EINDHOVEN UNIVERSITY OF TECHNOLOGY
Eindhoven University of Technology

MASTER

A study of doping atoms in III/V semiconductors by DFT and X-STM

Verstijnen, Tom J.F.

Award date:
2021

Link to publication

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Applied Physics

A study of doping atoms in III/V semiconductors by DFT and X-STM

Author:
T.J.F. Verstijnen

Supervisors:
D. Tjeertes, MSc
prof. dr. P. M. Koenraad

Eindhoven, February-December 2020

Abstract

The combination of experimental and theoretical techniques for the study of (doped) semiconductor materials is very useful. Here we present the combined use of cross-sectional scanning tunneling microscopy (X-STM), an experimental technique, and density functional theory (DFT), a theoretical technique. Two different problems are studied using these techniques. The first problem regards isoelectronic doping of GaAs with Bi and/or N. The pair formation of these isoelectronic dopants is studied. As well as the effect of co-doping with N on Bi incorporation in GaAs. From DFT, we find that the first nearest neighbour Bi-Bi pair configuration is energetically favourable to the second nearest neighbour configuration and this is corroborated by X-STM measurements. The co-doping of GaAs with N and Bi does not have a significant effect on the Bi incorporation. The second problem considered is the doping of AlAs with Si donors. The effect of these donors on the band structure of AlAs is studied using DFT calculations. DFT is also used to simulate X-STM images of these donors and these simulated images are compared to actual X-STM measurements. A surface layer Si donor is found to introduce an impurity band at the Fermi level into the system, to push up the conduction band approximately 0.2 eV and to split states in the conduction band. The simulation of X-STM images is presented for a surface and second layer Si donor in AlAs and compared to real X-STM images. The short range behaviour of the Si donor is successfully predicted in both cases. In the experimental measurements, the long range effects of the surface layer donor seem to be related to charge effects and these are not included in the DFT calculations.

Contents

1 Introduction **3**
1.1 Motivation . . . 3
1.2 Bi and N doping of GaAs . . . 4
1.3 Si donors in AlAs . . . 6

2 Theory **8**
2.1 Density functional theory . . . 8
2.1.1 Convergence tests . . . 11
2.2 Zinc blende (110) surface brillouin zone . . . 12
2.3 Scanning tunneling microscopy . . . 16
2.3.1 The square barrier problem . . . 16
2.3.2 Modified bardeen approach . . . 17

3 Method **19**
3.1 DFT methodology . . . 19
3.1.1 Software and hardware used . . . 19
3.1.2 Convergence tests . . . 19
3.2 STM experiments . . . 22
3.2.1 Sample and tip preparation . . . 22
3.2.2 STM system setup . . . 22

4 Results **24**
4.1 Isoelectronic doping of GaAs with Bi and N . . . 24
4.1.1 DFT simulation of bulk GaAs doped with Bi and N . . . 24
4.1.2 DFT simulation of Bi-Bi pairs in bulk GaAs . . . 26
4.1.3 DFT simulation of N-N pairs in bulk GaAs . . . 27
4.1.4 DFT simulation of Bi-N pairs in GaAs . . . 28
4.1.5 X-STM measurements of Bi doped GaAs . . . 31
4.1.6 DFT simulation of on Bi-Bi pairs in the growth surface of GaAs . . . 37
4.1.7 X-STM measurements of GaAs co-doped with Bi and N . . . 41
4.2 Si donors in AlAs . . . 46
4.2.1 DFT simulation of Si donors in AlAs . . . 46
4.2.2 X-STM measurements on Si donors in AlAs . . . 53

5 Conclusion **59**
5.1 Isoelectronic doping of GaAs with Bi and N . . . 59
5.2 Si donors in AlAs . . . 60

References **61**

1 Introduction

1.1 Motivation

Semiconductor materials are everywhere nowadays. Everyone walks around with plenty semiconductor chips and devices on them at all times. It is therefore no surprise that the behaviour and characteristics of these semiconductor materials are to be understood in as much detail as possible. Modifying the characteristics of these semiconductor materials locally allows us to create useful devices such as light emitting diodes (LEDs) or transistors. One of the main techniques available to us in order to influence the behaviour of these materials is doping of these semiconductor materials. Doping III/V semiconductors in particular can have a wide range of different results, allowing us to fine-tune the behaviour we want in a material. Two different ways of doping a semiconductor material are considered using experimental and theoretical techniques in this study.
In 1965 Moore predicted that the amount of transistors that fit on an integrated circuit would double every two years [1]. This prediction has held true until now and the size of these transistors is therefore reaching the size of individual dopants [2]. Doping control has therefore reached the atomic scale. This means that advanced techniques are needed to investigate semiconductor materials on the atomic scale. Both a theoretical and an experimental technique are presented in this study and applied to two differently doped semiconductor materials.
The theoretical technique utilised in this study is Density Functional Theory (DFT). This is an *ab initio* technique that is well equipped to describe materials with many nuclei and electrons. This technique was proven to be valid by Hohenberg and Kohn in 1964 [3]. The technique is based on describing a material in terms of the electron density. DFT relies on supercells of atoms subjected to periodic boundary conditions to describe real materials. The supercells used to simulate materials usually have a size in the order of a few hundred atoms. DFT is used to calculate the ground state of the material considered and from that state many interesting characteristics can be calculated: ground state energy, forces on different atoms, the band structure and the Density Of States (DOS) of different materials [4]. Therefore DFT provides us with great insight as to what is happening at the atomic scale in semiconductors.
The experimental technique utilised in this study is Scanning Tunneling Microscopy (STM). This technique was invented by Binning and Rohrer in 1981[5] and probes a material surface with an atomically sharp tip. This probing technique is based on the tunneling of electrons from the tip to the material that is scanned and vice versa. The lateral resolution of STM has been on the scale of Angström [6]. This high resolution means that STM can be used to image doped materials on the atomic scale.
This study aims to combine the two techniques described here, STM and DFT. This is done since the PSN group at the TU/e has a lot of experience with (X-)STM measurements [7] [8] [9], but is still in the early stages of combining the techniques [10] [11]. There is a lot to gain from combining these two techniques since they both compensate the weaknesses of the other nicely.
A limitation of DFT is the finite size of the supercell, caused by a limitation to computational power. A finite supercell will introduce an artificial periodicity in your system. This makes investigating a single dopant in a system difficult for instance, since the artificial periodicity will always introduce additional dopants. STM can be of great use to check the shape of the calculated DOS around these dopants without the influence of other nearby dopants. Another limitation of DFT is that it underestimates band gaps [4] and the size of these band gaps can also be checked experimentally by performing spectroscopy measurements with an STM. Even though STM also has its limitations accurately estimating band gaps due to tip induced band bending, these techniques can nicely complement each other here.
A limitation of the STM on the other hand is that an STM probes the DOS of a material rather than the physical shape of said material. This means that electronic effects on the surface can be hard to distinguish from the physical structure of a surface. DFT can be used to find the relaxed surface structure of the material investigated with STM and this can in turn be used to interpret the STM results.

As mentioned before two different types of semiconductor doping are studied here. These types are doping by introducing donors in the system and isoelectronic doping.
The first doping method discussed is the introduction of donors in the material. Doping a material with

donors is done by introducing atoms in the material that have more valence electrons than the atoms they are replacing. This can for instance be accomplished by introducing group IV atoms on group III atomic sites in III/V semiconductors. The doping of a material with donors can influence the characteristics of semiconductors in interesting ways. It can introduce new carriers in the system by influencing the electron density for instance. This in turn affects the conductivity of a material [12]. Doping with donors can also influence the band structure of the material by tuning the Fermi level [13] or even bend the bands of a material through for instance p-n junctions, when combined with acceptor doping.
A less often used way of doping a material is through isoelectronic doping. This type of doping is achieved by replacing a type of atoms in a material with atoms that have the same amount of valence electrons. Through isoelectronic doping the size, [14] and shape [15] of the band gap of a material can be tuned.

The two different doping problems discussed in this study are Si donors in AlAs and the isoelectronic doping of GaAs with Bi and N.
Si donors in AlAs are investigated since interesting characteristics of Si donors in GaAs were already found in the PSN group at the TU/e [16] and Si donors in AlAs could provide similar rich results. Completely similar behaviour is not expected since AlAs, unlike GaAs, is an indirect semiconductor and has a vastly different donor binding energy which will be discussed in section 1.3.
Isoelectronic doping of GaAs with Bi and N is of interest for optical communications, since the band gap of GaAs can be manipulated by doping with Bi or N. Both dopants reduce the band gap of GaAs but in different ways. Doping with Bi moves the top of the valence band up [17], while doping with N shifts the bottom of the conduction band down [18]. Specifically the pair formation of these dopants in GaAs is studied here. A more detailed description of the interesting features of GaAs doped with Bi and N will be given in section 1.2.
The two materials mentioned here present two very different challenges for both STM and DFT. In the case of the Si doped AlAs, a more general study of the local behaviour of these Si donors is to be performed. This means that the shape of the Si donors and effect of these donors on the DOS and band structure are of interest. Both STM and DFT are well suited to image the DOS around a Si donor and DFT can be used to calculate the band structure of the material. The pair formation of Bi and N dopants in GaAs can also be studied by both techniques. Using DFT, the energetically favourable pair configurations can be found, while STM measurements can be used to find these dopant pairs in GaAs samples.

The final objective of this study is twofold. First, we look to obtain a better understanding of the investigated doped semiconductor materials. In particular we will investigate the structural effects of isoelectronic doping and the electronic effect of doping with donors. The second objective is to achieve greater insight in the limitations and strengths of the combined use of DFT and STM.

1.2 Bi and N doping of GaAs

As mentioned in section 1.1, GaAs doped with Bi (and/or N) is an interesting material regarding applications in optical communication. This stems from the fact that Bi doping of GaAs can greatly influence the band gap of GaAs. A Bi concentration of 3-4% can already reduce the band gap by up to 25 % [19]. The incorporation of high amounts of Bi into GaAs is limited by weak Ga-Bi reactivity and the strong tendency for Bi to segregate [20]. These highly Bi doped GaAs structures can be used to create longer wavelength devices on GaAs substrates [21]. Because of the fact that highly Bi doped GaAs materials are this interesting, the growth of these materials has been researched extensively.
During the Molecular Beam Epitaxy (MBE) growth of Bi:GaAs, Bi acts as a surfactant layer on top of the growth surface due to its relatively large atomic radius and high vapour pressure [20]. Also, the growth surface of Bi:GaAs exhibits complex behaviour during the growth, different from the behaviour of regular (001) GaAs during growth. The growth surface used for most GaAs based optoelectronic applications worldwide is the (001) surface [22]. This growth surface undergoes different surface reconstructions during the growth of Bi:GaAs, for instance the (2×1) and (1×3) surface reconstructions, that also influence the Bi incorporation [23]. The main parameters used to optimise this growth are the As_2:Ga beam equivalent

pressure ratio and the growth temperature [20].
Another way of influencing the Bi incorporation into GaAs is co-doping with N. This has a few additional advantages over the techniques mentioned before. The co-doping with N is expected to make Bi incorporation easier since N doping induces the opposite strain to Bi doping. Bi doping induces compressive strain [24] while N doping induces tensile strain [25]. An additional advantage of doping with both N and Bi is improved carrier confinement over solely Bi doping. When GaAs is doped with Bi, the band gap is decreased by shifting up the valence band maximum as mentioned in section 1.1. This gives rise to hole confinement but no confinement of the electrons is enforced. This leads to low efficiency in optoelectronic devices. When GaAs is doped with N however, the band gap is lowered by lowering the conduction band minimum, leading to electron confinement. Combining these effects leads to electron and hole confinement in the same area. This in turn increases the efficiency of recombination.
The last method of influencing Bi incorporation during growth discussed here is light assisted MBE, as proposed by Beaton et al. [26]. They propose a model that describes the influence of the surfactant layer on the growth of the Bi doped GaAs. They also state that subjecting the material to laser pulses with energy above the band gap during growth can increase the range of optimal growth conditions. They found that the illuminated sample areas had a lower Bi content and suggested two possible explanations for this. First, the energy imparted by the light could lower the energy needed for Bi atoms to escape the surface again once incorporated. Second, the photogenerated carriers could influence the surface reconstruction in which the Bi atoms can easily incorporate into the surface.
When GaAs is isoelectronically doped with Bi and/or N, the clustering and pair formation of these dopants shows some interesting characteristics. The Bi pairs in $GaAs_{1-x}Bi_x$ are for instance known to influence the hole mobility of the material [27]. The extend to which dopants tend to cluster together in a certain material is also important to understand. When there is very strong clustering in a material, the material is made up out of areas with very high and very low dopant density. This means that the dopants would behave like they are present in a material with a much higher doping concentration than the avarage doping concentration.
The PSN group in Eindhoven has performed a study on both the incorporation of Bi in InP and the incorporation of N in GaAs.
Krammel et al. [28] found that Bi dopants in InP tend to cluster through a study of the nearest Bi-Bi neighbour pair occurrences. The result of this pair occurrences study is presented in figure 1.

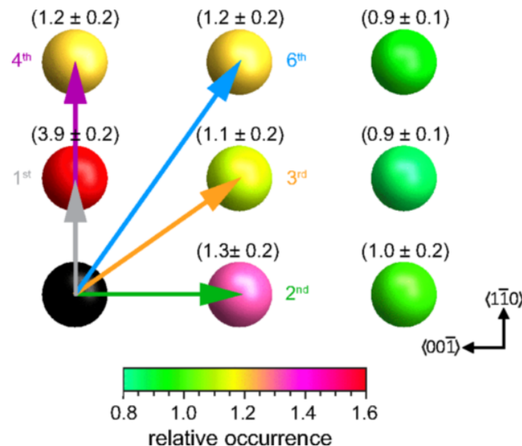


Figure 1: Relative occurrence of several different Bi-Bi nearest neighbour pair configurations in Bi:InP. The occurrence of the 1st, 2nd, 3rd, 4th and 6th nearest neighbour Bi-Bi pairs are presented relative to a random distribution of Bi atoms. (Taken from [28])

As can be seen in figure 1, the relative occurrences of the 1st, 2nd, 3rd, 4th and 6th nearest neighbour pairs

are all above 1. This means that all these pair configurations occur more often than would be expected from a random Bi distribution. The 1st nearest neighbour is even found close to four times more often than expected. This relative occurrence study indicates that Bi atoms tend to cluster in InP.
A similar study was also done by Plantenga, R. C. (2019) [29] on N doping of GaAs. The resulting pair distribution of N in GaAs is given in figure 2.

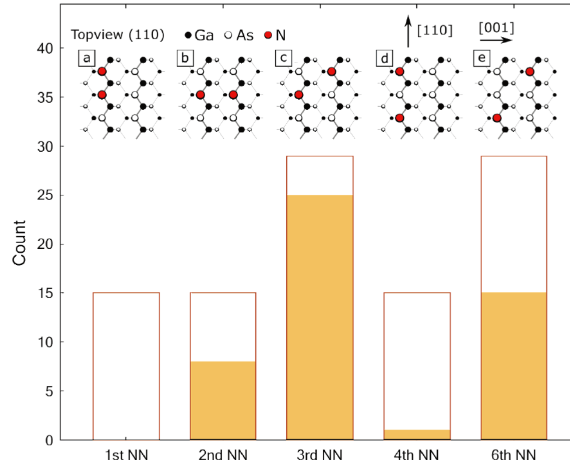


Figure 2: Occurrence of several different N-N nearest neighbour pair configurations in N:GaAs. The red boxes show the expected occurrence of each nearest neighbour pair if the N distribution was random. (Taken from [29])

The results shown in figure 2 indicate that N atoms tend to anti-cluster in GaAs. Especially the first nearest neighbour configuration seems unfavourable, since it is not observed at all.
As mentioned in section 1.1, we will study Bi and N co-doping of GaAs here. The influence of the presence of N atoms on the Bi clustering will be studied and this material will also provide a first indication of Bi-N pair formation in co-doped materials.

1.3 Si donors in AlAs

As mentioned in section 1.1, there is interest in Si doped AlAs within the PSN group. This stems from the interesting results found for Si doped GaAs. The Si dopants in GaAs showed switching within a bi-stable state and tip induced ionisation rings for instance [16]. Now we look at Si doped AlAs since there are a few striking differences between GaAs and AlAs and checking whether similar behaviour of the Si dopants is observed can increase our understanding in what happens around these dopants on the atomic scale. One of the main characteristics that differentiates AlAs from GaAs is the fact that AlAs is an indirect semiconductor. Increasing the Al concentration in an $Al_xGa_{1-x}As$ material moves the conduction band minimum from the Γ point to the X point as shown in figure 3.

Looking at figure 3, it is clear that the conduction band minimum shifts from the Γ point to the X point at an Al fraction of approximately 41%.
Another interesting aspect of this new conduction band minimum at the X point is that this point is three fold degenerate with the x- and y-direction, $X_{x,y}$ having the same effective mass but this effective mass being different from the one of the z-direction, X_z [31]. This degeneracy can be lifted by strain in the system. When a layered system of GaAs and AlAs is formed for instance, the biaxial compression lifts this degeneracy through compressive strain. This strain is caused by the lattice mismatch in the case of this layered structure. Due to the compressive strain, the $X_{x,y}$ valleys become the lowest point of the conduction

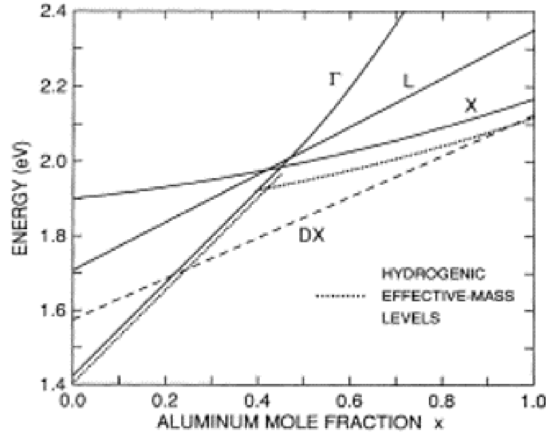


Figure 3: Schematic diagram showing energies of the three conduction-band minima, the hydrogenic levels, and the DX level as a function of the composition of $Al_xGa_{1-x}As$. (adapted from [30])

band and are split of from the X_z minimum [32].
The donor binding energy and therefore bohr radius of donors in AlAs also differs from the one of GaAs. The donor binding energy can be calculated within the hydrogenic model using

$$E_d = 1Ry \times \frac{m^*}{m_0 \epsilon_r^2} = 13.6\ eV \times \frac{m^*}{m_0 \epsilon_r^2}. \quad (1)$$

Here m_0 is the vacuum electron mass, m^* is the electron mass in the material and ϵ_r is the relative dielectric constant. Now using the fact that the electron mass and relative dielectric constant of GaAs are $0.067m_0$ [33] and 12.90 [34] respectively we can calculate the donor binding energy of GaAs to be 5.5 meV. For AlAs the electron mass and relative dielectric constant are $0.15m_0$ [33] and 10.06 [35] respectively and therefore AlAs has a donor binding energy of 20 meV. This means that electrons are more tightly bound to donors in AlAs. This is also reflected in the bohr radius around donors for these two materials. These can be calculated using

$$a_r = a_0 \times \frac{m_0 \epsilon_r}{m^*} = 0.529\ \mathring{A} \times \frac{m_0 \epsilon_r}{m^*}. \quad (2)$$

Again using the aforementioned values of the electron mass and dielectric constant we find the bohr radius of GaAs and AlAs to be 102 Å and 35 Å respectively. This means that electrons bound to donors in AlAs are kept much closer to the donor atom than in GaAs. When electrons are held more closely to the donor atom, they are exposed to less surface of the background material present. This means that the effects of the donor atom become more and more important relative to the background lattice when the bohr radius gets smaller. Since the bohr radius is indeed smaller for AlAs than for GaAs, we expect a stronger influence of the Si atom on the electronic behaviour near the position of the donor atom.
This smaller bohr radius of Si donors in AlAs makes it a suitable material to investigate with DFT. One of the limiting factors of DFT is the size of the supercell. A donor with a smaller bohr radius is expected to interact with a smaller environment and therefore a smaller supercell can more effectively be used to describe the donor behaviour in a real material.

2 Theory

2.1 Density functional theory

Density functional theory (DFT) is a powerful *ab initio* technique often used in simulation of materials with many electrons and nuclei. These calculations are often very laborious since there are many degrees of freedom involved: three for each electron. The power of DFT lies in the fact that the equations regard the electron density instead of the n individual electrons. This decouples the Schrödinger equation from an equation with $3n$ variables into n separate equations with 3 variables. This makes the calculations much more manageable. Another advantage of switching to the electron density is that it is measurable directly, contrary to the electron orbitals. The orbitals $\phi_i(r)$ and the electron density $\rho(r)$ are directly related however via

$$\rho(r) = \sum_{i}^{Occ} |\phi_i(r)|^2. \tag{3}$$

Where the sum runs over the occupied electronic orbitals. The relation shown in equation 3 means that information about the electron orbitals can still be extracted from the calculated electron density.

Hohenberg-Kohn
The DFT technique was expected to be a powerful tool for a while when Hohenberg and Kohn found the proof for its validity in 1964 [3] through two theorems:
The first theorem states that the external potential U_{ext} is a unique functional of the electron density.
The second theorem states that the ground state energy can be found via the variational principal. Meaning that the electron density that minimises the energy is the ground state electron density of a system.
These two theorems together lay the foundation for DFT calculations since the ground state can now be found by varying the electron density. And the electron density corresponding to this grounds state can be used to find the total energy or any other property of a given system, defined by a given U_{ext}.

Kohn-Sham
Even though these two theorems verified the use of the DFT method to describe many electron systems, this technique still needed to be developed further. The poor description of the kinetic energy of electrons as a function of the electron density was one of the main reasons for this. Kohn and Sham (1965) found a way to simplify these many electron systems though. They developed a so-called "divide and conquer' strategy. This strategy meant that they grouped all the energy contributions from electrons interaction with each other directly into one electronic contribution. Starting from the total energy, E, of

$$E = E_{kin} + E_{ext} + E_H + E_x. \tag{4}$$

Where E_{kin} is the kinetic energy of the electrons, E_{ext} is the external energy from the electrons interacting with the nuclei, E_H is the Hartree energy e.g. the energy from the interaction with the mean electron density and E_x is the exchange energy from electrons interacting with each other. Kohn and Sham rephrased this total energy equation as

$$E = E_{kin}^{non} + E_{ext} + E_H + E_{xc}. \tag{5}$$

Where E_{kin}^{non} is now the non-interacting part of the kinetic energy of the electrons and E_{xc} is the exchange correlation energy. In this rephrased version of the total energy, all the energy considerations regarding the interaction between electrons are condensed into this exchange energy. This is visually represented in figure 4.

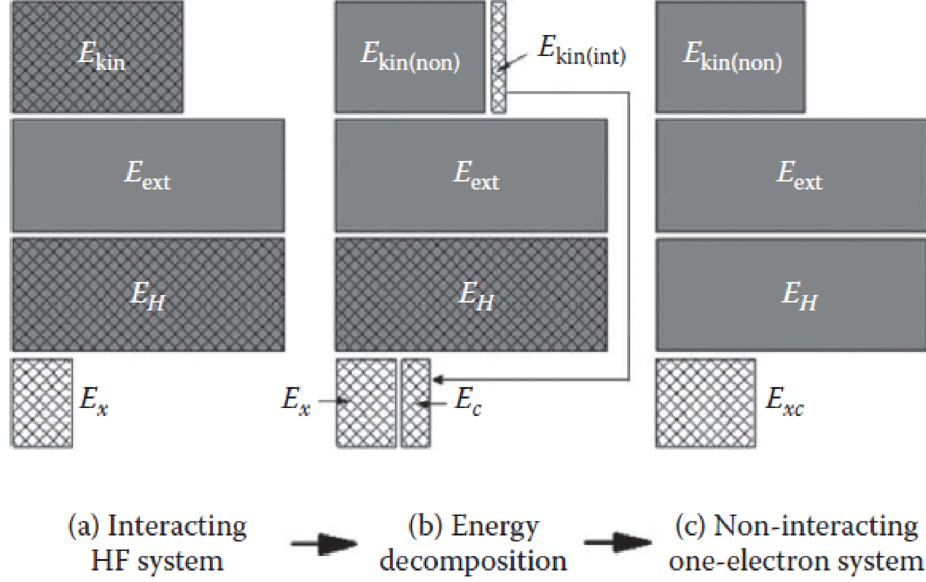


Figure 4: Visual representation of the rephrased energy as presented by Kohn-Sham. The kinetic energy of the electrons, E_{kin}, the energy from the electron-nuclei interaction, E_{ext}, the Hartree energy, E_H, exchange energy, E_x, correlation energy, E_c, and exchange correlation energy, E_{xc}, are all indicated in the figures. (figure taken from [4])

Variational principle approach

Now that we have an equation for the total energy we can use the variational principle to find the ground state. This means finding the minimum of E with respect to the variables involved. Now taking into account that both

$$E = \int \Psi^* \hat{H} \Psi d\mathbf{r} \tag{6}$$

and

$$\int \Psi_i^* \Psi_j d\mathbf{r} = \delta_{ij} \tag{7}$$

we can use a Lagrange multiplier, λ, to find the minimum energy. This gives us

$$0 = \delta\Big(\int \Psi^* \hat{H} \Psi d\mathbf{r} \Big) = \delta\Big[\int \Psi^* \hat{H} \Psi d\mathbf{r} - \lambda\Big(\int \Psi_i^* \Psi_j d\mathbf{r} - 1 \Big) \Big]. \tag{8}$$

Now using equation 6 we can find the Kohn and Sham equations (KS equations) via

$$0 = \frac{\delta E[\rho(\mathbf{r})]}{\delta \phi_i^*(\mathbf{r})} = \frac{\delta E_{kin}^{non}}{\delta \phi_i^*(\mathbf{r})} + \Big[\frac{\delta E_{ext}}{\delta \rho(\mathbf{r})} + \frac{\delta E_H}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})} \Big] \frac{\delta \rho(\mathbf{r})}{\delta \phi_i^*(\mathbf{r})} - \sum_j \lambda_{ij} \phi_j(r). \tag{9}$$

Now inserting the potentials for the energy derivatives and rearranging some terms we find

$$0 = \Big(-\frac{1}{2}\nabla^2 + U_{ext} + U_H + U_{xc} - \lambda_i \Big) \phi_i(\mathbf{r}) = \Big(-\frac{1}{2}\nabla^2 + U_{eff} - \epsilon_i \Big) \phi_i(\mathbf{r}). \tag{10}$$

Where λ is exchanged with ϵ, the energy of each orbital, since equation 10 can be recognised as the normal Schrödinger equation since it can be rewritten as

$$\Big[-\frac{1}{2}\nabla^2 + U_{eff}\Big]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \rightarrow \hat{H}_{KS}\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \quad (11)$$

This is the final eigenvalue problem that is solved in the DFT calculations, the KS equations. Through the years it has been discovered that ϵ_i is much more than just a Lagrange multiplier. It turns out that these eigenvalues are actually the energies of each KS electron so their sum can be treated as the total energy of a system.

XC energy approximations

This now leaves us with the problem of the form of the exchange correlation energy E_{XC}. There are several different ways of approximating the energy of the electrons interacting with each other. One example is the Local Density Approximation (LDA). In the LDA approach, the space is divided up into small areas in which a constant electron density is assumed. When this is the case, the exchange correlation energy can be calculated within such a homogeneous area and the energies of this area can be summed up to the total energy of the system. Making this approximation of course has some drawbacks since this does not describe inhomogeneous systems that well. These limitations include overestimating the cohesive energy and therefore underestimates the lattice parameter, calculating the adsorption energies too high and the diffusion barriers too low, underestimating the spin and orbital moments, calculating the band gap too small and failing to describe transition metals well.
There are numerous other approximations for E_{xc}: GGA, PBE, PW91 and all sorts of hybrid functionals to name a few [36] [37]. In the work considered here LDA is the method used to calculate the E_{xc}.

Pseusopotentials

Another approximation we can make when performing DFT calculations is found by taking a look at the potentials around nuclei present in the system. These potentials and the wave functions of the electrons arising from them can be very complex and difficult to compute. This is where the pseudopotential approach comes in.
The first step of the pseudopotential-approach is realising that we have two types of electrons in our system: core electrons and valence electrons. The core electrons are the electrons that are so localised near the nuclei that the rest of the material does not influence them. These atoms are therefore considered to be part of the core and for each core electron the charge of the core is simply lowered by $1e$. This is called the frozen core approximation. Now we are left only with the valence electrons to be taken into account in DFT calculations. These electrons are the ones that actually interact with the rest of the system and therefore the interesting ones.
The valence electrons we are left with have very complicated looking wave functions inside the core area of their corresponding nuclei. This is due to the fact that the electron states around a nucleus need to be orthogonal to each other and therefore the higher lying energy levels, those of the valence electrons, have a lot of nodes inside the core area. We consider the core electrons to simply be frozen into the core however. So these nodes inside the core area are no longer strictly necessary. This means that we can smooth out the wave function and therefore the potential landscape inside the core area and this is what the pseudopotentials do. The potentials and wave functions and their derivatives do need to have the same form outside the core region. This process is given schematically in figure 5. These smoother wave functions can also be more easily expanded into plane waves, which greatly reduces computation time. Moreover, the error introduced by using pseudopotentials is often only around 1%. There are many different types of pseudopotentials but the pseudopotentials used in this study are ultrasoft pseudopotentials. Ultrasoft pseudopotentials have a "softer" nature which means in practice that the peak in wave function at the core region boundary is shifted to the right and made lower. This lowering of the peak of the wave function makes the wave function even easier to be expanded into plane waves and this can reduce computation time by a factor of up to 10.

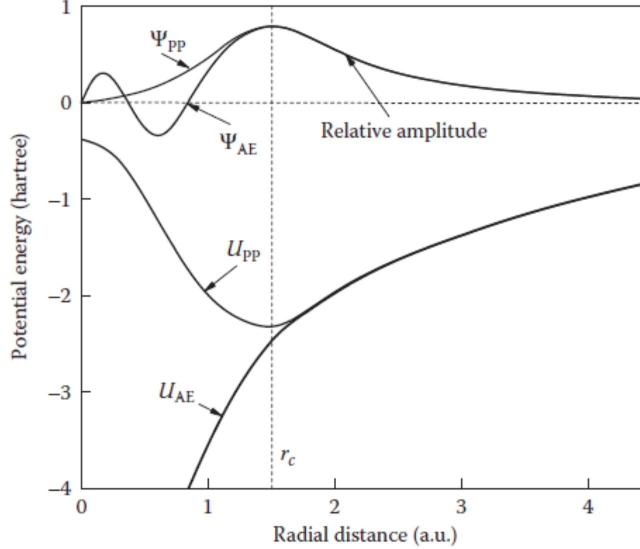


Figure 5: Graphical example of the real wave function and pseudopotential, Ψ_{AE} and U_{AE}, and their pseudized counterparts, Ψ_{PP} and U_P. (taken from [4])

2.1.1 Convergence tests

When performing DFT calculations, the goal is to get as accurate a result as possible within as little computation time as possible. There are three areas highlighted in this section that are suitable for saving computation time while keeping the accuracy within a set margin. These are a plane wave energy cutoff, a electron kinetic energy cutoff and the calculated k-point grid within the brillouin zone.

Plane wave energy cutoff
When starting the calculation, each wave function is expressed as a function of a set of plane waves (if you use the plane wave basis method). This means that each wave function can be written as

$$\Psi_k(r) = \frac{1}{\Omega} \sum_G c_{k,G} * e^{i(k+G)*r}. \tag{12}$$

Where k is the wave vector and G is the reciprocal lattice vector. This means that in principle all possible plane waves corresponding to the value of $(k + G)$ should be taken into account. This means an infinite number of plane waves. This can be circumvented however by realising that the contribution to the wave function of high values of $|k + G|$ are very small and therefore a plane wave energy cutoff given by

$$\frac{\hbar^2 |k + G|^2}{2m} < E_{wfc} \tag{13}$$

is introduced. This value of E_{wfc} has to be tested in practice so that the accuracy is sufficient but using this energy cutoff will greatly reduce the computation time.

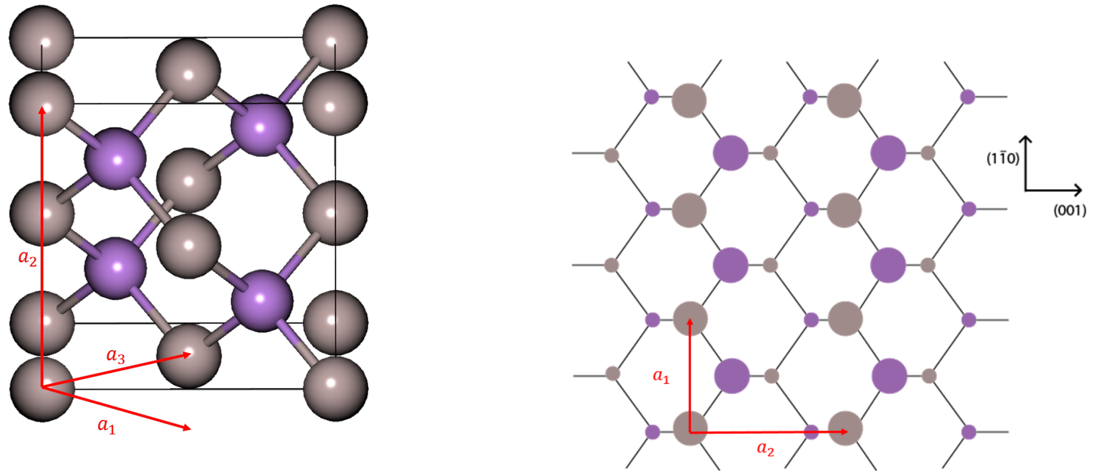
Kinetic energy cutoff
A similar energy cutoff can be used with regard to the kinetic energy of the electrons. This is the case because the charge density is also calculated in reciprocal space so again it is expanded in terms of plane waves. However, it has to be taken into account that, in order to calculate the charge density from the wave functions, we have to square the wave functions, see equation 3. So the reciprocal lattice of the charge density has to at least contain all reciprocal lattice vectors up to $2G_{wfc}$, where G_{wfc} is the reciprocal lattice vector

that corresponds to the largest G allowed according to equation 13. Similar accuracy tests can be done with respect to a $E_{cut,\rho}$ as long as the minimum value due to the value of E_{wfc} is taken into account. How the minimum value of E_ρ depends on the value of E_{wfc} depends strongly on the type of pseudopotential used. For ultrasoft pseudopotentials the rule of thumb is $E_\rho > 8 * E_{wfc}$ and for norm conserving pseudopotentials a rule of thumb is $E_\rho > 4 * E_{wfc}$.

K-point sampling
A similar trick can be done with respect to the brillouin zone sampling. A lot of quantities that are interesting to look at when doing DFT calculations, for instance total energy and number of electrons, require a summation over all k-points in the brillouin zone. This would in theory mean an infinite number of k-points to sum over leading again to an infinite computation time. Again a test can be done where we look at accuracy of the total energy or another parameter with regard to the number of k-points in the brillouin zone we use. When a sufficient amount of k-points is found, this can again greatly reduce the computation time.

2.2 Zinc blende (110) surface brillouin zone

As seen in section 2.1, one of the convergence tests in DFT finds an appropriate sampling mesh of the Brillouin Zone (BZ). Also, a good understanding of the BZ is needed in DFT band structure calculations. Therefore BZ of a (110) surface of a zinc blende material is discussed here. This surface BZ is used to investigate the band structure of a (110) zinc blende surface. For this it is necessary to know what high symmetry points exist in such a BZ and what the shape of this BZ is. The conventional unit cell with the corresponding vectors used that span the (110) surface and its normal are shown in figure 6.



(a) Conventional zinc-blende conventional unit cell with sides of length a.

(b) Top view of the (110) surface of a zinc blende material.

Figure 6: Illustrations of the lattice vectors a_1, a_2 and a_3 used to describe the (110) surface of a zinc blende material.

The vectors a_1, a_2 and a_3 shown in figure 6 are given by

$$\vec{a}_1 = \frac{a}{2}(1, -1, 0), \tag{14}$$

$$\vec{a}_2 = a(0, 0, 1) \tag{15}$$

and

$$\vec{a}_3 = c(1,1,0). \quad (16)$$

Here a is the lattice constant of the material and c is an arbitrary constant. The value of c is arbitrary since the periodicity perpendicular to the (110) surface is not taken into account. The non-importance of the length of vector $\vec{a}_3$ is also reflected in the fact that it does not show up in figure 6b. Vectors $\vec{a}_1$, $\vec{a}_2$ and $\vec{a}_3$ now make up the surface unit cell which can be used to calculate the surface BZ. In order to find this BZ we calculate the primitive lattice vectors $\vec{b}_1$, $\vec{b}_2$ and $\vec{b}_3$. For this we need the volume of the surface unit cell, V, given by

$$V = |\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)|. \quad (17)$$

The volume of the surface unit cell is found to be a^2c. Now we can use this volume and the lattice vectors to find the primitive lattice vectors through

$$\vec{b}_1 = \frac{2\pi}{V}(\vec{a}_2 \times \vec{a}_3), \quad (18)$$

$$\vec{b}_2 = \frac{2\pi}{V}(\vec{a}_3 \times \vec{a}_1) \quad (19)$$

and

$$\vec{b}_3 = \frac{2\pi}{V}(\vec{a}_1 \times \vec{a}_2). \quad (20)$$

From equations 18, 19 and 20 we can now calculate the primitive lattice vectors to be

$$\vec{b}_1 = \frac{2\pi}{a}(-1,1,0), \quad (21)$$

$$\vec{b}_2 = \frac{2\pi}{a}(0,0,-1) \quad (22)$$

and

$$\vec{b}_3 = \frac{\pi}{c}(-1,-1,0). \quad (23)$$

The (110) BZ is now made up of vectors $\vec{b}_1$ and $\vec{b}_2$. Notice that $|\vec{b}_1| > |\vec{b}_2|$ due to the length of the vector being a factor of $\sqrt{2}$ larger. Also, $\vec{b}_1$ and $\vec{b}_2$ are perpendicular to each other. This means that the surface BZ is a rectangle with width $\frac{4\pi}{\sqrt{2}a}$ and length $\frac{2\pi}{a}$.
Now we consider the actual surface used in the calculations following later. The (110) surface of that cell is given in figure 7.

Now we have a new set of unit vectors $\vec{a}'_1$, $\vec{a}'_2$ and $\vec{a}'_3$ given by

$$\vec{a}'_1 = 2a(1,-1,0), \quad (24)$$

$$\vec{a}'_2 = 4a(0,0,1) \quad (25)$$

and

$$\vec{a}'_3 = c(1,1,0). \quad (26)$$

Note that $\vec{a}'_3$ still has an arbitrary length. We can now calculate the new volume using equation 17, this new volume is $16a^2c$. Using this new volume and equations 18, 19 and 20 we now find vectors $\vec{b}'_1$, $\vec{b}'_2$ and $\vec{b}'_3$ to be

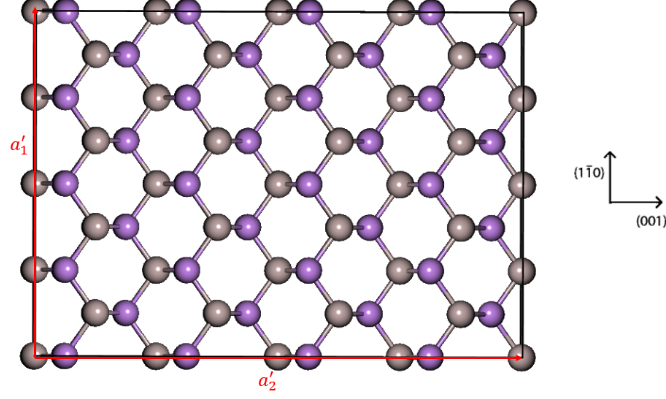


Figure 7: Top view of the surface unit cell used in DFT calculations with surface unit cell vectors $\vec{a}_1'$ and $\vec{a}_2'$ given schematically.

$$\vec{b}_1' = \frac{\pi}{2a}(-1, 1, 0), \tag{27}$$

$$\vec{b}_2' = \frac{\pi}{2a}(0, 0, -1) \tag{28}$$

and

$$\vec{b}_3' = \frac{\pi}{c}(-1, -1, 0). \tag{29}$$

Notice that the BZ is again given by a rectangular cell since $\vec{b}_1'$ and $\vec{b}_2'$ make up this BZ, $|\vec{b}_1'| > |\vec{b}_2'|$ and they are perpendicular to each other. We can now draw this surface BZ and identify the high symmetry points Γ, X, X' and M.

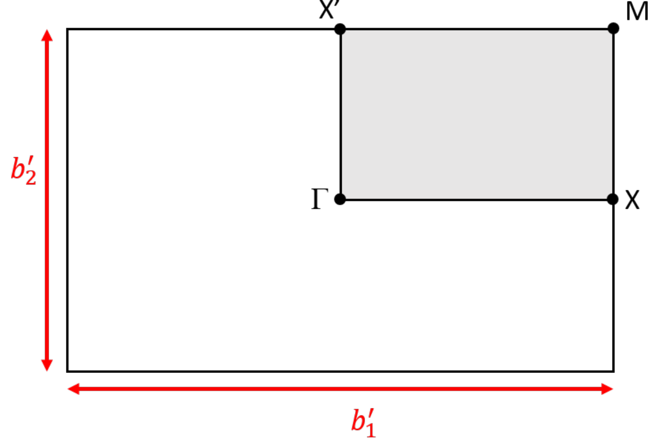


Figure 8: Surface BZ of (110) zinc blende material with the high symmetry points indicated. The irreducible BZ is highlighted in grey.

Lastly, some high symmetry points of the bulk BZ are folded onto points of this surface BZ. Some examples are: $X \to \Gamma$, $L \to X'$ and $W \to M$ [38]. We take a detailed look at the case $X \to \Gamma$. The X

point is situated at $\frac{2\pi}{a}(1,0,0)$ in the 3D BZ. This point is now projected onto the gamma point in the (110) surface BZ. In order to understand this we first need to revisit the meaning of c in our calculated vector $\vec{b}_3'$. This c could be given any value since there is no periodicity perpendicular to the (110) plane in a (110) surface. This means that we could assign any value to this c, lets take the value of a here. This turns our vector $\vec{b}_3'$ into $\frac{\pi}{a}(-1,-1,0)$. Now we can express the bulk BZ X point as

$$X = \frac{2\pi}{a}(1,0,0) = -2\vec{b}_1' - \vec{b}_3'. \tag{30}$$

The fact that the X point can be expressed in terms of the reciprocal lattice vectors with integer indices means that the X point is folded onto the Γ point.

2.3 Scanning tunneling microscopy

Scanning Tunneling Microscopy (STM) was invented in 1981 by Binning and Rohrer [5]. Ever since, it has been widely used to characterise and investigate surfaces of materials on the atomic scale. STM is a scanning probe technique based on tunneling current between a conducting sample and a sharp tip in close proximity. Since this tunneling current is highly sensitive to tip-sample distance, STM can be used to image material surfaces at the sub-nano scale.

There are two main operating modes of an STM, constant height and constant current mode. In constant height mode, the tip is held at a constant height above the surface and the surface is scanned by moving the tip over the sample. The resulting current as a function of tip position is then a measure for the height of different features on the sample surface. This is a very fast scanning technique since no feedback loop is needed to control the tip height. It is a relatively risky scanning mode however, since the tip-sample distance is only constant in theory. If the sample is tilted or features on the sample are larger in size than the tip sample distance, the tip could crash into the sample destroying the sharpness of the tip and/or the sample. This is why constant current measurements are utilised in this research.

Constant current mode is a scanning mode where the tip-sample distance is varied while the current between the tip and the sample is held constant by a feedback loop. This is a slower scanning method because the tip needs to be regulated more closely but safer since the tip will be less likely to crash into the surface. The tip movement is controlled by piezo-actuators since fine control of the tip is needed to control the tip movement on such a small scale.

2.3.1 The square barrier problem

The underlying process of STM is the tunneling of electrons through a classically forbidden region. In its most simplistic form this is given by the square barrier problem. This square barrier problem will be discussed here alongside its solution in the WKB approximation as well as its limitations and improvements to this model in the form of the Modified Bardeen Approach (MBA) [6].

The square barrier problem consists free electrons in space with a classically forbidden area of with W in the middle of this free space. It is given schematically in figure 9.

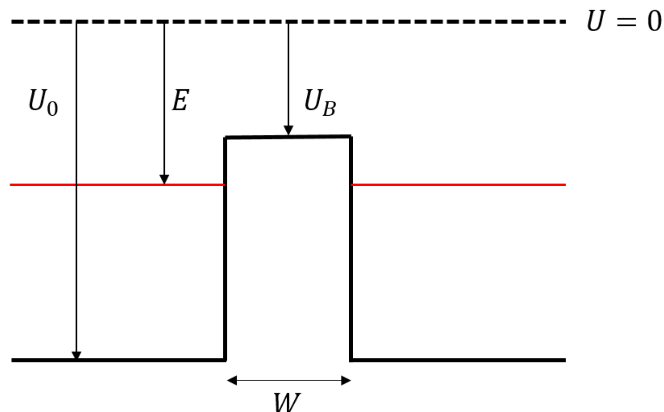


Figure 9: Schematic energy landscape of the square barrier problem. The particle energy, E, barrier height, U_B, barrier width, W, and reference energy, U_0, are given.

The schematic in figure 9 shows the square barrier problem schematically with classically allowed areas left and right of the barrier and a classically forbidden area inside the barrier. We now describe the incoming, reflecting and transmitting waves as $exp(iqx)$, $A * exp(-iqx)$ and $B * exp(iqx)$ respectively. Here q is defined as

$$q = \frac{\sqrt{2m(E - U_0)}}{\hbar}. \tag{31}$$

Here m is the electron mass and $\hbar$ is the reduced Planck constant. We can use these wave functions and the boundary conditions on the barrier edges to describe the transmission, T which is defined as

$$T = \frac{|B|^2}{1 + |A|^2}. \tag{32}$$

This leads to the transmission coefficient of

$$T = \left[1 + \frac{1}{2}\left(\frac{\kappa}{q} + \frac{q}{\kappa}\right)^2 sin^2(\kappa W)\right]^{-1}. \tag{33}$$

Here κ is given by

$$\kappa = \frac{\sqrt{2m(U_B - E)}}{\hbar} = 0.51\sqrt{U_B - E}. \tag{34}$$

For a thick barrier in the semi classical limit equation 33 becomes

$$T = 2\left(\left(\frac{\kappa}{q} + \frac{q}{\kappa}\right)^{-2} e^{-2\kappa W}. \tag{35}$$

When we look at equation 35 we see that the transmission probability decays exponentially with the barrier width. This is why the tunneling current is so sensitive to small changes in tip-sample distance.

2.3.2 Modified bardeen approach

The Modified Bardeen Approach (MBA) is a more advanced model used to describe the physics behind STM. In this approach, we consider not the simplistic energy landscape shown in figure 9, but instead we consider the overlap of 3D wave functions of the tip and sample. This approach is a perturbation approach where both wave functions are considered perturbed due to the presence of the other. From the overlap between these wave functions, the tunneling rate and tunneling current can then be found.
An intuitive description of the steps taken in the MBA are presented here. The different states with subscript μ and ν mark states and energies corresponding to the sample and tip respectively. In time dependent perturbation theory the transition probability, w of an electron between the tip and sample state is given by Fermi's golden rule:

$$w_{\mu\nu} = \frac{2\pi}{\hbar}|M_{\mu\nu}|^2\delta(E_\nu - E_\mu), \tag{36}$$

where M is the matrix element given by

$$M_{\mu\nu} = \int_{\Omega_T} \chi_\nu^* U_T \psi_\mu d\tau. \tag{37}$$

Here χ_μ are the tip wave functions and ψ_μ the wave function of the sample.
Looking at equations 36 and 37 we can see that the transition rate sets the requirement that the energy of the sample and tip state has to be the same for the transition to take place and the rate of this transition is then determined by the overlap of the wave functions given by the matrix element. We can now also calculate the tunneling current, I, using this matrix element by integrating over all the states in the tip and the sample when a bias, V, is applied

$$I = \frac{4\pi e}{\hbar}\int_{-\infty}^{\infty}[f(E_F - eV + \epsilon) - f(E_F + \epsilon)] \times \rho_S\rho_T|M_{\mu\nu}|^2\delta\epsilon. \tag{38}$$

Here f is the Fermi dirac distribution, E_F the Fermi energy and ρ_S and ρ_T are the density of states of the sample and tip respectively.
This now leaves the matrix element given in equation 37 as the determining factor for the tunneling current and therefore the behaviour of electron tunneling in STM measurements.
These matrix elements can be calculated by transforming the integral seen in equation 37 to a surface integral using Green's theorem and expanding the tip wave functions in spherical harmonics. The resulting matrix elements are actually very elegant and they are all functions of the sample state and its derivatives. For instance the matrix element for the s-wave tip state, M_S is given by

$$M_S = \frac{2\pi C\hbar^2}{\kappa m}\psi(r_0). \tag{39}$$

Here C is a constant and $\psi(r_0$ is the magnitude of the sample state at center of the tip. To obtain the p-wave tip states we now only need to differentiate the sample state with respect to the orientation of the p-wave we are interested in and we find the corresponding matrix element, M_p,

$$M_{p_x} = \frac{2\pi C\hbar^2}{\kappa m}\frac{\partial\psi(r_0)}{\partial x}. \tag{40}$$

This methodology of finding matrix elements by differentiating the sample state with respect to the different coordinates is where the name derivative rule stems from.
We can now calculate all the matrix elements that describe the tunneling behaviour using the wave function that makes up the sample state. This way the MBA gives us great insight in the physical process of tunneling in STM measurements.

3 Method

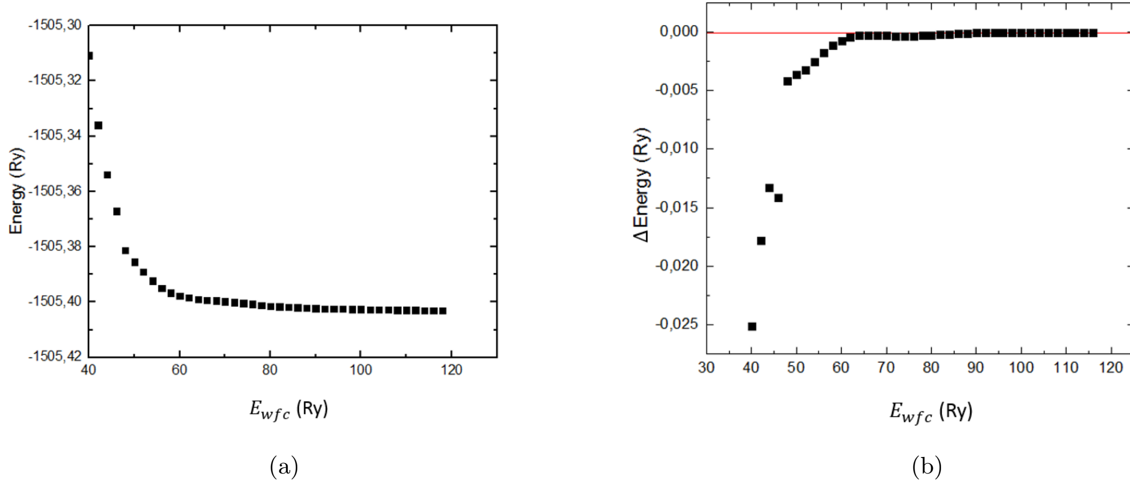
3.1 DFT methodology

3.1.1 Software and hardware used

The DFT calculations were performed on the Cartesius supercomputer situated in Amsterdam. The program used to execute the DFT calculation is Quantum-Espresso version 6.0 [39] [40]. The core electrons of the different atoms are treated within a pseudopotential method with ultrasoft pseudopotentials [41] and exchange-correlation energy functional is computed with the local density approximation (LDA). The images of atomic structures and simulated STM images are generated using XCrysden version 1.5.60 [42].

3.1.2 Convergence tests

The convergence tests shown below are performed for a (110) surface supercell of AlAs with a Si donor in the surface layer consisting of 126 atoms total. This surface supercell is 7 atomic layers thick and has 7 Å of vacuum above and below it. Since it is a surface supercell, the bottom three atomic layers are fixed as to simulate bulk while the top 4 atomic layers are free to relax. The first convergence test performed is the convergence of the cut off energy of the plane waves describing the wave functions, E_{wfc}. The results of this convergence test are shown in graph 10.



(a)

(b)

Figure 10: Results of the E_{wfc} convergence test of the Si doped AlAs supercell. The total energy of the system as a function of E_{wfc} is given in a) while the change in total energy is given in b). The red line in b) indicates the energy difference at which convergence has been achieved.

The total energy of the system as a function of E_{wfc} is given in figure 10a and the change in energy for each E_{wfc} is given in figure 10b. In figure 10b, the convergence threshold of 10^{-4} Ry is also given at the red line. By looking at figure 10b the E_{wfc} is determined to be 90 Ry. Since we are using ultrasoft pseudopotentials, the rule of thumb is to use a kinetic energy cutoff, E_ρ, of at least 8 times the E_{wfc}. To check whether the total energy is indeed converged at this E_ρ, we perform a convergence test on E_ρ.

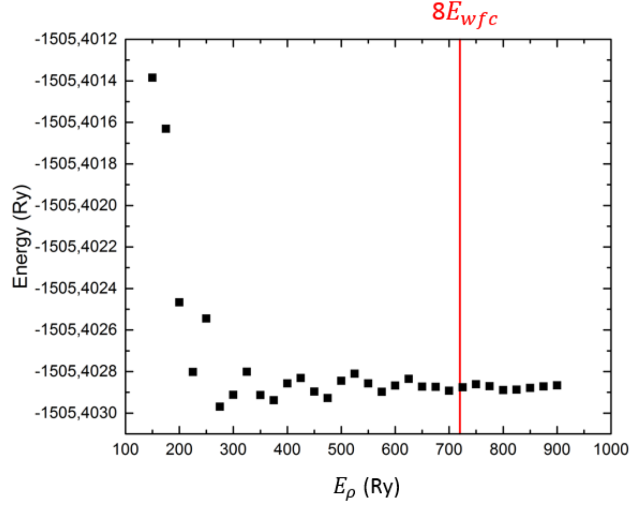


Figure 11: Total energy of the system as a function of E_ρ.

As can be seen in figure 11, the energy did indeed converge for an E_ρ of eight times E_{wfc}. Because of this we set the value of E_ρ to 750 Ry.

Since we use a finite number of k-points to sample our Brillouin Zone (BZ) in order to not have an infinite computation time, we need to find an appropriate density of this k-grid. For this we use a Monkhorst-Pack grid [43]. A Monkhorst-Pack grid is an unbiased method of choosing a set of k-points to sample the BZ. An $X * X * X$ Monkhorst-Pack grid represents a grid of evenly spaced k-points of dimensions $X_x * X_y * X_z$. We can test the density of such a grid for our system in a convergence test. For the example of a (110) surface of Si doped AlAs such a convergence test looks like what is shown in figure 12.

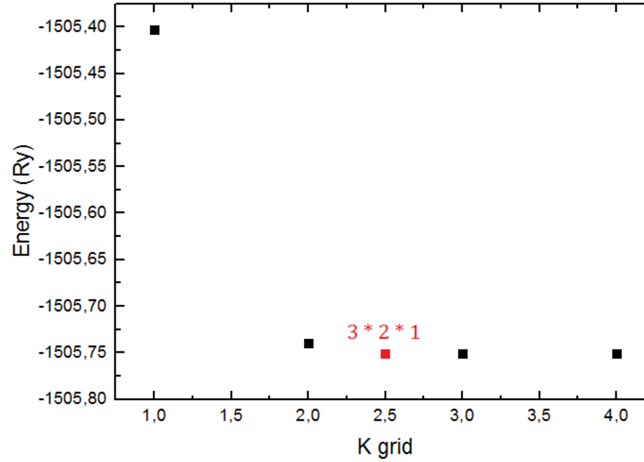


Figure 12: Total energy of the system as a function of the k-point sampling. Here the points on the x-axis, X, represent an $X * X * X$ grid used to sample the Brillouin zone.

Again the system shows a convergence for denser k-grids. In this case, we chose the k-grid of $3 * 2 * 1$, which is indicated in red in figure 12.

The last test to be performed is the lattice constant test. This is less of a convergence test and more of a system test since we now look for the lattice constant which results in the lowest total energy for our supercell. This lattice constant can vary slightly from the experimental lattice constant for the materials investigated due to the overestimation of cohesive energies in DFT. The results of the lattice test of the Si doped AlAs supercell are presented in figure 13.

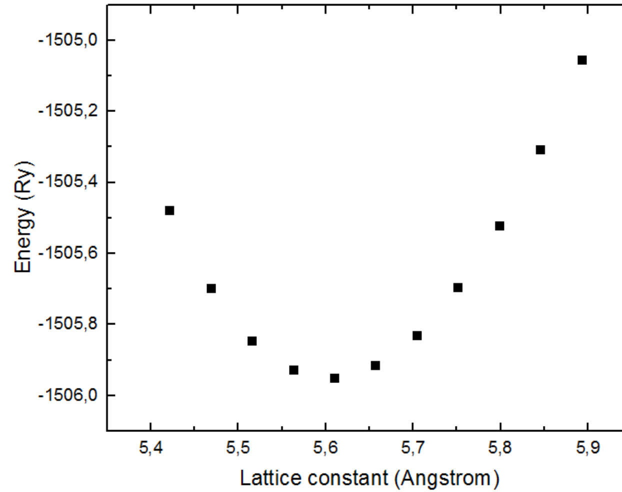


Figure 13: Total energy of the Si doped AlAs supercell as a function of the lattice constant of the AlAs.

In figure 13 the total energy of the supercell is given as a function of the lattice constant of the AlAs. The lattice constant that yields the lowest energy is approximately 5.61 Å, while the experimentally determined lattice constant of AlAs is 5.66 Å[44]. These lattice constants vary slightly on the one hand due to DFT overestimating cohesive energies and on the other hand due to it being heavily Si doped AlAs (1 in every 126 atoms is Si).

Now the testing phase is completed and the non self-consistent calculations can be performed. These include but are not limited to relaxation calculations, Local Density of States (LDOS) calculations and band structure calculations. All of which will be addressed later on in this work.

3.2 STM experiments

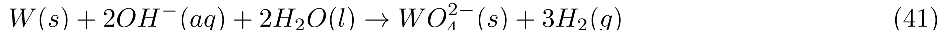
In this paragraph we describe the two main components of performing X-STM measurements: the preparation of the samples and tips used and the operation of the STM itself.

3.2.1 Sample and tip preparation

Both good sample and tip conditions are vital in order to generate good X-STM images. Therefore, many steps are taken in preparation of a measurement regarding the quality of samples and tips.

Sample preparation
The samples used are cut out of larger wafers into 4x8 mm rectangles. To ensure that the conductivity in the STM itself will be good, contacts are evaporated onto these 4x8 mm samples using thermal evaporation. These contacts are made up of Ge, Ni and Au. Ge is used as an ohmic contact, Ni is a buffer material and Au is used to improve the electrical contact to the sample holder.
The next step in the sample preparation is the polishing of the samples. To ensure our samples cleave properly, we make them as this as possible. Therefore we polish our samples mechanically using aluminum oxide (Al_2O_3). The samples are polished until they are 120-140 μm thick. This thickness has proven to be thin enough so that it can be cleaved easily in site but not so thin that it breaks unintentionally. These samples are then clamped into sample holders with small pieces of indium on each side and the sample holder is heated up. This melts the indium and thereby improves the electrical contact between sample holder and sample.
The final sample preparation steps are done inside the STM itself. First, the samples are baked at 180 °C for 25 minutes for additional cleaning. Then finally, in the STM chamber at 77K, the sample is cleaved. This exposes the (110) surface and this surface is clean since this cleaving took place at ultra-high vacuum at 77K.

Tip preparation
The tip preparation is done using chemical etching of a tungsten (W) tip. For this a W wire is cut into a 17 mm piece which is submerged into a 2 M KOH solution. This wire in the solution acts as an electrode in a redox reaction. The other electrode is made up of a platinum coil also submerged in the same KOH solution and the circuit is completed. Now a voltage is applied between the two electrodes and the following redox reaction provides us with the etching of the tip:



$$W(s) + 2OH^-(aq) + 2H_2O(l) \rightarrow WO_4^{2-}(s) + 3H_2(g) \tag{41}$$

This reaction causes the W wire to etch away. This primarily happens at the meniscus of the wire at the KOH solution surface. This is where the tip is therefore etched away the fastest and where the wire will break. This break forms a nice tip usable in STM measurements. The tip is first investigated with an optical microscope to make sure it isn't visibly blunt and if deemed a good tip, will be transferred into the STM.
In the STM some last tip preparation steps are performed. First, the tip is also baked at 230 °C for 30 minutes for additional cleaning. The tip is then sputtered with argon to remove any remaining oxides and to ensure tip sharpness. After this the tip is transferred into the STM chamber and used for STM measurements. During these measurements voltage pulses through the tip can be used to improve the tip quality.

3.2.2 STM system setup

The STM system used is a commercial Omicron LT-STM. The schematic overview of the STM system used in this study is given in figure 14.

The system shown in figure 14 is made up of three main chambers. These are the load lock, preparation chamber and STM chamber.
The load lock is where new samples and tips can be put into the system and old ones can be taken out. This section of the STM is pumped using a turbomolecular and scroll pump in series. The load lock connects to

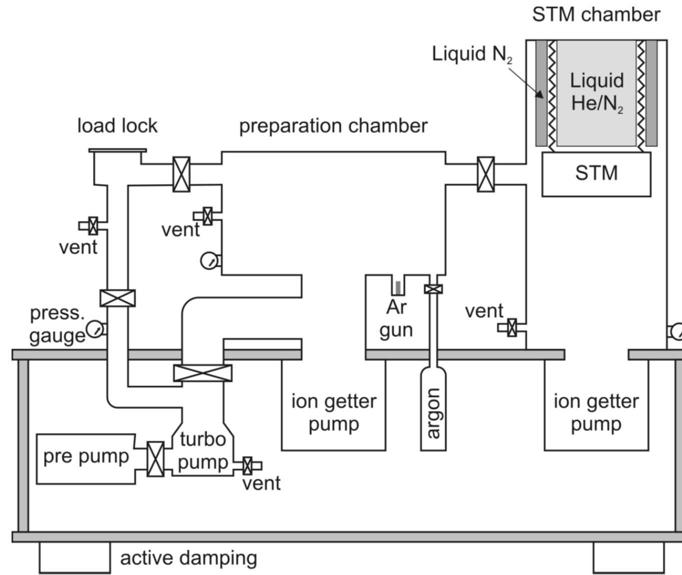


Figure 14: Schematic overview of the Omicron LT-STM system used to perform X-STM measurements in this study. (Taken from [16]

the preparation chamber through a connecting tube with a gate valve.
The preparation chamber is where the additional tip and sample preparation steps mentioned in section 3.2.1 are performed. This preparation chamber is normally pumped using an ion-getter pump. This chamber is pumped with the aforementioned turbomolecular pump when tips are being sputtered. This is done to ensure the ion-getter pump does not get polluted. The contact-resistance heater and ion gun that are needed for the final preparation steps for the tips and samples are also located in this chamber. The pressure in this preparation chamber is $1 \cdot 10^{-9}$ mbar.
The final section is the STM chamber. This chamber has a fridge where cleaved samples can be stored at 77K, a stage where additional samples and tips can be parked and the actual STM setup. The STM itself is surrounded by two concentric cryogenic shields. The outer one of which is connected to a liquid nitrogen bath. The inner one is connected to a separate bath containing either liquid nitrogen for 77K measurements or liquid helium for 4K measurements. In this study only 77K measurements are presented so both baths were filled with liquid nitrogen. The STM is suspended in springs to reduce vibrations. The vibrations during measurements are reduced further by the STM being mounted on a platform with a separate foundation to the rest of the building, active damping on the STM system as a whole, the setup being present in a closed of space with sound dampening materials on the walls and by the scroll and turbomolecular pumps being switched off. The pressure in the STM chamber is kept below $5 \cdot 10^{-11}$ mbar.

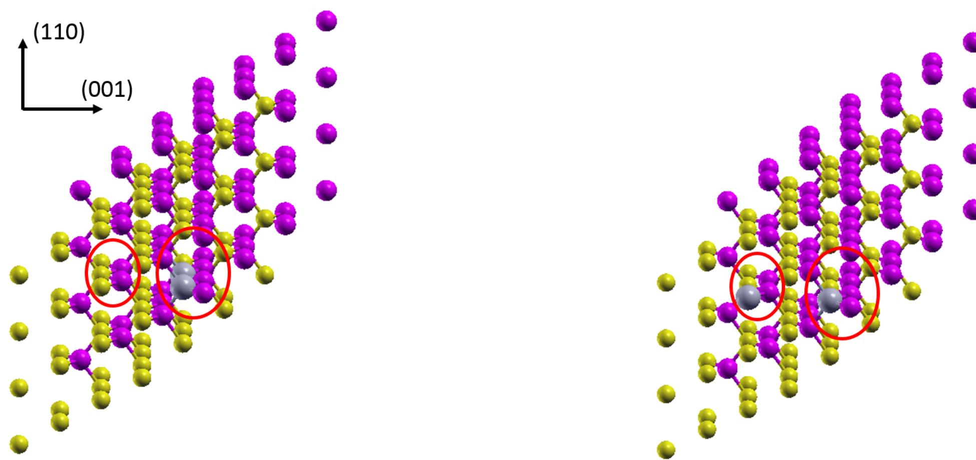
4 Results

4.1 Isoelectronic doping of GaAs with Bi and N

As mentioned in chapter 1, one of the problems we studied concerns the isoelectronic doping of GaAs with Bi and/or N. This is studied theoretically using DFT and experimentally using X-STM.
Using DFT, the pairing of Bi and N atoms in bulk GaAs is investigated. This is done using the nearest neighbour pair configurations. The energies of simulations with either a first or second nearest neighbour configuration are compared and from this it is determined whether the system favours pairing or anti-pairing. The fact whether the system favours pairing or anti-pairing is then used as an indication for whether the dopants tend to cluster or anti-cluster respectively. This can be done since the interaction between dopants is expected to be either attractive or repulsive and to change only in magnitude with distance between dopants, and not in sign.
Since we have results from literature on the behaviour of N and Bi atoms in semiconductors [29] [28] we can see whether that matches the theoretical results found here.

4.1.1 DFT simulation of bulk GaAs doped with Bi and N

DFT is used to theoretically investigate the pair characteristics of Bi and N atoms in GaAs. Pair energies of first and second nearest neighbour Bi-Bi, Bi-N and N-N within bulk GaAs are compared. By comparing these energies it can be found whether the system tends to cluster or anti-cluster. This is done by considering a 128 atom supercell of bulk GaAs where 2 As atoms are replaced with either Bi or N atoms depending on which pair is investigated. The bulk supercell is given in figure 15 to illustrate the difference between the first and second nearest neighbour supercell.

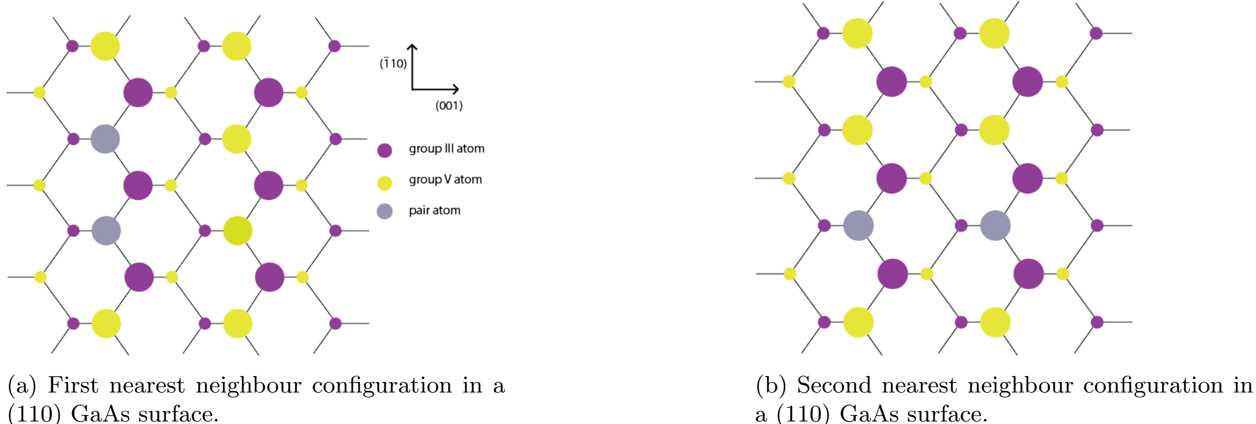


(a) Bulk GaAs supercell with 2 Bi atoms in a first nearest neighbour configuration.

(b) Bulk GaAs supercell with 2 Bi atoms in a second nearest neighbour configuration.

Figure 15: Ga atoms are given in pink, As atoms are yellow and Bi atoms are teal. The zig-zag rows of the (110) surface presented in figure 16 are highlighted in red.

This view of the supercell was chosen in figure 15 to clearly show the well-known zig-zag rows that stand out in the (110) surface of GaAs. In this (110) surface it is easy to see what the first and second nearest neighbour are. This is given in figure 16.
When comparing figures 16a and 15a it is clear that in 15a a first nearest neighbour pair is present since the Bi atoms are on the position of the closest As atoms in a zig-zag row. While in figure 16b and 15b, the Bi atoms are situated in neighbouring zig-zag rows. On these different supercells the convergence tests

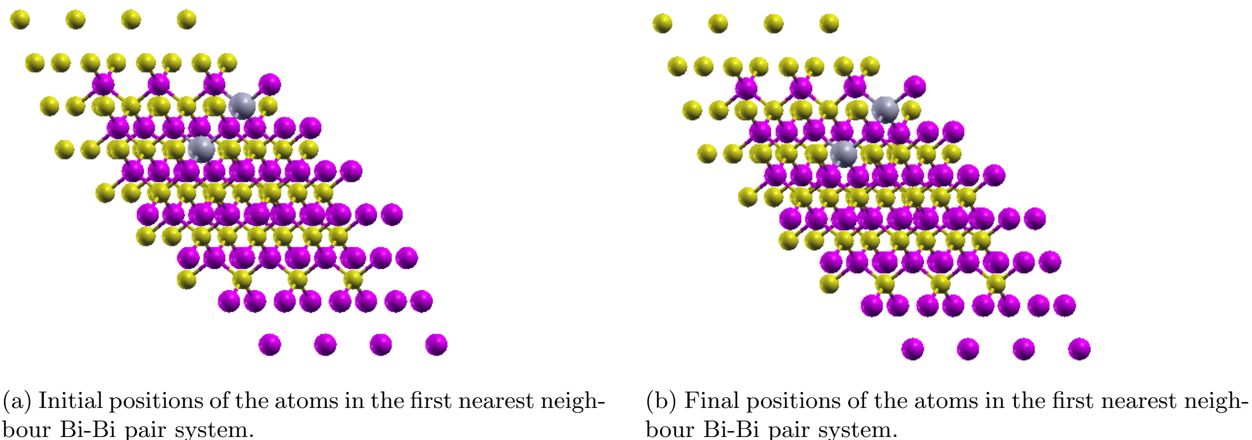


(a) First nearest neighbour configuration in a (110) GaAs surface.

(b) Second nearest neighbour configuration in a (110) GaAs surface.

Figure 16

mentioned in chapter 3 are performed except for the lattice constant test. The lattice constant test is not performed because the equilibrium lattice constant of the first and second nearest neighbour system would vary slightly. This is a problem since the total energies of the first and second nearest neighbour system are compared after relaxation and this is only valid when the total volume of the supercell remains the same. This is not the case when the lattice constant changes so the lattice constant of a pure GaAs system is chosen for all bulk supercells here. This bulk GaAs lattice constant is found by lattice testing a pure GaAs bulk supercell and is found to be 5.62 Å. Which is close to the experimental value of 5.65 Å[45]. DFT is known to underestimate lattice constants due to overestimation of cohesive energies. For all systems here the E_{wfc} is found to be 110 Ry while the E_{ρ} is found to be 900 Ry. The sampling of the brillouin zone is done using a 2x2x2 Monkhorst-Pack grid. The force convergence threshold used for the relaxation calculations is $5 \cdot 10^{-5}$ hartrees/bohr. This is the force on each individual atom. The initial and relaxed state for the first nearest neighbour Bi-Bi pair system are given in figure 17.



(a) Initial positions of the atoms in the first nearest neighbour Bi-Bi pair system.

(b) Final positions of the atoms in the first nearest neighbour Bi-Bi pair system.

Figure 17

Figure 17 is given as an illustration that not much relaxation takes place in the system apart from the local bonds around the impurity atoms. A zoomed in image of the initial and relaxed system around the Bi atoms is given in figure 19. The fact that most of the relaxation happens around the Bi atoms makes sense since we use the equilibrium lattice constant found from lattice testing bulk GaAs so the GaAs is not expected to relax. This means that all the interesting relaxation happens around the impurity atoms so from now on

the images will be focused on that area.
In DFT, the relaxation of a system can be simulated by using the force on each atom that is found when a self consistent calculation is performed. By moving each atom along the force vector an amount of distance relative to the magnitude of the force, a relaxation step can be performed. After such a step, we can perform a new self consistent calculation and repeat the previous steps until the total force on the system is below a certain threshold. In figure 18, the total force of the system as a function of relaxation steps is given.

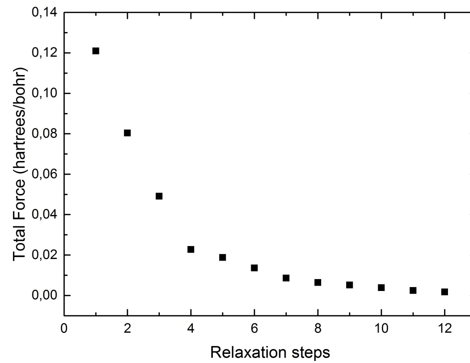
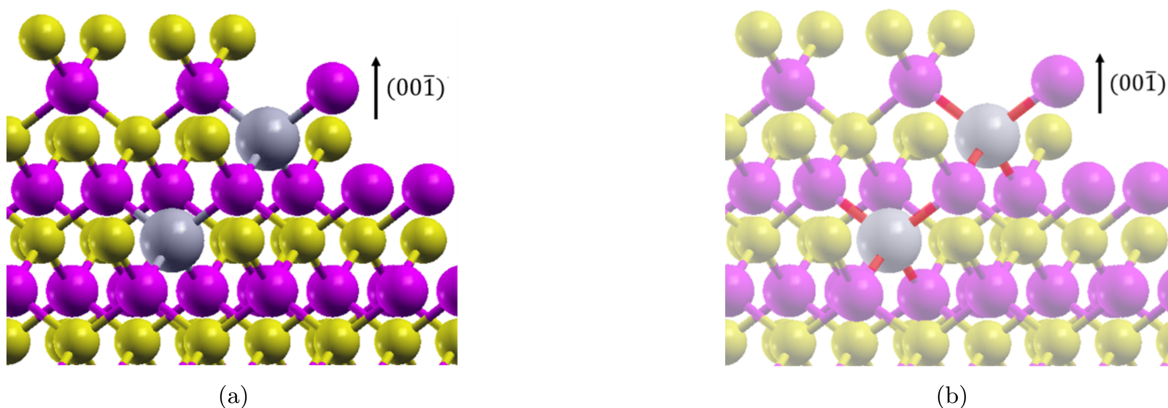


Figure 18: Total force of on the supercell of the first nearest neighbour Bi-Bi pair system as a function of the relaxation steps.

The total force plotted in figure 18 is calculated by adding up all forces on the individual atoms. As can be seen in figure 18, the relaxation is gradual and the system seems to relax to an equilibrium state. This relaxation behaviour is important to check due to the fact that the system can behave weirdly during relaxation. Now relaxation calculations can be performed for the first and second nearest neighbour pairs of Bi-Bi, Bi-N and N-N.

4.1.2 DFT simulation of Bi-Bi pairs in bulk GaAs

First we will look at the Bi-Bi pair results. A zoomed in view of the initial and relaxed positions of the first and second nearest neighbour Bi-Bi pairs are given in figures 19 and 20.



(a) (b)

Figure 19: Initial, a), and relaxed, b), atomic positions of the first nearest neighbour Bi-Bi supercell. The unrelaxed bond lengths are indicated in red in the relaxed supercell in b).

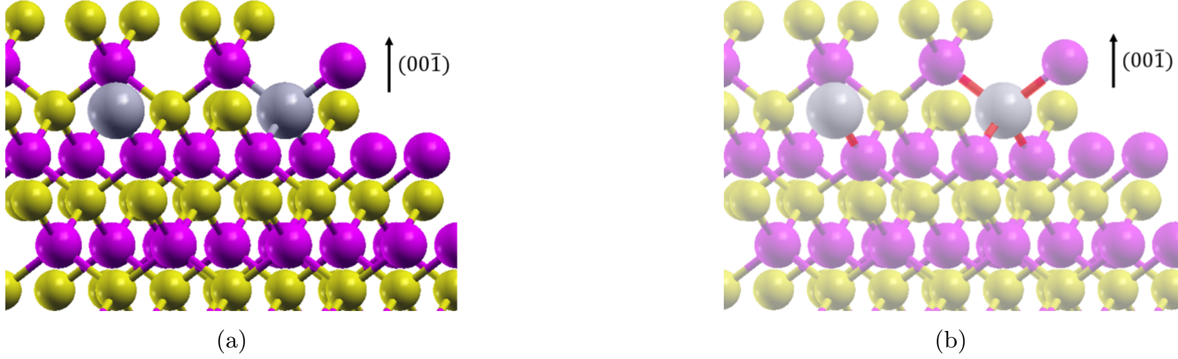


Figure 20: Initial, a), and relaxed, b), atomic positions of the second nearest neighbour Bi-Bi supercell. The unrelaxed bond lengths are indicated in red in the relaxed supercell in b).

It is visible in figures 19 and 20 that after relaxation the Ga atoms that neighbour the Bi impurities are pushed away from these Bi atoms. In the first nearest neighbour case this amounts to an average increase in distance between the Bi atom and its neighbouring Ga atoms of 7.25%. In the second nearest neighbour case this leads to an average increase of the distance by 7.38%. The difference between these two cases is mostly due to the Ga atom situated in between the two Bi atoms in the first nearest neighbour case. In the first nearest neighbour system, this middle Ga atom is pushed away by both Bi atoms, resulting in a bond length that is approximately 0.01 Å longer than the bonds from the Bi atoms to other Ga atoms.
The second nearest neighbour supercell has a total energy which is 30.7 meV lower than the energy of the first nearest neighbour supercell. The total energies of both cells are given in table 1. This energy difference causes the second nearest neighbour configuration to be favourable over the first nearest neighbour one.

4.1.3 DFT simulation of N-N pairs in bulk GaAs

Now we consider the N-N pair systems. Again the initial and final states of the first and second nearest neighbour systems are given in figures 21 and 22. Contrary to the Bi systems, the neighbouring Ga atoms now relax towards the N atoms present. They are also displaced further. The average decrease in bond length between the N and neighbouring Ga atoms is 14.67% and 15.65% for the first and second nearest neighbour pair systems respectively. Again the difference between the first and second nearest neighbour average bond length can be explained by looking at the Ga atom between the two N atoms in the first nearest neighbour supercell. This atom is now pulled towards both N atoms, resulting in an equilibrium position that is further away from both of the N atoms. In this case resulting in a bond length that is approximately 0.1 Å longer than the other three N-Ga bonds. The energies of the relaxed systems are again given in table 1. Similarly to the Bi-Bi pairs, the second nearest neighbour case is energetically favourable. However, the energy difference is significantly larger in the case of N-N pairs, 234.8 meV. So the N atoms would have a much higher probability of anti-clustering than the Bi atoms.

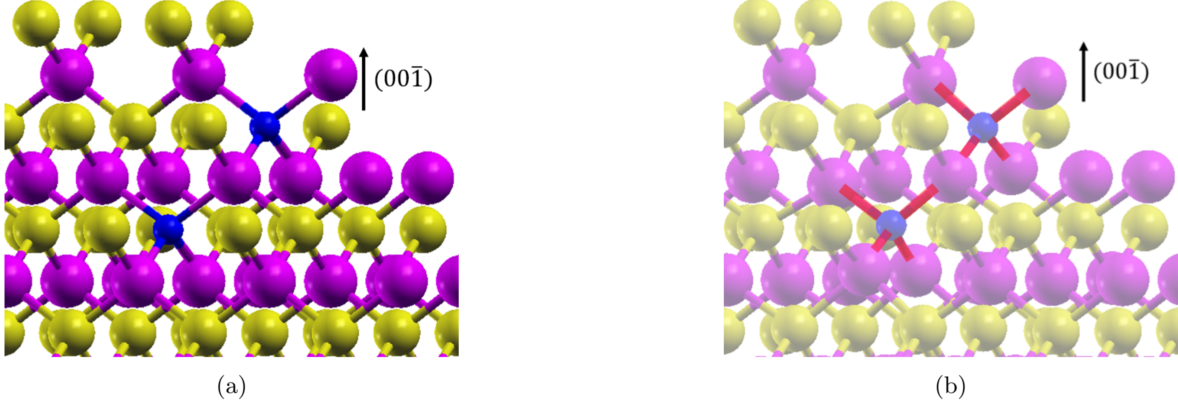


Figure 21: Initial, a), and relaxed, b), atomic positions of the first nearest neighbour N-N supercell. The N atoms are given in blue. The unrelaxed bond lengths are indicated in red in the relaxed supercell in b).

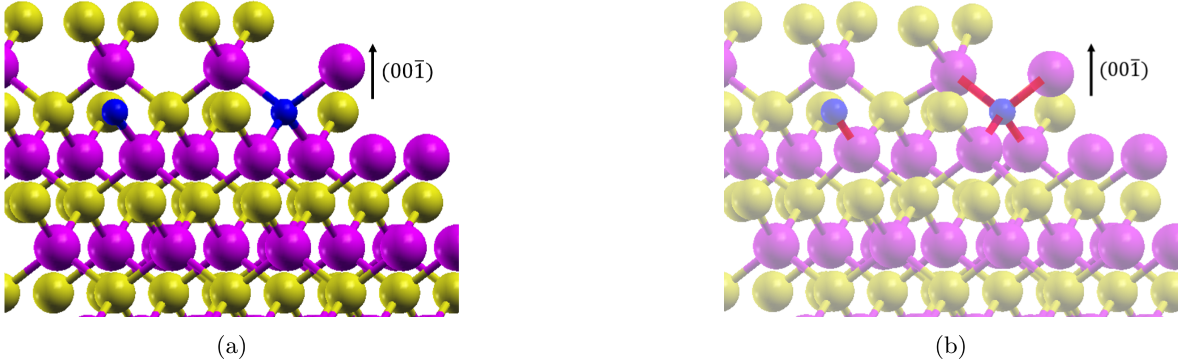


Figure 22: Initial, a), and relaxed, b), atomic positions of the second nearest neighbour N-N supercell. The unrelaxed bond lengths are indicated in red in the relaxed supercell in b).

4.1.4 DFT simulation of Bi-N pairs in GaAs

Lastly, we consider the case of a Bi-N pair in the GaAs bulk supercell. The initial and final atomic positions of the GaAs supercell around these pairs are presented in figure 23 and 24. As expected, the Ga atoms around the Bi atoms are pushed away from the Bi atom in both the first and second nearest neighbour configuration, while the Ga atoms around the N atoms move closer to these N atoms. In the case of the first nearest neighbour supercell, the N-Ga bond length decreases by 16.16% on average and the Bi-Ga bond length increases by 8.3% on average. In the case of the second nearest neighbour supercell, the N-Ga bond length decreases by 15.85% while the Bi-Ga bond length increases by 7.9%. We see now however, that the first nearest neighbour supercell is energetically favourable over the second nearest neighbour one. The energy difference between the two is 110.0 meV, this is a significant energy difference. In order to explain the fact that the first nearest neighbour supercell has a lower total energy we need to look in detail at the relaxation in this supercell.

In order to illustrate why the total energy of the first nearest neighbour supercell is significantly lower than that of the second nearest neighbour supercell, we compare its relaxation to that of the first nearest neighbour N-N pair supercell.
The relaxation of one atom in particular explains the lower energy of the first nearest neighbour Bi-N supercell compared to the second nearest neighbour variant. This atom is the Ga atom present in between the N

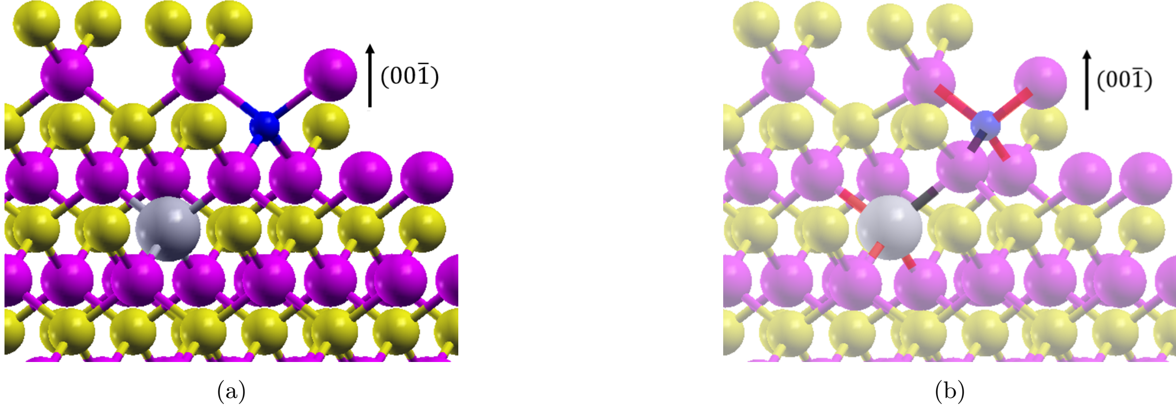


Figure 23: Initial, a), and relaxed, b), atomic positions of the first nearest neighbour Bi-N supercell. The unrelaxed bond lengths are indicated in red in the relaxed supercell in b). The unrelaxed bond lengths between the connecting Ga atom and the N and Bi atom are indicated in black in b).

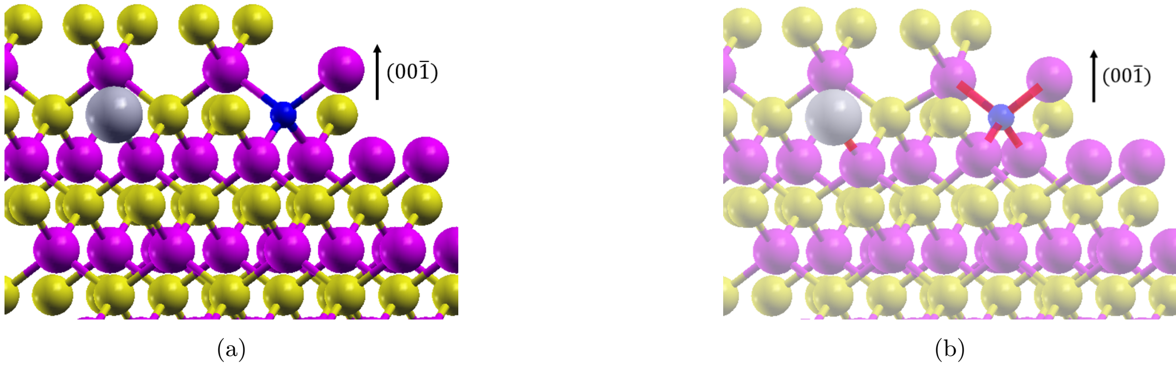


Figure 24: Initial, a), and relaxed, b), atomic positions of the second nearest neighbour Bi-N supercell. The unrelaxed bond lengths are indicated in red in the relaxed supercell in b).

and Bi dopants in the first nearest neighbour supercell, the bonds between this Ga atom and the N and Bi dopants are indicated in black in figure 24b . This Ga atom has both a N-Ga and Bi-Ga bond connected to it and will be referred to as the connecting Ga atom from here. The connecting Ga atom is also present in the first nearest neighbour supercell of N-N pairs. In the case of the N-N supercell, the connecting Ga atom has two N-Ga bonds however.

Both supercells are relaxed and bond lengths are allowed to change. We now consider the bond lengths of the bonds between the N atom and the connecting Ga atom in both the N-N and Bi-N first nearest neighbour supercells.

In the case of the N-N supercell, the bond between the N atom and the connecting Ga atom is 5.0% longer than the bonds between the N atom and other Ga atoms in the system. This is caused by the fact that in the N-N supercell, the connecting Ga atom is also being pulled towards the N atom on the other side of the connecting Ga atom. The N atoms are both pulling on the connecting Ga atom but in opposite directions. This results in an increase of the total energy of the supercell through an increase in N-Ga bond length between both N atoms and the connecting Ga atom.

Now we consider the Bi-N first nearest neighbour supercell. In this supercell, the bond between the N atom and the connecting Ga atom is 0.8 % shorter than the bonds between the N atom and other Ga atoms. This difference is caused by the fact that the connecting Ga atom now has a Bi atom on the other side that repels the Ga atom. The N atom pulls the connecting Ga atom closer while the Bi atom on the opposite

side pushes it away. This results in both a longer Bi-Ga bond length and a shorter N-Ga bond length. Thus the total energy of the system is reduced.

The results of the bulk pair calculations of Bi-Bi, N-N and Bi-N pairs are summarised in table 1. When we compare the results found there using DFT calculations to the results from Krammel et al. and Plantenga et al. shown in section 1.2, the results for N-N pairs match well but the results for Bi-Bi pairs do not. In the case of N doped GaAs, the results found here using DFT and the experimental results of Plantenga et al shown in figure 2 indicate that it is not energetically viable for N atoms in GaAs to form first nearest neighbour pairs. Krammel et al. however, found that Bi tends to cluster while the results found here seem to indicate that Bi anti-clusters.

	1st **nearest neighbour**	2nd **nearest neighbour**	ΔE **(meV)**	**Pairing/Anti-pairing**
Bi-Bi	-12480.14982091 Ry	-12480.15208042 Ry	-30.7	Anti-pairing
N-N	-12210.31502149 Ry	-12210.33228260 Ry	-234.8	Anti-pairing
Bi-N	-12345.25530258 Ry	-12345.24721970 Ry	110.0	Pairing

Table 1: Energies of bulk GaAs supercells with different impurity pairs. When the 1st nearest neighbour pair is energetically favourable the system is indicated as a pairing system and when the 2nd nearest neighbour state is energetically favourable the system is indicated as anti-pairing. Energy difference between 2nd and 1st nearest neighbour is given in meV.

Since the theoretical results predict Bi anti-clustering in GaAs while we know from Krammel et al. [28] that Bi dopants cluster in InP, X-STM measurements are used to check the behaviour of Bi dopants in GaAs. The results of these experiments are shown in section 4.1.5.

4.1.5 X-STM measurements of Bi doped GaAs

Since the DFT results for Bi-Bi pairing in GaAs shown in section 4.1.2 do not match with the experimental results found by Krammel et al. [28] for Bi-Bi pairing in InP, we now use X-STM to check the behaviour of Bi-Bi pairing in GaAs experimentally. In order to investigate the behaviour of Bi dopants in GaAs with X-STM, sample GaAs 35-2104 was grown by Anthony Rice and Kirstin Alberi of the National Renewable Energy Laboratory (NREL). The schematic overview of the different layers in this sample is given in figure 25. This sample will from here be referred to as the Bi doped sample.

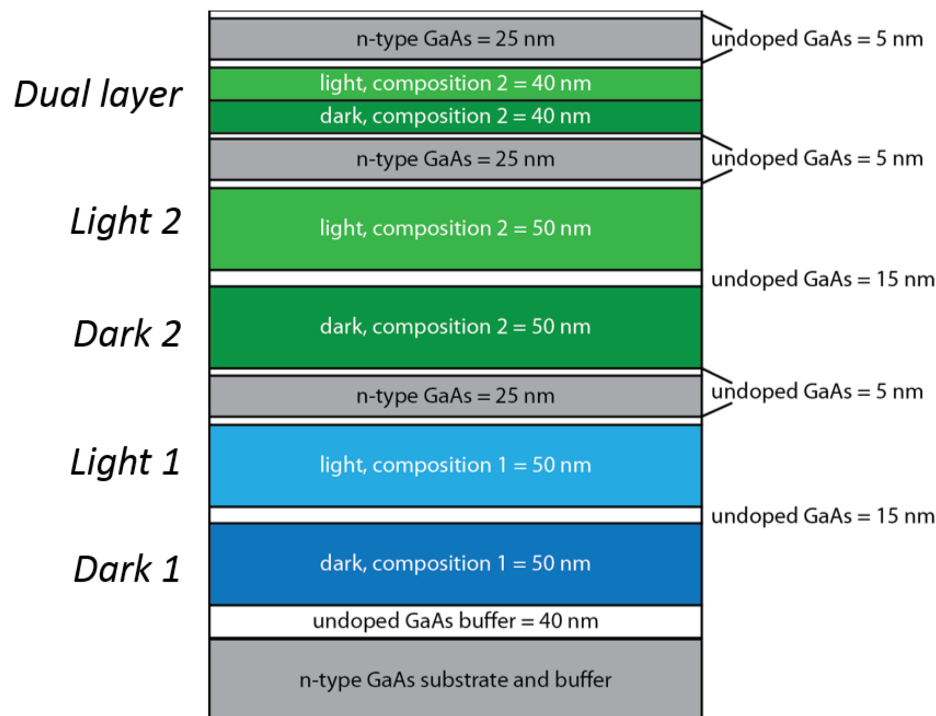


Figure 25: Schematic overview of the different layers in GaAs 35-2104. Sample regions with composition 1 and 2 have an intended Bi concentration of 0.4% and 0.7% respectively. The "light" sample regions are grown with light assisted MBE while the "dark" regions are grown using regular MBE.

The Bi sample shown in figure 25 has several different regions of interest. The sample regions with compositions 1 and 2 have an intended Bi concentration of 0.4% and 0.7% respectively. In these regions, the pairing of Bi atoms in GaAs can be investigated. The "light" and "dark" regions can be used to study the effects of light assisted MBE on the Bi incorporation since the "light" regions are grown using light assisted MBE. The different Bi doped sample regions will be referred to as (from top to bottom): the dual layer, "light 2", "dark 2", "light 1" and "dark 1".

An X-STM study of the Bi doped GaAs sample described before will now be discussed. Several 50x80 nm overview images are made of of the different Bi containing layers of this sample. The Bi pairs and concentration were then investigated by counting the Bi atoms present. A typical overview image used to count the Bi pairs is given in figure 26.

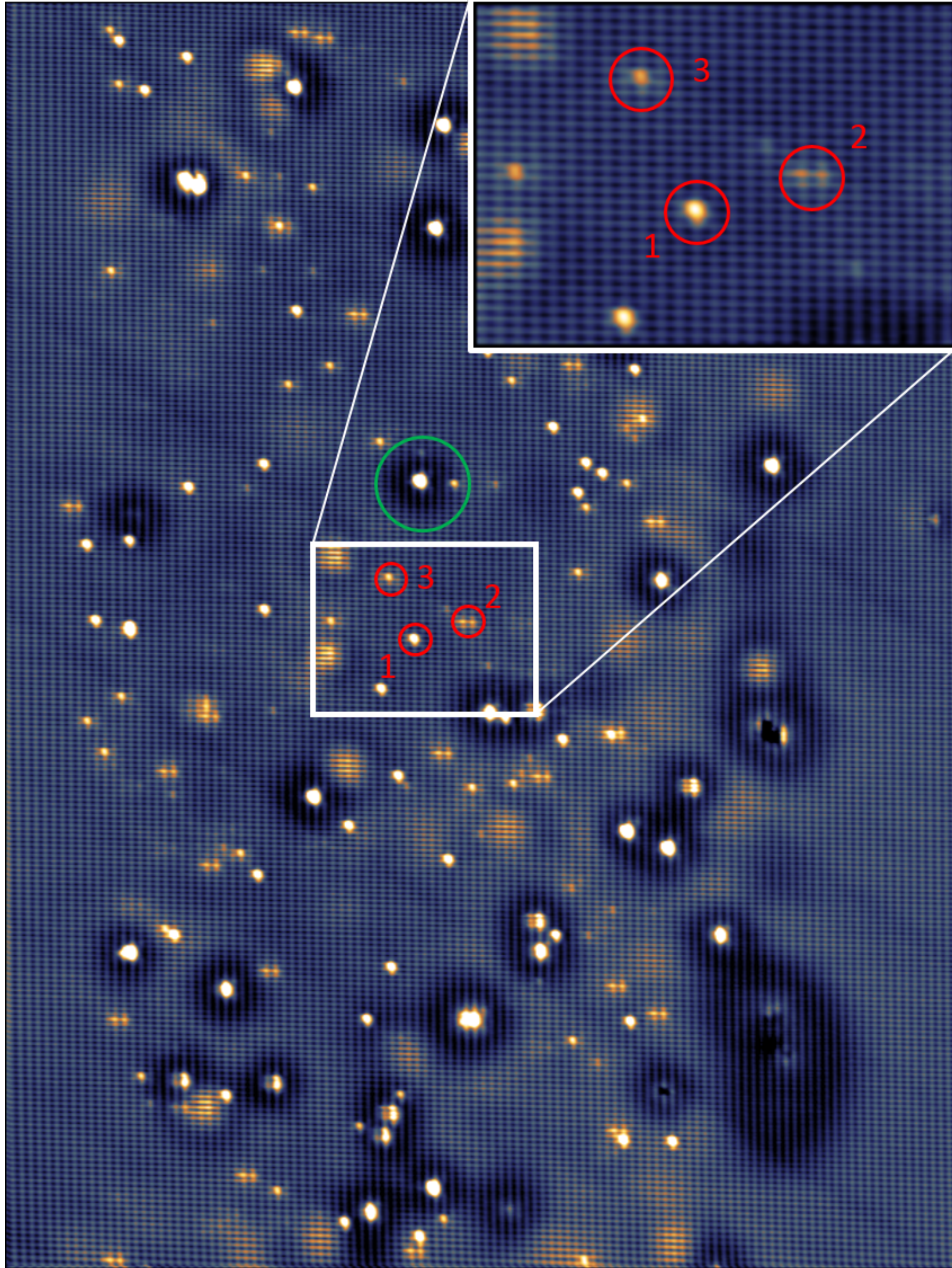


Figure 26: 60x80 nm filled state X-STM image of the Bi doped sample. First, second and third layer Bi atoms are labeled with 1, 2 and 3 respectively. A Si dopant is highlighted in green. The inset shows a zoomed in view of the area with the marked Bi features. (-3V, 50pA, 77K)

A surface (first) layer Bi atom is marked in figure 26 with the number 1, a second layer Bi atom below the surface with the number 2 and a third layer Bi atom below the surface is marked with number 3. Also, several Si dopants are visible, one such dopant is highlighted with a green circle. This particular overview was made in the dark 1 region of the Bi doped sample. In total 0.16 ± 0.01 μm^2 of area across the "light 1", "dark 1", "light 2", "dark 2" and dual layer regions are investigated in terms of Bi concentration and Bi pair occurrence. A schematic of all the Bi-Bi pair configurations considered in this work are given in figure 27. While experimental examples from X-STM images are given in figure 28.
The fifth nearest neighbour pair configuration is not discussed here since this is not a surface pair configuration. The fifth nearest neighbour of a surface Bi atom lies in the second layer below the surface. This makes the fifth nearest neighbour pair more difficult to recognise in X-STM images and is therefore omitted from the counting.

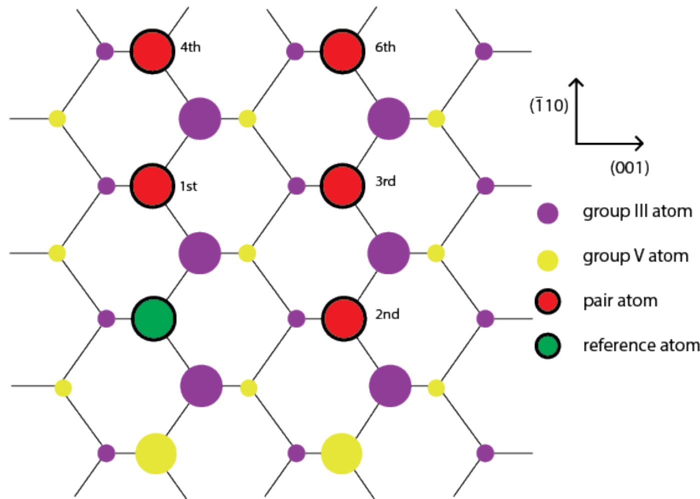


Figure 27: Schematic illustration of the first through sixth nearest neighbour configuration of (110) surface Bi-Bi pairs.

When comparing figures 28 and 27, keep in mind that these are filled state images so the group V atomic sites are visible. This means that the Ga atoms do not show up in figure 28 and we are instead looking at the sub-lattice of As atoms with Bi atoms present on As sites. The surface layer Bi atoms show up as a bright contrast localised on one such an As site. This makes first layer Bi atoms very suitable candidates when it comes to counting. Due to the localised nature of these first layer Bi atoms, the pairs are distinguishable and show up in a form very similar to what is given schematically in figure 27. The first nearest neighbour surface Bi pairs show up as two neighbouring bright spots in the ($\bar{1}$00) direction without a row of corrugation in between them while the fourth nearest neighbour configuration does have this one row of separation between the Bi atoms. The second nearest neighbour configuration shows two Bi atoms side by side in the (001) direction while the third and sixth nearest neighbour pairs have the second Bi atom also translated one and two corrugation rows in the ($\bar{1}$10) direction respectively.
First we consider the concentration of Bi atoms in the different sample regions. The growers provided us with an XRD measurement of this sample and gave an estimation of the Bi concentrations in the different regions. The results of this XRD measurement are given in figure 29.

The estimations of the Bi concentration made by the growers from the XRD measurement shown in figure 29 were a Bi concentration of approximately 0.05% in the composition 1 regions and 0.15% in the composition 2 regions. The Bi counting performed in this study yielded the Bi concentrations found in table 2.

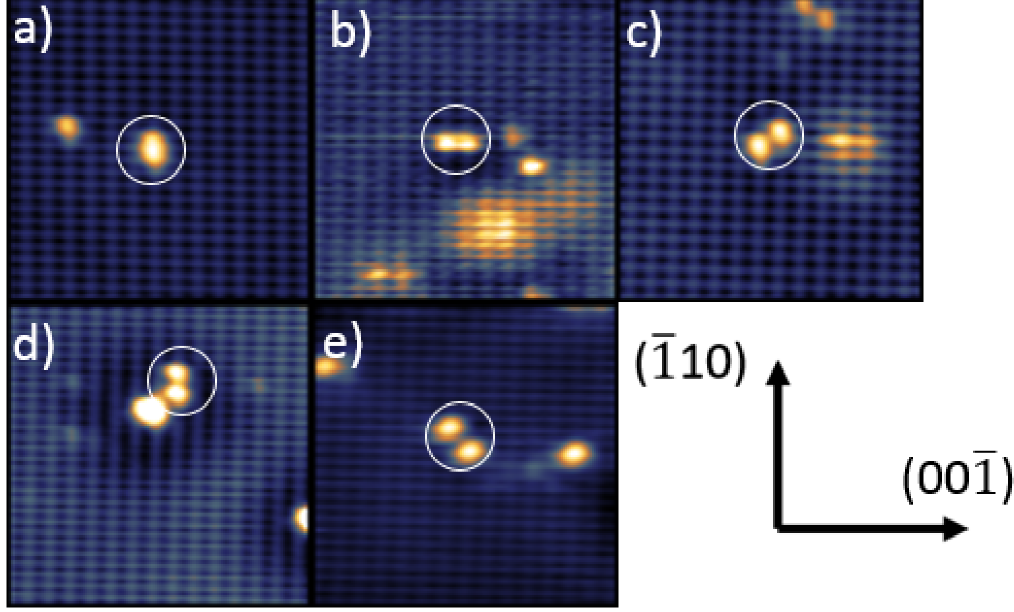


Figure 28: 7.5x7.5 nm filled state X-STM images zoomed in on a first, second, third, fourth and sixth nearest neighbour surface Bi-Bi pair in figures a) to e) respectively. (-3V, 50pA, 77K)

	Bi concentration		
Sample region	**STM**	**XRD**	**Intended**
Dark 1	$0.22 \pm 0.01\%$	0.05%	0.4%
Light 1	$0.18 \pm 0.01\%$	0.05%	0.4%
Dark 2	$0.45 \pm 0.02\%$	0.15%	0.7%
Light 2	$0.44 \pm 0.04\%$	0.15%	0.7%

Table 2: Bi concentrations of the different sample regions of the Bi doped GaAs sample as found from the Bi counting in X-STM measurements and the XRD measurement. The concentration intended by the growers is also given.

The uncertainties in table 2 are caused by three main factors. First a square root of the number of found Bi atoms is taken as a baseline uncertainty due to the Poisson statistics of the Bi atoms throughout the material. Second, due to the Si atoms present in the sample as seen in figure 26 some Bi atoms could be masked. To take this into account, an additional 1% of the total Bi atoms is assumed to be hidden in this way due to approximately 1% of the surface being covered by the Si atoms. Lastly, the measured surface area of the sample has an uncertainty of a few nm in both height and width. If we now compare the expected concentrations from XRD to our measured concentrations we find that there is a discrepancy. The sample regions with composition 1 and 2 have a Bi concentration 4 and 3 and times higher than the concentration found with XRD respectively. This can be explained by the fact that the result of the XRD measurement performed on this sample was hard to fit to a Bi concentration.
We can also compare the actual Bi concentration found in the Bi doped sample to the intended concentration as shown in table 2. The sample regions with composition 1 and 2 have an intended Bi concentration 2 and 1.57 times higher than the actual Bi concentration. This can be caused by the fact that this sample ran hotter than expected during the growth, which can influence the Bi incorporation. The ratio of the Bi composition in the different sample regions is still very similar. The intended Bi concentration ratio between regions with composition 1 and 2 was 0.51 while the actual ratio is 0.45± 0.06.
The effects of exposure to light during the MBE growth of the material on the Bi incorporation can also be

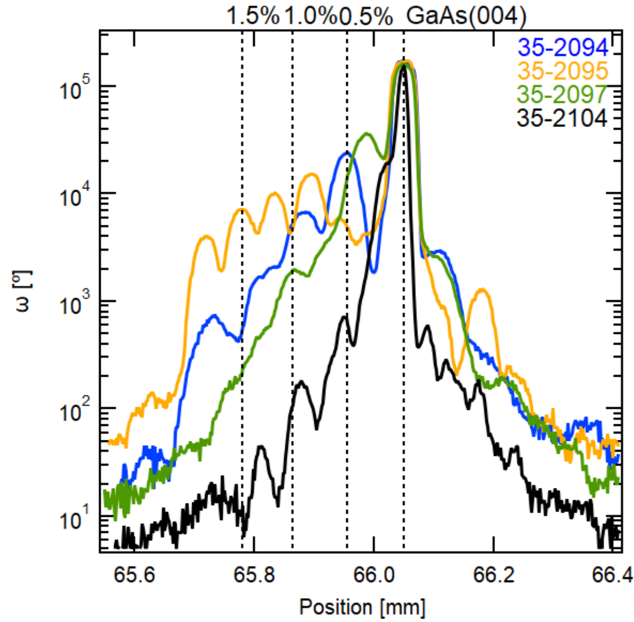
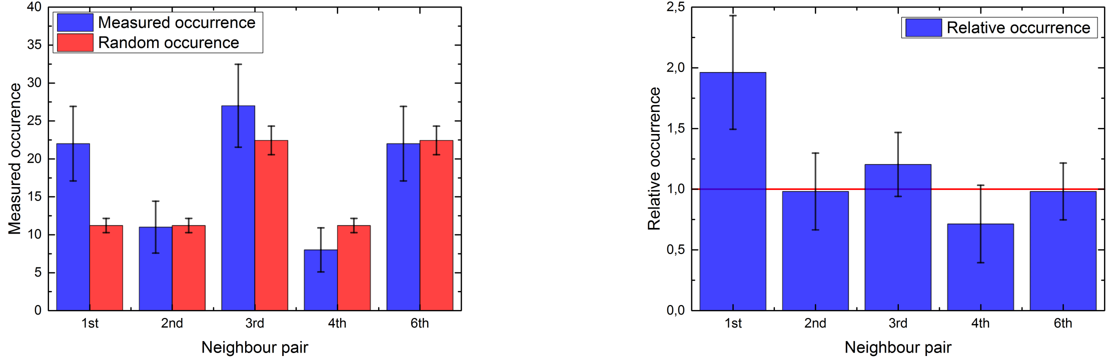


Figure 29: Results of the XRD measurement of the Bi doped GaAs sample used.

seen in table 2. Both compositions show a decrease in Bi concentration in the sample regions grown with light assisted MBE. The difference is within the margin of error for the sample regions with Bi composition 2, but not for the sample regions with Bi composition 1. This decrease in Bi concentration due to exposure to light during growth is also what was observed in literature as discussed in section 1.2.

Now we consider the occurrence of the surface layer Bi pairs as shown in figure 28. The results of the surface layer Bi-Bi pair counting are shown in figure 30 as well as the expected occurrence of each nearest neighbour pair if the Bi atoms were randomly distributed.
The results presented in figure 30 need some further explaining. First, the error in the measured pair occurrence is once again built up from the Poisson statistics of the Bi atoms and an additional 1% error for miscounting and Si covered areas.
Second, the calculation of the random occurrence as shown in figure 30 will be discussed. This random occurrence is calculated from the Bi concentration, the total group V atomic positions present in the considered area and the amount of configurations in which a pair can show up. The concentration used to calculate the random occurrence is the average concentration over all different sample regions that were considered for the counting of the pairs, 0.28 ± 0.01 %. This average concentration is a weighted average of the concentrations shown in table 2, where the concentration of each different sample region is given a relative weight corresponding to the measured surface area of that sample region. The random occurrence of Bi pairs is now calculated by considering the chance a Bi atom is present at a certain lattice site, 0.28 ± 0.01 %, and multiplying this by the chance that a Bi atom is present on the nearest neighbour site, 0.28 ± 0.01 %. This probability now needs only to be multiplied with the number of lattice sites present in 0.16 ± 0.01 μm^2 of GaAs and we find the probability for the occurrence of this Bi pair configuration. Now there are different positions that lead to equivalent nearest neighbour pairs. For the first, second and fourth nearest neighbour pairs there are two different positions where the second Bi atom can be placed with respect to the first one which lead to one of these neighbour pairs. To illustrate, we look at the first nearest neighbour pair. Looking at figure 27, a first nearest neighbour pair can be formed with the reference atom by either placing a second Bi atom in the $(\bar{1}10)$ or the $(1\bar{1}0)$ direction. Similar arguments can be made regarding the second and fourth nearest neighbour pair. This means that the probability for the pair configuration found earlier



(a) Absolute value of the measured Bi pairs. The absolute amount of expected Bi pairs if the distribution was random is also given.

(b) Relative occurrence of the Bi pairs for each pair configuration. The expected occurrence for a random distribution is set at one for each pair configuration.

Figure 30: Results of the surface layer Bi-Bi pair counting in the Bi sample. The first through sixth nearest neighbour pair occurrences are presented in absolute and relative form. The Bi pairs found throughout all the different sample regions are considered in the results shown here.

needs to be multiplied by two to find the expected random occurrence of the first, second and fourth nearest neighbour pairs. The third and sixth nearest neighbour pair however, have four different positions where the second Bi atom can be placed relative to the reference atom that lead to that neighbour pair. This means that for these nearest neighbour pairs, the probability for the pair configuration needs to be multiplied by four instead. This is reflected in the double random occurrence indicated for these two nearest neighbour pairs in figure 30.
The Bi pairs found in all the different sample regions are considered in the results shown in figure 30. This means that Bi pairs from "light" regions, "dark" regions and regions with different concentrations are combined into these results.

When comparing the occurrences of the nearest neighbour pairs shown in figure 30 to their random occurrences, it is clear that the second, third and sixth nearest neighbour pairs match within their range of uncertainty with their random occurrence. The fourth nearest neighbour pairs occur less than expected and the first nearest neighbour pairs occur approximately twice as often as expected from a random distribution of Bi atoms.
The fact that first nearest neighbour Bi pairs appear twice as much as to be expected from a random distribution of Bi dopants in GaAs matches with the findings of Krammel et al. [28] for Bi doped InP. This does not match with the DFT calculations shown in section 4.1.2 though. This leads us to believe that the Bi atoms do not cluster in bulk but rather cluster during the growth of the GaAs, after which they remain in these positions. We did already see that the energy difference between the first and second nearest neighbour Bi-Bi pair in bulk GaAs is only 30.7 meV. So rearranging of the Bi atoms in bulk GaAs is not likely. To investigate this using DFT, a different supercell is needed. The formation of Bi-Bi pairs in GaAs during the growth of the material is studied using DFT in section 4.1.6.
Another observation is the fact that a low number of fourth nearest neighbour pairs is found. This could be linked to the higher occurrence of the first nearest neighbour pairs. As shown in figure 27, both the first and fourth nearest neighbour atom are positioned along the $(\bar{1}10)$ direction. This means that it is easier for Bi atoms in the fourth nearest neighbour configuration to switch to the first nearest neighbour configuration. Both the first and fourth nearest neighbour dopant are also situated on top of the (001) surface during growth, which could also facilitate the switching between these two pair configurations.

4.1.6 DFT simulation of on Bi-Bi pairs in the growth surface of GaAs

In order to further theoretically investigate the reason behind the high amount of first nearest neighbour Bi-Bi pairs found in GaAs, the possibility of Bi atoms to cluster during growth of the material is studied here. The growth surface of the Bi doped GaAs that is investigated with STM is the (001) surface, which is a common growth surface for GaAs [22] [46] [47]. Here we investigate the first and second nearest neighbour Bi pairs in a (110) GaAs surface as indicated in figure 16, but in an As terminated (001) supercell of GaAs. The first nearest neighbour system is shown in figure 31. This is a 160 atom As terminated (001) surface GaAs supercell with 2 Bi atoms on As sites. There is 6 Å vacuum above and below the slab to simulate a surface and the bottom 4 atomic rows are fixed to simulate bulk below the surface. The E_{wfc} used is 110 Ry, the E_ρ used is 900 Ry and the Brillouin zone is sampled with a 2x2x1 Monkhorst-Pack grid. The lattice constant used is 5.65Å. Again the lattice constant is not tested since the volume of the first and second nearest neighbour cell would then be different and energies could not be compared easily. The force convergence threshold used in the relaxation calculations will be discussed later.

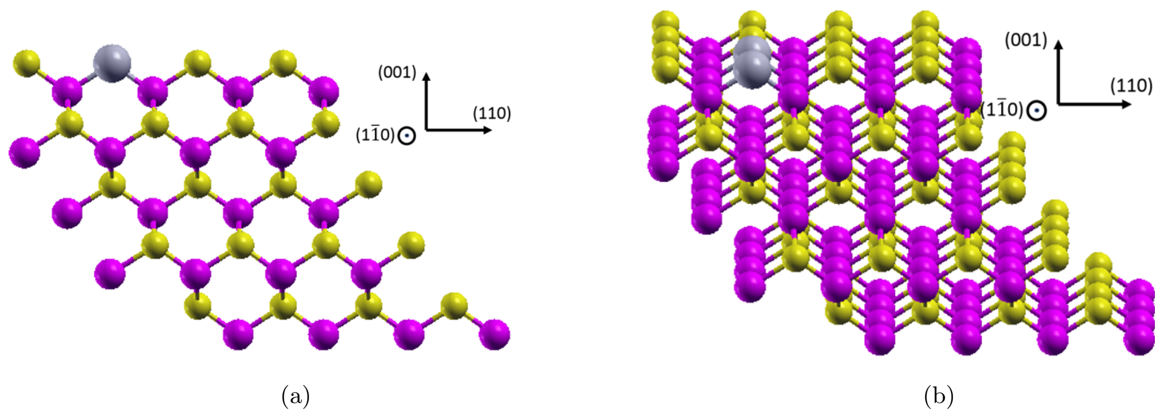


Figure 31: Side view (a) and tilted side view (b) of the supercell of the first nearest neighbour Bi-Bi pair system in a (001) GaAs surface.

For the second nearest neighbour system, the same values are used for E_{wfc}, E_ρ , the lattice constant and the k-grid dimensions as for the first nearest neighbour system. The second nearest neighbour cell is presented in figure 32.

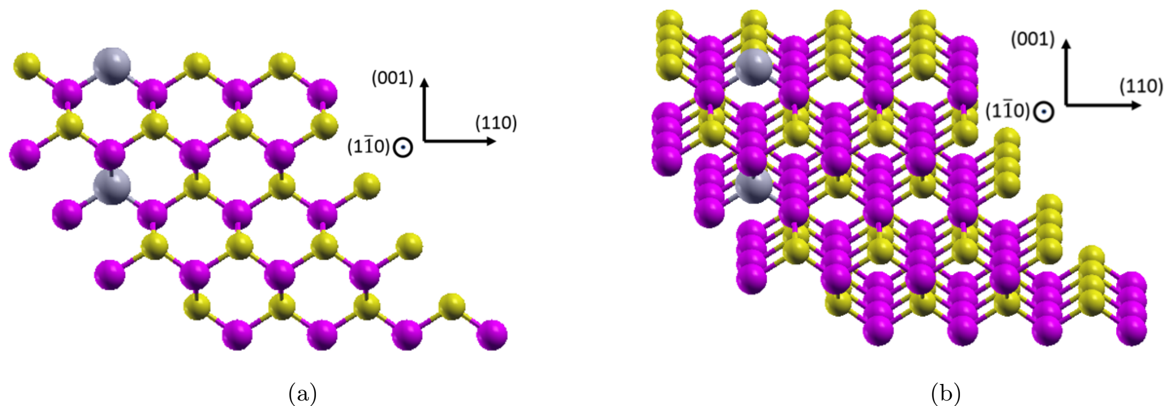


Figure 32: Side view (a) and tilted side view (b) of the supercell of the second nearest neighbour Bi-Bi pair system in an As terminated (001) GaAs surface.

When comparing figures 31 and 32 to figure 16 it is important to remember that in figures 31 and 32 the (110) direction is to the right so the (110) surface is a vertical cross section of figures 31 and 32 through the Bi atoms. The familiar zig zag row around the Bi atoms in the top layer is harder to recognize since the Ga half of this zig zag row is the next layer to be grown onto the terminated As atoms.
Now relaxation calculations are performed on these two systems. This is far from trivial since the (001) Bi doped GaAs surface is known to undergo various surface reconstructions that are highly dependent on the growth conditions and the growth temperature in particular [23] [48]. A major difficulty when describing temperature dependent surface reconstructions with DFT is that DFT is in essence a zero temperature method, so a strong temperature dependent phenomenon will be difficult to describe. These surface reconstructions will not be explicitly be considered here since it is beyond the scope of this work but some indications towards their existence will be discussed.
The total force of this supercell as a function of relaxation steps is given in figure 33. The atomic positions of the supercell after different amounts of relaxation steps are also given in figure 34.

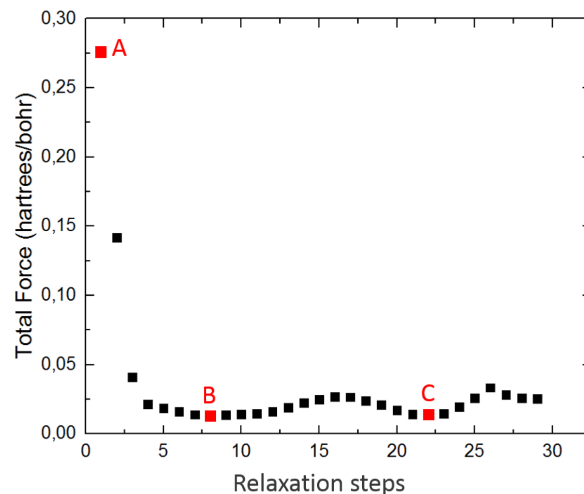


Figure 33: Total force of the first nearest neighbour Bi-Bi pair system in a (001) GaAs surface as a function of the relaxation steps. The atomic positions as marked on A, B and C are given in figure 34.

The total force of the system as a function of relaxation steps is given in figure 33 to illustrate that this is not a relaxation to a certain equilibrium. The system goes to a minimum force and then oscillates around this minimum total force by distorting the surface atoms in a major way. This is shown in figure 34 by comparing the surface atoms in situations A, B and C. From situation A to situation B we deal with a gradual relaxation, as seen in figure 33. The surface As atoms and Ga atoms bonded to the Bi atoms move away slightly from these Bi atoms, as seen with bulk Bi atoms. From situation B to C however, the surface As atoms in the $(\bar{1}10)$ and $(1\bar{1}0)$ direction move significantly closer to the Bi atoms. This is a major distortion of the system and the first indication that this growth surface shows complicated behaviour. The different shapes of the surface in situation B and C while the total energy remains similar leads to an active growth surface. This is in line with the temperature dependent surface reconstructions seen during growth of Bi doped GaAs. However, these surface reconstructions are not considered explicitly in this work. So the final position of these supercells is considered situation B, where the total energy is significantly lowered, but the surface is not completely restructured yet.

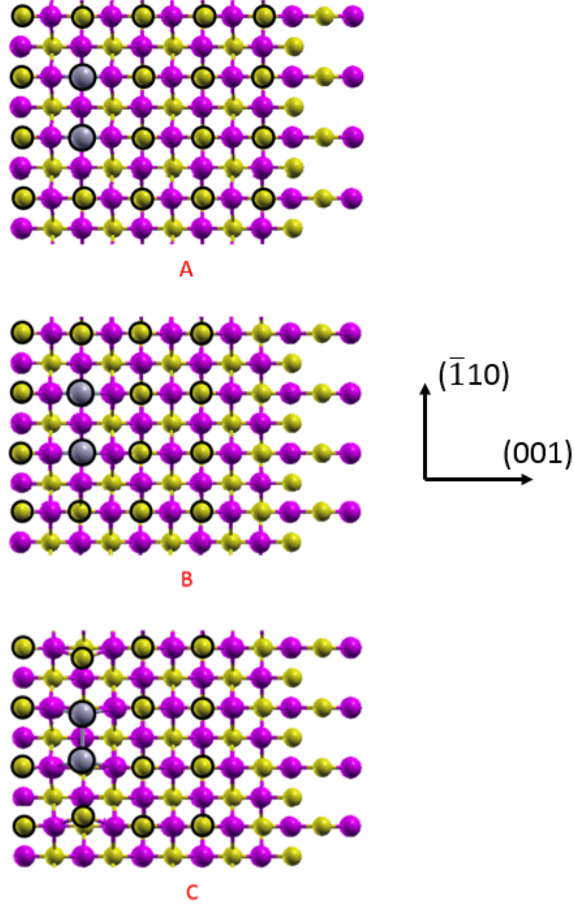


Figure 34: Atomic positions of the (001) surface slab GaAs supercell from a top view after the amount of relaxation steps indicated in figure 33. The top layer As and Bi atoms are highlighted.

	1st **nearest neighbour**	2nd **nearest neighbour**	**Δ E (eV)**	**Pairing/anti-pairing**
(001) **surface Bi-Bi**	-1.5527.90324874 Ry	-15527.80492454 Ry	1.34	Pairing

Table 3: Energies of the (001) GaAs slab supercells with either a first or second nearest neighbour Bi-Bi pair present.

The total energies of the relaxed first and second nearest neighbour supercells are given in table 3.
As can be seen in table 3, the first nearest neighbour system has a significantly lower total energy than the second nearest neighbour system. The energy difference is quite big when compared to the energy difference between the first and second nearest neighbour Bi-Bi pair in bulk GaAs as presented in table 1. This means that Bi atoms are very likely to cluster together, at least in first nearest neighbour pairs, during (001) growth of GaAs.

This result also matches the experimental result found by Krammel et al. [28]. They find a very high amount of first nearest neighbour Bi-Bi pairs in InP. Using a simulated growth surface, we find that the first nearest neighbour configuration is energetically very favourable, 1.36 eV, over the second nearest neighbour configuration. This leads to Bi dopants clustering together during the growth of GaAs on the growth surface. The energy difference between the first and second nearest neighbour configuration in bulk GaAs, section

4.1.2, is only 30.7 meV in favour of the second nearest neighbour configuration. This means that Bi dopants are likely to cluster together during the growth of GaAs and they are unlikely to anti-cluster once they are embedded in bulk GaAs. Therefore it makes sense that we experimentally observe a very high amount of first nearest neighbour Bi-Bi pairs in GaAs in section 4.1.5.

Some things are important to keep in mind when considering the DFT results shown here.
As mentioned before, we do not take into account the surface reconstructions of the (001) surface of GaAs during growth. The different surface reconstructions can greatly influence the orientation of Bi atoms in such a surface with respect to each other, which can in turn influence the energies of the Bi-Bi pairs.
Also, apart from the distance between the Bi atoms there is another difference when comparing the supercells describing the first and second nearest neighbour configuration presented here. Namely the type of atoms exposed to the vacuum. In the case of the first nearest neighbour supercell two Bi atoms are exposed to the vacuum as shown in figure 31. While in the second nearest neighbour supercell, only one Bi atom is exposed to the surface. Since the presence of Bi atoms induces strain in GaAs, having these Bi atoms present at the surface will likely reduce this strain and therefore reduce the total energy of the system.

The next step is to investigate the behaviour of Bi pairs in GaAs co-doped with Bi and N and to study whether the Bi and N dopants cluster or anti-cluster. From the DFT calculations presented in section 4.1.4, the conclusion was drawn that the Bi and N dopants should cluster together since their first nearest neighbour configuration was energetically favourable over their second nearest neighbour configuration.
This will now be studied experimentally using X-STM measurements on a GaAs sample that is co-doped with Bi and N.

4.1.7 X-STM measurements of GaAs co-doped with Bi and N

One of the results of section 4.1.4 is that for Bi-N pairs in GaAs, the first nearest neighbour configuration is favourable over the second nearest neighbour configuration. The Bi and N dopants are therefore expected to cluster together in GaAs co-doped with Bi and N. Another interesting characteristic of GaAs co-doped with Bi and N is the influence of N dopants on the incorporation of Bi. The presence of N could influence the Bi incorporation as mentioned in section 1.2.
In order to study these effects experimentally a GaAs sample co-doped with Bi and N is investigated using X-STM.
This is sample 35-2120 GaAsBiN again grown by Anthony Rice and Kirstin Alberi (NREL) and will from here be referred to as the co-doped sample. The structure of the co-doped sample is given schematically in figure 35.

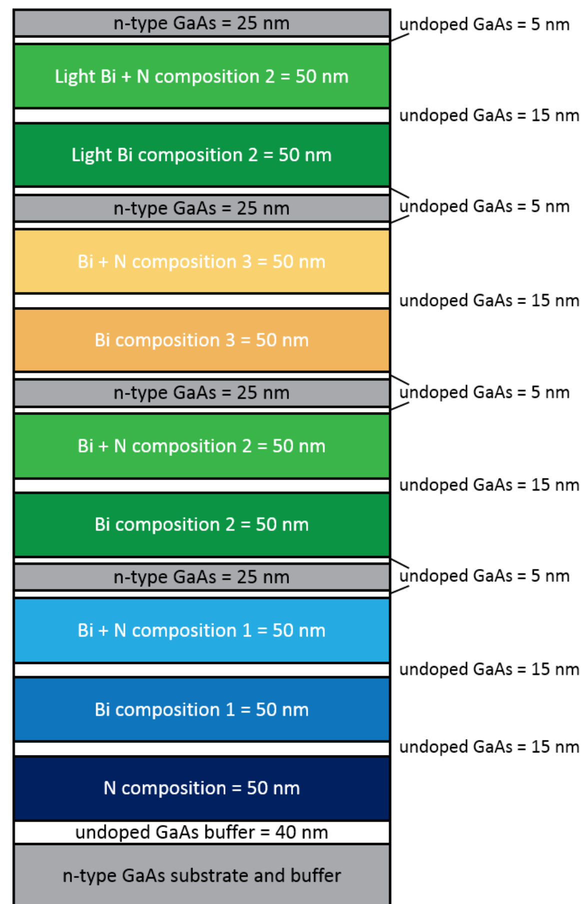


Figure 35: Schematic overview of the different layers in GaAs 35-2120. Sample regions with composition 1, 2 and 3 have an expected Bi concentration of 0.4%, 0.8% and more than 1% respectively. The N concentration is constant throughout the N doped regions and is expected to be 0.25%. The "light" sample regions are grown with light assisted MBE.

The Bi concentration is varied in the different sample regions similarly to the Bi doped sample. XRD was again used by the growers to estimate the Bi and N concentrations in the different layers. This XRD measurement estimates the Bi concentrations to be 0.4%, 0.8% and more than 1% for the regions with Bi compositions 1, 2 and 3 respectively. The N concentration is the same in all N doped regions and is estimated with XRD at 0.25%.

Next we look at the results of the X-STM measurements of the co-doped sample. A typical X-STM image of a co-doped region within this sample, in this case Bi + N composition 2, is given in figure 36.

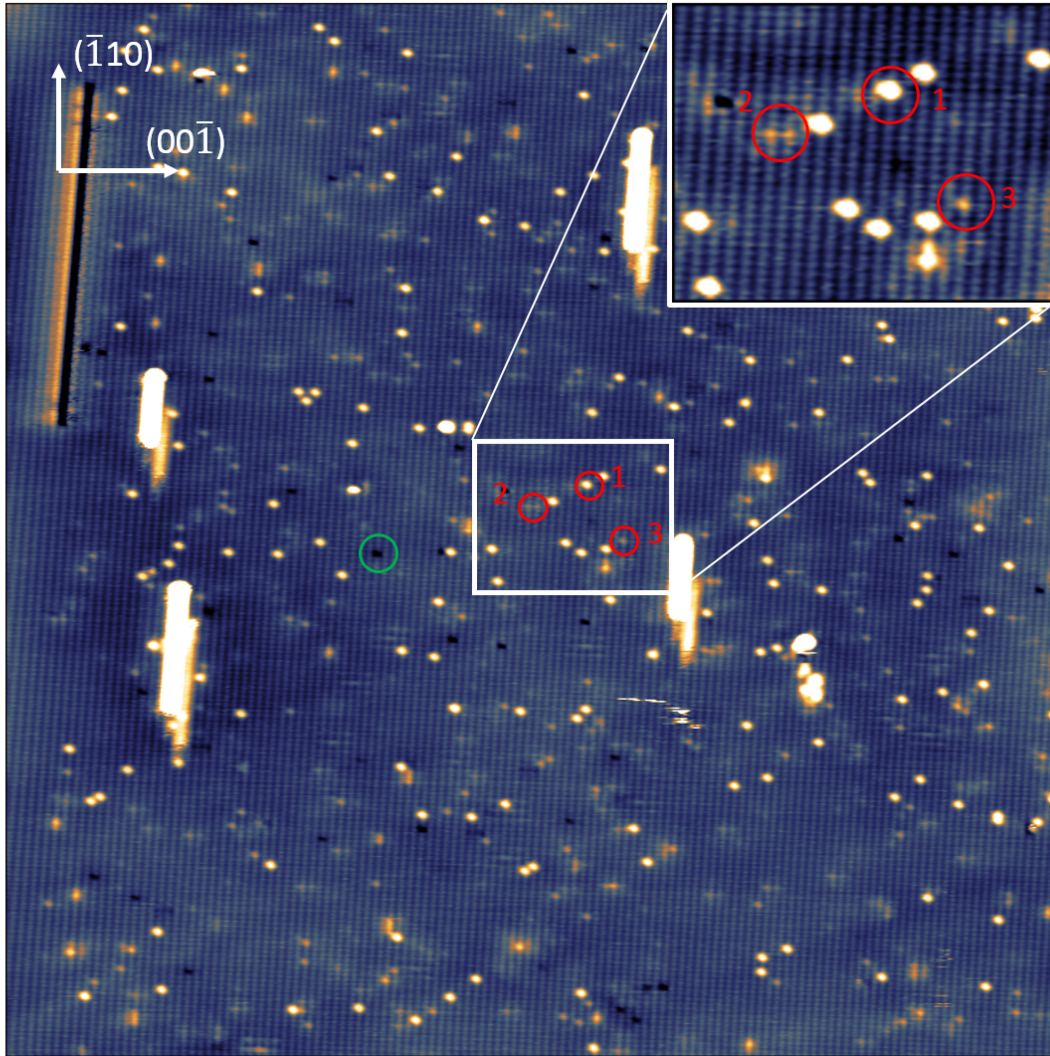
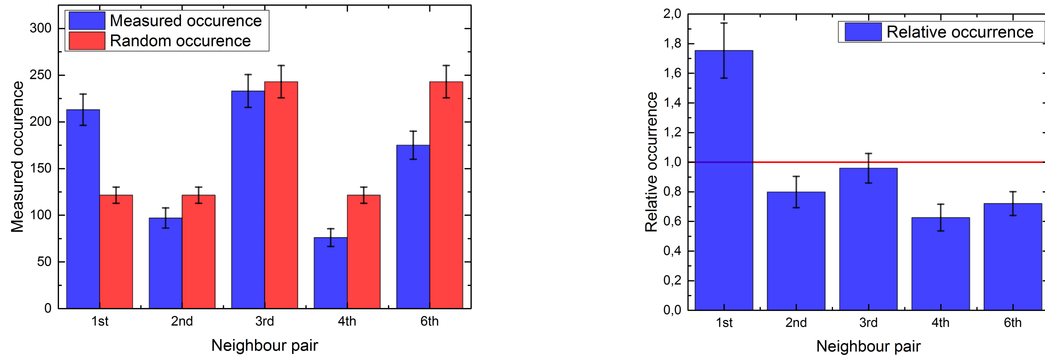


Figure 36: 80x80 nm filled state X-STM image of sample region Bi + N composition 2 of the co-doped sample. Bi dopants in different atomic layers are highlighted with a red circle and a N dopant is highlighted in green. (-3V, 50pA, 77K)

As can be seen in figure 36, the tip used to generate this X-STM image was very sensitive to the surface layer Bi atoms. The second and third layer Bi atoms are suppressed, as seen when figures 26 and 36 are compared. This is helpful when an analysis based on surface layer Bi atom counting is performed. Also, the N atoms are clearly visible in figure 36, as a depression in the contrast as given at the green circle. This feature corresponding to a N atom at the surface strongly resembles its calculated [10] and previously observed forms [29]. Now a similar pair counting of Bi pairs in the nearest neighbour configurations shown in figures 28 and 27 can be performed. In total 0.34 ± 0.01 μm^2 of surface area has been analysed in the search for Bi pairs and a total of 794 Bi pairs were considered in the following results. The surface layer Bi pair distribution obtained from considering the entirety of the co-doped sample, including Bi doped and co-doped areas, is given in figure 37.



(a) Absolute value of the measured Bi pairs. The absolute amount of expected Bi pairs if the distribution was random is also given.

(b) Relative occurrence of the Bi pairs for each pair configuration. The expected occurrence for a random distribution is set at one for each pair configuration.

Figure 37: Results of the surface layer Bi-Bi pair counting in the Bi and N co-doped sample. The first through sixth nearest neighbour pair occurrences are presented in absolute and relative form.

The random occurrences and uncertainties in figure 37 are calculated in a similar fashion to the ones obtained for figure 30. The only difference being that the expected number of Bi pairs for a random distribution of Bi dopants is calculated for each sample region separately and then added up to obtain the total number of expected pairs. Therefore there was no average concentration used here, rather the individual concentrations of each sample region given in table 4. A detailed description of the calculation of the uncertainties is given alongside the presentation of figure 30. When comparing figures 30 and 37, mostly similar behavior is observed. The first nearest neighbour configuration occurs close to twice as often as expected whereas the other configurations except for the third nearest neighbour one are observed less than expected within the margin of error. The fourth nearest neighbour one again being the least observed variant.
Each pair configuration has a lower relative occurrence than in the Bi doped sample however. This could indicate that the presence of N dopants lowers the amount of Bi pairs present in the material. On average the relative occurrence of nearest neighbour Bi-Bi pairs decreases by 17.7% from the Bi doped sample to the co-doped sample.

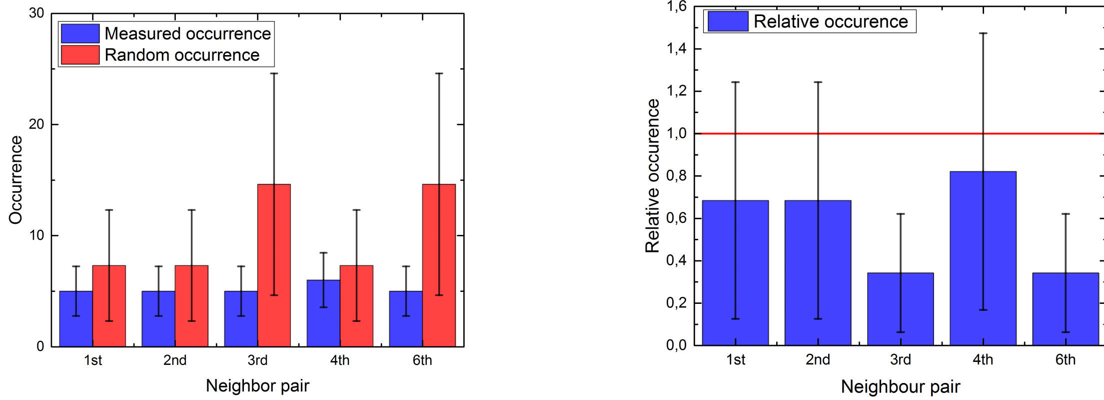
The effects of co-doping GaAs with Bi and N on the Bi concentration are also investigated. As mentioned before, Bi is hard to incorporate into GaAs in high amounts and nitrogen could help the incorporation of bismuth due to opposite lattice strain. This is investigated in the co-doped sample, by comparing the Bi concentration in the Bi doped regions and the co-doped regions. This sample will be used for such a comparison since there are regions with the same (intended) Bi concentration next to each other where one of the two is co-doped with nitrogen. As mentioned before, the growers predicted Bi concentrations of 0.4%, 0.8% and 1-1.4% for compositions 1 through 3 respectively and predicted a N concentration of 0.25% in all N doped regions. From counting the Bi dopants in our X-STM measurements we came to the Bi concentrations shown in table 4.
Along this counting of Bi atoms, the amount of N dopants present is also studied. This yielded a N concentration of 0.06% in all sample regions with N present.
The statistics presented in table 4 are based on the observation of surface Bi atoms in each different sample region. The amount of Bi dopants used to calculate the concentrations presented in table 4 ranges from a minimum of 756 dopants for Bi + N composition 1 to a maximum of 1918 dopants for Bi + N composition 2. In total 9437 Bi dopants were counted in order to ensure the results presented here are statistically relevant. The effects of light assisted MBE growth of these structures can also be studied for this sample. As we saw

	Bi concentration	
Sample region	**STM**	**XRD**
Bi composition 1	0.8 ± 0.1%	0.4%
Bi + N composition 1	0.47 ± 0.05%	0.4%
Bi composition 2	0.64 ± 0.06%	0.8%
Bi + N composition 2	0.67 ± 0.04%	0.8%
Bi composition 3	0.61 ± 0.06%	1-1.4%
Bi + N composition 3	0.67 ± 0.06%	1-1.4%
Light Bi composition 2	0.66 ± 0.06%	0.8%
Light Bi + N composition 2	0.60 ± 0.05%	0.8%

Table 4: Bi concentrations in the different sample regions of the co-doped sample.

in section 4.1.5, the layers grown with light assisted MBE showed a decrease in Bi concentration compared to the regularly grown layers. In the case of the co-doped sample, this effect is less pronounced. There is even a slight increase in Bi concentration when the sample regions with Bi composition 2 and "Light" Bi composition 2 are compared. Since the concentrations of both illuminated layers lie within the margin of error of their dark counterparts, no significant effect of the light assisted MBE is observed in this sample.
There is a large discrepancy in concentration between the sample regions with Bi composition 1 and Bi + N composition 1 as seen in table 4. The Bi concentration in the co-doped sample region Bi + N composition 1 is 0.32 percentage points lower than in Bi composition 1. This is probably caused by the fact that there were some complications during the growth of this layer. The N plasma used to supply the N dopants went out during the start of the growth of this layer. This resulted in a N deficiency at the start of this layer and a small excess of N at the end of the layer.
The Bi concentrations obtained from dopant counting in X-STM measurements again varies strongly from the concentration predicted from XRD measurements, like in table 2. This recurring discrepancy between STM results and XRD estimations of the Bi concentrations is most likely caused by the fact that it is challenging to properly estimate Bi concentrations with XRD in materials with a large number of layers with varying concentrations.
Apart from the sample regions with composition 1 described before, the whole sample shows a similar Bi concentration. The effects of the N doping on the Bi incorporation in GaAs are negligible in this sample. Since the N concentration in this sample is only 0.06% rather than the expected 0.25% predicted by the XRD measurement, the effects of the N doping are expected to be very small. In order to further investigate the effects of N co-doping on the Bi incorporation in GaAs a different structure with a higher N concentration should be studied.

Lastly, the Bi-N pairing in the co-doped sample is investigated. From DFT results in section 4.1.4 we expect that Bi and N dopants appear often in the first nearest neighbour configuration. Therefore we expect the Bi and N dopants to cluster together in the co-doped sample. A nearest neighbour pair counting of the Bi-N pairs is performed and the results are presented in figure 38.
The random occurrences and errors presented in the graphs shown in figure 38 are again calculated in a similar fashion as described at figure 30. The average N concentration used to calculate the random occurrence is 0.06% while for the Bi concentration the concentrations in each different co-doped layer shown in table 4 is taken. The results shown in graph 38 consider all co-doped layers of the co-doped sample.
A few things stand out when we look at the results of the Bi-N pair counting presented in figure 38.
First, the margin of error of both these graphs is large. This is caused by the fact that the N concentration is very low. This is also reflected in the measured occurrence shown in figure 38a. The amount of nearest neighbour Bi-N pairs found throughout all co-doped layers is only 26. A lower amount of Bi-N pairs found leads to a relatively high error margin of this counting.
Second, all nearest neighbour pairs show up less than expected, including the first nearest neighbour pair. This means that despite an energy difference found in DFT of 110.0 meV between the first and second



(a) Absolute value of the measured Bi-N pairs. The absolute amount of expected Bi-N pairs if the distribution was random is also given.

(b) Relative occurrence of the Bi-N pairs for each pair configuration. The expected occurrence for a random distribution is set at one for each pair configuration.

Figure 38: Results of the surface layer Bi-N pair counting in the Bi and N co-doped sample. The first through sixth nearest neighbour pair occurrences are presented in absolute and relative form.

nearest neighbour configuration, no increase in first nearest neighbour occurrence is found.

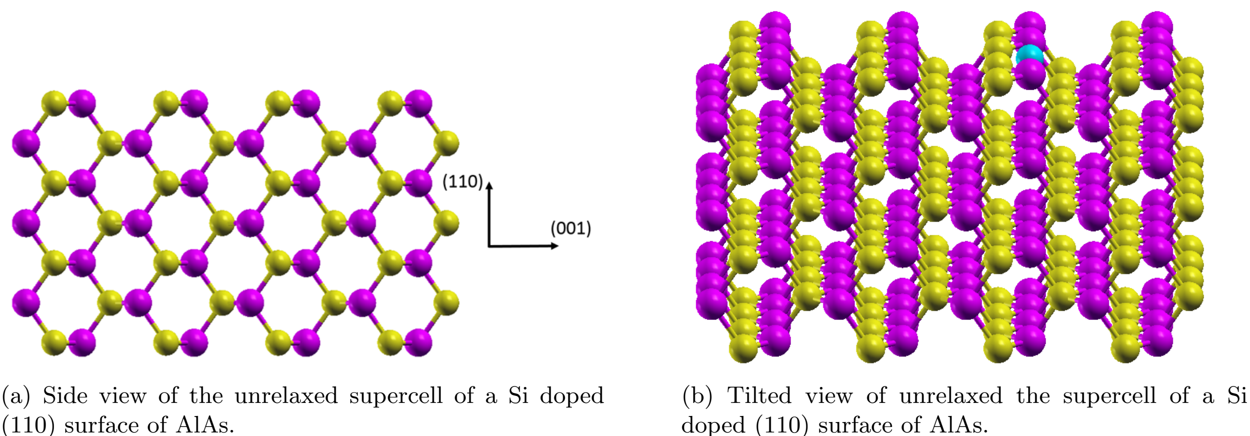
Given the fact that the margin of error of the Bi-N pair investigation is large and that its results do not agree with the DFT calculation performed, it is suggested that a new measurement is to be conducted. In this measurement a sample with a higher N concentration, preferably a concentration similar to the Bi concentration in the co-doped sample, is measured using X-STM. A co-doped sample with a higher N concentration will yield statistically valid results that can either prove or disprove the hypothesised clustering of Bi and N dopants in GaAs. With a more highly doped N sample, the effect of the co-doping with N on the Bi concentration in GaAs will also be more pronounced.

4.2 Si donors in AlAs

The second problem discussed in this report concerns Si donors in AlAs. As mentioned in section 1.3, interesting behaviour is expected around these donors from previous results of Si donors in GaAs. This problem will also be studied theoretically using DFT and experimentally using X-STM. We first discuss the DFT calculations here. The influence of Si donors on the AlAs band structure and the simulation of X-STM images will be discussed. These results will then be compared to the experimental results of X-STM measurements on Si donors in AlAs.

4.2.1 DFT simulation of Si donors in AlAs

In order to investigate the behavior of a Si donor in a (110) surface of AlAs theoretically, DFT calculations are performed on AlAs surface supercells. These supercells are 7 atomic layers thick and consist of 224 atoms total. In order to compare a clean (110) AlAs surface to a Si doped (110) AlAs surface, two different cells are compared in the following calculations. One has 112 Al and 112 As atoms and the other has 111 Al atoms, 112 As atoms and 1 Si atom present in the surface layer of the supercell. The Si doped supercell is given in figure 39.



(a) Side view of the unrelaxed supercell of a Si doped (110) surface of AlAs.

(b) Tilted view of unrelaxed the supercell of a Si doped (110) surface of AlAs.

Figure 39: Unrelaxed Si doped AlAs supercell. Al atoms are given in pink, As atoms are yellow and the Si donor is given in light blue.

The supercell shown in figure 39 has 7 Å of vacuum above and below the atoms so a total of 14 Å of vacuum is present between the slabs of material. The convergence tests discussed in section 3.1.2 are performed and the E_{wfc} is found to be 90 Ry, an E_ρ of 750 Ry is used and lattice constant of 5.605 Å is found for the AlAs lattice. The k-point sampling of the Brillouin Zone is a 3x2x1 Monkhorst-Pack grid. An important difference with the Bi doped GaAs systems discussed before is that in this supercell, both sides of the supercell are left free to relax. The only fixed atomic layer is the middle one, so buckling is expected to show up on both the top and bottom of the supercell. This is done since the band structure of the system is of interest here and will be investigated. The band structure will show strong dispersionless bands corresponding to the unrelaxed vacuum-AlAs interface at the bottom of the supercell if this interface is not left free to relax. The fully relaxed Si doped supercell is given in figure 40.

As can be seen in figure 40, there is indeed buckling on both sides of the supercell. The Si atom relaxes into the material slightly more than the surface Al atoms Now that the relaxed structures are found, band structure calculations can be performed for both the pristine AlAs supercell and the Si doped supercell. The pristine AlAs supercell band structure will be discussed first and is given in figure 41.

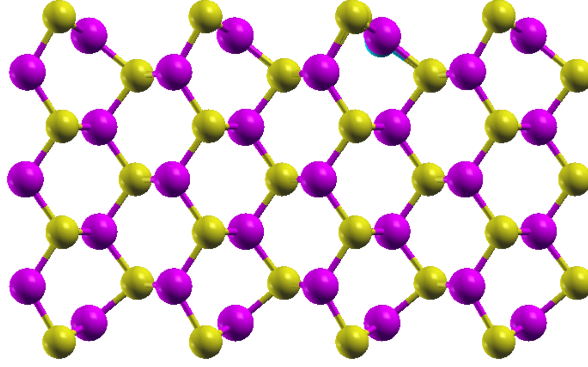


Figure 40: Fully relaxed Si doped (110) AlAs surface supercell with relaxation on both AlAs-vacuum interfaces. Characteristic buckling is visible on both sides of the supercell.

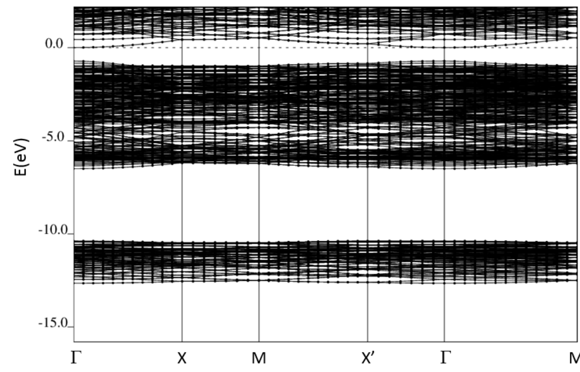
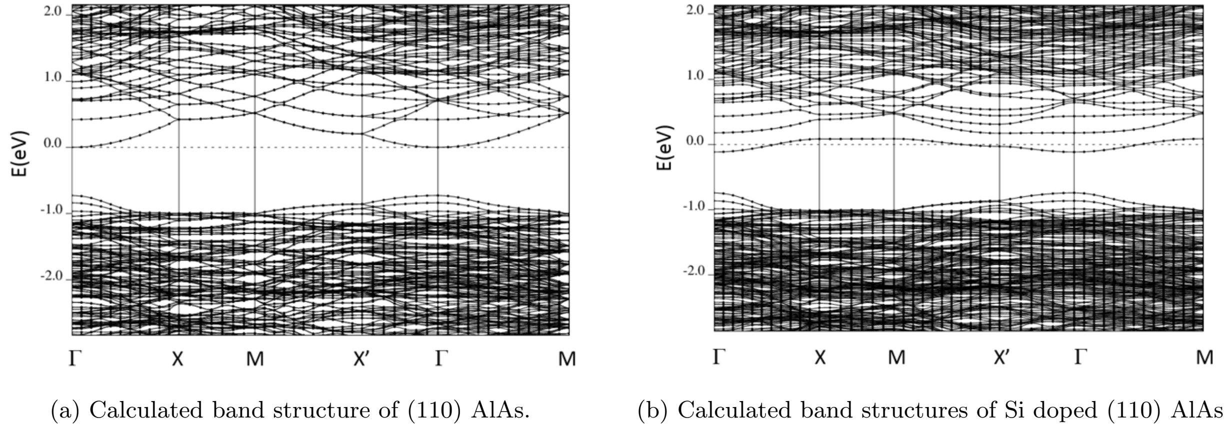


Figure 41: Calculated band structure of (110) AlAs. The Fermi energy is 0 eV in this band structure.

The path through the Brillouin Zone (BZ) along which the band structure is calculated in figure 41 is discussed in detail in section 2.2. The band structure in figure 41 is similar to (110) surface AlAs band structures found in literature [49].
The band structure in figure 41 shows a direct band gap at the Γ point. Bulk AlAs has an indirect band gap with the valence band maximum at the Γ point and the conduction band minimum at the X point. This difference is caused by the fact that, as discussed in section 2.2, in a (110) surface the X point of the bulk BZ is folded onto the Γ point of the surface BZ.
Now we focus on the most interesting part of the band structure of both the pristine AlAs and the Si doped AlAs, the band gap region. We focus on this part of the band structure since the region around the Fermi energy, which lies at 0 eV for the structure shown in figure 41, is the region that STM measurements are sensitive to. The band structures of both the undoped (110) AlAs and Si doped (110) AlAs around the Fermi energy are given in figure 42.
Important to note is that, alongside the energy range being smaller, the band structures are shifted in figure 42 so that the Fermi energy for both structures is at zero energy. This is done to more easily compare these band structures. The Fermi energy of the undoped structure is 0.8249 eV while the Fermi energy of the Si doped structure is 0.8652 eV. The similarities and differences of the structures shown in figure 42 will now be discussed.
First, the shape of the top of the valence band is very similar for both structures, and a similar behavior of the conduction band is visible. The bottom of the conduction band remains at the Γ point for instance. On the other hand, the impurity state caused by the Si donor is clearly visible near the bottom of the conduction band. From the comparison of the band structures in figure 42 it can be concluded that the state around 0 eV



(a) Calculated band structure of (110) AlAs.

(b) Calculated band structures of Si doped (110) AlAs.

Figure 42: AlAs band structures around the Fermi energy, which is at 0 eV.

in figure 42b is caused by the Si donor. Since the Si impurity is a donor, we indeed expect this state to show up near the bottom of the conduction band. The conduction band itself is also affected by the presence of the Si donor. Many states near the bottom of the conduction band have their degeneracy lifted for instance. This is a phenomenon that occurs more often in (calculated) band structures of doped semiconductors [50]. This behavior of the states splitting makes sense when the following is considered. The electron states in the conduction band are states that allow the electrons to move through the material. Now we provide the material with an additional periodic potential by implementing a Si impurity in every supercell. This situation is similar to that of the nearly free electron model that describes the formation of gaps in the electron dispersion. In this nearly free electron model, imposing a periodic potential (the material) on free electrons causes gaps in the electron dispersion. These gaps are caused by the lifting of degeneracies at the Brillouin Zone edges and center. This provides us with a nice analogy for what is happening with the states in our conduction band: the periodic potential from the Si atoms lifts the degeneracy of states near the conduction band minimum the same way a periodic potential in a material lifts the degeneracy of free electrons.

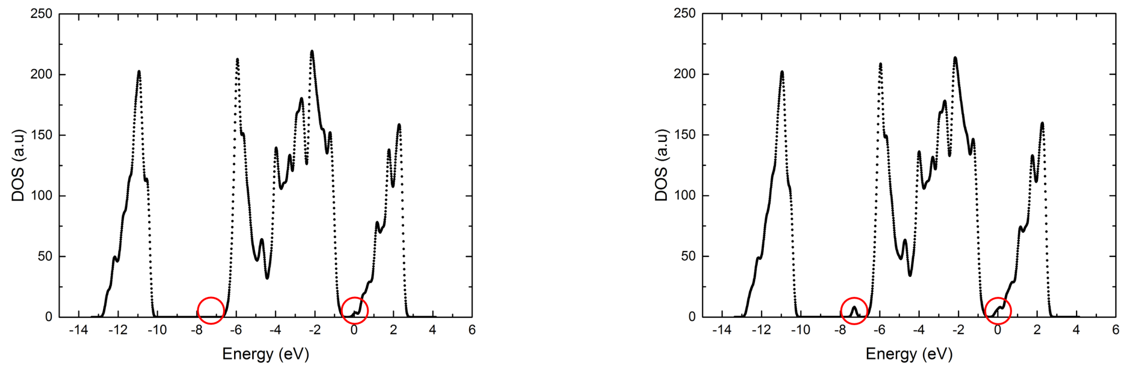
The Si donor also generates another clear impurity band at about 7.5 eV below the top of the valence band. To show this second impurity band, the density of states (DOS) for a large energy range is given for both the undoped and Si doped (110) AlAs supercells in figure 43.

The impurity state caused by the Si donor at the Fermi energy in figure 42b is again visible at 0 eV in figure 43b . While the increase in DOS due to the presence of the second impurity band is visible at approximately -7.5 eV.

Now that we have an understanding of what the DOS and band structure of our supercells look like, we can simulate STM images of the surfaces that they describe using Quantum-espresso. This simulation of the STM images is done by integrating the DOS from the Fermi energy to the applied bias voltage (Tersoff and Hamann [51]). The simulated STM images shown are the integrated DOS given for a certain horizontal plane in our supercell. This means that the STM tip is considered an infinitesimal small dot that scans a perfectly flat plane in constant height mode, that is only sensitive to the DOS in that one perfectly flat plane. In order to choose a proper bias voltage it is therefore necessary to have a good understanding of the band structure and DOS, since the energy range of the states of interest needs to be known.

Undoped (110) AlAs

First, we consider the simulated STM images for the undoped (110) AlAs supercell. This is done as a first check of the procedure since the behavior and STM imaging of clean (110) AlAs or III/V semiconductors in general is well-studied. To investigate the shape of the surface wave functions in this undoped (110) AlAs, we look at the DOS in the plane of the topmost atomic layer. These would correspond to constant height

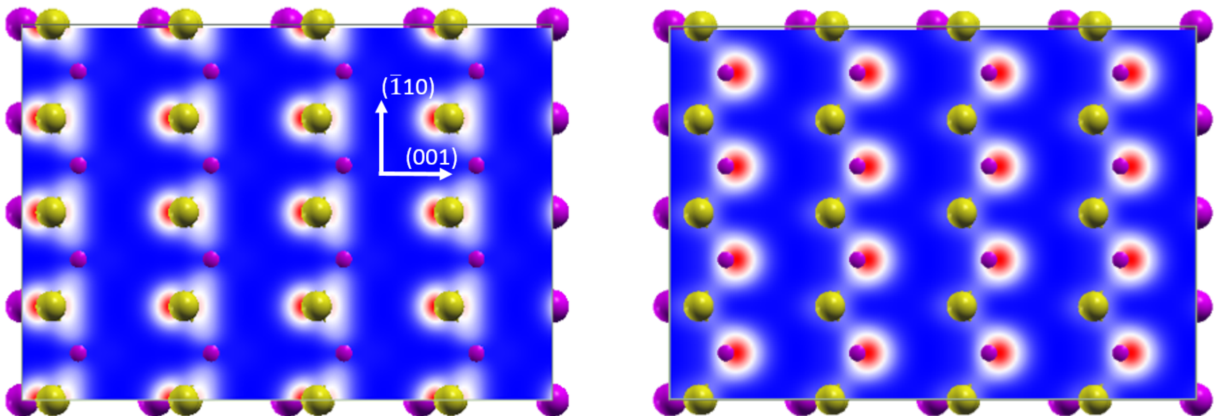


(a) Calculated density of states of the (110) AlAs supercell.

(b) Calculated density of states of the Si doped (110) AlAs supercell.

Figure 43: DOS of the AlAs and Si doped AlAs supercells with the areas influenced by the Si dopant highlighted in red. An energy of 0 eV corresponds to the Fermi energy.

STM images made at a height within the surface layer. This is experimentally impossible but these images can give us some nice insight in the shape of the surface states and the orbitals that make up these states. The simulated filled state, with a bias of -1V, and simulated empty state image, with a bias of +1V, are given in figure 44.



(a) STM image simulated at a tip bias of -1V, or filled state image.

(b) STM image simulated at a tip bias of +1V, or empty state image.

Figure 44: Simulated STM images through the surface layer of undoped (110) AlAs.

The isolines around either the As atoms and the Al atoms for filled and empty state images respectively are given in figure 45. These images with isolines are zoomed in variants of the ones presented in figure 44.
For GaAs, it is well-known that the states at the valence band maximum are made up of p-like wave functions centered on the As atoms while the conduction band minimum is made up of s-like wave functions centered on the Ga atoms [52]. Now for our undoped AlAs (110) surface we see similar behavior.
First we consider the filled state case, which deals with the top of the valence band. These images are given in figures 44a and 45a. With a tip bias of -1V with respect to the Fermi energy we deal only with the top of the valence band as can be seen in figure 42a. Now we can see clearly from figure 44a that the surface state

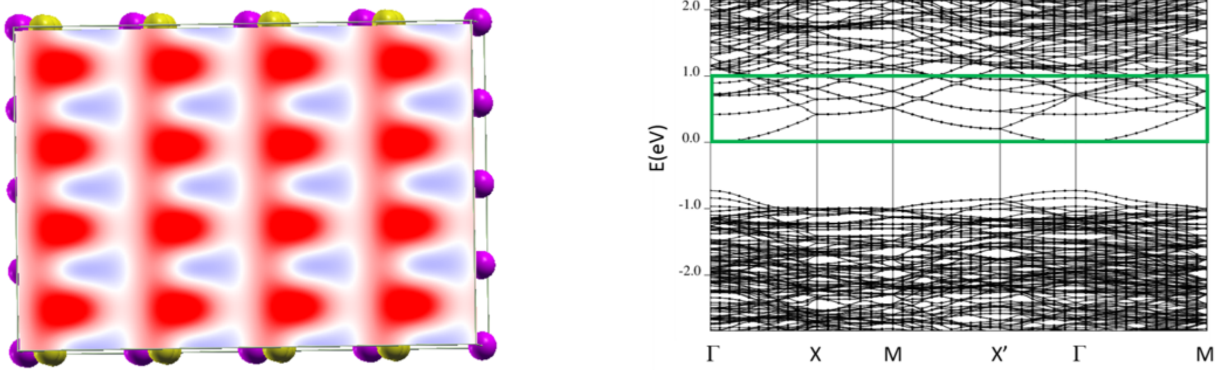
(a) Zoomed in image of 44a on an As atomo with isolines.

(b) Zoomed in image of 44b on an Al atom with isolines.

Figure 45: Simulated STM images through the surface layer of undoped (110) AlAs zoomed in on the As atoms in a) and the Al atoms in b).

is indeed centered on the As atoms and in figure 45a we can see that this state does indeed show p-type behavior.
A similar argument can be made for the empty state case. Images 44b and 45b deal with the empty state case. A tip bias of +1V deals with the bottom of the conduction band as can bee seen in figure 42a. Now the surface state is instead centered on the Al atoms, as visible in 44b, and shows s-like behavior, as visible in 45b.
This means that, despite GaAs and AlAs being very different material in terms of for example band gap size and type, their surface state behavior around the band gap is very similar.

The images shown in figure 44 are not very representative of an actual STM image however, since this is a LDOS image through the surface layer, where the tip could never be without being destroyed. To get a more realistic image, the STM image of a tip 4 Å above the surface layer was also simulated for the undoped (110) AlAs surface. The result of this simulation is given in 46a.



(a) Simulated STM image of clean (110) AlAs 4 Å above the surface.

(b) Band structure of clean (110) AlAs with the bands taken into account in 46a highlighted.

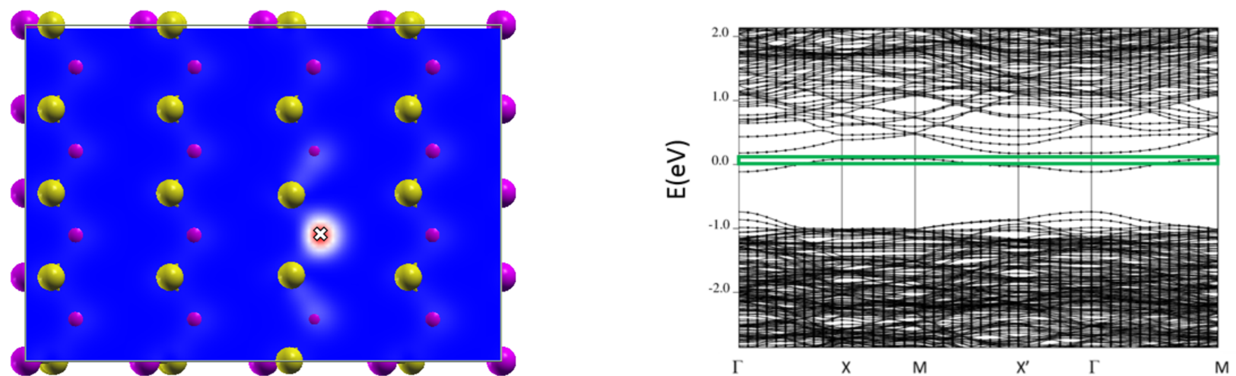
Figure 46: Simulated STM image and energy interval of interest in the band gap of (110) AlAs.

The image shown in figure 46a looks more like a typical STM image we can obtain. The highest DOS is still centered on the Al atoms, as expected for empty state imaging, but it is more spread out then what we saw in figure 44b. As an illustration the band structure with the bands of interest highlighted for an empty state image with a bias voltage of +1V is given in figure 46.
A study of which surface states are sampled at a specific bias voltages using X-STM has been performed

within the PSN group before [53]. Here it was shown that by changing the bias voltage slightly, we can sample different surface states and because of this the corrugation can switch directions.
This switching of corrugation we do not observe in our simulated STM images and we do not expect to observe this in the first place. The reason for this is the way the DFT calculations generate an STM image. The DFT calculations integrate the DOS from the Fermi level to the applied voltage and all the states present in this interval contribute evenly to generate the eventual DOS contour.
When an actual STM measurement is performed and is translated into an STM image, not all states contribute evenly however. We will discuss this in detail for empty state imaging.
In the case of empty state imaging, a positive bias voltage is applied to the tip. The electrons can now tunnel from the tip into the empty states present in the sample. The energy range with states available for tunneling in the sample lies between the Fermi level of the sample and the applied bias voltage of the tip.
The electrons tunneling into these states from the tip do not all see an equal tunneling barrier however. This barrier is much higher for electrons tunneling into states near the Fermi level than for electrons tunneling into states near the applied bias. This is due to the fact that electrons near the applied bias have a higher energy than the ones near the Fermi level for empty state imaging.
The fact that electrons near the Fermi energy see a much higher barrier means that it is much less likely for them to tunnel through the barrier than electrons with an energy near the applied bias. This leads to the tunneling current in empty state imaging mostly being generated by electrons with an energy near the applied bias. Therefore, the states near the applied bias determine the image generated by the STM more than the states near the Fermi level for empty state imaging.
Certain states can now be selected to image with STM by tuning the applied bias in such a way that those states will be most likely to tunnel through the barrier. This is why the corrugation can change by slightly tuning the tip bias for example.
Since all states between the Fermi level and the applied bias have the same weight in determining images like 46a, they will inherently differ from actual STM images.

Si doped (110) AlAs

Now we consider the Si doped (110) AlAs system. Looking at figures 42a and 42b, it is expected that the band visible around the Fermi energy, 0 eV, in figure 42b is the impurity state caused by the Si atom. This can be tested by simulating an STM image through the surface layer, where the Si atom is present, and check whether the DOS does indeed show that the Si causes this state. This simulated STM image through the surface layer with a tip bias voltage of +0.1V and the corresponding band structure and area of interest are shown in figure 47.

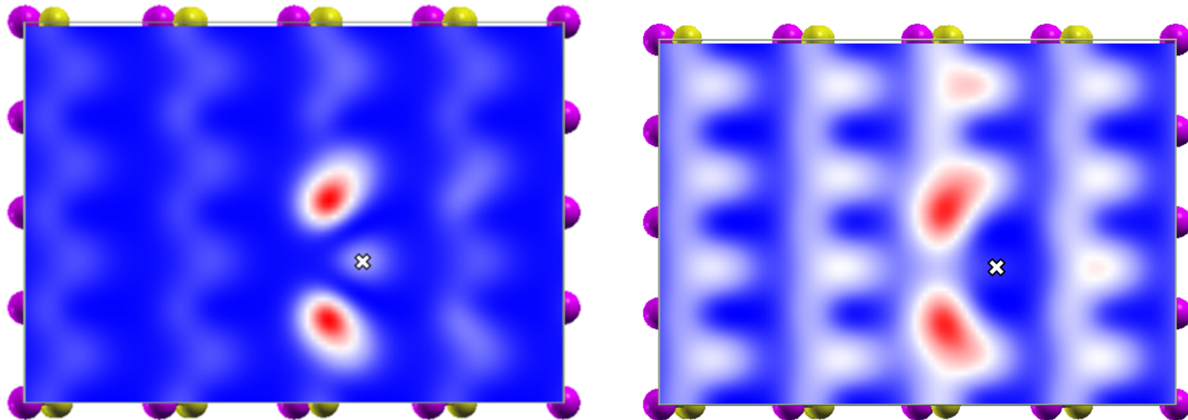


(a) Simulated STM image of Si doped (110) AlAs in the surface layer. The Si donor is situated at the cross.

(b) Band structure of Si doped (110) AlAs with the bands taken into account in 47a highlighted.

Figure 47: Simulated STM image and energy interval of interest in the band gap of (110) Si doped AlAs.

As can be seen in figure 47, this band is indeed made up of states corresponding to the Si impurity. Since only a very small energy interval is considered here, as can be seen in figure 47b, the states corresponding to the AlAs lattice are barely visible. The conduction band is pushed up by the presence of the impurity state as can be seen when comparing figures 42a and 42b, so with a tip bias of +0.1V we do not yet strongly sample the conduction band. From this we can conclude that the state visible around the Fermi energy is indeed an impurity state caused by the Si donor. To get a better understanding of the shape of the wave function corresponding to this impurity state, we show two STM images calculated at the same tip bias but at different heights above the surface in figure 48.



(a) Simulated STM image of Si doped (110) AlAs 2 Å above the surface. The Si donor is situated at the cross.

(b) Simulated STM image of Si doped (110) AlAs 4 Å above the surface. The Si donor is situated at the cross.

Figure 48: Comparison of simulated STM images of Si doped (110) AlAs at different heights.

As can be seen in figure 48, the surface state corresponding to the Si donor extends into the vacuum above the surface as a two lobed structure centered on top of the neighbouring As atoms. When the distance above the surface is 4 Å, a typical scanning height for STM, the background of the surface state regarding the AlAs surface increases in relative intensity, as seen when comparing 48a and 48b. This is caused by the fact that this color mapping is relative, so when the value of the DOS of the Si impurity state decreases further away from the surface, the visibility of the background DOS increases.

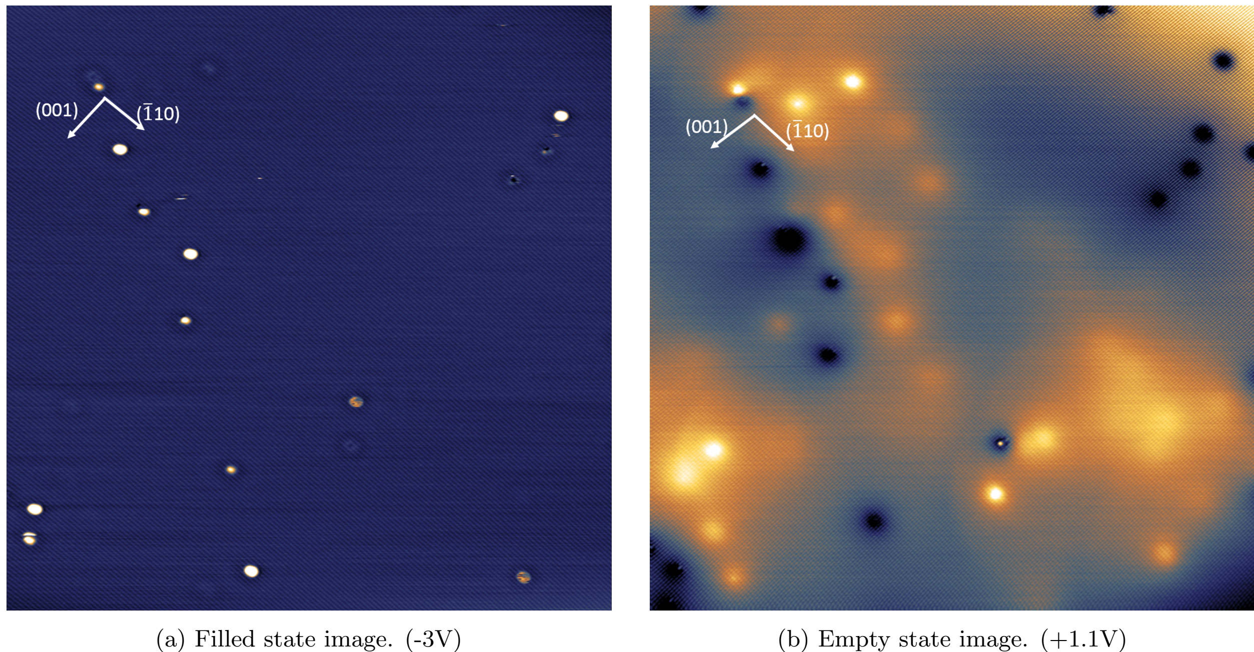
Since we are only looking at a small energy interval here in order to consider the shape of the Si state in particular, we are not taking into account that in realistic STM images a much larger bias is applied. This means that the states present in the conduction band would also contribute to the generated STM image. Because of this, the realistic empty state image of a first layer Si donor in a (110) AlAs doped would look most like the superposition of figure 46a and figure 48b.

The images shown in figure 46a, 47a, 48a and 48b are simulated DOS planes at a certain height above the surface. Therefore they correspond to constant height STM images. Constant height STM imaging is difficult to perform due to the risk of tip crashing however. STM imaging in this study is therefore performed in constant current mode. This is another reason the actual STM images might not fully resemble the simulated images shown in this section.

Now we will perform STM measurements on Si donors in AlAs. These experimental results will then be compared to the theoretical results obtained here.

4.2.2 X-STM measurements on Si donors in AlAs

The results of X-STM measurements on Si donors in AlAs will now be discussed. These measurements are performed to get a better insight in the behaviour of these donors. We presented the theoretically expected band structure and local density of states around these donors in section 4.2.1. These theoretically expected results will now be compared to X-STM measurements. Another reason for performing X-STM measurements on Si donors in AlAs is to see how and if this compares to Si donors in GaAs, which our group has extensively researched [54] [55] [56] as discussed in section 1.3.
A typical filled and empty state image of the same area of (110) Si doped AlAs is given in figure 49.



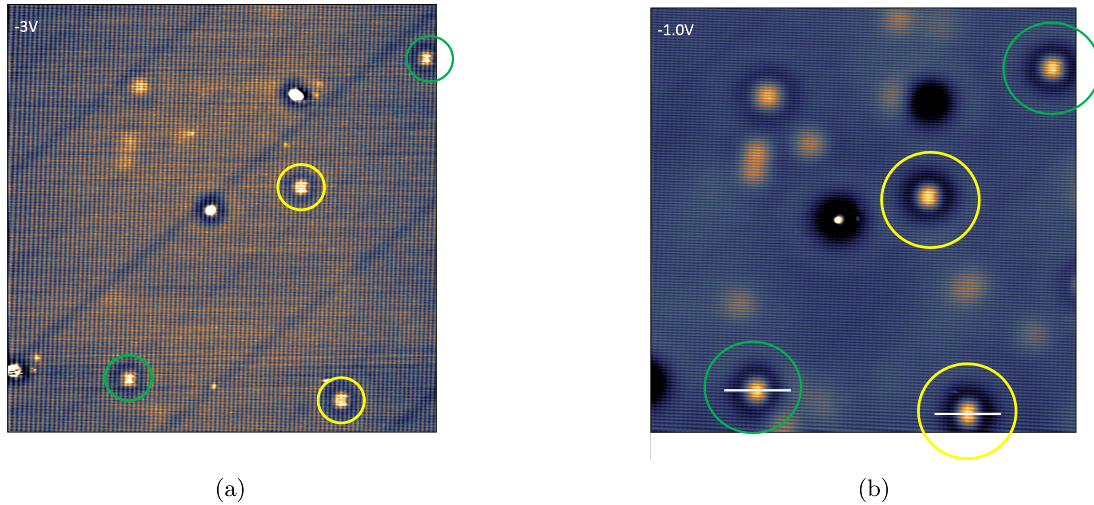
(a) Filled state image. (-3V)

(b) Empty state image. (+1.1V)

Figure 49: 80x80 nm X-STM images of Si doped AlAs. The filled and empty state image of the same area is given. (30pA, 77K)

In figure 49, a filled and empty state images of the same area in Si doped AlAs are presented. As can be seen, different features show up more prominently when switching bias, but some features show up in both 49a and 49b. This is to be expected since switching tip bias samples states from either the conduction or valance band. These states can be made up of either the Al or As sub-lattice for instance and features can show up in different forms as well. Images like the ones presented here were used to determine which features correspond to the Si donors in different layers below the surface.
First we identified which features belong to the Si donors. This was done by looking at the behaviour we expect around Si donors, like charge rings [56] and contrast switching [55] for instance. The depth of these features below the surface was then investigated. This was done using the magnitude of the Friedel oscillations around the different features identified as a Si donor.
Friedel oscillations in materials are fluctuations in electron density arising from a defect or impurity in the material. When filled state images of surfaces are made using X-STM, these oscillations are clearly visible since there can be charge accumulation at the surface. This charge accumulation in filled state imaging of the surface is caused by downward tip induced bend bending of the band structure of the sample. When the tip induced bend bending is strong enough, the conduction band states of the sample are pulled below the Fermi level. This causes an accumulated electron gas in the surface of the material.
The Si donors in the material are now treated as local impurities by the electron gas at the surface and are screened. This screening causes oscillation of the electron density around the Si donor, the Friedel

oscillations.
These Friedel oscillations can be used as a nice tool here since a donor closer to the surface is expected to have their charge oscillations show up more prominently than deeper lying Si donors. We expect stronger charge oscillations around Si donors closer to the surface since the oscillations decay in amplitude when moving away from the donor atom. As an illustration, the suggested second and third layer Si features are presented in figure 50.



(a) (b)

Figure 50: 50x50 nm filled state images of the Si doped AlAs with the second, yellow, and third, green, layer Si donor features highlighted. The height profiles of the features corresponding to a second and third layer Si atom given in figure 51 are taken along the paths indicated by the white lines in b). (30pA, 77K)

As can be seen when comparing the green and yellow highlighted features in figure 50, the yellow features show a stronger charge oscillation around them. A detailed height profile of the features corresponding to a second and third layer Si donor are given in figure 51.

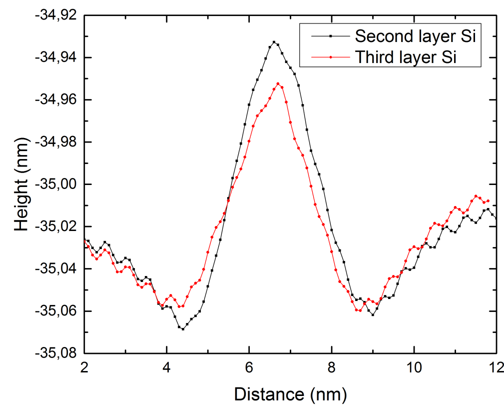


Figure 51: Height profiles of the features corresponding to a second and third layer Si donor. The paths along which the profiles are taken is indicated in figure 50b.

As can be seen in figure 51, the charge oscilation is indeed stronger for the features circled in yellow in figure 50b. The difference in magnitude of the charge oscillation is clearly visible at 4.5 nm in figure 51, where the second layer Si donor profile is significantly lower than that of the third layer Si donor.
From this we conclude that the features highlighted in yellow correspond to second layer Si donors. These second layer Si donor features are very similar to features observed in Si doped GaAs that were also attributed to the Si donors by Feenstra et al. [57]. These features found by Feenstra en al. and the features discussed here as second layer Si donors are given side-by-side in figure 52.

(a) Second layer Si donor in AlAs as observed in our measurements. (-2.7V, 30pA)

(b) Si donor in GaAs as observed by Feenstra et al. [57]. (-2.2V)

Figure 52: Comparison of the second layer Si donor observed in AlAs and the Si donor feature observed by Feenstra et al.

When comparing figures 52a and 52b, these are clearly very similar. Both the internal shape of the feature and the behavior around the feature are almost identical and this therefore strengthens our statement that these are indeed Si donors. However, Feenstra et al. state that the decrease in state density around the dopant, which we also observe, is caused by the repulsive interaction between the donor core and the valence band electrons. As mentioned before, we believe these charge oscillations around the Si donors to be Friedel oscillations in the electron accumulation layer at the sample surface. This is also corroborated by van der Wielen et al. [58]. They find similar oscillations around Si donors in GaAs and attribute these oscillations to Friedel oscillations. The fact that these charge oscillations are not found in empty state imaging, where there is no charge accumulation layer, is additional evidence that these oscillations are indeed Friedel oscillations around the Si donors.
We now use the magnitude of the Friedel oscillations to find out which features correspond to Si donors at different depths below the surface. We found the features that correspond to Si donors located in the second through fifth layer below the (110) AlAs surface. These are given in figure 53.

After the analysis of the second through fifth layer Si donors in the material, multiple features that could correspond to Si donors located in the surface layer remained. Since the surface can also contain adsorbants or vacancies, multiple features were found to correspond to surface layer effects. The feature corresponding to the surface layer Si donor was eventually found by considering the expected amount of surface layer Si atoms from the donor concentration. This provided us with an expected amount of surface layer Si donor features found in our measurement and allowed us to eliminate all but one surface related feature as possible candidates. The resulting surface layer Si donor feature is given in figure 54.
As can be seen in figure 54a, there seems to be an internal structure within the feature in filled state imaging. This structure is attributed to contrast switching when measuring surface layer Si donors, similar to what is observed for surface Si donors in GaAs [16]. This switching is attributed to the bistability of surface layer Si donors in GaAs. When the X-STM tip is close to the surface Si donor in filled state imaging, the tip could remove an electron from the Si. The Si donor then has two possible states. Either one of the electrons is taken away by the tip or this electron remains on the Si donor. The switching between those two states shows up as the contrast switching visible inside the feature shown in figure 54a.

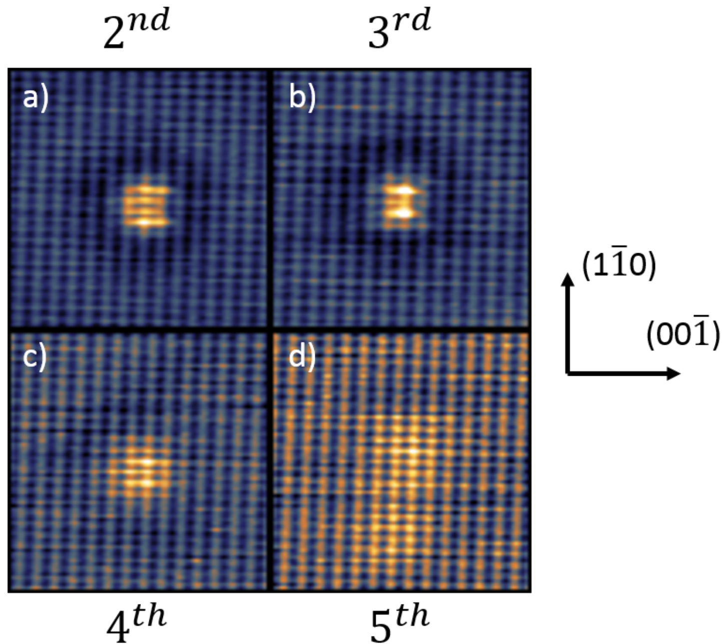
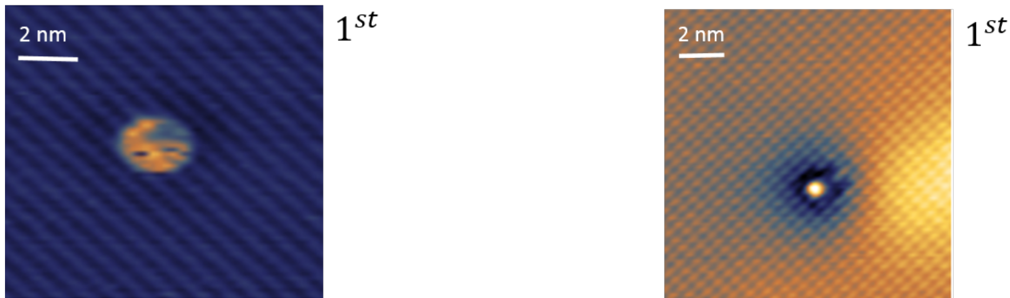


Figure 53: 7.5x7.5 nm images of a second, third, fourth and fifth layer Si donor in figures a)-d) respectively. (-3V, 30pA, 77K)



(a) Filled state image of the surface layer Si atom. (-3V, 30pA, 77K)

(b) Empty state image of the surface layer Si atom. (+1.1V, 30pA, 77K)

Figure 54: Surface layer Si in AlAs feature in both filled and empty state imaging.

Now that the different layer Si atom features are identified in our X-STM imaging. We can compare these to the simulated STM images shown in section 4.2.2. The problem studied with DFT is the surface layer Si donor in (110) AlAs and the resulting simulated STM image is given in figure 48b. Both the STM measured surface layer Si donor and the DFT calculated images are given side by side in figure 55.
The STM images simulated using DFT are compared to X-STM images for positive tip biases. This is done since the X-STM measurement will resemble the calculations better in empty state imaging. In filled state imaging, so a negative tip bias, there will be an accumulation of charge at the surface due to tip induced band bending as mentioned before. The measured wave function of the Si donor at the surface will then be influenced by this large amount of charge present in the surface layer. For empty state imaging, no such accumulation of charge is present in the surface and therefore the shape of the wave function of the Si donor measured with X-STM will resemble the simulated wave function more closely.

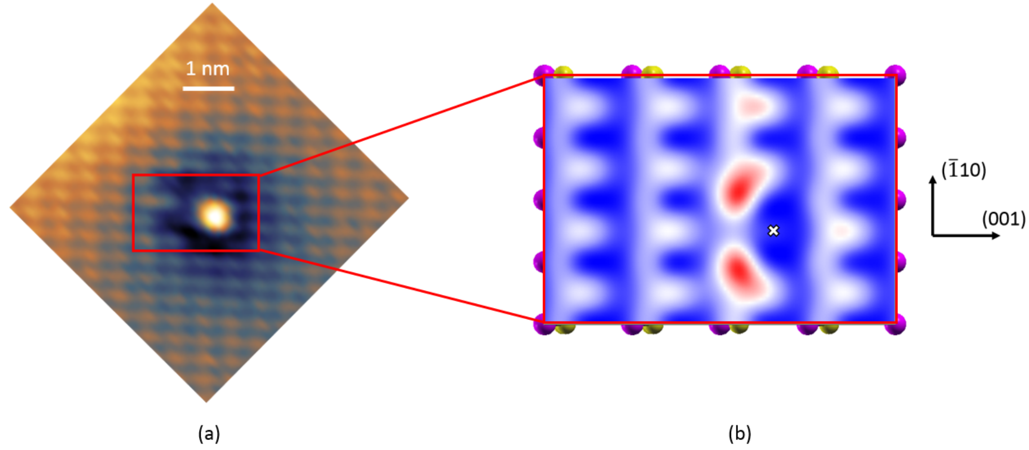


Figure 55: Feature corresponding to a surface layer Si donor in (110) AlAs as found experimentally using X-STM, (a) (+1.1V, 30pA, 77K), and theoretically using DFT, (b). The simulated image in (b) corresponds to a tip 4 Å above the surface with the Si donor situated at the white cross.

By comparing figures 55a and figure 55b we can see that a feature corresponding to a first layer Si atom observed in X-STM is much more extended than the feature observed in the simulations. Moreover, most of this feature is characterised by a lowering of the DOS around the feature. This could be caused by either a physical lowering of the structure or electronic effects. The bright feature in the middle of the surface layer Si donor in X-STM is centered on a group III lattice site and extends over both neighbouring group V sites , which matches with our simulated results. The simulated feature also shows two clear lobes while these lobes seem to be merged together in the feature observed with X-STM.
The difference between experiment and simulation should be taken into account here. First, the X-STM image shown in figure 55a is made at a bias voltage of +1.1V while the simulated image is generated at a bias voltage of +0.1V to isolate the impurity state. Also, the size of the supercell used is limiting here. The dimensions of the simulated image shown in figure 55b are 2.24x1.59 nm. While a rather large supercell containing 224 atoms is used in the simulation, this means that the cell is smaller than the size of the feature visible in figure 55a. Because of this, the large reduction in DOS around the Si donor feature relative to the background DOS of the AlAs surface would be very difficult to observe in the simulation. Lastly, the images generated by DFT are constant-height images while the X-STM images are constant current.

A collaborating group in Brazil has worked in parallel with us on simulating these STM images. Their method is similar to the one presented in this work. They used a supercell with larger dimensions and a different DFT code, SIESTA [59]. They have worked on Si donors in different layers and their results for the second layer Si donor and the empty state X-STM image measured in Eindhoven are presented in figure 56.

As can be seen in figure 56, in this case the simulated image matches really well with the X-STM image. The main difference being that two of the central lobes making up the simulated image in figure 56b are merged into one bigger bright contrast in figure 56a.
The feature is in both cases extended over the central 2 group III atomic sites in the $(00\bar{1})$ direction and 4 atomic sites in the $(1\bar{1}0)$ direction. The lattice shown in figure 56 is the group III sub-lattice since we consider empty state imaging here. The Si atom is not situated directly on one of the group III sites here since it is a second layer Si atom. These Si atoms lie in between the surface group III sites.
The strongest contrast is found at the same position in both figure 56a and 56b. In the experiment and

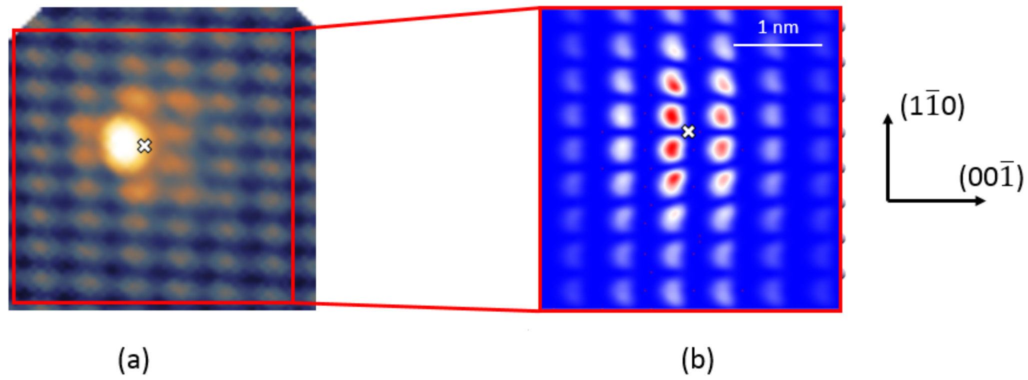


Figure 56: Feature corresponding to a second layer Si donor in (110) AlAs as found experimentally using X-STM, (a) (+1.1V, 30pA, 77K), and theoretically using DFT, (b) . The simulated image in (b) corresponds to a tip 2 Å above the surface with the Si donor situated at the white cross. Figure in (b) adapted from Marcos Gonçalves de Menezes with permission.

simulation the brightest spot of the feature are the two group III lattice sites left of the Si donor. Similarly to what we saw for a first layer Si donor, the simulation shows distinguishable lobes while these lobes seem to be merged in the X-STM measurement.
Why this feature matches better with the simulated image than the surface Si feature is not entirely certain. One reason could be that the surface layer Si atom induces strong charge effects on the surface and can easily interact with the tip. This means the tip in proximity to the surface Si atoms could influence the Si atoms themselves. This is not taken into account in the simulated images.

5 Conclusion

In this study the combination of DFT and X-STM when researching doped III/V semiconductors has been investigated. Both isoelectronic doping of GaAs and Si donors in AlAs were discussed and different aspects of the materials were studied using DFT and X-STM.

5.1 Isoelectronic doping of GaAs with Bi and N

In the case of isoelectronic doping of GaAs, the pair distribution of Bi-Bi, N-N and Bi-N pairs and the effect of N dopants on the incorporation of Bi atoms in GaAs was investigated.

Nearest neighbour pair distribution

In terms of the Bi-Bi pair distribution, the DFT calculations and X-STM measurements show a similar result. They agree that the Bi pairs favour the first nearest neighbour configuration over the second through sixth nearest neighbour configurations. This was found to be the case in both purely Bi doped GaAs and GaAs co-doped with Bi and N. A deficiency of especially fourth nearest neighbour pairs was also found in the X-STM measurements, the supercell used in DFT was too small to simulate the fourth nearest neighbour configuration unfortunately. The deficiency of fourth nearest neighbours could be caused by the Bi dopants more easily moving from the fourth nearest neighbour configuration to the first nearest neighbour one due to both neighbouring atoms being situated along the $(\bar{1}10)$ direction, as shown in figure 27. The first and fourth nearest neighbour pairs are also both exposed on the (001) growth surface of GaAs, making it easier for Bi atoms to move from the fourth nearest neighbour configuration to the first nearest neighbour configuration during growth. Using DFT, is was found that the formation of the first nearest neighbour pairs is likely to occur in the (001) growth surface of GaAs, since this configuration is energetically very favourable over the second nearest neighbour configuration in the growth surface.

The Bi-N pairs in GaAs were also investigated using DFT, the result of which is summarised in table 1. Here it was found that the first nearest neighbour Bi-N pair in GaAs has a much lower energy than the second nearest neighbour pair. This would also suggest that one should observe a larger amount of first nearest neighbour Bi-N pairs. This was not confirmed by X-STM measurements on a co-doped GaAs sample. The lack of these Bi-N pairs could be caused by for instance Bi-N bonding issues or strain effects. Another reason for the low amount of Bi-N pairs found is a very low N concentration, 0.06%, in the sample investigated.

Bi concentration

The Bi concentrations of different sample regions in Bi doped and co-doped GaAs samples were also researched. Here two main topics were of interest. First, the effect of the N dopants on Bi incorporation. Second, the effects of light assisted MBE growth.

In the measured samples there seems to be no significant effect of the N dopants on the Bi incorporation. These results are summarised in table 4. We did find a big difference between the regions with Bi composition 1 and Bi + N composition 1, but these regions were not grown according to the growth plan. For every other composition, no significant difference was found between the Bi concentration in the co-doped and singly doped layers. Considering the amount of first nearest neighbour Bi-N pairs was also lower than expected, the effect of N doping could overall be smaller than expected. The reason for the seemingly low amount of Bi-N pairs could very well be related to the fact that we see little to no effect of the presence of N atoms on the Bi concentration in GaAs.

However, the initial challenge was the incorporation of a large amount of Bi into GaAs and to see whether N could help with this incorporation. Since we are looking at relatively low concentration in this study does not mean that N does not have an effect on Bi incorporation for higher concentrations of Bi.

The effects of light assisted MBE on the Bi incorporation are relatively small in the samples investigated in this study. The Bi doped sample shows a small decrease in Bi concentrations in regions grown with light assisted MBE. The co-doped sample does not show a significant difference in Bi concentration between sample regions grown with light assisted MBE and regularly grown sample regions.

5.2 Si donors in AlAs

When considering Si donors in AlAs, DFT was used to get a better understanding of the LDOS around the Si donor, the donor state and its influence on the band structure. Also, this system was used as a proving grounds for methods of simulating STM images with DFT.

Band structure considerations

The presence of a Si donor was found to affect the band structure of clean (110) AlAs in three main ways. First, it generates a clear impurity band that is nearly flat at the Fermi level. This impurity band is given in figure 42. Second, the degeneracy of several states in the conduction band is lifted. This state degeneracy lifting is analogous to the band splitting at the edge and center of the Brillouin Zone when free electrons are subjected to a periodic potential in materials in the nearly free electron model. Another effect of the Si donor on the band structure is that the conduction band is shifted up approximately 0.2 eV.

STM image simulation

For a surface layer Si donor, STM images were simulated and compared to X-STM measurements. These simulated empty state STM images showed that there is a strong increase in DOS on the two As atoms neighbouring the Si donor. These simulated STM images correspond to a constant height image at a typical scanning height and are shown in figure 48b. This image was then compared to the empty state STM image of a surface layer Si donor shown in figure 54b. This image also showed an increase in LDOS extended over 2 group III sub-lattice sites, but the "v-shape" seen in figure 48 is not observed. The two lobes of the "v-shape" seeem to have merged together. The decrease in DOS around the feature seen in figure 54b is also not observed in DFT but the surface area of this decreased DOS is larger than the surface area of the supercell used in DFT so this could never be observed in such a supercell.

A collaborating group in Brazil worked in parallel on a similar system and the simulated STM image they found for a second layer Si donor matches very well with the X-STM measurements as shown in figure 56.

The results found for Si donors in AlAs show that DFT can indeed be used to successfully simulate STM images but that some important things need to be considered when comparing these simulated images to real STM measurements: simulations are constant height images while X-STM measurements are usually constant current mode, the size of the feature needs to be considered when constructing the supercell and X-STM samples different states at a given tip bias than DFT. With the current method of simulation, DFT integrates all states between the Fermi energy and the tip bias to generate an X-STM image. While in real X-STM images, the highest energy states that are allowed to tunnel will largely determine the image generated.

References

[1] R. R. Schaller, “Moore’s law: past, present and future,” *IEEE spectrum*, vol. 34, no. 6, pp. 52–59, 1997.

[2] Z. Durrani, M. Jones, F. Abualnaja, C. Wang, M. Kaestner, S. Lenk, C. Lenk, I. W. Rangelow, and A. Andreev, “Room-temperature single dopant atom quantum dot transistors in silicon, formed by field-emission scanning probe lithography,” *Journal of Applied Physics*, vol. 124, no. 14, p. 144502, 2018.

[3] P. Hohenberg and W. Kohn, “Inhomogeneous electron gas,” *Physical review*, vol. 136, no. 3B, p. B864, 1964.

[4] J. G. Lee, *Computational materials science: an introduction.* Crc Press, 2016.

[5] G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, “Surface studies by scanning tunneling microscopy,” *Physical review letters*, vol. 49, no. 1, p. 57, 1982.

[6] C. J. Chen and W. F. Smith, “Introduction to scanning tunneling microscopy,” *AmJPh*, vol. 62, no. 6, pp. 573–574, 1994.

[7] P. Offermans, P. M. Koenraad, J. H. Wolter, M. Beck, T. Aellen, and J. Faist, “Digital alloy interface grading of an InAlAs/InGaAs quantum cascade laser structure studied by cross-sectional scanning tunneling microscopy,” *Applied physics letters*, vol. 83, no. 20, pp. 4131–4133, 2003.

[8] P. Koenraad, M. Bozkurt, J. Garleff, V. Grant, R. Campion, T. Foxon, E. Marega, and G. Solomon, “Mn doping of InAs quantum dots studied by X-STM.,” *APS*, pp. Q22–003, 2009.

[9] P. Koenraad, C. Krammel, R. Plantenga, V. Kortan, M. Flatté, F. Tilley, M. Roy, P. Maksym, and T. Kita, “Individual iso-electronic N and Bi centers in GaAs studied by Scanning Tunneling Microscopy,” *APS*, vol. 2016, pp. X7–012, 2016.

[10] F. Tilley, M. Roy, P. Maksym, P. Koenraad, C. Krammel, and J. Ulloa, “Scanning tunneling microscopy contrast of isovalent impurities on the GaAs (110) surface explained with a geometrical model,” *Physical Review B*, vol. 93, no. 3, p. 035313, 2016.

[11] D. Tjeertes, T. Verstijnen, A. Gonzalo, J. M. Ulloa, M. Sharma, M. Felici, A. Polimeni, F. Biccari, M. Gurioli, G. Pettinari, *et al.*, “N-nH complexes in GaAs studied at the atomic scale by cross-sectional scanning tunneling microscopy,” *Physical Review B*, vol. 102, no. 12, p. 125304, 2020.

[12] E. G. Víllora, K. Shimamura, Y. Yoshikawa, T. Ujiie, and K. Aoki, “Electrical conductivity and carrier concentration control in β-Ga 2 O 3 by Si doping,” *Applied Physics Letters*, vol. 92, no. 20, p. 202120, 2008.

[13] Z. Ren, A. Taskin, S. Sasaki, K. Segawa, and Y. Ando, “Fermi level tuning and a large activation gap achieved in the topological insulator Bi 2 Te 2 Se by Sn doping,” *Physical Review B*, vol. 85, no. 15, p. 155301, 2012.

[14] Y. Zhang, A. Mascarenhas, H. Xin, and C. Tu, “Scaling of band-gap reduction in heavily nitrogen doped GaAs,” *Physical Review B*, vol. 63, no. 16, p. 161303, 2001.

[15] Z.-F. Wu, P.-F. Gao, L. Guo, J. Kang, D.-Q. Fang, Y. Zhang, M.-G. Xia, S.-L. Zhang, and Y.-H. Wen, “Robust indirect band gap and anisotropy of optical absorption in B-doped phosphorene,” *Physical Chemistry Chemical Physics*, vol. 19, no. 47, pp. 31796–31803, 2017.

[16] A. P. Wijnheijmer, *Manipulation and analysis of a single dopant atom in GaAs.* PhD thesis, PhD Thesis, Eindhoven University of Technology, 2011.

[17] K. Alberi, O. Dubon, W. Walukiewicz, K. Yu, K. Bertulis, and A. Krotkus, “Valence band anticrossing in $GaBi_xAs_{1-x}$,” *Applied Physics Letters*, vol. 91, no. 5, p. 051909, 2007.

[18] M. Weyers, M. Sato, and H. Ando, “Red shift of photoluminescence and absorption in dilute GaAsN alloy layers,” *Japanese Journal of Applied Physics*, vol. 31, no. 7A, p. L853, 1992.

[19] S. Francoeur, M.-J. Seong, A. Mascarenhas, S. Tixier, M. Adamcyk, and T. Tiedje, “Band gap of $GaAs_{1-x}Bi_x$, 0¡ x¡ 3.6%,” *Applied physics letters*, vol. 82, no. 22, pp. 3874–3876, 2003.

[20] R. Lewis, M. Masnadi-Shirazi, and T. Tiedje, “Growth of high Bi concentration $GaAs_{1-x}Bi_x$ by molecular beam epitaxy,” *Applied Physics Letters*, vol. 101, no. 8, p. 082112, 2012.

[21] K. Bertulis, A. Krotkus, G. Aleksejenko, V. Pačebutas, R. Adomavičius, G. Molis, and S. Marcinkevičius, “GaBiAs: A material for optoelectronic terahertz devices,” *Applied physics letters*, vol. 88, no. 20, p. 201112, 2006.

[22] V. P. LaBella, M. R. Krause, Z. Ding, and P. M. Thibado, “Arsenic-rich GaAs (0 0 1) surface structure,” *Surface science reports*, vol. 60, no. 1-4, pp. 1–53, 2005.

[23] M. Masnadi-Shirazi, D. Beaton, R. Lewis, X. Lu, and T. Tiedje, “Surface reconstructions during growth of $GaAs_{1-x}Bi_x$ alloys by molecular beam epitaxy,” *Journal of crystal growth*, vol. 338, no. 1, pp. 80–84, 2012.

[24] Z. Batool, K. Hild, T. Hosea, X. Lu, T. Tiedje, and S. Sweeney, “The electronic band structure of GaBiAs/GaAs layers: Influence of strain and band anti-crossing,” *Journal of Applied Physics*, vol. 111, no. 11, p. 113108, 2012.

[25] S. Zhang and S.-H. Wei, “Nitrogen solubility and induced defect complexes in epitaxial GaAs: N,” *Physical Review Letters*, vol. 86, no. 9, p. 1789, 2001.

[26] D. A. Beaton, A. Mascarenhas, and K. Alberi, “Insight into the epitaxial growth of high optical quality $GaAs_{1-x}Bi_x$,” *Journal of Applied Physics*, vol. 118, no. 23, p. 235701, 2015.

[27] R. Kini, A. Ptak, B. Fluegel, R. France, R. Reedy, and A. Mascarenhas, “Effect of Bi alloying on the hole transport in the dilute bismide alloy $GaAs_{1-x}Bi_x$,” *Physical Review B*, vol. 83, no. 7, p. 075307, 2011.

[28] C. Krammel, M. Roy, F. Tilley, P. Maksym, L. Zhang, P. Wang, K. Wang, Y. Li, S. Wang, and P. Koenraad, “Incorporation of Bi atoms in InP studied at the atomic scale by cross-sectional scanning tunneling microscopy,” *Physical Review Materials*, vol. 1, no. 3, p. 034606, 2017.

[29] “Plantenga, R. C. (2019). Probing impurities at the atomic scale in semiconductor structures. Eindhoven: Technische Universiteit Eindhoven. .”

[30] P. Mooney, “Deep donor levels (DX centers) in III-V semiconductors,” *Journal of Applied Physics*, vol. 67, no. 3, pp. R1–R26, 1990.

[31] M. Shayegan, E. De Poortere, O. Gunawan, Y. Shkolnikov, E. Tutuc, and K. Vakili, “Two-dimensional electrons occupying multiple valleys in AlAs,” *physica status solidi (b)*, vol. 243, no. 14, pp. 3629–3642, 2006.

[32] P. Bhattacharya, R. Fornari, and H. Kamimura, *Comprehensive Semiconductor Science and Technology: Online Version*. Newnes, 2011.

[33] W. Nakwaski, “Effective masses of electrons and heavy holes in GaAs, InAs, A1As and their ternary compounds,” *Physica B: Condensed Matter*, vol. 210, no. 1, pp. 1–25, 1995.

[34] W. Moore and R. Holm, "Infrared dielectric constant of gallium arsenide," *Journal of applied physics*, vol. 80, no. 12, pp. 6939–6942, 1996.

[35] C. Authors and editors of the volumes III/17A-22A-41A1a, "Aluminum arsenide (AlAs) dielectric constants," *Group IV Elements, IV-IV and III-V Compounds. Part a-Lattice Properties*, pp. 1–4, 2001.

[36] P. Verma and R. J. Bartlett, "Increasing the applicability of density functional theory. IV. Consequences of ionization-potential improved exchange-correlation potentials," *The Journal of chemical physics*, vol. 140, no. 18, p. 18A534, 2014.

[37] U. H. Bakhtiar, R. Ahmed, R. Khenata, M. Ahmed, and R. Hussain, "A first-principles comparative study of exchange and correlation potentials for ZnO," *Materials science in semiconductor processing*, vol. 16, no. 4, pp. 1162–1169, 2013.

[38] G. P. Srivastava, *The physics of phonons*, pp. 14–17. Routledge, 2019.

[39] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials," *Journal of Physics: Condensed Matter*, vol. 21, no. 39, p. 395502 (19pp), 2009.

[40] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. D. Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, "Advanced capabilities for materials modelling with QUANTUM ESPRESSO," *Journal of Physics: Condensed Matter*, vol. 29, no. 46, p. 465901, 2017.

[41] "We used the pseudopotentials Ga.pz-dn-rrkjus_psl.0.2.UPF, As.pz-n-rrkjus_psl.0.2.UPF, Bi.pz-dn-rrkjus_psl.1.0.0.UPF, Al.pz-n-rrkjus_psl.0.1.UPF, Si.pz-n-rrkjus_psl.0.1.UPF and N.pz-n-rrkjus_psl.0.1.UPF from http://www.quantum-espresso.org."

[42] A. Kokalj, "XCrySDen—a new program for displaying crystalline structures and electron densities," *Journal of Molecular Graphics and Modelling*, vol. 17, no. 3-4, pp. 176–179, 1999.

[43] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations," *Physical review B*, vol. 13, no. 12, p. 5188, 1976.

[44] ""Aluminum arsenide (AlAs), crystal structure, lattice parameters, thermal expansion: Datasheet from Landolt-Börnstein - Group III Condensed Matter · Volume 41A1β: "Group IV Elements, IV-IV and III-V Compounds. Part b - Electronic, Transport, Optical and Other Properties" in SpringerMaterials (https://doi.org/10.1007/10832182_80)"." Copyright 2002 Springer-Verlag Berlin Heidelberg.

[45] W. M. Haynes, *CRC handbook of chemistry and physics*. CRC press, 2014.

[46] J. Harris, B. Joyce, and P. Dobson, "Oscillations in the surface structure of Sn-doped GaAs during growth by MBE," *Surface Science Letters*, vol. 103, no. 1, pp. L90–L96, 1981.

[47] J. Steele, R. A. Lewis, J. Horvat, M. J. B. Nancarrow, M. Henini, D. Fan, Y. Mazur, M. Schmidbauer, M. Ware, S.-Q. Yu, *et al.*, "Surface effects of vapour-liquid-solid driven Bi surface droplets formed during molecular-beam-epitaxy of GaAsBi," *Scientific reports*, vol. 6, p. 28860, 2016.

[48] A. Duzik, J. C. Thomas, A. van der Ven, and J. M. Millunchick, "Surface reconstruction stability and configurational disorder on Bi-terminated GaAs (001)," *Physical Review B*, vol. 87, no. 3, p. 035313, 2013.

[49] R. P. Beres, R. E. Allen, J. P. Buisson, M. A. Bowen, G. F. Blackwell, H. P. Hjalmarson, and J. D. Dow, "Bound and resonant surface states at the (110) surfaces of AlSb, AlAs, and AlP," *Journal of Vacuum Science and Technology*, vol. 21, no. 2, pp. 548–550, 1982.

[50] K. Hoang, S. Mahanti, and M. G. Kanatzidis, "Impurity clustering and impurity-induced bands in PbTe-, SnTe-, and GeTe-based bulk thermoelectrics," *Physical Review B*, vol. 81, no. 11, p. 115106, 2010.

[51] J. Tersoff and D. R. Hamann, "Theory of the scanning tunneling microscope," *Phys. Rev. B*, vol. 31, pp. 805–813, Jan 1985.

[52] J. R. Chelikowsky and M. L. Cohen, "Self-consistent pseudopotential calculation for the relaxed (110) surface of GaAs," *Physical Review B*, vol. 20, no. 10, p. 4150, 1979.

[53] G. De Raad, D. Bruls, P. Koenraad, and J. Wolter, "Interplay between tip-induced band bending and voltage-dependent surface corrugation on GaAs (110) surfaces," *Physical Review B*, vol. 66, no. 19, p. 195306, 2002.

[54] K. Teichmann, M. Wenderoth, S. Loth, R. Ulbrich, J. Garleff, A. Wijnheijmer, and P. Koenraad, "Controlled charge switching on a single donor with a scanning tunneling microscope," *Physical review letters*, vol. 101, no. 7, p. 076103, 2008.

[55] J. Garleff, A. Wijnheijmer, C. vd Enden, and P. Koenraad, "Bistable behavior of silicon atoms in the (110) surface of gallium arsenide," *Physical Review B*, vol. 84, no. 7, p. 075459, 2011.

[56] A. Wijnheijmer, J. Garleff, K. Teichmann, M. Wenderoth, S. Loth, and P. Koenraad, "Single Si dopants in GaAs studied by scanning tunneling microscopy and spectroscopy," *Physical Review B*, vol. 84, no. 12, p. 125310, 2011.

[57] R. M. Feenstra, G. Meyer, F. Moresco, and K. Rieder, "Low-temperature scanning tunneling spectroscopy of n-type GaAs (110) surfaces," *Physical Review B*, vol. 66, no. 16, p. 165204, 2002.

[58] M. Van Der Wielen, A. Van Roij, and H. Van Kempen, "Direct Observation of Friedel Oscillations around Incorporated S i Ga Dopants in GaAs by Low-Temperature Scanning Tunneling Microscopy," *Physical review letters*, vol. 76, no. 7, p. 1075, 1996.

[59] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, "The SIESTA method for ab initio order-N materials simulation," *Journal of Physics: Condensed Matter*, vol. 14, no. 11, p. 2745, 2002.