A Density Functional Theory Study on the properties of Dopants in Silicon Nanostructures

J.T.L. Poulton

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Department of Physics and Astronomy
University College London

March 5, 2020
I, J.T.L. Poulton, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.
Impact Statement

The work presented in this thesis revisits a topic that has been subject to much previous study, dopants and defects in silicon. Yet the work here is an attempt to understand the unique properties of structures that have only recently become of interest as fabrication techniques have seen significant advances in precision allowing for deterministic doping and device manufacturing on an atom-by-atom basis. These technological advances bring with them a wide range of exciting applications, from the construction of truly identical transistors to on demand photon sources, information storage and silicon based solid state quantum computing.

In order to further the theoretical understanding of the electronic nature of the dopant structures that could make some of these technologies possible a range of systems have been considered in this thesis. These include the study of the electronic structure of group V dopants that have potential applications in qubit construction and quantum information processing. This work will hopefully aid in future experimental characterisation of the respective dopant structures. Studies of group III dopants were also performed, resulting in the identification of a previously unknown dopant complex existing in a given state, the properties of which could in principle be tuned. These findings should guide future fabrication processes involving these materials.

In addition, a careful characterisation of a novel method designed to improve the efficiency of orbital based density functional theory codes was performed. It is hoped that the findings from this study will be utilised in future theoretical work and ensure large scale simulations can be performed both more accurately and efficiently. It should be noted that while the technological promise of the material and tech-
nologies that have motivated the work within this thesis is compelling, the work itself represents an early stage of research. Therefore the scope for direct impact will be limited and much further work, which is undoubtedly forthcoming, will be needed to realise this.
Abstract

Advances in experimental techniques over the last decade have resulted in the realisation of atomically precise doping of silicon and the creation of single dopant devices. Therefore, a greater theoretical understanding of deep dopant structures and the resulting defect complexes are required. In this thesis, we apply a variety of computational methods based on density functional theory (DFT) to the study of dopants and defect centres in bulk silicon and silicon nanostructures. The calculations are performed using the plane wave based code VASP[1] and the linear scaling code Conquest[2]. Firstly, we discuss the effects of self-compensation in p-type doped bulk silicon in the case study of single aluminium dopants and aluminium dopant pairs in bulk silicon. We consider the formation of dimer type complexes and determine whether such complexes will be electrically active. We then investigate the formation of dopant vacancy pairings in silicon using bismuth dopants, which are not commonly used in the semiconductor industry but may be used to fabricate solotronic devices or for quantum information processing type applications. The DFT simulations performed are compared with experimental data. The incorporation of phosphorus in silicon is also studied within this by analysing phosphorus δ-doped layers. The effects of dopant depth and the diffusion of delta-doped layers upon the electronic structure are presented. Furthermore, this thesis includes some developmental work attempting to improve upon the efficiency multi-site support functions[3] when used in conjunction with the linear scaling methods within the Conquest code.
Acknowledgements

Firstly, I would like to thank my supervisor, Professor David Bowler, it has been a pleasure working with him and I could not be more grateful for the infinite patience he has shown while providing guidance throughout my PhD. I would also like to thank my co-supervisor Professor Andrew Fisher and all members of the ADDRFS research group for providing invaluable insights and asking thought provoking questions. On that note it would be amiss not to mention the members of the NIMS group, particularly Miyazaki-Sensei and Nakata-San with whom I spent some of the most enjoyable periods of this PhD.

In addition, I would like to thank the members of the Bowler group; Shereif Mujahed, Kane Shenton and Jack Baker for countless meandering methodological discussions which, although often quite random in nature, were undoubtedly helpful. As well I would like to extend my gratitude to all past and present members of the A9 and A22 offices (notably Louis Fry-Bourdaix, Francesco Silvestri, Giulia Lucarelli, Konstantinos Konstantinou, Katie Porsch and Alex Broad) for providing an enjoyable environment in which to work. I hope that all of these people will remain lifelong friends. In particular, I owe thanks Callum Gray for providing copious amounts of coffee and climbing based distractions whenever needed.

Finally, I would like to dedicate this thesis to my family who have been endlessly supportive of my academic pursuits and without whom this work would not have been possible.
# Contents

1 Introduction .......................... 16
   1.1 Implantation of Dopants ............... 18
       1.1.1 Molecular Beam Epitaxy ......... 18
       1.1.2 Ion Implantation ............... 18
   1.2 Precision doping ...................... 19
   1.3 Delta doping ......................... 21
   1.4 Materials Modelling .................. 22
   1.5 Thesis Outline ....................... 23

2 Many Body Theory and DFT methodology .... 25
   2.1 The Many Body Problem ............... 25
   2.2 The Born-Oppenheimer Approximation .... 26
   2.3 The Traditional Approach to DFT ....... 27
       2.3.1 Hohenberg-Kohn Theorems ........ 27
       2.3.2 The Kohn-Sham ansatz ......... 28
       2.3.3 Achieving Self-Consistency ...... 30
   2.4 Exchange-Correlation ................. 30
   2.5 Pseudopotentials ..................... 32
   2.6 Basis Sets ........................... 33
       2.6.1 Planewaves ..................... 34
       2.6.2 Pseudo-atomic Orbitals ........ 34
   2.7 Practical Limitations of Basis sets in DFT .... 35
   2.8 Linear scaling DFT and CONQUEST Methodology .... 36
## Contents

2.8.1 Localisation of the Density Matrix .......................... 36
2.8.2 Imposing the eigenvalue range of K .......................... 38
2.8.3 The Search for the Ground State Energy ....................... 39
2.8.4 Real Space Grid ............................................. 44
2.8.5 The CONQUEST Routine ..................................... 45

3 Developments in CONQUEST ........................................ 47

3.1 Filter Diagonalisation and on-site support functions ............... 47
  3.1.1 Background .................................................. 47
  3.1.2 Theory ...................................................... 49
3.2 Computational Details and Results ................................ 52
  3.2.1 Diagonalisation .............................................. 52
  3.2.2 Energy-Volume Curve Tests ................................. 55
  3.2.3 Order N ...................................................... 59
3.3 Conclusions ....................................................... 65
3.4 Returning to Self consistency ..................................... 66
  3.4.1 Theory ...................................................... 66
  3.4.2 Practical implementation within CONQUEST ................. 69
  3.4.3 Results ...................................................... 71
  3.4.4 Conclusions ................................................ 76

4 Self compensation and active Aluminium dopants in Si .......... 78

4.1 Background ....................................................... 78
4.2 Methods ........................................................ 79
  4.2.1 Computational Details ........................................ 79
  4.2.2 Structural Models ........................................... 80
  4.2.3 Electronic Structure .......................................... 81
4.3 Results and Discussion ........................................... 82
  4.3.1 Structural Features .......................................... 82
  4.3.2 Electronic Structure .......................................... 84
4.4 Conclusions ....................................................... 91
## 5 A structural study of single Bismuth dopants in crystalline Si

5.1 Background .................................. 93
5.2 Methods ........................................ 94
   5.2.1 Computational Details ...................... 94
   5.2.2 Structural Models ......................... 95
5.3 Results and Discussion ....................... 96
   5.3.1 Relative Energetic Stabilities ............... 96
   5.3.2 Introducing Additional Vacancies ............ 101
5.4 Conclusions .................................. 103

## 6 Surface effects on P δ-layers in Si slabs

6.1 Background .................................. 105
6.2 Methods ........................................ 108
6.3 Results and Discussion ....................... 110
   6.3.1 The effects of vacuum ...................... 110
   6.3.2 The influence of dopant densities .......... 112
   6.3.3 Dopant density effects on electronic structure parameters . 115
   6.3.4 Surface effects on electronic structure parameters ........ 115
   6.3.5 Disordered δ-layers ....................... 116
6.4 Conclusions .................................. 119

## 7 Conclusions and Outlook

Appendices

A Appendix A -OSSF Supplementary Information 125
   A.1 The Choice of Filtration Function .............. 125

B Appendix B -P δ-layers Supplementary Information 127

Bibliography 129
## List of Figures

1.1 Atomically precise doping schematic a) Hydrogen(blue) terminated Si(grey), b) Controlled removal of H using STM c) Deposition of P, taking advantage of dangling bonds ........................................ 21  
1.2 a) An individual P dopant place in Si from , b) A quantum dot consisting of multiple P dopants constructed using STM lithography . 21  
1.3 A schematic showing the length and time scales accessible using different materials modelling methods ........................................ 23  
2.1 A plot of the pseudised wavefunction and all electron wavefunction for the 1s orbital of Si generated using Pseudo-dojo ............ 33  
3.1 Volume energy curves for bulk Si using the primitive, OSSF and MSSF basis sets, fitted using the Birch-Murnaghan equation .... 55  
3.2 Binding energy curves for crystalline Ge and C using the primitive, OSSF and MSSF basis sets, fitted using the Birch-Murnaghan equation ................................................................. 58  
3.3 Binding energy curves for crystalline SrTiO$_3$ and PbTiO$_3$ using the primitive, OSSF and MSSF basis sets, fitted using the Birch-Murnaghan equation ................................................. 60  
3.4 An atomic diagram of a P dimer adsorbed onto the Si(001) surface with ................................................................. 64  
3.5 The density matrix residual against the number of iterations in the self consistency cycle for the contracted and primitive basis sets . 65
List of Figures

3.6 Atomic diagrams of the 128 and 256 electron crystalline Si structures extended in the z-dimension, produced using the Jmol visualisation package ............................................ 72

3.7 Front and side views Si nanowires used to test density convergence with modified residual definitions ......................................................... 74

3.8 The convergence behaviour for the different residual definitions against the number of iterations in the SCF cycle ..................................... 75

3.9 The convergence behaviour of energy difference for the different residual definitions against the number of iterations in the SCF cycle 76

4.1 A 2D plot of total charge density across the \{10\bar{1}\} plane containing the dopant for the Si$_{215}$Al$_1$ structure in the $S = \frac{1}{2}$ and $S = 0$ spin states 84

4.2 PDOS for both undoped Si and the Si$_{215}$Al$_1$ structure in the $S = 0$ and the $S = \frac{1}{2}$ spin state ...................................................... 85

4.3 A 2D plot of total charge density across the \{10\bar{1}\} plane containing the dopant for the Si$_{214}$Al$_2$ structure in the $S = 1$ and $S = 0$ spin states 88

4.4 PDOS for the Si$_{214}$Al$_2$ structure in the $S = 0$ spin state .......... 89

4.5 PDOS for the Si$_{214}$Al$_2$ structure in the $S = 1$ spin state .......... 90

4.6 Partial charge density isosurface at a value of $2.58 \times 10^{-3}$eV/Å, for the highest occupied state of of Si$_{510}$Al$_2$ in the $S = 0$ and the $S = 1$ spin up states ................................................................. 90

4.7 2D ELF plots across the \{10\bar{1}\} plane containing both neighbouring Aluminium dopants in S=0 and S=1 states. ....................... 91

5.1 Atomic models of the different configurations of Bi dopants considered with and without vacancies at nearest neighbouring substitutional lattice sites ............................................. 96

5.2 Subsections of the total charge density across the \{10\bar{1}\} and \{101\} planes for Bi and vacancy complexes following atomic relaxation . . 100

5.3 An atomic model of a subsection of the relaxed structure containing a Bi dopant with two neighbouring substitutional vacancies ........... 102
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>STM images (yet to be published) of Bi doped Si at biases of -1.0 and +1.0eV courtesy of Dr. Manuel Siegl (formerly of UCL)</td>
</tr>
<tr>
<td>6.1</td>
<td>A schematic of a P δ-doped Si(001) simulation cell</td>
</tr>
<tr>
<td>6.2</td>
<td>The bandstructures for the doped and undoped 80 Si layer structures without an explicit vacuum along the Γ-M path. The doped structure (b) includes the band splitting with minima at the Γ and Δ points labelled 1Γ, 2Γ and 1Δ, the splitting from 1Γ to the CBM is indicated by a red arrow.</td>
</tr>
<tr>
<td>6.3</td>
<td>The bandstructures and atomic schematics in the x-y plane for ordered and disordered 1/4ML δ-layers</td>
</tr>
<tr>
<td>6.4</td>
<td>The bandstructure and atomic schematic of a 1/4ML δ-layer extended in the z-dimension</td>
</tr>
</tbody>
</table>
## List of Tables

3.1 Energy and forces for H$_2$O molecule using both primitive support functions and OSSFs ........................................ 53

3.2 Bulk modulus and lattice constant for crystalline Si using both primitive support functions and OSSFs .......................... 55

3.3 Bulk modulus and lattice constant for crystalline Ge/C using both primitive support functions and OSSFs .................... 56

3.4 Bulk modulus and Equilibrium volume for SrTiO$_3$/PbTiO$_3$ using both primitive support functions and OSSFs ............. 57

3.5 The DFT energies for 8 atom Si using OSSF O(N) ................. 61

3.6 The relaxed neutral vacancy formation energies for 64, 216 and 512 atom crystalline bulk Si with a vacancy at a substitutional site using OSSF O(N) ................................................. 62

3.7 The computational run times for 64, 216 and 512 atom crystalline bulk Si with a vacancy at a substitutional site using OSSF O(N) .................................................... 63

3.8 The binding energies and computational run times for a P-dimer on the Si(001) surface using O(N) with a primitive SZP basis and contracted DZDP basis ................................................. 65

3.9 The HF and total DFT energy differences for the elongated crystalline Si cells for the different residual definitions upon convergence ....................................................... 72

3.10 The total number of iterations in the SCF cycle for the elongated crystalline Si cells for the different residual definitions upon convergence ...................................................... 73
List of Tables

3.11 The HF and total DFT energy differences for the elongated crystalline Si cells for the different residual definitions upon convergence

4.1 The separations between adjacent Al dopants and the nearest neighbouring Si atoms, shown as $\text{Si}_{NN}$, as well as the distance to the next nearest neighbouring Si atoms denoted here as $\text{Si}_{NNN}$. The bond angles and relaxed lattice constants for spin polarised and spin free simulations are included.

4.2 The formation energies of the different spin polarised and spin free structures with pairs of dopants occupying separated and adjacent substitutional sites labelled as 'sep' and 'adj' respectively.

5.1 The relative energetic stabilities of the structures containing Bi dopants.

5.2 Vacancy formation energies with pure Si and a Bi doped Si cell.

5.3 The relative energetic stabilities of the structures containing Bi dopants using plane wave methods.

5.4 The formation energy of a Bi dopant for Si crystalline structures with multiple vacancies neighbouring the dopant.

6.1 The electronic structure parameters, including the band-splittings at the $\Gamma$-point, of the 1/4 ML $\delta$-layers for Si(001) surfaces.

6.2 The CBM and Fermi energy values for the 20 layer H-passivated Si(001) converged using different SCF tolerance values.

6.3 The electronic structure parameters, including the band-splittings at the $\Gamma$-point, of the 1/4, 1/8 and 1/16ML $\delta$-layers for Si(001) surfaces.

6.4 The electronic structure parameters, including the band-splittings at the $\Gamma$-point, of the 1/4, 1/8 and 1/16ML $\delta$-layers at depths of approximately 1, 2, 3 and 4nm in Si(001).

6.5 The electronic structure parameters, including the band-splittings at the $\Gamma$-point, of the disordered 1/4 ML structures.
List of Tables

A.1 The binding energies for 8 atom Si using OSSF with finltration functions defined using the FD and tanh functions . . . . . . . . . . 126

B.1 Properties of 8 atom Bulk Si cell depending basis set quality . . . . 127
Chapter 1

Introduction

Semiconductor devices have become a ubiquitous component of modern life with their continued development key for recent technological advances in computing, telecommunications and more generally electronic systems as a whole. The development referred to here is the miniaturisation of semiconductor devices that has taken place over the last few decades, a trend first formally predicted by G.E. Moore in 1965[4]. Moore stated that the complexity of semiconductor devices, which is to say the density of transistors in such devices, would double every year, a claim he subsequently revised to two years[5], this eventually came to be known as Moore’s law. This law was adhered to, more or less, for over fifty years and became a benchmark for progress as much as a measure of it. The ambition to achieve the high component density required by Moore’s law has resulted in a continuous miniaturisation of devices and although Moore’s law is now widely considered to be dead it has led to the current research of reliable, precise nano-scale doping of semiconductors. A branch of research which has the grand, but very realisable aim of manipulating material properties by controlling the positions of single atoms.

In order to achieve this degree of control one must begin by introducing the above mentioned atoms to a material through doping, a process which has proved a crucial tool for manipulating the electronic, optical and magnetic properties of semiconductors. The origins of doping can be traced back as far as 1930 when the influence of impurities on the conductivity of semiconductors was first noted by Gudden[6, 7]. Later the understanding of this influence became more sophisticated with the de-
velopment of the theory of p-type and n-type doping. This is essentially the idea of introducing additional holes (p-type) or electrons (n-type) via dopants, which was introduced in works on rectifiers based on copper oxide and selenium\[7, 8\]. However due to the difficulties of manufacturing doped devices it was not until the development of the first transistor at Bell laboratories in the mid-20\textsuperscript{th} century\[9\] that doping became a widely used technique.

The material that has since almost exclusively been used as the basis of semiconductor devices and therefore subject to this doping is silicon. It has become dominantly used for a number of reasons; firstly the relative abundance of Si has made it an economically favourable option to other viable materials such as germanium, secondly Si has a stable oxide form (SiO\textsubscript{2}), which can serve as a dielectric. The existence of this oxide is beneficial for a number of reasons; it has a wide band gap making it a useful insulator, it can be used as a blocking material in ion implantation, detailed later, and it has few mechanical or electrical defects at the interface with Si\[10, 11\]. Si has further advantages, amongst which is the high maximum dopant concentration of the material, which is of the order of 1 \times 10^{21} \text{ atoms per cm}^{-3}[12, 13, 14].

All of these factors have contributed to the sheer popularity of Si and this in turn has resulted in the development of sophisticated extraction and purification methods for the crystalline form, including but not limited to the Czochralski process\[15\]. This process typically obtains ‘nine-9’ or 99.9999999% purity levels resulting in a nearly defect free material.

Hence Si is the logical, and perhaps obvious choice for testing techniques for atomically precise doping, the aim of which is to introduce single dopants at the exact position desired within an atomic lattice. This will be detailed below as the primary motivation of this thesis is to further the understanding of the outcomes of introducing dopants to Si with such precision. However, it seems necessary to firstly consider the experimental methods, and their limitations, that predated and in turn led to the advent of atomically precise doping.
1.1 Implantation of Dopants

Several techniques exist for the doping of Si, with the most commonly used techniques yielding greatly varying results. Often this because there are different desired outcomes, such as surface concentrations and dopant profiles. The doping itself is typically performed with B, P, Ga, As or Sb, although in theory any group III or group V element could act as a p or n-type dopant respectively, assuming it can be reasonably implanted. The methods outlined below are intended to illustrate these differences and these sections are by no means intended as an all encompassing review of doping methodologies.

1.1.1 Molecular Beam Epitaxy

Molecular beam epitaxy (MBE) is an atomic layer by layer growth technique, when used for doping this involves growing a layer of dopant rich material upon a usually undoped substrate such as Si\cite{16}. MBE is performed with the substrate ultra-high vacuum chamber, molecular beams are generated by heating either gaseous or solid sources of the elements desired for doping in effusion cells. The term beam here meaning that the atoms from individual sources do not interact until reaching the substrate as a result of the long mean free path of the atoms.

This method allows precise control of the thickness of each atomic layer deposited on the substrate which can be monitored using reflection high-energy electron diffraction. Furthermore MBE allows for the creation of precise doping profiles with high reproducibility, at concentrations up to \(10^{20} \text{ cm}^{-3}\) in Si. This allows for the fabrication of multilayered structures based on Si, including quantum wires, as this method does produce relatively few defects (\(\approx 10^{-7} \text{ cm}^2\))\cite{17}.

1.1.2 Ion Implantation

In simple terms, ion-beam implantation (IBI) is a doping technique that involves firing dopant ions into the chosen substrate usually with energies of tens or hundreds of keV, depending on the dopant and desired penetration depth\cite{18}. The ions
are typically angled upon firing to avoid any ion channelling effects where the ions travel through open channels in a diamond structure, if no such effects occur then the penetration depth will follow a Gaussian distribution[19]. Although in Si a dopant will often occupy substitutional sites following implantation the complexes formed by ion implantation can be more complicated, including interstitial point defects and vacancies. As many of the dopant atoms may not be situated in electrically active substitutional sites immediately after implantation, post-implantation processing to anneal out the defects and activate the dopants is then required, and much research has been devoted to minimising the diffusion of implanted dopants in this process[20].

This technique can be used for, amongst other things, the fabrication of integrated circuits[21], producing the p-n junction for photovoltaic devices[22] or adjusting the threshold of a metal oxide semiconductor field-effect transistor (MOSFET)[23]. In addition focused ion beams allow precision of the order of nanometres[24] when fabricating such features, this has been extended further to single ion implantation[25], which allows the deposition of single ions with high spatial accuracy and will be covered in more detail in the next section.

1.2 Precision doping

The need for ever shrinking semiconductor devices combined with the ability to produce extremely high purity Si has launched research conducted into the manufacture of devices with one, and only one, dopant. These solitary dopant optoelectronic devices, or solotronic devices[26], represent the logical limit of Moore’s law where the miniaturisation of semiconductor devices has gone as far as the discrete nature of matter will allow.

The motivation for such devices does extend beyond pushing Moore’s law to its limits however, some proposed devices require accurate three-dimensional dopant incorporation, profiling and modelling. For example, it was realised that single impurities could act as the qubits as a basis for solid state quantum computation, by utilising a single phosphorus dopant’s nuclear spin[27], or a single dopant’s bound
1.2. Precision doping

Though the uses of single dopants extend far beyond qubits, they can also be applied as non-classical light sources in quantum information science for quantum key distribution systems, quantum repeaters, single electron transistors as well as for quantum lithography, multivalent logic and local sensing. In fact, single defect devices have already shown some of their potential, with nitrogen-vacancy centres in diamond being used as a non-classical light source.

In addition to the development of potential solotronic devices listed above, there are other advantages to be gained from the further study and utilisation of single dopants, as a specific dopant will have reproducible quantized properties. However, a reliance on the properties of a solitary dopant generates new challenges: device behaviour depends on the local environment, i.e. strain, electric, magnetic and optical fields, and the dopant position within the device. Hence there is a strong need for precision when introducing the solitary dopant, and fabrication techniques that have atomic level precision have been realised, two of which are briefly detailed below.

As mentioned earlier one such approach is single ion implantation, this doping technique was recently demonstrated with high spatial accuracy by Pacheco et al. The study used a low energy ion beam at around 100 keV, focused by a series of lenses and apertures. The beam was pulsed at 16ns to ensure discrete implantation of dopants. For the example of antimony in Si the accuracy of the donor position could be localised to a sphere of radius of approximately 10nm. A similar outcome can be achieved by pre-arranging an array of ions in a magnetic optical trap rather than pulsed implantation. There are also other influences on the precision than how finely focused the beam is, including ion straggling and annealing, however we will not cover these in detail here.

An alternative technique, and the one favoured by experimentalists based at UCL, for the atomically precise placement of dopants in Si has recently been demonstrated for phosphorus, using a scanning tunnelling microscope (STM). The single atom manipulation capabilities of STM are used to selectively
remove, or desorb, individual hydrogen atoms from a H-passivated Si surface then phosphene gas is used to selectively deposit P onto the desorbed area, taking advantage of the increased reactivity of the newly formed dangling Si bonds compared to the H terminated surface, a simple schematic of which is shown in Figure 1.1.

This process, not only allows the atomically precise placement of P but can also be used to create larger structures, enabling the fabrication of whole devices with atomic scale precision, examples of both cases are illustrated in Figure 1.2.

### 1.3 Delta doping

In addition to the numerous possible solotronic applications the ability to precisely control and place dopants in Si granted by the advances in fabrication techniques outlined above gives a new approach to delta($\delta$) doping. This is the process by which dopant atoms are introduced in what should ideally be an atomically abrupt...
monolayer region known as a $\delta$-layer and had previously relied on the interruption of epitaxial growth for fabrication[39]. Although the ideal monolayer is not achieved even with precision doping due to diffusion effects.

Inside a silicon $\delta$-layer, as many as one in four silicon atoms can be substituted by a phosphorus donor, which is referred to as a doping density of $\frac{1}{4}$ ML[40]. At such high doping densities, the donor electrons behave similarly to a two-dimensional electron gas. Furthermore within these $\delta$-layers structures such as nanowires[41] and tunnel junctions[42] can be constructed. However there are significant gaps in the understanding of the physics of two-dimensional highly doped structures. For example, low-level structures inherent to devices can exhibit varying degrees of confinement, and determining the location of energy levels in these structures will prove crucial to the successful operation of any constructed devices. A part of this thesis will be dedicated to furthering the theoretical understanding of two dimensional P $\delta$-layers in Si.

1.4 Materials Modelling

The theoretical understanding mentioned above is derived from computational simulations performed using density functional theory (DFT)[43]. The field of materials modelling has grown in popularity, now commonly used in support of experimental findings, as the codes used to perform such simulations become more stable, accurate and computationally inexpensive. It is a large and varied field, both in terms of the systems that are studied and the physical theories that are employed; ranging from extremely ordered, periodic systems such as the crystalline condensed matter to disordered matter such as biological systems.

The systems considered in this thesis fall into the former category and do present a number of interesting theoretical challenges. The disruption on the crystalline lattice of Si caused during the process of precision doping can be extremely long-ranged; this therefore requires the use of large system sizes in order to sufficiently approach the dilute limit in order to accurately calculate fundamental dopant prop-
properties such as the formation energy, the lattice relaxation around a given dopant defect structure, and the relative stability of different dopant configurations.

The length scales considered fall perfectly into the realms of DFT, opposed to other material modelling techniques, a rough schematic of typical scales is shown in Figure 1.3. It is for this reason, combined with the high levels of accuracy of DFT due to its *ab initio* nature, that it is the chosen methodology for the work presented in this thesis.

### 1.5 Thesis Outline

In this thesis, we apply a variety of computational methods based on density functional theory to the study of dopants and defect centres in bulk silicon and silicon nanostructures. The calculations are performed using the plane wave based code VASP and the local orbital code CONQUEST.

Firstly we discuss some developmental work in Chapter 3, investigating improvements upon the efficiency multi-site support functions when used in conjunction with the linear scaling methods within the CONQUEST code.

In chapter 4, we discuss the effects of self-compensation in p-type doped bulk sil-
icon in the case study of single aluminium dopants and aluminium dopant pairs in bulk silicon. We consider the formation of dimer type complexes and determine whether such complexes will be electrically active.

We then, in chapter 5, investigate the formation of dopant vacancy pairings in silicon using bismuth dopants, which are not commonly used in the semiconductor industry but may be used to fabricate solotronic devices or for quantum information processing type applications. The DFT simulations performed are compared with experimental data.

Furthermore in Chapter 6, the incorporation of phosphorus in silicon is also studied by analysing phosphorous δ-doped layers, the effects of dopant depth and the diffusion of δ-doped layers upon the electronic structure are presented. Overall, we aim to show that large-scale DFT calculations can be employed to accurately simulate many properties of interest for point defects in semiconductors, thus providing both new insight into the electronic structure of bulk defects and dopants, as a theoretical understanding of such features becomes increasingly important with the advent of precision doping methodologies.
Chapter 2

Many Body Theory and DFT methodology

The following chapter outlines the basic principles of many body quantum mechanics and the reasoning behind the workhorse of much modern condensed matter research, density functional theory (DFT). The introduction of such concepts may seem superfluous, given the current ubiquity of DFT in the condensed matter community it will undoubtedly be familiar to anyone likely to be reading this thesis. Nonetheless I believe this summary will provide a convenient reference for the reader as the concepts of DFT will be utilised and discussed throughout this work. Furthermore although the ideas covered here are well established, the chapter should ensure there is no ambiguity of convention or notation by defining the terms used in chapters to follow.

2.1 The Many Body Problem

One of the key postulates of quantum mechanics is that the behaviour of a system of electrons and atomic nuclei can be described in its entirety by the wavefunction of the system. The wavefunction itself is governed by the Schrödinger equation[44], given below in its non-relativistic time independent form:

\[ \hat{H} \Psi(r, R) = E \Psi(r, R), \]  

(2.1)
where $\hat{H}$ is the Hamiltonian of our system of electrons and nuclei, $E$ the total energy and $\Psi(r,R)$ is the wavefunction. The Hamiltonian describes the interactions in our system, using natural units ($\hbar = e = m_e = 1/(4\pi\varepsilon_0) = 1$) it can be written as follows:

$$\hat{H} = -\frac{1}{2} \sum_i \nabla_i^2 - \frac{1}{2} \sum_I \frac{1}{M_I} \nabla_I^2 + \sum_{I,J} \frac{Z_I}{|r_I - r_J|} + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|r_I - r_J|} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|r_i - r_j|}$$

(2.2)

with the lowercase subscripts indicating electrons and the uppercase subscripts indicating nuclei. The position of each electron is given by $r_i$, the mass, position and charge of each nucleus are given by $M_I, R_I$ and $Z_I$ respectively.

It would then appear, given this information, that we have all we need to solve the Schrödinger equation and by doing so we will understand exactly how our system behaves. However it is by no means that straightforward. Although the Schrödinger equation does have an exact solution for the classic example of the isolated hydrogen atom this is one of a rare few cases, often limited to the one-dimensional[45].

In fact as our wavefunction $\Psi$ has $3N_{ion} + 3N_e$ degrees of freedom for a system with $N_{ion}$ nuclei and $N_e$ electrons there is no general way of arriving at an exact solution. This, in a nutshell, is the many body problem. The best we can hope for here is finding a good approximation to the solution of the Schrödinger equation and much of 20th century physics has been dedicated to exactly this. One such approach is the Born-Oppenheimer approximation[46].

## 2.2 The Born-Oppenheimer Approximation

The Born-Oppenheimer(BO) approximation simplifies the many body problem by separating the nuclear and electronic degrees of freedom. This is done by assuming that the electrons are in a well defined ground state when considering the timescales of nuclear motion, an adiabatic approximation that is justified by the fact the nucleus typically has a mass of the order of $10^4$ times greater than the mass of an electron. This is essentially a quasi-classical approximation in which the nuclei are treated as fixed point charges which form a fixed external potential in which the electrons move.

This allows us to consider the nuclear kinetic term separately to the rest of the
2.3. The Traditional Approach to DFT

Although the BO approximation simplifies the challenge of solving the many body problem we are still left with electron-electron interactions which themselves have a many body nature. An exact solution to even this simplified problem therefore still lays out of reach.

Nonetheless many practical approaches have been developed to solving the many electron Schrödinger equation given to us by the BO approximation. Although these approaches are too numerous to list here we shall look in more detail at the method chosen to perform the calculations in this thesis, density functional theory (DFT). The key idea behind DFT is that the many electron Schrödinger equation can be exactly reformulated so that the electronic density is our fundamental variable rather than the many electron wavefunction. This approach formally begins with two ingeniously simple theorems proposed by Hohenberg and Kohn[43].

2.3.1 Hohenberg-Kohn Theorems

The Hohenberg-Kohn (HK) theorems are as follows:

**Theorem I:** The external potential $V_{\text{ext}}(r)$ for any system of interacting particles is uniquely determined, up to an additive constant, by the ground state charge density.

**Theorem II:** It is possible to define a universal energy functional $E[\rho(r)]$ of the electron density which is minimised and equal to the ground-state energy for the exact ground state density.

The corresponding proofs can be found in any electronic structure course or textbook if the reader is interested. The energy functional from theorem II has the form:

$$E[\rho(r)] = \int V_{\text{ext}}(r) + F[\rho(r)]d^3r,$$

for some unknown functional $F[\rho(r)]$. A great deal of effort has gone into finding
the exact form of the functional $F$ as this would reduce all ground state problems exactly to minimisations over the charge density, thus giving exact solutions to the many electron Schrödinger equation, which would most likely be a Nobel prize worthy achievement.

However the exact form of $F[\rho(r)]$ is, at the time of writing still unknown. There are a range of difficulties associated with finding the exact form including, but not limited to, kinetic energy contributions. This is why most modern DFT methodology relies on a remapping of the many body problem to that of a system of non-interacting particles, an approach proposed by Kohn and Sham[49].

### 2.3.2 The Kohn-Sham ansatz

The Kohn-Sham (KS) ansatz states that we can construct an auxiliary system of non-interacting particles, in a modified effective potential, which has the same density as our fully interacting system of interest. The density of this system is expressed in terms of single particle orbitals, $\psi_n(r)$ such that:

$$\rho(r) = 2 \sum_{n}^{\text{occ}} |\psi_n(r)|^2$$  \hspace{1cm} (2.4)

for $N$ particles when summing over the occupied states. This allows us to reformulate the unknown $F[\rho(r)]$ from the HK theorems into something more manageable:

$$F[\rho(r)] = E_k[\rho(r)] + E_{\text{Har}}[\rho(r)] + E_{XC}[\rho(r)]$$  \hspace{1cm} (2.5)

where $E_k$ is the kinetic energy of the non-interacting KS system:

$$E_{k}^{KS} = \frac{1}{2} \sum_{n}^{\text{occ}} \int \psi_n(r)^* \nabla^2 \psi_n(r) d^3r$$  \hspace{1cm} (2.6)

and $E_{\text{Har}}$ is the Hartree energy, which is equivalent to the classical Coulomb interaction:

$$E_{\text{Har}} = \frac{1}{2} \int \frac{\rho(r)\rho(r')}{|r-r'|} d^3r d^3r'.$$  \hspace{1cm} (2.7)
2.3. The Traditional Approach to DFT

The $E_{XC}$ term represents the exchange correlation (XC) energy which is a term defined to take into account the energy difference between a KS system and the equivalent fully interacting system of electrons. Hence the XC functional is defined to be:

$$E_{XC}[\rho(r)] = E_{kin}[\rho(r)] - E_{KS}[\rho(r)] + E_{ee}[\rho(r)] - E_{Har}[\rho(r)],$$  \hspace{1cm} (2.8)

where $E_{ee}$ and $E_{kin}$ simply represent the potential and kinetic energy functional of the non-KS or interacting system, the very functionals that are not exactly known. Hence there is a need to approximate the XC functional and we shall return to this shortly. However, it is first worth clarifying the usefulness of the KS ansatz as it may seem that the full Schrödinger equation has just been rephrased, given we have made no further approximations than BO, without getting any nearer to a solution. Yet it is worth noting that for the majority of physical systems $E_{KS}^{\text{kin}}$ and $E_{Har}$, which can be calculated exactly, make up nearly the entire total energy and therefore the Kohn-Sham equation is quite a useful equation.

The KS equation can be arrived at using defined equations 2.5 and 2.3 to give,

$$\left(\frac{1}{2}\nabla^2 + V_{KS}(r)\right)\psi_n(r) = \varepsilon_n\psi_n(r),$$ \hspace{1cm} (2.9)

noting that $\psi_n(r)$ now represents the single particle KS wavefunctions and $V_{KS}$ is the effective potential experienced by the non-interacting electrons, defined in the equation below in terms of the density, external ($V_{\text{ext}}$) and exchange correlation ($V_{XC}$) potentials.

$$V_{KS}(r) = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} + V_{XC}(\mathbf{r}) + V_{\text{ext}}(\mathbf{r})$$  \hspace{1cm} (2.10)

The reader may have noticed that there is some inter-dependency here, we need the density to obtain $\psi$ but the density is itself defined in terms of $\psi$. In order to escape the circular nature of this problem one may use some initial estimate for the density to solve for $V_{KS}$. 


2.3.3 Achieving Self-Consistency

This initial estimate of the density, denoted here as \( \rho^{in} \), can be used to solve the KS equation and the determined orbitals are then used to define a new output density, \( \rho^{out} \), as shown in the rewritten forms of the KS equations shown below:

\[
\frac{1}{2} \nabla^2 + V_{KS}[\rho^{in}(\mathbf{r})] \psi_n(\mathbf{r}) = \varepsilon_n \psi_n \tag{2.11}
\]

\[
\rho^{out}(\mathbf{r}) = 2 \sum_{n}^{occ} |\psi_n(\mathbf{r})|^2 \tag{2.12}
\]

The output density is then used to solve the KS equations again and we begin an iterative process to find the ground state electron density (\( \rho^{sc} \)). At which point the output density is equal to the input density (i.e. \( \rho^{out} = \rho^{in} = \rho^{sc} \)), the KS equation is solved exactly and this is what is meant by “achieving self consistency”.

In reality however this exact solution is generally not reached, the iterative procedure essentially defines a sequence \( \{\rho_1^{in}, \rho_2^{in}, ..., \rho_n^{in}\} \), where in the simplest possible case \( \rho^{out}_n = \rho^{in}_{n+1} \), and the process is terminated once \( \rho_n^{in} \approx \rho^{sc} \) within some specified tolerance. There are various approaches to achieving this\([50, 51, 52, 53]\), referred to in the literature as self consistent field (SCF)\(^1\) processes\([49, 43]\), we will return to these concepts in Chapter 3.

2.4 Exchange-Correlation

It can be seen above then that the KS ansatz essentially maps the unknowns of the universal function \( F[\rho(\mathbf{r})] \) to the XC term. This is a worthwhile process as the ansatz gives a very efficient way of finding the state but at this point it is likely worth clarifying the XC term and so as a reference definitions of the exchange and correlation are provided.

**Exchange**: The exchange energy arises from the concept of exchange symmetry in quantum mechanics, no physical observable should change when exchanging two identical particles but the wavefunction will change sign if we are considering

\(^1\)The field referred to here is the three dimensional real scalar field that is the electron density
Correlation: The correlation, as coined by Löwdin[54], is defined as the difference between the energy determined by the approximations in the Hartree-Fock[55, 56] method and the energy of the exact solution to the non-relativistic solution Schrödinger equation. The correlation term arises as a result of the collective behaviour of electrons screening and reducing Coloumbic interactions.

The reason the KS ansatz utilises this XC term, is that it typically is a very small contribution to the total energy for a given system. Therefore we do not need a particularly accurate consideration of the XC term in order to arrive at useful results. Nonetheless the XC term is an important one, needed to make accurate bond predictions for example[57] and there are numerous ways of approximating the term, generally speaking the more complex the approximation the more accurate and universal it is. Each approximation is often referred to as a rung on Jacob’s ladder[58], which in biblical scripture ascends to heaven but in our case extends to a universal, accurate non-empirical approximation of the XC term. The simplest approximation and one of the most frequently used being the local density approximation (LDA)[49, 59, 60] with the next simplest being generalised gradient approximation (GGA)[61].

The LDA approach is a simple one, we integrate over the system with the XC energy at each point in space equated to the energy of a homogeneous electron gas of the same density. This is shown in equation 2.13 with $\epsilon$ representing the XC energy density at each point in space. This makes it possible to determine the exchange energy analytically while correlation energy is often determined using quantum Monte Carlo methods[62]. This performs well for systems with a slowly varying nearly homogeneous electron density, however not as well for systems where this is not the case such as systems with defects or even molecules.
2.5. Pseudopotentials

\[ E_{\text{XC}}^{\text{LDA}}[\rho(r)] = \int \rho(r)\varepsilon(\rho(r)) d\mathbf{r} \]  

Therefore a natural extension to this is to consider the local density gradients. This is precisely what is done in the GGA approach, as shown in equation 2.14. This approach is generally found to provide better approximations to ground state energies and bond geometries\[63, 64, 65\], although it is not without its shortcomings, such as increased lattice constants and underestimates of bulk modulii when compared to LDA\[66\]. There are further extensions to this approach known as meta-GGAs, which utilise second order derivatives, however these are covered in greater detail elsewhere\[67\].

\[ E_{\text{XC}}^{\text{GGA}}[\rho(r)] = \int \rho(r)\varepsilon(\rho(r), \nabla\rho(r)) d\mathbf{r} \]  

2.5 Pseudopotentials

The reader may note at this point, that although the exchange-correlation terms have been defined we have not touched upon the potential or the orbitals in the Kohn-Sham equation. For completeness, this will be done very briefly in the following two sections.

It is necessary to consider the potential as another key approximation is used here in the form of a “pseudopotential”. The pseudopotential is an effective ionic potential acting only on the valence electrons and acts to replace the strong Coulomb potential of a nucleus (of the form of \(-\frac{1}{r}\)) and the effects of tightly bound core electrons. The core states are explicitly removed for an atomic calculation utilising a pseudopotential, referred to as the “frozen core approximation” as they will remain unchanged. This is justified as the contribution of the core electrons to the total energy, although large, is only very weakly affected by the atom’s chemical environment. This removes the need to consider wavefunctions for the core electrons, generally characterised as some region within the radius \(r_c\), this is useful as the core induces a high amount of curvature in the wavefunction for the valence states. By removing the core the intention is to make our wavefunction as smooth as possible so that in can be accurately described by fewer basis functions, which will greatly
reduce the overall computational cost.

There are many methods for generating pseudopotentials, including accurate *ab initio* techniques which give us “norm-conserving” and “ultrasoft” pseudopotentials. The “norm-conserving” approach\cite{69, 70} requires the pseudised wavefunction, derived from our pseudopotential, to maintain the same amount of charge within the core radius as the true wavefunction for the valence states. These have the benefit of being largely transferable between different chemical environments. The “ultrasoft” method developed by Vanderbilt\cite{71}, relaxes the normalisation constraint replacing the charge density with the core region with an auxiliary function allowing for the use of a larger $r_c$, which in turn makes the process more computationally efficient.

### 2.6 Basis Sets

Using the pseudopotentials above we can start considering numerically solving the Kohn-Sham equation given in equation 2.9. There are multiple options available to achieve this, we can either use a real space method such as finite differences\cite{72, 73} to describe the single particle orbitals and the total electronic density, or we can employ a basis set representation. The latter is the focus of this section.

The choice of basis set we make is crucial in determining the both the accuracy and

---

**Figure 2.1**: A plot of the pseudised wavefunction and all electron wavefunction for the 1s orbital of Si generated using Pseudo-dojo\cite{68}
speed of a calculation, and generally speaking there is usually a trade-off between the two.

2.6.1 Planewaves

A popular choice for a basis set is plane waves (PWs), shown in equation 2.15, these are well suited to performing calculations on crystalline solids as the periodicity of the basis functions can be easily exploited. Their delocalised nature and independence from atoms’ movements are also helpful attributes as they allow for forces to be calculated without using Pulay corrections, and easily be obtained using fast Fourier transforms (FFT)[74, 75]. Furthermore, the basis can be systematically converged, by using an ‘energy cutoff’ which is directly proportional to the largest vector length in $k$-space, represented by $G_{\text{max}}$ in equation 2.15.

$$\psi(r) = \sum_{|G| < G_{\text{max}}} c_{n,G}e^{iGr}$$

The PW basis set is used for the work presented in Chapter 3 of this thesis.

2.6.2 Pseudo-atomic Orbitals

An alternative approach to the plane wave basis set is atomic orbitals, or in the case of the work presented in this document, pseudo-atomic orbitals (PAOS)[76]. In many ways these can be considered to be diametrically opposed to the plane wave basis, as it is a real space method and a more physically intuitive one that can be understood in terms of individual bonds and local chemistry. Although, of course both approaches must reproduce the same electronic structure when taken to convergence.

We generate PAOs by solving the Schrödinger equation for an atom using the pseudopotentials outlined above. This will give us a basis of the form equation 2.16, consisting of a radial function $R_{l\zeta}$ multiplied by a spherical harmonic $Y_{l}^{m}$, where $n$; $l$ and $m$ are the principle, azimuthal and magnetic quantum numbers respectively.

$$\phi_{\alpha}(r) = \sum_{lm\zeta} c_{i\alpha}^{lm\zeta} R_{l\zeta}(r) Y_{l}^{m}(\hat{r})$$

(2.16)
2.7. Practical Limitations of Basis sets in DFT

Here $i$ and $\alpha$ represent the $i$th atom and $\alpha$th support function (to be covered in greater detail in section 2.8.1), $\zeta$ is simply the number of basis functions. Although, unlike planewaves there is no way of systematically improving PAOs, the accuracy can, generally speaking, be improved by increasing the number of basis functions, in equation 2.14 this corresponds to a larger number of zetas ($\zeta$s), alternatively one could use a larger maximum angular momentum. This, in practical terms, means either more basis functions associated with the valence electrons or the addition of a polarisation orbital, which is represented by the addition of the $l + 1$ angular momentum channel to the basis. To clarify, the terminology we use with the CONQUEST code, and for that matter throughout this thesis, is single zeta (SZ) for one radial function per valence electron, double zeta (DZ) for two radial functions, triple zeta for three and so on, with ‘P’ denoting the addition of a polarisation orbital e.g an SZP basis set. In CONQUEST we choose the polarisation orbital to be the next highest orbital momentum channel[76], so a SZ basis set for Si would have 4 basis functions for the 4 valence electrons, whereas SZP would incorporate s,p, and d orbitals giving a total of 9 basis functions.

2.7 Practical Limitations of Basis sets in DFT

It should be apparent from the above that the choice of basis set is a significant one for anyone wishing to perform electronic structure calculations, and should be carefully considered. As touched upon in the above section the choice will come with some trade-off of desirable properties.

The ‘convergence’ behaviour of PAOs and planewaves has already been outlined, with the former proving more difficult to systematically reduced to a desired precision. Furthermore, it was stated that no Pulay terms were needed to calculate forces when using PWs, this is not the case for PAOs. This may raise question; why would a PAO be used at all? However, there is a final practical consideration to take into account and that is the scalability of the basis set.

The scalability is dependent on two things, the first is modern parallel processing. In simple terms, as plane waves are delocalised in real space the unit cell can not easily
be divided into sections to be handled by each processor. The second factor that determines scalability is the suitability of the basis set for linear scaling methods[77]. For both linear scaling and modern parallel processing it is advantageous to use localised basis sets such as PAOs.

2.8 \textbf{Linear scaling DFT and CONQUEST Methodology}

These linear scaling methods are so called as the memory and computational requirements scale linearly with the number of atoms in the system, whereas the traditional approaches to DFT scale with the cube of the number of atoms. These methods allow DFT to be extended to systems of the order of millions of atoms[2] bridging the gap between simulations and experiment. Numerous codes have implemented such methods[78, 79, 80, 81], which are covered in greater detail elsewhere[77], here we outline the CONQUEST methodology as it will be relevant in later chapters.

2.8.1 \textbf{Localisation of the Density Matrix}

From the standard formulation of DFT it is apparent a density matrix is required that has eigenvalues between zero and one. It is also important that the information stored in the density matrix is proportional to \(N\) with \(N\) being the number of occupied orbitals if a linear scaling approach can be formulated. To achieve this we use the principle of ‘nearsightedness’[82, 83] of electrons which results from a loss of quantum phase coherence between distant points in non-metallic systems. In other words properties such as the electronic density can be evaluated with high accuracy by considering the orbitals to be localised to some region in space and vanishing outside of this region. This is equivalent to only considering interactions between points with a certain proximity. This can be done as the density matrix for non-metallic systems decays as the separation between points increases,

\[
\rho(r,r') \rightarrow 0 \quad \text{as} \quad |r-r'| \rightarrow \infty. \tag{2.17}
\]
This decay property is used to ensure the information stored in the matrix scales linearly with system size by imposing a cutoff $R_c$ radius beyond which all components of the density matrix are equal to zero,

$$\rho(r, r') = 0 \text{ if } |r - r'| > R_c. \tag{2.18}$$

This clearly will result in the loss of some accuracy but such a density matrix with the condition of idempotency\(^2\) enforced and subject to a cutoff radius will give an upper bound for the ground state energy of the system. For this cutoff to be applied to a density matrix dependent on two vector positions then $\rho$ is written in a separable form, this simplifies the problem from working with a six dimensional function $\rho(r, r')$. To achieve this we used a basis of localised functions or support functions such that

$$\psi_n(r) = \sum_{i\alpha} c_i^n \phi_{i\alpha}(r), \tag{2.19}$$

so $\rho$ can be written in the form

$$\rho(r, r') = \sum_{ij, \alpha\beta} \phi_{i\alpha}(r) K_{i\alpha, j\beta} \phi_{j\beta}^*(r'). \tag{2.20}$$

Where $K_{i\alpha, j\beta}$ is a symmetric density matrix in the basis of $\phi_{i\alpha}$. The heart of this CONQUEST methodology\(^8\) is thus that the support functions, while in one sense playing the role of basis functions, in another they are flexible functions, which should themselves be represented in terms of basis functions. This is done either using B-splines (blip functions)\(^8\) or pseudo-atomic orbitals (PAOs)\(^7\), as defined in section 2.6.2., throughout this work the latter approach was preferred.

Using these basis functions the total energy of a system can then be expressed in terms of $\rho$, and so the energy will be minimised with respect to the matrix $K$ and the basis $\phi$. The important detail to note is that the spatial localisation imposed here using a cutoff radius on the elements of the matrix will result in the matrices involved

---
\(^2\)A matrix $M'$ is classified as idempotent if it satisfies the condition $M'^2 = M'$, in practice we only require the density matrix to be near idempotent such that the eigenvalues of the density matrix lie within the range $[0,1]$ and $\rho^2 \approx \rho$. A more detailed explanation of the reasoning behind this is provided in section 2.8.3.
being sparse, with the number of components of the matrix becoming proportional to the number of atoms in the system. Given that this sparsity can be exploited in the computational approach then this lays the framework for linear scaling DFT.

### 2.8.2 Imposing the eigenvalue range of K

In this framework the groundstate is determined by varying the support functions and the matrix $K$. It must be ensured that the values of the coefficients of $K$ are confined to the space for which the eigenvalues of $\rho$ are between zero and one. This can be done by expressing $\rho$ in a form which automatically satisfies these conditions, to do so we use the McWeeny purification[85] as part of a modified Li-Nunes-Vanderbilt(LNV)[86] scheme. For this $\rho$ is written in terms of an auxiliary function $\sigma$

$$\rho = 3\sigma^2 - 2\sigma^3 \tag{2.21}$$

where $\sigma$ can in turn be written as a separable function in the support function basis such that

$$\sigma(r, r') = \sum_{ij, \alpha \beta} \phi_{i\alpha}(r)L_{i\alpha\beta}\phi_{j\beta}(r') \tag{2.22}$$

Here $L$ is an auxiliary density matrix. This then gives a matrix relation for $K$,

$$K = 3LSL - 2LSLSL \tag{2.23}$$

where $S$ is the overlap matrix of support functions, assuming that the basis is non-orthogonal,

$$S_{i\alpha j\beta} = \int d\mathbf{r}\phi_{i\alpha}(\mathbf{r})\phi_{j\beta}(\mathbf{r'}). \tag{2.24}$$

This results in an iterative process, the eigenvalues of $K$ and $L$ obey the same relationship as the eigenvalues of $\rho$ and $\sigma$ so $\lambda_K = 3\lambda_L^2 - 2\lambda_K^3$. So if there is some initial matrix $L$ with eigenvalues in the range $[-1/2, 3/2]$ then this will ensure that the eigenvalues of $K$ will be in the range $[0, 1]$, then replacing $L$ with this new $K$ will cause subsequent matrices to have eigenvalues closer to the values of zero and one and converges quadratically to idempotency. So each iteration drives $K$ closer to idempotency, although in practice several iterations are required to achieve full
idempotency so near idempotency within a certain precision is accepted. As long as the initial matrix eigenvalues in the range \([-1/2, 3/2]\) then near idempotency of \(K\) is ensured and this in turn will make the density matrix \(\rho\) nearly idempotent. It must be said though that the initialisation is not a trivial process\[87\]. However whilst this ensures near idempotency of the matrix it is not necessarily the solution for the density matrix that will provide the ground state energy as the cutoff applied to \(\rho\) makes this technique non-variational. To find a good approximation of the groundstate energy some further work needs to be done.

2.8.3 The Search for the Ground State Energy

As shown above there are two sets of parameters with respect to which we may minimise the energy, the coefficients of the basis functions and the elements of the \(K\) matrix. As already outlined an idempotent, or nearly idempotent, density matrix is required but there is also a second constraint, we must have the correct electron number. The correct electron number will ensure the minimisation is variational as the calculations will give an energy that is above the exact DFT ground state. Furthermore if the electron number has changed following the search for the ground state energy then the solution can not be said to be representative of the system considered.

With these constraints in mind CONQUEST implements a two stage technique in order to find the ground state energy. This follows on from work done by both LNV\[86\] and Palser and Manolopoulos(PM)\[87\] which both rely on the McWeeny purification process as the fundamental algorithm.

2.8.3.1 Stage One: Modified PM

The PM work makes up the first stage of the two stage technique and is used to find the initial density matrix. It was found that if the initial density matrix commutes with the Hamiltonian and has eigenvalues within the range \([0, 1]\) then the repeated use of the McWeeny purification will be guaranteed to give the exact ground state density matrix. PM also gives a form for the initial density matrix \(\rho_{\text{init}}\) using the
2.8. Linear scaling DFT and CONQUEST Methodology

spectrum of the initial Hamiltonian,

\[ \rho_{\text{init}} = \frac{1}{2} \zeta (\mu I - HS) + \frac{1}{2} I. \] (2.25)

Here \( I \) is simply the identity matrix, \( \mu \) is the chemical potential and \( \zeta \) is defined as;

\[ \zeta = \min\left( \frac{1}{H_{\text{max}} - \mu}, \frac{1}{H_{\text{min}} - \mu} \right), \] (2.26)

with \( H_{\text{min}} \) and \( H_{\text{max}} \) representing the lower and upper bounds of the eigenvalue spectrum of the Hamiltonian.

At the ground state commutation with the Hamiltonian could be verified by expressing the ground state energy in terms of the density matrix and the Hamiltonian such that

\[ E_0 = Tr(H \rho) \] (2.27)

where \( \rho \) is the ground state density matrix, the trace of which is equal to the number of occupied states. This in turn is given by

\[ \rho = \theta(\mu S - H), \] (2.28)

where \( \theta \) is the Heaviside step function and \( \mu \) is the chemical potential. In the linear scaling regime \( \rho \) can not be straightforwardly expressed in terms of the Hamiltonian eigenstates as \( H \) can not be diagonalised, this means it is far from trivial to evaluate the groundstate density matrix. As previously outlined, in order to ensure linear scaling \( \rho \) is subject to a spatial or numerical cutoff\(^{88, 89}\) and this truncation must be imposed at each iteration of the McWeeny purification, resulting in truncation errors and the loss of monotonic convergence to the groundstate energy\(^3\).

Therefore the use of the first stage alone would not be ideal, as the lack of monotonic convergence makes the decision as to when to stop the algorithm somewhat heuristic and the ground state energy produced using this method will be dependent

\(^3\)In practice the energy does drop initially, although the monotonic convergence is not maintained after a certain number of steps
on the initial density matrix. In order to avoid these problems we use this just to provide an initial density matrix and move on to the second stage after an initial density matrix is found.

2.8.3.2 Stage Two: LNV

The LNV method also uses the purification techniques of McWeeny but in a slightly different manner, it is a variational minimisation of $L$ with respect to $E$ and not an iterative process. The grand potential is minimised with respect to the density matrix where

$$\Omega = \text{Tr}(H-\rho)$$

(2.29)

with $\Omega$ being the grand potential. This is subject to the condition that $\rho$ is nearly idempotent and can be written in terms of the purification as

$$\Omega = \text{Tr}(H-)(3\sigma^2 - 2\sigma^3)$$

(2.30)

using equation (2.21). This provides a variational approach with monotonic convergence to a single minimum as it is in a cubic form. Alternatively the LNV method is used to minimise the energy where

$$E = \text{Tr}H(3\sigma^2 - 2\sigma^3).$$

(2.31)

In the LNV scheme it is proposed that the minimisation is carried out with a constant chemical potential $\mu$, however as has already been stated we need to keep the electron number $N_e$ constant, so in CONQUEST the grand potential is minimised subject to the constraint that the number of occupied states held constant. Although obviously variations with respect to the elements of the density matrix and the support functions $\phi$ will cause the electron number to change. To correct this effect we use the Lagrangian multiplier technique.

If we consider the grand potential in terms of the auxiliary density matrices $K$ and $L$ defined above in the McWeeny method then we can perform a line minimisation
with respect to the density matrix $L$. During this minimisation the search direction is projected so that is is tangential to the local surface of constant $N_e$. This ensures that the minimisation will cause only a small change in $N_e$ and if an appropriate value for the chemical potential $\mu$ has been chosen then this correction can be carried out without losing the reduction in $\Omega$. So expressing the grand potential in terms of the density matrix $K$ defined above and with the correcting term to keep the electron number constant we have

$$\Omega = \text{Tr}[K(H - \mu S)] + \mu N_e$$

$$= \text{Tr}[KH] - \mu (\text{Tr}[KS] - N_e)$$

(2.32)

and the minimum is found using the derivative of the energy $E$ with respect to $L$,

$$\frac{\partial \Omega}{\partial L} = \frac{\partial E}{\partial L} - \mu \frac{\partial N_e}{\partial L}.$$  

(2.33)

The minimisation will continue varying $L$ such that at the $n^{th}$ iteration $L_{n+1} = L_n + \beta$ for a search direction $\beta$ until the change in energy is below a specified threshold and a minimum is found. In order to ensure the correct electron number is maintained as $L$ varies, given $N_e = \text{Tr}[LS]$, a scaling factor is applied at each step such that $L \rightarrow N_eL/\text{Tr}[LS]$. The search directions are usually determined using schemes such as conjugate gradients. This provides a variational method for finding the energy minimum.

This variation with respect to $L$ is used in CONQUEST as the innermost loop of a multiple loop approach to finding the groundstate energy. The above steps form the innermost loop in which the groundstate density matrix is found by varying $L$ whilst the electron density and support functions $\phi$ are held constant. If the inner loop alone was used the final result would be non-self-consistent as it would be found that the final $L$ matrix $L_n$ would not be mutually consistent with the Hamiltonian $H$, as $L$ has varied but $H$ (which itself is dependent on $L$) has not been updated. This therefore is a non-self consistent ab-initio tight binding approach, in order to have a full ab-initio approach self consistency must be imposed. This is achieved by using
a middle loop in which the electron density residual is systematically reduced, that is to say the difference between the input and output density matrices in the steps above is reduced. The outer loop is formed by recalculating $H$ when carrying out the minimisation of the grand potential with respect to the support functions, this loop terminates when the change in the total energy is below a specified threshold. This combination of all of these loops gives self consistent energy minimisation in CONQUEST. In order to provide some clarification to these loops a flowchart is provided below which may be helpful to refer to when reading through the above.
2.8.4 Real Space Grid

It is worth mentioning that central to the implementation of the method described in the previous two sections is the use of a regular cubic real space grid. The grid spans the entire system under consideration and the support functions are mapped onto this grid in order to perform some integrals. That is to say that the support functions are represented by their values at grid points, with non-zero values only at points within the localisation region. This means that some potentials, i.e. the exchange-
correlation ($V_{XC}$), the Hartree potential ($V_{Ha}$), and the local pseudopotential ($V_{PS}$) are calculated using sums over localised grid points. Fast Fourier transforms are used to calculate $V_{Ha}$ in reciprocal space and then perform a transformation back to real space.

The action of the fast Fourier transforms are not quite linear scaling, in fact they scale as $N \log(N)$ for an N particle system, the difference for the overall efficiency of the code is negligible but nonetheless the intention is to perform as few calculations using the grid as possible.

### 2.8.5 The CONQUEST Routine

In order to summarise the method outlined above and how they are implemented in CONQUEST below is a step by step description of how a simulation will be carried out.

1. Firstly construct the Hamiltonian using Hamiltonian operator and support functions such that $H_{i\alpha j\beta} = \langle \phi_{i\alpha} | \hat{H} | \phi_{j\beta} \rangle$.

2. Similarly construct the overlap matrix $S_{i\alpha j\beta}$ using the sum over the localised region, with PAOs the S matrix is found analytically and with a very fine mesh.

3. Then build an initial $K_0$ matrix for the McWeeny purification using the formulae set out in section 4 and perform a fixed number of iterations of the purification until obtaining a nearly idempotent $K_{fin}$.

4. Use the final matrix $K_{fin}$ from the McWeeny process and set it to the initial matrix ($L_0^{LNV}$) for the LNV minimisation, making corrections to $N_e = 2Tr[L_0^{LNV} S]$ to give the correct number of occupied states or electron number.

5. Then carry out the minimisation with respect to $L$ and $\phi$ and reduce the residuals as part of the LNV process as outlined above.
6. Once the LNV process is complete we have the correct density matrix and will still have the correct number of occupied states. These can then be used to determine the total energy of the system.
Chapter 3

Developments in CONQUEST

At the time of starting this PhD and during the process of writing this thesis the CONQUEST code has been in development. Among the advantages of working with a code in development is that it is, by nature, constantly evolving and this allows a certain freedom to make additions and to suggest alterations. This chapter outlines the developments and updates that have been made to the CONQUEST code while modelling the systems in this thesis. These updates will feature in an imminent open source general release of the CONQUEST code, it is intended that this chapter may provide some insight to an interested user.

3.1 Filter Diagonalisation and on-site support functions

3.1.1 Background

As outlined in Chapter 2 the CONQUEST code incorporates a linear scaling (O(N)) approach, where computational cost scales proportionally to the number of atoms in the system \( N \), to DFT alongside a more conventional approach where the computational costs scales cubically with \( N \), referred to here as ‘exact diagonalisation’. Both of these approaches utilise the above mentioned PAOs (see section 2.6.2). In CONQUEST the radial functions which constitute our PAOs are described by numerical values on regular grids. In contrast to exponential basis functions such as Gaussian functions, PAOs are exactly zero beyond a given cutoff. Although it is not possible
to improve the accuracy of PAOs systematically, the accuracy of calculations with PAOs is improved, in all tested cases at least, by increasing the number of radial functions for each spherical harmonic function[90].

That is to say that the PAOs used to perform more accurate calculations will have several radial functions for each spherical harmonic function and are referred to as ‘multiple-ζ’ PAOs, while the PAOs in which only one radial function is used for each spherical harmonic are known as ‘single-ζ’ (SZ) PAOs. A greater number of radial functions in turn means that a greater number of support functions (SFs), in what we call the ‘primitive’ case the number of PAOs and SFs are equal. However the computational cost scales with the cube of the number of PAOs for both the diagonalisation and O(N) methods, and so there is an accuracy-time trade-off implicit here.

There is however a way around this, by using linear combinations of our PAOs we can reduce the number of SFs required, this is what we call a ‘contracted basis’. The concept of contracted basis sets has been applied to Gaussian basis sets for decades[91, 92] however, it is only recently that this has been extended to numerical orbitals as the number of codes using such a basis has grown. Applying this contraction approach to numerical orbitals was first proposed by Ozaki and Kino[93, 94], and has gone on to be recently implemented with success in the CONQUEST code in the form of multi-site support functions (MSSFs)[3].

These MSSFs use linear combinations of the PAOs on individual atoms and the neighbouring atoms within a finite specified region. These essentially correspond to local molecular orbitals, and are therefore free from the limitation from the atomic orbital symmetry and allow us to reduce to the basis size to the minimal number of SFs while maintaining a high level of accuracy[3]. This contracted basis has improved the efficiency of calculations performed in CONQUEST when combined with exact diagonalisation.

However at the time of writing MSSFs have not been successfully combined with the O(N) approach, as it has not been possible to invert the S matrix as is required, an iterative process which requires the matrix to be sparse[84]. This maybe because
of the spatial extent of the MSSFs, and in order for O(N) to function correctly it has been suggested that the SFs should cover as small a region spatially as possible. Hence as an alternative approach I have investigated the possibility of using linear combinations of the PAOs on individual atoms without incorporating the neighbouring atoms. This method is known as onsite SFs (OSSFs), as they apply the same contraction principles of MSSFs on a single atomic site however unlike the MSSFs they are not contracted to the minimal number of SFs but instead the number associated with the single-\(\zeta\) plus polarisation basis (SZP). The theory of OSSFs and the results obtained using this methodology are outlined in greater detail below.

### 3.1.2 Theory

To recap from Chapter 2 our SFs are used to write our density matrix in a separable form as given below.

\[
\rho(r, r') = \sum_{ij, \alpha\beta} \phi_{i\alpha}(r) K_{i\alpha j\beta} \phi_{j\beta}^*(r').
\]  

(3.1)

Here the way in which K is calculated will differ depending on whether we are using the exact diagonalisation approach or O(N). In the former, K is calculated as the sum of the outer products of Kohn-Sham eigenvectors, whereas in O(N) K is calculated using the density matrix minimisation method\[85, 86\]. The definition of the conventional SFs however does not differ, given below for \(\mu\) PAOs (\(\chi_{i\mu}\)) for atom \(i\), with the coefficients \(c\) optimised to minimise the total energy of the system\[76\].

\[
\phi_{i\alpha}(r) = \sum_\mu c_{i\alpha, i\mu} \chi_{i\mu}(r)
\]  

(3.2)

For the MSSFs there would be an additional sum over the neighbouring atoms of the target atom \(i\), however with OSSFs we are solely interested in contracting our basis for an individual atom and therefore do not require a sum over neighbouring atoms. In order to perform this contraction we use an approach known as local filter diagonalisation (LFD), an idea proposed by Rayson and Briddon\[95, 96\]based on the filter diagonalisation algorithm of Neuhauser\[97\] and the Fermi operator projection method of Goedecker\[98\]. This is so called as we ‘filter’ out unwanted
high energy components of the basis to consider a smaller eigenspectrum of the Hamiltonian.

The first step of this ‘filtering’ process is to define a localised region of space, or subspace, within a ‘LFD radius’ \( r_{LFD} \) centred on the atom we are considering. The extent of \( r_{LFD} \) will have an influence on the accuracy of the resulting contracted basis set and as such it is usually set to extend over the majority of the simulation cell.

Yet regardless of the value of \( r_{LFD} \) once this subspace is determined, we can define both a subspace Hamiltonian \( H_S \) and a subspace overlap matrix \( S_S \). These subspace matrices are constructed from the original Hamiltonian and overlap matrices defined by our primitive basis. By selecting the coefficients of the original Hamiltonian that correspond to the atoms with our LFD radius we arrive at \( H_S \) which will have dimensions smaller than the original Hamiltonian, which has dimensions equal to the total number of PAOs for all atoms within the simulation cell. These subspace matrices can be diagonalised\(^1\) to give the eigenvalue equation shown below,

\[
H_S C_S = \varepsilon_S C_S S_S,
\]

for some eigenvalues \( \varepsilon_S \) and eigenvectors \( C_S \).

This subspace eigenvalue problem can be used to determine the coefficients of our new contracted basis set, the contraction coefficients. In order to achieve this we project the subspace eigenvectors onto a ‘trial vector’ \( t \), the number of trial vectors is equal to the dimensions of our desired contracted basis, for the OSSF case the dimensions are the same as the SZP basis for the target atom\(^2\). The choice of the coefficients of the trial vector can be completely arbitrary, so long as the vectors are localised on the target atom. In this study the trial vector coefficients are constructed using the PAOs for the given target atom \( i \). The PAOs are assigned a coefficient of 1 or 0 in the trial vector, with the largest PAO for each \( l \) and \( n \) value assigned a value of 1, with every other PAO assigned a value of zero.

\(^1\)In the CONQUEST code SCALAPACK routines are used to perform this task

\(^2\)For example Si would have 9 trial vectors, as the SZP basis consists of 9 PAOs, corresponding to 1s, 3p and 5d orbitals
Having defined our trial vectors we can determine the initial value for the contraction coefficients, represented here as $D$ using equation 3.4.

$$D = C_S f(\epsilon_S) C_S S_S t$$  \hspace{1cm} (3.4)$$

The amplitude of the contributions of our subspace eigenvectors, $C_S$, is determined in Eq. 3.4 by the ‘filtration function’ shown here as $f(\epsilon_S)$. It is at this point that high energy contributions from the eigenvectors are ‘filtered’ out. The filtration function is simply a very high temperature Fermi-Dirac (F-D) function, this is chosen for computational convenience and the temperature has no relation to the electronic temperature of the calculation for which the resulting basis set is used. Furthermore any step-like function could be used to perform the task of filtration as long as it gives a low weighting to the high energy contributions from eigenvectors. Alternatives to the F-D function were tested including the $tanh$ (a short example is shown in Appendix A) and $erf$ (Gauss error) functions, the advantage being that confusion could be avoided regarding the temperature term within the F-D function. However upon testing the respective functions there were found to be no advantages gained in terms of accuracy or computational efficiency and hence the F-D function has remains in use in the most recent version of the code.

The initial value of $D$ now allows us to define the contracted or filtered basis $\phi^{filt}$ as given in equation 3.5.

$$\phi^{filt}_{i\alpha}(r) = \sum_{\nu<\mu} D_{i\alpha,\nu} \chi_{\nu}(r)$$  \hspace{1cm} (3.5)$$

This then allows us to find $K$ in terms of the contracted basis, in order to obtain a new charge density.

$$\rho(r,r') = \sum_{ij,\alpha\beta} \phi^{filt}_{i\alpha}(r) K^{filt}_{i\alpha,j\beta} \phi^{* \ filt}_{j\beta}(r').$$  \hspace{1cm} (3.6)$$

This new charge density is then used to rebuild the Hamiltonian $H$ in the primitive basis and a self-consistent-field (SCF) calculation using these matrices and the
methods outlined in Chapter 2. This calculation will be variational with \( D \). Once the SCF calculation has converged we use the converged Hamiltonian to return to Equation 3.3 and recommence filtration at which point the coefficients (\( D \)) of our contacted basis are updated. This two-step procedure, involving the SCF calculations and the subsequent update of the filtration coefficients \( D \), is repeated until both the energy and density converge.

It should be noted here that it would be possible to optimise the filtration coefficients \( D \), following the filtration process itself, using a numerical optimisation method such as the conjugate gradient method and this would be a logical extension to this work. However time restrictions have meant it has not been possible to implement this at the time of writing.

Therefore it is the process outlined above, without numerical optimisation, through which a CONQUEST calculation is performed using the contracted basis set, or OSSFs. The results included in this chapter will illustrate that this method, and the changes implemented from the standard MSSF procedure, namely the coefficients of the trial vector and the use of a larger number of SFs in a localised basis, are not only viable but provide an accurate and efficient technique.

3.2 Computational Details and Results

3.2.1 Diagonalisation

3.2.1.1 \( \text{H}_2\text{O} \)

In order to assess the performance of the onsite support functions we first considered an isolated \( \text{H}_2\text{O} \) molecule in order to test how accurately we can reproduce key properties obtained using primitive support functions, such as total energies and forces.

For this assessment (and all others in this section) we generated three basis sets using the CONQUEST code, they are listed here in order of the number of SFs associated with each; SZP with one radial function for the valence electrons and one for polarisation, DZDP with two radial functions for valence and for polarisation and lastly TZTP with three radial functions for all angular momenta. In general
terms the more radial functions in the basis set has then the more accurate we would expect the basis set to be\cite{99}, and this is reflected for the primitive case. When using OSSFs we contract the DZDP and TZTP basis set to the same number of SFs as the SZP basis set\footnote{The constraint of maintaining the point group symmetry results in a limitation in reducing the number of support functions\cite{76}, it is not possible to reduce to the minimal numbers of supports using OSSFs, this is not the case for MSSFs}. This raises the question of whether the higher accuracy of the TZTP and DZDP basis sets is maintained when contracting the number of SFs.

The simulations of H$_2$O were run with the above mentioned basis sets using both the primitive and OSSF versions of DZDP and TZTP. A $10 \times 10 \times 10 \text{Å}^3$ simulation box was used allowing us to discount any influence of periodic images of the molecule. An integration grid cutoff of 150 Ha and a LFD range of 5 bohr. The total energies, forces and bond lengths resulting from these simulations are shown in Table 3.1 The structure was optimised using the TZTP basis set, and the energies were converged to a tolerance of $1 \times 10^{-6}$ for each basis set.

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>Total Energies (Ha)</th>
<th>Maximum forces(Ha/a$_0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZP</td>
<td>-17.63832</td>
<td>0.06300</td>
</tr>
<tr>
<td>DZDP (OSSF)</td>
<td>-17.68854</td>
<td>0.03018</td>
</tr>
<tr>
<td>DZDP</td>
<td>-17.69304</td>
<td>0.02837</td>
</tr>
<tr>
<td>TZTP (OSSF)</td>
<td>-17.70464</td>
<td>0.00021</td>
</tr>
<tr>
<td>TZTP</td>
<td>-17.71053</td>
<td>0.00003</td>
</tr>
</tbody>
</table>

\textbf{Table 3.1:} The total DFT energy and total forces for H$_2$O using both the primitive and OSSF basis sets.

As is illustrated the OSSF basis calculations provide close approximations to the energies produced when using their primitive counterparts. Moreover the OSSF basis sets are closer in energy to their primitive counterparts than the primitive SZP basis is, suggesting that the OSSF method provides a more accurate basis of the same dimensions. A good agreement can also be observed between the forces obtained with each of the contracted OSSF basis sets and their primitive counterparts, with differences of the order of $1 \times 10^{-3}$Ha/a$_0$. 
3.2. Computational Details and Results

3.2.1.2 Bulk Silicon

Further comparison of the basis sets can be performed with the consideration for energy-volume curves, this is a standard approach and it is repeated here for the case of crystalline 8 atom silicon. Again the SZP, DZDP and TZTP basis sets are used with their OSSF counterparts. The MSSF cases are also provided here for comparison. Again a $6 \times 6 \times 6$ Monkhorst-Pack $k$-mesh and an integration grid cutoff of 150 Ha with a LFD range of 10 bohr.

The resulting energy-volume curves are shown in Figure 3.1 and again good agreement is seen between the OSSF basis sets and their primitive counterparts, although the overlap of the OSSF curves suggests there is an upper limit to their accuracy. The MSSF curves do show closer agreement to the results from the primitive DZDP basis set which gives a glimpse of the accuracy of the MSSF method. The fact that OSSFs are outperformed here is not a concern. Firstly, the OSSFs provide a much higher level of accuracy than the primitive SZP basis both in terms of the total energy and the lattice constant. This is sufficient as we are merely attempting to establish accuracy before extending our test to the $O(N)$ regime, which is currently limited to basis sets no larger than the primitive SZP basis and where currently MSSF methodology is not tenable. Secondly, the results could be improved by optimising the basis, a numerical optimisation method such as the conjugate gradient method could be used to adjust the coefficients of the OSSFs after they are constructed, as has been done successfully in other contraction methods[3]. By doing so the agreement between results obtained using the primitive and OSSF basis sets would be improved. However, due to time constraints no optimisation has been performed here and the limitations in accuracy may simply be a consequence of the choice of trial vectors.

A more detailed comparison can be made using the data in Table 3.2 showing the bulk modulii and lattice constants for the variety of basis sets. It can be seen again that both OSSF methods and MSSF methods reproduce the primitive basis set results with a high level of accuracy, with the lattice constant varying by no more than 1% for either contraction method. Furthermore, the bulk modulus varies
Figure 3.1: Volume energy curves for bulk Si using the primitive, OSSF and MSSF basis sets, fitted using the Birch-Murnaghan equation[100, 101]. Black crosses are used to indicate sample datapoints for the primitive basis simulations by approximately 2 – 4% from the primitive solution using the OSSF and MSSF approaches.

Table 3.2: Bulk modulus ($B_0$) and lattice constant for crystalline Si using both the primitive support functions and OSSFs as basis sets.

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>$B_0$ (GPa)</th>
<th>Lattice constant(Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZP</td>
<td>78.53682</td>
<td>5.56281</td>
</tr>
<tr>
<td>DZDP (OSSF)</td>
<td>85.45173</td>
<td>5.49343</td>
</tr>
<tr>
<td>DZDP (MSSF)</td>
<td>87.61880</td>
<td>5.49285</td>
</tr>
<tr>
<td>DZDP</td>
<td>88.14753</td>
<td>5.48752</td>
</tr>
<tr>
<td>TZTP(OSSF)</td>
<td>88.41297</td>
<td>5.48756</td>
</tr>
<tr>
<td>TZTP(MSSF)</td>
<td>88.95005</td>
<td>5.48910</td>
</tr>
<tr>
<td>TZTP</td>
<td>86.97125</td>
<td>5.46798</td>
</tr>
</tbody>
</table>

3.2.2 Energy-Volume Curve Tests

It would be inadequate to test the quality of the basis produced by the OSSF using solely Si. Therefore further tests have been performed using two more elemental semiconductors in the forms of carbon and germanium, allowing us to test the contracted basis for a range of band gap sizes. In addition, we have considered the perovskite oxides SrTiO$_3$, PbTiO$_3$ in the cubic phase, to consider the effects a different bonding (ionic in this case) may have as well as the influence of a charged
environment. Semi-core states are also present in the primitive basis sets for these simulations but there is little reason to believe that this will have an impact on the final results.

For all the examples considered comparisons are made with the MSSF method. However it is important to note that the MSSF method can be optimised by adjusting the a feature known as the ‘multisite radius’ which determines the number of atoms that contribute PAOs to the contracted basis. Generally speaking the accuracy can be improved by increasing the radius size, this is discussed in greater detail elsewhere[3]. For the tests shown here the multisite radius has been set equal to the largest side length of the simulation cell, as this provides a reasonable compromise between accuracy and efficiency.

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>Ge (B_0) (GPa)</th>
<th>Ge Lattice constant(Å)</th>
<th>C (B_0) (GPa)</th>
<th>C Lattice constant(Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZP</td>
<td>57.49800</td>
<td>5.76489</td>
<td>415.12238</td>
<td>3.60320</td>
</tr>
<tr>
<td>DZDP (OSSF)</td>
<td>68.25127</td>
<td>5.74172</td>
<td>448.40003</td>
<td>3.59238</td>
</tr>
<tr>
<td>DZDP (MSSF)</td>
<td>70.16301</td>
<td>5.68537</td>
<td>426.70344</td>
<td>3.59188</td>
</tr>
<tr>
<td>DZDP</td>
<td>70.59008</td>
<td>5.69032</td>
<td>452.07820</td>
<td>3.56449</td>
</tr>
<tr>
<td>TZTP(OSSF)</td>
<td>69.17876</td>
<td>5.73851</td>
<td>482.38479</td>
<td>3.57593</td>
</tr>
<tr>
<td>TZTP(MSSF)</td>
<td>65.68155</td>
<td>5.70003</td>
<td>497.99316</td>
<td>3.60288</td>
</tr>
<tr>
<td>TZTP</td>
<td>64.99272</td>
<td>5.68278</td>
<td>452.95461</td>
<td>3.56189</td>
</tr>
</tbody>
</table>

Table 3.3: Bulk modulus \(B_0\) and lattice constant for the elemental semiconductors Ge and C using both the primitive support functions and OSSFs as basis sets.

For the elemental semiconductors the calculations use a k-point mesh of \(9 \times 9 \times 9\) with a integration grid cutoff of 300Ha. There is good agreement in the lattice constants for all contracted basis sets, as shown in Table 3.3, with the OSSF contracted basis lattice constant differing by less than 1% for either C or Ge. The MSSF contracted basis provides a similar level of accuracy. The bulk modulus also shows a reasonable level of agreement between the primitive and contracted OSSF basis, in the worst case scenario (TZTP for C) it differs by 6%, and typically overestimates \(B_0\) for TZTP, but for all other cases the values differ by less than 5%. The MSSF and OSSF methods seems to provide similar levels of accuracy with re-
spect to $B_0$ and the lattice constants at least, furthermore there seems to be no clear hierarchy of accuracy. The calculations using basis sets without either contraction method do exhibit convergence for both the bulk modulus and lattice constant as we increase the number of support functions (with the number increasing from SZP to DZDP and in turn TZTP). This is not the case for the MSSF and OSSF results presented in Table 3.3, although it should be stated that the MSSF calculations presented here are merely for comparison and have not been fully optimised to give an accurate result as possible. Similarly the OSSF method is, as of yet, lacking any numerical optimisation in the generation of the basis set, it is fully expected that the accuracy of the OSSF results utilising the contracted TZTP basis sets can be improved upon.

This lack of a hierarchy of accuracy is further illustrated in Figure 3.2, where it can be observed that binding energy curves do not follow a strict pattern, with no contraction method guaranteed to produce a closer estimate to the primitive TZTP basis. It should be noted however that both techniques could be further optimised, as already mentioned a numerical optimisation method could be introduced to the OSSF method and an optimised basis may alter the hierarchy of accuracy.

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>$\text{PbTiO}_3$</th>
<th>$\text{SrTiO}_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B_0$ (GPa)</td>
<td>$V_0$(Å$^3$)</td>
</tr>
<tr>
<td>SZP</td>
<td>189.17862</td>
<td>61.66259</td>
</tr>
<tr>
<td>DZDP (OSSF)</td>
<td>186.81761</td>
<td>60.61224</td>
</tr>
<tr>
<td>DZDP (MSSF)</td>
<td>179.08013</td>
<td>62.83250</td>
</tr>
<tr>
<td>DZDP</td>
<td>188.99455</td>
<td>60.78282</td>
</tr>
<tr>
<td>TZTP(OSSF)</td>
<td>194.89697</td>
<td>61.14628</td>
</tr>
<tr>
<td>TZTP(MSSF)</td>
<td>179.47401</td>
<td>62.82480</td>
</tr>
<tr>
<td>TZTP</td>
<td>190.32613</td>
<td>60.61592</td>
</tr>
</tbody>
</table>

Table 3.4: Bulk modulus ($B_0$) and Equilibrium volume ($V_0$) for the perovskite oxides $\text{SrTiO}_3$, $\text{PbTiO}_3$ using both the primitive support functions and OSSFs as basis sets.

For the perovskite oxides the results were similar, again the calculations used a k-point mesh of $9 \times 9 \times 9$ with a integration grid cutoff of 400Ha. Due to the non-cubic nature of the simulation cells we compare equilibrium volumes rather
3.2. Computational Details and Results

Figure 3.2: Binding energy curves for crystalline Ge and C using the primitive, OSSF and MSSF basis sets, fitted using the Birch-Murnaghan equation.
than lattice constants but there is still good agreement between the primitive and contracted basis sets. The contracted basis sets equilibrium volumes do not vary by more than 4% for either SrTiO$_3$ or PbTiO$_3$ from the primitive TZTP solution, however the MSSF method does overestimate $V_0$ for SrTiO$_3$. The values for $B_0$ show marginally worse agreement, as illustrated in Table 3.4, however still do not vary by more than 6% from the primitive TZTP solution. Notably the OSSF contracted TZTP basis sets overestimate the bulk modulus for both cases. Furthermore the OSSF contracted TZTP basis set is shown to be less accurate in estimates for both $B_0$ and $V_0$ than the DZDP basis set contracted using the same method, as is reflected in the energy-volume curves shown in Figure 3.3. This is likely to be an artefact of the LFD process, were numerical optimisation utilised than one would expect the usual hierarchy of basis sets to be restored, i.e. for the contracted TZTP to provide more accurate estimates for $B_0$ and $V_0$ than the contracted DZDP basis.

The above tests provide a useful benchmark for the accuracy of the basis set produced by the OSSF. However, as previously stated, the desired outcome of the OSSF methodology is to make large basis set linear scaling simulations accessible by using a contracted basis set. The section below includes results for linear scaling simulations.

3.2.3 Order N

3.2.3.1 Bulk Silicon

Again we commence by considering the case of 8 atom crystalline Si. Linear scaling calculations have been performed using the primitive SZ and SZP basis as well as the contracted DZDP and TZTP OSSF basis sets. The SZP basis set is the largest for which O(N) calculations can currently be performed and although this will allow for a point of comparison with the contracted basis sets it is not possible to provide comparison using either the primitive DZDP or TZTP basis.

For the simulations combining diagonalisation and the primitive number of support functions a $6 \times 6 \times 6$ Monkhorst-Pack k-mesh and an integration grid cutoff of 150 Ha with a LFD range of 10 bohr. In addition the $L$ matrix used in linear scaling simulations (see section 2.8.2) is subject to a real space cutoff which, for the O(N)
3.2. Computational Details and Results

Figure 3.3: Binding energy curves for crystalline SrTiO$_3$ and PbTiO$_3$ using the primitive, OSSF and MSSF basis sets, fitted using the Birch-Murnaghan equation.
Table 3.5: The DFT energies for 8 atom Si using diagonalisation with the primitive number of supports functions (NSF) and using O(N) with onsite support functions for DZDP and TZTP

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>Diag. Primitive NSF Energy (Ha)</th>
<th>O(N) NSF = 9 Energy (Ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZP</td>
<td>-33.698124</td>
<td>-33.679536</td>
</tr>
<tr>
<td>DZDP</td>
<td>-33.820683</td>
<td>-33.793137</td>
</tr>
<tr>
<td>TZTP</td>
<td>-33.834045</td>
<td>-33.794028</td>
</tr>
</tbody>
</table>

calculations, has been set at 18 bohr, using a greater cutoff is observed to have little impact on the final results.

The results shown in Table 3.5 primarily illustrate that the OSSF methodology is compatible with the linear scaling techniques implemented in CONQUEST. This is a key result as it will allow for O(N) calculations to be performed using basis sets of greater accuracy than the SZP basis allowing for a more accurate consideration of bond geometries and electronic structure. Furthermore, the results in Table 3.5, although simple, give an indication that the accuracy of the contracted basis sets obtained using OSSS is maintained when performing O(N) calculations. A more accurate groundstate energy is determined using O(N) for both the OSSF contracted TZTP and DZDP than the primitive SZP basis set. The groundstate energy obtained using the full primitive basis sets with the diagonalisation method and the contracted basis sets in O(N) only differ by the order of $1 \times 10^{-2} \text{ Ha}$ or approximately 0.2%. A similar difference is seen between the O(N) and diagonalisation calculations using the non-contracted SZP basis set, indicating that the OSSF method does not introduce significant errors. This shows that, for the case of crystalline Si at least, that a reliable groundstate energy can be determined using the OSSF method in conjunction with O(N) calculations. However, the groundstate energies alone can not determine the quality of the method, more insight is required, some of the further tests performed are given below.

3.2.3.2 Silicon with a vacancy

As the method utilises the PAOs of neighbouring atoms to construct the localised basis for our target atom it seems logical to consider the impact defects may have upon the efficacy of this technique. As a test case we have considered supercells of
Table 3.6: The vacancy formation energies for 64, 216 and 512 atom crystalline bulk Si with a vacancy at a substitutional site using OSSF O(N) and diagonalisation with the primitive NSF for the TZTP basis set.

<table>
<thead>
<tr>
<th>Vacancy Formation Energy (Ha)</th>
<th>63 atom</th>
<th>215 atom</th>
<th>511 atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZP O(N) NSF=9</td>
<td>0.1445</td>
<td>0.1424</td>
<td>0.1406</td>
</tr>
<tr>
<td>DZDP O(N) NSF=9</td>
<td>0.1259</td>
<td>0.1238</td>
<td>0.1223</td>
</tr>
<tr>
<td>TZTP O(N) NSF=9</td>
<td>0.1281</td>
<td>0.1262</td>
<td>0.1245</td>
</tr>
<tr>
<td>TZTP Diag. NSF=27</td>
<td>0.1242</td>
<td>0.1231</td>
<td>0.1218</td>
</tr>
</tbody>
</table>

crystalline Si ranging in size from 64 to 512 atoms each with a vacancy introduced at a substitutional site, such that the defect concentrations range from approximately $1.8 \times 10^{-22} \text{cm}^{-3}$ to $1.5 \times 10^{-21} \text{cm}^{-3}$.

The simulations utilised an integration grid cutoff of 150 Ha with a LFD range of 10 Bohr and the $L$ matrix was subject to a cutoff of 18 Bohr. For the simulations using the primitive TZTP basis set with the diagonalisation method a $2 \times 2 \times 2$ Monkhorst-Pack k-mesh was used. The atomic positions have been optimised using a quasi-Newton ionic relaxation algorithm which minimises both energy and interatomic forces. The convergence criteria for the atomic forces were set to 0.01 eV/Å.

The corresponding results shown in Table 3.6 illustrate that the OSSF method produces a good approximation, for the range of defect concentrations considered, to the vacancy formation energy given by the TZTP primitive basis set, which in turn shows good agreement with the literature[102, 103]. This is a useful result when studying dopants and defects. The contracted DZDP and TZTP basis sets provide more accurate approximations to the vacancy formation energy than the primitive SZP basis set for all structure sizes considered here, differing from the primitive TZTP solution by no more than 1.1%. However it should be noted that these levels of accuracy come with an associated increase in the total computational time as shown in Table 3.7. This time increase is a cost attributed to the local filter diagonalisation process and shows no correlation to the system size. For the largest simulations shown here an increase in computational time of approximately 2.56 seconds
Table 3.7: The computational for 64, 216 and 512 atom crystalline bulk Si with a vacancy at a substitutional site using OSSF O(N) averaged over 10 calculations.

<table>
<thead>
<tr>
<th>Basis/Method</th>
<th>63 atom</th>
<th>215 atom</th>
<th>511 atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZP O(N) NSF=9</td>
<td>3388.229</td>
<td>9704.381</td>
<td>25480.455</td>
</tr>
<tr>
<td>DZDP O(N) NSF=9</td>
<td>4762.127</td>
<td>11126.851</td>
<td>26792.311</td>
</tr>
<tr>
<td>TZTP O(N) NSF=9</td>
<td>4888.681</td>
<td>11211.162</td>
<td>26877.103</td>
</tr>
</tbody>
</table>

per atom is shown. For larger simulations the contraction of the basis will result in an improvement in efficiency as the cost of local filter diagonalisation becomes less significant proportion of the total calculation time. Hence, using a contracted basis will result in a net reduction of total computational time, evidence of this is provided at a later stage of the chapter.

3.2.3.3 P dimers on Si(001)

The improvement in accuracy of linear scaling simulations that comes as a result of the ability to use larger basis sets is undoubtedly beneficial but there is also another significant benefit associated with contracted basis sets, an improvement in efficiency. The linear scaling examples considered so far could all be simply modelled with conventional KS DFT approaches and in these cases linear scaling methods are not the optimum choice for an efficient calculation. However at scales where linear scaling calculations are required, an OSSF basis will have an impact on both the efficiency and accuracy. In order to illustrate this we consider the example of a P dimer adsorbed to a Si surface with the dangling bonds at the base terminated using H, using a simulation cell consisting of 3072 atoms (illustrated in Figure 3.4 with all H atoms removed for clarity), which is much closer to experimental scales than our previous examples. Similar models have previously been used to predict energies of different surface reconstructions and provide theoretical reproductions of scanning tunnelling microscopy data[104, 105].

The simulations were performed using a primitive SZP basis set and a contracted OSSF DZDP basis. An integration grid cutoff of 50Ha was used for both with an L matrix cutoff of 18 Bohr, for which the groundstate energy of crystalline Si converges. For the contracted basis an LFD range of 12 Bohr was used for both the
3.2. Computational Details and Results

Figure 3.4: An atomic diagram of a P-dimer, shown in red and circled, adsorbed onto the Si(001) surface

P and Si atoms and a smaller range of 8 Bohr was used for the H atoms. These values have been shown to give accurate bond lengths and converged energies for H₂O molecules as well as Bulk P and Si.

The total energies and computational run times, included here in Table 3.8, were averaged over five separate simulations. As is shown the calculations using the contracted basis set are notably more efficient than those using a primitive basis set, with the former taking approximately 48% of the total computational time of the former, with the calculations run using the same number of computational nodes.

This increase in speed largely comes as a result of a reduction in the number of iterations in the self consistency cycle (as shown in Figure 3.5), with the contracted basis taking fewer iterations than the primitive basis to produce a self consistent charge density. Another factor is the time required for matrix multiplication, as it is proportional to the cube of the number of support functions it will also be reduced. However, some other operations are still performed in the primitive basis and given further development there contracted basis sets could be made more efficient. Furthermore, the total binding energy produced using the contracted basis set only differs from the value produced by the primitive SZP basis by 0.3% indicating that it is unlikely the contracted basis has resulted in a large error in the energy.
3.3 Conclusions

The introduction of onsite support functions allows for the reduction of the number of support functions while retaining a high level of accuracy, producing forces and energies comparable to both the primitive basis sets and the multi-site support functions. This reduction in the number of support functions allows for more efficient calculations than the primitive basis sets, both for simulations utilising full...
Returning to Self consistency

3.4.1 Theory

In addition to the work on basis contraction we also considered the approach to self consistency in the CONQUEST code. When investigating large scale simulations (examples of which are shown in Chapter 6), it was discovered that the convergence of the SCF residual and the total DFT energy were not clearly related, while the values of residual indicated convergence this did not always coincide with the system being at an energy minimum, as would be expected. The result of which was that the use of a uniform SCF residual across simulations of different scales would not result in a uniform precision of total energy. In order to ensure consistent levels of convergence, in terms of the energy, between different large scale simulations we considered the formulation of the residual within the CONQUEST code, in order that a fixed SCF residual may be used which is independent of the system size being simulated.

4This would certainly be necessary for those wishing to close the gap between simulated structures and experimental scales or for those wishing to use DFT for the study of biological systems such as proteins.
3.4. Returning to Self consistency

Before considering the results of these modified residuals it is worth returning to the SCF approaches in Chapter 2 so that some context can be provided for the original formulation of the residual in the CONQUEST code. As has been previously mentioned, SCF techniques exist in a variety of forms, some of the more popular methods, at least amongst the commonly used plane wave electronic structure implementations[1] incorporate a density mixing approach.

A density mixing approach uses an iterative method to reach the self consistent charge density $\rho^{sc}$ (see Chapter 2), but combines the input and output density at the current iteration, or in some cases the entire history of densities, in such a way that the subsequent density is closer to convergence. It should be noted that self-consistency can be equivalently treated by defining a converging sequence of potentials, wavefunctions or even point charges in tight-binding approaches[106]. However, in the CONQUEST code we have chosen to utilise a density based method. The combination of input and output densities utilised can be expressed by defining some function $f$ such that,

$$\rho_{n+1}^{in} = f(\rho_i^{in}, \rho_i^{out}) \forall i \in [1,n].$$  \hfill (3.7)

Ideally we would arrive at a fixed point such that,

$$f[\rho^{sc}] = \rho^{sc}$$  \hfill (3.8)

however as previously mentioned this in reality is not always achievable. Instead we define a ‘residual’ (R) which we attempt to minimise so that,

$$R[\rho^{sc}] = f[\rho^{sc}] - \rho^{sc} \approx 0.$$  \hfill (3.9)

This is set to be approximately equal to zero as we attempt to achieve convergence within some tolerance or threshold, generally set to be of the order of what is considered an acceptable energy precision. There are a variety of numerical methods to solve the non-linear equation 3.9, and hence minimise the residual, for some defi-
section of $f$, by far the simplest of which is the ‘fixed point update’, where the SCF process is treated as a fixed point iteration, and the output density is fed back in as the input density, i.e. $\rho_{n+1}^{in} = \rho_n^{out}$. However this is generally unsuitable and often does not result in convergence for typical DFT systems unless the initial guess for the density is already close to the self consistent solution $\rho^{sc}$. In simple terms, if the condition,

$$\frac{\Delta \rho_n^{out}}{\Delta \rho_n^{in}} < 1$$

(3.10)

is satisfied, for some small change $\Delta \rho$, then convergence will be achieved. If this is not the case however then there will be a strong divergence from the fixed point solution $\rho^{sc}$, an occurrence known as ‘charge sloshing’. A slightly more sophisticated method is the linear mixing approach in which a damping parameter, $\alpha$, is included such that,

$$\rho_{n+1}^{in} = \rho_n^{in} + \alpha(\rho_n^{out} - \rho_n^{in}) = \rho_n^{in} + \alpha R[\rho_n^{in}]$$

(3.11)

so that the residual here defines not only the iterative error but also the direction of steepest descent of our iterative method. This is again heavily dependent on the initial guess, and hence certain systems, such as metallic systems, can take large amounts of time to converge[107]. Generally speaking we wish to reduce the number of SCF iterations required and in turn the amount of time needed to reach convergence. The most successful methods in achieving self consistency efficiently include but are not limited to Anderson mixing[111], Pulay mixing[52], also known as the direct inversion in the iterative subspace (DIIS) method and the Broyden update[112, 113]. The latter two are more popular in electronic structure codes following the developments of Vanderbilt and Louie[114, 53].

A variation upon the Pulay mixing technique is used in the CONQUEST code this is in keeping with the methodology in the majority of DFT codes. For the purposes of this chapter the property of interest for these methods is the point at which the iterative method terminates. For the Broyden methodology, this is the point at which

---

5 At least in the absence of any preconditioning, such as the Kerker[50, 108] method which can help prevent numerical instabilities and charge sloshing effects. These techniques are covered in more detail elsewhere[109, 110]
the residual is minimised with the residual in this case it is expressed in the form of the Frobenius norm[115]. The Frobenius norm $||\cdot||_f$ is defined for some generic matrix $A$ as,

$$||A||_f = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2} \quad (3.12)$$

for the coefficients of the matrix $a_{ij}$, which for example could represent the values of charge density when projected onto a grid. In the Broyden update the Frobenius norm of the Jacobian, $J[\rho]$ is used to define the quantity which the method aims to minimise where the Jacobian $(J)$, by considering approximate updates to $J$ in the form of,

$$J_R[\rho_n^{in}] = \frac{\partial R[\rho_n^{in}]}{\partial \rho_n^{in}}, \quad (3.13)$$

so that we arrive at a near exact Jacobian and in turn a converged charge density.

### 3.4.2 Practical implementation within CONQUEST

The use of the Frobenius norm in conjunction with our residual within CONQUEST means in practical terms we are trying to minimise the following quantity which is denoted as $R_0$,

$$R_0 = \frac{\sqrt{GPV \ast (R_i \cdot R_i)}}{N_e}, \quad (3.14)$$

where the vector $R$ is given by,

$$R = \rho_i^{out} - \rho_i^{in} \quad (3.15)$$

for the densities at a given iteration of the self-consistency cycle. $N_e$ simply denotes the number of electrons in the simulation cell being considered and $GPV$ is the grid point(see section 2.8.4) volume for the simulation cell.

This, of course is a perfectly reasonable choice for the residual in-keeping with common methodology of many codes[90], however it does result in the aforementioned difficulty of converging consistently across different atomic scales, perhaps due to its dependence on the number of electrons meaning the degree convergence will be directly dependent on the size of the simulation cell considered. This is manifest in the differences between the Harris-Foulkes (HF)[116] energy and the
DFT total energy, which increase with simulation cell size.

The HF energy functional gives a non-self-consistent approximation to the energy determined by self-consistently solving the KS equations\(^6\), referred to here as the total DFT energy. It is dependent on the input density \(\rho^{in}\) at a given SCF iteration and is defined as follows,

\[
E_{HF} = \sum_{\text{no.of bands}} \epsilon_j - \frac{1}{2} \int d^3 r V_{XC}[\rho^{in}](r)\rho^{in}(r) - \frac{1}{2} \int d^3 r V_{HF}[\rho^{in}](r)\rho^{in}(r) + E_{XC}[\rho^{in}]
\] (3.16)

where \(\epsilon\) is the band energy and all other definitions are given in Chapter 2. The HF energy varies much more slowly than the DFT energy with changes in the density. The difference between the HF energy and total DFT energy varies with the second order of the error of the approximate electron density, i.e. \(O((\rho^{sc} - \rho^{in})^2)\). Hence the difference between the two energies is a good indicator of whether a calculation has converged. In order to overcome the problematic issue of HF and DFT energy differences in larger simulation cells, a number of alternative methods of calculating the residual have been defined. Testing was performed to determine the respective performances of the alternative method, firstly in order to ensure that the ability to converge has not been affected, it should not be as we are merely adjusting the termination criteria of the SCF cycle. Secondly to check the number of SCF iterations required to achieve convergence. Subsequently two suitable candidates for the residual definition emerged, one of which removes the dependence on the number of electrons in the system while the other retains it. The first of these is defined as follows,

\[
R^{glob}_0 = GPV \sum_{i=1}^{N} |R_i|,
\] (3.17)

where a sum is performed over all \(N\) SCF iterations for the absolute value of \(R\). The superscript ‘glob’ indicates it is a ‘global’ residual independent of the simulation cell considered. By summing over the absolute value of \(R\), we have a simple solution

\(^6\)Although it is worth noting that the functional itself can give a good description of the binding-energies, equilibrium lattice constants, and bulk-modulus for covalently bonded systems like Ge and is even used as the primary energy functional of some tight-binding codes\(^{117}\)
Returning to Self consistency

for a measure of the magnitude which is not dependent on system size, which is not true of the Frobenius norm. The second alternative residual is defined similarly maintaining the sum over the absolute value of $R$,

\[
R_0^{loc} = \frac{GPV \sum_{i=1}^{N} |R_i|}{N_e},
\]  

(3.18)

where ‘$loc$’ is used to indicated it is a ‘local’ residual dependent on the system size of each simulation. Here we maintain the system size independent measure of the magnitude of R, however the ‘local’ residual allows us to consider the charge density difference per electron in the system which can be useful in certain circumstances, for example when considering systems with defects or doped systems with free charge carriers.

3.4.3 Results

As alluded to earlier a useful tool when determining SCF convergence, and one that is used in the tests performed here, is the magnitude of the difference between the Harris-Foulkes (HF) energy and DFT energies. The HF and DFT energy differences are presented below for two Si based systems using the respective definitions of $R_0$ given above.

3.4.3.1 Extended crystalline Si

In order to illustrate the effects of changing electron numbers on the different methods of calculating $R_0$ we consider crystalline Si that has been extended in one dimension to give extended slabs of Si, these structures are illustrated in Figure 3.6.

These structures contain from 8 to 64 Si atoms, or 32 to 256 valence electrons. They were modelled using the PBE functional and a grid cutoff of 150 Ha. The 8 atom structure was modelled using a $2 \times 2 \times 8$ Monkhorst-Pack mesh, with the final number of k-points in this mesh being inversely proportional to the number of the atoms considered, so halving when the number of atoms doubles. Each structure was converged to a self-consistency tolerance of $1 \times 10^{-7}$, such that at convergence the value of one, but not necessarily all, of the residuals defined in Equation 3.14, 3.17 or 3.18 falls below this threshold. The Harris-Foulkes and total DFT energy
3.4. Returning to Self consistency

Figure 3.6: Atomic diagrams of the 128 and 256 electron crystalline Si structures extended in the z-dimension, produced using the Jmol visualisation package.

differences for each structure and residual given above are shown in Table 3.9. As can be observed using the standard $R_0$ term as defined in Equation 3.14 the HF and DFT energy differences scale proportionally with the number of electrons in the system considered, which is equivalent to achieving self-consistency to a lesser precision.

<table>
<thead>
<tr>
<th>Residual Definition</th>
<th>32 electrons</th>
<th>64 electrons</th>
<th>128 electrons</th>
<th>256 electrons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0$</td>
<td>$1.2579 \times 10^{-5}$</td>
<td>$2.5052 \times 10^{-5}$</td>
<td>$5.0103 \times 10^{-5}$</td>
<td>$1.0103 \times 10^{-4}$</td>
</tr>
<tr>
<td>$R^{glob}_0$</td>
<td>$8.1038 \times 10^{-8}$</td>
<td>$1.5234 \times 10^{-9}$</td>
<td>$3.6678 \times 10^{-9}$</td>
<td>$7.0938 \times 10^{-8}$</td>
</tr>
<tr>
<td>$R^{loc}_0$</td>
<td>$5.6687 \times 10^{-8}$</td>
<td>$2.831 \times 10^{-7}$</td>
<td>$9.1751 \times 10^{-7}$</td>
<td>$5.4203 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

Table 3.9: The HF and total DFT energy differences for the elongated crystalline Si cells for the different residual definitions upon convergence.

The data in Table 3.9 shows that the HF and DFT energy differences increase with the system size, for the standard definition of the residual $R_0$, although the increase in energy differences is not quite linear with the total number of electrons. The same is not true for the modified definitions of the residuals with $R^{loc}_0$ and $R^{glob}_0$, producing no obvious relationship between the system size and the magnitude of the HF, DFT energy differences at this level of convergence, this is perhaps due to numerical instabilities when considering orders of this magnitude. Nonetheless, it
is obvious that there is a significant reduction in the magnitude of the energy differences for the global and local residual definitions in comparison to the standard definition, meaning in turn that the self consistent charge density solution to the KS equations has been determined to a higher level of precision. This increased precision does not come for free however, for the global and local definitions of the residual there is an increase, by at least a factor of roughly 3, in the number of iterations required in the SCF cycle to achieve self consistency, as shown in Table 3.10. This indicates that the new definitions of the residual must be utilised with different tolerances.

<table>
<thead>
<tr>
<th>Residual Definition</th>
<th>32 electrons</th>
<th>64 electrons</th>
<th>128 electrons</th>
<th>256 electrons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0$</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>$R_{\text{glob}}$</td>
<td>21</td>
<td>20</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>$R_{\text{loc}}$</td>
<td>18</td>
<td>18</td>
<td>17</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 3.10: The total number of iterations in the SCF cycle required to reach a self consistent charge density for the elongated crystalline Si cells for the different residual definitions upon convergence to a self consistency tolerance of $1 \times 10^{-7}$

This increase in iteration number will in turn have an impact on the total computational time, however there is not a uniform increase in time for the modified definitions of the residual for each system. This is because the modified residuals do not have a system size independent number of SCF iterations at this level of convergence, unlike the original definition $R_0$. In fact there is no clear relation between the number of iterations and the system size for $R_{\text{glob}}$ and $R_{\text{loc}}$, yet for all cases a greater number of iterations than $R_0$ are required to reach the same self consistency tolerance.

### 3.4.3.2 Si nanowires

The above is not a system that would occur in reality, in order to provide both a more realistic and larger scale test of the residuals we consider silicon nanowires, a much studied structure with a wide range of applications including photovoltaics[119].
3.4. Returning to Self consistency

The structures are again extended in a single dimension representing different lengths of nanowires as shown in Figure 3.7. The structures contain from 612 to 2448 Si atoms, or 2448 to 9792 valence electrons, they were modelled using the same exchange-correlation functional and grid cutoff as the extended Si crystalline cells. A DZDP basis set was utilised in conjunction with the multi-site support function method. The HF and DFT energy differences are shown in Table 3.11, the tolerances are set independently for each given residual definition. For $R_0$ and $R_0^{loc}$ we use a self-consistency tolerance of $1 \times 10^{-6}$, whereas for $R_0^{glob}$ a tolerance of $1 \times 10^{-4}$ is utilised as the convergence behaviour will differ depending on the residual definition as depicted in Figure 3.8. The difference in convergence behaviour is further illustrated in Figure 3.9 as we plot the HF and DFT energy differences for the 1224 atom Si nanowire. The fact that the energies do not converge in a uniform manner for the different definitions of the residual verifies that altering the residual definition does more than simply adjust the stopping criteria of a calculation.

The behaviour of the energy differences with respect to system size for the standard definition of the residual $R_0$ remains the same for the nanowires as for the extended crystalline Si structures. The reduction in the magnitude of energy
3.4. Returning to Self consistency

![Graph showing convergence behaviour for different residual definitions against the number of iterations in the SCF cycle for the 612 atom Si nanowire](image)

**Figure 3.8:** The convergence behaviour for the different residual definitions against the number of iterations in the SCF cycle for the 612 atom Si nanowire

<table>
<thead>
<tr>
<th>Residual Definition</th>
<th>HF – DFT (Ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>612 atoms</td>
</tr>
<tr>
<td>$R_0$</td>
<td>0.03574</td>
</tr>
<tr>
<td>$R_{glob}^0$</td>
<td>$4.6478 \times 10^{-6}$</td>
</tr>
<tr>
<td>$R_{loc}^0$</td>
<td>0.00016</td>
</tr>
</tbody>
</table>

**Table 3.11:** The HF and total DFT energy differences for the elongated crystalline Si cells for the different residual definitions upon convergence

The energy differences relating to the global definition of the residual $R_{glob}^0$ again bear no clear relation to the system size, which is the desired outcome. However, the same is not true for the local definition of the residual, $R_{loc}^0$, as a larger tolerance has been used here than for the example of the extended crystalline Si structures. As a result a clear proportionality between the energy differences obtained using this residual and the system size can be observed, as would be expected. Nonetheless the fact that both definitions of the modified residual result...
3.4. Returning to Self consistency

Figure 3.9: The convergence behaviour of the HF and DFT energies for the different residual definitions against the number of iterations in the SCF cycle for the 1224 atom Si nanowire

in lower magnitudes of energy differences at convergence than the default definition of the residual indicates that a more robust and reliable level of convergence have been achieved.

3.4.4 Conclusions

By defining two alternative definitions of the residual that are independent of the Frobenius norm we have been able to develop a method of achieving a precise and system independent level of self consistency. Users of the CONQUEST code will now have the choice of using the alternative residual definitions, $R_0^{loc}$ and $R_0^{glob}$, when performing calculations. The definition of $R_0^{glob}$ has no dependence on the number of electrons in the system and as a result a uniform tolerance can be used across different scales of simulations, this will be simpler for the user but more importantly ensure that the convergence of the SCF residual ensures that the energy has also converged to a minimum. This comes at the cost of efficiency, as the mod-
ified residual definitions reach lower magnitudes of energy differences the number of SCF iterations required is typically greater, this work could feasibly be extended to address such issues. If such issues were overcome this formulation of the residual would represent a more robust choice in terms of achieving self consistency than more commonly used definitions dependent on the Frobenius norm and this being the case this work could straightforwardly be adapted to other DFT codes.
Chapter 4

Self compensation and active
Aluminium dopants in Si

4.1 Background

The process of doping has proved a crucial tool for manipulating the physical properties of semiconductors and has been in widespread use for decades. Although a relatively low dopant concentration ($10^{13}$ atoms cm$^{-3}$) is sufficient to alter the material properties of a substrate, the decreasing dimensions of modern devices has required a corresponding increase in dopant concentration\[120\].

Silicon technology is approaching scales where device characteristics are determined not only by the dopant concentration but also by the location of individual dopant atoms. It is expected that variability in device performance will be caused by the statistical nature of dopant placement\[121\], which will impose limits in scalability prior to the physical limits associated with lithography and quantum effects. Hence precision doping is likely to become an essential tool. The need for precision has resulted in the development of patterned atomic layer epitaxy (PALE), which uses STM based lithography of H-passivated Si(001), and has been utilised to fabricate a FET transistor based on a single P atom using phosphine as a precursor\[30, 122, 123, 124\]. The progress in precision doping and the PALE process has not extended to acceptor dopants. Although boron is often used as a p-type dopant, diborane is not a suitable candidate for a precursor in the PALE
process due to the inability to selectively deposit single boron atoms\cite{125, 126}. Furthermore boron would induce strain when used for delta-doping due to its small size relative to Si, resulting in fast diffusion rates\cite{127} which would not give the precise dopant profile desired. However the amine alanes are known to be viable precursors for the thin film deposition of aluminium\cite{128}, therefore Al may be a suitable candidate for the precision doping of Si. This precision doping could allow for fully compensated co-doping, complementing P doping, which can be used to effectively tune the dopant populations, electronic properties, and magnetic properties\cite{129}. Alternatively, pairs of Al may self-compensate, as doped semiconductors can form intrinsic defects such as vacancies, anti-sites, and interstitials\cite{130, 131} which reduce the formation energy of these dopant structures. This self-compensation could potentially improve the carrier concentration and conductivity \cite{132} depending on whether the individual dopants are electrically active as acceptors.

This work is motivated by the expectation that Al will become a viable candidate for precision doping in Si\cite{133}. There is therefore a need to determine whether a self-compensated effect could occur in highly doped samples where Al dopants occupy neighbouring sites but form a complex where the dopants are electrically inactive. It is however plausible that bonded complexes could form in heavily doped materials.

\section*{4.2 Methods}

\subsection*{4.2.1 Computational Details}

All calculations were performed using density functional theory, as implemented in version 5.4.1 of the Vienna ab initio simulation package (VASP) software\cite{1}. The exchange correlation potential has been approximated by the Perdew-Burke-Enzerhof (PBE) generalised gradient approximation (GGA)\cite{61} functional. The VASP projector-augmented-wave (PAW)\cite{134} potentials for Al and Si were used with a 250 eV energy cut-off, which will incorporate all plane waves from the basis sets and allow consistency across all calculations. These potentials describe both core and valence electrons and the files (POTCAR) were dated 4/5 January 2001
and 15th June 2001, respectively\footnote{The author is aware of the findings of the ‘delta project’\cite{delta_project}, this project provided invaluable research benchmarking the precision and reproducibility of numerous DFT codes and methods. The precision of the potentials used here (although nearly two decades old at the time of publication) is not explicitly called into question by the ‘delta project’ and they are still provided in the latest stable release of the VASP code (5.4.4) at the time of writing.}.

All lattice parameters and atomic positions have been optimised using a quasi-Newton ionic relaxation algorithm which minimises both energy and interatomic forces. The convergence criteria for the atomic forces and for the electronic structure were set to 0.01 eV/Å and $1 \times 10^{-6}$ eV respectively. These parameters yield relative energies reliable to within 0.02 eV when the Brillouin zone sampling mesh is set appropriately.

The Monkhorst-Pack\cite{Monkhorst-Pack} k-point sampling employed in these calculations for the 64, 216 and 512 atom supercells detailed below, were $(4 \times 4 \times 4),(3 \times 3 \times 3)$ and $(2 \times 2 \times 2)$ respectively. The energy values were found to converge for these Brillouin zone samplings and give consistent sampling of the Brillouin zone.

For analysis and the production of all two-dimensional charge density and electron localisation function plots shown in this work the visualisation software VESTA\cite{VESTA} was used.

### 4.2.2 Structural Models

In this study cubic supercells of tetrahedral bulk silicon are constructed and are subject to periodic boundary conditions. The experimental bulk Si lattice parameter of 5.431 Å\cite{Experimental_lattice_parameter} was used as a starting value for all structures considered prior to the relaxation of lattice parameters and atomic positions. The aluminium dopants were placed at substitutional sites within the silicon structures with the total number of atoms modelled ranging from 64 to 512 atoms. The different cell sizes are used to explore the effect of system size on dopants, to adjust the isolation of the dopants and determine whether there is a non-negligible influence of the periodic images of the dopants which could affect bonding. For the electronic structure, a single reliable cell size will be used. Both single Al dopants and pairs of Al dopants are considered, with the pairs occupying either adjacent or non-adjacent substitutional sites. The different ratios of Si and Al atoms will be denoted as $\text{Si}_n\text{Al}_m$ throughout,
with \( n \) and \( m \) representing the number of silicon and aluminium atoms respectively.

### 4.2.3 Electronic Structure

#### 4.2.3.1 Density of States

In order to investigate the self compensation of the Al-Al pairing and the localisation of the associated holes the projected density of states (PDOS) will be considered. The PDOS is produced by projecting the orbitals onto spherical harmonics which have non-zero values within spheres around each atom, these spheres have radii equal to the Wigner-Seitz radius[139] of each species. The partial occupancies for the PDOS are determined using Gaussian smearing techniques, using a smearing width of 0.02 eV.

#### 4.2.3.2 Electron Localisation Function

In order to investigate the presence of bonds between Al dopants, we have used the electron localisation function (ELF)[140], which is a function of the spatial coordinates that has a high value in areas of concentrated electron density and can give a useful quantitative representation of chemical bonds[141]. It is used here to investigate the presence of bonding between Al dopants. The electron localisation function \( \eta(r) \) is dependent on the filled orbitals and density \( \rho \), it is defined as follows:

\[
\eta(r) = \frac{1}{1 + \left( \frac{D}{D_h} \right)^2}, \quad (4.1)
\]

\[
D = \frac{1}{2} \sum_{i=1}^{N} \nabla|\psi_i|^2 - \frac{1}{8} \frac{|
abla \rho|^2}{\rho}, \quad (4.2)
\]
where $D(r)$ is the probability of an electron being in the proximity of an electron of the same spin. For a homogeneous electron gas this probability has the value $D_h$:

$$D_h = \frac{3}{10} (3\pi^2)^{\frac{2}{3}} \rho^{\frac{5}{3}},$$  \hspace{1cm} (4.3)$$

here the standard definition of electron density has been used where $\rho$ is:

$$\rho = \sum_{i=1}^{N} |\psi_i|^2,$$  \hspace{1cm} (4.4)$$

for the Kohn-Sham orbitals $\psi_i(r)$. As shown above, the ELF is inversely proportional to $D(r)$. Therefore a low probability results in a high ELF which represents a localised pair of electrons. The ELF values are bound by 0 and 1 so a perfectly localised orbital has a value of 1, a homogeneous electron gas will have a value of 0.5 and values will tend to 0 as electrons become completely uncorrelated.

### 4.3 Results and Discussion

#### 4.3.1 Structural Features

The simulations containing a single substitutional Al dopant have identical structural features with a constant Si-Al bond length of 2.43 Å for all cell sizes considered for both the spin polarised and spin unpolarised cases in which the single unpaired electron occupies the $S = \frac{1}{2}$ or $S = 0$ spin state respectively. The $S = 0$ state has been modelled for comparison with the spin free pairs, and to represent metallic silicon. This indicates an increase in bond length around the dopant atom from the equilibrium Si-Si separation of 2.37 Å in Si bulk, although there are no non-zero forces present due to the successful atomic relaxation performed. This is consistent with results which have been previously reported in the literature\cite{142} where Si-Al bond lengths ranging from 2.42 Å to 2.50 Å in spherical Si nanocrystallite structures have been presented.

In the cases where pairs of Al dopants are initially positioned in adjacent substitutional sites, there is an expansion of the Al-Al separation following ionic relaxation for both the $S = 1$ and $S = 0$ spin states. The magnitude of this expansion is de-
dependent on whether spin polarisation is included in the simulation despite there being an even number of electrons present. When spin is introduced and the holes associated with the dopants are both forced to occupy either a spin up or spin down state then the Al-Al separation is comparatively smaller than the case where there is no net spin, the complete set of separation values is shown in Table 4.1. In both cases, the dopant separations are greater than the bulk silicon separations of 2.37 Å found in these simulations.

**Table 4.1:** The separations between adjacent Al dopants and the nearest neighbouring Si atoms, shown as Si$_{NN}$, as well as the distance to the next nearest neighbouring Si atoms denoted here as Si$_{NNN}$. The bond angles and relaxed lattice constants for spin polarised and spin free simulations are included.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Net Spin</th>
<th>Al-Al Separation(Å)</th>
<th>Al-Si$_{NN}$(Å)</th>
<th>$\angle$Si$<em>{NN}−Al−Si</em>{NN}$(°)</th>
<th>Si$_{NN}$-Si(Å)</th>
<th>Lattice Constant(Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si$_2$Al$_2$ (adjacent dopant)</td>
<td>1</td>
<td>2.61</td>
<td>2.43</td>
<td>112.71</td>
<td>2.35</td>
<td>5.4807</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2.85</td>
<td>2.42</td>
<td>115.82</td>
<td>2.36</td>
<td>5.4809</td>
</tr>
<tr>
<td>Si$_{10}$Al$_2$ (adjacent dopant)</td>
<td>1</td>
<td>2.62</td>
<td>2.42</td>
<td>112.84</td>
<td>2.35</td>
<td>5.4695</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2.83</td>
<td>2.41</td>
<td>115.60</td>
<td>2.36</td>
<td>5.4696</td>
</tr>
</tbody>
</table>

The separation for the $S = 1$ case is in good agreement with Al-Al bond lengths in small aluminium clusters previously obtained from DFT simulations[143] as well as the experimental Al$_2$ dimer bond length of 2.70 Å[144]. The non-spin-polarised separations show closer agreement to values obtained for bond lengths within bulk aluminium cells of 2.86 Å[145]. This disparity can not be attributed to a deformation of the cell size with the introduction of spin, as both the spin polarised simulations and those without spin polarisation have similar lattice constants following cell relaxation. The tetrahedral symmetry is preserved beyond the next nearest neighbours of the dopants in simulations with either spin configuration. However, the local tetrahedral symmetry is broken, as shown by the bond angles in Table 4.1, with the $S = 1$ case showing a greater distortion from the tetrahedral angle of $109.28^\circ$. This is as a result of Si-Al bond lengths that are marginally smaller than the bond lengths observed in pure Si bulk. These bonds lengths are also comparable to those obtained for the simulations containing single Al dopants, measured at 2.43 Å.
4.3.2 Electronic Structure

4.3.2.1 Single Substitutional Aluminium Dopants

The presence of Al-Si bonding is investigated by observation of the total charge density, as shown for the dopant plane in Figure 4.1. For clarity and consistency the 216 atom structures have been used to produce all plots containing a single atomic plane. The charge density in the regions between the Al dopant and the neighbouring Si atoms is measured at approximately 95% of the magnitude of the charge density in the regions occupied by Si-Si bonds. This indicates that the Al will have bonded to its neighbouring Si atoms forming an unoccupied dopant state hence the preservation of tetrahedral symmetry. The regions of such high charge density are consistent with electron localisation and therefore further analysis of bond with the ELF is unnecessary.

![Figure 4.1: A 2D plot of total charge density across the \{10\bar{1}\} plane containing the dopant for the Si\(_{215}\)Al\(_1\) structure in the \(S = \frac{1}{2}\) and \(S = 0\) spin states](image)

From the analysis of the density of states, shown in Figure 4.2, it can be seen that the incorporation of the Al dopant results in the presence of a shallow acceptor level in the gap for both the \(S = \frac{1}{2}\) and \(S = 0\) simulations. The bandgap is equal for both the \(S = \frac{1}{2}\) and \(S = 0\) states. The projected DOS shows that the main effect of the Al doping is to move the Fermi level into the valence bands, as would be expected for a p-type dopant, while introducing new peaks in the valence band; a clear example is seen at the Fermi level in Figure 4.2(ii). There is a clear difference
4.3. Results and Discussion

between the spin polarised (i-iii) and non-spin polarised (iv-vi) cases, both in terms of the position of the Fermi level and the detailed shape of the DOS. The Al atom shows subtle differences between the spin channels as well as the shift of the Fermi level.

Figure 4.2: The total and atom projected DOS for both undoped Si and the Si$_{2}$Al$_{1}$ structure in the $S = \frac{1}{2}$ spin state, shown as i)-iii) and the same Si$_{2}$Al$_{1}$ structure in the $S = 0$ state shown as iv)-vi)

4.3.2.2 Substitutional Aluminium Dopant Pairs

The single dopant Si$_{2}$Al$_{1}$ model is representative of a dopant concentration of $2.3 \times 10^{20}$ cm$^{-3}$, this beyond the metal-insulator transition for both boron[146] and phosphorous[147] both of the order of $1 \times 10^{18}$ cm$^{-3}$. At dopant concentrations of this order it is probable that a sample will contain dopants at nearby and even adjacent sites. This possible proximity could result in the dopants self-compensating, to determine whether such an effect is present we consider simulations of adjacent and non-adjacent substitutional Al pairs, with the individual dopants in non-adjacent pairs being placed at the greatest separation possible within the super cell.

A key quantity of interest when considering dopants or defects is the formation
energy $E^f_{\text{defect}}$. In general terms this provides a quantitative value for the relative energetic stability of doped structures in comparison to pure undoped structures. It can be helpful to consider this as the energy required to ‘embed’ a defect in a perfect crystal (in this case crystalline silicon), through a combination of adding, removing and rearranging atoms. In this manner we can define the formation energy of the Al dopants in the structures considered here following the standard formalism proposed by Zhang and Northrup[148]. This is shown in Equation 5,

$$E^f_{\text{defect}} = \frac{E_{\text{doped}} - \left( \frac{N_{\text{Si,doped}}}{N_{\text{Si,bulk}}} \right) \ast E_{\text{bulk}} - N_{\text{Al}} \mu_{\text{Al}}}{N_{\text{Al}}}$$  (4.5)

where $E_{\text{bulk}}$ is the energy of the optimised bulk undoped cell, $E_{\text{doped}}$ is the energy of the optimised doped cell and $N_{\text{Si,doped}}$ and $N_{\text{Si,bulk}}$ are the total numbers of Si atoms in each type of structure. $N_{\text{Al}}$ is simply the number of substitutional aluminium dopants introduced to a given doped structure. The remaining term $\mu_{\text{Al}}$ represents the chemical potential of the Al dopants, however for the purposes of this calculation this is not straightforward to define.

For a monatomic crystal such as pure silicon, the chemical potential of the species is well defined: it is simply equal to the energy per atom in the bulk crystal. However for the case of doped systems it is necessary to determine the of partitioning the energy of the simulation cell between its constituent species. Nonetheless we can simplify this problem by placing an upper bound on the chemical potential of our dopant ($\mu_{\text{Al}}$) by noting that for an Si-Al structure to form, neither of the Si or Al chemical potentials can be higher that the energy per atom of its species’ elemental configuration. Therefore, $\mu_{\text{Al}} \leq \mu_{\text{Al}}^{\text{Bulk}}$, where $\mu_{\text{Al}}^{\text{Bulk}}$ is the energy per atom of crystalline Al, which has been determined using a static simulation of a 64,216 and 512 atom supercells with the convergence criteria outlined in section 4.2.1. This gives an upper estimate for the chemical potential and therefore the formation energies presented here provide a lower bound for formation energies that may be determined experimentally. In order to determine the chemical potential of the dopant to
4.3. Results and Discussion

a higher level of accuracy we need to answer the following questions: (i) what reservoir is the dopant being removed to or taken from, and (ii) how much of the total energy of that reservoir can be attributed to the dopant species alone. This requires an in depth consideration of the doping process itself and hence we have restricted this study to considering an upper bound of $\mu_{Al}^2$.

The formation energies or relative energetic stabilities shown in Table 4.2 demonstrate that for larger systems the substitutional impurities have greater stability, likely due to the reduced interaction with the periodic images of the dopants. It is interesting to note that there is variation in the formation energies of pairs of dopants at adjacent and non-adjacent sites within the two largest supercells considered here. There is only a small increase in stability associated with the dopants occupying adjacent substitutional sites for the $S = \frac{1}{2}$ state however this is not necessarily prohibitive to the formation of such a structure, as there is a small driving force encouraging the Al atoms to form pairs. It can be noted at this point that electrostatic repulsion between dopants is effectively screened in semiconductors beyond a distance of approximately 1nm.

<table>
<thead>
<tr>
<th>Structure</th>
<th>$E_{\text{defect}}^f S = 0$ (eV)</th>
<th>$E_{\text{defect}}^f S \neq 0$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si$_{63}$Al$_1$</td>
<td>1.51</td>
<td>1.48</td>
</tr>
<tr>
<td>Si$_{62}$Al$_2$ (sep)</td>
<td>1.55</td>
<td>1.46</td>
</tr>
<tr>
<td>Si$_{62}$Al$_2$ (adj)</td>
<td>1.55</td>
<td>1.38</td>
</tr>
<tr>
<td>Si$_{1215}$Al$_1$</td>
<td>1.49</td>
<td>1.47</td>
</tr>
<tr>
<td>Si$_{1214}$Al$_2$ (sep)</td>
<td>1.49</td>
<td>1.44</td>
</tr>
<tr>
<td>Si$_{1214}$Al$_2$ (adj)</td>
<td>1.53</td>
<td>1.41</td>
</tr>
<tr>
<td>Si$_{511}$Al$_1$</td>
<td>1.49</td>
<td>1.49</td>
</tr>
<tr>
<td>Si$_{510}$Al$_2$ (sep)</td>
<td>1.46</td>
<td>1.43</td>
</tr>
<tr>
<td>Si$_{510}$Al$_2$ (adj)</td>
<td>1.53</td>
<td>1.42</td>
</tr>
</tbody>
</table>

The relative energetic stabilities are spin dependent with the spin-polarised

$^2$It should also be noted the $\mu_{Al}$ will have some temperature and pressure dependence although these terms are negligible for bulk solids generally speaking. More information on determining the chemical potential of dopants is provided in other sources[149, 150]
structures being energetically favourable for both the single Al dopant and the pairs of Al dopants. Although given the high dopant densities considered we would perhaps expect the metallic state to be more stable. It is also evident that the energetic stability of each Al dopant is lower when pairs of dopants are introduced rather than a single Al dopant, this is due to the lack of partial occupancy in the multiple dopant case. Furthermore these formation energies show good agreement with the formation energies of approximately 1.50eV for Al dopants in spherical Si nanocrystallite structure in the literature[142] also produced using the VASP software.

The formation energies indicate that in addition to the differences in separations between the adjacent Al dopants in the two possible spin states that there is a change in the electronic structure and indeed this is what is observed. As shown in Figure 4.3 the charge density between adjacent dopants in the $S = 1$ spin state is of a greater magnitude than in the same region for the $S = 0$ state. For the $S = 1$ state the charge density between the Al dopants has a value of approximately 60% of the density in the region between Si-Si bonds, opposed to a value of 41% of the Si-Si bond density for the $S = 0$ case.

![Figure 4.3: A 2D plot of total charge density across the \{10\overline{1}\} plane containing the dopant for the Si$_{214}$Al$_2$ structure in the $S = 1$ and $S = 0$ spin states](image)

The differences in the electronic structure of the two spin states are also reflected in the DOS, as shown in Figures 4.4 and 4.5. The adjacent dopant structure in both spin states exhibits the creation of shallow acceptor states, as is the case for the single Al dopant. However the Al PDOS in the $S = 1$ state, which is the
4.3. Results and Discussion

Figure 4.4: The total and atom projected DOS for the Si_{14}Al_{2} structure in the $S = 0$ state with the aluminium dopants in adjacent sites.

The same both dopants, shows a greater spin splitting in the acceptor state than the other simulations considered here with clear separate Al peaks. This $S = 1$ acceptor state also shows a greater energy separation from the valence band maximum.

There are also differences observed between the two spin configurations of adjacent dopants when considering the VBM. The band decomposed partial charge densities for all k-points show that for the model containing adjacent dopants in the $S = 1$ spin state there is a localisation of charge between the Al dopants in the highest energy occupied state, as depicted in Figure 4.6(b). The isosurface strongly resembles a localised dopant state. In comparison, the highest energy occupied band for the $S = 0$ spin state, shown in Figure 4.6(a), is highly delocalised. The localisation of charge in the occupied band for the $S = 1$ case combined with the high charge total density in the region between the adjacent Al dopants is suggestive of a bond. Furthermore with the majority of observable charge being more localised in the $S = 1$ state than the $S = 0$ state shows that the Al dopants are more likely to be active in the higher net spin configuration.
4.3. Results and Discussion

Figure 4.5: The total and atom projected DOS the Si\textsubscript{214}Al\textsubscript{2} structure in the $S = 1$ state with the aluminium dopants in adjacent sites.

Figure 4.6: The isosurface, at a value of $2.58 \times 10^{-3}$ eV/Å, for the partial charge density for the highest energy fully occupied state of Si\textsubscript{214}Al\textsubscript{2} in the $S = 0$ and the $S = 1$ spin up state with adjacent Al dopants shown in purple and labelled, the Si is shown in gray.

4.3.2.3 Electron Localisation Function for Substitutional Aluminium Pairs

The ELF plots for both the spin polarised and non-spin polarised simulations of Al dopants in adjacent substitutional sites are shown in Figure 4.7. These have been scaled to the maximum and minimum possible values of the ELF which are 0 and 1. As outlined in section 2.3 these values represent the probability of pairs of electrons
occupying a particular region of space. While the change in bond length for the $S = 0$ and the $S = 1$ states may not be clear in Figure 4.7, the property of interest is the higher ELF value of the region between the dopants in the spin polarised simulation. The values of the ELF within this region range from 0.92 to 0.93, one might associate an ELF value above 0.7 with a weakly bonded system. For the same simulation, the ELF values between the Si-Si bonds and Al-Si bonds are approximately 0.89 and 0.88 respectively, further evidence that a bond is present between the adjacent Al dopants. For comparison simulations performed using the same convergence criteria for alane (AlH$_3$) and water (H$_2$O) had ELF values of 0.99 in the regions where bonding would be expected to occur. As a result we are confident that Al-Al bonds are present in the $S = 1$ state, but there are no such bonds in the $S = 0$ state.

![Figure 4.7: 2D ELF plots across the \{10\bar{1}\} plane containing both neighbouring Aluminium dopants in S=0 and S=1 states.](image)

4.4 Conclusions

We have used DFT to investigate the structure and energetics of aluminium dopants within bulk silicon including multiple dopants occupying adjacent substitutional sites.

The interaction between the acceptor dopants has been studied using a combination of the ELF and visualisations of the densities of states. For the simple case of a sin-
gle Al dopant we observe the presence of three bonds with neighbouring Si atoms and no gap state present in the DOS, indicating the dopant exists in a saturated and electrically inactive state.

For the adjacent dopants the data presented shows that a bond is formed between the Al dopants in the $S = 1$ state. This is supported by a high ELF value in the region between the two acceptors, a non-negligible charge density in the same region and the separation of the Al dopants is consistent with the bond lengths cited from previous force field and DFT simulations. However the same analysis shows that no such bond is formed between the Al dopants in the $S = 0$ state.

The adjacent Al dopants therefore will not be electrically active and therefore will be self-compensating unless they occupy a high spin state in which case they will become bonded. Hence if amine alanes are viable candidates for precision doping then aluminium dopants may be used for co-doping devices. Furthermore, this raises the possibility of a precision doped structure which can be switched from the $S = 0$ to the $S = 1$ state using a combination optical pumping and magnetic fields, given the low 0.11eV energy difference between the states for the bonded Al complex.
Chapter 5

A structural study of single Bismuth dopants in crystalline Si

5.1 Background

The dawn of atomically precise doping combined with the ability to image single atoms\cite{151} offers numerous opportunities in quantum engineering and has led to renewed interest in dopants that previously have had little or no use within the semiconductor industry. An example of this is bismuth, Bi is almost insoluble in Si due to the vast difference in covalent radius\cite{152} making the incorporation of Bi into Si film problematic\cite{153}. Hence Bi has historically been considered too impractical to be used as a dopant for industrial applications. However, these issues do in turn make Bi an ideal test subject for overcoming difficulties associated with single atom manipulation. This is an intriguing possibility as, at the time of writing no method for targeting a specific atom within a three dimensional crystal exists. There are other properties which make Bi an appealing candidate for quantum technologies, for example it has the highest binding energy (70 meV)\cite{154} of group V donors in Si, the largest nuclear spin ($I = 9/2$), and the largest hyperfine coupling strength (1.4754 GHz)\cite{155}. These characteristics have resulted in Bi doped Si being identified as a viable candidate for donor based quantum computing\cite{156, 157, 158, 159}, for example in the optically driven Stoneham-Fisher-Greenland\cite{160} approach adapted from Kane’s original proposal\cite{27}. Bi
would be ideal for such an implementation, the properties listed above mean Bi based qubits would be able to operate at higher temperatures, and create a larger auxiliary state space for quantum data storage[159]. Furthermore Bi has recently been observed to have much longer spin dephasing times[156] than a P dopant in Si, of the order of 1ms at 10K. However the introduction of a Bi dopants into the Si lattice is likely to result in lattice distortions and point defects[161, 162], potentially affecting these dephasing times of the spins, which will in turn impact the coherence times of any qubit architecture constructed using Bi dopants. Hence, the success of Bi based quantum computing schemes will depend on the material quality after incorporation of the Bi atoms into Si and therefore a strong theoretical understanding of the structure is required.

5.2 Methods

5.2.1 Computational Details

The majority of calculations were performed using density functional theory, as implemented in the local orbital code CONQUEST. The exchange correlation potential has been approximated by the Perdew-Burke-Enzerhof (PBE) generalised gradient approximation (GGA)[61] functional. For the calculations performed using CONQUEST a double-$\zeta$ plus polarisation orbital basis of PAOs has been employed. The PAOs for the basis sets were generated using the radial cutoff method as implemented in the Siesta code[90], the cutoffs for the PAOs were chosen to give an optimised bulk lattice parameter. These pseudopotentials incorporate scalar relativistic which are necessary for heavy atoms with many electrons such as Bi and have proved essential in previous simulations involving this element[163]. In addition basis sets using the PBE functional and the projector-augmented-wave (PAW)[134] potentials for Bi and Si were used with a 300 eV energy cut-off were used for a comparative calculation performed in version 5.4.1 of VASP[1].

The lattice parameters and atomic positions for all structure considered have been optimised using a quenched molecular dynamics[164] which minimises both the groundstate energy and inter-atomic forces. The convergence criteria for the atomic
forces and for the electronic structure were set to $2 \times 10^{-4} \text{ Ha/Å}$ to $1 \times 10^{-7} \text{ Ha}$ respectively. These parameters yield relative energies reliable to within 0.01 eV when the Brillouin zone sampling mesh is set appropriately. Monkhorst-Pack\cite{136} k-point sampling employed in these calculations for the 64, 216 and 512 atom supercells detailed below, were $(6 \times 6 \times 6),(4 \times 4 \times 4)$ and $(3 \times 3 \times 3)$ respectively. The energy values were found to converge for these Brillouin zone samplings and give consistent sampling of the Brillouin zone. A grid cutoff of $150\text{ Ha}$ was used for all structures, having observed convergence in the groundstate energy of the order of $1 \times 10^{-5} \text{ Ha}$ for this value of the cut-off for the most densely doped structures. The self-consistency tolerance value, at which the iterations of the SCF cycle terminate, was set at $1 \times 10^{-7} \text{ Ha}$.

5.2.2 Structural Models

In order to consider the possible structures formed by introducing a Bi dopant to Si, cubic supercells of tetrahedral crystalline Si are constructed and are subject to periodic boundary conditions. The experimental bulk Si lattice parameter of 5.431 Å\cite{138} was used as a starting value for all structures considered prior to the relaxation of lattice parameters, which were optimised for each individual structure. Due to the large stresses induced by the presence of any Bi dopants in Si, as a result of the vast differences in covalent radii, vacancies are highly likely to be introduced and as such we consider the presence of vacancies in our structural models. The Bi dopants were placed either at a substitutional lattice site within the Si structures with and without an additional vacancy at one of the nearest neighbouring lattice sites, the structures without a vacancy are referred to here as ‘Structure A’ and those with a vacancy are referred to as ‘Structure B’. Alternatively the dopant is placed at the midpoint of two substitutional lattice sites with each site replaced by a vacancy this is referred to here as ‘Structure C’. The different atomic structure are depicted in Figure 5.1. The total number of atoms modelled ranging from 64 to 512 atoms. The different cell sizes are used to explore the effect of system size on dopants, to adjust the isolation of the dopants and determine whether there is a non-negligible influence of the periodic images of the dopants which could affect the Si-Bi bond-
5.3 Results and Discussion

5.3.1 Relative Energetic Stabilities

In order to determine the relative likelihoods of the structures listed above forming when a Bi dopant is introduced to the Si lattice it is instructive to consider the relative energetic stabilities of the structures formed. A direct comparison of the energies of these structures is not possible, because the simulation cells contain...
5.3. Results and Discussion

different numbers of atoms, instead the following quantity can be calculated:

\[ E_{\text{rel}}^f = E_{\text{dop}} - \frac{n_2}{n_1} E_{\text{undoped}} - \mu_{\text{Bi}} \]  

(5.1)

where \( E_{\text{rel}}^f \) is the formation energy or relative energetic stability of the Bi dopant, with the equation used following the approach of Zhang and Northrup[148]. Here \( E_{\text{dop}} \) and \( E_{\text{undoped}} \) are the total energies of the doped cell and undoped Si cells respectively. For structures B and C the undoped cell will contain a vacancy so that we are considering the formation energy of the dopant rather than the dopant-vacancy complex. The number of Si atoms prior to the introduction of the dopants is denoted by \( n_1 \) while \( n_2 \) is simply the total number of Si atoms present excluding the total number of dopants and vacancies. The chemical potential of the Bi atom \( \mu_{\text{Bi}} \) is estimated by placing an upper bound on the chemical potential of our dopant (\( \mu_{\text{Bi}} \)) by noting that for an Si-Bi structure to form, neither of the Si or Bi chemical potentials can be higher than the energy per atom of its species’ elemental configuration. Therefore we approximate \( \mu_{\text{Bi}} \) as \( E_{\text{Bi}} \), where \( E_{\text{Bi}} \) is the energy of an individual Bi atom within a bulk Bi crystalline structure calculated using the experimental lattice constant and same convergence criteria outlined in section 5.2.1.

The values determined using the CONQUEST code and Equation 5.1 for our simulated structures are given in Table 5.1.

<table>
<thead>
<tr>
<th>Total Atoms and Vacancies</th>
<th>Dopant Formation Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Structure A</td>
</tr>
<tr>
<td>64</td>
<td>4.007</td>
</tr>
<tr>
<td>216</td>
<td>3.759</td>
</tr>
<tr>
<td>512</td>
<td>3.452</td>
</tr>
</tbody>
</table>

Table 5.1: The dopant formation energies for the structures containing Bi dopants (shown in purple) at a substitutional site (Structure A), at a substitutional site with a neighbouring vacancy (Structure B) and between a vacancy and a substitutional site (Structure C). The dopants are placed in Si cells initially containing 64, 216 and 512 atoms.

As is expected there is a decrease in the formation energy of the dopant, which in turn indicates an increase in the relative energetic stabilities of the dopant, as the
5.3. Results and Discussion

simulations cells increase in size. This suggests that lower dopant concentrations result in more stable dopant configurations when Bi is introduced to Si. This is likely due to the influence of the periodic images of the dopant affecting the energetic stabilities of the structures in the smallest cells considered here. This is a consequence of the periodic boundary conditions within our simulations. It is possible that if the sizes of the simulation cells were taken to more extreme scales then further reduction in the formation energy would be observed for Structure A, nonetheless we have obtained a reasonable agreement to the experimental formation energy of a Bi dopant in Si measured at 3.5 eV[165].

The small scales of the differences in formation energies for the different structures considered here, ranging from 0.127 to 0.174 eV, indicate that it is plausible that any of these structures may form when a Bi dopant is introduced. However, it should be noted that for all simulation cells the structures including a vacancy (Structures B and C) are shown to be energetically favourable, with a reduction in the formation energy and increase in relative energetic stability of at least 0.090eV. Furthermore, the structure featuring Bi initially between two substitutional sites, has a marginally greater stability then the model at which the dopant is simply introduced at the substitutional site. If time allowed it would be of interest to extend this study to consider diffusion barriers and dopant mobility to explore the possibility of transitions between the structures considered here. The difference in formation energies between the two structures extends to 0.041eV in the largest simulation cell considered in these simulations.

In addition we can compare the formation energies of the vacancy in a Bi doped Si cell to the vacancy formation energies of an undoped cell by introducing a vacancy to structure A following the relaxation of the doped structure. The results of which are calculated using a variation of equation 5.1 and are shown in Table 5.2. The reduction in magnitude of the vacancy formation energy from the undoped Si cell to the Bi doped structure is likely a result of the strain induced by the Bi dopants. It follows that the Si-Si bonds in the respective simulations cells are more stable when strain from the Bi dopant is not present, as would be expected, and therefore
in simple terms there are greater energy requirements to form a vacancy. This indicates there is an increased likelihood of vacancies forming in Si when Bi dopants are present, in agreement with previous DFT simulations performed by Hudak et al.[166] who show a reduction in vacancy formation energies in crystalline Si of approximately 2.0 eV in the presence of a Bi dopant.

<table>
<thead>
<tr>
<th>Total Atoms and Vacancies</th>
<th>$E_{\text{vac}}^f$ Pure Si</th>
<th>$E_{\text{vac}}^f$ Structure A</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>3.424</td>
<td>1.543</td>
</tr>
<tr>
<td>216</td>
<td>3.348</td>
<td>1.863</td>
</tr>
<tr>
<td>512</td>
<td>3.317</td>
<td>1.968</td>
</tr>
</tbody>
</table>

Table 5.2: Vacancy formation energies with pure Si and a Bi doped Si cell (Structure A). The vacancies are introduced to simulations cells initially containing a total 64, 216 and 512 atoms.

These energy trends are reflected when performing plane wave based simulations of the same structures. As outlined above the largest structures considered here, containing 512 atoms and vacancies were also simulated using the VASP code. The corresponding Bi formation energies within these structures are shown in Table 5.3.

<table>
<thead>
<tr>
<th>Total Atoms and Vacancies</th>
<th>Structure A</th>
<th>Structure B</th>
<th>Structure C</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>3.640</td>
<td>3.629</td>
<td>3.610</td>
</tr>
</tbody>
</table>

Table 5.3: The relative energetic stabilities of the structures containing Bi dopants at a substitutional site (Structure A), at a substitutional site with a neighbouring vacancy (Structure B) and between two vacancies (Structure C). The dopants are placed in Si cells initially containing 512 atoms and modelled using the plane wave based DFT code VASP.

Although the absolute magnitudes of the formation energies differ from the simulations performed using CONQUEST, the overall trend remains the same with structures including a vacancy appearing to be marginally energetically favourable when introducing a Bi dopant.

The difference in formation energies of the dopant in structures B and C can be
attributed to the different bonding configurations present in each structure following atomic relaxation. A subsection of the charge densities for each structure is illustrated on Figure 5.2.

**Figure 5.2:** Subsections of the total charge density across the {101̅} and {101} planes for structures B and C following atomic relaxation, with contours at intervals of 0.15eV/Å³. The corresponding position of the Bi dopant within the atomic lattice has been denoted Bi.

(a) Structure B {101̅} plane containing Bi dopant
(b) Structure C {101} plane containing Bi dopant
(c) Structure B {101} plane featuring nearest neighbour Si atoms of Bi dopant
(d) Structure C {101} plane featuring nearest neighbour Si atoms of Bi dopant

By visualising the charge densities of the structures incorporating vacancies following atomic relaxation, it can be observed that stable dopant complexes are
formed with the Bi dopant remaining at the substitutional or interstitial site. Furthermore, the presence of a non-negligible charge density, at approximately 84% of the value in areas where Si-Si bonds are present, between the Bi dopant and the nearest neighbouring Si atom to the vacancy site indicate that a bond has been formed across the vacancy. This is depicted in the \{10\overline{1}\} plane in Figure 5.2(b), whereas the charge density within the same vacancy region for structure B (shown in Figure 5.2(a)) indicates no such bond is formed in this structure. In structure B however bonds do appear to form between the Bi dopant and the nearest neighbouring Si atoms. A depiction of the regions of charge density between Bi and the nearest neighbouring atoms in shown in Figure 5.2(c), again the charge density is at approximately 84% of the value in the areas of the Si-Si bonds. A similar likelihood of bonding is shown for structure C between Bi and the nearest neighbouring Si atoms, shown in Figure 5.2(d), indicating that these bonds are unlikely to be broken despite the likely presence of a bond across the vacancy. These findings indicate that stable Bi dopant complexes incorporating a single vacancy can exist with and without a bond across the vacancy. Ideally further analysis would be performed using the electron localisation function, however this is not currently a feature of the CONQUEST code. The presence of a stable Bi structure with a dangling bond would be consistent with previous DFT simulations performed by Kirkham et al\cite{167, 168}.

### 5.3.2 Introducing Additional Vacancies

The difference in the dopant formation energies observed for structures with vacancies present (Structures B and C) could be attributed to the presence of the vacancies relieving strain. Therefore it is possible that increasing the number of vacancies present may further improve the stability of the simulated structure, and that the introduction of a Bi dopant into a Si cells induces the presence of multiple corresponding vacancies. In order to verify the undoped Si cell was modified to include multiple vacancies so that the Bi dopant was introduced at a substitutional site that has two neighbouring Si atoms replaced by vacancies (i.e. one additional vacancy to structure B and is referred to as Structure D), an atomic schematic of
5.3. Results and Discussion

Figure 5.3: An atomic model of a subsection of the relaxed structure containing a Bi dopant (shown in purple) with two neighbouring substitutional vacancies. This structure is shown in Figure 5.3.

<table>
<thead>
<tr>
<th>Relative Energetic Stabilities (eV)</th>
<th>62 atom</th>
<th>214 atom</th>
<th>510 atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure D</td>
<td>4.292</td>
<td>4.056</td>
<td>3.803</td>
</tr>
</tbody>
</table>

Table 5.4: The relative energetic stability of the structures containing a Bi dopant with multiple neighbouring vacancies for structures containing a total of 62, 214 or 510 atoms.

The formation energies of the dopant within these structures for the different sizes of Si simulations cells considered is shown in Table 5.3. The data clearly shows that this structure is notably less stable than the structures with one or fewer vacancies. The formation energy increases by as much as 0.458 eV in comparison to Structure B (as shown in Table 5.1) for the smallest structure considered here and an increase in formation energy can also be observed when comparing to the simple substitutional case (Structure A). This increase is perhaps due to the presence of a dangling bond with only 4 Si atoms in close proximity to the Bi dopant. Therefore it seems unlikely that more than one vacancy would form at an atomic neighbouring site when a Bi dopant is introduced to Si.
5.4 Conclusions

This study presents three plausible dopant complexes that can be formed when a Bi dopant is introduced into crystalline Si, results that are likely to be reflected in simulations of surfaces when they approach the bulk limit. It has been shown that dopant complexes incorporating a single vacancy are energetically favourable to those with no vacancies present, as would be expected. The formation energies of the Bi dopant in single vacancy complexes have also been shown to be lower than complexes with multiple vacancies, perhaps a slightly less intuitive result.

There are numerous potential extensions to this study. The simulation cells could be extended to larger scales and perhaps modelled using the aforementioned multi-site support functions[3] to determine more rigorously if the energetic stability of the Bi structures converges at a given scale. These larger cells could also be used to model the incorporation of vacancies at sites further from the dopant, such as at next nearest neighbouring sites, or models incorporating multiple Bi dopants. These simulations would be ideal candidates for the recently developed \textit{ab initio} random structure searching techniques[169], which could potentially discover additional plausible Bi dopant complexes.

It would also be of interest to simulate these structures below the Si(001) surface as this would more closely model previous experimental work[161]. Preliminary calculations consisting of approximately 1400 atoms have been performed, however at the time of writing it had not been possible to successfully atomically relax these structures in order to minimise the forces, a computationally expensive process with large scale simulations. If successful this would allow for the production of simulated STM structures which may give greater insight into the dopant complexes formed, and allow for the verification of experimental data produced at UCL in which the Bi complex appears as a positively charged feature, as shown in Figure 5.4. It would be also be possible to perform DFT simulations of Bi complexes with net positive charge and again preliminary calculations have been performed but this work lacks the proper correction terms[170] that are required for reliable results, which would be a straightforward extension.
In addition, semi-core states using the density functional semi-core pseudo potential approximation[171] could be incorporated for the Bi dopants, however there is no reason to believe this will alter the hierarchy of dopant complex formation energies presented here.
Chapter 6

Surface effects on P δ-layers in Si slabs

6.1 Background

The advent of precision doping has made the potential realisation of solotronics, i.e. optoelectronic technologies based upon single dopants, possible. It has also brought about the possibility of exciting advances in monolayer based nanoelectronics. This is through the use of patterned atomic layer epitaxy (PALE)[172], which uses STM based lithography of H-passivated silicon or germanium surfaces to remove selected H atoms exposing dangling bonds in chosen patterns. Hence PALE gives the ability to precisely control and place individual phosphorus atoms as they adsorb to the exposed dangling bonds of Si or Ge [36, 38, 173]. Once P or the desired dopant is deposited then it is ‘overgrown’ with Si until the dopant is at the desired depth. One advantage of this precise control is that it can be used to fabricate a patterned δ-layer, an atomically abrupt monolayer region. Prior to this it had been possible to epitaxially grow a δ-layer however the methods used lacked any control over the position at which the impurity atoms were placed and hence the dopants would be at arbitrary positions[39]. The patterning advance has resulted in various prototype P-doped Si devices are currently being developed,[174, 175, 176, 177, 178, 179] in which patterned δ-layers are used to form conducting leads and gate electrodes. Furthermore these monolayers have been proposed as a method of fabricating a
solid state quantum computer with the P donors acting as qubits[27, 180], as they have been shown to exhibit long spin lifetimes in isotropically pure Si[181, 158]. These applications have attracted renewed interest in monolayers, bringing the theoretical understanding into focus.

In order to fully realise the potential of these devices and fabrication techniques, it is crucial that there is a solid theoretical understanding of their electronic structure and in turn their transport properties. However, at the time of writing, there are still significant gaps in the understanding of the physics of pseudo-two-dimensional dopant structures. For example, there are varying degrees of confinement exhibited by low-level structures, such as \( \delta \)-layers, inherent to these devices and determining the location of energy levels in these constituent structures will prove crucial to the successful operation of quantum devices as the conduction band valley structure (the valleys referring to minima in the electronic bandstructure\(^1\)) in Si is highly sensitive to changes in confinement in highly doped systems[38]. The approximately V-shaped potential resulting from the introduction of the \( \delta \)-layer causes lowering and discretisation of the conduction band, an effect known as ‘band-splitting’. This causes the formation of impurity bands potentially below Fermi level which consequently gives partially occupied states resulting in the overall structure having semi-metallic or metallic character.

Therefore there is ample motivation for further theoretical study of \( \delta \)-layers particularly when considering there have been relatively few theoretical and computational works studying these specific pseudo-two dimensional electron structures, particularly in the \textit{ab initio} regime. The first such study was performed by Qian et al.[182] using the planar Wannier orbital method to obtain the electronic structure of a single layer of P \( \delta \)-doped Si, this has since been followed by studies of P doped Si using the tight-binding method[183, 184, 185], and effective mass theory[186, 187]. However while large scales of simulations, perhaps of the order of tens of thousands of atoms close to that of realistic experimental scales, are accessible to these methods they rely on empirical parameters rather than \textit{ab initio} principles, there is an accuracy

\(^1\)Si is known as a ‘multi-valley’ semiconductor as it has six minima in the conduction band structure of the first Brillouin zone [75]
cost associated with this. There have on the other hand been previous DFT studies performed, such as that of Carter et al.\cite{188} using a mixed pseudopotential approach, as well as Drumm et al. who utilised a plane wave based code\cite{189} and the slightly more recent work by Budi et al. \cite{190}. Although using \textit{ab initio} methods in the form of DFT these studies use models which lack the explicit consideration of a vacuum or surface, raising the question of the significance of an interface, and the proximity of the $\delta$-layer to that interface, may have upon the electronic structure. It has already been suggested through experimental data\cite{191}, produced using angular resolved photo-emission spectroscopy (ARPES) that the depth of the encapsulation of the $\delta$-layer, and therefore its proximity to the surface, has an effect on the discretisation of the conduction band minima. The study considered 1/4 ML P-doped $\delta$-layers at depths from 1nm to 4nm below a clean non-H-passivated Si surface with an oxide termination. This work indicated the the energy splitting of the states below the conduction band were dependent on depth, with lesser magnitudes of energy splittings observed at greater depths. Furthermore there was the suggestion that the number of impurity bands formed may be dependent on surface effects, however findings were inconclusive due to the lesser intensity of the ARPES spectra for the 3 and 4nm $\delta$-doped samples. The primary concern of this chapter to determine the influence of an explicit surface-vacuum interface upon the electronic structure of P $\delta$-layers in Si, and verify the existing experimental studies.

A secondary aim will to be consider the effects of imperfect $\delta$-layers. Although the ideal $\delta$-layer is thought of as an entirely two-dimensional entity that exists in a single atomic plane within a substrate, this in reality is not necessarily the case. There are guaranteed imperfections in any fabrication process and in this case these may result in disorder within the $\delta$-layer or some broadening of the $\delta$-layer. This means that rather than the ideal two dimensional monolayer the dopants differ in depth within the Si slab resulting in the $\delta$-layer becoming a non-ideal three dimensional entity. This broadening has been observed to occur during the overgrowth process, when the overgrowth is performed at temperatures higher than approximately 540 K then a broadening effect can occur\cite{192}. This will result in P dopants in
closer proximity to the Si surface than perhaps desired, leading to the unintentional creation of a ‘thick δ-layer’. Therefore in this chapter we will consider the effects on the electronic structure of these disordered and broadened non-ideal δ-layers.

6.2 Methods

In this study, we utilised the Kohn-Sham based DFT methodology as implemented in the Conquest code with PAOs acting as our localised basis sets. Norm-conserving Hamann pseudopotentials[193] and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional[61] with the DZP basis set was used in this study (the choice of basis is explained in Appendix B). The generalised gradient approximation has previously been used for P-doped Si systems with excellent agreement with experiments[194, 195]. For the Si structures themselves we used a c(2 × 2) unit cell whose sides are parallel to the (110) and (110) planes, as the basis for our doped Si(001) surfaces. This unit cell has four sites per atomic plane, and provides the minimum size necessary to model 1/4 monolayer (ML) doping, which is the saturation coverage for a δ-layer, i.e. the maximum dopant concentration experimentally realisable in this type of structure.

In order to consider the influence of the vacuum, structures incorporating 40 and 80 Si layer surfaces were used to provide a fixed separation between the periodic images of the δ-layers. The layers were placed at substitutional sites at a depth of 20 and 40 Si layers in the respective cells, a schematic is provided in Figure 6.1 to aid visualisation.

These structures were modelled with and without an explicit vacuum of 20Å. In order to more closely represent the experimental...
tal conditions where H passivisation is used and an oxide may have formed on the surface, the Si slabs were terminated with hydrogen atoms at the top (monohydride) and bottom (dihydride) surfaces in order to terminate dangling bonds. The dihydride at the base of simulation cell represents the bulk type bonding structure that would be observed in the experimental δ-layer sample. The bottom two Si layers were kept fixed at bulk positions and the rest of the structure was optimised for each system. The structures were optimised, using a self consistency tolerance of $1 \times 10^{-7}$ and a force tolerance of 0.01eV/Å using the FIRE[196] molecular dynamics method. These simulations were performed using a Monkhorst-Pack[136] reciprocal space mesh of $6 \times 6 \times 1$ and a grid cutoff of 150 Ha. For consistency all simulations shown were performed using spin polarisation, for the cases where an odd number of P dopants are considered the additional electron is modelled occupying a spin up state.

In order to consider the influence of dopant depth and concentration, the 80 Si layer models including an explicit vacuum were expanded by a factor of two in the $x$ and $y$ dimensions to give 16 Si atoms per atomic layer and a total of 1328 atoms in the doped and undoped structures. In order to model systems of this scale the multi-site support function (MSSF) method was used[3]. These structure sizes allow for dopant densities of $1/16$ML to $1/4$ML and of course the extreme case of 1 ML where every Si atom in plane is substituted for a P atom however such a structure is entirely fictional. In addition the dopants were placed at different depths for the $1/16$ML structures considering δ-layers buried approximately 1, 2, 3 and 4 nm below the uppermost Si layer. The structures were again optimised, using a self consistency tolerance of $1 \times 10^{-5}$ and a force tolerance of 0.01eV/Å with the FIRE molecular dynamics. The Kerker preconditioning method[50] was also used in order to achieve self-consistency. These simulations were performed using a reciprocal space mesh of $3 \times 3 \times 1$ and a grid cutoff of 150 Ha.
6.3 Results and Discussion

6.3.1 The effects of vacuum

The presence of a group V δ-layer within Si has been observed both theoretically\cite{190, 188} and experimentally\cite{197} to result in the characteristic ‘band-splitting’ effect mentioned above. This effect manifests itself in the formation of impurity bands below the conduction band minima of the equivalent undoped structure. These nondegenerate states include what are often referred to as 1Γ and 2Γ as they are the lowest lying states at the Γ-point when observing the bandstructure. It is for this reason when considering the electronic structure our main focus, in this chapter at least, will be on the bandstructure.

For the 40 and 80 Si layer hydrogen passivated structures we consider both undoped and 1/4ML doped structure with and without a vacuum. The introduction of the dopant is observed to result in the aforementioned ‘band-splitting’ effect and this is illustrated in Figure 6.2 for the 1/4 ML dopant case. The bandstructure shown in Figure 6.2(b) is consistent with previous studies in the literature\cite{198, 185} which do not include the presence of an explicit vacuum in simulations.

The splitting effect results in the presence of multiple impurity bands. These are conduction bands pulled into the bulk band gap by the doping potential, two bands have minima at the Γ-point and are there referred to as 1Γ and 2Γ as previously mentioned, the third has a minimum approximately one third of the distance between the Γ and M high symmetry points, this is labelled as 1∆ in Figure 6.2. The 1Γ and 2Γ bands correspond to the out-of-plane minima of bulk silicon. All of the discrete bands are partially occupied. This bandstructure shows qualitative agreement with a previous conduction band model based on density functional theory utilising Wannier orbitals\cite{182}. The band splitting effect illustrated above is present for both the 40 and 80 layer structures with and without vacuum when a δ-layer is introduced. However the magnitude of the splitting varies depending on the structure considered, this is detailed in Table 6.1, where the ‘splitting’ values are given relative to the conduction band minimum(CBM), i.e. the lowest energy
6.3. Results and Discussion

Figure 6.2: The bandstructures for the doped and undoped 80 Si layer structures without an explicit vacuum along the Γ-M path. The doped structure (b) includes the band splitting with minima at the Γ and Δ points labelled 1Γ, 2Γ and 1Δ, the splitting from 1Γ to the CBM is indicated by a red arrow.
unoccupied state, of the equivalent undoped structure.

<table>
<thead>
<tr>
<th>Supercell</th>
<th>Energy Levels (eV)</th>
<th>40 L w/o vac.</th>
<th>40 L w. vac.</th>
<th>80 L w/o vac.</th>
<th>80 L w. vac.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_F$</td>
<td></td>
<td>-0.048</td>
<td>-0.132</td>
<td>-0.044</td>
<td>-0.108</td>
</tr>
<tr>
<td>$2\Gamma$</td>
<td></td>
<td>-0.511</td>
<td>-0.425</td>
<td>-0.508</td>
<td>-0.304</td>
</tr>
<tr>
<td>$1\Gamma$</td>
<td></td>
<td>-0.592</td>
<td>-0.495</td>
<td>-0.591</td>
<td>-0.377</td>
</tr>
</tbody>
</table>

**Table 6.1:** The electronic structure parameters, including the band-splittings at the $\Gamma$-point, of the 1/4 ML $\delta$-layer for Si(001) surfaces with and without vacuum (indicated as ‘w. vac.’ and ‘w/o vac.’ respectively). All values are given relative to the CBM of the corresponding structure.

For the examples considered here there is a clear influence of the vacuum on the ‘band-splitting’ values induced by the presence of a 1/4 ML P $\delta$-layer. The structures without an explicit vacuum have a ‘band-splitting’ between the $1\Gamma$ and $2\Gamma$ states at least 14% larger than the values given by the models with a vacuum. The calculated band minima for the structures without a vacuum show good agreement with values previously determined using the SIESTA code[198] which state $1\Gamma$, $2\Gamma$ values of 0.54 and 0.42eV respectively. For the models with a vacuum present we also see a much great dependence of the band splittings on the number of Si layers in the surface. The models without vacuum have similar band splittings for both the 40 and 80 layer structures. In addition the Fermi level for the doped structures is shifted much closer to the conduction band minimum for the structures without an explicit vacuum than those in which a vacuum is included. All further results are based upon structures with an explicit vacuum and will therefore reflect the influences of the vacuum of the electronic structure parameters.

### 6.3.2 The influence of dopant densities

The above examples show the influence of the presence of a vacuum and explicit surface interface, however they are limited to a fixed dopant density. In order to consider the influences of different dopant densities the simulation cells were expanded as outlined above. However the expansion was accompanied by problems in reaching a self consistent solution for the charge density matrix. It was only possible to achieve convergence at a self consistency tolerance of $1 \times 10^{-5}$, in an
6.3. Results and Discussion

6.3.2.1 Charge Sloshing and Kerker preconditioning

A common issue in DFT calculations of surfaces and semi-metallic systems is the so-called ‘charge-sloshing’ effect. This is due to an imbalance between the long range and short range distributions of charge perhaps caused by the presence of a vacuum or an area of high defect concentration or a combination of the two. This results in large amounts of charge being localised at extreme points of the simulation cell, rather than an even distribution, and this large amount of charge ‘sloshing’ around the cell as the simulation progresses. As a result it becomes difficult to obtain a self-consistent solution for the charge density, this is a new challenge that comes when DFT simulations of this nature are extended to these large scales. The simulation cells used to model P δ-layers in Si with an explicit vacuum are highly prone to this effect. The simulations are highly anisotropic given the metallic 2D layer parallel to the surface-vacuum interface as depicted in the schematic Figure 6.1.

As previously mentioned a way of addressing charge sloshing is to use the Kerker preconditioning method. This effectively damps the long range contributions of the charge by introducing an additional term into the density update process within the SCF cycle such that at each step of the cycle $i$,

$$
\rho_{in}^{i+1}(q) = \rho_{in}^{i}(q) + G\Delta\rho_{in}^{i}(q).
$$

(6.1)

where $\rho(q)$ is the Fourier transform,

$$
\rho(q) = \frac{1}{\Omega} \int \rho(r)e^{-iq\cdot r}d^3r,
$$

(6.2)

and $\Delta\rho$ is the change in charge density at each iteration of the SCF cycle, such that $\Delta\rho_{in}^{i} = \rho_{out}^{i-1} - \rho_{in}^{i}$ (see section 3.4 for further details). The Fourier transform is defined for a volume $\Omega$ and $G$ which is given by,

$$
G = \alpha - \frac{q^2}{q^2 + q_0^2}.
$$

(6.3)
6.3. Results and Discussion

<table>
<thead>
<tr>
<th>SCF Tolerance</th>
<th>CBM (eV)</th>
<th>$E_F$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 \times 10^{-5}$</td>
<td>-3.8811</td>
<td>-4.1303</td>
</tr>
<tr>
<td>$1 \times 10^{-6}$</td>
<td>-3.8811</td>
<td>-4.1304</td>
</tr>
<tr>
<td>$1 \times 10^{-7}$</td>
<td>-3.8811</td>
<td>-4.1304</td>
</tr>
<tr>
<td>$1 \times 10^{-8}$</td>
<td>-3.8811</td>
<td>-4.1304</td>
</tr>
</tbody>
</table>

Table 6.2: The CBM and Fermi energy values for the 20 layer H-passivated Si(001) converged using different SCF tolerance values.

In this equation $q_0$ acts as a screening parameter, the value determining the contribution of long distance charges. There is no ideal *a priori* choice for $q_0$, an educated guess of $q_0 \approx \sqrt{4k_f \pi}$ can be made using the Thomas-Fermi approximation. Similarly $\alpha$ is simply a damping parameter which is defined such that in the short range real space limit ($q >> q_0$) we recover a standard Pulay mixing approach\(^2\).

In the Conquest code the value of $q_0$ is determined by user input and values ranging from $1 \times 10^{-4}$ to 0.1 were tested for the P $\delta$-layer system. This was done in an attempt to achieve a self consistent solution for the charge density that had converged below a value of $1 \times 10^{-5}$, however these tests were unsuccessful.

As it had not proved possible to converge these simulations below a self consistency tolerance of $1 \times 10^{-5}$ the effect of the value of the tolerance on the conduction band minimum was investigated. To do so a smaller version of the systems outlined above was used, a 20 layer H-passivated Si(001) surface with 4 atoms per atomic layer and an explicit vacuum. This system was converged to using self consistency tolerances ranging from $1 \times 10^{-5}$ to $1 \times 10^{-8}$, the effect on the CBM is shown in Table 6.2.

As is illustrated the self consistency tolerance has no significant impact on the nature of the conduction band minimum while the Fermi energy only alters by the order of $1 \times 10^{-4}$ over the range of tolerances considered. As these are key properties of interest when investigating the band splitting associated with as P $\delta$-layer a pragmatic approach was taken and the structures containing more than 1000 atoms were converged to a tolerance of $1 \times 10^{-5}$, which although below the normal desired level of accuracy should not impact the results shown.

\(^2\)The short range limit in which Pulay mixing is recovered and methods of determining an optimum value of $\alpha$ are covered in greater detail elsewhere[199]
6.3. Results and Discussion

6.3.3 Dopant density effects on electronic structure parameters

The dopant densities considered were 1/16ML, 1/8ML and 1/4ML with the focus again being on the energy levels of the bands split from the CBM at the Γ point. In the standard case the dopants are uniformly distributed (or as close to uniform as possible for the 1/8ML case) in the x – y plane. For all of the dopant densities the δ-layer was placed at the midpoint of the cell at a depth of approximately 5.37nm from the uppermost Si atomic layer of the surface.

The results in Table 6.3 illustrate a dependency of the energy differences between

<table>
<thead>
<tr>
<th>δ-layer Dopant Density</th>
<th>Energy Levels (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E$_F$</td>
</tr>
<tr>
<td>1/16ML</td>
<td>-0.076</td>
</tr>
<tr>
<td>1/8ML</td>
<td>-0.092</td>
</tr>
<tr>
<td>1/4ML</td>
<td>-0.107</td>
</tr>
</tbody>
</table>

Table 6.3: The electronic structure parameters, including the band-splittings at the Γ-point, of the 1/4 ML, 1/8 and 1/16ML δ-layers. All values are given relative to the CBM of the corresponding structure

the 1Γ and 2Γ states on the dopant concentration with the energy difference broadening as the dopant concentration increases, in an approximately linear relationship. This pattern is consistent with previous effective mass calculations performed by Scolfaro et al[187]. Furthermore there is a monotonic increase in the separation of the Fermi level from the CBM with increasing dopant concentrations up to the maximum 1/4ML concentration considered here which is again consistent with previous effective mass calculations [186]. However it has been noted in previous tight binding simulations[183] that at higher dopant concentrations the behaviour of the Fermi level will no longer be monotonic and be more oscillatory in nature, nonetheless these concentrations have not been considered here due to the difficulty associated with realising such structures experimentally.

6.3.4 Surface effects on electronic structure parameters

In order to consider the influence of surface proximity on the ‘band-splitting’ an ideal 1/16 ML δ-layer was simulated at depths of 1.088, 2.101, 3.132 and 4.210nm
below the uppermost Si atomic layer. The 1/16 ML δ-layer was chosen as it has the lowest concentration of dopants and is least likely to distort the cell as moved in the z-dimension, isolating any changes in the electronic structure parameters to surface effects. The surface consists of 80 Si layers and is approximately 10.88 nm in depth, the base of the surface should have negligible effect on the δ-layers in these models. The results of these simulations are shown in Table 6.4.

| δ-layer Dopant Depth (nm) | $E_F$ | $1\Gamma$ | $2\Gamma$ | $|1\Gamma - 2\Gamma|$ | $|E_F - 1\Gamma|$ | $|E_F - 2\Gamma|$ |
|--------------------------|-------|-----------|-----------|----------------|----------------|----------------|
| 1.088                    | -0.076| -0.146    | -0.118    | 0.028         | 0.070          | 0.042          |
| 2.101                    | -0.083| -0.153    | -0.128    | 0.025         | 0.070          | 0.045          |
| 3.132                    | -0.083| -0.154    | -0.125    | 0.029         | 0.071          | 0.042          |
| 4.210                    | -0.075| -0.139    | -0.111    | 0.028         | 0.064          | 0.036          |

Table 6.4: The electronic structure parameters, including the band-splittings at the $\Gamma$-point, of the 1/16ML δ-layer at depths of approximately 1, 2, 3 and 4 nm in Si(001). All values are given relative to the CBM of the corresponding structure.

Interestingly, the Fermi level does appear to oscillate with the depths of the δ-layer as do the $1\Gamma$ and $2\Gamma$ energy levels themselves. However, the energy differences between the $1\Gamma$ and $2\Gamma$ states do not vary by significant amounts, ranging from a minimum of 0.025 eV to a maximum of 0.029 eV. This represents a large percentage change in the values of the ‘band splitting’ but only an absolute change of the order of $1 \times 10^{-3}$ eV. Furthermore, there is similar variation in the energy difference between the Fermi level and the $1\Gamma$ and $2\Gamma$ states, with the absolute differences varying by no more than 0.009 eV over the range of depths considered. In addition, there appears to be no obvious trend between the depths considered and the energy levels. This data would therefore suggest there is little dependence on the ‘band splitting’ values associated with a δ-layer and the proximity of that δ-layer to a surface interface.

### 6.3.5 Disordered δ-layers

In addition to the effects of dopant densities and surface proximities, the influence of disordered dopants was considered for the 1/4ML case. Intentionally or otherwise
the PALE process may not result in the uniform distribution of dopants modelled above. There may be disorder in the x-y plane with dopants not distributed equidistantly, to simulate this the 1/4ML model at a depth of 5.37nm is modified so that the dopants are distributed randomly. A schematic of both cases, and plots of their bandstructures, are shown in Figure 6.3. As well as the disorder in the x – y plane

![Figure 6.3](image)

Figure 6.3: The bandstructures and 2D atomic schematics in the x-y plane for ordered and disordered 1/4ML δ-layers in models including an explicit vacuum. P atoms are shown in red with the grey shaded areas showing the repeat unit of the periodic cell.

there is also the possibility that the dopants may be disordered in the z dimension. Although the proximity of an ideal δ-layer to a surface does not have a significant influence the structure parameters, as exhibited by the data above, the case of a
non-ideal layer should also be considered. This is simulated by modelling dopants deposited at different depths with respect to the surface, which is equivalent to a non-ideal $\delta$-layer that may result from diffusion effects during Si overgrowth in the final stage of the fabrication process. The simulation cell is based upon the same Si surface as the calculations for different dopant densities and utilises the same convergence criteria. In order to mimic diffusion as 1/4ML dopant density, comprising of four individual P dopants, has been modelled with each dopant place at a different depth relative the uppermost Si-vacuum interface, this is depicted in Figure 6.4. The dopants were placed at depths of 4.53, 4.80, 5.08, 5.36nm, so that the $\delta$–layer extends over a range of 0.83nm of the z-dimension in the Si surface, with the dopants at approximately equal separations. This is an extreme level of disorder, presented here as a limiting case.

The energy levels for the Fermi energy, $1\Gamma$ and $2\Gamma$ states for the disordered models are included in Table 6.5, the energy levels of the same states for the ideal
6.4 Conclusions

The electronic structure parameters, including the band-splittings at the Γ-point, of the 1/4 ML extended in the z-dimension labelled diffused and the 1/4 ML case with randomly distributed dopants in the x−y plane labelled disordered. The values for the ideal 1/4 ML are included for reference.

| δ-layer Dopant Density     | $E_F$  | 1\(\Gamma\) | 2\(\Gamma\) | $|1\Gamma−2\Gamma|$ |
|----------------------------|--------|-------------|-------------|-----------------|
| 1/4ML                      | -0.107 | -0.381      | -0.304      | 0.077           |
| 1/4ML disordered           | -0.104 | -0.364      | -0.307      | 0.057           |
| 1/4ML diffused             | -0.097 | -0.283      | -0.234      | 0.049           |

\(\delta\)-layer are given here again for reference. As is shown there is a clear reduction, of 0.028 eV in the energy differences between the 1\(\Gamma\) and 2\(\Gamma\) states for the \(\delta\)-layer extended in the z-dimension in comparison to the ideal case. This indicates that the separation of the dopants has a more significant effect on the electronic structure parameters than the separation of the \(\delta\)-layer and the surface vacuum interface.

There is a similar result for disorder in the x−y plane with a clear narrowing, a 0.020 eV decrease, of the energy difference between the 1\(\Gamma\) and 2\(\Gamma\) states. For both cases the result is more than likely to be reflected in experimental data as unlike the simulations here manufactured \(\delta\)-layers will not be ideal with disordering in x−y plane and in depth likely to be present. If time allowed it would be worth further investigating the effects of diffusion on different dopant densities as well as a greater range of separations between the dopants in the z-dimension. Furthermore a greater range of dopant separations in the x−y plane could be considered, to see if an exact relationship between these separations and the electronic structure parameters can be determined.

6.4 Conclusions

It has been shown that the predictions of DFT simulations for the electronic structure parameters associated with the ‘band splitting’ effect induced by the presence of Group V \(\delta\)-layers in Si surfaces are dependent on the presence of an explicit vacuum in a simulation cell. Given that result it would be of interest to further extend the simulations to include an oxide interface as this will more closely model the experimental reality and may have an impact on the electronic structure param-
eters. Furthermore it has been shown that the magnitude of the energy splittings of states below the conduction minimum are dependent on both the dopant density and level of the disorder of dopants within a $\delta$-layer. It has also been shown that the proximity of a $\delta$-layer to an interface has little effect on the electronic structure of the surface as a whole. Although given more time it may be worthwhile exploring a greater range of depths or perhaps interactions between multiple embedded $\delta$-layers which could be modelled by simply extending the dimensions of the simulation cells considered here.
Chapter 7

Conclusions and Outlook

In this thesis, the properties of several doped systems in silicon and silicon nanostructures have been studied using large scale DFT simulations, while presenting some of the difficulties faced for the simulation sizes considered. This work has tackled both theoretical issues relating to the accuracy and efficiency of the calculations performed, while presenting novel results for doped structures which are of technological relevance given the growing interest in solotronics and precision doping. To best further the understanding of such devices, a combination of both theoretical and experimental studies will be required. Hence, we have attempted to relate this work to experimental studies whenever possible. While only experimental studies can investigate the full complexity of real silicon based devices, theoretical work, such as the studies in this thesis, is vital when considering structural details as part of the more complex whole which are beyond the resolution of experimental methods. The usefulness of such theoretical and computational methods have been demonstrated in the studies of the following systems; (i) self compensation effects when considering pairs of group III dopants in Si for the case of Al, (ii) the structural complexes formed by the group V dopant Bi in Si and (iii) the electronic structure of P δ-layers in Si surfaces with a particular emphasis on surface interface effects.

For the study of self compensation of Al dopants in Si we considered interactions of pairs of Al using a combination of the electron localisation function and visualisations of the densities of states. It was found that for adjacent Al dopants a bond
is formed between the dopants when they are in the $S = 1$ spin state, whereas no such bond is formed in the $S = 0$ spin state. This finding is supported by a high ELF value in the region between the two acceptors, a non-negligible charge density in the same region and a separation of the Al dopants which is consistent with the Al-Al bond lengths presented in previous force field and DFT simulations. This presents the possibility of an electrically inactive dopant complex with tunable properties if a combination of magnetic fields and optical pumping are used to switch between the $S = 0$ and $S = 1$ states. It would be of interest to see if such tunable complexes are formed when considering other commonly used group III dopants, for example B, and this would provide a straightforward extension to the study.

For the case of the group V dopant Bi in Si it was found that structural complexes incorporating a vacancy in close proximity to the substitutional dopant are more energetically stable. However, complexes incorporating multiple vacancies were found to be intrinsically comparatively unstable. Furthermore the difficulties in extending such structures to large scale simulations, of the the order of thousands of atoms, were highlighted. Notably issues were encountered in processes as simple as atomic relaxations. Nonetheless many possibilities for extending the study were explored, including an accurate consideration of the net charge of the structure.

The study would also lend itself well to the recently developed ab initio random structure searching techniques. Another logical extension would be to consider the structure of Bi dopants under the surface, for which simulated scanning tunnelling microscopy images could be used to aid interpretation of their experimental counterparts and it would be of interest to discover if the relative stabilities within the bulk extend to the surface.

The CONQUEST code has already been shown to perform accurately for surface simulations incorporating Group V dopants and including and excluding an explicit vacuum. In Chapter 6 it was shown that the explicit presence of a vacuum in a surface simulation cell has an influence of the electronic structure parameters of P $\delta$-layers, an influence which is observed in the characteristic ‘band-splitting’ effect observed for such structures. Furthermore, it was shown that disorder of the $\delta$-
layer, either as a consequence of diffusion or the efficacy of dopant deposition, has a similar effect on the electronic structure parameters, an observation that may aid the interpretation of ARPES data produced by experimental teams based at UCL. The effects of the vacuum and disorder had not been explicitly considered in previous DFT studies of these structures and it would not have been possible without the efficiency of the local orbital based code that was used, notably the multi-site support function (MSSF) feature. A feature that will be useful in further extensions of this work which can be used a greater range of dopant disorder, dopant densities and \( \delta \)-layer depths.

It is also possible that these simulations may be performed using the recently developed onsite support function (OSSF) methodology outlined in Chapter 3, a modification of the MSSF technique utilised successfully within this thesis. It was shown the OSSF method can be used in conjunction with linear scaling techniques allowing for calculations containing potentially million of atoms to be performed more rapidly and with more accurate basis sets than had been previously possible. Further methodological developments were shown in this chapter including the definition of alternative residual used in the self consistency process. These alternative definitions of the residual that are independent of (and more dimensionally consistent than) the Frobenius norm has been used to develop a robust method of achieving a precise and system independent level of self consistency. This allows for the use of a universal self consistency tolerance across different scales of the process, simplifying calculations for any future user and could straightforwardly be adapted to other DFT codes.

These methodological developments will enable some of the numerous extensions suggested here to the studies outlined in this thesis so that the novel findings included here can be built upon. Although the findings within this thesis are incremental in nature it is hoped that the results have been of some use to experimental collaborators and may be used as a point of reference for future theoreticians. Furthermore, the applications that may emerge in part using this work and the extensions of this work will encompass a wide range of topics from quantum information
processing to single electron transistors and precision doping. On the whole, it has been a fascinating experience to study dopants within silicon nanostructures through the prism of precision doping methodologies, a field which may be well trodden but still contains exciting Physics to be further explored.
Appendix A

Appendix A - OSSF Supplementary Information

A.1 The Choice of Filtration Function

In Chapter 3 it was stated that the choice of the ‘filtration function’ within the local filter diagonalisation (LFD) process, which removes high energy components of the trial vectors in order to construct the contracted basis within the OSSF method, is insignificant. The filtration process will perform equally with any step-like function used to define the filtration function. In order to illustrate the arbitrariness of the filtration function we return to the case of simple 8 atom crystalline silicon.

We compare the groundstate energies produced by using the contracted basis set for two different definitions of the filtration function, namely one based upon the Fermi-Dirac (FD) function (as is the case for all results shown within Chapter 3) such that,

\[ f_{FD}(\varepsilon_S) = \frac{1}{e^{\varepsilon_S - \mu/kT} + 1} \]  

(A.1)

and a filtration function based upon the tanh function which gives,

\[ f_{tanh}(\varepsilon_S) = \frac{\tanh(\varepsilon_S - \mu/kT) + 1}{2} \]  

(A.2)

The simulations are performed using a $6 \times 6 \times 6$ Monkhorst Pack Mesh and
A.1. The Choice of Filtration Function

Table A.1: The binding energies for 8 atom Si using the onsite support functions constructed using the FD and tanh functions for the DZDP and TZTP basis sets

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>Primitive NSF Energy (Ha)</th>
<th>OSSF FD Filt. Energy (Ha)</th>
<th>OSSF Tanh Filt. Energy (Ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DZDP</td>
<td>-33.820683</td>
<td>-33.801264</td>
<td>-33.801227</td>
</tr>
<tr>
<td>TZTP</td>
<td>-33.834045</td>
<td>-33.165981</td>
<td>-33.165967</td>
</tr>
</tbody>
</table>

A 150 Ha energy cutoff for the integration grid and converged to a self consistency tolerance of $1 \times 10^{-6}$.

As is shown in Table A.1 the difference between the groundstate energies produced using the FD function and tanh function as filtration functions of $1 \times 10^{-5}$, illustrating for the case of 8 atom Si at least that the choice of filtration function has little impact on the final groundstate energy.
Appendix B

Appendix B - P δ-layers

Supplementary Information

In order to determine the choice of basis set for the P δ-layer simulations several factors were taken into account. Firstly, the smallest possible basis set was required in order to reduce the computational cost as much as possible for the simulations not using the MSSF method. Secondly, the bandstructure needed to be qualitatively correct, which is the case for all basis sets larger than and including the SZP basis. Lastly, the accuracy of the basis set needed to be verified, in order to do so an 8 atom Si bulk cell was modelled using a $12 \times 12 \times 12$ Monkhorst-Pack mesh and a 150 Ha grid cutoff. The lattice constants ($a_0$), bulk modulii ($B_0$) and band structure parameters including the band gap $E_{\text{gap}}$ and the energy of the conduction band minimum at the $\Gamma$-point (denoted here as $E_{\Gamma}$) were compared for the primitive SZP, DZP and TZTP basis sets. In addition we compare to the experimental values of the lattice constant[139], bulk modulus[200] and bandstructure parameters[201].

Table B.1: Properties of 8 atom Bulk Si cell for a range of basis set qualities, all electronic structure parameters are given relative to the valence band maximum

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>$a_0$(Å)</th>
<th>$B_0$(GPa)</th>
<th>$E_{\text{gap}}$(eV)</th>
<th>$E_{\Gamma}$(eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SZP</td>
<td>5.541</td>
<td>82.81</td>
<td>0.85</td>
<td>2.51</td>
</tr>
<tr>
<td>DZP</td>
<td>5.458</td>
<td>90.96</td>
<td>0.64</td>
<td>2.60</td>
</tr>
<tr>
<td>TZTP</td>
<td>5.437</td>
<td>91.70</td>
<td>0.62</td>
<td>2.75</td>
</tr>
<tr>
<td>Expt.</td>
<td>5.431</td>
<td>102</td>
<td>1.12</td>
<td>3.40</td>
</tr>
</tbody>
</table>
The results are shown in Table A.2.
Given the accuracy of the bulk module and lattice constant values produced by the DZP basis, with the latter varying by less than 1% from the experimental value the choice was made to use the DZP basis for the P δ-layer simulations. This decision was aided by the strong agreement of the electronic lattice parameters of the TZTP and DZP basis sets, and the fact that the latter incorporates less than half the number of support functions of TZTP (13 compared to 27) which represents significant savings in computational time.
Bibliography


[177] Frank J. Ruess, Lars Oberbeck, Michelle Y. Simmons, Kuan Eng J. Goh, Alex R. Hamilton, Toby Hallam, Steven R. Schofield, Neil J. Curson, and


