Computational Modelling of Oxygen Defects and Interfaces in Monoclinic HfO_2

Samuel Rowan Bradley

Department of Chemistry

University College London

2015

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Engineering (Eng.D.) of University College London

Supervisor: Prof. Alexander L. Shluger

I, Samuel Rowan Bradley, 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 thesis.

Samuel R. Bradley

September 2015

Abstract

Hafnia (HfO_2) is of particular interest for microelectronics applications. It is currently used as a high-k insulating dielectric in the latest generations of transistors and as the insulator in new metal-insulator-metal resistive memory devices (ReRAM). Oxygen defects within HfO_2 are thought to have a significant impact on the performance of transistors and are seen to be a leading factor in dielectric breakdown. This same breakdown phenomenon is exploited in HfO_2 ReRAM devices to form reversibly switchable conducting filaments.

In this work computational modelling is used to look at the properties of oxygen defects in monoclinic HfO_2, to explain their mechanisms of formation and diffusion, both in the bulk material and at interfaces.

The properties of oxygen vacancy aggregates ranging from 2 to 4 vacancies are analysed. It is shown that the most stable aggregates form the largest void within the bulk lattice and the larger aggregates have a higher binding energy per vacancy. Negatively charged aggregates are found to be stable whereas positively charged aggregates are not. The properties of Frenkel pairs is also calculated leading to the development of a model for the formation of conductive filaments in HfO_2; by generation of oxygen vacancy aggregates, via Frenkel pair formation during electron flooding conditions, and oxygen interstitial out-diffusion.

A Pt/HfO_2 interface model is created and its structural and electronic properties are analysed. The stability of oxygen defects at the interface is calculated. It is shown that at low oxygen concentrations, vacancies may be more stable at the interface however at high oxygen concentrations, oxygen interstitial ions are unlikely to penetrate the Pt electrode and so, remain within HfO_2. The stability of oxygen defects across a Si/SiO_2/HfO_2/TiN interface are calculated and a model of oxygen scavenging process, from SiO_2 and into defective TiN, is suggested.

Contents

1 Introduction **14**

1.1 Motivation 14

1.2 Memory Scaling 16

1.2.1 Memory Devices 17

1.3 Transistors 18

1.3.1 High-k Dielectrics 21

1.4 Resistive RAM 22

1.4.1 Memristive Systems 23

1.4.2 Memory Applications 24

1.4.3 Switching Mechanisms 25

1.4.4 MIM ReRAM 28

1.5 Summary 30

2 Theoretical Background and Methods **31**

2.1 Introduction 31

2.2 Wave Function Methods 33

2.3 Density Functional Theory 34

2.3.1 Approximations in DFT 36

2.3.2 Pseudopotentials 38

2.3.3 Basis Sets 38

2.4 Conjugate Gradient Method 42

2.5 Nudged Elastic Band Method 43

2.6 Defect Energetics 44

2.7 Summary 46

3 HfO_2 **48**

3.1 Introduction . . . 48

3.1.1 Atomic Layer Deposition of HfO_2 . . . 49

3.1.2 Bulk Properties of HfO_2 . . . 50

3.2 $HfCl_4$ . . . 51

3.2.1 Pople Basis Sets . . . 52

3.2.2 Dunning Basis Sets . . . 52

3.2.3 Structural Properties of $HfCl_4$ . . . 53

3.2.4 Vibrational Properties of MX_4 . . . 55

3.2.5 Electronic Properties of $HfCl_4$ . . . 57

3.2.6 $HfCl_4$ Cl Dissociation . . . 58

3.2.7 $HfCl_4$ Summary . . . 59

3.3 Bulk HfO_2 . . . 59

3.3.1 HfO_2 Structure . . . 59

3.3.2 Dielectric Properties of HfO_2 . . . 61

3.3.3 Polarons in HfO_2 . . . 63

3.4 Bulk Hafnia Summary . . . 64

4 Oxygen Interstitials in HfO_2 **66**

4.1 Methods of Calculation . . . 67

4.2 Reference O_2 Molecule . . . 68

4.3 Oxygen Interstitial Diffusion . . . 69

4.4 Summary . . . 72

5 Vacancies in HfO_2 **73**

5.1 Introduction . . . 73

5.2 Single Vacancies in HfO_2 . . . 74

5.3 Oxygen Vacancy Aggregates in HfO_2 . . . 77

5.3.1 Structural Properties of Oxygen Vacancy Aggregates . . . 78

5.3.2 Electronic Properties of Oxygen Vacancy Aggregates . . . 84

5.4 Summary . . . 87

6 Frenkel Defect Formation in HfO_2 90

6.1 Introduction . . . 90

6.2 Formation of Nearest Neighbour FD Pairs . . . 91

6.3 Formation of FD Pairs Next to a Pre-existing Vacancy . . . 94

6.4 Summary . . . 98

7 The Pt/HfO_2 Interface 99

7.1 Introduction . . . 99

7.2 Interface Building . . . 101

7.3 Platinum . . . 103

7.4 Pt(111) Surface . . . 104

7.5 HfO_2(001) Surface . . . 106

7.6 Pt/HfO_2 Interface . . . 109

7.7 Summary . . . 115

8 Oxygen Scavenging in the Si/SiO_2/HfO_2/TiN Stack 117

8.1 Introduction . . . 117

8.2 The TiN/HfO_2 Interface . . . 119

8.3 The Si/SiO_2/HfO_2 Interface . . . 121

8.4 O Scavenging Mechanism . . . 124

8.5 Summary . . . 126

9 Conclusions and Further Work 127

List of Figures

1.1 Schematic structure of a p-MOSFET 19

1.2 Band diagrams showing tunnelling mechanisms through a metal-insulator-metal system. The dashed line represents the Fermi energy, red arrows show the motions of electrons. a) Direct tunnelling in a low electric field, b) Fowler-Nordheim tunnelling in a high electric field, c) trap-assisted tunnelling via neutral charge traps in the insulator band gap 21

1.3 Band gap v. dielectric constant of various insulators.[1] 22

1.4 An $I - V$ schematic (Lissajous figure) of the pinched hysteresis loop typical of memristive systems for a periodic input.[2] 24

1.5 A schematic of a typical crossbar array of memristive cells. Each cell is connected via an array of vertical and horizontal metallic contacts separated by an insulating material. A represents an ammeter.[2] . . . 25

1.6 Typical $I - V$ curves for bipolar and unipolar switching devices. cc represents a compliance current.[2] 26

1.7 The $I-V$ characteristics of a Cu/SiO_2/Pt electrochemical metallisation cell. The insets show the dynamics of metallic filament formation.[3] . 27

2.1 An illustration of the all-electron (blue) and pseudoelectron (red) potentials and their corresponding wave functions. The point at which the potentials match is shown as r_c.[4] 39

2.2 Forces $F_i^{S\parallel}$ and $F_i^{\perp}$ act on F_i^{NEB} to minimise images towards the MEP on the potential energy surface.[5] 45

3.1 Monoclinic HfO_2 (Hf: green, O: red), [001] view. 50

3.2 The T_d structure of $HfCl_4$ (Hf: dark green, Cl: blue-green). 53

3.3 Total Hartree-Fock energies for the $HfCl_4$ molecule calculated with various basis sets on the Cl atoms. 54

3.4 Schematic of the four Raman active vibrational modes of T_d MCl_4: (ν_1) Twist, (ν_2) scissor, (ν_3) symmetric stretch, (ν_4), asymmetric stretch. . 55

3.5 Qualitative molecular orbital diagram for $HfCl_4$ 57

3.6 MCl_4 (M = Si, Hf) molecular orbitals (MOs): (a) Polarized bonding MO, (b) HOMO showing the Cl 2p lone pairs (c) antibonding LUMO. 58

3.7 The total and projected densities of states of m-HfO_2 calculated with HSE06. 61

4.1 Oxygen interstitial diffusion pathways from the starting configuration (left) to the final configuration (right) via the barrier point (centre). a) O^0_{int} within the 3C sublattice ([100] view) b) O^{2-}_{int} within the 3C sublattice ([100] view) c) O^0_{int} between the 4C and 3C sublattice ([001] view). (green balls: Hf, red balls: 3-coordinated O, blue balls: 4-coordinated O, purple ball: oxygen interstitial). 71

5.1 (a) Conductive-AFM topography, and leakage current profile at +6V after the (b) 1st scan, (c) 2nd scan, (d) 3rd scan of a HfO_2/SiO_x sample.[6] 74

5.2 The Kohn-Sham energy levels for neutral and charged 3C (blue) and 4C (black) oxygen vacancies. a) PW91, b) HSE. 75

5.3 The lowest energy neutral m-HfO_2 di-vacancy structures. (a) $D(C)^0_{44}$, (b) $D(F)^0_{33}$, (c) $D(G)^0_{34}$. Left: projection view along the [001] direction, right: projection view along the [010] direction. (Green balls: Hf, red balls: 3C oxygen, blue balls: 4C oxygen, red squares: 3C oxygen vacancy site, blue squares: 4C oxygen vacancy site). 81

5.4 The binding energies of the calculated di-vacancies with respect to the vacancy – vacancy separation. Squares: D_{33}, circles: D_{44}, triangles: D_{34}. 82

5.5 The lowest energy neutral m-HfO$_2$ tri-vacancy structure, $T(D)^0_{443}$. Left: projection view along the [001] direction, right: projection view along the [010] direction (coloured as in Fig. 5.3) 82

5.6 The lowest energy neutral m-HfO$_2$ tetra-vacancy structure, $Q(D)^0_{4443}$. Left: projection view along the [001] direction, right: projection view along the [001] direction (coloured as in Fig. 5.3). 83

5.7 Total electronic densities of states (black) and Fermi energy (blue) of the HfO$_2$ di-vacancy, $D(I)^0_{34}$. 85

5.8 The one electron densities (dark blue) for the HfO$_2$ di-vacancy $D(I)^0_{33}$ (a) lower energy di-vacancy state (b) higher energy di-vacancy state. . 85

6.1 The band-decomposed partial charge density associated with the highest occupied molecular orbital in NN FD $[2V_3]^{2-} + O^{2-}_{int}$, calculated using HSE. (green balls: Hf, light red balls: O, purple ball: O$^{2-}_{int}$, blue: highest occupied molecular orbital). 94

6.2 Barriers of formation for the HfO$_2$ NN FD pair, di-vacancy NN FD pair and charged di-vacancy NN FD pairs. 96

6.3 (a) - (b) Formation of a single NN FD pair. (c) - (d) Formation of di-vacancy NN FD pair in the presence of a pre-existing vacancy (large green balls: Hf, small red balls: O, blue squares: vacancy, highlighted red ball: O$^{2-}_{int}$). 96

6.4 Schematic of the Frenkel pair formation mechanism where blue squares represent vacancies and red circles represent O^{2-}. (a) An oxygen atom leaves its lattice site forming an oxygen Frenkel pair . (b) The O$^{2-}_{int}$ ion diffuses away leaving behind a V^0. (c) The vacancy's NN then leaves its lattice site, forming a divacancy Frenkel pair. (d) The second O$^{2-}_{int}$ ion diffuses away. (e) This leaves behind $2V^0$ which can trap electrons and form another NN FD. 97

7.1 The close packed FCC structure of platinum. 103

7.2 Total electronic density of states of bulk platinum. 105
7.3 (111) $2\times\sqrt{3}$ surface of FCC platinum. 105
7.4 Projected density of states of the Pt(111) surface onto (a) the surface layers, (b) the fourth deepest layers. 107
7.5 The (001) surface of monoclinic HfO_2. (a) unrelaxed, (b) relaxed. . . 107
7.6 The band structure of the highest occupied and lowest occupied bands of bulk monoclinic HfO_2 (black) and the monoclinic (001) surface (red) projected along the $B-\Gamma-Y$ direction of the Brillouin zone. 108
7.7 (a) The Pt(111) $2\times\sqrt{3}$ surface. (b) The HfO_2(001) 1×1 surface (blue: Pt, green: Hf, red: O). 108
7.8 The Structure of the Pt(111) $2\times\sqrt{3}$/HfO_2(001) 1×1 interface model (blue: Pt, green: Hf, red: O). 110
7.9 Total electronic densities of states (black) and the states projected onto the Pt atoms (blue), Hf atoms (green) and O atoms (red) adjusted so the Fermi level, $E_F = 0$. The CBM is taken as the energy of the lowest Hf state above E_F and the VBO taken as the energy of the highest O state below E_F. 111
7.10 Layer projected CBM and VBM of the Pt(111) $2\times\sqrt{3}$/HfO_2(001) 1×1. 111
7.11 Interface internal energy energy versus O chemical potential for interface terminations of varying O concentration. Labels correspond to the number of interfacial O atoms between the top layer of the metal and bottom layer of Hf ions. $\mu_O = 0$ is defined as the energy of a free isolated O_2 molecule. Vertical dashed lines represent the limits above and below which α-PtO and metallic Hf would be favourable. 112
7.12 Band offsets of the stable interface termination structures. Left to right: O deficient, stoichiometric, O rich. 113
7.13 Incorporation from HfO_2 O lattice sites, a, c, d, e, and O interstitial site, b, into Pt interstitial O site, a. (Light blue: Pt, silver: Hf, red: O, green: defect sites). 114

8.1 (a) Illustration of the structure of the Si/SiO_2/HfO_2/TiN stack. HRTEM image of three gate stacks with (b) HRTEM of the stack with ~5 nm TiN layer and (c) ~2 nm TiN on top of imec-clean/1.8 nm HfO_2.[7] . . 118

8.2 Interface internal energy versus O chemical potential for interface terminations of varying O concentration. Labels correspond to the number of interfacial O atoms between the top layer of the metal and bottom layer of Hf ions. $\mu_O = 0$ is defined as the energy of a free isolated O_2 molecule. Vertical dashed lines represent the limits above and below which TiO_2 and metallic Hf would be favourable 120

8.3 Incorporation from HfO_2 O lattice sites, a, c, d, e, and O interstitial site, b, into TiN interstitial O site, a. (Dark blue: Ti, pink: N, silver: Hf, red: O, green: defect sites.) 121

8.4 Formation energies of 43 neutral oxygen vacancies at various positions between the Si and HfO_2 parts of the stack. The right-most defects are directly at the HfO_2/SiO_2 interface and the left-most directly at the SiO_2/Si interface.[8] 122

8.5 Formation energies of 28 oxygen interstitials within the HfO_2 layer at the SiO_2/HfO_2 interface. 124

8.6 Summary of the band offsets between the components of the Si/SiO_2/HfO_2/TiN stack. Mid-gap red lines show the neutral oxygen vacancy levels in the oxides. 125

List of Tables

3.1 Functions used in the Dunning basis sets. 53

3.2 Bond lengths and Raman active vibrational frequencies of $HfCl_4$ using different basis sets to describe the Cl atoms 54

3.3 Bond lengths, 1st ionisation energies and Raman active vibrational frequencies of $HfCl_4$ using different basis sets to describe the valence electrons of Hf atoms and a 6-311G* basis on the Cl atoms. 55

3.4 Bond lengths and Raman vibrational frequencies of tetrahedral MCl_4 molecules. 56

3.5 Optimised structural parameters and band gaps of bulk monoclinic HfO_2. 60

3.6 Dielectric tensors of m-HfO_2. $k = (k_{xx} + k_{yy} + k_{zz})/3$. 62

4.1 Total energy of the O_2 molecule in eV, calculated in a large, triclinic unit cell. 68

4.2 Formation energies in eV and nearest neighbour O–O distances of interstitial O ions in m-HfO_2 calculated using the GGA and hybrid density functionals, in a 96 atom periodic cell. The energies are referenced to an ideal lattice cell of the same size and charge. Energies are in eV and O–O distances are in Å. 69

5.1 Formation energies, in eV, of O vacancies in m-HfO_2. 75

5.2 Binding energies of the calculated neutral vacancy aggregates. D, indicates a di-vacancy where each configuration is labelled by a letter in parentheses. Subscripts denote the Hf coordination of the constituent vacancies and superscripts indicate the overall charge of the unit cell. 79

5.3 Binding energies of the calculated neutral vacancy aggregates. T and Q indicate tri- and tetra-vacancies respectively. Each configuration is labelled by a letter in parentheses. Subscripts denote the Hf coordination of the constituent vacancies and superscripts indicate the overall charge of the unit cell. 80

5.4 Formation energies of oxygen vacancies and di-vacancies (eV) calculated using equation 5.1 and the HSE06 functional. D denotes a di-vacancy, subscripts denote the coordination of the constituent vacancies. 86

6.1 Formation energies and adiabatic barriers for formation and recombination of FD pairs (eV). $E_{for,D}$ corresponds to NN FD pair formation energy, $E_{bar,f}$ corresponds to the barrier energy for FD pair formation, and $E_{bar,r}$ the barrier energy for recombination. 92

7.1 Lattice parameter (a_0), bulk modulus (B_0), and energy (E_0) of a bulk FCC platinum crystal. 104

7.2 Energetic cost, in eV, to transfer O from various sites in HfO_2 into the Pt electrode (Fig. 7.13). 115

8.1 Energetic cost in eV to transfer O from various sites in HfO_2 into the TiN electrode (Fig. 8.3). 120

List of Publications

The following publications are derived from work presented in this thesis:

1. S. R. Bradley, K. P. McKenna and A. L. Shluger, "The behaviour of oxygen at metal electrodes in HfO_2 based resistive switching devices", *Microelectron. Eng.* **109**, 346–350 (2013).

2. S. R. Bradley, G. Bersuker and A. L. Shluger, "Modelling of oxygen vacancy aggregates in monoclinic HfO_2: can they contribute to conductive filament formation?", *J. Phys. Cond. Mat.* **27**, 41, 415401,(2015).

3. S. R. Bradley, G. Bersuker and A. L. Shluger, "Electron injection assisted generation of oxygen vacancies in monoclinic HfO_2", *Phys. Rev. Applied* **4**, 064008 (2015).

Acknowledgements

The completion of this thesis would not have been possible without the incredible support of many people.

First and foremost I would like to thank my supervisor Prof. Alex Shluger for his tireless dedication, mentorship, and the many hours of interesting and valuable discussions. I would also like to thank Dr Keith McKenna for his supervision in my first two years and help getting this work started. I am extremely grateful to the Shluger group for huge amounts of technical support and advice. In particular I would like to thank Sanliang Ling and Matthew Watkins for assistance with interface building and analysis, Samuel Murphy and Thomas Durrant for help calculating charge corrections, and Al-moatasem El-Sayed for an uncountable number of things.

I would like to thank EPSRC, the M3S doctoral training centre, Sematech and Gennadi Bersuker, for sponsorship and the guidance they gave to this work. The calculations performed here made use the UK's national high-performance computing services, HECToR and Archer, and were accessed via our membership of the UK's HPC Materials Chemistry Consortium, which is funded by EPSRC (EP/F067496). This work also gratefully acknowledges access to the Chinook supercomputer at EMSL, located at Pacific Northwest National Laboratory.

I would also like to thank my office mates of Wilkins 232 for their patience and of Physics A9 for taking me in when I had nowhere to go.

Infinite thanks go to my ever loving family and wonderfully supportive friends without whom this work would've been truly impossible. Finally, I must thank my darling girlfriend, Sally, for her boundless kindness and understanding.

1

Introduction

1.1 Motivation

In this thesis computational modelling is used in an attempt to develop mechanisms related to oxygen defects within hafnia (HfO_2). These mechanisms involve defect creation and annihilation, and the motions of the defects across the oxide and at interfaces.

When I started my thesis in 2010, research on new memory devices known as resistive RAM (ReRAM) was developing quickly. These new devices were starting to be seriously considered as a viable replacement for the current state of the art consumer memory, flash. Although good switching characteristics had been demonstrated by experimental ReRAM cells, very little was known about the mechanisms behind how these devices operate. The research consortium, Sematech, who part-funded this work, were interested in using their expertise in transistor reliability problems and applying it to this new problem.

It was thought that defects and interfaces, features implicated in the dielectric breakdown of transistors, were likely to play a crucial role, since the materials used as dielectrics and gate contacts in transistors were the same as those used in ReRAM

cells. Initially the structure of the interfaces was thought to be important, and so some steps were made to consider how these interfaces form during deposition processes, and the character of the molecular precursors and surfaces. However, further research showed that the conductive region through HfO_2 was likely to be made up of an oxygen deficient filament, which can be reversibly broken and reformed within the oxide. The question of how these filaments form within HfO_2, and how to model those processes on an atomic scale, has become the primary focus of this work.

Firstly, in chapter 3 various exchange-correlation density functionals and basis sets are evaluated. The atomic layer deposition (ALD) precursor molecule, $HfCl_4$, and similar tetrachloride molecules, are calculated using several basis sets and pseudopotentials. The properties of bulk m-HfO_2 are then calculated using various exchange-correlation functionals using plane-wave approaches.

Secondly, a model of conductive filament (CF) formation and charge trap generation within bulk HfO_2 is developed. In order to do this, the properties of various defects are calculated. The formation energies and diffusion barriers for the O interstitial are calculated in chapter 4. The motion of these interstitial ions then forms the leading factor for out-diffusion of O ions in the CF formation model. The stabilities and electronic properties of oxygen vacancies and vacancy aggregates are calculated in chapter 5. These vacancy aggregates form the conductive region of the conductive filament. Finally, the feasibility of the generation of Frenkel defects is examined in chapter 6. The formation of these Frenkel defects provides a low energy pathway for the generation of oxygen vacancies under electron flooding conditions.

The structure and properties of metal/HfO_2 interfaces is then considered. In chapter 7 a model of the Pt/HfO_2 interface is developed. This is done by first calculating the properties of the relevant Pt and HfO_2 surfaces, then calculating the lowest energy structure for the interface and the properties of defects at that interface. Finally, in chapter 8 the energetics of various oxygen defect processes throughout a Si/SiO_2/HfO_2/TiN stack are summarised. The formation energy of oxygen interstitials at the HfO_2/SiO_2 interface are calculated and a model for the oxygen scavenging

process in a Si/SiO_2/HfO_2/TiN stack is proposed.

Background

1.2 Memory Scaling

The current boom in smartphones and tablet computers is creating ever more demand for smaller low power microelectronics. Particularly important are the latest generations of processors and memory storage devices. The requirements of these new technologies are extremely difficult; they must be fast and efficient, yet more powerful and store more data, whilst being physically small, cheap and reliable.

A popular idea driving this trend is "Moore's Law". This was an observation made by Gordon E. Moore, a co-founder of *Intel*, that the number of transistors on an integrated circuit doubled every one to two years.[9,10] As this prediction became widely accepted in the semiconductor industry it has become a target which guides research, development and policy.[11] This has led an exponential advancement, not only in processors, but also other aspects of microelectronics, particularly memory storage. However, as these devices advance to keep up with Moore's law increasingly difficult barriers need to be overcome.

In modern semiconductor device fabrication, 2 year targets are made to keep up with Moore's law. These targets are referred to as nodes, and named for the half distance between two identical features in an integrated circuit, which is roughly equivalent to the size of one memory cell. In 2014 the 14 nm node target was met, and the the 10 nm node is aimed to be reached in 2016–2017.

Devices which are so small need to be designed and fabricated on the nano scale. At this size, atomic scale processes become extremely important and the motions of individual atoms and electrons within the materials must be well understood. Small effects can have a large impact on both device performance and reliability. This is very difficult to do, and needs a full scale understanding, from engineering and

measurement of macroscale effects, down to theoretical modelling of atomic scale processes. New devices can then be built with a full understanding and, hopefully, a good degree of control to allow this device scaling to continue into the future.

In order to continue scaling, broadly speaking two approaches can be taken. The first is to improve upon the current technologies by using better materials and manufacturing processes. The second is to design new technologies which perform the same function as the old e.g. storing memory, but at a higher density on the chip.

In this chapter these two cases are examined, one of which is the ongoing scaling of transistors using high-k materials, the other is the development of new resistive switches, which can be used to store memory at a higher density.

1.2.1 Memory Devices

A binary system is simply a system which can be in two distinguishable states, usually labelled as 1 and 0, and with each unit referred to as a bit. Binary is easily implemented into digital circuits, and so, is the primary way in which data is processed, transmitted, and stored in computers.

Early mechanical data processing machines such as Babbage's "Analytical Engine" or Jacquard's loom used the position of a gear or lever or the presence or absence of a hole in a punch-card or piece of tape to store a bit of data. The first electrical devices with discrete logic, such as elevators or telephone switches, would register bits as the state of electrical relays that were either open or closed. Modern computers work largely the same way; reading and writing binary data and then processing the binary code into a useful process. However, they can use a variety of ways to store digital memory. These largely fall into two types; volatile and non-volatile memory.

Volatile memory requires power to maintain stored information, if the power is removed then the information is lost. In general this memory is much faster than non-volatile memory. It is used for random access memory (RAM) drives and is referred to as "dynamic-RAM" or DRAM. This increase in speed is generally attributed to RAM chips being made of solid-state components that are much more easily accessed

than their non-volatile counterparts. The increase in speed is at the cost of capacity, with modern consumer units 500 times smaller than non-volatile memory of similar cost.

The original dynamic RAM used in the 1940s cryptanalytic computer "*Aquarius*" used an array of capacitors which were either charged or discharged. Over time the charge would drain from the capacitors so a pulse was periodically applied to maintain the information. The principle of dynamic RAM has since remained relatively similar with the memory being made of an array of cells consisting of a transistor attached to a capacitor. The transistor acts as a basic switch and is either on or off, keeping the capacitor charged or discharged.

Non-volatile memory retains data even when the power is removed from the device. Classic examples include punch-cards, magnetic tape and optical storage on CDs and HDDs. The most common current example is flash memory, where electrical charge is stored in capacitors controlled by floating gate transistors. Flash memory uses solid state memory cells, so can be accessed very quickly, however, because the floating gate is electrically insulated by a layer of oxide on all sides, electrons stored in the capacitor are trapped until an electric field is applied to change the bit.

Since the basic principle behind memory storage is very simple, there are a huge number of ways it can be done, each with advantages and disadvantages. This simplicity allows for a lot of innovation and development in memory storage devices.

1.3 Transistors

The ongoing scaling of microelectronics has been largely driven by developments in metal-oxide-semiconductor field-effect transistors (MOSFET). MOSFETs are electronic devices used in switching and amplification, and form the building blocks of digital complementary MOS (CMOS) logic which is used to operate microprocessors and RAM. Ideally they act as a 3 terminal switch, whereby the source (S) and drain (D) terminals are either connected or isolated depending on the voltage applied at the

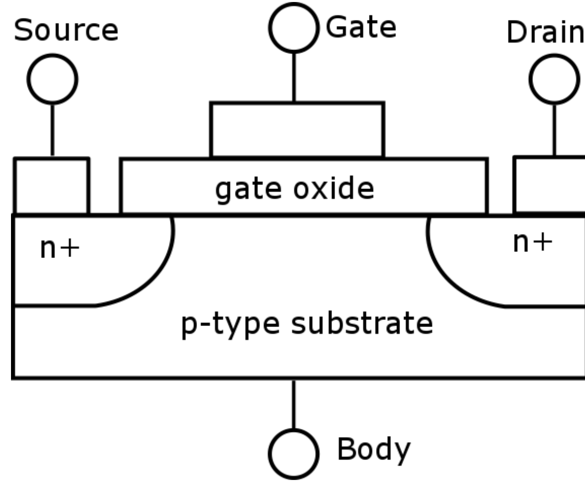


Figure 1.1: Schematic structure of a p-MOSFET

gate (G) terminal (Fig. 1.1). Properties of the insulating gate oxide layer, which separates the metallic gate electrode and the semiconducting channel region, are critical in the defining transistor characteristics. The charge induced in the channel region, Q, is a product of the capacitance of the gate oxide, $C_o x$, and the potential across the the gate capacitor, V:

$$Q = C_{ox} V. \tag{1.1}$$

The gate oxide can be considered as a parallel plate capacitor whose capacitance per area is proportional to its physical thickness, t_{ox}, and the permittivity of free space, ϵ_0, and inversely proportional to its dielectric constant, k_{ox}:

$$C_{ox} = \frac{k_{ox}\epsilon_0}{t_{ox}}. \tag{1.2}$$

The drain-source current operating in the saturation region of the channel, I_{ds}, is expressed as

$$I_{ds} \approx \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_{gs} - V_t)^2, \tag{1.3}$$

where μ is the channel mobility, W and L are the channel region width and length respectively, V_{gs} is the gate-source potential and V_t is the threshold voltage.[12]

Traditionally the insulating layer has been made from silica (SiO_2), which can

be easily grown on top of the semiconducting silicon wafers, with excellent control and quality. Silica also has a very wide insulating band gap of 9 eV, forming a huge potential barrier to current leakage between the source and gate electrodes. Continued scaling has meant that the physical dimensions of transistors has decreased while the channel current has had to remain static. In order to achieve this, the capacitance of the gate oxide must increase through either reduction of the physical thickness of the oxide layer or by using an insulator with a higher dielectric constant.

Over the past 40 years, improved manufacturing techniques have helped reduce the thickness of the silica layer. However, there is a practical limit at which the oxide no longer acts as an insulator.

In 1988, Nagai *et al.* found they could reduce the thickness of the gate oxide and, provided the leakage current was significantly less than the channel current, the devices were still functional. This was achieved by reducing the gate length.[13] In spite of this, a SiO_2 gate oxide $\sim$ 1 nm thick is considered the practical scaling limit. Muller *et al.* used electron-energy-loss spectroscopy (EELS) to measure the electronic structure of thin SiO_2 gate oxides.[14] They measured the Si/SiO_2 interfacial states that were a result of the spillover of the Si conduction band. The spacial extent of these states places a fundamental limit of $\sim$ 0.7 nm SiO_2 thickness.

Furthermore, gate leakage currents, measured in low applied electric fields, have been found to increase after a device is stressed with a high field. These stress induced leakage currents (SILC) are related to the generation of neutral electron traps caused by hot-electron transport. The neutral traps which form, act as stepping stones for charge carriers through the oxide, in a process known as trap-assisted tunnelling (TAT)(Fig. 1.2 c). In TAT a percolation path can develop when a sufficient density of charge traps form across the oxide layer, leading to breakdown of the insulator. In a high field, Fowler-Nordheim (F-N) tunnelling takes place. F-N tunnelling electrons first tunnel into the oxide conduction band before entering the anode contact (Fig. 1.2 b). F-N stress can lead to SILC when the tunnelling electrons hit the gate anode, transferring kinetic energy to a hole which tunnels back into the oxide. These holes

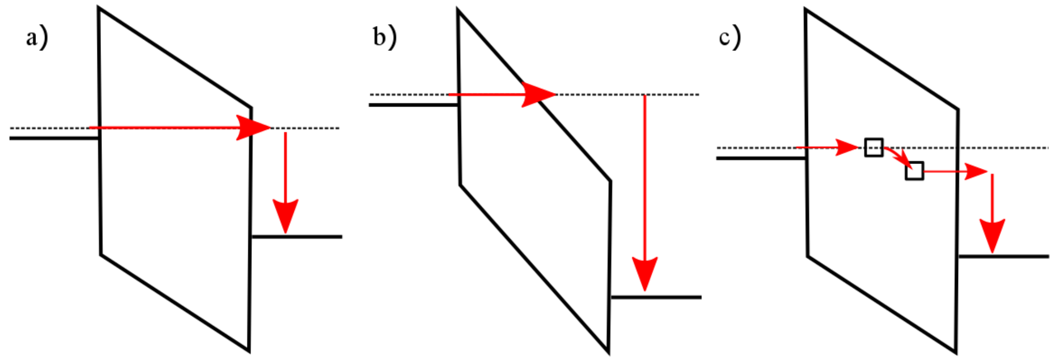


Figure 1.2: Band diagrams showing tunnelling mechanisms through a metal-insulator-metal system. The dashed line represents the Fermi energy, red arrows show the motions of electrons. a) Direct tunnelling in a low electric field, b) Fowler-Nordheim tunnelling in a high electric field, c) trap-assisted tunnelling via neutral charge traps in the insulator band gap

can then go on to form charge traps. Charge traps increase the amount of current flowing into the gate oxide which can lead to thermal damage and to more charge traps. These charge traps can then build up over time to form conductive filaments across the insulator, causing the transistor to breakdown.[15,16]

1.3.1 High-*k* Dielectrics

Over the past decade alternative "high-k" dielectrics have been developed. These materials have a higher dielectric constant than SiO_2 (i.e. > 3.9) and so, according to eq. 1.2, can give the same capacitance whilst being physically thicker, thus raising the tunnelling barrier.

However, these materials must meet a range of specifications in order to be viable as gate dielectrics. First of all they must be wide band gap insulators. Fig. 1.3 shows that this is a problem because the dielectric constant tends to vary inversely with the band gap in many materials. Secondly, the dielectric must form a high quality, stable interface with Si to be incorporated into current transistor fabrication. They must also have a band offset relative to the conduction band of Si of > 1 eV.[1]

HfO_2 is a good match for all these criteria and is now used on an industrial scale. HfO_2 is an insulating material with a band gap of around 6 eV, a valence band offset relative to Si of 3.3 eV, and conduction band offset of 1.4 eV, giving a significant

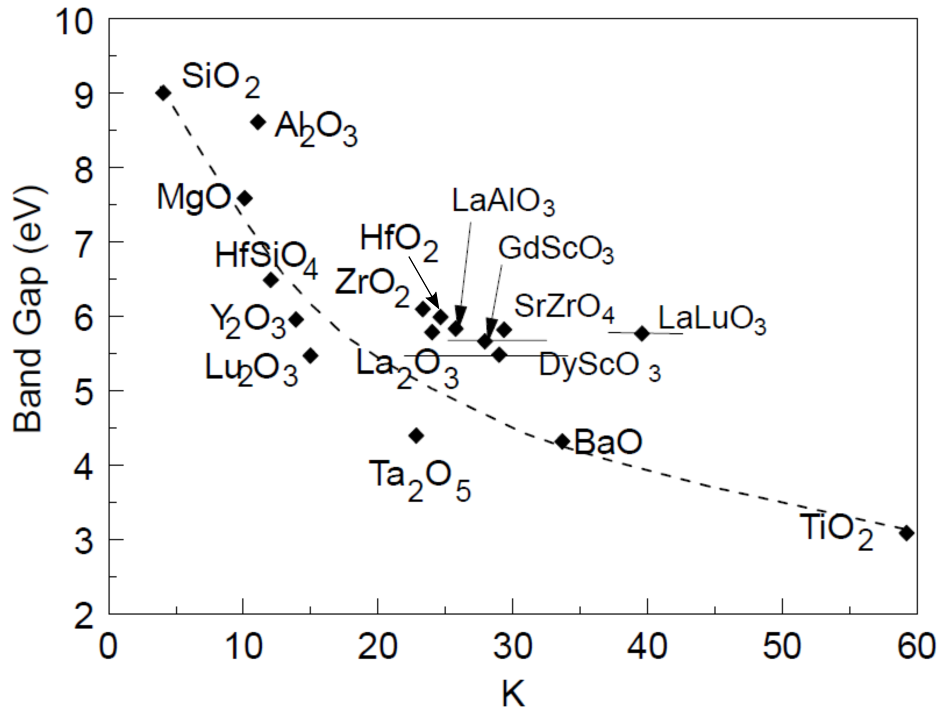


Figure 1.3: Band gap v. dielectric constant of various insulators.[1]

potential barrier to a leakage current.[17] It is also thermally stable and, unlike zirconia, forms a stable interface with Si.[18] Importantly hafnia has a high dielectric constant of 25.[19]

HfO_2 has allowed the post-SiO_2 scaling of transistors to become industrially viable, however, continued scaling is still hindered by breakdown occurring within thin films. SILC remains an issue within these devices and TAT is thought to be the main cause of breakdown, but the mechanism behind the production of charge traps within HfO_2 is still unclear.

These reliability issues mean that there is still a fundamental limit to the scaling of even high-k transistors and scaling in line with Moore's law using these devices cannot go on forever.

1.4 Resistive RAM

Resistive RAM has been proposed as a new technology to replace transistors in memory storage devices. They store bits by exploiting an effect known as "memristance", and by switching memory cells between two discrete, high resistance and low resis-

tance states. Because each memory cell is a two terminal switch, the density of the cells on a chip can be much higher than with three terminal transistors. This means the ongoing scaling of memory storage devices can continue into the foreseeable future using this type of device.

1.4.1 Memristive Systems

The memristor was first proposed by Leon Chua in 1971.[20] He reasoned that were three passive fundamental circuit elements (resistor, capacitor and inductor) connecting pairs of the 4 fundamental circuit variables (current, I, charge, q, voltage, V, and flux linkage, φ). Furthermore, the variables are interrelated since the charge is the time integral of the current ($dq = Idt$) and the flux is the time integral of the voltage ($d\varphi = Vdt$). From these relations he reasoned that there should be a fourth circuit element, which was termed the "memristor" (a contraction of *memory resistor*) with a property called "memristance", M, to provide a functional relationship between charge and flux:

$$d\phi = Mdq. \tag{1.4}$$

The memristor is defined as a two-terminal element in which the flux between the terminals is a function of the electric charge passed through the device. Since flux is simply the time integral of voltage, and charge the time integral of current, memristance can be written,

$$M(q(t)) = \frac{V(t)}{I(t)}, \tag{1.5}$$

which is analogous to Ohm's Law if M is not charge-dependent:

$$R(t) = \frac{V(t)}{I(t)}. \tag{1.6}$$

Therefore, at any given time a memristor acts as an ordinary resistor.

However, the value of M at any time depends on the time integral of the memristor

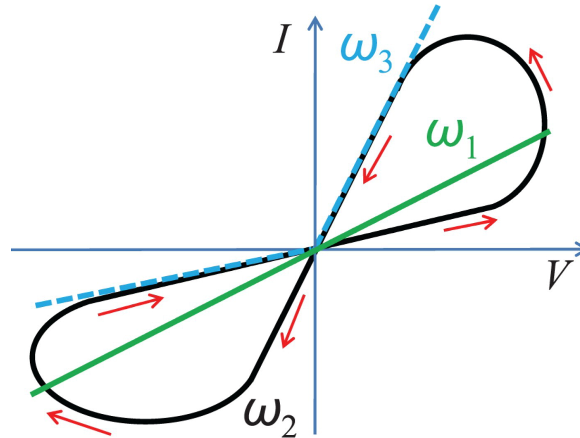


Figure 1.4: An $I-V$ schematic (Lissajous figure) of the pinched hysteresis loop typical of memristive systems for a periodic input.[2]

current, from $t = -\infty$ to $t = t_0$. Therefore, the value of the resistance is dependent on the complete history of the current which has passed through that particular memristor. The $I - V$ characteristic of such a non-linear relation between flux and charge for a sinusoidal input, results in a frequency-dependent Lissajous figure which is unique to a non-linear memristor, as shown in Fig. 1.4. The pinched-hysteresis loop shape of the figure, and the dependence of the lobe area with respect to the frequency of the forcing signal, is used to define a memristor experimentally.[2] This definition can, in reality, be applied to a broad range of systems termed "memristive systems".

1.4.2 Memory Applications

Memristive properties have been demonstrated in a wide array of materials including binary oxides (TiO_2,[21] SiO_2,[22] NiO, CoO,[23] HfO_2[24]), perovskite type oxides,[25] sulphides,[26] semiconductors[27] and organics.[28] The physical mechanisms of the switching process are not well understood in all cases, however, there are many similarities in terms of fabrication and switching operation between the systems. Experimentally, a memristive system is usually set up using an array of capacitor-like cells. Each cell consists of an insulating layer sandwiched between two metal contacts. These cells are arranged into a crossbar architecture, which allows each cell to be addressed independently by a particular word (horizontal) and bit (vertical) line. However, current between the lines can flow through different paths so often individual access devices,

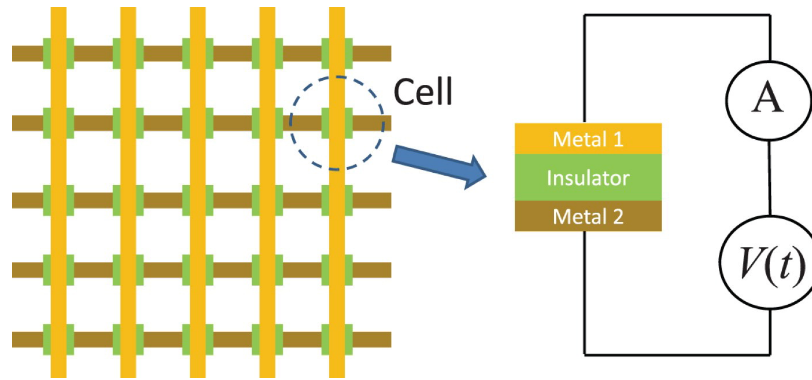


Figure 1.5: A schematic of a typical crossbar array of memristive cells. Each cell is connected via an array of vertical and horizontal metallic contacts separated by an insulating material. A represents an ammeter.[2]

such as transistors or diodes, are needed (Fig. 1.5).

Memristors are generally fabricated in a highly resistive state and so an electroforming step is used to put the cell into a state in which resistive switching operation can begin. During this electroforming process a high-voltage pulse is applied to the cell causing a soft breakdown in the insulating layer, dramatically reducing the measured resistance. This is thought to occur via the formation of metallic filaments across the cell, connecting the two metal electrodes.[29] This filament formation has been shown experimentally in both NiO[30] and TiO_2,[31] however the microscopic nature of the filament remains unclear. In the dielectric breakdown of SiO_2 investigated by Li *et al.* a deficiency of oxygen was observed along the breakdown path of 50-60%. This breakdown was caused by large currents passing through the oxide, and suggests a significant localized rearrangement of the atomic structure occurs during the electroforming process.[32]

1.4.3 Switching Mechanisms

There are two types of resistive switching shown in Fig. 1.6. In bipolar resistive switching, both voltage polarities are required to switch the device from the low resistance, ON, state to the high resistance, OFF, state and back. The switch usually occurs once the voltage overcomes a certain threshold. This means that a relatively high voltage is necessary to cause the switch but a low voltage applied across the cell

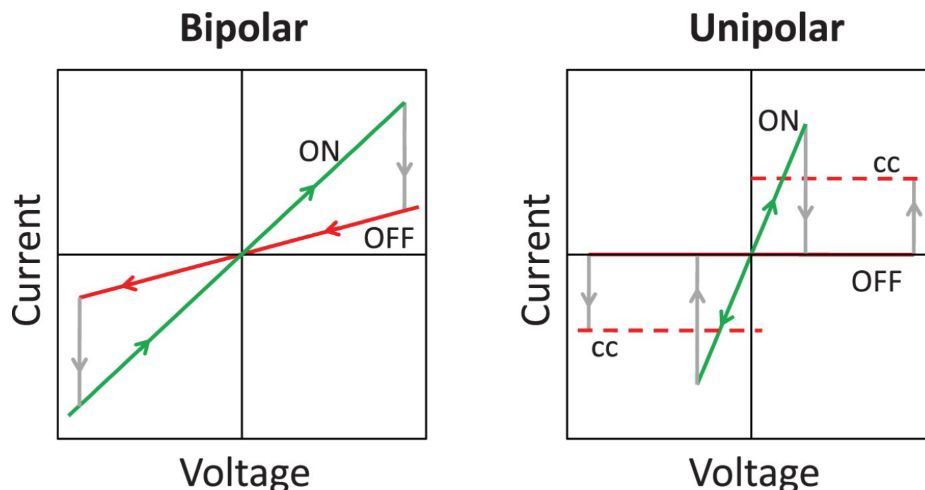


Figure 1.6: Typical $I - V$ curves for bipolar and unipolar switching devices. *cc* represents a compliance current.[2]

will not change the resistance. This bipolar regime often applies to electrochemical or ionic systems.

In 2006, Strukov *et al.* demonstrated a TiO_2 based bipolar resistive switch. Although resistive switching systems had been previously demonstrated, this was the first to be directly connected to Chua's memristive theory.[21] The cell consisted of a Pt/TiO_2/Pt stack in which the TiO_2 was split into 2 layers, one doped to create oxygen vacancies and the other undoped. It was suggested that the small voltages applied across the nanoscale cell yielded a very high electric field across the oxide. This caused an ionic drift of the positive vacancies and a shift in the barrier between the doped and undoped layers, which then led to a change in the resistance. This effect was then reversible by switching the voltage polarity. The proposed mechanism was later challenged by Wu *et al.* who suggested that the change in conductivity was due to the reduction of Ti^{IV} oxide to the much more conductive Ti^{III} oxide, analogous to a solid state redox reaction.[33]

Within the same system Jeong *et al.* suggested that the mechanism of the bipolar switching was in fact due to an asymmetric electroforming process led to oxygen-related electrochemical reactions, including vacancy formation and annihilation, at one of the TiO_2/Pt interfaces.[34] These electrochemical reactions, driven by the applied switching voltage, caused a variation in the density of oxygen vacancies at the interface

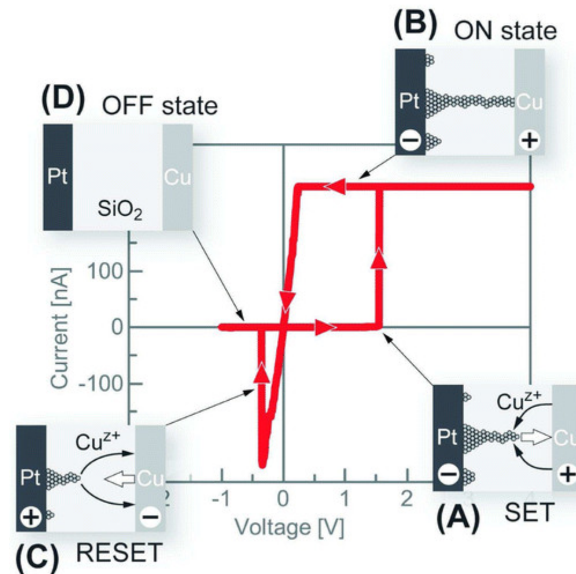


Figure 1.7: The $I-V$ characteristics of a Cu/SiO$_2$/Pt electrochemical metallisation cell. The insets show the dynamics of metallic filament formation.[3]

and a related change in the Schottky barrier height leading to the change in resistance of the cell.

A different type of bipolar memristor uses a nanoionic-based cell, comprised of a thin insulator layer, sandwiched between two electrodes. The insulating material (e.g. SiO$_2$) acts as a solid electrolyte with one electrode made of a relatively inert metal (e.g. Pt) while the other is electrochemically active (e.g. Cu).[3] A negative bias is applied to the inert electrode forcing a flow of metal ions, from the now positive, active electrode, into the electrolyte and towards the inert electrode. These metal ions form a metallic filament connecting the two electrodes, reducing the resistance of the cell. When the polarity of the bias is reversed the flow of ions is reversed and the filament is destroyed. The $I-V$ curve of this process is shown in Fig 1.7.

Memristive systems can also exhibit unipolar switching in which the switching is produced by two voltages. The systems begin with an electroforming step creating a weak conductive filament through an insulating layer. In the reset process (ON to OFF) a voltage is applied and this filament is partially destroyed due to the large amount of heat released.[35] The set process is driven by a higher voltage than the reset process, reforming the conductive filament. A compliance current is used to prevent hard breakdown of the insulating layer.[31]

1.4.4 MIM ReRAM

Resistive random access memory (ReRAM) is one of the applied, commercial uses for memristive cells. ReRAM cells are being considered to replace both flash memory and DRAM for low power electronic devices. As such, ReRAM devices must have the specifications to compete with existing technologies e.g. fast switching speeds, low power consumption, high density on a chip and long lifetimes. Furthermore, ReRAM cells must be able to fit in with current microelectronic fabrication techniques, and into full scale circuits, without expensive modifications. These criteria narrow the number of candidates for ReRAM materials significantly. The current material frontrunners include binary oxide insulators such as TiO_x, TaO_x, HfO_x and SiO_x, along with Pt or TiN electrodes. These materials are arranged in a capacitor-like metal-insulator-metal (MIM) stack in the ReRAM cell.

SiO_2 is of particular interest since the processing technology required to make silica based devices has been widely used for decades and the cells would be extremely cheap to manufacture. In 2012 a Si-rich, SiO_2 resistive switching device was reported by Mehonic *et al.*.[22] Scanning tunnelling microscopy (STM) investigations on the cell suggested the formation of a conductive pathway 10 nm across. It was suggested that this pathway was likely to be a column of a highly non-stoichiometric, suboxide phase of SiO_2. This column was then switched via field-driven filament formation and current-driven filament destruction. This switching process is thought to be caused by the motion of oxygen atoms and vacancies into and away from this Si-rich region.

A major issue with ReRAM has been the endurance, or the number of switching cycles a memristor can undergo before permanent breakdown. However in 2011 a high-endurance TaO_x based ReRAM cell was demonstrated by Lee *et al.*.[36] The cell comprised of a Pt base electrode with a thick, medium resistivity TaO_2 base layer and a localized 10 nm highly resistive Ta_2O_{5-x} layer above it, capped with a second Pt electrode. This asymmetric stack localized the resistive switching in a narrow layer, reducing the switching current drastically and resulted in very high endurance of 10^{12} switching cycles.

ReRAM with a switching speed of <10 ns was demonstrated by Lee *et al.* in 2010.[37] This cell was comprised of a TiN/Ti/HfO_x/TiN stack. It was suggested that annealing after deposition of the thin Ti layer allowed the Ti to absorb oxygen atoms from the HfO_2 layer leaving behind a large amount of oxygen vacancies which are central to the resistive switching process.

Despite recent advances in ReRAM technology the underlying switching mechanism remains unclear. A HfO_2 based ReRAM device was demonstrated by Goux *et al.*, by sandwiching the oxide between various combinations of TiN and Pt electrodes.[24] The proposed switching mechanism in this case involved a set process whereby absorption of oxygen atoms into the electrode left behind a large concentration of oxygen vacancies to form a conductive pathway. On reversal of the bias polarity, these oxygen atoms would be recovered by the oxide due to an ionic drift. This would result in the conductive pathway being broken close to the interface.

ReRAM is a promising technology but this discrepancy between various proposed switching mechanisms means that technological development is significantly more difficult. The microscopic nature of the conductive filaments can only be probed before and after the switching event meaning the dynamic process is difficult to deduce. Furthermore, it is particularly difficult to directly address the metal/oxide interface and, as such, the suggested role which it plays in the switching mechanism, varies from simply being an inert metal contact, to being the focus of the switching phenomena. It is in these cases where computational modelling is particularly useful.

As mentioned above, the generally accepted mechanisms for the electroforming and switching processes in MIM ReRAM involve the motion of oxygen ions. Deposited HfO_2 is a polycrystalline wide band gap material in which oxygen vacancies are known to be intrinsic defects. Broqvist *et al.* used hybrid density functional theory (DFT) to show that the HfO_2 oxygen vacancy is responsible for the (Poole-Frenkel-type) TAT conduction, by matching calculated defect energy levels with those of specific experiments.[38] This supports the idea that conducting filaments can be made solely of oxygen defects, as electrons conduct via theses oxygen vacancies. McKenna *et*

al. also showed that oxygen vacancies favourably segregate towards the HfO_2 grain boundaries, making grain boundaries a likely site for the formation of a conducting filament.[39] However, conducting filaments have been shown to be significantly larger than grain boundaries so bulk processes must be important in their formation.[6] DFT has also shown that negatively charged interstitial oxygen atoms have a particularly low barrier to diffusion of 0.3 eV within bulk HfO_2. This means an ionic drift of charged oxygen interstitial ions is highly likely under an external bias and may be responsible for part of the set/reset mechanism.[40]

1.5 Summary

The development of transistor technology has driven the microelectronics industry forwards for the past 40 years, however physical scaling limits are now being reached and new materials and manufacturing techniques must now be considered. The use of high-k dielectrics, such as HfO_2, has led to the latest generation of transistors, but reliability still remains a significant problem. In particular the generation of charge traps, and the role of defects in the gate oxide layer and at interfaces in the breakdown process.

ReRAM has been shown to be a promising emerging technology for the next generation of nanoscale memory storage devices but questions remain as to the exact switching mechanism. In the case of MIM ReRAM there is a significant amount of evidence suggesting the formation of conductive filaments through the insulating layer, after the electroforming process. There is also evidence that the conduction within this filament is due to trap assisted tunnelling between oxygen vacancies within this filament. Furthermore, extended defects such as grain boundaries and the metal oxide interface are thought to play a central role both in filament formation and in the switching process itself.

2

Theoretical Background and Methods

2.1 Introduction

Computational modelling has become an important field within materials science, to analyse the properties of materials which are difficult to find with experiment, to give a deeper understanding of experimental results, and to predict important mechanisms. These modelling techniques are often based on first principles, or "*ab initio*", and, in theory, can derive the properties of materials through only the input of the atoms' locations and atomic numbers.

The simulation of these atoms are performed by solving the many-body Schrödinger equation for all the nuclei and electrons in the system. The time-independent Schrödinger equation for a many-nuclei, many-electron system is:

$$\hat{H}\Psi(r,R) = E_t\Psi(r,R), \tag{2.1}$$

where E_t and Ψ are the energy and the wave function of a system respectively, with electrons and nuclei at coordinates r and R, respectively. In this equation the Hamiltonian, $\hat{H}$, in atomic units, is given by the following general form:

$$\hat{H} = -\frac{1}{2}\sum_{i}\nabla_i^2 - \sum_{i,I}\frac{Z_I}{|r_i - R_I|} + \frac{1}{2}\sum_{i,j\neq i}\frac{1}{|r_i - r_j|} - \sum_{I}\frac{1}{2M_I}\nabla_I^2 + \frac{1}{2}\sum_{I,J\neq I}\frac{Z_I Z_J}{|R_I - R_J|}. \tag{2.2}$$

From left to right the terms represent the kinetic energy of the electrons, the Coulomb interaction between electrons and nuclei, the Coulomb interaction between electrons, the kinetic energy of the nuclei, and the Coulomb interactions between nuclei. Lower-case subscripts denote the identity of electrons and upper-case subscripts denote nuclei. M and Z represent the mass and charge of the nuclei respectively. The wave function, Ψ, of a many-body system, contains the coordinates of both electrons and nuclei and, due to the Pauli exclusion principle, must be anti-symmetric with respect to the exchange of the coordinates of two electrons.

The numerical solution of the many-body Schrödinger equation is impractical for systems containing more than a few electrons and nuclei. In practice, the problem is often solved using a number of approximations. The first is the Born-Oppenheimer (adiabatic) approximation, in which the motions of the electrons and nuclei are separated. Because the electrons are much lighter and faster than the nuclei, the approximation is made that they react instantaneously to any nuclear motion.[41] Based on this apporoximation the electronic part of the Schrödinger equation is separated from the nuclear part:

$$\hat{H}_e\Psi = E\Psi. \tag{2.3}$$

Ψ is now the many-electron wave function and depends parametrically on the positions of the nuclei. E is the electronic energy, and $\hat{H}_e$ is the electronic Hamiltonian, which only contains the first three terms of equation 2.2, i.e.

$$\hat{H}_e = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}, \tag{2.4}$$

where $\hat{T}$, $\hat{V}_{ne}$ and $\hat{V}_{ee}$ represent the electronic kinetic energy, the electron-nuclei inter-

action and the electron-electron interaction respectively.

The solution to the simplified many-electron Schrödinger equation is not possible analytically, and so numerical solutions are necessary. The Hartree-Fock (HF) method, and post-HF methods such as configuration interaction (CI), solve the equation based on the single electron wave function. These methods are effective but scale poorly, and so become computationally expensive for large systems. In 1964 density functional theory (DFT) was developed by Kohn *et al.*.[42,43] DFT is based on the electron density of the system, which is a function of three spatial coordinates, and substantially less expensive to compute than the electronic wave functions.

In this chapter, the theory and background behind the computational and analytical methods used throughout this thesis are introduced. Firstly, wave function methods are briefly introduced followed by the fundamentals of DFT. Then the background behind the conjugate gradient and nudged elastic band simulation methods are discussed. Finally, the analytical method behind calculation of point defect formation energies is explained.

2.2 Wave Function Methods

In the Hartree approximation,[44] the exchange and correlation effects of electron-electron interactions are neglected, and the many-electron wave function is written as the product of all the single-electron wave functions:

$$\Psi = \Psi_{HP} = \prod_{i=1}^{N} \phi_i(r_i), \tag{2.5}$$

where N is the number of electrons in the system. In this situation, $\hat{V}_{ee}$ is the classical electrostatic interaction:

$$\hat{V}_{ee} = \frac{1}{2} \sum_{i=1}^{N} \sum_{i \neq j}^{N} \frac{1}{|r_i - r_j|}. \tag{2.6}$$

Equation 2.5 violates the Pauli exclusion principle. To overcome this the HF ap-

proximation is made,[45,46] which expresses the wave function as an anti-symmetrizing combination of Hartree products using a Slater determinant:

$$\Psi = \Psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(r_1) & \phi_2(r_1) & \dots & \phi_N(r_1) \\ \phi_1(r_2) & \phi_2(r_2) & \dots & \phi_N(r_2) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_1(r_N) & \phi_2(r_N) & \dots & \phi_N(r_N) \end{vmatrix}, \tag{2.7}$$

where $\phi_i(r_j)$ refers to the electron j in single-electron orbital ϕ_i.

The HF approximation prohibits electrons with the same spin states to remain in one orbital, however it allows two electrons with opposite spin to remain at the same point. In other words, electron exchange is treated exactly, but electron correlation is omitted entirely.

A solution to the missing correlation effect is the use of CI which calculates the wave function as a linear combination of Slater determinants. This combination includes the interaction of the ground state determinant with every excited state determinant. However, this method becomes extremely expensive very quickly, scaling with N^6, compared to N^4 with HF and $\leq N^2$ with DFT.

2.3 Density Functional Theory

The density functional theory proposes that, the ground state energy of an interacting many-body system is a functional of electron density $\rho(r)$, and the ground state energy, E_0, corresponds to the ground state electron density $\rho_0(r)$. Once ρ_0 is known it is possible to calculate all other ground state properties.

According to the Hohenberg-Kohn theorem,[43] the ground state energy is:

$$E_0 = E[\rho_0] = \bar{T}[\rho_0] + \bar{V}_{ne}[\rho_0] + \bar{V}_{ee}[\rho_0], \tag{2.8}$$

where ρ_0 is the ground state electron density and bars represent averages. In equation 2.8, $\bar{V}_{ne}[\rho_0]$ is:

$$\bar{V}_{ne}[\rho_0] = \int \rho_0(r)v(r)dr, \tag{2.9}$$

where $v(r)$ is an external potential imposed by all the nuclei at point r.

The kinetic energy and electron-electron interaction, $\bar{T}[\rho_0]$ and $\bar{V}_{ee}[\rho_0]$, are grouped into one functional $\bar{F}[\rho_0]$:

$$E_0 = E[\rho_0] = \int \rho_0(r)v(r)dr + \bar{T}[\rho_0] + \bar{V}_{ee}[\rho_0] = \int \rho_0(r)v(r)dr + \bar{F}[\rho_0]. \tag{2.10}$$

In 1965, Kohn and Sham proposed a practical method for finding E_0 of a system of interacting electrons.[42] This method uses a fictitious reference system of N non-interacting electrons, which has the same ground state as the interacting system. In this ground state the electrons occupy the lowest energy orbitals, and have the electron density given by:

$$\rho(r) = \sum_{i=1}^{N} |\psi_i^{KS}(r)|^2, \tag{2.11}$$

where $\rho(r)$ is summed over all the occupied Kohn-Sham orbitals, ψ^{KS}. The ground state energy of this system is:

$$E_0^{KS} = E^{KS}[\rho_0] = T_s[\rho_0] + \int \rho_0(r)v(r)dr + E_H[\rho_0] + E_{XC}[\rho_0]. \tag{2.12}$$

The kinetic energy of the fictitious electrons, $T_s[\rho_0]$, is:

$$T_s[\rho_0] = -\frac{1}{2}\sum_{i=1}^{N} \langle \psi_i | \nabla_i^2 | \psi_i \rangle . \tag{2.13}$$

$E_H[\rho_0]$, the Hartree term, is the classical Coulomb energy between the electrons:

$$E_H[\rho_0] = \frac{1}{2}\int\int \frac{\rho_0(r)\rho_0(r')}{|r-r'|}. \tag{2.14}$$

The difference between $T_s[\rho_0] + E_H[\rho_0]$, in the non-interacting fictitious system,

and $\bar{T}[\rho_0] + \bar{V}_{ee}[\rho_0]$, in the true interacting system, is the exchange-correlation energy, E_{XC}:

$$E_{XC}[\rho_0] = \bar{T}[\rho_0] - T_s[\rho_0] + \bar{V}_{ee}[\rho_0] - E_H[\rho_0]. \tag{2.15}$$

For each non-interacting electron, i, the Kohn-Sham equation applies:

$$\left[-\frac{1}{2}\nabla_i^2 - \sum_I \frac{Z_I}{|r_i - R_I|} + \int \frac{\rho(r_j)}{|r_i - r_j|} dr_j + V_{XC}\right] \psi_i^{KS} = \epsilon_i^{KS}\psi_i^{KS}. \tag{2.16}$$

With the exchange-correlation potential, V_{XC}:

$$V_{XC}(r) = \frac{\delta E_{XC}[\rho(r)]}{\delta\rho(r)}. \tag{2.17}$$

Equation 2.16 can be simplified to:

$$\left[-\frac{1}{2}\nabla_i^2 + V_{eff}(r_i)\right] \psi_i^{KS} = \epsilon_i^{KS}\psi_i^{KS}, \tag{2.18}$$

where $V_{eff}(r)$ is the effective potential, containing the sum of the external potential, Hartree term and the exchange correlation potential.

In order to find the ground state of the system, it is necessary to calculate the Kohn-Sham orbitals ψ_i^{KS}. However, to calculate ψ_i^{KS} it is necessary to know the exchange-correlation potential, which itself, depends on ψ_i^{KS}. This means the Kohn-Sham equations must be solved self-consistently.

After knowing the ground state energy of the electronic system it is necessary to add the Coulomb interaction between the nuclei to obtain the total ground state energy of the system.

2.3.1 Approximations in DFT

To solve the Kohn-Sham equations, some approximations are made. These include approximating the exchange-correlation, using a pseudopotential for the external po-

tential, and using basis sets for the Kohn-Sham orbitals.

In equation 2.15 $\bar{T}[\rho_0]$ and $\bar{V}_{ee}[\rho_0]$ are unknown so V_{XC} must be approximated. In most solids, the electron density changes very slowly with respect to position so the electrons can be regarded as a homogeneous gas. In these cases the local density approximation (LDA) can be made:

$$E_{XC}^{LDA}[\rho(r)] = \int \epsilon(\rho(r))\rho(r)dr, \tag{2.19}$$

where $\epsilon(\rho(r))$ is the exchange plus the correlation energy of a uniform electron gas with density $\rho(r)$. The exchange part of the density can be shown to have the simple analytical form:

$$\epsilon_x(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \rho(r)^{1/3}. \tag{2.20}$$

The correlation part of the density is taken from quantum Monte Carlo calculations.[47]

In many solid systems LDA has proven to give very accurate results. However, in systems with more inhomogeneous electron densities it is less appropriate. The generalised gradient approximation (GGA) is a commonly used improvement on the LDA which takes into account the gradient of the density at the same coordinate.

Semilocal functionals systematically underestimate the band gap of materials with respect to experiment. Hybrid density functionals, which include a portion of Hartree-Fock exchange, are commonly used to help overcome this problem. However, including this exchange component drastically increases the computational demand in periodic systems. An alternative to this is a screened hybrid functional developed by Heyd, Scuzeria and Ernzerhof (HSE).[48–50] In the HSE functional the $1/r$ Coulomb potential is replaced by a screened potential to accelerate the spatial decay of the Hartree-Fock exchange term. This drastically reduces the computational cost in extended systems.

2.3.2 Pseudopotentials

In DFT the Kohn-Sham equation is applied to all electrons, including the valence and core electrons, to calculate the total energy of the system. However, the physical properties of, and interactions between, compounds and materials are mostly governed by the interactions between the valence electrons of atoms. The core electrons often do not contribute significantly to chemical bonding, and the calculation of their wave functions is computationally expensive. Therefore, pseudopotentials are used to describe these electrons. Pseudopotentials exploit the fact that core electrons are less significant in molecular properties by effectively freezing them with respect to the valence electrons, and replacing them, and the strong ionic potential due to the nucleus, with a weaker pseudopotential. This pseudopotential acts on a set of pseudo wave functions rather than the true valence wave-functions.

Pseudopotentials are ideally constructed so the scattering and phase shift properties associated with the pseudo wave functions are identical to those of the nucleus and core electrons for the valence wave function, but in such a way that there are no radial nodes within the core region (Fig. 2.1). The pseudopotential also has the benefit in that the relativistic effects of the core electrons can be taken into account while the potential is being constructed.

Generally pseudopotentials take the form:

$$V_{NL} = \sum_{lm} \Big| lm\rangle V_l \langle lm \Big|, \tag{2.21}$$

where lm are the spherical harmonics and V_l is the pseudopotential for angular momentum l. This operator, acting on the electronic wave function, decomposes it into spherical harmonics, each of which is multiplied by the pseudopotential V_l.[4]

2.3.3 Basis Sets

Solving the Kohn-Sham equations usually requires the single-electron orbitals to be written as a linear combination of known functions, called basis functions. A general

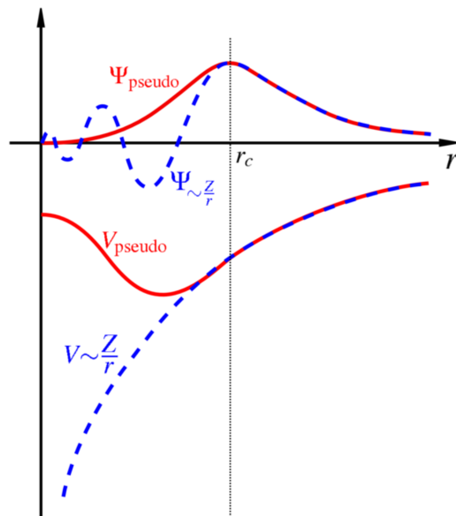


Figure 2.1: An illustration of the all-electron (blue) and pseudoelectron (red) potentials and their corresponding wave functions. The point at which the potentials match is shown as r_c.[4]

approach to this involves projecting the single-electron orbital onto a linear combination of basis orbitals:

$$\psi_i(r) = \sum_{\alpha=1}^{M} c_{i\alpha}\phi_\alpha(r), \tag{2.22}$$

α labels the basis functionals, $c_{i\alpha}$ is an expansion coefficient, and the sum is over all the basis functions up to the size of the basis set, M. Based on this generalised basis representation, the Kohn-Sham equations can be shown to become a generalised linear eigenvalue problem:

$$Hc = \Lambda Sc, \tag{2.23}$$

where Λ is the diagonal matrix of the eigenvalues, matrix c is a column eigenvector, i.e. the expansion coefficients. H and S are the expansion and overlap matrices respectively and are given by:

$$H_{\alpha\beta} = \left\langle \phi_\alpha|\hat{H}|\phi_\beta \right\rangle = \int \phi_\alpha^*(r)\hat{H}\phi_\beta(r)dr, \tag{2.24}$$

$$S_{\alpha\beta} = \langle \phi_\alpha \phi_\beta \rangle = \int \phi_\alpha^*(r)\phi_\beta(r)dr. \tag{2.25}$$

For molecular systems, the electron density is localised around the nuclei, and so the most efficient approach is to expand the single electron wave functions in terms of the atomic orbitals.

Gaussian basis sets are atomically centred functions and allow for a compact description of the wave function. The density represented as a Gaussian basis set is written as:

$$\rho(r) = \sum_{i=1}^{N} |\psi_i(r)|^2, \tag{2.26}$$

where $\rho(r)$ is the density, N is the number of electrons and $\psi_i(r)$ is the ith molecular orbital. The molecular orbital can be approximated by a linear combination of atomic orbitals (LCAO), with each molecular orbital described by:

$$\psi = \sum_j C_{mi}\psi_j, \tag{2.27}$$

where C_{mi} is a mixing coefficient of the atomic orbital and ψ_j is the atomic orbital which is a contracted Gaussian that takes the form:

$$\psi = \sum_\mu C_{\mu n} g_\mu, \tag{2.28}$$

where g_μ is a primitive Gaussian. Typically, the mixing coefficients of the molecular orbitals, C_{mi}, are varied until the ground state density is obtained, while the mixing coefficients of the contracted Gaussian, $C_{\mu n}$, are kept fixed.

Gaussian basis sets have the advantage that there are efficient algorithms which can be used to analytically calculate the matrix integrals. However, they suffer from the finite size of the basis set.

For periodic systems such as solids, plane waves are a better choice for representing the single-electron wave function, due to Bloch's theorem. Crystal systems contain

a huge number of atoms and so can effectively be treated as infinite. For a periodic crystal, the effective potential field (described in equation 2.18) is periodic, and so the Kohn-Sham orbital, according to Bloch's theorem, can be written as:

$$\psi_{i,k}(r) = e^{ik\cdot r}u_{i,k}(r), \tag{2.29}$$

where k is the wave vector in the first Brilouin zone, and $u_{i,k}$ is periodic:

$$u_{i,k}(r) = u_{i,k}(r+R). \tag{2.30}$$

R is the lattice vector in real space. By expanding $u_{i,k}$ into a Fourier series, the Kohn-Sham orbitals can be obtained:

$$\psi_{i,k}(r) = \frac{1}{\sqrt{\Omega}}\sum_{G} c_{ik}(G)e^{i(k+G)\cdot r}. \tag{2.31}$$

$\sqrt{\Omega}$ represents the crystal volume, $c_{ik}(G)$ is the Fourier coefficient, G is the reciprocal lattice vector and $e^{i(k+G)\cdot r}$ represents a plane wave. The Kohn-Sham orbitals in different unit cells of the crystal differ only by a phase factor of $e^{ik\cdot R}$. From this we can map the wave function, in infinite real space, onto a wave function in a unit cell within k space. This means the infinite periodic crystal can be calculated by just calculating one unit cell.

Since this method is theoretically modelling an infinite crystal, a number of approximations must be made in order to make the calculations feasible. In equation 2.31, the Kohn-Sham orbital is a series of infinite terms. In practice, this basis set must be truncated at some point, this truncation point is called the cutoff energy. The higher in energy this cutoff is the more accurate, and computationally expensive, the calculation will be. Similarly, k is a continuous term within the first Brillouin zone, however, only a few k-points need to be sampled for reasonable results. To do this a Monkhorst-Pack sampling method is used to choose an evenly spaced set of k-points within the Brillouin zone.[51]

There is also a Gaussian plane waves method (GPW) which is designed so that the

electron density of the system is represented in both Gaussian and plane wave bases. The method uses an atom-centred Gaussian-type basis to describe the wave functions, but uses an auxiliary plane wave basis to describe the density.[52] This method takes the time-advantageous parts of each type of basis set for different elements of the calculation, whilst maintaining accuracy.

2.4 Conjugate Gradient Method

Various algorithms can be used in order to optimise the geometry of structure. The conjugate gradient method (CG) uses an iterative algorithm which minimises the energy of the system. Optimisation methods such as this follow the forces applied on the atoms until their magnitude drops below a certain pre-set value.[5,53]

The conjugate gradient method is based on the steepest descent (SD) method. In SD a force vector, F_j, acting on a body is followed from an initial position, R_j, until there is zero force acting on the body:

$$R_{j+1} = R_j + \alpha F_j. \tag{2.32}$$

α is an adjustable parameter which is set to be less than $1/k_{max}$, where k_{max} is the maximum curvature of the system. The SD method is the simplest optimiser, and converges slowly for stiff systems.

The CG method follows conjugate search directions, rather than just following the force. The first step in the algorithm calculates the gradient of the force, as in the SD method. In each following step the preconditioned gradient is conjugated to the previous search direction in the following procedure:

Initialise the search direction, d_0 along the initial force F_0,

$$d_0 = F_0, \tag{2.33}$$

calculate the step size λ,

$$R_{j+1} = R_j + \lambda d_j, \tag{2.34}$$

evaluate the new conjugate search direction,

$$d_{j+1} = F_{j+1} + \gamma d_j, \tag{2.35}$$

$$\gamma = F_{j+1} \cdot F_{j+1} - F_j / |F_j^2|. \tag{2.36}$$

The CG method is only marginally more complex than the SD method, however, it has the advantage that because the search directions are based on gradients, the process makes good uniform progress towards the solution with every step.

2.5 Nudged Elastic Band Method

It is often necessary to find the minimum energy pathway (MEP) by which a reaction proceeds, and the height of the energetic barrier that needs to be overcome for the reaction to occur. This usually involves finding the lowest energy route between two energetic minima along a potential energy surface. The activation barrier is the saddle point along this reaction coordinate. The reaction coordinate is the rearrangement of atoms in the system necessary to go from one minimum to the other. This can be used to describe chemical reactions, changes in molecular conformations or diffusion processes in solids. The activation barrier, at the saddle point of the MEP, is related to the probability and the rate at which the reaction will occur.[54]

There are a number of ways to calculate the MEP.[55] One of the more reliable and efficient methods is the nudged elastic band (NEB) method.[5] The NEB method finds the lowest energy pathway between two stable states, the initial state and the final state, by minimising a chain of intermediate images between those two states, along the reaction coordinate. These image configurations are connected by spring forces to keep them equally separated along the coordinate. When the NEB is converged

these images describe the reaction mechanism with a resolution equal to the number of images.

The images are chosen as a linear interpolation between the initial and final states, else, if there is a known intermediate state, an interpolation via that geometry. The images along the NEB are minimised by a force projection scheme, in which potential forces act perpendicular to the band, and parallel forces act along the band. To make these projections a tangent along the path, $\hat{\tau}$, is used as the unit vector to the higher neighbouring image. The NEB force, F_i^{NEB}, on an image, i, contains two components – the potential force perpendicular to the NEB, $F_i^{\perp}$, and the spring force parallel to the band, $F_i^{S\|}$:

$$F_i^{NEB} = F_i^{\perp} + F_i^{S\|}, \tag{2.37}$$

$$F_i^{\perp} = -\nabla(R_i) + \nabla(R_i)\cdot \hat{\tau}_i\hat{\tau}_i, \tag{2.38}$$

$$F_i^{S\|} = k(|R_{i+1} - R_i| - |R_i - R_{i-1}|)\hat{\tau}_i. \tag{2.39}$$

R_i is the position of the ith image and k is the spring constant.

The NEB method has many desirable qualities. Given enough images it converges to the MEP, always giving one continuous path. It only requires the calculation of the potential energy and the first derivative of the energy with respect to the coordinates. It is also inherently parallel, so is easily parallelised across computing nodes.

2.6 Defect Energetics

The formation energy of a point defect depends on the chemical potential, as well as the Fermi level if the defect is charged. The largest possible range of chemical potentials is given by the stability limits of a material, with respect to the pure constituents of that material, e.g. for a binary metal oxide, MO_2:

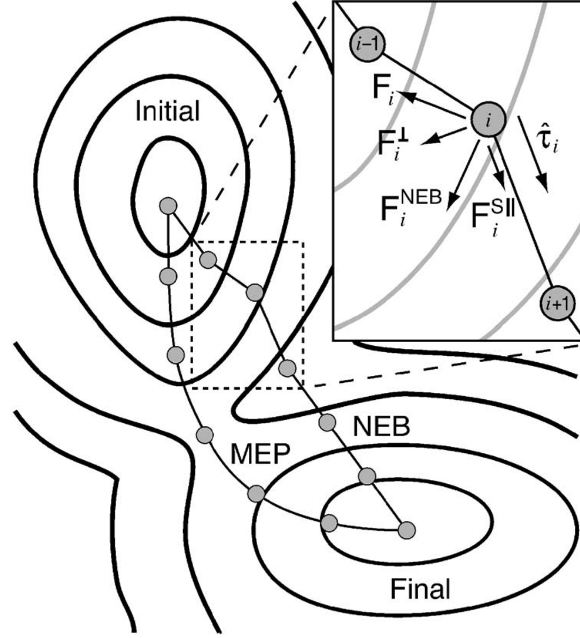


Figure 2.2: Forces $F_i^{S\|}$ and $F_i^{\perp}$ act on F_i^{NEB} to minimise images towards the MEP on the potential energy surface.[5]

$$\mu_M^0 + \Delta E_f^{MO_2} < \mu_M^{MO_2} < \mu_M^0, \tag{2.40}$$

$$\mu_O^0 + \frac{1}{2}\Delta E_f^{MO_2} < \mu_O^{MO_2} < \mu_O^0, \tag{2.41}$$

where $\mu_M^{MO_2}$ and $\mu_O^{MO_2}$ are the chemical potentials for M and O in MO_2 respectively, μ_M^0 is the chemical potential for M in pure metal, μ_O^0 is the chemical potential for O in an O_2 gas, and $\Delta E_f^{MO_2}$ is the formation enthalpy of MO_2.

The formation energy of a charged defect, $E_{for}(D)$, can be calculated using:

$$E_{for}(D) = E^q(D) - (E_0 \pm \mu_x) + q(E_V + E_F). \tag{2.42}$$

E_D^q is the total energy of the relaxed defective supercell, E_0^q is the energy of the perfect cell, μ_x is the elemental chemical potential, with a positive sign for interstitials and a negative sign for vacancies, E_V is the valence band maximum (VBM) of the perfect supercell, and E_F is the Fermi level, referenced to the VBM.

In the calculations presented throughout this thesis, the Fermi level is set by the

location of the highest occupied molecular orbital, and does not change with respect to the defective supercell and the perfect supercell, i.e. the number of electrons in the system is conserved throughout the formation process. This means that the electron chemical potential term, q, in equation 2.42 can be neglected.

Furthermore, rather than use a range of chemical potentials for the oxygen reference, the energy of a single O_2 molecule is used. The calculation of this oxygen reference is discussed in section 4.2

Hence, the formation energy of an oxygen defect, $E_{for}(D)$, used throughout this thesis, is calculated as,

$$E_{for}(D) = E_D^q - (E_0^q \pm nE_O) \tag{2.43}$$

Where E_D^q is the total energy of the relaxed defective supercell, E_0^q is the energy of the perfect cell and n is the number of defect atoms. E_O is the energy of half an O_2 molecule calculated in an asymmetric unit cell and is added for interstitial defects and subtracted for vacancies. The charge of the defect and ideal lattice, q, must be conserved, else a charge correction needs to be applied. All the calculations must also be carried out using the same methods and functionals.

This formation energy effectively describes the energy required to remove O atoms from the crystal lattice to an infinite non-interacting distance or, conversely, add atoms from an infinite distance into the lattice.

2.7 Summary

Throughout this thesis computational modelling is used to calculate various properties of solid state systems. The Vienna "*ab initio*" Simulation Package (VASP) is the most commonly used code in this work. VASP is largely based on the DFT methods introduced in section 2.3; using GGA and hybrid potentials and plane wave basis sets. In addition to VASP, the electronic structure program Gaussian, which uses Gaussian basis sets, is used in section 3.2, and the CP2K code, which uses Gaussian plane

wave (GPW) basis sets is used in section 8.3. The CG algorithm is used for geometry optimisations, and the NEB method is used to find low energy diffusion pathways and the energetic barriers to diffusion. Many of the conclusions in this thesis are based on the formation energies of various defects; these energies are calculated using equation 2.43.

3

HfO_2

3.1 Introduction

Hafnia has become an important material as a high-k insulating dielectric in new generations of transistors, and a leading material for use in ReRAM applications.

Hafnia, HfO_2, is the oxide of hafnium. Hafnium makes up about 5.8 ppm of the earth's crust, however, it is not found as a free element but in solid solution with zirconium compounds such as zircon, $ZrSiO_4$, with 1-4% of the zirconium ions replaced by hafnium ions.[56] The main source of hafnia is from the hafnium by-products produced during the purification of zirconium containing minerals.

Zirconium and hafnium are mostly mined from heavy mineral sands ore deposits found particularly in Brazil, Malawi and Western Australia. Zirconium has a low neutron capture-cross section and high temperature and chemical stabilities making it the ideal material for cladding nuclear fuel rods. This property drove zirconium to be a commercially valuable material and encouraged much of the research into its chemistry and properties in the second half of the 20th century. Hafnium absorbs thermal neutrons 600 times more strongly than zirconium, so its concentration in the zirconium cladding material must be less than 100 ppm. This means that the removal

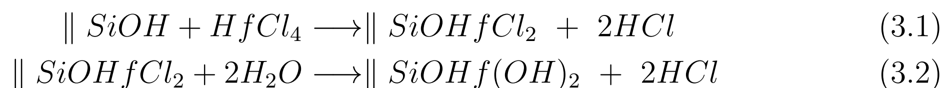
of hafnium is an important step in the zirconium production process.[57,58]

The separation of Hf and Zr is very difficult as both elements have very similar chemical properties, particularly ionic radii (Zr: 72 pm, Hf: 71 pm). They do, however, have different melting points (Zr: 1857°C, Hf: 2222°C), boiling points (Zr: 4200°C, Hf: 4450°C), and different solvent solubilities.[59] These properties mean the most common methods of Zr/Hf separation involve solvent extraction and fractional separation.[60,61] The end product of the fractional separations is the tetrachlorides, $ZrCl_4$ and $HfCl_4$.[62] $HfCl_4$ is then used as an atomic layer deposition (ALD) precursor for the production of HfO_2 thin films.

3.1.1 Atomic Layer Deposition of HfO_2

Atomic layer deposition (ALD) is one of the most common techniques that is used for growing HfO_2 thin films for transistors and ReRAM devices.[24,63] During ALD material precursors are injected into a reaction chamber and settle onto a substrate. The reactants are introduced in stages in order for the reaction to take place on the substrate surface, forming very thin films of material on the substrate after each stage. This method gives high quality films and interfaces with precise control over the thickness.

ALD of HfO_2 onto an Si substrate has been demonstrated with $HfCl_4$ and H_2O precursors in a binary self-limiting reaction sequence:



$$\| SiOH + HfCl_4 \longrightarrow \| SiOHfCl_2 \; + \; 2HCl \tag{3.1}$$
$$\| SiOHfCl_2 + 2H_2O \longrightarrow \| SiOHf(OH)_2 \; + \; 2HCl \tag{3.2}$$

Dkhissi *et al.* used quantum mechanical computational modelling to investigate the mechanism of the ALD of HfO_2. They found it involves a series of steps: The injected $HfCl_4$ dissociatively adsorbs onto the hydroxylated SiO_2 surface, forming a bond between the Hf atom and a surface hydroxyl oxygen, whilst one chlorine atom bonds to the hydrogen forming the HCl. The hafnium atom of the remaining $HfCl_3$ bridges to form an oxane bridge with the oxygen of a neighbouring hydroxide,

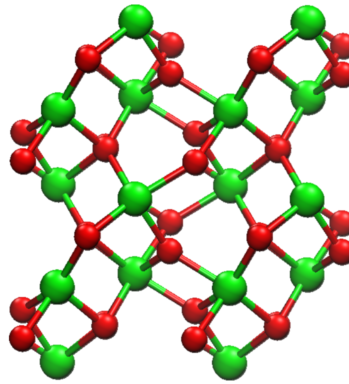


Figure 3.1: Monoclinic HfO_2 (Hf: green, O: red), [001] view.

passivating the surface with $HfCl_2$, eliminating a second HCl molecule and terminating the first half-cycle. After injection of H_2O, the bridging $HfCl_2$ transforms to $Hf(OH)_2$ leaving a hydroxylated hafnia surface with which the next batch of injected $HfCl_4$ can react.[64] This sequence can then be repeated until the film is the desired thickness.

3.1.2 Bulk Properties of HfO_2

HfO_2 is a polymorphic crystal with 3 solid phases, of which the monoclinic phase (space group $P2_1/c$) is the most stable at room temperature (Fig. 3.1). At around 2000 K hafnia transitions into the tetragonal phase ($P4_2/nmc$). This phase is metastable at room temperature.[63] At 2870 K cubic hafnia ($Fm3m$) forms however this phase is not observed under any conditions in which a hafnia based device will operate. At high pressure (> 30 GPa) a high density cotunnite ($Pnma$) phase forms which is then stable at ambient conditions.[65] Amorphous HfO_2 is grown using low temperature ALD (< 500 K)[66] and electron beam evaporation.[67] Deposited HfO_2 thin films tend to form polycrystalline monoclinic and tetragonal structures when annealed at high temperatures.[68–70]

HfO_2 has two advantages which have made it a very desirable material for use in transistors. Firstly, it is a good insulator with a wide band gap. The band gap has been measured using a variety of techniques including XPS,[71] XAS,[72,73] EELS,[74] SE,[75] UPS and IPS,[76] giving values of 5.1 – 6.0 eV. This variation in band gap measurement is partly due to the thickness of the HfO_2 film measured, with films of

30 Å giving the widest band gap of 6.0 eV. Films grown using similar ALD growth techniques and conditions give similar band gaps of around 5.7 eV, however thermal treatment can lead to further variation in the measured value.

Secondly, HfO_2 has a high dielectric constant, k. As previously mentioned in section 1.3 this high-k value allows a thicker insulating layer of HfO_2 whilst retaining the charge induced in the channel of the transistor, reducing the likelihood of current leakage and maintaining performance. The dielectric constant of poly-crystalline m-HfO_2 has been measured as 22 - 24 by measuring the film thickness and capacitance.[77–79] Lin *et al.* measured a dielectric constant of 23 however they note that the constant varied between 16 and 23 depending on processing conditions and the film thickness.[80] The high dielectric constant is related to a soft mode within the phonon spectrum of HfO_2. This soft mode relates to a transition between the monoclinic and tetragonal phases of HfO_2 and is discussed further in section 3.3.2.

In this chapter the ALD precursor $HfCl_4$ and bulk monoclinic HfO_2 are used as test cases for computational methodologies. In section 3.2 properties of molecular $HfCl_4$ are calculated and used to analyse various basis sets and pseudopotentials using the Hartree-Fock (HF) self consistent field method implemented in the Gaussian 03 code.[81] It must be noted that these results were included here for completeness, as they were originally used to provide guidance for the choice of basis sets in further local basis codes (e.g. Crystal). However, as shown in later chapters of this thesis, it was necessary to look into electron localisation of bulk crystals, for which plane wave codes are a better choice. In section 3.3 properties of bulk monoclinic HfO_2 are calculated using various different potentials using the periodic plane-wave code VASP.[82,83]

3.2 $HfCl_4$

In this section the electronic structure code, Gaussian, is used to calculate the energy and molecular properties of $HfCl_4$. A variety of basis sets are used to describe the

Cl atoms, and a variety of pseudopotentials are used to describe the Hf atoms. The basis sets used can be split into two categories, Pople and Dunning.

3.2.1 Pople Basis Sets

In 1970 Pople *et al* suggested using different Gaussian-type functions (GTF) to describe the core and valence shells of an atom.[84] For example, the 3-21G basis set uses a single contracted GTF (CGTF) that is a linear combination of 3 primitive GTFs to describe the each inner-shell atomic orbital and a CGTF of 2 primitive Gaussians and a single Gaussian to describe the valence atomic orbitals. The orbital coefficients and exponents were chosen in a way that would minimise the energy of the atomic ground states. The varying of inner and outer weightings gave the basis sets directional variability allowing for anisotropy and showed good improvement in a variety of properties, including hydrogenation energies, over earlier basis sets.[84]

The 6-31G* basis set adds a set of d-type or f-type Gaussian polarization functions to the basis set which allows for a better description of polarization and hybridization in bonds. Diffuse functions can also be added to improve the accuracy of calculations involving anions, molecules with lone pairs or hydrogen bonded dimers which have significant electron density far from the atomic nuclei.[85]

3.2.2 Dunning Basis Sets

Dunning *et al* developed a series of basis sets designed in such a way that they can be systematically increased in size and their results can be empirically extrapolated to the complete basis set limit. These sets are known as correlation-consistent polarized valence N-zeta (cc-pVNZ), in which V indicates that they act only on the valence electrons, and N describes the number of polarization functions used, as in table 3.1. The basis sets can be augmented with core functions to improve the basis sets of geometric calculations (aug-cc-pVNZ). Pseudopotential basis sets can be included for large atoms (cc-pVNZ-PP).[86]

Basis set	H-He	Li-Ne	Na-Ar
cc-pVDZ	[2s1p]	[3s2p1d]	[4s3p2d]
cc-pVTZ	[3s2p1d]	[4s3p2d1f]	[5s4p3d1f]
cc-pVQZ	[4s3p2d1f]	[5s4p3d2f1g]	[6s5p4d2f1g]

Table 3.1: Functions used in the Dunning basis sets.

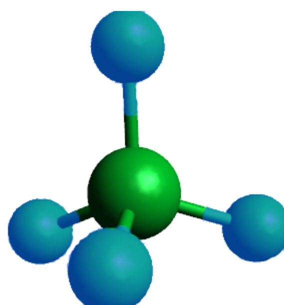


Figure 3.2: The T_d structure of $HfCl_4$ (Hf: dark green, Cl: blue-green).

3.2.3 Structural Properties of $HfCl_4$

$HfCl_4$ has a tetrahedral (T_d) symmetry and in every case the geometry of the molecule was minimised into the correct tetrahedral structure shown in Fig. 3.2.

Table 3.2 shows that the inclusion of the additional polarization functions improved the accuracy of the calculated bond lengths and vibrational spectra in both Pople and Dunning basis sets. This shows that the polarization of the Hf–Cl bond is significant. Fig. 3.3 shows that the increase in the size of the basis set also gave a decrease in energies however, the addition of polarization functions did not reduce the energy as significantly in $HfCl_4$ despite the more ionic Hf–Cl bond. This is because the polarization functions were only added to the chlorine atoms reducing their impact on the molecular wave-function.

Several basis sets describing the Hf atomic orbitals were tested, using the LANL2 pseudopotential to describe the core electrons. The LANL2TZ gave the best results underestimating the equilibrium bond length by just 0.0014 Å and the first ionization energy by 0.08 eV, as well as giving the closest reproduction of the experimental Raman vibrational frequencies. The inclusion of f-polariation functions

Method/ Basis Set	Bond Length /Å	Vibrational Frequency/ cm^{-1}			
		ν_1	ν_2	ν_3	ν_4
Experimental	2.33	102	112	382	390
HF/ 6-31G	2.36	101	104	370	381
HF/ 6-31G*	2.33	101	110	392	398
HF/ 6-31+G	2.36				
HF/ 6-311G	2.35	99	103	369	378
HF/ 6-311G*	2.32	100	103	393	395
HF/ 6-311+G	2.36	98	102	362	376
HF/ cc-pVDZ	2.33	106	111	388	392
HF/ aug-cc-pVDZ	2.33	101	111	393	399
HF/ cc-pVTZ	2.32	106	113	402	403
HF/ aug-cc-pVTZ	2.33	101	109	395	399

Table 3.2: Bond lengths and Raman active vibrational frequencies of $HfCl_4$ using different basis sets to describe the Cl atoms

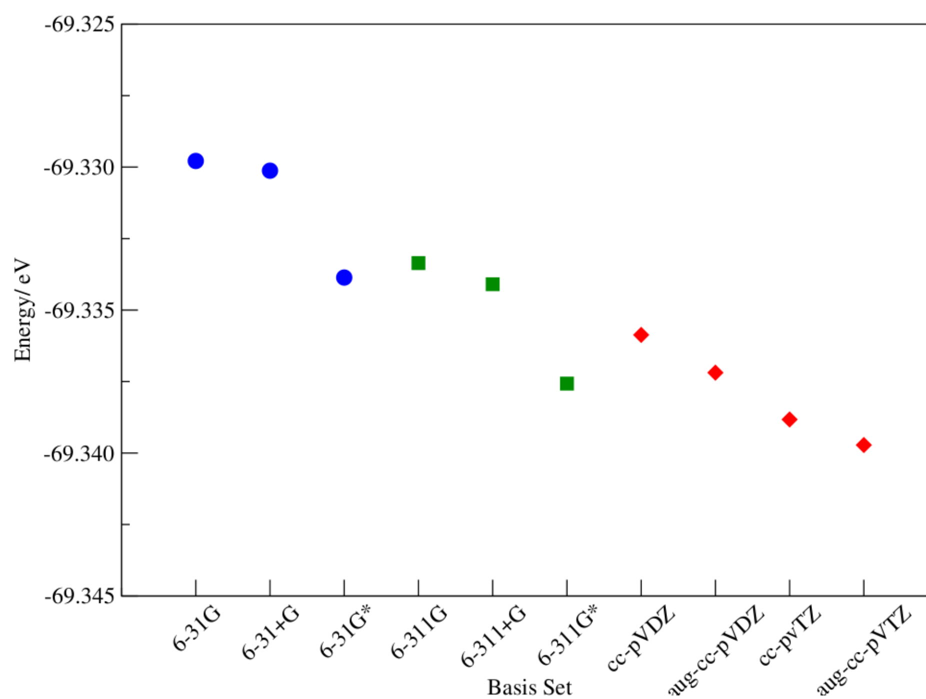


Figure 3.3: Total Hartree-Fock energies for the $HfCl_4$ molecule calculated with various basis sets on the Cl atoms.

Method/ Basis Set	Bond Length /Å	I_1/ eV	Vibrational Frequency/ cm^{-1}			
			ν_1	ν_2	ν_3	ν_4
Experimental	2.33	12.03	101.5	112	382	390
HF/ LANL2DZ[90]	2.32	11.76	103	107	384	385
HF/ LANL2TZ[91]	2.32	11.95	102	107	387	391
HF/ LANL2TZ(f)[91]	2.36	12.99	102	108	392	393
HF/ cc-pVDZ	2.33	11.95	103	111	392	397

Table 3.3: Bond lengths, 1st ionisation energies and Raman active vibrational frequencies of $HfCl_4$ using different basis sets to describe the valence electrons of Hf atoms and a 6-311G* basis on the Cl atoms.

in the LANL2TZ(f) basis set increased the first ionisation energy considerably. The pseudopotential is optimised in conjunction with the basis set. This means that variations in the basis set alters the pseudopotential, which in turn, affects the molecular Hamiltonian. Because of this, the total energies of the structures are not directly comparable. All electron calculations were attempted using the ANO-RCC[87,88] and cc-pVTZ[89] basis sets however these proved impractically memory intensive. Calculations using the cc-pVDZ basis set were successful however the results were very similar to those calculated with the LANL2TZ effective core potential but were 30% more time consuming.

3.2.4 Vibrational Properties of MX_4

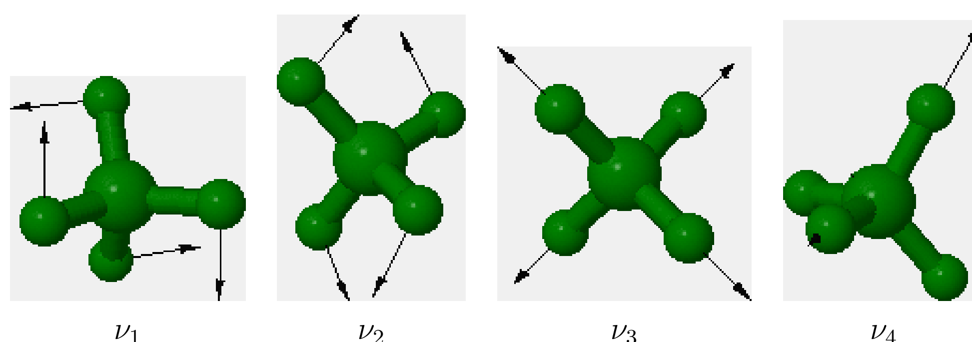


Figure 3.4: Schematic of the four Raman active vibrational modes of T_d MCl_4: (ν_1) Twist, (ν_2) scissor, (ν_3) symmetric stretch, (ν_4), asymmetric stretch.

Molecule	Bond Length /Å	Raman Active Vibrational Frequencies/ cm^{-1}			
		ν_1	ν_2	ν_3	ν_4
$SiCl_4$	2.03	159	238	446	656
$TiCl_4$	2.17	123	148	410	513
$ZrCl_4$	2.34	99	113	383	425
$HfCl_4$	2.33	106	110	385	388
$SiBr_4$	2.20	96	139	267	533
$TiBr_4$	2.34	75	92	254	413
$ZrBr_4$	2.52	63	79	238	322
$HfBr_4$	2.52	66	78	242	276

Table 3.4: Bond lengths and Raman vibrational frequencies of tetrahedral MCl_4 molecules.

The T_d MX_4 molecules (M = Si, Ti, Zr, Hf; X = Cl, Br) give four Raman active vibrational modes. These Raman active modes were calculated using the 6-31G** basis set to describe all atoms and the LANL2TZ pseudopotential was used to describe the Hf and Zr cores, and the results are shown in Table 3.4.

The frequencies are separated into two sets, the lower frequency pair are attributed to distortions of the bond angles; ν_1 corresponding to the MX_2 twist and ν_2, to the MX_3 scissor. The higher frequency vibrations correspond to variations in bond length, with ν_3 and ν_4 corresponding to the symmetric and asymmetric M-X stretching frequencies respectively (Fig. 3.4). Table 3.4 shows the frequency difference between ν_3 and ν_4 decreases down the group IV metals but is large in $SiCl_4$. This decrease is due to the relative weights of the halide and metal atoms; when the halide is larger than the metal the asymmetric stretch frequency is increased relative to the symmetric stretch as shown in $SiCl_4$, $SiBr_4$ and $TiBr_4$, conversely in the case of $HfCl_4$ these frequencies are almost degenerate.

The classical vibrational frequency is related to the bond strength via Hooke's Law:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{3.3}$$

Where k is the force constant, which correlates with bond strength, and μ is the

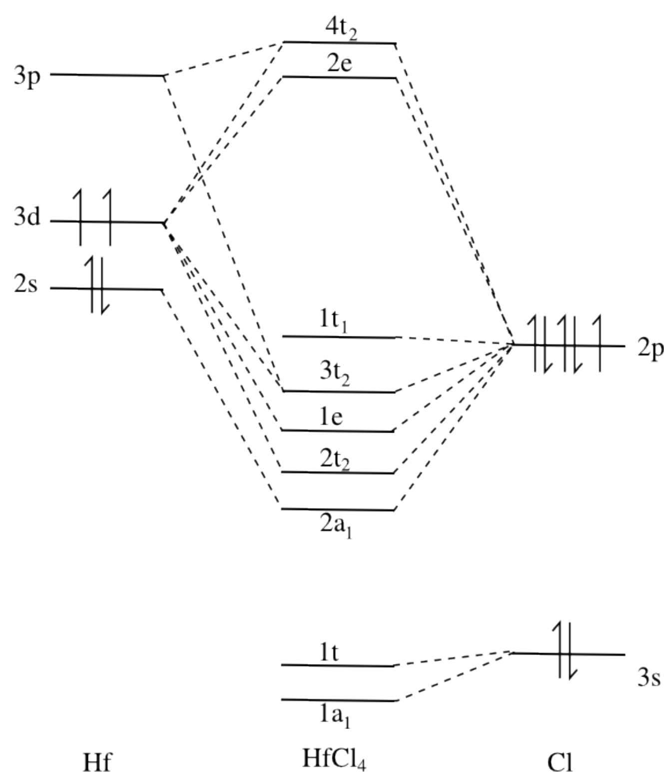


Figure 3.5: Qualitative molecular orbital diagram for $HfCl_4$

effective mass of the atoms. The symmetric stretch can therefore be used to compare relative bond strengths. The Si–X bond is stronger than the Hf–X bond due to a stronger polarization. The Hf–X and Zr–X bonds are calculated to be almost identical in strength, this is to be expected since Hf and Zr are isoelectronic and exhibit very similar properties as mentioned in section 3.1.

3.2.5 Electronic Properties of $HfCl_4$

The molecular orbital diagram in Fig. 3.5 shows the possible atomic contributions to the $HfCl_4$ MOs. The HOMO was found to be the non-bonding lone pairs of the chlorine atoms. The $1a_1$ and 1t orbitals are the chlorine 3s orbitals. The $2a_1$ and $2t_2$ orbitals are associated with the Hf–Cl σ bonding and the 1e and $3t_2$ are associated with the Hf–Cl π bonds. A previous DFT study found a similar sequence of MOs which is confirmed by the HeI photoelectron spectrum of $HfCl_4$.[92]

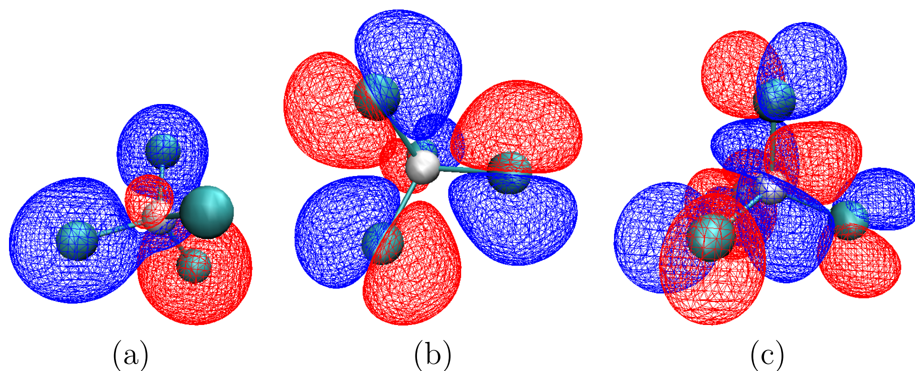
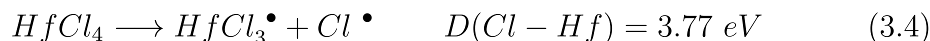


Figure 3.6: MCl_4 (M = Si, Hf) molecular orbitals (MOs): (a) Polarized bonding MO, (b) HOMO showing the Cl 2p lone pairs (c) antibonding LUMO.

3.2.6 $HfCl_4$ Cl Dissociation

The dissociation energy of chlorine atoms from $HfCl_4$ was calculated using the unrestricted HF energies of $HfCl_4$, $HfCl_3$, Cl_2, OH and $HfCl_3OH$ molecules in their lowest energy configurations. The calculated dissociation reaction energies are shown below:



$$HfCl_4 \longrightarrow HfCl_3^{\bullet} + Cl^{\bullet} \qquad D(Cl - Hf) = 3.77\ eV \tag{3.4}$$

$$HfCl_3OH \longrightarrow HfCl_3^{\bullet} + OH^{\bullet} \qquad D(HO - Hf) = 4.07\ eV \tag{3.5}$$

The dissociation energy for a chlorine atom from $HfCl_4$ was found to be 3.77 eV. The $HfCl_3$ radical formed a trigonal planar C_{3v} structure and Mulliken analysis of the spin density showed the radical electron sitting on the central cation. $HfCl_3OH$ was found to form a tetrahedral C_s structure however Mulliken analysis showed the dipole formed by the addition of the OH group made little difference to the electronic distribution around the Hf and Cl atoms relative to the electronic distribution of $HfCl_4$.

The overall reaction energy of $HfCl_4$ involving the dissociation of Cl and association of OH, is -0.3 eV, however, the model used here to calculate this reaction is not realistic. Firstly, all calculations were carried out in the gas phase at 0 K. The $HfCl_3$ molecule is only identified when it is adsorbed onto a surface, such as in the ALD mechanisms described above, even then it is only an intermediate to forming a siloxane bridge and a chemisorped $HfCl_2$ species, and so a gas phase dissociation is unlikely

to occur in nature.

3.2.7 $HfCl_4$ Summary

The structural, vibrational, and electronic properties of $HfCl_4$ were investigated using gas quantum mechanical calculations at the Hartree-Fock level, utilising a variety of basis sets. The more extensive basis sets tend to describe the electronic orbitals most accurately and polarization functions are particularly important to describe the highly polarized nature of the Hf–Cl bond. Therefore it is suggested that the 6-311G* basis set and LANL2TZ pseudopotential be used for reasonable accuracy in modelling these molecules at a reasonable computational cost.

However, it must be noted that these calculations have been limited to the Hartree-Fock level and a greater increase in calculation accuracy can be achieved using post-HF level methods such as MP2, configuration interaction or density functional theory, although these same basis sets may give more accurate results using these methods, they may also lead to unreasonable computational expense.

3.3 Bulk HfO_2

The properties of bulk m-HfO_2 are calculated using VASP, and a variety of functionals, and compared with previous experimental measurements and calculations. Firstly, the structural and electronic properties were calculated. These calculations were carried out using a 12 atom unit cell. The hybrid functionals used a $3\times3\times3$ k-point mesh and the other calculations used a $7\times7\times7$ k-point mesh. In all VASP calculations a plane wave cut-off energy of 400 eV was used.

3.3.1 HfO_2 Structure

The calculated lattice parameters, shown in Table 3.5, are in good agreement with the experimental values, overestimating by less than 1% in all cases. Previous studies have found the LDA functional to underestimate the lattice parameters and GGA func-

	a/ Å	b/Å	c/ Å	β	Band Gap/ eV
Experiment[93]	5.117	5.175	5.291	99.22	5.80[94]
LDA[95]	5.106	5.165	5.281	97.92	3.95[96]
PW91[97]	5.128	5.191	5.297	99.71	3.80
LDA	5.057	5.192	5.196	98.91	3.82
PW91	5.136	5.193	5.317	99.63	3.97
PBE	5.094	5.174	5.266	99.49	4.13
PBEsol	5.094	5.174	5.266	99.50	4.13
Hartree-Fock	5.123	5.197	5.228	98.93	14.79
PBE0	5.013	5.178	5.149	98.51	6.99
HSE03	5.026	5.174	5.154	98.65	5.80
HSE06	5.021	5.180	5.176	98.85	6.15
B3LYP	5.090	5.178	5.243	99.25	6.38

Table 3.5: Optimised structural parameters and band gaps of bulk monoclinic HfO_2.

tionals to overestimate. Similarly the band gap is characteristically underestimated by all the semi-local DFT methods; with PW91 calculation underestimating by 32%. The PW91 band gap calculated by Jaffe *et al.*, used ultrasoft pseudopotentials unlike the PAW pseudopotentials used here.[97]

The PBEsol[98] functional gave exactly the same results as the PBE calculations. This in line with the findings of Gabor *et al.* who found that the PBEsol functional made very little improvement for bulk material characteristics other than in alkali metals and alkali halides.

Hartree-Fock, due to the lack of electron correlation, overestimates the band gap hugely. The hybrid functionals, HSE03,[48–50] HSE06,[99] PBE0,[100] and B3LYP[101,102] incorporate a small amount of exchange, calculated using HF, to correct the GGA band gap.

The HfO_2 total densities of states (DoS) along with the projections of the Hf and O states, is shown in Fig. 3.7. This agrees excellently with that calculated by Musgrave *et al.*.[103] The HfO_2 DoS is split into 4 distinctive bands. The deepest band corresponds to the core-like Hafnia f-states, confirming that the f-states are unlikely

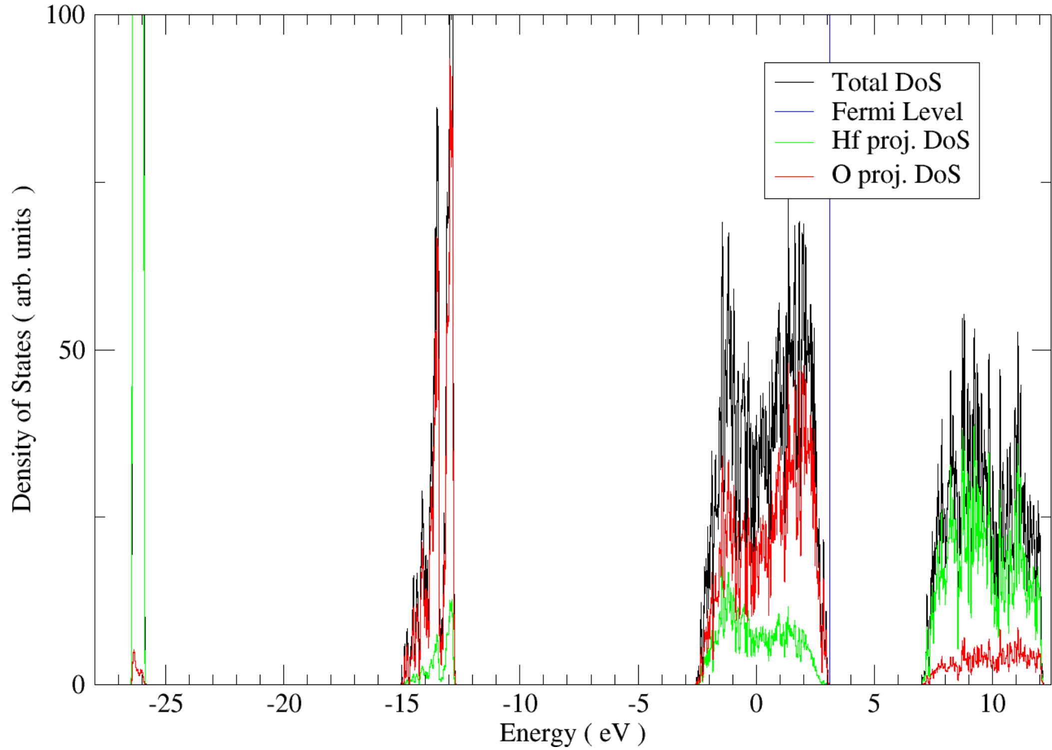


Figure 3.7: The total and projected densities of states of m-HfO_2 calculated with HSE06.

to be involved in any bonding. The next deepest band corresponds to the core 2s states of the O atoms. The 6 eV wide valence band was shown by Musgrave *el al.* to be comprised mostly of oxygen 2p states but with some contribution from the 5d states of Hf. This mixing of states suggests some covalency in the Hf-O bonds. Finally, the conduction band is composed of the Hf 5d states.

3.3.2 Dielectric Properties of HfO_2

A dielectric material is an insulator which can be polarized when an electric field is applied. When placed into an electric field the internal structure of the material becomes polarized, with the positive ions shifting towards the field and the negative ions shifting in the opposite direction. This creates a charge difference across the whole material without the conductance of electrons.

The dielectric constant, k, also known as relative permittivity, of a material indicates how easily it is polarized by an applied electric field. k is a dimensionless ratio between the complex frequency-dependent absolute permittivity of the material $\epsilon(\omega)$,

Functional	k_{xx}	k_{xy}	k_{xz}	k_{yy}	k_{yz}	k_{zz}	k
LDA	15.39	0.00	1.29	14.00	0.00	10.47	13.29
PW91	14.76	0.00	1.27	13.32	0.00	10.14	12.74
PBE	14.81	0.00	1.25	13.35	0.00	10.17	12.78

Table 3.6: Dielectric tensors of m-HfO_2. $k = (k_{xx} + k_{yy} + k_{zz})/3$.

and the permittivity of free space ϵ_0:

$$k = \frac{\epsilon(\omega)}{\epsilon_0}. \tag{3.6}$$

k is related to the electric susceptibility, χ_e, by:

$$k = 1 + \chi_e. \tag{3.7}$$

The electric susceptibility is defined as the constant of proportionality which relates the dielectric polarizability, P, induced by an electric field, E:

$$P = \epsilon_0 \chi_e E. \tag{3.8}$$

As mentioned in section 3.1.2 the dielectric constant has been measured in crystalline films of HfO_2 to be $\sim$23. The dielectric constant has also been calculated in the three crystal phases of HfO_2. Due to the anisotropy of the tetragonal and monoclinic phases, the dielectric constant is a 2nd rank tensor. This tensor can be calculated using density functional perturbation theory (DFPT) in which the frequency response of phonons in the lattice, as a result of perturbation, are calculated.[104]

The dielectric tensors of m-HfO_2 were calculated in 12 atom unit cells using a 7×7×7 k-point mesh. Previous papers have calculated the dielectric tensor in various HfO_2 phases. The one most closely resembling the experimental value is the cubic phase. The monoclinic phase is significantly lower and anisotropic.[95,105]

The dielectric constant is related to the longitudinal optical (LO) phonons and scales as Z^{*2}/ω_λ^2 where Z^* is a mode effective charge tensor which is related to the

polarizability and ω_λ is the frequency of the LO phonon. Therefore, a high dielectic constant comes from a highly polarizable low frequency phonon.[95] Zhao *et al.* found that in HfO_2 there is an intense LO phonon in the mid IR range at 286, 117 and 384 cm^{-1} for the cubic, tetragonal and monoclinic phases respectively, and calculated dielectric constants of 29, 70, and 16 for c-HfO_2, t-HfO_2, and m-HfO_2 respectively.

The dielectric constant calculated here is lower than those previously calculated, however, ultrasoft pseudopotentials were used in those calculations. The measured experimental dielectric is therefore attributed to the presence of metastable t-HfO_2 polycrystalline m-HfO_2 film.

Cockayne *et al.* calculated k of defective m-HfO_2, finding it was increased by 1-2% with neutral vacancies and suppressed by 1-3% by positively charged vacancies.[106] The increase arises from an increase in the dielectric response due to the electron pair occupying the easily polarized vacancy state. The decrease is attributed to phonon hardening which reduces the dielectric response.

3.3.3 Polarons in HfO_2

Polarons are quasiparticles formed when a charge carrier (electron or hole) enters a solid and interacts with the crystal lattice. The ions surrounding the polaron adjust their positions slightly, either attracted towards, or repulsed by, the polaron depending on their relative charge states. This change in position can lead to a local polarization in the lattice centred on the charge carrier.

Polarons can act like particles travelling through the lattice however, it is possible that they can be trapped. These trapping states may occur as pre existing defect states, or the polaron may self-trap in a perfect lattice if a potential well is created by the lattice-polarization induced by the charge carrier.

In order to explain defect formation mechanisms, during the ReRAM electroforming process, it was necessary to consider how injected electrons might localise within bulk HfO_2. Muñoz Ramo *et al.* calculated the self-trapped electron polaron in bulk HfO_2 using B3LYP.[107] This polaron state was found to be 0.32 eV below the CB, and

located on three Hf ions and a single O ion.

A repeat of this calculation was attempted using HSE06. It was found that any electrons added to the system would not self-trap, even when distortions were added to the lattice before optimization, in order to encourage trapping at a preferable site. Instead any electrons added to the system remained at the bottom of the HfO_2 CB and delocalized across the Hf ions. As shown in Table 3.5 HSE06 has a smaller band gap than B3LYP by 0.23 eV. Since the polaron state was so close to the conduction band, the narrower band gap may mean that the polaron state is covered over by the CB in the HSE06 calculation.

As stated in section 3.1.2, the measured band gap of m-HfO_2 is around 5.7 eV. Since both these functionals seem to overestimate the band gap, and the polaron is only found in B3LYP, it is unlikely to be an active defect. Also this state is so shallow that, at any practical operating temperature, it is unlikely that a polaron will trap there for any meaningful period of time.

3.4 Bulk Hafnia Summary

The properties of HfO_2 were successfully calculated using DFT and a variety of functionals. In this system GGA calculations ran quickly and produced good structural parameters, however, the band gaps were characteristically underestimated by over 30%. HSE06 functionals were found to run in a reasonable amount of time and produced excellent structural properties and electronic properties.

The dielectric constant of m-HfO_2 was also calculated using LDA and GGA functionals. k was calculated to be significantly lower than experimental values but in good agreement with previous theoretical calculations. This difference is attributed to the polycrystalline nature of the experimental HfO_2 films which are likely to include both monoclinic and tetragonal phases, the tetragonal phase having a much higher dielectric constant, raising the overall measured value.

HSE06 calculations were attempted to find electron polaron, however, the electron

did not localise in these calculations. This is due to the previously calculated B3LYP polaron state being very close to the conduction band. With the smaller band gap of HSE06, this shallow polaron state is covered up by the conduction band, so any electrons added to the system delocalise within the CB.

4

Oxygen Interstitials in HfO_2

The electrical characteristics of HfO_2 depend strongly on the composition of the oxide. The properties of HfO_2 based transistors, such as threshold voltage shift, gate leakage, bias temperature instabilities, dielectric breakdown etc., are particularly affected by oxygen defects.[108–112] The switching properties of HfO_2 MIM ReRAM devices depends on the composition of metal-rich conductive filaments (CF) in the oxide. These CFs are thought to be made up of oxygen deficient regions which form by the expulsion of oxygen ions from a specific region within the insulator layer.[113,114] Similar processes have been assumed in Lu_2O_3,[115] SiO_2[22,116] and TiO_2,[117] ReRAM cells.

The HfO_2 films for use in ReRAM cells are designed to be oxygen deficient before the electroforming process. Oxygen scavenging layers can be put into the MIM stacks to remove oxygen from the insulator and create more vacancies. Similar processes are used to reduce the thickness of the oxide layer in the transistor fabrication process. In this process, the stacks are annealed at high temperatures, during which the O ions migrate from the HfO_2 layer, into the scavenging layer.[118] This migration process involves the motion of both oxygen interstitials and oxygen vacancies.[119] Point defects can mediate the rate of oxygen diffusion across the HfO_2 layer and so may control the rates of interface layer growth or the formation and destruction of CFs through the

oxide.

Foster *et al.* calculated the barriers to interstitial oxygen (O_{int}) diffusion to be 0.8, 0.3, and 0.6 eV for the neutral, 1-, and 2- cases respectively. Given that the negative vacancy is not very stable, and the neutral vacancy is not very mobile in HfO_2,[120,121] at a high electron chemical potential, interstitial ions are likely to be the most mobile defect.

In this chapter, the results of previous calculations of O atom and ion diffusion in m-HfO_2 are re-evaluated using the GGA PW91 density functional.[40] In the calculations presented here, a larger 325 atom unit cell is used and the possible directions for O_{int} diffusion are extended with respect to ref. 40. Oxygen incorporation energies are also calculated using the hybrid HSE06 functional in a 96 atom periodic cell.

4.1 Methods of Calculation

These calculations were carried out using VASP and the Perdew-Wang 91 (PW91) GGA functional.[122,123] To optimize the geometries of the defect systems, 320 – 324 atom, 3×3×3, supercells of m-HfO_2 were used. All calculations were carried out in the gamma point and geometries were optimized using the conjugate-gradient algorithm until the maximum residual forces were less than 0.06 eVÅ^{-1}.

As shown in section 3.3 the hybrid functional HSE06[99] gives a much more accurate band gap (5.6 eV), and was used here to interstitial incorporation energies in HfO_2. Due to computational constraints, a 94 – 96 atom unit cell was used in these calculations.

A Lany-Zunger (LZ) correction based on an anisotropic point charge was applied to remove the spurious interactions between the charged defects and their periodic images.[124,125] This correction depends on the dielectric constant of HfO_2 and the size of the cell and was calculated to be 0.02 eV in this system.

Defect formation energies, $E_{for}(D)$ were calculated using, equation 2.43 and the oxygen references calculated in section 4.2 using the same functionals and potentials.

Functional	PP	E (O_2)	$\frac{1}{2}$E (O_2)
PW91	O	-9.77	-4.89
PBE	O	-9.86	-4.93
PBEsol	O	-9.06	-4.53
PBE0	O	-39.97	-19.99
HSE06	O	-27.53	-13.77
PW91	O soft	-8.99	-4.50
PBE	O soft	-9.35	-4.67
PBEsol	O soft	-9.09	-4.54
PBE0	O soft	-36.92	-18.46
HSE06	O soft	-25.65	-12.82

Table 4.1: Total energy of the O_2 molecule in eV, calculated in a large, triclinic unit cell.

4.2 Reference O_2 Molecule

In order to calculate the formation energy of a defect, using equation 2.43, the energy of an O_2 molecule needs to be calculated. It is important that the methods used to calculate this molecule are the same as those used to calculate O defects in the HfO_2 lattice for which the molecule is a reference. There are a number of functionals available in VASP along with two different oxygen pseudopotentials; standard and soft. Here, five were functionals were used and the O energies calculated here will be used as references throughout this thesis.

To properly describe the triplet ground state of O_2, the calculations were carried out using spin polarised calculations. The calculations are also carried out in a large triclinic cell in order to prevent interactions between periodic images and to break the symmetry of the charge density. The results of these calculations are shown in Table 4.1

Table 4.1 shows that the energy of the O_2 molecule varies widely depending on the functional which is used. Soft potentials gave higher energies due to the lower plane wave cut-off energy. The oxygen references used in all the following calculations use the same functionals as those used to calculate the O atoms in the calculation.

	Charge	PW91[126] $E_{for,D}$	d_{O-O}	HSE $E_{for,D}$	d_{O-O}
3C	0	1.69	1.50	2.00	1.48
3C	1-	- 0.79	2.05	- 0.81	2.01
3C	2-	- 4.16	2.35	- 4.55	2.35
4C	0	1.69	1.50	2.76	1.41

Table 4.2: Formation energies in eV and nearest neighbour O–O distances of interstitial O ions in m-HfO_2 calculated using the GGA and hybrid density functionals, in a 96 atom periodic cell. The energies are referenced to an ideal lattice cell of the same size and charge. Energies are in eV and O–O distances are in Å.

4.3 Oxygen Interstitial Diffusion

The formation energies and O–O nearest neighbour (NN) distances of interstitial O ions in m-HfO_2 are shown in Table 4.2. Three charge states of oxygen interstitial can be found within the 3C oxygen sublattice.[126] The O – O NN distance increases depending on the charge state of the interstitial ion, In the neutral case, the oxygen ions are close, with an O–O distance of 1.5 Å; this is about 0.3 Å longer than the equilibrium bond length of a free O_2 molecule. At the high Fermi level position assumed here, the interstitial oxygen can trap one or two extra electrons and moves further away from the lattice oxygen as the bond order is reduced. The negative formation energies for negatively charged O ions reflect the energy gain due to O atom electron affinity in the electrostatic potential at the O lattice site. In the 2- case, the interstitial oxygen becomes a closed shell ion and has the same Mulliken charge as the lattice oxygens, so separates until it is around 2.4 Å away from both of its oxygen NNs. This O–O distance can be used as a good indicator of the charge state of the interstitial ion.

Within the 4C oxygen sublattice, only the neutral oxygen interstitial is stable and remains in its interstitial position after geometry optimisation. This neutral 4C interstitial oxygen, like in the 3C case, has a NN distance of about 1.5 Å. The 4C O^{-}_{int} and O^{2-}_{int} are unstable, and any attempts to optimise their structures resulted

in the oxygen ions moving into 3C charged interstitial positions. This is due to the closer packing of the O ions in the 4C sublattice. This means that when the electron localizes on the interstitial oxygen ion, and the O–O bond lengthens, there is not sufficient space for the larger interstitial ion to relax into, within that sublattice.

An interstitial defect state is located 0.1 eV above the valence band of HfO_2. This low lying state means that any electrons added to the system will localise on the interstitial oxygen ions, to form a negatively charged interstitial.

The diffusion barriers of O^0_{int}, O^-_{int} and O^{2-}_{int} were calculated using GGA in the 325 atom periodic cell, and the nudged elastic band method. The barrier energies are calculated as 0.87, 0.37 and 0.51 eV for the O^0_{int}, O^-_{int} and O^{2-}_{int}, respectively. Despite using a larger unit cell, the results are in excellent agreement with those mentioned above by Foster *et al.*, and follow the same interstitialcy mechanism.[40] To check the effect of using a hybrid functional, O^{2-}_{int} diffusion within the 3C sublattice was also calculated using HSE06 where we find a very similar barrier energy of 0.49 eV.

The structure of HfO_2 during this diffusion process is shown in figure 4.1. In each case the barrier point occurs when the interstitial oxygen is almost equidistant between the lattice oxygen ions in the diffusion pathway. The neutral interstitial has the highest barrier for diffusion and involves the breaking and remaking of the O–O bond. In the 1- case, the electron is fully localized onto the interstitial ion. The charged interstitial can then move more independently from its oxygen nearest neighbour, and the O–O bond is lengthened.

In the 2- case, the interstitial ion maintains an even distance of around 2.4 Å with the nearest oxygen neighbours throughout the diffusion process. Due to the larger Coulomb interaction introduced by the second electron, the lattice displacements around the ion are slightly larger than in the singly charged case, which accounts for the 0.1 eV rise in the barrier energy, with respect to the 1- case.

The diffusion barrier for O^0_{int} moving between the 3C and 4C sublattices was also calculated to be 1.43 eV. The neutral oxygen interstitial follows a similar pathway to the 3C case above, however, the distortions through the closer packed 4C sublattice

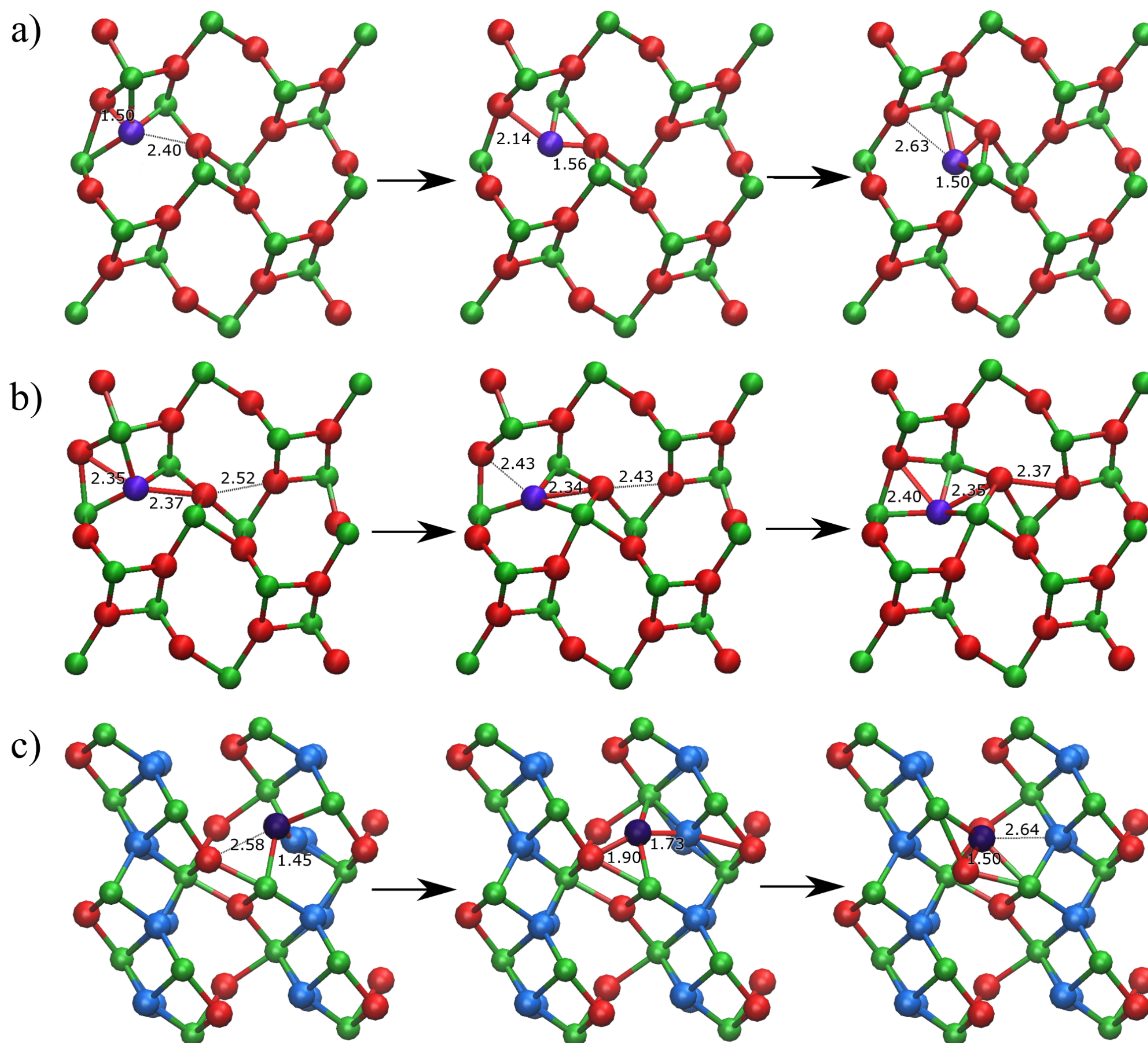


Figure 4.1: Oxygen interstitial diffusion pathways from the starting configuration (left) to the final configuration (right) via the barrier point (centre). a) O^0_{int} within the 3C sublattice ([100] view) b) O^{2-}_{int} within the 3C sublattice ([100] view) c) O^0_{int} between the 4C and 3C sublattice ([001] view). (green balls: Hf, red balls: 3-coordinated O, blue balls: 4-coordinated O, purple ball: oxygen interstitial).

are larger.

When a high bias is applied during the ReRAM electroforming process, not only is there a high density of electrons in the system, but also an electric field across the film. Under these conditions the oxygen interstitial ions will trap electrons and become charged and mobile, diffusing quickly through the oxide towards the anode. It is possible, therefore, that these defects are the primary way in which oxygen ions can migrate through m-HfO_2, and that a concerted diffusion of these ions from a region may lead to a relative O deficiency and CF formation.

4.4 Summary

The properties of the oxygen interstitial in m-HfO_2 was calculated using PW91 and HSE. Three charge states were found to be stable within the 3C O sublattice with the O–O NN distance increasing depending on the charge state of the interstitial. Only O^0_{int} was found to be stable within the 4C sublattice.

The barriers to diffusion were also calculated and it is shown that the negatively charged interstitials are very mobile in comparison HfO_2 oxygen vacancies, with a 0.4 eV barrier to diffusion for O^{1-}_{int} along the 3C sublattice. These results suggest that these defects could act as deep electron traps and, having trapped electrons, these negatively charged interstitials would be very mobile within the 3C sublattice.

5

Vacancies in HfO_2

5.1 Introduction

As mentioned in section 1.4.4 conductive filaments are thought to be formed of O deficient regions. The formation of these oxygen deficient regions is thought to begin at grain boundaries (GB) in poly-crystalline HfO_2. The role of the GB in the CF formation process has been well studied.[24,127] McKenna *et al.* found that oxygen vacancies are significantly more stable close to a GB than in the bulk oxide, suggesting mobile vacancies will tend to aggregate there.[39] This aggregation of vacancies is thought to be the initial driving force behind CF formation and is followed by generation of oxygen vacancies in the surrounding crystal.[29,70,128] Conductive-AFM measurements of the electron conductance through HfO_2 have shown that switchable conductive paths begin in small regions centred on GBs and then extend outwards forming conductive regions up to 20 nm in diameter (Fig. 5.1).[6,70,114,129] The size of the conductive region implies large oxygen vacancy aggregates must form within the crystalline part of the grain, however, the atomic scale mechanisms behind the formation of such an aggregate in HfO_2 is unclear.

In this chapter atomic scale modelling is used to look at the properties of oxygen

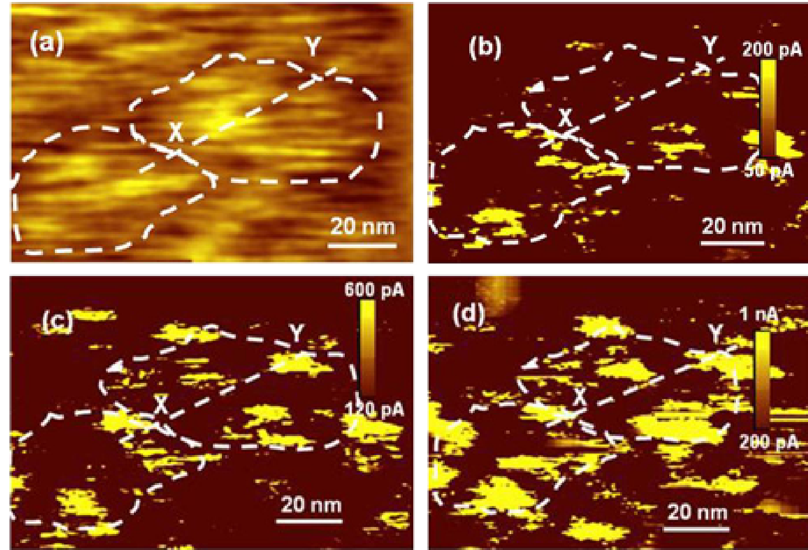


Figure 5.1: (a) Conductive-AFM topography, and leakage current profile at +6V after the (b) 1st scan, (c) 2nd scan, (d) 3rd scan of a HfO_2/SiO_x sample.[6]

vacancies and vacancy aggregates in m-HfO_2. In section 5.2, neutral and charged oxygen vacancies are considered. In section 5.3, the structure and properties of oxygen vacancy aggregates, from 2 to 4 vacancies in size, are evaluated to establish how feasible a large scale aggregate in crystalline HfO_2 is.

These calculations are carried out using the same methods as described in section 4.1.

5.2 Single Vacancies in HfO_2

Fig. 3.1 shows that HfO_2 contains a sublattice of 3 coordinated (3C) and 4 coordinated (4C) O atoms in alternating layers. The formation energy of the 3C and 4C O vacancies were calculated using PW91 and HSE and are shown in Table 5.1.

The 4C vacancies are more stable in the negatively charged and neutral states whereas the 3C vacancies are more stable in the positively charged states. This difference in stability is explained by the motions of the surrounding Hf ions. When an electron is removed from the system the surrounding Hf ions displace away from the positively charged vacancy site by 0.1 - 0.2 Å, when electron is added the Hf ions displace towards the vacancy by a similar amount. Since the partial charge at the 4C

Defect	$E_{for,D}$ PW91	$E_{for,D}$ HSE
V_3^0	6.62	6.64
V_3^{1-}	6.51	6.60
V_3^{2-}	6.29	6.58
V_3^{1+}	3.03	2.81
V_3^{2+}	-2.90	-0.84
V_4^0	6.50	6.52
V_4^{1-}	6.29	6.20
V_4^{2-}	6.04	5.89
V_4^{1+}	3.36	3.15
V_4^{2+}	-2.22	-0.20

Table 5.1: Formation energies, in eV, of O vacancies in m-HfO_2.

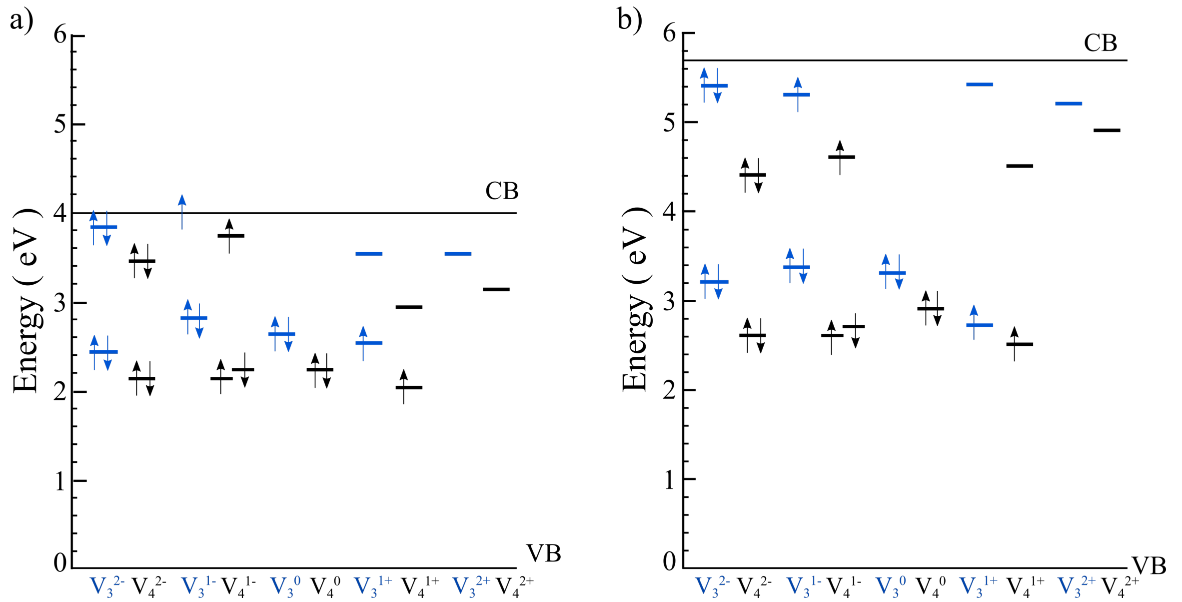


Figure 5.2: The Kohn-Sham energy levels for neutral and charged 3C (blue) and 4C (black) oxygen vacancies. a) PW91, b) HSE.

vacancy site is more positive than at the 3C site, the negative vacancy is more stable and the positive vacancy is less stable.

This is in agreement with previous studies.[40,120] Broqvist *et al.* calculated the formation energy of the oxygen vacancy with respect to the electron chemical potential, μ. They show that for increasing μ the most stable states are V_3^{2+}, V_3^{1+}, V_4^0, V_4^{1-}, with transition energies at $E_{+2/+1}$ = 3.7 eV, $E_{+1/0}$ = 4.0 eV, and $E_0/1$ = 5.5 eV.[120] This implies that during an electroforming process, where a high bias is applied across the HfO_2 layer and μ is high, neutral and singly negatively charged vacancies will be the most stable.

The Kohn-Sham energy levels of the vacancies are shown in Fig. 5.2. The levels of the 4C vacancies are 0.3 – 0.4 eV lower than 3C vacancies for the PW91 calculations and 0.3 – 0.7 eV lower in the HSE calculations. The HSE states are all higher in energy than the relevant PW91 states. States in the middle of the band gap increase by at least 0.6 eV, while those closer to the CBM move up by around 2.5 eV, in line with the 2.7 eV change in the band gap. This result compares well with previous PBE0 and B3LYP results in which the states close to the CB change in line with the change in the band gap and mid-gap states change very little with the different functionals.[120,130]

Capron *et al.* calculated the activation energies for the diffusion of oxygen vacancies in m-HfO_2. For V^0 the diffusion pathway was found to be almost isotropic with $\approx$ 2.4 eV activation energy for long-range diffusion. For V^{2+} the activation energy is 0.69 eV, and corresponds to the vacancy travelling down one-dimensional chains in the [110] and [$1\bar{1}0$] directions. In this process the oxygen ion, moving in the opposite direction to the vacancy, maintains two links to neighbouring Hf ions as it changes position.

5.3 Oxygen Vacancy Aggregates in HfO_2

The formation energy of oxygen vacancy aggregates was calculated using equation 2.43, where n is the number of vacancies in the aggregate.

In order to compare the stability of aggregates, and calculate the extent to which the defects interact, it is useful to consider the binding energy between the vacancies. These binding energies can be calculated in two ways. The simplest way is to take the formation energy of the vacancy cluster, calculated using equation 2.43, and deduct the formation energies of the constituent defects:

$$E_{bind,A} = E_{for,A} - E_{for,V_y} - E_{for,V_x} - ... \tag{5.1}$$

Here, A is a vacancy aggregate, and V is a vacancy within the aggregate with co-ordination number y, x etc.. Since the vacancy formation energies are calculated individually in separate calculations, the binding energy is the change in energy due to vacancies being simultaneously brought from an infinite separation into the aggregate. A negative E_{bind} would indicate that the aggregate is more stable than its parts, whereas a positive E_{bind} would describe an unstable aggregate which is unlikely to form at all. This method effectively describes an aggregation process, regardless of the size of the aggregate and the way in which vacancies come together via diffusion.

A second way to calculate the binding energies is to consider a sequential process, beginning with a single vacancy then adding more vacancies one by one as the aggregate grows. Here, this is described as an attachment energy:

$$E^q_{att,D} = D^q_{x,y} - ([V_x + V_y]^q), \tag{5.2}$$

$$E^q_{att,T} = T^q_{w,x,y} - ([D_{x,y} + V_w]^q), \tag{5.3}$$

$$E^q_{att,Q} = Q^q_{v,w,x,y} - ([T_{w,x,y} + V_v]^q), \tag{5.4}$$

where E_{att} is the attachment energy of the di-vacancy, D, tri-vacancy, T, or tetra-vacancy, Q. V indicates a non-interacting vacancy site with a Hf coordination number

of v, w, x or y and with overall charge of the system n.

In this method E_{att} represents the change in energy when a distant "non-interacting" vacancy is brought close enough to interact with another vacancy or vacancy cluster. This method has some advantage over the method described by equation 5.1 in that it describes a more realistic, additive process regardless of the size of the aggregate, and the binding energies can be directly compared.

In order to calculate this additive process reference calculations were carried out in cells where V was within the same supercell but over 10 Å away from the other vacancy or vacancy cluster. In this way V can be considered structurally isolated from the other vacancies.

5.3.1 Structural Properties of Oxygen Vacancy Aggregates

Table 5.3 shows the binding energies of various multiple vacancy clusters. Each cluster is made up of a different configuration of neutral O vacancies within the m-HfO_2 lattice. These results show that neutral di-vacancies are stable with a binding energy which varies depending on the aggregate's size and shape. The neutral di-vacancy binding energies are comparable to those calculated by Zhang *et al.* despite their using the PBE functional.[131]

The calculation of attachment energies simply corresponds to a change in the reference when compared to the binding energy. Due to the finite size of the supercell the "non-interacting" vacancies are shown to be weakly interactive with binding energies of -0.06 and -0.03 eV. As such $E_{bind,A}$ is a more useful method for measuring the interaction of the vacancies in the aggregate.

The lowest energy di-vacancy structures for each of the vacancy coordination combinations is shown in Fig. 5.3. In each case the lowest energy structure is one in which the two vacancies are not screened from one another by a Hf cation and therefore, form the largest void possible within the structure. Fig. 5.4 shows that there is no correlation between the binding energies of the di-vacancies and the vacancy – vacancy separation, instead, the leading factors seem to be the vacancy configuration

Name	d_{V-V} (Å)	$E_{for,D}$ (eV)	$E_{bind,A}$ (eV)	E_{att} (eV)
V_3		6.60		
V_4		6.52		
$V_4 + V_4$	12.08	12.99	-0.06	
$V_3 + V_3$	11.96	13.18	-0.03	
$V_3 + V_4$	11.46	13.07	-0.06	
$D(A)^0_{44}$	2.70	13.03	-0.01	0.05
$D(B)^0_{44}$	2.58	13.00	-0.05	0.01
$D(C)^0_{44}$	2.54	13.03	-0.01	0.05
$D(D)^0_{44}$	2.88	12.93	-0.11	-0.06
$D(E)^0_{33}$	2.74	13.17	-0.03	0.00
$D(F)^0_{33}$	2.86	13.04	-0.17	-0.14
$D(G)^0_{33}$	2.88	13.18	-0.03	0.00
$D(H)^0_{33}$	2.71	13.04	-0.16	-0.14
$D(I)^0_{34}$	2.56	12.92	-0.21	-0.15
$D(J)^0_{34}$	2.86	12.95	-0.18	-0.12

Table 5.2: Binding energies of the calculated neutral vacancy aggregates. D, indicates a di-vacancy where each configuration is labelled by a letter in parentheses. Subscripts denote the Hf coordination of the constituent vacancies and superscripts indicate the overall charge of the unit cell.

Name	d_{V-V} (Å)	$E_{for,D}$ (eV)	$E_{bind,A}$ (eV)
$T(A)^0_{444}$	2.58	19.42	-0.15
$T(B)^0_{444}$	2.54	19.47	-0.10
$T(C)^0_{444}$	2.70	19.48	-0.09
$T(D)^0_{443}$	2.56	19.20	-0.45
$T(E)^0_{333}$	2.73	19.59	-0.22
$T(F)^0_{333}$	2.60	19.63	-0.18
$T(G)^0_{333}$	2.76	19.47	-0.34
$T(H)^0_{334}$	2.43	19.34	-0.39
$Q(A)^0_{4444}$	2.58	25.96	-0.14
$Q(B)^0_{4443}$	2.56	25.75	-0.43
$Q(B)^0_{4443}$	2.53	25.86	-0.31
$Q(B)^0_{4443}$	2.56	25.64	-0.54
$Q(E)^0_{3333}$	2.88	26.00	-0.42
$Q(F)^0_{3333}$	2.44	25.96	-0.45
$Q(G)^0_{3334}$	2.86	25.63	-0.71
$Q(H)^0_{3334}$	2.56	25.74	-0.59

Table 5.3: Binding energies of the calculated neutral vacancy aggregates. T and Q indicate tri- and tetra-vacancies respectively. Each configuration is labelled by a letter in parentheses. Subscripts denote the Hf coordination of the constituent vacancies and superscripts indicate the overall charge of the unit cell.

and coordination. The lowest energy D_{44} in particular does not consist of the closest vacancy pair. The di-vacancies containing a 3C vacancy site are lowest in energy. This is due to the closer packing of the 4C sites and so a D_{44} vacancy cannot create as big a void in the lattice.

The neutral oxygen tri-vacancies demonstrate behaviour similar to the di-vacancies and have generally small binding energies. However, all the energies are slightly more negative than those in the di-vacancies, showing that the addition of a third vacancy to the aggregate is more favourable than the second. This depends, however, on the overall structure of the aggregate as there is still a 0.36 eV variation of E_{bind} in the considered structures.

Fig. 5.5 shows the tri-vacancy structure with the most negative $E_{bind,T}$ of -0.45 eV. The vacancies in this tri-vacancy, like in the di-vacancy case, are not screened from one another by a Hf cation and form a trigonal structure and the largest void

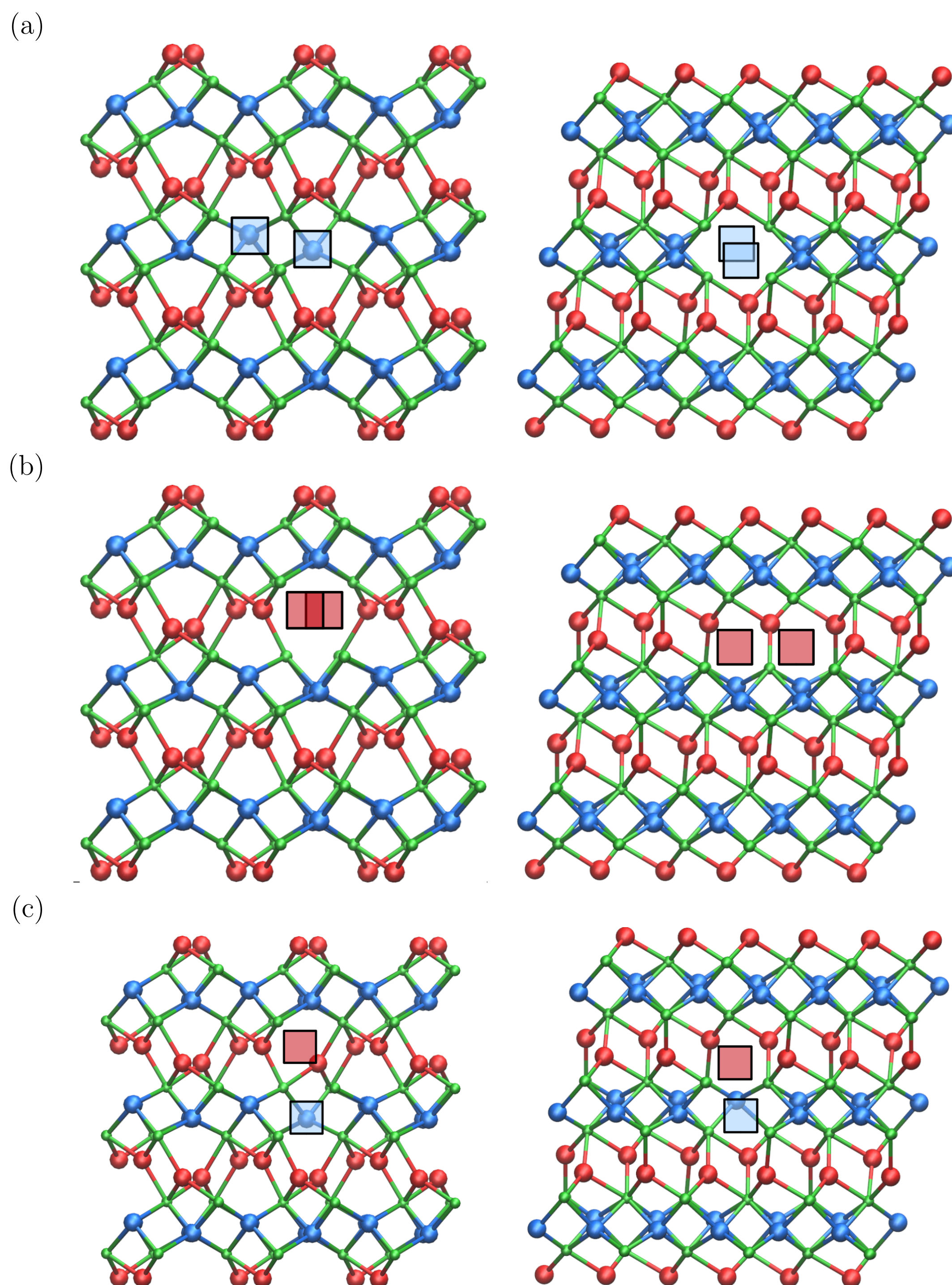


Figure 5.3: The lowest energy neutral m-HfO_2 di-vacancy structures. (a) $D(C)^0_{44}$, (b) $D(F)^0_{33}$, (c) $D(G)^0_{34}$. Left: projection view along the [001] direction, right: projection view along the [010] direction. (Green balls: Hf, red balls: 3C oxygen, blue balls: 4C oxygen, red squares: 3C oxygen vacancy site, blue squares: 4C oxygen vacancy site).

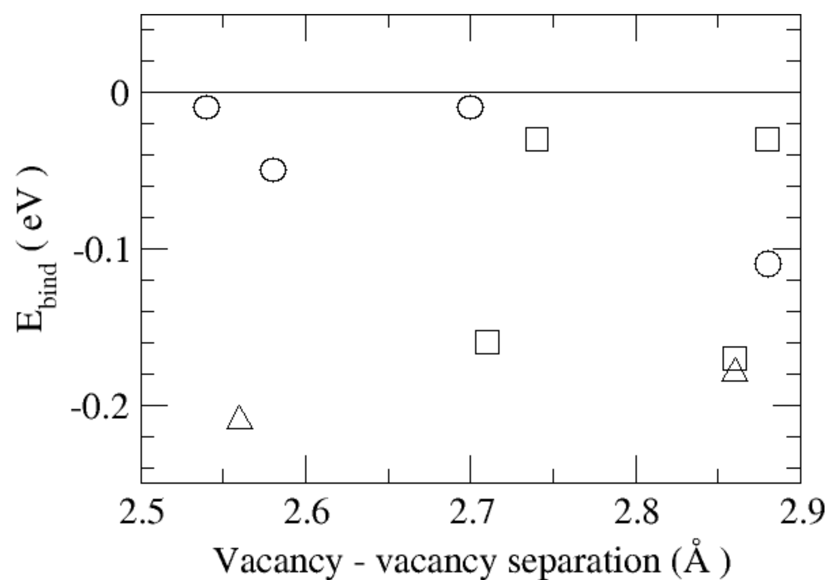


Figure 5.4: The binding energies of the calculated di-vacancies with respect to the vacancy – vacancy separation. Squares: D_{33}, circles: D_{44}, triangles: D_{34}.

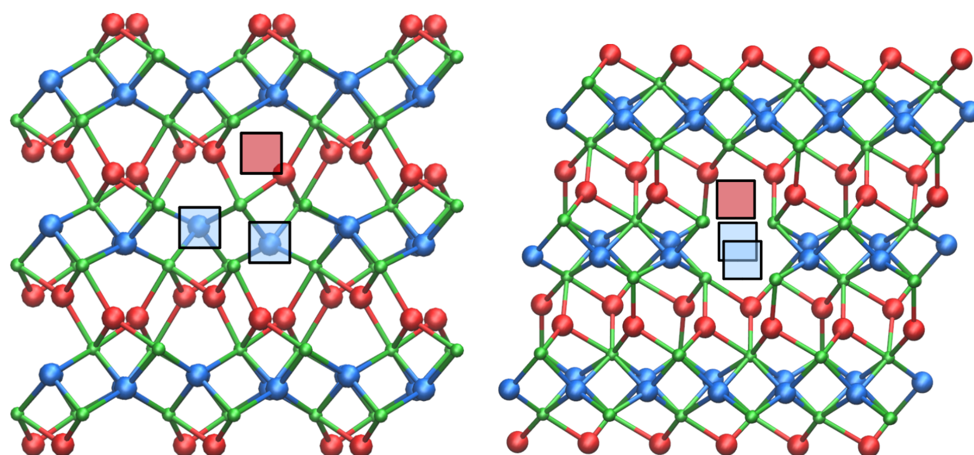


Figure 5.5: The lowest energy neutral m-HfO_2 tri-vacancy structure, $T(D)^0_{443}$. Left: projection view along the [001] direction, right: projection view along the [010] direction (coloured as in Fig. 5.3)

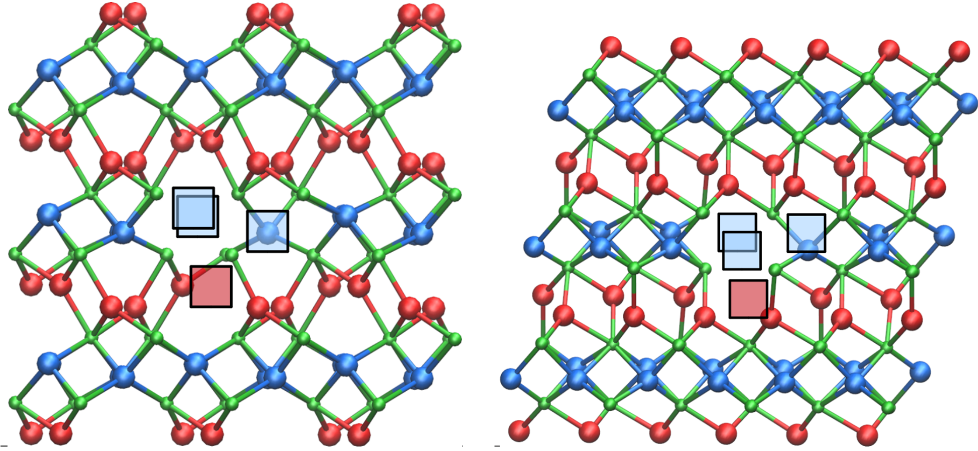


Figure 5.6: The lowest energy neutral m-HfO_2 tetra-vacancy structure, $Q(D)^0_{4443}$. Left: projection view along the [001] direction, right: projection view along the [001] direction (coloured as in Fig. 5.3).

possible within the m-HfO_2 lattice.

In the case of neutral tetra-vacancies, the trends shown by di- and tri-vacancies are further enhanced. Fig. 5.6 shows one of the lowest energy neutral m-HfO_2 tetra-vacancy structures, which has an E_{bind} of -0.54 eV. Again this configuration forms a large void and the vacancy sites are not screened from one another by Hf cations. Other favourable tetra-vacancy configurations show similar characteristics, whereby the constituent vacancies are largely unscreened from one another. Configurations forming a three-dimensional void have lower energies than those forming two-dimensional structures within a single sublattice. This suggests that, despite the layered structure of HfO_2, there is no clear directionality to how the vacancy clusters will form.

The largest binding energy per vacancy is -0.11, -0.15 and -0.18 eV for di-, tri- and tetra-vacancy aggregates, respectively. This implies that the attachment of vacancies to an existing aggregate may become more favourable as the aggregate grows. However, this is strongly dependent on the overall structure of the aggregate. It is also possible that a less favourable aggregate may re-organize into a lower energy structure allowing this energetic gain to continue.

There is little data in the literature to compare the binding energies of the HfO_2 vacancy aggregates and those calculated in other oxides. The binding energies of the

di-vacancies are, however, comparable to those of alkaline earth diatomic molecules. These diatomics have similar bonding and antibonding ordering of highest occupied molecular orbitals and a formal bond order of zero and, as a consequence, have small binding energies. In particular, they are very similar to the reported properties of Be_2 measured using laser-induced fluorescence.[132] It was found that Be_2 has an E_{bind} of -0.10 eV and bond length of 2.45 Å which is very similar binding energy and interatomic distance to the calculated HfO_2 di-vacancies.

The entropic contribution to the free energy, due to the number of ways isolated vacancies can be accomodated in the lattice, may outperform the gain in binding energy at low vacancy concentrations. In this case a stable vacancy aggregate may not form.[133] However, the concentration of oxygen vacancies in HfO_2 films in CMOS devices usually exceeds 10^{18} cm^{-3}[134] and needs to be greater than 4% for conductive filaments to form in ReRAM devices.[135] High oxygen deficiency is often induced by gettering of Ti or Hf layers inserted into the stack.[136–138]

The kinetics of neutral oxygen vacancy aggregation has been studied only in MgO.[139] This study has demonstrated that the defect interaction energy is the key factor in aggregate formation and that the vacancy binding energies as small as 0.03 eV are sufficient for aggregates to form once the vacancy diffusion activation energy (about 3.5 eV) can be overcome. In MgO stable aggregates are formed within several minutes when the vacancy concentration exceeds 10^{18} cm^{-3} and the temperature is above 1500 K.[139] Based on the calculated neutral vacancy diffusion barriers of 2.4 eV in HfO_2,[121] film temperatures of the order of 1000 K are required to accelerate the vacancy aggregation. Such high temperatures, however, are likely to lead to breaking of the loosely bound aggregates and stimulate aggregation at lower energy sinks, such as grain boundaries and dislocations.

5.3.2 Electronic Properties of Oxygen Vacancy Aggregates

The neutral di-vacancy and its electronic structure were calculated using HSE. Fig. 5.7 shows there are two occupied di-vacancy states in the middle of the band gap. The

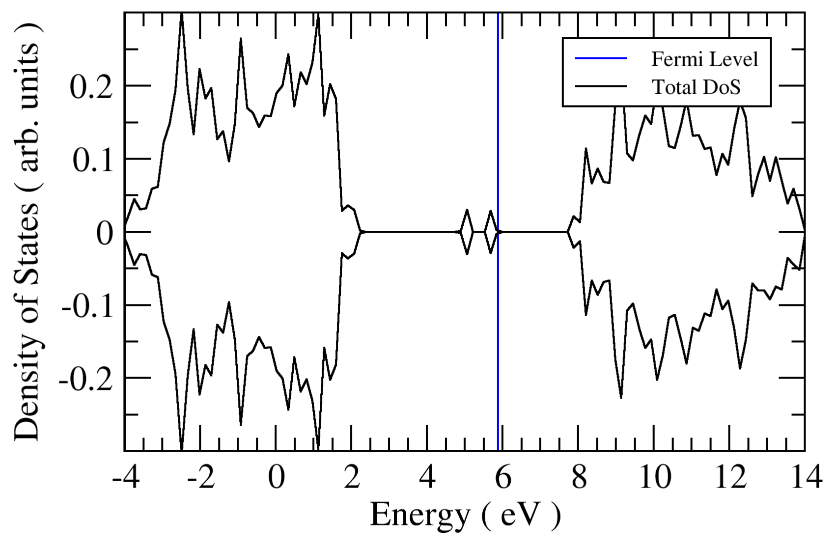


Figure 5.7: Total electronic densities of states (black) and Fermi energy (blue) of the HfO_2 di-vacancy, $D(I)^0_{34}$.

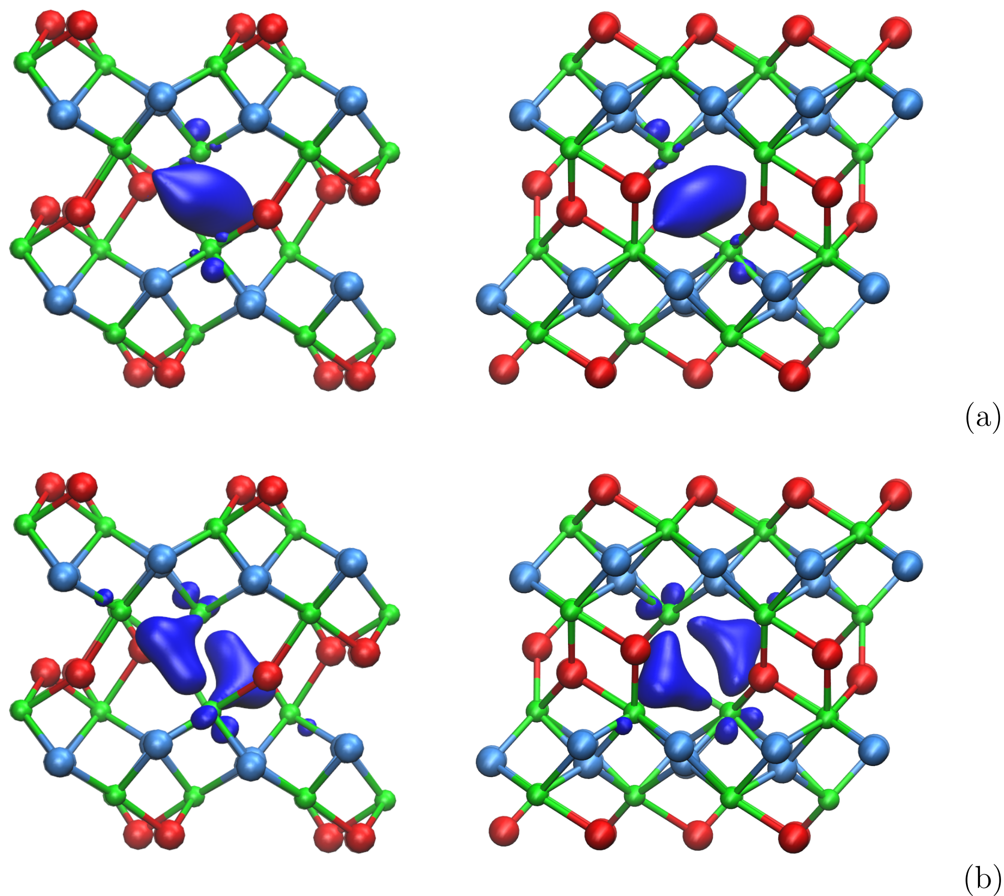


Figure 5.8: The one electron densities (dark blue) for the HfO_2 di-vacancy $D(I)^0_{33}$ (a) lower energy di-vacancy state (b) higher energy di-vacancy state.

Defect	$E_{bind,A}$			
	$2V^0$	$V^0 + V^{1-}$	$V^{1-} + V^{1-}$	$V_3^0 + V_4^{2-}$
D_{33}	-0.16	0.13	0.27	
D_{44}	0.13	0.04	-0.04	
D_{34}	0.06	0.38		-0.37

Table 5.4: Formation energies of oxygen vacancies and di-vacancies (eV) calculated using equation 5.1 and the HSE06 functional. D denotes a di-vacancy, subscripts denote the coordination of the constituent vacancies.

lower energy state is at 2.8 eV above the valence band whereas the second vacancy state appears at 3.4 eV above the valence band. Fig. 5.8 shows the isosurface of the band-decomposed partial charge density associated with these vacancy states. Both states are a combination of Hf *d*-orbitals located inside both vacancies and can be described as two di-vacancy states. The lower and higher energy states (Fig. 5.8 (a) and (b)) are bonding and antibonding, respectively, implying the di-vacancy has an overall formal bond order of zero, and this may explain why the binding energy between these di-vacancies is so small.

The neutral and charged di-vacancy were calculated using HSE and the method described by equation 5.1. The results are summarized in Table 5.4. The binding energies of the neutral di-vacancies are comparable to those calculated using PW91 and PBE, varying by around 0.1 eV.

The highest occupied states which form when the di-vacancies are negatively charged lie between the di-vacancy states and the conduction band (CB). In D^{1-} this state is singly occupied and is located 0.4 eV below the CB. For D^{2-} this state is a triplet state with two singly occupied states located 1 eV and 1.2 eV below the CB. In the lowest energy D_{34}^{2-} the deeper state is 1.5 eV below the CB. This lower state is due to the deeper well created by the surrounding positive Hf cations, this creates a region of slightly positive charge, lowering the electronic levels and the overall energy of the structure.

In the case of D_{33}^{n-} and D_{44}^{n-}, electrons localize on the Hf atoms between the two

vacancy sites, and are distributed evenly between the vacancies. In the case of D_{34}^{n-} the electrons localize onto the 4C vacancy site.

This effect is also seen in the case of the lowest energy tetra-vacancy structure. The tri- and tetra-vacancies are too large to be properly calculated in the smaller cell so only PW91 calculations, in the larger supercell, were attempted for these systems. In the PW91 calculations, the CB is artificially lowered and covers up the negatively charged vacancy states, so that any extra electrons added to these systems delocalize in the HfO_2 CB. However, in the case of the lowest energy tetra-vacancy, the highest occupied states of the negatively charged vacancy are in the band gap. For the 1- and 3- cases a doublet state forms, for the 2- and 4- cases a singlet state forms. The lower energy level descends from 0.5 eV below the CB in the 1- case to 0.7 eV below the CB in the 4- case, while a higher energy state appears for the 3- and 4- cases at 0.2 eV below the CBM. These states do not appear in tetra-vacancy configurations with higher formation energies. This is because, in these tetra-vacancies with low formation energies, a very deep well is created by the surrounding Hf cations which can trap electrons. This deep well lowers the electronic states and lowers the overall energy of the aggregate.

An attempt was made to calculate positively charged and doubly positively charged di-vacancies. However, every configuration where the vacancy – vacancy distance was within 4 Å was found to be unstable. In each case the vacancies would repel each other and spontaneously separate until they were sufficiently screened by at least one Hf cation.

5.4 Summary

To calculate the binding energies for negatively charged di-vacancies we assume that neutral O vacancies can trap one or two extra electrons and neutral and negative vacancies can diffuse in the vicinity of each other. The individual vacancies are calculated in separate cells in different charge states. When they are brought together, the

electron density is re-distributed, forming occupied bonding and anti-bonding states described above. The binding energies of negatively charged di-vacancies are then calculated as the difference in the total energy of an aggregate and individual vacancies. These are shown in Table 5.4. The binding energy varies depending on the constituent vacancies in the aggregate and the initial state at infinite separation. In the D_{33} case, the positive E_{bind} indicates a slightly repulsive interaction. The only attractive vacancies are doubly negatively charged di-vacancies containing a 4C vacancy. These binding energies are generally small indicating charged vacancies do not interact strongly with one another.

Our results demonstrate that the aggregation of neutral oxygen vacancies in *m*-HfO_2 is feasible and that vacancy aggregates, if created, will be stable at room temperature. The binding energy per vacancy becomes larger as the aggregate grows, implying the addition of a vacancy to an aggregate larger than a tetra-vacancy may be even more favourable. However, this depends on the structure of the aggregate as a whole, so aggregate reorganizations may need to occur for this process to continue. The negative di-vacancy was also shown to be reasonably stable, and the positive di-vacancy is unstable.

Within the context of a ReRAM conductive filament forming process, it is expected that the oxide will be under bias. However, high vacancy diffusion barriers in the neutral charge state and small binding energies suggest that the process of aggregation of these vacancies via thermally activated diffusion is slow and inefficient under the voltage/temperature conditions of practical interest. Positive vacancies can diffuse more efficiently and may get converted into a neutral state after merging with other already aggregated vacancies. Such a cluster can capture injected electrons very effectively, and it shares trapped electrons between all linked nearby vacancies. An alternative mechanism may involve generation of new vacancies in the vicinity of pre-existing vacancies.

The results for charged vacancies and di-vacancies show that these defects can trap electrons and remain stable. This is particularly important for trap-assisted

tunnelling and supports the idea that the CF is composed of an oxygen deficient region. Furthermore, the instability of positively charged di-vacancies may lead to a break-up of the CF if the bias induced across the insulator is reversed or the system is flooded with holes.

6

Frenkel Defect Formation in HfO_2

6.1 Introduction

As discussed in previous chapters the performance of transistors is strongly affected by electron transport through the insulating dielectric film.[140–142] This leakage through the insulator can self-accelerate and eventually lead to dielectric breakdown.[8,16,143–146] Within HfO_2 these breakdown processes are thought to be caused by trap-assisted tunnelling through pre-existing defects and newly generated defects, such as oxygen vacancies.[147–149] As discussed in section 1.4.4 the electroforming step in ReRAM operation has been well studied, and has been shown to lead to oxygen-vacancy generation.[70,114,150–152,152–154] Accumulation of these vacancies could lead to the eventual formation of a conductive filament (CF) through the dielectric, which can sustain large electron current.

The experimental and theoretical results demonstrate that CF formation in m-HfO_2 can proceed via aggregation of O vacancies at grain boundaries.[6,39,70,155] However, the calculated barriers for vacancy diffusion in the bulk of m-HfO_2 are quite high; 2.38 eV and 0.69 eV for neutral and doubly positively charged vacancies, respectively.[121] Also, as shown in section 5.2, the formation energy of 3-coordinated (3C)

and 4-coordinated (4C) neutral O vacancies is high, 6.6 eV and 6.5 eV, respectively, and the FD pair formation energy is even higher for charged defects. Furthermore, the positively charged vacancy will be unstable during an electroforming process due to the high applied bias. On the other hand, in chapter 4, it is shown that diffusion of oxygen interstitials is quite fast, with barriers on the order of 0.5 eV.[40] Therefore, the formation of Frenkel defects (FDs) – O vacancies and O interstitial atoms or ions, accompanied by O interstitial out-diffusion is a possible alternative method of CF formation which does not rely on neutral vacancy diffusion. It is often assumed that new oxygen vacancy/ion pairs are generated when a Hf–O bond is weakened as a result of electronic stress caused by a strong electric field.[150,156]

In this chapter, the possibility that FD pairs can be formed by electron injection, particularly close to pre-existing vacancies, is investigated. In MIM ReRAM structures the Fermi level is determined by the metal electrodes (TiN, Pt). At positive bias, electrons can be injected from the electrodes and into defect states via F-N tunnelling (see Fig. 1.2 b). The results in chapter 5 show that positively charged and even neutral oxygen vacancies in m-HfO_2 have large affinities to electrons from the conduction band and neutral vacancies can trap up to two extra electrons, forming negatively charged vacancies, with the electrons localized on the surrounding Hf ions.[120,157]

In this chapter the thermodynamic stability of FD pairs in m-HfO_2 under electron flooding conditions is studied. It is assumed that up to two electrons are injected into CB of HfO_2 within the simulation cell from a metal electrode. It is shown that these extra electrons can weaken Hf–O bonds and reduce the formation energies, and barriers to formation, of FD pairs.

6.2 Formation of Nearest Neighbour FD Pairs

Firstly, FD formation in perfect m-HfO_2 lattice is considered, using a neutral 324 atom periodic cell and PW91. The methods used in these calculations are consistent

n_{elec}	Defect	$E_{for,D}$	$E_{bar,f}$	$E_{bar,r}$
0	$V_3 + O_{int}$	Unstable		
1	$[V_3 + O_{int}]^-$	Unstable		
2	$V_3^0 + O_{int}^{2-}$	1.19	1.96	0.77
2	$V_4^0 + O_{int}^{2-}$	2.16	2.64	0.48
0	$2V_3 + O_{int}$	Unstable		
1	$[2V_3 + O_{int}]^-$	Unstable		
2	$[2V_3]^0 + O_{int}^{2-}$	0.96	1.25	0.29
3	$[2V_3]^- + O_{int}^{2-}$	0.71	1.23	0.52
4	$[2V_3]^{2-} + O_{int}^{2-}$	0.29	1.12	0.83

Table 6.1: Formation energies and adiabatic barriers for formation and recombination of FD pairs (eV). $E_{for,D}$ corresponds to NN FD pair formation energy, $E_{bar,f}$ corresponds to the barrier energy for FD pair formation, and $E_{bar,r}$ the barrier energy for recombination.

with those described in section 4.1.

Starting with the stable O interstitial atom structures, oxygen atoms are removed from different nearest neighbour (NN) oxygen sites to form an NN FD pair. The Fermi level position in these calculations is at the top of the HfO_2 valence band. Vacancy positions in both the 3C and 4C sublattice were considered. However, since the charged interstitial is unstable in the 4C sublattice, only interstitial positions in the 3C sublattice were considered. The results of these FD calculations are summarised in Table 6.1.

Removing an O atom creates a neutral O vacancy, which donates two electrons to interstitial O atom, creating a pair of charged defects, $[V^{2+} + O_{int}^{2-}]$. In every case the NN FD pairs recombined, restoring the perfect lattice. This recombination is caused by a strong Coulomb attraction between the two defects.

It is then assumed that the Fermi level is at the bottom of the HfO_2 conduction band and extra electrons are injected in m-HfO_2 from an electrode and propagate in the conduction band. In MgO it has been shown that thermally activated formation of FD pairs can take place as a result of thermal fluctuations and displacement of an

O ion from its site into interstitial position.[158] When one extra electron is trapped a V^{1+} vacancy is formed. However, the NN FD pair $[V^{1+} + O^{2-}_{int}]$ is still unstable and recombines.

When two extra electrons are trapped V^0 is formed and the resulting $[V^0 + O^{2-}_{int}]$ NN FD pair is stable. Therefore, there is no Coulomb attraction between the neutral vacancy and the interstitial O ion. The distance between the interstitial O and the nearest neighbour O ions is 2.45 Å, indicating that this is O^{2-}_{int}.

The formation energy for the $[V^0_3 + O^{2-}_{int}]$ pair, calculated as the energy difference between the perfect periodic cell with two extra electrons in the conduction band and the final FD pair state, is particularly low at 1.19 eV. This low energy results from the energy gained through localizing the two electrons from the HfO_2 conduction band into the positively charged oxygen vacancy. These electrons are localized on the Hf ions surrounding the vacancy, as shown in Fig. 6.1. However, the $[V^0_4 + O^{2-}_{int}]$ formation energy is 0.97 eV higher than the V^0_3 case, indicating that the structure of the defect must have a significant impact on the overall formation energy.

The density of states of the FD pair, calculated using HSE06, shows an interstitial state 0.1 eV above the valence band maximum (VBM) and a vacancy state 3.0 and 3.2 eV above the VBM for the 3C and 4C neutral vacancies, respectively. This is in good agreement with previous calculations.[120,126,159] Formation energies for 4C and 3C NN FD pairs calculated using HSE06 in a 96 atom cell are by about 0.3 eV lower than those calculated using GGA. This reflects the fact that the HfO_2 band gap calculated using HSE06 is wider and the vacancy level is lower with respect to the bottom of the conduction band.

The barriers for formation of these defect pairs were calculated using GGA and the NEB method in a 324 atom cell (see Table 6.1). The barrier for the 3C FD pair is significantly lower than for the 4C FD pair. The barriers for formation and recombination of the NN FD pair are shown in Fig. 6.2. The recombination barrier for the $[V^0_3 + O^{2-}_{int}]$ pair is about 0.8 eV, which is much higher than the O^{2-}_{int} diffusion barrier of 0.5 eV, suggesting that the defects can efficiently separate via O^{2-}_{int} diffusion.

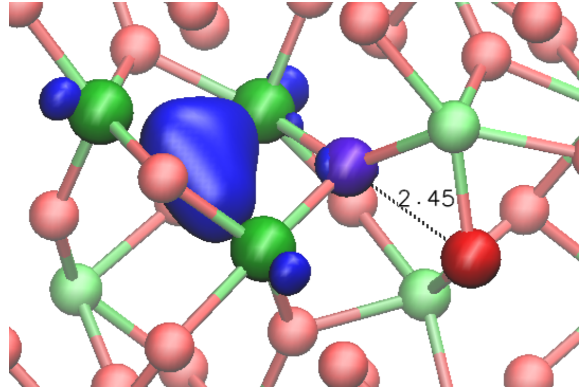


Figure 6.1: The band-decomposed partial charge density associated with the highest occupied molecular orbital in NN FD $[2V_3]^{2-} + O_{int}^{2-}$, calculated using HSE. (green balls: Hf, light red balls: O, purple ball: O_{int}^{2-}, blue: highest occupied molecular orbital).

The mechanism of formation of this FD pair corresponds to two electrons effectively assisting thermal fluctuations of O ion,[158] and pushing it out of its site into the interstitial position. The electrons then localise in the vacancy, thus gaining extra energy with respect to their energy at the bottom of the conduction band. However, the rate of such process in the perfect lattice is low as it requires a correlated action of two electrons. Therefore we consider whether a similar process could take place at a pre-existing O vacancy, which can trap two extra electrons and thus, significantly increase the cross section of such reaction.

6.3 Formation of FD Pairs Next to a Pre-existing Vacancy

In order to check whether pre-existing oxygen vacancies could facilitate the formation of new vacancies, the formation energy of a NN FD pair consisting of two O vacancies and one interstitial O ion is calculated. Beginning with a stable NN FD pair, a second vacancy was created nearby the existing vacancy. This process is outlined in Fig. 6.3 and can be described as follows.

First, we consider a neutral O vacancy in the 323 atom unit cell using the PW91 functional and add two extra electrons. These electrons are delocalized in the con-

duction band of HfO_2 and do not form a doubly negatively charged O vacancy, due to the low CB in GGA. This state is the starting point and the energetic reference for further calculations, just as the ideal lattice with two extra electrons was the reference in the NN FD pair calculations.

We then form a second O vacancy near the pre-existing one, and a NN interstitial O ion, by removing an O atom. As in the single vacancy case above, if none, or only one electron, are added to the system, an O^{2-}_{int} forms, leaving behind positively charged vacancies. The Coulomb interaction between these charged defects is so large that they recombine during the geometry optimisation. However, if two extra electrons are added, the two vacancies are neutralized and the O^{2-}_{int} ion cannot recombine, creating a stable system of two neutral O vacancies and an interstitial O^{2-}_{int} ion. The cost of forming a $[V^0_3 + O^{2-}_{int}]$ pair next to an existing neutral vacancy is only 0.96 eV. Furthermore, the barrier for the formation of this defect system was calculated to be 1.25 eV, 0.71 eV lower than that for the formation of the NN FD pair from the ideal lattice (see Table 6.1 and Fig. 6.2).

Therefore, the formation of the FD is significantly more likely to occur nearby a pre-existing oxygen vacancy, under electron injection conditions. As one can see in Table 6.1, the addition of more electrons to the system leads to creation of a charged di-vacancy and charged interstitial and further lowers the formation energy, and the barrier to formation, of this defect system.

The barrier profiles for several charge states are summarised in Fig. 6.2. One can see that the recombination barriers for charged FD pairs also increase, reflecting their repulsion. The system of two vacancies and an interstitial O ion is too large to be calculated in 96 atom cell. Therefore we were unable to perform HSE06 calculations for this system, but can assume that the defect formation energies can be lowered by at least 0.3 eV, as described above.

The results presented here demonstrate an initial trend to forming clusters of O vacancies and charged O_{int} ions at the periphery of the vacancy cluster facilitated by electron injection into the system. The formation energies of these defects are

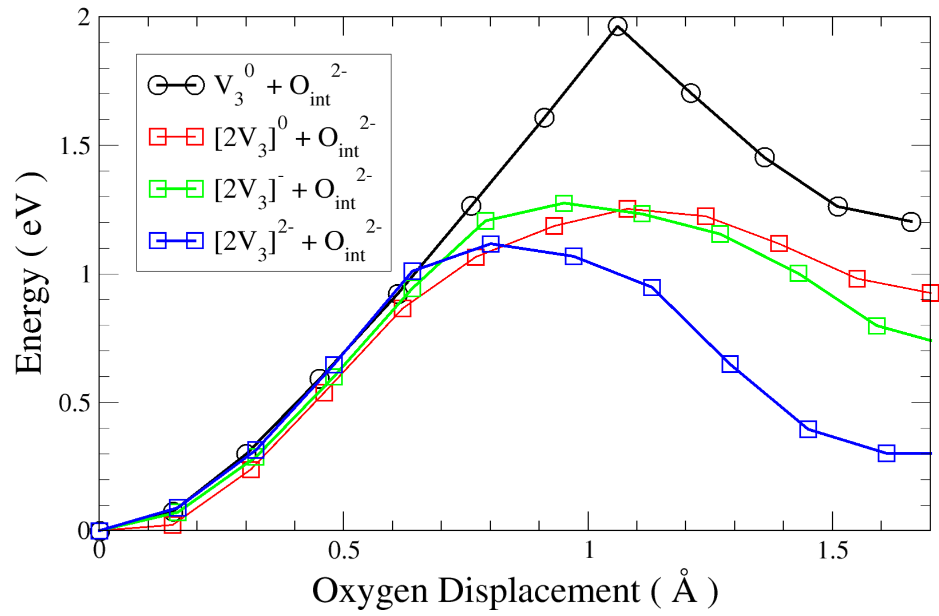


Figure 6.2: Barriers of formation for the HfO_2 NN FD pair, di-vacancy NN FD pair and charged di-vacancy NN FD pairs.

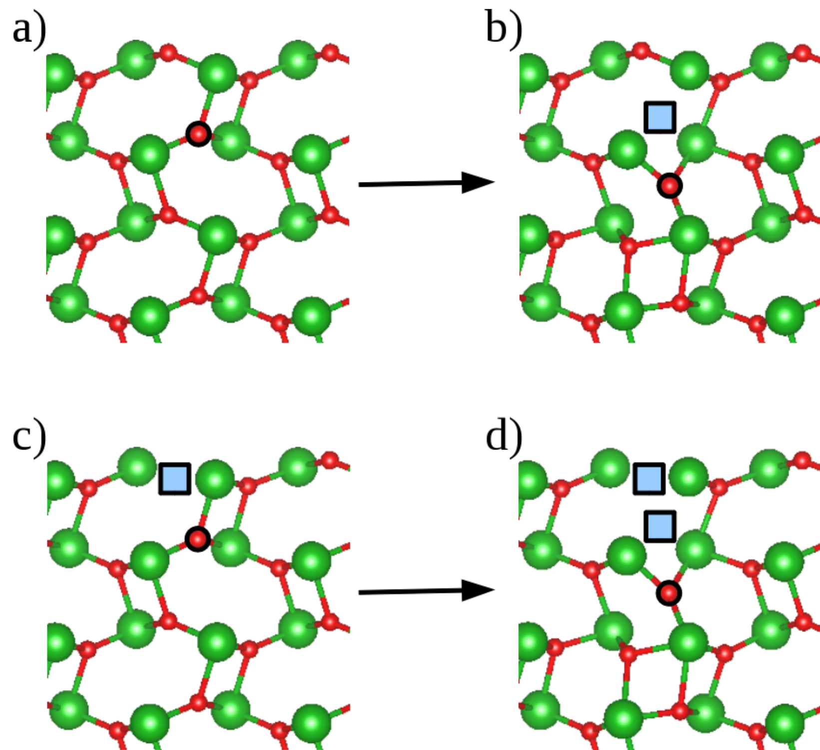


Figure 6.3: (a) - (b) Formation of a single NN FD pair. (c) - (d) Formation of di-vacancy NN FD pair in the presence of a pre-existing vacancy (large green balls: Hf, small red balls: O, blue squares: vacancy, highlighted red ball: O_{int}^{2-}).

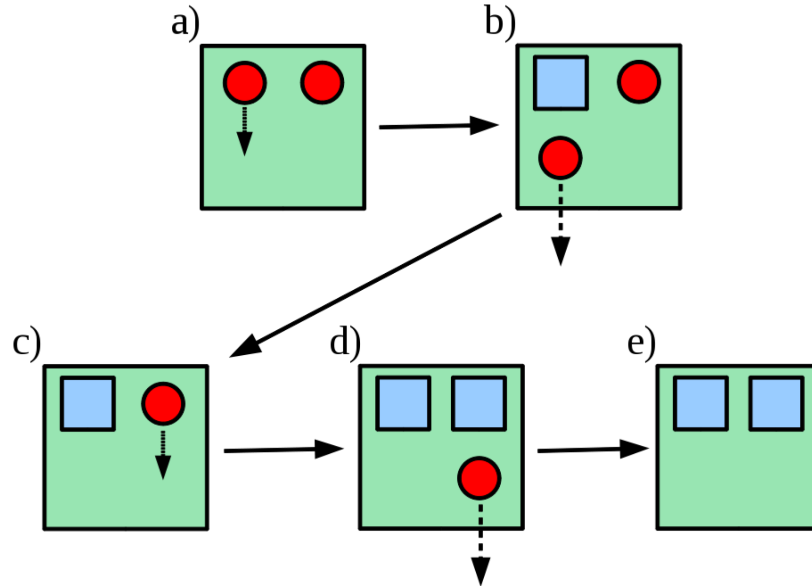


Figure 6.4: Schematic of the Frenkel pair formation mechanism where blue squares represent vacancies and red circles represent O^{2-}. (a) An oxygen atom leaves its lattice site forming an oxygen Frenkel pair . (b) The O^{2-}_{int} ion diffuses away leaving behind a V^0. (c) The vacancy's NN then leaves its lattice site, forming a divacancy Frenkel pair. (d) The second O^{2-}_{int} ion diffuses away. (e) This leaves behind $2V^0$ which can trap electrons and form another NN FD.

lowered by the additional gain due to attraction of O vacancies. The mechanism is summarised in Fig. 6.4.

This mechanism requires a concerted action of two electrons. The probability of such a process is not high for low electron current densities, but increases as the current density increases.[160,161] However, since the probability of trapping two electrons is low, it acts as a limiting factor in this process, preventing breakdown from happening too quickly, restricting the size of the CF and allowing it to be reversibly switched. The fact that this mechanism produces charged O ions is in agreement with voltage driven ion migration in bipolar devices. As shown in chapter 4 charged interstitial ions are likely to be the fastest diffusing oxygen defect in HfO_2, especially under high positive bias and electron injection conditions. Such a bias favours neutral and negatively charged vacancies which have diffusion barriers in excess of 2.4 eV.[121] This suggests that effective CF formation in ReRAM cells can proceed through creation of new O vacancies in the vicinity of pre-existing vacancies, complementary to an aggregation via diffusion processes.

6.4 Summary

We used DFT calculations to consider a mechanism by which oxygen vacancies and interstitial ions are formed in m-HfO_2 aided by electron injection. The results demonstrate that such a process is thermodynamically possible in the perfect lattice and near pre-existing O vacancies and requires overcoming activation barriers feasible at the temperatures of ReRAM electroforming. The pre-existing vacancy acts as an electron trap, and the di-vacancy which forms is stable, further lowering the formation energy of the defect. Furthermore, since the binding energy per vacancy in oxygen vacancy aggregates increases as the aggregate grows (see chapter 5), it is likely that the formation of NN FD pairs next to a larger vacancy aggregate will require even lower formation energy. Fig. 6.4 shows how the proposed vacancy formation process can lead to a propagation process within the bulk oxide.

More generally, the results of this work show that formation of FD pairs in oxides can be aided by electrons in the conduction band. These carriers can be produced by irradiation or tunnelling from an electrode. Understanding the electronic mechanisms of defect creation in materials under irradiation and carrier injection, is crucial for a wide range of technological applications. Injected electrons can play a significant role in the formation of O vacancy aggregates and stable oxygen interstitials.

7

The Pt/HfO_2 Interface

7.1 Introduction

Developments in microelectronics have led to a large amount of interest in metal-oxide interfaces. As the size of transistors is scaled down it has become necessary to replace the conventional gate materials, e.g. polycrystalline silicon (poly-Si), with a metal.

Despite not being an excellent conductor, highly doped poly-Si had a number of advantages over metal gate materials and was the standard gate material for many years. The threshold voltage of the transistor is modified by the work function difference between the channel material and gate material. Since poly-Si has the same band gap as the underlying Si channel, its work function can be easily tuned by changing the level of doping, to achieve low threshold voltages for both PMOS and NMOS devices. Also, it could be deposited before high temperature processes, during which many metals would melt. This meant a one step etching process could be used rather than the multi-step processes used for metal gates. However, as transistors were scaled down the poly-Si depletion effect led to undesirable variations in the threshold voltage. In this effect, randomly scattered charge carriers in poly-Si are separated, by the applied positive electric field. Electrons in poly-Si move closer toward the gate

terminal but don't start to flow due to the open circuit configuration. This separation of charges leads to a depleted region at the poly-Si/SiO_2 interface, which has a direct effect on the threshold voltage and channel formation.[162]

Metal gates are now usually used in conjunction with high-k dielectrics, such as HfO_2. To improve the performance of transistors made with such a new configuration, we need a detailed understanding of the interface between metal gates and gate dielectrics, including detailed atomistic structures of the interface and of defects at the interface. We also need to understand how these features affect the electronic structure of the metal-oxide interface, and affect device performance.

The metal-oxide interface is also extremely important in ReRAM. As explained in section 1.4, HfO_2 has been shown to be a promising material for this application.[70,163,164] HfO_2 ReRAM is thought to operate via formation of CFs made up of oxygen deficient regions. TiN and Pt appear to be the most promising materials for the metal electrodes in the MIM cells and Pt has been suggested to be a particularly beneficial material for the anode.[24] Furthermore, these materials have Fermi levels within the band gaps of the binary oxides, giving a reasonably large conduction band offset.

Although these devices have been demonstrated successfully, questions still remain around the mechanisms behind CF formation. As previously explained, the leading model is via the formation of oxygen deficient regions under an applied bias. This model implies that during the forming process oxygen ions must diffuse away from the growing filament. However, the mechanism behind the motion of the oxygen ions and the final location of the ions expelled from the CF region is not well understood.

During the forming process the CF is thought to grow via diffusion of oxygen ions beginning at one electrode.[70] It has also been suggested that during forming oxygen may diffuse into the metal electrodes.[24,165] If oxygen can be reversibly stored in the electrode, it may serve as a reservoir for re-oxidation and reduction of the metallic filament (i.e. reset and set). Therefore, the behaviour of oxygen near electrodes may play a critical role in determining the performance and reliability of ReRAM devices.

Directly probing oxygen ion dynamics near a buried interface in a device is very difficult and the role the electrode plays in these devices remains unclear. Addressing these issues is seen to be an important step to implementing ReRAM technology. In this chapter atomic scale simulations of an interface are used to see how defects can affect the structure and electronic properties of these devices.

Different techniques have been developed to prepare high-quality metal and oxide films, such as molecular beam epitaxy (MBE) and ALD.[166] In most cases both the grown film and the metal substrate experience strain due to the lattice mismatch. Different mechanisms of strain release have been suggested for systems with different lattice mismatch,[166] which include formations of grain boundaries, dislocation networks, and low-coordinated surface sites, such as kinks, steps, corners, and vacancies. The formations of such defects have a significant impact on the electronic properties of metal oxides, and might further cause failures in real microelectronic devices.[167,168]

Previous theoretical studies have been carried out on metal/HfO_2 interfaces using various different metals, and have mostly focussed on calculating the band offset and the stability of the interfaces.[169–173] The stoichimetry of the Pt/HfO_2 interface has been shown to have a variety of O terminations related to the O chemical potential.[174] There are fewer studies for defective interfaces. The segregation of oxygen vacancies towards the Pt/HfO_2 interface driven by the formation of Hf-Pt bonds has been demonstrated.[175] The enhanced stability of charged vacancies near high work function metals leading to the stabilization of extended Frenkel pair defects has also been discussed.[176]

7.2 Interface Building

Modelling an interface for nanoelectronic components is a trade-off between creating a model which realistically mimics a device and the computational cost needed to minimise and calculate the properties of the model. The major approximation made is in using a periodic slab model. This immediately limits the simulation cell to

one which periodically repeats and so an epitaxial interface must be formed. For an interface in which the two materials have a significantly large lattice mismatch ($>$ 2%) the strain induced by that mismatch would be released elsewhere by an extended defect region. However within small unit cells it is not possible to model the the large regions necessary to relieve this strain. Therefore, it is important to produce an interface in which the strain, and the overall energy of the interface, is minimised.

Based on this principle, one of the most popular methods for constructing metal/oxide interface models is the best coincidence approach. In this approach the interfacial surface of each material is chosen to be as close in parameters to one another as possible. Once these surfaces are chosen, they are rotated and shifted with respect to one another, to find the lowest energy interface structure. In this method the rotation required is determined by the space groups of the two surfaces, and the amount of lattice mismatch between two materials. This method is very straightforward, and this method is used here to explore the interface structure between Pt and HfO_2.

Once a reasonable interface structure has been found, the second challenge is to select a theoretical method which gives accurate band offsets at the metal-oxide interface, since both the energy levels of defects, and Fermi level of the metal substrate sit in the band gap of the metal oxide. This means that if two energy levels get close to each other, charge might transfer from the defect energy level to the metal Fermi level, or from the metal Fermi level to defect energy level, depending on their relative positions. Since the metal-oxide interface system which we are dealing with is a heterogeneous system, in which we have metal as a conductor and oxide as an insulator, the choice of a good theoretical method for description of the electronic structure of the interface is even more challenging, as there is currently no uniform method which can describe both materials properly.

Generally for an accurate description of the electronic structures of metal and oxide, we have two choices, we can either use GGA or a non-local hybrid density functional, in which GGA is augmented by a certain fraction of the Hartree-Fock exchange. As it has been widely accepted, GGA is computationally cheap and gives

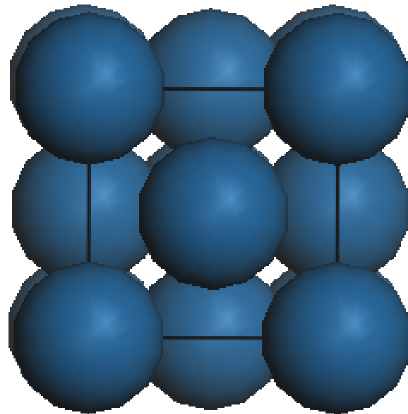


Figure 7.1: The close packed FCC structure of platinum.

good descriptions for metals. However, GGA underestimates the band gaps of semiconductors and insulators (see chapter 3.3). On the other hand, it has been reported that some hybrid functionals, e.g. B3LYP, fail for metals because they cannot attain the homogeneous electron gas limit.[177] Although non-local hybrid functionals give much better band gaps compared with pure GGA, the fourth order scaling, as a result of the calculation of Hartree-Fock exchange, means we cannot use them for a very large system.

7.3 Platinum

Platinum is a desirable material for a metal electrode in both transistors and ReRAM. It has a very high conductivity, excellent chemical stability and can form high quality interfaces with a variety of materials.

Bulk platinum forms a close packed, face centred cubic (FCC) structure (space group $Fm-3m$)(Fig. 7.1). Initially three functionals, LDA,[178] PBE,[179] and PW91[122] were tested for calculating the properties of bulk Pt.

For metals a dense k-point mesh must be used to account for the sharp discontinuity in k-space between occupied and unoccupied states. In the bulk Pt calculations a 15×15×15 k-point mesh is used along with a first order Methfessel-Paxton smearing method to avoid convergence problems because of band-crossings above and below the Fermi level.

	a_0/ Å	B_0	E_0 / eV
Experiment	3.920[181]	278[182]	5.84[182]
LDA	3.90	320	7.45
PW91	3.98	260	6.04
PBE	3.97	270	6.05

Table 7.1: Lattice parameter (a_0), bulk modulus (B_0), and energy (E_0) of a bulk FCC platinum crystal.

Table 7.1 shows the lattice constants (a_0), bulk moduli (B_0) and cohesive energies (E_0). The LDA lattice constant of 3.90, is just 0.5% short of the experimental value. The PW91 and PBE functionals do overestimate the lattice constant however the values of B_0 and E_0 agree far more closely with experiment than those calculated with the LDA functional. Although the PBE functional performed better than the PW91, in this case the PW91 was used for all subsequent calculations. This is because there are a number of previous studies on the Pt/ HfO_2 interface which use this functional, showing that it performs well for that system, and allowing for direct comparison.[176,180]

The electronic density of states (DoS) of bulk platinum was investigated using the PW91 functional. The total DoS is shown in Fig. 7.2, with the Fermi level corresponding to 0 eV. The bulk DoS is comparable with that calculated using the LDA by Kokalj *et al.*. The total DOS is dominated by *d*-states with a contribution by the *s*-states at the lower *d*-band energies, which is suggested to be indicative of hybridisation of an empty *s*-band with the valence *d*-band.[183] The *p*-states have only a very slight contribution although the *p*-band does span the *d*-band.

7.4 Pt(111) Surface

Like all FCC crystals the (111) surface of Pt is the most stable (Fig. 7.3). This is due to the close packed nature of the surface in which each atom is 9 coordinated whereas the (100) and (110) surfaces 8 and 7 coordinated respectively. It was found by Kokalj

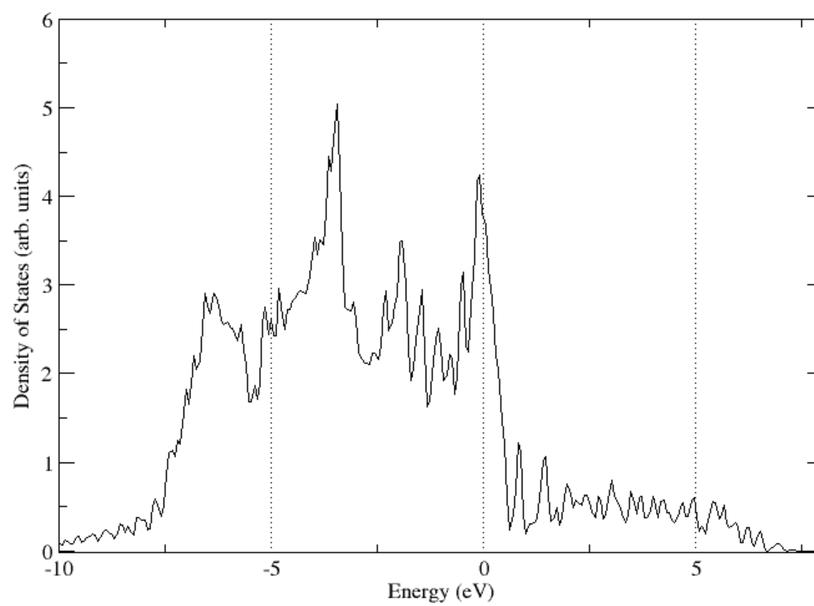


Figure 7.2: Total electronic density of states of bulk platinum.

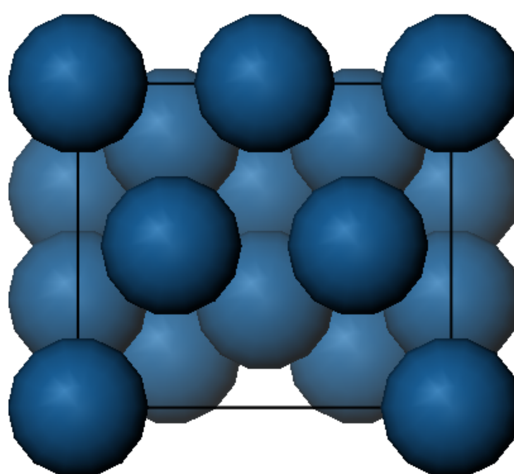


Figure 7.3: (111) $2\times\sqrt{3}$ surface of FCC platinum.

el al. that the surface states of the Pt(111) surface decayed within 4 monolayers.[183] This means the fourth layer is bulk-like and so, an 8 layer 1×1 slab was used to probe the electronic and structural properties of the Pt(111) surface. For these calculations a 7×7×1 k-point mesh was used. Figure 7.4 shows the sum of the total DOS projected onto the surface and fourth deepest layers. Comparison between the surface, 4th layer and bulk DOS is difficult as they all share similar, complex features, however, the surface DOS does show good agreement with that calculated by Kokalj *el al.*[183] with a slightly narrower and sharper *d*-band. The 4th layer DOS has lost the sharp nature of the surface DOS and does resemble the bulk structure. It is reasonable, therefore, that an 8 layer Pt electrode can be used to model both interface and bulk properties of the Pt/HfO_2 interface.

7.5 HfO_2(001) Surface

m-HfO_2 has eight possible surface directions, the most stable of which is the ($\bar{1}11$) surface.[103] However, in order to model the Pt/HfO_2 interface in a reasonably sized supercell, the interface must be epitaxial. It is therefore preferable to use the rectangular (001) surface unit cell, as opposed to the rhombohedral ($\bar{1}11$) surface unit cell, in order to increase the likelihood of finding a periodic match to the Pt(111) surface. Despite being higher in energy, the (001) surface has been observed by X-ray diffraction in HfO_2 crystal growth studies and so is a reasonable choice.[184]

The (001) surface was cleaved from the bulk structure and its geometry optimised using an 1×1×8 supercell to match the dimensions of the earlier platinum surface. Geometry relaxation of the surface reduced the rumpling of the 2-coordinated terminating oxygen atoms. The final structure corresponded well with those calculated by Musgrave *et al.*.[103] The surface electronic structure was analysed through the DOS and projected band structure. The electronic states of the (001) surface show the emergence of deep surface states corresponding to the 2s states of oxygen and on the edge of the conduction band corresponding the the 5d states of Hf, narrowing the

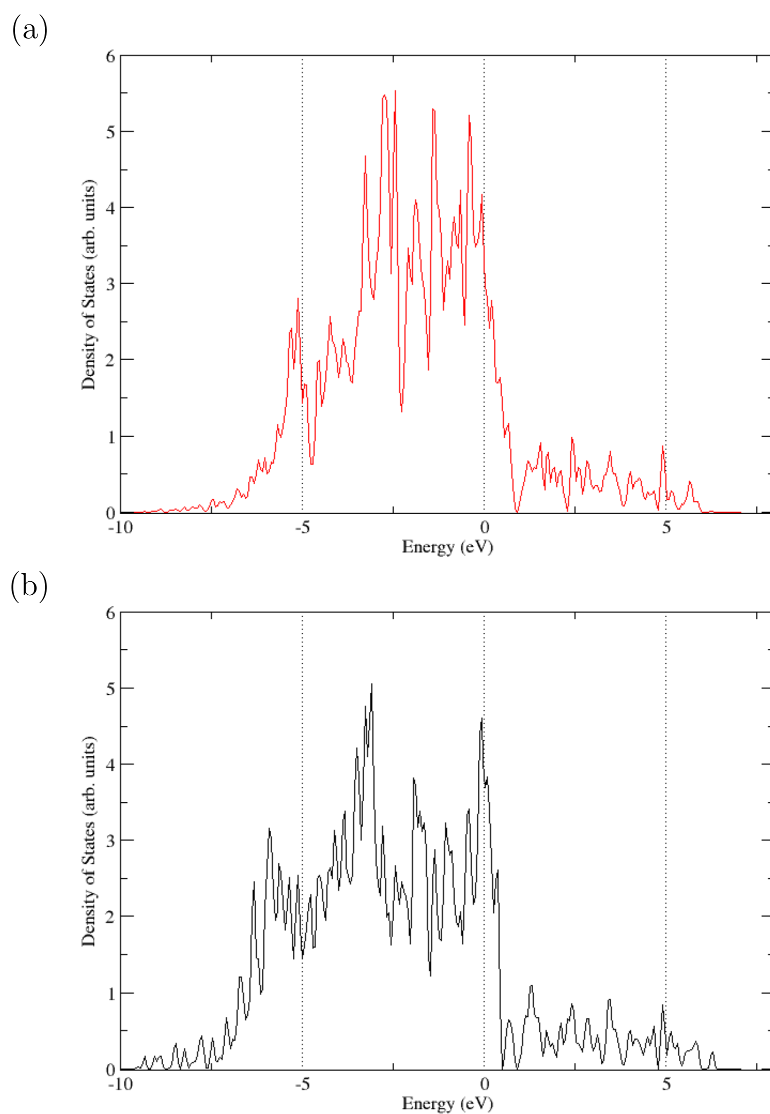


Figure 7.4: Projected density of states of the Pt(111) surface onto (a) the surface layers, (b) the fourth deepest layers.

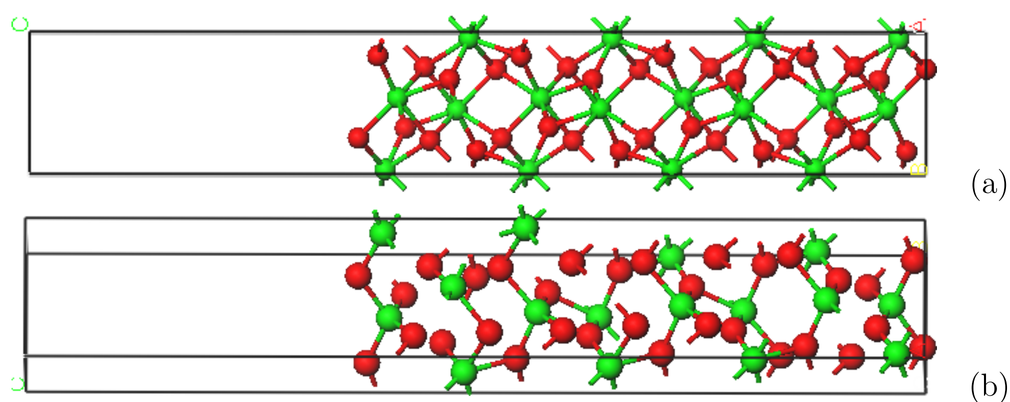


Figure 7.5: The (001) surface of monoclinic HfO_2. (a) unrelaxed, (b) relaxed.

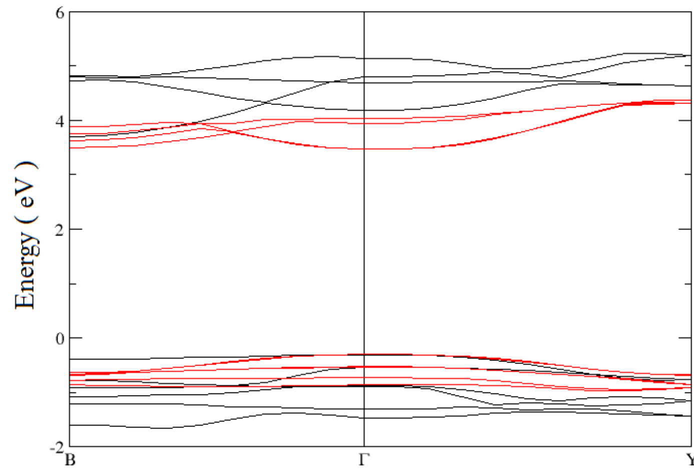


Figure 7.6: The band structure of the highest occupied and lowest occupied bands of bulk monoclinic HfO_2 (black) and the monoclinic (001) surface (red) projected along the $B - \Gamma - Y$ direction of the Brillouin zone.

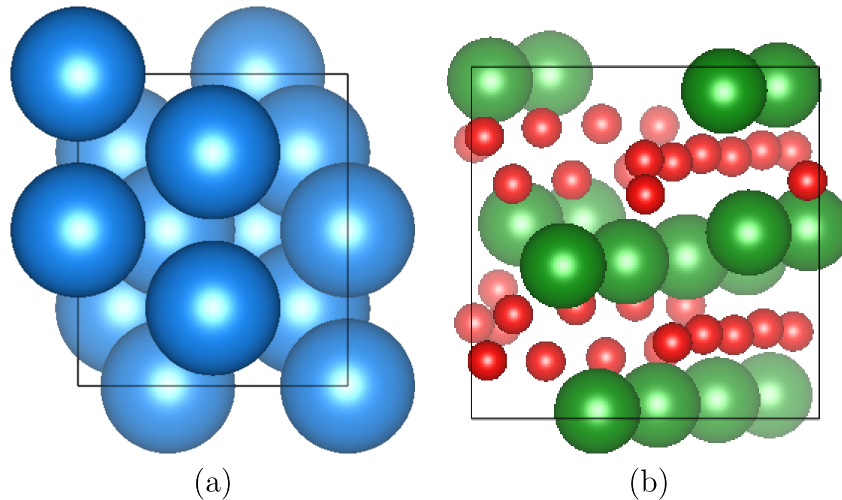


Figure 7.7: (a) The Pt(111) $2\times\sqrt{3}$ surface. (b) The HfO_2(001) 1×1 surface (blue: Pt, green: Hf, red: O).

band gap by 0.5 eV. This narrowing is confirmed in the projected band structure shown in figure 7.6. The figure shows the highest occupied and lowest unoccupied states of the HfO_2(001) surface (red) and bulk m-HfO_2 (black). The surface states are located in the Γ point and narrow the difference between the lowest conduction band states in the B and Γ points.

7.6 Pt/HfO_2 Interface

The Pt(111)/HfO_2(001) interface model was constructed using the best coincidence approach and rotating the HfO_2 layer with respect to the Pt layer to find the best configuration in which two lattices have the smallest lattice mismatch. Previous studies have been conducted using the Pt(111) $2\times\sqrt{3}$/HfO_2(001) 1×1 interface.[173,176,180] The HfO_2(001) 1×1 surface has a rectangular shape, the smallest rectangular cell which can be made from the FCC surface is the $2\times\sqrt{3}$ surface shown in figure 7.7. A 6.6% and -7.1% strain is then added the Pt cell in the x and y directions respectively to match the lattice parameters of HfO_2. Although the strain is unphysically large, this configuration has the smallest strain possible for such a small cell. The calculated work function of the Pt(111) surface reduced from 5.80 to 5.76 eV when the strain was applied but remained within the experimental margin of error.[185] The oxygen terminated HfO_2(001) surface was also found to be the most stable at the Pt/ HfO_2 interface using first principles thermodynamics.[173]

The thickness of the Pt and HfO_2 layers must also be considered. Goux *et al.* produced a Pt/HfO_2/Pt resistive switch composed of a 10–90 nm Pt base electrode, a 5 nm HfO_2 layer and a thick top electrode.[24] It was noted that the thickness of the Pt layer did not affect the operation of the switch and so, following on from previous calculations, an 8 layer electrode should be sufficient to model both the interface and bulk characteristics of the electrode. 5 nm is infeasibly thick for periodic DFT calculations, however, the HfO_2 layer must be large enough to be able to model the motion of defects without them interacting with their periodic images. The distance over which this interaction occurs is unknown therefore this criterion must be converged to minimise the interaction.

The Pt(111) $2\times\sqrt{3}$/HfO_2(001) 1×1 interface geometry was optimised using the following procedure. For this interface the construction of two almost equivalent interfaces within one unit cell is not possible. Therefore two, 4 monolayer thick, HfO_2 layers were placed either side of the 8 monolayer thick Pt layer, in order to create a symmetric slab. A 15 Å vacuum gap was used to allow the layers to fully relax

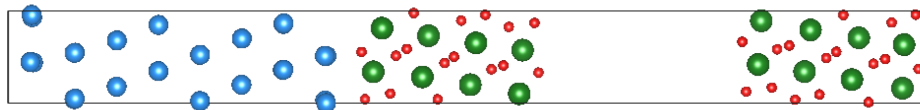


Figure 7.8: The Structure of the Pt(111) $2\times\sqrt{3}$/HfO_2(001) 1×1 interface model (blue: Pt, green: Hf, red: O).

perpendicular to the interfacial plane and placed between the two HfO_2 layers to prevent any surface polarization effects. The atomic coordinates were then relaxed whilst retaining the unit cell lattice parameters. In order to find an optimum interface geometry, a grid search method was used whereby the relaxed slabs were shifted relative to one another in the x,y plane in a 10×10 grid. An initial survey of the shifted geometries was used to eliminate those which were equivalent to one another (by rotation or translation) and the remaining interface structures were then re-optimised. The resulting lowest energy interface was used in further calculations.

The lowest energy structure of the Pt(111)/HfO_2(001) interface is shown in figure 7.8. In the resultant structure, the two interfacial oxygen atoms are positioned at 2.11 Å and 2.25 Å on top of the Pt atoms and Hf atoms of the first layer in a bridging site 2.82 Å and 3.38 Å from the nearest Pt surface atoms. This arrangement is in good agreement with previously reported structures of this interface which used computationally more expensive simulated annealing to optimise the structure.[186]

The projected density of states of the interface calculated using the PW91 functional is shown in figure 7.9.

Fig. 7.10 shows the layer projected CBM and VBM across the whole interface model. The first layer of HfO_2 is filled with the Pt metal states and so does not have a band gap. This is due to the penetration of the metal states into the oxide layer in an effect known as metal induced gap states. The central 2 layers, 4 monolayers deep, show the bulk BG of HfO_2

Fig. 7.11 shows the internal energy of interfaces for a wide range of O chemical potential, μ_O, calculated using the thermodynamic formalism of Scheffler et al. modified to describe the oxygen potential across an interface.[187] The dashed lines show the limits between which the two phases which comprise the interfaces remain sta-

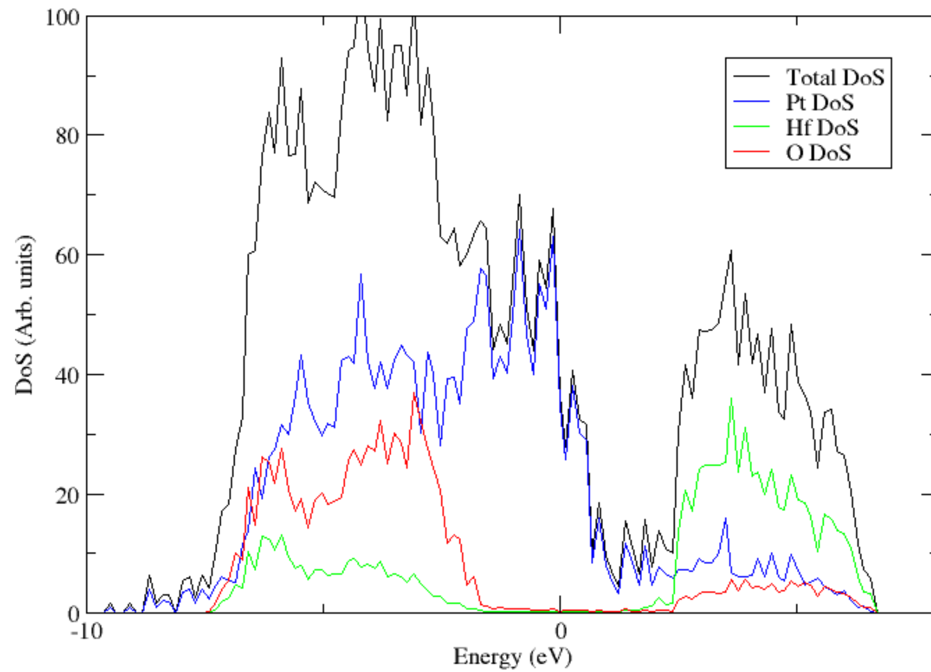


Figure 7.9: Total electronic densities of states (black) and the states projected onto the Pt atoms (blue), Hf atoms (green) and O atoms (red) adjusted so the Fermi level, $E_F = 0$. The CBM is taken as the energy of the lowest Hf state above E_F and the VBO taken as the energy of the highest O state below E_F.

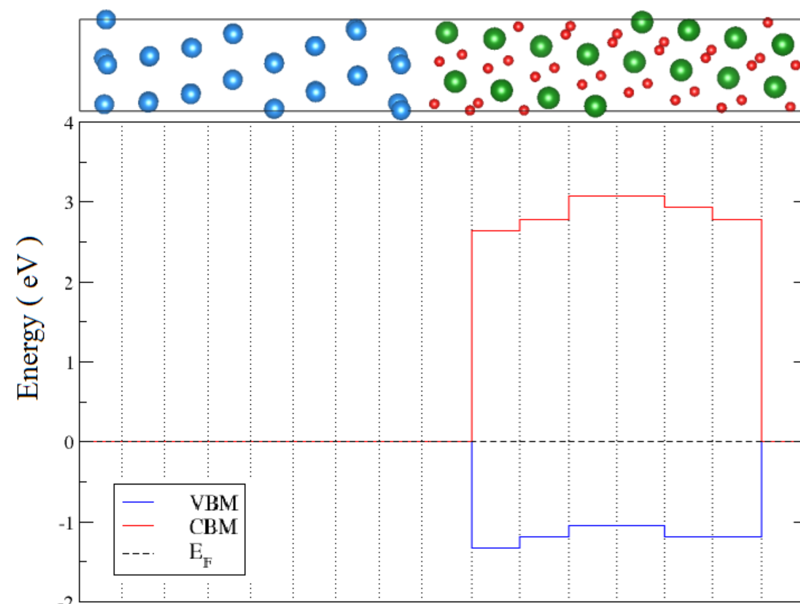


Figure 7.10: Layer projected CBM and VBM of the Pt(111) $2\times\sqrt{3}$/HfO_2(001) 1×1.

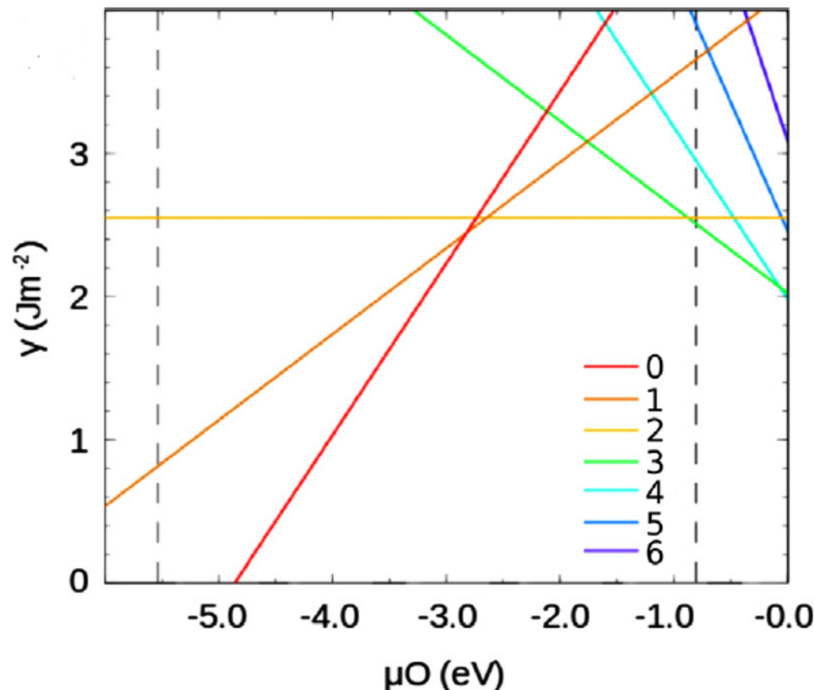


Figure 7.11: Interface internal energy energy versus O chemical potential for interface terminations of varying O concentration. Labels correspond to the number of interfacial O atoms between the top layer of the metal and bottom layer of Hf ions. $\mu_O = 0$ is defined as the energy of a free isolated O_2 molecule. Vertical dashed lines represent the limits above and below which α-PtO and metallic Hf would be favourable.

ble. Below the lower limit ($\mu_O < -5.5$ eV) hcp-Hf becomes more stable than HfO_2, while the upper limit corresponds to oxidation of the electrode forming α–PtO_2 (for $\mu_O > -0.8$ eV), respectively. Positive gradients indicate O-poor interfaces i.e. containing O vacancies within the interfacial layer. Negative gradients indicate O-rich interfaces i.e. containing O interstitial atoms within the interfacial layer. It is shown that the interfacial O concentration is very sensitive to the O chemical potential. As the oxygen chemical potential is reduced, the most stable interface terminations go from containing a high concentration of O interstitials to a high concentration of O vacancies. At low μ_O, a metallic Hf-Pt interface is favoured and only at very high μ_O is it favourable to have extra oxygen atoms at the interface relative to the stoichiometric case.

If a bias is applied across the cell the motions of oxygen ions will change the O chemical potential at the interface. This result implies that if the bias is applied and μ_O becomes low, a substoichiometric interface layer is energetically favoured and so

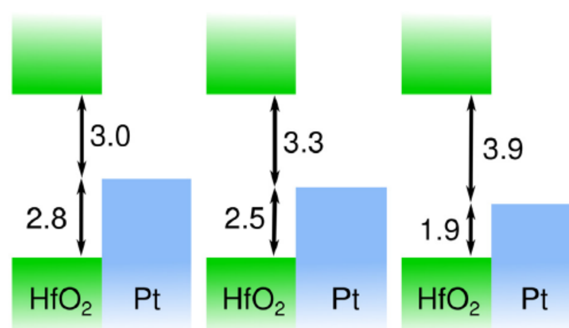


Figure 7.12: Band offsets of the stable interface termination structures. Left to right: O deficient, stoichiometric, O rich.

could act as a sink for oxygen vacancies.

The metal-oxide band offsets for the stable interfaces are shown in figure 7.12. These band offsets were calculated using HSE06, correcting the HfO_2 band gap from 3.48 eV to the experimental value of 5.8 eV. The Pt Fermi level is located in the middle of the HfO_2 band gap and reduces with increasing interfacial oxygen concentration. This results from polarisation effects and compression of the metallic wavefunctions due to reduction of the metal-oxide interfacial separation. The stoichiometric band offsets are similar to those calculated by Fonseca *et al.* of +3.5 eV and -2.3 eV using the GW approximation to correct the band gap error.[186] Experimentally the band offsets are measured as +2.48 and -3.32 eV.[188] However, these calculations are of the polycrystalline Pt/HfO_2 system therefore the measurements are for the band offsets averaged over a number of Pt/HfO_2 surface pairings rather than specifically the Pt(111)/HfO_2(001) calculated in all the theoretical work. This may suggest an O rich interface in the experimental cell.

At the anode, diffusion of oxygen into the electrode can facilitate nucleation of a metallic filament and initiate forming. Oxygen ions driven towards the cathode by the electric field may also penetrate into the electrode.[24] Therefore, the energies to move oxygen from various sites near the interface into the electrode were calculated using a 2×2×1 supercell of the interface structure. This 320 atom unit cell reduced the density of the defects in the interface model preventing interaction between vacancies in periodic images. Energies for oxygen incorporation were then calculated by taking

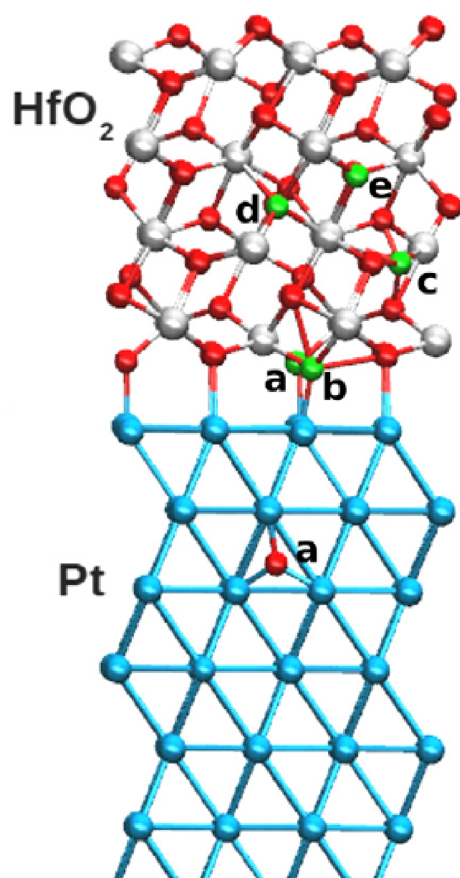


Figure 7.13: Incorporation from HfO_2 O lattice sites, a, c, d, e, and O interstitial site, b, into Pt interstitial O site, a. (Light blue: Pt, silver: Hf, red: O, green: defect sites).

Site	Bonds	Pt (interstitial)
a	1Pt/2Hf	-1.5
b	1Pt/2Hf	3.5
c	3Hf	1.7
d	4Hf	2.5
e	3Hf	2.6

Table 7.2: Energetic cost, in eV, to transfer O from various sites in HfO_2 into the Pt electrode (Fig. 7.13).

an O atom from a particular site in the HfO_2 layer (Fig. 7.13 c-g) and placing it into an interstitial position in Pt (Fig. 7.13 a or b). Table 7.2 shows the calculated incorporation energies.

Oxygen incorporation is shown to be unfavourable in every case, except when the vacancy and interstitial were very near one another (i.e. Fig. 7.13 (a)). However, this is likely to be due to the strain relief offered by an interfacial vacancy. Furthermore, the location of the Pt Fermi level in the middle of the HfO_2 band gap close to the defect levels led to subtle changes in the charge transfer between specific defects and the metal electrode. This is particularly noticeable in the difference in energies between the 3 and 4 coordinated oxygen vacancies. In this case the 3C O defect state lies slightly higher in the HfO_2 band gap than the 4C state while the Pt Fermi level sits between the two. This results in the 4C O vacancy remaining neutral whilst the 3C vacancy loses an electron to the metal electrode and becomes positively charged. This causes a much larger perturbation in the HfO_2 lattice, raising the incorporation energy.

7.7 Summary

The stable structures and band offsets of the Pt/HfO_2 interface was calculated using DFT. The concentration of oxygen at the interface is sensitive to the O chemical potential, but only at high O concentrations are O atoms likely to be stored at the

Pt interface. Furthermore, the band offset is shown to be sensitive to the effects of defects close to the interface, decreasing as the O concentration increases.

These band offset predictions allow us to make some conclusions on the nature of the mobile oxygen species in resistive switching devices near to different electrodes. In these systems the electron chemical potential is fixed by the Fermi level of the electrode, which determines the charge state of the defects in the oxide. Close to the Pt electrode oxygen vacancies are predicted to be neutral or in the singly positively charged state. The diffusion barrier of 2.1 eV for the positively charged vacancy suggests that oxygen is unlikely to diffuse via the vacancy mechanism near to Pt electrodes.

Calculations of incorporation energies for oxygen ions into the electrode show that Pt presents a significant barrier to oxygen diffusion. Therefore, it is unlikely that an incorporation process at this electrode is responsible for the set/reset behaviour of ReRAM.

8

Oxygen Scavenging in the Si/SiO_2/HfO_2/TiN Stack

8.1 Introduction

The thickness of the gate oxide, and the quality of the interfaces within a transistor stack, has a direct impact on device performance. Equation 1.2 shows that, to improve the gate capacitance, the oxide needs to be as thin as possible. Post-deposition annealing is often used to both reduce the number of defects at interfaces, and to thin the oxide layer.

The equivalent oxide thickness (EOT), is a measure of the thickness of a dielectric oxide layer, with respect to the thickness of SiO_2 required to give the same dielectric effect.

$$EOT = t_{high-k}\frac{k_{SiO_2}}{k_{high-k}}. \tag{8.1}$$

A typical high-k/metal gate stack structure contains a silicon dioxide (SiO_2) interfacial layer (IL), between the Si channel and the high-k material.[189] An apparent method for scaling these stacks, is to reduce the EOT by reducing the thickness of

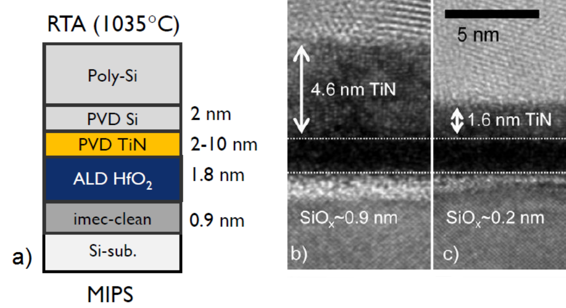


Figure 8.1: (a) Illustration of the structure of the Si/SiO_2/HfO_2/TiN stack. HRTEM image of three gate stacks with (b) HRTEM of the stack with ∼5 nm TiN layer and (c) ∼2 nm TiN on top of imec-clean/1.8 nm HfO_2.[7]

the SiO_2 IL. This is achieved by capping the stack with a material with a high oxygen affinity, and annealing the stack at high temperatures. The metal layer then absorbs oxygen ions from the IL in a process known as remote IL oxygen scavenging.

Fig. 8.1 shows high resolution transmission electron microscopy (HRTEM) measurements, made by Ragnarsson *et al.*, on a Si/SiO_2/HfO_2/TiN stack. These results show that the thinner TiN metal layer led to a reduced the EOT.[7] It has also been shown that the reduction in EOT is independent of the starting thickness of the SiO_2 layer.[189] Hopstaken *et al.* used SIMS, and isotopic labelled oxygen, to characterise the transport of O through a Si/HfO_2/TiN stack. It was demonstrated that annealing at higher temperatures increased the amount of labelled oxygen in the TiN layer, and decreased the amount in the oxide layer.[190]

This O scavenging process is also used in ReRAM devices, to create an oxygen deficient insulating layer, increasing the number of vacancies available for CF formation. For example, a Hf layer can be deposited on top of the HfO_2 layer in an MIM ReRAM stack. The Hf metal is essentially a vacancy rich version of HfO_2, and the chemical potential of oxygen is high in this system. Upon annealing the Hf layer is then oxidised by the O ions from the HfO_2 layer, equilibrating the number of oxygen atoms throughout the stack, and creating a substoichiometric oxide. Thin AlCu and Ti capping layers have been also incorporated as oxygen scavengers, generating vacancies within the oxide films and enhancing resistive switching properties.[191]

Remote IL-scavenging is a technologically relevant problem, and a significant

amount of experimental and theoretical data relating to it has been collected. However, a satisfying mechanism for the O scavenging in the $Si/SiO_2/HfO_2/TiN$ has yet to be proposed.

In this chapter, a model for the oxygen scavenging process is developed in the $Si/SiO_2/HfO_2/TiN$ stack. Firstly, in section 8.2, the feasibility of O scavenging at the TiN/HfO_2 interface is examined. Then in section 8.3, the results of various calculations on the $Si/SiO_2/HfO_2$ interface are collected. Finally, in section 8.4, these results are combined to propose a scavenging mechanism through the stack.

8.2 The TiN/HfO_2 Interface

As suggested above, the TiN layer is expected to absorb oxygen atoms from the HfO_2 layer in the scavenging process. In order to look at this a TiN/HfO_2 interface model was developed.

As discussed in section 7.2, for the purpose of periodic calculations, the interface needs to be commensurate. To match with the calculations in section 7.6, the (111) surface of TiN and (001) surface of m-HfO_2 were used to form the TiN(111)($5 \times \sqrt{2}$)/m-HfO_2(001)(3×1) interface. As in the Pt/HfO_2 interface, the lattice parameters are fixed to the HfO_2 layer, inducing a small strain in TiN (+2.7% in x-direction and -0.7% in y-direction). Similarly, the lowest energy structure was found by a grid search method; shifting the layers with respect to one another across a 10×10 grid, and choosing the lowest energy structure for further calculations. However, in this cell a vacuum gap was not necessary since it was possible to construct two almost equivalent interfaces within 15 Å. The z unit cell parameter was optimised to find the equilibrium interfacial separation. Within the lowest energy interface structure, the distribution of Ti-O bonds at the TiN/HfO_2 interface is 2.0–2.5 Å with an average of 2.2 Å. The Hf-O bonds near the interface are 2.0-2.2 Å with average of 2.1 Å.

Fig. 8.2 shows the interface internal energies across a range of O chemical potentials, as described in section 7.6. In this case, the upper limit corresponds to oxidation

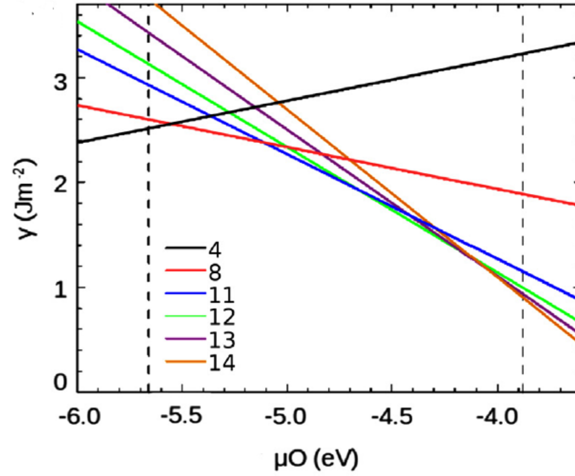


Figure 8.2: Interface internal energy versus O chemical potential for interface terminations of varying O concentration. Labels correspond to the number of interfacial O atoms between the top layer of the metal and bottom layer of Hf ions. $\mu_O = 0$ is defined as the energy of a free isolated O_2 molecule. Vertical dashed lines represent the limits above and below which TiO_2 and metallic Hf would be favourable

Site	TiN (interstitial)	TiN (N vacancy)
a	6.5	1.0
b	4.6	-0.9
c	5.0	-0.5
d	7.7	2.2
e	7.0	1.5

Table 8.1: Energetic cost in eV to transfer O from various sites in HfO_2 into the TiN electrode (Fig. 8.3).

of the electrode forming TiO_2 ($\mu_O < -3.9$ eV). As in the Pt/HfO_2 case, the interfacial O concentration is sensitive to the O chemical potential. For the TiN/HfO_2 interface, oxygen rich interfaces are favoured over a wide range of μ_O, and so it is possible that this interface may act as a reservoir for interstitial O atoms.

For the TiN interface, incorporation of an interstitial O atom into an interstitial position in TiN costs at least 4.6 eV. However, the system can gain energy (~0.9 eV) by moving oxygen from HfO_2, into a nitrogen vacancy in TiN. Therefore, if there are no N vacancies near the interface, or if they are already fully occupied, oxygen interstitials accumulate at TiN/HfO_2 interface and do not penetrate easily.

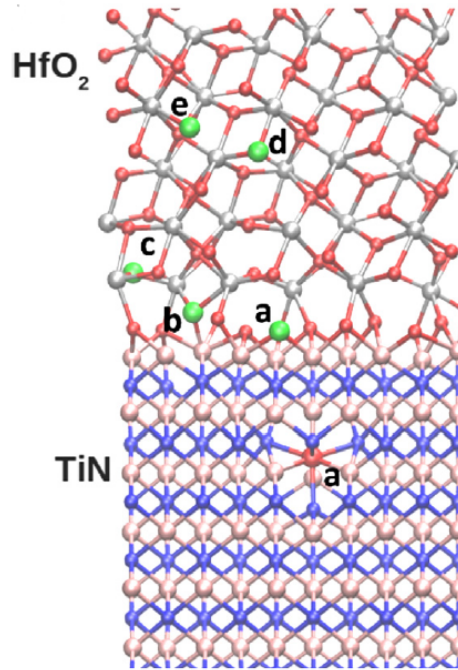


Figure 8.3: Incorporation from HfO_2 O lattice sites, a, c, d, e, and O interstitial site, b, into TiN interstitial O site, a. (Dark blue: Ti, pink: N, silver: Hf, red: O, green: defect sites.)

The TiN/HfO_2 band offset has been shown to be dependent on the stoichiometry of the interface.[192] TiN has small conduction band offset, so oxygen vacancies are neutral in this system, and are predicted to be relatively immobile, with a barrier to diffusion of 2.7 eV.[121] The small conduction band offset also means that oxygen interstitials are doubly negatively charged. These charged interstitials have a low activation barrier to diffusion through the HfO_2 layer.

The results here suggest that at high oxygen concentrations, not only will O interstitial ions accumulate at the TiN/HfO_2 interface but if N vacancies are available the O ions will diffuse into them. Thinner TiN layers are likely to be more defective than the thicker layers. This may mean that the reduction of EOT when the TiN layer is thinner is due to O ions being scavenged by more defective TiN.

8.3 The Si/SiO_2/HfO_2 Interface

The Si/SiO_2/HfO_2 interface model used here was developed by Gavartin *et al.*.[193] This interface structure was obtained using VASP code and GGA and was built using

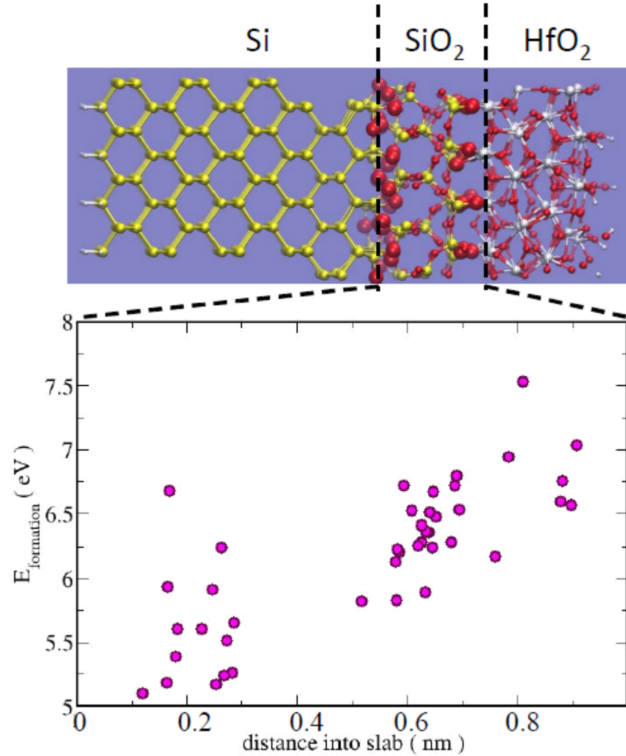


Figure 8.4: Formation energies of 43 neutral oxygen vacancies at various positions between the Si and HfO_2 parts of the stack. The right-most defects are directly at the HfO_2/SiO_2 interface and the left-most directly at the SiO_2/Si interface.[8]

a different procedure to that described in section 7.2.

First, Si/SiO_2 interface was generated in a large 15×15×40 Å supercell. The bottom Si (001) layer was passivated by hydrogen and kept constrained during subsequent geometry optimizations. The rest of the system was relaxed using short (∼5 ps) simulated temperature annealing followed by static relaxation. Second, a block of m-HfO_2 was added at the top of the structure, and the top surface of this HfO_2 block was hydroxilated. It is notable that the initial interface structure had many O ions bonded to one Hf and one Si ion, however, during subsequent annealing the interface became further oxidised. Finally, defect states in the structure were identified and manually 'healed' by addition of oxygen atoms, and the whole system relaxed again. This procedure was repeated until most of the defects in the system were eliminated.

The stability of oxygen vacancies across this interface model was calculated by Watkins *et al.*.[8] To do this they employed *ab initio* simulations using the CP2K program suite.[52] The Quickstep DFT module was used to carry out hybrid density func-

tional calculations using the PBE0-TC-LRC-ADMM hybrid density functional,[194,195] containing 25% exact exchange. The primary basis sets were the DZVP-MOLOPT-SR-GTH basis distributed with the code along with the corresponding GTH pseudopotentials (with 12, 4, 6, and 1 electrons treated as valence for Hf, Si, O and H, respectively).[196–198] The auxiliary Gaussian basis for the ADMM method was pFIT3 as detailed in reference 194, and a FIT3 basis for Hf was optimized using a variant of the procedure outlined in the same paper. The number of Si layers was increased with respect to the original structure to converge the band gap of the Si substrate (to 1.2 eV). The geometric structure of the whole stack has been optimized using the hybrid density functional and basis sets described above.

Fig. 8.4 shows the formation energy of various vacancies calculated throughout the SiO_2 layer. These results suggest that although there is a range of formation energies depending on the vacancy site, the formation energy of vacancies at the Si/SiO_2 interface are significantly lower than those at the SiO_2/HfO_2 interface. Due to this, it is possible that vacancies within HfO_2 will migrate through the SiO_2 layer to and reduce the EOT.

The formation energy of oxygen vacancies across the $Si/SiO_2/HfO_2$ interface were also calculated by Capron *et al.*.[121] They found that the formation energy of the neutral O vacancy decreased by 1.22 eV going from the HfO_2 to the SiO_2 layer, while the formation energy of the doubly positively charged vacancy increased dramatically across the same interface. This suggests that for any electron chemical potential within the Si band gap, oxygen vacancies will become neutral in SiO_2.

The stability of oxygen interstitials was investigated using the same methods as Watkins *et al.*. Fig. 8.5 shows the formation energies of oxygen interstitials in HfO_2 at the SiO_2/HfO_2 interface. As in the vacancies case, there is a spread of formation energies depending on the location of the defect. However, interstitial ions at the interface generally have lower formation energies than those inside the slab. The O–O NN distance is ~2.4 Å for all of the interstitial ions, indicating they are all O^{2-}_{int}, trapping electrons from the Si layer.

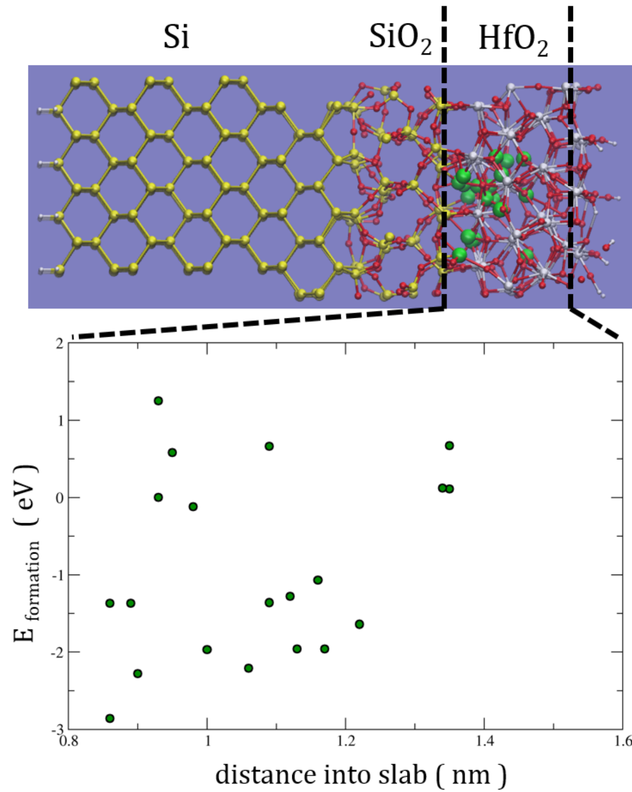


Figure 8.5: Formation energies of 28 oxygen interstitials within the HfO$_2$ layer at the SiO$_2$/HfO$_2$ interface.

Tang *et al.* used GGA to calculate the diffusion tendencies of O interstitials in the Si/HfO$_2$ interface. They found that there are strong thermodynamic and kinetic driving forces for O interstitials to migrate towards and get trapped at the interfacial Si layers. These results imply that any O interstitial ions are likely to remain at the Si side of the HfO$_2$ layer, and not migrate into the TiN metal.

8.4 O Scavenging Mechanism

For the oxygen scavenging process to occur, O ions must move from the Si/SiO$_2$ interface and into the TiN electrode, or conversely, O vacancies must migrate from the TiN side and to the Si/SiO$_2$ interface.

The following data can be used construct a theoretical model for remote IL-scavenging through the Si/SiO$_2$/HfO$_2$/TiN stack. The incorporation energy for O from HfO$_2$ into an N vacancies in TiN is -0.5 eV. Diffusion of O^{2-}_{int} through the HfO$_2$

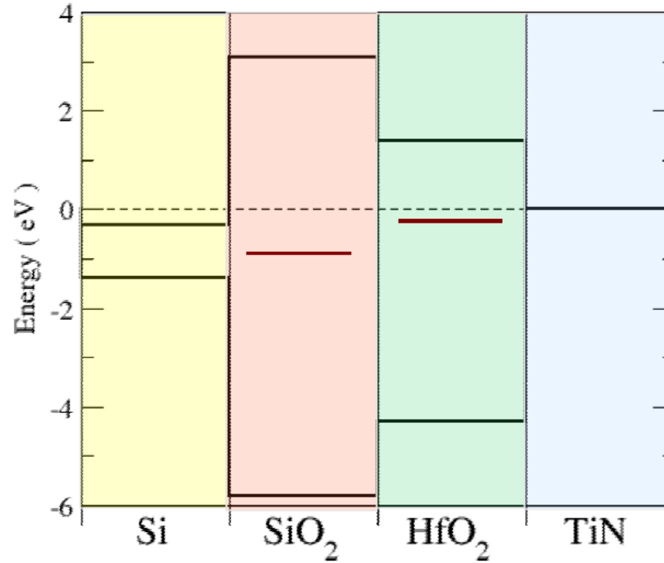


Figure 8.6: Summary of the band offsets between the components of the Si/SiO_2/HfO_2/TiN stack. Mid-gap red lines show the neutral oxygen vacancy levels in the oxides.

has a +0.7 eV barrier. Diffusion of O^0_{vac} through HfO_2 has a +2.4 eV barrier. The incorporation energy for O^0_{vac} into SiO_2 from HfO_2 is -1.2 eV.[199] The vacancies in the stack have the lowest formation energies at the Si/SiO_2 interface, ranging from 5.1 eV to 7.5 eV. The interstitials in the HfO_2 layer have the lowest formation energies at the SiO_2/HfO_2 interface with a minimum of -2.8 eV, but, according to ref. 200, are unlikely to migrate from there. Fig. 8.6 summarises the band offsets of the relevant materials using data from ref. 8 and ref. 192. In this system the Fermi energy is set by the metal, TiN, and lies above the vacancy levels in both HfO_2 and SiO_2, meaning the vacancies will not be positively charged.

During the high temperature process negatively charged interstitial ions will migrate quickly through the HfO_2 layer to the interfaces on either side. At the HfO_2/TiN interface, O ions will be absorbed into N vacancies within TiN, creating O^0_{vac} in HfO_2. There is then a consistent thermodynamic drive for the vacancies to migrate from the TiN side of the stack to the Si/SiO_2 interface, with a maximum barrier of 2.7 eV within the HfO_2 layer.

Finally, the formation energies of FDs at the SiO_2/HfO_2 interface range from 2.96–7.82 eV. Therefore, it is plausible that at high temperatures FD pairs may form at the

SiO_2/HfO_2 interface. In this case the O interstitial would be trapped at the interface, and the O vacancy would migrate to the Si/SiO_2 interface, reducing the EOT.

8.5 Summary

A mechanism for the remote IL-scavenging of O ions by TiN, in a $Si/SiO_2/HfO_2/TiN$ stack has been suggested. This mechanism proceeds primarily via neutral O vacancy diffusion during a high temperature, post-deposition annealing process.

The absorption of HfO_2 interfacial O ions, into nitrogen vacancies within TiN, was found to be favourable. This is confirmed experimentally by the concentration of O in TiN increasing after high temperature anneals in SIMS experiments. Furthermore, the measured EOT is reduced by using more defective TiN layers.[189,190] Calculations on the $Si/SiO_2/HfO_2$ model show that although the O interstitial would be the most mobile O defect through the HfO_2 layer, it is likely to get trapped at the SiO_2/HfO_2 interface.

From these results, it can be concluded that the O migration process is more likely to be initiated by O vacancy creation at the HfO_2/TiN interface, followed by the migration of these neutral O vacancies through HfO_2 and SiO_2 layers, to the Si/SiO_2 interface. The result of this process is the remote scavenging of O ions from the Si/SiO_2 interface, into the TiN layer, reducing the thickness of the SiO_2 interlayer and the EOT.

9 Conclusions and Further Work

In this work computational modelling has been used to investigate the properties of interfaces and defects in HfO_2. These findings provide an atomic scale insight into the mechanisms of conductive filament formation, and dielectric breakdown, in ReRAM and transistors. The results of this work can provide a better understanding for the continued development of these devices. The findings are summarised below.

The energetics of oxygen interstitials and vacancies, in bulk m-HfO_2 have been calculated. Negatively charged interstitial ions have been found to be particularly mobile within the 3-coordinated sublattice of m-HfO_2. Oxygen vacancies, particularly neutral vacancies, have been found to be less mobile. This suggests that processes involving the migration of O ions through m-HfO_2 are more efficient, via an O interstitial mechanism. This is particularly true for the electroforming process involved in dielectric breakdown, during which the high applied bias will mean the defects are neutral or negatively charged.

The stability of O vacancy aggregates was investigated. It was found that neutral vacancy aggregates were stable, and the binding energy per vacancy increased as the size of the aggregate increased. This supports the idea that conductive filaments are composed of vacancy-rich regions within m-HfO_2. Furthermore, it suggests that as

the aggregate grows it will remain stable. The shape of the aggregates was shown to have a significant effect on its stability, with the most stable aggregates being those which form the largest voids within the monoclinic lattice structure. The result of this is that CF growth is unlikely to be anisotropic.

The formation energy of nearest neighbour Frenkel defects within m-HfO_2 was calculated. These defects were found to be unstable and will recombine, unless electrons are added to the system to neutralise the positively charged vacancy. When electrons are added to the system, the resulting FD has a low barrier to formation. This barrier is further lowered when the NN FD forms next to a pre-existing vacancy.

The findings of chapters 4, 5, and 6 can be combined to provide an atomic scale mechanism for conductive filament formation during the electroforming process in bulk m-HfO_2. In this mechanism, a high bias is applied across the oxide, flooding the system with electrons. These electrons can be trapped at pre-exiting O vacancy sites. In a concerted process, two electrons trap at an oxygen vacancy and the charge of the vacancy repels the vacancy's oxygen nearest neighbour. As the O moves away from its lattice position, the trapped electrons occupy the vacancy left behind, forming a stable Frenkel pair consisting of O^{0}_{vac} and a O^{2-}_{int}. The negatively charged interstitial then diffuses away towards the anode, leaving behind a stable di-vacancy. The di-vacancy can then trap more electrons, to form more NN FDs, and a larger vacancy aggregate, eventually forming a conductive filament through the oxide.

A Pt/HfO_2 interface was constructed, and its properties investigated. It was found that the Fermi level of Pt lies in the middle of the HfO_2 band gap. Furthermore, both O vacancies and O interstitials are sable at the interface in O–poor or O–rich conditions, respectively. Despite this, the Pt electrode acts as a significant barrier and does not readily absorb O ions.

A mechanism for O scavenging in the Si/SiO_2/HfO_2/TiN stack was suggested by combining collected data on the properties of defects at the separate interfaces. The absorption of HfO_2 interfacial O ions, into nitrogen vacancies with TiN, was found to be favourable. It was found that, although the O interstitial would be the most

mobile O defect through the HfO_2 layer, it is likely to get trapped at the SiO_2/HfO_2 interface. This suggests that during a high temperature anneal, interfacial O atoms will be absorbed into nitrogen vacancies within TiN. The resulting neutral vacancies will have a thermodynamic drive to migrate through the HfO_2 and SiO_2 layers, to the Si/SiO_2 interface. The result of this process is the scavenging of O ions from the Si/SiO_2 interface, into the TiN layer, reducing the thickness of the SiO_2 layer.

The mechanisms presented here are based on thermodynamic parameters calculated during static, atomic scale calculations. Moving forwards, the mechanisms proposed here could be tested within large scale systems, using these parameters within macroscale kinetic Monte Carlo simulations. Furthermore, although these calculations were primarily focussed on HfO_2, the concepts can be more generally applied to other materials. In particular, defect creation in materials under irradiation and carrier injection is crucial for a wide range of technological applications and injected electrons may play a crucial role in the formation of defects in many more systems.

Bibliography

[1] J. Robertson, *ECS Transactions*, 2009, **19**, 579–591.

[2] Y. V. Pershin and M. Di Ventra, *Advances in Physics*, 2011, **60**, 145–227.

[3] C. Schindler, G. Staikov and R. Waser, *Appl. Phys. Lett.*, 2009, **94**, 072109.

[4] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, *Rev. Mod. Phys.*, 1992, **64**, 1045–1077.

[5] D. Sheppard, R. Terrell and G. Henkelman, *J. Chem. Phys.*, 2008, **128**, 134106.

[6] K. Shubhakar, K. L. Pey, N. Raghavan, S. S. Kushvaha, M. Bosman, Z. Wang and S. J. O'Shea, *Microelectron. Eng.*, 2013, **109**, 364–369.

[7] L.-Å. Ragnarsson, T. Chiarella, M. Togo, T. Schram, P. Absil and T. Hoffmann, *Microelectron. Eng.*, 2011, **88**, 1317–1322.

[8] D. Veksler, G. Bersuker, M. B. Watkins and A. L. Shluger, Reliability Physics Symposium, 2014 IEEE International, 2014, pp. 5B.3.1–5B.3.7.

[9] G. E. Moore, *Electronics*, 1965, **8638**, 114–117.

[10] G. E. Moore, Electron Devices Meeting, 1975 International, 1975, pp. 11–13.

[11] H. van Lente and A. Rip, in *Getting New Technologies Together: Studies in Making Sociotechnical Order*, ed. C. Disco and B. van der Meulen, de Gruyter, Berlin, 1998, ch. 7, p. 206.

[12] A. Kawamoto, *Ph.D. thesis*, Stanford Universtity, 2001.

[13] K. Nagai and Y. Hayashi, *Electron Devices, IEEE Transactions on*, 1988, **35**, 1145–1147.

[14] D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt and G. Timp, *Nature*, 1999, **399**, 758–761.

[15] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur and H. S. P. Wong, *Proceedings of the IEEE*, 2001, **89**, 259–288.

[16] D. J. DiMaria and E. Cartier, *J. Appl. Phys.*, 1995, **78**, 3883–3894.

[17] J. Robertson and P. W. Peacock, in *Materials Fundamentals of Gate Dielectrics*, ed. A. A. Demkov and A. Navrotsy, Springer, Dordrecht, 2005, pp. 179–214.

[18] M. Gutowski, J. E. Jaffe, C.-L. Liu, M. Stoker, R. I. Hegde, R. S. Rai and P. J. Tobin, *Appl. Phys. Lett.*, 2002, **80**, 1897–1899.

[19] M. Copel, M. Gribelyuk and E. Gusev, *Appl. Phys. Lett.*, 2000, **76**, 436.

[20] L. O. Chua, *Circuit Theory, IEEE Transactions on*, 1971, **18**, 507–519.

[21] B. D. Strukov, G. S. Snider, D. R. Stewart and R. W. Williams, *Nature*, 2008, **453**, 80–83.

[22] A. Mehonic, S. Cueff, M. Wojdak, S. Hudziak, O. Jambois, C. Labbe, B. Garrido, R. Rizk and A. J. Kenyon, *J. Appl. Phys.*, 2012, **111**, 074507.

[23] I. H. Inoue, S. Yasuda, H. Akinaga and H. Takagi, *Phys. Rev. B*, 2008, **77**, 035105.

[24] L. Goux, X. P. Wang, Y. Y. Chen, L. Pantisano, N. Jossart, B. Govoreanu, J. A. Kittl, M. Jurczak, L. Altimime and D. J. Wouters, *Electrochem. Solid-State Lett.*, 2011, **14**, H244.

[25] Y. B. Nian, J. Strozier, N. J. Wu, X. Chen and A. Ignatiev, *Phys. Rev. Lett.*, 2007, **98**, 146403.

[26] T. Tamura, T. Hasegawa, K. Terabe, T. Nakayama, T. Sakamoto, H. Sunamura, H. Kawaura, S. Hosaka and M. Aono, *Japn. J. Appl. Phys.*, 2006, **45**, L364–L366.

[27] Y. Dong, G. Yu, M. C. McAlpine, W. Lu and C. M. Lieber, *Nano Lett.*, 2008, **8**, 386–391.

[28] Y.-S. Lai, C.-H. Tu, D.-L. Kwong and J. S. Chen, *Appl. Phys. Lett.*, 2005, **87**, 122101.

[29] A. Sawa, *Materials Today*, 2008, **11**, 28–36.

[30] D. C. Kim, S. Seo, S. E. Ahn, D.-S. Suh, M. J. Lee, B.-H. Park, I. K. Yoo, I. G. Baek, H.-J. Kim, E. K. Yim, J. E. Lee, S. O. Park, H. S. Kim, U.-I. Chung, J. T. Moon and B. I. Ryu, *Appl. Phys. Lett.*, 2006, **88**, 202102.

[31] D.-H. Kwon, K. M. Kim, J. H. Jang, J. M. Jeon, M. H. Lee, G. H. Kim, X.-S. Li, G.-S. Park, B. Lee, S. Han, M. Kim and C. S. Hwang, *Nature Nanotech.*, 2010, **5**, 148–153.

[32] X. Li, C. H. Tung and K. L. Pey, *Appl. Phys. Lett.*, 2008, **93**, 072903.

[33] J. Wu and R. L. McCreery, *J. Electrochem. Soc.*, 2009, **156**, 29–37.

[34] D. S. Jeong, H. Schroeder and R. Waser, *Phys. Rev. B*, 2009, **79**, 195317.

[35] R. Waser and M. Aono, *Nat. Mater.*, 2007, **6**, 833–840.

[36] M.-J. Lee, C. B. Lee, D. Lee, S. R. Lee, M. Chang, J. H. Hur, Y.-B. Kim, C.-J. Kim, D. H. Seo, S. Seo, U.-I. Chung, I.-K. Yoo and K. Kim, *Nat. Mater.*, 2011, **10**, 625–630.

[37] H. Y. Lee, Y. S. Chen, P. S. Chen, T. Y. Wu, F. Chen, C. C. Wang, P. J. Tzeng, M.-J. Tsai and C. Lien, *Electron Device Letters, IEEE*, 2010, **31**, 44 –46.

[38] P. Broqvist and A. Pasquarello, Symposium H Characterization of Oxide/Semiconductor Interfaces for CMOS Technologies, 2007.

[39] K. McKenna and A. Shluger, *Appl. Phys. Lett.*, 2009, **95**, 222111.

[40] A. Foster, A. Shluger and R. Nieminen, *Phys. Rev. Lett.*, 2002, **89**, 1–4.

[41] M. Born and R. Oppenheimer, *Annalen der Physik*, 1927, **389**, 457–484.

[42] W. Kohn and L. Sham, *Phys. Rev.*, 1965, **140**, A1133.

[43] P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.

[44] D. R. Hartree, *Proc. Cambridge Phil. Soc.*, 1928, **24**, 89–110.

[45] V. Fock, *Z. Phys.*, 1930, **61**, 126–148.

[46] J. C. Slater, *Phys. Rev.*, 1930, **35**, 210–211.

[47] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.*, 1980, **45**, 566–569.

[48] J. Heyd, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.*, 2003, **118**, 8207–8215.

[49] J. Heyd and G. E. Scuseria, *J. Chem. Phys.*, 2004, **121**, 1187–1192.

[50] J. Heyd, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.*, 2006, **124**, 219906.

[51] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B*, 1976, **13**, 5188–5192.

[52] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comp. Phys. Comm.*, 2005, **167**, 103–128.

[53] J. R. Shewchuk, *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, Carnegie Mellon University technical report, Carnegie Mellon University, Pittsburgh, PA, USA, 1994.

[54] R. A. Marcus, *J. Chem. Phys.*, 1966, **45**, 4493–4499.

[55] M. L. McKee and M. Page, in *Computing Reaction Pathways on Molecular Potential Energy Surfaces*, John Wiley & Sons, Inc., 2007, pp. 35–65.

[56] W. A. Deer, R. A. Howie and J. Zussman, *Rock-Forming Minerals: Orthosilicates, Volume 1A:*, Geological Society, 1982.

[57] J. H. Schemel, *ATSM Manual on Zirconium and Hafnium*, American Society for Testing and Materials, 1977.

[58] M. Taghizadeh, M. Ghanadi and E. Zolfonoun, *J. Nucl. Mater.*, 2011, **412**, 334–337.

[59] N. N. Greenwood and A. Earnshaw, in *Chemistry of the Elements (Second Edition)*, ed. N. N. Greenwood and A. Earnshaw, Butterworth-Heinemann, Oxford, Second Edition edn., 1997, pp. 954–975.

[60] E. H. Huffman and L. J. Beaufait, *J. Am. Chem. Soc.*, 1949, **71**, 3179–3182.

[61] B. G. Schultz and E. M. Larsen, *J. Am. Chem. Soc.*, 1950, **72**, 3610–3614.

[62] L. Poriel, A. Favre-Règuillon, S. Pellet-Rostaing and M. Lemaire, *Sep. Sci. Technol.*, 2006, **41**, 1927–1940.

[63] R. Liu, in *Materials Fundamentals of Gate Dielectrics*, ed. A. A. Demkov and A. Navrotsy, Springer, Dordrecht, 2005, pp. 1–36.

[64] A. Dkhissi, G. Mazaleyrat, A. Esteve and M. Djafari Rouhani, *Phys. Chem. Chem. Phys.*, 2009, **11**, 3701–3709.

[65] S. Desgreniers and K. Lagarec, *Phys. Rev. B*, 1999, **59**, 8467–8472.

[66] J. Aarik, A. Aidla, A.-A. Kiisler, T. Uustare and V. Sammelselg, *Thin Solid Films*, 1999, **340**, 110–116.

[67] K. Cherkaoui, S. Monaghan, M. A. Negara, M. Modreanu, P. K. Hurley, D. O'Connell, S. McDonnell, G. Hughes, S. Wright, R. C. Barklie, P. Bailey and T. C. Q. Noakes, *J. Appl. Phys.*, 2008, **104**, 064113.

[68] M. Ritala, M. Leskelä, L. Niinistö, T. Prohaska, G. Friedbacher and M. Grasserbauer, *Thin Solid Films*, 1994, **250**, 72–80.

[69] M.-H. Cho, Y. S. Roh, C. N. Whang, K. Jeong, S. W. Nahm, D.-H. Ko, J. H. Lee, N. I. Lee and K. Fujihara, *Appl. Phys. Lett.*, 2002, **81**, 472–474.

[70] G. Bersuker, D. C. Gilmer, D. Veksler, P. Kirsch, L. Vandelli, A. Padovani, L. Larcher, K. McKenna, A. Shluger, V. Iglesias, M. Porti and M. Nafría, *J. Appl. Phys.*, 2011, **110**, 124518.

[71] R. Puthenkovilakam and J. P. Chang, *J. Appl. Phys.*, 2004, **96**, 2701–2707.

[72] S. Toyoda, J. Okabayashi, H. Kumigashira, M. Oshima, K. Ono, M. Niwa, K. Usuda and H. Hirashita, *J.Electron Spectrosc. Relat. Phenom.*, 2004, **137-140**, 141–144.

[73] G. Lucovsky, J. G. Hong, C. C. Fulton, Y. Zou, R. J. Nemanich and H. Ade, *J. Vac. Sci. Technol. B*, 2004, **22**, 2132–2138.

[74] N. Ikarashi and K. Manabe, *J. Appl. Phys.*, 2003, **94**, 480–486.

[75] W. J. Zhu, T. Tamagawa, M. Gibson, T. Furukawa and T. P. Ma, *Electron Device Letters, IEEE*, 2002, **23**, 649–651.

[76] E. Bersch, S. Rangan, R. A. Bartynski, E. Garfunkel and E. Vescovo, *Phys. Rev. B*, 2008, **78**, 085114.

[77] M. Balog, M. Schieber, M. Michman and S. Patai, *Thin Solid Films*, 1977, **41**, 247–259.

[78] K. Kukli, J. Ihanus, M. Ritala and M. Leskela, *Appl. Phys. Lett.*, 1996, **68**, 3737–3739.

[79] E. P. Gusev, E. Cartier, D. A. Buchanan, M. Gribelyuk, M. Copel, H. Okorn-Schmidt and C. D'Emic, *Microelectron. Eng.*, 2001, **59**, 341–349.

[80] Y.-S. Lin, R. Puthenkovilakam and J. P. Chang, *Appl. Phys. Lett.*, 2002, **81**, 2041–2043.

[81] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M.

Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03, Revision C.02*, 2004, Gaussian, Inc., Wallingford, CT, 2004.

[82] G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169–11186.

[83] G. Kresse and J. Furthmüller, *Comp. Mater. Sci.*, 1996, **6**, 15–50.

[84] W. J. Hehre, R. F. Stewart and J. A. Pople, *J. Chem. Phys.*, 1970, **51**, 2769.

[85] P. C. Hariharan and J. A. Pople, *Theoretic. Chim. Acta.*, 1973, **28**, 213–222.

[86] I. N. Levine, *Quantum Chemistry*, Pearson Prentice Hall, New Jersey, 2008.

[87] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov and P.-O. Widmark, *The Journal of Physical Chemistry A*, 2004, **108**, 2851–2858.

[88] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov and P.-O. Widmark, *The Journal of Physical Chemistry A*, 2005, **109**, 6575–6579.

[89] D. E. Woon and T. H. Dunning, *J. Chem. Phys.*, 1993, **98**, 1358–1371.

[90] P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 299–310.

[91] L. E. Roy, P. J. Hay and R. L. Martin, *J. Chem. Theory Comput.*, 2008, **4**, 1029–1031.

[92] N. Kaltsoyannis, *Chem. Phys. Lett.*, 1997, **274**, 405–409.

[93] J. Adam and M. D. Rogers, *Acta Crystallogr.*, 1959, **12**, 951.

[94] J. Robertson, *Rep. Prog. in Phys.*, 2006, **69**, 327.

[95] X. Zhao and D. Vanderbilt, *Phys. Rev. B*, 2002, **65**, 233106.

[96] H. Jiang, R. I. Gomez-Abal, P. Rinke and M. Scheffler, *Phys. Rev. B*, 2010, **81**, 085119.

[97] J. E. Jaffe, R. A. Bachorz and M. Gutowski, *Phys. Rev. B*, 2005, **72**, 144107.

[98] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, *Phys. Rev. Lett.*, 2008, **100**, 136406.

[99] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 224106.

[100] J. P. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys.*, 1996, **105**, 9982–9985.

[101] A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

[102] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

[103] A. B. Mukhopadhyay, J. F. Sanz and C. B. Musgrave, *Phys. Rev. B*, 2006, **73**, 115330.

[104] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller and F. Bechstedt, *Phys. Rev. B*, 2006, **73**, 045112.

[105] V. Fiorentini and G. Gulleri, *Phys. Rev. Lett.*, 2002, **89**, 266101.

[106] E. Cockayne, *Phys. Rev. B*, 2007, **75**, 1–8.

[107] D. Muñoz Ramo, A. L. Shluger, J. L. Gavartin and G. Bersuker, *Phys.Rev.Lett*, 2007, **99**, 155504.

[108] E. Cartier, B. P. Linder, V. Narayanan and V. K. Paruchuri, 2006 International Electron Devices Meeting, Vols 1 and 2, 2006, pp. 57–60.

[109] V. Narayanan, V. K. Paruchuri, E. Cartier, B. P. Linder, N. Bojarczuk, S. Guha, S. L. Brown, Y. Wang, M. Copel and T. C. Chen, *Microelectron. Eng.*, 2007, **84**, 1853–1856.

[110] M. Jo, H. Park, M. Chang, H.-S. Jung, J.-H. Lee and H. Hwang, *Microelectron. Eng.*, 2007, **84**, 1934–1937.

[111] E.-A. Choi and K. J. Chang, *Appl. Phys. Lett.*, 2009, **94**, 122901.

[112] H. J. Lim, Y. Kim, I. S. Jeon, J. Yeo, B. Im, S. Hong, B.-H. Kim, S.-W. Nam, H.-K. Kang and E. S. Jung, *Appl. Phys. Lett.*, 2013, **102**, 232909.

[113] K.-L. Lin, T.-H. Hou, J. Shieh, J.-H. Lin, C.-T. Chou and Y.-J. Lee, *J. Appl. Phys.*, 2011, **109**, 084104.

[114] M. Lanza, *Materials*, 2014, **7**, 2155–2182.

[115] X. Gao, Y. Xia, B. Xu, J. Kong, H. Guo, K. Li, H. Li, H. Xu, K. Chen, J. Yin and Z. Liu, *J. Appl. Phys.*, 2010, **108**, 074506.

[116] P. Lorenzi, R. Rao, G. Romano and F. Irrera, *Adv. Cond. Matt. Phys.*, 2015, 136938.

[117] B. Magyari-Kope, S. G. Park, H.-D. Lee and Y. Nishi, *J. Mater. Sci.*, 2012, **47**, 7498–7514.

[118] K. Nakajima, A. Fujiyoshi, Z. Ming, M. Suzuki and K. Kimura, *J. Appl. Phys.*, 2007, **102**, 064507.

[119] P. McIntyre, *ECS Trans.*, 2007, **11**, 235–249.

[120] P. Broqvist and A. Pasquarello, *Appl. Phys. Lett.*, 2006, **89**, 262904.

[121] N. Capron, P. Broqvist and A. Pasquarello, *Appl. Phys. Lett.*, 2007, **91**, 192905.

[122] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1992, **46**, 6671.

[123] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1993, **48**, 4978.

[124] S. Lany and A. Zunger, *Phys. Rev. B*, 2008, **78**, 235104.

[125] S. T. Murphy and N. D. M. Hine, *Phys. Rev. B*, 2013, **87**, 094111.

[126] A. S. Foster, F. Lopez Gejo, A. L. Shluger and R. M. Nieminen, *Phys. Rev. B*, 2002, **65**, 174117.

[127] G. Bersuker and D. Gilmer, in *Advances in Non-volatile Memory and Storage Technology*, ed. Y. Nishi, Woodhead Publishing, Cambridge, 2014, ch. 2, p. 288.

[128] J. Robertson and R. Gillen, *Microelectron. Eng.*, 2013, **109**, 208–210.

[129] S. Brivio, G. Tallarida, E. Cianci and S. Spiga, *Nanotechnology*, 2014, **25**, 385705.

[130] J. L. Gavartin, D. Muñoz Ramo, A. L. Shluger and G. Bersuker, *ECS Transactions*, 2006, **3**, 277–290.

[131] W. Zhang and Z. F. Hou, *Physica Status Solidi B*, 2013, **250**, 352–355.

[132] V. E. Bondybey, *Chem. Phys. Lett.*, 1984, **109**, 436–441.

[133] R. Grau-Crespo, K. C. Smith, T. S. Fisher, N. H. de Leeuw and U. V. Waghmare, *Phys. Rev. B*, 2009, **80**, 174117.

[134] G. Bersuker, C. S. Park, J. Barnett, P. S. Lysaght, P. D. Kirsch, C. D. Young, R. Choi, B. H. Lee, B. Foran, K. van Benthem, S. J. Pennycook, P. M. Lenahan and J. T. Ryan, *J. Appl. Phys.*, 2006, **100**, 094108.

[135] Y. Dai, Y. Zhao, J. Wang, J. Xu and F. Yang, *AIP Advances*, 2015, **5**, 017133.

[136] A. Padovani, L. Larcher, G. Bersuker and P. Pavan, *Electron Device Letters, IEEE*, 2013, **34**, 680–682.

[137] A. Padovani, L. Larcher, P. Padovani, C. Cagli and B. De Salvo, Memory Workshop (IMW), 2012 4th IEEE International, 2012, pp. 1–4.

[138] Y. Guo and J. Robertson, *Microelectron. Eng.*, 2015, **147**, 339–343.

[139] V. N. Kuzovkov, A. I. Popov, E. A. Kotomin, M. A. Monge, R. González and Y. Chen, *Phys. Rev. B*, 2001, **64**, 064102.

[140] S. A. Lombardo, J. H. Stathis and G. Bersuker, *ECS Transactions*, 2009, **19**, 177–194.

[141] B. Kaczer, R. Degraeve, N. Pangon and G. Groeseneken, *Electron Devices, IEEE Transactions on*, 2000, **47**, 1514–1521.

[142] X. Li, K. L. Pey, M. Bosman, W. H. Liu and T. Kauerauf, *Appl. Phys. Lett.*, 2010, **96**, 022903.

[143] Z. Xu, M. Houssa, R. Carter, M. Naili, S. De Gendt and M. Heyns, *J. Appl. Phys.*, 2002, **91**, 10127–10129.

[144] S. Chatterjee, Y. Kuo, J. Lu, J.-Y. Tewg and P. Majhi, *Microelectron. Reliab.*, 2006, **46**, 69–76.

[145] N. A. Chowdhury, D. Misra, G. Bersuker, C. Young and R. Choi, *J. Electrochem. Soc.*, 2007, **154**, G298–G306.

[146] D. P. Ioannou, S. Mittl and G. La Rosa, *IEEE Transactions on Device and Materials Reliability*, 2009, **9**, 128–134.

[147] M. Houssa, *High-k Gate Dielectrics*, Taylor & Francis, 2003.

[148] S. Kar, *High permittivity gate dielectric materials*, Springer, Berlin, 2013, p. 489.

[149] T. Grasser, *Bias Temperature Instability for Devices and Circuits*, Springer, 2013.

[150] A. Padovani, L. Larcher, O. Pirrotta, L. Vandelli and G. Bersuker, *Electron Devices, IEEE Transactions on*, 2015, **62**, 1998–2006.

[151] F. Nardi, S. Larentis, S. Balatti, D. C. Gilmer and D. Ielmini, *Electron Devices, IEEE Transactions on*, 2012, **59**, 2461–2467.

[152] B. Traore, P. Blaise, E. Vianello, E. Jalaguier, G. Molas, J. F. Nodin, L. Perniola, B. De Salvo and Y. Nishi, Reliability Physics Symposium, 2014 IEEE International, 2014, pp. 5E.2.1–5E.2.5.

[153] L. Goux, P. Czarnecki, Y. Y. Chen, L. Pantisano, X. P. Wang, R. Degraeve, B. Govoreanu, M. Jurczak, D. J. Wouters and L. Altimime, *Appl. Phys. Lett.*, 2010, **97**, 243509.

[154] L. Goux, N. Raghavan, A. Fantini, R. Nigon, S. Strangio, R. Degraeve, G. Kar, Y. Y. Chen, F. De Stefano, V. V. Afanas'ev and M. Jurczak, *J. Appl. Phys.*, 2014, **116**, 134502.

[155] M. Lanza, K. Zhang, M. Porti, M. Nafría, Z. Y. Shen, L. F. Liu, J. F. Kang, D. Gilmer and G. Bersuker, *Appl. Phys. Lett.*, 2012, **100**, 123508.

[156] J. McPherson, J.-Y. Kim, A. Shanware and H. Mogul, *Appl. Phys. Lett.*, 2003, **82**, 2121–2123.

[157] J. L. Gavartin, D. Muñoz Ramo, A. L. Shluger, G. Bersuker and B. H. Lee, *Appl. Phys. Lett.*, 2006, **89**, 082908.

[158] J. L. Gavartin and A. L. Shluger, *Phys. Rev. B*, 2001, **64**, 245111.

[159] K. Xiong, J. Robertson, M. C. Gibson and S. J. Clark, *Appl. Phys. Lett.*, 2005, **87**, 183505.

[160] P. N. Murgatroyd, *J. Phys. D: Appl. Phys.*, 1970, **3**, 151.

[161] A. Rose, *Phys. Rev.*, 1955, **97**, 1538–1544.

[162] N. D. Arora, R. Rios and C.-L. Huang, *Electron Devices, IEEE Transactions on*, 1995, **42**, 935–943.

[163] H.-Y. Lee, P. S. Chen, C.-C. Wang, S. Maikap, P.-J. Tzeng, C.-H. Lin, L.-S. Lee and M.-J. Tsai, *Jpn. J. Appl. Phys.*, 2007, **46**, 2175.

[164] V. V. Afanas'ev, A. Stesmans, L. Pantisano, S. Cimino, C. Adelmann, L. Goux, Y. Y. Chen, J. A. Kittl, D. Wouters and M. Jurczak, *Appl. Phys. Lett.*, 2011, **98**, 132901.

[165] H. Y. Lee, Y. S. Chen, P. S. Chen, P. Y. Gu, Y. Y. Hsu, S. M. Wang, W. H. Liu, C. H. Tsai, S. S. Sheu and P. C. Chiang, Electronic Devices Meeting (IEDM), 2010 IEEE International, 2010, pp. 19.7.1–19.7.4.

[166] S. A. Chambers, *Surf. Sci. Rep.*, 2000, **39**, 105–180.

[167] P. V. Sushko, A. L. Shluger and C. R. A. Catlow, *Surf. Sci.*, 2000, **450**, 153–170.

[168] P. V. Sushko, J. L. Gavartin, and A. L. Shluger, *J. Phys. Chem. B*, 2002, **106**, 2269–2276.

[169] K. Xiong, J. Robertson, G. Pourtois, J. Pétry and M. Müller, *J. Appl. Phys.*, 2008, **104**, 074501.

[170] C. L. Hinkle, R. V. Galatage, R. A. Chapman, E. M. Vogel, H. N. Alshareef, C. Freeman, E. Wimmer, H. Niimi, A. Li-Fatou, J. B. Shaw and J. J. Chambers, *Appl. Phys. Lett.*, 2010, **96**, 103502.

[171] P.-Y. Prodhomme, F. Fontaine-Vive, A. van der Geest, P. Blaise and J. Even, *Appl. Phys. Lett.*, 2011, **99**, 022101.

[172] A. A. Demkov, *Phys. Rev. B*, 2006, **74**, 085310.

[173] H. Zhu, C. Tang and R. Ramprasad, *Phys. Rev. B*, 2010, **82**, 1–10.

[174] A. V. Gavrikov, A. A. Knizhnik, A. A. Bagaturyants, B. V. Potapkin, L. R. C. Fonseca, M. W. Stoker and J. Schaeffer, *J. Appl. Phys.*, 2007, **101**, 014310.

[175] E. Cho and S. Han, *Microelectron. Eng.*, 2011, **88**, 3407–3410.

[176] O. Sharia, K. Tse, J. Robertson and A. A. Demkov, *Phys. Rev. B*, 2009, **79**, 125305.

[177] J. Paier, M. Marsman and G. Kresse, *J. Chem. Phys.*, 2007, **127**, 024103.

[178] J. P. Perdew and A. Zunger, *Phys. Rev. B*, 1981, **23**, 5048.

[179] J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

[180] E. Cho and S. Han, *Microelectron. Eng.*, 2011, **88**, 3407–3410.

[181] R. W. G. Wyckoff, *Crystal Structures 2nd edition*, Interscience, New York, 1963.

[182] C. Kittel, in *Introdution to Solid State Theory*, Wiley, New York, 6th edn., 1986, pp. 55–57.

[183] A. Kokalj and M. Causa, *J. Phys.: Condens. Matter*, 1999, **11**, 7463–7480.

[184] J. Aarik, A. Aidla, H. Mändar, V. Sammelselg and T. Uustare, *J. Crystal Growth*, 2000, **220**, 105–113.

[185] G. N. Derry and Z. Ji-Zhong, *Phys. Rev. B*, 1989, **39**, 1940–1941.

[186] L. Fonseca, P. Prodhomme and P. Blaise, *Journal Integrated Circuits and Systems*, 2007, **2**, 94–103.

[187] K. Reuter and M. Scheffler, *Phys. Rev. B*, 2001, **65**, 035406.

[188] W. J. Zhu, T.-P. Ma, T. Tamagawa, J. Kim and Y. Di, *Electron Device Letters, IEEE*, 2002, **23**, 97–99.

[189] T. Ando, *Materials*, 2012, **5**, 478.

[190] M. J. Hopstaken, J. Bruley, D. Pfeiffer, M. Copel, M. M. Frank, E. Cartier, T. Ando and V. Narayanan, *ECS Transactions*, 2010, **28**, 105–113.

[191] F. De Stefano, M. Houssa, V. V. Afanas'ev, J. A. Kittl, M. Jurczak and S. A., *Thin Solid Films*, 2013, **533**, 15–18.

[192] S. R. Bradley, K. P. McKenna and A. L. Shluger, *Microelectron. Eng.*, 2013, **109**, 346–350.

[193] J. Gavartin and A. Shluger, *Microelectron. Eng.*, 2007, **84**, 2412–2415.

[194] M. Guidon, J. Hutter and J. VandeVondele, *J. Chem. Theory Comput.*, 2009, **5**, 3010–3021.

[195] M. Guidon, J. Hutter and J. VandeVondele, *Journal of Chemical Theory and Computation*, 2010, **6**, 2348–2364.

[196] J. VandeVondele and J. Hutter, *J. Chem. Phys.*, 2007, **127**, 114105.

[197] S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B*, 1996, **54**, 1703–1710.

[198] M. Krack, *Theor. Chem. Acc.*, 2005, **114**, 145–152.

[199] W. L. Scopel, A. J. R. da Silva, W. Orellana and A. Fazzio, *Appl. Phys. Lett.*, 2004, **84**, 1492–1494.

[200] C. Tang and R. Ramprasad, *Phys. Rev. B*, 2007, **75**, 241302.