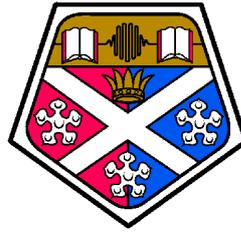# University of Strathclyde

Department of Physics
Semiconductor Spectroscopy and Devices

Native defects in the Group III Nitrides

PhD Thesis

Simon Kraeusel

May 30, 2013

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed: Date: May 30, 2013

Contents

List of Figures

List of Tables

As first author

Global search for stable screw dislocation cores in III-N semiconductors.
Kraeusel, Simon; Hourahine, Benjamin
In: Physica Status Solidi A, Vol. 209, No. 1, 01.2012, p. 71-74.

Contributions

Defect evolution and interplay in n-type InN.
Rauch, Christian; Tuomisto, Filip; Vilalta-Clemente, Arantxa; Lacroix, Bertrand; Ruterana, Pierre; Kraeusel, Simon; Hourahine, Benjamin; Schaff, William J.
In: Applied Physics Letters, Vol. 100, No. 9, 2012, p. 091907.

Electron channeling contrast imaging studies of nonpolar nitrides using a scanning electron microscope
G. Naresh-Kumar, C. Mauder, K. R. Wang, S. Kraeusel, J. Bruckbauer, P. R. Edwards, B. Hourahine, H. Kalisch, A. Vescan, C. Giesen, M. Heuken, A. Trampert, A. P. Day, and C. Trager-Cowan
In: Applied Physics Letters, Vol. 102, No. 14, 2013, p. 142103.

Abstract

The promise of the broad range of direct band gaps of the (Al,Ga,In)N system is limited by the crystal quality of current material. As grown defect densities of InN, when compared with the more mature GaN, are extremely high and InN is strongly influenced by these defects. This is particularly important due to the unusual position of the charge neutrality level of InN, leading to both the well known surface charge accumulation and difficulties in p-type doping. While impurities and native defects clearly impact on the bulk carrier density in InN, the effects of threading dislocations on the electrical properties are still in dispute. Issues such as whether the dislocation line is charged or contains dangling bonds remain open.

In this work an empirical Stillinger-Weber inter-atomic potential method is employed in a systematic global search for possible dislocation core reconstructions for screw and edge dislocations in GaN. The global optimisation of the dislocation cores is performed for a wide variety of core stoichiometries ranging from Ga rich to N rich. The most promising optimised core configurations are subsequently investigated using density functional theory for GaN and InN, in order to discuss relative stability under a wide range of growth conditions and their influence on the electronic properties of the bulk material.

Introduction

1.1   Motivation

The group-III nitrides are direct bandgap semiconductors that are ideal for opto-electronic devices as the bandgap range of InN, GaN, AlN and their alloys encompasses the entire visible spectrum as it extends from the infrared (InN) into the far ultraviolet (AlN). Gallium rich InGaN quantum well structures are already in wide commercial use as high efficiency light emitting diodes (LEDs) and laser diodes (LDs). These devices are used for lighting, sensors, data storage (Blu-ray™ technology), biosensing, sterilisation, etc. In addition the narrow band gap (<0.7 eV) of InN could al-

Figure 1.1: Band gaps of several semiconductor compounds, drawn versus lattice constant. The visible spectrum has been included as colored background. Courtesy of [1]

low the development of the highest efficiency solar cells and production of LEDs with tuneable colour temperatures [3]. Moreover, due to an electron mobility

1

of around >3000 cm^2/Vs [4] and very high saturation velocities, InN is an ideal material for the development of high electron mobility devices capable of operating in the Terahertz range. However, even though high efficiency LEDs can be constructed from InGaN alloys, there still exists a very strong efficiency droop for the indium rich end. This results in the current impossibility of achieving true white light LEDs, as currently available devices use the blue emitting gallium rich InGaN structures in combination with a yellow phosphor coating. The ability to create efficient lighting at arbitrary wavelengths in the visible light spectrum not only increases a devices white light colour rendering quality, but also enables their use in many more lighting applications, which can reduce a countries total elictricity consumption by up to 15% [5].

A major source of the efficiency droop in the InGaN alloys are unintentional defects, such as point defects and extended defects. The effects of impurities and native point defects have been studied in order to better control growth conditions to minimise their impact on the efficiency of devices[6]. Due to the large lattice mismatch of the usual growth substrates such as sapphire GaN, InN and their alloys feature very high densities of extended defects. While these defects seem to have little effect on the efficiency of gallium rich structures, their influence is thought to be much greater in the In rich end. Extended defects such as dislocations and stacking faults act as scattering centres for electrons due to the lattice distortion itself and can act as sources of non-radiative electron hole recombination centres. In addition dislocations can have a major impact on the electronic properties, depending on their core structure, as questions such as whether dislocation lines are charged or metal filled are still in dispute.

In order to study the influence of dislocations on GaN and InN, a two fold approach was taken in this thesis. The first step is to systematically determine all feasible dislocation core structures across the range of nitrogen to group-III rich conditions, while the second step is to determine the electronic properties of the most stable solutions and their influence on the semiconductor. The structural investigation incorporates a global optimisation of the nuclei positions in the dislocation core of screw and edge type perfect dislocations. This is done with a random search algorithm previously developed by Pickard and Needs [7], where structural relaxations were performed exclusively with density functional

theory (DFT). In contrast, within this work, the local structural relaxations were performed with a very fast empirical Stillinger Weber type interatomic potentials method. A more reliable DFT atomic position relaxation was not considered, due to the extreme amounts of computational efforts required. However the subsequent investigation of the optimised structures was performed using DFT.

1.2   Thesis structure

This thesis is divided into 7 chapters, where the first 3 serve to introduce the applied techniques as well as the material system, the results of this work are then presented in the following 3 chapters while the impact and implications of these are discussed in the last chapter. The theoretical and material background is strictly divided into chapters 2 (theory) and 3 (material) which serve to introduce all implemented concepts that are used for the generation and discussion of the results. Chapter 4 presents the results for a global optimisation applied to the fully stoichiometric screw dislocation core in GaN and discusses all arising features and difficulties of the application of the optimisation technique on this particular core. Chapter 5 then presents a comparative study of optimised screw dislocation cores for a wide range of stoichiometries in GaN as well as an additional investigation of the electronic structure of a smaller selection of stable cores for GaN and InN using Density Functional Theory. The specific application of the global optimisation to every core stoichiometry is not discussed in chapter 5, but the relevant data can be found in appendix B. Additional structural data for each lowest energy (optimised) core can be found in appendix C. Chapter 6 then continues to present results of the global optimisation technique applied to edge dislocation cores in GaN, while chapter 7 serves to present a summary, conclusion and possible future work.

Theoretical framework

This Chapter introduces concepts used to derive the results presented later in this thesis. Brief reviews of Density Functional Theory (DFT), Stillinger Weber (SW) type interatomic potentials, global optimisation, the fundamentals of elasticity theory and an overview of the basics of dislocation models are given. With each topic the specific approximations, functionals, parameters or assumptions that have been adopted for this work, as well as the external computational codes used, are presented and discussed.

2.1 Density Functional Theory

Density Functional Theory (DFT) is a widely used method to obtain electronic properties in solid state physics and quantum chemistry. With the use of more refined approximations, DFT is able to predict material properties with a satisfactory accuracy and lower computational cost compared to traditional ab-initio methods such as Hartree Fock theory.

DFT builds on two major principles to solve many body quantum mechanical systems, the Hohenberg Kohn theorem and the Kohn Sham equations [8]. The following sections give a short review of these principles and lay out specifics of the implementation of the AIMPRO code that has been used in this thesis.

2.1.1   Many body Hamiltonian

Henceforth electron coordinates are indexed by lower case letters (i,j,k) while atomic nuclear positions and charges are labeled with upper case letters (I,J,K). Similarly, single particle wavefunctions are described by ψ or ϕ and Ψ or Φ represent the many-body case.

The time independent Schrödinger equation for a single particle in a general (time independent) potential $V(\mathbf{r})$ takes the form

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{2.1}$$

where the first term describes the kinetic energy, which when combined with the potential forms a Hamiltonian that is equal to a constant energy term E. The resulting differential equation can be solved for the wavefunction $\psi(\mathbf{r})$, that describes a probability density of finding the particle at a point $\mathbf{r}$.

However in order to explore a solid composed of nuclei and electrons the wavefunction for a many-particle system needs to be determined. For such system a many-body Schrödinger equation can be constructed and simplified by applying the Born-Oppenheimer approximation. It states that effects of the nuclear motion can be neglected with the assumption that electron masses are sufficiently small compared to those of the nuclei. The kinetic energy for the nuclei can thus be neglected and the inter-nuclear coulomb potential becomes a constant. Consequently the N-body Schrödinger equation takes the following form in atomic units[1]:

$$\left(-\frac{1}{2}\nabla^2 + \frac{1}{2}\sum_i\sum_j V_{ee}(\mathbf{r}_i,\mathbf{r}_j) + \sum_i V_{ext}(\mathbf{r}_i)\right)\Psi(\mathbf{r}_1,...,\mathbf{r}_N) = E\Psi(\mathbf{r}_1,...,\mathbf{r}_N), \tag{2.2}$$

here V_{ee} represents the inter electron Coulombic potential while V_{ext} stands for the electron-nuclear Coulomb interaction:

$$V_{ee}(\mathbf{r}_i,\mathbf{r}_j) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad \text{and} \quad V_{ext}(\mathbf{r}_i) = \sum_I \frac{Z_I}{\mathbf{r}_i - \mathbf{R}_I}. \tag{2.3}$$

---

[1]   Energy in Hartrees $H = \hbar/(ma_0^2) \simeq 27.21\text{eV}$, distances in Bohr radii $a_0 = 4\pi\epsilon_0\hbar^2/(me^2) \simeq 0.5292\text{Å}$ and $e = m = \hbar = 4\pi\epsilon_0 = 1$

A many body electron wavefunction, such as the solution to equation (2.2) is fermionic in nature and is antisymmetric with respect to interchange of any electron pair. Methods to solve the many body Schrödinger equation, such as Hartree Fock theory usually explicitly construct such a wavefunction by combining many single particle functions with the correct symmetry. This can be done with a determinant construction and is then called a Slater determinant.

2.1.2 Hohenberg-Kohn theorem

The many-electron Schrödinger equation of a given system can also be solved purely from the electron density instead of requiring the more complex interacting electron wavefunction. A proof that this is possible is given by the Hohenberg Kohn theorem, which states that the ground state electron density of a non-degenerate system is unique and determines all of the information contained within the wavefunction[9].

The definition of the ground state results in the following statement as, due to the variational principle, any wavefunction either represents a ground state or is higher in energy than the ground state energy E_0:

$$\langle\Psi|\hat{T} + \hat{V}_{ee} + \hat{V}_{ext}|\Psi\rangle \geq E_0, \tag{2.4}$$

where the kinetic energy term has been abbreviated with T and the Bra-ket notation has been used, in which $|\Psi>$ represents a vector with the vector product $\langle\Psi|\hat{T}|\Psi\rangle = \int \Psi^*(\mathbf{r})T\Psi(\mathbf{r})$ including an acting operator denoted by $\hat{}$. The electron density n for a many body wavefunction is defined as

$$n(\mathbf{r}) = N \int \Psi^*(\mathbf{r}\mathbf{r}_2...\mathbf{r}_N)\Psi(\mathbf{r}\mathbf{r}_2...\mathbf{r}_N) \, d\mathbf{r}_2...d\mathbf{r}_N. \tag{2.5}$$

Assuming a fermionic wavefunction (antisymmetry) and identifying the local nature of the external potential $\hat{V}_{ext}$ one can write for the last term of equation (2.4)

$$\begin{aligned}\langle\Psi|\hat{V}_{ext}|\Psi\rangle &= \int \Psi^*(\mathbf{r}_1...\mathbf{r}_N) \sum_i V_{ext}(\mathbf{r}_i)\Psi(\mathbf{r}_1...\mathbf{r}_N) \, d\mathbf{r}_i...d\mathbf{r}_N \\ &= \int n(\mathbf{r})V_{ext}(\mathbf{r}) \, d\mathbf{r} = V_{ext}[n].\end{aligned} \tag{2.6}$$

Where $V_{ext}[n]$ is a functional of the density n as it maps a function space to $\mathbb{R}$. The remaining part of equation (2.4) can also be written as a functional of n by defining

$$F[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle. \tag{2.7}$$

Where $\Psi \to n$ indicates that the minimisation is with respect to all Ψ that generate the given density n. Together equations (2.6) and (2.7) represent the Energy functional

$$
\begin{aligned}
E[n] &= \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle + \int n(\mathbf{r}) V_{ext}(\mathbf{r}) \, d\mathbf{r} \tag{2.8} \\
&= F[n] + V_{ext}[n]. \tag{2.9}
\end{aligned}
$$

The Hohenberg Kohn Theorem makes two statements with respect to the energy functional. The first is equivalent to the inequality (2.4) and states

$$E[n] = F[n] + V_{ext}[n] \geq E_0 \tag{2.10}$$

and the second is that

$$F[n_0] + V_{ext}[n_0] = E_0, \tag{2.11}$$

where n_0 is the ground state energy, constructed from the ground state wavefuncton Ψ_0.[2] As an important result the inverse is also true, a non degenerate ground state wavefunction can be constructed from the ground state charge density n_0. This was proven first by [9] and later by many more including [11], with stronger approaches incorporating degenerate ground states. If a degenerate ground state is present there is no unique mapping from n_0 to Ψ_0. However, a more general functional for F can be constructed to acquire all degenerate ground states [12]. In addition the inequality (2.10) also provides a simple implementation of a variational principle in order to solve for n.

2.1.3   Kohn-Sham equations

The Kohn-Sham (KS) equations are the second central pillar of DFT. They provide an elegant way of determining the charge density of a given system without

---

[2] In fact the equality between the true ground state wavefunction Ψ_0 and the wavefunction Ψ_0^{min} that constructs $F[n_0]$ in the minimisation needs to be shown[10]

any initial approximations[3]. The key argument to deriving the KS equations is that the true electron density can also be constructed using a wavefunction made up from hypothetical non-interacting single electron wavefunctions, called Kohn-Sham orbitals (KSO), from here on denoted as ϕ. Even with non-interacting single particle wave functions, the antisymmetry of the total wavefunction is still required and is thus usually constructed using a slater determinant of the KSOs. The error stemming from the interaction of the electrons from the electron-electron Coulomb potential can be accounted for by an adjusted potential [13]. Therefore the charge density can be equivalently described by either the many-body wavefunction or the sum of the KSO's:

$$
\begin{aligned}
n(\mathbf{r}) &= N \int \Psi^*(\mathbf{r}\mathbf{r}_2...\mathbf{r}_N)\Psi(\mathbf{r}\mathbf{r}_2...\mathbf{r}_N)\,d\mathbf{r}_2...d\mathbf{r}_N \\
&= \sum_i f_i \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}),
\end{aligned}
\tag{2.12}
$$

where at the same time here, a finite temperature T has been introduced with the occupation probability f_i defined though the Fermi-Dirac function

$$
f_i = 2f_F(\epsilon_i) = \frac{1}{1 + \exp(\frac{\epsilon_i - \epsilon_F}{k_B T})}.
\tag{2.13}
$$

The Fermi energy ϵ_F is defined as to ensure the correct number of electrons: $N = \sum_i 2f_F(\epsilon_i)$.

The first step to obtaining a useable method for the KSO's is to use the variational principle with the functional derivative from equation (2.10). This results in the ground state condition taking the form

$$
\frac{\delta E[n]}{\delta n} = \mu = \frac{\delta F[n]}{\delta n} + \frac{\delta V_{ext}}{\delta n} = \frac{\delta F[n]}{\delta n} + V_{ext},
\tag{2.14}
$$

where μ is a Lagrange multiplier and has the physical interpretation of the increase of energy when an additional electron is to be included into the system (chemical potential). Functional derivatives with respect to a function f are here denoted by $\delta/\delta f$.

Notably the external potential V_{ext} in equation (2.14) remains valid even with the introduction of a density consisting of non-interacting particles. Therefore only

---

[3] Approximations are still being made, they are just swept into a specific term, the exchange correlation functional

the functional $F[n]$ defined in equation 2.7 needs to be reexamined:

$$F[n] = T[n] + V_{ee}[n], \tag{2.15}$$

where both functionals (kinetic energy and electron interaction) can be restructured to account for non-interacting particles. This has been done before in other methods such as the Thomas Fermi model [14] for Jellium conditions and Local Density Approximations. However, the key point is that here the difference with respect to the true functional is repackaged as the exchange-correlation functional and recognised as a change to the potential, creating an effective potential. This idea results in the formulation of the Hohenberg-Kohn-Sham Functional

$$F^{HKS}[n] = T_s[n] + E_H[n] + E_{xc}[n]. \tag{2.16}$$

The above functional incorporates two altered functionals [8], the first being the classical Coulomb potential for the electron-interaction $E_H[n]$ which takes the form

$$E_H[n] = \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{2.17}$$

And the second is the kinetic energy functional $T_s[n]$ for the non-interacting KS states. In order to get the real electron density, the difference to the complete kinetic and electron repulsion energy must be included and is incorporated into the exchange correlation energy functional E_{xc}.

At this point the total energy of the system can be represented through the rewritten functional and becomes

$$\frac{\delta E^{HKS}}{\delta n} = \frac{\delta T_s}{\delta n} + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + V_{ext}(\mathbf{r}) = \mu, \tag{2.18}$$

where the exchange-correlation potential is defined through the functional derivative

$$V_{xc} = \frac{\delta E_{xc}}{\delta n}. \tag{2.19}$$

In equation (2.18) all contributions except the single particle kinetic energy are local operators (diagonal in a real space representation) and therefore represent single particle potentials and can be combined into an effective potential:

$$V_{eff}(\mathbf{r}) = V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + V_{ext}(\mathbf{r}) \tag{2.20}$$

Using this effective potential the KS equations can be constructed, which is essentially a Schrödinger equation for each hypothetical non-interacting electron wavefunction ψ_λ called Kohn Sham Orbitals:

$$\left(-\frac{1}{2}\nabla^2 + V_{eff}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}).\tag{2.21}$$

The non-interacting part of the total kinetic energy can be constructed then as

$$T_s = \sum_i f_i\langle i|\hat{T}|i\rangle = \sum_i f_i \int \phi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\phi_i(\mathbf{r})d\mathbf{r},\tag{2.22}$$

with f_i representing the above defined Fermi factor stating occupancy.
With these definitions the functional is then:

$$E^{HKS}[n] = T_s[n] + E_H[n] + E_{xc}[n] + E_{ext}[n] + E_{zz},\tag{2.23}$$

where E_{zz} is a constant representing the energy due to the nuclear coulomb repulsion. The only unknown at this point is the exchange correlation functional E_{xc}, for which approximate forms can be found.

2.1.4   Exchange correlation functionals

The KS equations are the Hamiltonian for a set of non-interacting single particle states in an additional potential that compensates for the neglected many-body effects in the full Schrödinger equation. This difference between the non-interacting single particle equations and the full Schrödinger equation is known as the exchange-correlation. The exchange-correlation is usually included approximately as a functional of the total charge density in DFT calculations.

Local density approximation

The earliest and simplest approximation of the exchange correlation functional is the local density approximation (LDA) which treats the inhomogeneous electron case as locally uniform:

$$E_{xc}[n] = \int n(\mathbf{r})\varepsilon_{xc}(n(\mathbf{r}))d\mathbf{r}\tag{2.24}$$

or in terms of the potential

$$V_{xc}(\mathbf{r}) = \varepsilon_{xc}(n(\mathbf{r})) + \left[n \frac{d\varepsilon_{xc}(n)}{dn} \right]_{n=n(\mathbf{r})}. \tag{2.25}$$

Several functional forms of ϵ_{xc} have been developed, a common solution can be found in [15]. Although the accuracy of the LDA is surprisingly good, a major source of error stems from the self-interaction problem. Within the LDA every electron (incorrectly) interacts with itself as well as all others.

An improvement to this problem can be achieved by a slight generalisation of the LDA, the local spin density approximation. It treats the two spin states separately and is commonly referred to as 'LDA'.

$$E_{xc}^{LSD} = \int n(\mathbf{r})\varepsilon_{xc}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}))d\mathbf{r} \tag{2.26}$$

The functional form of ε_{xc} that is being used in this thesis is from [16], which splits it into exchange and correlation energies for various situations. The exchange energy follows a dependency of $n^{4/3}$ and correlation $n^{5/3}$ with separate densities for up and down spins.

General gradient approximation

In order to further refine the approximation for the exchange correlation functional the general gradient approximation includes information of the electron density gradient, a general form of which is

$$E_{xc}^{GGA}[n] = \int f(n(\mathbf{r}), \nabla n(\mathbf{r}))d\mathbf{r}. \tag{2.27}$$

several implementations of this idea can be found in the literature. However, the General Gradient Approximation is not used within this work, here LDA is applied exclusively.

2.1.5   Periodic boundary conditions

While the above DFT method is capable of calculating approximate charge densities for any atoms and in principle any number of atoms, further considerations apply for periodic structures such as crystals. These are defined through a unit cell containing a set of atoms and symmetry operations such as rotations and

translations. Hence it is sufficient to only solve for the charge density of the unit cell with the knowledge of the crystal symmetry in order to study the whole crystal. This can be done utilising Bloch's theorem [8], which is presented in this subsection.

A general description for the lattice repeat can be written as a linear combination of translations along lattice vectors to an identical position to any other unit cell:

$$\mathbf{L} = n_1\mathbf{a_1} + n_2\mathbf{a_2} + n_3\mathbf{a_3} \tag{2.28}$$

Where $n_i \in \mathbb{N}$ and $\mathbf{a}_i$ are the lattice vectors defining the symmetry of the crystal. With this in mind the reciprocal lattice can be defined as

$$\mathbf{b}_i = 2\pi\epsilon_{ijk}\frac{\mathbf{a}_j \times \mathbf{a}_k}{\mathbf{a}_i \cdot (\mathbf{a}_j \times \mathbf{a}_k)}. \tag{2.29}$$

The reciprocal lattice vectors b_i are essentially a Fourier transform of the lattice vectors and similarly define a unit cell in Fourier space or reciprocal space, called the Brillouin Zone (BZ). Specifics on the BZ for the materials under study in this thesis can be found in section 3.2.

Bloch's theorem states that any single particle wavefunction in a periodic potential $V(\mathbf{r} + \mathbf{L}) = V(\mathbf{r})$ can be represented by the product of a plane wave-like phase factor with the wave vector $\mathbf{k}$ and a periodic function $u_{i\mathbf{k}}(\mathbf{r} + \mathbf{L}) = u_{i\mathbf{k}}(\mathbf{r})$:

$$\psi_{i\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{i\mathbf{k}}(\mathbf{r}_r). \tag{2.30}$$

Which results in only a phase shift of the wavefunction under translations by $\mathbf{L}$:

$$\psi_{i\mathbf{k}}(\mathbf{r} + \mathbf{L}) = e^{i\mathbf{k}\cdot\mathbf{L}}\psi_{i\mathbf{k}}(\mathbf{r}). \tag{2.31}$$

This also means that the charge density is not affected and displays the same periodicity as the lattice ($\rho(\mathbf{r}+\mathbf{L}) = \rho(\mathbf{r})$). The construction of a charge density in a system with multiple non-interacting wavefunctions is done by integrating over the first BZ and summation over all occupied particle wavefunctions. However each k-point wavefunction needs to be determined separately and therefore it is common that the BZ is only sampled across the irreducible BZ, which is the minimal section of the first BZ according to the lattice symmetry operations such as rotations and inversion. A weighing factor $\omega_{\mathbf{k}}$ is included to account for this

reduction to the irreducible BZ.

$$\rho(\mathbf{r}) = \sum_i \sum_{\mathbf{k} \in \mathrm{BZ}} \omega_{\mathbf{k}} |\psi_{i\mathbf{k}}(r)|^2. \tag{2.32}$$

Various sampling schemes can be adopted, however the BZ sampling in this work is performed exclusively using the scheme proposed by Monkhorst and Pack [17]. It follows the construction of a regular grid of $(I \times J \times K)$ points.

$$\mathbf{k}_{ijk} = u_i \mathbf{b}_1 + u_j \mathbf{b}_2 + u_k \mathbf{b}_3 \tag{2.33}$$

with $u_i = (2i - I - 1)/2I$, $u_j = (2i - J - 1)/2J$ and $u_i = (2i - K - 1)/2K$.

As the charge density obeys the same periodicity as the external potential, Bloch's theorem can also be used for the effective potential in the KS equations:

$$\left(-\frac{1}{2}\nabla^2 + V_{eff}^k(\mathbf{r}) \right) \phi_{i\mathbf{k}}(\mathbf{r}) = \epsilon_{i\mathbf{k}} \phi_{i\mathbf{k}}(\mathbf{r}). \tag{2.34}$$

Therefore each KSO is also a function of k and the charge density is now similarly defined through the summation

$$\rho(\mathbf{r}) = \sum_{\mathbf{k} \in \mathrm{BZ}} \omega_{\mathbf{k}} \sum_i^{N_{\mathbf{k}}} f_{i\mathbf{k}} |\phi_{i\mathbf{k}}(r)|^2. \tag{2.35}$$

The density of the sampling scheme is now dependent on the size of the BZ and the k dependence of the KSO's. However, the size of the BZ is inversely proportional to the lattice constant, resulting in the possibility of lower sampling point densities for larger real-space cells. The actual quality of the sampling for a particular system in principle needs to be determined for each property separately.

2.1.6   Pseudopotentials

The use of pseudopotentials is a very important step in the ab-initio modelling of large scale systems. The central idea is that only valence electrons should be included in the charge density, as core electrons can be considered to be chemically inactive. The absence of core electrons is then compensated by a modified nuclear potential. This idea ties in nicely with DFT, as the KS equations already

rely on a modified effective potential.

The pseudopotential approach parameterises the external potential for each chemical species. This parameterisation is commonly done by utilising all electron ab-initio methods to ensure consistency, with several methods proposed for constructing these pseudopotentials and their parameterisation. However, high accuracy and transferability is obtained by constructing pseudopotentials with norm-preserving states. A popular example for such a norm-preserving pseudopotential is that of Bachelet, Hamann and Schlüter [18].

Norm-conservation is defined through a cut-off radius, beyond which the pseudo valence electron wavefunction agrees with the respective wavefunction obtained by an all-electron calculation. In addition the integrated charge must be identical for the two wavefunctions within the cutoff distance [14]. The norm conservation thus results in the equality of the total integrated charge for each wavefunction with that for the all electron wavefunction. The functional shape of the resulting pseudo-wavefunction is also chosen to be smooth at the cutoff boundary including it's derivatives.

The construction of such a norm preserving potential is done by performing an all electron calculation on a single (radially symmetric) atom and calculating the KSO's. The effective all-electron potential is

$$v_{eff}(r) = -\frac{Z}{r} + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}[n(\mathbf{r})], \qquad (2.36)$$

where r is the distance from the nucleus. The effective pseudo potential is chosen to be non singular at $(r = 0)$; A function $f(x)$ is introduced that is equal 1 at $x = 0$ and tends to zero as $x > 1$:

$$v_{eff,l}^{ps}(r) = v_{eff}(r) \left(1 - f\left(\frac{r}{R_{cl}}\right) \right) + c_l f\left(\frac{r}{R_{cl}}\right). \qquad (2.37)$$

the parameter R_{cl} is the cutoff radius, which has to be chosen carefully to ensure accuracy, while c_l is adjusted to match the lowest energy eigenvalue for each angular momentum l with the all electron case. Note that this dependence on the angular momentum makes the potential non-local. The ground state pseudo wavefunction of this initial modified potential must match its all electron equivalent (outside of R_{cl}). The pseudopotential capable of generating this normalised wavefunction can be found by inverting the Schrödinger equation at the energies of the matched eigenvalues. The resulting potential contains then a modified

electron-ion potential as well as part of the exchange correlation and Hartree potential and can usually be separated into a non-local and local part.

The particular implementation used within this work have been developed by Hartwigsen, Goedecker and Hutter [19] and includes relativistic separable, dual-space Gaussian pseudopotentials as well as non-linear core corrections (NLCC). NLCC are particularly important for Group-III elements (such as Ga and In) as here the semi-core d states overlap more strongly with the valence p states. Therefore the exchange-correlation potential is non linear and

$$v_{xc}(n_c, n_v) = v_{xc}(n_c) + v_{xc}(n_p) \tag{2.38}$$

can no longer be assumed. See Louie et al. [20] for a discussion of a method to produce NLCCs.

2.1.7   AIMPRO code implementations

All DFT calculations in this thesis have been performed using the AIMPRO code, which has been developed at the University of Newcastle and is in use at several groups in Europe [21]. It uses Gaussian basis functions to expand the KS states, enabling efficient simulations. Many other DFT codes instead use a simpler plane wave expansion, which allows for easy control of the accuracy of the solution, however the expansion of KSOs as a sum of Gaussian functions requires fewer terms and therefore allow for a much smaller set of functions to solve the KS equations[22], enabling the convergence to be achieved more rapidly. In addition AIMPRO is capable of including pseudopotentials and allows a choice of either LDA [16] or GGA [23] functionals for the estimation of the exchange correlation energy. The results presented here have been obtained solely within LDA. These features enable the AIMPRO code to perform very efficient DFT-calculations for structures containing up to hundreds of atoms on a desktop machine. However, recent advances have resulted in the development and implementation of the basis filtration principle [24] within the AIMPRO code. Usual local basis codes using direct diagonalisation to solve the eigenvalue problem require $O(N^3)$ steps, where N is the number of basis functions. The filtration algorithm effectively reduces this by dynamically building a smaller contracted basis, consisting of 4 basis functions per atom. For an original basis of 16 functions per atoms this results in a theoretical speed gain of $\sim (\frac{16}{4})^3 = 64$ [24], while retaining the

accuracy of the original basis.

In order to solve the KS equations, the AIMPRO code applies a direct diagonal-isation method. Specifically, it expands each KSO into basis functions[4] φ_i with coefficients $c_i^{\lambda s}$. This takes the usual form (s denoting the spin)

$$\phi_{\lambda s} = \sum_{i=1}^{N} c_i^{\lambda s} \varphi_i(\mathbf{r}). \tag{2.39}$$

The quality of the above expansion is only limited by the suitability of the basis functions and their quantity N. The normalisation of a KSO is then expressed as

$$\int |\phi_{\lambda s}(\mathbf{r})|^2 \, d\mathbf{r} = \sum_{ij} c_i^{\lambda s} c_j^{\lambda s} S_{ij} = 1, \tag{2.40}$$

where S_{ij} denotes the overlap matrix, defined through:

$$S_{ij} = \int \varphi_i(\mathbf{r}) \varphi_j(\mathbf{r}) d\mathbf{r}. \tag{2.41}$$

The spin dependent charge density in this notation is then represented as

$$n_s(\mathbf{r}) = \sum_{ij} b_{ij}^s \varphi_i(\mathbf{r}) \varphi_j(\mathbf{r}), \qquad b_{ij}^s = \sum_{\lambda} f_{\lambda s} c_i^{\lambda s} c_j^{\lambda s} \tag{2.42}$$

With b_{ij}^s being the spin dependent charge density matrix. The self consistent minimisation of the KS equations can now be done with respect to the quantities $c_i^{\lambda s}$

$$\frac{\delta}{\delta c_i^{\lambda s}} \left[E[b_{ij}^s] - \sum_{\lambda',s'} \epsilon_{\lambda' s'} \left(\sum_{ij} S_{ij} c_j^{\lambda' s'} c_j^{\lambda' s'} - 1 \right) \right] = 0. \tag{2.43}$$

The differentiation can be done formally using

$$\frac{\delta E}{\delta c_i^{\lambda s}} = \sum_{jk} \frac{\delta E}{\delta b_{jk}^s} \frac{\delta b_{jk}^s}{\delta c_i^{\lambda s}} = \sum_j H_{ij}^s c_j^{\lambda s} \tag{2.44}$$

---

[4] The basis functions are typically Gaussian within AIMPRO and will therefore be treated in this section as real valued for simplicity as opposed to the usual complex approach. Important to note is that for periodic Bloch functions the complete complex treatment is required.

where the Hamiltonian matrix $\hat{\mathbf{H}}$ is defined through

$$H_{ij}^s = \frac{\delta E}{\delta b_{ij}^s}. \tag{2.45}$$

Therefore a matrix equation has been constructed

$$\sum_j H_{ij}^s c_j^{\lambda s} = \epsilon_{\lambda s} \sum_j S_{ij} c_j^{\lambda s} \tag{2.46}$$

which has the simple form of a generalised eigen-problem

$$(\hat{\mathbf{H}} - \hat{\mathbf{S}}\boldsymbol{\epsilon})\mathbf{c} = 0 \tag{2.47}$$

and can be solved through common linear algebra methods [25]. The self consistency cycle itself consists of

1. building the Hamiltonian $\hat{\mathbf{H}}$ in a chosen basis,

2. calculating the overlap matrix and formulating the eigenvalue problem in equation (2.47),

3. performing the diagonalisation and updating the density matrix $\hat{\mathbf{b}}$ as well as constructing the charge density with equation (2.42).

Having found a new density matrix at the end of a cycle, step (1) can be repeated starting a new cycle until the difference in the solution (commonly with respect to the total energy) is sufficiently small.

Choice of basis set

In order to solve the eigenvalue problem posed by the KS equations it is necessary to use a suitable basis set spanning that function space. A very common and orthogonal choice are plane waves, resulting in the expansion becoming:

$$\phi_{\lambda s}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\lambda s}(\mathbf{G}) \exp(i\mathbf{G}\mathbf{r}), \tag{2.48}$$

where $\mathbf{G}$ is a reciprocal space vector and the summation is done up to a cutoff value. The advantage being that the quality of the basis set is easily adjustable by increasing the cutoff energy and a straightforward incorporation of periodic

boundary conditions.

The basis set used in the AIMPRO code however are Cartesian Gaussian in nature, with the form:

$$\varphi_i(\mathbf{r}) = (x - R_{ix})^{n_1}(y - R_{iy})^{n_2}(z - R_{iz})^{n_3} \exp\left(-\alpha_i(\mathbf{r} - \mathbf{R}_i)^2\right). \qquad (2.49)$$

The basis functions here are real space representations and are centred around the atom positions $\mathbf{R}_i$, where the α_i are atomic type dependent coefficients that need to be optimised for each species and bonding type. The coefficients n_1, n_2, n_3 are integers and vary the nature of the basis function between s-, p-, d-, ... type.[5] Linear combinations of these Gaussians can be chosen to transform like spherical harmonics under rotations [22]. The Gaussian basis set requires only a few functions per atom, is real valued and has easily determined overlap matrices, for example two s-type orbitals:

$$S_{ij}^{00} = \left(\frac{\pi}{\alpha_i + \alpha_j}\right)^{3/2} \exp\left(-\frac{\alpha_i \alpha_j}{\alpha_i + \alpha_j}(\mathbf{R}_i + \mathbf{R}_j)^2\right) \qquad (2.50)$$

and similarly for higher momenta. Therefore utilising these orbitals requires little computational time and memory. In addition a Gaussian basis set is more suitable for pseudopotential calculations as core states in an all-electron calculation deviate too strongly from a Gaussian form requiring many functions to expand them.

However, contrary to the KSO orbital construction, the charge density is represented as a plane wave expansion within AIMPRO and thus is dependent on a cutoff frequency.

Filtration principle

Basis set filtration is a method to drastically reduce computational time for each self consistency step in a DFT calculation. It does this by redefining the eigenvalue problem in a subspace consisting of fewer states, and more importantly a reduction in basis functions, by projecting only a subset of local eigenfunctions that contribute to each chosen diagonalisation. This requires an original basis set of localised orbitals such as Gaussian orbitals and a restructuring of the eigen-

---

[5] A s-nature results when all n coefficients are zero, if only one coefficient deviates to be equal 1 the basis function acquires a p-type and the six choices of $\sum_i n_i = 2$ incorporate a combination of 5 d-states and 1s state.

value problem [24].

The filtration of the basis set is performed by defining a function of the Hamilton operator $\mathbf{F}(\mathbf{H})$, chosen to suppress eigenvalue-eigenfunction pairs according to a Fermi-Dirac function $f(x)$ operating at high temperature (this is separate from the electron temperature introduced earlier in the charge density construction). In a non-orthogonal basis this becomes

$$\mathbf{c}f(\boldsymbol{\epsilon})\mathbf{c}^T\mathbf{S}|t\rangle = \mathbf{F}\mathbf{S}|t'\rangle, \tag{2.51}$$

where F acts on a test function, filtering out components so that the remaining ones lie predominantly in the desired subspace. This filtration is repeated until the space is sufficiently spanned, creating a smaller diagonalisation problem.

Until this point filtration has not introduced any approximations. However there have also been no computational savings, as the construction of the filtration is as costly as the diagonalisation of the original problem. Therefore the filtration must be constructed locally for each atom position ($\mathbf{F_I}$, with $I = 1, .., M$) and only approximately includes all basis functions by restricting its support to surrounding atoms within a cutoff distance, the filtration radius. From here subspace Hamiltonians and density matrices can be constructed and diagonalised separately. Subsequently the full charge density can be reconstructed. This filtration radius is therefore a quality parameter and by increasing it to infinity the original solution is achieved.

2.2   Interatomic potentials

Interatomic potentials utilise the fact that, in principle a system's total energy is a function of atomic types and positions. These potentials can also be used for Molecular Dynamics calculations under the Born-Oppenheimer approximation. Under this constraint, forces can be calculated from the gradient of the determined potential.

While ab-initio methods such as DFT can provide such a potential, interatomic potential methods use a purely empirical approach to the construction of the energy of a given system. Clearly the most prominent advantage to such a formulation its great speed. However electronic interactions and charge transfer are commonly not included at all and can therefore not be studied, only structural information and total energy is to be gained. Further disadvantages are the low transferability as these potentials can only give accurate results in situations similar to those where the potential parameters were fitted for. This quality is also strongly dependent on the functional complexity of the potentials themselves.

Several interatomic potential forms have been developed, the earliest and simplest being pair potentials such as hard spheres, Lennard-Jones or Morse potentials. Pair potentials calculate the total energy with

$$E(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = \sum_I \sum_{J>I} U_2(r_{IJ}), \tag{2.52}$$

where $r_{IJ} = |\mathbf{r}_I - \mathbf{r}_J|$ is the distance between two atoms. A simple example of which is the Lennard-Jones potential:

$$U_2(r_{IJ}) = 4A \left(\left(\frac{B}{r_{IJ}} \right)^{12} - \left(\frac{B}{r_{IJ}} \right)^6 \right), \tag{2.53}$$

with A and B being parameters to be fitted for different atom types. This pair potential approximates the short range 'hard sphere' repulsion due to the Pauli principle with the polynomial r^{-12} and an attractive part due to van der Waals attraction with $-r^{-6}$ and is as such very suitable functional description of gases and few solids [26]. However in a crystal where directional bonds play a crucial role it is necessary to add an angular dependent term to the energy potential. Three body potentials include this information about the configuration of atoms and are hence applicable to covalent bonded materials [26]. A popular example

is the Tersoff potential that builds on a Morse Pair potential (an exponential description of attraction and repulsion) and adds a three body term on the attractive part that includes the bond angle between three atoms. The potential of choice in this thesis however is one developed by Stillinger and Weber[27] and has been used here exclusively for structural optimisations on GaN systems.

2.2.1 Stillinger Weber potentials

The Stillinger Weber (SW) Type interatomic potential is a three body potential that builds on a Lennard-Jones type two body pair potential. It is capable of describing directional bonds and has been originally parameterised for group 4 elements such as silicon and carbon[27]. The potential itself is a sum of two and three body parts:

$$E(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = \sum_{IJ} U_2(r_{IJ}) + \sum_{IJK} U_3(r_{IJ}, r_{IJ}, \theta_{IJK}). \tag{2.54}$$

The two body potential incorporates positive repulsive, negative attractive and exponential terms. The last serves as a decay over the cutoff distance a_{IJ},

$$U_2(r_{IJ}) = A_{IJ}\epsilon_{IJ} \left[B_{IJ} \left(\frac{\sigma_{IJ}}{r_{IJ}} \right)^{q_{IJ}} - \left(\frac{\sigma_{IJ}}{r_{IJ}} \right)^{p_{IJ}} \right] g(r_{IJ}, \sigma_{IJ}, a_{IJ}), \tag{2.55}$$

where g is the cutoff function defined through

$$g(r, \sigma, a) = \begin{cases} \exp\left(\frac{\sigma}{r - a\sigma}\right) & \text{if } r < a \\ 0 & \text{if } r \geq a. \end{cases} \tag{2.56}$$

The three body term U_3 is similarly defined, in that it utilises the same cutoff function, however scaled additionally with a parameter γ. It ranges from zero to positive numbers depending on the angle θ_{IJK} between the three atoms, where this angle is measured as the smallest angle between r_J and r_K with r_I as the vertex.

$$\begin{aligned} U_3(r_{IJ}, r_{IK}, \theta_{IJK}) &= \lambda_{IJK}\epsilon_{IJK} \left(\cos\theta_{IJK} - \cos\theta_{IJK}^0 \right)^2 \\ & \qquad g(r_{IJ}/\gamma_{IJ}, \sigma_{IJ}, a_{IJ}/\gamma_{IJ}) \qquad (2.57) \\ & \qquad g(r_{IK}/\gamma_{IK}, \sigma_{IK}, a_{IK}/\gamma_{IK}) \qquad (2.58) \end{aligned}$$
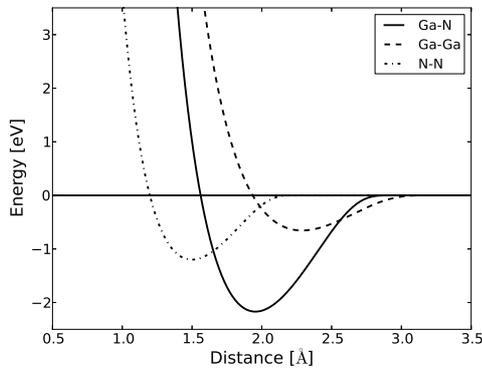
(a) SW Pair potential part (U_2)
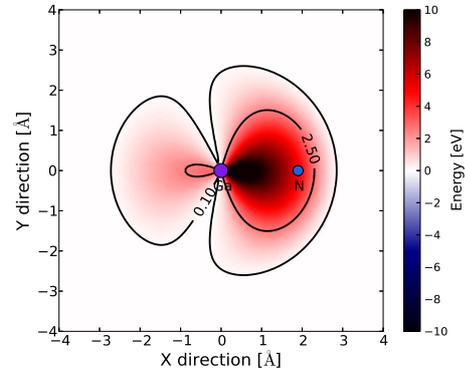
(b) Single SW three body potential (U_3) for Ga-N-N
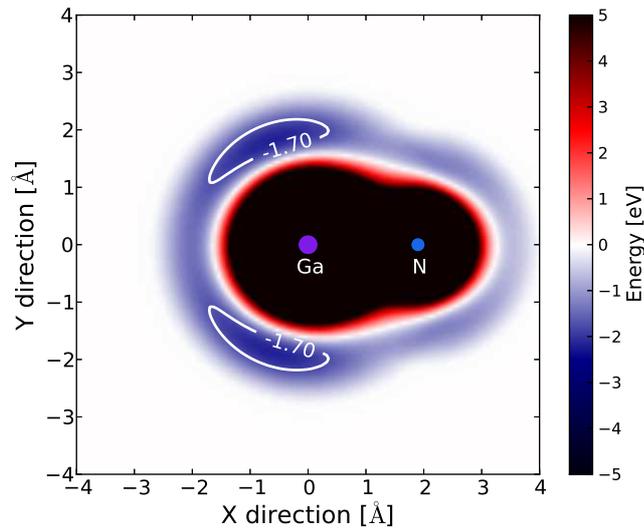
(c) Complete SW potential for Ga-N-N

Figure 2.1: Figures depict the individual components and the complete SW type interatomic potential. The radial part (b) and complete potential (c) are those acting on a nitrogen atom due to one gallium atom at $(0,0)$ and one nitrogen atom at $(1.9, 0)$. As there are only two atoms drawn in figures (b) and (c), a rotational symmetry develops with rotational axis along the Ga-N direction (x-direction) and as a result the complete potential can be viewed in a 2D plot. Parameters for the potential were taken from [2].

The parameters γ and λ are included to scale the three body to two body potential and hence regulate the angular potential strength and range, while θ_{0IJK} determines the optimum bonding angle. At an angle of $109.47°$ ($\cos(\theta^0_{IJK}) = -1/3$) the potential favours tetrahedral bonds. A depiction of the two and three body potential as well as the full potential can be found in figure (2.1).

The SW potential with various parameterisations and slight modifications has been used before in extended defect studies in group-III Nitrides. Examples include dislocation core configurations from [28], [29] and [30] or stacking faults and partial dislocations from [31]. In addition ab-initio or semi-empirical methods (such as DFTB) have been used to compare predictions and confirm the applicability of the SW-potential.

Various SW-parameters for the group-III nitrides can be found in the literature, which are usually fitted to lattice parameters and elastic constants. The parameters used in this work were developed by Béré et al [2] for GaN and were an optimisation of the parameters from Aïchoune et al. [32]. Aïchoune et al. originally used a modified version of the SW-potential and optimised the parameters to reproduce the experimental lattice parameters, experimental elastic constants and ab-initio results, then claiming that usual Inversion Domain Boundaries (IDB) are unfavourable in GaN. However, more recent experimental results proved the existence of these IDBs. Therefore Béré et al. included slight modifications to the parameters and was able to recover to the original form of the SW-potential by reducing interactions to zero smoothly before the second neighbour atomic distance. These parameters then differ in the prediction of the IDB stability and match experimental data. Additional comparisons have been made to ab-initio modelling with extended defects, confirming their applicability [2]. These parameters are listed in table (2.1) and have been used for all interatomic potentials calculations in this work. It is important to note that all parameters are given pairwise, i.e. do not involve a third atom. Therefore $\lambda_{IJK} = \lambda_{IK}$ and $\epsilon_{IJK} = \sqrt{\lambda_{IJ}\epsilon_{IJ}\lambda_{IK}\epsilon_{IK}}/\lambda_{IK}$ are assumed.

The LAMMPS code

All structural optimisations utilising the SW-potential were performed using the LAMMPS code [37]. LAMMPS is a classical molecular dynamics code, and an acronym for Large-scale Atomic/Molecular Massively Parallel Simulator. It in-

(a) Stillinger Weber potential parameters from [2]

	ϵ	σ	λ	a	A	B
Ga-N	2.17	1.695	32.5	1.8	7.917	0.72
Ga-Ga	1.2	2.1	32.5	1.6	7.917	0.72
N-N	1.2	1.3	32.5	1.8	7.917	0.72

(b) Wurtzite lattice parameters and elastic constants

	a	c	c_{11}	c_{12}	c_{13}	c_{33}	c_{44}	c_{66}	B
SW	3.19	5.21	354	140	124	370	97	107	206
exp.[1]			391	144	108	399	103	124	210
exp.[2]			359	129	92	389	98	115	194
exp.	3.19	5.18							

(c) Zinc-blende lattice parameters and elastic constants

	a	c_{11}	c_{12}	c_{44}	B
SW	4.52	318	149	123	205
calc.[3]		296	154	206	201
exp.[4]	4.49				

Table 2.1: Parameters and comparison for the Stillinger Weber potential for Ga, N and GaN taken from [2]. The units for ϵ are given in eV while σ is in Å, all other quantities in table (a) are dimensionless. Elastic constants in table (b) and (c) are in GPa, lattice constants in Å.

[1] experimental data from [33]
[2] experimental data from [34]
[3] calculated from wurtzite by [35]
[4] experimental data from [36]

cludes a formulation of the SW-potential and accepts SW-parameters for any
interaction. The actual optimisations were performed with periodic boundary
conditions using large supercells and a conjugate gradient optimiser.

2.3   Optimisation algorithms

The general definition of an optimisation is the selection of a best element from a group with regard to pre-defined criteria. In many cases including this one the pre-defined criteria is a real valued cost function or energy function. In addition the group of elements in which the solution is to be found is often a subspace of $\mathbb{R}^n$ and specified by a set of constraints. Elements of this group are called feasible solutions.

$$
\begin{aligned}
f, g_i \quad &: \quad \mathbb{R}^n \to \mathbb{R} \\
\text{minimize} \quad &f(x), \qquad \quad x \in \mathbb{R}^n \\
\text{subject to} \quad &g_i(x) \ \leq \ 0, \quad i = 1, \ldots, m
\end{aligned}
\tag{2.59}
$$

where f is the objective cost function and g_i are constraint functions that can be equalities. With this definition it is possible to classify optimisation problems into linear, non-linear a subset of non-linear, convex problems; if the following is true for both constraint and objective function:

$$
\begin{aligned}
f(\alpha x + \beta y) \ &= \ \alpha f(x) + \beta f(y) \qquad \text{linear} \\
f(\alpha x + \beta y) \ &\neq \ \alpha f(x) + \beta f(y) \qquad \text{non-linear} \\
f(\alpha x + \beta y) \ &\leq \ \alpha f(x) + \beta f(y) \qquad \text{convex}
\end{aligned}
\tag{2.60}
$$

where $\alpha, \beta \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$. A convex minimisation problem has only a single minimum that can be readily found with a variety of optimisation algorithms, however all other problems may have several local minima and a global optimisation technique needs to be applied to find the optimal solution[38].
The following sections will introduce only a few select optimisation methods and their application to find local or global minima.

2.3.1   Local optimisations

Many optimisation algorithms exist that are well defined global optimisations for convex problems such as the simplex and conjugate gradient algorithm. The simplex algorithm is defined for linear cost functions and works by systematically (iteratively) reducing the feasible solution set by changing the constraint conditions until an optimal solution is found. The efficiency of the simplex algo-

rithm is usually of polynomial order, however for all its implementations, worst case scenarios have been shown to exist where the efficiency drops to exponential order. On the other hand the conjugate gradient algorithm very efficiently solves for quadratic cost functions by using local gradient information in addition. It is usually much faster than the more simple gradient descent method, where the main difference lies in the assumption of a purely quadratic function in a standardized form. However for general non-linear problems, where many local extrema can exist, neither of these methods can guarantee to produce an optimal solution and are hence treated here as local optimisations.

In this work the prominent optimisation problem consists of a potential energy function to be optimized that uses atomic positions as arguments. This function can be computed with either DFT or SW potentials. Some of the possible local optimisations for the atomic positions are a downhill simplex (Nelder-Mead), simple gradient or a nonlinear conjugate gradient method[25]. Where in this work primarily the conjugate gradient option has been chosen for its superior efficiency.

2.3.2   Global optimisations

In order to find the true global extremum for a general non-linear optimisation problem, global optimisation techniques need to be applied, as there are commonly a large quantity of local extrema (potentially of exponential order). Optimisations for non-linear functions is a field of active research that has already produced a wide variety of methods where generally a heuristic or stochasitc approach can be taken. Stochastic approaches include Monte Carlo methods and simulated annealing, while heuristic approaches include evolutionary and swarm-based algorithms. Only a very select few methods will be briefly introduced here and compared to the approach taken within this work. The random search algorithm used here, however could be classed as a stochastic method as it randomly searches through the feasible set of solutions to find initial tries.

Simulated annealing

Simulated annealing is a stochastic metaheuristic global optimisation technique, it therefore produces a probable global solution to a given system subject to a quality parameter, the annealing schedule [39]. The method iteratively searches

for the global minimum of a given system by looking at the neighbouring states of an arbitrary starting state, computing the probability of moving to these states and deciding on the optimal choice. This is done by using a probablity function $P(E, E', T)$ for moving from a current configuration with energy E to a new neighbouring configuration with energy E' at temperature T, which is usually a Boltzmann factor $exp(-(E' - E)/k_B T)$. The important feature is that the probability to change into another state is always positive and goes to 1 for any choice with higher temperature. The algorithm then works by assuming a random starting point at a sufficiently high temperature that is lowered with every iteration according to the annealing schedule until a the configuration is frozen as all probabilities for movement tend to 0. The annealing schedule is therefore the manner of lowering the temperature, where the probability of finding a global optimal solution approaches 1 as the annealing schedule is extended. However, the time required to reach a high probability of finding the globally optimal solution can be exceedingly long and confirmation of actually finding the optimal solution cannot be given.

Genetic algorithm

A genetic algorithm is a type of evolutionary algorithm where the evolutionary process in biology is taken as an example for a metaheuristic optimisation, meaning that an approximate global solution can be found. This has been perfomed on structural optimisation before for example by Hartke et al. [40]. In this method the geometry of a particular solution is encoded into a bit string, which is the equivalent of a genome of an individual. In order find the global optimimum a random starting population is created and according to an energy "fitness" function parents are chosen from which children are constructed by combining the genomes into that of an entire new generation. This is repeated until a best solution has been found, subject to the convergence of the fitness criteria of the energy function. A guarantee of actually finding an optimal solution however cannot be given.

Particle swarm optimisation

The particle swarm optimisation technique is a more recent development and has been applied to structural optimisations before [41] as well as a very wide area of other topics [42]. It is an iterative method and works by introducing

an initial population of particles at random positions in the feasible solution space, where each particle is assigned a velocity that determines its change in position per iteration. The method is inspired by flocks of birds and hence uses two separate influences on the velocity of each particle, those are the individuals particle's social neighbourhood and its own history. The selection of which parts of neighbourhood and history are used is determined by the cost function. A common description of the evolution of the i$^{\text{th}}$ particles velocity $\mathbf{v}_i(t)$ at time t is:

$$\mathbf{v}_i(t+1) = \omega\mathbf{v}_i(t) + AR_1(\mathbf{x}_{ni}(t) - \mathbf{x}_i(t)) + BR_2(\mathbf{x}_{hi}(t) - \mathbf{x}_i(t)), \qquad (2.61)$$

where $\mathbf{x}_i$ is the i$^{\text{th}}$ particle's position, $\mathbf{x}_{ni}$ is the particles position with the best cost function of the i$^{\text{th}}$ particles social neighbourhood, $\mathbf{x}_{hi}$ is the position with the best cost function of the i$^{\text{th}}$ particles hisory, ω, A, B are constant parameters and R_i, R_2 are random coefficients. With each iteration all particle positions are updated and the swarm will tend towards the global minimum. Other variations of the velocity construction include a global swarm position and current global swarm best particle position. However the performance of the particular particle swarm optimisation flavour including its parameterisation depends strongly on the landscape of the problem it is applied to[43].

Random search

The random search method applied in this work is a "brute force" method, meaning that a complete sampling of the feasible solution space is performed. Specifically, this method randomly samples the solution space and performs local optimisations from each random starting point. As a result this approach features an exponential increase of computational cost with increasing dimensionality of the solution space, as the number of sampling points necessary for a constant resolution increases exponentially. Heuristic approaches such as the above particle swarm optimisation and Simulated Annealing tend to be more efficient than this complete sampling, however the clear disadvantage of each heuristic approach is that it cannot guarantee that an actual global minimum has been found.

This random search global optimisation technique for structural optimisations has been developed and applied previously by Pickard et al. [26][44], where the crystal structure of solid hydrogen was investigated initially. Within this

approach, the atomic positions inside a region of interest are randomised and subsequently relaxed using DFT in conjunction with a local conjugate gradient optimiser multiple times. In fact this randomisation/relaxation cycle needs to be repeated as many times as it takes to let the space of feasible solutions to be mapped so thoroughly that a lowest energy structure has been found with a sufficient degree of certainty. This can be achieved provided that the lowest energy structures occur disproportionately more often, since the structural relaxation algorithm will always tend towards the local minima. This behaviour becomes apparent when the resulting distribution of relaxed total energies is being investigated: It shows a normal-like distribution (or Gamma distribution) with clear spikes towards the lowest limit. In other words the mapping from the space of possible starting positions to the set of local minima of the total energy surface has more connections or larger basins towards the global minimum.

The exponential increase in computational time for this search method stems from the fact that it is based around a complete investigation of the feasible solution space through random sampling of points. However by using local optimisations at each random point the exponential cost behaviour also implies an exponential increase of local minima with increasing variables or dimensions. In the case of an energy function f with N distinguishable particles in a volume V with constant shape and $\Omega(N)$ local minima, the general assumption of exponential behaviour can be expressed as:

$$\lim_{N \to \infty} \left(\frac{\ln(\Omega(N))}{N} \right) = \alpha, \qquad \alpha > 0. \tag{2.62}$$

The exponential rate parameter α can depend on the particle density N/V and is specific to the energy/cost function, i.e. depends on the material and the approximations made. A simple argument for the validity of equation (2.62) is is that if a system of N particles is divided into M subsystems with N/M particles large enough to have independent stable solutions, then the number of local minima of the complete system is:

$$\Omega(N) = \Omega^M(N/M), \tag{2.63}$$

with the solution $\Omega(N) = \exp(\alpha N)$. However, this simple argument assumes that the function Ω is continuous and scaleable to all system sizes. A more

detailed discussion of this behaviour has been made by Stillinger [45], where the exponential increase of local minima is described with upper and lower bound scenarios.

2.4 Elasticity Theory
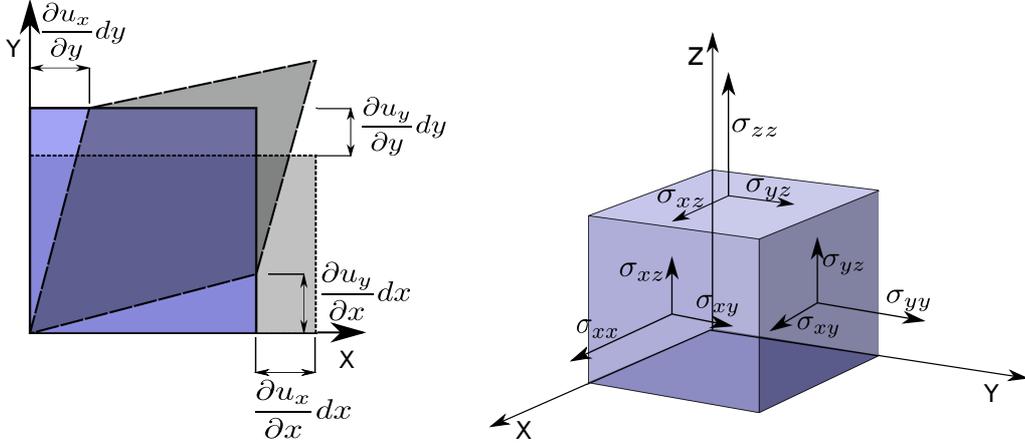
(a) Graphical representation of strain deformation of an area.

(b) Graphical representation of stress on a volume element. Stress components are symmetric and have thus been labeled.

Figure 2.2: Schematics of strain and stress

In this section a brief review of elasticity theory will be given. The focus will remain with linear isotropic elasticity theory and its application on the displacement fields of dislocations. Anisotropy will only be discussed briefly for the hexagonal wurtzite symmetry. For details on the influence of symmetries on linear elasticity theory or more precisely its influence on the shape of the elastic stiffness tensor, please refer to Appendix A.

Elasticity theory describes the relationship between stress and strain in materials that return to their initial state after deforming (i.e. elastic deformations). Stress stands here for an applied force on a surface area element and strain is the resulting deformation of the volume element. A graphical representation of each can be seen in figure 2.2. Classical linear elasticity theory is then the restriction to the regime where stresses are small enough so that the resulting strain can be expressed by a linear transformation. Materials are therefore approximated as continuous media rather than a discrete selection of atoms.

Stress has six independent components and is expressed by a symmetric matrix:

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{pmatrix} \tag{2.64}$$

As seen in figure 2.2(b) each stress component (σ_{ij}) stands for a force component applied to one of the volume's surface areas (defined through surface normal). For instance σ_{xy} is the force in y direction on the volume's $x = 0$ surface. This results in the following identity, implying no net internal torque:

$$\sigma_{ij} = \sigma_{ji}. \tag{2.65}$$

With the inclusion of an external force $\mathbf{f}$ the net force must still be zero and therefore

$$\partial_j \sigma_{ij} + f_i = 0 \tag{2.66}$$

holds true. For the above expression, and henceforth, Einstein's summation rule is being used. Similarly the strain is represented by the same matrix notation:

$$\epsilon = \begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{pmatrix} \tag{2.67}$$

As can be seen in figure 2.2(a), the components for strain can be expressed with the spatial derivative of the displacement field $\mathbf{u}(\mathbf{r})$. One possibility would be:

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \tag{2.68}$$

the above definition of the strain components is symmetric in nature and therefore results in the neglect of all antisymmetric (rotational) parts of the deformation. Rotational parts ω of a displacement u can be described by the antisymmetric part:

$$\omega_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right). \tag{2.69}$$

In addition, it is worthwhile to note for the strain tensor that all volume changing compressive/tensile elements are on the diagonal, while the volume preserving shear strains are on the off diagonal.

As mentioned above the relationship between strain and stress in *linear* elasticity theory is described by a *linear* transformation. The strain and stress components therefore have the relation:

$$\sigma_{ij} = c_{ijkl} \epsilon_{kl} \tag{2.70}$$

where **c** is called the stiffness tensor and is material dependent. The inverse of the stiffness tensor **s** is then called the elastic compliance tensor and can be used to describe the strain [6]

$$c_{ijkl}s_{klmn} = \delta_{im}\delta_{jn} \quad \Rightarrow \quad \epsilon_{ij} = s_{ijkl}\sigma_{kl}, \tag{2.71}$$

where δ is the Kronecker delta.

From the above equations and given the linear relationship between stress and strain it is possible to deduce that the stiffness tensor has the following identities:

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{klij} \tag{2.72}$$

This decreases the number of independent elements of the stiffness tensor to 21. Therefore the linear transformation above can be represented using a matrix by rewriting σ and ϵ as vectors [7]. In the most common Voigt notation this results in the following scheme:

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & . & . & . & c_{16} \\ c_{21} & c_{22} & & & & . \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & . \\ c_{61} & . & . & . & . & c_{66} \end{pmatrix} \begin{pmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{12} \end{pmatrix}, \tag{2.73}$$

where $\gamma_{ij} = 2\epsilon_{ij}$ to preserve convention (this way, matrix components do not have to be scaled when incorporating all strain and stress components).

Also, it is useful to remind oneself that the energy density can be calculated by

---

[6] It is interesting to note that with included (isotropic) thermal expansion α the relationship between stress and strain becomes

$$\sigma_{ij} = c_{ijkl}\left(\epsilon_{kl} - \alpha\Delta T\delta_{kl}\right)$$

and

$$\epsilon_{ij} = s_{ijkl}\sigma_{kl} + \alpha\Delta T\delta_{ij},$$

where ΔT is the difference in temperature and α is material dependent.

[7] Using only the independent elements of the symmetric σ and ϵ tensors to create a 6 dimensional vector $(3(3+1)/2 = 6)$. This means that **c** and **s** will be symmetric 6×6 matrices with 21 independent indices $(6(6+1)/2)$

force times distance over volume, i.e.

$$\mathcal{U} = \frac{1}{2}\sigma_{ij}\epsilon_{ij}. \tag{2.74}$$

2.4.1   Isotropic elasticity theory

In this subsection the elastic stiffness tensor and its elements are analysed for materials with properties independent of direction. In this case the elements of the stiffness tensor do not change under any orthogonal transformation, i.e. length preserving unit transformations such as rotations and reflections. This leads to dependencies between the matrix elements and some of them becoming zero (details in Appendix A). In terms of shape of the stiffness tensor, an isotropic material is very close to the cubic system. For a cubic system the stiffness tensor becomes:

$$\mathbf{c}_{cub} = \begin{pmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{44} \end{pmatrix} \tag{2.75}$$

If, in addition to the stiffness tensor above, complete isotropy is considered, the number of independent elements in the stiffness tensor decreases to two with the relationship

$$2c_{44} = c_{11} - c_{12}. \tag{2.76}$$

In fact for anisotropic materials the factor

$$A = \frac{2c_{44}}{c_{11} - c_{12}} \tag{2.77}$$

is called the anisotropy ratio and a deviation from 1 is a measure of anisotropy. Since at this point only two elastic constants remain independent, it is common to define

$$\mu = c_{44} \quad = \quad \frac{1}{1}(c_{11} - c_{22}) \tag{2.78}$$

$$\lambda \quad = \quad c_{12} \tag{2.79}$$

$$\Rightarrow \quad \lambda + 2\mu \quad = \quad c_{11}. \tag{2.80}$$

Where λ and μ are called the Lamé constants. μ specifically is called the shear modulus, as it is the ratio of shear stress to shear strain. Expressing the elastic constants tensor with Lamé constants then yields:

$$\mathbf{c}_{iso} = \begin{pmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}. \tag{2.81}$$

or the same in index notation:

$$\sigma_{ij} = \lambda\epsilon_{kk}\delta_{ij} + 2\mu\epsilon_{ij} \tag{2.82}$$

At this point, the relations between specific stresses and resulting strain can be quantified for each material with several constants. Those include the Young's modulus (E, ratio of simple tensile stress to strain), the bulk modulus (B, ratio of pressure to volume change) and Poisson's ratio (ν, transverse contraction to elongation in simple tension). These can be expressed for isotropic materials:

$$B = \frac{\frac{1}{3}(\sigma_{11}+\sigma_{22}+\sigma_{33})}{\epsilon_{11}+\epsilon_{22}+\epsilon_{33}} = \lambda + \frac{2}{3}\mu \tag{2.83}$$

$$E = \frac{\sigma_{11}}{\epsilon_{11}} = \frac{\mu(2\lambda+2\mu)}{\lambda+\mu}, \qquad\qquad \wedge \quad \sigma_{22} = \sigma_{33} = 0 \tag{2.84}$$

$$\nu = \frac{\frac{1}{2}(\epsilon_{22}+\epsilon_{33})}{\epsilon_{11}} = \frac{\lambda}{2(\lambda+\mu)}, \qquad\qquad \wedge \quad \sigma_{22} = \sigma_{33} = 0 \tag{2.85}$$

and in turn the Lamé constants can then be expressed as

$$\lambda = 3\frac{E\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{E}{2(1+\nu)}, \tag{2.86}$$

Planar conditions

Solving for strain and stress fields can be largely simplified if a displacement field is present that is independent of one dimension. In this special case a differential

equation for a scalar field can be constructed in order to solve for all strain and stress components.

If the independent dimension is chosen to be along the x_3 axis the general displacement field reads:

$$\mathbf{U}(\mathbf{x}) = \begin{pmatrix} u_1(x_1, x_2) \\ u_2(x_1, x_2) \\ 0 \end{pmatrix}. \tag{2.87}$$

This also implies that all derivatives of the displacement field and by extension stress and strain fields with respect to x_3 will be zero. Therefore equation (2.66) without external forces will simplify to

$$\frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{12}}{\partial x_2} = 0 \tag{2.88}$$

$$\frac{\partial \sigma_{12}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} = 0 \tag{2.89}$$

The two equations above can be solved with the scalar field ψ:

$$\frac{\partial^2 \psi}{\partial x_2^2} = \sigma_{11}, \qquad \frac{\partial^2 \psi}{\partial x_1^2} = \sigma_{22}, \qquad \frac{\partial^2 \psi}{\partial x_1 \partial x_2} = \sigma_{12} \tag{2.90}$$

Similarly, looking at the definition for the strain field (2.68) under planar conditions one can construct

$$\frac{\partial^2 \epsilon_{11}}{\partial x_2^2} + \frac{\partial^2 \epsilon_{22}}{\partial x_2^2} = 2\frac{\partial^2 \epsilon_{12}}{\partial x_1 \partial x_2} \tag{2.91}$$

Using the elastic compliance tensor for isotropic materials (equations (2.71) and (2.81)) to rewrite the above in terms of σ_{ij} and inserting then the definition for ψ will lead to the following differential equation:

$$\frac{\partial^4 \psi}{\partial x_1^4} + 2\frac{\partial^4 \psi}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 \psi}{\partial x_2^4} = 0 \tag{2.92}$$

$$\nabla^2(\nabla^2 \psi) = 0 \tag{2.93}$$

$$\nabla^4 \psi = 0 \tag{2.94}$$

By solving the above equation for ψ (the Airy stress function) one can construct the general solution for all internal strain and stress effects under planar conditions. This can readily be applied to the edge dislocation as its displacement

field satisfies planar conditions. By qualitatively selecting relevant parts of the solution for the Airy stress function the displacement field for the edge dislocation can be constructed.

2.4.2 Elasticity theory for hexagonal systems

Hexagonal systems such as wurtzite InN and GaN are anisotropic (as are cubic), however they are usually described by a transverse isotropic symmetry. This means that all rotations and reflections in two dimensions are symmetries. At this point the x_3 axis is chosen to be the axis of symmetry, which is the axis around which all rotations are symmetries and as a result is orthogonal to the symmetry plane. Under these conditions the stiffness tensor can be shown to take the form (for a complete analysis see Appendix A):

$$
\mathbf{c}_{hex} =
\begin{pmatrix}
c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\
c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\
c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\
0 & 0 & 0 & c_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & c_{44} & 0 \\
0 & 0 & 0 & 0 & 0 & c_{66}
\end{pmatrix},
\tag{2.95}
$$

where the additional relationship $c_{66} = \frac{1}{2}(c_{11} - c_{12})$ reduces the number of independent matrix elements to five.

In the case of a hexagonal system the elastic moduli described above for isotropic materials are not as simply expressed, as there is now a unique direction. For the Young's modulus we separate $E_T = \frac{\sigma_{33}}{\epsilon_{33}}$ and $E_P = \frac{\sigma_{11}}{\epsilon_{11}} = \frac{\sigma_{22}}{\epsilon_{22}}$ as the planar and transverse part. Similarly the Poisson's ratio and shear modulus have separate values for $xz = yz$ and xy directions.

2.5   Theory of Dislocations

The following discussion of dislocations fol-
lows the description given in [46] but does
not go into great detail. It utilises the prin-
ciples of linear elasticity theory outlined in
the previous section as a description of the
long range deformation of general disloca-
tions.

A general line defect is a one dimensional
defect that can be classified with a line
direction and a Burgers vector. The line
direction describes the path of the defect

Figure 2.3: Schematic of the integra-
tion path for a Burgers circuit

through the material while the Burgers vector is a measure of the crystal de-
formation that is caused by the defect. The Burgers vector $\boldsymbol{\mathcal{B}}$ can be defined
using the displacement field $\mathbf{U}(\mathbf{r})$ that represents the deformation from the per-
fect crystal. The local Burgers vector is hence given by the line integral around
the defect line $\mathbf{L}$, which is taken using a curve C in a right handed sense:

$$\boldsymbol{\mathcal{B}} = \oint_C \frac{\partial \mathbf{U}}{\partial s} ds. \tag{2.96}$$

A line defect where $\mathcal{B} \neq 0$ is then called a dislocation. In a crystal, it follows
that the length of the Burgers vector is constrained by the nature of the crystal
lattice. This enables the definition of a perfect dislocation as a line defect with
a Burgers vector that is a linear combination of integer lattice vectors, while a
partial dislocation is a line defect with a Burgers vector containing fractions of a
lattice vector.

As a perfect dislocation is only classified by its Burgers vector and line direction,
two limiting cases can be identified in the extrema of the relationship of their rela-
tive direction. When line direction and Burgers vector are orthogonal the perfect
dislocation is called an edge dislocation, while the parallel case is described as a
screw dislocation. Consequently any other perfect dislocation can be decomposed
into its edge and screw components, including the long range strain field within
linear elasticity. Although pure edge and screw type dislocations are commonly
present in materials a "linear combination" of those can also be found, called the
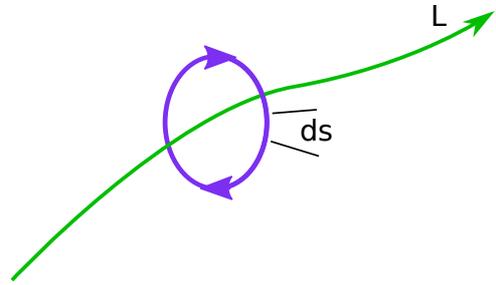mixed dislocation.

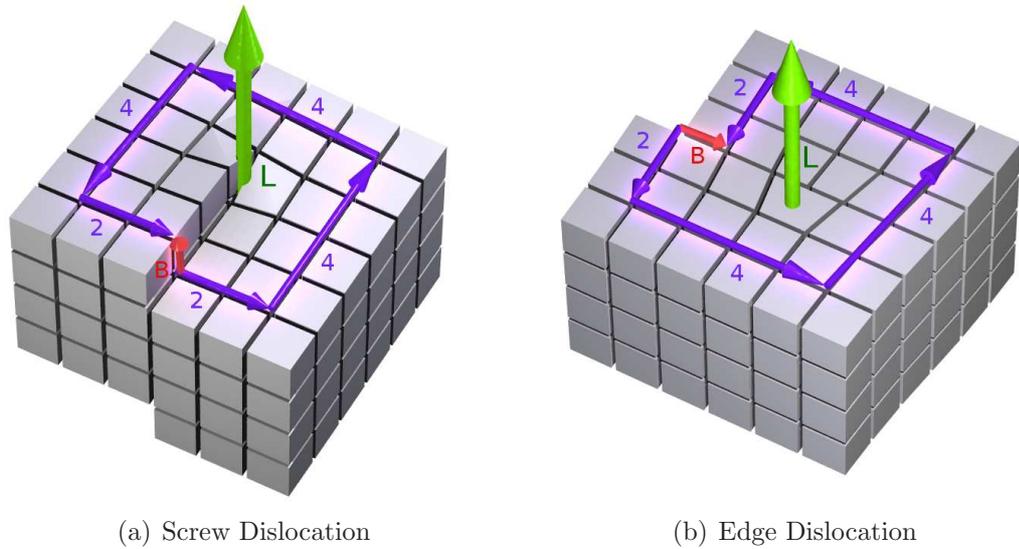(a) Screw Dislocation   (b) Edge Dislocation

Figure 2.4: Schematic of a screw and edge dislocation. Burgers vector is shown in red, line direction in green and Burgers circuit in purple.

2.5.1   Screw dislocations

A screw dislocation is a perfect dislocation with a Burgers vector $\mathcal{B}_s$ parallel to its line direction, therefore

$$\mathbf{L} \times \mathcal{B}_s = 0 \tag{2.97}$$

holds true. A schematic of a screw dislocation in a cubic lattice can be found in figure 2.4(a), where the line direction, Burgers vector and Burgers circuit are depicted.

At this point isotropic elasticity theory (section 2.4) is applied to a straight screw dislocation for a quantitative description of the resulting crystal deformation. The deformation necessary to produce such a screw dislocation in polar coordinates can be deduced from figure 2.4(a) to be a smoothly increasing displacement from 0 to $\mathcal{B}_s$ with the angle θ (0 to 2π) around the dislocation centre. Therefore the displacement field $U(\mathbf{r})$ for a screw dislocation in x_3 direction will only consist of one non-zero component:

$$u_3 = \frac{\mathcal{B}_s}{2\pi}\theta = \frac{\mathcal{B}_s}{2\pi}\tan^{-1}\left(\frac{x_2}{x_1}\right), \tag{2.98}$$

wherer $\mathcal{B}_s$ is the length of the Burgers vector. From the singular nature of the displacement field at $\mathbf{x} = 0$ it can already be seen that linear elasticity theory
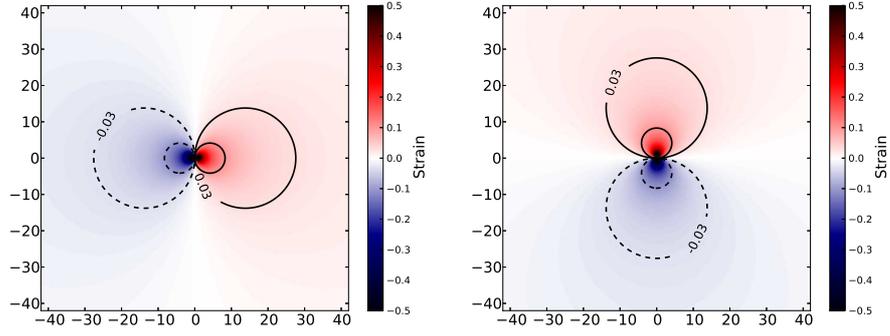
Figure 2.5: Non-zero strain components for a screw dislocation in an isotropic medium with a Burgers vector length of 5.2Å.

will not be applicable close to the dislocation centre. In fact the threshold inside which elasticity theory breaks down and the atomic configuration in a real cell is not being described correctly by elasticity theory any more is marked by the so called the Dislocation Core Radius R_c. Therefore elasticity theory is only applicable to regions away from the dislocation centre and will only be used in the analysis chapters to find R_c itself and verify dislocation centre positions.

From equations (2.68) and (2.98) it can be determined that for a screw dislocation the strain components become:

$$\epsilon_{11} = \epsilon_{22} = \epsilon_{33} = \epsilon_{12} = 0 \tag{2.99}$$

$$\epsilon_{23} = \frac{\mathcal{B}_s}{4\pi} \frac{x}{x^2 + y^2} = \frac{\mathcal{B}_s}{4\pi} \frac{\cos(\theta)}{r} \tag{2.100}$$

$$\epsilon_{13} = -\frac{\mathcal{B}_s}{4\pi} \frac{y}{x^2 + y^2} = -\frac{\mathcal{B}_s}{4\pi} \frac{\sin(\theta)}{r} \tag{2.101}$$

and with the elastic constants for isotropic media (2.81) the stress components become:

$$\sigma_{11} = \sigma_{22} = \sigma_{33} = \sigma_{12} = 0 \tag{2.102}$$

$$\sigma_{23} = \frac{\mathcal{B}_s\mu}{2\pi} \frac{x}{x^2 + y^2} = \frac{\mathcal{B}_s\mu}{2\pi} \frac{\cos(\theta)}{r} \tag{2.103}$$

$$\sigma_{13} = -\frac{\mathcal{B}_s\mu}{2\pi} \frac{y}{x^2 + y^2} = -\frac{\mathcal{B}_s\mu}{2\pi} \frac{\sin(\theta)}{r} \tag{2.104}$$

It can be seen that the elastic stresses and strains do not have any tensile or compressive parts, consisting solely of shear components that only vary in the plane orthogonal to the line direction.

Elastic and dislocation core energy

As now all strain and stress components of the screw dislocation have been determined, the energy density of the strain field as defined in equation (2.74) can be derived and becomes:

$$\mathcal{U}(\mathbf{r}) \;=\; \frac{\mathcal{B}_s^2 \mu}{8\pi^2}\left(\frac{\sin^2(\theta)}{r^2}+\frac{\cos^2(\theta)}{r^2}\right) \tag{2.105}$$

$$=\; \frac{\mathcal{B}_s^2 \mu}{8\pi^2}\frac{1}{r^2} \tag{2.106}$$

The elastic energy density can be integrated in a cylindrical volume around the dislocation line in order to derive the total dislocation energy. However, as mentioned above the singular nature of the strain field results in a diverging, unphysical situation close to the dislocation centre. Therefore it is to be expected that in a real crystal a certain dislocation core radius exists at which the linear elasticity theory becomes invalid and a more realistic atomistic theory needs to be applied inside. As a result the energy of the strain field is only included in the total energy of a real dislocation for the region outside the dislocation core radius R_c. The integration of the elastic strain energy density then becomes:

Figure 2.6: Schematic of the integration path for a Burgers circuit

$$E_{el}/L \;=\; \int_{R_c}^{R}\mathcal{U}(r)2\pi r\,dr d\theta \tag{2.107}$$

$$=\; \frac{\mathcal{B}_s^2 \mu}{4\pi}\ln\left(\frac{R}{R_c}\right) \tag{2.108}$$

Where L is the integration cylinder length in the line direction. The elastic energy E_{el} is also divergent as $R \rightarrow \infty$ and therefore means that the elastic
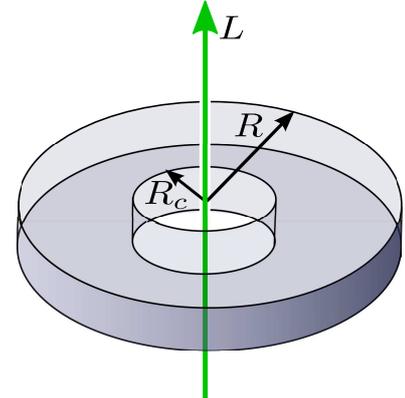
energy cannot be used as a measure to describe a dislocation. However, within a cylinder of radius R around a screw dislocation the total energy is:

$$E^{screw}/L = E_{core}/L + E_{el}(R)/L. \tag{2.109}$$

And thus in elasticity theory the dislocation core radius R_c and dislocation core energy E_{core} are defined as the minimum distance from the dislocation line to where the elastic energy describes the crystal well and as difference of the elastic energy E_{el} outside the core radius and the total dislocation energy E^{screw}.

The above description of the screw dislocation elastic fields and energy has been for isotropic media only. However for hexagonal or even orthorhombic systems there will be a difference in the primitive Burgers vector lengths for various directions and as a result the energy density in the dislocation strain field will be directionally dependent. The forces on atoms will also have differences in direction, since the two shear stress components in the case of hexagonal materials have different shear moduli if the line direction is not perpendicular to the isotropic plane. However, these issues will not be discussed here further, for an analysis of dislocations in anisotropic media refer to [46].

2.5.2   Edge dislocation

An edge dislocation is a perfect dislocation with a Burgers vector $\boldsymbol{\mathcal{B}}_e$ orthogonal to the line direction $\mathbf{L}$:

$$\mathbf{L} \cdot \boldsymbol{\mathcal{B}}_e = 0. \tag{2.110}$$

This means that topologically speaking the edge dislocation is obtained by inserting a half-plane of lattice sites into the crystal perpendicular to both line direction and Burgers vector which terminates at the dislocation line. This can be seen in the graphical represention of an edge dislocation in figure 2.4(b), where a Burgers circuit is drawn around the dislocation line, revealing the Burgers vector for a cubic lattice structure.

The displacement field of the edge dislocation can be found by acknowledging the applicability of the constraints of planar stress. Which is to say that the displacement field acts only in two dimensions and is independent of the third. All non-zero components lie in the plane orthogonal to the line direction of the edge dislocation. Therefore solving for the Airy stress function (2.94) will provide a general solution for stresses in the two dimensional case without external forces,

including the stress distribution for the edge dislocations.

Solving for ψ becomes much easier when considering another scalar function defined as:

$$\phi = (\sigma_{11} + \sigma_{22}) = \nabla^2 \psi. \tag{2.111}$$

This results in equation (2.94) simplifying to the Laplace equation:

$$\nabla^4 \psi = \nabla^2 \phi = 0 \tag{2.112}$$

which reads in polar coordinates as:

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} \right) \phi = 0 \tag{2.113}$$

The standard solution to this can be found readily by realising that the ϕ is separable in its variables. However, only part of the solution corresponds to the long range field of the edge dislocation. By qualitativeley analyzing the geometry of the edge dislocation and only incorporating long range strains (details can be found in [46]), the relevant part of the general solution can be identified to be:

$$\phi_e = \frac{\beta}{r} \sin\theta, \tag{2.114}$$

where β is an unknown constant at this point. Which yields for ψ through integration:

$$\psi_e = \frac{\beta}{2} r \ln r \sin\theta \tag{2.115}$$

The constant β can now be found by producing the strain function with the above solution for ψ and integrating along a Burgers circuit. The result should therefore be equal to the Burgers vector, which means that

$$\beta = -\frac{\mu \mathcal{B}_e}{\pi(1-\nu)}. \tag{2.116}$$

Inserting β into the equations for the stress (see section 2.4.1 for details) will then produce:

$$\sigma_{11} = -\frac{\mu \mathcal{B}_e}{2\pi(1-\nu)} \frac{x_2(3x_1^2 + x_2^2)}{r^4} \tag{2.117}$$

$$\sigma_{22} = \frac{\mu \mathcal{B}_e}{2\pi(1-\nu)} \frac{x_2(x_1^2 - x_2^2)}{r^4} \tag{2.118}$$

$$\sigma_{12} = \frac{\mu \mathcal{B}_e}{2\pi(1-\nu)} \frac{x_1(x_1^2 - x_2^2)}{r^4} \tag{2.119}$$

$$\sigma_{33} = -\frac{\mu \nu \mathcal{B}_e}{\pi(1-\nu)} \frac{x_2}{r^2} \tag{2.120}$$

$$\sigma_{13} = \sigma_{23} = 0 \tag{2.121}$$

and for the strain, by use of the compliance tensor:

$$\epsilon_{11} = -\frac{\mathcal{B}_e}{4\pi(1-\nu)} \left((1-2\nu)\frac{x_2}{r^2} + \frac{x_2 x_1^2}{r^4} \right) \tag{2.122}$$

$$\epsilon_{22} = \frac{\mathcal{B}_e}{4\pi(1-\nu)} \left((1+2\nu)\frac{x_2}{r^2} - \frac{2x_2^3}{r^4} \right) \tag{2.123}$$

$$\epsilon_{12} = \frac{\mathcal{B}_e}{4\pi(1-\nu)} \frac{x_1(x_1^2 - x_2^2)}{r^4} \tag{2.124}$$

$$\epsilon_{33} = \epsilon_{13} = \epsilon_{23} = 0 \tag{2.125}$$

and finally, the displacement field can be found by integrating the strain tensor and setting the constant of integration for u_1 with the boundary values of $u_1(x_1, 0) = 0$. The constant of integration for u_2 is set to $C = b/3\pi(1-\nu)$ in order to create a symmetric expression in x_1 and x_2. This is done in analogy to [46, 47]

$$u_1(x_1, x_2) = \frac{\mathcal{B}_e}{2\pi} \left(tan^{-1}\left(\frac{x_2}{x_1}\right) + \frac{1}{2(1-\nu)}\frac{x_1 x_2}{r^2} \right) \tag{2.126}$$

$$u_2(x_1, x_2) = -\frac{\mathcal{B}_e}{2\pi} \left(\frac{1-2\nu}{2(1-\nu)} ln(r) + \frac{1}{4(1-\nu)}\frac{x_1^2 - x_2^2}{r^2} \right). \tag{2.127}$$

The above solutions for strain, stress and displacement field around an edge dislocation can be seen in Figure 2.7.

(a) Displacement field $\mathbf{u}(\mathbf{r})$

(b) Strain ϵ_{11}

(c) Strain ϵ_{12}
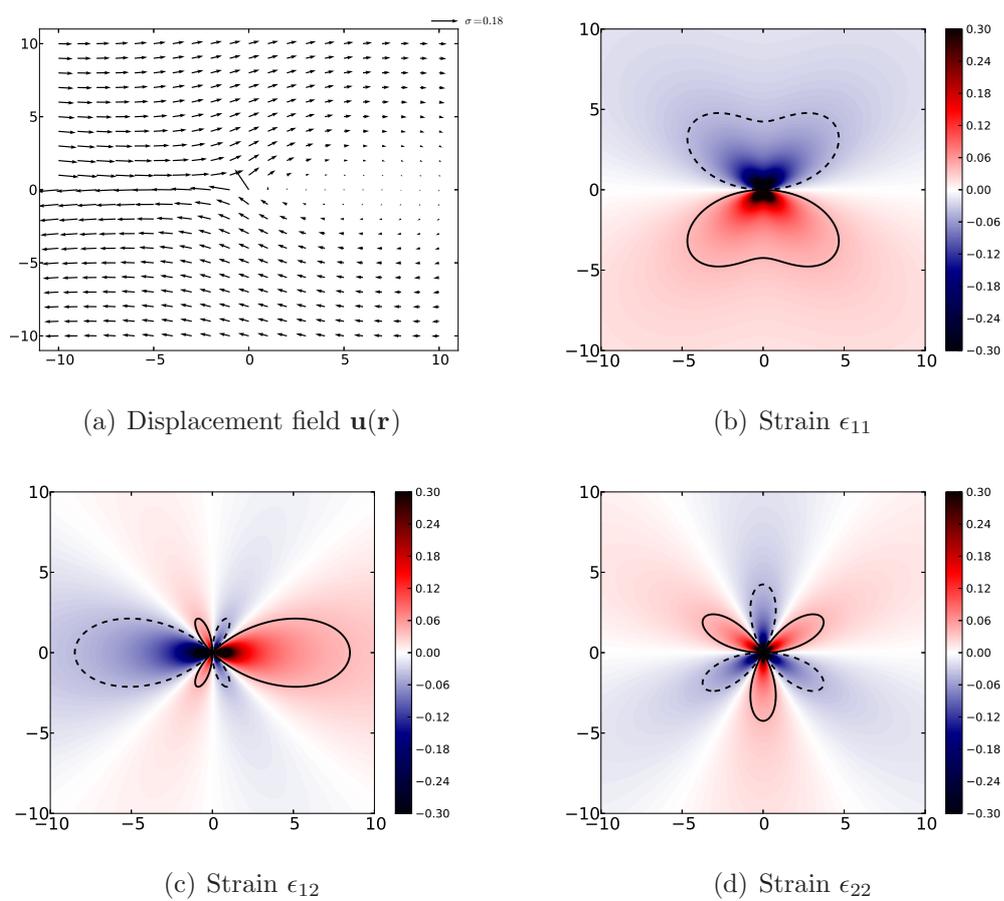
(d) Strain ϵ_{22}

Figure 2.7: Strain components and displacement field of an edge dislocation

Elastic and dislocation core energy

The elastic energy density for the edge dislocation and its subsequent cylindrical integration can be done in analogy to the screw dislocation arriving at the form for the total dislocation energy:

$$\frac{E^{edge}(R)}{L} = \frac{E_{core}}{L} + \frac{\mu \mathcal{B}_e^2}{4\pi(1-\nu)} \log\left(\frac{R}{R_c}\right).$$

(2.128)

2.5.3 Mixed dislocations

Mixed dislocations are perfect dislocations that have a Burgers vector $\boldsymbol{\mathcal{B}}_m$ that can be decomposed into edge and screw dislocations.

$$\boldsymbol{\mathcal{B}}_s = (\boldsymbol{\mathcal{B}}_m \cdot \mathbf{L})\mathbf{L}$$

(2.129)

$$\boldsymbol{\mathcal{B}}_e = \mathbf{L} \times (\boldsymbol{\mathcal{B}}_m \times \mathbf{L})$$

(2.130)

Within *linear* elasticity any elastic field of a dislocation that is a linear combination of edge and screw dislocations can be described by a simple addition of their component strain and stress fields. Due to the fact that mixed dislocations will not be discussed within the results of this work, no detailed description of their elastic fields will be given here.

Elastic and dislocation core energy

It is possible to describe a general mixed dislocation total energy with the expression

$$\frac{E^{mixed}(R)}{L} = \frac{E_{core}}{L} + \frac{k(\beta)\mathcal{B}_m^2}{4\pi} \log\left(\frac{R}{R_c}\right),$$

(2.131)

where β is the angle between the Burgers vector and the line direction of the given dislocation. $k(\beta)$ is then an energy factor and depends on β and the elastic medium. Assuming isotropic condition it takes the form

$$k(\beta) = \mu\left(\cos^2\beta + \frac{\sin^2\beta}{1-\nu}\right)$$

(2.132)

This expression can be found by following the same procedure as for the screw dislocation, however including a linear combination of edge and screw dislocations. The somewhat lengthy construction is also not shown here but can be

found in [46] or [47].

2.5.4   Partial dislocations and stacking faults

Partial dislocations are dislocations with a Burgers vector length of a fraction of the underlying crystal lattice vectors. Therefore a lattice mismatch has to exist along the dislocation line during the Burgers circuit, creating a stacking fault. This can occur in crystals where the primitive unit cell is composed of multiple atoms, allowing for atomic neighbours that are not in their native configuration. For specific examples of partial dislocations in the wurtzite lattice please refer to section 3.3.2.

In an infinite crystal, a stacking fault is a two dimensional structure that is bounded by partial dislocations at either site. These partial dislocations together have then a total Burgers vector of integer lattice vector repeats. Therefore they can be created by a dissociation of a perfect dislocation. However in real crystals partial dislocations and therefore stacking faults can be created in various ways including errors during growth that can also be bounded by the surface.

Group III Nitrides fundamentals

This chapter covers the properties and applications of the Group-III nitrides and provides a review of the literature on these topics. First the fundamental properties such as the crystal lattice and mechanical behavior and growth challenges are discussed. Then the electronic properties are discussed with an emphasis on GaN and InN including surface properties and point defects such as native point defects and dopants. Last but not least the most common dislocation types in wurtzite material are presented including their occurrence and effect in GaN and InN.

3.1 Mechanical properties

3.1.1 Crystal structure

The group-III nitrides are usually associated with indium, gallium and aluminium nitride, however the group-III also includes boron and thallium. While boron nitride can form as a polymorphic solid under normal conditions with semiconducting properties and can even be used to construct boron nitride nanotubes, thalium nitride (Th_3N) is only stable in ammonia solutions. As a result henceforth only indium, gallium and aluminium nitride will be considered. These group-III nitrides are composed of two atomic types and can therefore form hetero struc-

ture semiconductors. In addition, a tetrahedral coordination is found to be most stable, which means that all three compounds are commonly found in a wurtzite or zincblende structure, with wurtzite being the more stable configuration.

The zincblende structure features a cubic lattice where all atoms have the same tetrahedral coordination making this similar to a diamond cubic configuration, but where next nearest neighbours are of alternating atomic type (point symmetry group T_d). The wurtzite structure is also composed of tetrahedrally coordinated atoms with alternating atomic type, however there are two distinct bonding sites present in a unit cell in a hexagonal lattice (space group C_{6v}). Figure 3.1 displays the wurtzite lattice including its lattice vectors and the major planes: polar c-plane (0001) as well as non-polar a-plane ($1\bar{1}00$) and m-plane ($2\bar{1}\bar{1}0$). While the zincblende structure can be fully classified through its lattice constant, the wurtzite structure is further specified by the ratio of its two lattice constants c/a and the internal parameter u, that determines the nitrogen group-III bond distance in the c-direction. The ideal wurtzite structure provides identical bond lengths for all atoms, in this case $c/a = 2\sqrt{2/3}$ and $u = 3/8$ and each atomic type features a hexagonal close packed sublattice.

First principles DFT-LDA (section 2.1) can predict structural parameters well for the group-III nitrides, even though the $3d$ and $4d$ electrons in GaN and InN bonding respectively are not included in the pseudopotential method actively, but are commonly included into pseudopotentials via NLCCs (see section 2.1.6). The interaction of the outer d electrons in the group-III atoms and the p-shell electrons of nitrogen create a repulsion term that is often incorrectly handled in even an all-electron DFT calculation [48] and results in a shift of the energies of the valence band. However total structural energies and therefore structural parameters are not affected as strongly. A comparison of LDA-DFT predicted structural parameters can be found in table 3.1.
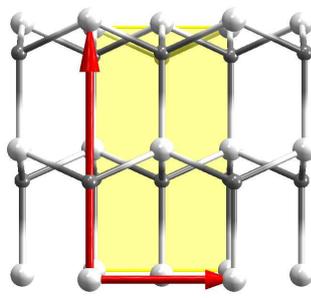
The hardness of the materials, as specified by the bulk modulus B, reduces the heavier the group-III atom is. In addition, InN in the wurtzite structure has the largest lattice constants of the group-III nitrides, but is still the most dense at $6.81\,\text{g/cm}^3$.
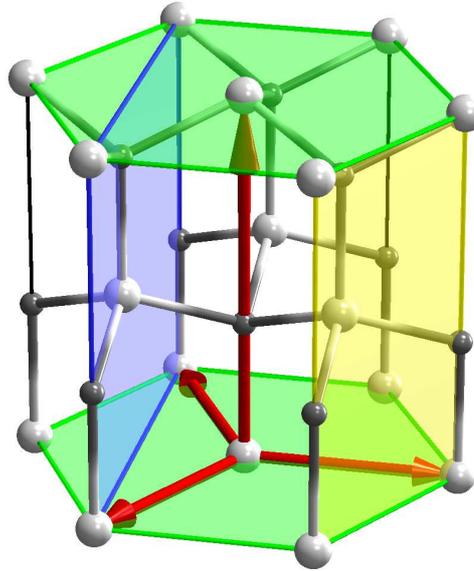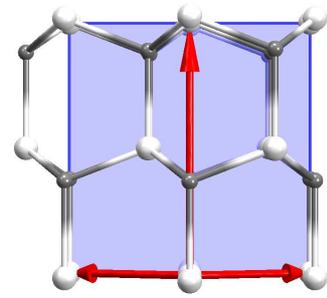
(a) Non-primitive unit cell of a wurtzite crystal
depicting planes and lattice vectors
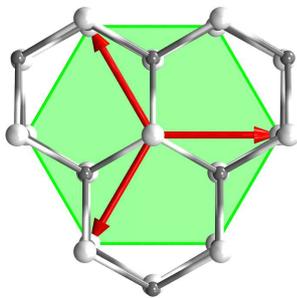
(b) c-plane (0001)   (c) m-plane (1$\bar{1}$00)   (d) a-plane (2$\bar{1}\bar{1}$0)

Figure 3.1: Depiction of the wurtzite lattice. Atoms are represented by a ball and stick model, lattice vectors are represented by red arrows, green surface represents c-plane, blue surface a-plane and yellow surface m-plane.

Table 3.1: Structural parameters for AlN, GaN and InN. Computed with LDA-DFT, lattice constants are given in Å, bulk moduli in *Mbar*.

	zinc-blende		wurtzite				
	a	B	a	c	c/a	u	B
InN	5.0	-	3.54	5.71	1.61	0.375	-
[1]	5.0	1.4	3.54	5.76	1.63	0.377	1.4
exp[2]	-	-	3.538	5.703	1.612	-	-
GaN	-	-	3.18	5.21	1.64	0.376	-
[1]	4.52	1.9	3.19	5.22	1.64	0.376	2.0
exp[2]	-	-	3.189	5.185	1.626	-	-
AlN[1]	4.31	2.1	3.06	4.94	1.61	0.378	2.1
exp[2]	-	-	3.111	4.980	1.601	-	-

[1] Ref [48]

[2] Ref [49]

3.1.2   Growth

GaN and InN crystals are usually grown using molecular beam epitaxy (MBE) or vapour deposition methods such as MOCVD and HVPE, on sapphire or silicon carbide substrates. This creates a large lattice mismatch of up to 10% for InN which in turn causes creation of high structural defect densities. The sapphire growth supports growth for multiple directions. For c-plane growth this means that interfaces are created at the polar c-plane, resulting in the formation of strain fields that result in the creation of electric fields due to the piezoelectricity of the wurtzite structure on these polar surfaces. However, the magnitude of these strain fields depends on the lattice mismatch, misfit dislocation density and growth temperatures.

3.2   Electronic properties

The following discussion of electronic properties is limited to the properties of wurtzite GaN and InN. As discussed in section 2.1 any crystal can be described by its primitive unit cell and its lattice constants, which can be used to define the Brillouin zone. The Brillouin zone for any hexagonal structure is displayed in figure 3.2 where the irreducible zone is marked including special points. The energy momentum dispersion of electrons in any infinite hexagonally ordered structure
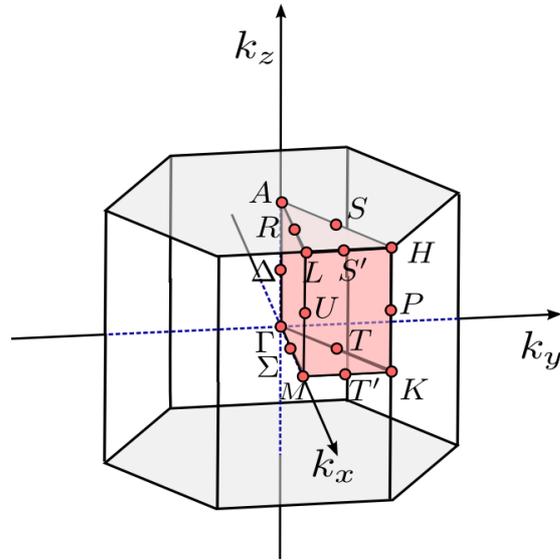
Figure 3.2: Brillouin zone for hexagonal symmetry including irreducible zone with special points marked

can be described by following a complete path around the irreducible zone, such a description is called a band structure. The integration of the band structure over the Brillouin zone produces the density of states.

The lack of a centre of a symmetry in the wurtzite structure of the group-III nitrides provides the potential for piezoelectricity along the polar directions, which results in the creation of polarisation when the material is strained. Therefore, a two dimensional electron gas (2DEG) forms on interfaces with other materials [3], depending on the band offset. The formation of a very narrow width 2DEG and the high electron mobility ($>$ 3000 cm²/Vs [4]) can be used for high-frequency electronics. The symmetry of the hexagonal cell also results in either polar or non-polar structures depending on the crystal orientation. In polar quantum well structures a spatial separation of electron and hole wavefunction is usually present prohibiting efficient recombination[50].

3.2.1 Band structure

The band structure and density states of InN and GaN can be found in figure 3.3. The band structure displays the E-k dispersion in reciprocal space of the material for a specific path through the Brillouin zone and the density of states displays

an integration of the electron states over all k-points. The band gap generally refers to the difference in energy between the conduction band minimum and the valence band maximum. In the case of the group-III nitrides these two points both occur at $\mathbf{k} = 0$, the γ-point, making them direct band gap semiconductors.

Group-III nitrides and their alloys can have band gaps that easily span the entire visible spectrum and are therefore ideal for all applications in optoelectronics. However, the band gap of InN has been in discussion for several years. The latest theoretical [51] and experimental [52] results both point to a band gap of 0.69 eV, which is now widely accepted. The origin of this discussion stems from the difficulty growing high quality InN and the challenges to theory in predicting excited states energies accurately. The low dissociation temperature of InN and the high equilibrium vapour pressure of N^2 over InN create a difficult environment for crystal growth. Only once high quality (large electron mobility and reduced free electron concentration) thick InN films had been grown, could the experimental band gap estimates converge from the original ~ 1.9 eV [53] to the recent lower value.

While predicting lattice parameters with good accuracy, DFT-LDA usually underestimates the bandgap of semiconductors. In the case of InN it has predicted metallic behaviour with a negative band gap [48]. For GaN it predicts a band gap of around 2.1 eV[48]. In order to compensate for the improper approximation of the exchange correlation functional, alternative methods have been applied by several groups. Before the new experimental narrow band gap was found, self-interaction corrected pseudopotentials were applied, yielding band gap estimations around 1.6 eV [54]. With the new experimental value, semi-empirical LDA methods were used to adjust and produce band gaps of 0.85 eV by using external delta-function potentials [55] and 0.81 eV by adjusting the atomic pseudo-potentials for relaxation-corrected self interaction [56]. A first principles result of 0.69 eV has also been obtained by employing a GW method based on Kohn-Sham states [51].

3.2.2 Surface properties

Surface states appear in all crystals due to the incomplete bonding arrangement and hence large quantity of defect states at the surface. As a result the Fermi level is strongly influenced as it tends towards the Fermi-level stabilisation energy [57]

(a) Band structure for InN

(b) Band structure for GaN

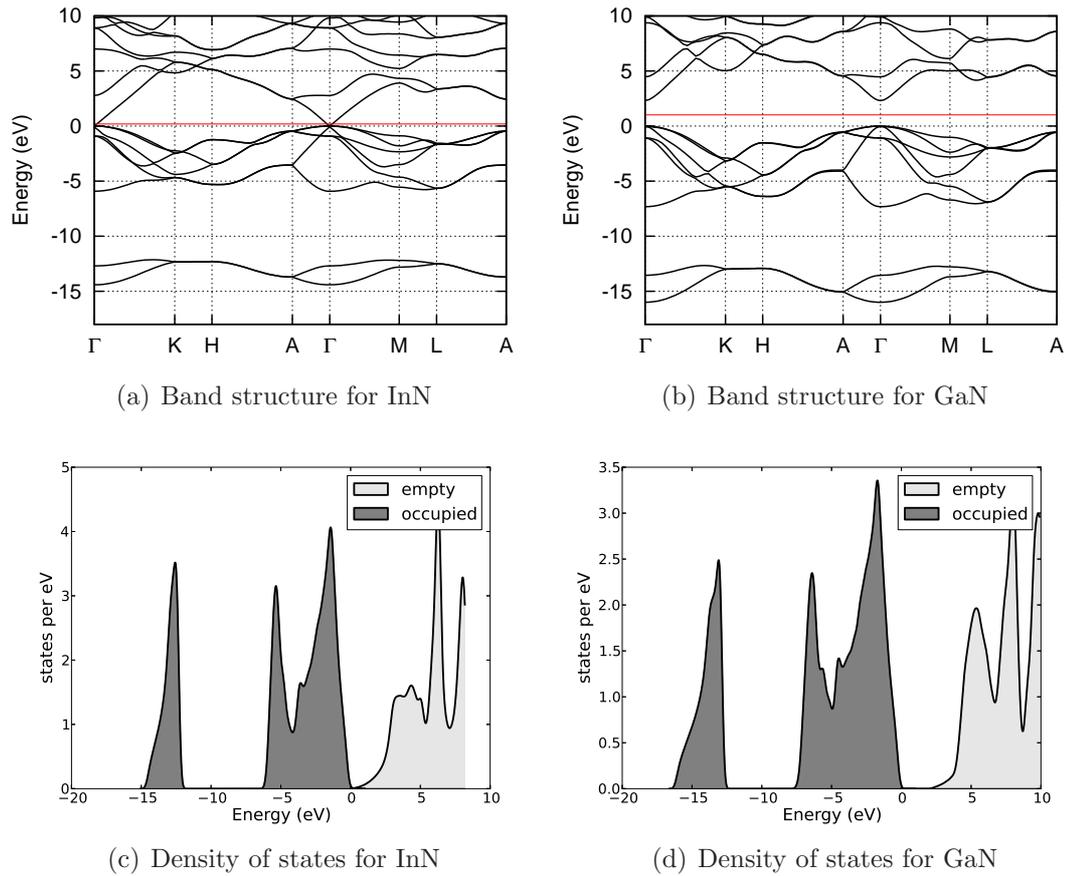(c) Density of states for InN

(d) Density of states for GaN

Figure 3.3: Band structures and density of states for InN and GaN computed with DFT-LDA, the valence band edge has been set to zero, the Fermi energy has been marked by a red line in the band structures.

and is pinned by these states. Depending on the relative position of the defect statest the Fermi level pinning at the surface can act to create Schottky-barriers when contacted. Due to this surface Fermi pinning effect and the expected position of the Fermi stabilisation energy within the conduction band, a charge accumulation layer is expected to form in InN [58]. DFT calculations present evidence that this charge accumulation exists for InN on polar surfaces while it remains negligible on non-polar surfaces [59, 60]. The charge accumulation layer is present for both n-type and p-type InN, creating an n-p junction on the surface. This complicates the analysis of Hall-measurements of p-type InN [3].

3.2.3   Vacancy defects

While the properties of the ideal bulk are important, a real material will contain defects, including both deliberate and accidental doping and also native defects such as vacancies. Of the two types of vacancy defects, the nitrogen vacancy has been shown by ab-initio methods to be energetically more favourable in equilibrium bulk conditions. It also acts as a source of electrons, either acting as an n-type dopant or to compensate p-type doping. For InN the formation energy of the nitrogen vacancy varies with the level of background doping and is lower for p-type material [6]. It was also found that in n-type conditions the neutral charge state is favoured, which results in the formation of "nitrogen vacancy clusters"[61] and regions with higher In-content creating metallic regions [62].

Doping

The intentional incorporation of impurities in order to control the electronic behaviour of a material is called doping. These impurities are usually single atoms creating point defects but can be clusters of these. In the case of GaN common choices are silicon or oxygen for n-type doping and magnesium to p-type dope.

In the case of InN many sources act as n-type dopants. For instants hydrogen interstitial or substitutional sites have been shown to act as donors. While being inferior electron donors, interstitial hydrogen is a shallow single donor, substitutional hydrogen has lower formation energy and acts as a double donor[63]. Even nitrogen vacancy defects can also act as electron donors and compensate p-type doping. Given these sources InN has a tendency for high background elec-

tron densities. However, the realisation of optoelectronic devices depends on the ability to p-type dope and thus is of major interest. Mg doped InN has been shown experimentally by low temperature photoluminescence to act as a p-type material [64].

3.3   Extended defects

Extended defects such as topological defects represent strain relaxation mechanisms and are present in GaN and InN in very high concentrations, due to large lattice mismatch of the common growth substrates and the resulting strain during growth. These in grown dislocations are called threading dislocations if their line direction is not purely orthogonal to the growth direction. On the other hand misfit dislocations are found perpendicular to the growth direction and usually at interfaces and act as strain relaxation.

The precise origin or threding dislocationsin in GaN has been a source of controversy [65]. While it is clear that they commonly form during growth on interfaces, the notion that they form on island boundaries as the islands coalesce during the initial stages of growth has not been supported unanimously. Other results suggest that threading dislocations instead form near the interfaces independent of island boundaries[66], where most likely lattice error sites perform as nucleation sites.

3.3.1   Screw dislocations

Screw dislocations are perfect dislocations with a Burgers vector $\mathcal{B}$ in line direction $\mathbf{L}$. See section 2 for details. As the formation energy is proportional to $|\mathcal{B}|$ the screw dislocation is most commonly found in $[2\bar{1}\bar{1}0]$ and in $[0001]$ directions with Burgers vectors of $\frac{1}{3}[2\bar{1}\bar{1}0]$ and $[0001]$ respectively.

Most wurtzite GaN and InN material is grown in c-direction, therefore the threading screw dislocation is energetically unfavourable as it's Burgers vector is along the longer lattice repeat. In addition the strain field of the screw dislocation consists solely of shear components, which do not influence the in-plane compressive and tensile strains from the substrate mismatch. In current state of the art InN and GaN samples there are only about 2-8% screw dislocations. However their intrinsic properties are still very important as they are not directly caused by the

interface lattice mismatch during growth. The question of their dislocation core configuration and it's influence on the electronic properties will be discussed in chapters 4 and 5.

3.3.2 Edge dislocations

In wurtzite GaN and InN pure edge dislocations are commonly found with dislocation lines along the $[0001]$ or $[01\bar{1}0]$ direction, as these provide an orthogonal Burgers vector in the c-plane with the smallest lattice repeat in the crystal $(\boldsymbol{\mathcal{B}} = \frac{1}{3}[2\bar{1}\bar{1}0])$. This results in the smallest strain field and hence features the least strain energy. Therefore, as most GaN and InN samples are grown in c-direction the threading edge dislocations are usually most common, followed by threading mixed dislocations. However, depending on growth conditions these ratios can vary widely. In Section 6 the threading edge dislocation core will be studied in detail.

Screw dislocations in GaN

This chapter investigates the fully sto-ichiometric core structure of the screw dislocation in GaN with a line direction along **c**. The first section details modelling parameters for screw dislocations while the later section discusses an applied global optimisation of the dislocation core structure. The performed global search allows a systematic investigation into possible dislocation cores as predicted by Stillinger Weber (SW) type interatomic potentials. An emphasis on strain related issues is given as well as a discussion of the optimisation performance. Electronic structure results for the predicted dislocation cores are not presented here but can be found in Chapter 5, where a comparative study of dislocation cores across a wide stoichiometry range is given.
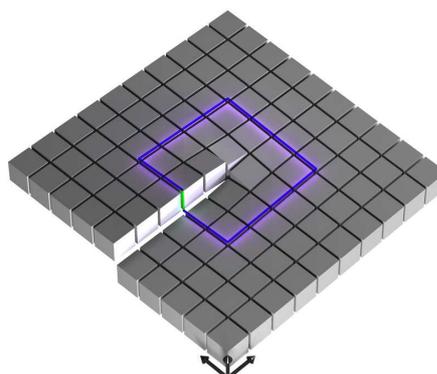
Figure 4.1: Schematic of a single screw dislocation in a cubic system.

4.1 Dislocation modelling in periodic boundary conditions

(a) Dipole configuration   (b) Quadrupole configuration
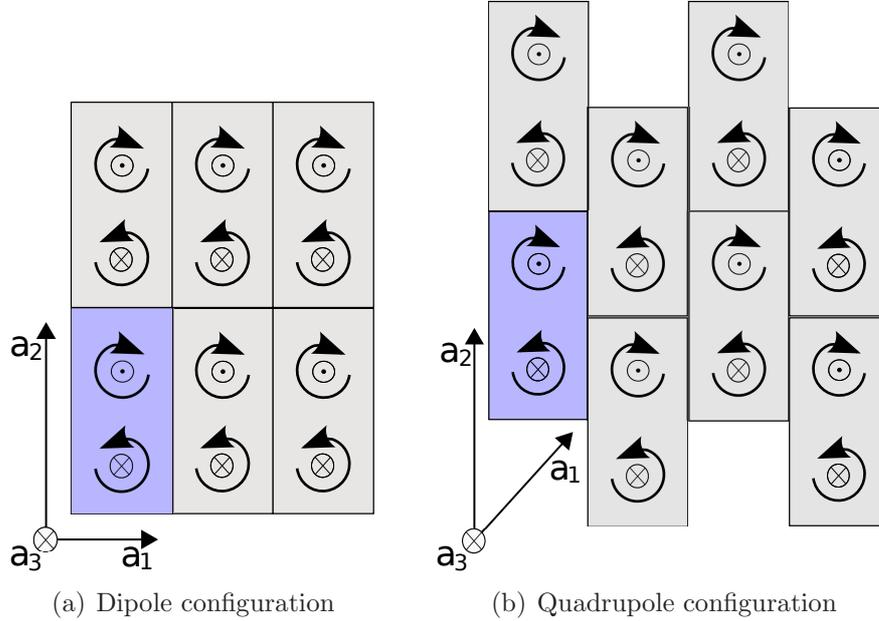
Figure 4.2: Comparison of periodic boundary conditions. Shows a plan view of screw dislocations in $\pm a_3$ direction, ↻ and ↺ mark dislocation centres and distinguish the sign. Dipole unit cells and lattice vectors a_1, a_2, a_3 are marked

The nature of the lattice distortion that defines a dislocation means that it is intrinsically impossible to model a single dislocation with periodic boundary conditions (PBC). The Burgers vector is a measure of the lattice mismatch created by a dislocation and therefore only unit cells where the net Burgers vector is zero can be modelled in PBC. This can be done by creating a structure with multiple dislocations, the simplest of which is a dipole configuration. Another solution is the use of semi-periodic systems, where a dislocation is modelled in a cylindrical cluster type structure, where only the dislocation line direction is treated with PBC and the surface of the resulting column can be treated in various ways. The clear advantage to such an approach is the relatively small number of atoms to be modelled and that there is no interaction of surrounding dislocations with the dislocation under study. However, with a lack of PBC the electronic structure investigation is limited to a set of discrete levels. In addition the treatment of the cluster surface has a significant influence on the electronic properties of the

60

material, in order reduce dangling bonds on the surface a hydrogen termination is commonly implemented[67]. As for structural disadvantages, the surface of such a cluster needs to be frozen during structural relaxation in order to simulate strain effects of the bulk and not of the undesired surface. As a result the long range strain field and with it the exact dislocation line position is artificially fixed. Therefore, in the case of core structure changes for the same cluster, the strain field cannot adapt to the new configuration and it's exact core position, as will be discussed in detail later for the supercell approach.

For this work the supercell approach has been adopted for both structural optimisation and electronic structure calculations. With computational costs in mind it is best to choose a dipole configuration as the unit cell, however there are a number of ways that the PBC can be realised. Orthorhombic boundary conditions would result in an array of dipoles, while the more commonly used PBC result in an array of quadrupoles, see figure 4.2 for details. Historically either choice has been used in a range of studies [68, 69, 70, 71, 72], while the dipole array has been found to produce a mismatch at the boundaries of the cells previously [73]. Therefore the array of quadrupoles condition has been used extensively for all types of dislocations. However, a study by Lehto et al.[67] presents a general solution for the boundary inconsistencies. They conclude that while the array of dipoles represents the minimum energy state for the edge type dislocation, in contrast a quadrupolar PBC for the screw type dislocation appears to be more favourable, due to the difference in the strain fields. Therefore in this work, the PBC producing an array of quadrupoles are used.

The supercell structure used in this work for electronic structure calculations was created by repeating a usual wurtzite primitive unit cell 7 times in the $[2\bar{1}\bar{1}0]$ and 8 times in the $[01\bar{1}0]$ direction, resulting in a 448 atom cell with a single lattice layer in the $[0001]$ direction [1]. Subsequently the displacement field for two screw dislocations (equation (2.98)) with opposing Burgers vectors was applied to the atomic coordinates, with a core separation of about 22Å along the $[01\bar{1}0]$

---

[1] This creates an orthorhombic unit cell of size

$$
\begin{aligned}
7a \times 8 \cdot 2\sin(1/3\pi)a \times c \quad &= \quad 22.4\text{Å} \times 44.34\text{Å} \times 5.2\text{Å} \qquad &&\text{for GaN} \\
&= \quad 24.5\text{Å} \times 48.50\text{Å} \times 5.7\text{Å} \qquad &&\text{for InN,}
\end{aligned}
$$

as the repetition of the wurtzite cell along $[01\bar{1}0]$ incorporates 2 unit cells

direction. The supercell lattice vectors were then taken to be

$$
\begin{aligned}
a_1 &= 7[2\bar{1}\bar{1}0] + 4[01\bar{1}0] + \frac{1}{2}[0001] \\
a_2 &= 8[01\bar{1}0] \\
a_3 &= [0001].
\end{aligned}
\tag{4.1}
$$

This choice of non-orthorhombic lattice vectors creates an array of quadrupoles as depicted in figure 4.2(b). The addition of half a Burgers vector ([0001]) to a_1 serves to compensate for the distortion of the unit cell by the screw displacement field and ensures a lattice match at the opposite site.

4.1.1 Strain effects of screw dislocations in periodic boundary conditions

As outlined above a single dislocation cannot be described within PBC. A dipole configuration however can be used with various PBC, where the array of dipole structure is unfavourable. Therefore an array of quadrupoles is used. A schematic of such a quadrupole condition can be found in figure 4.2(b). These PBC result in an effective array of straight screw dislocations arranged in groups of 4, where direct neighbours have Burgers vectors opposite in sign. Such a situation could create a strong strain influence of neighbouring dislocations on any particular dislocation core. However, in order to model an unperturbed dislocation core, the creation an environment free of external strain is of high importance. Therefore the influence these conditions create on a single dislocation core region was modelled using the analytical elastic field generated by a screw type displacement field (see section 2.4 for details) in a quadrupolar arrangement under PBC.

Figure 4.3 shows several plots for the two non-zero strain components of such a quadrupolar array. The dislocation core positions are clearly marked by the diverging strain field. In addition it shows the strain at a missing dislocation in an array of quadrupoles, in order to gain a quantitative perspective of the amount of strain influence of the dislocation neighbours on a single dislocation core region in the modelled cell. In order to simulate this only one dislocation in the centre array is missing, in an otherwise complete arrangement of quadrupoles. For a single cell under PBC this scenario would not be possible to model, due to

the non zero Burgers vector for the central cell and, as per definition the images
in PBC being identical to the unit cell. However, in order to simulate such a
unique situation the PBC of a normal simulation was represented by repeating
a quadrupole 25 times in a square grid and centring on the site of interest. In
addition typical simulation values have been chosen for this computational repre-
sentation, i.e. a Burgers vector length of about 5Å with a dislocation separation
of about 25Å. Comparing this to usual dislocation core radii of about 6Å, it can
be easily seen that in such a setup the surrounding dislocations have almost no
distortion effect on the small core region. This demonstrates clearly the very
low strain influence of the surrounding dislocations on the core region of each
dislocation. Thus a quadrupolar array with sufficiently large cell sizes provides
an excellent environment to model unperturbed dislocation cores.
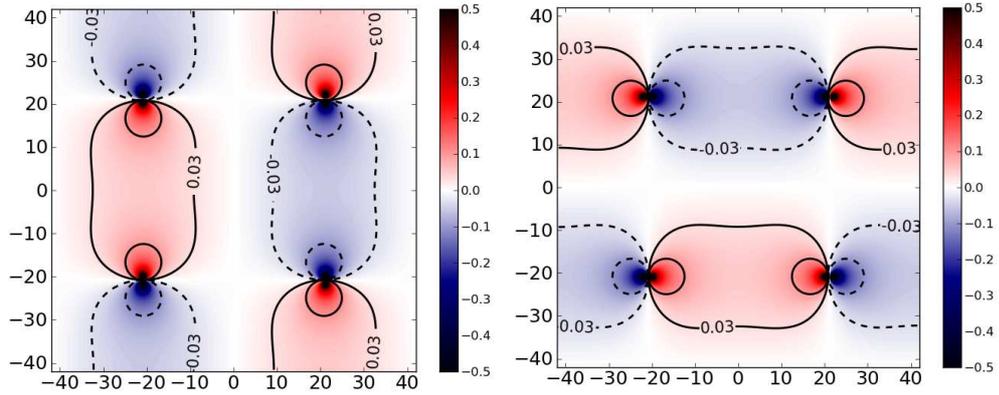
(a) ϵ_{13} strain component of a complete screw dislocation quadrupole in PBC.

(b) ϵ_{23} strain component of a complete screw dislocation quadrupole in PBC.
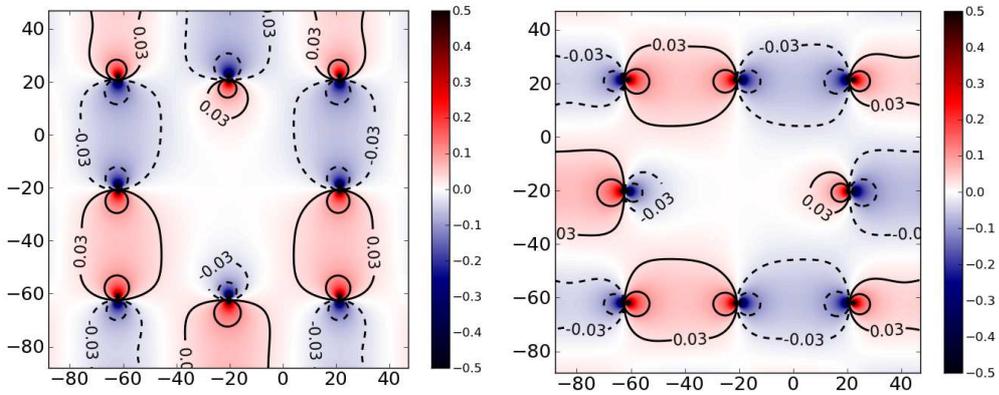
(c) ϵ_{23} strain component of an incomplete screw dislocation quadrupole in PBC.

(d) ϵ_{23} strain component of an incomplete screw dislocation quadrupole in PBC.
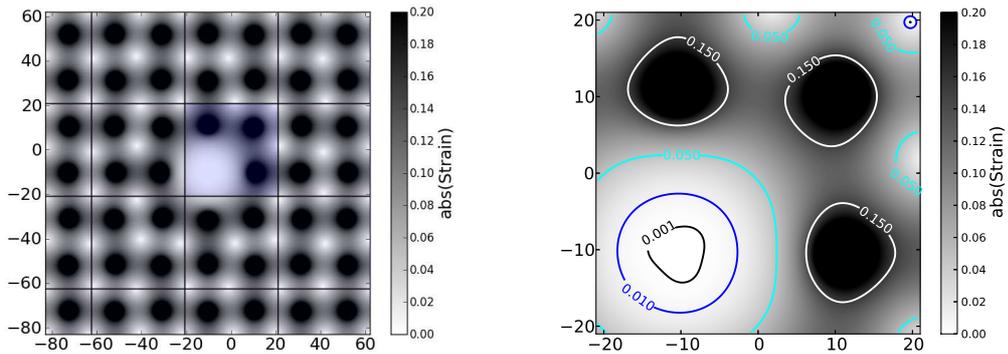
(e) abs(ϵ) of an incomplete screw dislocation quadrupole in PBC.

(f) abs(ϵ) of an incomplete screw dislocation quadrupole in PBC (zoomed).

Figure 4.3:   Depiction of the two strain components of an array of screw dislocation quadrupoles. Figures (a) and (b) centre on a complete unit cell. Figures (c)-(d) show the same section of a repeated quadrupolar cell, where one dislocation has been removed.

4.2 Global optimisation for the stoichiometric dislocation core

Finding the optimal atomic configuration for a dislocation core is a non-linear, non-convex optimisation problem, as the energy function defining the optimal solution as a global minimum is non trivial depending on the underlying approximations made. In general it can be seen that there exist many stable local minima or even meta-stable saddle point configurations for any structural optimisation consisting of a group of atoms. As discussed in section 2.3 such a problem can be solved by employing any one of a wide variety of optimisation techniques, such as particle swarm optimisations, simulated annealing or genetic algorithms.

In order to employ a systematic approach, the global optimisation through randomisation scheme was adopted to investigate possible core structures for the screw dislocation. Details for the optimisation algorithm can be found in section 2.3. The application of which on the fully stoichiometric screw dislocation core will be discussed in this subsection. As detailed above a quadrupolar array of dislocations in PBC was used, however for each configuration only a single core of the structure has been optimised. Specifically all atomic positions within a cylindrical volume around the dislocation centre with a radius of 3.2Å have been randomised and relaxed repeatedly. In the chosen structure that was described in the previous section for GaN this volume encloses twelve atoms, six nitrogen atoms and six group-III atoms.

The random nature of the configuration space sampling of the optimisation procedure results in an exponential increase of sampling tries in order to reach a converged result as the number of optimised atoms increases. This means that when all 12 atoms are included only after up to 10^5 tries will a core structure be fully investigated and all lowest energy structures found with sufficient certainty. Therefore using a DFT approach for the structural optimisation is far out of reach for currently available computational resources. As a result a SW type interatomic potential method has been used for the structural relaxation during the global optimisation. As is discussed in section 2.2 a parameterisation for only GaN was used because of the sensitivity of the approach to the quality of the relaxation and the excellent performance of the parameters.

The typical dislocation core radius of a screw dislocation in [0001] direction in GaN is around 6Å as discussed in section 3.3, which means that the volume of the dislocation core is much greater than the optimised cylindrical volume with a radius of only 3.2Å. Therefore it could be argued that the size of the randomisation volume does not suffice. However, by definition the dislocation core radius is only the point at which classic linear elasticity theory fails and a more basic atomistic theory is to be applied. It does not mean that the material changes completely from a bulk-like tetrahedral bond arrangement exactly at the core radius. In fact I am unaware of any claim in the literature that this would be the case. In addition the structural optimisation, unlike the randomisation, is not restricted to the smaller centre volume and distortions are allowed to occur during relaxation which could result in alien core configuration larger than the randomisation volume. However, this does mean that such cores are less likely to be found as they fall into only an edge of the configuration space sampling. With a sufficiently large quantity of structural optimisation tries this is not a problem.

The randomisation and relaxation procedure has been fully automated using a combination of Python [74] scripts and the LAMMPS code [37]. A Python controller program has been developed that sets up local work folders to start a variable amount of relaxation threads. For each thread a template quadrupolar dislocation array structure is taken and in the above mentioned volume around one of the dislocation centres all atomic positions are randomised.[2] The randomised version of the template is then placed in a thread specific working folder and a LAMMPS conjugate gradient structural relaxation is started. After each relaxation the procedure is repeated until convergence or an upper limit of the step amount is reached. The large quantity of randomisation tries and the specific mult-threaded implementation results in a nearly perfectly parallel scalability of the global optimisation. Convergence of the result is reached when the unique structure count for the lowest energy structure approaches a constant fraction of the total amount of structures. Distinct structures are identified by sorting the total energies into bins with a width of 10^{-6}eV, which is chosen to be sufficiently larger than the structural relaxation tolerance, but small enough to uniquely dis-

---

[2]   An additional constraint has been added to prevent atoms being initially positioned within 0.2Å of each other before relaxation.

tinguish structures.

As the initial and final structure, as well as the total relaxed energy, needs to be stored for analysis for each relaxation step a large quantity of data is being accumulated. Therefore the controller program parses each completed LAMMPS output and stores only the relevant data in a binary datafile. A number of further Python scripts were created for accessing the data files and subsequent analysis of the structures.

4.2.1 Energy distribution

The total energy of all relaxed structures can be displayed as an energy distribution. Convergence of the energy distribution, showing a high count of stable dislocation core structures, has been found to occur after $\approx 5 \times 10^4$ randomisation and relaxation steps. Figure 4.4 shows such a distribution after 9×10^4 relaxation steps. The structural energies have been displayed after binning all energies into narrow 10^{-6} eV bins, showing the amount of times an individual structure has been found during the global optimisation, as that bin width is only marginally larger than the selected accuracy of each local SW optimisation. However, after applying a gaussian broadening to the obtained energies (see figure 4.4(c)) it is obvious that the number of structures per energy resemble a statistical distribution[3] with a mean of -3801 eV and a variance of 18 eV. A clear indication of the high probability of finding low energy structures can be seen in the divergence from the normal distribution of the tail, showing a high count of the lower energy structures.

In order to visualise individual core reconstructions a differential displacement map (DDM) was created. A DDM displays atom-sites (2 distinct sites in wurtzite GaN, distinguishable by colour) as circles and the displacement of a neighbouring site in a pre-defined direction as arrows, where size implies magnitude and scaled to bond-length. As a result a DDM is an excellent method to display a specific strain component in combination with the atomic configuration and its deformation orthogonal to the viewed plane and hence presents a more atomistic

---

[3]   The structure distribution is close to a normal distribution and a fitting procedure has been attempted, although a gamma distribution is expected to be more accurate as this would account for the existence of a lowest energy structure behaviour.
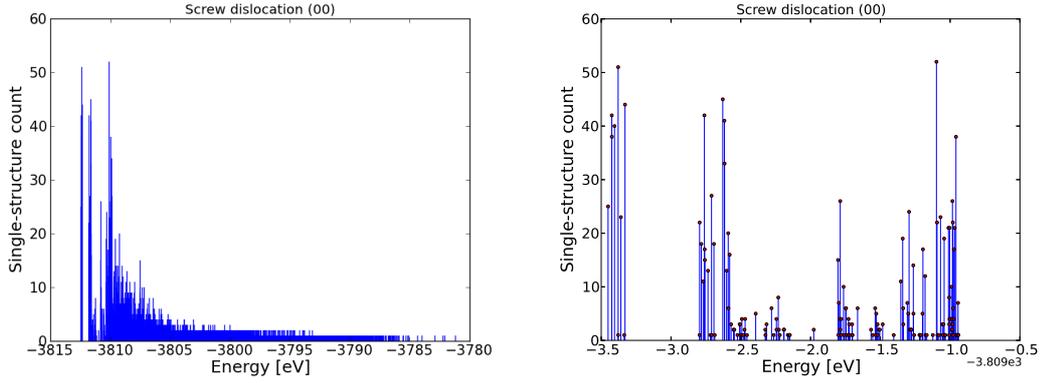
(a) All total energies binned into narrow $\sim 10^{-6} eV$ bins. The bin height represents the number of times an individual structure has been found during the random search.

(b) Lowest energy structures from (a)
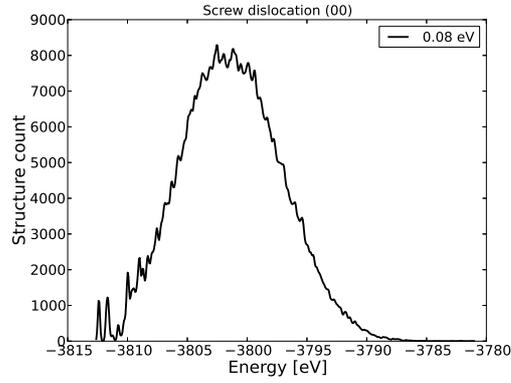
(c) All total energies from (a) displayed with a 0.08 eV Gaussian broadening

Figure 4.4: Energy distribution of the total energy of the fully stoichiometric core after 10^5 randomisation/relaxation steps shown with individual structure counts and gaussian broadened.

representation of the continuum strain model. Figure 4.5 shows DDMs for the energetically lowest two unique structures in the energy distribution in a c-plane view for displacements in [0001] direction. The dislocation centre is marked as a diamond symbol.

The lowest energy structure found by this optimisation procedure is identical to a structure obtained by applying the screw dislocation displacement field with a dislocation core centre at the bond site, therefore confirming similar results by Béré et al. [29]. This structure is not displayed here for brevity but can be found

as a ball and stick model in appendix C. In addition a DDM of this lowest energy structure is depicted in figure 4.5(a). However the same structure can be found in multiple variations in the structure energy distribution as there are several equivalent bond sites to centre the dislocations that will create a slight separation in energy due to change in position relative to the strain field of the other non-optimised dislocations of the quadrupole. This convolution of the energy distribution with equivalent structures at different positions is analysed quantitatively in section 4.2.3.

The next lowest energy structure is a new location for the dislocation line which is centred near to an atom site, but in between two bond sites (see figure 4.5(b) for details). However, in addition to the difference in dislocation line position to the most stable structure, atoms positioned at the site closest to the dislocation line direction also show displacements in the c plane. This structure is separated from the lowest energy case by only about 0.7eV and exhibits a similar convolution in the energy distribution ranging around 0.1eV.

A channel centred dislocation line position has been found as well at 0.8eV above the lowest energy structure. This core reconstruction was found with a very small structure count, suggesting that the SW potentials method predicts this configuration to be meta stable.

4.2.2 Influence of the strain field and locating dislocation cores

As mentioned briefly above, the strain field of the quadrupolar dislocation array can influence the total energy of a structure depending on the dislocation centre position of the optimised dislocation core. In order to quantitatively investigate this behaviour the strain at each atom position $\mathbf{r}_J$ has been calculated. As strain is the spatial deviation from the bulk (see section 2.4), this is done by comparing the tetrahedron formed by the four first neighbours to each atom to the expected tetrahedron in an unstrained bulk material. The deviation is evident in the vectors describing the relative position of the four surrounding atoms $\mathbf{u}_{Jn} = \mathbf{r}_{N_n} - \mathbf{r}_J$, where $\mathbf{r}_{N_n}$ is a nearest neighbour position to $\mathbf{r}_J$ and $n = 1...4$. From the definition of strain in equation (2.68) it is clear that the relationship between

69

(a) DDM for the lowest energy struc-
ture

(b) DDM for the second lowest unique
structure
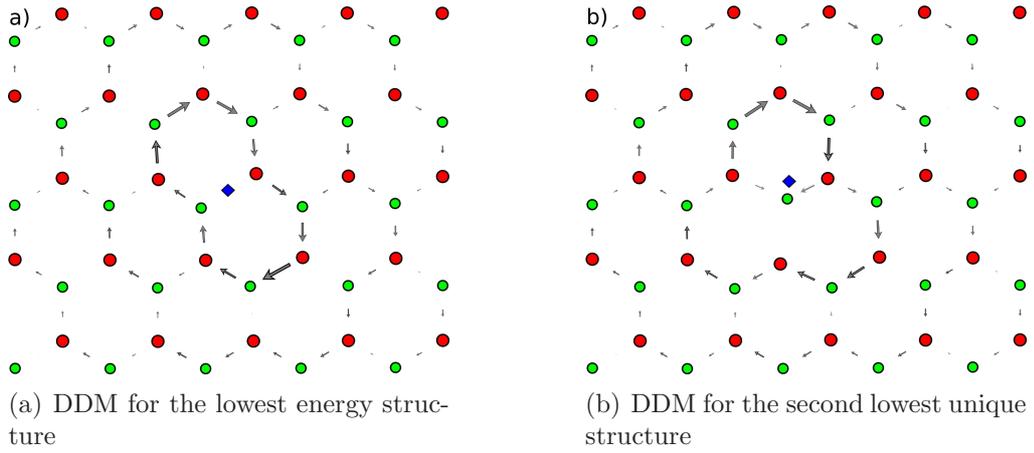
Figure 4.5: Differential displacement map (DDM) for the lowest energy (a) and
second lowest (b) unique structures obtained for the fully stoichiometric core.
The blue diamond symbol marks the dislocation line position as obtained by
computational to analytical strain fitting.

strained $\mathbf{u}^s_{Jn}$ and unstrained $\mathbf{u}_{Jn}$ distances for every n at position $\mathbf{r}_J$ is

$$\mathbf{m}_J\,\mathbf{u}_{Jn} = \mathbf{u}^s_{Jn}, \qquad \text{with} \qquad \mathbf{m}_J = \mathbf{I} + \boldsymbol{\epsilon}_J, \tag{4.2}$$

where $\mathbf{I}$ is the identity matrix. The matrix $\mathbf{m}_J$ incorporates the strain tensor
$\boldsymbol{\epsilon}_J$ at position $\mathbf{r}_J$ in this definition. In addition, an implicit approximation was
introduced with the assumption that the strain is constant for all distance vectors
$\mathbf{u}_{Jn}$ of a single tetrahedron (i.e. $\boldsymbol{\epsilon}_J$ is independent of n). However this case
approaches the limit of the continuum definition of the strain, since there must
be a physically distorted geometric shape that can be compared to the norm
in order to find strain and the tetrahedron formed by the next nearest bonding
atoms represent the smallest possible shape. As a consequence, the strain within
a single tetrahedron at centre-position $\mathbf{r}_J$ can be determined using a least squares
minimisation of the euclidean norm over all n:

$$\text{minimize}\left\{\sum_n ||\mathbf{u}^s_{Jn} - \mathbf{m}_J\,\mathbf{u}_{Jn}||^2\right\}. \tag{4.3}$$

Only for strongly distorted centre atoms of a given dislocation core can this al-
gorithm not detect the strain, as there is no atomic configuration present that is
comparable to the bulk. Within the lowest energy structure in the given distri-

bution there are 2 atoms exhibiting this behaviour.

The computed strain at each atomic position can then be displayed by inter-polating the irregular data to a regular grid and averaging along one axis to give a two dimensional projection. Figure 4.6 shows the computed strains for the low-est energy structure of the performed and above discussed global optimisation. Core regions and the optimised core are marked. It is evident that strain within the optimised core is identical to its counterpart in the diagonally opposite posi-tion within the supercell.

(a) computational strain ϵ_{12} (b) computational strain ϵ_{23}
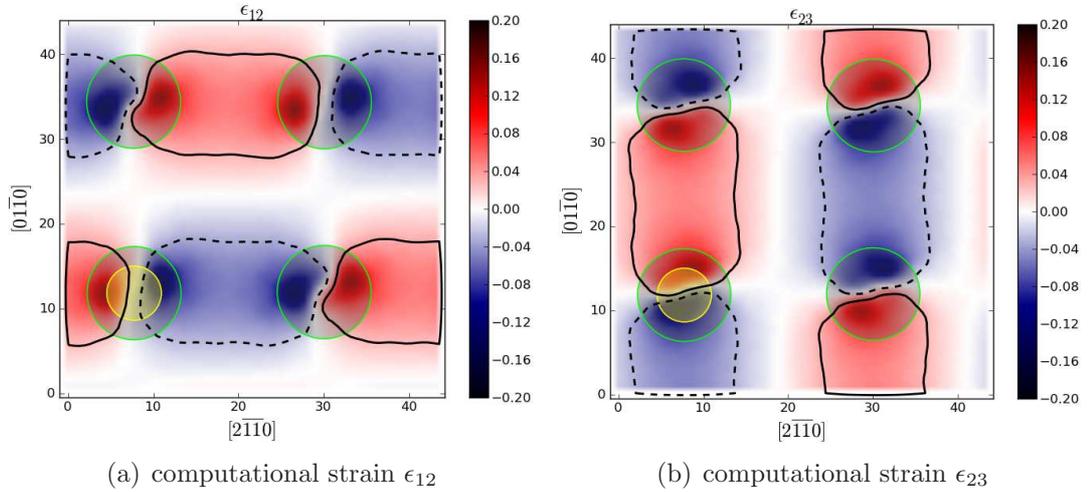
Figure 4.6: Computational strain for the lowest energy structure with a fully sto-ichiometric core. Green circles indicate core regions, while the yellow circle marks the optimised core. The strain has been obtained by analysing local deviations of the crystal structure from the bulk.

In order to determine the position of the dislocation line centre after optimisa-tion, the computed strain can be compared to the analytical strain solution from linear elasticity theory for a quadrupolar array of dislocations in PBC. This is done by a least squares optimisation of the dislocation centre position and Burg-ers vector, as these are the only free parameters in the linear elasticity model of a dislocation. By definition the strain within the core region should not be used in this parameter optimisation, therefore only strains outside a fixed cylindrical vol-ume of the expected dislocation line position, using a generous radius of $\sim 10\text{Å}$, were included. The Burgers vector was optimised as a test as it should always

be equal to the lattice vector of the total structure in line direction (c direction in this case). For all parameter optimisations performed the obtained Burgers vector was equal to the expected value to within $< 10^{-1}$Å.

4.2.3   Strain effects on the energy distribution

A specific dislocation core reconstruction will not only be found in a single energy bin, due to a convolution of the energy distribution with the breaking of the lattice translational symmetry by the induced strain field of the quadrupole. This is primarily a result of the randomisation in the search method, as the subsequent relaxation process does not guarantee to produce every example of a particular dislocation core at the same lattice position. In fact, the core position will vary up to one full lattice repeat in the c plane. Figure 4.7 is a graphical representation of the spatial arrangement across the strain broadened bins for two unique structures. It was created by determining the dislocation centre positions for the 500 most stable structures of the energy distribution. The dislocation line position is given as the deviation from its location in the quadrupole for the optimised core. It is important to note that within a single energy bin, two separate locations can be found provided the sites are related by a mirror symmetry around the zero point. The size of each point in the figure represents the structure count at that position and hence the likelihood of finding a dislocation line there.

Given the energy convolution, it is still possible to quickly identify new higher energy core reconstructions as the total energy only changes by around 30 meV/Å due to changes in location with respect to the quadrupolar strain. The energy difference between unique core structures is usually greater, specifically in the presented case for the two lowest unique structure the separation amounts to 0.7 eV. However, there is no reason why this should be the case for all structures; and indeed a general trend can be observed in the energy distribution of higher energy/less stable structures being closer together in energy. This is apparent in the similarity of the energy distribution to a Gaussian normal distribution, where closer to the mean more structures will be found with smaller energy separation.

In order to produce a clearer energy distribution a deconvolution with respect to dislocation line position can be attempted. With the ability to determine the geometric deviation from a quadrupole for each structure the resulting change for

(a) Lowest energy structure   (b) Second lowest energy structure
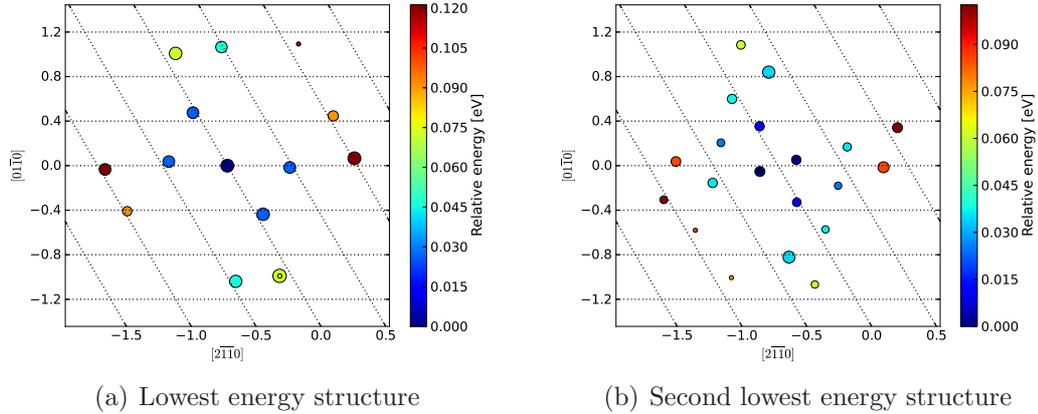
Figure 4.7: Spatial arrangement of dislocation core centres of a unique core with respect to its most stable position in the quadrupole strain field. The size of each point represents the structure count and hence the likelihood of finding the centre at the particular position. Distances are given in units of the a lattice vector. Both structures have been presented in section 4.2.1 including their DDMs.

the analytic elastic strain energy can be determined and used to adjust the total energy of every structure in the energy distribution individually. However, such a deconvolution has not been attempted as there is no information to be gained from such a process, since only a few distinct low energy structures are to be considered actual candidates with likely real crystal application and those are easily extracted from the data with the tools at hand. In addition, the inherent error introduced in the dislocation line position calculation would result in a larger uncertainty in the corrected energy for a particular structure, hence the energy bins would have to be larger in order to sort structures correctly. With larger energy bins it is also more likely to sort two energetically degenerate structures into the same bin, making a thorough investigation more difficult. The likelihood of exactly finding energetically degenerate distinct structures seems negligible. However, to my knowledge, there is no way of estimating the likelihood of finding such a case.

4.2.4 Dislocation core radii and energies

The classification and analysis of dislocation cores in the literature usually focusses on the dislocation core radius and energy (see definitions in section 2.5.1). The dislocation core radius can be found by noting the point of divergence of

the cylindrically integrated elastic energy density from the constant slope in a logarithmic plot [75][28][76]. With the knowledge of the core radius, the core energy is easily determined in an atomistic theory by integrating the atom resolved energies within the now defined boundaries of the dislocation core (see equation 2.109). However within this definition of the dislocation core energy, a quantitative comparison between two separate core structures is impossible as the surrounding strain energy is excluded at a different point, due to the difference in core radius. It is therefore prudent to compare cores with a core energy defined through a generous upper limit core radius that is kept equal for all structures. In fact within this work, core structures are compared only based on their total energy difference to the lowest energy fully stoichiometric structure, which is the same approach as assuming a constant radius, up to a constant of integration. A complication within this approach is however the strain induced energy differences for cores at different positions. This effect is being neglected due to the relatively low impact (less than 0.1eV).

The search for the dislocation core radius is further complicated by the broken radial symmetry of the long range elastic fields due to the quadrupolar arrangement of dislocations. Figure 4.8 displays the energy density for the used quadrupolar array in PBC computed for the analytical screw dislocation strain field, and determined directly from the relaxed atomic positions. As can be seen the energy density does not drop off with radial symmetry as it would for a single dislocation. Therefore integrating over a cylindrical volume will not show a deviation from the constant logarithmic slope as neatly. As a consequence the dislocation core radius needs to be determined from the difference between the analytical and computational energy densities. As can be seen in Figure 4.8, the divergence starts being significant at about 6.9Å.

In order to find the dislocation core radius under the present conditions it is essential to locate the dislocation core position (using the method outlined in section 4.2.2). This provides the possibility to calculate the analytical strain field of any individual geometry, which can then be used to compare it to the computed strain at each atom position and note the point of divergence, i.e. the core radius. As the computational strain can only be determined at each atomic position it forms a discrete irregular three dimensional grid. This irregular grid

74

(a) Analytical

(b) Computational

(c) Difference around a single core (analytical-computational)

(d) Analytical and computational strain at each atom site from dislocation core centre

(e) Averaged difference of the analytical and computational strain density as a function of distance from dislocation core centre
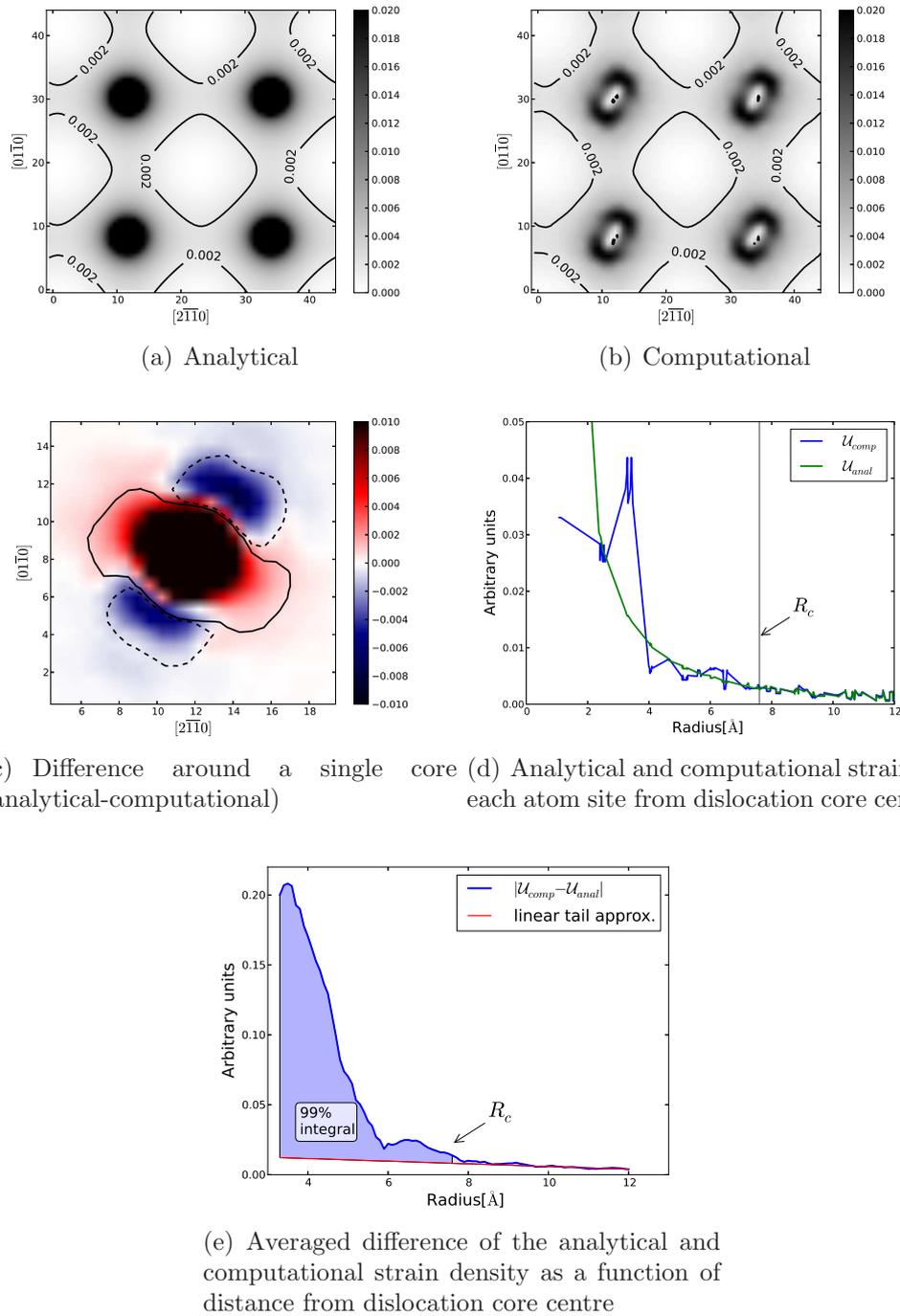
Figure 4.8: Difference of the analytical and computational strain energy density of the dislocation quadrupole in PBC, computed with the analytical screw strain/stress fields and determined directly from the relaxed atomic positions.

can be used directly to determine the core radius, or it can be interpolated and projected to the c-plane, removing noise and creating a continual data set. The interpolation method proved to be less influenced by noise and hence was used to determine the radius. With this continual data set the absolute value of the difference of the analytical and computational strain energy density can be averaged along the angular coordinate of a radial coordinate system centred on the dislocation line position. This creates a radial measure of the discrepancy between the analytical and computational strain. Figure 4.8(e) shows the strain energy difference determined by this approach. Furthermore, this difference does not only show the strain mismatch in the dislocation core, but all other influences such as slight inaccuracies in the determined dislocation core position. A mismatch in a single core position would result in an added strain energy density difference that extends beyond the dislocation core boundary and follows a $1/r^2$ behaviour. Therefore such an influence can be removed by analysing the behaviour of the strain difference in the outer region. However, only the linear part can be determined reliably from the tail, as the error in a quadratic coefficient can be too large and more than one dislocation core position will deviate slightly. Subsequently a radial integration of the corrected strain difference was performed between $r = 3.2$ and 12Å. The dislocation core radius was then determined to be the point where 99% of the integrated difference is bounded (see figure 4.8(e)).

4.3 Discussion

4.3.1 Core optimisation

In the example presented here, 87444 randomisation and local optimisation steps were performed in order to obtain a converged energy distribution. The resulting structures were sorted into 81248 very narrow energy bins, which means that most bins only contain a single structure. This implies a non-converged result, however looking at the energy/structure count relationship of the energy bins in figure 4.9, it is evident that all low energy bins feature a very high structure count of up to 52 in a single bin. It is important

Figure 4.9: Energy bin size over total structure energy for all energy bins created during the optimisation of the fully stoichiometric core.

to note that the appearance of single structure bins at the lowest energy range in figure 4.9 represent the same unique structures as the surrounding high counting bins, but located at less likely positions away from the quadrupole equilibrium[4]. Therefore, the convolution of the energy distribution due to the strain effects on the dislocation line position *does* influence the measure of when an optimisation is converged. Unfortunately it is difficult to monitor this automatically during the optimisation procedure, however due to the simplicity of the approach the optimisation can be resumed at any point until a global minimum has been found and verified.

With the help of a range of automated scripts structural properties of the optimisation region, such as dislocation core position and radius, could be determined for the large amount of found stable, low energy structures. As this enables an analysis and identification of the relevant part of the large quantity of structural data, and since the results (see figures 4.5(a) and C.5(d) for a DDM and structure) confirm previous findings about structural properties of probable best cores[75],

---

[4] See section 4.2.3 for details. The only exemption found for low energies is the channel centred dislocation core at a relative total energy of 0.8eV above the minimum

the optimisation procedure can be seen as successful. Therefore, applying this technique on non-stoichiometric cores can provide a systematic approach to solving the question of the nature of dislocation core structure in any growth situation.

The lowest 500 energy bins in the optimisation of the full core, representing approximately 40 unique stable structures, were investigated for core position and radius. No clear dependence of the core radius to the relative dislocation core energy could be determined. In fact, the core radii consistently vary between 7 and 8.5Å without a clear dependence on core stability. In addition it is to be expected that higher energy local minima in the optimisation do not represent realistic core structures, but artefacts of inadequate approximations within the SW inter-atomic potentials approach. A theory incorporating electronic structure that accounts for non-tetrahedral bonding situations needs to be applied for more realistic structures. However, as lowest energy cores are seemingly represented and predicted well within SW and only these represent relevant structures, a more in-depth investigation into higher energy structures would be purely academic.

4.3.2   Optimisation performance

The random brute force sampling optimisation approach used in this work and described in detail above, was chosen for its simplicity and the fact that it provides a measure of certainty of finding the actual global minimum. The performance, as measured in the amount of randomisation and relaxation steps required to be sure that the minimum has been found, can only be described as slow. However, with the use of SW-type inter-atomic potentials the computational cost becomes manageable on a current workstation machine for the presented example of a fully stoichiometric dislocation core, where 12 centre atoms are included in the optimisation.

Clear limits of the random search method include the exponential behaviour of the computational cost to the dimensionality of the feasible solution space. Due to the nature of the complete sampling of the solution space this property cannot be changed, making larger randomisation areas impossible and hence prohibiting the investigation into dislocation core structures with more than one lattice repeat in line direction. The only performance optimisations to the method that can be done involve the increase of the amount and strength of the search constraints.

This effectively reduces the size of the feasible solution space and could greatly reduce the number randomisation/relaxation steps needed to reach convergence for any system size. However, this step could mean that the global minimum is being excluded from the search space and hence care needs to be taken. In this work no constraints other than a restriction of a minimum nearest neighbour distance have been included. However restrictions of bond angle arrangements or a more homogeneous space allocation in the random arrangement could be considered, but were not implemented as a stable structure for a dislocation core is not expected to completely follow the crystal symmetry or specific bond arrangements. A study of the influence of additional constraints can be found in [44] for various structural compositions and situations.

A possible improvement to the randomisation procedure that can be implemented without risking a biased result is the reduction in equivalent structures due to permutations. As the randomisation region for the example of a fully stoichiometric core includes 12 atoms of two types, there exist 924 [5] equivalent configurations for every random structure. A function could be implemented to dismiss new random points if they are close to any previous try, including all permutations. However, such a step would not increase the likelihood of finding a global or more stable solution to one higher in energy, nor would it decrease the number of local minima to be searched, it merely would decrease the total sampling space. However adding such a distance measure in the configuration space and excluding close sampling points could deny repeated sampling of small local minima energy basins. In combination with the previously established premise that a trend exists that the lower the local minimum the larger the energy basin/sampling chance, this could increase the optimisation performance without imposing a structural bias. On the other hand this thought could be extended into diverging from a random search to a structured, rasterised sampling of the solution space, with a raster separation defined through the distance measure in configuration space.

---

[5] Permutation of N sites occupied by two types is:

$$\frac{N!}{N_{Ga}!N_N!} = \frac{12!}{6!6!} = 924$$

CHAPTER 5

Non-stoichiometric screw dislocation cores

This chapter incorporates a discussion of the modelling of screw dislocations in InN and GaN. As in chapter 4 the results presented are produced by implementing a global search algorithm for the screw dislocation core. However, here a wide selection of non-stoichiometric conditions ranging from nitrogen rich to group-III rich as well as open core structures are included in the search space. The core configuration search is performed using an empirical interatomic potentials method (a Stillinger Weber (SW) type potential) for GaN only. A selection of the most favourable optimised core structures and their electronic properties are then investigated for both GaN and InN with DFT within LDA.

The first section in this chapter gives a comparative study of the structural properties of all of the studied core compositions that were optimised using the approach introduced in chapter 4. A discussion of the electronic properties of the most likely core configurations and compositions is given in the last section of this chapter. A more complete presentation of the data, including cases omitted for clarity from this chapter, can be found in appendix B and C.

5.1 Structural optimisation

In chapter 4 it has been established that the applied global optimisation technique shows that the proposed energetically most favourable structures in the literature [75] for the fully stoichiometric cores are indeed the lowest energy structures. This rather unspectacular result emphasises that the proposed method is well suited to the system under examination, as structurally unrealistic dislocation cores were all found to be higher in the energy distribution, and the relative energy separation is large enough to distinguish structures reasonably well. However, the strength of the global optimisation method lies in straightforward application to more complex non-stoichiometric core reconstructions, which is done by simply placing varying amounts of atoms inside the randomisation volume for each configuration during the optimisation.

5.1.1 Optimisation region

As part of this work, a wide range of stoichiometries were considered and optimised using the same procedure as described in Chapter 4 for the fully stoichiometric core configuration. For brevity the resulting total energy distributions are not discussed or shown here, however details can be found in appendix B where table B.1 lists a summary of all optimisations performed and figures B.1-B.8 show a graphical representation of all obtained energy distributions.

In order to identify core configurations, a labelling convention has been adopted that describes a core composition by the number of vacancies it contains. The fully stoichiometric core is then labelled as (00) which is itself an abbreviation for $(0V_{Ga}, 0V_N)$. Similarly (-11) represents a structure containing an additional gallium atom and a nitrogen vacancy [1]. An additional bond angle convention has been adopted, where the type of angle is given by three atomic type indices. The

---

[1] A similar labelling scheme has been introduced by Northrup et al. [77], where a core region also contains 6 gallium atoms and 6 nitrogen atoms. The labelling is then done by specifying the amount of atoms by type in the core region, i.e. the (66) core there corresponds to the (00) core label used here and the (65) core corresponds to the (01) core here:

$$(n\,m)_{\text{Northrup}} = (\{6-n\}\{6-m\})_{\text{here}} \qquad (5.1)$$

A different naming scheme was chosen for this work in order not to imply a constant core region size. As a result the (00) core is always the fully stoichiometric core, independent of randomisation region size

81

acute angle is given between the latter atom types, with first index representing the atomic type at the vertex.

Comparing energies between structures of different stoichiometries is achieved by determining the chemical potentials for Ga and N in unstrained bulk GaN. With the chemical potential the thermodynamic Grand Potential Ω can be constructed, if the entropy is neglected, due to the relatively low temperatures in solids. The Grand Potential can be used to easily compare structural energies as it incorporates the difference of atoms in different structures with the inclusion of the chemical potential. In the case of SW interatomic potentials, the chemical potential is found by taking atom resolved energies. A reference point for the energy has been chosen as the lowest energy structure for the fully stoichiometric core configuration. It is important to note that this type of comparison of energies is nearly identical (up to a constant) to using a dislocation core energy for a fixed (independent of core configuration and composition) dislocation core radius. Figure 5.2 shows a plot of these relative energies for the lowest energy structures of all core configuration considered.

As a consequence of the limitation of the randomisation volume, it is only possible to optimise core compositions within a certain range. This range is bounded by the (66) fully open core, $(-x6)$ Ga filled cores and the $(6-x)$ N filled cores. Where a choice of restricting x to a maximum of 6 has been chosen. However, this composition area still incorporates 91 different core compositions, without even considering adding additional atoms to the core, i.e. $nV_{Ga} + mV_N < 0$ with $n, m \in \mathbb{N}$. Adding compositions that follow $nV_{Ga} + mV_N = -1$ would increase the number of possible core optimisations to 103. As a consequence of
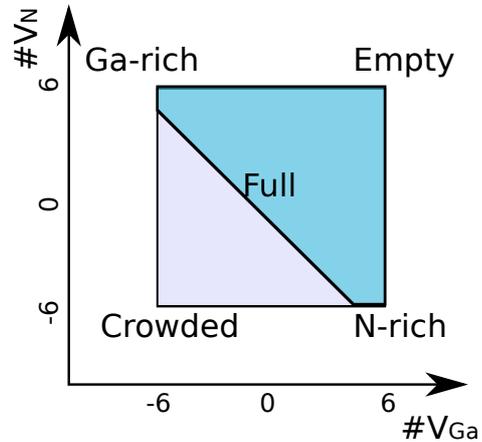
Figure 5.1: Range of optimised core stoichiometries. The darker region marks the range of stoichiometries considered for optimsiation.

computational expenses and time only about a quarter (24) of the possible compositions were optimised.[2] However, emerging trends in the relative energies from the sampled set can be used to dismiss or include cores from the remaining set.

5.1.2 Core structure analysis

Figure 5.2(b) displays the relative energies of all stoichiometric cores from the full core (00) to a completely open core containing no randomised atoms (66). Interestingly the lowest energy is not the full core but the (11) core with an energy difference of -0.27eV, making this particular configuration the most favourable dislocation core. From this point the relative energy increases linearly with decreasing GaN pairs. Figures 5.2(c) to 5.2(e) display core compositions each with a constant total number of atoms contained, but with varying gallium to nitrogen ratios. These dependencies show a more quadratic behaviour in the relative energy with the minimum being close to the stoichiometric centre. However, a clear trend favouring N rich conditions can be observed for all diagonals, by noting that the projected minimum of a square energy dependence would be away from the stoichiometric middle into the nitrogen rich side. This results in mirror configurations such as (01) and (10) always favouring the composition with more nitrogen atoms contained.

The relative energy trends shown in figure 5.2 can be explained by examining the structural properties of the optimised cores. Table 5.1 lists these properties for the optimum structures of all stoichiometries, determined by a script running through all atoms within a cylindrical volume of 10Å radius around the dislocation line. The script determines distances and angles to nearest neighbours for all atoms within the PBC. Restrictions to neighbour distances were determined according to the type dependent binding distances of the SW parameters (see figure 2.1), with cutoff distances of 2Å for N-N, 2.7Å for Ga-Ga and 2.5Å for Ga-N bonds.

With this statistical data, the trends in figure 5.2(b) can be analysed more carefully. The SW potential favours the (11) core as, unlike the (00) core, a geometry

---

[2] The computational time for optimising the complete core region can be estimated to be $\approx$20 cpu years.
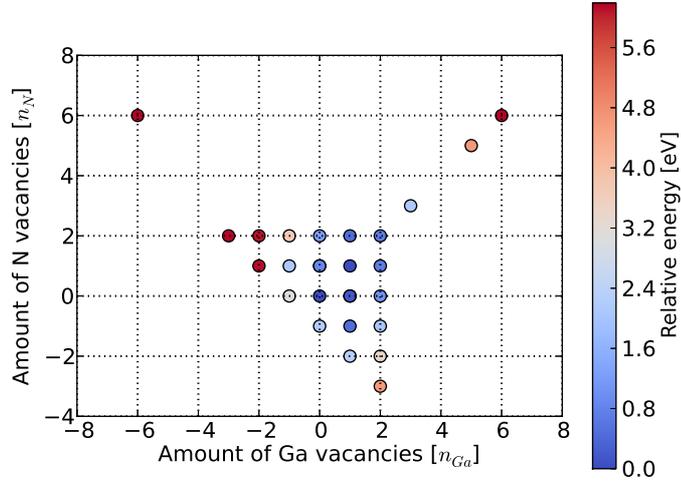
(a) 2D representation of all investigated structures.
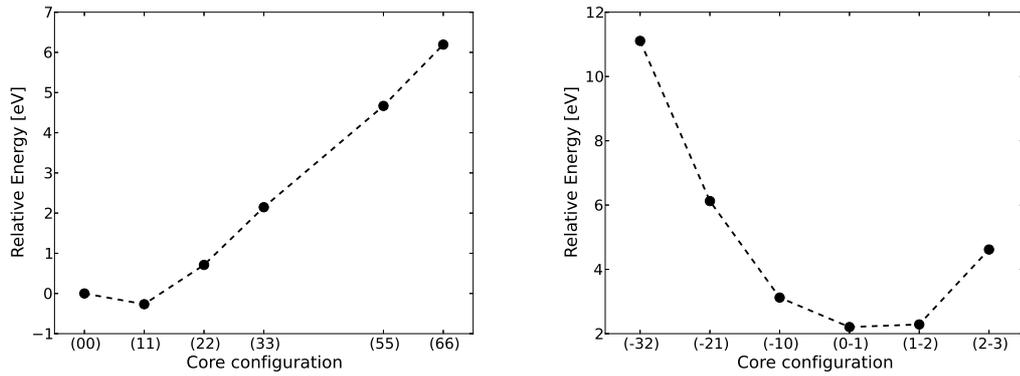
(b) Major diagonal from (a); $nV_{Ga} - mV_N = 0$

(c) Diagonal from (a); $nV_{Ga} + mV_N = -1$
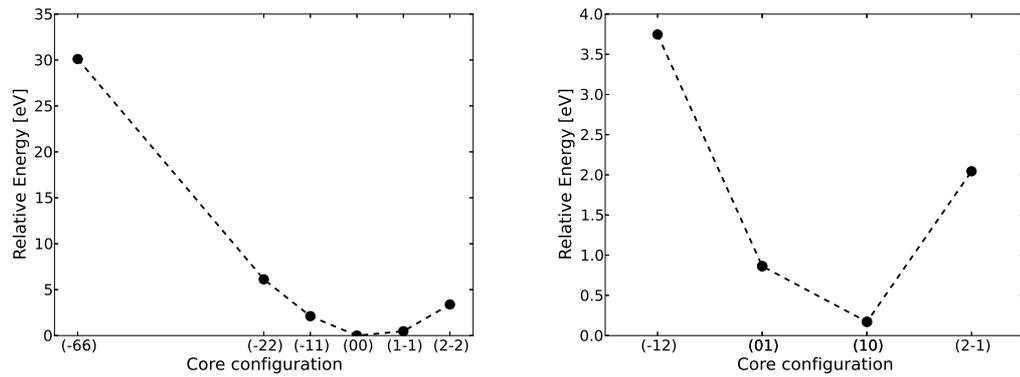
(d) Diagonal from (a); $nV_{Ga} + mV_N = 0$

(e) Diagonal from (a); $nV_{Ga} + mV_N = 1$

Figure 5.2: Comparison of total energies of all investigated core configurations. Chemical potentials for Ga and N were taken from bulk. Relative energies were computed by compensating the difference in atomic composition with the chemical potentials of Ga and N and are shown relative to the (00) total energy.

84

is possible where no homonuclear bonds exist[3], instead 2 atoms are surrounded by 5 nearest neighbours. However, as the angular configuration of the N-Ga-Ga and Ga-N-N feature a wider deviation in bonding angles than the (00) core, meaning bonding angles deviate more strongly from the tetrahedral ideal (109.49°), making the energetic lead of the (11) core relatively low at -0.27Å. As the core composition becomes emptier with the (22) core it can be constructed without homonuclear bonds but as a consequence includes atoms with only three neighbours. The (33) core then features the first fully "open core" as there are no more atoms around the centre, increasing the amount of atoms with 3 neighbours.

Looking at the relative energies across compositional changes (figures 5.2(c)-5.2(e)), nitrogen rich cores are being favoured as the N-N interaction is lower in energy than the Ga-Ga interaction in the SW pair parameters. With this in mind it could be argued that since the perfect bonding angle Θ^0_{IJK} within the parameters is constant for all interaction types, the most stable core configuration is symmetric under atomic type exchange. However, this is not the case. The difference in energy and bond length of the N-N and Ga-Ga bonds means that the most stable cores can be unique in structure and vacancies or interstitials will form at different places.

The fact that SW potentials favour nitrogen rich conditions is however not necessarily physical and should be considered with caution. This could be due to the simple case of comparing energies with the chemical potentials at a too extreme nitrogen rich case, as a result of inaccuracies in the SW parameterisation. However, it can also be argued that the SW potentials method itself is insufficient by looking at the extremes of the present situation. It shows that a gallium filled core (-66) is highly unfavourable while a nitrogen rich core (2-2) is much less unstable. This can be explained as due to the insensitivity of the SW potential towards bonding environments beyond immediate neighbours. In a N rich core more N-N homonuclear bonds will exist, in addition these same atoms will still contribute towards all N-Ga-N and N-Ga-Ga bonds without restriction in the SW potentials method. This would not necessarily be true for a real situation, therefore a more rigorous ab-initio method should be used to confirm the SW results predictions.

---

[3] Ga-N bonding has the lowest energy in the used SW parameters, with N-N the second lowest and Ga-Ga the highest. See figure 2.1 for details

Table 5.1: List of properties of investigated core configurations across a wide compositional range. Energies are given in eV, core radii R_c and distances in Å, angles in degrees. Columns 4 and 5 show a count of atoms with 5 and 3 nearest neighbours in the structure, columns 6 through 9 show a count of homonuclear bonds and average bond distances d, while columns 10 through 21 show bond angle counts, averages and variance measure.

			nn		Ga-Ga		N-N		Ga-Ga-N			N-N-Ga			N-Ga-Ga			Ga-N-N		
Core	rel. E	R_c	5	3	#	d	#	d	#	θ	var	#	θ	var	#	θ	var	#	θ	var
(00)	0.00	6.5	0	0	1	2.19	1	1.62	6	103.1	15.0	6	113.2	11.0	138	109.1	8.8	138	109.6	8.6
(01)	0.86	6.6	0	2	2	2.26	0	0.00	10	107.6	15.4	0	0.0	0.0	141	109.4	8.3	136	109.5	7.9
(02)	1.25	7.6	0	2	3	2.30	0	0.00	14	108.2	16.5	0	0.0	0.0	138	109.4	8.0	122	109.5	7.4
(10)	0.17	6.5	0	2	0	0.00	1	1.56	0	0.0	0.0	5	112.7	12.9	145	109.5	8.9	150	109.3	9.0
(11)	−0.27	6.7	2	0	0	0.00	0	0.00	0	0.0	0.0	0	0.0	0.0	160	109.2	11.2	160	109.2	11.4
(12)	0.48	6.9	3	1	2	2.47	0	0.00	14	97.0	24.0	0	0.0	0.0	148	109.2	10.7	141	109.5	9.7
(20)	1.05	8.2	0	0	0	0.00	2	1.56	0	0.0	0.0	12	112.5	13.9	132	108.8	9.8	150	108.3	15.4
(21)	0.71	8.4	2	0	0	0.00	2	1.67	0	0.0	0.0	12	111.8	13.8	148	109.0	10.7	154	109.2	10.7
(22)	0.71	8.7	0	2	0	0.00	0	0.00	0	0.0	0.0	0	0.0	0.0	141	109.4	9.6	141	109.3	9.7
(33)	2.15	8.1	0	10	0	0.00	0	0.00	0	0.0	0.0	0	0.0	0.0	129	109.4	6.6	129	109.4	6.6
(0-1)	2.21	6.5	2	0	0	0.00	3	1.83	0	0.0	0.0	16	120.8	29.3	139	108.9	9.3	144	109.4	9.6
(-10)	3.12	10.5	2	3	3	2.47	0	0.00	16	114.1	28.0	0	0.0	0.0	144	109.3	9.5	130	109.3	9.1
(-12)	3.75	10.8	4	3	7	2.43	0	0.00	34	104.9	28.9	0	0.0	0.0	146	108.9	11.4	120	109.8	7.2
(-66)	30.10	7.5	8	0	33	2.42	0	0.00	70	103.7	21.2	0	0.0	0.0	126	108.5	12.7	92	110.6	6.5
(-11)	2.11	9.1	4	2	5	2.38	0	0.00	24	103.5	20.9	0	0.0	0.0	148	109.1	9.8	131	109.8	8.0
(1-1)	0.47	6.7	2	0	0	0.00	3	1.63	0	0.0	0.0	16	110.2	14.7	139	109.2	8.7	152	108.8	11.5
(-22)	6.12	10.9	7	2	12	2.44	0	0.00	53	100.9	22.9	0	0.0	0.0	140	108.7	10.8	113	110.3	7.0
(2-2)	3.39	10.4	2	2	0	0.00	7	1.63	0	0.0	0.0	34	110.3	14.6	118	108.8	7.2	134	108.9	11.2
(-21)	6.12	10.4	6	2	8	2.43	0	0.00	42	106.9	20.8	0	0.0	0.0	142	109.1	9.8	119	109.6	8.6
(2-1)	2.04	10.7	1	3	0	0.00	5	1.63	0	0.0	0.0	24	110.7	14.5	120	108.8	7.7	130	109.1	10.3
(1-2)	2.29	9.3	2	3	0	0.00	5	1.57	0	0.0	0.0	22	109.5	14.7	136	109.4	7.8	143	108.8	11.7
(-32)	11.11	11.8	4	1	11	2.37	0	0.00	46	108.1	20.9	0	0.0	0.0	142	109.0	8.9	113	109.8	6.4
(2-3)	4.62	10.3	1	3	0	0.00	9	1.57	0	0.0	0.0	35	110.0	11.7	116	109.0	7.0	130	109.0	9.3

(a) (00)   (b) (01)
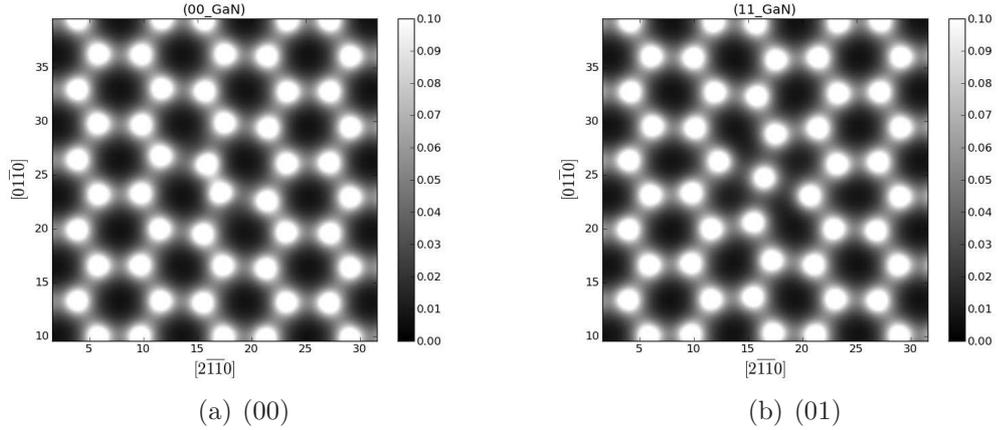
Figure 5.3: Electron density projection of (00) and (11) on the c-plane. The electron density has not been corrected for core electrons.

5.2 Electronic structure for the screw dislocation

This section presents a study of the electronic structure of several optimised core configurations that were presented in section 5.1. In the first stage the GaN quadrupole structures are re-relaxed using a conjugate gradient optimisation with a DFT energy functional. Notable differences to the SW potential predictions are discussed. Subsequently the electronic structure is discussed based on the calculated electron density and the density of states. Finally total energies over a range of chemical potential conditions are compared for all selected structures to find the most probable cores.

The second part of this section shows results for InN core structures. These cores were created using initial GaN structures, scaling the lattice parameters for InN and re-relaxing the core structures. The same analysis is presented for these cores as for the GaN cores.

5.2.1 GaN - Structures

In order to obtain electronic structure information, DFT in combination with the LDA, was used to solve for the electron density of several core structures. A GGA calculation was not performed as structural accuracies are not increased nor is the

87

band gap problem changed for InN, therefore no significant gain of information in the results is to be expected with the application of a GGA functional. As introduced in section 2.1 the AIMPRO code was used, which implements a basis set filtration method and a local wave function basis set to enhance efficiency. As presented in section 4.1 arrays of quadruples are constructed using unit cells containing a dislocation dipole in combination with lattice constants creating a non-orthorhombic symmetry. This results in the modelling of up to 448 atoms per structure. A convergence test of the Monkhorst and Pack k-point sampling was performed to confirm a reliable total energies and density of states.

The structures used here were selected from the dislocation core optimisation presented in the previous section. However this would mean that only one core per unit cell is the desired core configuration, while the remaining one is (00) configuration, which results in a combined influence of both dislocation cores on the electronic structure of the total structure. In order to prevent such an unwanted influence, new unit cells were constructed containing a dipole of mirrored copies of the optimised core. This could be done by using the mirror symmetry of the $(2\bar{1}\bar{1}0)$ plane of the underlying wurtzite structure. This operation changes the parity of the screw dislocation while retaining its core configuration, hence a unit cell containing two opposite screw dislocations could be created. However, as the dislocation line position varies for each core structure, it is not always precisely on the chosen mirror plane. As a result the mirror operation creates a deformation of the rectangular quadrupole into a rhombohedral configuration. This creates a slight total energy change, which will be neglected in the further analysis as it does not influence the core structures and can be estimated to be less than 0.1eV using the SW predictions from chapter 4.2.

The structures chosen for the DFT analysis are located around the global relative energy minimum as predicted in 5.1 ((00),(11),(01),(10),(-11),(1-1)) as well as the three lowest energy structures of the (00) core and several core configurations ranging into the gallium filled core structures ((-12),(-22),(-66)). Ball and stick models for several of the studied cores can be found in figure 5.4.

All selected structures were relaxed using a local conjugate gradient optimisation with DFT. Major structural changes from the SW predictions could only be
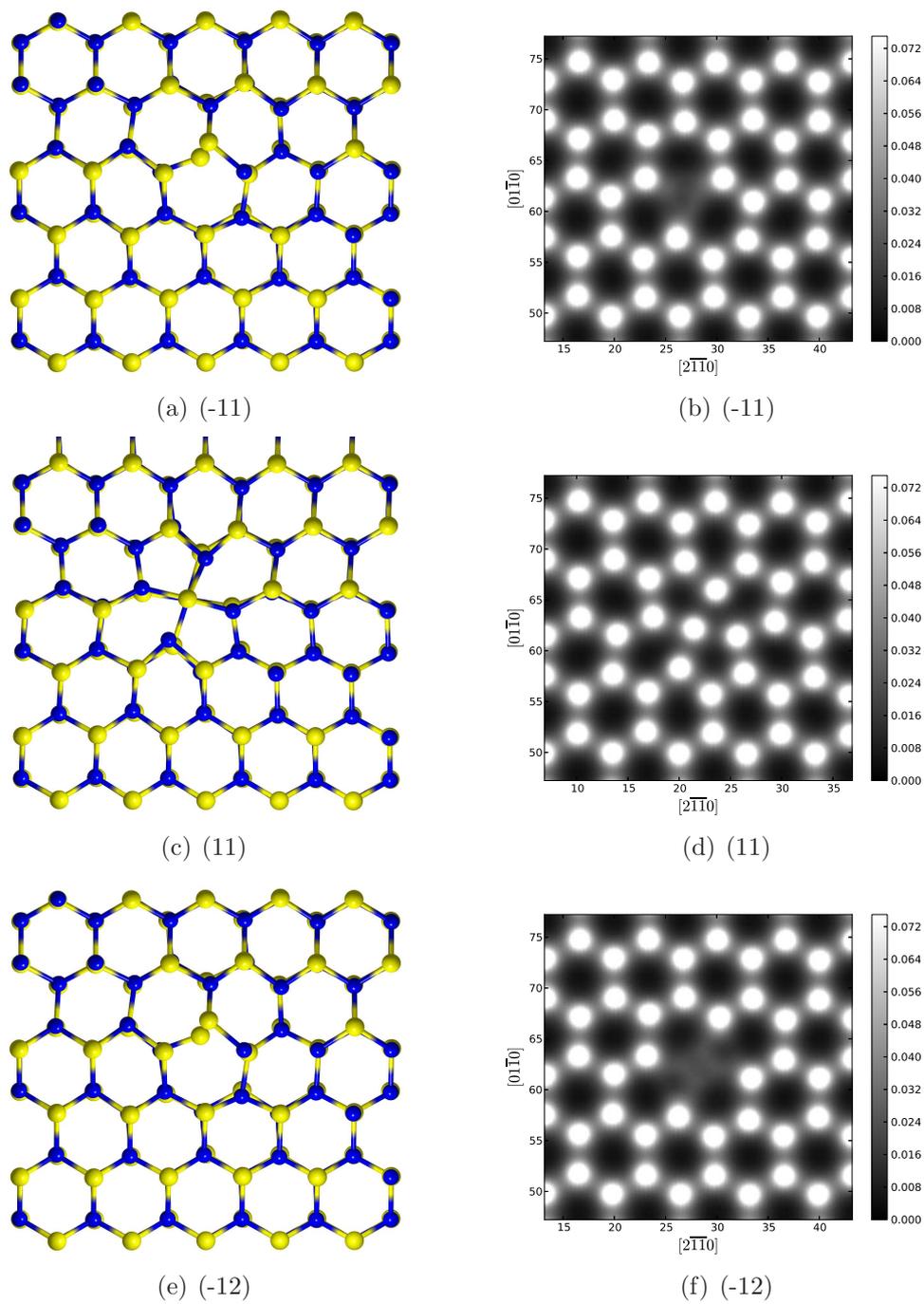
(a) (00)

(b) (01)

(c) (10)

(d) (11)

(e) (-11)
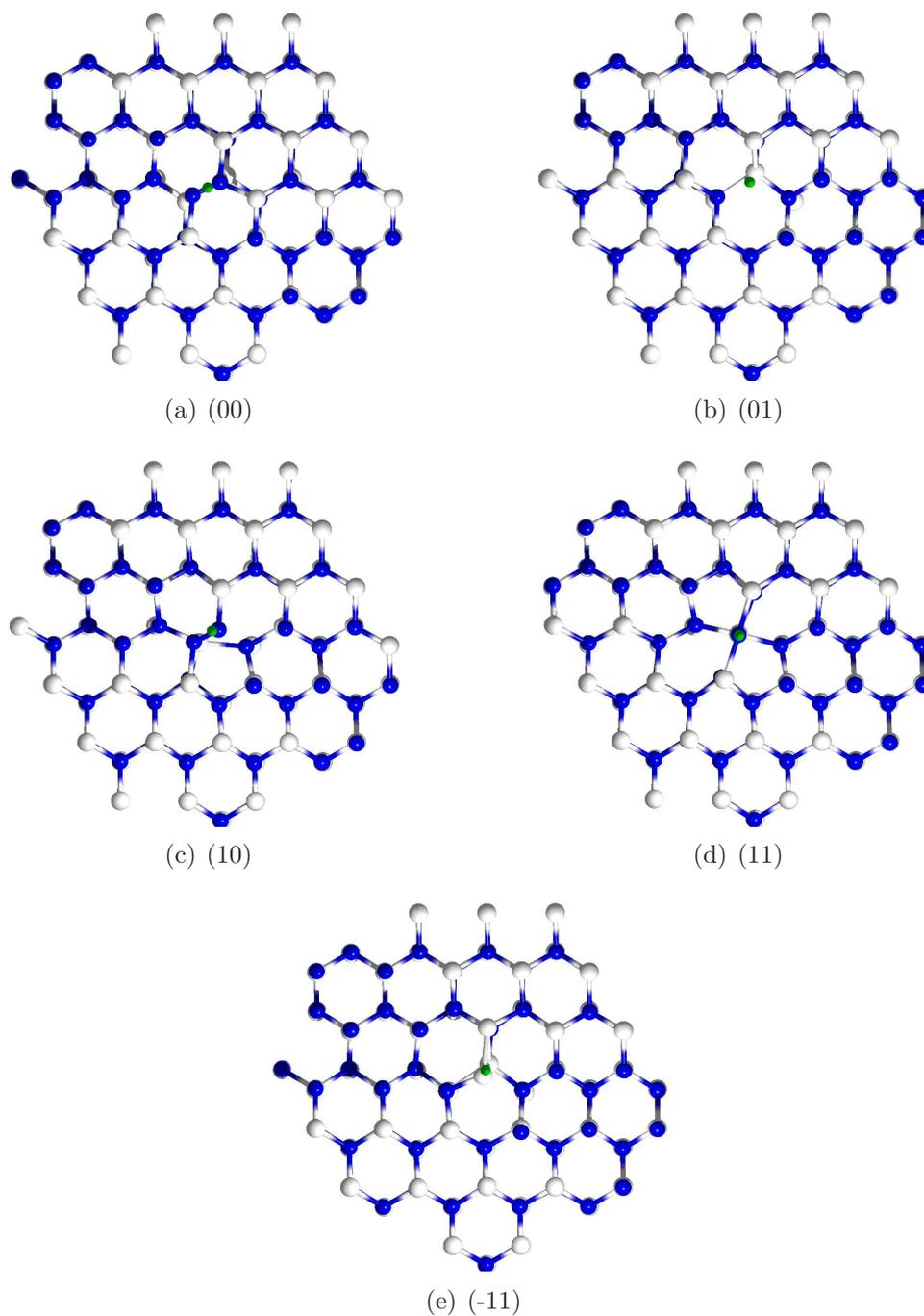
Figure 5.4: Core structures for various stoichiometries, where all atoms were re-relaxed using a conjugate gradient minimiser and DFT. Gallium is marked as white balls nitrogen as blue and the dislocation line position as green circle.

89

observed for (1-1) where the additional nitrogen atoms departed from a tetrahe-
dral bonding and favoured a proximity to other nitrogen atoms. As can be seen
from figure 5.4(a), in comparison with figures C.1 through C.12, the DFT relaxed
structures feature only small changes. More precisely both the dislocation line
position as well as the structure remains the same for the (00) core and also for
the (01) core, however the (10) configuration slightly changes the position of the
nitrogen at the gallium vacancy site to be more closely bonded to its nitrogen
neighbour. At the same time the gallium at the neighbour position moves closer
to the dislocation line position. This deviation creates a core which more closely
resembles the (11) configuration.[4] The structural changes for the (11) core are
more visible but are in fact minor. The DFT relaxed structure simply moves the
central GaN site into a perfect centre position, creating a more symmetric core
with the line position exactly on top of this central site. The (-11) core shows no
obvious structural changes from the SW predictions.

5.2.2   GaN - relative energies

The total energy for all selected structures was computed using DFT under LDA
with a local basis set, non linear core corrected pseudo-potentials and a converged
k-point grid. However, in order to compare total energies of two structures with
different stoichiometries the thermodynamic Grand Potential

$$\Omega = U - TS - \sum_i \mu_i n_i, \tag{5.2}$$

needs to be determined [78]. The following procedure is used in analogy to
Northrup et al. [79][77][80]. Equation (5.2) states that the total structure energy
U needs to be corrected by an entropy term TS and the sum of the chemical
potential of its constituents. The Grand potential compensates the difference in
composition of different structures by the inclusion of the chemical potential of
the constituents, thus creating an energy term that can be directly compared for
all obtained core structures and directly relates to the relative likelihood of core
formation in a real solid. The entropy term in equation 5.2 can be neglected here
due to the low influence of the entropy in solids at relatively low temperatures
(including room temperature [79]), while the chemical potentials for gallium and

---

[4]   This behaviour is much more pronounced for the InN (10) core, see section 5.2.4

Table 5.2: Total energy comparison for all structures investigated by DFT expressed with their thermodynamic potential relative to the (00) structure at equilibrium position. (00)' and (00)" denote higher energy core configurations as predicted by the random optimisation and SW potentials.

Core	(00)	(00)'	(00)"	(01)	(10)	(11)	(22)	(-11)	(-12)	(-22)	(-66)
$\Omega[eV]$	0.00	-0.07	0.03	0.27	0.88	-0.14	0.95	-0.39	-0.70	0.84	13.85

nitrogen needs to be determined.

The chemical potentials of Ga and N in a GaN crystal are not easily determined as they can vary depending on the condition of the thermodynamic bath. However the variables can be constrained, if the system is in contact with a bulk GaN reservoir. In that case the corresponding constraint on the chemical potentials

$$\mu_{\text{GaN}} = \mu_{\text{Ga}} + \mu_{\text{N}} \tag{5.3}$$

holds true with μ_{GaN} being set by the calculated energy of a GaN formula unit. μ_{GaN} can be calculated readily for the two limiting cases the chemical potentials cannot exceed, allowing chemical potentials to be expressed solely as a function of μ_{Ga}, as equation (5.3) holds true in all cases. These two limiting cases are the chemical potential of bulk gallium and that of molecular nitrogen. Therefore the range of values for μ_{Ga} represents growth conditions from gallium to nitrogen rich. The bulk gallium chemical potential $\mu_{\text{Ga(bulk)}}$ was determined by computing the structural energy for the stable configuration of α-Ga as specified by [81], while the nitrogen molecule potential $\mu_{\text{N(mol)}}$ was computed similarly from a N_2 DFT molecular calculation. With these two limiting points, the heat of formation of GaN ΔH can be expressed as:

$$\Delta H = \mu_{\text{Ga(bulk)}} + \mu_{\text{N(mol)}} - \mu_{\text{GaN}} = 1.68eV \tag{5.4}$$

and is equal to the range μ_{Ga} can have. The value for the formation enthalpy, ΔH, agrees with similar LDA calculations in the literature [82].

The determined thermodynamic potential for each investigated structure is shown as a function of the chemical potential range in figure 5.5. Here $\mu_{\text{Ga(GaN)}}$ was chosen as a reference for the centre range so as to describe equilibrium conditions.
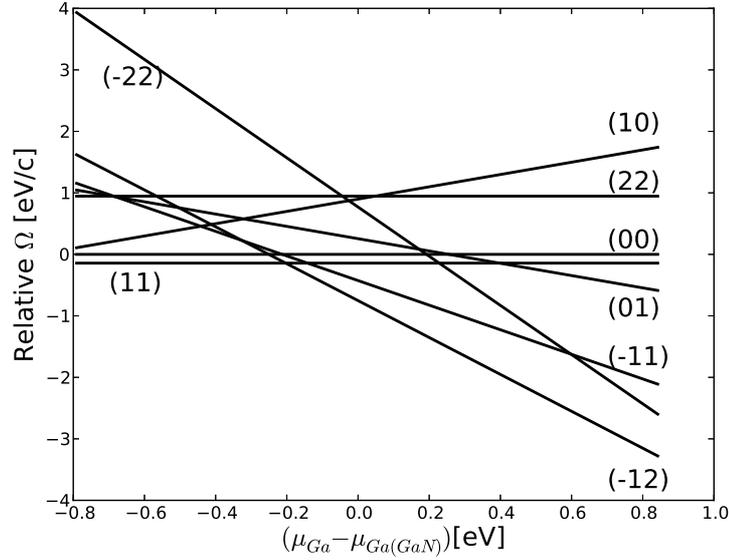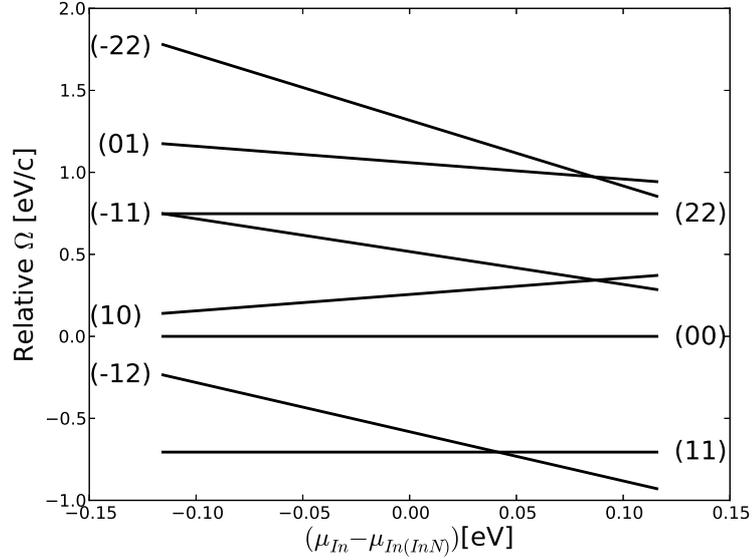
Figure 5.5: Relative dislocation core energies as calculated by DFT plotted for varying gallium chemical potential. The (00) core has been taken as reference.

For the fully stoichiometric cores only the lowest energy core were shown, as the energy difference is too small to be clear on the scale used. All other cores were drawn relative to the (00) cores. The relative thermodynamic potential for all structures at the equilibrium position can be found in table 5.2.

From figure 5.5 it is notable that, as the SW potentials optimisation predicted, the (11) core is found to be more stable than the normal (00) core. However the gallium rich cores are shown to be much more favourable. In fact for most of the growth condition range, the (-12) core is found to be the most stable core configuration. Only under nitrogen rich conditions do the stoichiometric cores represent the lowest energy state. The completely gallium filled core (-66) containing 12 gallium atoms is very unfavourable for all growth conditions and has not been included.

5.2.3 GaN - density of states

The density of states (DOS) for several of the investigated structures are displayed in figure 5.6, where the defect DOS has been drawn behind the GaN bulk DOS to emphasise states within band gaps. Notable for figures 5.6(a),5.6(b) and 5.6(c)
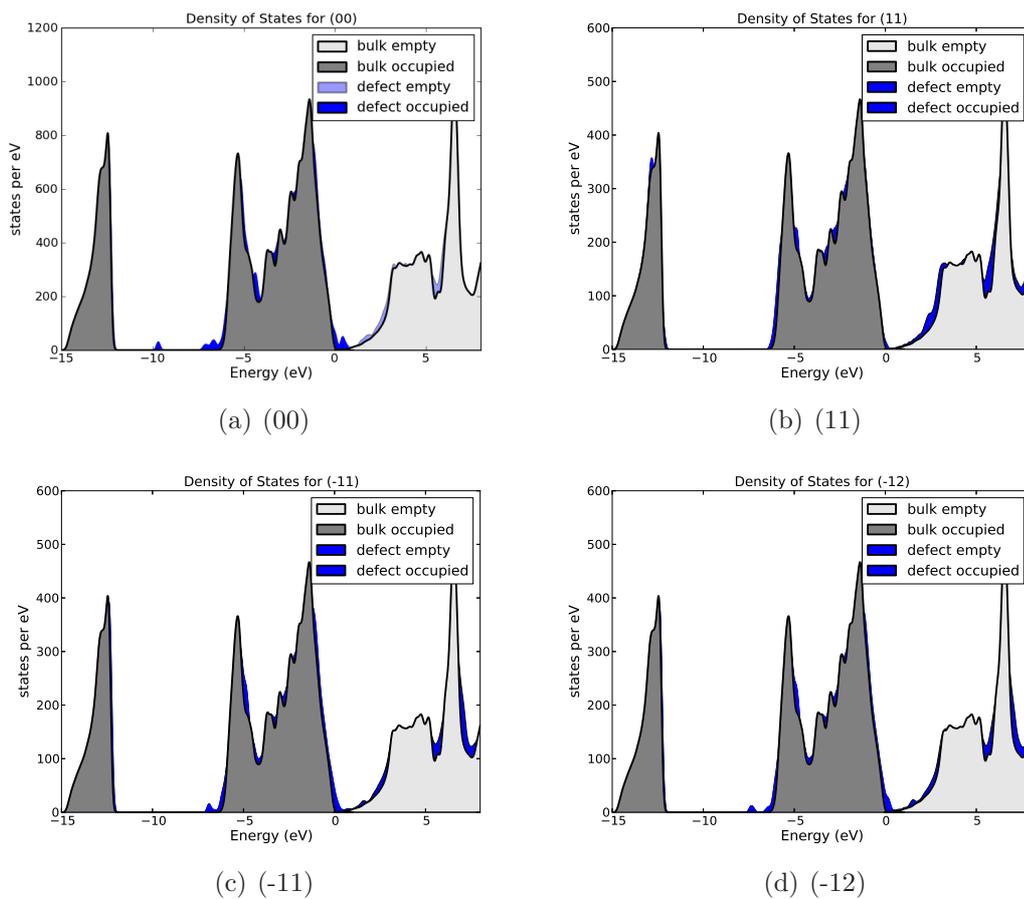
(a) (00)

(b) (01)

(c) (10)

(d) (11)

(e) (-11)

(f) (-66)
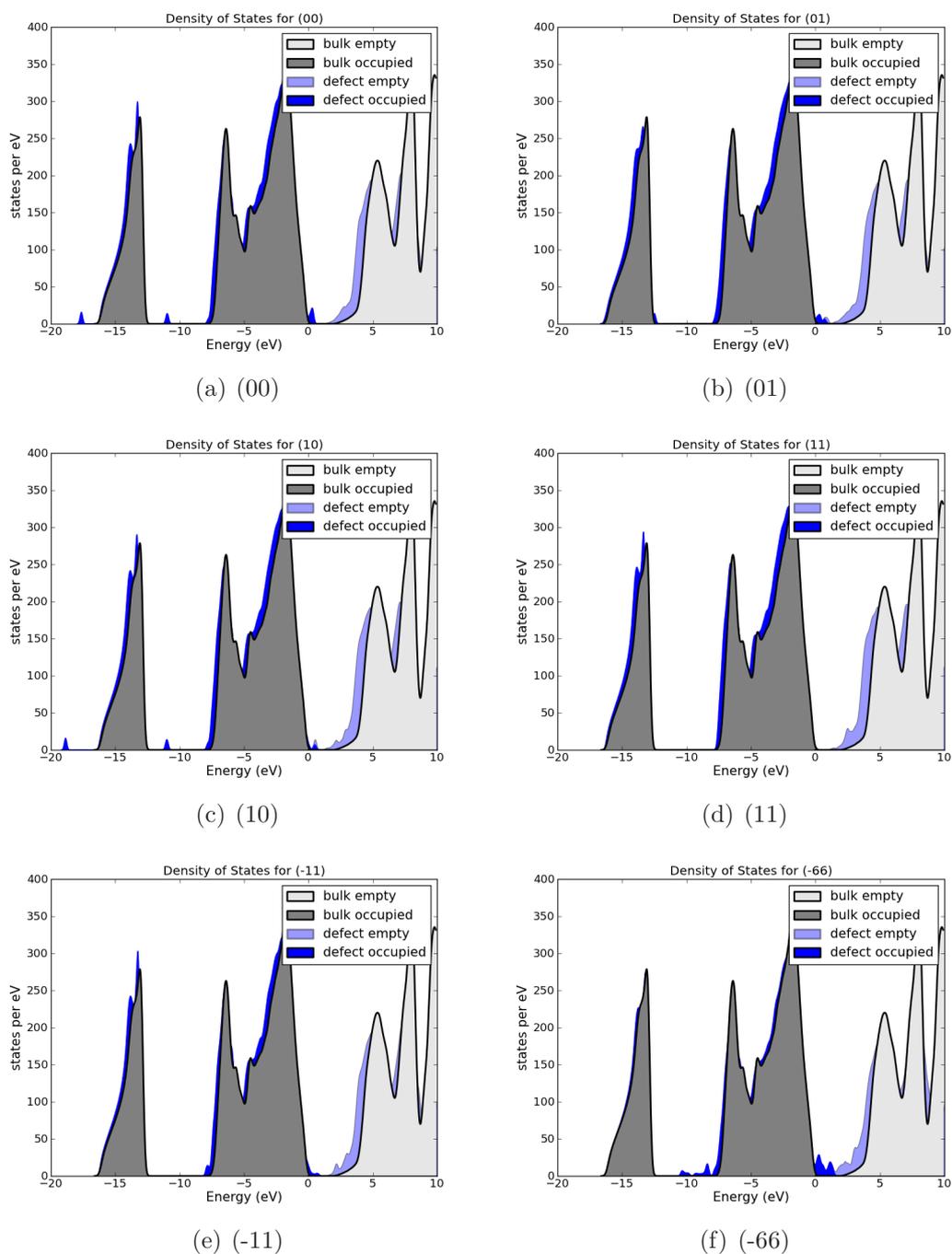
Figure 5.6: Density of states comparison of bulk and various core configurations. The valence band edge of the bulk is set as zero point.

are the position and height of the gap states. These imply a strong influence of the homo-nuclear bonds in these structures, in particular the N-N on states in the gap. The most notable difference is the complete lack of states above the valence band for the (11) core in figure 5.6(d). There are however in all cases lower band gap for all dislocation structures, with a metallic behaviour for cores more rich in gallium than the (-11) core, where the gap has been decreased to only about 0.4eV.

5.2.4   InN - structure

A structural and electronic investigation of screw dislocation cores was also performed for InN using the stable core configurations found for GaN. This analysis was undertaken under the assumption that the global optimisation of dislocation cores in GaN produces core configurations that are structurally sound for general wurtzite materials. The main difference between a GaN and InN bulk wurtzite structure is the size of the lattice repeat, which is about 10% larger in InN. This type of scaling can therefore transform a GaN cell into InN by changing the atomic type of the group-III atoms. It would produce valid copies of bulk material and long range dislocation strain fields, as these scale with the Burgers vector, which is by definition proportional to the lattice vectors for any dislocation. The dislocation core however can incorporate homonuclear bonds and bonding situations that do not translate well from GaN to InN, as the nitrogen bond length does not scale and bulk indium is most stable in a different configuration than gallium.[5] Therefore different stable core configurations could be viable for InN and GaN, which is not further considered, instead in order to deal with non-scaling relative atomic positions a local optimisation (conjugate gradient) was performed subsequently.

A selection of investigated structures can be seen in figure 5.7, where for each structure a projection of the electron density has been drawn in addition, using only the valence electrons without correcting for the pseudo potentials. As a result indium positions cannot be seen directly, only a diffuse rise in electron density around their position is visible. As a result the metallic single atomic channel

---

[5]   Bulk indium crystallises in a body centred tetragonal lattice, while gallium has various stable forms, but favours under normal conditions α-Ga, which is an eight atom unit cell in an orthorhombic lattice configuration.

(a) (-11)

(b) (-11)

(c) (11)

(d) (11)

(e) (-12)

(f) (-12)

Figure 5.7: Core structures for various stoichiometries, where all atoms were re-relaxed using a conjugate gradient minimiser and DFT. Presented with ball and stick model as well as an electron density projection on the c-plane. Indium is marked as yellow balls, nitrogen as blue.

in the (-11) core structure is emphasised nicely (5.7(b)). Structural differences after the atomic position relaxation are small throughout all investigated structures, except for the (10) structure which shows a visibly more central location of the single most central indium atom, when compared to the equivalent GaN structure.

5.2.5   InN - relative energies

In order to compare all investigated structures the thermodynamic potential Ω was obtained in the same manner as for GaN. In order to determine indium rich conditions, the chemical potential for indium was determined by calculating the binding energy of indium bulk in the body centred tetragonal lattice with LDA-DFT. The same molecular nitrogen chemical potential could be used for the nitrogen rich conditions as in the GaN case. With these values, including the chemical potential of an InN pair in a wurtzite lattice, the enthalpy ΔH of InN was determined to be 0.23eV.

The enthalpy of InN is much lower than for GaN, indicating that a change in growth conditions will have a lower impact on the thermodynamic potential of structures with different stoichiometries. Figure 5.8 displays the found relative thermodynamic potentials of the investigated structures. It can clearly be seen that a change in growth conditions has a much lower impact on the ordering of the various structures. This is shown as a function of the chemical potential of indium μ_{In}. The relative ordering of the structures is different from the GaN case as the group-III rich structures are not favoured as strongly, in fact the (10) core is consistently lower in energy than the (01) core. In addition the (11) core is lower in energy compared to the (00) core by 0.7eV making this structure the most feasible core for most of the growth condition range. Only for the last third of the indium rich end is the (-12) core predicted to be lower in energy. As in the GaN case the relative energy for the (22) core is much higher than that of the (11) and (00) cores, confirming also the SW-potential predictions.

5.2.6   InN - density of states

The densities of states for the InN core structures were computed and presented in analogy to the GaN case. Here also the valence band maximum of InN bulk

Figure 5.8: Relative dislocation core energies as calculated by DFT plotted for varying gallium chemical potential. The (00) core has been taken as reference.

was chosen as the reference energy. It is important to note that LDA-DFT predicts a zero band gap width for InN as was shown in section 3. Still, defect states appear above the bulk valence band edge features, suggesting that with a band gap present these would lie inside or even cover it. Especially the distance of filled defect states to filled bulk states should be predicted with reasonable reliability within these calculations as DFT is a ground state theory.

The densities of states of several low energy configurations are shown in figure 5.9. As in the GaN case and as can be seen in figure 5.7, the stoichiometric cores differ clearly in the number of homo-nuclear bonds, meaning that the (11) core is a configuration featuring a 5 fold coordinated centre InN pair without any homo-nuclear bonding. As a result there appear no additional defect states immediately above the valence band in the (11) core. The same is true for the (22) core, however this configuration has a much higher relative energy and is therefore not shown here. On the other hand the (-11) and (-12) cores are Indium rich and feature a single-atomic indium wire at the dislocation core centre. This results in additional states above the valence band edge, making the InN bulk "more" metallic.

(a) (00)

(b) (11)

(c) (-11)

(d) (-12)

Figure 5.9: Density of states comparison of bulk and various core configurations. The valence band edge of the bulk is set as zero point.

5.3   Discussion

The results in this section are divided into two major parts, the first is the global optimisation across a wide core composition range and the second are electronic structure investigations performed with DFT on the optimised cores in GaN and InN. The global optimisation on each configuration was performed in analogy to the (00) case presented in chapter 4. Looking at the predictions of core stability directly from the global optimisation it can be seen that the lowest energy core is the (11) core. However looking at the more general trends of the relative energies it was shown that SW predicts nitrogen rich cores to be more stable. Empty cores as well as the extreme nitrogen and gallium rich cores were predicted to be unfavourable. However, all relative energies were given by using the chemical potentials for Ga and N in GaN bulk as given by the atom resolved energies. Only for the DFT total energies were chemical potentials over a wide range considered, meaning that relative energies with different N to Ga ratios can shift with varying conditions.

In section 5.2 the most stable cores from the optimisation procedure are re-relaxed using DFT. It can clearly be seen that structural changes to the optimised cores are small, with the largest deviations being in the (11) core: The dislocation core centre and a GaN pair are shifted to an exact symmetry centre, creating two 5 atom rings and two heavily strained 6 atom rings, when viewed in plan view (see figure 5.4(d)). The relative energies of the three lowest energy structures of the (00) structures as produced by the optimisations, do not follow the same behaviour as in the SW case. However, all relative energies are extremely small in comparison to the difference in energy to other core configurations. A similar study of a range screw dislocation cores in GaN has been presented by Northrup et al. [77], where the gallium filled cores are predicted to be most stable over the complete range of growth conditions. However, the results shown here include core structures not considered by Northrup et al., such as the (11), (-12) and (-11) cores, in addition a wider range is predicted here for the GaN enthalpy ΔH, resulting in "more" nitrogen-rich conditions. This creates a small range where gallium filled cores are unlikely. The value of 1.68eV for ΔH calculated here agrees with similar LDA-DFT predictions from [82], however their GGA-DFT calculations predict a much lower value of 1.12eV. However, it is still clear that there is a trend of very stable gallium filled cores, even into the Nitrogen rich

condition range. From the presented results it can be projected that cores like the (-1,3),(-2,3) or even (-4,6) are the most stable configurations over the complete enthalpy range.

InN has a much lower enthalpy than GaN. The calculations performed here predict a value of 0.23 eV, which does not agree as well with the results (0.3eV) presented in [82], but is closer to experimental results (0.21eV). A much lower value results in a much lower reordering of relative core stability as growth conditions do not affect the core stoichiometry as heavily. As a result it can be seen in figure 5.8 that for most of the presented chemical potential range the most stable stoichiometric core (11) is the overall energetically best core configuration, while the Indium rich configurations only become favourable under very Indium rich conditions. Again, more Indium rich core configurations could have lower energies, however looking at the relative ordering of the other cores, the InN material does not tend towards metal rich compositions.

The electronic properties of the various cores investigated show clearly that as expected the more metal rich cores introduce defect states above the valence band all across the band gap. This means that they can act as conductive paths and non-radiative recombination centres. The core configuration with the least influence on the band gap is the (11) core as it does not feature any homonuclear bonds. This means that for the InN case c-screw dislocations should be electrically inactive as here the (11) core is the most stable of the investigated configurations across most growth conditions.

Edge dislocations in GaN

This chapter contains a discussion of the application and results of a global optimisation of the edge dislocation core structure in GaN. As in chapter 4, a random search algrithm was used that utilizes Stillinger Weber type interatomic potentials to generate a cost function. In addition to the stoichiometric core, a selection of non-stoichiometric conditions ranging from nitrogen rich to Group-III rich conditions as well as open core structures were considered. Unlike chapters 4 and 5, the effects of dislocation climb on core composition become apparent and are discussed.
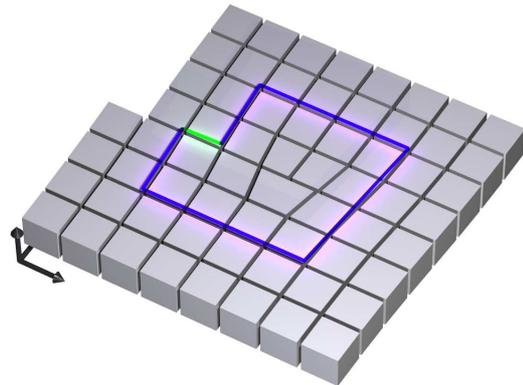
Figure 6.1: Schematic of an edge dislocations in a cubic system. The purple line represents a Burgers circuit, the green line shows the Burgers vector

The first section in this chapter details the geometry of the investigated structures and provide a discussion on the feasibility of this approach. The following sections shows results in detail for the global optimisation of the fully stoichiometric core, while a comparative study of all studied core compositions is performed

(a) unit cell schematic   (b) abs(ϵ)
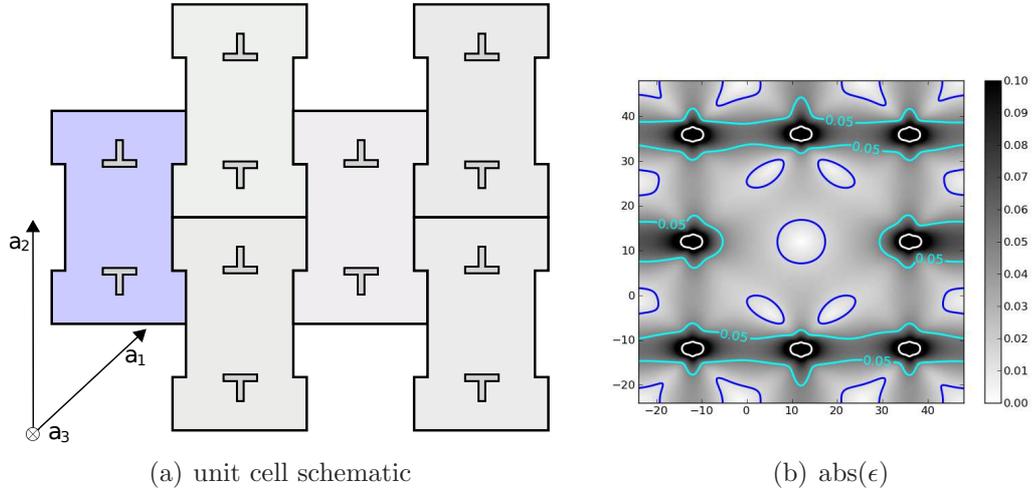
Figure 6.2: Unit cell schematic for the edge dislocation optimisation (a) and absolute value of the strain influence on investigated core region (b)

subsequently.

6.1   Optimisation parameters

The optimisation procedure was performed for the edge dislocation in analogy to the screw dislocation case (chapter 4). Similarly a supercell structure was created with a choice of boundary conditions to construct an array of dislocation quadrupoles. The size of the the unit cell has identical dimensions, the dislocations however were included by applying the displacement field for an edge dislocation as presented and derived in equation (2.127) to the constructed unit cell. The chosen quadrupole arrangement is presented as a schematic in figure 6.2(a). The resulting strain influence from the dislocation quadrupole arrangement on the investigated dislocation core was determined as for the screw dislocation case. The resulting absolute value of the strain on the dislocation core region can be seen in figure 6.2(b). Due to asymmetry of the strain in the c-plane for an edge dislocation with a line direction along **c**, the absolute value of the strain in figure 6.2(b) differs from the arrangement seen for the screw case (figure 4.3(f)). However, the fact remains that the strain due to the surrounding dislocations on a single dislocation core is minimal for the cell sizes chosen.

The application of the edge dislocation displacement field can yield a range of

core configurations depending on the position of the dislocation line on the c-plane within a unit cell. After relaxation, this can lead to a 5/7 core, 8-core or 4-core configuration [71]. The initial line position chosen within this work resulted in a 5/7-core configuration arrangement, where the dislocation core consists of a 5 atomic site ring and a 7 site ring (see figure 6.5(a) for an example).

As with the screw dislocation, the global optimisation was performed using a randomisation radius of around 3.2Å surrounding the dislocation line position, which results in a cylindrical randomisation region containing 12 atoms. As a result usual optimzation performances were similar and up to 10^5 randomisation relaxation steps were completed for each core configurations.

6.2 Fully stoichiometric core

The fully stoichiometric core was defined for the screw dislocation as the dislocation core that was randomised with a full set of atoms (12 atoms, 6 Ga and 6 N) inside the randomisation region, i.e. the (00) core. However, the situation for the edge dislocation is more complex as the Burgers vector lies inside the plane of the simulated sheet of atoms, resulting in a half plane of inserted material perpendicular to the Burgers vector. A movement of the edge dislocation core along this half plane is called dislocation climb and will result in the stoichiometrically constant[1] filling or emptying (depending on direction) of the core region. As a result the edge dislocation can form its most stable dislocation core with a change of centre position by gaining or losing GaN pairs. Therefore the fully stoichiometric core for the edge dislocation will here be defined as the combination of all possible stoichiometric cores, i.e. (00),(11),..,(nn); with n being the number of GaN pairs inside the randomisation region.

The movement of an edge dislocation can be classified into slip and climb. It is a slip movement if the dislocation line position shifts along the Burgers vector and climb if it is perpendicular to it. Figure 6.3(a) displays a schematic of dislocation climb and slip for the investigated core region in the present simulation environment. In addition a strain energy analysis has been performed by simulating the analytical elastic energy for dislocation movement along slip and climb

---

[1] Meaning a change of equal numbers of Ga and N atoms present in the core region.
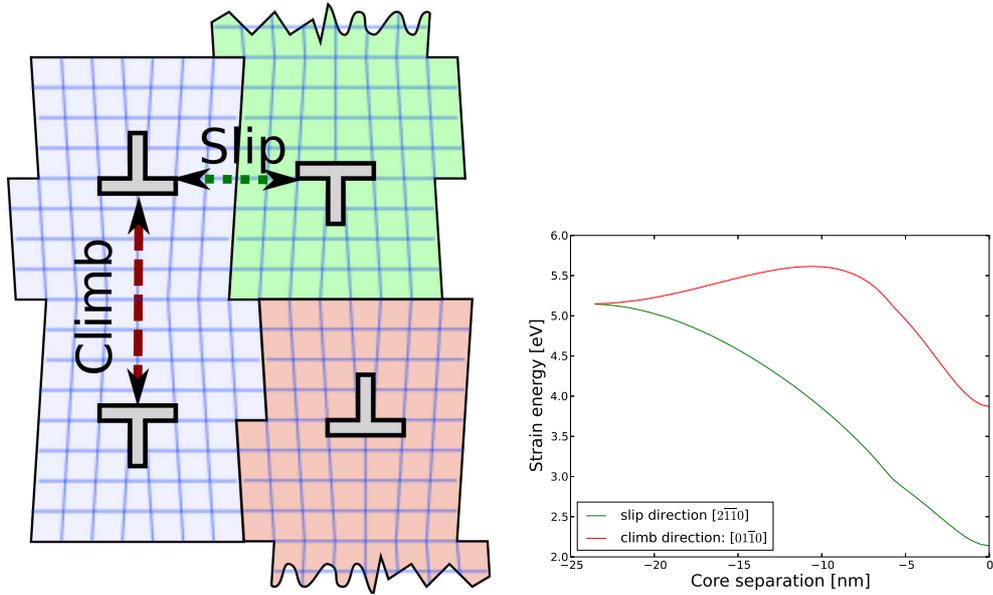
(a) Schematic showing possible simula-
tion cell (blue) plus PB neighbours (green
$(+\mathbf{a}_1)$ and red$(+\mathbf{a}_1 - \mathbf{a}_2)$) with added dis-
location climb and slip directions

(b) Strain energy per simulation cell for dislo-
cation slip and climb movement

Figure 6.3: Schematic and simulation for dislocation slip and climb movements.

directions under present simulation conditions (figure 6.3(b)). The simulation
clearly shows an energy barrier for climb movement and an immediate drop in
total elastic energy for slip movement.

As seen for the screw dislocation optimisation procedure, core configurations are
expected to form around the starting position with distances of up to one lattice
constant. Therefore the energy barrier in climb direction, due to the strain field
surrounding the dislocation under investigation, will result in an ordering of the
structure distribution to show cores centred on the equilibrium position. Slip
movement on the other hand has no effect on core composition, but features a
quick decline in total energy, creating the possibility of core movements during
relaxation. Such a movement of the dislocation core along the $[2\bar{1}\bar{1}0]$ direction
for the chosen geometry leads to an annihilation with the next nearest neighbour
dislocation featuring opposite Burgers vector.

The structure energy distributions created by the optimisation algorithm for the
(00), (11), (22) and (33) cores are depicted in figure 6.4. It is immediately pos-
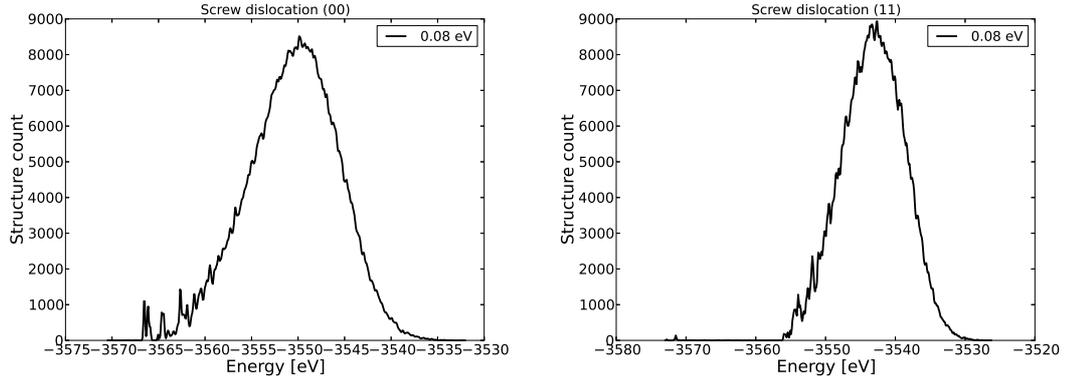
(a) Structure energy distribution for the (00) optimisation

(b) Structure energy distribution for the (11) optimisation
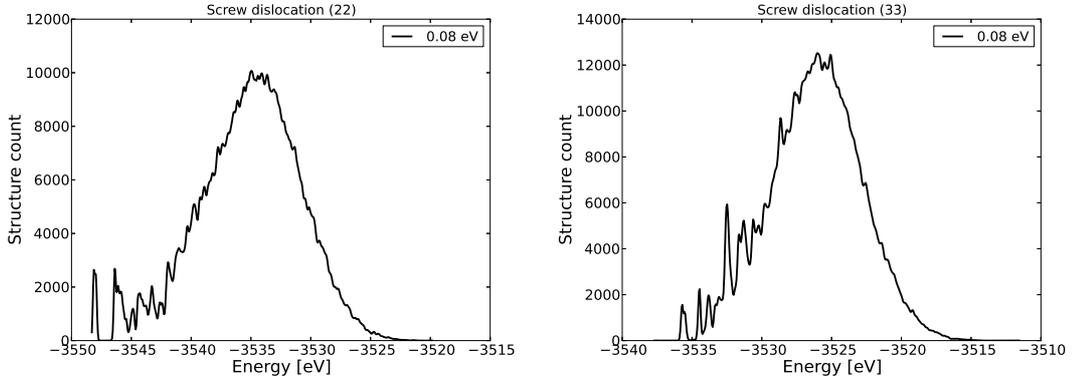
(c) Structure energy distribution for the (22) optimisation

(d) Structure energy distribution for the (33) optimisation

Figure 6.4: Structure energy distributions for the global optimisation of the stoichiometric edge dislocation core.

sible to see that aside from the main distribution, a very long low energy tail exists. The structures present in the tail however, are all geometries that have undergone slip movement and annihilated with the opposing dislocation. However these unwanted slip configurations are all very low in probability as the structure count per bin is low compared to the structure count of the genuine low energy structures. Furthermore the (11) optimisation features a much higher number of slip-annihilated structures as compared to the other stoichiometric optimisations.

The genuine lowest energy structures in the (00) optimisation procedure are all variations of the original 5/7 cores with different core positions. The next lowest genuine structure is approximately 2eV above this position and features a

larger core size containing a "double" 5/7-core, where the two atomic layers in c-direction in the cell both feature a 5 and a 7 site ring that are displaced in a slip direction by one lattice constant. However, the lowest energy structures in the optimisation results for the (00) core do not contain 4 cores or 8 cores that are expected to exist, as they have been shown to be stable by experiment and theory results in the literature. In addition both cores can be directly created with the edge dislocation displacement field acted on a perfect wurtzite lattice and a subsequent SW-optimisation, suggesting that they exist as local minima in the SW paramterised interatomic potential cost function and are therefore expected to be represented within the optimisation structure distribution for the stoichiometric core. Instead both core types are not present in the (00) optimisation but can be found readily as the most stable structures in the (11) optimisation. Figure 6.5 displays structure models for the mentioned cores.

The situation that 5/7 and 8-,4-cores appear mutually exclusive in the (00) and (11) optimisation respectively can be explained with the chosen boundary conditions and their relationship with respect to their dislocation centre position. Relative to a lattice unit cell, the 8-core and 4-core centre positions are separated only along the slip plane while the 5/7 cores centre is positioned away along the climb direction. The distance apart in climb direction results in half of a unit cell (a pair of Ga and N atoms) being incorporated into the 5/7 core, which are not present in the 8- or 4-core. However, as the total number of atoms in the simulated supercell remains constant, the other half unit cell is then not present at the connection to the boundary condition, which needs to fit to the other opposite core in the simulated unit cell, which has been chosen and confined to be a 5/7-core. Or in other words, the 5/7 core contains a single GaN atom pair more than the 4- and 8- core and as a result a supercell containg both an 8-core and a 5/7-core needs to contain an uneven amount of GaN pairs, even though this would not be possible in a supercell constructed from bulk with dislocation displacement fields and PBC. Hence in the present optimisation only when half a unit cell[2] is removed or added from the randomisation procedure will the creation of an 8- or 4-core be possible.

---

[2] Due to the possibility in dislocation climb with integer steps of the lattice vector, it is more general to say that uneven amounts of GaN pairs need to be removed or added.

(a) Most stable structure for the (0a) core

(b) Second stable structure for the (0a) core

(c) Most stable structure for the (0b) core

(d) Other stable structure for the (0b) core
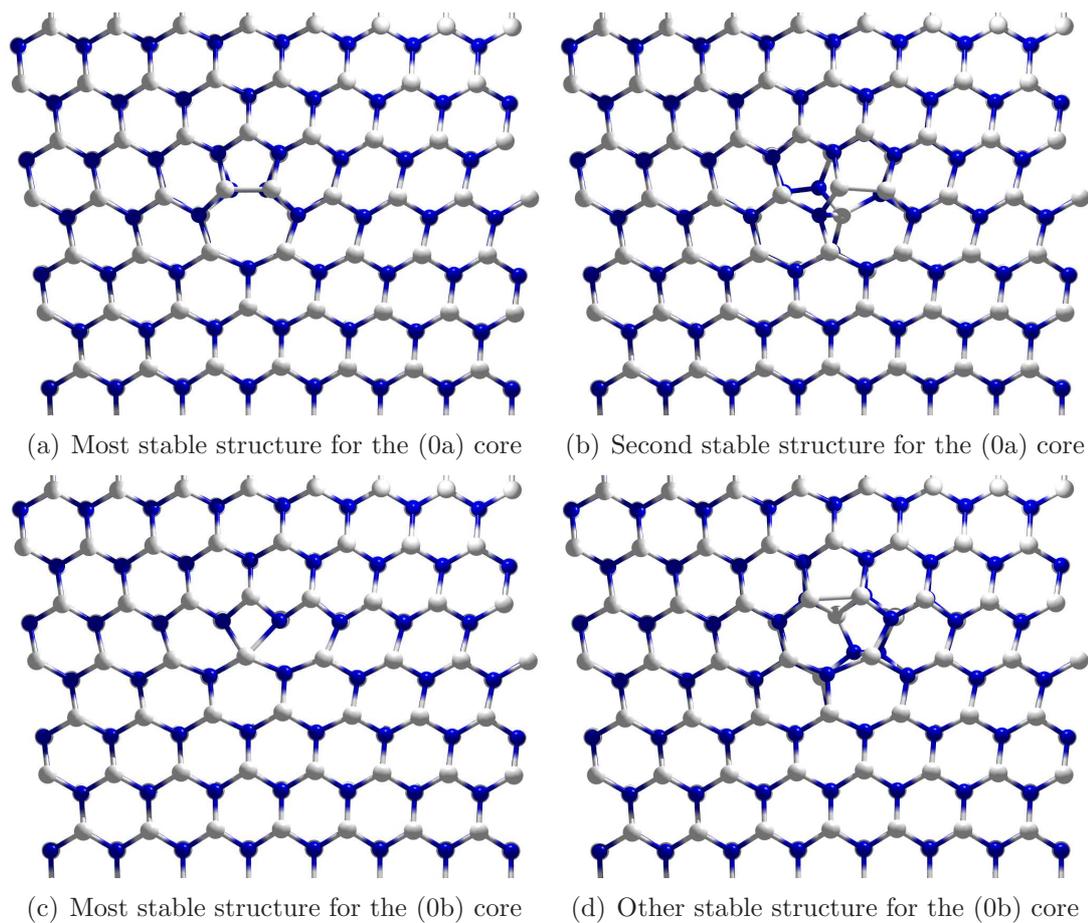
Figure 6.5: Ball and stick model for stable stoichiometric edge dislocation core structures.

The necessity of preserving complete unit cells in the dislocation climb is further evidenced by the appearance of a shifted double 5/7 core in the (11) optimisation. The shifted double 5/7 core is composed of two 5/7 ring structures, split between the two atomic layers in c-direction and separated by one lattice site in the climb direction.

The high count of slip annihilated structures during the (11) optimisation can be linked to the predominance of 8- and 4-cores and their high mobility along the slip direction. As 4- and 8-cores are directly related to each other by a shift of dislocation centre position along the slip direction within one unit cell, they constitute two repeating local minima for the motion of the core along the path to annihilation. The slip movement during structural relaxation causing the annihilation process is aided by the decrease in total elastic energy as two dislocation cores approach each other (figure 6.3(a)).

All other performed stoichiometric core optimisations ((22) and (33)) show very similar results. In addition, a clear trend of the edge dislocation core reconstructions towards filled cores can be observed. Both the (22) and (33) optimisation produce open core structures at their randomisation site. However the lowest energy structures in both cases are dislocation cores moved in climb direction as to reach a filled state, despite the accompanying increase in elastic energy of the simulated supercell.

6.3   Composition range

As for the stoichiometric case, the non-stoichiometric cores cannot be easily classified through their amount of randomized atoms (or lack thereof). Instead the composition range discussed for the screw dislocation case can be reduced to purely the difference between the number of Ga to N atoms. Therefore the two dimensional composition range $(N_{V_{Ga}}, N_{V_N})$ can be projected into $(N_{V_{Ga}} - N_{V_N})$ through the equality of $(N_{V_{Ga}}, N_{V_N}) = (N_{V_{Ga}} + C, N_{V_N} + C)$, with $C \in \mathbb{Z}$. However it is best to keep in mind that, due to the boundary conditions and manner of optimisation, two separate optimisations need to be perfomed for each class: $(N_{V_{Ga}} - N_{V_N}) = (N_{V_{Ga}} + C, N_{V_N} + C), (N_{V_{Ga}} + C + 1, N_{V_N} + C + 1)$. For instance the class of optimisations (10),..,(n+1,n)=(1) is complete when optimisations only for

(10) and (21) have been performed. Therefore dislocation cores will be labelled as $((N_{V_{Ga}} - N_{V_N})a)$ and $((N_{V_{Ga}} - N_{V_N})b)$, where a denotes the case where the lowest vacancy count is even and b when it is odd.

Optimisation procedures for non-stoichiometric cores were performed for the cases (01),(10),(12),(21),(23) and (32). These reduce to the two edge stoichiometries (1a),(1b) and (-1a),(-1b), where only the a cases are contained twice. As expected from the stoichiometric case, the lowest energy structures in the (23) and (32) optimisations case map onto the (01) and (10) results and will not be discussed further. The lowest energy structures emerging from the remaining optimisations can be found in figure 6.6, where all structures are shown in the same segment of the simulated cell.[3]

The most stable (-1a) [4] structure features a 5/7 core with a nitrogen vacancy at the centre of the dislocation removing the nitrogen homonuclear bond, while the most stable (-1b) structure features a mix between 8 and 5/7 cores. This can be understood as either a 5/7 core with a gallium interstitial or an 8-core with a nitrogen vacancy. The three distinct most stable structures for the (1a) dislocation core can be found in figures 6.6(c)-6.6(e) and show several interesting mixtures of 4-,8- and 5/7-cores. The lowest energy structure shows a 5/7 core arrangement and a shifted 8-core, where each configuration is separated by the two atomic layers in c-direction. The rise[5] in dislocation climb for the 8-core (upper layer) means that a gallium vacancy is inherent to the structure but not obvious. The next stable structure is less subtle, and represents a 5/7-core with a gallium atom removed at the homonuclear bond site, while the next higher structure features the same strategy as the lowest energy structure, but with a 4-core, 5/7-core combination. Similarly the (1b) core uses a combination of 4 core and 5/7 core with the half 5/7 core shifted one lattice site higher than the (1a) case, making up the GaN pair difference.

---

[3] As a result relative shifts in climb direction can be observed easily.
[4] one more Ga than N atom per simulated cell
[5] dislocation climb will be referenced here as "upwards" if it results in the addition of atoms in the core and "downwards" if they are removed. However, the general trend towards filled cores hinders a downward movement.

(a) Edge dislocation core (-1a)

(b) Edge dislocation core (-1b)

(c) Edge dislocation core (1a)

(d) Edge dislocation core (1a)+0.11eV

(e) Edge dislocation core (1a)+0.28eV

(f) Edge dislocation core (1b)
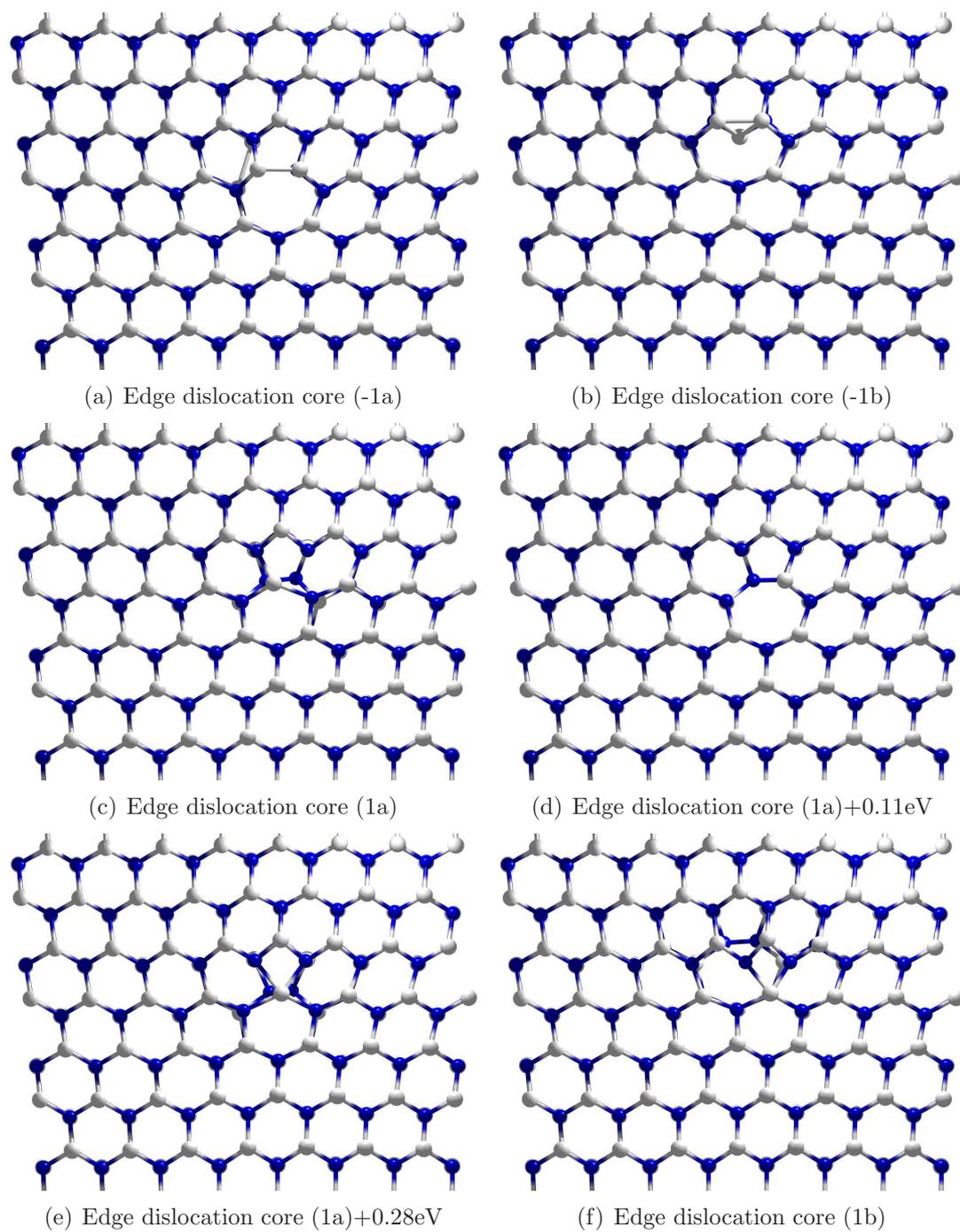
Figure 6.6: Ball and stick model for stable non-stoichiometric edge dislocation core structures.

6.4   Conclusion

The global optimisation performed on the edge dislocation core was able to find all core configurations for the stoichiometric case ((0a) and (0b) core) that have been predicted by previously published results in the literature for GaN [71][83] and InN [72] and observed by experiment[84]. The 5/7 core is predicted here (SW interatomic potentials) to be the most stable configuration, while 8- and 4- core have nearly identical structural energies, with low diffusion barriers separating the two. More complex core configurations have been found in addition, but are expected to be much less favourable, such as the double 5/7 core (figure 6.5(b)). However very similar structures were proposed as configurations for mixed dislocation cores[28][85], where a double 5/7 core becomes a double 5/6-core by the addition of a screw displacement field. Additional optimisation runs ((22) and (33)) have been performed and shown that the dislocation climb energy barrier is low compared to the gain in structural energy by forming full dislocation cores away from the original randomisation volume. This suggests that open cores are generally unfavourable making the additional optimisation results redundant.

The implications of dislocation climb and slip movement were discussed. Dislocation climb movement effectively reduces the amount of optimisation procedures for each stoichiometry to two cases, while slip movement has caused the appearance of anomalous annihilation structures in the optimisation distributions due to the very low diffusion barrier between the 4- and 8-core. However, it stands to reason that the dislocation slip behaviour can be changed by adjusting the geometry of the simulated supercell. This procedure has not been attempted, but should not pose a problem if a more general dislocation displacement field is used, as suggested by Lehto et al.[67].

Two non-stichiometric cases have been investigated, the (1) and (-1) cores. While the (-1) lowest energy structures feature usual 5/7- and 8-cores with N vacancies, the (1) core tends to contain split 5/7- and 4-cores. This can be seen as splitting the core along the c-direction and moving one part along the climb direction or as removing one gallium atom at the homonuclear bonding site at the 5/7 core and resulting slight relaxations of the surrounding atoms.

The SW type interatomic potentials favour the stoichiometric cores, where the

5/7-core is the most stable and the (0b),(1a) and (-1a) are all of comparable energy. However, in order to compare the stability of structures with varying stoichiometries, electronic structure theories such as DFT need to be employed to construct more meaningful chemical potentials. In addition the relative energy for the (0) cores have already been compared in the litearture by Lymperakis et al. [71] and it can be seen that the ordering of formation energy shifts to a favoured 8-core when DFT is employed. This is possibly due to the more complex bonding arrangement with dangling bonds at the 3-fold coordinated centre atoms.

CHAPTER 7

Summary, conclusion and future work

The question of the influence of dislocations on a material is not easily answered, mainly due to the number of unknown conditions such as the effects of growth conditions on the creation of dislocations and the dislocation core composition itself. In addition their intrinsic geometry results in the inability to model single dislocations in an otherwise bulk material, which results in large models and computationally costly investigations. While the long range strain field of a dislocation can be accurately described by elasticity theory, the modelling of the dislocation core is more complex and requires the application of atomistic theories. The dislocation core structure can diverge strongly from the surrounding bulk and depends on a wide range of influences such as growth conditions, that govern the relative chemical potentials of the constituents. In addition and as a result of the nature of dislocation geometry, it is necessary to model a considerable amount of atoms per simulation to provide low external influences, making slow but accurate electronic theories impractical, even with currently available computing resources. This work was aimed at solving the questions of the influence of dislocation cores on the material by performing a systematic investigation into a wide range of conditions. This was done by implementing a multi-scale approach, where an initial structural investigation was performed by a fast interatomic potentials method and a subsequent less inaccurate density functional theory calculation on the resulting optimised structures.

The systematic approach was realised by implementing a global search method that performs a stochastic sampling of the search space and can reliably present global minima for the chosen cost-functions. Although other global search methods can be faster, it is feasible to use such a "brute force" method, as the implementation of a very fast interatomic potentials method as cost function cannot be circumvented. Since no assumptions are made of the core structure previous to optimisation, any condition inside can be simulated. This fact was used to extend this method to optimise screw dislocation cores for a large range of stoichiometries. In this manner core structures were found, presented here and published[86] that, to my knowledge, have not been proposed in the literature previously, such as the (11) core. The range of stable cores could then be effectively be compared in relative stability by using the most stable structures in DFT calculations. Similar comparisons of core stoichiometries in GaN have been performed previously by Northrup et al. [77], however the core structures used there were not optimised previously. In addition results for GaN were translated to InN with the assumption that no significant change in core geometry should occur and similar DFT calculations were performed, concluding that relative core energies are more important than possible growth condition for determining the most stable core in InN due to the lower InN enthalpy.

The applicability of the dislocation core search for a wide variety of conditions was demonstrated not only for the screw dislocation but also for the edge dislocation. While a wide range of cores for the screw dislocation could be investigated, the edge dislocation showed more complex behaviour, due to more pronounced dislocation slip and the existence of dislocation climb movements during relaxation. Dislocation slip movement did occur for (0b)-type cores during structural relaxation and subsequent core annihilation was found to introduce unwanted features into the optimisation distribution. However, these features were few and could be easily identified and thus did not pose a problem for the optimisation procedure as a whole. In addition possible solutions were presented to circumvent slip movement during relaxation altogether. Also, due to the mechanisms of dislocation climb, edge dislocation cores can be filled or emptied by a change in position, which results in a the appearance of same core structures in various optimisation procedures. However, this trend was found to effectively reduce the

amount of structures to be investigated and can be seen as a positive side effect. Still, no additional optimisations or subsequent DFT calculations were performed as of now. Yet, with the results presented it is shown that the presented approach of investigating core structures is equally applicable if not more effective for the edge case.

Similarly to the edge dislocation core optimisation no further complications are expected for mixed dislocation core studies, since these can be treated as linear combinations of screw and edge dislocation simulations, or even partial disloca-tion cores, provided that reasonable sized simulation supercells can be constructed that contain dislocation cores with minimal external strain influences. In addi-tion preliminary investigations into basal plane screw type dislocation cores have already been performed, but results were not presented within this thesis.

The clearest limitation to the presented method is the necessity of using compu-tationally cheap inter-atomic potential methods for the structural investigation. As a result all found structures are by default only the most stable structures according to the theory used to calculate their total energy. Unfortunately even by drastically reducing the amount of necessary structural relaxation steps, this cannot be prevented as more costly electronic theories are too slow by several or-ders of magnitude and their application would be accompanied by the inclusion of too many constraints on the search space. As a consequence, in this work, only SW-type potentials were used with parameterisations for GaN that have shown a good performance in previous studies of extended defects. The effects of the changing the quality of the parameters within the SW-type potentials method were not tested directly in conjunction with the applied optimisation technique. However, it stands to reason that a slight deviation from the used parameters will only shift relative energies of stable cores in the obtained energy distribution and a subsequent analysis of lowest energy structures with DFT will compensate for this. In addition equally well performing Tersoff type potentials could be used instead of or even in combination with the SW-type potentials to increase the quality of the optimisation results. The quality check of the found optimised cores is performed within this work with DFT in the LDA, which does produce

structural energies well and can be considered an appropriate[1] measure for the chosen structures. However, the investigation into the electronic properties can be improved with the application of GGA or even hybrid functionals.

Another limitation of the chosen randomisation optimisation approach is the exponential scaling of computational time with the number of optimised atoms, which results in the inability to model core structures extending two or more lattice repeats along the line direction. These situations could be solved by using more efficient heuristic global optimisations such as evolution or swarm based optimisation algorithms.

---

[1] The AIMPRO implementation of LDA-DFT produces results with a reasonable tradeoff between quality and computational performance for the quantity of structures and their large sizes.

Symmetry operations on elastic constants

A.1 Transforming elastic constants

Starting from the basic relationship between stress and strain in the uncompressed form:

$$\sigma_{ij} = c_{ijkl}\epsilon_{kl}. \tag{A.1}$$

With isotropic can be defined through invariance with rotation. A general coordinate transformation can be used on the strain:

$$\epsilon'_{ij} = T_{il}T_{jm}\epsilon_{lm} \tag{A.2}$$

$$\sigma'_{ij} = T_{il}T_{jm}\sigma_{lm}, \tag{A.3}$$

where $\mathbf{T}$ is a unitary orthogonal matrix satisfying $T_{il}T_{jl} = \delta_{ij}$. It is then evident with (A.2) and (A.3) and the orthogonality of $\mathbf{T}$ that the elastic constants transform as:

$$c^*_{ijkl} = T_{im}T_{jn}T_{ko}T_{lp}c_{mnop}, \tag{A.4}$$

The interesting bit is now that under a symmetry transformation (i.e. the underlying material does not change properties under such a transformation) the elastic constants remain unchanged as well. Therefore such a coordinate transformation

T^{sym} will yield the relationship

$$c_{ijkl} = T_{im}T_{jn}T_{ko}T_{lp}c_{mnop}. \tag{A.5}$$

This relationship can now be used to find relationships between the elastic constants under the constraints of a symmetry transformation. Therefore it is here prudent to examine some general orthogonal transfomrations. For instance a reflective transformation can be described using the identity matrix ($\mathbf{I}$) and the normal vector ($\mathbf{n}$) to the reflection plane

$$\mathbf{T}^{refl} = \mathbf{I} - 2\mathbf{n}\mathbf{n}^T. \tag{A.6}$$

choosing $n = (\cos\theta, \sin\theta, 0)$ will therefore enable a parameterisation for all reflection planes incorporating the x_3 axis. Therefore $\mathbf{T}$ will take the form

$$\mathbf{T}^{refl}(\theta) = \begin{pmatrix} -\cos(2\theta) & -\sin(2\theta) & 0 \\ -\sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{T}^{refl}(0) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (A.7)$$

where it is useful to recognize that

$$T_{ij}^{refl}(0) = \begin{cases} \delta_{ij} & i \neq 1 \\ -\delta_{ij} & i = 1 \end{cases}. \tag{A.8}$$

Let us similarly define a rotation around the x_3 axis as:

$$\mathbf{T}^{rot,1}(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix}, \quad \mathbf{T}^{rot,3}(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$\tag{A.9}$$

Using the above defined general transformations we can now consider a few special cases of symmetries.

A.2 Cubic symmetry

Cubic symmetry is the most commonly found symmetry in materials, it is described by the possibility of rotations and reflections with respect to the axis of

an orthogonal basis. Therefore all rotations with 90° are symmetries under (A.5) as well as reflections with respect to the surfaces $x_1 = 0$, $x_2 = 0$ and $x_3 = 0$. Therefore with $\mathbf{T}^{refl}(0)$, being symmetry, one can write:

$$c_{ijkl} = T_{im}^{refl}(0)T_{jn}^{refl}(0)T_{ko}^{refl}(0)T_{lp}^{refl}(0)c_{mnop} \tag{A.10}$$

$$c_{ijkl} = (-1)^N \delta_{im}\delta_{jn}\delta_{ko}\delta_{lp}c_{mnop} \tag{A.11}$$

$$c_{ijkl} = (-1)^N c_{ijkl}, \tag{A.12}$$

with N being the amount of ones appearing in the indices i, j, k, l. Therefore, for $N = 1, 3$ c_{ijkl} will be zero. Similarly, using $\mathbf{T}^{refl}(90)$:

$$c_{ijkl} = (-1)^M c_{ijkl}, \tag{A.13}$$

where M is the amount of twos. At this point it is not necessary to consider a reflection with respect to the $x_3 = 0$ surface, as no new relations will be found. This fact is apparent when reminding oneself that if there is an uneven amount of ones and twos in 4 indices all uneven amounts of 3's are covered as well... Considering the relations in (2.72) one can easily see that only the elastic constants with indices $1222, 1333, 1233, 1323, 1322, 1223, 1113, 1112, 2333, 2311, 2113, 2223$ will be set to zero. Resulting in the elastic constant matrix taking the form:

$$
\begin{pmatrix}
c_{1111} & c_{1122} & c_{1133} & 0 & 0 & 0 \\
c_{2211} & c_{2222} & c_{2233} & 0 & 0 & 0 \\
c_{3311} & c_{3322} & c_{3333} & 0 & 0 & 0 \\
0 & 0 & 0 & c_{2323} & 0 & 0 \\
0 & 0 & 0 & 0 & c_{1313} & 0 \\
0 & 0 & 0 & 0 & 0 & c_{1212}
\end{pmatrix}
=
\begin{pmatrix}
c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\
c_{21} & c_{22} & c_{23} & 0 & 0 & 0 \\
c_{31} & c_{32} & c_{33} & 0 & 0 & 0 \\
0 & 0 & 0 & c_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & c_{55} & 0 \\
0 & 0 & 0 & 0 & 0 & c_{66}
\end{pmatrix}.
\tag{A.14}
$$

If now also 90° rotations are considered one can quickly find the remaining relations between the elastic constants. Again only two of the three possible rotations are needed to find all dependencies. Considering therefore the above defined rotations (A.9, $\theta = 90°$), in combination with A.5, we get:

$$c_{1111} = c_{2222} = c_{3333}, \qquad c_{1313} = c_{1212} = c_{2323}, \qquad c_{1122} = c_{1133} = c_{2233} \tag{A.15}$$

One can see now that only 3 independent matrix elements are left, resulting in one of the simplest possible elastic constants. The final form of the elastic constant tensor in a material with cubic symmetry will therefore be

$$
\mathbf{c}_{cub} = \begin{pmatrix}
c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\
c_{12} & c_{11} & c_{12} & 0 & 0 & 0 \\
c_{12} & c_{12} & c_{11} & 0 & 0 & 0 \\
0 & 0 & 0 & c_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & c_{44} & 0 \\
0 & 0 & 0 & 0 & 0 & c_{44}
\end{pmatrix}.
\tag{A.16}
$$

A.3   Isotropy

Taking a further step in simplicity from the cubic system results in an anisotropic material. Here all directions are equal, therefore all rotations are a symmetry transformation, not only 90° ones. Using a lot of tedious arithmaticy we can find one additional dependency for the case c_{1112} under transformation with $\mathbf{T}^{rot,3}(\theta)$:

$$
0 \stackrel{(A.14)}{=} c_{1112} = T_{1i}^{rot,3}(\theta)T_{1j}^{rot,3}(\theta)T_{1k}^{rot,3}(\theta)T_{l2}^{rot,3}(\theta)c_{ijkl}
\tag{A.17}
$$

$$
0 = 2(\cos^2\theta - \sin^2\theta)C_{66} - \cos^2\theta C_{11} + \sin^2\theta C_{22} + (\cos^2\theta - \sin^2\theta)C_{21}
$$

$$
\Rightarrow c_{66} = c_{44} = \frac{1}{2}(c_{11} - c_{12}),
\tag{A.18}
$$

where (A.15) has been used. This step has now reduced the amount of independent constants to two.

A.4   Hexagonal symmetry or transverse isotropy

The hexagonal symmetry is usually being described by transverse isotropy, in this case the material is isotropic in only 2 dimensions. We will choose here the x_3 axis to be the axis of symmetry, i.e. the axis around which all rotations will be symmetries and the axis that is part of all possible reflection planes.

Therefore $\mathbf{T}^{refl}(0)$, $\mathbf{T}^{refl}(90)$, and $\mathbf{T}^{rot,3}(\theta)$ are symmetry transformations. This means all considerations up to (A.14) are incorporated and we can start from there.

Using all elastic constants with indices 1222, 1333, 1233, 1323, 1322, 1223, 1113,

1112 equal 0, we can write

$$\mathbf{T}^{refl}(\theta) = \begin{pmatrix} -\cos(2\theta) & -\sin(2\theta) & 0 \\ -\sin(2\theta) & cos(2\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{A.19}$$

$$= \begin{pmatrix} -\cos(\theta) & -\sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{A.20}$$

$$= \mathbf{T}^{refl}(\theta/2)\mathbf{T}^{rot}(\theta) \tag{A.21}$$

for 1112 :
$$0 = 2c'C_{66} - \cos^2(\theta)C_{11} + \sin^2(\theta)C_{22} + c'C_{21} \tag{A.22}$$

for 1113 :
$$0 = \sin^2(\theta)C_{24} + \cos^2(\theta)C_{14} + 2\cos^2(\theta)C_{56} \tag{A.23}$$

for 1222 :
$$0 = -\sin^2(\theta)C_{11} + \cos^2(\theta)C_{22} - \cos(\theta)C_{12} - 2\cos(\theta)C_{66} \tag{A.24}$$

for 1333 :
$$0 = \cos(\theta)C_{35} + \sin(\theta)C_{34} \tag{A.25}$$

for 1233 :
$$0 = -C_{13} + C_{23} \tag{A.26}$$

for 1323 :
$$0 = -C_{55} + C_{44} \tag{A.27}$$

for 1322 :
$$0 = \cos^2(\theta)C_{24} + \sin^2(\theta)C_{14} - 2\cos^2(\theta)C_{56} \tag{A.28}$$

for 1223 :
$$0 = -\cos^2(\theta)C_{14} + \cos^2(\theta)C_{24} - \cos^2(\theta)C_{56} + \sin^2(\theta)C_{56}, \tag{A.29}$$

As a result of the above relations for the elastic constants, the tensor takes the form

$$\mathbf{c}_{hex} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{pmatrix} \tag{A.30}$$

with $c_{66} = \frac{1}{2}(c_{11} - c_{12})$. This will set the number of independent elastic constants to 5.

APPENDIX B

List of total energy distributions

This appendix lists all performed core optimisations of the screw dislocation in the c-direction. A wide selection of core stoichiometries have been considered and an optimisation procedure has been done for each. Core stoichiometries have been labeled according to the amount of vacancies in the core, where the first digit represents Ga vacancy and the second N. Hence (00) represents a full dislocation core and (-11) represents a core with an additional Ga and a N vacancy. Details on the optimisation procedure can be found in section 4.2, while section 5 discusses the lowest energy structure of all optimisations presented here. Additional structural plots for the lowest energy structures can be found in appendix C.

An overall summary of the optimisation parameters can be found in table B.1. For each structure the number of performed optimisation steps is listed as well as the count of how often the lowest energy unique structure has been found. This lowest structure count however does not include bin counts for the same unique structrue that has been reformed at different positions and hence undergone an energy symmetry breaking due to the boundary conditions (see section 4.2.2 for details) and can therefore be misleading. However, this can still be a measure of convergence in combination with the distribution of the lowest energy strucures.

Table B.1: Summary of all screw dislocation optimisations. Energies are given in eV.

Core	steps	E_0 bin	E_0	E_{lowest}	average	variance	rel. E
(00)	87444	25	-3812.45	-3781.19	-3801.32	18.0	0.00
(01)	100000	43	-3807.25	-3781.22	-3797.71	17.2	0.86
(02)	100000	22	-3802.52	-3776.29	-3793.57	15.4	1.25
(10)	100000	11	-3807.94	-3779.45	-3798.29	20.2	0.17
(11)	81156	6	-3804.04	-3777.29	-3794.57	18.6	-0.27
(12)	100000	160	-3798.95	-3774.84	-3790.31	16.0	0.48
(20)	100000	19	-3802.72	-3777.93	-3793.90	16.9	1.05
(21)	100000	7	-3798.72	-3774.99	-3790.22	15.9	0.71
(22)	100000	103	-3794.38	-3769.36	-3786.11	14.1	0.71
(33)	100000	13	-3784.26	-3765.17	-3777.11	9.3	2.15
(0-1)	56390	4	-3814.59	-3785.95	-3804.50	16.8	2.21
(-10)	54769	2	-3813.67	-3785.61	-3803.85	16.2	3.12
(-12)	59263	15	-3804.36	-3780.93	-3796.31	14.4	3.75
(-66)	61882	9	-3782.35	-3772.59	-3780.46	1.3	30.10
(-11)	100000	108	-3810.34	-3783.52	-3800.37	16.3	2.11
(1-1)	100000	303	-3811.98	-3783.55	-3801.65	20.2	0.47
(-22)	100000	104	-3806.33	-3781.12	-3798.17	13.4	6.12
(2-2)	100000	2	-3809.06	-3783.64	-3800.18	16.2	3.39
(-21)	99998	57	-3810.67	-3785.27	-3802.08	14.6	6.12
(2-1)	99998	13	-3806.07	-3780.91	-3797.26	16.7	2.04
(1-2)	99998	25	-3814.50	-3786.53	-3804.39	18.7	2.29
(-32)	99998	2	-3805.69	-3783.34	-3799.09	10.5	11.11
(2-3)	99998	3	-3812.17	-3786.59	-3802.80	15.2	4.62
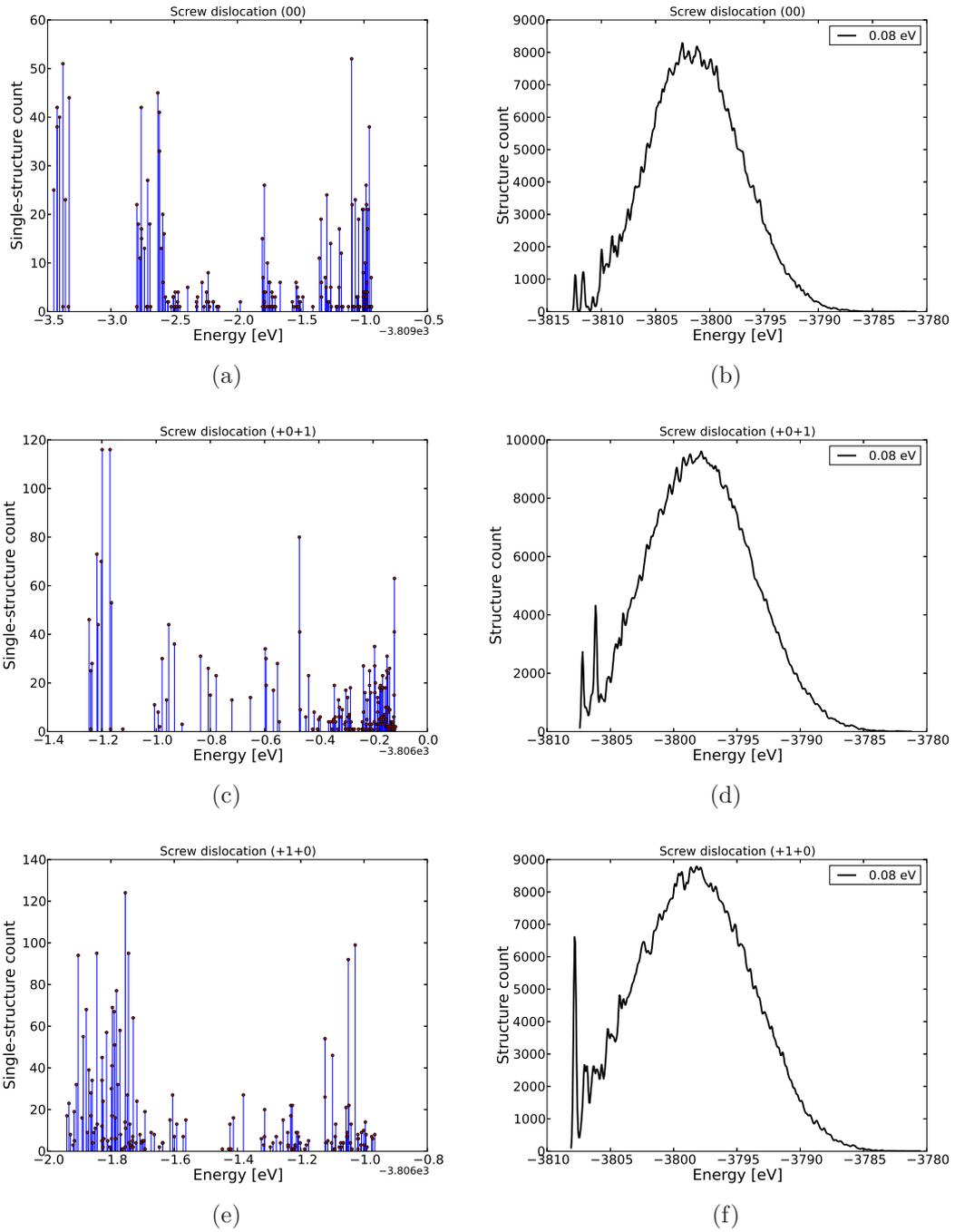
Figure B.1: Optimisation distributions for (00),(01),(10) cores
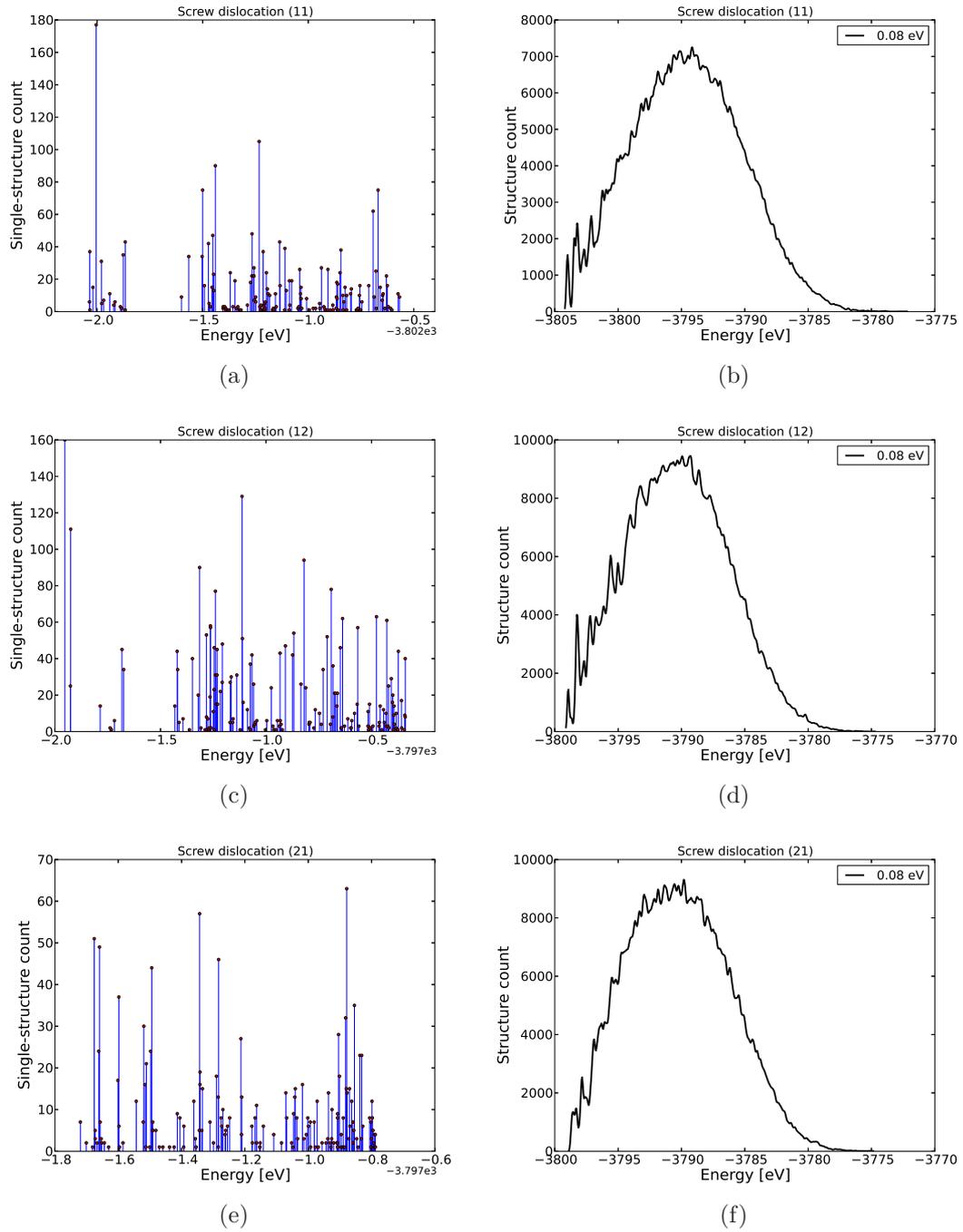
Figure B.2: Optimisation distributions for (11),(12),(21) cores
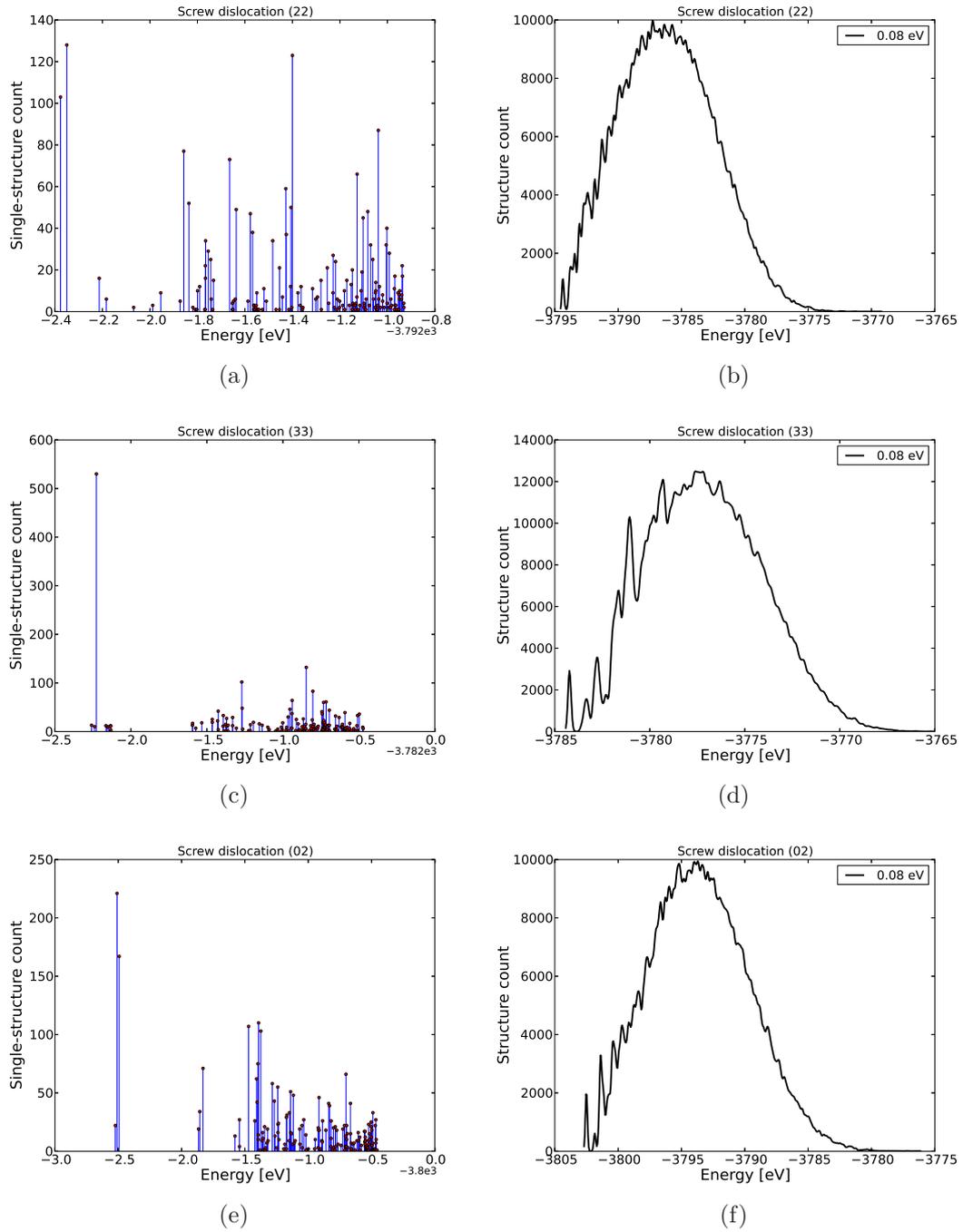
Figure B.3: Optimisation distributions for (22),(33),(02) cores
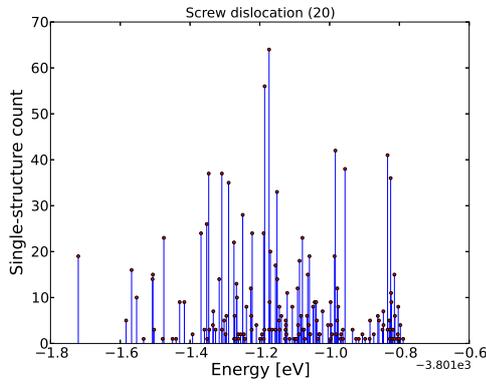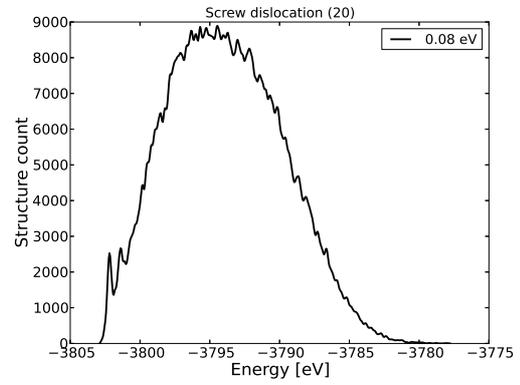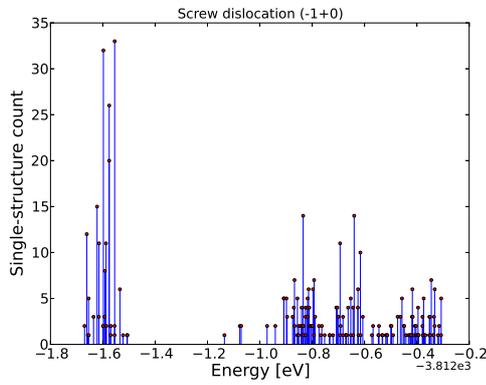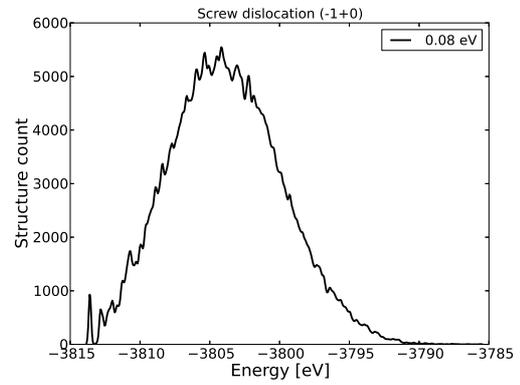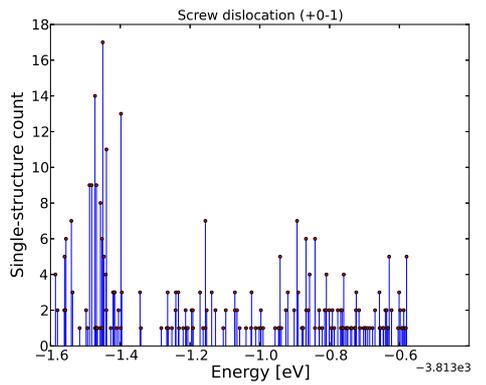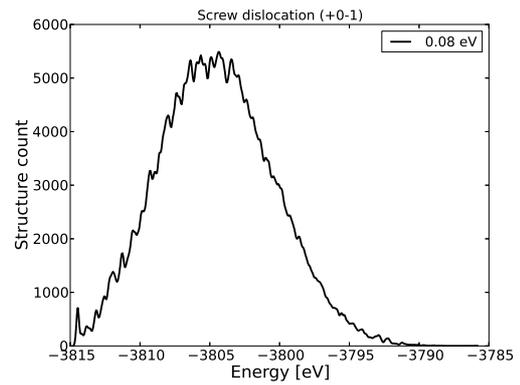
Figure B.4: Optimisation distributions for (20),(-10),(0-1) cores
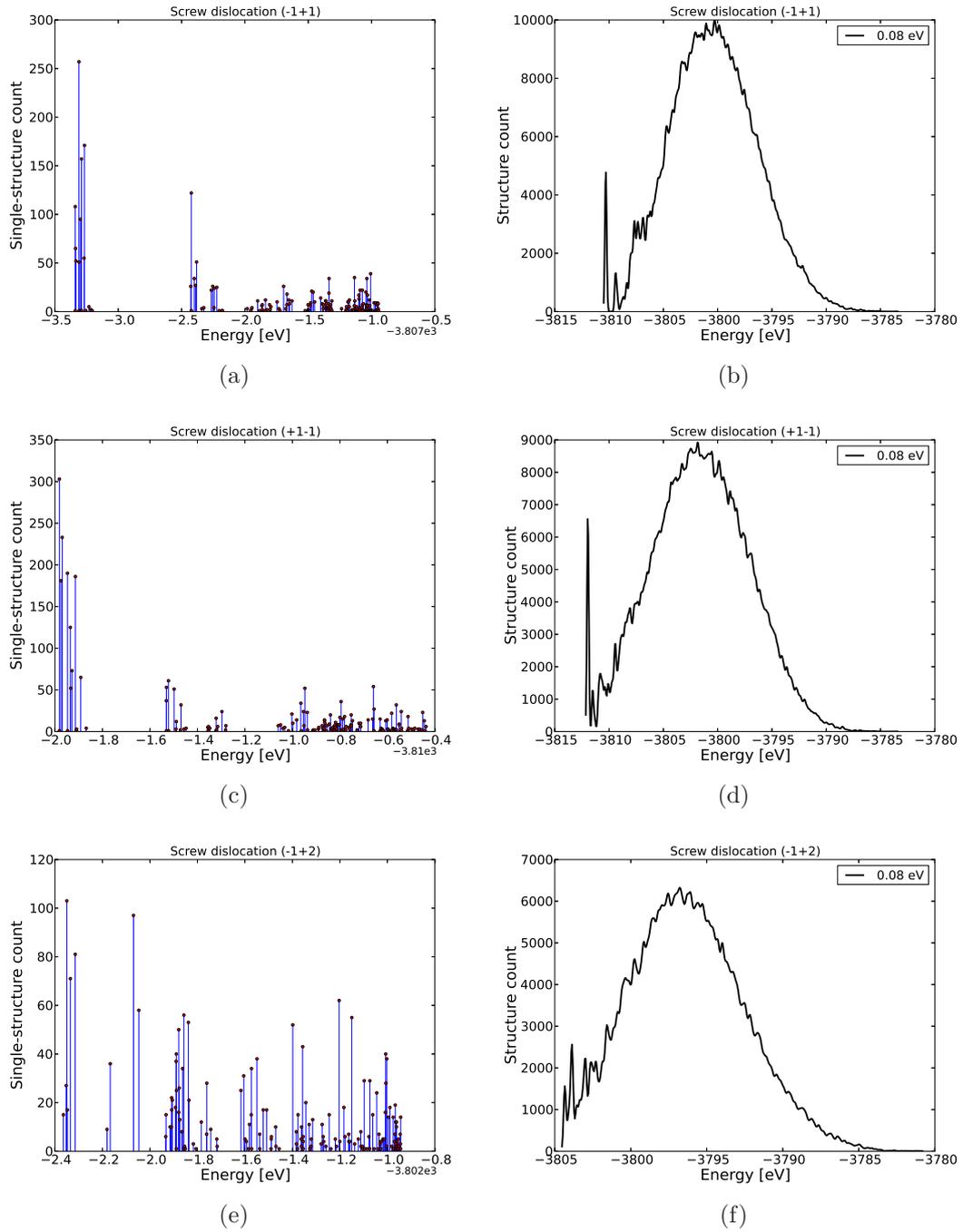
Figure B.5: Optimisation distributions for (-11),(1-1),(-12) cores
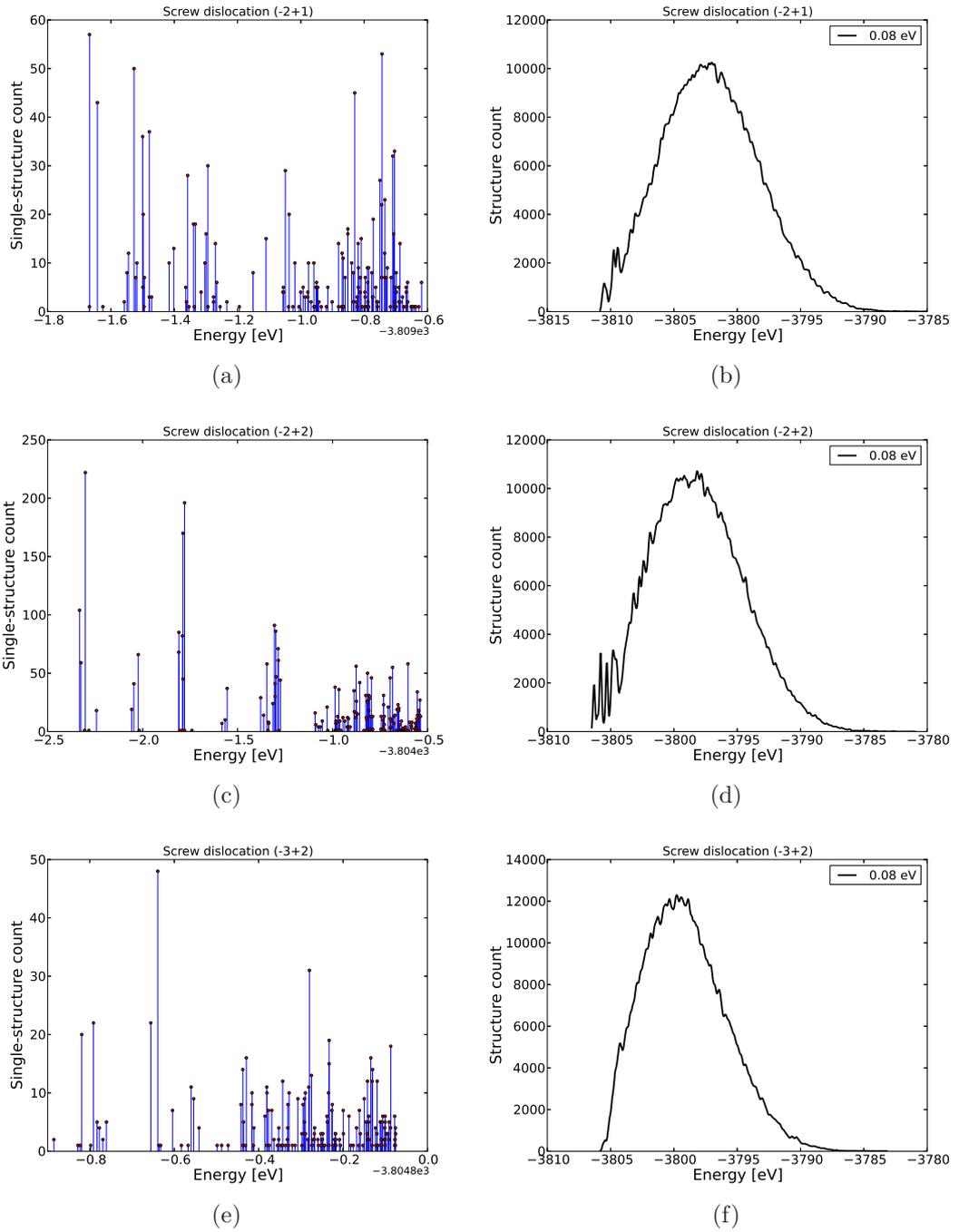
Figure B.6: Optimisation distributions for (-21),(-22),(-32) cores
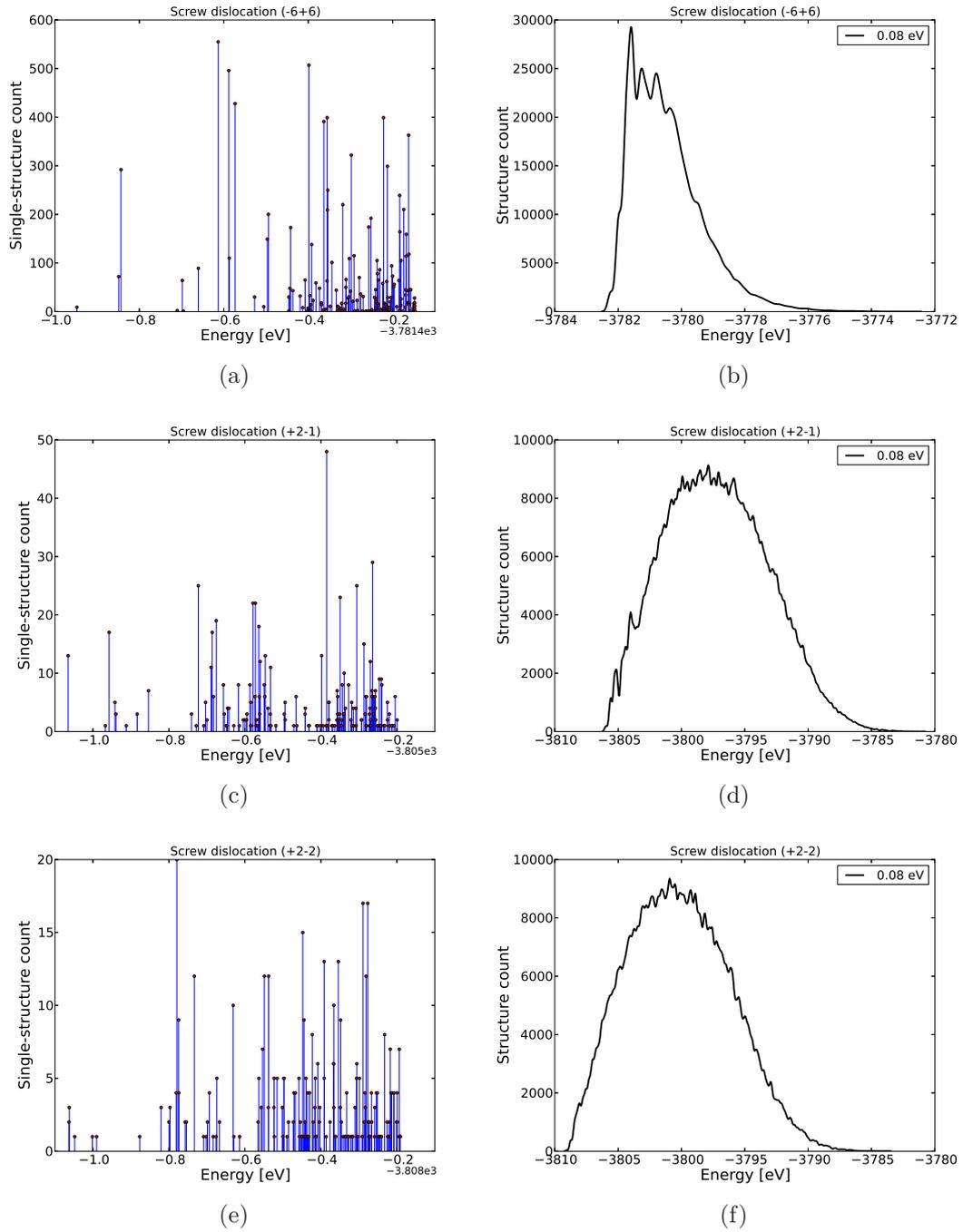
Figure B.7: Optimisation distributions for (-66),(2-1),(2-2) cores

(a)

(b)

(c)

(d)
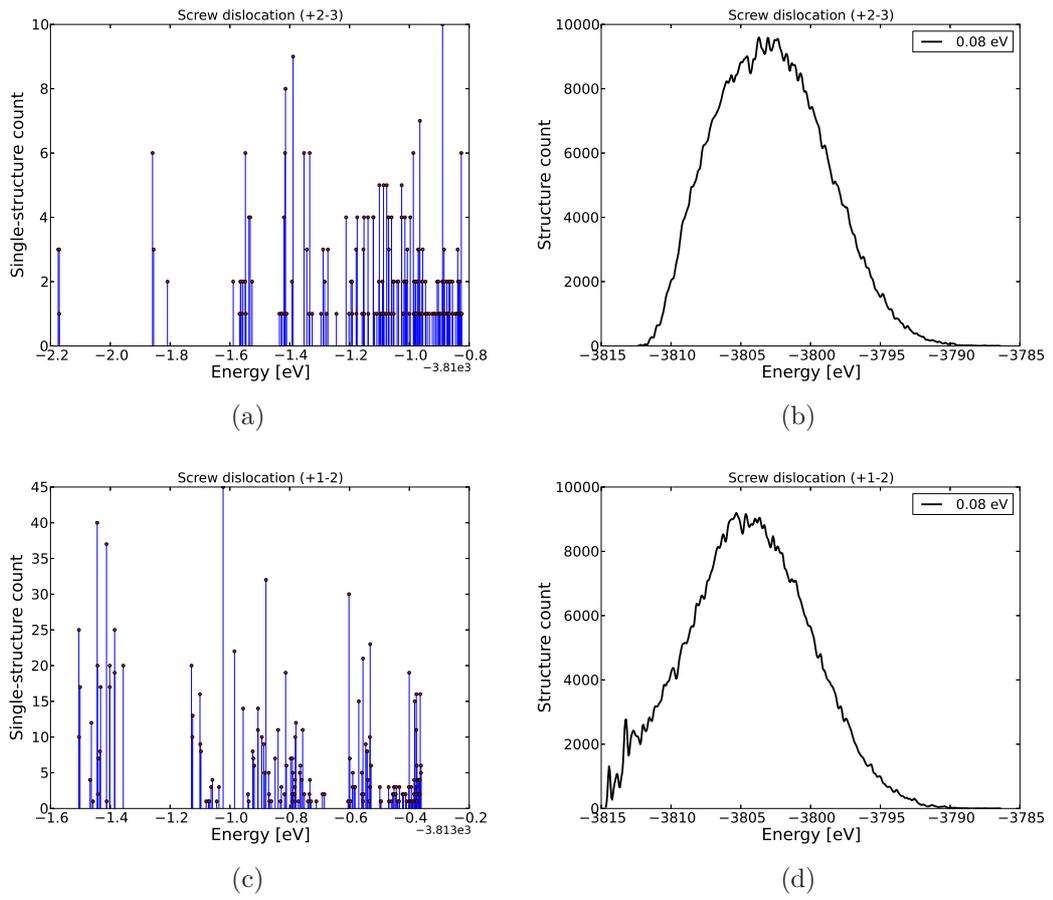
Figure B.8: Optimisation distributions for the (2-3) and (1-2) cores

131

APPENDIX C

Optimized screw dislocations

In this appendix structural details are given for all lowest energy screw dislocations taken from the global optimisations for a wide range of core stoichiometries. Details of the optimisations can be found in appendix B. An overall discussion of these structures can be found in section 5.

For each dislocation core a structural representation is given by a ball-and-stick model, in addition the strain energy density is plotted as a function of distance from the dislocation core centre for each atomic position and as an average for all angles at a specific radius. The energy density average is used to determine the dislocation core radius R_c through integration. Details can be found in section 4.2.2. Bonds in the ball-and-stick model have been drawn with a length cutoff according to atomic type and the SW parameterisation; This is done in the same manner as the bond angle/distance statistic accumulation in table 5.1. Nitrogen positions are represented by blue spheres while gallium is shown with white spheres. The dislocation core centre for the optimised core is marked by a green dot, however the centre of the image is in all cases the original dislocation centre according to the unoptimised dislocation quadrupole. As a result the presented part of the total structure is constant.
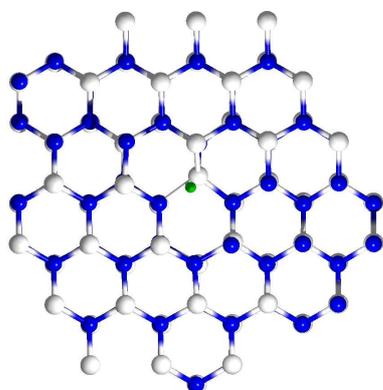
(a) (01) Core structure
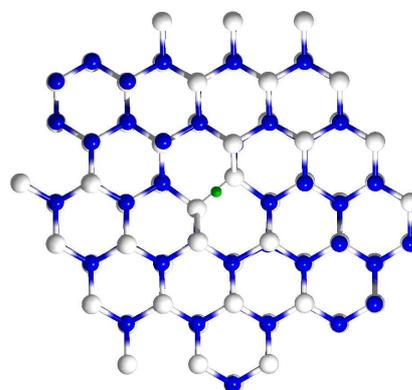
(d) (02) Core structure
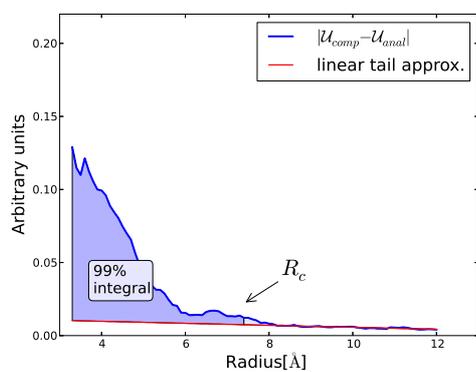
(b) (01) R_c retrieval
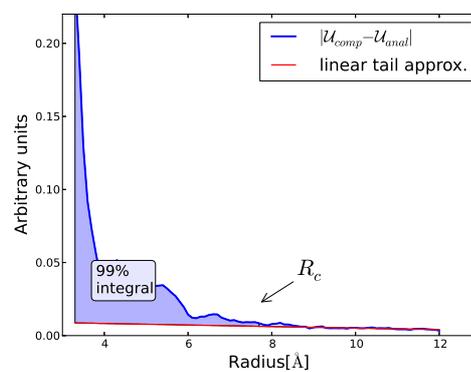
(e) (02) R_c retrieval
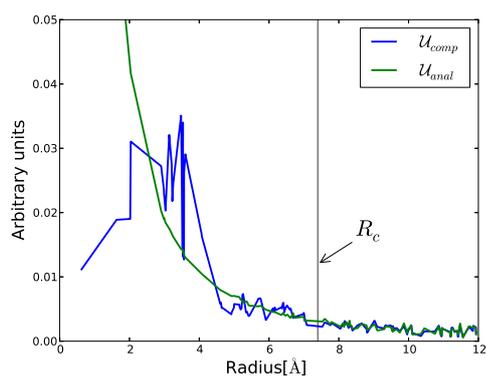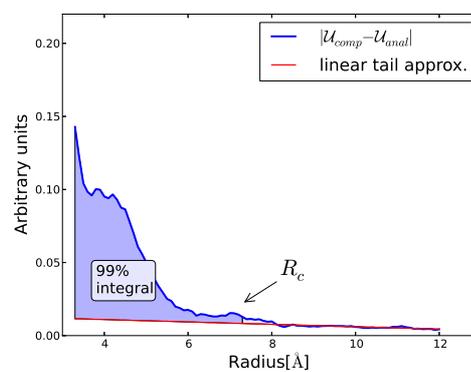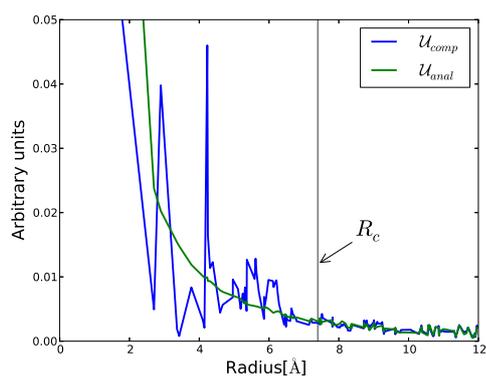

(c) (01) Strain energy
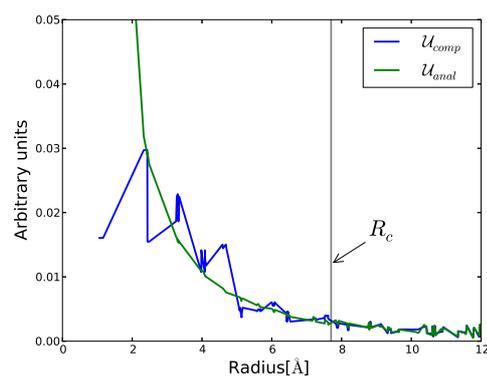
(f) (02) Strain energy

Figure C.1: Core structure and radius for (01) and (02)
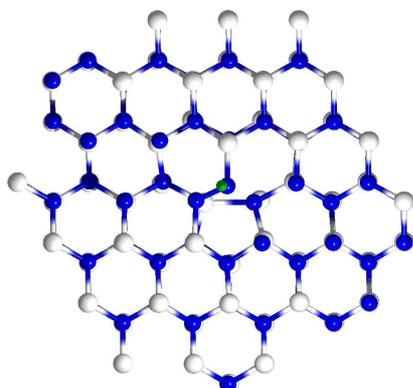
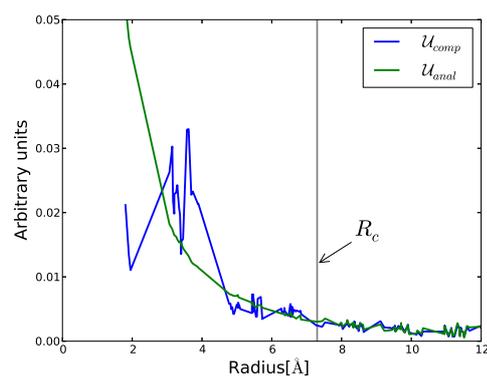(a) (10) Core structure
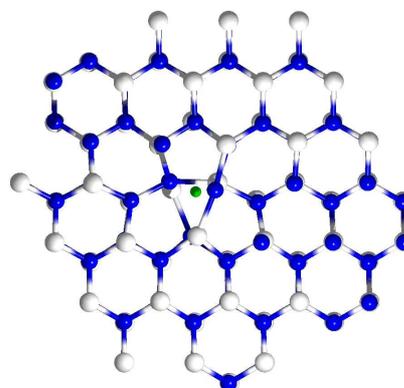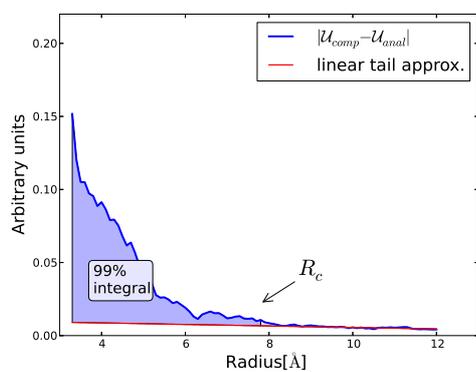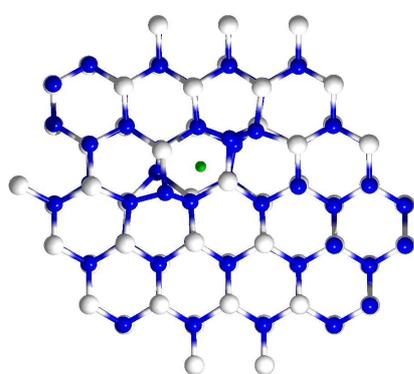
(d) (11) Core structure
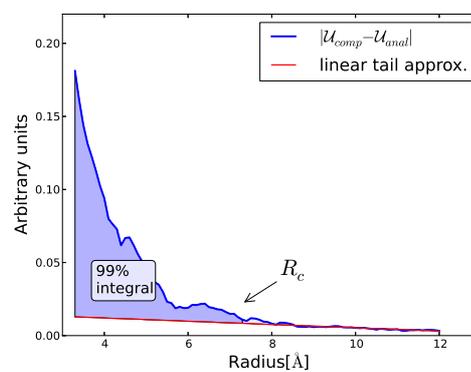
(b) (10) R_c retrieval
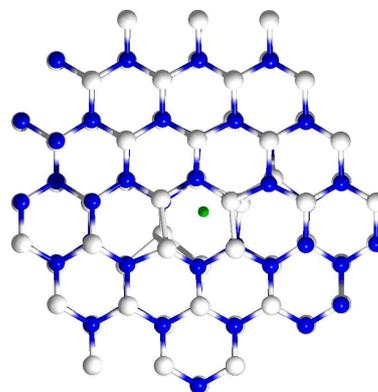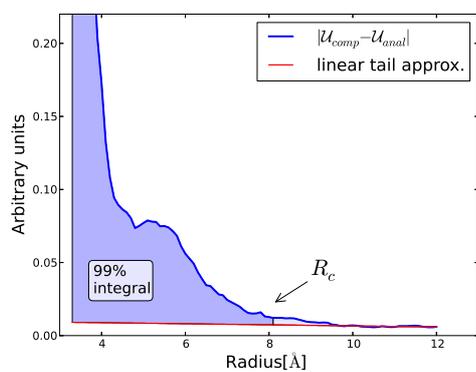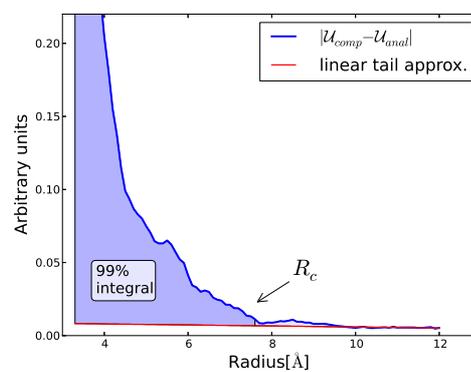
(e) (11) R_c retrieval
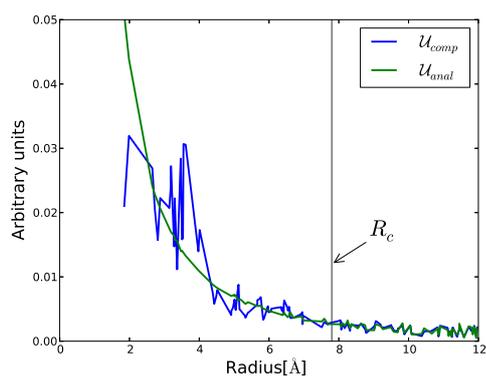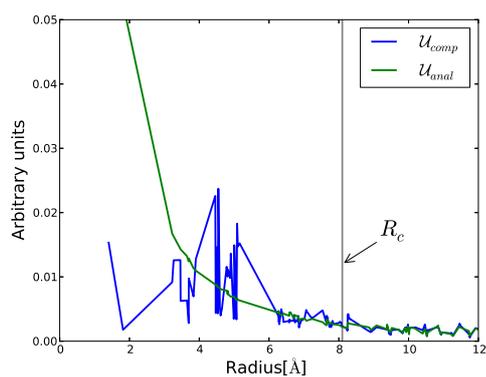
(c) (10) Strain energy
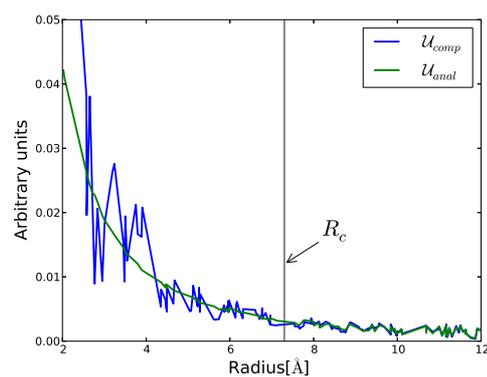
(f) (11) Strain energy
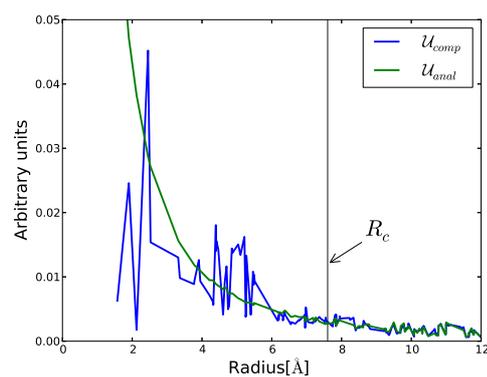
Figure C.2: Core structure and radius for (10) and (11)
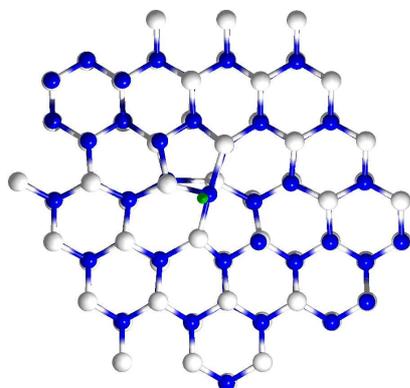
(a) (12) Core structure
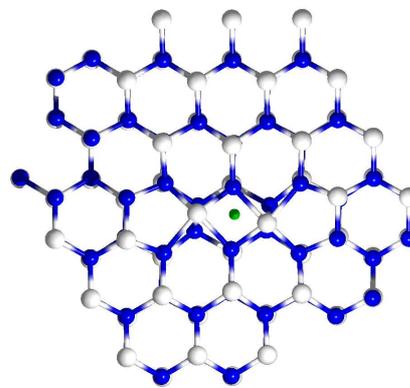
(d) (20) Core structure
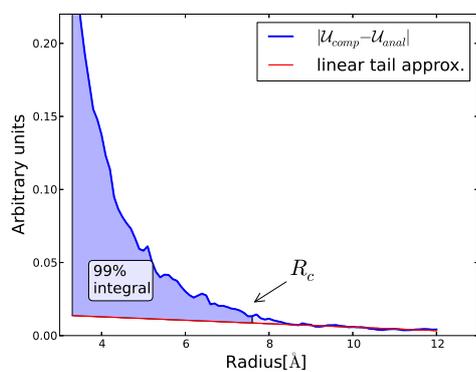
(b) (12) R_c retrieval
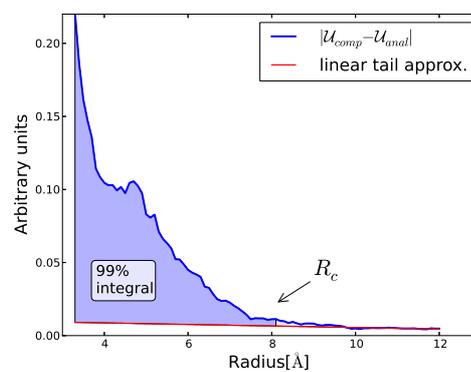
(e) (20) R_c retrieval
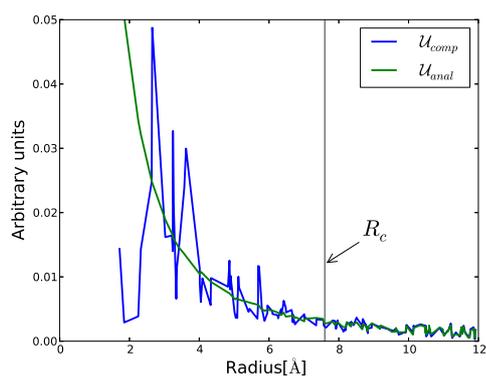
(c) (12) Strain energy
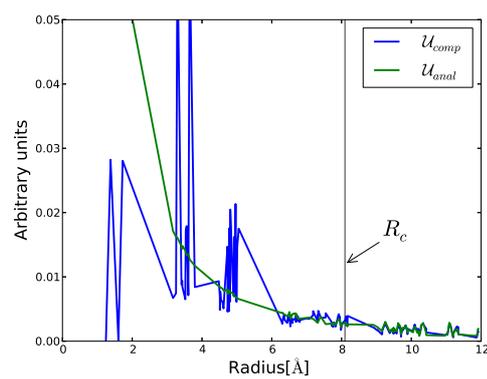
(f) (20) Strain energy

Figure C.3: Core structure and radius for (12) and (20)

135
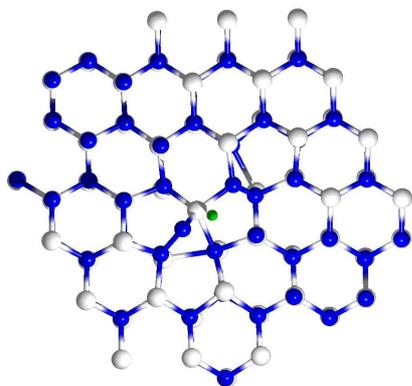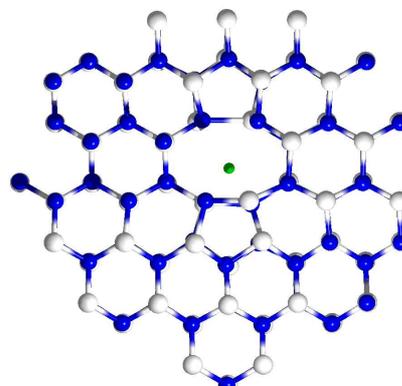
(a) (21) Core structure
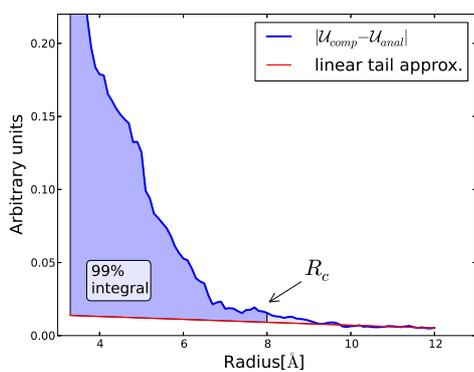
(d) (22) Core structure
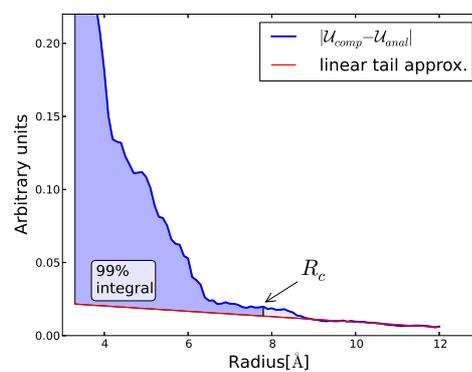
(b) (21) R_c retrieval

(e) (22) R_c retrieval
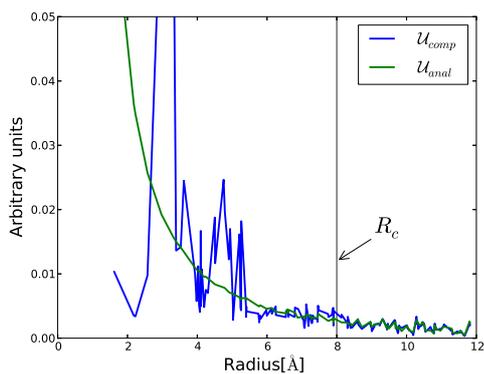
(c) (21) Strain energy
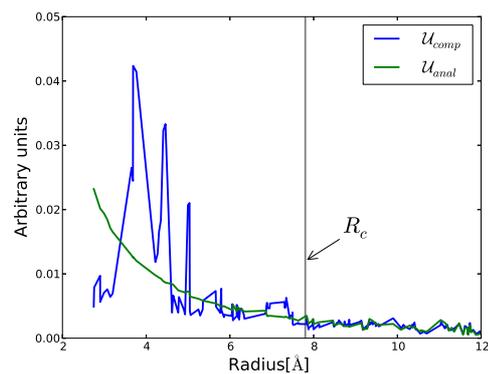
(f) (22) Strain energy
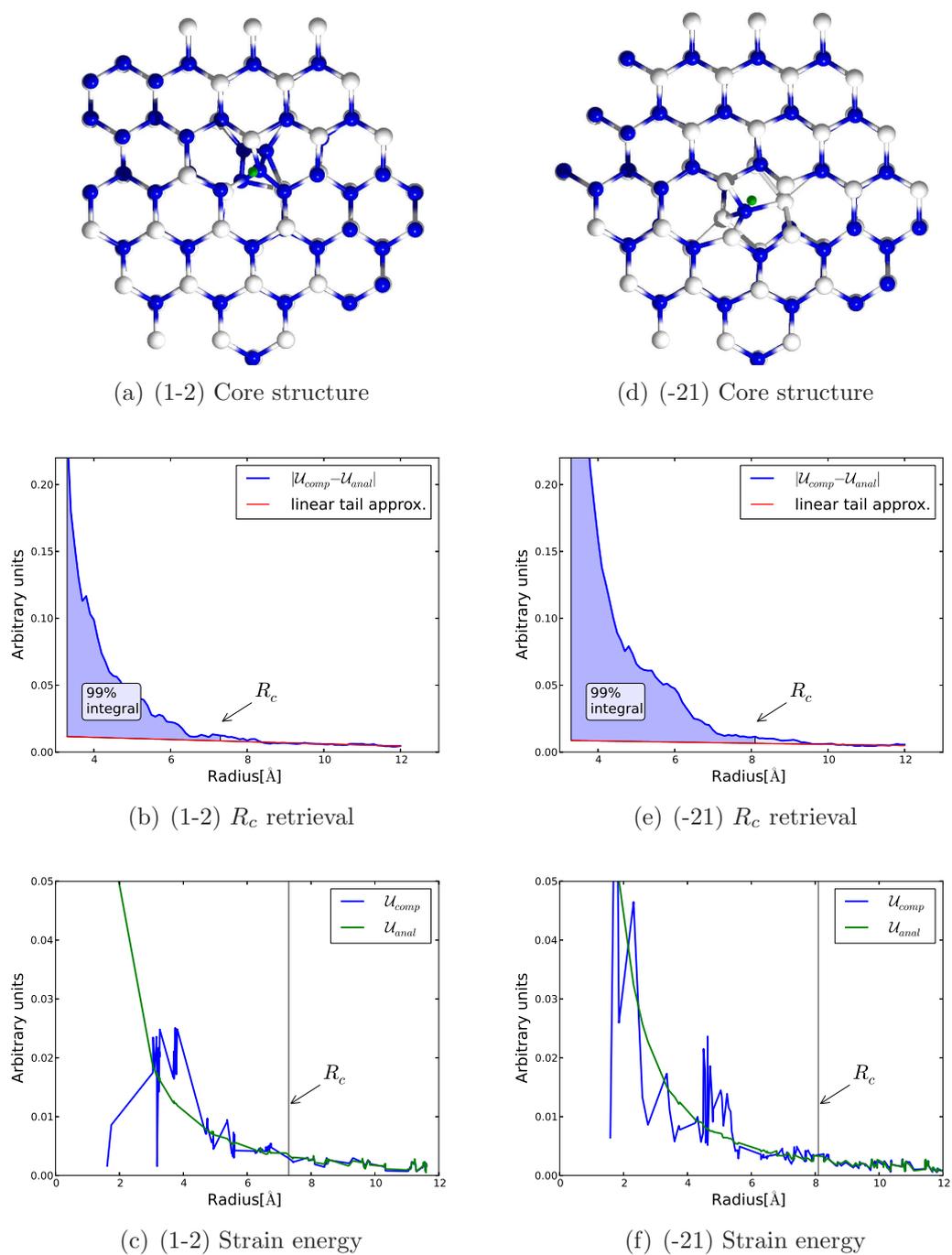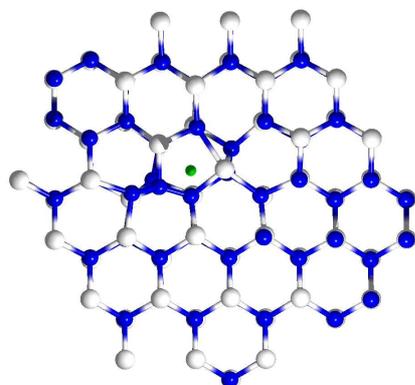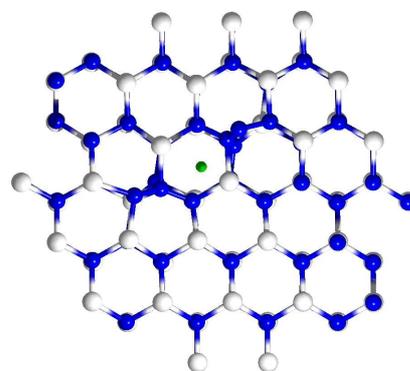
Figure C.4: Core structure and radius for (21) and (22)

136
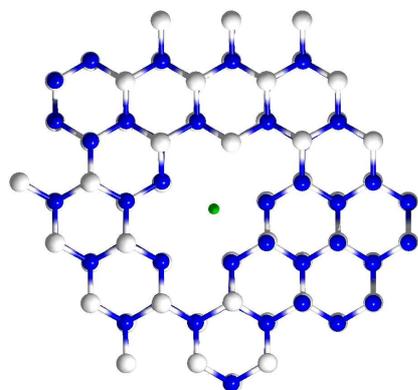
(a) (33) Core structure
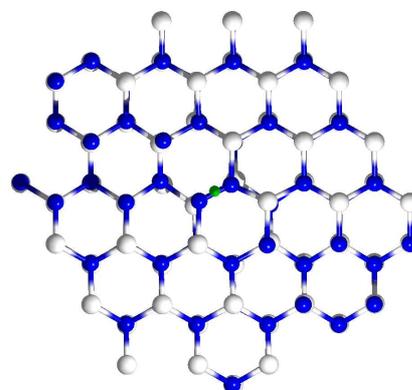
(d) (00) Core structure
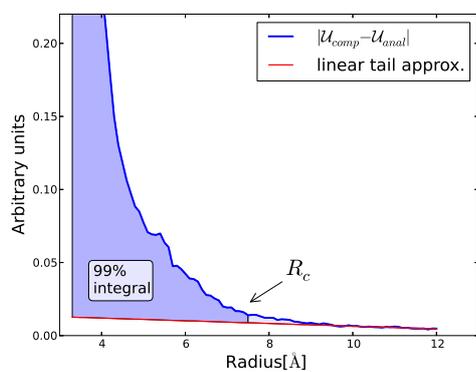
(b) (33) R_c retrieval
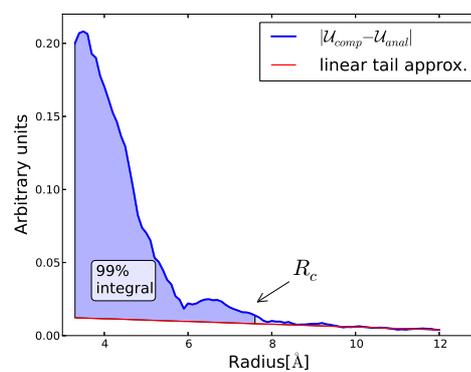
(e) (00) R_c retrieval
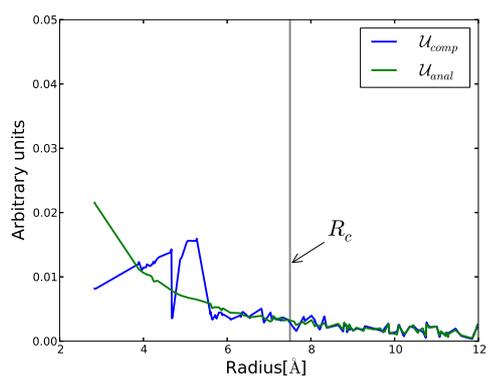
(c) (33) Strain energy
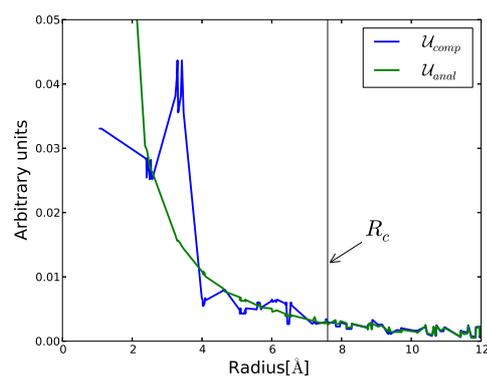
(f) (00) Strain energy

Figure C.5: Core structure and radius for (33) and (00)
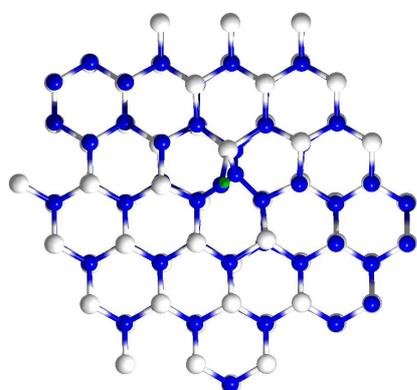
(a) (1-1) Core structure
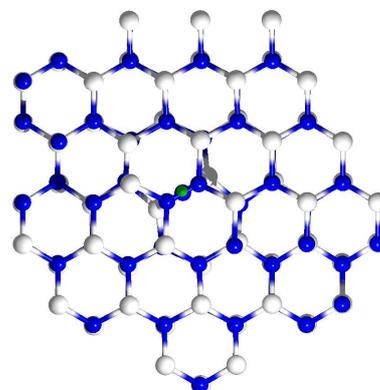
(d) (0-1) Core structure
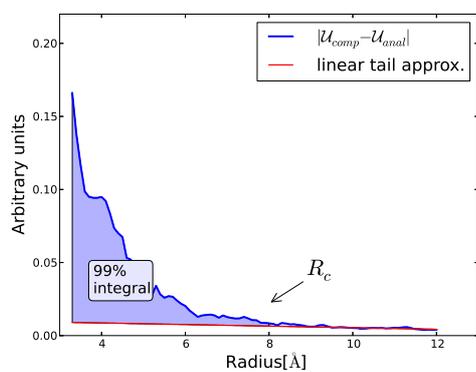
(b) (1-1) R_c retrieval
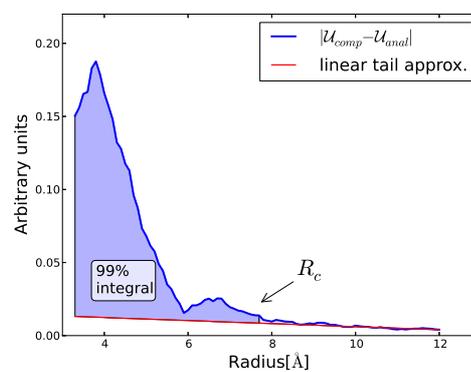
(e) (0-1) R_c retrieval
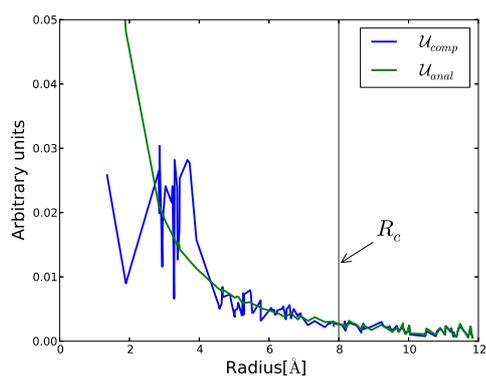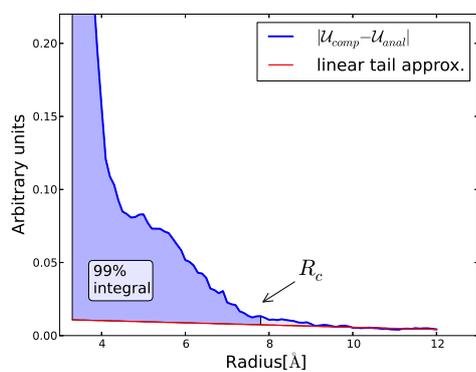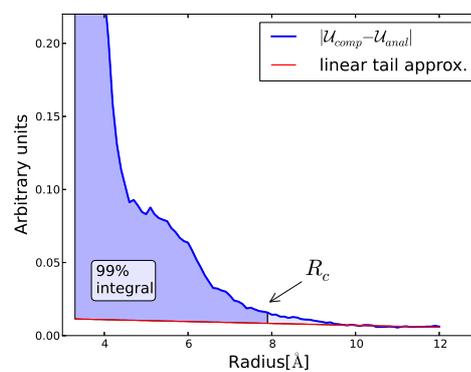

(c) (1-1) Strain energy
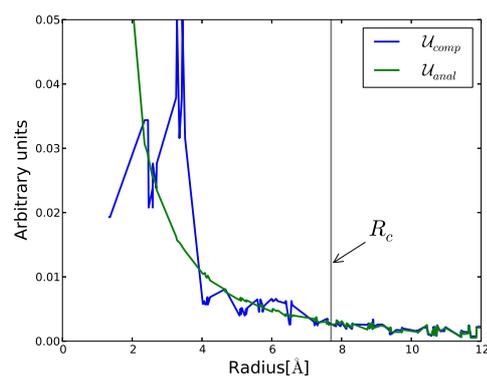
(f) (0-1) Strain energy

Figure C.6: Core structure and radius for (1-1) and (0-1)
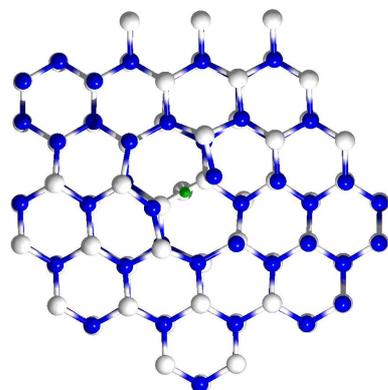
(a) (-10) Core structure
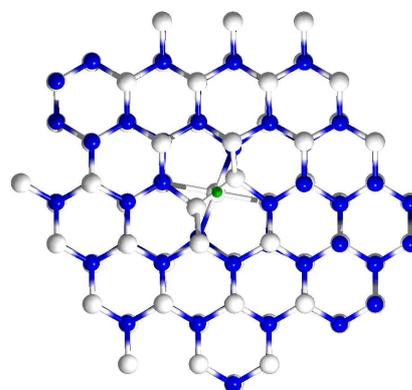
(d) (-12) Core structure
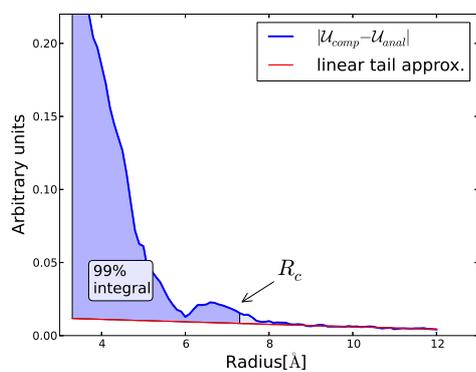
(b) (-10) R_c retrieval
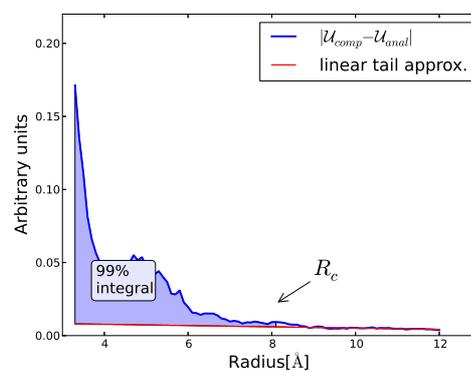
(e) (-12) R_c retrieval
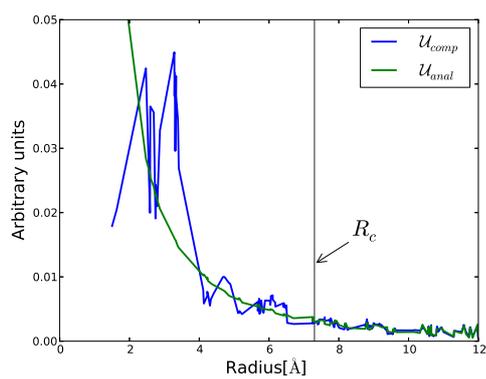
(c) (-10) Strain energy
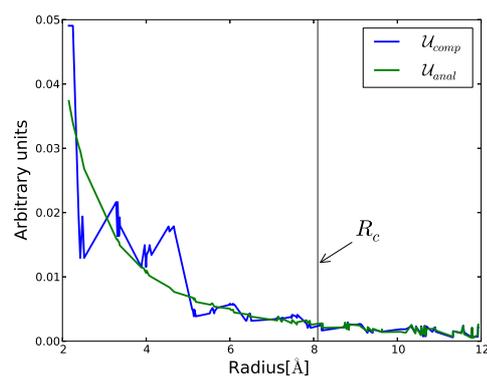
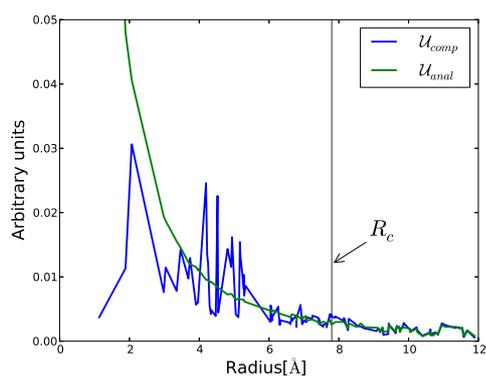(f) (-12) Strain energy

Figure C.7: Core structure and radius for (-10) and (-12)
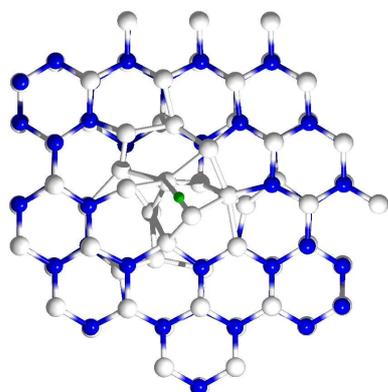
(a) (-66) Core structure
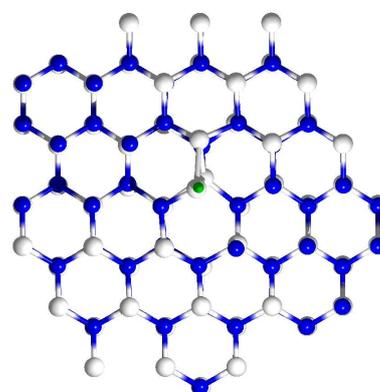
(d) (-11) Core structure
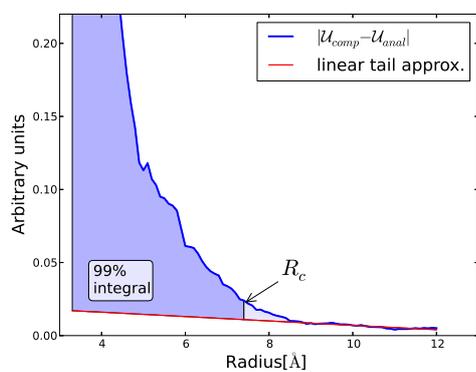
(b) (-66) R_c retrieval

(e) (-11) R_c retrieval

(c) (-66) Strain energy

(f) (-11) Strain energy

Figure C.8: Core structure and radius for (-66) and (-11)

(a) (2-2) Core structure

(d) (-22) Core structure

(b) (2-2) R_c retrieval

(e) (-22) R_c retrieval

(c) (2-2) Strain energy

(f) (-22) Strain energy

Figure C.9: Core structure and radius for (2-2) and (-22)

141

(a) (1-2) Core structure

(d) (-21) Core structure

(b) (1-2) R_c retrieval

(e) (-21) R_c retrieval

(c) (1-2) Strain energy

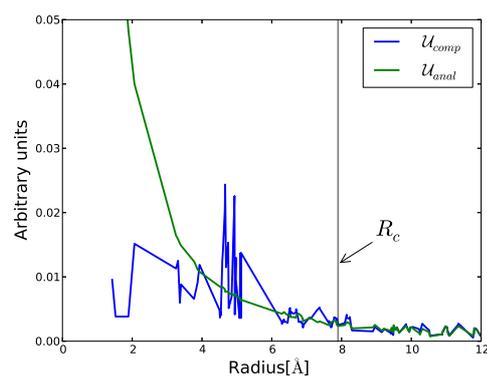(f) (-21) Strain energy

Figure C.10: Core structure and radius for (1-2) and (-21)

(a) (2-1) Core structure
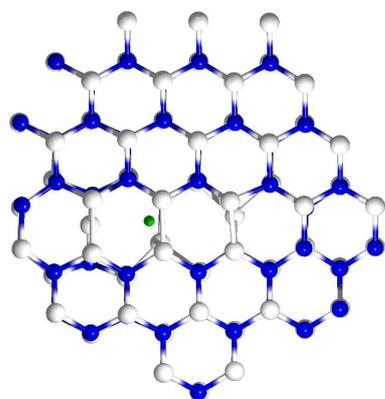

(d) (2-3) Core structure

(b) (2-1) R_c retrieval

(e) (2-3) R_c retrieval
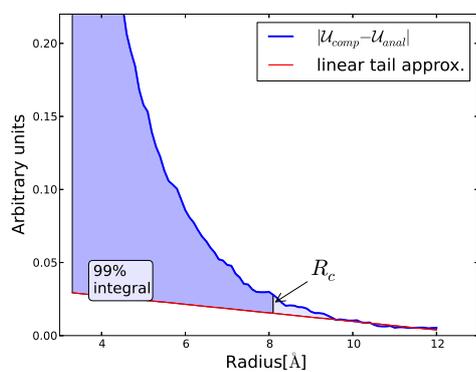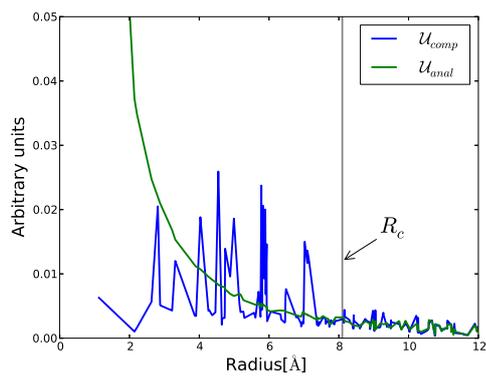

(c) (2-1) Strain energy

(f) (2-3) Strain energy

Figure C.11: Core structure and radius for (2-1) and (23)

143

(a) (-32) Core structure

(b) (-32) R_c retrieval

(c) (-32) Strain energy

Figure C.12: Core structure and radius for (-32)

Bibliography

[1] Ya-Hong Xie. Prof. Ya-Hong Xie's group, Semiconductor Materials Lab, Los Angeles, California, USA. Available at http://www.seas.ucla.edu/smrl/GaN.html, April 2013.

[2] A. Béré and A. Serra. On the atomic structures, mobility and interactions of extended defects in GaN: dislocations, tilt and twin boundaries. *Philosophical Magazine*, 86(15):2159–2192, 2006.

[3] Junqiao Wu. When group-III nitrides go infrared: New properties and perspectives. *Journal of Applied Physics*, 106(1):011101, 2009.

[4] Xinqiang Wang, Shitao Liu, Nan Ma, Li Feng, Guang Chen, Fujun Xu, Ning Tang, Sen Huang, Kevin J. Chen, Shengqiang Zhou, and Bo Shen. High-Electron-Mobility InN Layers Grown by Boundary-Temperature-Controlled Epitaxy. *Applied Physics Express*, 5(1):015502, 2012.

[5] Michelle A. Moram. Light-emitting diodes and their applications in energy-saving lighting. *Proceedings of the ICE - Energy*, 164:17–24(7), 2011.

[6] C. Stampfl, C. G. Van de Walle, D. Vogel, P. Krüger, and J. Pollmann. Native defects and impurities in InN: First-principles studies using the local-density approximation and self-interaction and relaxation-corrected pseudopotentials. *Phys. Rev. B*, 61(12):R7846–R7849, Mar 2000.

[7] Chris J. Pickard and Richard J. Needs. Structure of phase III of solid hydrogen. *Nat Phys*, 3(7):473–476, 07 2007.

145

[8] Mike Finnis. *Ineratomic forces in condensed matter*. Oxford University Press, New York, 2003.

[9] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[10] R. O. Jones and O. Gunnarsson. The density functional formalism, its application and prospects. *Reviews of Modern Physics*, 61(3):689, 1989.

[11] Mel Levy. Electron densities in search of Hamiltonians. *Phys. Rev. A*, 26:1200–1208, Sep 1982.

[12] Klaus Capelle. A Bird's-Eye View of Density-Functional Theory. *Brazilian Journal of Physics*, 36(4A):1318, 2006.

[13] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[14] Richard M. Martin. *Electronic Structure*. Cambridge University Press, 2004.

[15] J. P. Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048–5079, May 1981.

[16] John P. Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45:13244–13249, Jun 1992.

[17] Hendrik J. Monkhorst and James D. Pack. Special points for Brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976.

[18] Giovanni B. Bachelet and M. Schlüter. Relativistic norm-conserving pseudopotentials. *Phys. Rev. B*, 25:2103–2108, Feb 1982.

[19] C. Hartwigsen, S. Goedecker, and J. Hutter. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B*, 58:3641–3662, Aug 1998.

[20] Steven G. Louie, Sverre Froyen, and Marvin L. Cohen. Nonlinear ionic pseudopotentials in spin-density-functional calculations. *Phys. Rev. B*, 26:1738–1742, Aug 1982.

[21] AIMPRO. Aimpro.abinitio website , http://aimpro.ncl.ac.uk/, April 2010.

[22] P.R. Briddon and R. Jones. LDA Calculations Using a Basis of Gaussian Orbitals. *physica status solidi (b)*, 217(1):131–171, 2000.

[23] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77(18):3865–3868, Oct 1996.

[24] M. J. Rayson and P. R. Briddon. Highly efficient method for Kohn-Sham density functional calculations of $500 − −10\,000$ atom systems. *Phys. Rev. B*, 80:205104, Nov 2009.

[25] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipies: the art of scientific computing.* Cambridge University Press, 1986.

[26] Yoshiyuki Kawazoe Kaoru Ohno, Keivan Esfarjani. *Computational Materials Science.* Springer Verlag, 1999.

[27] Frank H. Stillinger and Thomas A. Weber. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B*, 31:5262–5271, Apr 1985.

[28] I. Belabbas, A. Béré, J. Chen, S. Petit, M. Akli Belkhir, P. Ruterana, and G. Nouet. Atomistic modeling of the $(\mathbf{a} + \mathbf{c})$-mixed dislocation core in wurtzite GaN. *Phys. Rev. B*, 75:115201, Mar 2007.

[29] Antoine Béré and Anna Serra. Atomic structure of dislocation cores in GaN. *Phys. Rev. B*, 65:205323, May 2002.

[30] J Chen, P Ruterana, and G Nouet. Multiple atomic configurations of the a edge threading dislocation in GaN. *Materials Science and Engineering: B*, 82(1–3):117 – 119, 2001.

[31] J. Kioseoglou, G. P. Dimitrakopulos, Ph. Komninou, and Th. Karakostas. Atomic structures and energies of partial dislocations in wurtzite GaN. *Phys. Rev. B*, 70:035309, Jul 2004.

[32] Aïchoune, V Potin, P Ruterana, A Hairie, G Nouet, and E Paumier. An empirical potential for the calculation of the atomic structure of extended

defects in wurtzite GaN. *Computational Materials Science*, 17(2–4):380 – 383, 2000.

[33] A. Polian, M. Grimsditch, and I. Grzegory. Elastic constants of gallium nitride. *Journal of Applied Physics*, 79(6):3343–3344, 1996.

[34] N. Nakamura, H. Ogi, and M. Hirao. Elastic, anelastic, and piezoelectric coefficients of GaN. *Journal of Applied Physics*, 111(1):013509, 2012.

[35] Kwiseon Kim, Walter R. L. Lambrecht, and Benjamin Segall. Electronic structure of GaN with strain and phonon distortions. *Phys. Rev. B*, 50:1502–1505, Jul 1994.

[36] T. Lei, M. Fanciulli, R. J. Molnar, T. D. Moustakas, R. J. Graham, and J. Scanlon. Epitaxial growth of zinc blende and wurtzitic gallium nitride thin films on (001) silicon. *Applied Physics Letters*, 59(8):944–946, 1991.

[37] S. Plimpton. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J Comp Phys*, 117:1–19, 1995.

[38] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[39] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

[40] Bernd Hartke. Global geometry optimization of clusters using genetic algorithms. *The Journal of Physical Chemistry*, 97(39):9973–9976, 1993.

[41] Jian Lv, Yanchao Wang, Li Zhu, and Yanming Ma. Particle-swarm structure prediction on clusters. *The Journal of Chemical Physics*, 137(8):084104, 2012.

[42] Riccardo Poli. Analysis of the Publications on the Applications of Particle Swarm Optimisation. *Journal of Artificial Evolution and Applications*, 2008.

[43] Yuhui Shi and Russell Eberhart. Parameter selection in particle swarm optimization. 1447:591–600, 1998. 10.1007/BFb0040810.

[44] Chris J Pickard and R J Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

[45] Frank H. Stillinger. Exponential multiplicity of inherent structures. *Phys. Rev. E*, 59:48–51, Jan 1999.

[46] John (Price) Hirth and Jens Lothe. *Theory of Dislocations.* Krieger Publishing Company, Malabar, Florida 32950,US, second edition, 1982.

[47] W.T. Read. *Dislocation in Crystals.* McGraw-Hill, New York, 1953.

[48] C. Stampfl and C. G. Van de Walle. Density-functional calculations for III-V nitrides using the local-density approximation and the generalized gradient approximation. *Phys. Rev. B*, 59(8):5521–5535, Feb 1999.

[49] M A Moram and M E Vickers. X-ray diffraction of III-nitrides. *Reports on Progress in Physics*, 72(3):036502, 2009.

[50] J. S. Speck and S. F. Chichibu. Nonpolar and Semipolar Group III Nitride-Based Materials. *MRS Bulletin*, 34:304–312.

[51] Patrick Rinke, M. Winkelnkemper, A. Qteish, D. Bimberg, J. Neugebauer, and M. Scheffler. Consistent set of band parameters for the group-III nitrides AlN, GaN, and InN. *Physical Review B (Condensed Matter and Materials Physics)*, 77(7):075202, 2008.

[52] J. Wu, W. Walukiewicz, W. Shan, K. M. Yu, J. W. Ager III, S. X. Li, E. E. Haller, Hai Lu, and William J. Schaff. Temperature dependence of the fundamental band gap of InN. *Journal of Applied Physics*, 94(7):4457–4460, 2003.

[53] T. L. Tansley and C. P. Foley. Optical band gap of indium nitride. *Journal of Applied Physics*, 59(9):3241–3244, 1986.

[54] Dirk Vogel, Peter Krüger, and Johannes Pollmann. Structural and electronic properties of group-III nitrides. *Phys. Rev. B*, 55(19):12836–12839, May 1997.

[55] Su-Huai Wei, Xiliang Nie, Iskander G. Batyrev, and S. B. Zhang. Breakdown of the band-gap-common-cation rule: The origin of the small band gap of InN. *Phys. Rev. B*, 67(16):165209, Apr 2003.

[56] F. Bechstedt and J. Furthmuller. Do we know the fundamental energy gap of InN? *Journal of Crystal Growth*, 246(3-4):315 – 319, 2002.

[57] W. Walukiewicz. Mechanism of Fermi-level stabilization in semiconductors. *Phys. Rev. B*, 37:4760–4763, Mar 1988.

[58] D. Segev and C. G. Van de Walle. Origins of Fermi-level pinning on GaN and InN polar and nonpolar surfaces. *EPL (Europhysics Letters)*, 76(2):305, 2006.

[59] Chris G. Van de Walle and David Segev. Microscopic origins of surface states on nitride surfaces. *Journal of Applied Physics*, 101(8):081704, 2007.

[60] P. D. C. King, T. D. Veal, C. F. McConville, F. Fuchs, J. Furthmüller, F. Bechstedt, P. Schley, R. Goldhahn, J. Schörmann, D. J. As, K. Lischka, D. Muto, H. Naoi, Y. Nanishi, Hai Lu, and W. J. Schaff. Universality of electron accumulation at wurtzite c- and a-plane and zinc-blende InN surfaces. *Applied Physics Letters*, 91(9):092101, 2007.

[61] X. M. Duan and C. Stampfl. Nitrogen vacancies in InN: Vacancy clustering and metallic bonding from first principles. *Physical Review B (Condensed Matter and Materials Physics)*, 77(11):115207, 2008.

[62] Christian Rauch, Ilja Makkonen, and Filip Tuomisto. Identifying vacancy complexes in compound semiconductors with positron annihilation spectroscopy: A case study of InN. *Phys. Rev. B*, 84:125201, Sep 2011.

[63] Anderson Janotti and Chris G. Van de Walle. Sources of unintentional conductivity in InN. *Applied Physics Letters*, 92(3):032104, 2008.

[64] Xinqiang Wang, Song-Bek Che, Yoshihiro Ishitani, and Akihiko Yoshikawa. Growth and properties of Mg-doped In-polar InN films. *Applied Physics Letters*, 90(20):201913, 2007.

[65] S. E. Bennett. Dislocations and their reduction in GaN. *Materials Science and Technology*, 26(9):1017–1028, 2010-09-01T00:00:00.

[66] M. A. Moram, C. S. Ghedia, D. V. S. Rao, J. S. Barnard, Y. Zhang, M. J. Kappers, and C. J. Humphreys. On the origin of threading dislocations in GaN films. *Journal of Applied Physics*, 106(7):073513, 2009.

[67] Niklas Lehto and Sven Öberg. Effects of Dislocation Interactions: Application to the Period-Doubled Core of the 90° Partial in Silicon. *Phys. Rev. Lett.*, 80:5568–5571, Jun 1998.

[68] J. Bennetto, R. W. Nunes, and David Vanderbilt. Period-Doubled Structure for the 90° Partial Dislocation in Silicon. *Phys. Rev. Lett.*, 79:245–248, Jul 1997.

[69] Wei Cai, Vasily V. Bulatov, Jinpeng Chang, Ju Li, and Sidney Yip. Anisotropic Elastic Interactions of a Periodic Dislocation Array. *Phys. Rev. Lett.*, 86:5727–5730, Jun 2001.

[70] X. Blase, Karin Lin, A. Canning, S. G. Louie, and D. C. Chrzan. Structure and Energy of the 90° Partial Dislocation in Diamond: A Combined *Ab Initio* and Elasticity Theory Analysis. *Phys. Rev. Lett.*, 84:5780–5783, Jun 2000.

[71] L. Lymperakis, J. Neugebauer, M. Albrecht, T. Remmele, and H. P. Strunk. Strain Induced Deep Electronic States around Threading Dislocations in GaN. *Phys. Rev. Lett.*, 93:196401, Nov 2004.

[72] E. Kalesaki, J. Kioseoglou, L. Lymperakis, Ph. Komninou, and Th. Karakostas. Effect of edge threading dislocations on the electronic structure of InN. *Applied Physics Letters*, 98(7):072103, 2011.

[73] J. R. K. Bigger, D. A. McInnes, A. P. Sutton, M. C. Payne, I. Stich, R. D. King-Smith, D. M. Bird, and L. J. Clarke. Atomic and electronic structures of the 90° partial dislocation in silicon. *Phys. Rev. Lett.*, 69:2224–2227, Oct 1992.

[74] Python. Python Reference Manual, PythonLabs, Virginia, USA, 2001. Available at http://www.python.org, April 2010.

[75] I. Belabbas, J. Chen, and G. Nouet. A new atomistic model for the threading screw dislocation core in wurtzite GaN. *Computational Materials Science*, 51(1):206 – 216, 2012.

[76] A. T. Blumenau. *The Modelling of Dislocations in Semiconductor Crystals.* PhD thesis, University of Paderborn, December 2002.

[77] John E. Northrup. Screw dislocations in GaN: The Ga-filled core model. *Applied Physics Letters*, 78(16):2288–2290, 2001.

[78] L. D. Landau and E. M. Lifshitz. *Statistical Physics*, volume 1. Elsevier Butterworth-Heinemann, 3 edition, 1980.

[79] John E. Northrup. Energetics of GaAs island formation on Si(100). *Phys. Rev. Lett.*, 62:2487–2490, May 1989.

[80] John E. Northrup. Theory of intrinsic and H-passivated screw dislocations in GaN. *Phys. Rev. B*, 66:045204, Jul 2002.

[81] M. Bernasconi, Guido L. Chiarotti, and E. Tosatti. *Ab initio* calculations of structural and electronic properties of gallium solid-state phases. *Phys. Rev. B*, 52:9988–9998, Oct 1995.

[82] Agostino Zoroddu, Fabio Bernardini, Paolo Ruggerone, and Vincenzo Fiorentini. First-principles prediction of structure, energetics, formation enthalpy, elastic constants, polarization, and piezoelectric constants of AlN, GaN, and InN: Comparison of local and gradient-corrected density-functional theory. *Phys. Rev. B*, 64:045208, Jul 2001.

[83] J. Kioseoglou, Ph. Komninou, and Th. Karakostas. Core models of a-edge threading dislocations in wurtzite III(Al,Ga,In)-nitrides. *physica status solidi (a)*, 206(8):1931–1935, 2009.

[84] Y. Xin, S. J. Pennycook, N. D. Browning, P. D. Nellist, S. Sivananthan, F. Omnes, B. Beaumont, J. P. Faurie, and P. Gibart. Direct observation of the core structures of threading dislocations in GaN. *Applied Physics Letters*, 72(21):2680–2682, 1998.

[85] I. Belabbas, A. Béré, J. Chen, P. Ruterana, and G. Nouet. Investigation of the atomic core structure of the $(\vec{a} + \vec{c})$-mixed dislocation in wurtzite GaN. *physica status solidi (c)*, 4(8):2940–2944, 2007.

[86] Simon Kraeusel and Benjamin Hourahine. Global search for stable screw dislocation cores in III-N semiconductors. *Physica Status Solidi A*, 209(1):71–74, 2012.

Acknowledgments