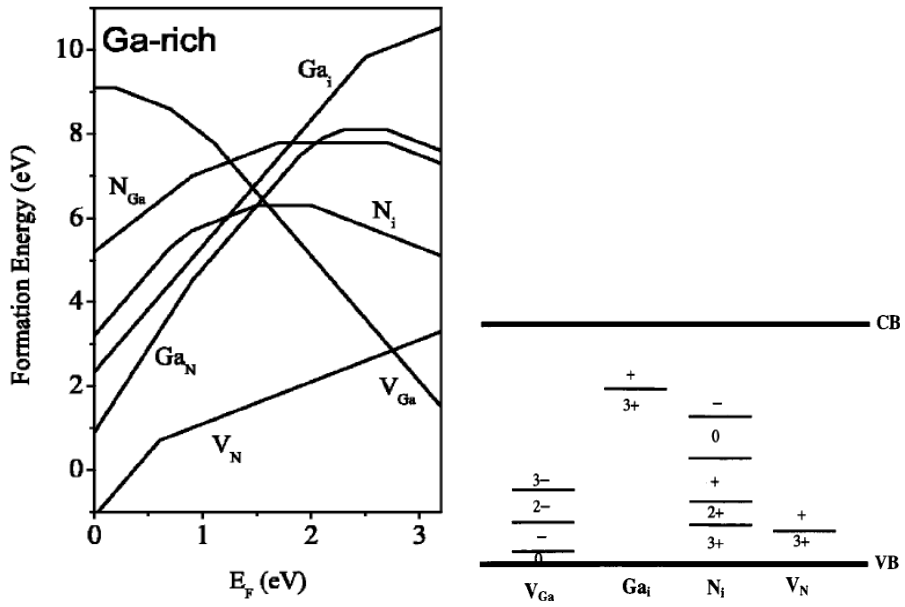

Figure 2.5: Example of formation energy diagram for GaN intrinsic defects on the left and the corresponding ionization levels on the right. From Van de Walle and Neugebauer [16].

always put two electrons into a neutral defect rather than two electrons to two defects (one for each). In the formation energy diagram, there would then be just $\epsilon(0/2-)$ level. This kind of defects are called "negative-U", where U is related to the electron-electron interaction (in the Hubbard model). If it is negative then the electrons have, for some reason, an attractive interaction. Usually this is because of a large geometric changes occurring during the ionization.

Up to this point, the methodology has been totally independent of any numerical details and exact apart for the difficulty of estimating some parameters. Next, the two most severe problems faced during the total energy calculations are discussed: periodic-image interaction and DFT-LDA band-gap problem.

2.3.3 Periodic-image interaction

Ideally, we would like to calculate the formation energy of a single defect in an infinite crystal or in a crystal with some concentration of randomly distributed defects. The former is challenging and the latter is even more challenging. On the other hand, within a periodic supercell approach, it is easy to calculate a formation energy of a periodic array of defects. As the supercell size is increased, it is possible to extrapolate in to the limit of a single defect in an infinite crystal, although this can be quite tedious process. In any case, it is important to understand how the periodic images of the defects interact with each other, and how this affects the formation energies.

There are mainly three aspects that need to be considered:

- Geometric relaxation of defect and/or supercell lattice constant.

- Charge neutrality of the supercell and the Coulomb interaction of the charged defects.
- Overlap of the defect potentials and wavefunctions.

The first aspect is of least concern as the geometric changes do not propagate very far. Moreover, atoms can be locked to the bulk positions at the edges of the supercells. At the experimental defect concentrations the effect on the lattice constant is vanishingly small. At the concentrations used in the supercell calculations, the lattice constant relaxation may have a noticeable effect. In order to model the low concentration of defects in the calculations, the lattice constant is often fixed to that of the host crystal.

We can not simply add an electron to a supercell, since the total charge of the model would be infinite and also the sum of the Coulomb interactions would diverge. Therefore, the supercell is made charge neutral by adding a homogeneous background charge. This is sometimes justified as having an extra free electron in the CB or hole in the VB as they also have very delocalized wavefunctions. In reality, the free charge carriers would screen the defect charge, which is not taken into account in this uniform background charge model. A periodic array of defect charges offset with a background charge would be expected to have rather different total energy to a single defect. The correction to total energy can be estimated, for example for a Gaussian charge distribution, by comparing the self-energy of an isolated charge and the sum of interactions in the periodic lattice [18]. The magnitude of these terms depend on the defect charge distribution and the size of the supercell [19, 20]. Of course, if the added electron was also homogeneously distributed, then the defect and background charge cancel each other exactly. On the other hand, for point charges one ends up with the simple Madelung sum $E_{Mad} = q^2\alpha/L\epsilon$, where α is the Madelung constant which depends on the structure of the lattice, and ϵ is the dielectric function. Reality is probably somewhere in between these two extrema. Therefore, the problem in applying such corrections lie in the estimation of the shape of the "extra charge" distribution from the rest of the electron density. Independent of the charge distribution shape, these interactions give $1/L$, $1/L^3$, and higher-order dependent terms to the total energy, although $1/L$-term can be corrected according to Paier et al. [21]. As the supercell grows bigger, the density of the homogeneous background charge vanishes, but the density of defect charge should approach some localized distribution. Knowing the asymptotic dependence, extrapolation schemes would be expected to work fairly well as has been demonstrated in Refs. [22, 23].

For the last point, the potential that would be created by an isolated defect now overlaps with its periodic images This is illustrated in Figure 2.6. The total potential of periodic defects is quite different to the isolated defect and more homogeneous. This is of course closely related to the total energy considerations in the last paragraph, but it is worth highlighting that the wavefunction solutions are also going to be quite different. The periodic potential gives a group of solutions forming an impurity band. Alternatively, one could think that the wavefunctions of isolated defects are overlapping and creating the band. In any case, if the potential and wavefunction overlap is significant it will be seen as dispersion of the defect states in the band structure. As mentioned in the beginning of the section, the size of the wavefunctions can be in the order of hundreds of lattice constants for "shallow" defects and in the order of a few for

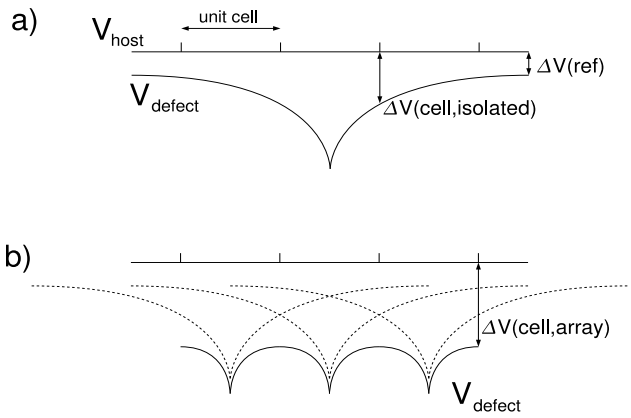

Figure 2.6: (a) An isolated defect potential and (b) the overlap of potentials. Also shown are the corresponding ΔV terms.

"deep" defects. However, the achievable supercell lattice constant with a DFT calculation is only a few lattice constants and the wavefunction overlap can be expected. That is, especially for shallow defects, the wavefunction overlap is significant and as the potential of a shallow defect is only a small perturbation to the host crystal, the dispersion actually follows that of the host band. In this case, in the calculation of ionization levels, it might be better to take the electron equally from all over the highest valence band and not from the VBM, because the defect state with its similar dispersion is also occupied equally all over the **k**-space. Especially, in the special **k**-point calculation, it is sufficient to average over these special **k**-points.

Finally, let's revisit the ΔV term. Evaluating the VBM energy in a defect supercell might be difficult sometimes if the defect states are close to VBM states. Adjusting host VBM to defect supercell energy reference via ΔV should help solve this problem. However, the periodic interactions of the defects make the estimation of ΔV quite unreliable. As illustrated in Figure 2.6, if the potential has not yet converged to bulk value, then the ΔV term is going to overcorrect. This already occurs with isolated defect and is more pronounced in the periodic array of defects. Moreover, even the geometry might not have converged at the edge of the computational supercell. In the study by Castleton et al. [23], it looked like this method gives very good results after all; as if the excessive potential correction would cancel other errors from the periodic image interactions. The applicability of this correction requires more studies. There is still the problem on whether to evaluate the electrostatic potentials on the interstitial sites, nuclei positions, or somewhere else. In our test calculations, however, different locations had hardly any effect on ΔV, and in the 216-atom supercell this term is already rather small.

2.3.4 DFT-LDA band-gap problem

The ionization levels can not be directly read from the position of the KS states, because the effects of electronic and geometric relaxation upon the adding of an electron would be disregarded. Still, just by looking at where the KS states lie, one can get a good first guess on the position of the ionization levels.

As stated before, the DFT-LDA underestimates the band gap. Then, a question artises: what happens to the states, both the KS states and the ionization levels, in the band gap? The studies performed with the GW method suggest that the shallow defects tend to follow the band edges and deep defects are correctly positioned even if the gap is too small [24]. That is, shallow levels should be referenced to CBM and deep levels to the corresponding band edge. However, all of this depends on the "shallowness" of the defect. Furthermore, the order of the ionization levels should still appear correct.

These problems are further mixed with the number of **k**-points used in the calculation. If the calculations are performed at some special **k**-points on which the CBM is higher than at the Γ-point, then the deep levels which would be inside the CB at Γ-point, now lie in the gap. If the charge state would be changed in this case, then at Γ-point we would change the occupation of the CBM and at special **k**-points we would change the occupation of the defect. Therefore, using only few **k**-points far from Γ-point can be both an advantage as well as a disadvantage (and quite suspicious).

Some authors have used the average band edges over the **k**-space to relieve this problem [25, 26]. This could be justified either by considering that ionization of a periodic array of defects (infinite number) requires corresponding amount of electrons which means completely emptying the highest valence band, or calculating ionization energy for the VBM ($\epsilon(0/+)$), which gives the same result. The advantage is that the band gap is much bigger, but unfortunately it depends on the supercell size.

Overall, the results for deep defects can be expected to be more reliable within the DFT supercell method. The supercell size scaling is useful, if not to extrapolate the results, then at least to check the convergence.

2.4 Band offsets

The real applications of semiconductors involve combining different semiconductor materials requiring understanding of the physics of the interfaces. In the thesis, the emphasis is on the effects of interface on the electronic properties which is intimately related to band offsets: how is the band-gap difference of the constituents divided between the valence and conduction bands (cf. Figure 2.7).

There is a comprehensive review on the subject by Franciosi and Van de Walle [27]. Here, after a quick overview of the basic concepts and common computational approaches, the applicability of the methods on the dilute nitrides will be discussed in more depth.

2.4.1 Basic concepts

This phenomenon is closely related to the valence- and conduction-band alignment in an n-p homojunction. The situation is illustrated in figure 2.8a. Before

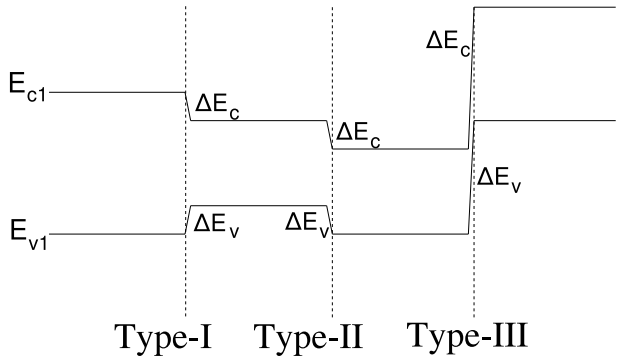

Figure 2.7: The three types of heterojunction band alignment.

connection, the Fermi-energies are at different levels. At connection, electrons flow from higher energy on n-side to lower energy on p-side, creating charge imbalance as they go, and similarly for holes. Charge imbalance, in the form of a dipole, creates an electric field opposing the charge transfer, which can also be illustrated with an electrostatic potential. The potential also shifts the band edges creating the band bending. This bending occurs at rather large distances (from a few Å to some μm) compared to the actual band offset just at the interface, usually within only a few monolayers.

For the heterojunction band offsets, in the first, simplest approximation the conduction band offset would be taken just from the difference of the electron affinities $q(\chi_1 - \chi_2)$ ([28]) (Anderson affinity rule), as shown in figure 2.8b. In this case, there is no dipole at the interface. This works well for some materials and not so well for others. For example, this results in about 230 meV for AlAs/GaAs interface, whereas the experimental value is closer to 530 meV.

However, a charge imbalance can occur not only due to doping, but also due to different "chemical" and electrostatic environment for atoms at different sides of the interface. Let's consider the AlAs/GaAs interface again. If, for example, on one side of an arsenic atom is aluminium and on the other side gallium, then there is likely to be charge transfer depending on the bond strengths. On the other hand, the Coulomb interactions due the charges on the AlAs side are different from the interactions due the charges on the GaAs side. It is convincingly shown in Ref. 29 that the latter is the dominant factor in at least at the AlAs/GaAs interface. This was done by forming a model interface from the charge densities of the bulk Wigner-Seitz cells and finding the same interface potential independent of the interface orientation.

The combination of both Fermi-level alignment and band offset is shown in figure 2.8c. At the size scale of the potential change due to Fermi-level alignment, the band-offset potential occurs very abruptly. Notice, that the band offset should be independent of the Fermi-energies. The potential step at

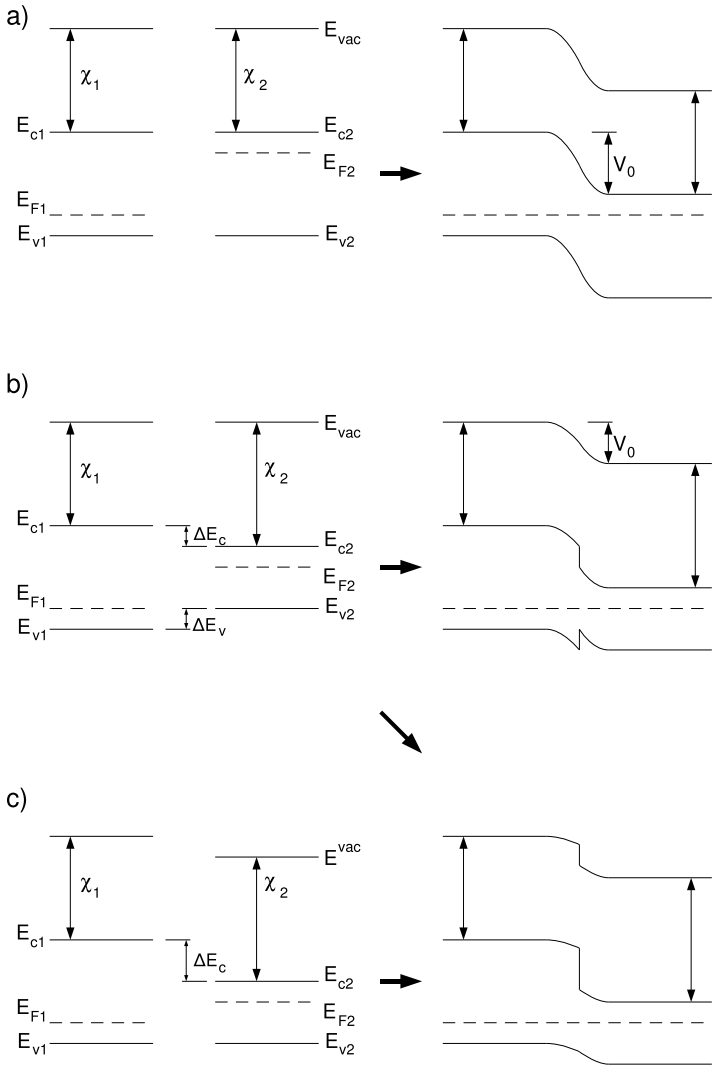

Figure 2.8: Band diagram for homojunction (a), heterojunction without interface dipole (b) and heterojunction with interface dipole (c).

a heterojunction interface can now be written as

$$\begin{aligned} \Delta V &= V_F + V_i \\ V_i &= V_{\text{es}} + V_{\text{ct}} + V_{\text{ref}}, \end{aligned} \tag{2.28}$$

where V_F is the potential step due to the alignment of the Fermi-levels and occurs also in homojunctions. V_i is the interface-specific term, which consist of the electrostatic term V_{es}, the charge transfer term V_{ct} and the alignment of the energy reference V_{ref}, which is needed later in the computational approach.

A one-dimensional illustration might not hold very well knowing that Coulombic interaction has a long range and a physical sample must have boundaries. Often in semiconductors, the area of the sample is much larger than the depth and the one-dimensional approximation then holds fairly well.

Whenever the theory gives a material-specific parameter for the band alignment, it also leads to the transitivity of the band offsets i.e., $E_{BO}(A/C) = E_{BO}(A/B) + E_{BO}(B/C)$ and also to the commutativity i.e. $E_{BO}(A/B) = -E_{BO}(B/A)$. In other words, the band offsets are a result of an intrinsic bulk property. For common-atom isovalent systems, the transitivity is found to hold within the experimental uncertainty [27], as well as within the computational error bars [30, 31]. In systems such as InAs/GaSb the band offset can be slightly different for the two inequivalent interfaces (In-Sb and Ga-As), but this is less than 0.1 eV. Transitivity can also be broken in heterovalent systems such as Ge/GaAs, but these interfaces will not be considered here, because they result in interface states (monopole charges at the interface) and consequently electric fields. Also, in rare cases, piezoelectric fields might be spontaneously generated in QW.

So far, a single parameter that could be used for alignment hasn't been found, although due to transitivity one could just reference all the materials to some common reference material. During the years, such levels have been sought for by several authors [28, 32, 33] with varying success. From the model of Baldereschi [29], it would seem that even if the band-offset really would be a bulk property, it is not a single parameter, since it is mostly governed by the charge distribution. Following along these lines, several authors have also tried to construct the charge density from simple building blocks. These are called the model-solid theories [30, 33, 29].

It has also been found in several studies that the interface orientation has a negligible effect on the band offset, especially in the case of isovalent systems. Similarly to Baldereschi et al. [29], Lambrecht and Segall [34] pointed out that it is possible to just rigidly shift both sides of the interface. On the contrary, they show how the dipole can be extracted from the bond polarities i.e., the charge transfer is the dominant effect, which leads to dependency on the interface orientation. The reason for the apparent contradiction could be that the sp^3-orbital energies are related to the depth of the atomic potential (electronegativity), which is then reflected to the potential difference across the interface. Indeed, in the method by Harrison and Tersoff [33] the sum of the sp^3 hybrid energies is used for the alignment.

2.4.2 Band-offset determination from first principles

If we are to accurately model the interface, we have to first accurately model the bulk, which leads us to the realm of first-principle methods. This also allows

accurate reproduction of electronic structure at the interface. Fortunately, in this case, we do not have to worry separately about all the different causes for the interface dipole. Unfortunately, we can not separate the causes.

In order to fully accommodate all the possible effects, usually the interface is built-in to the supercell in the way described in section 2.2.1. Then, the most straightforward method would be to look at the local densities of states [35]. This is practically troublesome, and this is not really the band offset, but the difference of eigenvalues of confined states which depends on the superlattice geometry. A better way is to combine the information about the interface from one calculation and information about the two bulk constituent systems from separate calculations [31, 34, 36, 29]. This method gives the "ideal" band-offset without the effects of confinement of states in a superlattice.

In the natural band-offset study [31] this was done by the help of core-levels: the VBM w.r.t. core level in an unstrained material and core-level difference from an interface calculation are used to find the VBO. The interface calculation is strained, but the core-level states should be less affected by the strain than the valence states. The authors estimate the errors to be within 50 meV due to this. This method requires an all-electron method.

It is more common to calculate the VBM w.r.t. an average electrostatic potential from bulk calculations and the potential difference across the interface from an interface calculation. The potential step is then used to align the bulk VBM as shown in Figure 2.9a (remember from section 2.1.1, a constant shift of the electrostatic potential is directly reflected to a constant shift in the band structure). The effects of strain (especially the changes in the band edges) are straightforward to take into account in the calculations. This method combines the physically valid framework with the computational affordability, since it is well-suited for DFT and also allows the use of pseudopotentials. A more thorough review of this method can be found from paper IV and also in Ref. 27.

True DFT should give the correct electron density and therefore also the correct electrostatic potential. Unfortunately, the true DFT eludes us; instead, we have LDA-DFT. LDA-DFT should still give a very good approximation of the density and subsequently of the electrostatic potential. Also, for example in Ref. 37, it was shown that GW potential is very similar to the DFT potential. Independent of whether the VBM or the midgap is correct in exact DFT, in LDA-DFT the VBM is not correct. There are some results showing that, compared to LDA, GW approximation lifts the VBM w.r.t. to the average potential differently on different materials. In many semiconductor interfaces, the shifts are similar and in case of band offset there will be a very fortunate error canceling [38], but in general this can not be expected For an example of such case, see the Si/SiO_2 interface in [37]. Also, even in the exact DFT, the band gap is severely underestimated and conduction band offsets can not be directly estimated. However, if the experimental (or GW) band gaps are known, then CBO can be calculated from them.

If the midgap is considered to be more correct in the DFT-LDA, then the computational band offset has to be further corrected. If the LDA band gap is underestimated w.r.t. experimental gap by ΔE_A and ΔE_B, then the VBO correction becomes $1/2(\Delta E_B - \Delta E_A)$. This is illustrated in Figure 2.9b. Often, the LDA band-gap errors are very similar which makes the correction small. If band-gap errors are dissimilar, this correction can give marked improvement

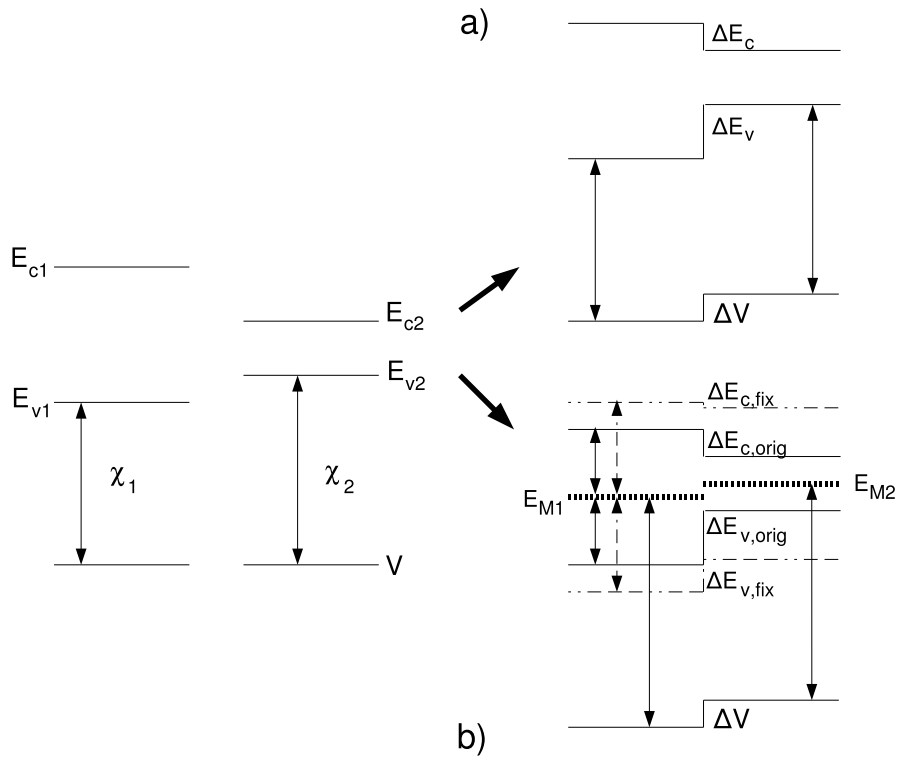

Figure 2.9: Diagrammatic figure of band alignment via alignment of the electrostatic potential (a). The use of calculated midgap and experimental band gap for further correction is shown in (b). The dot-dashed line presents the band edges, which have been obtained by adding half of the experimental gap to the calculated (and aligned) midgap level.

(e.g. in the case of the Si/SiO_2 in Ref. 37).

Even though the total potential (electrostatic plus XC-potential) looks a bit different to the bare electrostatic potential, the changes across the interface are fully included in the electrostatic potential. The reason is that the XC-interaction is very short-ranged compared to Coulomb interaction.

The relativistic effects are usually omitted in the DFT calculations, which is reflected in the lack of spin-orbit (SO) splitting of the valence band maximum. Including the spin-orbit coupling (SOC) in the calculations usually lifts the VBM higher by about 1/3 of the SO split. Again, as a large part of the errors tend to cancel, this is not a huge issue, and in fact, the correction can be estimated rather reliably after the calculations. This is also presented in paper IV for the AlAs/GaAs band offset, where increase of about 13 meV was found. Moreover, enabling spin-orbit coupling in the calculations has no effect on the potential step.

There's one more error source, which is rarely discussed. If the lattice constant difference, and the subsequent strain, does not match with the experimental difference, then there is going to be unphysical strain present in the calculation. Testing the generated error is difficult, but it should only result in some additional VBM split, which should be of small size.

Finally, the AlAs/GaAs interface is revisited as the comparison to literature might seem to partly contradict our results. The LDA band offsets are: [30] 0.37 eV, [39] 0.41 eV, [40] 0.41 eV and [41] 0.40 eV, whereas [31] 0.51 eV, our paper IV 0.52 eV (0.53 eV w/ SOC). On the other hand the GW correction was shown to change the offsets to [40] 0.53 eV and [41] 0.52 eV, because GW raises the GaAs VBM more than AlAs VBM by about 0.12 eV. Adding the GW fix to our results would increase the VBO too much. However, most of the LDA calculations and the GW corrections were done without $3d$ electrons (or not mentioned). Including the gallium $3d$ electrons in to the calculation also raises the GaAs VBM [39]. Apparently due to this, the GW correction would be very similar for GaAs calculated with the d electrons and for AlAs. Our calculations without gallium $3d$ electrons yield a clearly smaller band offset of about 0.45 eV.

2.4.3 Method for the dilute alloys

For the dilute systems, the method described in the previous section is suitable with minor modifications. GaAsN already requires a rather large supercell and GaAsN/GaAs interface even larger. Also the geometric relaxation in this system is considerable. In paper IV, we use the standard method modified in a way that the electrostatic potential is determined from the atomic cores (also employed by Jaffe et al. [42]), which enables us to overcome many of these problems.

In principle, if the average of the potential is shifted, then one could take only one point (any point) from the electrostatic potential and compare that between the interface calculation and the bulk calculation as long as the point is sufficiently bulk-like. In practice, due to charge redistribution and geometric relaxations, a well-defined reference point in space should be chosen. An example of such a point is the potential in the core regions of the ions. The equation

for the band offset between the materials X and Y can be written as

$$\begin{aligned}\Delta E_v &= (E_v[X] - V_c[X])^b - (E_v[Y] - V_c[Y])^b + (V_c[X] - V_c[Y])^i \\ &= (E_v[X] - E_v[Y])^b + (V_c[X]^i - V_c[X]^b - V_c[Y]^i + V_c[Y]^b) \\ &= \Delta E_v^b + \Delta V_c^i, \end{aligned} \tag{2.29}$$

where E_v is the valence-band maximum and $V_c[X](V_c[Y])$ is the electrostatic potential at the core of a given type of an atom (anion or cation) located in the material $X(Y)$. Superscripts b and i stand for the bulk and interface calculation, respectively. See Figure 2.10a for illustration. Finally, ΔE_v^b is the lineup of the VBM between the bulk constituents X and Y, and ΔV_c^i aligns the energy reference by combining the electrostatic potentials at the core of some appropriately chosen anions or cations from the bulk and interface calculations.

This can be justified alternatively by starting from the method used by Wei and Zunger [31] (which is actually based on the method used to experimentally determine the band offset from the core-level photoemission spectroscopy [43]), and replacing the reference core-level by the electrostatic potential at the core. The justification is, that if the electron is localized very close to the nucleus and there is no overlap of the wavefunctions, then the electrostatic potential at the nucleus contains all the interactions to this electron. The electrostatic potential at each ion ("the core potential") is calculated by placing a test charge at each ion and calculating in the usual way

$$V_c(\mathbf{R}_n) = \int V(\mathbf{r})\rho_{test}(|\mathbf{r} - \mathbf{R}_n|)d^3\mathbf{r}, \tag{2.30}$$

i.e., we estimate the test-charge distribution weighted average of the electrostatic potential. The distribution somewhat models a typical core electron. More precisely, the integral is calculated over a spatially extended spherically symmetric region whose radius is related to the PAW core radius. The norm of the test charge distribution is constrained to one.

In paper IV, we compare the traditional averaging method and our core-potential method with the help of the AlAs/GaAs test case. Results are shown in Figure 2.10. The ΔV_c^i values agree to within a few meV.

Although this method can be applied to any material system, with or without a lattice-matched interface, it is especially useful in calculating the band offset of the nitrogen dilute and low-concentration GaAsN alloy with respect to GaAs, as can be seen from the following points.

- A laterally wide computational cell is required for the dilute nitride material systems, but on the other hand, it is known that the macroscopic average of the electric dipole-related electrostatic potential becomes saturated rapidly away from the interface, and therefore only a few atomic layers of the material is needed in the longitudinal direction to estimate the interface potential.

- The calculation of the averaged electrostatic potential over conventional 8-atom unit cells is ill-defined since in the dilute nitride systems the atomic positions around the nitrogen atoms are strongly displaced from their ideal positions in the zinc-blende lattice structure.

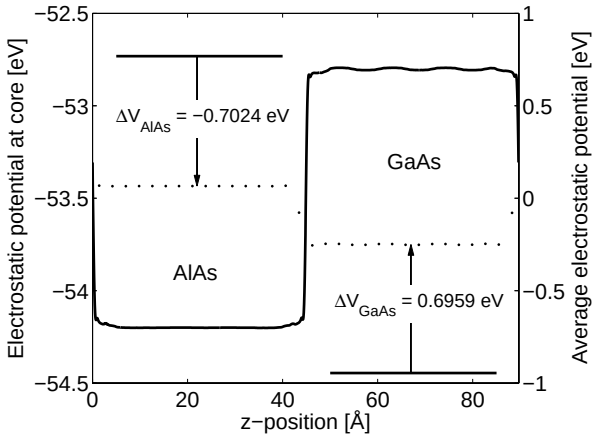

Figure 2.10: The comparison of the different methods for extracting the interface potential step from paper IV. The solid line is obtained with the potential averaging method (scale on the right) and the dots are the potentials at the atomic cores with the bulk values denoted by straight lines (scale on the left).

- The valence-band offset in the GaAsN/GaAs systems is known to be very small, so reading the band offset from the layer-projected local density of states (LDOS) is practically impossible, particularly due to the relatively small thickness of these layers.
- Finally, the direct comparison of potentials [36, 44], would require that the planar averages of both the bulk and superlattice electrostatic potentials have been computed using the same **k**-point mesh, and that identical lattice geometries have been used in the both cases.

In the previous section, it was shown how DFT-LDA resulted in several uncertainties. The similarity of GaAsN and GaAs fortunately increases the error canceling even further. Alloying of only few percent of nitrogen, results in only a few percent error in the computational to experimental lattice constant change from GaAs (-0.04 Å) to GaN (-0.06 Å). The spin-orbit split in GaAsN is known to be nearly constant for $0 < x < 3\%$, being 0.33 eV for $x = 2.2\%$, from the electroreflectance experiments by Perkins *et al.* ([45]). Via the same reasoning the differences in GW corrections to VBM can be expected to be small.

Chapter 3

Dilute nitrides and alloys

GaAsN with gradually increasing nitrogen concentration shows a few compositional regions with distinct properties. The dramatic band-gap reduction, being perhaps the most important reason for the popularity of $GaAs_{1-x}N_x$, shows four separate regions of different scaling ([46, 47]):

- $x \lesssim 10^{-5}$: The region of isolated isoelectronic nitrogen traps.
- $10^{-5} \lesssim x \leq 1.5 \times 10^{-3}$: The ultradilute region, with band-gap scaling (bowing) similar to conventional alloys. Isolated cluster states starts to emerge.
- $10^{-3} \leq x \leq 2.5 \times 10^{-2}$: The dilute region with slower band-gap scaling and marked interaction of the clusters.
- $x \geq 2.5 \times 10^{-2}$: High concentration limit which is experimentally problematic due to strong clustering and even phase decomposition (miscibility gap).

The isolated defects are first discussed in section 3.1. This includes the substitutional nitrogen, nitrogen clusters, and the interstitial defects. This is important also because many of the GaAsN properties can be thought of as "defect-like" behavior pertaining up to higher concentrations. This section also includes the results from paper Iand paper IIon nitrogen interstitial defects. Then, the alloying effects are covered in section 3.2. The involved physical origins are the same, even if some effects are more dominant at some concentration regions. The results from paper IIIare included here.

3.1 Nitrogen defects in GaAs

In more detail, the interesting properties of isolated nitrogen and clustered nitrogen defects in GaAs are listed here as presented by Kent and Zunger [47] (and references therein).

1. *Localized, single-impurity levels appear near the band gap.* Absorption and photoluminescence excitation (PLE) of GaP and GaAs show the "N_x center" due to anion-substitutional isolated nitrogen. In GaP this level

appears as in impurity-bound exciton at $E_{\mathrm{CBM}} - 33$ meV below the conduction band minimum (CBM), whereas in GaAs it appears as a sharp resonance at $E_{\mathrm{CBM}} + 180$ meV above the CBM. In contrast, in conventional isovalent alloys such as GaAs:P or GaAs:In the ensuing perturbation potential $V_{\mathrm{As}} - V_{\mathrm{P}}$ or $V_{\mathrm{Ga}} - V_{\mathrm{In}}$ is too weak to create a bound state.

2. *Anomalously small pressure dependence of single impurity states is observed.* Shallow, effective-mass like impurity levels are constructed from the wave function of the single nearest host crystal state. Consequently, when pressure is applied, such impurity levels change their energy at the same rate as the energetically nearest host crystal staet. In contrast, the impurity levels in dilute GaPN and GaAsN have anomalously small pressure coefficients: In GaAsN the nitrogen level moves with pressure to higher at much slower rate than the Γ_{1c} CBM of GaAs. Thus the impurity level becomes deeper as pressure is applied, emerging eventually as a discrete level into the gap at a pressure ≈ 2.2 GPa. Such small pressure coefficients of electronic state are usually indicative of localization whereby the wave function is constructed from many band of the host crystal, rather than from the nearest host crystal state.
3. *Sharp photoluminescence (PL) lines appear due to impurity clusters.* Even random substitution of impurities onto the atomic sites of a host crystal creates, by chance, impurity pairs and higher-order cluster. In conventional isovalent III-V alloys, such pairs give rise to broad resonances, within the valence and conduction continua, but no gap levels. In contrast, in GaPN and GaAsN, the N-N pairs form discrete levels inside the band gap extending in GaP down to $E_{\mathrm{CBM}} - 160$ meV and in GaAs down to $E_{\mathrm{CBM}} - 10$ or $E_{\mathrm{CBM}} - 80$ meV. These levels and their associated phonon replica are observed in emission after excitation into higher states. Excitation is then efficiently transferred to the deeper pair levels via tunneling (at low temperature) or via thermal excitation to the mobile states, followed by hopping into the pair levels (at higher temperatures).
4. *Redshift between absorption/PLE and emission is observed.* Already at a concentration of 0.05–0.1% nitrogen in GaAs, the emission lines are redshifted with respect to absorption. At higher concentrations the shift increases in energy, in contrast with high structural quality random, direct-gap III-V alloys, where absorption and emission occur at the same energy.
5. *Composition pinning of the impurity pair energy levels is seen.* The sharp emission lines from the pair levels remain initially at a fixed energy as the nitrogen composition increases both in GaPN and in GaAsN. This surprising pinning suggests that the impurity do not interact with each other. This behavior is craracteristic of deep transition-metal impurities in semiconductors, but not of hydrogenic impurities, which readily broaden into band and shift in energy as their concentration increases. In GaAsN and GaPN, as the concentration increases further, the PL from pair states acquires an asymmetric line shape with a sharp cutoff at high energy and a lingering tail at low energy. The carrier decay time in these tails is anomalously long. At yet higher concentration all of the pair/cluster lines disappear into a single, broad emission line. This behavior contrasts sharply with conventional III-V alloys such as InGaAs, where the emission line is featureless at all alloy compositions.

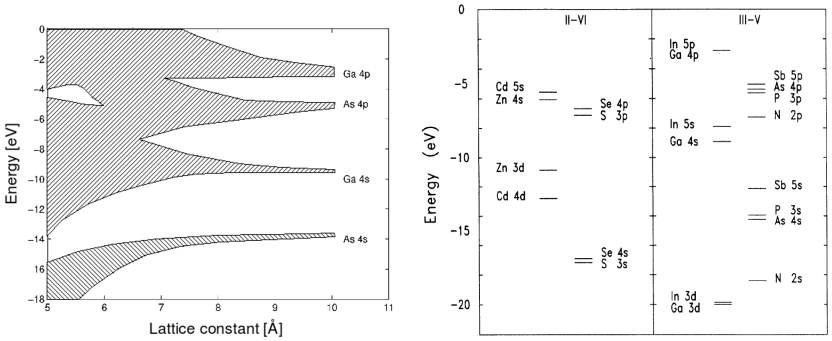

Figure 3.1: The allowed states and gaps in GaAs as a function of lattice constant and the free-atom LDA eigenvalues for some elements involved in the formation of III-V and II-VI compounds from [48].

The first two become apparent mostly at the ultradilute region ($x < 0.01\%$) and the next three at the dilute region ($0.01 < x < 0.1\%$).

3.1.1 Substitutional nitrogen

The formation of the valence and conduction bands of GaAs from the s and p states of gallium and arsenic is shown in Figure 3.1a. Even if not obvious from the figure, the valence band maximum mostly originates from the As4p orbitals and conduction band minimum from the Ga4s orbitals. How this happens might be easier to understand by first forming hybrid sp^3-orbitals for both Ga and As. As the sp^3-orbitals are brought closer together, they form the bonding and antibonding states; valence and conduction bands in a crystal. The As sp^3-orbital is lower in energy than the Ga sp^3 and therefore the valence band is more As-like and conduction band more Ga-like.

For the formation of GaAsN, we should also take into account the nitrogen states (Figure 3.1b). Not only are the N2p states mixing with the valence and conduction bands, but also N2s states are rather close to the Ga3d states. It is difficult to say where the N2p derived states would end up. In order to track the nitrogen contribution onto the band structure, the nitrogen local DOS is calculated and shown in Figure 3.2. The N2s states form two discrete bonding and antibonding peaks around a group of Ga3d states. This shows that Ga3d states should not be frozen in the PAW cores. The upper band structure is correctly reproduced without Ga3d, but they are still needed for e.g. finding the lattice constant. Apart from these states, there is nitrogen s-character only just at the CBM. This will be discussed in more detail in the following section. There is no nitrogen localization in the As4s band, but some contribution of N2p in the valence and conduction bands.

From the experiments, we know that for the isolated nitrogen there is a nitrogen derived defect state at about 180 meV above the CBM. Nature of a defect state inside a band is quite different from a defect state inside a gap. They usually mix heavily with the host states and can be difficult to see in experiments. However, the band gap can be increased by applying pressure or

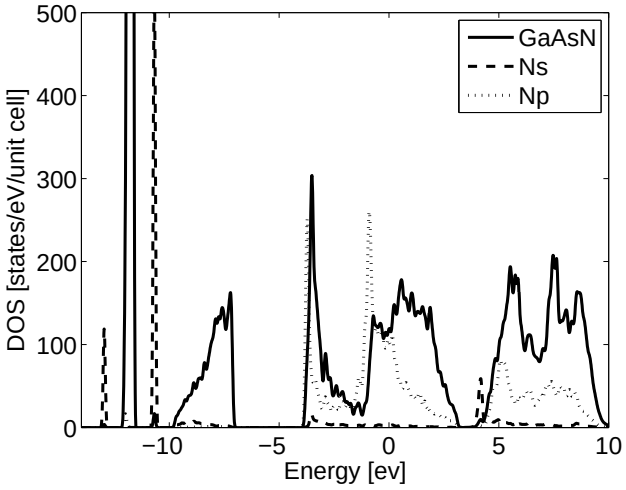

Figure 3.2: The calculated GaAsN (0.9% nitrogen concentration) total DOS (solid) and the nitrogen local DOS s (dashed) and p (dotted) components. The local DOSes are 100x magnified for clarity. The VBM is at about 3 eV.

changing the composition which will eventually bring out the defect state, since its pressure dependence is much smaller. Then, the defect state energy inside the band can be found by extrapolation, as shown in figure 3.3.

The properties in the list items (1) and (2) are rare, but not exceptional and are actually already explained in the list. They can still be understood on the basis of the familiar deep and shallow defects. The only difference is that now the defect state is inside the band and forms resonant states with the bulk states. For most isovalent substitutional defects, the potential is rather shallow. For nitrogen in GaAs, the strong (deep) potential is created by both the large electronegativity difference between N and As (3.04 and 2.18, respectively) and the large size mismatch (covalent radii 0.71 Å and 1.19 Å, respectively). It is generally known, that deep defects do not follow the pressure dependence of any single host valley as is also the case here; item (2).

The nitrogen resonant state localization has been calculated in several articles [50, 51, 47]. O'Reilly and Lindsay mention that over 50% of the wavefunction is localized on the nitrogen and the four neighbouring Ga atoms. On the other hand, according to Wang, 70% is within a sphere of 10 Å radius, but still 30% is farther than that. Also, Figure 3 in [47] shows how similar the nitrogen wavefunctions are in GaAs and GaP even though the other one is a resonance state and the other is a gap state.

The deep nature of substitutional nitrogen is also reflected in the discrete nature of the pair and cluster states. Concentration of nitrogen can be rather large, before the wavefunctions start overlapping, but when they do, the interaction can be quite strong. Moreover, the interaction is strongly dependent on the distance and orientation. A pair of nitrogens can be considered forming bonding and antibonding states and due to the strong interaction, the energy of the bonding state can drop below CBM [47]. Similarly for larger nitrogen

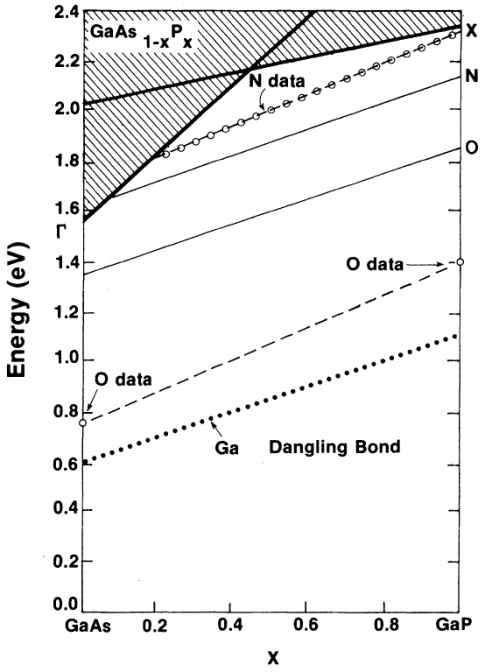

Figure 3.3: Single nitrogen substitutional defect levels in GaAsP from Hjalmarson et al. [49].

clusters. The list items (3), (4) and (5) are all evidencing the deepness of these cluster states.

3.1.2 Interstitial nitrogen

There is always some non-substitutional nitrogen present in the sample and increasingly at higher concentrations. Nitrogen is rather small which would bode well for the interstitial nitrogen presence. However, it is also rather reactive and especially the N_2 molecule is very stable. It is found out that the energetically favored defects are mostly (N-As) and (N-N) split-interstitials, and they have also been observed experimentally (also in paper VI). The presence of a (N-N) split interstitial is easy to understand as the second nitrogen would alleviate the high strain created by the first substitutional nitrogen and already studied by several groups [52, 53, 14].

In paper I, we aimed at a comprehensive study of all the pure interstitial defects. All the computational supercells had the same stoichiometry making the comparisons straightforward. Unfortunately, the small size of supercell (64 atoms) probably caused rather large error bars. Perhaps the most reliable results are in the last column of the Table I for PAW-LDA with $2\times2\times2$ **k**-points. With this method, the energy of a free N_2 molecule is about -1.74 eV per nitrogen (compare to Figure 1). The interstitial nitrogen was found to occur mostly as N_2 molecule. Moreover, the band structures show that these defects do not produce any states inside the gap.

The split-interstitials were not explicitly included in the study of paper I and also the charge state was always neutral. In paper II, we extended the previous study with the split-interstitials and used the formation energy methodology along with larger supercells. All the defects are still traditional deep defects for which the formation energy approach with DFT works rather well. The formation energy diagram is shown in figure 3.4. Concerning the chemical potential choice of this figure, it is worth noting that $\mu_N = \frac{1}{2}\mu_{N_2}$ is about 0.85 eV *higher* than $\mu_N = \mu_{\mathrm{GaN}} - \mu_{\mathrm{Ga}}$ in the arsenic-rich limit (and about 1.64 eV in Ga-rich limit), in contradiction to what is claimed in paper II. The former choice should still be valid for modeling as-grown samples [14]. In any case, by using either choice consistently results in clearly higher formation energy for arsenic-substitutional defects in As-rich limit compared to Ga-rich limit; as is intuitively correct. For the geometries, cf. Figure 1 of paper II.

The split-interstitials emerge as the lowest energy configurations. The eGN configuration of paper I is the same as the (N-As) split-interstitial. The interstitial nitrogen molecules appear energetically close to (N-As) in paper I, but clearly higher in paper II. However, the comparison among these two papers is not always straightforward. First of all, the energies in paper I are *per nitrogen atom*. Then, the energy difference of interstitial N_2 molecules is about $2 \cdot 0.33 = 0.66$ eV in paper I and about 0.60 eV in paper II (or 0.64 eV from a 64-atom supercell), in fairly good agreement. Similarly, the energy difference of neutral tGN and eGN is about 1.6 eV in paper I and about 2.12 eV in paper II (1.88 eV from a 64-atom supercell). Finally, knowing that moving the N_2 out of the crystal releases about 3.47 eV energy and using this as the chemical potential for the nitrogen, then the formation of eGN requires only half of this $0.33 - 1.47 = -1.14$ eV. The same choice for chemical potential was used in the Figure 2a of paper II, in which the neutral eGN formation is about 1.12 eV

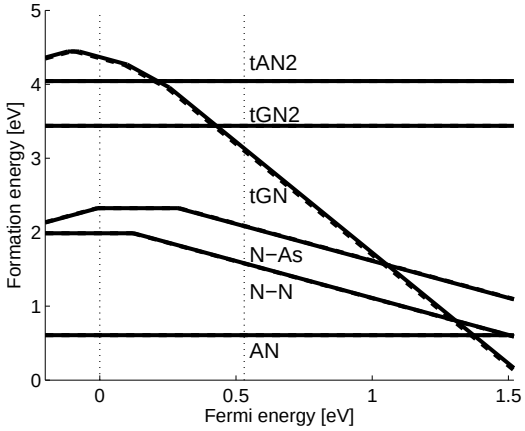

Figure 3.4: The formation energies of the nitrogen interstitial defects in GaAs(N) in arsenic-rich conditions and $\mu_N = \frac{1}{2}\mu_{N_2}$. Figure from paper II.

below tGN2. In conclusion, the results of the paper I and paper II agree rather well.

Next, the electronic structure is discussed in more detail. Starting from the N-N split interstitial defect case, as it has been discussed before in [14, 52, 53], there really is a singly occupied antibonding defect state in the gap. Moreover, we observe another nitrogen localized a bit below the VBM. These states are shown in the Figure 3.5a. The defect states intersect the valence band states, but are not strongly mixed with them. The defect states are fairly localized with the dispersion of about 0.5 eV, which even if not insignificant, it is still much smaller than the dispersion of the VBM. In Ref. 53, the upper state is described as $pp\pi^*$ and the partial charge density of this state indeed looks like $pp\pi^*$ (not shown). Nitrogen dimer has all the bonding orbitals filled and V_{As} has a high-lying singly occupied state, from which the electron would be donated to the nitrogen dimer. The model of N_2 and an arsenic vacancy explains the observed states, although it disregards the mutual modifications of the dimer and vacancy orbitals. Addition of an electron to this state increases the N-N bond length as expected from antibonding state, but it also decreases the N-Ga bond length suggesting bonding to gallium (see the bond lengths in Figure 1 of paper II). Another way would be to consider the N_2 orbitals and gallium dangling bonds. The symmetry of the system makes the application of either pure p-orbitals or sp^3-orbitals for nitrogen difficult.

In the case of N-As defect, the situation is very similar to the N-N case. There are two defect states close to or below the VBM as shown in Figure 3.5b. By looking at the localization of the states, we find that the upper state is localized to arsenic and lower state on nitrogen. Nitrogen has higher electronegativity, so the upper state should be antibonding and the lower state should be bonding. The changes in the bond lengths are similar to the N dimer. Formation energy diagrams for both the N-N and N-As defects should therefore look rather similar, as they do. Although (N-N) pair should inflict less strain on the crystal, (N-N) and (N-As) can compete energetically depending on the chemical potentials of

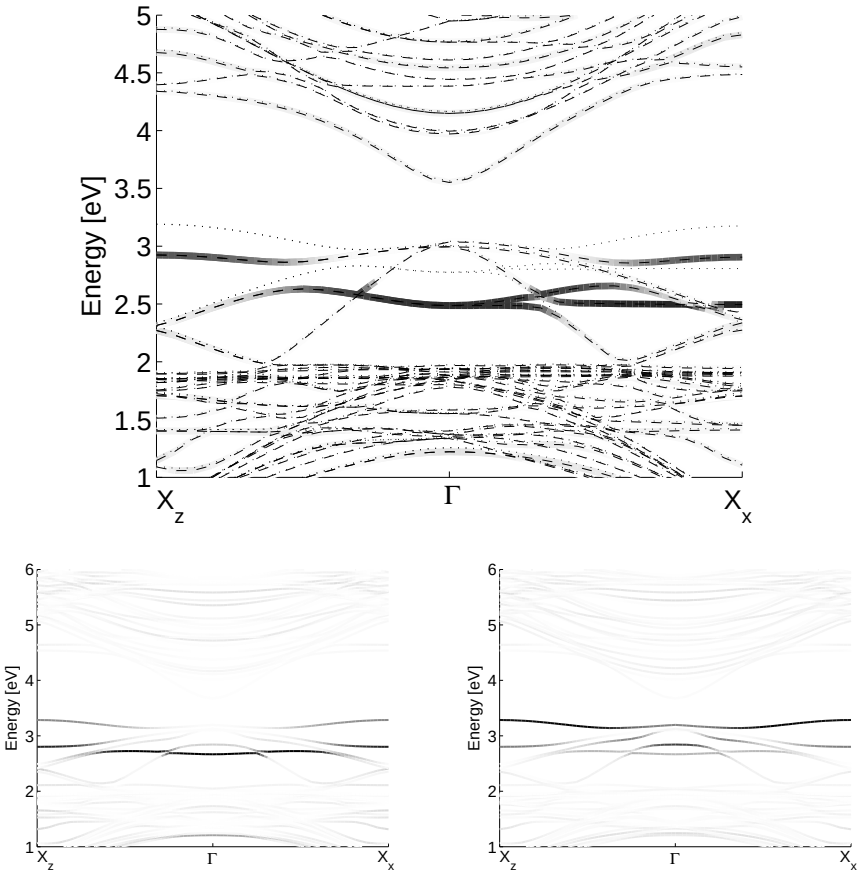

Figure 3.5: Up: The band structure of the (N-N) split interstitial defect, where X_z is parallel to the direction of the axis of the dimer and X_x is perpendicular to that. Dashed and dotted lines show the band structures for the two spin components and the dark solid parts denote high nitrogen localization (N localization for only one spin-component is shown for clarity). Down: The nitrogen (left) and arsenic (right) projected band structure of the (N-As) split interstitial defect.

nitrogen and arsenic

The (N-N) and (N-As) split interstitials are quite similar to (N-Si) split interstitial studied by Janotti et al. [54]. Similarly, bonding and antibonding states were found, only now the antibonding state is completely empty due to the lesser valency of silicon. The charge density isosurfaces in Figure 3 of Ref. 54 are somewhat similar to those found in (N-N) and (N-As).

The only true single interstitial defect, tGN, has all the atomic p-states more or less near the VBM depending on the occupation (occupied by three electrons at the neutral charge state). Naturally, up to another three electrons can be added to these states with a rather low energy, which subsequently brings down the formation energy at high Fermi-energy values.

Interestingly, tGN2 and tAN2 defects have no states near the gap. In the band structures in paper I, extra structure is found higher in the conduction band. Indeed, the nitrogen local density of states (not shown) also reveals nitrogen-localized states at about 1 eV above CBM and also about 1 eV below VBM. However, as one can also see from Figure 3.4, these states have no effect on the formation energy curves (i.e., no ionization levels inside the gap).

In summary, the only interstitials that can be expected to be found at meaningful concentrations are the split-interstitials. In n-doped samples they can have formation energies close to that of the substitutional nitrogen.

3.2 Alloying GaAs with nitrogen

At concentrations higher than 0.1% of nitrogen, a number of peculiar alloy properties are observed. The list from Ref. 47 continues:

6. *The band gap shows huge, composition-dependent optical bowing.* In conventional $A_xB_{1-x}C$ isovalent III-V alloys the band gap $E_g(x)$ changes with respect to the composition-weighted average of the constituents as

$$\Delta E_g(x) = E_g(x) - [(1-x)E_{AC} + xE_{BC}] = -bx(1-x) \tag{3.1}$$

   where the bowing coefficient b is constant and usually < 1 eV. The reasons for band-gap bowing are well understood. In contrast, in GaPN and GaAsN the bowing is huge and composition dependent, being largest at small x: ≈ 26 eV at $x < 1\%$ and ≈ 16 eV at $x > 1\%$.
7. *The electron mass is anomalously heavy but decreases with concentration.* In pure GaAs the electron effective mass is $0.066m_e$. Small amounts of nitrogen increase it to $\approx 0.4m_e$ or $(0.12\text{–}19)m_e$, but subsequent addition of nitrogen appears to reduce the electron mass. As the Fermi energy moves further into the conduction band, the effective mass becomes higher. In sharp contrast, in conventional alloys the mass changes monotonically with composition.
8. *The reduction in band gap with increased temperature slows down with nitrogen addition:* Band gaps are always reduced as temperature is increased. However, in conventional alloys the temperature coefficient is close to the concentration-weighted average of the constituents. This reduction in PL energy with increased temperature slows down dramatically with small addition of nitrogen to GaAs and GaP. Furthermore, the intensity of the PL lines of conventional alloys decreases with increasing

temperature, but this decrease is acceleraetd by nitrogen addition, especially at low temperatures.

9. *The energy of the PL lines are blueshifted as the excitation power increases,* indicating occupation of previously empty states (so excitation must now occupy higher-energy states). This is known to occur in alloys containing localized, quantum-dot like clusters.
10. *The emission decay time becomes longer with decreasing emission energy.* In other words, the states that are deeper in the gap (lower emission energies) have weaker dipole transition elements (or equivalently, less Γ character and more off-Γ character).
11. *New, high-energy bands appear in reflectivity at higher concentration:* Electro-reflectance measurements detect a new composition-dependent band edge, called E_+, at about 0.4–0.6 eV above the band edge, called E_-. Whereas the lowest state E_- decreases with x ("bowing"), E_+ is seen to increase with x. Klar *et al.* demonstrated via resonant Raman scattering that E_+ is derived from nitrogen-inducted Γ-L mixing, not from isolated nitrogen impurities. At yet higher energies, around 3 eV in GaAsN, one observes the E_1 transition, which is known to encompass a large volume of the Brillouin zone containing the L_{1v}-L_{1c} transition. One branch of this transition is nearly composition independent according to Sham *et al.*, but changes somewhat according to Leibiger et al.
12. *The conduction-band L_{1c} state appears split.* Kozhevnikov *et al.* observed that the L_{1c} band of GaAs is split into an upper branch with energy increasing with nitrogen concentration, and a lower branch with decreasing energy. They interpreted these to be $t_2(L_{1c})$ and $a_1(L_{1c})$.

These properties hold for up to a few percent of nitrogen. At even higher concentrations, the nitrogen is no longer miscible (soluble) and the excess nitrogen would result in a phase separation [55, 56]. This miscibility gap holds for up to $x \approx 0.95$, the dilute arsenide GaN limit. As discussed in section 2.2.3, the nitrogen solubility to GaAs is actually much lower than this, but the nitrogen substitution onto the surface enhances the practical obtainable nitrogen concentration [14]. Therefore, in principle nitrogen would like to phase separate in the GaAsN samples, but the nitrogen diffusion is very slow. Low temperature growth is required for GaAsN, in which case the nitrogen surface mobility is low favoring miscibility. Also gallium and arsenic mobility is low which results in gallium vacancies and other native defects. This is not really advantageous for nitrogen diffusion and the native defects act as centers for non-radiative recombination [57]. Although detrimental to lasers, such short non-radiative lifetimes would be desirable e.g. for SESAMs [58, 59]. The growth then governs the nitrogen distribution, which is to a good approximation random substitutional. In bulk, nitrogen atoms would like to stay away from each other (cf. paper III, [47] and many others). Local nitrogen configuration stability is mostly affected by the lattice strain. In general, the clustering of nitrogen is not energetically favored, although some studies suggest that [001]-oriented nitrogen chains are energetically favored [60] or that the growth environment with hydrogen promotes [110] chains [61].

3.2.1 Electronic structure

Let's start again by considering a conventional alloy with weak impurity potential. Substitutional aluminium in GaAs also creates a resonance inside the band, but the resonance is weak and quite delocalized. As the concentration is increased (but already at a very low concentration), the resonant states overlap forming something akin to an impurity band (inside a band), but heavily hybridized with the delocalized host states. The material gradually obtains AlAs character and becomes AlGaAs alloy. Even if there would be a pair of aluminium relatively close to each other, the interaction is small and it's mixed with other "higher-order" pair and cluster states resulting in the broadening/tail of the band edges.

In the case on nitrogen, the localized resonant states start to overlap only when they are relatively close to each other. When they do, the interaction is strong and strongly dependent on the distance and orientation of nitrogen atoms. As a result, instead of an impurity band, discrete cluster states are taking a more dominant role. At the same time, there appears strong local fluctuations in the potential, which also results in the unconventional perturbation of the host states. Adding to this the interaction of the perturbed host states with the nitrogen localized states brings forth a variety of interesting phenomena.

It was found by Neugebauer and Van de Walle [62] and Wei and Zunger [63], that in majority the large band-gap reduction in GaAsN alloy owes to the large cation displacements around nitrogen, and the nitrogen localization of the CBE owes to the electronegativity difference. This sounds credible, but as a quick check, in my calculations in a 64-atom supercell, the LDA band gap of unrelaxed GaAsN (but with the GaAsN lattice constant) is about 100 meV below that of GaAs. The geometric relaxation then leads to further reduction of about 300 meV. In reality both of these effects are coupled in forming the strong "central-cell" potential, as already discussed.

It is useful to consider the situation from the perspective of the perturbation theory with the CB states as the basis on which a perturbative potential acts. A substitutional impurity breaks the translational symmetry of the crystal, which has two consequences: First, the wavefunctions can not really be labeled with the wave vector k any more. In a supercell construction, there is always a periodic array of defects. In this case, some states fold into the Γ-point, which at the limit of infinitely large supercell (where all states fold into Γ-point) equals to the complete lack of k-labeling. Second, some degeneracy of the states, stemming from the translational symmetry, is lifted. A change in the symmetry representation of a state might not seem important in itself, but the states of different symmetry are affected differently by the perturbing potential, which then splits the degeneracy.

From the perturbation theory, the first order correction to the energy of a state ϕ_n by the perturbing potential ΔV is

$$E_n^{(1)} = \langle \phi_n | \Delta V | \phi_n \rangle$$

and the second-order correction is

$$E_n^{(2)} = \sum_{k \neq n} \frac{|\langle \phi_k | \Delta V | \phi_n \rangle|^2}{E_n^0 - E_k^0}$$

The states interact only if they have a matching symmetries so that the matrix element is non-zero; especially, if both have the a_1 symmetry. Therefore, a strong impurity potential causes both a first-order shift, which can be estimated from orbital energies, and enables through the second-order correction the interaction of several host states and the impurity state. Often, the first-order correction is sufficient, but this is not the case with GaAsN.

A zinc-blende crystal has T_d point symmetry and it stays T_d with the introduction of nitrogen (if the origin is at the impurity). However, as mentioned, the loss of translational symmetry splits the states: the X_1-derived triplet into a $a_1(X_1)$ singlet and a $e_2(X_1)$ doublet, the X_2-derived triplet is not split and has $t_2(X_3)$ symmetry, the L_1-derived qudruplet is split into a $a_1(L_1)$ singlet and a $t_2(L_1)$ doublet, and Γ_1 is just $a_1(\Gamma_1)$. [64] The nitrogen impurity state is denoted $a_1(N)$. Naturally, in alloy these symmetry considerations are only approximate.

As the states of a same symmetry interact, they can also mix their wavefunction "characters" (the projections to host crystal states). For example, when the host $a_1(\Gamma_1)$ and $a_1(L_1)$ interact, the resulting wavefunction can be approximated as their linear combination. L valley and the extrapolated nitrogen-impurity level are energetically relatively close to CBM: about 300 meV and 180 meV, respectively. This enhances the interaction and the Γ-L mixing becomes strong.

In order to illustrate the band repulsion, we could look at the states with increasing concentration. On the other hand, states originating from different valleys of the host band structure have different pressure dependence. The band repulsion can therefore be illustrated also as a function of pressure, as shown in Figure 3.6. The states of the same symmetry repel and the $\Gamma/L/X$-characters mix accordingly.

The repulsion of some a_1 states i.e., the strength of some matrix elements, could be small. In the extreme, it could be sufficient to consider the repulsion of only two states. A model, where only the repulsion of the CBE and the nitrogen resonant state is considered, "the band-anticrossing model" (BAC), provided a very good reproduction of the measured CBM dependence [66]. Next, I provide a quick derivation of the BAC (for more details, cf. Atkins and Friedman [67], chap. 6.1). Let the CBE and nitrogen state be denoted by ϕ_c and ϕ_N, and they are the eigenstates of H_0 with energies E_c and E_N, respectively. The perturbation in this two-state basis can be written as

$$\Delta V = \begin{pmatrix} V_{cc} & V_{cN} \\ V_{Nc} & V_{NN} \end{pmatrix}$$

where $V_{cc} = \langle\phi_c|V|\phi_c\rangle$ etc., and $V_{Nc} = V_{cN}^*$. V_{Nc} is the repulsion term giving the band anticrossing. To solve $(H_0+V)\phi = E\phi$, we need to find the non-trivial solutions of the determinant

$$\begin{vmatrix} E - E_c - V_{cc} & -V_{cN} \\ -V_{Nc} & E - E_N - V_{NN} \end{vmatrix} = 0$$

which results in the energies

$$E_\pm = \frac{1}{2}(E_c + V_{cc} + E_N + V_{NN}) \pm \frac{1}{2}\sqrt{(E_c + V_{cc} - E_N - V_{NN})^2 + 4|V_{cN}|^2} \quad (3.2)$$

In the equation derived in [66], it was assumed that the first-order correction is zero i.e., $V_{cc} = V_{NN} = 0$. This approximation was not discussed in the original

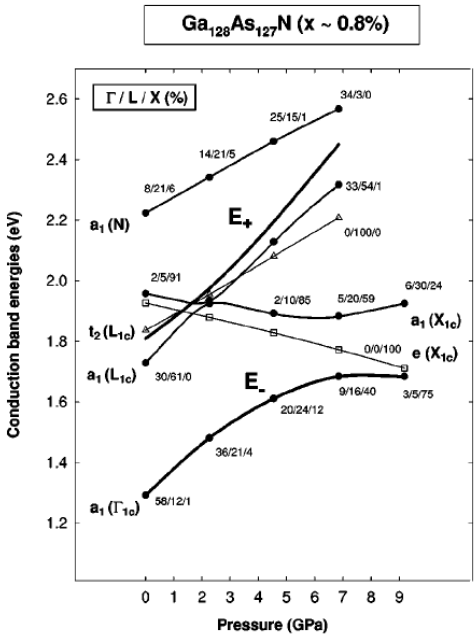

Figure 3.6: The pressure dependence of selected states in GaAsN from the calculations of Mattila et al. [65].

article, but clearly from the comparison to experiments, it seems to work well. The final form is then

$$E_{\pm} = \frac{1}{2}\left(E_c + E_N \pm \sqrt{(E_c - E_N)^2 + 4|V_{cN}|^2}\right) \tag{3.3}$$

It is maybe worth pointing out, that applying the expansion

$$\sqrt{a+x} \approx a^{1/2} + \frac{x}{2a^{1/2}} + \ldots$$

which is valid with small perturbations, leads to same form as from the conventional perturbation theory with weak perturbation

$$E_- = E_c + V_{cc} + \frac{|V_{cN}|^2}{E_c + V_{cc} - E_N - V_{NN}} \tag{3.4}$$

$$= E_c + \frac{|V_{cN}|^2}{E_c - E_N} \tag{3.5}$$

The assumptions used in the BAC model appear somewhat justified. Especially, since it gives good results on the energy dependence of the E_- states on composition and pressure, and fairly good results on the energy dependence of E_+ and the effective mass. Then again, it also fails on producing several properties and thhe first-principles studies showed that the situation is more complicated in reality.

In the more complete picture, we have to worry about three things: a) the nitrogen resonant state, b) the nitrogen pair/cluster states, and c) the perturbed

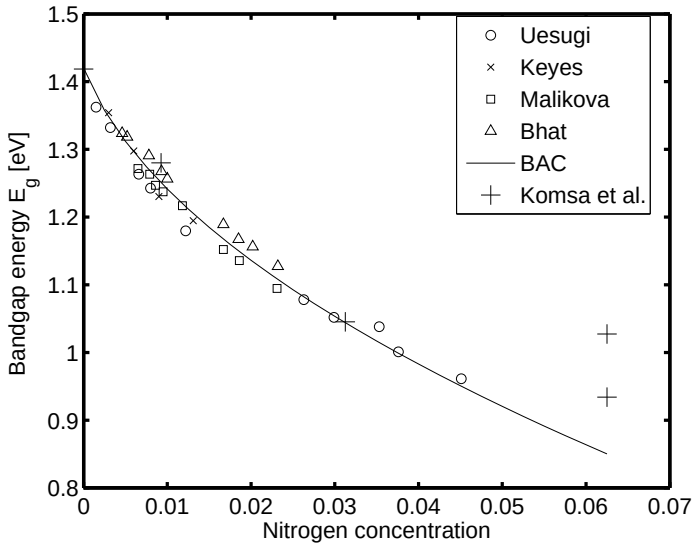

Figure 3.7: The band gap of GaAsN as a function of nitrogen concentration including the experimental results of several groups [70, 71, 72, 73], our calculated results and the line denoting the BAC theory. In our DFT results the calculated change of the band gap relative to GaAs is applied to the experimental band gap of GaAs.

host states. All these states are interacting in a complex manner and discussing each state separately would serve no purpose here. In this thesis the band edges are of importance. Therefore, discussing the behavior of the experimentally relevant E_- and E_+ transitions seems more fruitful.

3.2.2 The origins of E_- and E_+ transitions

The E_- transition/state starts from the GaAs CBE. As the nitrogen concentration is increased, the state drops dramatically down in energy with non-linear bowing (first observed by Weyers et al. [68], Kondow et al. [69]), the PL intensity decreases, and the effective mass increases. The band gap decrease is shown in Figure 3.7.

It is agreed that the band bowing does not require cluster states [47, 74]. The $a_1(N)$ state disappears already at 1% concentrations, so a model where a nitrogen state at a constant energy is the principal cause for the bowing seems lacking. At 0.1%–1% concentration range, there should exists a group of nitrogen localized levels at the nitrogen resonance energy, although broadened. At above 1% there are no purely nitrogen localized states near the CBE. In the calculations by Mattila ([65]), the $a_1(N)$ state was found to rise to much higher energies. It seems certain though, that the GaAs Γ_1 state gradually obtains both the L-character, suggesting interaction with $a_1(L_1)$ state, and nitrogen localization. As a side-note, O'Reilly and Lindsay found an effective nitrogen state which could be used to obtain the bowing (the extended BAC model).

As CBE plunges down in energy, the cluster states which might lie below CBM, are taken over. The hybridization of CBE states with the cluster states and increasing cluster-cluster interactions cause strong local fluctuations in the potential and subsequently in the perturbed host states.

Transition matrix element $\langle\psi_i|\widehat{p}|\psi_{\mathrm{VBM}}\rangle$ from conduction band state i to VBM is directly proportional to the percentage of the Γ-character [65]. The loss of the Γ-character, is the dominant factor in decreasing the PL intensity, although other effects might also play a part. Gain of the L-character on the other hand means that the wavefunction is delocalized in reciprocal space and localized in real space (to N and neighboring Ga atoms). This results in the decrease of the electron diffusion length, and subsequently the mobility [74]. On the other hand, excitons are more easily bound which increases the exciton lifetime [47].

The delocalization in reciprocal space also means a higher effective mass. Interestingly, cluster states have effect on the effective mass. If the CBE energy coincides with cluster states, the hybridization leads to sudden increase in the effective mass (cluster states have naturally very low effective mass). [75]

At low concentration, the nitrogen state emerged from the CB with the application of pressure. At higher concentration, there is no single state emerging, but the pressure coefficient decreases markedly close to that of a nitrogen resonant state. In this case the CBE gradually changes to more nitrogen-localized as can be seen from Figure 3.6.

The origin of the E_+ transition has been a subject of rather more controversy. According to BAC model, this state originates from the nitrogen $a_1(N)$ state and other models suggest it originates from the $a_1(L_1)$ state. The difficulty in settling this debate was in the energetic proximity of the two states. Very recently, [76, 77] were able to measure the state at sufficiently low concentrations with μ-PL and found the extrapolation to $a_1(L_1)$ as already anticipated by several other authors [64, 78, 79]. The transition from the top of the valence band to $a_1(L_1)$ and $a_1(X_1)$ are forbidden in pure GaAs, but due to mixing they obtain more and more Γ-character enhancing the oscillator strength. Rather contradictory, this peak was found to be rather sharp in the modulation spectroscopic studies [78], but almost 100 meV wide in the μ-PL studies [76]. Kent and Zunger [47] found a group of nitrogen localized L-like states near the E_+ energy originating from the hybridization of the $a_1(L_1)$ with other states. They defined E_+ as the upper edge of this band. Szwacki and Bogusławski [64] also points out that $a_1(L_1)$ state is induced by the presence of N, but is not localized on N. This again, is in contradiction with the BAC model, which moreover would give the E_+ transition strength dependence incorrectly.

3.3 GaAsN electronic structure within DFT

The computational studies of GaAsN has been hindered mostly by the large number of atoms required in a supercell. Due to this, most of the earlier studies were performed with the tight-binding methods [80, 81] or empirical pseudopotentials [82, 65, 47], but also DFT [63, 83, 50].

The small supercell in DFT is still a serious shortcoming. The situation at the 0.1% and 1% concentrations are quite different. At the moment, 0.1% case is quite impossible to achieve. Also, in order to model the random distribution of

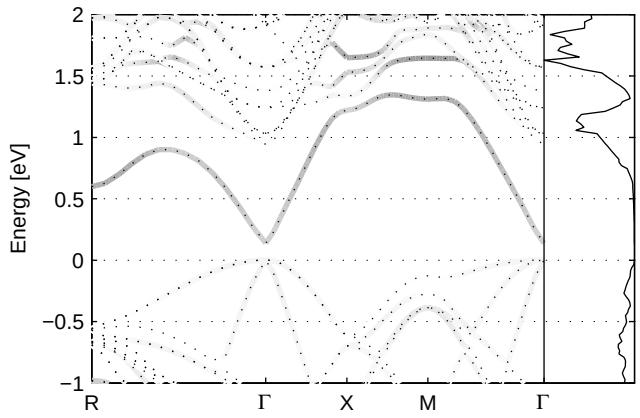

Figure 3.8: Nitrogen localized band structure and local DOS from a calculation of 64-atom supercell with one N atom. Figure from paper III.

nitrogen atoms i.e., the fluctuations in the material, a large supercell is needed. In this context, it is worth mentioning the charge-patching by Wang [50], where the electron density of a big supercell is patched from the electron densities of smaller supercells and then only a few eigenvalues near the gap are solved with the folded spectrum method. This method is quite limited for studying random configurations, but should work well for isolated nitrogen atoms or clusters embedded in GaAs.

In many instances, the best we can get is to put one nitrogen atom in a supercell. This results in a periodic lattice of nitrogen impurities. Also, pure impurity and cluster states are absent. The band bowing can still be obtained fairly well as was shown in figure 3.7 with the results from our DFT calculations. Due to the strong dependence on the geometry, the nitrogen positions and the shape of the supercell now affect the results. At high concentrations (more than few percents), the band gap fluctuations can become meaningful. This is especially the case in our periodic model, which is actually closer to a lattice of nitrogen pairs/clusters. This was demonstrated in our paper III, where over 200 meV changes in the band gap were found at 6% nitrogen.

The band structure from a (small) supercell calculation looks rather peculiar after all the discussion about a_1 states in the previous section. Band structure of 64-atom supercell is shown in figure 3.8. (Also [83] and [79] E_+ can be found).

Now, the CBE is *a nitrogen localized band*. The nitrogen resonant state is probably somewhere higher in energy (and forming a band) [65]. At the zero-concentration limit, this band should become the lowest conduction band. At the Γ-point, the next lowest states originate from the folded L-states of GaAs; both the $a_1(L_1)$ and the $t_2(L_1)$. In addition, there is structure of the E_- band in the L-direction, which naturally has L-character. Also the changes in the N localization match with other models and experiments. Interestingly, the calculated dielectric functions [79] show that the E_+ transitions might originate both from the folded L-states of GaAs and from the E_- band near the supercell L-point, although mostly from the former. Moreover, it seems likely that the latter states are connected to the "L-like" states below the E_+ found by Kent and

Zunger [47], and interpreted to originate from the interaction with the $a_1(L_1)$.

So far, the nitrogen localization of both the E_- and E_+, and their energy dependence is fairly well reproduced, disregarding the alloy fluctuations. Since it seems that the E_- state is fairly correctly obtained, then, not surprisingly, so is the pressure dependence of this state [83].

In paper III, we have calculated the density of states, and compared to those obtained from the ballistic electron emission microscopy (BEEM) experiments [84]. The features and tendencies matched well: the CBE and $t_2(L_1)$ are shown in the spectra as they have large DOS (with delocalized wavefunction character), and similarly peaks were found in our calculated DOS. They also found a weak peak, which was assigned to $a_1(L_1)$.

Valence band structure should be almost as correct as that of GaAs. There is also a small splitting of the VBM due to the symmetry breaking, but it's only about 3 meV for 1% nitrogen concentration and about 20 meV for 3% nitrogen according to Kent and Zunger [47]. My DFT calculations give similarly about 4 meV for 1% and a little bit smaller value of 6 meV for 3 % nitrogen. This is in addition to the splitting due to the strain. As one can see from Figure 3.8, there are no nitrogen localized bands near the VBM which could contribute to a strong VBM splitting.

Chapter 4

The local environment effects of nitrogen

The unconventional properties of GaAsN material are promising for many applications. Unfortunately, there are also some serious limitations to its real-world usage. There is no substrate on which it could be grown without strain, and both the electrical properties (e.g. charge carrier mobility) and the optical properties (e.g. PL intensity) are poor. Therefore, GaAsN needs to be alloyed and naturally also doped. In this chapter, the behavior of nitrogen in an alloy and the interaction of nitrogen with defects is studied. In the case of defects, the interaction occurs in two ways: nitrogen can affect the defect formation and defects can change the nitrogen (alloy) properties. In the case of quaternary alloys, we are mostly interested on the alloy properties as a function of nitrogen local geometry. Finally, we take a look at the interfaces of (In)GaAsN alloys. This chapter includes results from paper IV and paper V concerning the band offsets, paper VI on the short-range ordering, and paper VII on defects.

4.1 Defects in GaAsN

As shown in the previous chapter, even a very dilute concentration of nitrogen can change the "alloy-properties" of the material, such as the band gap or the effective mass. In this case, it is possible to have the defect concentration comparable to nitrogen concentration, and subsequently, the defects in interaction with nitrogen, can also affect these properties. This is in contrast to conventional alloys.

Naturally, in GaAsN the CBM decrease can also change the defect properties: if the charge states which are stable at the high Fermi-level positions in GaAs, might reside inside the CB in GaAsN. Isolated hydrogen donor is seen to behave like this according to the calculations by Janotti et al. [85].

The formation of GaAs native defects might experience changes in the vicinity of nitrogen. In general, in paper VII we found that the defect formation energy converges very fast to that of pure GaAs when moving away from nitrogen. Quite interestingly, we found the gallium vacancy formation to be more favorable near nitrogen. This can be understood simply by the relief of strain associated with the three remaining Ga-N bonds. Gallium vacancy in GaAsN

was also studied in [86] with similar results.

Concerning the defect formation in the presence of other dopants, previously, there has been abundance of studies of hydrogen in GaAsN and of donors (mostly Si) in GaAsN. Hydrogen is innately involved in the MOCVD and gas-source MBE growth of GaAsN and has been seen to cause several unusual properties. Hydrogenation removes the nitrogen induced band-gap reduction and when hydrogen is removed by annealing the sample, the typical GaAsN band gap returns. [87] It also compensates, even overcompensates, the lattice strain due to nitrogen. [88, 61] Hydrogen seems to first passivate the nitrogen clusters, which can be used to remove the deep cluster-states. [61] Moreover, hydrogenation has been used to fine-tune the nitrogen concentration.[75] The nature of the hydrogen-nitrogen complexes has been debated quite a lot, [85, 86, 26] but it seems like a hydrogen molecule bonded to nitrogen is the dominant configuration [89]. Silicon in GaAs is a dopant which can behave either as a donor or an acceptor depending on the substitutional site. Donors have been found to form split-interstitial type complexes with nitrogen. Again, the nitrogen atoms which were part of these complexes were "deactivated" i.e., they did not take part on the band-gap reduction. [90, 91, 54]

The mechanism behind the disappearance of the band-gap lowering has not been discussed much in the literature. Supposedly, especially the disappearance of the tetrahedral symmetry is the culprit for the lack of nitrogen interaction with the host bands (the CBE band consisted of Ga-*s* and N-*s*). To put this into terms of the discussion in section 3.2.1, the perturbing potential of the defect is such (due to the symmetry), that the matrix elements are vanishing. Moreover, the defect is no longer a simple substitutional isovalent impurity, but more as a molecule in a substitutional site. This would suggest more localized defect states, and indeed it seems that these defects are "pure" deep defects. The energies of the defect states are also changed, maybe inside the gap or high in the CB, which is reflected on the magnitude of state repulsion.

4.1.1 Beryllium doping of GaAsN

Beryllium in GaAs is almost completely located substitutionally in the gallium site providing p-type doping and is possibly the most common dopant for this. In paper VII, we present the results of the studies on the beryllium doping of GaAs(N).

The formation energy diagrams are shown in Figure 4.1. It is clear that the gallium-substitutional beryllium is clearly the lowest energy configuration in pure GaAs. Beryllium interstitial next to substitutional beryllium (Be_{Ga}-Be_I) is the second lowest, and could slightly compensate the p-type carrier concentration, although their energy difference depends on the beryllium chemical potential.

The situation is more interesting in the case of GaAsN. The energy of (Be_{Ga}-Be_I) now seriously competes with Be_{Ga} and also several (Be-N) split-interstitial defects appear relatively low in energy. The former would explain the compensation found in Be doped GaAsN samples. The dissociation or formation of (Be-N) could also affect compensation, but in solid-source MBE grown samples, the carrier concentration did not change during a thermal annealing suggesting that (Be-N) defects are either very stable or their concentration is small.

As with several other defect-nitrogen complexes, we also found the nitrogen

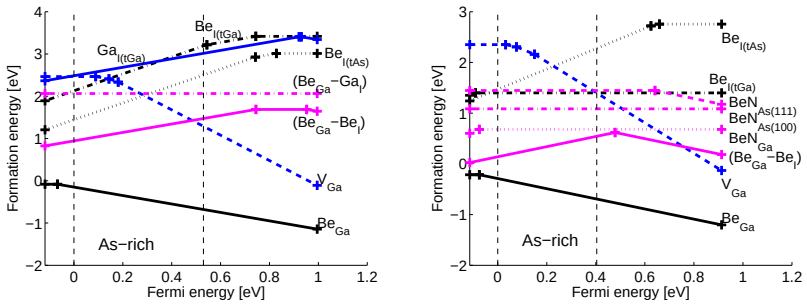

Figure 4.1: The formation energies of Be doped GaAs and GaAsN. Figure from paper VII.

deactivation due to Be doping. Especially, (Be-N) split interstitials were very effective at cancelling the band-gap reduction due to nitrogen alloying, although Be_{Ga} at the nitrogen nearest neighbor site also produced some cancellation.

Moreover, the small size of both the beryllium and nitrogen, and the subsequent strain at the substitutional sites, makes the defect clustering around N_{As} and Be_{Ga} quite likely, and consequently decreases diffusion, although the energy landscape of Be_I sites indicated decreased barriers for diffusion.

4.2 Quaternary alloys

As with defects, additional alloy components can change the alloy properties considerably also depending on how they are organized around the nitrogen. In essence, the fourth component can affect GaAsN via two mechanisms: the conventional alloying effects and the change in the nitrogen alloying effects.

The most common fourth component is indium, forming InGaAsN. The lattice constant can now be matched with that of the GaAs substrate and indium can relieve the local strain around nitrogen atoms. Since InAs band gap is well below that of GaAs, alloying with indium would be expected to further decrease the band gap, which allows the use of InGaAsN in long wavelength optoelectronic applications. Further alloying with aluminium could be used to increase the band gap.

Alternatively, in the anion matrix, phosphorus would increase the band gap, but also decrease the lattice constant. Also, as the nitrogen states pop out of the conduction band, the properties of the alloy become quite different to GaAsN. Antimony should behave rather similarly to In.

Almost all GaAsN-based quaternaries have some rather unusual properties, but my work has concentrated on the InGaAsN system due to its technological importance and the abundance of the available experimental and computational studies.

4.2.1 Nitrogen short-range order in InGaAsN

The band-gap of InGaAsN fluctuates strongly, which is related to the number of gallium and indium atoms in the neighborhood of nitrogen atoms. Such fluc-

tuations are missing in InGaAs alloys. Already before convincing experimental evidence, Kim and Zunger [92] found with Monte Carlo simulations that the number of In-N bonds in InGaAsN should be higher than in a perfectly random alloy. That is, indium is driven into the nitrogen neighborhood. They also found out a blue-shift related to this ordering. The decrease of Ga-N bonds was reported by Kurtz et al. [93] and Klar et al. [94] were the first to find a discrete "set of band gaps" from different NN configurations, which were then successfully matched to the results from TB calculations. Further studies showed, however, that the amount of short-range ordering is much smaller than predicted by Kim and Zunger [95].

In paper VI, InGaAsN systems are studied with x-ray diffraction and Raman spectroscopy. Vibrational modes related to Ga-N and In-N bonds in Ga_3In_1-N and Ga_4-N configurations can be discerned in the spectra. The changes in concentrations can be estimated from the Raman peak intensities. For the help in the interpretation, DFT calculations were also performed. We used the following simple method to calculate the local vibrational modes: we fixed all the other atoms in the crystal except for the nitrogen atom or nitrogen dimer, which were displaced to a high-symmetry direction or to the direction of the normal mode. The potential curve was fitted to either harmonic potential or Morse-potential for which the energy levels can be analytically calculated. By shifting the calculated mode frequencies up by 35 cm^{-1} we were able to reproduce the most of the Raman frequencies surprisingly well. The shift can be related to the "recoil" of the surrounding crystal, which was disregarded in our simple model. Moreover, we could assign the lower frequency peaks to nitrogen dimer in the gallium and arsenic-sites and the small shift in the former upon annealing was suggested to originate from the indium atom moving farther away from the dimer, which is the energetically favored configuration.

This and the other similar Raman-spectroscopic studies of the local vibrational modes of the nitrogen-centric clusters [96, 97] and PL/ER/XAS studies [98] have given reliable data on the concentration of each clusters types at given temperatures. In a summary, as-grown samples tend to have indium atoms rather randomly distributed somewhat depending on the growth temperature. Upon annealing at moderate annealing temperatures, concentration of Ga_3In_1-N clusters tends to be the most common [93, 92, 99]. At high temperatures, even Ga_0In_4-N clusters can become substantial, although simultaneously other sample properties might be degraded (such as PL due to In diffusion) [98].

The preference of Ga-N bonds during the growth and preference of In-N bonds after annealing can be understood simply by noting that: a) the cohesive energy of GaN is higher than InN energetically favoring Ga-N bonds at the surface and b) the strain relief related to In-N bonds in bulk starts to favor In-N bonds over Ga-N bonds.

Many of the above studies rely on the induced blue-shift on separating the effects of different nitrogen clusters. This blue-shift is also computationally verified by several authors [94, 100, 99]. In [101], the band gap of InGaAsN as a function of both the indium and the nitrogen concentration was studied. It was found out that the band-gap decrease with the introduction of nitrogen is less dramatic at higher In-concentrations, and also more linear. An explanation was suggested in [100]: The nitrogen defect potential is partly created by the lattice strain. The strain is alleviated when the nearest neighbor atoms are indium instead of gallium, making the potential shallower. Consequently, the

alloy becomes more conventional as seen by the increasingly linear band-gap dependence. From the perturbation theory point of view there is of course the first-order perturbation which brings down the band gap, but furthermore the second-order mixing (band repulsion) becomes smaller as the perturbing potential becomes smaller. All in all, the In introduction makes the material more conventional, and the potential becomes shallower. Interestingly, the mechanism is quite the opposite to that found with defects near nitrogen where the localization increases, but the result is still the same: band gap lowering decreases!

As an interesting side note, McGee et al. [102] bring forth another complicating issue: nitrogen also tends to segregate forming quantum dot-like areas. This would of course affect on the interpretation of changes causing the blue-shift; is it due to localization in QD or the short-range ordering?

4.3 Band alignment of (In)GaAsN interfaces

If the band gap fluctuates strongly as a function of nitrogen local environment, then what happens to the band alignment concurrently? In paper IV and paper V we have studied this question by employing the first-principles method described in section 2.4.

4.3.1 GaAsN/GaAs interface

The experimental results by Pan et al. [103] showed that the N concentration does not affect the valence-band offset (VBO) much, although the In concentration does. Also, all throughout the thesis, we have seen how nitrogen has generated large changes only in the conduction band whereas the valence-band has been largely unaffected. This means that the band-offset is small, but it is not clear if the alignment is type-I or type-II and how large it is. In fact, both alignments have been obtained experimentally, although it tends to be accepted nowadays that the alignment is of type-I. See the introduction in paper IV for further discussion.

In paper IV, we studied the GaAsN/GaAs interface. Previous calculations have been using either $\boldsymbol{k} \cdot \boldsymbol{p}$ method [104], tight-binding (to get parameters for extended BAC) [80], and semiempirical pseudopotentials [82]. All of these studies had only estimated the VBM energy in a global energy reference and do not take into account the effects of interface dipole or the strain.

Our model-supercell has 256 atoms, consisting of two 64-atom supercells of GaAsN and two 64-atom supercells of GaAs. The construction is shown in top of Figure 4.2. The nitrogen-nitrogen distance over the "GaAs layer" in our calculation is then six lattice constants or about 34 Å. This corresponds to the average nitrogen-nitrogen distance of a GaAsN alloy with about 0.1% nitrogen concentration. In [47] 0.1% concentration was found to be within the impurity-like region i.e., the nitrogen resonant states do not overlap. Indeed, looking at the electrostatic potential in Figure 4.2, it seems that the potential is well converged to the bulk value; within about one or two unit cells. Notice however, that this is still clearly larger than in the conventional semiconductor alloys, where one or two monolayers is enough (cf. Figure 2.10). It is also interesting to see that the electrostatic potential on the GaAsN region really forms something

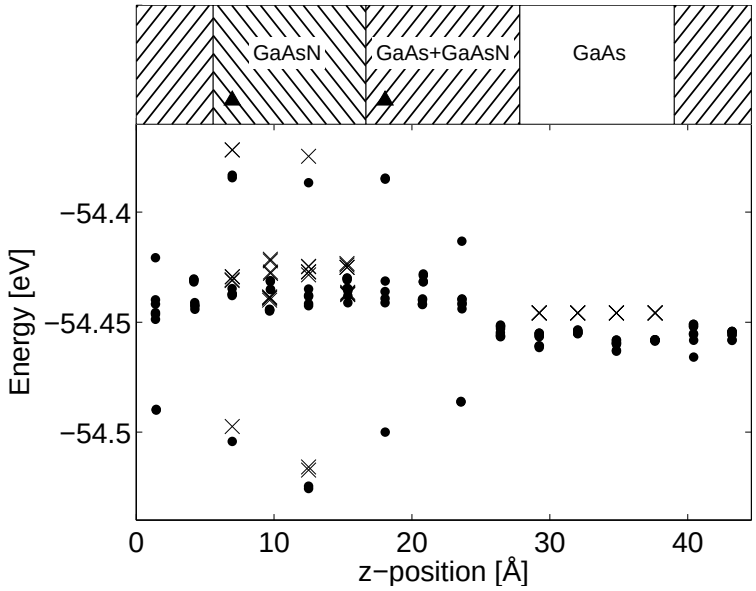

Figure 4.2: The electrostatic potential across a GaAsN/GaAs interface from arsenic cores. Crosses are from the bulk calculations and dots from the interface calculations. The supercell construction is illustrated at the top. Figure from paper IV.

of a plateau (if the few values of large displacement are disregarded). The nitrogen atoms are interacting within this region.

For the results, simply put, we find 35 meV type-I valence-band offset for GaAsN/GaAs interface with about 3% nitrogen. The valence band maximum of the GaAsN bulk lattice-matched to GaAs has a 48 or 65 meV split with or without spin-orbit coupling, respectively. In either case, the split is larger than the band offset. This also strongly suggests that the VBO mostly originates from the VBM splitting due to the anisotropic strain. The small size of nitrogen compared to arsenic, plays the major role again, although this time in a different manner: the change in the lattice constant is large resulting in large VBM split. This is also supported by the fact that the electrostatic potential changes very little in going from the bulk calculation to the interface calculation i.e., from the electrostatics point of view both regions are very similar.

If this is the case, then it seems likely that small changes in the nitrogen positions are not going to change the situation much. Still, the VBM position could fluctuate somewhat in concert with the band gap fluctuations. In our calculations of two different positions for two nitrogen atoms in a 216-atom supercell (positions [110] and [332] by using the notation of paper III) gave the VBM change of 8 meV. These are rather extreme cases and in more realistic models, the fluctuations should be within this range. The reason for this can also be understood from Figure 3.2: nitrogen weight on the VBM is very small.

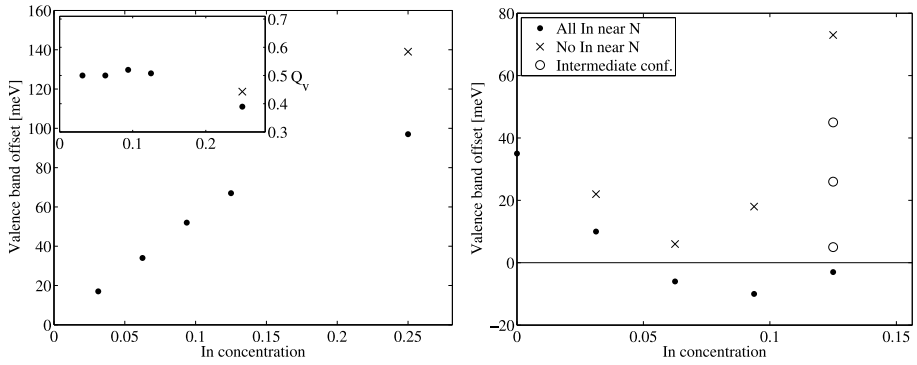

Figure 4.3: The band offsets of the InGaAs/GaAs (a) and the InGaAsN/GaAs (b) interfaces as a function of indium concentration. Figure from paper V.

4.3.2 InGaAsN/GaAs interface

In paper V, we extend the study to the InGaAsN/GaAs interfaces. The above considerations of VBM "stability" do not hold any more. It is known that in InGaAs grown on GaAs the band gap change is divided roughly in half to VBO and CBO. Even though the band gap fluctuation in InGaAsN is a well-studied phenomenon the VBM fluctuation is not.

We used the same method and model as in GaAsN article. The results are shown in Figure 4.3. In InGaAs/GaAs interface, the VBO changes almost linearly up to more than 10% of indium. Moreover, the calculations show that the atomic configuration (even segregation) does not have a large effect on the band offset with conventional alloys like InGaAs. That is, both the electrostatics and VBM are largely unaffected by small changes in the indium distribution.

In the case of InGaAsN/GaAs interface, the VBO is strongly dependent on the number of indium/gallium atoms in the nitrogen nearest neighbor sites. The Figure 4.3b shows the results from configurations where all the indium atoms are at the nearest neighbor sites and the configurations where all the indium atoms are off of the nearest neighbor sites. Moreover, at 12.5% concentration all the configurations with 1–4 indiums at the nearest neighbor sites are shown showcasing a rather smooth transition.

When all the indium atoms are at the nitrogen nearest neighbor sites, the VBO is much smaller At the same concentration on which the strain is at minimum, the VBO also drops down to its minimum value close to zero, and even slightly type-II in our calculations. This reminds us of the GaAsN/GaAs interface, where the band offset was very small and mostly generated by the strain-induced splitting of the VBM. It is as if the indium effect on the VBM had disappeared leaving only the effect of relieving strain. When indium atoms are moved out of the nitrogen vicinity, the VBO increases of similar magnitude as is the VBO of the InGaAs/GaAs interface with the same indium concentration. The strain at the two cases is of course different making direct comparison difficult.

With all indium near nitrogen, the band gap is much larger as discussed in the previous section, making also the CBO smaller in this case. The lowest conduc-

tion band has largest contributions from the s-orbitals of the nitrogen and the surrounding cations. When more indium is brought next to nitrogen, the CBE gains quickly more In4s-character and also loses the strong nitrogen-localization due to the loss of local strain. The wavefunctions become less "deep" and the material becomes conventional. We find these changes of the indium/nitrogen character of the CBE also in our calculations. This can be also seen in Figure 4.4.

This doesn't explain why the VB alignment is also changing. We found out that, when all indium is near nitrogen, the VBO follows closely that of the GaAsN (apart from the effects of strain). This suggests that the conventional alloying effects, which push the VBM up in InGaAs, as described in the beginning of section 3.2.1, effectively disappear. It is clear from the table I of paper V that it is only the VBM (already in the bulk phase) that is pushed up and the interface potential is keeping relatively constant. One possibility would be that some delocalized indium-related states below the VBM in the band structure become more localized near nitrogen which reduces the interaction with the VBM. These indium-related states could come either from the In5s or from the In4d orbitals (deeper states do not form bands). Similar to Ga3d, In4d states are also strongly interacting with the nitrogen, but it seems that the band formation of these states is still too small. Indeed, calculations with Ga3d and $In4d$ electrons frozen in the PAW core resulted in the same VBM shifts. There are several marked changes in the valence band of the indium-projected band structures in Figure 4.4. However, there is no clear delocalized-to-localized transition visible and it is difficult to conclude what causes the largest contribution.

In any case, the localization vs. delocalization effects are playing an interesting game: indium reduces localization of the originally nitrogen-localized state (CBE) and nitrogen would seem to enhance the localization of the originally delocalized indium-related states. The former shifts the CBM up and the latter shifts the VBM down.

Similar to the GaAsN/GaAs interface, the electrostatic potential in Figure 4.5 shows clear plateaus. It seems that the supercell is big enough for the potential convergence.

One has to keep in mind, that the studied configurations are extreme cases. In real samples, there is not going to be only one type of configuration, but a distribution of configurations resulting either in an average VBM or a "fluctuating" VBM, depending on the localization of the VBM states. Unfortunately, no studies exist on the VBM localization, but since it is mostly localized on As4p, it is unlikely to be so largely affected by the interplay of N/In/Ga. Therefore, I suggest the VBO should be taken as a configuration-weighted average.

In our calculations, the band gap closed at high indium concentrations. In the Table I of paper V, negative band gaps are listed for a few configurations. These are obtained by looking at the atomic projections of the bands. If a state, in Γ-point, has mostly Ga-s character it is interpreted as a CBM state and if it has mostly As-p character it is interpreted as a VBM state. Band crossing and hybridization are causing errors to VBM energies, but judging from the band structures (not shown) and the tendencies in band gap and VBO it seems that the errors are not dramatic. Naturally, the extent of these errors is difficult to estimate.

Finally, summing up the results in this chapter, it could be interesting to speculate about the behavior of defects in InGaAsN. If nitrogen is considerably

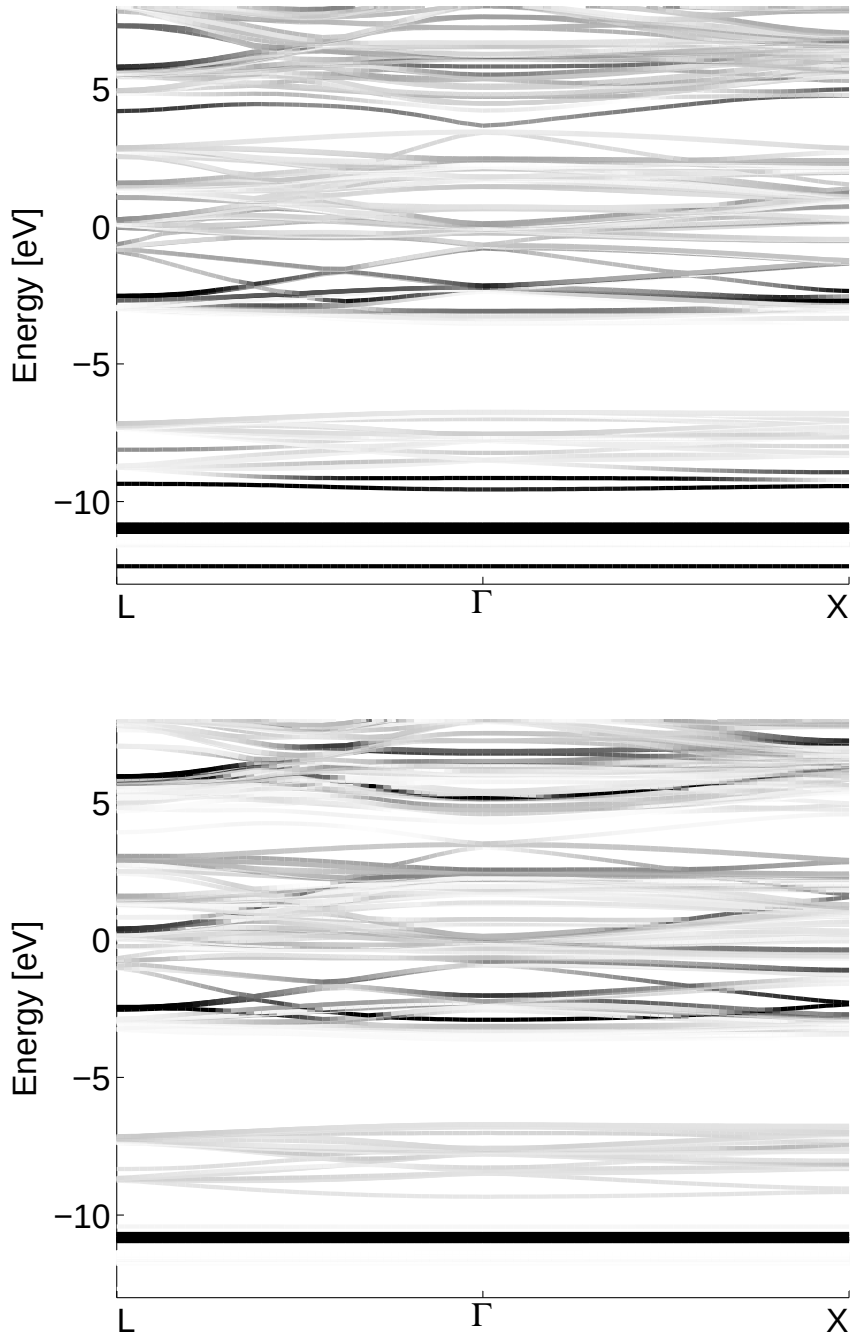

Figure 4.4: Indium projected band structure of all indium near nitrogen and no indium near nitrogen (indium non-segregated). VBM is at about 3 eV.

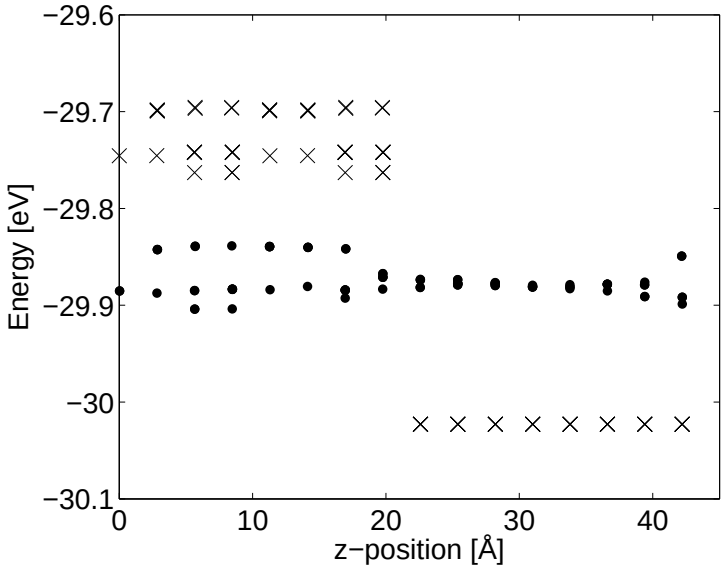

Figure 4.5: The electrostatic potential of the InGaAsN/GaAs interface from the gallium cores. This is the In_4N_1 configuration, with all indium far from nitrogen and segregated. Crosses are from the bulk calculations and dots from the interface calculations.

present in the (N-N) split interstitial mode, it decreases the indium clustering with nitrogen, which could increase the VBO. The interstitials do not contribute to the band gap reduction, so the band gap would be quite big and consequently CBO would be small. In the case of beryllium, it is hard to say what is its effect on the nitrogen-indium clustering as Be also occupies cation sites and (Be-N) split interstitials are further complicating matters.

Conclusions

Overview of the GaAsN material, its interactions with dopants and other alloy components, and effect on the band offsets has been presented mainly from the computational point of view.

The main result of each paper in the context of this thesis is summarized here. First, paper I and paper II aim at providing comprehensive overview of the competition of the interstitial defects and the split-interstitial defects in GaAsN. The formation of split-interstitials is found comparable to the formation of substitutional nitrogen, depending on the conditions. The comparison of calculated and experimental density of states is presented in paper III. The valence-band offsets and the non-abrupt interface potential of GaAsN/GaAs interface is showcased in paper IV, and extended to InGaAsN/GaAs interfaces in paper V. The latter revealed an interesting strong dependence of the valence-band offset on the geometry of nitrogen and indium. The effects of the nitrogen/indium interplay is also studied in paper VI, with the help of Raman spectroscopy and corresponding DFT calculations. Finally, in paper VII, the defect formation in the presence of nitrogen and the effect of defects on the nitrogen alloying properties are investigated.

All throughout the thesis, the competition in the localization vs. delocalization of the states has played a central role. The degree of localization was modified by the presence of beryllium, indium, and other nitrogen atoms (or even interfaces, in a sense), which, in turn, has a strong effect on the properties of GaAsN.

The resarch for this thesis also leaves a few open questions for further research: The surfaces of (In)GaAsN are very poorly studied, let alone the interaction of beryllium and nitrogen on the surfaces during growth. The subject of defects in the vicinity of interfaces is uncharted territory. The midgap correction scheme is still arguable. Finally, combining GW with the finite-size scaling method in the defect calculations should fix both the periodic-image interaction and the LDA band-gap problems.

Bibliography

[1] Jorge Kohanoff. *Electronic Structure Calculations for Solids and Molecules*. Cambridge University Press, 2006.

[2] J. M. Thijssen. *Computational Physics*. Cambridge, 2000.

[3] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos. Iterative minimization techniques for *ab initio* total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64: 1045, 1992.

[4] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864, 1964.

[5] W. Kohn and J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133, 1965.

[6] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566, 1980.

[7] C.-O. Almbladh and U. von Barth. Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues. *Phys. Rev. B*, 531:3231, 1985.

[8] A. Savin, C. J. Umrigar, and Xavier Gonze. Relationship of Kohn-Sham eigenvalues to excitation energies. *Chem. Phys. Lett.*, 288:391, 1998.

[9] John P. Perdew and Mel Levy. Physical content of the exact kohn-sham orbital energies: Band gaps and derivative discontinuities. *Phys. Rev. Lett.*, 51:1884, 1983.

[10] G. Kresse and J. Hafner. *Ab initio* molecular dynamics for liquid metals. *Phys. Rev. B*, 47:558, 1993.

[11] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mat. Sci.*, 6:15, 1996.

[12] G. Kresse and J. Furthmüller. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54: 11169, 1996.

[13] P.W. Atkins. *Physical Chemistry*. Oxford University Press, 1994 (1978).

[14] S. B. Zhang and S.-H. Wei. Nitrogen solubility and induced defect complexes in epitaxial GaAs:N. *Phys. Rev. Lett.*, 86:1789, 2001.

[15] P. Y. Yu and M. Cardona. *Fundamentals of Semiconductors*. Springer-Verlag, Berlin, 2001.

[16] Chris G. Van de Walle and Jörg Neugebauer. First-principles calculations for defects and impurities: Applications to III-nitrides. *J. Appl. Phys.*, 95:3851, 2004.

[17] T. Mattila and Alex Zunger. Deep electronic gap levels induced by isovalent P and As impurities in gan. *Phys. Rev. B*, 58:1367, 1998.

[18] Juha Markus Lento. *First-principles calculations of native defects in tetrahedrally-coordinated isovalent semiconductors*. PhD thesis, Helsinki University of Technology, 2003.

[19] G. Makov and M. C. Payne. Periodic boundary conditions in *ab initio* calculations. *Phys. Rev. B*, 51:4014, 1995.

[20] D. Segev and S.-H. Wei. Design of shallow donor levels in diamond by isovalent-donor coupling. *Phys. Rev. Lett.*, 91:126406, 2003.

[21] Joachim Paier, Robin Hirschl, Martijn Marsman, and Georg Kresse. The Perdew-Burke-Ernzerhof exchange-correlation functional applied to the G2-1 test set using a plane-wave basis set. *J. Chem. Phys.*, 122:234102, 2005.

[22] C. W. M. Castleton and S. Mirbt. Finite-size scaling as a cure for supercell approximation errors in calculations of neutral native defects in InP. *Phys. Rev. B*, 70:195202, 2004.

[23] C. W. M. Castleton, A. Höglund, and S. Mirbt. Managing the supercell approximation for charged defects in semiconductors: Finite-size scaling, charge correction factors, the band-gap problem, and the *ab initio* dielectric constant. *Phys. Rev. B*, 73:035215, 2006.

[24] P. Deák, Th. Frauenheim, and A. Gali. Limits of the scaled shift correction to levels of interstitial defects in semiconductors. *Phys. Rev. B*, 75:153204, 2007.

[25] Su-Huai Wei. Overcoming the doping bottleneck in semiconductors. *Comput. Mat. Sci.*, 30:337, 2004.

[26] A. Amore Bonapasta, F. Filippone, and P. Giannozzi. Nitrogen passivation by atomic hydrogen in $GaAs_yN_{1-y}$ and $In_xGa_{1-x}As_yN_{1-y}$ alloys. *Phys. Rev. B*, 68:115202, 2003.

[27] A. Franciosi and C. G. Van de Walle. Heterojunction band offset engineering. *Surf. Sci. Rep.*, 25:1, 1996.

[28] R. L. Anderson. Germanium-gallium arsenide heterojunctions. *IBM J. Res. Dev.*, 4:283, 1960.

[29] A. Baldereschi, S. Baroni, and R. Resta. Band offsets in lattice-matched heterojunctions: A model and first-principles calculations for GaAs/AlAs. *Phys. Rev. Lett.*, 61:734, 1988.

[30] C. G. Van de Walle and R. M. Martin. Theoretical study of band offsets at semiconductor interfaces. *Phys. Rev. B*, 35:8154, 1987.

[31] S.-H. Wei and A. Zunger. Calculated natural band offsets of all II-VI and III-V semiconductors: Chemical trends and the role of cation *d* orbitals. *Appl. Phys. Lett.*, 72:2011, 1998.

[32] J. Tersoff. Theory of semiconductor heterojunctions: The role of quantum dipoles. *Phys. Rev. B*, 30:4874, 1984.

[33] W. A. Harrison and J. Tersoff. Tight-binding theory of heterojunction band lineups and interface dipoles. *J. Vac. Sci. Technol. B*, 4:1068, 1986.

[34] W. R. L. Lambrecht and B. Segall. Theory of semiconductor heterojunction valence-band offsets: From supercell band-structure calculations toward a simple model. *Phys. Rev. Lett.*, 61:1764, 1988.

[35] J. M. Bass, M. Oloumi, and C. C. Matthai. A method for determining band offsets in semiconductor superlattices and interfaces. *J. Phys.: Condens. Matter*, 1:10625, 1989.

[36] D. M. Bylander and L. Kleinman. *Ab initio* $(GaAs)_3(AlAs)_3(001)$ superlattice calculations: Band offsets and formation enthalpy. *Phys. Rev. B*, 36:3229, 1987.

[37] R. Shaltaf, G.-M. Rignanese, X. Gonze, Feliciano Giustino, and Alfredo Pasquarello. Band offsets at the Si/SiO_2 interface from many-body perturbation thery. *Phys. Rev. Lett.*, 100:186401, 2008.

[38] Xuejun Zhu and Steven G. Louie. Quasiparticle band structure of thirteen semiconductors and insulators. *Phys. Rev. B*, 43:14142, 1991.

[39] Su-Huai Wei and Alex Zunger. Role of *d* orbitals in valence-band offsets of common-anion semiconductors. *Phys. Rev. Lett.*, 59:144, 1987.

[40] S. B. Zhang, D. Tománek, Steven G. Louie, Marvin L. Cohen, and Mark S. Hybertsen. Quasiparticle calculation of valence band offset of AlAs-GaAs(001). *Solid State Commun.*, 66:585, 1988.

[41] K. Karlsson, R. J. Needs, A. Qteish, and R. W. Godby. Tailoring materials for quantum wells: band offsets at (001)-oriented $GaAs/(AlAs)_n(GaAs)_m$ interfaces. *J. Phys.: Condens. Matter*, 2:5265, 1990.

[42] J. E. Jaffe, M. Dupuis, and M. Gutowski. First-principles study of non-commutative band offsets at α-Cr_2O_3/α-$Fe_2O_3(0001)$ interfaces. *Phys. Rev. B*, 69:205106, 2004.

[43] E. A. Kraut, R. W. Grant, J. R. Waldrop, and S. P. Kowalczyk. Precise determination of the valence-band edge in x-ray photoemission spectra: Application to measurement of semiconductor interface potentials. *Phys. Rev. Lett.*, 44:1620, 1980.

[44] D. M. Bylander and L. Kleinman. Comparison of dipole layers, band offsets, and formation enthalpies of GaAs-AlAs (110) and (001) interfaces. *Phys. Rev. Lett.*, 59:2091, 1987.

[45] J. D. Perkins, A. Mascarenhas, Yong Zhang, J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz. Nitrogen-activated transitions, level repulsion, and band gap reduction in $GaAs_{1-x}N_x$ with $x < 0.03$. *Phys. Rev. Lett.*, 82:3312, 1999.

[46] T. Taliercio, R. Intartaglia, B. Gil, P. Lefebvre, T. Bretagnon, U. Tisch, E. Finkman, J. Salzman, M.-A. Pinault, M. Laügt, and E. Tournié. From GaAs:N to oversaturated GaAsN: Analysis of the band-gap reduction. *Phys. Rev. B*, 69:073303, 2004.

[47] P. R. C. Kent and Alex Zunger. Theory of electronic structure evolution in GaAsN and GaPN alloys. *Phys. Rev. B*, 64:115208, 2001.

[48] V. Fiorentini, M. Methfessel, and M. Scheffler. Electronic and structural properties of GaN by the full-potential linear muffin-tin orbitals method: The role of the *d* electrons. *Phys. Rev. B*, 47:13353, 1993.

[49] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow. Theory of substitutional deep traps in covalent semiconductors. *Phys. Rev. Lett.*, 44:810, 1980.

[50] L.-W. Wang. Large-scale local-density-approximation band gap-corrected GaAsN calculations. *Appl. Phys. Lett.*, 78:1565, 2001.

[51] E. P. O'Reilly, A. Lindsay, S. Tomić, and M. Kamal-Saadi. Tight-binding and k·p models for the electronic structure of Ga(In)NAs and related alloys. *Semicond. Sci. Technol.*, 17:870, 2002.

[52] J. E. Lowther, S. K. Estreicher, and H. Temkin. Nitrogen-related complexes in gallium arsenide. *Appl. Phys. Lett.*, 79:200, 2001.

[53] P. Carrier, S.-H. Wei, S. B. Zhang, and S. Kurtz. Evolution of structural properties and formation of N-N split interstitials in $GaAs_{1-x}N_x$ alloys. *Phys. Rev. B*, 71:165212, 2005.

[54] A. Janotti, P. Reunchan, S. Limpijumnong, and C. G. Van de Walle. Mutual passivation of electrically active and isovalent impurities in dilute nitrides. *Phys. Rev. Lett.*, 100:045505, 2008.

[55] Y. Qiu, S. A. Nikishin, H. Temkin, V. A. Elyukhin, and Yu. A. Kudriavtsev. Thermodynamic considerations in epitaxial growth of $GaAs_{1-x}N_x$ solid solutions. *Appl. Phys. Lett.*, 70:2831, 1997.

[56] Y. Qiu, S. A. Nikishin, H. Temkin, N. N. Faleev, and Yu. A. Kudriavtsev. Growth of single phase $GaAs_{1-x}N_x$ with high nitrogen concentration by metal-organic molecular beam epitaxy. *Appl. Phys. Lett.*, 70:3242, 1997.

[57] I. A. Buyanova, W. M. Chen, and B. Monemar. Electronic properties of Ga(In)NAs alloys. *MRS Internet J. Nitride Semicond. Res.*, 6:2, 2001.

[58] M. Le Dû, J.-C. Harmand, O. Mauguin, L. Largeau, L. Travers, and J.-L. Oudar. Quantum-well saturable absorber at 1.55 μm on gaas substrate with a fast recombination rate. *Appl. Phys. Lett.*, 88:201110, 2006.

[59] Mircea D. Guina, Pietari Tuomisto, Oleg G. Okhotnikov, Saulius Marcinkevicius, Kenichiro Mizohata, and Juhani Keinonen. Semiconductor saturable absorbers with recovery time controlled through growth conditions. *Proceedings of SPIE*, 6451:645113, 2007.

[60] O. Rubel, K. Volz, T. Torunski, S. D. Baranovskii, F. Grosse, and W. Stolz. Columnar [001]-oriented nitrogen order in Ga(NAs) and (GaIn)(NAs) alloys. *Appl. Phys. Lett.*, 85:5908, 2004.

[61] Mao-Hua Du, Sukit Limpijumnong, and S. B. Zhang. Hydrogen-mediated nitrogen clustering in dilute III-V nitrides. *Phys. Rev. Lett.*, 97:075503, 2006.

[62] Jörg Neugebauer and Chris G. Van de Walle. Electronic structure and phase stability of $GaAs_{1-x}N_x$ alloys. *Phys. Rev. B*, 51:10568, 1995.

[63] S.-H. Wei and A. Zunger. Giant and composition-dependent optical bowing coefficient in GaAsN alloys. *Phys. Rev. Lett*, 76:664, 1996.

[64] N. Gonzalez Szwacki and P. Bogusławski. GaAs:N vs GaAs:B alloys: Symmetry-induced effects. *Phys. Rev. B*, 64:161201, 2001.

[65] T. Mattila, Su-Huai Wei, and Alex Zunger. Localization and anticrossing of electron levels in $GaAs_{1-x}N_x$ alloys. *Phys. Rev. B*, 60:R11245, 1999.

[66] W. Shan, W. Walukiewicz, J. W. Ager III, E. E. Haller J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz. Band anticrossing in GaInNAs alloys. *Phys. Rev. Lett.*, 82:1221, 1999.

[67] P.W. Atkins and R.S. Friedman. *Molecular Quantum Mechanics*. Oxford University Press, 1997 (1970).

[68] Markus Weyers, Michio Sato, and Hiroaki Ando. Red shift of photoluminescence and absorption in dilute GaAsN alloy lasers. *Jpn. J. Appl. Phys*, 31:L853, 1992.

[69] Masahiko Kondow, Kazuhisa Uomi, Kazuhiko Hosomi, and Teruo Mozume. Gas-source molecular beam epitaxy of GaN_xAs_{1-x} using a N radical as the N source. *Jpn. J. Appl. Phys*, 33:L1056, 1994.

[70] K. Uesugi, I. Suemune, T. Hasegawa, T. Akutagawa, and T. Nakamura. Temperature dependence of band gap energies of GaAsN alloys. *Appl. Phys. Lett.*, 76:1285, 2000.

[71] B. M. Keyes, J. F. Geisz, P. C. Dippo, R. Reedy, C. Kramer, D. J. Friedman, S. R. Kurtz, and J. M. Olson. Optical investigation of GaNAs. *AIP Conf. Proc.*, 462:511, 1999.

[72] L. Malikova, F. H. Pollak, and R. Bhat. Composition and temperature dependence of the direct band gap of $GaAs_{1-x}N_x$ ($0<x<0.0232$) using contactless electroreflectance. *J. Electron. Mater.*, 27:484, 1998.

[73] R. Bhat, C. Caneau, L. Salamanca-Riba, W. Bi, and C. Tu. Growth of GaAsN/GaAs, GaInAsN/GaAs and GaInAsN/GaAs quantum wells by low-pressure organometallic chemical vapor deposition. *J. Cryst. Growth*, 195:427, 1998.

[74] E. P. O'Reilly, A. Lindsay, and S. Fahy. Theory of the electronic structure of dilute nitride alloys: beyond the band-anti-crossing model. *J. Phys.: Condens Matter*, 16:S3257, 2004.

[75] A. Polimeni, F. Masia, G. Pettinari, R. Trotta, M. Felici, M. Capizzi, A. Lindsay, E. P. O'Reilly, T. Niebling, W. Stolz, and P. J. Klar. Role of strain and properties of N clusters at the onset of the alloy limit in $GaAs_{1-x}N_x$. *Phys. Rev. B*, 77:155213, 2008.

[76] P. H. Tan, X. D. Luo, Z. Y. Xu, Y. Zhang, A. Mascarenhas, H. P. Xin, C. W. Tu, and W. K. Ge. Photoluminescence from the nitrogen-perturbed above-bandgap states in dilute $GaAs_{1-x}N_x$ alloys: A microphotoluminescence study. *Phys. Rev. B*, 73:205205, 2006.

[77] B. Fluegel, A. Mascarenhas, A. J. Ptak, S. Tixier, E. C. Young, and T. Tiedje. e_+ transition in $GaAs_{1-x}N_x$ and $GaAs_{1-x}Bi_x$ due to isoelectronic-impurity-induced perturbation of the conduction band. *Phys. Rev. B*, 76:155209, 2007.

[78] S. Francoeur, M. J. Seong, M. C. Hanna, J. F. Geisz, A. Mascarenhas, H. P. Xin, and C. W. Tu. Origin of the nitrogen-induced optical transitions in $GaAs_{1-x}N_x$. *Phys. Rev. B*, 68:075207, 2003.

[79] V. Timoshevskii, M. Côté, G. Gilbert, R. Leonelli, S. Turcotte, J.-N. Beaudry, P. Desjardins, S. Larouche, L. Martinu, and R. A. Masut. Experimental and theoretical studies of the E_+ optical transition in GaAsN alloys. *Phys. Rev. B*, 74:165120, 2006.

[80] A. Lindsay and E. P. O'Reilly. Theory of enhanced bandgap nonparabolicity in GaN_xAs_{1-x} and related alloys. *Solid State Commun.*, 112: 443, 1999.

[81] N. Shtinkov, P. Desjardins, and R. A. Masut. Empirical tight-binding model for the electronic structure of dilute GaNAs alloys. *Phys. Rev. B*, 67:081202, 2003.

[82] L. Bellaiche, S.-H. Wei, and A. Zunger. Localization and percolation in semiconductor alloys: GaAsN vs GaAsP. *Phys. Rev. B*, 54:17568, 1996.

[83] E. D. Jones, N. A. Modine, A. A. Allerman, S. R. Kurtz, A. F. Wright, S. T. Tozer, and X. Wei. Band structure of $In_xGa_{1-x}As_{1-y}N_y$ alloys and effects of pressure. *Phys. Rev. B*, 60:4430, 1999.

[84] M. Kozhevnikov, V. Narayanamurti, C. V. Reddy, H. P. Xin, C. W. Tu, A. Mascarenhas, and Y. Zhang. Evolution of $GaAs_{1-x}N_x$ conduction states and giant $Au/GaAs_{1-x}N_x$ schottky barrier reduction studied by ballistic electron emission spectroscopy. *Phys. Rev. B*, 61:R7861, 2000.

[85] A. Janotti, S. B. Zhang, S.-H. Wei, and C. G. Van de Walle. Effects of hydrogen on the electronic properties of dilute GaAsN alloys. *Phys. Rev. Lett.*, 89:086403, 2002.

[86] A. Janotti, S.-H. Wei, S. B. Zhang, S. Kurtz, and C. G. Van de Walle. Interactions between nitrogen, hydrogen, and gallium vacancies in $GaAs_{1-x}N_x$ alloys. *Phys. Rev. B*, 67:161201, 2003.

[87] G. Baldassarri H. v. H., M. Bissiri, A. Polimeni, M. Capizzi, M. Fischer, M. Reinhardt, and A. Forchel. Hydrogen-induced band gap tuning of (InGa)(AsN)/GaAs single quantum wells. *Appl. Phys. Lett.*, 78:3472, 2001.

[88] G. Bisognin, D. De Salvador, A. V. Drigo, E. Napolitani, A. Sambo, M. Berti, A. Polimeni, M. Felici, M. Capizzi, M. Güngerich, P. J. Klar, G. Bais, F. Jabeen, M. Piccin, S. Rubini, F. Martelli, and A. Franciosi. Hydrogen-nitrogen complexes in dilute nitride alloys: Origin of the compressive lattice strain. *Appl. Phys. Lett.*, 89:061904, 2006.

[89] Mao-Hua Du, Sukit Limpijumnong, and S. B. Zhang. Hydrogen pairs and local vibrational frequencies in H-irradiated $GaAs_{1-y}N_y$. *Phys. Rev. B*, 72:073202, 2005.

[90] K. M. Yu, W. Walukiewicz, J. Wu, D. E. Mars, D. R. Chamberlin, M. A. Scarpulla, O. D. Dubon, and J. F. Geisz. Mutual passivation of electrically active and isovalent impurities. *Nat. Mater.*, 1:185, 2002.

[91] Jingbo Li, Pierre Carrier, Su-Huai Wei, Shu-Shen Li, and Jian-Bai Xia. Mutual passivation of donors and isovalent nitrogen in GaAs. *Phys. Rev. Lett.*, 96:035505, 2006.

[92] K. Kim and A. Zunger. Spatial correlations in gainasn alloys and their effects on band-gap enhancement and electron localization. *Phys. Rev. Lett.*, 86:2609, 2001.

[93] S. Kurtz, J. Webb, L. Gedvilas, D. Friedman, J. Geisz, J. Olson, R. King, D. Joslin, and N. Karam. Structural changes during annealing of GaInAsN. *Appl. Phys. Lett.*, 78:748, 2001.

[94] P. J. Klar, H. Grüning, J. Koch, S. Schäfer, K. Volz, W. Stolz, W. Heimbrodt, A. M. Kamal Saadi, A. Lindsay, and E. P. O'Reilly. (Ga,In)(N,As)-fine structure of the band gap due to nearest-neighbor configurations of the isovalent nitrogen. *Phys. Rev. B*, 64:121203, 2001.

[95] G. Ciatto, F. D'Acapito, L. Grenouillet, H. Mariette, D. De Salvador, G. Bisognin, R. Carboni, L. Floreano, R. Gotter, S. Mobilio, and F. Boscherini. Quantitative determination of short-range ordering in $In_xGa_{1-x}As_{1-y}N_y$. *Phys. Rev. B*, 68:161201, 2003.

[96] H. Ch. Alt, Y. V. Gomeniuk, and G. Mussler. Splitting of the local mode frequency of substitutional nitrogen in (Ga,In)(As,N) alloys due to symmetry lowering. *Semicond. Sci. Technol.*, 21:1425, 2006.

[97] A. M. Teweldeberhan and S. Fahy. Effect of indium-nitrogen bonding on the localized vibrational mode in $In_yGa_{1-y}N_xAs_{1-x}$. *Phys. Rev. B*, 73: 245215, 2006.

[98] V. Lordi, H. B. Yuen, S. R. Bank, M. A. Wistey, J. S. Harris, and S. Friedrich. Nearest-neighbor distributions in $Ga_{1-x}In_xN_yAs_{1-y}$ and $Ga_{1-x}In_xN_yAs_{1-y-z}Sb_z$ thin films upon annealing. *Phys. Rev. B*, 71: 125309, 2005.

[99] A. Jenichen, C. Engler, G. Leibiger, and V. Gottschalch. Growth and annealing of GaInAsN: density-functional calculations on the reactions of surface and bulk structures. *Surf. Sci.*, 574:144, 2005.

[100] V. Lordi, V. Gambin, S. Friedrich, T. Funk, T. Takizawa, K. Uno, and J. S. Harris. Nearest-neighbor configuration in (GaIn)(NAs) probed by x-ray absorption spectroscopy. *Phys. Rev. Lett.*, 90:145505, 2003.

[101] J.-Y. Duboz, J. A. Gupta, Z. R. Wasilewski, J. Ramsey, R. L. Williams, G. C. Aers, B. J. Riel, and G. I. Sproule. Band-gap energy of $In_xGa_{1-x}N_yAs_{1-y}$ as a function of N content. *Phys. Rev. B*, 66:085313, 2002.

[102] W. M. McGee, R. S. Williams, M. J. Ashwin, T. S. Jones, E. Clarke, J. Zhang, and S. Tomić. Structure, morphology, and optical properties of $Ga_xIn_{1-x}N_{0.05}As_{0.95}$ quantum wells: Influence of the growth mechanism. *Phys. Rev. B*, 76:085309, 2007.

[103] Z. Pan, L. H. Li, Y. W. Lin, B. Q. Sun, D. S. Jiang, and W. K. Ge. Conduction band offset and electron effective mass in GaInNAs/GaAs quantum-well structures with low nitrogen concentration. *Appl. Phys. Lett.*, 78:2218, 2001.

[104] J. Wu, W. Walukiewicz, K. M. Yu, J. D. Denlinger, W. Shan, J. W. Ager, III, A. Kimura, H. F. Tang, and T. F. Kuech. Valence band hybridization in N-rich $GaN_{1-x}As_x$ alloys. *Phys. Rev. B*, 70:115214, 2004.