Structure and Mechanisms of Formation of Point defects in HfO$_2$, MgO and hexagonal boron nitride

Jack William Strand

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

Department of Chemistry
University College London

June 4, 2019
I, Jack William Strand, 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.
Abstract

In this thesis, density functional theory (DFT) methods are used to model a range of defects and defect processes in three functional dielectric materials: MgO, HfO$_2$ and hexagonal boron nitride (hBN).

My results demonstrate that a novel implementation of time-dependent DFT in CP2K provides accurate optical properties of oxygen vacancies in MgO.

In amorphous (a-) HfO$_2$ the existence of self-trapped holes and electrons is predicted. These trapped states are found to be energetically deeper than their crystalline counterparts and are separated, on average, by 8 Å. Calculated optical spectra of electron traps agree well with exhaustive photo-depopulation spectroscopy experiments.

It is then shown that the average formation energies of oxygen vacancies and interstitials in a-HfO$_2$ are close to those in monoclinic (m-) HfO$_2$, however they follow a distribution. My calculations of optical spectra of oxygen vacancies in a-HfO$_2$ demonstrate that the characteristic blue luminescence of HfO$_2$ is likely due to the oxygen vacancy in its +2 state. It is also found that a 3.6 eV luminescence can be caused by a radiative tunnelling transition between a hole and a +1 charged oxygen vacancy.

Next, oxygen Frenkel pairs (FPs) in HfO$_2$ are studied. A barrier of 6.6 eV must be overcome to generate a FP in m-HfO$_2$. Charging (by injection of electrons) decreases this barrier by over 4 eV. Similar barrier reduction, due to carrier localisation, is found for FP generation in a-HfO$_2$, however with a broader energy distribution. It is demonstrated that both formation energies and barrier heights are reduced when FPs are created adjacent to vacancies or vacancy clusters.
Finally, a range of intrinsic defects in hBN layers are modelled. In particular, it is predicted that divacancies stabilise N-N and B-B bridges between layers. These bridges sufficiently lower formation energies of Frenkel pairs.
Acknowledgements

There are a number of people I would like to thank for their support throughout my doctoral research. Without them this work would not have been possible.

First and foremost, I would like to thank my supervisor Prof. Alex Shluger for his dedicated mentoring and supervision. I would also like to thank all the members of the Shluger group, who have provided me with invaluable discussions and assistance throughout my doctoral research. In particular I would like to thank David Gao, Thomas Durrant and Julian Gaberle. There are also many previous members of the group from whom I received a great deal of helpful advice, in particular Al-moatasem El-Sayed, Oliver Dicks and Mathew Watkins.

I would also like to thank the EPSRC and the M3S Doctoral Training Center for their support and sponsorship. I would also like to thank Sematech and Gennadi Bersuker for support and a very many helpful discussion. I have also made extensive use of the UK National Supercomputing Service, Archer, which made this work possible.

I am also grateful to Moloud Kaviani, who produced the original models of amorphous hafnia, and who conducted the initial work on electron trapping.
Impact Statement

This research project was motivated by a collaboration with Sematech and the desire to understand the behaviour of hafnium dioxide based microelectronic devices. It later expanded outwards to do similar investigations on other materials. The results presented in this thesis are therefore the output of three main projects which aimed to investigate defect properties in three materials. The commonality between all these materials is that they are wide band gap insulators and have practical application in nanoscale devices. All throughout this thesis, modelling techniques are use to simulate, and make predictions about, defects. These modelling techniques are often combined with experimental investigation, and help to provide new understanding and insight that experimental characterisation could not achieve alone.

In the first Chapter, the optical absorption spectra of oxygen vacancies in MgO are studied. Optical absorption spectra can help to characterise and identify defects in real samples. Furthermore, this chapter verified the reliability of a novel implementation of time dependent density functional theory in the CP2K software package. This should therefore be of use to other members of the materials modelling community who may wish to do similar calculations on other systems.

In the following three chapters, the defects of hafnium dioxide are studied. Since hafnium dioxide is a commonly applied gate dielectric, and since defect related processes are responsible for its degradation and reliability issues, modelling of defects in hafnium dioxide is of paramount importance. The results of these three chapters demonstrate that, when in its amorphous phase, hafnium dioxide can trap electrons and holes into very deep, stable and localised states. These states then play a role in the dielectric breakdown of hafnium dioxide, a major reliability is-
Acknowledgements

sue in microelectronics. Understanding the mechanism of dielectric breakdown is extremely important if we are to continue to develop microelectronic technology. Furthermore, the breakdown mechanism presented, and the discovery of these deep trap states, may be fairly general and not specific to hafnium dioxide. The impact of these results may, therefore, be even wider.

2D materials have seen an explosion of interest in recent years, mostly owing to the excitement about the properties of graphene. Hexagonal boron nitride is a layered material, similar in structure to graphite, and can also be reduced down to a 2D form which is iso-structural to graphene. It also has a wide band gap, and is therefore being investigated as a dielectric to be applied in 2D nanoscale devices. In the final chapter, a wide range of defects in hexagonal boron nitride are studied. The results of this chapter indicate which defects could produce reliability problems in devices due to the charging of the dielectric film. These instabilities have long been known to exist in other devices, and understanding their origin is a crucial step in producing the future generation of nanoscale transistor technology.
2.3.3 Auxiliary Density Matrix Method .................................. 49

2.4 The PBE0-TC-LRC Functional ........................................ 52

2.5 Time Dependent Density functional theory .......................... 53
  2.5.1 Runge-Gross Theorem .............................................. 54
  2.5.2 The Time Dependent Kohn-Sham Equations ...................... 54
  2.5.3 Time-Dependent XC Functionals - The Adiabatic Approximation .............................................. 55
  2.5.4 Implementation of TDDFT in CP2K ............................... 55

2.6 Practical Considerations in DFT Calculations ..................... 56
  2.6.1 Periodic Boundary conditions .................................... 56
  2.6.2 Defect Calculations ................................................ 58
  2.6.3 Charged Defect Corrections ....................................... 59
  2.6.4 Potential Alignment ................................................ 61
  2.6.5 Nudged Elastic Band Calculations .............................. 61

3 Oxygen Vacancies in MgO ................................................. 63
  3.1 Introduction & Background ........................................... 63
    3.1.1 Previous Studies on Oxygen Vacancies in MgO ................. 63
    3.1.2 Details of Calculations ......................................... 66
  3.2 Results .............................................................. 67
    3.2.1 Bulk MgO ......................................................... 67
    3.2.2 Geometric And Electronic Structure of Oxygen Vacancies in MgO .............................................. 69
    3.2.3 Optical Absorption of O Vacancies .......................... 70
    3.2.4 Photoluminescence of O Vacancy ........................... 73
  3.3 Conclusions .......................................................... 74

4 Carrier Trapping in Pristine HfO$_2$ .................................. 77
  4.1 Introduction & Background .......................................... 77
    4.1.1 Previous Calculations for Carrier Trapping in Monoclinic HfO$_2$ .............................................. 78
## Contents

4.1.2 Modelling the Structure of Amorphous HfO$_2$ .......................... 79
4.2 Results .................................................................................. 82
  4.2.1 Electronic Structure of Bulk HfO$_2$ ........................................ 82
  4.2.2 Trapped Holes in HfO$_2$ ...................................................... 86
  4.2.3 Double Hole Traps .............................................................. 90
  4.2.4 Optical Excitations of Electron Traps in HfO$_2$ ..................... 93
  4.2.5 Conclusions ....................................................................... 97

5 Oxygen Vacancy and Interstitial defects in amorphous HfO$_2$ 99
  5.1 Introduction & background ...................................................... 99
    5.1.1 Previous studies on O vacancies in HfO$_2$ ......................... 100
  5.2 Results .................................................................................. 101
    5.2.1 Vacancy Defects in the monoclinic phase ......................... 101
    5.2.2 Vacancy defects in the amorphous phase ......................... 103
    5.2.3 Optical absorption and luminescence of O vacancies ......... 108
    5.2.4 Interstitial defects in the monoclinic phase ..................... 110
    5.2.5 Interstitial defects in the amorphous phase ..................... 112
  5.3 Summary & Conclusions ......................................................... 113

6 Vacancy generation in HfO$_2$ ................................................... 115
  6.1 Introduction & Background ................................................... 115
    6.1.1 Previous Work on Vacancy Generation in HfO$_2$ ............. 117
  6.2 Results .................................................................................. 118
    6.2.1 Frenkel pair formation in m-HfO$_2$ .................................... 118
    6.2.2 Frenkel pair formation in a-HfO$_2$ .................................... 122
    6.2.3 Interaction Between Oxygen Vacancies ......................... 124
  6.3 Conclusions .......................................................................... 127

7 Hexagonal Boron Nitride - Structure and Defects 130
  7.1 Previous Studies on hBN ....................................................... 131
  7.2 Methodology ....................................................................... 133
  7.3 Results .................................................................................. 135
7.3.1 Bulk hBN .................................................. 135
7.3.2 The Stone-Wells Defect ................................. 137
7.3.3 Vacancy Defects in hBN ............................... 138
7.3.4 Divacancies in hBN ................................... 140
7.3.5 Interstitials Defects in hBN ........................... 144

7.4 Conclusion .................................................. 147

8 General Conclusions and Future Work ....................... 149

Bibliography ..................................................... 152
List of Figures

1.1 A schematic of a n-MOSFET. A p-doped substrate of silicon contains two (highly) n-doped regions, termed the source and the drain. If a voltage, $V > V_{Th}$ is applied to the gate, this produces an inversion layer in the substrate between the source and the drain regions, termed the ‘n-channel’. This increases the conductance between the source and the drain. .................................................. 28

1.2 The impact of defects on the function of a dielectric in a MOSFET device. Defects, indicated by red dots, can be charged. Charging can affect the conduction through the n-channel. Also, electron trapping at a defect allows for trap assisted tunnelling (TAT) through the layer. This produces a small current through the insulating layer. If the defect concentration is high enough, a ’conductive filament’ can exist, which conducts electrical current through the layer. Such filaments are responsible for dielectric breakdown (see Chapter 6). . 30

3.1 (A) Experimentally measured optical absorption spectra of MgO. The peaks corresponds to the first optical absorption peak of the $V^0_O$ and $V^{+1}_O$ defect. (B) Optical absorption results with the separate peaks more clearly resolved. Irradiation causes the absorption line to shift down in energy slightly, as a results of the $V^0_O \rightarrow V^{+1}_O$ photoconversion process. Both images are retrieved from [1]. . . . . 65
3.2 Schematic of the typical optical transitions in oxides with defects. Type-I is a band-to-band transition. Type II is a VB to defect-induced unoccupied resonant state in CB transition. Type III represents the VB to unoccupied defect state in the bandgap transition. Type IV is an occupied defect state into defect-induced state in the CB transition, and type V is an occupied defect state into unoccupied defect state in the bandgap transition. A type VI transition is from a shallow, occupied defect state into the defect-induced resonant state in the CB.

3.3 The $\Delta$-SCF method in DFT. One begins in the singlet ground state. Then the triplet is calculated. This corresponds to the vertical transition energy at stage 1. The nuclear coordinates are then optimised while the electronic wavefunction is kept in the triplet state. Once fully relaxed, the system is again recalculated in a singlet state, but with the geometry held constant. This corresponds the the vertical transition energy at stage 3. This final vertical transition energy gives an approximation of the luminescence energy.

3.4 (A) The bulk 1000 atom MgO simulation cell. MgO has a face centred cubic (FCC) crystal structure with a lattice parameter of 4.21 Å. (B) The projected density of states (PDOS) of the 1000 atom MgO bulk simulation cell. The calculated CBM state is 7.8 eV above the VBM, however the density of this state is too small to be seen in the PDOS plot.

3.5 (A) The $V^0_O$ defect in MgO. Blue surface indicates the iso-surface of the KS defect state. (B) The $V^{+1}_O$ defect in MgO. The top image shows the occupied KS defect state (blue iso-surface). The bottom image shows the unoccupied KS defect state (orange iso-surface). (C) The $V^{+2}_O$ defect in MgO. The orange surface indicated an iso-surface of the unoccupied KS defect state. All iso-surface values are 0.001.
3.6 An isosurface of the $|\psi|^2$ of the localised states introduced into the conduction band by the oxygen vacancy. The states have a $p$-like character. Iso-value = 0.004. . . . . . . . . . . . . . . . . . . . . . 71

3.7 Optical absorption spectra of oxygen vacancies in MgO. (A) is the $V^{0}_O$ defect in MgO. There is strong absorption peaking at 4.85 eV. This corresponds to a transition between an $s$-like and a $p$-like state. (B) is the $V^{+1}_O$ defect in MgO. The lower energy absorption peaks at 4.70 and comes from excitation of an alpha spin electron in a gap-state into CB states. The higher energy absorption peaks at 5.26 eV and comes from excitation of beta spin electrons in VB states into the unoccupied state in the gap. (C) is the $V^{+2}_O$ defect in MgO. Blue delta functions are the original, untampered excitation energies. The red lines give the spectrum smoothed out with Gaussians of standard deviation $\sigma=0.2$ eV. . . . . . . . . . . . . . . . . . . . . 72

3.8 The $|\psi|^2$ of the high energy triplet state of the $V^{0}_O$ defect. (A) gives the HOMO-1 state, visualised from each of the crystallographic directions. (B) shows the HOMO state, shown from the same crystallographic directions. All iso-surfaces have value = 0.004. . . . . . . . . . . . . . 74

3.9 The local modes associated with the $s \rightarrow p$ transition of the $V^{0}_O$ defect in MgO. Orange circles represent the NN Mg ions. (A) Is the displacement of the high energy triplet state. There are extensions along the 3-fold rotation axis of the vacancy. The displacements are formed from a a linear combination of the $A_{1g}$-symmetric and $T_{2g}$-symmetric local modes. (B) The local mode of $A_{1g}$ symmetry. (C) The three local modes of $T_{2g}$ symmetry. . . . . . . . . . . . . . . . 75

4.1 A schematic of the process used to generate models of a-HfO$_2$. The melting of the crystal is simulated, followed by a rapid cooling. This produces a periodically repeated amorphous cell, which is then reoptimised using a DFT approach. . . . . . . . . . . . . . . . . 80
4.2 The O-Hf radial distribution functional (RDF) of atomistic models of the tetragonal (t-) and amorphous (a-) phase of HfO$_2$. The tetragonal phase has sharp and distinct lines, characteristic of crystals. The amorphous phase has a much wider spread of O-Hf distances. 81

4.3 (A) Shows that 324-atom periodic simulation cell of m-HfO$_2$ with the VBM state indicated by orange surfaces. Oxygen ions are coloured red and Hf ions are coloured cyan. The VBM is composed of $p$-like states from the 3C oxygen sublattice. (B) Shows the CBM (blue transparent surface). The CBM has predominant Hf $d$-character. The perspective has been aligned such that layers of 3C and 4C oxygen ions (which have been indicated) are alternating. Iso-surfaces of both images is 0.015. 83

4.4 (A) The VBM (orange surface) of tetragonal (t-) HfO$_2$. The lack of translational symmetry indicates that the VBM is not at the $\Gamma$-point. (B) Shows the CBM state (blue) of t-HfO$_2$. The iso-surface in both images is 0.01. 84

4.5 (A) A typical 324-atom model of a-HfO$_2$. The transparent orange surface shows an iso-surface (of value = 0.01) of the VBM state. Disorder causes the band edge states to become semi-localised around precursor structures. (B) The IPR spectrum of a typical 324-atom model of a-HfO$_2$. High values of IPR indicate localisation, low values indicate delocalisation. Amorphisation leads to the localisation of band edge states. 85

4.6 The IPR spectrum of monolinic HfO$_2$. Compared to the amorphous IPR, the crystal IPR spectrum shows no significant build of localisation at the band edges. 86
4.7 (A) The 3-coordinated oxygen ion and its three hafnium neighbours in monoclinic HfO$_2$. Bond lengths (in Å) are indicated. (B) The 3C self-trapped hole in m-HfO$_2$. A single hole state, shown by the orange iso-surface (value = 0.025) is trapped onto an oxygen ion, causing extension of the O-Hf bonds. 

4.8 (A) The geometry of two adjacent oxygen atoms and their nearest neighbour hafnium ions in t-HfO$_2$. (B) The local geometry of the STH in t-HfO$_2$. Square modulus of the hole state is shown by the orange iso-surface (iso-surface value = 0.025). Bond lengths, in Å, are indicated. Direction of the major displacements are indicated by arrows.

4.9 (A) The local atomic structure around a typical precursor in a-HfO$_2$, with bond lengths indicated. (B) The local configuration of a hole trap at the same site after trapping and relaxation. Most of the O-Hf distances have been increased due to Coulomb repulsion. The $|\psi|^2$ of the hole state is indicated by the orange iso-surface (iso-surface value=0.015). All bond lengths are given in Å.

4.10 Trapping energy distribution of electron and hole traps.

4.11 (A) The double hole trap in m-HfO$_2$. Bond lengths (in Å) are indicated. The transparent orange surface shows an iso-surface (iso-surface value=0.01) of the $|\psi|^2$ of the double trap state. (B) is the projected density of states of a m-HfO$_2$ simulation cell containing the double hole trap. The state can be seen high in the band gap.

4.12 A schematic of the ideal process which defines the binding energy. The binding energy is defined as the energy released when a single self-trapped hole moves and traps onto a second self trapped hole, with which it was well separated. The products of this reaction are a double STH and a region of pristine m-HfO$_2$, which are infinitely separated.
4.13 The local structure of a typical double STH in a-HfO$_2$. The orange transparent surface indicates the $|\psi|^2$ of the (double occupied) hole state (iso-surface value=0.015). Major structural distortions are caused and an O-O dimer, with a bond length of 1.46 Å, is formed. All bond lengths are indicated in Å. 93

4.14 (A) is the intrinsic electron trap in a-HfO$_2$. Ionic displacements are indicated by arrows and given in Å. The blue surface is an iso-surface of the $|\psi|^2$ of the STE KS state. (B) is the double electron trap in a-HfO$_2$. Both iso-surface values are = 0.005. Images reproduced with permission from [2]. IOP Publishing Ltd. 94

4.15 The energy distribution of electron traps in a-HfO$_2$. The top panel shows the predicted optical absorption energies of single and double electron traps based on TDDFT modelling. The bottom panel shows the results of EPDS measurements. Both spectra show a distribution centered around approximately 3 eV. Images reproduced with permission from [2]. IOP Publishing Ltd. 96

5.1 The 3C oxygen vacancy in m-HfO$_2$. Hf ions are colored in cyan and oxygen ions a colored in red. Each configuration is labeled with its charge state. The translucent surfaces give the $|\phi|^2$ of the defect orbitals. Where applicable, the higher eigenvalue state (closer to the CBM) has been colored in yellow, and the lower valued eigenvalue state (closer to the VBM) has been coloured in blue. All iso-values = 0.005. The energy of these states are indicated in Fig. 5.2. 102

5.2 The KS states of the oxygen vacancy in m-HfO$_2$. Levels are shown (in eV) relative to the CBM. 103

5.3 Formation energy diagram of the oxygen vacancy in HfO$_2$ in its 5 charge states. Crossing points indicate the charge transition levels. 104
5.4 The formation energy of oxygen vacancies in amorphous HfO$_2$ as a function of the system Fermi level. The values are similar to those of the monoclinic phase. Unlike crystalline phases, however, there is a ‘width’ to each line. This width corresponds to the spread of formation energies possible due to the statistical nature of amorphous systems.

5.5 (A) Histogram of the set of neutral oxygen interstitial defect configurations analysed in the a-HfO$_2$ models. (B) Histogram of the set of the neutral oxygen vacancy formation energies analysed in the in a-HfO$_2$ models.

5.6 The oxygen vacancy in amorphous HfO$_2$. The charge state of each configuration is indicated. Since it is an amorphous system, each O vacancy is unique, and so only one configuration among many is shown. Isosurfaces indicate $|\phi|^2$ of the defect states. Where relevant, the higher energy eigenvalue KS state is colored yellow, and the lower blue. All iso-values are $= 0.005$.

5.7 The KS defect level positions for oxygen vacancies in amorphous HfO$_2$. The level positions are given (in eV) relative to the CBM. The spread of level positions are indicated by the bars. In some charge states, levels are spread over 1 eV.

5.8 (A) The optical absorption spectrum of the $V^0_O$ defect in m-HfO$_2$. (B) The optical absorption spectrum of the $V^{+1}_O$ defect in m-HfO$_2$. (C) The optical absorption spectrum of the $V^{+2}_O$ defect in m-HfO$_2$. In all three graphs, the original spectrum calculated by TDDFT is given by the blue delta functions. The red line give the spectrum which has been greated by smoothing the TDDFT spectrum, using Gaussians of $\sigma = 0.2$. 

5.9 The triplet state of the $V^{+2}_{O}$ defect. The hole has localised onto a STH state and is separated from the $V_O$ defect, which is now in the +1 charge state. The separation between the hole and the vacancy is variable, however the total energy of the defect is weakly dependant on this separation. Both iso-vales are $= 0.006$. . . . . . . . . . . . . . . . . . . 109

5.10 The neutral oxygen vacancy in the 324-atom cell (A), and close up with other atoms hidden (B). Similar to the hole bipolaron, a O-O dimer is formed. Bond lengths are indicated and are in Å. . . . . . . . 111

5.11 The $O^{-1}_I$ defect in m-HfO$_2$. (A) shows the full supercell and (B) shows the local structure. The NN ions are distorted strongly, and $O_I$ forms sits 2.0 Å from its NN O ion. The NN O ion is displaced out of its crystal position and is symmetrically identical to the interstitial ion. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 111

5.12 The -2 charged oxygen interstitial defect. (A) is the entire super cell, where the position of the interstitial in the lattice is clear. (B) shows the local structure of the oxygen interstitial, with bond lengths indicated in Å. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112

5.13 The incorporation energy of the oxygen interstitial defect in amorphous HfO$_2$. Width of lines corresponds to the range of possible energies. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 113


6.2 The percolation path breakdown model. An electrode-dielectric-electrode structure, with some initial defect concentration, undergoes electrical stress. This produces a build up of randomly positioned defects. Eventually, a complete pathway, connecting the two electrodes, is formed and the dielectric is broken. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 116
6.3 (A) The bulk structure of m-HfO$_2$. This is the configurations of the system before the FP creation has occurred. Important atoms, which will take part in the reaction, are highlighted. (B) The final configuration of the neutral (+2/-2) FP in m-HfO$_2$. Oxygen ion 2 displaces oxygen ion 1, and oxygen ion 1 dislocates and moves into an interstitial position. The vacancy is at the location previously occupied by oxygen ion 2. The vacancy and interstitial therefore have an oxygen ‘barrier’ between them.

6.4 (A) The energy profile of the O$_{\text{Bulk}}$ → O$_{\text{I}}^{-2}$ + V$_{O}^{+2}$ process in m-HfO$_2$. (B) The energy profile of the -2 FP production process in m-HfO$_2$. Reaction coordinate refers to the replica index, with 0 being the initial (bulk) state and 9 being the final (FP) state.

6.5 (A) This initial state of the -2 charged FP generation process in m-HfO$_2$. The electrons are in the CBM state. (B) Final configuration of the -2 FP. Oxygen ion 1 dislocates, leaving behind a neutral vacancy and creating a -2 charge oxygen interstitial.

6.6 Schematic of the conversion of bipolarons into FPs. Hf ions are grey and O ions are red. (a) Shows the hole injection case. Injected holes are localised onto hole bipolaron states. Frenkel defects are created when an oxygen at a hole bipolaron site displaces away. The Frenkel defect is stabilised by hole transfer to the interstitial, resulting in a V$_O^{+2}$O$_1^0$ Frenkel pair. (b) shows the case of electron injection. An oxygen atom neighbouring an electron bipolaron site is dislocated from its position. The Frenkel defect is stabilised via the transfer of electrons into V$_O^{+2}$ to create a V$_O^0$. Images originally published in [4].

6.7 Histogram of the Frenkel pair formation energies in a-HfO$_2$. ’+2’ Refers to a FP formed at a double hole trap site and ‘-2’ refers to a Frenkel pair formed at a double electron trap site.
6.8 A typical oxygen di-vacancy in a-HfO$_2$. The pink and blue surfaces show the $|\psi|^2$ of the defects states. In this configuration, the vacancy states combine and look similar to bonding and antibonding orbitals. Both iso-values are $= 0.0025$.

6.9 (A) Histogram of the distribution of formation energies of oxygen divacancies. Formation energy is referenced to pristine amorphous structure. (B) Histogram of the distribution of formation energies of oxygen divacancies. Similar to (A), the formation energy is referred to the pristine bulk.

6.10 Schematic of the two types of oxygen tri-vacancy in a-HfO$_2$. (a) shows the the linear type of configuration. (b) shows the triangular type of configuration. On average, triangular type tri-vacancies are lower in energy.

7.1 The impact of defects on the function of a dielectric in a MOSFET device. Defects, indicated by red dots, can be charged. Charging can affect the conduction through the n-channel. Also, electron trapping at a defect allows for trap assisted tunnelling (TAT) through the layer. This produces a small current through the insulating layer. If the defect concentration is high enough, a 'conductive filament' can exist, which conducts electrical current through the layer. Such filaments are responsible for dielectric breakdown (see Chapter 6).

7.2 The bulk structure of hexagonal boron nitride from a top-down (A) and side-one (B) view. The layers are organised into a honeycomb structure, and are stacked in an AB fashion. Layer separation is given in Angstroms.

7.3 The PDOS of a $6 \times 6 \times 4$ cell of bulk hexagonal boron nitride. The predicted band gap is 5.7 eV.

7.4 (A) A hBN layer. The Stone Wells defect is formed by rotating one BN unit by 90 degrees. (B) The fully relaxed Stone Wells defect.
7.5 The PDOS of the SW defect. There is an occupied state near the VBM and an unoccupied state at the CBM. 

7.6 The boron vacancy in its three charge states. The N-N distances are given in Å. The coloured surfaces indicate spin density, with yellow indicating positive spin density and green indicating negative spin density. In the neutral and -2 charge states the defect geometry has $C_{2v}$ symmetry. The -1 charge state has $D_{3h}$ symmetry. The spin density iso-value is 0.03.

7.7 (A) Formation energy plot of the boron vacancy in its charge states. (B) Formation energy plot of the nitrogen vacancy in its charge states. The Fermi level has, in each case, been defined so that the neutral case of each defect is equal to zero.

7.8 (A) The $V_{N}^{-1}$ defect in hBN. The spin state is singlet. (B) The $V_{N}^{0}$ defect in hBN. The spin state is doublet. Yellow surfaces show the spin density. (C) The $V_{N}^{+1}$ defect in hBN. The spin state is singlet. All B-B distances are given in Å. The iso-surface value is 0.02.

7.9 (A) The in-plane boron divacancy in the singlet spin configuration. Surfaces indicate spin density, with yellow representing positive (or up) spin and green representing negative (or down) spin. In the divacancy, an N atom, previously bonded to two B atoms, migrates to form a bond with another two N atoms adjacent to the vacancies. (B) Shows the PDOS of the in-plane boron divacancy in the singlet spin configuration. Iso-surfaces have a value of 0.025.

7.10 (A) The in-plane boron divacancy in the triplet spin configuration. Surfaces indicate spin density. In the divacancy, an N atom, previously bonded to two B atoms, migrates to form a bond with another two N atoms adjacent to the vacancies, identical to the singlet case. (B) Shows the PDOS of the in-plane boron divacancy in the triplet spin configuration. Isosurfaces have a value of 0.025.
7.11 (A) The inter-plane boron divacancy. The ground state of this divacancy configuration has a singlet spin arrangement. (B) The PDOS of the inter-plane boron divacancy. The states are in very similar places to the isolated boron vacancy. This similarity emerges because of the minimal interaction between the boron vacancies between layers, thus the interplane divacancy behaves as if it is two noninteracting boron vacancies. Iso-surface values are 0.03.

7.12 The N-bridged boron divacancy. (A) shows an orthographic view where the bridging between layers is clear. (B) gives a perspective view, showing the spin density (with different colouring representing different spin). Iso-surface values are 0.03.

7.13 (A) The -1 charged N divacancy. (B) The Neutral charged vacancy. (C) The +1 charged vacancy. Iso-surface is 0.02.

7.14 The neutral boron interstitial. (A) gives a far of few with the two layers adjacent to the B\textsubscript{I} visible. (B) Is the B\textsubscript{I} and its NN with bond lengths indicated in Å. The spin density iso-surface value is 0.01.

7.15 The formation energy of the B interstitial is its possible charge states. Crossover point indicate the charge transition levels. The Fermi level is set in each case so that the B\textsubscript{0}\textsubscript{I} and N\textsubscript{0}\textsubscript{I} formation energy is 0.

7.16 (A) The neutral N interstitials and its two adjacent layers. The nitrogen interstitial sits closer to one layer, and is separated from in nearest neighbour nitrogen by 1.5 Å. (B) Shows the surrounding structure of the N interstitial. The orange surface shows an isosurface = 0.01 of the spin density.

7.17 The charge transition levels of the single intrinsic defects in hBN. The band gap of hBN is 5.8 eV.
List of Tables

3.1 The lattice parameters (Å) and band gap values (eV) for bulk MgO. The optical band gap (BG) is calculated using TDDFT. Cell sizes are given in number of atoms. ........................................ 68

3.2 Displacements of the NN and NNN in the O-vacancy in MgO. All displacements are given in Å. .......................................................... 70

6.1 Summary of the defects in a-HfO₂. “Spread” is quantified as the energy difference between the lowest and highest value is the range of the given value. All energies are given in eV. In CTLs, a dash represents that further charging is not possible. ......................... 128

6.2 Summary of the defects in m-HfO₂. All energies are given in eV. For the CTLs, a dash indicates that further charging is not possible. .... 128
Chapter 1

Introduction

1.1 Motivation and Impact

Microelectronic devices have played a major role in technology throughout the 20th and 21st centuries. Such devices have typically relied on the creation of structures which feature layering of semiconductors, insulators, and metals. For example, in the transistor design known as a metal-oxide-semiconductor field effect transistor (MOSFET, see Fig. 1.1), one can find n-type semiconductors, p-type semiconductors, a metal electrode and an insulator (a dielectric). The focus of this thesis is on insulating materials - wide band gap materials used in the 'gate dielectric', which can be seen in Fig 1.1. Dielectric materials are not just employed in MOSFET design, however. They have applications in other nanoscale devices and in proposed future devices (section 1.4). The three materials investigated in this thesis have all been, or are expected to be, used as a dielectric in nanoscale devices. Their functionality relies on their insulating properties, and often the functionality of a device relies on the neutrality of the dielectric layer. The essence of this thesis, therefore, is to investigate the properties of these materials, and understand when such materials fail in devices. The dielectric properties are strongly affected by the presence of grain boundaries, dislocations and point defects. Defects are not only responsible for breaking the insulating properties of dielectrics (‘dielectric breakdown’, see chapter 6), but can also be responsible for charging of the dielectric. In a MOSFET, the purpose of the gate dielectric is to allow control of the conducting channel via
1.1. Motivation and Impact

an electric field. If the insulating properties of the dielectric layer are lost, there can be no significant separation of charge at the gate and therefore the ability to control the n-channel via an $E$-field is lost.

HfO$_2$ has recently been applied as a replacement to SiO$_2$ as the typical gate oxide used in MOSFET devices [5]. The advantage of HfO$_2$ over SiO$_2$ is rooted in the fact that HfO$_2$ has a significantly higher dielectric constant. This has allowed continued ‘down-scaling’ of transistor devices (section 1.3.2), which has been responsible for the rapid growth in computer power over the latter half of the 20th century. Owing to the reduced dimensions of the HfO$_2$ layer, and to the fabrication process, it is common for the structure of the HfO$_2$ layer to be amorphous. This carries with it certain advantages. In an amorphous material, there are no crystal grains and therefore no grain boundaries. Grain boundaries in polycrystalline dielectrics have been linked to degradation processes in devices, and so amorphous dielectrics have been seen as somewhat preferable. However, the energetics and dynamics of defects has a dependence on the microscopic structure of the material, and thus such properties may be expected to differ between the crystal and amorphous phase. For this reason, both crystalline and amorphous phases of HfO$_2$ are studied in this thesis.

HfO$_2$ and SiO$_2$ are not the only insulators employed in nanoscale devices. There are many insulating materials with wide band gaps and other properties which make them appropriate for applications as dielectrics. These include MgO, Al$_2$O$_3$, TiO$_2$, MoS$_2$, hexagonal Born Nitride (hBN) and many more. The problems that plague the application of such materials are fairly common between them - point defects and grain boundaries reduce reliability. The details of such processes, however, differ between materials. For this reason, characterisation and modelling of individual defect structures within each material is required. The research topic of this thesis has been motivated by collaboration with Sematech on the understanding of breakdown and resistive switching (discussed later). Defects and defect-related processes of three highly practical insulators - MgO, HfO$_2$ and hBN - are studied.
1.2 Structure of Results Sections

The results of this thesis are presented as follows. Firstly, a new implementation of time dependent density functional theory is tested on oxygen vacancies in MgO. MgO has long been treated as an archetypical insulating oxide. Within the field of computational solid state physics and materials science, MgO is often one of the many materials used as a test system for new methodological approaches. MgO is also, however, a practical material in its own right. It is currently being used as a dielectric oxide in magneto-tunnel junction (MTJ) devices [6, 7]. In the Chapters 4 to 6, defects in both crystalline and amorphous HfO$_2$ are studied. First, the phenomena of intrinsic charge trapping is studied. This looks into how pristine (defect-free) HfO$_2$ may charge when exposed to excess holes or electrons. Then, oxygen based defects (vacancies and interstitials) are studied. Oxygen vacancies are believed to be at the root of many degradation processes in HfO$_2$. For this reason they are investigated in this thesis. Finally, the generation of these defects under an applied electric field is studied. In the final chapter, hexagonal Boron Nitride (hBN) is investigated. hBN has recently been subject to a growth of scientific interest. The discovery of graphene has led to much excitement over the possibility of 2D-material based devices. Since hBN is isostructural to graphite and, like graphite, can exist as a multi-layer or single layer 2D material, it is being considered as a possible material for the dielectric layer in graphene based devices. In the final chapter of this thesis, a wide range of defects in hBN are studied.

In this introduction, the properties of the MOSFET are discussed. The effects of down scaling transistor devices are discussed, and the need to replace the gate dielectric with a high dielectric constant material is explained. The effects that defects and ‘trap states’ can have on the material is then discussed. This is then put into the wider context of devices, and the particular degradation phenomena are explained. Then, newer technologies, which also motivate this thesis, will also be discussed. This includes the fast-developing technology known as Resistive Random Access Memory (ReRAM), which uses direct application of the ‘Memristor’, a electrical circuit element which is now being fully realised.
1.3 Semiconductor Devices

1.3.1 The MOSFET

Figure 1.1 shows a schematic of a n-MOSFET. The principle of a n-MOSFET is that a sufficient voltage applied to the gate will induce an inversion layer in the p-doped region, which then greatly increases conduction between the (n doped) source and drain. A p-MOSFET works on a similar principle, but with an n-doped substrate and p doped source and drain. The physics is analogous, so here the operation mechanisms will be discussed in terms of n-MOSFETs (henceforth called MOSFETs).

An inversion layer occurs when band bending (induced by an external field) causes the majority carrier to switch from holes to electrons. Thus, an n-channel appears which connects the two (highly) n-doped regions, the source and drain. This gives the basic operating principle of a transistor - a conducting channel can be turned on via an applied voltage. The voltage required to induce an n-channel is called the threshold voltage, $V_{Th}$. The current through the n-channel can be approximated as [8]

$$I_D = \mu_n C_{ox} \frac{W}{L} \left( (V_{GS} - V_{Th})V_{DS} - \frac{V_{DS}^2}{2} \right)$$

(1.1)

where $I_D$ is the drain current, $W$ is the channel width, $L$ is the channel length, $C_{ox}$ is the oxide capacitance density, $V_{GS}$ is the potential difference between the gate and
source, and $V_D$ is the voltage of the drain. The drain current increases linearly with $V_D$ until a maximum is reached when $V_D = V_{GS} - V_{Th}$, yielding

$$I_{D,Sat} = \frac{W}{L} \mu_n C_{Ox} \frac{(V_G - V_{Th})^2}{2}. \quad (1.2)$$

A reduction in channel length, or an increase to the gate capacitance will cause an increase to $I_{D,Sat}$. $(V_{GS} - V_{Th})$ cannot be increased easily, as a high $V_{GS}$ will lead to a problematically large $E$ field across the dielectric, and $V_{Th}$ cannot be reduced below approximately 200 mV, since (at room temperature) $k_B T \approx 25$ meV [8, 9]. This would lead to thermal fluctuations significantly affecting the $V_{Th}$ value.

### 1.3.2 Scaling of Transistor Technology

The increasing demand for greater computing power has ultimately manifested itself as a continued down scaling of the MOSFET. Reduction of the transistor dimensions allows for a greater device density on a given CPU [10]. A proportionate decrease of the channel width and length will not decrease the drive current (equation 1.2). It is also desirable to reduce the gate oxide thickness, since this will increase $C_{ox}$. An increase in the capacitance increase the current through the channel, making reductions to $V_{GS}$ practical. Thus, increasing the device density on a silicon wafer requires a reduction of all three dimensions of the MOSFET. For many decades, this process has allowed for a remarkable increase in computer power. The famous Moore’s law stated that the number of transistors on a Si wafer doubles every 2 years [11]. This continuous down-scaling cannot, however, continue forever. The reduction of the gate oxide thickness has been, in particular, a problem for the continuation of Moore’s law. As the gate oxide thickness is reduced to the nanometer regime, electron tunnelling effects become significant. The function of the gate oxide requires it to remain electrically insulating. Very thin oxides, however, will allow a significant ‘leakage current’ through via electron tunneling from the gate through to the Si substrate. To combat this problem, the gate oxide capacitance can be increased in a different way. The simple model for the capacitance-density of a
1.3. **Semiconductor Devices**

Figure 1.2: The impact of defects on the function of a dielectric in a MOSFET device. Defects, indicated by red dots, can be charged. Charging can affect the conduction through the n-channel. Also, electron trapping at a defect allows for trap assisted tunnelling (TAT) through the layer. This produces a small current through the insulating layer. If the defect concentration is high enough, a ‘conductive filament’ can exist, which conducts electrical current through the layer. Such filaments are responsible for dielectric breakdown (see Chapter 6).

A parallel plate capacitor can be written:

\[
C = \frac{\kappa \varepsilon_0}{d},
\]

where \(\kappa\) is the dielectric constant, \(\varepsilon_0\) is the permittivity of free space, and \(d\) is the capacitor thickness. As discussed above, decreasing the thickness increases the capacitance. Another way to increase capacitance is by increasing the dielectric constant. This is what has lead to the replacement of SiO\(_2\) as the gate oxide. The common replacement is HfO\(_2\), which has a dielectric constant of between 20 and 25 [12], compared to that of SiO\(_2\) which has a \(\kappa\) of \(\approx 3.9\) [13]. By using a gate oxide with a larger dielectric constant, a thicker oxide layer can be used to provide a given capacitance. This will reduce the electron tunnelling probability, and therefore reduce current through the layer. A new material, however, presents new challenges. Defects in HfO\(_2\) will have different properties to those in SiO\(_2\). Furthermore, there will be different atomistic processes at play, which can affect the oxide’s insulating properties.

### 1.3.3 The impact of defects on dielectric layers

Figure 1.2 shows a schematic of how defects in the gate dielectric can impede MOSFET performance. All of the problems are rooted in (some) defects’ ability to trap...
1.4 Memristors

There are exciting new technologies emerging which are based on the circuit element known as the ‘memristor’. The memristor was first proposed as a ‘missing
1.4. Memristors

circuit element’ in 1971 [20]. The four two-terminal basic circuit elements are defined in terms of how they relate the four basic circuit variables - \( I \) (current), \( V \) (voltage), \( Q \) (charge) and \( \lambda \) (flux linkage). The flux linkage is defined by

\[
\lambda = \int_{-\infty}^{t} V(t')dt'.
\]

(1.5)

The resistor is defined by a relationship between \( V \) and \( I \), the inductor is defined by a relationship between \( \lambda \) and \( I \), and the capacitor links \( Q \) and \( V \). The memristor is the circuit element which defines the relationship between \( \lambda \) and \( Q \). The memristor has a property called ‘memristance’, \( M \), which gives the relationship

\[
d\lambda = M(Q)dQ.
\]

(1.6)

The terms ‘memristor’ and ‘memristance’ earned their names since the resistance of a memristor depends on its history, that is, it has memory. Now, given the above definition of the flux linkage, and the fact that \( Q = \int_{-\infty}^{t} I(t')dt' \), one finds

\[
M(Q) = \frac{d\lambda}{dQ} = \frac{d\lambda}{dt} \frac{dt}{dQ} = \frac{V(t)}{I(t)}.
\]

(1.7)

Thus, the memristance is the charge-dependent resistance. This is similar to Ohm’s law, however the memristance depends on the time integral of the current through the memristor. Thus the resistance of a memristor is influenced by the history of its current.

1.4.1 ReRAM Devices

A realisation of the memristor component has been found in so-called resistive random-access-memory (ReRAM) devices. Their name is derived from their potential application as non-volatile random access memory. Memristors are naturally suited for memory storage applications, since the switching between high resistance and low resistance states can correspond to switching between 0 (‘OFF’) and 1 (‘ON’) states. In a memristor, a switching will be driven by current, or, more precisely, the time integral of the current through the memristor.
Resistive switching has been observed in a number of insulators [21, 22]. The basic observation is that, with appropriate voltage stressing, the insulator will switch to a low-resistance state and start conducting. Furthermore, again with an appropriate voltage (or current), the insulator can switch back to a high-resistance state. There are two broad categories of such switching. Bipolar switching occurs by reversing the polarity of the voltage. The ReRAM device is first set by applying a large enough voltage through the insulator, leading to the low resistance state. Reversal of the voltage polarity then switches the device back into the high resistance state. Alternatively, there is also unipolar switching. In this case, the OFF to ON voltage and ON to OFF voltages have the same polarity, but different value [23].

The mechanism of ReRAM devices is often understood to be the development of a conductive filament [24, 25]. The SET of an operation is the growth of the conductive filament in the insulator. The RESET operation ruptures the filament, breaking the conduction pathway and bringing the material back to a high resistance state. The precise composition of the resistive filament, and the mechanism of its growth and rupturing, varies from system to system and is, in general, the subject of controversy. Throughout this thesis, a particular model of ReRAM operation in HfO$_2$ is investigated in which the conductive filament is composed of oxygen vacancies. This mechanism, and others, is discussed in Chapter 6.

1.5 Summary

In this introduction, it has been shown that dielectric materials play an important role in microelectronics technology. Down-scaling in MOSFETS has led to the demand for high-κ materials, requiring modelling and investigation of their defects. These defects affect the function of devices, and therefore the modelling of such defects plays an important role. By using atomistic models to describe defects, one can calculate formation energies, charge transition levels, optical spectra and many more properties. All of these have a part to play - whether by offering predictions on the performance of a material in-device or by helping to interpret experimental data which is aiming to identify defects in lab samples, for example by calculating
optical absorption lines. Characterisation of defects will also play a role in the development of future technologies, such as ReRAM, which aim to advance computer technology beyond the end of Moore’s law.

1.6 Publications Derived from this Thesis


Chapter 2

Theory & Methodology

2.1 Condensed Matter Systems and the Many-Electron Problem

The properties of solid systems can be predicted and explained using the Schrödinger equation. Solving the Schrödinger equation means finding a many-body wavefunction which is a function of the position of each particle in the system. In a macroscopic system, however, the number of particles to be described is very large. For a macroscopic system, with mass on the order of \(10^{-3}\) kg, the number of atoms to consider will be on the order of \(N_A (\approx 10^{23})\). For each atom, both its nucleus and its electrons will need to be described with three position coordinates. Thus, the wavefunction depends on a huge number of variables. In a solid state system, the full (electrons and nuclei) equation to solve is

\[
\mathcal{H}\Psi_n(\{r\},\{R\}) = E_n\Psi_n(\{r\},\{R\}) \tag{2.1}
\]

where \(\mathcal{H}\) is the Hamiltonian and \(\{r\}\) and \(\{R\}\) are the set of the electronic and nuclear coordinates, respectively. The Hamiltonian, \(\mathcal{H}\), is expressed in atomic units as

\[
\mathcal{H} = -\frac{1}{2} \sum_i \nabla_i^2 + \frac{1}{2} \sum_{i,j \neq i} \frac{1}{|r_i - r_j|} - \sum_i Z_i \frac{1}{|r_i - R|} - \sum_i \frac{1}{2M_i} \nabla_i^2 + \frac{1}{2} \sum_{i,j \neq i} \frac{Z_i Z_j}{|R_i - R_j|} \tag{2.2}
\]
where the terms represent, from left to right, the electronic kinetic energy, the
electron-electron Coulombic repulsion, the Coulomb interaction between the elec-
trons and the nuclei, the kinetic energy of the nuclei, and the Coloumb interaction
between nuclei. Solving this equation analytically is impossible for real systems
and this necessitates a computational approach. Even with modern computer tech-
nology, however, this task is still incredibly difficult. A number of simplifications
must be made to reduce the complexity of this task.

2.1.1 Finding the Electronic Structure- Density Functional The-
ory

In this project calculations are made within the adiabatic approximation. The first
step of the adiabatic approximation is to separate the electronic and nuclear degrees
of freedom. This means that the full wavefunction is written

$$\Psi_n(\{r\}, \{R\}) = \Phi(\{r\}; \{R\})\chi(\{R\}),$$

(2.3)

where $\Phi(\{r\}; \{R\})$, the wavefunction of the electrons (the 'electronic wavefunc-
tion'), which depends parametrically on the nuclear coordinates. $\chi(\{R\})$ is the
wavefunction of the nuclei (the 'nuclear wavefunction'). The adiabatic approxima-
tion assumes that a system can be modelled as though the electrons are responding
to static nuclei. Nuclei can still be allowed to move, but the energy of the electrons
depends only on nuclear coordinates, and not on the momentum of the nuclei. This
approach ultimately means that two equations must be solved. Firstly, the electronic
Schrödinger equation

$$\mathcal{H}_e \Phi(\{r\}; \{R\}) = E_e \Phi(\{r\}; \{R\}),$$

(2.4)

where $\mathcal{H}_e$ is the electronic Hamiltonian and is expressed

$$\mathcal{H}_e = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{i,j \neq i} \frac{1}{|r_j - r_i|} - \sum_{i,j} \frac{Z_j}{|R_j - r_i|}.$$
The Schrödinger equation for nuclei is then expressed as

\[
\left( -\sum_i \frac{1}{2M_i} \nabla_i^2 + \frac{1}{2} \sum_{i,j \neq i} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} + \varepsilon_e(\{\mathbf{R}\}) \right) \chi(\{\mathbf{R}\}) = E_n \chi(\{\mathbf{R}\}) \tag{2.6}
\]

where it has been assumed that derivatives of \( \Phi \), with respect to the nuclear coordinates, are not very large. This assumption correspond to the situation in which the wavefunction varies smoothly with the nuclear coordinates. Therefore, the electronic wavefunction “adiabatically” follows the motion of the nuclei.

Solving the electronic Schrödinger equation for a system size of just a few hundred atoms and electrons is a computationally expensive task. In principle, exact solutions can be found using the configuration interaction (CI) approach. CI uses a linear sum of Slater determinants to represent the electronic wavefunction. This approach however is computationally intensive and is not practical for the system sizes studied in this project.

2.1.1.1 Representing the Electronic Wavefunction

Electrons are spin-\( \frac{1}{2} \) particles, or Fermions. This means that the electronic wavefunction is antisymmetric under particle exchange operations, and that a many-electron system should obey Fermi-Dirac statistics when in thermal equilibrium. The most important feature of this fact, for the purposes of this project, is ensuring that the wavefunction is antisymmetric under particle exchange. This has typically been done by the use of Slater determinants. A slater determinant makes use of the properties of determinants (from linear algebra) to ensure that a many-electron wavefunction is antisymmetric. It is written as the \( N \times N \) matrix,

\[
\Phi(\{\mathbf{r}\}) = \frac{1}{N!} \begin{vmatrix}
\psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \psi_1(\mathbf{r}_3) & \ldots & \psi_1(\mathbf{r}_N) \\
\psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \psi_2(\mathbf{r}_3) & \ldots & \psi_2(\mathbf{r}_N) \\
\psi_3(\mathbf{r}_1) & \psi_3(\mathbf{r}_2) & \psi_3(\mathbf{r}_3) & \ldots & \psi_3(\mathbf{r}_N) \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
\psi_N(\mathbf{r}_1) & \psi_N(\mathbf{r}_2) & \psi_N(\mathbf{r}_3) & \ldots & \psi_N(\mathbf{r}_N)
\end{vmatrix}. \tag{2.7}
\]
The $\psi(r)$ functions are the one electron orbitals of the system. This means the many-electron wavefunction as being composed of electrons each occupying a molecular orbital. If the $j$th electron is in molecular orbital $i$, this is represented as $\psi_i(r_j)$. Therefore, swapping the labels of two electrons (or two orbitals) would correspond to interchanging the row or columns of the Slater determinant. As is known from linear algebra, if determinant $\mathcal{D}'$ is equal to determinant $\mathcal{D}$ with two rows (or columns) interchanged, then $\mathcal{D}' = -\mathcal{D}$. Furthermore, if two electrons are in the same molecular orbital, this corresponds to the situation where two $\psi_i$ are identical, in which case two rows are identical and the determinant is equal to 0. Thus the Pauli exclusion principle is obeyed.

2.1.1.2 Density Functional Theory

A highly effective method for finding the electronic structure of a system was discovered by Walter Kohn and Pierre Hohenberg [31]. Writing the electronic Hamiltonian as

$$ \hat{\mathcal{H}}_e = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{i,j} \frac{1}{|r_j - r_i|} + V_{\text{ext}}, $$

(2.8)

Kohn and Hohenberg demonstrated that there is a one-to-one correspondence between the external potential of the system and the ground-state electronic density. In the case of condensed matter physics, the external potential is the electrostatic potential of the atomic nuclei. This means that knowledge of the external potential is sufficient to calculate the ground state electron density of the system, that is, there is a one-to-one correspondence

$$ V_{\text{ext}} \leftrightarrow n_{\text{GS}}(r), $$

(2.9)

where $n_{\text{GS}}(r)$ is the ground state electron density. Secondly, Kohn and Hohenberg also demonstrated that the correct ground-state electron density, $n_{\text{gs}}(r)$, minimises the electronic energy. One can therefore use variational methods to find $n_{\text{gs}}(r)$. Thus, a new approach was suggested in which one could express the total energy as a functional of the electron density, $E[n(r)]$. This approach was termed density

Functional theory (DFT) and has had a huge impact on computational materials science. The fundamentals of DFT are now well established and detailed descriptions can be found in [32, 33, 34]. The energy, \( E[n(r)] \), can be expanded out as

\[
E[n(r)] = T[n(r)] + U_{\text{ext}}[n(r)] + U_{\text{ee}}[n(r)],
\]

(2.10)

where \( T[n(r)] \) is the kinetic energy, \( U_{\text{ext}}[n(r)] \) is the energy of the interaction between the electrons and the external potential, and \( U_{\text{ee}}[n(r)] \) is the electron electron interaction. This can be alternatively expressed as

\[
E[n(r)] = \int n(r)v_{\text{ext}}(r) d^3r + F[n(r)]
\]

(2.11)

where

\[
F[n(r)] = T[n(r)] + U_{\text{ee}}[n(r)],
\]

(2.12)

where \( F[n(r)] \) contains the full many-body electron-electron interaction and electron kinetic energy.

In practice, DFT calculations use the Kohn-Sham formalism [35]. Kohn and Sham formulated DFT as an approach where one simultaneously solves a set 1-electron-like Schrödinger equations, one for each of the single electron orbitals \( \{\psi(r)\} \), termed the Kohn-Sham (KS) orbitals. These orbitals form a Slater determinant which is referred to as the KS wavefunction. The KS orbitals then give the lowest energy solution to the set of equations

\[
\left( \frac{\hbar^2}{2m_e} \nabla^2 + V_{\text{KS}} \right) \psi_j(r) = \varepsilon_j \psi_j(r) \quad \text{for} \quad j = 1, 2\ldots N,
\]

(2.13)

where \( N \) is the number of electrons. The KS wavefunction is the wavefunction of a system of noninteracting electrons which are acted upon by the KS potential, defined as

\[
V_{\text{KS}} = V_{\text{ext}}(r) + \int \frac{n(r')}{|r-r'|} d^3r' + V_{\text{XC}}[n(r)]
\]

(2.14)

where \( V_{\text{ext}}(r) \) is the external potential, which in a solid originates from the lattice
of nuclei. The integral gives the Hartree term, which is the Coulombic electron-electron interaction. The final term, $V_{XC}$, is the exchange-correlation potential, which is defined as

$$V_{XC}[n(r)] = \frac{\partial E_{XC}[n(r)]}{\partial n(r)} \tag{2.15}$$

where $E_{XC}[n(r)]$ is the exchange-correlation functional, which contains the many-body effects of the electron-electron interaction.

The KS wavefunction is not an approximation to the true, many-body wavefunction. However, the electron density of this noninteracting system, defined $n(r) = \sum_j N |\psi_j(r)|^2$, equals the electron density of the ground state with full many-body interactions. In addition, the ground state energy of the fictitious noninteracting system equals the ground state energy of the true many-body system. The KS wavefunction is thus a mathematical object used to find the correct $n(r)$ and $E_e$ of the ground state, rather than the correct wavefunction.

Density functional theory is, in principle, an exact theory which could be used to solve the many-electron problem. Unfortunately, the exact form of the exchange-correlation functional is not known. The definition of the exchange-correlation functional can be understood by making an expansion to the total electronic energy, $E[n(r)]$, in a similar way above. In the Kohn-Sham method, part of the kinetic energy is calculated straight from the orbitals,

$$T_{KS}[n(r)] = \sum_i < \psi_i | -\frac{1}{2} \nabla_i^2 | \psi_i > . \tag{2.16}$$

Any part of the many-body kinetic energy that is not included by this expression is considered to be part of the exchange correlation energy, $E_{XC}[n(r)]$. A similar idea is used for the electron-electron interaction, where one calculates the Hartree integral

$$E_H[n(r)] = \frac{1}{2} \int \frac{n(r)n(r')}{|r-r'|} d^3rd^3r' \tag{2.17}$$
and then the total energy is rewritten

\[ E[n(r)] = T_s[n(r)] + E_H[n(r)] + \int n(r)v_{\text{ext}}(r)d^3r + E_{\text{XC}}[n(r)]. \]  

(2.18)

Thus, the many-electron effects that are not captured by \( T_s \) or \( E_H \) are now part of the exchange-correlation functional. By reference to the earlier expansion for \( E[n(r)] \), the exchange-correlation energy can then be written as

\[ E_{\text{XC}}[n(r)] = (T[n(r)] - T_s[n(r)]) + (U_{ee}[n(r)] - E_H[n(r)]) \]  

(2.19)

Approximations must be made to \( E_{\text{XC}}[n(r)] \). The most basic approximation is the ‘Local Density Approximation’, or LDA. In the LDA, the exchange-correlation energy is calculated from the uniform electron gas. It is expressed as

\[ E_{\text{XC}}[n(r)] = \int d^3r n(r)V_{\text{LDA}}^{\text{XC}}(n), \]  

(2.20)

where \( V_{\text{LDA}}^{\text{XC}}(n) \) is the exchange-correlation energy per particle of a uniform electron gas of density \( n \). The LDA takes advantage of the fact that the exchange-correlation energy can be calculated exactly for a uniform electron gas. The approximation, however, is fairly crude and is particularly ill equipped to model systems with rapidly varying electron density. An improvement can be made by including gradient dependent terms, \( \nabla n(r) \). This is termed the generalised gradient approximation (GGA). GGA functionals have the form

\[ E_{\text{XC}}^{\text{GGA}} = \int d^3r n(r)V_{\text{XC}}^{\text{GGA}}(n, \nabla n). \]  

(2.21)

There are many different ways density-gradient dependance can be incorporated into an XC functional. This project uses an XC functional which is based on the PBE [36] GGA functional, which is itself a popular choice.
2.1.2 The Basis of the KS Orbitals

KS orbitals are usually expressed using a basis set. In principle, any function of space can be expressed as a linear combination of functions, \( f(\mathbf{r}) \), from a complete, orthonormal set \( \{f(\mathbf{r})\} \). The set of planewaves, \( e^{i\mathbf{k} \cdot \mathbf{r}} \), is one such set and therefore can be used as a basis set for a KS orbital by writing

\[
\psi_j = \sum_i A_i e^{i\mathbf{k}_i \cdot \mathbf{r}} \tag{2.22}
\]

where the sum runs for increasing \( \mathbf{k}_i \) up until a ‘planewave cutoff’. Appropriate cutoff values have to be found for each system and computational set-up, and this is done by increasing the cutoff wavelength until the total energy of the calculation is changing by a negligible amount. Plane waves offer a number of advantages. They form a complete set, and reflect the periodic nature of the crystal states. This project, however, makes use of the ‘Linear Combination of Atomic Orbitals’ (LCAO) method. In this approach, the KS orbitals are expressed as linear combinations of atomic-type (\( s, p, d \ldots \)) orbitals. That is,

\[
\psi_i(\mathbf{r}) = \sum_{k,j} C_{kj}^i \phi_k(\mathbf{r} - \mathbf{R}_j) \tag{2.23}
\]

where \( C_{kj}^i \) are the coefficients of the expansion, and \( \phi_k(\mathbf{r} - \mathbf{R}_j) \) is the \( k \)th atomic orbital, centered on the \( j \)th atom at \( \mathbf{R}_j \).

2.2 CP2K

In this project the CP2K software package [37] has been used to implement density functional theory. CP2K uses a joint Gaussian and planewave basis set [38]. The atom-centered Gaussian functions serve as a primary basis set to represent the KS orbitals, and the planewave representation is used as an auxiliary basis set to represent the electron density. By representing electron density with a plane wave basis, highly efficient fast Fourier transform (FFT) methods can be exploited to solve the Poisson equation. Using Gaussian functions to represent the wave functions also has its advantages - the localised nature of the function causes the overlap, KS and
density matrices to become sparser with increasing system size [39]. CP2K uses this sparsity to increase the efficiency of SCF calculations.

In this project calculations have used periodic boundary conditions. This takes advantage of the translational symmetry of crystal systems so that one need not simulate $\mathcal{O}(N_A)$ number of particles. One need only solve the KS equations for a unit cell with the periodic boundary conditions appropriately included in the Hamiltonian. In practice, this reduces the complexity of the task to solving an equation representing 100s of atoms and 1000s of electrons (depending on the choice of unit cell). It does, however, introduce extra difficulties which will be explained in section 2.6.1.

2.2.1 Calculation of energy in CP2K with a dual Gaussian-Planewave basis set

The electron density is represented in both a Gaussian and a planewave basis set. For the primary, Gaussian basis set

$$n(r) = \sum_{\mu, \nu} P^{\mu \nu} \phi_\mu(r) \phi_\nu(r), \quad (2.24)$$

where $P^{\mu \nu}$ is the density matrix element, and $\phi_\mu(r)$ are the contracted Gaussians. In the auxiliary plane wave basis, the electron density is represented by

$$\tilde{n}(r) = \frac{1}{\Omega} \sum_G \tilde{n}(G) \text{exp}(iG \cdot r) \quad (2.25)$$

where $G$ is the reciprocal lattice vector and $\Omega$ is the unit cell volume. The total energy functional can be written as


where

$$E^T[n] = \text{Kinetic Energy} = \sum_{\mu, \nu} P^{\mu \nu} \langle \phi_\mu(r) | -\frac{1}{2} \nabla^2 | \phi_\nu(r) \rangle \quad (2.27)$$
2.2. CP2K

\[ E^V(n) = \text{Electron-Ion interaction} = \sum_{\mu \nu} p^{\mu \nu} \langle \phi_\mu(r) | V_{\text{loc}}^{\text{PP}}(r) | \phi_\nu(r) \rangle \]
\[ + \sum_{\mu \nu} p^{\mu \nu} \langle \phi_\nu(r) | V_{\text{nl}}^{\text{PP}}(r, r') | \phi_\mu(r') \rangle \]

(2.28)

\[ E^H(n) = \text{Hartree Energy} = 2\pi \Omega \sum_G \tilde{n}^* (G) \tilde{n} (G) \frac{G^2}{G^2} \]

(2.29)

\[ E^{XC}(n) = \text{Exchange-Correlation Energy} = \int \rho^{xc}(r) dr \]

(2.30)

\[ E^{I-I}(n) = \text{Ion-Ion Interaction} = \sum_{I \neq J} \frac{Z_I Z_J}{|R_I - R_J|} \]

(2.31)

In equation 2.28 the interaction between the electrons and the ionic cores is represented by interaction between valence electrons and a pseudopotential (PP) (see section 2.2.1.1). The Hartree energy has been expressed in the plane wave representation. The kinetic and exchange-correlation energy are calculated using Gaussian representation of the KS orbitals. Linear scaling of the KS matrix construction is achieved through two steps. Terms involving a product of two Gaussian basis functions are analysed and removed if negligible. Such products are negligible once the Gaussian basis functions are sufficiently far apart. By this screening process, the number of non-zero terms scales linearly with system size. The second step involves the transformation from \( \tilde{n}(r) \) to \( \tilde{n}(G) \) using FFTs.

2.2.1.1 Pseudopotentials

All-electron calculations in DFT can be computationally expensive. Core electron wavefunctions oscillate rapidly near the atom center, leading to the need for very high quality basis sets. It is not necessary, however, to describe the core electrons in detail for a wide range of interesting phenomena. Typically it is only the valence electrons which play an important role in, for example, bond breaking and making. Therefore, it can be computationally efficient to replace the core electrons and nucleus with a ‘pseudopotential’. This removes the need to describe the core electrons with full quantum mechanical molecular orbitals. Furthermore, pseudopotentials also smooth out the molecular orbitals, reducing the number of nodes and so reducing the demand on the basis set (for example, by reducing the necessary planewave...
The rationale behind pseudopotentials can be understood by looking to the ‘orthogonalised plane wave’ approach [40]. In this approach valence wavefunctions are represented as planewave-like functions which are orthogonal with core states. One writes a core function as a Bloch state formed from a sum of localised one electron states on atoms (that is, the LCAO method)

$$\phi_{i}^{\text{Core}}(\mathbf{r}) = \sum_{l} \exp(i\mathbf{k} \cdot \mathbf{l}) c_{i}(\mathbf{r} - \mathbf{l})$$ (2.32)

were $c_{i}(\mathbf{r} - \mathbf{l})$ is a highly localised core state on a single atom (at position $\mathbf{l}$). This gives a reasonable form for the core electron molecular orbitals. The mathematical form of the velence orbitals, $\phi^{\text{Valence}}$ are now approximated. Two conditions are to be enforced - First, the core and valence states should be orthogonal:

$$\langle \phi_{k}^{\text{Valence}}(\mathbf{r}) | \phi_{k'}^{\text{Core}}(\mathbf{r}) \rangle = \delta_{kk'}$$ (2.33)

Secondly, the valence states should be close to a free-electron wave in regions between ions. A reasonable approximation to this is then the orthogonalised plane wave, $\phi_{k}^{\text{OPW}}(\mathbf{r})$:

$$\phi_{k}^{\text{OPW}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) - \sum_{j} \Lambda_{j} \phi_{j}^{\text{Core}}(\mathbf{r})$$ (2.34)

where $\Lambda_{j} = \langle \phi_{j}^{\text{Core}}(\mathbf{r}) | \exp(i\mathbf{k} \cdot \mathbf{r}) \rangle$. This works as a valence state since, if the core states are highly localised onto atoms they will be very small in regions between them, and therefore the plane wave term will dominate. The next step is to use the orthogonalised plane waves as a basis for the valence orbital, $\psi_{k}(\mathbf{r}) = \sum_{g} \alpha_{k - g} \phi_{k}^{\text{OPW}}(\mathbf{r})$. This molecular orbital could then be found using variational methods, and this it termed the orthogonalized planewave approach. The emergence of pseudopotentials can be seen if one attempts to use simple planewaves.
as a basis, that is if one writes

\[ \phi_{k\text{PW}}^\text{PW}(\mathbf{r}) = \sum_g \alpha_{g-k} \exp(i(\mathbf{k} - \mathbf{g}) \cdot \mathbf{r}) \tag{2.35} \]

where the same coefficients are being used. One can then express the valence orbital as

\[ \psi_k(\mathbf{r}) = \phi_k^\text{PW} - \sum_j <\phi_{jk'}^\text{Core}(\mathbf{r})|\phi_k^\text{PW}> \phi_{jk'}^\text{Core}(\mathbf{r}) \tag{2.36} \]

If one then substitutes this expression into the Schrödinger equation, \( \mathcal{H}\psi_k = E\psi_k \), the following expression emerges;

\[ \mathcal{H}\phi_{k\text{Valence}} + \sum_j (E - \varepsilon_j)\phi_{jk'}^\text{Core}(\mathbf{r}) <\phi_{jk'}^\text{Core}(\mathbf{r})|\phi_k^\text{PW}> = E\phi_k^\text{PW} \tag{2.37} \]

where \( \varepsilon_i \) is energy of eigenstate \( \phi_{ik'}^\text{Core}(\mathbf{r}) \). One is now in a position to define a pseudopotential,

\[ \gamma_{pp}\phi_{k\text{Valence}} = V(\mathbf{r}) + \sum_j (E - \varepsilon_j)\phi_{jk'}^\text{Core}(\mathbf{r}) <\phi_{jk'}^\text{Core}(\mathbf{r})|\phi_k^\text{PW}> \tag{2.38} \]

where \( V(\mathbf{r}) \) is the potential of the nuclei. Writing \( \mathcal{H}_0 = -\frac{1}{2}\sum_i \nabla^2_i + \frac{1}{2}\sum_{i,j\neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \)

the Schrödinger equation becomes

\[ (\mathcal{H}_0 + \gamma_{pp})\phi_k^\text{PW} = E\phi_k^\text{PW}. \tag{2.39} \]

This overview of orthogonalised plane waves provides the essentials element of pseudopotential methods. In this formulation of the problem one attempts to find \( \phi_k^\text{PW} \), termed the “pseudowavefunction”. The pseudowavefunction (which in this case is a sum of planewaves) is smoother in core regions than the real molecular orbital, \( \psi_k(\mathbf{r}) \). Further, \( \phi_k^\text{PW} \) solves a Hamiltonian formed from \( \mathcal{H}_0 + \gamma_{pp} \), where \( \gamma_{pp} \) is the pseudopotential term, which represents the nuclei and the core electronic orbitals. Finally, only the valence orbitals, represented by \( \phi_k^\text{PW} \), are sought.

The above demonstration, although illustrative, does not give a general formula for the pseudopotential, \( \gamma_{pp} \). General-use pseudopotentials are generated for
each atomic element. There is a great deal of research into how one can produce pseudopotentials which give high accuracy. In this project the pseudopotentials of Goedecker, Teter and Hutter (GTH) are used [41, 42]. These are “norm-conserving” and “separable” pseudopotentials. Norm-conserving pseudopotentials are required to satisfy two conditions. Firstly, inside the pseudopotential cut-off radius, $r_c$, the norm of each pseudo wavefunction must be equal to the corresponding true wavefunction. That is, if the pseudowavefunction is $\phi_k$ and the true (all-electron) wavefunction is $\psi_k$, 

$$\int_{r < r_c} d^3r \phi_{k,j}^* \phi_{k',j} = \int_{r < r_c} \psi_{k,i}^* \psi_{k',i}.$$  \hspace{1cm} (2.40)

Secondly, the pseudowavefunction and all electron wavefunctions must be identical outside the cutoff radius.

### 2.3 Hybrid functionals

#### 2.3.1 Band Gaps and Self-Interaction error in Density functional theory

LDA and GGA density functionals suffer from a wide range of problems. The most immediate problem when modelling insulators is that LDA and GGA functionals can severely underestimate the electronic band gap [43]. The band gap can be defined in a number of ways. Firstly, there is the KS band gap. This is defined as the eigenvalue energy of the lowest unoccupied KS molecular orbital (LUMO), or the conduction band minimum (CBM), minus the eigenvalue energy of the highest occupied KS molecular orbital (HOMO), or the valence band maximum (VBM). Expressed mathematically this is

$$E_{\text{BG}}^{\text{KS}} = \epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}.$$  \hspace{1cm} (2.41)

Alternatively one can define the ‘fundamental band gap’. Writing the electron affinity (EA) as $E(N) - E(N + 1)$ and the ionisation potential (IP) as $E(N - 1) - E(N)$,
the fundamental band gap can be expressed as

\[ E_{\text{BG}}^{\text{Fundamental}} = \text{IP} - \text{EA} = E(N + 1) + E(N - 1) - 2E(N). \] (2.42)

Finally, one can define an optical band gap. The optical band gap is the energy required to optically excite an electron from the VBM into the CBM. This leaves a hole in the valence band. Typically the optical band gap will be smaller than the KS band gap or fundamental band gap due to Coulomb interaction between the electron and hole.

It is found that LDA and GGA functionals severely underestimate all these quantities. This phenomenon has typically been named the self-interaction problem. At the GGA level, each electron has a spurious Coulombic interaction with itself. The ideal XC functional should cancel out this self-interaction, however in most basic approximations this cancellation does not happen. Note that in Hartree-Fock theory, the self interaction of the electron-electron interaction is canceled out by the exchange term.

Another problem is that the ground state electron densities predicted by LDA and GGA functionals tend to underestimate the localisation of electrons and holes. These functionals fail to predict the existence of polaronic states such as self-trapped holes. This phenomenon has been linked to the lack of piecewise linearity in local and semi-local DFT [44]. The ideal DFT functional should produce piecewise-linear behaviour in the total energy vs. electron number, or \( E(N) \), graph. The energy varies linearly for fractional electron numbers, except at integer values of \( N \) where the gradient jumps discontinuously. Thus the \( E(N) \) vs. \( N \) graph is a series of linear segments connected at integer values of \( N \) [45, 46]. The LDA functional, however, shows convex dependence of \( E \) on \( N \). Similar results are found for GGA functionals.

One solution to these problems in DFT is to mix in a fraction of Hartree-Fock exact exchange in the exchange-correlation (XC) functional. It has been found that this produces much more accurate predictions for band gaps [47, 48] and also avoids
2.3. Hybrid functionals

the delocalisation error.

2.3.2 Mathematical form of hybrid functionals

In a hybrid functional a percentage of the GGA-level exchange is replaced with a percentage of the exact exchange. Firstly, the DFT XC functional is split into its correlation and exchange parts,

\[ E_{XC} = E_{C}^{\text{DFT}} + E_{X}^{\text{DFT}}. \]  \hspace{1cm} (2.43)

Exchange in hybrid-DFT is calculated by taking a fraction of the DFT exchange, \( a \), and replacing it with that same fraction of exact exchange,

\[ E_{X}^{\text{Hybrid}} = (1-a)E_{X}^{\text{DFT}} + aE_{X}^{\text{HF}}. \]  \hspace{1cm} (2.44)

The full expression for a simple hybrid-DFT functional is then

\[ E_{XC}^{\text{Hybrid}} = (1-a)E_{C}^{\text{DFT}} + aE_{X}^{\text{HF}} + E_{C}^{\text{DFT}} \]  \hspace{1cm} (2.45)

Common examples of hybrid functionals are B3LYP [49, 50] and PBE0 [51, 52]. PBE0 mixes 25% Hartree-Fock exchange into the PBE XC functional. Expressed mathematically;

\[ E_{XC}^{\text{PBE0}} = 0.75E_{X}^{\text{PBE}} + 0.25E_{X}^{\text{HF}} + E_{C}^{\text{PBE}} \]  \hspace{1cm} (2.46)

2.3.3 Auxiliary Density Matrix Method

Calculation of the Hartree-Fock exchange energy is computationally expensive. If we express the molecular orbitals as

\[ \psi_i(r) = \sum_{\mu} C_{\mu i} \phi_{\mu}(r) \]  \hspace{1cm} (2.47)

then the Hartree fock exchange energy is calculated as

\[ E_{X}^{\text{HFX}}[P] = -\frac{1}{2} \sum_{\lambda \sigma} \sum_{j \nu} p^{\mu \sigma} P^{\nu \lambda}(\mu \nu | \lambda \sigma) \]  \hspace{1cm} (2.48)
where $P^{\mu \nu}$ are the density matrix elements and are obtained from the molecular orbital coefficients:

$$P^{\mu \nu} = \sum_i C^{\mu i} C^{\nu i}$$  \hspace{1cm} (2.49)

and $(\mu \nu|\lambda \sigma)$ is the electron repulsion integral,

$$(\mu \nu|\lambda \sigma) = \int \int \phi_{\mu}(\mathbf{r}) \phi_{\nu}(\mathbf{r}) g(|\mathbf{r}' - \mathbf{r}|) \phi_{\lambda}(\mathbf{r}') \phi_{\sigma}(\mathbf{r}') d\mathbf{r} d\mathbf{r}'$$  \hspace{1cm} (2.50)

where the standard expression for $g(|\mathbf{r}' - \mathbf{r}|)$ is the Coulomb operator $(\frac{1}{|\mathbf{r}' - \mathbf{r}|})$.

The auxiliary density matrix method (ADDM) [53] is employed to increase the computational speed of exact exchange calculation. This is achieved by replacing the primary Gaussian basis set with a smaller set of faster decaying basis functions (the auxiliary basis set). This decreases the computational cost to evaluate the electron repulsion integrals (ERIs). The auxiliary density matrix, $\hat{P}$, is constructed from the primary density matrix, $P$. Calculation of the exchange energy using the auxiliary basis set introduces an error. The ADMM approximation accounts for this error by assuming it can be accurately calculated using a GGA level of approximation. Stated mathematically,

$$E^{\text{HFX}}[P] \approx E^{\text{HFX}}[\hat{P}] - (E^{\text{GGA}}[P] - E^{\text{GGA}}[\hat{P}]).$$  \hspace{1cm} (2.51)

Calculation of the exact exchange, therefore, requires this extra step (calculating the error). For small systems, this extra step can outweigh the efficiency gains made by substituting for the auxiliary basis set. For larger systems, however, ADMM can reduce the computational cost of a hybrid functional calculation by orders of magnitude.

### 2.3.3.1 ADMM Basis Sets

The auxiliary basis set can be written as

$$\hat{\psi}_i(\mathbf{r}) = \sum_{\mu} \hat{C}^{\mu i} \hat{\phi}_\mu(\mathbf{r})$$  \hspace{1cm} (2.52)
such that the auxiliary density matrix can then be written
\[
\hat{\rho}^{\mu\nu} = \sum_i \hat{C}^{\mu i} \hat{C}^{\nu i}.
\] (2.53)

Then one needs to find an optimal set of auxiliary MO coefficients, \(\{\hat{C}^{\mu i}\}\). This can be done by requiring that the square difference between the molecular orbitals expressed in the primary basis and the molecular orbitals expressed in the auxiliary basis,
\[
M = \sum_j \int (\psi_j^2(r) - \hat{\psi}_j^2(r))^2 dr
\] (2.54)
is minimised. Substituting the expression for the molecular orbitals and collecting summations, this quantity can be expressed as
\[
M = \sum_j \left( \sum_{\mu\nu} C^{\mu j} C^{\nu j} \int \phi_\mu \phi_\nu dr + \sum_{\mu\nu} \hat{C}^{\mu j} \hat{C}^{\nu j} \int \hat{\phi}_\mu \hat{\phi}_\nu dr - 2 \sum_{\mu\nu} C^{\mu j} \hat{C}^{\nu j} \int \phi_\mu \hat{\phi}_\nu dr \right)
\] (2.55)

By writing
\[
S_{nn'} = \int \phi_n(r) \phi_{n'}(r) dr \quad \text{and} \quad \hat{S}_{nn'} = \int \hat{\phi}_n(r) \hat{\phi}_{n'}(r) dr
\] (2.56)
and
\[
Q_{nm} = \int \hat{\phi}_n(r) \phi_m(r) dr
\] (2.57)

\(M\) can be expressed as
\[
\sum_j \left( \sum_{\mu\nu} C^{\mu j} C^{\nu j} S_{\mu\nu} + \sum_{\mu\nu} \hat{C}^{\mu j} \hat{C}^{\nu j} \hat{S}_{\mu\nu} - 2 \sum_{\mu\nu} C^{\mu j} \hat{C}^{\nu j} Q_{\nu\mu} \right)
\] (2.58)
or, in matrix notation,
\[
M = \text{Tr} \left( C^T SC + \hat{C}^T \hat{S} \hat{C} - 2 \hat{C}^T QC \right)
\] (2.59)
Minimising $M$ with respect to $\hat{C}$ leads to the condition

$$\hat{C} = AC \tag{2.60}$$

where $A = \hat{S}^{-1}Q$. Further, it can be required that the auxiliary basis set be orthonormal. In this case, the expression to optimise becomes

$$M = \sum_j \int (\psi_j(r) - \tilde{\psi}(r))^2 dr + \sum_{k,l} \Lambda_{kl} \left( \int \tilde{\psi}_k(r) \tilde{\psi}_l(r) dr - \delta_{kl} \right) \tag{2.61}$$

where $\Lambda_{kl}$ are the Lagrange multipliers. This gives the condition

$$\tilde{C} = \hat{C} \Lambda^{-\frac{1}{2}} \tag{2.62}$$

with $\Lambda = \hat{C} \hat{S} \hat{C}$, with $\Lambda$ the matrix of Lagrange multiplies, $\Lambda_{kl}$. Using this condition reasonable approximations can be made for a auxiliary density matrix. There are a number of ways of doing this, which are described fully in [53].

### 2.4 The PBE0-TC-LRC Functional

The PBE0-TC-LRC functional is implemented in CP2K. A full description of its implementation can be found in [54], here a brief overview is given.

In a periodic system, with a given $k$-point mesh, the Hartree-Fock exchange is calculated as

$$E_X = -\frac{1}{2 N_k} \sum_{i,j} \sum_{k,k'} \int \int \psi_i^k(r) \psi_j^{k'}(r) g(|r - r'|) \times$$

$$\psi_i^k(r') \psi_j^{k'}(r') d^3r d^3r' \tag{2.63}$$

with, conventionally, $g(|r - r'|) = \frac{1}{|r - r'|}$. In a $\Gamma$-point calculation (see section 2.6.1), eqn. 2.63 reduces to

$$E_X^\Gamma = -\frac{1}{2} \sum_{i,j} \int \int \phi_i^0(r) \phi_j^0(r) g(|r - r'|) \times$$

$$\phi_i^0(r') \phi_j^0(r') d^3r d^3r' \tag{2.64}$$
In the PBE0-TC-LRC functional, the potential \( g(|r - r'|) \) is replaced with the truncated Coulomb operator,

\[
g_{\text{TC}}(|r - r'|) = \begin{cases} 
\frac{1}{|r - r'|}, & |r - r'| \leq R_c \\
0, & |r - r'| > R_c
\end{cases}
\]  

(2.65)

Beyond the cutoff radius, \( R_c \), a ‘long range correction’ (LRC) is applied to replace the lost long-range exchange interaction. This correction is calculated using a semilocal density functional. Specifically, the spherically averaged PBE exchange hole, \( \rho_{\text{SA}}^X(r, u) \), is used [55]. The long range correction to the exchange energy is then

\[
E_{\text{X}}^{\text{PBE-LRC}}[\rho] = \frac{1}{2} \int \rho(r)dr \int_{R_c}^{\infty} 4\pi u \rho_{\text{SA}}^X(r, u)du,
\]  

(2.66)

where \( u \) is the electron-electron distance. This then enables one to write the full expression for the exchange correlation energy as

\[
E_{\text{X}}^{\text{PBE0-TC-LRC}} = aE_{\text{X}}^{\text{HF-TC}} + aE_{\text{X}}^{\text{PBE-LRC}} + (1 - a)E_{\text{X}}^{\text{PBE}} + E_{\text{C}}^{\text{PBE}},
\]  

(2.67)

where \( E_{\text{X}}^{\text{HF-TC}} \) is equation 2.64 with \( g(|r - r'|) \) substituted with the truncated Coulomb interaction, \( g_{\text{TC}} \) and \( a \) is the fraction of exact exchange.

## 2.5 Time Dependent Density functional theory

The two Hohenberg-Kohn theorems provide the logical foundation for DFT. However, these theorems are only applicable in the stationary case. They do not prove that we are justified using a density-functional approach to calculate dynamic phenomena such as optical transitions. The foundational theorem of time-dependent DFT (TDDFT) is the Runge-Gross theorem. There are two major differences between the stationary and time-dependent case of DFT. Firstly, there is no variational principle for the time dependent Schrödinger equation. Secondly, in the time dependent case the wavefunction at some arbitrary time \( t \) will depend on its initial value at some earlier arbitrary time \( t_0 \). Therefore the initial state must always be considered.
2.5. Time Dependent Density functional theory

2.5.1 Runge-Gross Theorem

Runge and Gross demonstrated the validity of TDDFT in 1984 [56]. They proved that if two systems begin in the same state (electron densities are the same at time $t_0$) and are subject to time-dependent potentials that are different such that,

$$v(r, t) - v'(r, t) \neq c(t)$$  \hspace{1cm} (2.68)

their electron densities, $n(r, t)$ at time $t$ will be different. The Runge-Gross theorem applies only to potentials of the form

$$v(r, t) = \sum_{n=0}^{\infty} \frac{v_n(r)}{n!} (t - t_0)^n$$  \hspace{1cm} (2.69)

that is, those which can be expanded as a Taylor series in time. The Runge-Gross theorem proves that there is a one-to-one correspondence between the time dependent potential and the time dependent electron density, which allows one to write

$$v(r, t) = v[n, \Psi_0](r, t)$$  \hspace{1cm} (2.70)

where $\Psi_0$ is the initial state. One can simplify further if one chooses the ground state, $\Psi_{GS}$, as the initial state $\Psi_0$. By the Hohenberg-Kohn theorems, $\Psi_{GS}$ is also a functional of the electron density, and so the potential $v(r, t)$ becomes a functional of the electron density alone.

2.5.2 The Time Dependent Kohn-Sham Equations

Ground state DFT finds the electron density by solving a fictitious non-interacting system. An analogous process exists for TDDFT, where one calculates the time-dependent density of a non-interacting system. The time dependent density, $n(r, t)$, is expressed as

$$n(r, t) = \sum_{j=1}^{N} |\phi_j(r, t)|^2$$  \hspace{1cm} (2.71)

where $\{\phi_j(r, t)\}$ is the set of single-particle orbitals of the fictitious system (termed the time-dependent Kohn-Sham orbitals). These orbitals then satisfy the time-
dependent Kohn-Sham equations:

\[
 i \frac{\partial}{\partial t} \phi_j(\mathbf{r}, t) = \left( -\nabla^2 + v_{\text{KS}}(\mathbf{r}, t) \right) \phi_j(\mathbf{r}, t)
\]

(2.72)

for all \(j\). \(v_{\text{KS}}(\mathbf{r}, t)\) is the time-dependent Kohn-Sham potential.

### 2.5.3 Time-Dependent XC Functionals - The Adiabatic Approximation

The time dependent XC potential is a functional of the electron density and the initial state of the system. As explained above, if the initial state is the ground state the potential can just be regarded is a functional of the time dependent electron density. The functional is still highly complex, however. The XC potential will depend on the density at all other locations and at all previous times. The adiabatic approximation offers a simplification. In this approximation, the time-dependent XC potential is just a ground state DFT functional evaluated instantaneously at time \(t\), that is

\[
v_{\text{XC}}^{\text{Adiab}}(\mathbf{r}, t) = v_{\text{XC}}[n_0](\mathbf{r})|_{n_0(\mathbf{r})=n(\mathbf{r},t)}.
\]

(2.73)

The adiabatic approximation will be valid as long as there are not significant rapid fluctuations in the electron density.

### 2.5.4 Implementation of TDDFT in CP2K

CP2K uses the well established Linear Response formulation of the TDDFT within the adiabatic approximation. The TDDFT equations can be cast in the form of a non-Hermitian eigenproblem [57]:

\[
\begin{pmatrix}
A & B \\
B^* & A^*
\end{pmatrix}
\begin{pmatrix}
X_p \\
Y_p
\end{pmatrix} = \omega_p
\begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}
\begin{pmatrix}
X_p \\
Y_p
\end{pmatrix},
\]

(2.74)

where \((X_p, Y_p)\) is an eigenvector and \(\omega_p\) is a corresponding transition energy.

In terms of Kohn-Sham orbitals \(\{\phi\}\), the elements of the matrices \(A\) and \(B\) can (for a hybrid functional) be written in form [58]:

\[ A_{i\sigma,j\tau} = \delta_{ij} \delta_{ab} \delta_{\sigma\tau} (\epsilon_{i\sigma} - \epsilon_{j\sigma}) + (i\sigma a_{\sigma} | j\tau b_{\tau}) \]
\[ - \delta_{\sigma\tau} c_{\text{HF}} (i\sigma j\sigma | a_{\tau} b_{\tau}) + (1 - c_{\text{HF}}) (i\sigma a_{\sigma} | f_{\sigma\tau} | j\tau b_{\tau}), \] (2.75)
\[ B_{i\sigma,j\tau} = (i\sigma a_{\sigma} | b_{\tau} j\tau) - \delta_{\sigma\tau} c_{\text{HF}} (i\sigma b_{\sigma} | a_{\tau} j\tau) + (1 - c_{\text{HF}}) (i\sigma a_{\sigma} | f_{\sigma\tau} | b_{\tau} j\tau). \] (2.76)

Where \( c_{\text{HF}} \) is the fraction of Hartree-Fock exchange, and the terms for \( A_{i\sigma,j\tau} \), from left to right, are the orbital energy difference, electron-hole Coulomb, exact-exchange, and exchange-correlation (XC) terms, while notations \((i\sigma a_{\sigma} | j\tau b_{\tau})\) and \((i\sigma a_{\sigma} | f_{\text{sc,}\sigma\tau} | j\tau b_{\tau})\) stand for four-centre Electron Repulsion Integrals (ERIs) and XC-integrals over Kohn-Sham orbitals \(\{\phi\}\):

\[ (i\sigma a_{\sigma} | j\tau b_{\tau}) = \int \phi_{i\sigma}^*(r) \phi_{a\sigma}(r) \frac{1}{|r - r'|} \phi_{j\tau}^*(r') \phi_{b\tau}(r') drdr' \] (2.77)
\[ (i\sigma a_{\sigma} | f_{\text{sc,}\sigma\tau} | j\tau b_{\tau}) = \int \phi_{i\sigma}^*(r) \phi_{a\sigma}(r) f_{\text{sc,}\sigma\tau}(r,r',r',r') \phi_{j\tau}^*(r') \phi_{b\tau}(r') drdr'. \] (2.78)

\( f_{\text{sc,}\sigma\tau} \) is the exchange-correlation kernel, which can be defined as

\[ f_{\text{sc,}\sigma\tau} = \left. \frac{\delta v_{\text{xc}}[n(r,t)]}{\delta n(r',t')} \right|_{n_0(r)} \] (2.79)

Like many implementations of TDDFT, CP2K uses the Tamm-Dancoff approximation [59]. In this approximation, the matrix \( B \) is set to 0 and the Casida equation is reduced to

\[ AX_p = \omega_p X_p, \] (2.80)

which is a standard eigenvalue problem.

### 2.6 Practical Considerations in DFT Calculations

#### 2.6.1 Periodic Boundary conditions

All simulations in this thesis use periodic boundary conditions. In crystal structures one has to consider how they are to calculate the band structure of the material. This
calls for the need to sample the first Brillouin zone (henceforth called the Brillouin zone) at different \( k \) points. This is usually done within a Monkhurst-Pack [60] scheme. The Brillouin zone is represented by ‘\( k \)-point mesh’. A given band is represented discretely (an orbital at each \( k \)-point is calculated). Integration over the whole Brillouin zone is approximated as a weighted sum over each \( k \) point. A given \( k \)-point mesh is represented by the notation \( k_a \times k_b \times k_c \), with \( k_a \) being the number of \( k \)-points sampled along the \( a^* \) reciprocal lattice vector in the Brillouin zone, and likewise for the other terms.

CP2K is a Γ point code. This means that each band is only sampled at the point \( k = 0 \) and other \( k \)-points are not explored. Therefore, no dispersion relation \( (E(\mathbf{k})) \) is gained. Calculating the electronic structure of a periodic solid with sampling only at \( k=0 \) will lead to a very poor description of the electronic density. This problem is overcome by using ‘supercells’. Instead of simulating a single unit cell of a given material, the unit cell is repeated in each direction a number of times. Throughout this thesis the following notation will be used: a \( N_a \times N_b \times N_c \) supercell refers to a periodic simulation cell which has been constructed by repeating the unit cell \( N_a \) times in the \( a \) direction, \( N_b \) times in the \( b \) direction, and \( N_c \) times in the \( c \) direction.

This is a linear transformation, taking the basis vectors \( \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \) into new basis vectors \( \mathbf{a}_1', \mathbf{a}_2', \mathbf{a}_3' \). This can be written

\[
\mathbf{a}_j' = \sum_{i=1}^{3} T_{ji} \mathbf{a}_i \quad j = 1, 2, 3
\]  

For the primitive unit cell, the reciprocal lattice vectors can be expressed as [61]

\[
\mathbf{b}_i = 2\pi \sum_{k=1}^{3} (\mathbf{G}^{-1})_{ik} \mathbf{a}_k
\]  

where \( \mathbf{G} \) is the metric tensor, defined \( G_{ik} = \mathbf{a}_i \cdot \mathbf{a}_k \). In this work, all major results come from simulations of at least \( 3 \times 3 \times 3 \) supercells. This achieves a better description of the electron density due to an effect known as band folding. Expansion of the cell into a supercell linearly transforms not only the lattice vectors but also the reciprocal lattice vectors [62, 61]. The reciprocal lattice vectors after the linear
transformation $T$ are

$$b'_i = \sum_k (T^T)_{ik}^{-1} b_k$$  \hspace{1cm} (2.83)$$

It is the inverse of the (transposed) transformation matrix which transforms the reciprocal lattice vector. Thus, if the direct lattice is expanded, there will be a contraction of the Brillouin zone. This has the effect of “folding” certain $k$ points into the new Brillouin zone. If one expresses the $k$ point coordinates as a fraction of the reciprocal lattice vectors, $k = (k_1, k_2, k_3)$, then the coordinates will be transformed as

$$k' = k \cdot T^T$$  \hspace{1cm} (2.84)$$

Once the cell has been transformed and the Brillouin zone contracted, some $k$ points which are now outside the Brillouin zone (but weren’t before) become equivalent to points inside the Brillouin zone. These points are connected by

$$k' = k'_{BZ} + \sum_{j=1}^{3} m_{ij} b'_j,$$  \hspace{1cm} (2.85)$$

where $m_{ij}$ are zero or integers. The point $k'$, outside the transformed Brillouin zone, is equivalent to the point $k'_{BZ}$ inside it. For example, the gamma point ($k=(0,0,0)$) in a $2 \times 2 \times 2$ extension becomes identical to 8 other points, including $(1/2,0,0)$, $(1/2,1/2,0)$ and $(1/2,1/2,1/2)$ (expressed as fraction of the untransformed vectors, $b$). Thus, by symmetrically expanding the supercell, greater $k$-point sampling is achieved as many $k$ points become equivalent. Thus, when applied to a supercell, a given $k$-point mesh in fact samples a greater amount of the (primitive) Brillouin zone than if it had been applied to the primitive direct cell. It is therefore the case that one can achieve a good description of the electron density, even in the $\Gamma$-point approximation, by taking a large enough, symmetrically extended supercell.

### 2.6.2 Defect Calculations

This project is predominantly a study of defects in insulating systems. One of the important properties of a defect is its formation energy. In this thesis, the formation
energy of a defect is defined as

\[ E_{\text{Form.}} = E_{\text{Def.}} - E_{\text{Cryst.}} + \sum_i n_i \mu_i + q(E_{\text{VBM}} + E_{\text{Fermi}}) + V_{\text{Corr.}} \] (2.86)

Where \( E_{\text{Def.}} \) is the energy of the system containing a defect, \( E_{\text{Cryst.}} \) is the energy of the pure crystal, \( \mu_i \) is the chemical potential of an added or removed atom (or chemical species), \( n_i \) is the number of such atoms added or removed. \( q \) is the charge state of the defect, \( E_{\text{VBM}} \) is the energy of the top of the valence band, and \( E_F \) is the Fermi level referenced to the top of the valence band. \( V_{\text{Corr.}} \) is the set of corrections which must be applied to defect calculations in periodic boundary conditions, discussed later.

The formation energy depends on the choice of chemical potential, \( \mu_i \). The chemical potential depends on the environment with which atom types are exchanged. For example, in oxygen monovacancy calculations it is common to choose the gas phase of \( \text{O}_2 \) as a reference for the chemical potential. The formation energy will then correspond to the cost of creating an oxygen vacancy by removing oxygen atoms from the solid system and creating an \( \text{O}_2 \) molecule in a reservoir. Chemical potentials relate to particular conditions in which defect formation happens.

The charge dependent terms, \( q(E_{\text{VBM}} + E_{\text{Fermi}}) \), account for the energy cost to charge the defect. The Fermi level acts as a chemical potential for electrons. The value for the Fermi level depends on the reservoir with which electrons are exchanged. For example, the reservoir may be a metallic substrate on which an insulating solid sits. Charging of a defect in the insulator may be caused by transfer of electrons between the metal and the defect.

### 2.6.3 Charged Defect Corrections

It was stated above that all results presented in this thesis are calculated within periodic boundary conditions. This means that defects will be periodically repeated.
For example, placing a defect in a $3 \times 3 \times 3$ supercell will produce a defect which is repeated every 3 unit cell vectors (in each direction). Take, as an example, a cubic lattice with $a=4.2$ Å (the lattice vector of MgO). A $3 \times 3 \times 3$ supercell has a lattice vector of 12.6 Å in each direction and a volume $2000 \, \text{Å}^3 = 2 \times 10^{-21} \, \text{cm}^3$. Therefore, if the simulation cell contains a single defect, one is in fact simulating a system with a defect concentration of $5 \times 10^{20} \, \text{cm}^{-3}$. This is a very large concentration of defects. This is problematic if one is interested in the isolated defect. With a very high concentration, defects will interact with one another and this will affect their behaviour. For neutral defects, this interaction is not strong, and the defects can be regarded as having approached the isolated defect limit. For charged defects however, the long-range Coulomb interaction can affect defect formation energies even for the large supercells considered here.

Another related problem is that of the compensating jellium background. As stated above, a defect is in fact simulated as a periodic array of defects. If the defect is charged, this leads to an infinitely large number of charged defects, all interacting with each other. This gives an infinite interaction energy. To account for this, a compensating uniform background charge is included in the cell. For example, if simulating a +1 charged defect, a uniform -1 background charge is inserted into the simulation cell. This prevents the total energy from diverging, but again affects the properties of the defect.

Theses two effects need to be removed from the calculations. This is usually done in a post hoc fashion. Several approaches have been developed which take the DFT calculated defect (with all charge image and jellium interaction) and then correct the total energy [63, 64, 65, 66, 67]. The correction process will depend on the static dielectric constant, $\varepsilon_0$, as the dielectric response of the material screens the charged defects from one another. In this project the Lany-Zunger (LZ) method is used to correct the charged supercell calculations [64].
2.6.4 Potential Alignment.

Another issue is the problem of potential alignment. Creating defects in the simulation cell shifts the electrostatic potential. Thus, charging the cell introduces an arbitrary energy shift which needs to be calculated and removed from the final value for energy. In this project, this is achieved by averaging the electrostatic potential across planes of atoms, and then investigating the behaviour of the electrostatic potential far from the defect. It is found that, far from the defect, the pure cell electrostatic potential and charged-defect electrostatic potential do not approach each other, but are instead shifted. This is the shift of the electrostatic potential. This number is calculated for all defect calculations and the correction, $q\Delta V$, applied to the final energy. This process is said to “realign” the potentials.

2.6.5 Nudged Elastic Band Calculations

In order to study reaction rates, or phenomena such as defect generation and diffusion, it is necessary to investigate the system’s potential energy surface (PES) near the defect configuration. For example, an oxygen vacancy moving from one location to another requires the system to move through a PES, which will usually require the overcoming of an energetic barrier. The value of this energetic barrier is important for understanding what the rate for a given defect process is. To find the most physical pathway through the PES (between an initial and a final state), a minimum energy pathway (MEP) must be found. The climbing image nudged elastic band (CI-NEB) technique [68, 69, 70] achieves this.

Each pathway, or “band”, is a set of $N + 1$ configurations $[C_0, C_1, ..., C_N]$. The end points, $C_0$ and $C_N$, are the initial and final configurations and are fixed. In the NEB method, forces are calculated on the configuration which optimise the pathway to give the MEP. The forces calculated, however, are not the total force but are a force projection. Letting the tangent of the pathway at replica $i$ be $\tau_i$, only the component of the force, $\nabla E(C_i)$, perpendicular to $\tau_i$ is used. In addition, each
configuration, $C_i$ experiences a spring force defined

$$F_i^{\text{Spring}} = k(|R_{i+1} - R_i| - |R_i - R_{i-1}|),$$

(2.87)

where $k$ is the spring constant. A nudged elastic band is then understood to be a set of configurations along a pathway, between two fixed point, where each configuration is connected to its adjacent configurations by a fictitious spring. To ensure that the spring does not affect the converged MEP, only the component of $F_i$ parallel to $\tau_i$ is used. The total force experienced by each replica is

$$F_i = (F_i^{\text{Spring}} \cdot \tau_i) \tau_i + (\nabla E(C_i) - \nabla E(C_i) \cdot \tau_i).$$

(2.88)

The force on each replica is minimised, giving the MEP. There is no guarantee, however, that the highest energy replica will be at or close to the saddle point. To ensure this, the CI-NEB method is used. In CI-NEB, after a number of ordinary NEB iterations, the force on the highest energy image is instead given by

$$F_m = -\nabla E(C_m) + 2(\nabla E(R_m) \cdot \tau_i) \tau_i,$$

(2.89)

which is simply the full force due to the PES but with the component parallel to the pathway inverted. This causes the image to move up the PES and brings it closer to the saddle point.
Chapter 3

Oxygen Vacancies in MgO

3.1 Introduction & Background

In this chapter oxygen vacancies in MgO are studied. The main result is the demonstration that a novel implementation of time-dependant density functional theory (TDDFT) calculates the optical spectra of the MgO oxygen vacancy to within 0.3 eV of the experimentally detected peak. This gives assurance that the TDDFT functionality in CP2K is reliable. In addition to this, the photoluminescence (PL) of the oxygen vacancy is studied using the $\Delta$-SCF method. This offers a deeper insight into the excited states of the neutral oxygen vacancy.

3.1.1 Previous Studies on Oxygen Vacancies in MgO

MgO has long been studied as a model wide band gap oxide with numerous technological applications. Extensive experimental studies have been carried out on defects in MgO, and particularly the O-vacancy (also known as an F center) \[71, 72, 1, 73, 74, 75\]. Theoretical studies, however, still struggle to give reliable insights into the optical spectra of different charge states of this defect. For example, there is still disagreement on the exact positions of optical absorption peaks of neutral and positively charged vacancies. The first optical absorption peaks of the neutral and +1 charged oxygen vacancies ($V_0^0$ and $V_0^{+1}$ centers) occur at very similar energies: 5.01 and 4.96 eV, respectively \[1\] (See figure 3.1). These peaks are difficult to resolve both experimentally and in theoretical calculations. In \[1\], the peaks are resolved by arguing that continued irradiation of $V_0^0$ defects produces $V_0^{+1}$ de-
fects via photoconversion. This means that the experimentally detected peak should shift during irradiation, and this can be used to resolve the two charge states (Fig. 3.1(B)). *Ab initio* cluster model calculations of $V^0_O$ and $V^{+1}_O$ defects [76] predicted very similar optical absorption energies, however, the absolute value was closer to 6 eV, rather than the experimental 5 eV. It was argued that the main source of error in this study was the limited basis set size and that 3d polarization functions should be added to the oxygen basis set. Further calculations [77] using a larger basis set predicted the optical absorption energy for $V^0_O$ at 5.44 eV, in closer agreement with experiment. The optical excitation energy for the $V^{+1}_O$ defect, however, was not improved. It was therefore concluded that much larger basis sets are necessary for accurate prediction of the optical absorption spectrum. In both of these studies only excitations in the alpha spin channel were reported for the $V^{+1}_O$ center. However, there are, in fact, two types of optical excitation in the $V^{+1}_O$ center (see Fig. 3.2): a type-V transition into a higher state located close to the bottom of the conduction band (CB) in the $\alpha$-spin channel, and a type-III transition from the defect-induced valence band states into the unoccupied state associated with the $V^{+1}_O$ center in the $\beta$-spin channel.

The $\beta$ channel excitation was suggested in previous experimental studies [78, 79] as a possible origin of the 5 eV absorption peak. Interestingly, both the $\alpha$ and $\beta$ channel excitations (type V and III, correspondingly) in the $V^{+1}_O$ center are predicted to be close to 5 eV. Thus, all three optical absorption peaks associated with $V^0_O$ and $V^{+1}_O$ defects have very close energies. A qualitative difference between $\alpha$ and $\beta$ spin excitation for the $V^{+1}_O$ center, however, is that $\beta$ channel excitations should be associated with the release of holes into the valence band (VB). It has been shown that, when exciting MgO crystals with 5 eV light, a 2.3 eV absorption band also develops, which is associated with holes trapped on V-type centers [78, 79]. Recent periodic Density Functional Theory (DFT) calculations [80] account for different spin channel excitations in MgO using many-body perturbation theory in the $G_0W_0$ approximation and the Bethe-Salpeter approach. These calculations predicted the existence of a 3.6 eV absorption band in the $V^{+1}_O$ defect (in addition to the known
3.1. Introduction & Background

Figure 3.1: (A) Experimentally measured optical absorption spectra of MgO. The peaks correspond to the first optical absorption peak of the $V_{O}^{0}$ and $V_{O}^{1+}$ defect. (B) Optical absorption results with the separate peaks more clearly resolved. Irradiation causes the absorption line to shift down in energy slightly, as a result of the $V_{O}^{0} \rightarrow V_{O}^{1+}$ photoconversion process. Both images are retrieved from [1].

Figure 3.2: Schematic of the typical optical transitions in oxides with defects. Type-I is a band-to-band transition. Type-II is a VB to defect-induced unoccupied resonant state in CB transition. Type III represents the VB to unoccupied defect state in the bandgap transition. Type IV is an occupied defect state into defect-induced state in the CB transition, and type V is an occupied defect state into unoccupied defect state in the bandgap transition. A type VI transition is from a shallow, occupied defect state into the defect-induced resonant state in the CB.
5 eV band in the $\alpha$ channel), which could be used to distinguish between the $V_{\text{O}}^0$ and $V_{\text{O}}^{+1}$ defects. Thus the position and nature of optical absorption peaks of $V_{\text{O}}^0$ and $V_{\text{O}}^{+1}$ defects in MgO remain controversial.

MgO has a well-established photoluminescence (PL) band at 2.3 eV attributed to the $V_{\text{O}}^0$ center [81, 73, 75, 74, 82]. The life-time of this luminescence is much longer than that for the $V_{\text{O}}^{+1}$ center, which is caused by a single-electron transition in the doublet state. Therefore the 2.3 eV luminescence of the $V_{\text{O}}^0$ defect could be due to a partially-allowed $2s \rightarrow 1s$ type and/or triplet-to-singlet transition. The kinetics of this PL is, however, much more complicated than that of the $V_{\text{O}}^0$ defect in CaO, where the lowest relaxed excited state has been shown to have a tetragonal (100) triplet state with fast spin-lattice relaxation (see, for example, Ref. [83]). The semi-empirical[81] calculations suggested that tetragonal singlet and triplet states are both located very close to the bottom of the CB and that there are two minima in the relaxed excited state corresponding to $3T_{1u}$ and $3A_{1g}$ configurations, respectively. The predicted PL energies corresponding to transitions from these two minima are equal to 2.9 eV and 2.2 eV, respectively. The PL peaking at 2.9 eV has been observed experimentally [84] but has an excitation energy at about 7 eV. It has been concluded that the PL peaks at 2.3 eV and 2.9 eV may involve both transitions intrinsic to the $V_{\text{O}}^0$ center and those caused by electron transfer from impurities [81, 73]. The dynamics of similar PL of anion vacancies in alkali halides has been studied in detail [85]. In that case, the transition is allowed due to mixing of $2s$ and $2p$ terms, with $2s$ having the lower energy in the relaxed excited state [85].

3.1.2 Details of Calculations

MgO was modelled using the PBE0-TC-LRC functional with its fraction of exact exchange set to 32.5 % and the cutoff radius set to 6 Å. This set up produces a KS band gap which is in good agreement with experimental values (section 3.2.1). The bulk structure was calculated in $3 \times 3 \times 3$, $4 \times 4 \times 4$ and $5 \times 5 \times 5$ cells. All defect calculations use a $4 \times 4 \times 4$ (512 atom) simulation cell, unless otherwise stated.

Since forces in excited states (calculated by TDDFT) are not yet implemented in CP2K, luminescence calculations were made using the $\Delta$-SCF method (Fig. 3.3).
3.2. Results

3.2.1 Bulk MgO

The 1000-atom simulation cell of bulk MgO is shown in Fig. 3.4. The lattice parameter and band gap of each cell size is given in table 3.1. Lattice parameters and band gap are in good agreement with experimental results. The band gap energies are calculated as the difference of KS energies and using TDDFT. The TDDFT
3.2. Results

Figure 3.4: (A) The bulk 1000 atom MgO simulation cell. MgO has a face centred cubic (FCC) crystal structure with a lattice parameter of 4.21 Å. (B) The projected density of states (PDOS) of the 1000 atom MgO bulk simulation cell. The calculated CBM state is 7.8 eV above the VBM, however the density of this state is too small to be seen in the PDOS plot.

<table>
<thead>
<tr>
<th>Cell Size</th>
<th>216</th>
<th>512</th>
<th>1000</th>
<th>Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>4.21</td>
<td>4.21</td>
<td>4.21</td>
<td>4.211 [88]</td>
</tr>
<tr>
<td>KS BG</td>
<td>7.7</td>
<td>7.7</td>
<td>7.8</td>
<td>-</td>
</tr>
<tr>
<td>Optical BG</td>
<td>7.1</td>
<td>7.4</td>
<td>7.5</td>
<td>7.78 [72]</td>
</tr>
</tbody>
</table>

Table 3.1: The lattice parameters (Å) and band gap values (eV) for bulk MgO. The optical band gap (BG) is calculated using TDDFT. Cell sizes are given in number of atoms.

calculated optical band gap is lower than the KS band gap. This is because, unlike in GGA-based TDDFT, in hybrid TDDFT there is electron-hole interaction [89]. Excitation between delocalised band states creates an electron-hole pair which is confined within the simulation cell and thus has an artificially high electron-hole interaction energy. This explains why the difference between the KS and TDDFT band gap is greatest for small simulation cells where the electron-hole pair is more confined and thus the electron-hole interaction is greater.
3.2. Results

**Figure 3.5:** (A) The $V_0^0$ defect in MgO. Blue surface indicates the iso-surface of the KS defect state. (B) The $V_0^{+1}$ defect in MgO. The top image shows the occupied KS defect state (blue iso-surface). The bottom image shows the unoccupied KS defect state (orange iso-surface). (C) The $V_0^{+2}$ defect in MgO. The orange surface indicated an iso-surface of the unoccupied KS defect state. All iso-surface values are 0.001.

### 3.2.2 Geometric And Electronic Structure of Oxygen Vacancies in MgO

The removal of a neutral oxygen atom in the MgO periodic cell leaves behind a neutral oxygen vacancy, $V_0^0$. Two electrons localise on the vacancy, occupying two degenerate mid-gap levels (one state for each spin) which are located 4.45 eV below the CBM. The neutral vacancy is shown in Fig. 3.5(A). There is very little structural reorganisation since the Coulomb interaction between the vacancy and the surrounding ions is very similar to the Coulomb interaction between a O$^{−2}$ ion and its surrounding Mg$^{+2}$ ions in the bulk lattice.

The oxygen vacancy introduces a perturbation to the lattice which creates quasi-localised states in the conduction band (Fig. 3.2), some of which are plot-
3.2. Results

### Ionic Displacements

<table>
<thead>
<tr>
<th>MgO</th>
<th>ΔMg</th>
<th>ΔO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_O^0$</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>$V_O^{+1}$</td>
<td>0.11</td>
<td>0.03</td>
</tr>
<tr>
<td>$V_O^{+2}$</td>
<td>0.20</td>
<td>0.08</td>
</tr>
</tbody>
</table>

**Table 3.2:** Displacements of the NN and NNN in the O-vacancy in MgO. All displacements are given in Å.

Table 3.2. The states in Fig. 3.6 have a $p$-like form. In fact, many of the states introduced into the band by the vacancy look similar to hydrogenic orbitals. This is similar to the case in F-centers (anion vacancies) in alkali halides [90]. The ground state of the vacancy has an $s$-like character, the first excited state also has $s$-like character, and the three next excited states have $p$-like character.

Removing a single electron charges the vacancy and creates a $V_O^{+1}$-defect. The $V_O^{+1}$ defect is shown in Fig. 3.5(B). Two defect states are present. One is a mid-gap, singly occupied state similar to the defect state of the $V_O^0$ defect. The other is an unoccupied state with its KS energy eigenvalue higher in the band gap at 2.0 eV below the CBM. This splitting of the doubly-occupied state is the origin of the two excitation channels in the $V_O^{+1}$ defect. Charging the vacancy increases the Coulomb interaction between the vacancy and the surrounding ions. This increases the local atomic displacements. These displacements are given in Table 3.2.

Removing another electron creates a doubly positively charged oxygen vacancy, the $V_O^{+2}$ defect. This causes greater atomic displacements (See Table 3.2). The defect state of the $V_O^{+2}$ defect is shown in Fig. 3.5(C). There are two identical states (one for each spin direction), both of which are unoccupied. The KS level of this state is 2.6 eV below the CBM.

### 3.2.3 Optical Absorption of O Vacancies

Using TDDFT the optical absorption spectra of the oxygen vacancies can be calculated. To make the results more comparable with experiment, the spectrum has been smoothed out with Gaussian functions. If the set of excitation energies is $\{E_i\}$, and each excitation has oscillator strength $f_i$, then the smoothed spectrum, $F(E)$, is
3.2. Results

Figure 3.6: An isosurface of the $|\psi|^2$ of the localised states introduced into the conduction band by the oxygen vacancy. The states have a p-like character. Iso-value = 0.004.

calculated as

$$F(E) = \sum_i \frac{f_i}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(E - E_i)^2}{2\sigma^2}\right). \quad (3.1)$$

This form is chosen so that the integral of each excitation is normalised to its oscillator strength,

$$\int_{-\infty}^{\infty} dE \frac{f_i}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(E - E_i)^2}{2\sigma^2}\right) = f_i. \quad (3.2)$$

Unless stated otherwise, $\sigma$ is set to 0.2 eV.

The lowest energy transition of the MgO $V^0_O$ defect has zero oscillator strength. This can be understood if we analyse the $\Psi_i$'s which make up that excited state. The dominant $\Psi_i$ is a single excitation from the vacancy state into the LUMO. As described above, the LUMO has $s$-like character. Since both the defect state and LUMO state are dominantly $s$-like, this transition is forbidden by electric dipole selection rules. The next three higher energy transitions all have strong absorption lines and are equal in excitation energy. It is these optical absorptions which are responsible for the main excitation peak seen in $V^0_O$ defects in MgO (Fig. 3.7(A)). Analysis for these excited states show that each one is dominated by one of the three p-like states (see Fig. 3.6). Optical transitions of the $S \rightarrow P$ type are particularly
3.2. Results

Figure 3.7: Optical absorption spectra of oxygen vacancies in MgO. (A) is the $V_0^0$ defect in MgO. There is strong absorption peaking at 4.85 eV. This corresponds to a transition between an $s$-like and a $p$-like state. (B) is the $V_0^{+1}$ defect in MgO. The lower energy absorption peaks at 4.70 and comes from excitation of an alpha spin electron in a gap-state into CB states. The higher energy absorption peaks at 5.26 eV and comes from excitation of beta spin electrons in VB states into the unoccupied state in the gap. (C) is the $V_0^{+2}$ defect in MgO. Blue delta functions are the original, untampered excitation energies. The red lines give the spectrum smoothed out with Gaussians of standard deviation $\sigma = 0.2$ eV.

strong. This explains the high oscillator strength of these transitions. The peak in Fig. 3.7(A) is located at 4.85 eV. When compared to the experimental results from reference [1] (shown in Fig. 3.1) it can be seen that this is reasonably close to the experimental value of 5.03 eV.

It is found that the peak for absorption in the $V_0^{+1}$ defect is 4.70 eV, which is 0.15 eV lower than the $V_0^0$ peak. Experimentally (reference [1, 82]), this is also found to be the case, though the degree of difference is smaller. Reference [1] finds the positive vacancy absorption peak to be only 0.07 eV lower than the neutral, reference [82] finds the positive peak to be 0.09 eV lower. For both charge states, the peak energy is slightly underestimated by 0.2 to 0.3 eV.

For the $+1$ charge state peak, transitions are split into two types. There are transitions in the $\alpha$-spin channel (type IV) and transition in the $\beta$-spin channel (type III). In MgO this is expected to be particularly complicated since, as discussed above, the absorption energies in both spin channels have previously been predicted to be close to one another [78], potentially making the interpretation of optical absorption experiments difficult. In these calculations, however, the beta spin excitation is calculated to be higher in energy (Fig. 3.7(B)). It is therefore expected that the optical absorption spectrum of the $V_0^{+1}$ defect in MgO should show asymmetry in the 5 eV
peak. There will also be qualitative differences between the two types of excitation. Excitations in the alpha channel will excite an electron from the gap state into CB states, similar to the $V^0_O$ defect. Excitations in the beta channel, however, will release holes into the valence band. It is then possible for these holes to be trapped onto V-centres [78]. In reference [80] it is predicted that the beta absorption line should be much lower than the alpha spin absorption energy ($\approx 3$ eV). This absorption band is not found in these calculations. It is possible that the 3 eV absorption band in reference [80] emerges because of the small size used (64 atoms), which will lead to the atomic relaxations in the $V^{+1}_O$ centre being inadequately described. The effect of atomic relaxation on absorption energies can be tested by charging the $V^0_O$ centre to the +1 state but not allowing the structure to relax. Doing this brings the main $\beta$-spin absorption peak down to 3.5 eV. One can therefore conclude that failure to allow the defect to fully relax can change absorption energies by as much as 3 eV.

The $V^{+2}_O$ defect produces an unoccupied state high in the band gap (Type V). The optical absorption of the $V^{+2}_O$ defect is similar to optical absorptions for the $V^0_O$ and $V^{+1}_O$ defects. Therefore, all of the $V_O$ defects in MgO are predicted to have optical absorption lines near 5 eV. This implies that it will be difficult to distinguish between different charge states of the $V_O$ defect in MgO using optical experiments alone.

### 3.2.4 Photoluminescence of O Vacancy

As discussed above, the first transition of a $V^0_O$ defect is an $s \rightarrow p$ transition. The $V^0_O$ defect has cubic ($O_h$) symmetry. This means that the local modes associated with the $s \rightarrow p$ transition will have $A_{1g}, E_{1g}$ and $T_{2g}$ symmetry [91]. The photoluminescence (PL) of the vacancy in MgO is studied using $\Delta$-SCF. The lowest energy triplet state has $^3A_{1g}$ symmetry (see Fig. 3.8(B)). A higher energy triplet state is also found which has a geometry formed from a linear combination of $A_{1g}$ and $T_{2g}$ displacement modes (see Figs. 3.8(A) and 3.9(A)). The displacements of the high-energy triplet state are parallel to the 3-fold rotation axis of the vacancy octahedral. Since there are 4 such axes, there are 4 possible configurations of this state. This
3.3 Conclusions

A novel implementation of TDDFT in CP2K has been used to calculate the optical absorption spectrum of the oxygen vacancy in MgO. This spectrum is found to depend on charge state, and is in reasonable agreement (within 0.3 eV, see Fig. 3.1 and Fig. 3.7) with experimental data [1, 82]. In addition, the $V_{O}^{+1}$ defect is found to have an additional peak (due to excitation from the VB) which may explain the asymmetry seen in experimental spectra (Fig. 3.1). Compared to previous theoretical investigations (discussed in section 3.1.1), the method used in this chapter has been quite successful. This is due, in part, to the use of large cell size in the

**Figure 3.8:** The $|\psi|^2$ of the high energy triplet state of the $V_{O}$ defect. (A) gives the HOMO-1 state, visualised from each of the crystallographic directions. (B) shows the HOMO state, shown from the same crystallographic directions. All iso-surfaces have value $= 0.004$.

differs from the low energy triplet state, for which there is only one possible configuration. In the high energy triplet state, the HOMO axis is also aligned along the 3-fold rotation axis (see Fig. 3.8(B)). The calculated luminescence energy of the low energy state is 2.0 eV. The calculated PL energy from the high energy state is 2.9 eV. These calculations support the PL model suggested in earlier studies [81, 73] with surprising agreement of calculated PL energies.
3.3. Conclusions

**Figure 3.9:** The local modes associated with the \( s \rightarrow p \) transition of the \( V_0 \) defect in MgO. Orange circles represent the NN Mg ions. (A) Is the displacement of the high energy triplet state. There are extensions along the 3-fold rotation axis of the vacancy. The displacements are formed from a linear combination of the \( A_{1g} \)-symmetric and \( T_{2g} \)-symmetric local modes. (B) The local mode of \( A_{1g} \) symmetry. (C) The three local modes of \( T_{2g} \) symmetry.

Simulations. Such large cell sizes would not have been possible if a more intensive electronic structure method (such as Bethe-Salpeter) had been used. Therefore, it is demonstrated here that the combination of hybrid DFT with large super cells in the \( \Gamma \) point strikes a good balance between the need to use a large cell and the sophistication of the electronic structure calculations.

In addition, the photoluminescence properties of the \( V_0 \) defect in MgO were studied. Using the \( \Delta \)-SCF method, the structure of the excited states of a neutral vacancy were simulated. Two excited states were found. One (the lower energy state) was found to have \( A_{1g} \) symmetry and a PL peak of 2.0 eV. The second (the higher energy state), was found to be four-fold degenerate and is predicted to have a PL peak of 2.9 eV.

The major result of this chapter, therefore, is that the approach used to model optical processes provided reasonable results. This approach is used in chapters 4
and 5, which both include optical absorption predictions of certain defects. The rest of this thesis is focussed on HfO$_2$ and hexagonal BN, and will more directly consider the issues discussed in the Introduction.
Chapter 4

Carrier Trapping in Pristine HfO$_2$

4.1 Introduction & Background

Hafnium dioxide (HfO$_2$) is a wide band gap insulator with a number of different crystal phases. At standard pressure and temperature, the ground state of HfO$_2$ is monoclinic. HfO$_2$ can, however, exist in other crystalline phases. For example, the monoclinic-tetragonal phase transition of HfO$_2$ has been measured to occur at temperatures of 1500K to 2000K [92, 93, 94]. The monoclinic phase is shown in Fig. 4.3. O atoms are either 3 or 4 coordinated by Hf and Hf atoms are all 7 coordinated by O.

HfO$_2$ has a number of advantages which make it suitable as a gate dielectric. Firstly, it is a wide band gap insulator ($\approx$ 5.7 eV) and has a conduction band offset of $>$ 1 eV with Si [95]. This increases the barrier height for the process of electrons tunnelling from the Si layer to the gate. [96, 97]. As discussed in Chapter 1, its high dielectric constant means a large capacitance can be achieved with a thicker layer, again reducing the tunnelling current. If the direct tunnelling current is small, the primary source of leakage current will come from trap assisted tunnelling (TAT). This chapter, therefore, will begin the investigation into trap states of HfO$_2$ by studying the process of intrinsic trapping. Intrinsic trapping occurs without the assistance of a defect (such as a vacancy or impurity).

The most novel results of this chapter are the results for the amorphous phase. Amorphous gate insulators have been of growing interest for application as a gate
dielectric. The rationale comes from the understanding that grain boundaries, in polycrystalline gate dielectrics, act as a source of degradation problems. By using an amorphous film, grain boundaries are removed and this problem is avoided. In this chapter it will be shown that amorphous insulators can have their own, unique problems. It is demonstrated that electrons and holes can trap in pristine amorphous HfO₂ (without defects). These traps are stable at room temperature. This has major implications for the reliability and degradation of amorphous HfO₂ films in MOS devices.

4.1.1 Previous Calculations for Carrier Trapping in Monoclinic HfO₂

There have been a number of theoretical studies on electron and hole trapping in monoclinic (m-) HfO₂. Ramo et al. [98] calculated the structure of self-trapped holes (STHs) and electrons (STEs) in m-HfO₂ using the B3LYP hybrid functional [99]. Stable electron and hole trap states were both found. The STE has an electron localised on two Hf ions. There are two possible configurations of the STH owing to the two oxygen sub-lattices. The STH can be formed on either a 3-coordinated (3C) or 4-coordinated (4C) oxygen ion. In all cases the electron or hole localisation is facilitated by creating a distortion around an oxygen or hafnium ion, after which the system is relaxed. In both cases, nearest neighbour ions are stretched outwards by 0.1 to 0.2 Å.

The stability of a self-trapped electron or hole is related to its ‘trapping energy’. The trapping energy of a self-trapped electron is defined as

\[ E_{\text{Trapping}} = E_{\text{Delocalised } e^-} - E_{\text{STE}} \]  

(4.1)

and likewise for a hole. \( E_{\text{Delocalised } e^-} \) is the energy of the perfect crystal structure with one excess electron. \( E_{\text{STE}} \) is the energy of the structure with the electron localised and the lattice geometry allowed to fully relax. Positive values of \( E_{\text{Trapping}} \) indicate that self-trapping is energetically favourable. High values of trapping en-
ergy correspond to higher amounts of energy needed to thermally release the electron (hole) trap. Within a simple one electron model, the trapping energy can be split into two contributions [100],

\[ E_{\text{Trapping}} = -\left( \Delta E_{\text{Kinetic}} + \Delta E_{\text{Relaxation}} \right), \] (4.2)

where \( \Delta E_{\text{Kinetic}} \) and \( \Delta E_{\text{Relaxation}} \) are, respectively, the kinetic energy change and relaxation energy due to localisation. Localising an electron increases its kinetic energy, therefore \( \Delta E_{\text{Kinetic}} \) is positive. In addition, the lattice lowers its energy by responding to the localisation of charge (via polarisation). Thus, \( \Delta E_{\text{Relaxation}} \) is a negative quantity. The stability of a self trapped electron (or hole) is therefore a balancing act between the increase in kinetic energy (which acts to destabilise the state) and the stabilising effect of polarising the lattice.

Trapping energy has some dependence on the functional used. Ramo et al. [98] predicted trapping energies of 0.32 eV for the STE, 0.14 eV for the 4C STH, and 0.57 eV for 3C STH. McKenna et al. [101] studied self-trapped holes in m-

\[ \text{HfO}_2 \] using the cancellation of nonlinearity (CON) method [102]. The predicted trapping energies using CON are 0.18 eV for a 3C STH and -0.31 eV for a 4C STH. In the 4C case, the B3LYP calculations by Ramo et al. predict the STH to be stable, whereas the CON predictions by McKenna et al. give the STH to be metastable. Thus, trapping energy predictions can vary by hundreds of meV between different XC functionals. Therefore, for low trapping energies, it is possible for different DFT XC functionals to disagree as to whether trap states are stable or meta-stable.

4.1.2 Modelling the Structure of Amorphous HfO\(_2\)

The simulation of amorphous materials poses a number of challenges. Amorphous systems are defined by their lack of translational symmetry. In crystals, the whole structure can be represented by a periodic cell. In amorphous systems this is no longer possible. Furthermore, amorphous structure is statistical in nature. Owing to the lack of translational symmetry, the microscopic structure varies across space. It
4.1. Introduction & Background

Figure 4.1: A schematic of the process used to generate models of a-HfO₂. The melting of the crystal is simulated, followed by a rapid cooling. This produces a periodically repeated amorphous cell, which is then reoptimised using a DFT approach.

is important to emphasise, however, that the structure is not totally random. There is still short-range order governed by the requirements of the bonding in the system. For example, Fig. 4.2 shows the radial distribution function (RDF) of O-Hf distances in amorphous HfO₂. Medium to long-range distances follow a uniform distribution, however at short-range there is a significant peak corresponding to the O-Hf bond distance. Models of amorphous structure should reflect the fact that the microscopic structure of an amorphous material varies greatly. This means either using a very large number of atoms in a model or using multiple models to form a statistical ensemble.

The amorphous models of HfO₂ in this project were created by M. Kaviani [29] using the melt-and-quench method. The melt and quench method takes a large supercell of crystalline HfO₂ and, using molecular dynamics (MD), simulates its melting at high temperature. Starting at absolute zero and using a NPT ensemble, the temperature of the supercell is increased until the crystal is melted and equilibrated. The system is then cooled back down to absolute zero. Once solidified, the system exists in a disordered state. In the final stage, this amorphous model is re-optimised using the PBE0-TC-LRC functional. This process is repeated to produce 9 models. The a-HfO₂ models used in this project were 324 atom cells (equal to the number of atoms in a 3 × 3 × 3 monoclinic cell of HfO₂) of pseudoamorphous structure. The term pseudoamorphous refers to the fact that the amorphous structure is represented by a large, periodically repeated cell of disordered HfO₂. This periodicity should not negatively affect the results as long as the unit cell is large
Disordering the periodic supercell has two major effects on the microscopic structure. Firstly, the O-Hf bond lengths are no longer discrete, but instead form a distribution (see Fig. 4.2). Secondly, the coordination numbers of the oxygen and hafnium ions also form a distribution. For example, the Hf ion coordination number (with respect to oxygen) is 7 in m-HfO$_2$. In the amorphous models, the majority of Hf ions are 6-coordinated, and a small percentage of 5-coordinated Hf ions appear (likewise, there is a decrease in the average coordination number of the oxygen ions, and the emergence of some 2 coordinated oxygen ions). The distribution of bond lengths means there is a number of O-Hf bond lengths which are smaller (and some larger) than the crystal bond length. One other major effect of the disorder is to reduce the density. The average density of the a-HfO$_2$ cells used in this project are 9.6 g cm$^{-3}$. The distribution of bond lengths is shown in Fig. 4.2. The purpose of this chapter is to investigate the behaviour of excess electrons in amorphous HfO$_2$. In addition, results for STHs are also calculated. The chapter begins with the results of calculations of trapped electron and hole states in crystal HfO$_2$. The results of trapped carriers in amorphous HfO$_2$ are then presented, followed by calculations on the optical spectra of trapped states.
The existence and behaviour of self-trapped carriers in a-HfO$_2$ has a number of important consequences. Phenomena such as random telegraph noise and bias temperature instabilities in MOSFETS are linked to trapping and detrapping of charge. Furthermore, as discussed in later chapters, trapped electrons and holes may play a role in the generation of oxygen vacancies. This has impacts on the degradation of HfO$_2$ based dielectrics and on future ReRAM technologies.

The PBE0-TC-LRC functional is used to optimise the geometry and calculate the electronic structure of the 324-atom a-HfO$_2$ models (see section 2). A cutoff radius of 4 Å is used with an exact exchange of 25%. This gives an average (across the 9 models) KS band gap of 6 eV.

This chapter presents the results on intrinsic electron and hole trapping in HfO$_2$. The main result is the prediction that electrons and holes can trap onto “precursor” sites in pristine amorphous HfO$_2$ into states which are much energetically deeper than those found in crystalline HfO$_2$. Models of amorphous HfO$_2$ which were previously generated using the melt-and-quench method were used. Electron and hole traps were also calculated in crystal HfO$_2$ models for comparison. For self trapped electrons and holes in amorphous HfO$_2$, a number of different configurations were simulated so that distributions of important properties could be calculated. In the final part of this chapter, optical spectra of deep electron traps are calculated and compared with results from spectroscopic experiments.

4.2 Results

4.2.1 Electronic Structure of Bulk HfO$_2$

In this chapter, both the crystalline and the amorphous phases of HfO$_2$ are presented. In this initial section, the difference between the electronic structure of the two phases will be described.

4.2.1.1 Electronic Structure of Crystalline HfO$_2$

At room temperature and pressure, the monoclinic (m-) phase is the most stable crystal structure of HfO$_2$. The band gap of m-HfO$_2$ is calculated to be 6.6 eV,
which is higher than experimental values of 5.5-5.8 eV [96, 97]. The valence band maximum (VBM) and conduction band minimum (CBM) states are shown in Fig. 4.3. As discussed above, there are two oxygen sublattices. One sublattice is composed of oxygen ions which are 3 coordinated with respect to hafnium ions, the other 4 coordinated. The VBM state sits only on the 3C oxygen sublattice, and has a dominant $p$-character (see Fig. 4.3(A)). Hafnium ions are all identical in m-HfO$_2$, and the CBM state is delocalised over all hafnium ions equally. For the sake of comparison, one can also look at the tetragonal phase electronic structure. The 3x3x3 tetragonal simulation cell is shown in Fig. 4.4. The calculated KS band gap is 7.1 eV. Although this is an overestimate, it agrees with other work [103] which argues the tetragonal phase band gap of HfO$_2$ is higher than the monoclinic. All oxygen atoms are 4C by hafnium ions. All hafnium ions are 8C by oxygen.

4.2.1.2 Electronic Structure of Amorphous HfO$_2$

A helpful metric for the discussion of a-HfO$_2$ electronic structure is the “inverse participation ratio”, or IPR. The IPR of a KS state gives a measure of its degree
4.2. Results

Figure 4.4: (A) The VBM (orange surface) of tetragonal (t-) HfO$_2$. The lack of translational symmetry indicates that the VBM is not at the $\Gamma$-point. (B) Shows the CBM state (blue) of t-HfO$_2$. The iso-surface in both images is 0.01.

of localisation [104, 105, 106, 107, 108]. This method takes advantage of the atom-centered basis set used in CP2K to quantify the degree of localization of each eigenvector. Specifically, if the Kohn-Sham (KS) states are linear combinations of atom-centered basis functions, $\psi_n(r) = \sum_i^N C_{ni} \phi_i(r)$, where $\phi_i(r)$ are the basis functions, the IPR can be calculated as

$$IPR(\psi_n) = \frac{\sum_i^N C_{ni}^4}{(\sum_i^N C_{ni}^2)^2}.$$  \hspace{1cm} (4.3)

The IPR is calculated for each KS state in the valence and conduction band. According to the definition used in eq. 4.3, the IPR of a given KS state should vary between 0 and 1, and will be very small for delocalised states. For example, if a state is delocalised across all the $N$ basis functions, with the coefficients $c_{ni}$ equal to one another, the IPR will be $IPR(\psi_n) = \frac{1}{N}$, where $N$ is the total number of basis functions. The IPR spectrum of a typical a-HfO$_2$ model is shown in Fig. 4.5(B). It can
be seen that states near the valence and conduction band edges are highly localised (particularly in the VBM). Also, if one compares the crystal HOMO (Fig. 4.5(A)) to the amorphous HOMO (Fig. 4.3(A)), one can also see the localised nature of the amorphous-structure’s band edge states compared to that of the crystalline. One can also compare the crystal IPR (Fig. 4.6) to the amorphous IPR and see that the bands in crystal structure have a flat IPR spectrum, with no build up at the edges. Using the IPR spectrum one can also estimate the so-called ‘mobility edge’ (ME) [109]. The ME is the transition between semi-localised and delocalised states. In Fig. 4.5 the hole mobility edge is approximately 1 eV below the VBM. Below the hole mobility edge the IPR of the VBM is much smaller and ranges between 0.003 and 0.0035. This corresponds to delocalisation over approximately 300 basis functions. The localisation of band states in disordered insulators is a well known phenomenon [110, 111].
4.2. Results

Figure 4.6: The IPR spectrum of monolinic HfO$_2$. Compared to the amorphous IPR, the crystal IPR spectrum shows no significant build of localisation at the band edges.

4.2.2 Trapped Holes in HfO$_2$

4.2.2.1 Self Trapped Holes in the Crystalline Phase

The characteristics of STHs in crystalline HfO$_2$ are presented in this section. These calculations are used for comparison to STHs in a-HfO$_2$, which is one of the main results of this chapter. It is found that STHs are generally composed of a hole state which has been localised onto a single oxygen ion or an oxygen ion pair. Hole localisation distorts the surrounding structure and produces long range polarisation of the lattice. In crystals, the structure and stability of the STH depends on crystal phase.

These results imply that, if there are STHs in m-HfO$_2$, they will trap onto oxygen ions. The configuration for a STH on a 3C oxygen ion is shown in Fig. 4.7. The hole is localised onto a $p$-like state. Localisation of the hole leads to lattice polarisation, displacing nearby Hf ions by 0.1 Å. This polarisation extends throughout the whole lattice, however ions farther from the STH centre are less displaced. The trapping energy of the 3C STH is found to be 0.45 eV. This is higher than the results gained using CON by McKenna et al. [101], but lower than the 0.57 eV calculated by Ramo et al. using B3LYP [98]. The unoccupied Kohn-Sham hole state sits 2.0 eV above the valence band. As can be seen in Fig. 4.7 the degree of
Figure 4.7: (A) The 3-coordinated oxygen ion and its three hafnium neighbours in monoclinic HfO\textsubscript{2}. Bond lengths (in \textAA) are indicated. (B) The 3C self-trapped hole in m-HfO\textsubscript{2}. A single hole state, shown by the orange iso-surface (value = 0.025) is trapped onto an oxygen ion, causing extension of the O-Hf bonds.

localisation is high. Mulliken population analysis calculates the spin moment on the oxygen ion is 0.86. Calculations for a hole trapped on a 4C oxygen were attempted, however no stable configuration could be found.

The STH in tetragonal (t-) HfO\textsubscript{2} is shown in Fig. 4.8. Compared to the monoclinic phase, it has a smaller trapping energy of 0.34 eV, and a KS level at 2.97 eV above the VBM. Figures 4.8(A) and (B) show, respectively, the local structure of two adjacent oxygen ions before and after hole trapping. The hole is distributed asymmetrically. Mulliken analysis gives the spin moment on the right oxygen (Fig. 4.8(B)) to be 0.57 and the spin moment on the left oxygen to be 0.38. Hafnium ions are displaced between 0.1 and 0.4 \textAA. Comparing figures 4.7(B) and 4.8(B) shows that the STH structure varies between crystal phases. In the monoclinic case, the hole is localised onto a single oxygen as a one-center trap. In the tetragonal phase, the hole is localised over two oxygen atoms as a two-center hole trap. This is similar to the situation observed in silica glass [112], in which holes self trap into both one-center and two-center states. In the tetragonal hole trap, the two oxygen ions move closer together. This is because the trapped holes decrease the magnitude of the oxygen ion’s charge. This reduces the Coulombic repulsion between the two
4.2 Results

Figure 4.8: (A) The geometry of two adjacent oxygen atoms and their nearest neighbour hafnium ions in t-HfO$_2$. (B) The local geometry of the STH in t-HfO$_2$. Square modulus of the hole state is shown by the orange iso-surface (iso-surface value = 0.025). Bond lengths, in Å, are indicated. Direction of the major displacements are indicated by arrows.

oxygen ions and thus decreases their separation.

4.2.2.2 Hole trapping in Amorphous HfO$_2$

In a-HfO$_2$, simulations do not require any pre-emptive reorganisation of atoms to encourage trapping. Instead, holes trap readily onto certain ‘precursor’ sites. There are usually 3 to 4 precursors per 324-atom cell. Precursor sites are structural motifs which lower the electrostatic potential seen by the hole. These motifs are mostly 2 coordinated oxygen ions and oxygen ions with elongated O-Hf bonds. One consequence of this is the band edge states will be localised around precursor regions. This effect can be seen in Fig. 4.5(A). In crystalline systems the hole state is delocalised across the whole cell (see Figs. 4.3(A) and 4.4(A)).

Each 324-atom a-HfO$_2$ cell contains up to 4 precursors separated by an average distance of 8 Å. Precursors can be identified by projecting the localised states at the band edge onto atoms. Precursor atoms will be those whose basis functions contribute most to the partly-localised KS states at the band edges. One takes the set of coefficients, $\{C_{ni}\}$, and for each atom calculates $C_{\text{Atom}} = \sum_i C_{ni}^2$, where the sum is only over basis functions centred on that atom. $C_{\text{Atom}}$ will be highest for atoms which form a precursor.

A typical configuration of a STH in a-HfO$_2$ is shown in Fig. 4.9. The distribution of the hole is comparable to that of the STH in the t-HfO$_2$ phase, with the hole distributed over two adjacent oxygen ions. The spin density over the two oxygen
4.2. Results

Figure 4.9: (A) The local atomic structure around a typical precursor in a-HfO$_2$, with bond lengths indicated. (B) The local configuration of a hole trap at the same site after trapping and relaxation. Most of the O-Hf distances have been increased due to Coulomb repulsion. The $|\psi|^2$ of the hole state is indicated by the orange iso-surface (iso-surface value=0.015). All bond lengths are given in Å.

atoms is not equivalent as the hole predominantly occupies one oxygen atom. It can be seen from Fig. 4.9 that the hole has $p$ character, similar to the crystalline cases.

The statistical nature of amorphous structures gives a range of values for properties such as trapping energy and KS-level position. 15 different STH configurations in a-HfO$_2$ were calculated. The stability of these states depends on their trapping energy. The definition of a trapped carrier state in amorphous structure is similar to that of a crystal. The reference point, however, is not a delocalised state, but the initial (unrelaxed) state with an electron removed (or added). That is,

$$E_{\text{Trapping}} = E_{\text{Initial}} - E_{\text{Trapped}}.$$  \hspace{1cm} (4.4)

This is similar to the definition for crystal STEs and STHs, but with some conceptual differences. The initial state is not delocalised, but semi-localised around a precursor. Secondly, the trap is no longer a “self-trapped” state. In the self-trapping process, localisation is facilitated by the electron-phonon interaction [113, 114, 115]. In the case of trapping in amorphous HfO$_2$, a large part of the localisation happens due to the deepening of the electrostatic potential by precursors. Thus the state is not completely self-trapped. On average, the trapping energy of a hole in a-HfO$_2$
4.2. Results

Figure 4.10: Trapping energy distribution of electron and hole traps.

is 1.4 eV. This value for trapping energy is calculated with respect to +1 system in its initial configuration (no relaxation). This is much greater than the trapping energy numbers predicted in crystalline phases. STHs will, on average, be stable at room temperature. Furthermore, many STHs have a trapping energy greater than the average, as the spread of values is high. The full range is from 1.0 to 1.9 eV, as shown in Fig. 4.10. The average position of the KS-level is 3.2 eV (above the VBM), ranging from 2.5 eV to 3.7 eV.

Thus the trapped hole states in a-HfO$_2$ are energetically deeper than in the crystalline phases of HfO$_2$. In the amorphous state, the initial state of the hole is already part localised before it forms a STH. Therefore, the increase in kinetic energy due to localisation is reduced. Furthermore, the surrounding ions are displaced by greater amounts, implying greater polarisation. For example, in the monoclinic case distortions are $\approx 0.1$ Å whereas STHs in the amorphous models typically extend the surrounding Hf ions by 0.1 to 0.15 Å.

4.2.3 Double Hole Traps

In the previous sections, single hole trapping was discussed. As explained in section 4.1.1, the stability of a self-trapped hole depends on the balance between relaxation energy and the increase in kinetic energy. If the relaxation energy is great enough,
4.2. Results

Figure 4.11: (A) The double hole trap in m-HfO$_2$. Bond lengths (in Å) are indicated. The transparent orange surface shows an iso-surface (iso-surface value=0.01) of the $|\psi|^2$ of the double trap state. (B) is the projected density of states of a m-HfO$_2$ simulation cell containing the double hole trap. The state can be seen high in the band gap.

it becomes possible to trap a second hole onto a STH. This will only happen if the relaxation energy is large enough to compensate for both the increase kinetic energy and the positive Coulomb interaction between the two holes.

4.2.3.1 Double Holes Traps in Crystal HfO$_2$

The double STH in m-HfO$_2$ is shown in figure 4.11. The double hole trap is in a singlet state. The double STH stabilises the formation of an O-O dimer of bond length 1.4 Å. Similar results have been predicted for TiO$_2$ [116]. The formation of a dimer pushes the hole levels high into the band gap. Further, the hole distribution around the two oxygen atoms is symmetric. In fact, the double STH is similar in structure to the neutral oxygen interstitial (discussed later in Chapter 5).

An important property of double STHs is their binding energy. Binding energy is a measure of how stable a double STH is compared to two, well-separated holes. It can be calculated using the formula

$$E_{\text{Binding}} = 2E_{\text{Hole}} - (E_{\text{Double Hole}} + E_{\text{Crystal}}),$$

where $E_{\text{Hole}}$ and $E_{\text{Double Hole}}$ are the energies of the cell containing a single STH.
Figure 4.12: A schematic of the ideal process which defines the binding energy. The binding energy is defined as the energy released when a single self-trapped hole moves and traps onto a second self trapped hole, with which it was well separated. The products of this reaction are a double STH and a region of pristine m-HfO$_2$, which are infinitely separated.

and double STH, respectively. $E_{\text{Crystal}}$ is the energy of the pure, nondefective crystal cell. The reaction this equation represents has been illustrated schematically in Fig 4.12. Total charge and number of atoms are both conserved in this process. A positive binding energy indicates that holes are more stable when bound into a double hole trap than they are when separated at large distances. Using the definition in equation 4.5 the binding energy is calculated to be 0.7 eV, indicating greater stability for a double STH. This is similar to results in [116] for TiO$_2$.

4.2.3.2 Double Hole Trapping in amorphous HfO$_2$

Similar to m-HfO$_2$, a-HfO$_2$ can also trap two holes onto the same site. The structural distortions for a double STH in a-HfO$_2$ are quite large ( > 0.5 Å) and an O-O dimer is formed. The structure of the double hole trap in a-HfO$_2$ is similar to the double hole trap in m-HfO$_2$. The O-O separation of the dimer is 1.5 Å, and its Bader charge is -1.5 $|e|$. On average, the binding energy of a double hole trap in a-HfO$_2$ is 0.5 eV. Similarly to single STH in a-HfO$_2$, double STHs in a-HfO$_2$ have no barrier to formation. Thus holes injected into a-HfO$_2$ can spontaneously trap as
4.2. Results

**Figure 4.13:** The local structure of a typical double STH in a-HfO$_2$. The orange transparent surface indicates the $|\psi|^2$ of the (double occupied) hole state (iso-surface value=0.015). Major structural distortions are caused and an O-O dimer, with a bond length of 1.46 Å, is formed. All bond lengths are indicated in Å.

both single and double STH. The binding energy of a double STH predicts whether it is energetically favourable for holes to be separated or combined as a double STH. In a-HfO$_2$ the binding energy is positive.

### 4.2.4 Optical Excitations of Electron Traps in HfO$_2$

#### 4.2.4.1 Trapped Electrons in a-HfO$_2$

Similar to holes, electrons self-trap into deep, polaron-like states in amorphous HfO$_2$ [29]. Electrons trap onto precursor sites, which are either 5C Hf ions or Hf ions with elongated Hf-O bonds. In a STE, the electron localises over 2 or 3 hafnium ions. Upon localisation, the surrounding Hf ions are stretched outwards by 0.12 Å. Multiple configurations of STEs were analysed - the distribution of the trapping energies in plotted in Fig. 4.10. the occupied KS states were found to be distributed between 1.6 and 2.5 eV below the bottom of the conduction band. These states can trap a second electron to form double electron traps (figure 4.14(B)). The second electron is trapped over the same Hf ions, and so the double STE shows a similar distribution of electron density to the single STE. The second electron trapping is associated with further Hf-O bond stretching of, on average, 0.09 Å [29].
4.2. Results

Figure 4.14: (A) is the intrinsic electron trap in a-HfO$_2$. Ionic displacements are indicated by arrows and given in Å. The blue surface is an iso-surface of the $|\psi|^2$ of the STE KS state. (B) is the double electron trap in a-HfO$_2$. Both iso-surface values are $\approx 0.005$. Images reproduced with permission from [2]. IOP Publishing Ltd.

4.2.4.2 The Observed Negative Charging of a-HfO$_2$ Thin Films

Dielectric thin films are known to charge when placed under electrical stress. Exhaustive photodepopulation spectroscopy (EPDS) offers a way of probing the energy levels of trapped states in insulators [117]. In EPDS, a-HfO$_2$ is irradiated with low frequency light to excite the shallowest electron traps in the sample. If the electron is excited into a delocalised state, it will have a higher conductivity and be easily swept away by an applied (small) voltage. Thus, when the photon energy is enough to promote electrons into the ME, detrapping will occur. Detrapping is measured by analysing IV curves to ascertain the amount of trapped charge before and after the irradiation and taking the difference. Irradiation continues until charge detrapping stops, that is, once all the traps up to that energy level have been emptied. The frequency of the irradiating light is then increased by a small increment. This will allow deeper electron traps to then be liberated. By measuring the change in the charging of the sample (again by measuring IV curves), the density of traps in that small energy window can be deduced. This process is then repeated (the frequency is increased by another increment) and the DOS of trap levels across the
band gap can be measured.

4.2.4.3 Calculating the Optical Ionisation Spectra of STEs

When running EPDS experiments on a-HfO$_2$, therefore, it is the case that decharging will only occur when the photon energy is high enough to promote an electron from its trap state to the ME. Thus in order to compare theoretical and experimental results, optical transitions between the trapped state and the ME must be calculated. Transitions into ‘post ME’ states can be investigated using time dependent DFT (TDDFT) combined with IPR analysis of the electronic structure. Excited-state wavefunctions are represented in TDDFT as a linear combination of single-excitation determinants

$$\Phi_{\text{Exc}} = \sum_{i,a} C_{i}^{a} \Psi_{i}^{a},$$

where $\Psi_{i}^{a}$ is the Slater determinant with molecular orbital $i$ replaced with virtual orbital $a$. $\Phi_{\text{Exc}}$ is therefore a mixture of single excitations. $C_{i}^{a}$ is called the ‘mixing amplitude’ and indicates to what extent a given single-excitation contributes to the total excited-state wavefunction. The mixing coefficients follow the sum rule

$$\sum_{i,a} (C_{i}^{a})^2 = 1.$$ 

By analysing the mixing amplitudes of a given excitation, one can ascertain whether the electron has been excited past the ME. Firstly, the location of the ME must be identified. Using IPR, the virtual orbital at which there is an onset of delocalised states is located. This orbital is denoted $a = a_{\text{ME}}$. Associated with this orbital is the energy $E_{\text{ME}}$, which is its KS energy eigenvalue relative to the CBM. Then, for a given excitation, if the sum $\sum_{i,a} (C_{i}^{a})^2$ for $a > a_{\text{ME}}$ is greater than a chosen threshold value, the state is considered to be post ME. A natural choice for the threshold may be 0.5 (when the excited state is majority post ME), but in practice there is a transition point where the value of this sum jumps to above 0.8. This transition point is exactly where one would expect it to be based on IPR analysis alone, at $E_{\text{ME}}$ above the CBM. Thus, in amorphous structures, there is a threshold energy required to optically excited electrons into delocalised states. This can be calculated by locating the ME position with IPR analysis.
Figure 4.15: The energy distribution of electron traps in a-HfO$_2$. The top panel shows the predicted optical absorption energies of single and double electron traps based on TDDFT modelling. The bottom panel shows the results of EPDS measurements. Both spectra show a distribution centered around approximately 3 eV. Images reproduced with permission from [2]. IOP Publishing Ltd.

TDDFT calculations were carried out on a range of single and double trapped electron models in a-HfO$_2$. The results of these calculations were then compared with EPDS experiments [2]. The DFT/TDDFT and EPDS results are shown in Fig 4.15. Both the theoretical models and the EPDS spectrum give a wide range of optical absorption energies which center on 3 eV. The ab-initio spectrum was generated by applying a shift to the KS eigenvalue spectrum. TDDFT calculations on a large number of models is computationally expensive, so instead 5 TDDFT calculations were run on a sample of models. The sample was chosen by taking the configurations with the lowest, highest, and middle-three valued trapping energies. From this, it is established that excitation into the ME is accurately predicted by the KS energy differences. The spectrum can therefore be generated by taking the KS energy difference between the ME and the position of the electron trap (or bi-trap). The ME is calculated for each a-HfO$_2$ model-cell and is found (on average) to be 0.5 eV above the CBM. The agreement between the EPDS results and the predicted optical transition energies is good evidence that intrinsic electron and bi-electron traps are responsible for the observed negative charging of a-HfO$_2$ films.
4.2.5 Conclusions

In this chapter the PBE0-TC-LRC functional has been used with large periodic simulation cells to study carrier trapping in HfO\textsubscript{2}. The stability of hole and electron polarons in m-HfO\textsubscript{2} has been confirmed and, as expected, it is found that the trapping energy of these defects depends on the XC-functional used in the simulations. Similar to previous work, these polarons are predicted to be unstable at room temperature. Then, it has been shown that in amorphous HfO\textsubscript{2} electrons and holes can trap into deep polaron-like states. Trapping in a-HfO\textsubscript{2} happens onto structural precursor sites. These precursor sites occur at approximately 1 in every 100 atoms. They consist of extend Hf-O bonds and undercoordinated atoms, which cause a drop in the local electrostatic potential.

Using IPR analysis, the electronic structure of the a-HfO\textsubscript{2} models was probed. The valence and conduction bands are composed of a set of semi-localised states near the band edges, and delocalised states deeper in the bands. These transition point between these two subparts of the bands is called the ME. The ME was located, on average, to be 0.5 eV above the CBM and 1.0 eV below VBM.

Using TDDFT, the optical absorption spectra of the electron single and double traps is calculated. In particular, the energy required to excite the trapped electrons into the ME is calculated. These energies are found to be in reasonable agreement with EPDS experiments.

The results of this section have impacts for the application of a-HfO\textsubscript{2} as a gate dielectric in CMOS devices. As discussed in chapter 1, charging of defects in the gate oxide layer has been linked to bias temperature instabilities, manifested in a fluctuation of threshold voltage. According to the results in this thesis, however, it is possible that even highly pristine (stoichiometric, impurity free) amorphous oxide layers will undergo charging. Further, the concentration of the precursor sites, at which electron and holes are trapped, is quite high. This means that intrinsic trapping of carriers in a-HfO\textsubscript{2} could also facilitate trap assisted tunnelling through the oxide (process 1 in figure 1.2).

This thesis will now examine further effects that this charging behaviour can
4.2. Results

exhibit. In the next two chapters, oxygen vacancies and vacancy clusters on HfO$_2$ will be discussed. Ultimately (in chapter 6) it will be shown that these trapping sites can lead to the generation of oxygen vacancies, which play their own role in the degradation of MOS devices.
Chapter 5

Oxygen Vacancy and Interstitial defects in amorphous HfO$_2$

5.1 Introduction & background

Hafnium dioxide (HfO$_2$) is a high-$\kappa$ insulator with a band gap of around 5.7 eV [118, 119]. HfO$_2$ is currently used as the gate dielectric in many CMOS devices, and is also a candidate for the functional material in future Resistive RAM (ReRAM) and other non-volatile memory technologies. As discussed in chapter 1, RRAM devices function by using an applied field to control the growth and retraction of a conductive filament across an insulator. Experimental characterisation of conductive filament formation (called the ‘electroforming’ process) in HfO$_2$ has revealed the filament to be an oxygen deficient region [24]. In general, the presence of defects in materials such as HfO$_2$ fundamentally alters their functionality in applications. Defects can often act as charge trapping centers, leading to problems such as leakage current via trap-assisted tunnelling or to shifting of the MOSFET threshold voltage via charging. Characterisation of defects in functional oxides such as HfO$_2$ is therefore of paramount importance. In this chapter, a survey of the basic properties of oxygen defects in HfO$_2$ is given. This contributes to the understanding of how such defects play a detrimental role in some applications (MOSFETs), but also lays the foundations for chapter 6, which studies the process of defect creation in HfO$_2$. 
The oxygen vacancy in monoclinic (m-) m-HfO$_2$ will first be discussed, followed by results on the oxygen vacancy in amorphous (a-) HfO$_2$. The optical properties of oxygen vacancies in crystalline HfO$_2$ have also been calculated and will be discussed. Finally, the electronic and geometric structure of oxygen interstitial defects in m- and a-HfO$_2$ are presented.

5.1.1 Previous studies on O vacancies in HfO$_2$

There have been a number of previous publications tackling the subject of oxygen vacancies in HfO$_2$. Foster et al [120] studied vacancy and interstitial defects in m-HfO$_2$ using the Perdew and Wang “GGA-II” exchange correlation functional [121]. Oxygen vacancies were modelled by removing neutral oxygen atoms from a periodic simulation cell (containing 96 atoms), and then allowing the system to relax. It was found that for a neutral oxygen vacancy, the relaxation energy contributed very little to the overall formation energy. The formation energy of the V$_0^0$ defect was given as 9.3 eV (with the oxygen chemical potential being the energy of an isolated oxygen atom), and the relaxation energy was less than 0.1 eV. Positive charge states of the oxygen vacancy are simulated by removing an electron from the periodic cell. This causes displacement of the surrounding ions by 0.1 to 0.2 Å. The relaxation energy for the +1 state is larger than the neutral and is around 0.6 eV. Removing another electron from the cell gives the V$_0^{+2}$ defect. Nearest neighbour ions are displaced by a further 0.1 to 0.2 Å. The relaxation energy is around 0.8 eV. For the oxygen interstitial, Foster et al found the incorporation energy in the neutral state to be -1.3 eV for a three coordinated interstitial and -0.6 eV for a 4-coordinated interstitial. Again, the isolated oxygen atom was used as the oxygen chemical potential. The O$_I$ defect forms a dimer with its nearest neighbour atomic oxygen with a bond length of 1.5 Å. Extra electrons, added to the system, localise onto the oxygen dimer after the system is fully relaxed. The O-O length increases to 2.0 Å. The relaxation from the initial configuration is 2.3 eV. Adding a second electron causes the O-O bond to increase further, to 2.4 Å. This accompanies a relaxation energy of 1.53 eV, and the second electron is again localised over the O$_I^{-2}$ defect after relaxation.

The optical spectra of oxygen vacancies has been studied before. Ramo et al.
5.2. Results

[122] modelled the optical transitions of V\textsubscript{O} defect using both periodic boundary conditions and the embedded cluster method [123, 124, 125]. The periodic calculation were run using the CRYSTAL03 [126]. It was found that the V\textsuperscript{+2}_O and V\textsuperscript{+1}_O defects have optical absorption peaks at 4.94 and 4.67 eV, respectively. Both of these excitations are from the valence band into the gap state. The V\textsuperscript{+1}_O defects was found to have two additional optical absorption peaks, at 3.27 and 2.64 eV. The V\textsuperscript{0}_O defect had two optical absorption lines, at 3.41 and 2.45 eV.

Studies such as these two have, until now, been limited by computational restrictions. Simulations have typically been run either using the more crude (but computationally faster) GGA or LDA approaches, which risk giving the incorrect electronic structure for bulk and defects. Or, even where advanced functionals (such as hybrids) are used, small cell sizes are employed which affect the results. By using the PBE0-TC-LRC functional, and large simulation cell sizes, the results in this chapter should avoid these issues.

In this Chapters, Defects are modelled in m-HfO\textsubscript{2} using the same XC functional and a-HfO\textsubscript{2} models described in Chapter 4. Oxygen vacancy models are generated by removing an oxygen ion from a crystal site and then relaxing the geometry of the periodic supercell, interstitials by adding oxygen atoms and then relaxing. TDDFT calculations are run using the method described in Chapter 3.

5.2 Results

5.2.1 Vacancy Defects in the monoclinic phase

In the neutral charge state of the V\textsubscript{O} defects (V\textsubscript{O}\textsuperscript{0}) there are two electrons localised on the Hf ions surrounding the oxygen vacancy (see Fig. 5.1). Adding or removing electrons to the simulation cell charges the system. If, for example, one adds a single electron, it localises on the Hf ions surrounding the vacancy. This gives a V\textsubscript{O}\textsuperscript{−1} defect. In total, 5 charge states (for -2 to +2) can be found this way. As is common, there are gap states associated with the localised electrons. Fig. 5.2 shows the KS states of the oxygen vacancy in the 5 different charge states. For the charge states -2, -1, 0 and +1 there is an occupied state in the middle of the band gap. In
Figure 5.1: The 3C oxygen vacancy in m-HfO$_2$. Hf ions are colored in cyan and oxygen ions a colored in red. Each configuration is labeled with its charge state. The translucent surfaces give the $|\phi|^2$ of the defect orbitals. Where applicable, the higher eigenvalue state (closer to the CBM) has been colored in yellow, and the lower valued eigenvalue state (closer to the VBM) has been coloured in blue. All iso-values = 0.005. The energy of these states are indicated in Fig. 5.2.

the case of the +2 charge oxygen vacancy ($V_{O}^{+2}$), there are no occupied states in the gap but there is an unoccupied state close to the CBM.

The localisation of electrons around the Hf ions in the $V_{O}^{0}$ defect (Fig. 5.1) causes a slight reorganisation of atoms around the vacancy, of the order of 0.1 Å. This distortion of atoms increases as the vacancy is charged owing to greater Coulomb interaction between the vacancy and the surrounding ions. The negatively charged defects, plotted in Fig. 5.1, have higher energy eigenstates. This states have electron(s) localised over two Hf ions. Thus, the -1 and the -2 charged O vacancy in HfO$_2$ can be understood as a polaron or bipolaron (respectively) trapped at a neutral oxygen vacancy.

By inspecting Fig. 5.3, one can see the the doubly positively charged state of the oxygen vacancy is the most stable for the majority of Fermi levels (i.e. for
5.2. Results

Figure 5.2: The KS states of the oxygen vacancy in m-HfO$_2$. Levels are shown (in eV) relative to the CBM.

most of the band gap as one sweeps the Fermi level through it). There is a +2/0 charge transition level (CTL) at above 4 eV, where the stable configuration switches from +2 to neutral. There is then the 0/-2 charge transition at approximately 6 eV (just below the CBM) where the stable configuration switches from 0 to -2. This charging and discharging behaviour can be responsible for functionality problems with insulators in CMOS applications.

5.2.2 Vacancy defects in the amorphous phase

As discussed in Chapter 1, interfacial strain can cause oxide layers in CMOS devices to exist in the amorphous phase. Indeed, this is even seen as desirable as it removes the problems of grain boundaries, which are known to be detrimental to device performance.

Amorphous materials may themselves pose interesting new problems for application of insulating oxides in nanoelectronics. HfO$_2$ is not a natural glass former and it may undergo structural changes during operation. Defects may have different behaviour, different dynamics et cetera. Here the properties of oxygen vacancies in amorphous (a-) HfO$_2$ are reported.

The formation energy diagram of the V$_O$ defect in a-HfO$_2$ is plotted in Fig.
5.2. Results

Figure 5.3: Formation energy diagram of the oxygen vacancy in HfO$_2$ in its 5 charge states. Crossing points indicate the charge transition levels.

5.4. It can be seen that, unlike in Fig. 5.3, the lines are represented as ‘bands’. This originates in the fact that each oxygen atom in an amorphous oxide sits in a unique environment, leading to a spread of formation energies. The oxygen vacancies are sampled so as to reflect the coordination number distribution of oxygen atoms in a-HfO$_2$ discussed in Chapter 4. It is not found that formation energy significantly dependents on the coordination number of the deleted oxygen atom. Some typical configurations for the oxygen vacancy in amorphous HfO$_2$ are shown in Fig. 5.6. The configurations are similar to oxygen vacancies in m-HfO$_2$. A histogram showing how the distribution of formation energies looks can be found in Fig. 5.5.

One interesting difference between negative charging in the amorphous and crystalline simulation cells is that electrons need not trap at oxygen vacancy sites. As discussed above, the $V_{O}^{-2}$ defect can be interpreted as an electron bipolaron trapped on a neutral oxygen vacancy. In amorphous HfO$_2$, however, there are several intrinsic precursor sites where deep polaron trapping can occur (Chapter 4). Therefore, it can be energetically favourable for electrons to trap as isolated polarons or bipolarons, separated from the oxygen vacancy. This balance can be delicate, with some amorphous $V_{O}$ models favouring a separated vacancy-polinor...
5.2. Results

Figure 5.4: The formation energy of oxygen vacancies in amorphous HfO$_2$ as a function of the system Fermi level. The values are similar to those of the monoclinic phase. Unlike crystalline phases, however, there is a ‘width’ to each line. This width corresponds to the spread of formation energies possible due to the statistical nature of amorphous systems.

Figure 5.5: (A) Histogram of the set of neutral oxygen interstitial defect configurations analysed in the a-HfO$_2$ models. (B) Histogram of the set of the neutral oxygen vacancy formation energies analysed in the in a-HfO$_2$ models.
5.2. Results

Figure 5.6: The oxygen vacancy in amorphous HfO$_2$. The charge state of each configuration is indicated. Since it is an amorphous system, each O vacancy is unique, and so only one configuration among many is shown. Isosurfaces indicate $|\phi|^2$ of the defect states. Where relevant, the higher energy eigenvalue KS state is colored yellow, and the lower blue. All iso-values are $= 0.005$.

configuration and some favouring a polaron-at-vacancy. The latter case usually occurs when one of the oxygen vacancy’s adjacent Hf ions is also an electron trapping precursor. In Fig. 5.7 the KS defect levels are plotted for the case of oxygen vacancies in a-HfO$_2$. The positions are similar, but again with levels having a spread of values due to the uniqueness of each oxygen vacancy. The negatively charged vacancies have higher-energy occupied KS states in the gap. These states are energetically deeper than they are in the monoclinic phase. For example, the $V_{O}^{-2}$ high energy state is, on average, 2.0 eV below the CBM. For the monoclinic case it is 1.5 eV below the CBM. On average, however, the state is higher in energy than the isolated bipolaron (intrinsic double electron trap) state. This provides insight into the question of whether surplus electrons will trap at intrinsic precursor sites or at
5.2. Results

Figure 5.7: The KS defect level positions for oxygen vacancies in amorphous HfO$_2$. The level positions are given (in eV) relative to the CBM. The spread of level positions are indicated by the bars. In some charge states, levels are spread over 1 eV.

<table>
<thead>
<tr>
<th>$V_{O}^{-2}$</th>
<th>$V_{O}^{-1}$</th>
<th>$V_{O}^{0}$</th>
<th>$V_{O}^{+1}$</th>
<th>$V_{O}^{+2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>1.7</td>
<td>1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.7</td>
<td>3.0</td>
<td></td>
<td>3.4</td>
</tr>
</tbody>
</table>

Figure 5.8: (A) The optical absorption spectrum of the $V_{O}^{0}$ defect in m-HfO$_2$. (B) The optical absorption spectrum of the $V_{O}^{+1}$ defect in m-HfO$_2$. (C) The optical absorption spectrum of the $V_{O}^{+2}$ defect in m-HfO$_2$. In all three graphs, the original spectrum calculated by TDDFT is given by the blue delta functions. The red line gives the spectrum which has been created by smoothing the TDDFT spectrum, using Gaussians of $\sigma = 0.2$.

neutral oxygen vacancies. For many oxygen vacancies, this vacancy KS state has a higher energy compared to the intrinsic trapping state. This is not, however, true of all oxygen vacancy configurations. On average it is more favourable for electrons to trap at intrinsic precursor sites. The disorder, however, leads to a spread of energy levels which then produces a number of cases where this general rule is broken.
5.2. Results

5.2.3 Optical absorption and luminescence of O vacancies

5.2.3.1 Optical Absorption of Vacancies

For the $V_0^0$ center in m-HfO$_2$, the TDDFT calculations predict a broad absorption spectrum with an onset at 2.5 eV, and then with higher energy peaks at 3.0 eV and 3.5 eV. Unlike in MgO, the $V_0^0$ excitation in HfO$_2$ is from the vacancy state into states delocalised at the bottom of the conduction band. Hence much lower oscillator strengths of these transitions. The calculated spectra of the $V_0^0$ in HfO$_2$ are plotted in Fig. 5.8 (A). Note that the predicted transition energies are in good agreement with the results of the embedded cluster calculations [122] discussed at the start of this chapter.

In HfO$_2$, optical excitations in the $\alpha$ and $\beta$ spin channels have different energies. To illustrate why, one can look at the electronic structure of the $V_0^{+1}$ defect in HfO$_2$. The alpha HOMO-LUMO separation is calculated to be 3.6 eV, whereas the beta HOMO-LUMO separation is calculated to be 5.6 eV (very close to the optical band gap value). Thus, when irradiating HfO$_2$ with photons of sub-bandgap energy, one would expect the $V_0^{+1}$ spectrum to be dominated by $\alpha$ spin transitions (transitions from the vacancy state into the CB states, (type 4 in Fig. 3.2). It is therefore predicted that the first optical absorption peak occurs at 2.9 eV and corresponds to the transition of the alpha-spin electron from the gap-state into a quasi-local state in the CB. All the higher energy peaks up to band-gap energy also correspond to excitations of the $\alpha$-spin electron out of the gap-state and into higher energy CB states. Excitations in the $\beta$ spin channel are only attainable once the excitation energies are comparable with the band gap.

In HfO$_2$, the optical transition energies from the valence band into the $V_0^{+2}$ defect are comparable to the band gap energy and are predominantly of type II in Fig. 3.2. Optical experiments conducted on HfO$_2$ [127] have detected a 5.2 eV absorption peak which has been attributed to the presence of oxygen vacancies. The TDDFT calculations predict that the doubly positively charged oxygen vacancy ($V_0^{+2}$) has an optical absorption peak at 5.3 eV, which is in good agreement with this detected absorption band. We note that the position of the peak predicted in this
5.2. Results

Figure 5.9: The triplet state of the $V_{O}^{+2}$ defect. The hole has localised onto a STH state and is separated from the $V_{O}$ defect, which is now in the +1 charge state. The separation between the hole and the vacancy is variable, however the total energy of the defect is weakly dependant on this separation. Both iso-vales are = 0.006.

work is at higher photon energy than 4.94 eV calculated in Ref. [122] using TDDFT in a relatively small embedded cluster. This is consistent with our observation that full account of defect-induced lattice distortion for charged defects is important for predicting correct optical absorption energies.

5.2.3.2 Luminescence of the $V_{O}^{+2}$ defect

Using the $\Delta$-SCF method, the luminescence of the $V_{O}^{+2}$ defect can be calculated. By calculating the triplet state, one can simulate the relaxation of the system after the $V_{O}^{+2}$ defect has been excited. As discussed in 5.2.3 excitation of the $V_{O}^{+2}$ defect corresponds to promoting an electron from the valence band into the vacancy states. Alternatively, one can view this as a hole being excited out of the vacancy and into the valence band. When simulating this process and allowing the system to respond to this excitation, it is found that this excited state is initially quite delocalised across the cell, but then localises to form a hole polaron separate from the vacancy. The vacancy has one electron localised in it and relaxes into the $V_{O}^{+1}$ configuration. Thus, the excitation converts the $V_{O}^{+2}$ defect into the $V_{O}^{+1}$ defect, and creates a hole polaron away from the defect site. There are a range of stable sep-
5.2. Results

Separations between vacancy and polaron. The smallest separation found is 5 Å. This highest separation found is 11 Å. Higher separations cannot be found because of the periodic boundary conditions. Thus it is possible that the hole will sit even farther from the vacancy, on average. In fact, there is no significant dependence of total energy in the $V_{O}^{+1} - h^+$ separation, and both defects act as though they are independent. Therefore, they should be seen as two isolated defects. The hole polaron will migrate throughout the crystal and will not, in general, be correlated with the $V_{O}^{+1}$ defect. The recombination will emit a photon by a radiative tunnelling transition (RTT) [128]. The energy of this emission is calculated to be 3.7 eV.

5.2.4 Interstitial defects in the monoclinic phase

The interstitial oxygen defect can exist in three different charge states in m-HfO$_2$ - 0, -1 and -2. The configuration of the O$_I$ defect depends strongly on this charge state. As the charge state is decreased (becomes more negative), the distance between the O$_I$ and its nearest oxygen ion increase. For the neutral charge state, the nearest neighbour distance is 1.46 Å, for the -1 it is 2.0 Å, and for the -2 charge state it is 2.35 Å.

In the neutral configuration there is a peroxide linkage formed between the O$_I^0$ defect and its nearest neighbour, that is, the interstitial oxygen and its nearest neighbour form a bonded O-O dimer and overall have a formal charge of -2 (The two electrons are still present due to the -2 charge of oxygen ions in the system). There are no states in the gap either as the states associated with the linkage are deeper in the CB. This is similar to the case of the hole bipolaron discussed in section 4. In the case of the hole bipolaron, an O-O dimer is formed and stabilised by the fact that the two holes remove the anti-bonding orbitals which would usually be present. In the case of a neutral interstitial O$_I$ defect, a similar process happens. In this case, however, the holes are introduced by a change in stoichiometry (an excess oxygen atom), rather than by charging of the supercell. Adding electrons breaks the peroxide linkage and increases the O-O separation 2.0 Å.

In the -2 charge state the oxygen-oxygen distance is very close to that of the ionic O distance in clean m-HfO$_2$. Further more, Bader charge analysis gives similar
5.2. Results

Figure 5.10: The neutral oxygen vacancy in the 324-atom cell (A), and close up with other atoms hidden (B). Similar to the hole bipolaron, a O-O dimer is formed. Bond lengths are indicated and are in Å.

Figure 5.11: The $O_f^{-1}$ defect in m-HfO$_2$. (A) shows the full supercell and (B) shows the local structure. The NN ions are distorted strongly, and $O_f$ forms sits 2.0 Å from its NN O ion. The NN O ion is displaced out of its crystal position and is symmetrically identical to the interstitial ion.
results for the $O_I^{-2}$ defect and an oxygen ion in m-HfO$_2$. This is unsurprising since oxygen ions in m-HfO$_2$ have a formal charge of -2. Furthermore, the $O_I^{-2}$ defect has states very close to the VBM (0.15 eV above the VBM) and are practically indistinguishable from other VB states. Thus, in terms of charge and its contribution to the density of states, the $O_I^{-2}$ looks similar to the $O^{-2}$ ion in m-HfO$_2$.

### 5.2.5 Interstitial defects in the amorphous phase

Neutrally charged oxygen interstitials in the amorphous phase are very similar to those of the monoclinic phase. A peroxy linkage is formed, with a bond length of 1.45 to 1.5 Å, and no states are seen in the band gap. The major difference, as with all defect in amorphous oxides, is that the incorporation energy has a distribution of values rather than a few values characteristic of the crystal structure. This can be seen in Fig. 5.13, which shows the incorporation energy of the $O_I$ in a-HfO$_2$ in its three charge states, as a function of electron chemical potential. Interestingly the neutral charge state shows much more variation than the other two charge states, with the -1 and -2 charge states showing comparatively little range in incorporation energy.
The charged $O_I$ defects in a-HfO$_2$ are similar to their counterparts in m-HfO$_2$. The $O_I^{-1}$ defect is separated from its NN oxygen by approximately 2.1 Å, and the $O_I^{2-}$ defect is separated from its NN by approximately 2.3 Å. There is very little variation in the NN distances for the -1 and -2 charge states.

5.3 Summary & Conclusions

The structure and energetics of oxygen vacancies and interstitials has been calculated using the PBE0-TC-LRC functional in 324-atom cell of m-HfO$_2$. The results confirm previous theoretical results discussed in section 5.1.1, which relied on either GGA DFT or smaller cells, or both. Five charge states of the $V_O$ defect, ranging from -2 to +2, were found. For the $O_I$ defect, 3 charge states, ranging from -2 to 0, were found. In addition, the optical spectra of the oxygen vacancy has been calculated using TDDFT with a hybrid functional. The results demonstrate that the well-characterised blue luminescence in HfO$_2$ is most likely due to the $V_O^{+2}$ defect. Also, another photoemission process has been calculated which occurs via a radiative tunnelling transition between a trapped hole and a $V_O^{+1}$ center.

The structure of defects in the amorphous phase are also investigated. The results show that the O-based defects in a-HfO$_2$ have similar formation energies and CTLs to those in m-HfO$_2$. The difference between the two, however, is that in amorphous structure the formation energies and CTLs are sampled from a distribution,
leading to some variance in the formation energies and CTLs.

In the context of degradation and breakdown, these results have the following impact. The valence band offset of HfO$_2$ with the Si VBM is 2.5 eV [129]. Therefore, under zero bias conditions, the vacancies are likely to be positively charged. During device operation, however, a bias will be applied. It is possible then for some vacancies to exist as a neutral vacancy, and even for electrons to tunnel into vacancies to form $V_{O}^{-2}$ defects. Oxygen vacancies can therefore be responsible for leakage current via trap assisted tunnelling. Further, if there is an accumulation of oxygen vacancies it is possible to initiate dielectric breakdown (Process 3 in 1.2). This will be discussed in the next chapter.
Chapter 6

Vacancy generation in HfO$_2$

6.1 Introduction & Background

The phenomenon of dielectric breakdown has been one of the major parts of MOS-FET reliability concerns [130, 131, 132]. The process has been studied intensively in SiO$_2$, which was, for most of the history of MOS devices, the most commonly used dielectric oxide.

The process of breakdown can be separated into soft and hard breakdown (Figure 6.1). Both occur as a result of local failure in the oxide, but soft breakdown (SBD) is orders of magnitude smaller than hard breakdown (HBD). In addition, the

![Figure 6.1: Typic I-t characteristics of a HfO$_2$ insulator. The current increases with time, eventually hitting 'hard' breakdown. Retrieved from [3]](image-url)
post-breakdown IV curve of HBD is Ohmic, whereas IV curves after soft break- 
down are usually of the form of $I = A e^{BV}$ or $I = aV^b$ [133]. The creation of carrier-
trapping defects has been understood to be at the root of dielectric breakdown [134].
In particular, it has been the oxygen vacancy that has been attributed responsibility 
for BD in SiO$_2$ and HfO$_2$ [135, 136, 137]. One model which has been proposed 
is the simple ‘percolation path’ model [138]. In this model, defects in the oxide 
gradually build as current is passed through the layer. Eventually, at some critical 
density of defects, the current through the oxide abruptly jumps (breakdown). This 
occurs once a pathway from one electrode to the next has been formed. The process 
is shown schematically in figure 6.2. In the percolation model, defects are produced 
at random positions. Thus the ”time to dielectric breakdown” (TDDB) follows a 
statistical distribution.

The breakdown of an insulating oxide is not always unfavourable, however.
As discussed in Chapter 1, new technologies, based on the memristor, are emerging. 
These technologies (for example, ReRAM) have a resistance that depends on 
the previous voltage, or current, through the device. It is therefore important to un-
derstand the precise model of oxygen vacancy generation and accumulation, since 
this process will be at the heart of the property of memristance (see Chapter 1).

Oxygen vacancy generation in oxides is therefore an important process in 
MOSFETs and other technologies. Dielectric breakdown consists of the formation 
of an oxygen deficient filament. In this chapter, a model for the process of oxygen
vacancy generation in HfO$_2$ is presented. Similar calculations on crystalline HfO$_2$ have been done before, however these calculation used GGA levels of theory. All results presented here use a hybrid DFT formalism. Furthermore, the model is also applied to the case of a-HfO$_2$.

The aim is to understand how an electric field, applied to an insulating oxide, can induce the process of oxygen vacancy generation. An important feature of the mechanism presented in this chapter is carrier injection. Applying an electric field to an insulator causes its bands to bend. Then, with an appropriate alignment with the metal Fermi level, electrons (or holes) can be injected into the oxide. Two possibilities then exist. Electrons, for example, injected into the conduction band may just move quickly through the oxide via band transport. Alternatively electrons may become trapped at a defect site, either by direct tunnelling or by capture from the conduction band. This can happen at, for example, oxygen vacancies. It is also possible for electrons to become trapped into STE states (in crystals) onto intrinsic precursors (in amorphous structure) as discussed in chapter 4. Therefore, during device operation, it is possible that the gate oxides are becoming charged. Hence it is important to consider the effect that charging of thin-films may have on defect generation. In this chapter, calculations will show this effect to be very important. As will be shown in the results section of this chapter, the trapping of electrons is found to decrease the activation energy of oxygen Frenkel pairs by over 5 eV.

6.1.1 Previous Work on Vacancy Generation in HfO$_2$

There has already been some modelling of oxygen vacancy generation in HfO$_2$ in its crystalline phases. Öttking et al. [139] used the PBE functional on 96-atom periodic cells of m-HfO$_2$. Oxygen Frenkel pairs (FPs) were created in charged cells which contained bipolarons (+2 or -2). Barrier heights and formation energies were lower for 3-coordinated oxygen atoms. Creation of a FP from an electron bipolaron was calculated to have a formation energy of 1.5 eV and a barrier height of 1.73 eV. Creation from a hole bipolaron was calculated to have a formation energy of 0.89 eV and a barrier height of 1.29 eV. Each FP was modelled by an initial displacement of an oxygen atom by 2Å. This was similar to calculations by Bradley et al. [140],
where the same reactions were investigated, but within a 324 atom cell. These calculations should, in principle, be superior since longer-range distortions induced by the defects are included. Bradley et al. [140] also used a GGA functional, but PW91 rather than PBE. Creation of an oxygen FP from an electron bipolaron was calculated to have a formation energy of 1.19 eV and a barrier height of 1.96 eV.

6.2 Results

6.2.1 Frenkel pair formation in m-HfO$_2$

Oxygen FPs were modelled in 324-atom cells of m-HfO$_2$. Firstly, the creation of a neutral oxygen FP was investigated. This means that an oxygen ion in the bulk, O$_{\text{Bulk}}$, is displaced, creating a V$_{\text{O}}$ defect and an interstitial oxygen, O$_{\text{I}}$. There are three possible ways a neutral FP can be created,

\[
O_{\text{Bulk}} \rightarrow O_{\text{I}}^{0} + V_{\text{O}}^{0}, \tag{6.1}
\]

\[
O_{\text{Bulk}} \rightarrow O_{\text{I}}^{-1} + V_{\text{O}}^{+1}, \tag{6.2}
\]

\[
O_{\text{Bulk}} \rightarrow O_{\text{I}}^{-2} + V_{\text{O}}^{+2}. \tag{6.3}
\]

All three of these reactions create neutral FPs, however the charge distribution among the interstitial and vacancy differ. The stability of each of these pairs, as indicated by the formation energy, are also very different. By reference to the formation energy graphs in Chapter 5, one can see that the -2/+2 configuration will be stablist for infinite separation, with a formation energy of roughly 2 eV. Furthermore, the +2 charge state of the oxygen vacancy and the -2 charge state of the O$_{\text{I}}$ have been shown, in previous calculations, to have the lowest diffusion barriers [141, 142]. This is an important feature that the breakdown mechanism must have. Once an oxygen atom is dislocated, the vacancy and interstitial must be able to separate, either by diffusion of the vacancy or diffusion of the ion. Therefore the creation of a $V_{\text{O}}^{+2}$-O$_{\text{I}}^{-2}$ pair was investigated.

The calculated FP configuration for m-HfO$_2$ is shown in figure 6.3. The reac-
6.2. Results

Figure 6.3: (A) The bulk structure of m-HfO$_2$. This is the configurations of the system before the FP creation has occurred. Important atoms, which will take part in the reaction, are highlighted. (B) The final configuration of the neutral (+2/-2) FP in m-HfO$_2$. Oxygen ion 2 displaces oxygen ion 1, and oxygen ion 1 dislocates and moves into an interstitial position. The vacancy is at the location previously occupied by oxygen ion 2. The vacancy and interstitial therefore have an oxygen ‘barrier’ between them.

tion happens by one oxygen ion displacing another, the displaced ion then moving into an interstitial position. The vacancy and interstitial are separated by an oxygen “barrier”. This is necessary to stabilise the FP, otherwise the oxygen ion will immediately recombine with the vacancy. This tendency to recombine can be understood to come from the attractive Coulomb interaction between the oxygen and the vacancy. The formation energy for this FP is 5.5 eV. This is higher than for the case of infinite separation. The barrier height for this process is 6.6 eV (see Fig. 6.4). Such a high activation energy is not unexpected. HfO$_2$ is a stable material and only experiences dielectric breakdown under voltage stressing. It therefore should be the case that generation of vacancies under standard conditions is unlikely. Therefore, the impact of the electric field must be taken into account. The effect the electric field has on the barrier can be estimated by calculating the defect dipole at the initial and transition states of the reaction pathway. Let the defect barrier dipole be $\mathbf{d} = \mathbf{d}_{\text{Tr}} - \mathbf{d}_{\text{In}}$, where $\mathbf{d}_{\text{Tr}}$ and $\mathbf{d}_{\text{In}}$ are the dipoles of the transition and initial state, respectively. An external electric field will interact with the defect barrier dipole, giving an energetic contribution $U = -\mathbf{E} \cdot \mathbf{d}$. Thus, when the electric
6.2. Results

Figure 6.4: (A) The energy profile of the $\text{O}_{\text{Bulk}} \rightarrow \text{O}^{-2} + \text{V}_{\text{O}}^{+2}$ process in m-HfO$_2$. (B) The energy profile of the -2 FP production process in m-HfO$_2$. Reaction coordinate refers to the replica index, with 0 being the initial (bulk) state and 9 being the final (FP) state.

Field is appropriately aligned with the defect barrier dipole, the height of the energy barrier will be reduced. For the reaction shown in Fig. 6.3, the barrier dipole is calculated to be 7.11 eÅ (In CP2K, this is calculated from the Berry phase of the KS states [143, 144]). At a field strength of 1 V nm$^{-1}$ this gives a barrier reduction energy of 0.711 eV (if the field and barrier dipole are perfectly aligned). Thus the barrier would in fact be 5.9 eV. Typical breakdown fields of HfO$_2$ are 1 - 2 V nm$^{-1}$ [145, 146]. Thus, the barriers of this mechanism are still too high to claim responsibility for O vacancy generation at room temperature. Thus the reaction must proceed by a different mechanism.

6.2.1.1 Production of FPs facilitated by electron injection

As discussed above, the electric field can induce the injection of carriers into an oxide thin film. In the production of $\text{V}_{\text{O}}^{+2}$-$\text{O}^{2-}$ defect pairs the Coulomb interaction produced high barriers and formation energies for pairs with low separation, even though at large separation these pairs should be the most stable. It therefore seems reasonable to expect that the injection of carriers could help to reduce the barrier by neutralising one of the defects in the FP. For example, the injection of electrons
Figure 6.5: (A) This initial state of the -2 charged FP generation process in m-HfO$_2$. The electrons are in the CBM state. (B) Final configuration of the -2 FP. Oxygen ion 1 dislocates, leaving behind a neutral vacancy and creating a -2 charge oxygen interstitial.

would alter the reaction, giving

O$_\text{Bulk}$ + 2e$^-$ → V$_O^0$ + O$_I^{-2}$.

(6.4)

The vacancy is now neutral, rather than positive. There are, in fact, two possibilities. This arises from the fact that the electron can be either delocalised as a conduction band state, or they can be localised into a STE state. Since the STE state is very shallow, it will not be stable at room temperature. The calculations here therefore consider the electrons being in the CB in the initial state. The generation mechanism is shown in Fig. 6.5. The final state is more stable than the neutral case with a formation energy of -0.4 eV. The negative formation energy emerges because there is a large stabilisation effect when the two electrons drop from the conduction band minimum state into the mid-gap level of the vacancy. The barrier height is calculated to be 1.9 eV, much lower than in the neutral case. Thus, the electron injection effect alone can reduce the barrier for V$_O$ production by almost 5 eV.

The V$_O^0$-O$_I^{-2}$ defect pair will also have a dipole at the transition state. Calculating this in an identical way to the V$_O^{+2}$-O$_I^{-2}$ defect pair, it is found that the V$_O^0$-O$_I^{-2}$ has a dipole of 8.0 eÅ. At fields of 1 V nm$^{-1}$, this gives a barrier reduction energy of 0.8 eV (for the case of perfect alignment between the E-field and the defect dipole). This would bring the activation energy down to 1.1 eV - a much more reasonable number for processes which should be occurring at room temperature.
6.2. Results

Figure 6.6: Schematic of the conversion of bipolarons into FPs. Hf ions are grey and O ions are red. (a) Shows the hole injection case. Injected holes are localised onto hole bipolaron states. Frenkel defects are created when an oxygen at a hole bipolaron site displaces away. The Frenkel defect is stabilised by hole transfer to the interstitial, resulting in a \( V^{+2}_O - O^{0}_i \) Frenkel pair. (b) shows the case of electron injection. An oxygen atom neighbouring an electron bipolaron site is dislocated from its position. The Frenkel defect is stabilised via the transfer of electrons into \( V^{+2}_O \) to create a \( V^{0}_O \). Images originally published in [4].

6.2.2 Frenkel pair formation in a-HfO\(_2\)

There is an analogous reaction for FP production in a-HfO\(_2\). There is, however, an important difference. Unlike in the crystalline case, trapped electrons and holes are stable at room temperature. Therefore, the reaction proceeds from a doubly trapped electron (or hole), rather than relying of a transition of an electron from the CBM to the newly created valence state. Frenkel pairs in a-HfO\(_2\) were created by displacing an oxygen atom at the bipolaron site and moving it to a nearby oxygen. This leaves behind a neutral oxygen vacancy in its original position and creates an interstitial in the new position. Without trapped electrons or holes, neutral FPs can be formed in a similar way to the monoclinic case. Again, formation of FPs composed of a +2 charged O vacancy and -2 charged O interstitial is energetically more favourable then the other neutral pairs. The presence of a double electron trap facilitates
6.2. Results

Figure 6.7: Histogram of the Frenkel pair formation energies in a-HfO$_2$. ’+2’ Refers to a FP formed at a double hole trap site and ‘-2’ refers to a Frenkel pair formed at a double electron trap site.

creation of stable neutral O vacancies and interstitial O$^{-2}$ ions, which can separate via fast diffusion of O$^{-2}$ ions in the applied field. Hole traps, on the other hand, facilitates the formation of stable V$^{+2}$ vacancies and interstitial O atoms, both of which can diffuse relatively fast.

In particular, hole bipolarons in a-HfO$_2$ form an O-O dimer with a Bader charge of -1.5 $|e|$ and a bond length of 1.4 to 1.5 Å. This causes significant displacements of the network ions and the formation of stable FDs requires about 1.5 eV. The same applies to the case of the electron injection and electron bipolaron formation. Both of these processes are illustrated in Fig. 6.6.

The average FD formation energies in the electron injection case range from 0.8 to 2.5 eV (See Fig. 6.7). The FD formation energies are higher in the amorphous phase than in m-HfO$_2$. This stems from the qualitative difference between the FP formation in a-HfO$_2$ as compared to m-HfO$_2$. In the crystalline case, the initial state
6.2. Results

has electrons in the conduction band. Therefore there is a significant energy gain due to the localisation of these extra electrons in an oxygen vacancy in the process of FP formation [140] (however, the cross-section for formation of bipolarons is rather low). In the amorphous case, however, the second electron can tunnel directly into the single electron polaron state and the FP formation takes place from the already localised bipolaron state and the initial configuration thus has a lower energy level. As a result, the formation energies in the amorphous case are greater than in the crystalline case.

Nudged elastic band simulations were used to estimate the barrier heights to Frenkel defect formation. Barrier heights were calculated for a range of Frenkel defect configurations, which span the values of formation energy. The low energy Frenkel defect is predicted to have a barrier height of 1.5 eV whereas the high energy FP is calculated to have a formation barrier height of 2.8 eV (The distribution is given in Fig. 6.7). For the hole injection case, the formation energy is, on average, 1.6 eV (St. Dev. = 0.3 eV). The barrier heights to formation from holes are estimated to be between 2.1 and 3.0 eV. The corresponding barrier for FD formation from a hole bipolaron in m-HfO$_2$ is 1.9 eV.

The results demonstrate that both electron and hole injection can facilitate formation of O vacancies and interstitials in a-HfO$_2$ via lowering energies and barriers for formation of stable charged defects mobile in the electric field. The formation energies for the electron-induced FD formation are lower in m-HfO$_2$, but cross-sections for charge trapping at intrinsic sites are higher in a-HfO$_2$ due to the existence of intrinsic trapping sites. New O vacancies could be potentially formed near the pre-existing ones too, as has been suggested for m-HfO$_2$ in ref. [140]. This effect will be discussed next.

6.2.3 Interaction Between Oxygen Vacancies

As discussed above, breakdown is caused by the generation of an oxygen deficient filament through the dielectric. This means that as the dielectric degrades, vacancies will be generated close to already-existing vacancies. Thus, interaction between vacancies must be considered. To study the interaction between vacan-
cies in a-HfO$_2$, 15 different vacancy configurations were simulated. To produce di-vacancy models a subset of 9 vacancies was selected, and for each vacancy in this subset 3 di-vacancies were produced by deleting adjacent oxygen atoms. This produces 27 di-vacancy models in total. The choice of the original set of parent vacancies was made so as to reflect the distribution of coordination numbers of oxygen in the a-HfO$_2$ models.

The interaction between neutral vacancies in a di-vacancy can be quantified by calculating the difference between the ‘first’ and ‘second’ formation energy. The first formation energy is calculated using standard equation given in the methodology. The second formation energy, $E_{F}^{2nd}$, is calculated using $E_{Vac}$ in place of $E_{Bulk}$, that is $E_{F}^{2nd} = E_{Divac} - E_{Vac} + \mu_O$. In the case of a crystal, if $E_{F}^{2nd}$ is lower than $E_{F}$, the vacancies are said to be attractive. If $E_{F}^{2nd} > E_{F}$, the vacancies are said to be repulsive. However, in an amorphous systems there is a distribution of vacancy formation energies. As shown in Chapter 5, the formation energy of an oxygen vacancy has a spread. This spread comes from the variation of possible bond lengths and coordination spheres intrinsic to the amorphous structure. It is therefore the
6.2. Results

Figure 6.9: (A) Histogram of the distribution of formation energies of oxygen divacancies. Formation energy is referenced to pristine amorphous structure (E_{Bulk}). (B) Histogram of the distribution of formation energies of oxygen divacancies. Similar to (A), the formation energy is referred to the pristine bulk.

case that $E_{F}^{2nd}$ will differ from $E_{F}$ simply due to this statistical variation alone. In addition, there will also likely be a variation of the degree of interaction between vacancies (which form the divacancy). The distribution of second formation energies, therefore, will be a convolution of the distribution of formation energy due to the amorphous structure, and the distribution of interaction energies between the two vacancies. It is therefore the case that if, for example, $E_{F}^{2nd} < E_{F}$ for some adjacent vacancies then it may simply be due to the intrinsic variation, rather than the variation of vacancy-vacancy interaction. To rule out this possibility, it is necessary to take a statistical sample of divacancies and focus on the shift of the average second formation energies compared to the first.

On average, the second vacancy formation energy is 0.1 eV lower than the first. The standard deviation, however, is 0.4 eV. This means that some di-vacancy configurations will be of the repulsive type, even though the average case is attractive. A typical di-vacancy configuration in α-HfO$_2$ is shown in Fig. 6.8. Interaction between the vacancies is visualized by the plot of the di-vacancy doubly occupied electron states. Instead of two localized states adjacent to one another, the plot shows two states which are extended across the vacancies in a bond-like fashion. This type of molecular bonding type configuration corresponds to the majority
of cases (approximately 80%). Further stabilization is possible for tri-vacancy clusters. The interaction in tri-vacancies can be quantified in a similar manner to di-vacancies by taking the third vacancy formation energy and comparing it to the second. It is found that the third vacancy formation energy is, on average, 0.4 eV lower than the second. Oxygen tri-vacancies have two general configurations, they can be organized in a linear or triangular fashion (see Fig. 6.10). The standard deviation of the tri-vacancy formation energy distribution is 0.5 eV. This means that, similar to the di-vacancy case, there exists tri-vacancy aggregates that are of the repulsive type, despite the fact that, on average, tri-vacancies are attractive.

The implication of these results is that creation of di-vacancies or tri-vacancies is energetically cheaper than creation of a two or three isolated vacancies. Thus, one can infer a correlated mechanism of $V_O$ production.

6.3 Conclusions

In this chapter, the generation of oxygen vacancies in HfO$_2$ has been investigated. It has been shown that the direct effect of an E-field alone is not enough to account for oxygen vacancy generation. Instead, it is the field-induced injection of carriers which are mostly responsible for reducing the activation energy of oxygen FP production. The mechanism proceed by injecting electrons (holes) and creating a neutral (doubly positive) vacancy and doubly negative (neutral) oxygen interstitial. Similar calculations in a-HfO$_2$ show this mechanism to also apply in the amorphous case. Furthermore, the existence of binding in $V_O$ defects in a-HfO$_2$ implies that
### 6.3. Conclusions

#### Carrier Trapping Data

<table>
<thead>
<tr>
<th>Trap</th>
<th>Spin State</th>
<th>Trapping Energy</th>
<th>Trapping energy Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hole</td>
<td>Doublet</td>
<td>1.4</td>
<td>0.9</td>
</tr>
<tr>
<td>Electron</td>
<td>Doublet</td>
<td>1.0</td>
<td>0.7</td>
</tr>
</tbody>
</table>

#### Oxygen Vacancy Data

<table>
<thead>
<tr>
<th>Charge State (q)</th>
<th>Spin State</th>
<th>$q$ to $q-1$ CTL</th>
<th>Total CTL Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>+2</td>
<td>Singlet</td>
<td>3.9</td>
<td>0.8</td>
</tr>
<tr>
<td>+1</td>
<td>Doublet</td>
<td>3.6</td>
<td>0.8</td>
</tr>
<tr>
<td>0</td>
<td>Singlet</td>
<td>5.4</td>
<td>0.8</td>
</tr>
<tr>
<td>-1</td>
<td>Doublet</td>
<td>4.9</td>
<td>0.5</td>
</tr>
<tr>
<td>-2</td>
<td>Singlet</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

#### Oxygen Interstitial Data

| 0                | Singlet    | 1.6              | 0.8              |
| -1               | Doublet    | 1.2              | Negligable       |
| -2               | Singlet    | -                | -                |

#### Frenkel Pair Formation

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Formation E Average</th>
<th>Barrier Height</th>
<th>Barrier Height Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2e^- \rightarrow V_{0}^{0} + O_{l}^{-2}$</td>
<td>1.7</td>
<td>1.9</td>
<td>1.5 to 2.8</td>
</tr>
<tr>
<td>$2h^+ \rightarrow V_{0}^{+2} + O_{l}^{0}$</td>
<td>1.6</td>
<td>2.6</td>
<td>2.1 to 2.8</td>
</tr>
</tbody>
</table>

**Table 6.1:** Summary of the defects in a-HfO$_2$. “Spread” is quantified as the energy difference between the lowest and highest value is the range of the given value. All energies are given in eV. In CTLs, a dash represents that further charging is not possible.

#### Carrier Trapping Data

<table>
<thead>
<tr>
<th>Trap</th>
<th>Spin State</th>
<th>Trapping Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hole</td>
<td>Doublet</td>
<td>0.45</td>
</tr>
<tr>
<td>Electron</td>
<td>Doublet</td>
<td>0.23</td>
</tr>
</tbody>
</table>

#### Oxygen Vacancy Data

<table>
<thead>
<tr>
<th>Charge State (q)</th>
<th>Spin State</th>
<th>$q$ to $q-1$ CTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>+2</td>
<td>Singlet</td>
<td>4.3</td>
</tr>
<tr>
<td>+1</td>
<td>Doublet</td>
<td>4.4</td>
</tr>
<tr>
<td>0</td>
<td>Singlet</td>
<td>5.9</td>
</tr>
<tr>
<td>-1</td>
<td>Doublet</td>
<td>6.0</td>
</tr>
<tr>
<td>-2</td>
<td>Singlet</td>
<td>-</td>
</tr>
</tbody>
</table>

#### Oxygen Interstitial Data

| 0                | Singlet    | 2.15             |
| -1               | Doublet    | 3.0              |
| -2               | Singlet    | -                |

**Table 6.2:** Summary of the defects in m-HfO$_2$. All energies are given in eV. For the CTLs, a dash indicates that further charging is not possible.
conducting filaments in a-HfO$_2$ could be formed by a correlated creation mechanism, where formation energies of oxygen vacancies are reduced by the presence of already generated V$_O$ defects.

The previous three chapters have presented results on HfO$_2$. Table 6.2 summarises the results for the defects investigated for a-HfO$_2$. In the first chapter, it was shown that electrons and holes can trap into deep, polaron-like states in a-HfO$_2$ without the need for defects. Such trap states, referred to as intrinsic traps, should play an important role in the reliability of a-HfO$_2$ films in metal-insulator-metal (MIM) and metal-insulator-semiconductor (MOS) devices which employ a-HfO$_2$ as the insulating layer. Traps will facilitate leakage current via trap assisted tunneling. In the case of MOSFETS, they will also produce random telegraph noise (in the drain current) and threshold voltage shift. The subsequent chapter then focused on oxygen vacancy and interstitial defects. Since oxygen vacancies can charge, they will exhibit the same reliability and degradation problems that intrinsic traps do.

In addition, oxygen vacancies are known to play an important role in the breakdown of HfO$_2$ insulating films. For this reason, oxygen vacancy generation processes were studied in this chapter. It has been shown that the intrinsic traps discussed in Chapter 4, and the vacancies discussed in Chapter 5, are linked. Intrinsic traps, in a-HfO$_2$, play a crucial role in the mechanism of oxygen vacancy generation, and thus in the process of dielectric breakdown. These results imply that the breakdown characteristics of HfO$_2$ will be different between amorphous and crystalline phases. For example, in a-HfO$_2$, it is predicted in this chapter that breakdown can begin at intrinsic precursor sites. Thus, defect-free sample of a-HfO$_2$ should still undergo breakdown at typical breakdown field strengths of 5 to 10 MVcm$^{-1}$. m-HfO$_2$, however, will begin the breakdown process at already present vacancies or at grain boundaries, and it may be harder to break highly clean samples of m-HfO$_2$. 
Chapter 7

Hexagonal Boron Nitride - Structure and Defects

Hafnia has been used as a dielectric layer in standard MOS devices. There is growing interest, however, in newer designs for nanoscale electronic devices. The discovery of graphene [147, 148, 149], and other single layer and multi-layer 2D systems [150, 151], has provoked interest in the use of 2D and layered systems in device design [152, 153, 154]. Hexagonal boron nitride (hBN) is iso-structural to graphite. Nitrogen and boron atoms form a honeycomb structure and are organised into layers which are stacked in an ABAB fashion. In a similar way to graphite, hBN can also be fabricated in a mono- or multilayer form. One of the major differences between hBN and graphite (or graphene) is that hBN has good insulating properties. The band gap of hBN has been measured to be in the range of 5.5 - 6 eV [155]. Similar to the case of graphene, the $sp^2$ bonding character leads to high mechanical strength and thermal conductivity [156]. Also, hBN has been investigated as a useful substrate in graphene based technologies [157, 158, 159]. hBN is therefore a viable candidate as a dielectric layer in graphene based microelectronic devices.

Fig. 7.1, reused from Chapter 1, illustrates the detrimental impact of dielectric layer defects on devices. Like with other dielectrics, such as HfO$_2$ and MgO, carrier trapping at defects (process 2) in the dielectric layer can negatively affect the functionality of the device. In this chapter, this is explicitly investigated by calculating the
7.1 Previous Studies on hBN

Figure 7.1: The impact of defects on the function of a dielectric in a MOSFET device. Defects, indicated by red dots, can be charged. Charging can affect the conduction through the n-channel. Also, electron trapping at a defect allows for trap assisted tunnelling (TAT) through the layer. This produces a small current through the insulating layer. If the defect concentration is high enough, a 'conductive filament' can exist, which conducts electrical current through the layer. Such filaments are responsible for dielectric breakdown (see Chapter 6).

energetic position of the VBM and CBM with respect to the vacuum. This allows one to calculate the VBM offset with the electrode Fermi level (see section 7.3.1.1). From this, the charge state of a given defect can be calculated. It is therefore useful to characterise the structure and behaviour of defects, in their multiple charge state, in hBN. Also, there has been growing interest in hBN as a RRAM material. Studies have shown conductive filaments to form in hBN when under electrical bias [160, 161]. Some studies, such as [160], demonstrated the boron vacancies in particular play a role in the breakdown of hBN, via a mechanism like that of process 3 in Fig. 7.1. It is therefore clear that atomistic modelling of point defect in hBN can help to make predictions about the applicability and reliability of hBN in future nano-scale devices.

7.1 Previous Studies on hBN

Bulk h-BN was previously calculated by Blase et al. [162] in 1995. Using LDA and GW levels of theory, the pristine band structure was studied. As is typical of the LDA, the band gap was predicted to be 3.9 eV, an underestimation. Using the GW method, however, the band gap was more accurately found to be 5.4 eV. It was found that, although the LDA severely underestimates the band gap, the ordering of bands is correct. The VBM position was found to be at the K-point and the CBM
was found to be at the M point of the Brillouin zone. Thus hBN has an indirect and non-$\Gamma$ band gap. The agreement of band-ordering between LDA and GW occurs because the self-energy was weakly $k$ dependent. Thus using LDA KS orbitals in the quasiparticle spectrum of the GW calculation led to very little change to the $k$-dependance of the band structure. These results were verified using optical spectroscopy by Cassabois et al. [155], who demonstrated that bulk hexagonal boron nitride has an indirect band gap of 5.955 eV.

There have been a number of studies on defects in hBN and hBN-derived structures. Huang et al. [163] studied boron and nitrogen defects in the hBN monolayer using both the PBE and HSE DFT functionals. HSE is found to predict 4 (formal) charge states for the $V_{B}$ defect, ranging from -3 to 0. The $V_{B}^{0}$ undergoes a symmetry-lowering distortion from $D_{3h}$ to $C_{2v}$. The -1 charge state is found to have $D_{3h}$ symmetry and the -2 and -3 charge states are found to have $C_{3v}$ symmetry. In all cases, the KS defect states are localised around the surrounding N atoms. Calculations of the $V_{N}$ defect found that, in the neutral case, there is only one occupied gap state. This gap state is $\pi$-like in nature, similar to findings by A. Zunger et al. [164] for bulk hBN. Furthermore, there are negligible distortions and the $V_{N}^{0}$ defect has $D_{3h}$ symmetry. Another difference emerges to the $V_{B}$ defect when considering the charge states. There are only three possible charge states of the N-vacancy, ranging from -1 to +1. Thus the boron and nitrogen vacancy in monolayer hBN are found to act very differently to one another. The neutral defect formation energies for N rich-conditions were calculated as 7.65 eV and 8.47 eV for the $V_{B}^{0}$ and $V_{N}^{0}$ defects, respectively. To calculate the $V_{B}$ formation energy for N-rich conditions, reference [163] employed a thermodynamic argument: It is argued that if the BN sheet is in equilibrium, the $\mu_{B}$ and $\mu_{N}$ should satisfy the condition $\mu_{B} + \mu_{N} = \mu_{BN}$, where $\mu_{BN}$ is the total energy of a BN pair in a monolayer of BN. $\mu_{N}$ is selected to represent N-rich conditions, and so the boron chemical potential is calculated as $\mu_{B} = \mu_{N} - \mu_{BN}$. The authors also looked further into the layering effects of hBN on defect properties, however only using PBE owing to the computational expense of HSE. Some disagreement between the single layer and bulk properties were found.
For example, there is no stable -2 charge state for the V\textsubscript{B}. Most of the defect properties, however, were quite similar between the single layer and bulk calculations.

There have also been attempts to model defect generation processes. A. Zobelli et al. [165] modelled the effect of knock-on radiation damage by electron beams. It was argued that, although V\textsubscript{B} generation is slightly less energetically favourable than V\textsubscript{N} generation, V\textsubscript{N} creation has a lower cross section as the boron atom's lower mass allows for greater energy transfer during collision. Thus it is expected that under electron beam irradiation, boron vacancies may be created more readily than nitrogen vacancies. The creation of boron vacancies via electron beam irradiation has been seen in some experimental studies. C. Jin et al. [166] created hBN monolayers using an electron beam to reduce down multi-layer hBN. Using high resolution transition electron microscopy (HRTEM) a number of single and poly-defects were found in the layer, and were identified as being V\textsubscript{B} type.

There has also been extensive characterisation of hBN as a dielectric in MIM structures. C. Pan [160] et al. studied the nature of breakdown in Cu/hBN/Ti structures using electron energy loss spectroscopy (EELS). It found that breakdown occurs by creation of a conductive filament which is boron deficient and also contains a significant number of Ti. Thus, the generation of boron vacancies is a crucial step in the breakdown of hBN dielectrics.

### 7.2 Methodology

Bulk hBN is simulated in periodic boundary conditions using a \(6 \times 6 \times 4\) unit cell. This extension of the cell ensures that the correct LUMO and HOMO are both included in the description of the electronic structure (see section 2.6.1). Each cell contains 576 atoms. The electronic structure is calculated using the PBE0-TC-LRC functional with a \(2\,\AA\) cutoff. These is the standard set-up of the PBE0-TC-LRC functional and, combined with the appropriate supercell extension, predicts a band gap which is in good agreement with experiment (section 7.3.1). Periodic charge and potential alignment corrections and are applied using the method described in Chapter 2.
Materials such as hexagonal boron nitride and graphite are formed by layers held together by van der Waals forces. Density functional theory is well known to fail [167] at reproducing these interactions, and so additional methods are required to improve the interlayer interaction. The project uses the dispersion correction developed by Grimme et al. known as DFT-D3 [168]. This uses a force field to calculate the dispersion energy, \( E_{\text{disp}} \), which is then subtracted from the KS energy, that is, \( E_{\text{Total}} = E_{\text{KS}} - E_{\text{disp}} \). The force field uses atom-pairwise dispersion coefficients and contains both two-body and three-body terms. The two body term was calculated from the two-body dispersion energy expansion;

\[
E^{(2)} = -\sum_{\text{AB}} \sum_{n=6,8,10,...} s_n C_{n}^{\text{AB}} \frac{R_{\text{AB}}^{n}}{R_{\text{AB}}} f_{\text{damp}}(R_{\text{AB}}) \tag{7.1}
\]

where \( s_n \) is a scaling factor, adjusted for a particular XC functional, \( C_n^{\text{AB}} \) is an isotropic nth-order coefficient for the atom pair AB, \( R_{\text{AB}} \) is the internuclear distance between A and B, and \( f_{\text{damp}}(R_{\text{AB}}) \) is a damping term which prevents a singularity at low \( R_{\text{AB}} \) and also to avoid double counting of correlation at mid-range distances. Double accounting occurs because DFT XC functionals replicated dispersion effects well at short-range [169] but very poorly at long-range. If the dispersion correction is not damped for low and intermediate values of \( R_{\text{AB}} \), therefore, double counting of dispersion effects will occur. In DFT-D3, the scaling factor for \( n = 6 \) \( (s_6) \) is kept at one to ensure the correct asymptotic behaviour. Since the \( n > 6 \) terms are more short-ranged, they will have an impact of the correlation double-counting problems at short-range. The higher order \( C_n \) terms are therefore XC functional dependent (since the double counting of correlation depends on how correlation is described at the DFT level). In DFT-D3, the above sum is truncated at \( n=8 \), since higher order terms introduce too many instabilities. The main advantage of DFT-D3 is that it is computationally cheap, giving greater efficiency compared to non local van der Waals functionals [170].
7.3 Results

7.3.1 Bulk hBN

The optimised bulk hBN structure with AB stacking is plotted in Fig. 7.2. In Fig. 7.2 the optimised interlayer spacing of the honeycomb sheets has been indicated as 3.27 Å (giving $c$ as 6.54 Å). This is reasonably close to experimental results \[171\]. $a$ is calculated to be 2.49, which is also in good agreement with experiment. The B-N bond length is 1.4 Å. As discussed above, the VBM and CBM are believed to be at the M and K points, respectively. The 6x6x4 supercell extension was chosen to sample these points. The calculated DOS is shown in Fig. 7.3. The calculated KS band gap is 5.7 eV, which is in good agreement with the range of band gap values reported in the literature \[155, 172, 173\]. The HOMO and LUMO are both non-$\Gamma$ point states. It can be seen in Fig. 7.3 that the valence band is predominantly formed from N states and the conduction band is predominantly formed from B states. Both bands, however, show significant hybridisation. This is an indication of the covalent nature of the bonding in hBN.
7.3. Results

Figure 7.3: The PDOS of a $6 \times 6 \times 4$ cell of bulk hexagonal boron nitride. The predicted band gap is 5.7 eV

7.3.1.1 Alignment of the VBM and CBM

In a metal-insulator-metal (MIM) structure, the charge state of insulator defects depends on the position of the metal Fermi level and the charge transition levels of the defect. In theoretical calculations, the position of the metal Fermi level must be referenced to the position of the band edges of the insulator. This was done by simulating a 5-layer slab of hBN in a 3D PBC cell with a 15 Å of vacuum present. The VBM position can then be calculated relative to the vacuum potential. Then, one takes the metal work function as the metal Fermi level relative to the vacuum potential. This gives the VBM position and the metal Fermi level position referenced to the same value. Thus, the Fermi level offset (with the VBM) is calculated as

$$\varepsilon_{\text{Offset}} = \Phi_{\text{WF}} - \varepsilon_{\text{VBM}}$$  \hspace{1cm} (7.2)

The VBM position relative to the vacuum is calculated using a 5 layer slab of hBN. This gives the VBM position to be 6.9 eV below the vacuum, and the CBM to be 1.2 eV below the vacuum. This is in close agreement with ultraviolet photoelectron spectroscopy (UPS) results, which give a range of VBM levels of 6.98 to 7.35 eV below the vacuum level [174]. Typical electrode wavefunctions vary from 4 to 5 eV. For example, the workfunction of graphene ranges between 4.5 and 4.8 eV [175]. The workfunction of copper ranges between 4.46 to 4.94 eV [176, 177].
7.3. Results

7.3.2 The Stone-Wells Defect

hBN is isostructural to graphite. The Stone-Wales defect was first investigated for carbon based systems [178], and is composed of two adjacent carbon atoms which have been rotated by 90 degrees. Analogously, it may be expected that a Stone-Wales defect will exist in hBN, with a BN unit rotated by 90 degrees. Fig. 7.4 shows the Stone-Wells defect in hBN. The formation energy of the Stone-Wales defect is 5.4 eV, which is independent of chemical environment (no N or B atoms are lost or gained). The PDOS is shown in Fig. 7.5. Interestingly, this is very close to predicted formation energies of Stone-Wales defects in graphite of 5.4 [179] and 5.8 [180] eV.
7.3.3 Vacancy Defects in hBN

Removing a boron atom from hBN will also remove 3 valence electrons from the system. The neutral boron vacancy therefore introduces three holes into the system. These holes are located on the N atoms surrounding the vacancy. The PDOS and spin density of the boron vacancy are plotted in Fig. 7.6. The ground state of the boron vacancy is the doublet spin configuration, with the quartet being 0.3 eV higher in energy. In the neutral case, before relaxation, the surrounding N atoms form an equilateral triangle. After geometry optimisation, however, one N-N distance is longer than the other two. This asymmetric distortion reduces the point group from $D_{3h}$ to $C_{2v}$. The three unoccupied states in the gap are non-degenerate. The hole states themselves can be seen to have strong $p$ character, and are located 2.6, 3.1
and 3.9 eV above the VBM. Bader spin analysis gives a magnetic moment of 0.78 $\mu_B$ on two of the adjacent N atoms and -0.73 $\mu_B$ on one of the adjacent N atoms. Bader charge analysis gives charges of $+2 |e|$ on the surrounding N atoms.

The formal charge state of the boron vacancy ranges from -2 to 0. The 0/-1 and -1/-2 charge transition levels (CTLs) are found to be at a Fermi level of 2.1 eV and 5.2 eV, respectively (see Fig. 7.7). This indicates that the boron vacancy will, in thermal equilibrium, be in its -1 charge state when hBN is interfaced with graphene. In the -1 charge state, the ground state of the boron vacancy is a triplet state (see Fig. 7.6), with spin density distributed equally across all three nitrogen atoms. The nitrogen atoms also form an equilateral triangle, that is all N-N distances are equal. The N-N distances are also shorter (by at least 1.3 Å) than the N-N distances in the neutral $V_B$ defect. Each nitrogen atom has a magnetic moment of 0.67 $\mu_B$.

In the -2 charge state the $D_{3h}$ symmetry is broken, with the defect now having symmetry of the $C_{2v}$ group. The surrounding N atoms form an isosceles triangle. The spin configuration is a doublet. Thus the geometric structure and spin configuration is similar to the $V^0_B$ defect. Bader charge analysis shows that the charge of the defect is $-1.5 |e|$.

The isolated N vacancy ($V_N$) is shown in Fig. 7.8. The $V_B$ defect introduced three unoccupied states into the gap. The $V_N$ vacancy, however, introduces only
two KS states into the KS gap, one of which is occupied. Furthermore, unlike in the case of the \( V_B \) defect, these states are \( \pi \)-like in nature, with the KS orbitals being spread around the surrounding B atoms. The \( V_N \) defect has three different charge states: -1, 0 and +1. The \( V_N^{-1} \) geometry has \( D_{3h} \) symmetry and its spin configuration is a singlet Bader charge analysis gives a total charge of 0.8 \( |e| \) localised onto each adjacent B atom. The \( V_N^0 \) ground state is a doublet. The surrounding structure has asymmetric distortions, with one N-N distance being approximately 0.3 \( \text{Å} \) shorter than the other two. This occurs because the boron atom is displaced out of plane. The three NN B atoms have magnetic moments of 0.14, 0.15 and 0.48 \( \mu_B \). The \( V_N^{+1} \) is, like the \( V_N^{-1} \) defect, in a singlet spin configuration. The surrounding N atoms form an equilateral triangle, and the geometry has \( D_{3h} \) symmetry. The CTLs of the \( V_N \) defect are 3.6 eV for the +1/0 CTL and 4.6 eV for the 0/-1 CTL (Fig. 7.17).

### 7.3.4 Divacancies in hBN

Boron divacancies can exist in two main configurations - in-plane and inter-plane. The in-plane boron divacancy is plotted in Fig. 7.9. The singlet-triplet splitting in this vacancy complex is very small (order of meV), and thus the singlet and triplet states are considered defacto degenerate in these calculations. Interestingly, the binding energy, as calculated by the equation \((E_{\text{Divac}} + E_{\text{bulk}}) - 2E_{\text{Vac}}\), is 4.9 eV.

When forming a B divacancy, there is one N atom which has two bonds broken. As can be seen in Fig. 7.9(A), the N atom moves to a location previously occupied.
7.3. Results

Figure 7.9: (A) The in-plane boron divacancy in the singlet spin configuration. Surfaces indicate spin density, with yellow representing positive (or up) spin and green representing negative (or down) spin. In the divacancy, an N atom, previously bonded to two B atoms, migrates to form a bond with another two N atoms adjacent to the vacancies. (B) Shows the PDOS of the in-plane boron divacancy in the singlet spin configuration. Iso-surfaces have a value of 0.025.

by a B atom and forms bonds with two neighbouring N atoms (both of which were previously bonded to a shared B atom). These two N-N bonds greatly stabilize the divacancy. The length of these bonds is 1.38 Å. The lower gap states seen in Fig. 7.9(B) are all occupied in the neutral state. These states are mostly associated with N atoms and emerge from the N-N bonds formed by the displaced N atom. The higher states, close to the conduction band, are unoccupied in the neutral state.

The triplet spin density and PDOS can be seen in Fig. 7.10. The geometric configuration is virtually identical to the singlet case. The only difference is that in the triplet configuration, all the defect states have aligned spin. Furthermore, just as in the singlet case, the low energy defect states near the conduction band are occupied in the neutral configuration and the gap states near the conduction band are unoccupied.

Boron vacancies can also exist in an inter-plane configuration. The configuration and PDOS are plotted in Fig. 7.11. Interestingly, this configuration shows no binding energy, that is, the formation energy of an interplane boron divacancy is the same as if one were creating boron vacancies which are infinitely separated. It can also be seen in Fig. 7.11(B) that the vacancy defect levels are in similar positions to that of the isolated boron vacancy. One can then interpret this divacancy
7.3. Results

Figure 7.10: (A) The in-plane boron divacancy in the triplet spin configuration. Surfaces indicate spin density. In the divacancy, an N atom, previously bonded to two B atoms, migrates to form a bond with another two N atoms adjacent to the vacancies, identical to the singlet case. (B) Shows the PDOS of the in-plane boron divacancy in the triplet spin configuration. Isosurfaces have a value of 0.025.

Figure 7.11: (A) The inter-plane boron divacancy. The ground state of this divacancy configuration has a singlet spin arrangement. (B) The PDOS of the inter-plane boron divacancy. The states are in very similar places to the isolated boron vacancy. This similarity emerges because of the minimal interaction between the boron vacancies between layers, thus the interplane divacancy behaves as if it is two noninteracting boron vacancies. Iso-surface values are 0.03.

as two close but noninteracting single boron vacancies. The whole divacancy is in a singlet state, but each of its vacancies (which form the divacancy) is in a doublet configuration (with opposite orientation to one another).

The interplane boron divacancy can be negatively charged. This corresponds to adding an electron to either of the boron vacancies, and this then corresponds to
7.3. Results

Figure 7.12: The N-bridged boron divacancy. (A) shows an orthographic view where the bridging between layers is clear. (B) gives a perspective view, showing the spin density (with different colouring representing different spin). Iso-surface values are 0.03.

A $V_B^0 V_B^{-1}$ pair. The ground state spin configuration is a doublet. A second electron can be added which then localises around the other boron vacancy, to form a $V_B^{-1}$. $V_B^{-1}$ pair. The stable spin configuration of the -2 interplane divacancy is a singlet.

The configuration of the Boron interplane divacancy described above has two boron vacancies which are weakly interacting. Each vacancy has three holes distributed around its surrounding nitrogen atoms. A more stable configuration can be achieved by displacing two nitrogen atoms and forming a N-N bond which bridges the two layers. This configuration is 0.4 eV lower in energy than in the non-bridged interplane configuration. In Fig. 7.12(B) it can be seen that each vacancy is surrounded by spin density of one orientation, but each vacancy has oppositely oriented spin. The nitrogen dimer itself is in a singlet configuration. The triplet configuration is approximately 0.1 eV lower than the singlet configuration. The triplet configuration simply has one of the holes on the nitrogen atoms flipped.

The N-N bridge greatly distorts the surrounding structure. The neighbouring boron atoms are displaced by 0.5 Å. This distortion is energetically costly, however the formation of the bridge is, overall, energetically favourable due to the high stability of the N-N bond. The N-N bond length is 1.47 Å.

The neutral N divacancy is shown in Fig. 7.13(B). Removing two N atoms creates a B atom with two dangling bonds. This atom moves to form a bond with
Figure 7.13: (A) The -1 charged N divacancy. (B) The Neutral charged vacancy. (C) The +1 charged vacancy. Iso-surface is 0.02.

an adjacent B atom. The binding energy of the N divacancy is 2.9 eV. This is much lower than the binding energy of the B divacancy. This may be due to the fact that part of the stabilisation of the divacancy comes from the formation of a B-B bond. In the B divacancy, however, the stabilisation will come from a N-N bond, which is much stronger. The ground state of the neutral N divacancy is a triplet state, which is 1.4 eV lower than the energy of the singlet state. One boron atom has two boron neighbours, and has a magnetic moment of 1 µB. The rest of the spin density is distributed over a number of B atoms. The -1 charge state is also shown in Fig. 7.13(A). In the -1 charge state, the quartet is more stable than the doublet, by 0.95 eV. Spin density is spread over the boron atoms adjacent to the divacancy. Again there is a boron atom which is 2 coordinated with respect to boron which has a magnetic moment of 1 µB. Other boron atoms surrounding the vacancy have magnetic moments ranging from 0.2 to 0.4 µB. The +1 charge state is shown in Fig. 7.13(C). The groundstate of the N-divacancy is a spin doublet, with the quartet spin state being 2.6 eV higher. The spin density is mostly distributed around three boron atoms (See Fig. 7.13(C)). The largest proportion of the spin density is, again, focused on the B atom which is 2C with respect to boron. This boron atom has a magnetic moment of 0.4 µB.

7.3.5 Interstitials Defects in hBN

Fig. 7.14 shows the configuration of the neutral B interstitial (B₀¹ or B₁). The B₀¹ defect sits between two layers. Directly “above” (Fig. 7.14(A)) the interstitial sits another B atom, and below it a N atom. The boron atom is pulled out of the plane slightly, as can be seen in Fig. 7.14(A). The ground state spin configuration is a
7.3. Results

Figure 7.14: The neutral boron interstitial. (A) gives a far of few with the two layers adjacent to the $\text{B}_I$ visible. (B) Is the $\text{B}_I$ and its NN with bond lengths indicated in Å. The spin density iso-surface value is 0.01.

doublet. Mulliken population analysis gives a spin moment of 0.54 and a charge of 0.02 on the boron interstitial.

The $\text{B}_I^0$ defect introduces a total of 6 KS states into the band gap. Three of these are occupied and three are unoccupied. The $\text{B}_I$ defect can exist in the charge states ranging from -2 to +1. Charging outside this range simply occupies (or deoccupies) perturbed band states. Fig. 7.15 show the formation energy of the $\text{B}_I$. There are many CTLs in typical Fermi level regions. This means that $\text{B}_I$ defects could be responsible for trap-related instabilities. The neutral B interstitial atom has a charge of +0.9 $|e|$, compared to a Bader charge of +3 on B atoms in bulk hBN. The nitrogen atom directly beneath the $\text{B}_{\text{Int}}$ has a Bader charge of -3.9 $|e|$, compared to a Bader charge of -3 $|e|$ on N atoms in the bulk. Thus some electron density is donated from the $\text{B}_I$ to its nearest neighbour N atom. The -2 charge state of the $\text{B}_I$ defect has the boron interstitial with a charge state of -1 $|e|$. The remaining negative charge is spread over two of the adjacent B atoms. The ground state of the $\text{B}_I^{-2}$ defect has a doublet spin configuration. The B interstitial has a magnetic moment of 0.6 $\mu_B$. The -1 charge state of the $\text{B}_I$ has the boron interstitial in the -2 charge state. The surrounding atoms have a cumulative charge of +1 $|e|$. The ground state of the $\text{B}_I^{-1}$ is a spin singlet. In the +1 charge state of the $\text{B}_I$ defect, the singlet and triplet energies are too close to be distinguished using DFT (0.01 eV). In the singlet state, the charge of the boron interstitial is +1 $|e|$. In the triplet state, the charge of the
7.3. Results

Figure 7.15: The formation energy of the B interstitial is its possible charge states. Crossover point indicate the charge transition levels. The Fermi level is set in each case so that the $B^0_I$ and $N^0_I$ formation energy is 0.

Figure 7.16: (A) The neutral N interstitials and its two adjacent layers. The nitrogen interstitial sits closer to one layer, and is separated from in nearest neighbour nitrogen by 1.5 Å. (B) Shows the surrounding structure of the N interstitial. The orange surface shows an isosurface = 0.01 of the spin density.

boron interstitial is $1.6 \ |e|$. The magnetic moment of the boron interstitial is $1.3 \ \mu_B$.

The neutral N interstitial is shown in Fig. 7.16. Unlike the B$_I$ defect, which sits approximately halfway between the layers, the N$_I$ defect sits close to one layer. The N$_I$ sits 1.5 Å away from its nearest N neighbour. Bader analysis gives the N interstitial a magnetic moment of $0.9 \ \mu_B$. The Bader charge of the N interstitial is $-0.85 \ |e|$. The -1 charge state is in a singlet spin configuration. The N interstitial has a Bader charge of $-1.7 \ |e|$. Any extra electrons added to the cell go into the CBM, implying that the N$_I$ cannot be charged past -1.

The $N^{+1}_I$ defect is a triplet spin state. The N interstitial in the $N^{+1}_I$ defect has
7.4 Conclusion

A wide range of intrinsic point defects in hBN have been investigated in this chapter. The charge transition levels of some of these defects are shown in Fig. 7.17. As discussed in chapter 1, the presence of charge transition levels in the band gap leads to charging of defects and subsequent instabilities in devices. In Fig. 7.17 it can be seen that the $V_B$ defect has its 0/1 transition just below the mid-gap position (2.1 eV above the VBM). As mentioned above, one can estimate the charge state of a defect, neglecting band bending, by comparing the VBM position, calculated in section 7.3.1.1, to the work function of a metal electrode. For example, take the workfunction of Cu to be approximately 4.5 eV [176]. This will give the valence band offset of the hBN/Cu interface to be 2.4 eV, and thus one could expect the $V_B$ defect to be in the -1 charge state. In fact, most of the defects are likely to be charged for typical Fermi level positions.

Divacancies have also been investigated in this chapter. It is found that $V_B$ polyvacancies have higher binding energies than $V_N$ polyvacancies. This could be explained by the fact that the structure of B polyvacancies include N-N bonds, which are very strong and, therefore, highly stabilising. Finally, it has also been
found that molecular bridges can form between layers. Whether such bridges have a role to play in the dielectric breakdown of hBN should be the subject of further investigation.
Chapter 8

General Conclusions and Future Work

In this thesis, density functional theory has been used to model defects and defect-generation processes in a number of insulating materials typically used in nanoscale devices. These results provide insight into how defects may impact the function of MIM and MIS devices. The major results will now be summarised.

The oxygen vacancy in MgO has been studied. In particular, the optical absorption spectrum has been calculated using a TDDFT implementation in CP2K. The results show reasonable (within 0.3 eV) agreement with experiment. These theoretical calculations benefit from the use of hybrid DFT and large simulation cells. Further, the excited states of the $V^0_O$ were studied using the $\Delta$-SCF method. Two excited states were found, one of which had a luminescence of 2.0 eV, in good agreement with the experimentally detected 2.3 eV.

In chapter 4 intrinsic carrier trapping in HfO$_2$ was studied. The results for the crystalline phases of HfO$_2$ agree qualitatively with previous work. There was, however, disagreement in the value for trapping energy. This is unsurprising, however, since there is some dependency of the XC functional used. Results in the literature for hole trapping in m-HfO$_2$, for example, vary from 0.18 to 0.57 eV. This chapter also presented calculations predicting the existence of deep trap states in a-HfO$_2$. Both electrons and holes are predicted to trap onto precursor sites intrinsic to the a-HfO$_2$ structure, without the need for defects (such as vacancies or impurities).
Holes were found to have an average trapping energy of 1.4 eV. Finally, the calculated optical ionisation spectra of intrinsic electron traps agree with the results of exhaustive photodepopulation spectroscopy supporting the proposed model that intrinsic electron traps are responsible for charging of a-HfO$_2$ films.

In chapter 5, oxygen vacancy and interstitial defects in HfO$_2$ were investigated. Previous results for vacancies in m-HfO$_2$ were confirmed using larger simulation cells and a hybrid functional. Furthermore, vacancy and interstitial defects were calculated in a-HfO$_2$. A statistical sample of each defect was made. Averages of quantities such as formation energy were taken and found to be similar to the values for m-HfO$_2$. In some cases, however, the distribution of formation energy could be as wide as 1 eV. The optical spectra of the $V_O$ defect was also calculated. The $V_O^{\pm2}$ defect was predicted to have an absorption peak at 5.3 eV, which is in good agreement with the experimentally detected 5.2 eV. Similar to Chapter 3, the $\Delta$-SCF method was used to study the excited states and luminescence energies of the $V_O^{\pm2}$ defect. It is predicted that there should be a 3.6 eV photoemission peak which corresponds to a radiative tunnelling transition between a trapped hole and a $V_O^{\pm1}$ defect.

Chapter 6 presented the results of calculations on oxygen vacancy generation in HfO$_2$. To investigate $V_O$ generation, the mechanism for oxygen Frenkel pair generation was simulated using nudged elastic band simulations. It was found that, in the neutral bulk, an activation energy of 6.6 eV was required. Also, it was found that an applied electric field (of strength 10 MV cm$^{-1}$) was not enough to reduce the barrier to a value achievable at room temperature. Therefore, results of a different Frenkel pair generation mechanism were presented. This mechanism included the injection of two carriers (electrons or holes). This reduced the activation energy to below 2 eV. In a-HfO$_2$, this process happens at a double electron (or hole) trap. These results have a major impact on the understanding of dielectric breakdown in HfO$_2$. They predict that electron (or hole) injection is a necessary part of the oxygen vacancy generation process. This means that there should be behavioural differences between the breakdown of crystalline and amorphous HfO$_2$. 
In the final chapter, intrinsic point defects in hBN are investigated. Each defect was calculated in different charge states, and the spin configuration of the ground state was calculated. Using formation energy calculations, the charge transition levels of each defect were calculated. These were used to make a prediction about the possible charging of hBN insulators in MIM and MIS structures. It is highly likely that common defects such as $V_B$ will be positively charged when typical electrodes are used. Furthermore, many CTLs lie close to these Fermi level positions, which implies that they should be responsible for conduction via trap assisted tunnelling when the stack is placed under bias. Finally, it is demonstrated that B polyvacancies are highly binding due to the formation of N-N bonds.

There are a number of ways the work in this thesis could be taken further. Firstly, in two Chapters the $\Delta$-SCF method is used to study the configuration of excited states and their photoemission spectra. This method, however, is slightly limited. Firstly, it only calculated the singlet to triplet excitations. Although, in principle, the excited triplet and excited singlet state can be similar in configuration, this is not always true. It is also the case that the exchange interaction may alter energies somewhat. Additionally, one of the bigger shortcomings is that the excited state wavefunction will not be orthogonalised with the ground state. These problems could be overcome by calculating the forces in the excited state wavefunction with TDDFT. This capability, however, is not implemented in CP2K.

There is also the question of what precise predictions do these calculations make with regard to device functionality. There have been previous studies in which defect parameters, calculated using atomistic modelling, are used to calculated IV characteristics of MIM structures using multiphonon trap assisted tunnelling calculations. Furthermore, simulations of MIM structures which include defect generation processes also allow for a prediction of time to breakdown statistics [144]. A natural next step for this work would therefore to be use the parameters calculated for defects in HfO$_2$ and hBN to model the degradation of dielectric thin films and make such predictions. This would allow for a rigorous test of the model (for dielectric breakdown) proposed in this thesis.
Bibliography


[121] John P Perdew, John A Chevary, Sy H Vosko, Koblar A Jackson, Mark R Pederson, Dig J Singh, and Carlos Fiolhais. Atoms, molecules, solids, and sur-


[145] Laegu Kang, Byoung Hun Lee, Wen-Jie Qi, Yongjoo Jeon, Renee Nieh, Sundar Gopalan, Katsunori Onishi, and Jack C Lee. Electrical characteristics of


