Imperial College London

Computational Screening of Photoelectrochemical Electrodes

Liam Anthony Harnett-Caulfield

September 2024

Department of Materials

Imperial College London

Submitted in part fulfilment of the requirements for the degree of

Doctor of Philosophy of Imperial College London

and the Diploma of Imperial College London

Abstract

As the emphasis on carbon-free energy sources grows, so does the pressure to produce higher efficiency renewable and non-polluting energy harvesting methods. In recent years, photoelectrochemical water splitting has become an attractive method for generating hydrogen directly from water using solar energy. This is because, in principle, no such carbon is involved in the combustion process, with the only product being water. However, device development has been slowed by issues concerning the properties of the electrode semiconducting materials. Such electrode materials must simultaneously satisfy a crucial set of criteria to be considered economical. To date, no system has satisfied all of these criteria simultaneously. Such a multivariate problem is ideally suited for study via a computational materials screening and/or inverse materials design procedure.

Here, three examples of low-cost computational screening procedures that consider these criteria are investigated. Firstly, a low-cost means to screen the bulk component of the ionisation potential, based on past work by Frensley and Kroemer, was assessed in light of modern computational tools. Secondly, the formation of self-trapped small polarons was considered within a high throughput framework. Thirdly, special quasi-random structure models were used to rapidly predict the homogeneity and catalytic activity of a range of rutile oxide alloys. Additionally, each of these three examples was reviewed and summarised in terms of five suggested criteria for an ideal computational materials screening procedure.

It is hoped that the results and discussion of this work will provide the reader with an understanding of the advantages and limitations of these computational materials screening tools. These tools may be used in a modular fashion, with or without machine learning protocols, to accelerate the prediction of suitable candidate electrode materials for photoelectrochemical cell development.

Statement of Originality

The information disclosed in this thesis, unless otherwise stated, is entirely my own. All work that is not my own has been appropriately referenced within this thesis.

Copyright Statement

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

I would like to thank **Prof. Aron Walsh & Prof. James Durrant**

I would also like to acknowledge **Dr Michael Sachs** (of Imperial College London), **Prof. Ifan Stephens** (of Imperial College London), **Dr Soren Scott** (of Imperial College London), **Dr Kazuki Morita** (of Imperial College London), and **Dr Alexander Bagger** (of Imperial College London).

List of Tables

4.1 Tabulated data for Figure 4.34. 136

4.2 Tabulated data for Figures 4.37, 4.38, and 4.39. C denotes the fully converged data whilst Γ denotes the Γ-point data. The $BiVO_4$ electron, $BiVO_4$ hole, and Co_3O_4 electron self-trapping energies at the Γ-point are omitted due to the band state energies not being fully converged with respect to the total energy. These values could be obtained in future calculations by increasing the maximum number of electronic minimisation steps beyond the default values provided by VASP. 137

5.1 Thermodynamic properties of the RuTi SQS unit cells. 159

A.1 Tabulated data for Figures 3.3, 3.10a, 3.11. 198

A.2 Tabulated calculated data and reference[1,2,3] lattice parameter data for Figures 3.9 and 3.10a. 199

A.3 Tabulated data for Figures 3.8 and 3.9. 200

A.4 Tabulated data for Figures 3.11 and 3.11. 201

A.5 Tabulated data for Figure 3.15. 202

A.6 Tabulated reference[1,2,4,5,6,7,8,9,3] and calculated band gap data for Figure 3.10. 203

A.7 Supercell parameters that were used for self-trapping calculations displayed in Section 4.2. Listed are the number of primitive unit cells included in the supercells and their converged Monkhorst-Pack grid densities. Note that for magnetic systems, such as NiO and CuO, the primitive cells correspond to the *magnetic* primitive cell and not the structural primitive cell which ignores any magnetic ordering of the lattice. Non-magnetic primitive cell structure files were sourced from the *Materials Project* database[10]. Magnetic primitive cell structure files were sourced from the *MAGNDATA* database[11,12,13,14]. Note also that the non-localising or non-converging polaron calculations do not have converged **k**-point data because they were discarded from further convergence testing. 210

A.8 Tabulated data for Figure 5.6. 211

A.9 Continued: Tabulated data for Figure 5.6. *AVG*, *MIN* and *STD* give the average, minimum and standard deviation respectively of all 8 adsorption site energies. 211

A.10 Continued: Tabulated data for Figure 5.6. 212

A.11 Continued: Tabulated data for Figure 5.6. 213

A.12 Continued: Tabulated data for Figure 5.6. 214

A.13 Continued: Tabulated data for Figure 5.6. 215

A.14 Continued: Tabulated data for Figure 5.6. 216

A.15 Continued: Tabulated data for Figure 5.6. 216

A.16 Continued: Tabulated data for Figure 5.6. 217

A.17 Continued: Tabulated data for Figure 5.6. 218

A.18 Continued: Tabulated data for Figure 5.6. 219

A.19 Continued: Tabulated data for Figure 5.6. 220

A.20 Continued: Tabulated data for Figure 5.8. *RHS* and *LHS* denote the right-hand-side and left-hand-side values respectively for the volcano plot. 221

A.21 Raw DFT Energies for Figure 5.10. 222

A.22 Adsorption Energies for Figure 5.10. 223

A.23 Tabulated data for Figure 5.10. *RHS* and *LHS* denote the right-hand-side and left-hand-side values respectively for the volcano plot. 223

C.1 List of 7 lattice systems. When drawn from a common lattice point, a, b and c correspond to lattice constants in Euclidean space. α, β, and γ correspond to the angles between b and c, a and c, and a and b respectively. 270

List of Figures

1.1 A simplified diagram of the operation of a hydrogen fuel cell that illustrates the key reactions involved and the electrodes at which they occur. H^+ represents a proton, whilst e^- represents an electron. 27

1.2 A simplified diagram of the operation of a photoelectrochemical cell that illustrates the key reactions involved and the electrodes at which they occur. $h\nu$ represents the absorption of electromagnetic (solar) radiation which must generate a minimum voltage of 1.23V to facilitate the water splitting reaction. e^- represents an electron and h^+ represents a hole. The positioning of the OER below the HER further illustrates their relationship with the electron kinetic energy. Since both electrons and holes naturally seek to minimise their respective kinetic energies, a photoexcited electron will use the HER to move from a high-energy electronic state to a lower-energy electronic state. Conversely, due to its opposite charge, a hole will use the OER to move towards higher energy electronic states. This is commonly analogised as electrons *"sinking"* and holes *"floating"* across the electron energy spectrum. Thus, the HER is presented here as being higher in energy than the OER. 29

1.3 A sketch of the idealised relation between the wavelength of absorption and the *spectral response* of a photoabsorbing semiconducting material. The spectral response is the current generated by the device divided by power of the incident photons. As such, the spectral response ends abruptly at λ_{E_g} which is the wavelength corresponding to the band gap energy (E_g). This is via $\lambda_{E_g} = \frac{hc}{E_g}$, where h is Planck's constant and c is the speed of light[15]. Real (non-ideal) systems exhibit a less uniform dependency of the spectral response on the absorption wavelength. As such, the ideality of a material's spectral response may be quantified via the loss of area under this sketch. 34

1.4 Illustration of a refined model of a band gap in a semiconductor or insulator. The curvature of the band edges is known as the *dispersion*. In this figure, the valence band maximum and conduction band minimum are not in alignment in reciprocal space. This is known as an *indirect* band gap semiconductor. If they were aligned, it would be a *direct* band gap semiconductor. 35

1.5 A simplified diagram of the difference between a well-aligned and poorly-aligned semiconductor for water splitting. A well-aligned semiconductor (i.e. semiconductor A) must have band edges that straddle at least one of the hydrogen/oxygen evolution reaction redox potentials. 36

2.1 Illustration of the difference in the distribution of electronic states between metals, intrinsic semiconductors, and insulators. 46

2.2 Plot of $f(\epsilon)$ across 3 different temperatures. In this case, $E_F = 0.5$eV. This figure shows that the probability of an electron occupying higher energy states increases with the temperature. 48

2.3 Illustrative example of how the primitive unit cell is insufficient to describe the magnetic structure of copper (II) oxide. In the primitive cell (Left) the copper atoms at the corners of the unit cell have opposing magnetic moments (given by the red arrows). This contradicts the periodicity requirements of the unit cell and would not be useful for computational study. In contrast, the larger, magnetic primitive unit cell (Right) properly accounts for the long-range magnetic ordering and unit cell periodicity requirements but results in an increased computational demand. 50

2.4 Example of a routinely calculated bandstructure for chromium oxide (Cr_2O_3). The VBM and and CBM are given by the green and red dots respectively. The blue lines represent bands of electronic states that are composed primarily of oxygen p orbitals, whilst the red lines represent chromium d orbital dominated bands. The corresponding *density of states* (the number of quantum states per unit of energy[16]) up to 3 eV is presented to the right of the figure. This figure was produced using the *sumo* package[17]. *Gamma*, L, F and Z represent characteristic, high symmetry points, in reciprocal space, within the first Brillouin zone. It is worth noting that this particular bandstructure was produced using density functional theory, using the PBE functional, which tends to underestimate the size of the band gap. 55

2.5 Convergence testing data of the plane wave cutoff energy for rutile titanium dioxide. The point of this figure is to illustrate how various quantities plateau after a certain value. In this case, there is no need to increase the plane wave cutoff energy beyond 500 eV. Different quantities (*total energy*, surface energy, ionisation potential, etc.) may plateau at different cutoff energies. 66

2.6 Convergence testing data of the Monkhorst-Pack grid density for a 50:50 germanium-lead alloy (See the *Catalytic Activity* chapter for a further discussion of such compounds). Similarly to Figure 2.5, there is no need to increase the Monkhorst-Pack grid density beyond 5 units, after which the total energy is consistent within an acceptable margin of error. 67

2.7 Plot of the variation in the total energy across various lattice parameters. The lowest energy point interpolated from this data, via an Equation of State, would be used as the lattice parameter for further calculations. 68

2.8 Illustration of the interaction between a central defect (O) and its periodic images within a supercell of a square lattice composed of lattice points (X). The red arrows represent this interaction. The dashed lines designate the arbitrarily defined bounds of the unit cell. 69

3.1 Illustration of how the band gap, ionisation potential (IP), electron affinity (EA), vacuum potential (ϕ), conduction band minimum (CBM), and valence band maximum (VBM) relate to one another. 72

3.2 Illustration of how the electrostatic potential varies across the bulk, surface, and vacuum regions of a slab model. The interface dipole determines how the electrostatic potential varies across the boundary between the interface and vacuum regions. Thus, affecting the extent to which the vacuum potential differs from the bulk periodic potential. . . . 73

3.3 A plot of the data initially compiled by Tung[18]. The reference data was compiled by Pelatt et al[19]. Data points that are close to the *Reference* line indicate a stronger agreement with the reference data. It can be seen that the results of Frensley and Kroemer (*F&K*) appear to be no worse than those of Van de Walle and Martin (*VDW & M*) or Harrison. This figure was previously published by the author and their supervisor[20]. 75

3.4 The $F\bar{4}3m$ Zinc Blende primitive unit cell. O_h and T_d are marked in green and red respectively. This image was produced using the VESTA program[21,22,23]. This figure was previously published by the author and their supervisor[20]. 79

3.5 The $Fm\bar{3}m$ Rocksalt primitive unit cell. The interstice is marked in green. This image was produced using the VESTA program[21,22,23]. 84

3.6 The $Pm\bar{3}m$ primitive unit cell. The interstice is marked in green. This image was produced using the VESTA program[21,22,23]. 85

3.7 Illustration of the FFT grid sampling using the MacroDensity package[24,20]. Plot of the electrostatic potential across the [111] vector in the Zinc Blende lattice, which spans the O_h and T_d interstices. The interstices are positioned at the crests of the peaks in the potential. The array of coloured points in the lower portion illustrates the relationship of points in the sampling volume. The black points represent the sampling volume origin with 2x2x2 and 3x3x3 sampling volumes being represented in green and red respectively. This figure was adapted from work previously published by the author and their supervisor[20]. 87

3.8 Variation in the VBM according to the sampling volume. Each number (n) in the legend represents an n x n x n grid of FFT grid points. The closer each grid is to pointwise sampling, the less local variation there must be in the electrostatic potential about the interstices. The PBEsol functional was used in this case. However, the same qualitative trends were observed for each of the other functionals and more coarse FFT grids. This figure was previously published by the author and their supervisor[20]. 88

3.9 (Top) Difference in the VBM between a pointwise and 10x10x10 sampling volume versus the lattice parameter. (Bottom) Difference between the reference VBM[19] and the calculated VBM with single point sampling versus the lattice parameter. The data presented here was produced using the PBEsol functional but the same trend is observed for each of the other functionals. The data shows that the calculated VBM is less representative of the experimental result for smaller lattice parameter compounds. This figure was previously published by the author and their supervisor[20]. 89

3.10 (a) Difference in the calculated and reference valence band maxima versus the lattice parameter of the compound. (b) Comparison of the calculated band gaps with their reference values from Pelatt[19]. Data points that are close to the black line indicate a strong agreement. The data shows that the calculated HSE06 band gaps and ionisation potentials are more representative of both the band gap and the ionisation potential. This figure was previously published by the author and their supervisor[20]. 90

3.11 Comparison between this work and Frensley and Kroemer's original work[25,26,27,28]. Data points that are close to the black line indicate a strong agreement for that compound. This figure was previously published by the author and their supervisor[20]. 91

3.12 Plot of the band alignments for the compounds investigated in this chapter. The lower bars of each colour represent the VBM whilst the upper bars represent the CBM. This figure was previously published by the author and their supervisor[20]. 92

3.13 Plot of the band alignments for the compounds investigated in this chapter using the Standard Hydrogen Electrode as the reference potential. The red and blue markers represent the electron affinity and ionisation potential of each compound respectively. The dotted and dashed lines represent the HER and OER redox potentials according to the standard hydrogen electrode (-0.41 V and 0.82 V respectively[29]). Only the HSE06 values are included here, for simplicity and because they provide the best overall agreement with experimental data. 93

3.14 Electrostatic (Hartree) potential contour plots for the AlAs (1-01) plane, the MgO (1-01) plane, and the $CsPbI_3$ (101) plane. The interstices are marked using O_h, T_d and *Int.* These images were produced using the VESTA software[21,22,23]. This figure was previously published by the author and their supervisor[20]. 94

3.15 (Top) Comparison of the Frensley-Kroemer and Bader charge ionicities. The ionicity metrics are normalised such that 0.0 implies a fully covalent compound, whilst 1.0 implies a fully ionic compound. (Bottom) Comparison of the Frensley and Kroemer and Phillips ionicity. For both the top and bottom plots, the points that are close to the black line indicate a strong agreement. This figure was previously published by the author and their supervisor[20]. 95

4.1 Simplified illustration of Radiative, Non-radiative, and Auger recombination for an electron in the conduction band and a hole in the valence band. $h\nu$ represents the emission of a photon. $\omega(\mathbf{k})$ represents the transfer of electron and hole energy to a phonon. The state in the middle of the Non-radiative diagram represents the recombination of the electron and hole via recombination centres. 99

4.2 Illustration of small polaron transport based on a configuration coordinate diagram. Here, the polaron moves (or *"hops"*) from the donor site (A) to the acceptor site (B) upon receiving enough thermal energy to overcome the activation energy barrier. The parabolic shape of each site is due to a harmonic approximation of the available vibrational states at each site. 101

4.3 Illustration of small polaron formation based on Marcus theory. In this case, the polaron state is lower in energy than the band state, implying that self-trapping is likely. However, there may still be a thermal activation barrier for the self-trapping mechanism, which is represented by the crossing point of each parabola, as for Figure 4.2. In some cases, there may be no such barrier to self-trapping. These systems are described as being *"barrierless"*[30]. 102

4.4 Example of the electron band and polaron states for titanium dioxide using the bond distortion method. For the band state, it can be seen that the charge is shared evenly across the unit cell, indicating delocalisation. For the polaron state, it can be seen that the charge is highly localised on a d(xy) orbital of a (blue) titanium atom, with some participation of the d(z^2) orbitals of the nearest neighbour titanium atoms. The black arrows indicate the extent to which the attached atoms have moved during relaxation. As such, it can be seen that the nearest neighbour (red) oxygen atoms move outwards from the central titanium atom, likely due to electrostatic repulsion as a result of the excess charge density. 104

4.5 Illustration of the anatase and rutile TiO_2 crystal phases. The titanium atoms are given in blue. These images were produced using the Vesta[21,22,23] software. 106

4.6 Illustration of the Fe_2O_3 and Cr_2O_3 crystal structures. The green and yellow atoms represent the alternate spin states of the iron and chromium atoms that determine the magnetic structures. As such, it can be seen that the Fe_2O_3 and Cr_2O_3 magnetic structures differ in that the spin states are distributed across different Miller planes throughout each crystal. This image was produced using the Vesta[21,22,23] software. . . 107

4.7 Illustration of the $BiVO_4$ and Bi_2O_3 crystal structures. The bismuth atoms are given in purple. The vanadium atoms are the larger red atoms. This image was produced using the Vesta[21,22,23] software. 108

4.8 Illustration of the CdO and Co_3O_4 crystal structures. The cadmium atoms are given in purple. The green and yellow atoms represent the alternate spin states of the cobalt atoms that define the magnetic structure. Spin inactive cobalt atoms are given in blue. This image was produced using the Vesta[21,22,23] software. 109

4.9 Illustration of the Cu_2O and CuO crystal structures. The spin inactive copper atoms are given in blue. The green and yellow atoms represent the spin states of the copper atoms that define the magnetic structure. This image was produced using the Vesta[21,22,23] software. 110

4.10 Illustration of the In_2O_3 and MoO_3 crystal structures. The indium atoms are given in pink. The molybdenum atoms are given in grey. This image was produced using the Vesta[21,22,23] software. 111

4.11 Illustration of the NiO and SnO crystal structures. The green and yellow atoms represent the alternate spin states of the nickel atoms that define the magnetic structure. The tin atoms are given in grey. This image was produced using the Vesta[21,22,23] software. . . 112

4.12 Illustration of the $SrTiO_3$ crystal structure and WO_3 orthorhombic crystal structure. The strontium atoms are given in green and the titanium atoms are given in blue. The tungsten atoms are given in grey. This image was produced using the Vesta[21,22,23] software. 113

4.13 Illustration of the monoclinic and rutile VO_2 crystal phases. The green and yellow atoms represent the alternate spin states of the vanadium atoms that define the magnetic structures. These images were produced using the Vesta[21,22,23] software. 114

4.14 Illustration of the ZnO crystal structure. The zinc atoms are given in grey. This image was produced using the Vesta[21,22,23] software. 115

4.15 Fe_2O_3 polaron isosurfaces. It can be seen that the hole polaron is centred around an iron atom, with a substantial contribution from the iron d$(x^2 - y^2)$ orbital and nearest neighbour oxygen p orbitals in the d(xy) plane. There is little to no contribution from the oxygen p along the d(z^2) axis. The contributing oxygen atoms are shown to move inwards towards the metal centre, presumably to compensate for the positive charge. The electron polaron takes the form of a Fe-Fe dimer, where the iron d(xy) orbitals on nearest neighbour iron atoms share the added charge. This results in a shortening of the Fe-Fe distance in the dimer and a lengthening of the distance between oxygen atoms that bridge the two iron atoms. 118

4.16 Binding energies of Fe_2O_3 using Γ-point calculations. The horizontal axis represents the initial configuration of the system, whilst the description on each bar denotes the resulting configuration following a structural relaxation. *"Fe-centred"* means that the resulting polaron was centred around a single iron atom and/or that the calculation was initiated as such. *"Dimer"* means that the resulting polaron was centred around two iron atoms. *"Fe-Fe distance modified"* means that the localisation was induced by moving two nearest neighbour iron atoms closer together or further apart. *"Rattled"* means that the ASE rattle function was used before the structural relaxation. It should be noted that these calculations were performed during the benchmarking stage of this investigation, using Γ-point sampling and a smaller Fe_2O_3 unit cell. Thus, the precise values of the binding energies differ from those given later in this chapter. Nevertheless, the same lowest energy morphology was reproduced for the larger unit cell. 119

4.17 Cr_2O_3 polaron isosurfaces. It can be seen that both the electron and hole polarons are centred around a chromium atom. In both cases, the chromium d(z^2) orbital is shown to be involved in the binding of the unpaired spin. In the hole case, all six nearest neighbour oxygen p orbitals are involved, whereas in the electron case, the oxygen p orbitals along the chromium d(z^2) axis contribute noticeably more. The atomic displacements reflect this distinction with the hole polaron nearest neighbour oxygen atoms moving towards the central chromium atom in approximately equal measure, presumably to compensate for the positive charge. For the electron polaron, the d(z^2) axial nearest neighbours move away from the charge more strongly than the oxygen atoms in the d(xy) plane of the central chromium, presumably to accommodate the added negative charge. 120

4.18 TiO_2 (rutile) polaron isosurfaces. It can be seen that the hole polaron is centred around an oxygen atom, with the central oxygen p orbital and its nearest neighbours contributing the most to the binding of the unpaired spin. The eight nearest neighbour oxygen p orbitals contribute evenly to the polaron. The nearest neighbour titanium atoms also move outwards from the polaron centre. The electron polaron is centred around a titanium atom. The titanium d(xy) orbital contributes most strongly to its stabilisation, with the four nearest neighbour oxygen atoms in the d(xy) plane moving outwards from the polaron, likely due to the electrostatic repulsion. 120

4.19 TiO_2 (anatase) polaron isosurfaces. It can be seen that the hole polaron is centred around an oxygen atom with the stabilising orbital contribution coming primarily from the central oxygen p and four nearest neighbour oxygen p orbitals. The nearest neighbour titanium atoms move outwards from the polaron site whereas the nearest neighbour oxygen atoms move inwards, presumably to compensate for the removed electron. The central oxygen atom moved downwards, closer to the plane of the contributing nearest neighbours. This may imply a stabilisation interaction through oxygen p-oxygen p orbital overlap. No polaron formation could be observed for the excess electron, given the even distribution of charge density across the unit cell. 121

4.20 Two dimensional slice of the VO_2 (rutile) polaron charge density. It can be seen that both the electron and hole polarons are centred around a single vanadium atom. For the electron polaron, the d(xy) orbital appeared to contribute the most to the stabilisation of the polaron. For the hole polaron, there appears to be an absence of charge density relative to the surrounding atoms. 122

4.21 $BiVO_4$ polaron isosurfaces. It can be seen that the $BiVO_4$ electron polaron was centred around a single vanadium atom. The d(z^2) orbital of the central vanadium atom appears to accept the added charge. The hole polaron localises around a central bismuth atom with a substantial contribution from the neighbouring oxygen p orbitals, presumably to stabilise the added charge. 122

4.22 Co_3O_4 polaron isosurfaces. It can be seen that the Co_3O_4 electron polaron appeared as a dimer polaron. However, unlike Fe_2O_3, the two nearest neighbour atoms used the d(xy) and d(z^2) orbitals to bind the charge. Neighbouring oxygen p orbitals also appeared to contribute to the stabilisation of the electron polaron. The hole polaron appeared to be stabilised by the d(z^2) orbital of a central cobalt atom. 123

4.23 In_2O_3 polaron isosurfaces. It can be seen that the In_2O_3 hole polaron was centred around a single oxygen atom and was stabilised through neighbouring oxygen p orbitals. No substantial localisation was observed for the electron polaron. 123

4.24 $SrTiO_3$ polaron isosurfaces. The $SrTiO_3$ hole polaron was centred around a single oxygen atom and was stabilised through four neighbouring oxygen p orbitals in the same plane. These oxygen atoms moved towards the central atom accordingly. No substantial localisation was observed for the electron polaron, though there appears to be some asymmetry in the charge distribution. Perhaps indicating localisation under the right conditions. 124

4.25 WO_3 polaron isosurfaces. It can be seen that the hole polaron is centred around an oxygen atom, stabilised by nearest neighbour oxygen p orbitals. The electron polaron appeared to localise across a single plane of the WO_3 unit cell, but confirmation of this is limited by the bounds of the unit cell. 125

4.26 NiO polaron isosurfaces. The hole polaron appears to be centred around a nickel atom. The electron polaron appears to be centred around an oxygen atom and is stabilised via neighbouring nickel d(xy) orbitals. 125

4.27 CuO polaron isosurfaces. The hole polaron appears to be centred around a copper atom, with neighbouring oxygen atoms moving towards the central copper atom. A similar observation is made for the electron polaron, except that the nearest neighbour oxygen atoms move outwards from the central copper atom. 126

4.28 Bi_2O_3 polaron isosurfaces. It can be seen that the hole polaron is centred around a single bismuth atom, with surrounding oxygen p orbitals stabilising the added charge. No substantial localisation was observed for the electron polaron, as shown by the diffuse and even distribution of the charge density isosurface. 126

4.29 MoO_3 delocalised isosurfaces. For the excess hole, no substantial localisation was observed by virtue of the even distribution of charge density across the atoms in the unit cell. For the excess electron, there is a noticeable asymmetry in the charge density distribution. However, this does not correspond to a clearly definable small polaron morphology. 127

4.30 CdO delocalised isosurfaces. In the case of an excess hole, there is a noticeable asymmetry in the charge density distribution. However, this does not correspond to a definable small polaron morphology. The charge is otherwise evenly distributed across the unit cell. In the case of an excess electron, the charge density is evenly distributed across the unit cell, indicating no substantial localisation. 127

4.31 SnO delocalised isosurfaces. In both the electron and hole cases, the excess charge density is spread evenly and symmetrically throughout the unit cell, implying delocalisation. 128

4.32 Cu_2O delocalised isosurfaces. In both the electron and hole cases, the excess charge density is spread evenly and symmetrically throughout the unit cell, implying delocalisation. 128

4.33 ZnO delocalised isosurfaces. In both the electron and hole cases, the excess charge density is spread evenly and symmetrically throughout the unit cell, implying delocalisation. 129

4.34 Fully converged electron and hole self-trapping energies. The d-orbital occupancy is given at the top of each bar. The grey bar corresponds to the Co_3O_4 electron polaron calculation that was not pre-converged with a coarse **k**-point grid. 130

4.35 Band decomposed charge density of the highest energy, partially occupied states for the fully converged Co_3O_4 calculation. The cobalt orbitals involved in these states appear to be mixtures of d(xy) and d(z^2) orbitals. None of these morphologies correspond to those observed for the Γ-point calculations, nor those reported in past literature[31]. . . 131

4.36 (Left axis - blue line) Number of geometry optimisation (ionic) steps in the structural relaxation of the Fe_2O_3 electron polaron for the number of **k**-points. (Right axis - orange line) Number of computer cores, required to complete the calculation for a given number of **k**-points within a 24-hour interval. This figure shows that the number of required computer cores increased rapidly with the number of **k**-points. However, the majority of the ionic steps were accounted for with a single **k**-point. 132

4.37 (a) Comparison of the self-trapping energies using a single **k**-point with the fully converged **k**-point grids. The distance from the black line indicates the degree to which the two values are in agreement. (b) Comparison of the overall lattice distortions using a single **k**-point with the fully converged **k**-point grids. The *Lattice Distortion* quantity was obtained using the *getdelQ.py* code of the *carrier capture* utility[32]. The *getdelQ.py* code obtains the mass-weighted coordinate difference between two structural input files and is thereby a means to represent the overall distortion within a unit cell. The distance from the black line indicates the degree to which the two values are in agreement. . . . 134

4.38 Comparison of the fully converged polaron and band state total energies with their respective Γ-point energies. The *"Errors"* displayed in this figure are the differences between the fully converged total energies and their corresponding Γ-point total energies. 135

4.39 Comparison of the fully converged self-trapping energies with the overall lattice distortion. The distance from the black line indicates the degree to which the two values are in agreement. 137

5.1 Examples of the structural output for each intermediate for the OER adsorption calculation. 145

5.2 50:50 Rutile SQS template unit cell. This unit cell contains sixteen AO_2 formula units. 150

5.3 Plot of the volume scans for $Ti_{0.5}Pt_{0.5}O_2$ and $Sn_{0.5}Pt_{0.5}O_2$, wherein the internal parameters are held constant (static) and where they are allowed to relax at each volume (ISIF = 4 in the VASP syntax). Polynomial fits are also included to distinguish the 4 plots. 151

5.4 (Left) Transformed 8x supercell using the ASE supercell optimisation tool. (Right) Illustration of the slab adsorption site numbering. 152

5.5 Thermodynamic quantities for each of the 50:50 rutile mixtures. The row and column for a particular cell give the corresponding chemical species. The colour scale indicates the relative magnitude of the quantity in each cell which reflects the relative likelihood of each composition to mix homogeneously; green being a high likelihood of mixing and red being a low likelihood of mixing. 155

5.6 Plot of ΔE^{OOH} versus ΔE^{OH} for each of the 50:50 rutile mixtures. The black line represents the universal scaling relation of Equation 5.6. The *"Site"* numbering corresponds to the numbering of each adsorption site as presented in Figure 5.4b. This figure shows that the majority of the data from this work is in accordance with the aforementioned universal scaling relation. 156

5.7 The corresponding structure for the outlier at approximately 23 eV on the E^{HOO} axis in Figure 5.6, presented at two different angles. It can be seen that the oxygen atoms of the hydroperoxyl adsorbate have each bound to a separate metal ion as opposed to a single oxygen atom binding to a single metal atom. Moreover, the hydrogen atom appears to have dissociated from the adsorbate and has bound to an adjacent bridging oxygen atom of the slab model. 157

5.8 Volcano plot of the adsorption energies from this work. The black line is a linear plot of η to define the volcano. The red triangles correspond to the pure compounds and the blue circles correspond to the average adsorption energies of sites 1-8 for each of the mixtures (including the aforementioned outliers of Figure 5.6). The horizontal error bars correspond to the standard deviation of the mean for each mixture's adsorption energy. 158

5.9 SQS structures for the RuTi intermediate compositions generated via ICET[33]. 159

5.10 Volcano plot of the adsorption energies. The black line plots the ideal scaling relations that form the volcano. The red triangles correspond to the pure compounds. The blue circles correspond to the Ti adsorbate binding sites, whilst the green circles correspond to the Ru adsorbate binding sites. The error bars are the standard deviation between the two binding sites for each species (ruthenium or titanium). 160

C.1 An example of the Fourier series applied to a square wave function, represented by the dashed line, for 3 different values of n. 266

C.2 A 2-dimensional representation of a square lattice. Each X may represent a lattice point and/or an atomic position within the crystal structure. 269

C.3 The first and irreducible Brillouin zone for the square lattice of Figure C.2. The space enclosed by the solid box represents the first Brillouin zone, whilst the space enclosed by the dashed line represents the irreducible Brillouin zone. For reference, Γ corresponds to the coordinates (0,0,0) in reciprocal space. 271

Contents

1 Introduction **23**

1.1 The Ever Growing Demand for Energy 23

1.1.1 Non-Renewable Energy Sources 23

1.1.2 Renewable Energy Sources 24

1.2 Hydrogen Fuel Cells 26

1.3 Photoelectrochemical Water Splitting 28

1.3.1 Photoelectrochemical Cells 28

1.3.2 Criteria for an Efficient Photoelectrochemical Cell 32

1.3.3 Computational Screening and the Inverse Design of Materials 38

1.4 Motivation and Outline for Subsequent Chapters 41

2 Electronic Structure Theory and Methods **44**

2.1 Motivation 44

2.2 Electrons in Insulators, Semiconductors, and Metals 44

2.2.1 Electron Spin 48

2.3 The Electron Wavefunction 49

2.3.1 The Schrödinger Equation 51

2.4 Solutions to the Time Independent Schrödinger Equation in a Periodic Potential . . . 53

2.4.1 Crystal Momentum, Reciprocal Space, and the Blöch Theorem 53

2.4.2 Mapping Reciprocal Space 55

2.5 Approximations to the Time Independent Schrodinger Equation 56

2.5.1 The Born-Oppenheimer Approximation 56

2.5.2 Basis Sets 57

2.5.3 Pseudopotentials 57

2.5.4 Koopmans' and Janak's Theorems . . . 58
2.5.5 The Hartree Approximation . . . 58
2.5.6 The Hartree-Fock Approximation . . . 59
2.6 Density Functional Theory . . . 61
2.7 Practical Considerations . . . 64
2.7.1 Electronic Minimisation Algorithms . . . 65
2.7.2 Convergence Testing . . . 66
2.7.3 Geometry Optimisation Algorithms . . . 67
2.7.4 Data Parallelisation . . . 68
2.8 Applications to Photoelectrochemical Cells . . . 69

3 Band Alignment & Ionicity **71**

3.1 Introduction . . . 71
3.1.1 Band Alignment . . . 71
3.1.2 Computational Band Alignment . . . 72
3.1.3 Ionicity . . . 76
3.2 Motivation and Materials of Interest . . . 78
3.2.1 Zinc Blende Type Structures . . . 79
3.2.2 Diamond Type Crystals . . . 83
3.2.3 Rocksalt Type Structures . . . 84
3.2.4 Perovskite Type Structures . . . 84
3.3 Methods . . . 86
3.4 Results and Discussion . . . 87
3.4.1 Band Alignment . . . 87
3.4.2 Ionicity . . . 93
3.5 Conclusion . . . 96

4 Self Trapping **98**

4.1 Introduction . . . 98
4.1.1 Types of Polarons . . . 100
4.1.2 Small Polaron Transport . . . 100

4.1.3 Small Polaron Formation . . . 101
4.1.4 Modelling Small Polarons with DFT . . . 101
4.2 Motivation and Materials of Interest . . . 104
4.3 Methods . . . 116
4.4 Results and Discussion . . . 117
4.4.1 Converged Polaron Morphologies . . . 117
4.4.2 Non-Converged Polaron Calculations . . . 124
4.4.3 Non-Localising Polaron Calculations . . . 126
4.4.4 Converged Self-Trapping Energies . . . 129
4.4.5 Reducing the Computational Cost . . . 132
4.4.6 Further Reducing the Computational Cost . . . 133
4.5 Conclusions and Future Work . . . 138

5 Catalytic Activity **141**

5.1 Introduction . . . 141
5.1.1 Rutile Alloys . . . 142
5.1.2 Simulating Alloy Structures . . . 142
5.1.3 Special Quasi-random Structures . . . 143
5.1.4 Estimating the Thermodynamic Properties with Density Functional Theory . . . 144
5.1.5 Estimating the Catalytic Activity with Density Functional Theory . . . 144
5.2 Motivation and Materials of Interest . . . 147
5.3 Methods . . . 150
5.3.1 SQS Structure Optimisation . . . 150
5.3.2 Thermodynamic Property Calculations . . . 151
5.3.3 SQS Slab Generation . . . 151
5.4 Results and Discussion . . . 154
5.4.1 Thermodynamic Properties . . . 154
5.4.2 Adsorption Energies . . . 155
5.4.3 Intermediate Compositions . . . 158
5.5 Conclusions and Future Work . . . 161

6 Summary **163**

Bibliography **166**

A Data tables **197**

A.1 Band Alignment 198

A.2 GULP Output Data 204

A.2.1 Zinc Blende 204

A.2.2 Rocksalt 206

A.2.3 Perovskite 208

A.3 Self Trapping 210

A.4 Corrosion Resistance 211

B Clarifying Background Matter **224**

B.1 The Fermi-Dirac Distribution 225

B.2 The Plausibility of the Schrödinger Equation 226

B.3 Born-Von Karman Periodic Boundary Conditions 228

B.4 The Bloch Theorem 229

B.4.1 Imposing Born-Von Karman Boundary Conditions 230

B.4.2 Imposing Reciprocal Lattice Conditions 230

B.5 The Born Oppenheimer Approximation 232

B.5.1 The Hellmann-Feynman Theorem 233

B.6 Density Functional Theory 235

B.6.1 The First Hohenberg-Kohn Theorem 235

B.6.2 The Second Hohenberg-Kohn Theorem 236

B.6.3 The Kohn-Sham Equation 237

B.7 The Projector Augmented Wave Method 239

B.8 The Murnaghan Equation of State 241

B.9 The Madelung Potential 242

B.10 The Ewald Summation 243

B.11 Special Quasi-Random Structures 247

B.11.1 Monte Carlo SQS 248

B.12 Derivation of the OER Scaling Relation . . . 249
B.12.1 Chemical Potential Equations in Terms of DFT Energies . . . 249
B.12.2 The Hydrogen Bonding Approximation . . . 250
B.12.3 The Oxygen Molecule Approximation . . . 251
B.12.4 Step One of the Oxygen Evolution Reaction . . . 251
B.12.5 Step Two of the Oxygen Evolution Reaction . . . 252
B.12.6 Step Three of the Oxygen Evolution Reaction . . . 252
B.12.7 The Final Step of the Oxygen Evolution Reaction . . . 253
B.13 Experimental Methods . . . 254
B.13.1 X-Ray Diffraction . . . 254
B.13.2 Stern-Gerlach Experiment . . . 254
B.13.3 Transient Absorption Spectroscopy . . . 255
B.13.4 Cyclic Voltammetry . . . 256
B.13.5 Photoelectron Spectroscopy . . . 256
B.13.6 Sputter Deposition . . . 257
B.13.7 Chemical Vapour Deposition . . . 257
B.13.8 Inductively Coupled Plasma Mass Spectrometry . . . 257
B.13.9 Optical Absorption Spectroscopy . . . 257
B.13.10Tauc Analysis . . . 258
B.13.11Nanoparticle Synthesis . . . 258

C Mathematical Background **259**

C.1 Domains, Ranges, Functions and Functionals . . . 260
C.2 Matrices . . . 261
C.2.1 Diagonalisation . . . 262
C.3 Eigenvalues, Eigenvectors and Eigenfunctions . . . 264
C.3.1 Eigenbases and Matrices . . . 264
C.4 Taylor Series . . . 265
C.4.1 Maclaurin Series . . . 265
C.5 Fourier Series and Fourier Transforms . . . 266

C.5.1 Fourier Integral 267
C.5.2 Fourier Transforms 267
C.5.3 Inverse Fourier Transforms 267
C.5.4 Discrete and Fast Fourier Transforms 267
C.6 Crystal Structures and Lattices 269
C.7 Special Functions and Operators 272
C.7.1 Complex Numbers 272
C.7.2 Euler's Formula 272
C.7.3 The Delta Function 273
C.7.4 The Gaussian Function 274

Introduction

1.1 The Ever Growing Demand for Energy

A concern in modern society is how to balance the ever-increasing demand for energy with the environmental impact associated with the type of fuel. Over time, modern tastes have to come to reject non-renewable fuels, in favour of renewable (self-replenishing) and non-polluting fuels. However, energy demand is expected to increase substantially over at least the next 20 years[34,35]. As such, large amounts of public and private spending are focused on facilitating both endeavours through the research and development of renewable energy sources. The most established renewable and non-renewable energy sources are summarised below.

1.1.1 Non-Renewable Energy Sources

Coal Existing in the solid state, the composition of coal may vary according to where it is mined and how it is processed before combustion. It is primarily composed of carbon with variable amounts of hydrogen, sulphur, oxygen, and nitrogen. Coal is moderately *energy dense*, meaning that it produces a moderate amount of energy for the volume of fuel consumed. The combustion products of coal include various levels of *carbon dioxide*, *carbon monoxide*, *nitrogen oxides*, *sulphur dioxide* and heavy metals, depending on the composition and the completeness of combustion. Carbon dioxide contributes substantially to the *greenhouse effect* which impacts the Earth's climate. Carbon monoxide, sulphur dioxide, and several oxides of nitrogen are known to be toxic. Furthermore, sulphur dioxide and nitrogen oxides are believed to contribute to the occurrence of acid rain[36].

Petroleum Existing in the liquid state, and producing fewer nitrogen oxides and sulphurous oxides than coal, petroleum is primarily composed of hydrocarbons that burn more completely. Petroleum is also more energy-dense than coal, making it a more efficient alternative. Consequently, less toxic waste, per unit volume, is produced when burning petroleum, but the production of carbon dioxide and methane remains an issue[37].

Natural Gas Existing in the gaseous state, and even less polluting than petroleum, natural gas is composed of shorter, more volatile hydrocarbons (primarily methane). Due to the greater hydrogen-to-carbon ratio, natural gas is believed to produce less carbon dioxide per unit of energy[38].

Nuclear The most energy-dense of the non-renewable fuel sources. Nuclear power has the benefits of high-efficiency energy production per unit mass of fuel consumed. It relies on the principle of nuclear fission, wherein unstable isotopes of a given atom[a] split into more stable products, producing energy and at least one neutron. These neutrons collide with other unstable isotopes, causing them to split, producing further energy and further neutrons, resulting in a chain reaction. This chain reaction is controlled through the use of neutron-absorbing materials. The energy produced during nuclear fission is substantial and can be used to generate steam that powers a turbine. This is similar to coal, petroleum, and natural gas (fossil fuel) energy production. However, the production of nuclear waste remains a long-term issue. This is because radioactive nuclear waste products may take an incredibly long time to decay before they are considered safe for human exposure[39].

1.1.2 Renewable Energy Sources

Biofuel Biofuel producers endeavour to grow high energy density crops that may be used to synthesise a combustible fuel[40]. To be considered renewable, the mass of carbon produced during the combustion of the fuel must be absorbed by the crop during its growth, thereby maintaining the atmospheric concentration of carbon dioxide. The main disadvantage of biofuels is the use of vast areas of otherwise agricultural land that would be required to grow such crops. Furthermore, the primary role of biofuels is to maintain the balance of carbon dioxide. It does not necessarily reduce the emission of other pollutants that are created during the combustion process.

Hydroelectric Hydroelectric power uses a turbine to exploit the flow of water from one location to another, thereby converting the mechanical energy of the flowing water into electrical energy[41]. The advantages of this method are that it is a relatively robust and reliable high-power production system.

[a] Usually uranium or plutonium.

A disadvantage of hydroelectric power is the loss of land and ecosystems caused by the construction and running of the hydroelectric facility. Moreover, such facilities are geographically limited to bodies of water and structural failures within a hydroelectric facility can be economically and ecologically catastrophic.

Wind Wind power functions similarly to hydroelectric power but exploits the flow of wind currents as opposed to the flow of water[42]. Wind turbines arguably lessen the loss of land and ecosystems compared to hydroelectric power. However, they do not produce a reliable energy flow due to the inconsistencies in wind currents.

Geothermal Geothermal power exploits the heat generated within the Earth's crust[43]. An advantage of this method is that it is reliable and passive once established. The disadvantages of geothermal power include the limited number of locations in which a geothermal plant may be established, the release of greenhouse gases from the earth's crust, and the high cost of constructing a geothermal facility.

Solar Solar power involves the conversion of absorbed solar photons into electrical energy. The advantage of solar energy is that it is a relatively safe and passive means of energy production, with minimal ecological impact. This is especially true when a panel is mounted on an existing structure, such as the roof of a building. Furthermore, the capital costs are relatively low, allowing solar panel use to be implemented in a decentralised fashion. The main disadvantage of solar energy generation is that the absorbed energy is diffuse, requiring a large active surface area to be economical. Additionally, much like wind power, the flow of energy is subject to changes in the weather and energy production is geographically inconsistent. For instance, regions with the greatest photovoltaic power potential are at 15-30 degrees north (North Africa and the Middle East) and 15-30 degrees south (Australia, South America) whilst regions closest to the poles (the Scandinavian nations) have the lowest photovoltaic power potential[44].

As such, it is common practice to store the electrical energy generated by solar panels in some form of *battery*, for a steadier overall flow of power and an overall increase in efficiency. Generally speaking, a battery is an electrochemical device that is capable of storing energy in the form of chemical bonds and then releasing that energy, when desired, as electrical energy. A battery usually consists of several *galvanic cells* that exploit the energy released by a *redox* (*reduction* and *oxidation*) reaction to drive the movement of electrons in a circuit[16]. Broadly speaking, a galvanic cell consists of at least two, oppositely charged electrodes that are immersed in an electrolyte. Each of these electrodes is responsible for a respective oxidation and reduction reaction, known as *half-reactions*[16]. The overall electrochemical equation for the operating galvanic cell may be deconstructed into its component half-reactions.

The apparatus responsible for a given half-reaction is called the *half-cell*. The main issue with battery technology is the degradation of performance across a finite number of charging and discharging cycles. Moreover, the recycling of such degraded batteries is, as of yet, not economical nor ecologically conserving[45]. An alternative to battery storage is to feed back absorbed solar energy directly into the grid, though this would require significant changes to the existing energy infrastructure[46].

1.2 Hydrogen Fuel Cells

To meet the forecast energy demand projections, whilst minimising the ecological impact of using such energy, the generative capacity of renewable energy sources and the storage capacity of battery technology must increase. A potentially useful area to assist in this endeavour is to mitigate the disadvantages of solar energy and battery production with the energy-dense, combustible, and non-polluting *hydrogen* fuel. Hydrogen fuel cells are another example of a galvanic cell that facilitates the controlled combustion of hydrogen in the same manner as a fossil fuel. However, because hydrogen and oxygen are the only reagents in the combustion reaction, the only expected product is water.

$$2\,H_2 + O_2 \longrightarrow 2\,H_2O$$

Generally speaking, a hydrogen fuel cell includes a cathode, anode, and electrolyte. At the anode, an electrochemical reaction occurs that splits hydrogen molecules into *protons* (positively charged hydrogen ions) and *electrons* (negatively charged particles). The electrons produced at the anode are fed through an external circuit to the cathode whilst the protons move from the anode to the cathode via the electrolyte. At the cathode, the electrons and protons recombine with oxygen molecules to form water. A load may be placed in the external circuit between the anode and the cathode, allowing energy to be extracted from the system to do work[47]. This is illustrated in Figure 1.1.

Besides the hydrogen fuel cells themselves, which act as the short-term storage and delivery method of hydrogen energy, more robust long-term hydrogen storage solutions have also been considered. In general, the underground storage of hydrogen is preferred over the overground (surface) storage. This is due to the reduced fire risk, greater protection from military or criminal sabotage, and better integration of the technology with urban infrastructure. The nature of the underground storage vessel is an important consideration in this regard, with many current vessels exploiting natural formations or existing infrastructure. These include salt/hard rock caverns, aquifers, depleted oil and gas reservoirs, and depleted mines. Despite these options, *lined rock caverns* are increasing in popularity due to their

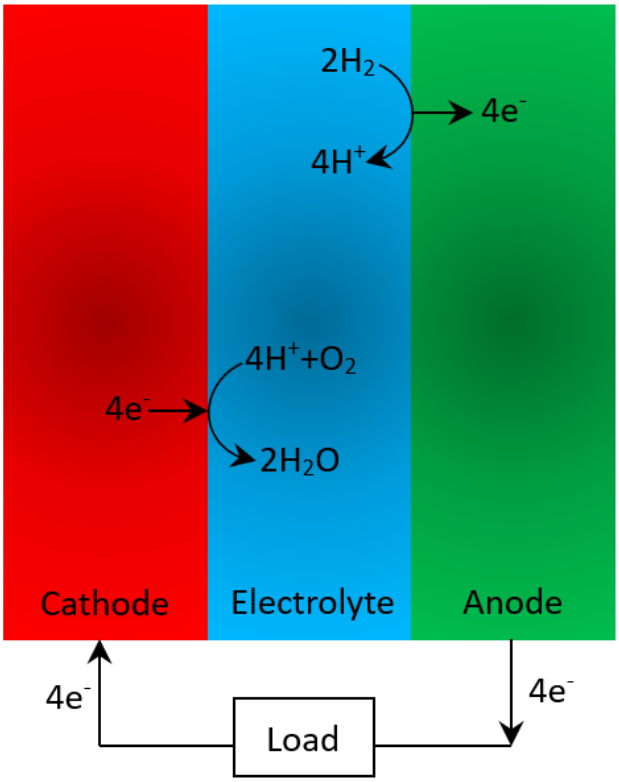


Figure 1.1: A simplified diagram of the operation of a hydrogen fuel cell that illustrates the key reactions involved and the electrodes at which they occur. H^+ represents a proton, whilst e^- represents an electron.

reduced permeability with respect to the stored hydrogen, reduced corrosion of the storage vessel, and the higher operating temperatures[48]. Lined rock caverns are formed by excavating a hard rock formation and then lining the cavern created with concrete, followed by steel or plastic. This provides a gas-tight and inert inner surface of the hydrogen storage vessel that remains structurally stable at high pressures[48].

Despite these measures, there are still further challenges associated with the storage of hydrogen. One of which is the *Joule-Thomson effect*, wherein the temperature of the hydrogen increases as it flows through valves and porous membranes. This is contrary to the usual behaviour of gases which decrease in temperature under the same conditions. This creates an atypical safety and efficiency challenge for hydrogen gas transport systems to and from the lined rock cavern. Another challenge concerning lined rock caverns is the possibility of *hydrogen embrittlement*, wherein the hydrogen gas corrodes the steel lining of the cavern, causing structural vulnerabilities within the lining over prolonged storage periods. This is often mitigated by using lower storage pressures, applying corrosion-resistant coatings, or using more corrosion-resistant steel alloys[49].

Beyond these challenges, the key issue with using hydrogen as a fuel is obtaining a sustainable source

of hydrogen. Unlike fossil fuels, which are mined or drilled and then extracted from the earth's crust, hydrogen is almost exclusively synthesised industrially. Ironically, the vast majority of hydrogen is produced through the combustion of fossil fuels[50], particularly the steam reforming of natural gas[29]. This defeats the purpose of using hydrogen as a fuel in the first place. However, if there were a means with which to efficiently split water into hydrogen and oxygen using renewable and non-polluting energy sources, then hydrogen fuel would become a much more attractive prospect.

1.3 Photoelectrochemical Water Splitting

The concept of water electrolysis (or *water splitting*[51]) is an established means with which to generate hydrogen and oxygen for use within a hydrogen fuel cell. The principle is essentially the reverse reaction of the hydrogen fuel cell and conveniently uses a similar apparatus.

$$2\,H_2O \longrightarrow 2\,H_2 + O_2$$

In this system, an anode and cathode are immersed in an aqueous electrolyte and a minimum voltage of 1.23V[b] is applied to both electrodes via an external circuit. In practice, an *overpotential* of 1.48 V is often required to account for loss mechanisms via heat production[52]. At the cathode, water is reduced, producing hydrogen gas. At the anode, water is oxidised, producing oxygen gas. These product gases typically bubble to the surface of the aqueous solution and can be readily extracted. Since this process is driven by an applied voltage, the energy that is required to produce the hydrogen must come from an external source. Devices that use solar energy to drive the water splitting reaction are called *photoelectrochemical cells*. In literature, a distinction is often made between *"photocatalytic"*, *"photoelectrochemical"*, and even *"artificial photosynthesis"*[29]. However, it is respectfully submitted that most of the device requirements are the same for such systems. Thus, for this work, the term *"photoelectrochemical"* will be applied generally.

1.3.1 Photoelectrochemical Cells

The primary appeal of photoelectrochemical water splitting is that, in principle, the reaction produces no controversial by-products (carbon dioxide, sulphur dioxide, nuclear waste, etc.). Furthermore, the entire process is stoichiometrically renewable and the main reagent (water) is in great abundance. By directly storing the absorbed energy into H-H and O-O bonds, water splitting may also have the

[b] which is the thermodynamic barrier for water splitting[52]

potential to improve the efficiency of the solar-generated power per photon because it bypasses the need for intermediate storage systems and their associated loss mechanisms.

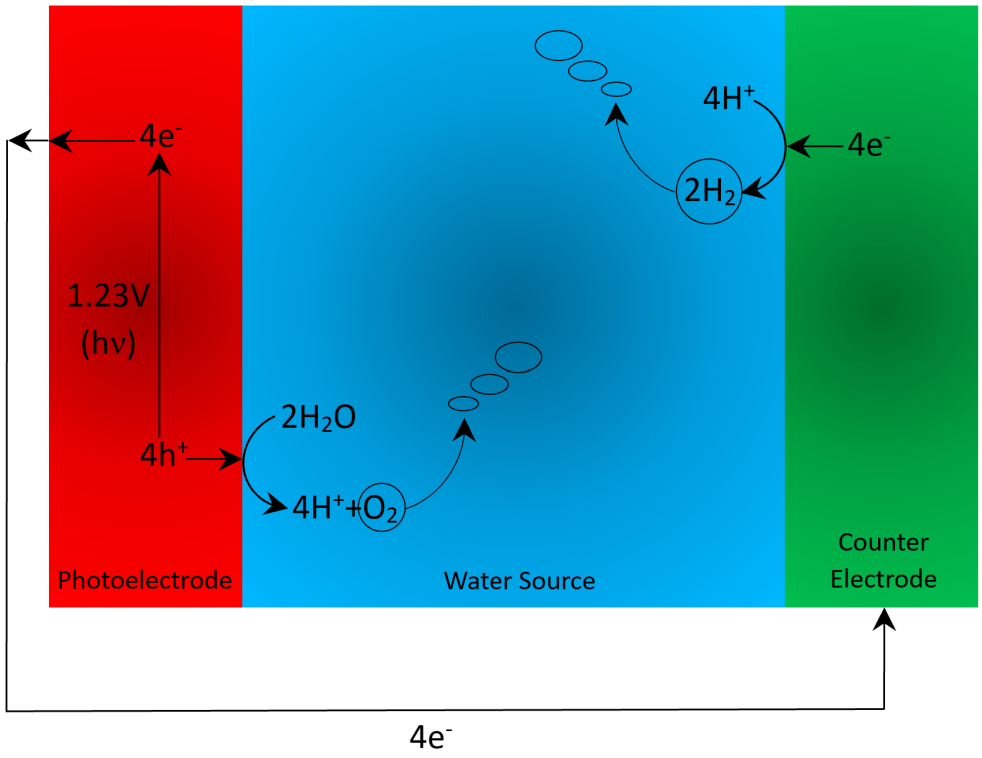


Figure 1.2: A simplified diagram of the operation of a photoelectrochemical cell that illustrates the key reactions involved and the electrodes at which they occur. $h\nu$ represents the absorption of electromagnetic (solar) radiation which must generate a minimum voltage of 1.23V to facilitate the water splitting reaction. e^- represents an electron and h^+ represents a hole. The positioning of the OER below the HER further illustrates their relationship with the electron kinetic energy. Since both electrons and holes naturally seek to minimise their respective kinetic energies, a photoexcited electron will use the HER to move from a high-energy electronic state to a lower-energy electronic state. Conversely, due to its opposite charge, a hole will use the OER to move towards higher energy electronic states. This is commonly analogised as electrons *"sinking"* and holes *"floating"* across the electron energy spectrum. Thus, the HER is presented here as being higher in energy than the OER.

As a tangential point, photoelectrochemical cells are not merely limited to the water splitting reaction. Most notably, a similar apparatus may be used to facilitate the conversion of carbon dioxide into liquid phase compounds[53] (often referred to as *carbon capture*[54]). Aside from possibly reducing the atmospheric carbon dioxide concentration, carbon capture also has the potential to use the captured carbon to make useful side products such as plastics, biofuels and alcohols[53]. A photoelectrochemical cell operates in a similar fashion for both water splitting and carbon capture applications, with the reagents, electrode properties, and required voltages differing according to whichever reaction is facilitated. This work will hereon focus solely on the problem of water splitting, but the methods employed may still apply to any similar photoelectrochemical process if properly reparameterised.

A photoelectrochemical cell usually comprises a semiconducting photoelectrode that absorbs solar

radiation and catalyses the water splitting reaction, a source of water that is in contact with the photoelectrode that serves as the reagent, a counter electrode to balance the electrochemical reaction, and a means to collect the product gases[55]. The first photoelectrochemical cell was demonstrated by Fujishima and Honda in 1972[56], using a titanium dioxide photoelectrode and a platinum counter electrode. During operation, electrons within the photoelectrode are excited to higher energy states via the absorption of electromagnetic (solar) radiation. In the absence of these electrons, effective positive charges are created within the lower energy states, known as *holes*. Holes pass through the photoelectrode material and contribute to the electrical current, much like an electron, but move in the opposite direction to the electron flow. Electrons and holes are referred to collectively as *charge carriers* because they are the particles that carry electrical charge within the system. The electrons are fed, via an external circuit, to the counter electrode where they are used to catalyse the conversion of protons to hydrogen gas. This is known as the *hydrogen evolution reaction* (*HER*). Meanwhile, the holes are used to catalyse the conversion of water to oxygen gas and protons, which replenish the supply of protons lost during the hydrogen evolution reaction. This is known as the *oxygen evolution reaction* (*OER*). The product gases (hydrogen and oxygen) may then bubble to the surface of the water source for collection. This is illustrated in Figure 1.2.

The Oxygen Evolution Reaction

The oxygen evolution reaction is believed to proceed via four key steps. The first step involves the dissociation of a proton (H^+) and electron (e^-) from the water molecule (H_2O), upon contact with a metal atom (M) on the electrode's surface. Following this step, the remaining hydroxyl species (OH) remains bound to the metal atom via the oxygen atom:[57,58,59]

Step 1: $\mathrm{M + H_2O \longrightarrow M(OH) + H^+ + e^-}$

The second step involves the dissociation of a second proton and electron from the water molecule, in a similar fashion to the previous step. The oxygen atom remains bound to the metal atom.

Step 2: $\mathrm{M(OH) \longrightarrow M(O) + H^+ + e^-}$

The third step involves a second water molecule binding to the remaining oxygen atom and dissociating a further proton and electron. This results in a hydroperoxyl group (OOH) remaining bound to the metal atom.

Step 3: $\mathrm{M(O) + H_2O \longrightarrow M(OOH) + H^+ + e^-}$

The fourth and final step is the dissociation of the metal-bound hydroperoxyl group into an oxygen molecule, a proton, and an electron.

Step 4: $\mathrm{M(OOH)} \longrightarrow \mathrm{M} + \mathrm{O_2} + \mathrm{H^+} + \mathrm{e^-}$

As implied by Figure 1.2, the electrons produced in this reaction would be collected by the photoelectrode and fed through the external circuit, whilst the protons would be released into the water source and would travel to the counter electrode due to the electrostatic attraction. The oxygen evolution reaction is believed to be the rate-limiting reaction of the water splitting process due to its requirement for a particular multi-electron and multi-proton transfer mechanism to result in the formation of the relatively weak O-O bond[29].

The Hydrogen Evolution Reaction

Assuming that the same protons and electrons of the oxygen evolution reaction are used to generate hydrogen, the hydrogen evolution reaction proceeds via two key steps[60]. The first step involves a proton from the water source and an electron from the external circuit combining at the surface of the counter electrode to form a hydrogen atom bound to a metal atom of the counter electrode (MH).

Step 1: $\mathrm{M} + \mathrm{H^+} + \mathrm{e^-} \longrightarrow \mathrm{MH}$

The second step may involve either a second proton and electron combining with the bound hydrogen atom, to produce a hydrogen molecule, or it may involve two adjacent bound hydrogen atoms combining to form a hydrogen molecule:

Step 2: $\mathrm{MH} + \mathrm{H^+} + \mathrm{e^-} \longrightarrow \mathrm{M} + \mathrm{H_2}$ or $2\,\mathrm{MH} \longrightarrow 2\,\mathrm{M} + \mathrm{H_2}$

It is noted that the above reactions represent a neutral or slightly acidic water source. However, the reaction may also proceed in alkaline solutions via a slightly different mechanism[60].

Photoelectrochemical Cell Design

Beyond the simplified picture of the apparatus described in Figure 1.2, photoelectrochemical cell designs can be further grouped into 3 varieties:[51]

1. A planar photoelectrode is simply immersed into an electrolyte, similar to a galvanic cell. This is the setup that was used by Fujishima and Honda[56].

2. Instead of a planar electrode, a mixture of H_2 and O_2 catalysing *nanoparticles* is added as a *colloidal suspension* to an aqueous electrolyte. A colloidal suspension comprises a *dispersed phase* (nanoparticles) within a *dispersion medium*[16]. Colloids differ from solutions in that they *scatter* (change the direction of) light, as opposed to allowing light to pass through unperturbed[16]. Bard[61,62,63] is believed to have first introduced nanoparticle-based photoelectrochemical cells[29].

3. The H_2 and O_2 catalysing components carry out their respective half-reactions (producing either hydrogen or oxygen) in independent electrolyte solutions. A means to circulate the H^+ ions between the two electrolyte solutions is provided to complete the electrochemical circuit. A clear advantage of this system is that the product gases may be more easily separated.

1.3.2 Criteria for an Efficient Photoelectrochemical Cell

Whatever the device design, for a photoelectrochemical cell to be economical, it must simultaneously satisfy several criteria. These include a sufficient *photovoltage* that is generated following solar irradiation, the *band gap* of the semiconducting material must be such that it can absorb a large range of solar wavelengths that are mostly within the visible range, the absolute band energies of the photoelectrode and counter electrode must closely align with the redox (reduction and oxidation) potentials of the hydrogen/oxygen evolution reactions, the photoelectrochemical cell components must be *corrosion resistant* under the working conditions (acidic, basic, chlorinated, etc.) over an extended period of time, the reaction with water at the surface of the photoelectrode must cost as little energy as possible (hereon referred to as *catalytic activity*), and the reaction with water at the surface of the photoelectrode must selectively produce only the desired products (hereon referred to a *catalytic selectivity*)[55]. Each of these criteria is elaborated on below.

The Photovoltage is the electromotive force that is generated upon irradiation of the photoelectrochemical cell electrode[64]. The electromotive force is the energy transferred to the electrical circuit per unit charge[16]. As such, the photovoltage determines the effectiveness of the electrode material for converting light energy into an electrical current. The photovoltage is primarily limited by loss mechanisms that occur during the operation of the photoeletrochemical cell. The cumulative effect of such loss mechanisms within the system causes the photovoltage to always be lower than that of the band gap[55c]. Mechanistically, the photovoltage is often reduced via the *recombination* of charge

[c] via $E_{pv} = QV_{pv}$, where E_{pv} is the energy transferred through the electrical circuit, Q is the charge transferred, and V_{pv} is the photovoltage.

carriers within the semiconducting material. This is the opposite process of the aforementioned *separation* of charge carriers, into electrons and holes, that is caused by the absorption of solar energy. Recombination is a problem for photoelectrochemical cells because if an electron and hole recombine within the material, the energy that they carry may be lost through the production of heat or light, as opposed to being used to do useful work. Many aspects of the photoelectrochemical cell may reduce the photovoltage. Notably including, but not limited to, any *intrinsic*/extrinsic defects of the electrode material, the dielectric properties of the electrode material, any unwanted material-material interactions[65], and any unwanted electrode-electrolyte interactions[66]. These aspects are briefly summarised below.

An intrinsic (or native) defect is a defect that is present in the material under the given operating conditions[d]. Intrinsic defects arise as a result of the balance between the lattice enthalpy (H) of the crystal and the configurational entropy (S), which is dependent on the temperature (T)[67].

$$G = H - TS \tag{1.1}$$

where G is the *Gibbs Free Energy*. As such, the formation of intrinsic defects can, to various extents, be controlled through appropriate annealing protocols. Examples of intrinsic defects are missing atoms (*vacancies*), 'misplaced' atoms in the interstitial sites of the crystal structure (*interstitials*), and the interfaces between two crystallite grains (*grain boundaries*) within the crystal structure. Extrinsic defects are caused by external influences, such as the presence of impurity atoms within the material. Such defects often act as non-radiative (heat producing) recombination centres within the bulk of the semiconductor.

The dielectric properties (i.e. dielectric constant/tensor) of the material have the effect of screening the electrostatic attraction between the electrons and holes within the material, reducing the chance of them recombining. As such, a reasonably high dielectric constant is often desirable, to ensure that the charge carriers remain separated after photogeneration.

Unwanted material-material interactions relate to how well the system is configured to support the transport of charge across material-material interfaces. As such this is a particularly important consideration for multi-layered PEC devices, where each layer serves a specific purpose in the device (e.g. photoabsorber layer, passivation layer, catalyst layer, etc.)[65]. As such, and only where desirable, charges would need to be able to pass unimpeded through the material-material interfaces between these layers. This depends on the precise device design and can be achieved through carefully selecting

[d] *"Operating conditions"* is a general term that includes (but is not limited to) the applied potential (voltage/bias), the pH of the electrolyte, the purity of the electrolyte, the quality of the electrode, and the temperature.

complementary charge transporting materials and by using robust device fabrication protocols.

At the electrode-electrolyte interface, particularly if the photoabsorbing material is immersed in the electrolyte, photogenerated charges may also directly catalyse an unwanted side-reaction within the electrolyte. For example, depending on the operating conditions, the photoexcited electrons in the photoelectrode of Figure 1.2 could react directly with an energetically aligned contaminant within the electrolyte, as opposed to passing through the external circuit[66]. Moreover, due to the increased likelihood of defects near the surface of a material, electronic states at the electrode surface may also act as recombination centres for photogenerated charges.

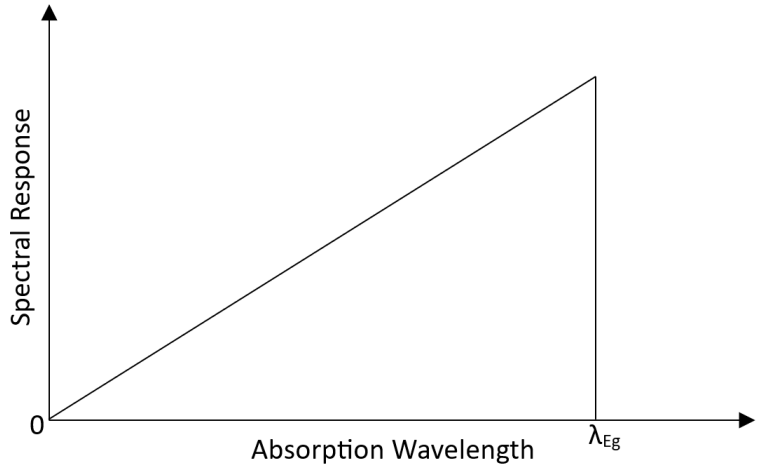


Figure 1.3: A sketch of the idealised relation between the wavelength of absorption and the *spectral response* of a photoabsorbing semiconducting material. The spectral response is the current generated by the device divided by power of the incident photons. As such, the spectral response ends abruptly at λ_{E_g} which is the wavelength corresponding to the band gap energy (E_g). This is via $\lambda_{E_g} = \frac{hc}{E_g}$, where h is Planck's constant and c is the speed of light[15]. Real (non-ideal) systems exhibit a less uniform dependency of the spectral response on the absorption wavelength. As such, the ideality of a material's spectral response may be quantified via the loss of area under this sketch.

The Band Gap is a substantial gap in energy between the allowable electronic energy bands of a semiconducting or insulating material. The size of this gap determines the ability of the material to absorb photons at particular wavelengths. For instance, photons with an energy below the band gap energy are not absorbed by the material and therefore do not contribute to the electrical current (See Figure 1.3). However, a band gap is essential for the separation of electrons and holes within the material[68]. Hence, it is desirable to select semiconductors with band gaps that are appropriately sized such that a large portion of the solar spectrum is absorbed. This is approximately 1.3 eV according to the *Shockley-Queisser Limit*[e]. Furthermore, in light of the overpotential and the particular device configuration, a band gap between 1.5 eV - 2.5 eV is often considered as acceptable for electrochemical water splitting catalysis[52,29].

[e]The Shockley-Queisser Limit considers the necessary conditions (including the optimal band gap) for the maximum theoretical efficiency obtainable for a single p-n junction solar cell[69].

Beyond the size of the band gap, the alignment of the electronic states at the edges of the gap is also an important consideration. The highest energy, occupied band of electronic states, is known as the *valence band*. The lowest energy, unoccupied band of electronic states, is known as the *conduction band*. A simplified illustration of this concept is given in Figure 1.4. For indirect band gaps, photoexcitation from the valence band maximum to the conduction band minimum is necessarily coupled to a *phonon* to conserve crystal momentum[70f]. Consequently, indirect band gap materials may present issues for photoelectrochemical cells because indirect transitions involve the coupling of an electron and a particular phonon to ensure the successful transfer of energy. Such transitions are therefore less likely to occur and are more dependent on the temperature[70]. A phonon is a quasiparticle[g] that serves to describe the quantised nature of lattice vibrations[16]. Hence, they are often used to rationalise mechanical and temperature-dependent phenomena. As mentioned, they may also couple to electronic transitions to facilitate more complex behaviour. Such phenomena are described as *electron-phonon coupling*.

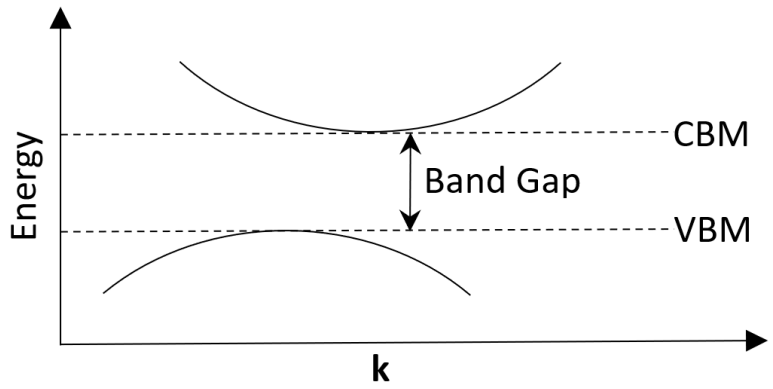


Figure 1.4: Illustration of a refined model of a band gap in a semiconductor or insulator. The curvature of the band edges is known as the *dispersion*. In this figure, the valence band maximum and conduction band minimum are not in alignment in reciprocal space. This is known as an *indirect* band gap semiconductor. If they were aligned, it would be a *direct* band gap semiconductor.

On the other hand, and despite absorbing photons more readily, direct band gap semiconductors are more likely to undergo *radiative recombination*. Radiative recombination may be considered to be the reverse mechanism of photoexcitation, wherein an electron and hole recombine and their excess energy is released as a photon. Thus, because a direct band gap semiconductor does not require phonon coupling to photoexcite charge carriers, there is also no phonon coupling to prevent said charge carriers from recombining and producing a photon. As such, the directness or indirectness of the band gap is a noteworthy consideration for photoelectrochemical cell development.

The Band Alignment is the alignment of the band edges with the respective water splitting half-

[f]The concept of crystal momentum is elaborated on in the next chapter.

[g]A quasiparticle being an effective particle that adequately describes the behaviour of a collection of interdependent particles.

reactions. In addition to the size of the band gap, this quantity must be determined because the charge carriers must be of similar energy to each respective half-reaction to efficiently catalyse them. This is illustrated in Figure 1.5 and further elaborated in the third chapter.

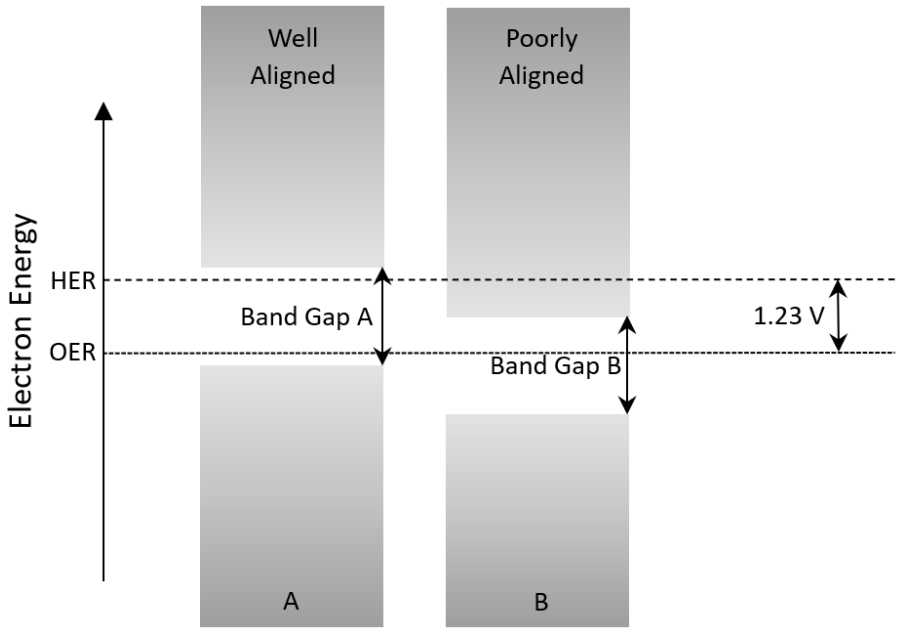


Figure 1.5: A simplified diagram of the difference between a well-aligned and poorly-aligned semiconductor for water splitting. A well-aligned semiconductor (i.e. semiconductor A) must have band edges that straddle at least one of the hydrogen/oxygen evolution reaction redox potentials.

The Corrosion Resistance is the ability of the electrode material to resist degradation under operating conditions. Corrosion of photoelectrochemical cell components usually involves the outer surface of the electrodes being rendered catalytically inactive. Often this is due to unwanted and irreversible reactions taking place on the surface of the electrode. For example, the direct electrolysis of seawater is known to corrode most electrode materials. This is usually attributed to the high concentration of chloride ions (Cl^-) within seawater which are believed to remove catalytically active metal ions from the electrode surface through successive coordination. Furthermore, these chloride ions may be substituted by hydroxide ions (HO^-) in solution. Thus, the chloride ions are reintroduced into the system for further corrosion[71,72], as shown via the below reaction steps.

Step 1: $\mathrm{Electrode{-}M + Cl^- \longrightarrow Electrode{-}M{-}Cl + e^-}$

Step 2: $\mathrm{Electrode{-}M{-}Cl + Cl^- \longrightarrow Electrode + M(Cl)_{\mathit{n}} + n(e^-)}$

Step 3: $\mathrm{M(Cl)_{\mathit{n}} + n(HO^-) \longrightarrow \ + M(HO^-)_{\mathit{n}} + n(Cl^-)}$

As such, it is desirable to limit these sorts of processes as much as possible, to ensure the longevity of the photoelectrochemical cell electrodes. The mechanisms involved in the corrosion process may vary according to the photoelectrochemical cell operating conditions and the composition of the electrode

material. For example, pure magnesium electrodes are known to be more susceptible to corrosion than iron-magnesium alloys due to the corrosion products of such alloys having relatively low formation enthalpies. Furthermore, said corrosion processes could be further suppressed by using phosphate and arsenate inhibitor species, within the electrolyte, which disrupted the usual iron-magnesium corrosion mechanisms. However, the efficacy of these disruption processes was found to be dependent on the pH of the electrolyte.[73,74].

The Catalytic Activity refers to the favourability of the desired reaction when compared to competing reactions. As such, it is desirable for the conversion of photons to chemical bonds to be as fast a process as possible. This may be quantified via *rate constants* and *activation energies*. The rate constant is a proportionality constant that relates the rate of a reaction to the concentration of reagents:

$$rate = k[A]^m[B]^n \tag{1.2}$$

where k is the rate constant, $[A]$ and $[B]$ are respective reagent concentrations, and m and n are the partial orders of the reaction which are determined upon reaching an understanding of the reaction mechanism[16]. The activation energy is the minimum energy required to initiate a chemical reaction[16]. To compare the catalytic activity of candidate materials, one can measure the rate of hydrogen and oxygen gas evolution and measure the *photocurrent density*. Correspondingly, two efficiency metrics can be derived[29]:

1) the Incident Photon to Current Efficiency (IPCE):

$$IPCE = 1240\frac{I_{ph}}{\lambda P_{in}} \tag{1.3}$$

where I_{ph} is the photocurrent density (current per unit area of illuminated electrode), λ is the incident wavelength, and P_{in} is the incident light intensity.

2) the Solar to Hydrogen Efficiency:

$$STH = \frac{\text{Energy of evolved } H_2}{P_{in}} = \frac{\text{Number of } H_2\text{molecules} \cdot \text{Energy of one molecule}}{P_{in}} \tag{1.4}$$

The Catalytic Selectivity refers to the ability of a catalyst to facilitate only the desired reaction. As such, corrosion may be considered to be an issue of catalytic selectivity. However, more broadly, side reactions may also occur that do not corrode the electrode but reduce the overall efficiency of the device because the absorbed energy is facilitating unwanted reactions. An example of this is the production of ozone, as opposed to oxygen, in many lead oxide (PbO_2) based photoelectrochemical

systems[75]. Catalytic selectivity may, for instance, be determined by identifying all of the products of the working system. The specific side reactions for a given system may vary depending on the working conditions such as the pH, electrode materials, and purity of the electrolyte.

1.3.3 Computational Screening and the Inverse Design of Materials

It is worth noting that the aforementioned criteria need not be satisfied with a single material but may be satisfied via a combination of photoelectrode materials, counter electrode materials, and other materials that each fulfil a specific function within the device. *Cocatalysts* are good examples of this approach, in which one material is responsible for photoabsorption, whilst another is responsible for facilitating the electrochemical water splitting reaction[52]. However, there is not yet a device that simultaneously satisfies all of the aforementioned criteria. Part of the challenge in finding suitable candidate materials is that it is not always obvious which properties of a material make it well-suited for a particular function. This may be achievable for a problem involving a single criterion (variable). For instance, fine-tuning the size of the band gap through systematic compositional changes is a common practice[76]. However, finding a suitable compromise across all six of the above criteria for the composition, bulk properties, and surface properties of the electrodes is a much greater challenge. Moreover, determining and comparing all such material properties reliably and systematically, via an experimental procedure, would be an incredibly cumbersome and costly endeavour. With this in mind, the *computational screening* of material databases and *inverse materials design* may be a means to hasten the discovery of new candidate electrodes.

There are a great variety of ways in which materials screening may be approached; from the simplest *pen and paper* algorithms to the most sophisticated neural network[h] models. Crucially, a materials screening procedure must analyse and rank candidate materials in terms of their appropriateness for a given function[77]. When considering traditional materials discovery and design, a new material with interesting properties would be synthesised. Those interesting properties would then be rationalised, through theory and experimentation, to determine how those properties arose. Finally, a suitable use for the new material would be suggested and trialled to solve a certain problem. Conversely, *inverse* materials design approaches the problem in reverse: a materials-related problem is expressed; the desired material properties are deduced according to the problem; theory and simulation are used to predict a suitable candidate, or set of candidates, for experimentation; and the experimentation procedure verifies whether or not the material can be synthesised and if it has the desired properties[78].

[h] A computing architecture that is based on the structure of the connections within the human brain[16].

If successful, a notable benefit of computational materials screening and inverse materials design is the reduced amount of physical experimentation that is required to find suitable candidate materials. This is because the theory and simulation procedure would help to exclude unlikely candidates from the pool of candidates by predicting their desired properties. This shifts the burden of research onto the theorists and simulation experts, but given the advances in computing power in recent years, the overall cost compared to a full experimental investigation is expected to be substantially lower. The computational screening of material databases would be a good first step in this regard because it involves analysing large numbers of known candidates to estimate their material properties. Such data may then be fed into a *machine learning* algorithm, within a *parallel computing* framework, for a more efficient and unbiased approach towards candidate selection.

Serial and Parallel Computing

Serial computing works by executing a set of predetermined instructions in sequence. Thus, the overall calculation is limited by the clock speed of the processor[79]. To overcome this limitation, *parallel computing* involves separating a calculation into separate components and then distributing each of these components across multiple computing *nodes*, which are an ordered array of computer *processors* that each carry out the calculation in a serial fashion. When optimised correctly, parallel computing may greatly reduce the duration of the calculation whilst producing the same result as a serial calculation. However, not all computations may be *parallelisable*. *Bernstein's conditions* are used as a test to determine whether or not a calculation is parallelisable[80]. These conditions consider two processes, *P1* and *P2*. These processes are only executable in parallel if three criteria are simultaneously satisfied:

- The input of P1 and output of P2 are independent.
- The input of P2 and output of P1 are independent.
- The outputs of P1 and P2 are independent.

It is also important to note that, if improperly optimised, a *parallelised* calculation may be less efficient than a serial calculation. The factors that determine the optimal parallelisation scheme are the parallelisability of the calculation data, the number of computer nodes and processors that are available, the amount of computer *random access memory*[i] that is available, and the speed at which computer

[i]The amount of data that a computer can manage whilst running the calculation.

nodes and processors may communicate to exchange or combine data. Given the wide variety of computing system architectures and calculations that may be performed, predicting the optimal settings for a given calculation is often not obvious. Thus, the parallelisation scheme for a given calculation and computing system is usually determined through *trial and error*, by experimenting with different parallelisation combinations, using a low-cost but representative calculation, and monitoring the efficiency.

Artificial Intelligence and Machine Learning

For a traditional serial or parallel computing system, the programmer is the entity that learns, reasons and problem-solves. The computing system merely executes a series of binary logical operations, following a predefined set of instructions that are composed by the programmer. In the broadest sense, *Artificial Intelligence* (AI) is the study of modelling such intelligent human processes[16] and aims to reverse engineer the human thought patterns of the programmer within the framework of a computing system. The origin of this concept may be attributed to a treatise published in 1854 entitled *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities*[81] wherein it is argued that even somewhat subjective concepts such as *"virtue"* may be deconstructed and categorised as *"habits"*, *"faculties"*, or *"passions"*. The validity of deconstructing *"intelligent"* or *"human"* thought in such a way remains a moral, philosophical, and practical question. For example, this *rational* theory of human thought is contradictory to the widely accepted economic *prospect theory*[82], which does not consider human thought to be entirely rational. Nevertheless, a computing system that is capable of detecting patterns, at various levels of abstraction, would be expected to *learn* and problem-solve without the need for a programmer. Moreover, said computing system would be able to complete its logical operations much faster than a human counterpart.

Machine learning is considered to be a sub-discipline of Artificial Intelligence. The term was first used in 1959 to describe a computer program that could play a board game more effectively than its programmer[83]. Broadly speaking, there are three branches of machine learning that vary in their effectiveness depending on the context in which they are used. These branches are *Supervised Learning*, *Unsupervised Learning*, and *Reinforcement Learning*.

Supervised Learning Supervised learning involves providing the computer with exemplary input and output data, known as *training data*, to *"train"* the computer to build a mathematical model of the system. As such, the more training data that is provided to the supervised learning model,

the more sophisticated and accurate the resultant mathematical model is expected to be. A typical example of supervised learning is *linear regression* wherein the data is mapped onto a function that accurately describes the overall trend in that data, despite many of the data points possibly deviating from that function. Supervised learning may be useful in the context of photoelectrochemical cells by using known, well-performing compounds and their properties as training data which may be used to suggest new compounds. For example, Oral et al[84] used a form of supervised learning to establish correlations between the band gap of the material and its dopants, synthesis methods, and thermal treatments.

Unsupervised Learning With unsupervised learning, no exemplary output data is provided, only input data. Consequently, unsupervised learning focuses on finding patterns and groupings within a dataset, without any human-imposed constraints or biases dictating which patterns and groupings are expected to be adhered to. As such, unsupervised learning is often useful for handling problems in which non-obvious patterns and associations may be discoverable within a dataset. An example of unsupervised learning is *cluster analysis*, wherein various data are grouped into *clusters* according to patterns that are identified by the unsupervised machine learning process. Unsupervised learning may be useful in the context of photoelectrochemical cells by finding patterns and associations between various material properties that may not yet be understood. For example, Wang et al[85] used unsupervised learning to suggest appropriate quaternary chalcogenige semiconductor compositions, for photovoltaic cells, without using any empirical data. This was motivated by the scarcity of said empirical data in light of the large composition and structure space for such materials.

Reinforcement Learning Reinforcement learning uses a cumulative rewards system to reinforce successful behaviour. It is most commonly associated with game-playing machines and autonomous robotics. As such, reinforcement learning may be useful for the optimisation of photoelectrochemical cell properties within a known framework. For example, past literature[86] has considered various reinforcement learning augmentations to molecular geometry optimisation calculations, for more efficient and precise crystal structure predictions.

1.4 Motivation and Outline for Subsequent Chapters

Such computational materials screening or inverse materials design procedures, whether or not machine learning is involved, could calculate the properties of a wide variety of candidate materials and estimate their effectiveness as photoelectrochemical cell electrodes. These estimates could then be used to refine

the selection of candidate materials to a more manageable figure for experimental study. With this in mind, and similarly to the photoelectrochemical cell criteria, it is respectfully submitted that a computational materials screening or inverse materials design procedure should be, at the very least:

Inexpensive The calculations require minimal computing power. This would allow the procedure to analyse a large number of candidates in a short period of time.

Automatable The implementation and analysis of the calculations require minimal human intervention. This would accelerate the output of data and reduce the chances of human error influencing the result.

Accurate The calculations produce results that, at minimum, *qualitatively* reflect the trends in the properties of known systems.

Understood in terms of its limitations The limitations of the screening procedure are known and easily identifiable, to avoid misleading results.

Modular The components of the screening procedure function independently. This would enable convenient experimentation with different approximations, standards, and machine learning implementations. Practically, this also makes the development of the procedure, including error correction, more manageable.

As such, it would be desirable to consider examples of how such a procedure may be established for a number of the challenges in photoelectrochemical cell development, in light of the available computational tools and the aforementioned criteria for a suitable photoelectrochemical cell electrode.

In this chapter, the general motivation behind this work has been reviewed. This includes the requirement for renewable and non-polluting energy sources in the face of an ever-growing energy demand, a comparison of the most prevalent renewable and non-renewable fuel sources, the appeal of hydrogen as a fuel source, how hydrogen may be generated using a photoelectrochemical cell, the challenges associated with developing said photoelectrochemical cells, and how these problems may be overcome using computational materials screening and inverse materials design procedures.

The next chapter introduces key electronic structure concepts that are necessary to understand the subsequent chapters. These subsequent chapters address three challenges concerning the computational screening of photoelectrochemical cell electrodes. These include the band alignment of semiconducting compounds, the intrinsic photovoltage loss through small *polaron* formation, and the catalytic activity

of hypothetically corrosion-resistant *rutile alloys*. In each of these chapters, at least one computational workflow is suggested. Each of these workflows is independent of one another (i.e. modular) but may be implemented together in a first-principles screening procedure or as feature calculators in a machine-learning algorithm. The final chapter contains a summary of the conclusions and suggested improvements concerning the described workflows.

The appendices are divided into three chapters. The first contains data tables that correspond to the provided figures. The second and third sets of appendices contain clarifying background matter and key mathematical concepts, which may be beneficial to the reader's overall understanding but are not immediately necessary for an understanding of this work.

Electronic Structure Theory and Methods

2.1 Motivation

The purpose of this chapter is to introduce key electronic structure concepts that are required for an understanding of the subsequent chapters. The section *Electrons in Insulators, Semiconductors and Metals* reviews how electronic states in such crystalline materials are distributed and outlines how the electronic structure of a crystalline material is understood within the framework of a periodic lattice. A more sophisticated understanding of such systems is then introduced in the *Electron Wavefunction* section. In particular, the role of the Schrödinger Equation. The *Solutions to the Time Independent Schrödinger Equation* section describes how the approximate solutions to the Schrödinger Equation may be obtained for a multi-electron, periodic crystal system. The *Hartree-Fock Method* section illustrates a practical approach to obtain such approximate solutions. The *Density Functional Theory* section outlines a modern and more efficient reinterpretation of the Schrödinger Equation. The *Practical Considerations* section outlines additional measures and protocols with which to make density functional theory more efficient. Finally, the *Application to Photoelectrochemical Cells* section briefly summarises how such density functional theory calculations may be used to predict material properties for the development of photoelectrochemical cells.

2.2 Electrons in Insulators, Semiconductors, and Metals

For a given material, mere knowledge of the crystal structure is insufficient to determine all of the properties of the system. In particular, the *electronic* properties are required to understand the pho-

tocatalytic and photoabsorption properties (such as the band gap, photovoltage, etc.) and thereby assess the suitability of the material as a photoelectrochemical cell electrode. To evaluate the electronic properties, a physical and mathematical formalism for the electrons inside that system must be established. In this regard, a good starting point is to develop an understanding of the *ground state* of the system.

The ground state is the lowest energy *quantum state* that the system can occupy (i.e. at *absolute zero* or *0 K*) whilst an *excited state* is any other quantum state that is higher in energy than the ground state[16]. Specifying a quantum state acknowledges that electrons have a *quantum* nature, meaning that a given electron may only possess energy across a discrete set of values. Other values are not permitted. Proof of the quantum nature of the electron may be inferred from the photoelectric effect[16]. In the ground state of a crystalline material, the lowest energy electronic states are fully occupied, whilst any higher energy states are fully unoccupied.

As mentioned in the previous chapter, the valence band and conduction band are separated by a substantial gap in energy known as the *band gap*. A material without a band gap is described as a *metal*. Merely for completeness, semimetals are a further possibility in which there is a very small overlap between the valence and conduction bands. This often results in reduced charge carrier concentrations when compared to metals[16]. However, semimetals are also able to transport both electrons and holes, unlike a metal which cannot separate such charges due to the lack of a band gap. Graphene is one such example of a semimetal[87,16]. In the general context of electronics, the size of the band gap determines whether or not the material is a semiconductor or an insulator. If the band gap is sufficiently small such that thermal energy at the operating conditions is sufficient to excite an electron from the valence band to the conduction band, the material is described as a semiconductor. If the gap is much larger than this, the material is described as an insulator (See Figure 2.1).

These differences in the band gap have significant effects on the observable properties of the material. For instance, the electrical conductivity of a metal decreases with an increase in temperature because heating the material causes the atoms within the lattice to move more frequently and to a greater extent. This impedes the electron's ability to pass freely and uniformly through the crystal lattice. In contrast, the conductivity of a semiconductor increases with an increasing temperature within most practical temperature ranges[88]. This is because the increase in conductivity, due to the aforementioned thermal excitation of electrons into the conduction band, has a greater contribution to the net conductivity than any reduction in conductivity due to the movement of atoms within the crystal structure[89].

However, this distinction between semiconductors and insulators is less useful for photoelectrochemical cells in which photoexcitation across the band gap covers a much greater energy range. In the past, this has resulted in the use of the term *"semi-insulator"*[90] for materials such as gallium arsenide, which are considered to be semiconductors in more recent literature. For this work, the term semiconductor will be used to describe materials that have a band gap that is within the useful range for a photoelectrochemical cell. This includes band gaps that are well-suited for photoabsorption and catalysis purposes.

The mobility may also be used as a quantity to distinguish semiconductors from insulators. The mobility is the ratio of the drift velocity (average velocity of the non-uniform movement of electrons) of a charge carrier to an applied electric field[16]. As such, the conductivity may be considered to be proportional to the mobility. This is important for semiconductors because, even if the semiconductor band gap is within the appropriate excitation range, if its mobility is too low, it may not be useful as a photoelectrochemical cell electrode.

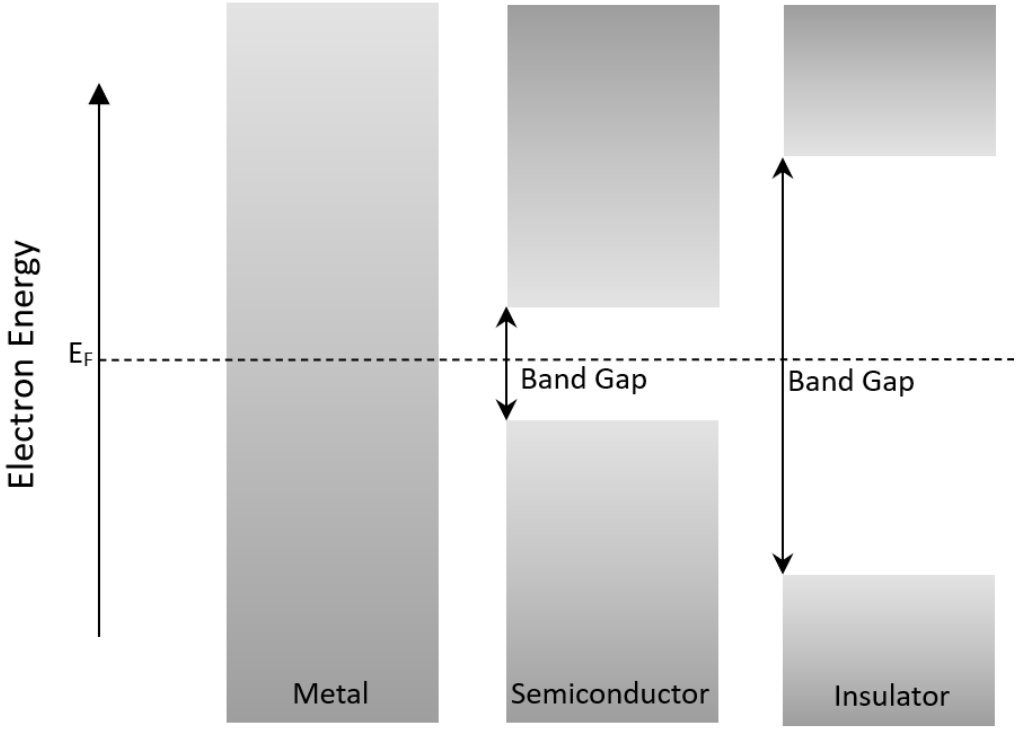


Figure 2.1: Illustration of the difference in the distribution of electronic states between metals, intrinsic semiconductors, and insulators.

It is worth noting that the above distinctions of insulator, semiconductor, and metal may be further obscured by intrinsic defects and extrinsic dopants. For instance, a heavily *doped* semiconducting material may behave more like a metal at high temperatures. Doping involves adding more electron-rich or more electron-deficient species to a semiconducting material, to change its electronic properties.

For instance, replacing a silicon atom in a pure silicon lattice (a semiconductor) with a phosphorus atom results in an *excess electron* being added to the silicon lattice. This is because phosphorus has one extra valence electron compared to silicon. Correspondingly, an additional electronic state is created, inside the band gap of silicon, due to the presence of the dopant. Furthermore, the excess electron is more likely to be available for conduction rather than for bonding, and consequently, the energy required to excite an electron to this higher energy state is lower than that of the undoped (*intrinsic*) semiconductor. As such, doping is routinely used to increase the conductivity of a semiconductor at a given temperature and may therefore enhance a photoelectrochemical cell's performance. Conversely, dopant states may also act as traps and recombination centres for the charge carriers[29]. Moreover, if a narrow band gap semiconductor is doped so heavily that electronic states on either side of the band gap are saturated with relaxed (unexcited) electrons, then the semiconductor may effectively behave as a metal.

For a wide band gap semiconductor, high dopant concentrations may lead to the formation of *mid-gap states*[29]. Mid-gap states are localised states that lie deep within the band gap and can act as effective valence/conduction bands, facilitating electronic transitions that are lower in energy than the band gap. Therefore, their presence may be detrimental or beneficial to device performance, depending on their alignment with the water splitting redox potentials.

Converse to electron-rich dopants, if an electron is removed from the lattice (i.e. by adding an electron-deficient impurity to the lattice), the conductivity may still increase. This is because when an electron is removed from a crystal, the absence of that electron creates a hole. Similarly, when an electron is excited from a state within the valence band to a state within the conduction band, a hole is created at the previously occupied state in the valence band. A neighbouring electron may then rapidly neutralise this charge, leaving its initial electronic state and thereby creating another hole. This is how a hole passes through a crystal much like an electron, exclusively occupying states within the valence band, and contributing to the electrical current. Formally speaking, a hole has a negative rest energy[a].

In certain cases, the conductivity of a material can be further modified by removing atoms from the lattice. A common example is the creation of oxygen vacancies. Often, such vacancies are expected to act as recombination centres, an undesirable trait. However, in certain cases, these oxygen vacancies create defect states within the band gap of the material. These defect states allow electrons to be promoted more easily into the conduction band[92], enhancing the conductivity. For instance, work by Scanlon et al[93] has shown that PbO_2 is a narrow band gap semiconductor with a metal-like conductivity

[a] i.e. E is negative for the mass-energy equivalence relation ($E = mc^2$)[91].

that was attributed to intrinsic oxygen vacancies in the material. It was concluded that the oxygen vacancies created defect states within the PbO_2 band gap. The electrons lost from the missing oxygen atoms were compensated for by electrons from the bulk of the material.

The above concepts are conveniently represented via the *Fermi level* (E_F). The Fermi level is the energy at which the probability of finding an electron in the lattice is exactly one-half[91], with 0 being impossible and 1 being certain. Therefore, the Fermi level serves as a reference point to define the proportion of electronic states that are occupied and unoccupied, without being an electronic state itself. For example, in an intrinsic semiconductor, the Fermi level lies somewhere within the band gap, depending on the temperature. However, if the semiconductor is doped with an excess electron, the Fermi level will shift towards the conduction band. This effect continues according to the dopant concentration. The relationship between the Fermi level and the temperature is given by the Fermi-Dirac Distribution[16]:

$$f(\epsilon) = \frac{1}{e^{\frac{\epsilon - E_F}{k_B T}} + 1} \tag{2.1}$$

where $f(\varepsilon)$ represents the probability that the electron will have energy ε, k_B is the Boltzmann constant, and T is the temperature. As such, as $T \to 0$, then the probability of finding a particle above E_F tends to zero. This is illustrated in Figure 2.2. A derivation is provided in Section B.1

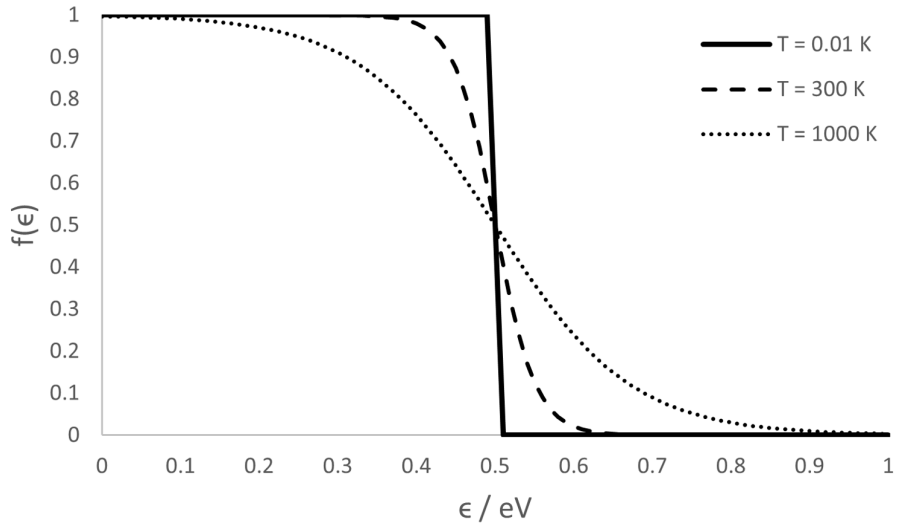


Figure 2.2: Plot of $f(\epsilon)$ across 3 different temperatures. In this case, $E_F = 0.5$eV. This figure shows that the probability of an electron occupying higher energy states increases with the temperature.

2.2.1 Electron Spin

Spin is the intrinsic and quantised angular momentum that an elementary particle, such as an electron, may possess[16]. Despite its name, it does not relate to any sort of observable *rotation* of the electron

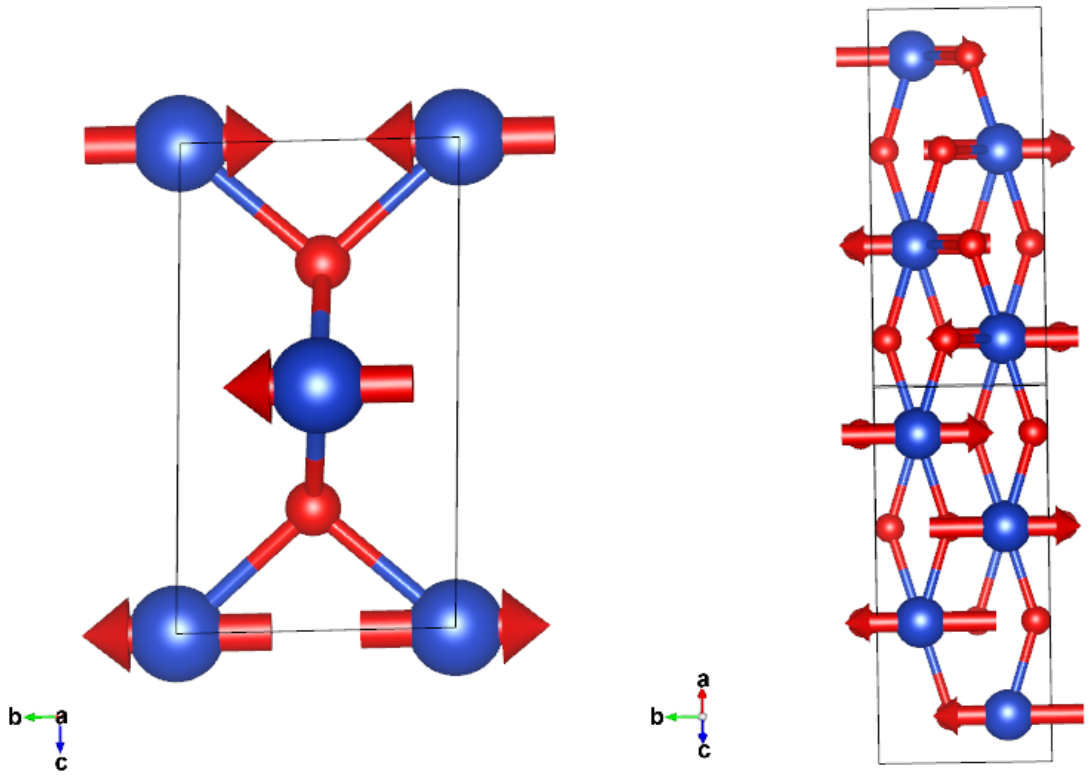
around an axis. This is at least in part due to the difficulty in defining the radius of an electron, due to its wave-particle character (See the next section). Despite this indefiniteness, the concept of spin provides a consistent theoretical framework for understanding multi-electron systems. This is particularly true for magnetic systems because it rationalises why a stationary material (i.e. a simple iron magnet) may produce a magnetic field. If the electrons in such a material possess some form of angular momentum, then that moving charge must therefore produce a magnetic field, despite the macroscopic material remaining stationary. The concept of spin is also important for distinguishing various classes of elementary particles. For example, electrons belong to a class of sub-atomic particles known as *fermions*, which are characterised by a half-integer spin. In the case of electrons, the spin may be either *up* (1/2) or *down* (-1/2). The concept of electron spin is supported experimentally by the Stern-Gerlach Experiment[94].

For multi-electron systems, a proper treatment of the electron spin is required for an accurate representation of the system. Particularly concerning the magnetic properties. The means with which this is accounted for is via the *Pauli Exclusion Principle*[95], which states that two identical fermions with a half-integer spin are not able to occupy the same quantum state[16]. As such, electrons in electronic states are often referred to as being *paired*, such that one up spin is paired with one down spin per electronic state[16]. *Unpaired* spins may only be permitted in singly occupied orbitals and are principally attributed as the cause of a material's magnetic properties. It is often the case that the arrangement of unpaired spins within a crystal lattice has a long-range order that is different from that of the crystal lattice. This ordering is referred to as the *magnetic structure*. In many cases, a *supercell*, which is larger and less symmetric than the crystallographic *primitive cell*[b], is required to correctly capture the magnetic structure, as shown in Figure 2.3.

2.3 The Electron Wavefunction

Due to their quantum nature, electrons have both wave and particle properties. For instance, the mass of an electron can be measured, implying particle properties. However, electrons are also known to diffract when passing through a double slit, implying wave properties[96]. This is known as wave-particle duality[16]. Hence, it is not immediately clear which mathematical formalism (particle or wave) to use when studying the behaviour of an electron for a particular property. For example, under a particle formalism, an electron could be modelled as an indivisible entity with a well-defined mass and position

[b]The smallest possible unit cell (repeating unit) that represents the spatial distribution of atoms within the crystal structure.



(a) Primitive Cell **(b)** Magnetic Primitive Cell

Figure 2.3: Illustrative example of how the primitive unit cell is insufficient to describe the magnetic structure of copper (II) oxide. In the primitive cell (Left) the copper atoms at the corners of the unit cell have opposing magnetic moments (given by the red arrows). This contradicts the periodicity requirements of the unit cell and would not be useful for computational study. In contrast, the larger, magnetic primitive unit cell (Right) properly accounts for the long-range magnetic ordering and unit cell periodicity requirements but results in an increased computational demand.

that passes through the lattice according to classical mechanics. Conversely, under a wave formalism, the electron could be modelled as a zero-mass disturbance that propagates through the lattice. In practice, neither of these cases has proven to be effective in encompassing all of the electronic properties of a system. Fortunately, an electron *wavefunction* ($\psi(x)$) can be used to mathematically represent such quantum particles within a system[97]. The wavefunction is a continuous, *wave-like* function to describe the behaviour of quantum entities[96]. Whilst the wavefunction is a function of the position (x) of the electron, it is defined and interpreted in terms of the *probability* of detecting the electron at a position x according to[96]:

$$P(x) = |\psi(x)|^2 \tag{2.2}$$

where $P(x)$ is the probability density of finding the electron at position x. The wavefunction differs from a classical wave in that the wavefunction need not be associated with a disturbance in a medium. Rather, the oscillatory nature of the wavefunction merely allows for a convenient mathematical formalism to make probabilistic predictions for quantum entities[96]. Thus, the wavefunction is merely a

descriptor of a particle subject to the laws of quantum mechanics[96]. When considered as a function of time (t), it fulfils the same function as the position operator $x(t)$ in classical mechanics. Therefore, it implicitly contains all of the information necessary to understand the electron. However, unlike the position operator, the wavefunction does not provide a *precise* location for the electron in Euclidean space. This is due to the *Heisenberg uncertainty principle*, wherein the certainty in the position and momentum of a quantum particle has an inverse relationship according to[96]:

$$\Delta x \Delta p_x \geq \frac{h}{2} \tag{2.3}$$

where p_x is the x-component of the electron's momentum. Despite this uncertainty, the wavefunction can still be bounded by a tangible quantity. For example, an electron inside an infinitely large crystal lattice must be somewhere within that crystal lattice, even if its precise location cannot be defined. Thus, it is certain that the *probability density* across the entire lattice is equal to 1. This concept is known as the *normalisation condition*[96]:

$$\int_{-\infty}^{\infty} P(x)dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1 \tag{2.4}$$

For the wavefunction to be useful for computation, a mathematical form must be provided. Whilst $|\psi(x)|^2$ is uniquely determinable via experimental data because it is directly proportional to the electron density, $\psi(x)$ is not uniquely defined. Instead, its form is determined by the conditions and constraints of the system. For instance, the wavefunction of a free particle confined to a 1-dimensional range has a particular wavefunction, whilst the wavefunction for an electron in a hydrogen atom has a separate and distinct form[96]. As such, for a multi-electron system, such as a crystalline solid composed of atoms, the practical goal is to determine the *multi-electron wavefunction* ($\Psi(x)$) of the system in its ground state.

2.3.1 The Schrödinger Equation

The ground state wavefunction, and in principle the entire multi-electron quantum state ($|\Psi\rangle$), can be determined by solving the *Time Independent Schrödinger Equation*. It is noted that a quantum state is an abstract concept that encapsulates all of the known properties of a quantum mechanical system, whilst the wavefunction (Ψ), merely serves as a representation of the observable properties, such as the electron position[16]. The Time Independent Schrödinger Equation is given as[98] :

$$\hat{H}\Psi = E\Psi \tag{2.5}$$

where $\hat{H}$ is the Hamiltonian operator, E is the total energy (kinetic + potential) of the system, and Ψ is the mathematical form of the wavefunction of the system (see Section B.2 for a Plausibility argument concerning the Schrodinger Equation). For a given quantum state, any observable property of said quantum state is the *expectation value* of an operator that integrates over all spatial coordinates[99]. For example, the *Total Energy* (E) of a multi-electron system is the expectation value of the Hamiltonian operator[99]:

$$E = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \langle\hat{H}\rangle = \langle\hat{T}\rangle + \langle\hat{V}_{e-e}\rangle + \int d^3r V_{e-n}(\mathbf{r})n(r) + E_{n-n} \tag{2.6}$$

where $\hat{T}$ is the kinetic energy operator, $\hat{V}_{e-e}$ is the electron-electron interaction potential operator, V_{e-n} is the electron-nuclear (external) potential, and E_{n-n} is the nuclear-nuclear interaction energy.

The Hamiltonian Operator

The Hamiltonian operator is a function that represents the rate of change of a dynamic system with respect to time. A multi-electron system is an example of such a system. In a classical (non-quantum) system, the Hamiltonian operator is a function of position coordinates (q_i) and momenta (p_i) of the system:

$$H = \sum p_i\dot{q}_i - L \qquad \text{where} \qquad p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.7}$$

where L is the Lagrangian function of the system which is the kinetic energy minus the potential energy and $\dot{q}_i$ is the first derivative of q_i with respect to time. Assuming that H or L are time-independent, H must therefore correspond to the *total energy* (kinetic + potential) of the classical system:

$$\frac{\partial H}{\partial t} = \frac{\partial L}{\partial t}, \qquad \frac{\partial H}{\partial p_i} = \dot{q}_i, \qquad \frac{\partial H}{\partial q_i} = -\dot{p}_i \tag{2.8}$$

As mentioned, the dynamics of a quantum system are not defined in the same way as a classical system. Nevertheless, in a quantum system, the Hamiltonian operator serves a similar purpose of characterising the dynamics of the system[16,100]. The Time-Independent Schrödinger Equation for a single electron is given as:

$$\left[\underbrace{-\frac{\hbar}{2m}\nabla^2}_{\text{Kinetic}} + \underbrace{V}_{\text{Potential}}\right]\psi = E\psi \tag{2.9}$$

where ψ is a single electron wavefunction, m is the mass of the electron, $\hbar$ is the reduced Planck constant, and ∇^2 is the Laplacian Operator. Since there is currently no universal wavefunction and no universal Hamiltonian operator, the Hamiltonian operator provides the mathematical form of the *level of theory* applied to the system. Due to the number of approximations involved in calculating the wavefunction of a multi-electron system, the mathematical form of the Hamiltonian operator may vary across applications. The selection of a particular mathematical form is usually based on a compromise between the required accuracy and the computational expense.

Time Dependent Schrödinger Equation

There also exists a more general, time-dependent version of the Schrödinger Equation, the *Time-Dependent Schrödinger Equation*, which describes how such a dynamic quantum system may evolve with time:

$$\hat{H}\Psi = i\hbar\frac{d\Psi}{dt} \tag{2.10}$$

where i is the imaginary unit ($\sqrt{-1}$) and $\hbar$ is the reduced Planck constant. Under normal operating conditions, such as room temperature, each of the electrons in a *real* system are only briefly in their ground state and dynamically exchange energy with their environment. However, the ground state is often a sufficient approximation to describe many real properties of that system and may further serve as the reference point from which to understand such excited states. Whilst it is possible to directly investigate excited state phenomena using the time-dependent Schrödinger equation[101], such investigations are often prohibitively expensive, complex, and unnecessary for a materials screening procedure. Furthermore, because the ground state is considered to be a *stationary state*, in which all properties of the system are independent of time[16], the use of the time-dependent Schrödinger equation is usually unnecessary.

2.4 Solutions to the Time Independent Schrödinger Equation in a Periodic Potential

2.4.1 Crystal Momentum, Reciprocal Space, and the Blöch Theorem

It is noted that the electronic structure of a semiconductor presented in Figure 2.1 is a simplification. For instance, the particularities of a given crystal structure have a profound effect on an electron's

ability to pass through the crystal unperturbed. A more sophisticated model of a semiconductor's electronic structure accounts for the *crystal momentum*. The crystal momentum considers the electron's momentum in light of the particular arrangement of the atoms that compose the crystal lattice. The total force applied to an electron moving through the lattice is equal to the rate of change of the crystal momentum[16]. Thus, for crystalline materials, it would be intuitive to bound the system in terms of its periodicity, by imposing periodic boundary conditions on the order of the primitive cell[102c], and to classify the distribution of electronic states via their crystal momentum.

To assist in this endeavour, it is often more instructive to consider the problem in *reciprocal space*, rather than *Euclidean* (*real*) space. Reciprocal space is the *Fourier transform*[d] of real space and the primitive cell of the reciprocal lattice is known as the *first Brillouin zone*. The reciprocal lattice of a given system may be defined in terms of a set of wave vectors[e] ($\mathbf{k}$) which relate to the crystal momentum ($\mathbf{p}$) via $\mathbf{p} = \hbar\mathbf{k}$[16]. This formalism is useful because solutions to the Time Independent Schrödinger Equation within a periodic potential (a crystal lattice) may then be obtained via the Blöch theorem[103]:

$$\psi_{\mathbf{k}} = u(\mathbf{r})e^{i\mathbf{kr}} = u(\mathbf{r})e^{i(\mathbf{k}+\mathbf{K})\mathbf{r}} \tag{2.11}$$

where $\psi_{\mathbf{k}}$ is the wavefunction described as a function of an electron wave vector$\mathbf{k}$, $u(\mathbf{r})$ is a function that represents the periodicity of the system, and $\mathbf{K}$ is a lattice vector of the system [f]. Some implicit assumptions of Equation 2.11 are that $\mathbf{k}$ is unique within the bounds of the first Brillouin zone and that the crystal momentum is conserved[99]. Thus, the distribution of electronic states within the crystalline system can be determined exactly and finitely in terms of Blöch wavefunctions. Moreover, as the size of the crystal tends to infinity, further values of $\mathbf{k}$ are added to the first Brillouin zone[104], and the spacing between values of $\mathbf{k}$ tends to zero[99]. Consequently, $\mathbf{k}$ can be treated as a continuous variable across a *path* in reciprocal space. This effective continuum of allowable electronic states is referred to as a *band*. Each value of $\mathbf{k}$ is further associated with a discrete set of eigenstates that account for the quantisation of electronic energy[99].

As such, the application of Blöch's theorem further allows for the distribution of allowable electronic states (as bands) within the crystal to be conveniently represented and interpreted in reciprocal space as $E(\mathbf{k})$, where E is the electron energy. Figure 2.4 is an example of such a representation for chromium oxide. Consideration of the crystal momentum also has consequences for the Fermi level. Since the

[c] See Section B.3 for further elaboration on this concept.
[d] See Section C.5.
[e] A wave vector is a vector quantity that describes the magnitude and direction of propagation of a wave[16].
[f] See Section B.4 for further elaboration of *Blöch's Theorem*.

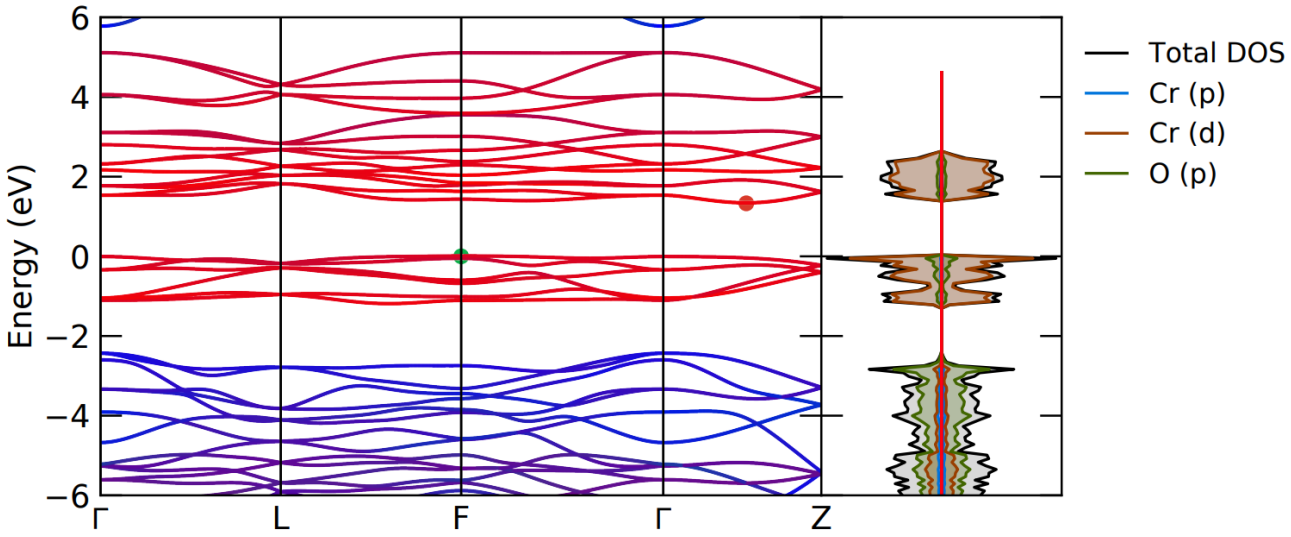


Figure 2.4: Example of a routinely calculated bandstructure for chromium oxide (Cr_2O_3). The VBM and and CBM are given by the green and red dots respectively. The blue lines represent bands of electronic states that are composed primarily of oxygen p orbitals, whilst the red lines represent chromium d orbital dominated bands. The corresponding *density of states* (the number of quantum states per unit of energy[16]) up to 3 eV is presented to the right of the figure. This figure was produced using the *sumo* package[17]. *Gamma*, *L*, *F* and *Z* represent characteristic, high symmetry points, in reciprocal space, within the first Brillouin zone. It is worth noting that this particular bandstructure was produced using density functional theory, using the PBE functional, which tends to underestimate the size of the band gap.

distribution of electronic states can be described as a function of $\mathbf{k}$, so too can the Fermi level. This is most commonly represented as a *Fermi surface* which is a three-dimensional surface, in reciprocal space, that separates the occupied and unoccupied states. Visualisation of the Fermi surface can be used to infer the electrical, magnetic, and thermal properties of the material.

2.4.2 Mapping Reciprocal Space

Despite the distribution of electronic states being an effective continuum within a given band, it is actually more computationally convenient to process the wavefunction as a discrete set of values during a calculation[105]. As such, the variation of $\psi_{\mathbf{k}}$ with $\mathbf{k}$ is approximated as a sum over a select number of $\mathbf{k}$ values, as opposed to a continuous function. The select number of spatially resolved $\mathbf{k}$ values that are considered for a calculation are referred to as ***k**-points*. The discretised approximation of $\psi_{\mathbf{k}}$ is given below[105,99]:

$$\overline{f}_i = \frac{1}{\Omega}\int_{IBZ} d\mathbf{k}\ \ f_i(\mathbf{k}) = \frac{1}{N_{\mathbf{k}}}\sum_{\mathbf{k}} f_i(\mathbf{k}) \tag{2.12}$$

where $f_i(\mathbf{k})$ is a function that defines the discrete set of states for a given $\mathbf{k}$; $\overline{f}_i$ is the average $f_i(\mathbf{k})$ for $N_{\mathbf{k}}$ values of $\mathbf{k}$; and Ω is the volume of the irreducible Brillouin zone. In practice, $f_i(\mathbf{k})$ can take the form of a step function wherein the *Linear Tetrahedron Method with Blöchl Corrections*[106] can be

used to efficiently interpolate across the irreducible Brillouin zone.

Such interpolation schemes are required to present a continuous and unambiguous representation of the Fermi surface, and therefore the properties of the material, without having to use an incredibly fine (and computationally expensive) mesh of **k**-points. The Linear Tetrahedron Method with Blöchl Corrections interpolates between **k**-points by representing the space between grid points using tetrahedra. Geometrically, any space and **k**-point grid may be precisely and efficiently sampled using tetrahedra, with the vertices of said tetrahedra being interpolated from the original **k**-point grid. The Blöchl corrections further allow for a simplification of the procedure for insulators, which further lowers the computational cost, as well as incorporating a more sophisticated means with which to interpolate across the vertices, since linear interpolation between vertices can lead to systematic errors[99]. However, this method often struggles to accurately represent metals and narrow band gap semiconductors[107]. In such cases, Gaussian smearing can be used in place of the tetrahedron method[108,109].

Regardless of the particular interpolation scheme, a common method with which to sample reciprocal space is using a uniform grid of **k**-points, known as *Monkhorst-Pack grids*[110]. The optimal Monkhorst-Pack grid density is the minimum value at which convergence to a desired level of precision is achieved for the property of interest. This is elaborated in Section 2.7.2.

2.5 Approximations to the Time Independent Schrodinger Equation

Even with an appropriate bounding of the crystalline wavefunction, via the Blöch theorem, and the ability to map reciprocal space and derive material properties, using suitable interpolation methods, a mathematical formalism for the Hamiltonian operator has not yet been outlined here. Moreover, for systems that contain three or more electrons (e.g. any crystalline system), the many-body problem is introduced. Consequently, there are no exact solutions to the Time Independent Schrödinger equation for such systems and approximations are used to make the mathematical basis for modern calculation methodologies more manageable.

2.5.1 The Born-Oppenheimer Approximation

The Born-Oppenheimer Approximation considers that, due to the great difference in mass between the atomic nucleus and the surrounding electrons, the nuclear motion is far slower than the electron motion. Thus, from the perspective of the surrounding electrons, the nuclei can be assumed to be fixed

in space. This enables a mathematical simplification wherein the nuclear motion is separable from the electron motion[111]. As such, the Hamiltonian operator for a multi-electron and multi-nuclear system:

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{n-e} + \hat{V}_{e-e} + \hat{V}_{n-n} \tag{2.13}$$

is reduced to:

$$\hat{H} = \hat{T}_e + \hat{V}_{n-e} + \hat{V}_{e-e} + \hat{V}_{n-n} \tag{2.14}$$

where $\hat{T}_e$ is the electron kinetic energy, $\hat{T}_n$ is the nuclear kinetic energy, $\hat{V}_{n-e}$ is the electrostatic potential between the nuclei and electrons, $\hat{V}_{e-e}$ is the electrostatic interaction between the electrons, and $\hat{V}_{n-n}$ is the electrostatic interaction between the nuclei (see Section B.5 for further elaboration of the Born-Oppenheimer Approximation). The relationship between the charge density and the electrostatic potential is given by *Poisson's Equation*[16]:

$$\nabla^2\phi(\mathbf{r}) = -\frac{\rho}{\varepsilon_0} \tag{2.15}$$

where ϕ is the electrostatic potential, ρ is electron density and ε_0 is the free space permittivity.

2.5.2 Basis Sets

A basis set considers that the multi-electron wavefunction may be composed of a linear combination of a finite number of *basis functions*. This can be as detailed as the Linear Combination of Atomic Orbitals (LCAO) approach[112], used to model molecular/crystalline orbitals,[g] or as simple as a linear combination of plane waves[114]. As such, an appropriate basis set may be chosen according to the accuracy and precision requirements of a particular calculation.

2.5.3 Pseudopotentials

Alongside a basis set, *pseudopotentials* are often used to represent the chemical species' contribution to the wavefunction of the system. The premise of the pseudopotential method is to replace the complex and computationally expensive parts of the atomic radial wavefunction with a simpler and cheaper function that requires fewer Fourier components for computation[115]. The replaced part usually describes the core electrons as they rarely influence chemical bonding and contribute to a great deal

[g] A well-known example are the Pople basis sets[113] (e.g. 6-31G, 3-21G) which are composed primarily of Gaussian functions.

of complexity within the atomic radial wavefunction. Projector Augmented Wave (PAW) pseudopotentials[116] utilise *projectors* to construct a separable pseudopotential operator. Projectors consider a pseudopotential operator such that each pseudopotential can be represented as a sum of *projection operators*. This reduces the computational expense because the corresponding matrix elements of the Slater determinant only require the products of projection operations[99] (see Section B.7 for an elaboration on the PAW method).

2.5.4 Koopmans' and Janak's Theorems

Up to this point, it has been implicitly assumed that *Koopmans' theorem* and *Janak's theorem* are suitable approximations for the systems in question. Both Koopmans' and Janak's theorems consider the distribution of electronic states for a neutral system to be the same as its charged counterpart. In reality, some degree of orbital relaxation occurs, to account for an added charge. However, in many cases, this effect may be negligible. The benefit of Koopmans' and Janak's theorems for the band gap is that the band gap of the DFT calculation may be assumed to correspond to the band gap of the material. Strictly speaking, if the orbital relaxation were to be accounted for, two separate calculations of the neutral and charged system would need to be performed and compared. Koopmans' theorem considers the problem from the perspective of Hartree-Fock theory, whilst Janak's theorem considers the problem from the perspective of DFT[117,118], both of which will be introduced shortly.

2.5.5 The Hartree Approximation

Whilst the Blöch wavefunction is considered to be a good representation of a single particle moving through a periodic potential, it provides no further description of how a *multi-electron* system in such a potential may behave. An early approximation was provided by Hartree[119,120,121] who considered a time-independent system of electrons that are uncorrelated (non-interacting). Since the electrons do not affect one another, the multi-electron Time Independent Schrödinger Equation can be represented via a combination of single electron Time Independent Schrödinger Equations:

$$\hat{H}_{eff}\psi_i^\sigma(\mathbf{r}) = \left[-\frac{\hbar^2}{2m_e}\nabla^2 + V_{eff}^\sigma(\mathbf{r}) \right] \psi_i^\sigma(\mathbf{r}) = \epsilon_i^\sigma \psi_i^\sigma(\mathbf{r}) \tag{2.16}$$

where $V_{eff}^\sigma(\mathbf{r})$ is the effective potential that acts on each electron, σ is the electron spin, $\mathbf{r}$ is the position of the electron in Euclidean space, ψ is the single electron wavefunction, i is the index of a given electron, m_e is the electron mass, $\hbar$ is the reduced Planck constant, ϵ is the energy, and ∇^2 is

the Laplacian operator. Moreover, because the electrons are independent of one another, the resultant single-electron wavefunctions could be combined into a multi-electron wavefunction via their product:

$$\Psi(\mathbf{r}) = \psi_1(\mathbf{r})\psi_2(\mathbf{r})...\psi_N(\mathbf{r}) = \Pi_i^N \psi_i(\mathbf{r}) \tag{2.17}$$

The ground state of the system may then be calculated by occupying only the lowest energy states. Practically, this is usually performed via the *Self Consistent Field* procedure. The Self-Consistent Field procedure is an iterative method that ensures that the *potential field* generated by the charge distribution from a given wavefunction is consistent with the initial field (i.e. is "self-consistent") of that wavefunction. The threshold for this self-consistency is determined by the user, in light of the precision allowed by the available computational resources. Furthermore, the aforementioned potential field requires a mathematical conception. This is commonly referred to as *electron correlation* which refers to the Coulombic influence the electrons have on one another[122]. In this case, Hartree utilised what is commonly referred to as the *Mean Field Approximation*, wherein the many body Coulombic electron-electron interactions are represented as an averaged effect[123]. However, this treatment has proven to be insufficient to describe so-called *strongly correlated* systems, such as transition metal oxides[124].

2.5.6 The Hartree-Fock Approximation

Unfortunately, the Hartree approximation does not sufficiently describe the multi-electron wavefunction because the product of single-electron wavefunctions does not guarantee a resultant multi-electron wavefunction that is *antisymmetric* upon the *exchange* of two electrons. This antisymmetry requirement is merely a rearticulation of the Pauli Exclusion Principle[95] but using the wavefunction formalism[16]. To solve this problem, Fock and Slater[125] demonstrated that the antisymmetry requirement may be satisfied by using a *Slater determinant* to reconstruct the multi-electron wavefunction, as opposed to a product[99]:

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1,\sigma_1) & \phi_1(\mathbf{r}_2,\sigma_2) & \cdots & \phi_1(\mathbf{r}_N,\sigma_N) \\ \phi_2(\mathbf{r}_1,\sigma_1) & \phi_2(\mathbf{r}_2,\sigma_2) & \cdots & \phi_2(\mathbf{r}_N,\sigma_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\mathbf{r}_1,\sigma_1) & \phi_N(\mathbf{r}_2,\sigma_2) & \cdots & \phi_N(\mathbf{r}_N,\sigma_N) \end{vmatrix} \tag{2.18}$$

where each $\phi_i(\mathbf{r}_i, \sigma_i)$ is a generic and orthogonal eigenfunction that is a function of its position r_i and spin state σ_N that further satisfies:

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \tag{2.19}$$

This is a computationally convenient means to represent a many-body wavefunction whilst simultaneously satisfying the Pauli exclusion antisymmetry requirements.

The Hartree-Fock Hamiltonian

Presented in a similar format to the Time-Independent Schrödinger Equation, the resultant mathematical form of a Hartree-Fock Hamiltonian operator is given by Equation 2.20[h]:

$$\left[\underbrace{-\frac{\hbar^2}{2m_e}\nabla^2}_{\text{Kinetic}} + \underbrace{V_{ext}(\mathbf{r}) + V_{Hartree}(\mathbf{r}) + \hat{V}_x^{i,\sigma}(\mathbf{r})}_{\text{Potential}} \right] \psi_i^\sigma = \varepsilon_i^\sigma \psi_i^\sigma(\mathbf{r}) \tag{2.20}$$

wherein:

$$\hat{V}_x^{i,\sigma}(\mathbf{r}) = -\sum_j \int d\mathbf{r}' \psi_i^\sigma(\mathbf{r}')^* \psi_i^\sigma(\mathbf{r}') \frac{\psi_j^\sigma(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'| \psi_i^\sigma(\mathbf{r})} \tag{2.21}$$

where $\hbar$ is the reduced Planck constant, m_e is the electron mass, $V_{ext}(\mathbf{r})$ is the external potential felt by an electron due to the atomic nuclei, $V_{Hartree}(\mathbf{r})$ is the electron correlation potential felt by each electron due to the presence of other electrons, $\hat{V}_x^{i,\sigma}(\mathbf{r})$ is the electron exchange operator, $\psi_j^\sigma(\mathbf{r})$ is the single electron wavefunction (i) with spin (σ), and ε_i^σ is the corresponding eigenvalue for this wavefunction. The relationship between E and ε_i^σ is given by the weighted sum of non-interacting electron energies, according to the Fermi-Dirac distribution[126,127]:

$$E = \sum_i^\sigma \frac{\varepsilon_i^\sigma}{e^{(\varepsilon_i^\sigma - \mu)/k_B T} + 1} \tag{2.22}$$

where k_B is the Boltzmann constant, T is the temperature, and μ is the chemical potential of the electrons. In the HF method, the ground state is usually obtained via the variational principle[128]. According to this principle, the mathematical form of the wavefunction is guessed, and its parameters are then varied until the lowest energy solution is obtained.

The most common inaccuracies of the HF method are often attributed to its treatment of electron correlation, which remains a challenge in the broader context of electronic structure theory[99]. Another issue concerning the HF method is its rapid increase in computational expense as the system size and

[h] A full derivation can be found on pages 61 - 64 of *Electronic Structure* by Martin[99].

complexity increase. This is because four-centre, two-electron integrals need to be evaluated in order to complete the calculation[129]. This is often referred to as N^4 *scaling*, where N represents the number of basis functions that are used which is directly proportional to the number of eigenstates[99].

2.6 Density Functional Theory

A commonly used alternative to the HF method is Density Functional Theory[130] (DFT), which retains many of the aforementioned approximations of the HF method. However, DFT initially guesses an electron density, not a wavefunction, and applies the variational principle accordingly. This generally results in scaling of the order of N^3 by virtue of matrix diagonalisation in three dimensions[129,99]. To justify this method, DFT relies on two theorems, known as the *Hohenberg-Kohn Theorems*[131].

The *First Hohenberg-Kohn Theorem* considers that if the spatially resolved external (nuclear) potential (V_{ext}) in the ground state uniquely determines the spatial distribution of the electron density ($\rho_0(\mathbf{r})$), so too can said external potential be uniquely described by that electron density. Consequently, all of the properties of the system may be determined via the ground state electron density.

The *Second Hohenberg-Kohn Theorem* considers that the energy and wavefunction, of such a system, are calculated exactly via a *universal functional* of the electron density ($E[n]$). In particular, the global minimum of $E[n]$ corresponds to the exact ground state energy (E_0) and density ($n_0(\mathbf{r})$)[131].

In this way, the Time Independent Schrödinger Equation can be replaced with an *ansatz*[i], such that the electron density is guessed and modified iteratively. Such an equation was first suggested by Kohn and Sham[132]:

$$\left[\underbrace{-\frac{\hbar^2}{2m_e}\nabla^2}_{\text{Kinetic}} + \underbrace{V_{ext}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' + V_{XC}^{i,\sigma}[n(\mathbf{r})]}_{\text{Potential}}\right]\psi_i^\sigma = \epsilon_i^\sigma \psi_i^\sigma(\mathbf{r}) \tag{2.23}$$

where $\int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'$ describes the electron-electron repulsion; whilst $V_{XC}^{i,\sigma}[n(\mathbf{r})]$ represents the exchange-correlation potential; and $(n(\mathbf{r}))$ is the electron density (see Section B.6 for a mathematical proof of the Hohenberg-Kohn Theorem and a derivation of the Kohn-Sham equation). In practice, DFT greatly reduces the computational overhead, due to the reduced number of calculations that need to be performed per electron in the system. As such, larger and more complex systems can be investigated with DFT. However, the universal functional suggested by the Second Hohenberg-Kohn Theorem has not

[i] A guess as to the mathematical form of the solution

yet been determined. Only approximate solutions are in use. Generally speaking, the main approximation concerns the exchange-correlation potential and its corresponding energy[99]. DFT functionals are normally selected according to a compromise between the computational expense and the accuracy/precision for a given calculation. As such, there is a wide variety of available functionals that each vary in their ability to accurately predict certain properties. For reference in the upcoming chapters, four different classes of DFT functional were used throughout this work:

The Local Density Approximation Using this approximation, the exchange-correlation energy density at each point in space is treated as a homogeneous electron gas[99]:

$$E_{xc}^{LDA} = \frac{1}{2}\int d^3r n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r})) \tag{2.24}$$

where $\epsilon_{xc}(n(\mathbf{r}))$ is the energy per electron at $\mathbf{r}$. A corresponding expression that accounts for spin also exists, known as the *Local Spin Density Approximation*[99]:

$$E_{xc}^{LSDA} = \int d^3r n(\mathbf{r})\epsilon_{xc}(n^{up}(\mathbf{r}), n^{down}(\mathbf{r})) \tag{2.25}$$

As a consequence, all solutions obtained via the LDA consider the electrons to be fully delocalised across the entire crystal lattice. Therefore, the local density approximation is often accurate for delocalised systems (metals) and inaccurate for strongly localised systems (semiconductors and insulators)[133].

The Generalised Gradient Approximation Since the local density approximation often fails to accurately capture the behaviour of semiconducting materials, a more sophisticated functional is required. To this end, the *generalised gradient approximation*[134] has been introduced. Using this functional, the magnitude of the gradient of the exchange-correlation energy density ($|\nabla n|$), in addition to its value at each point in space, is included in the functional[99].

$$E_{xc}^{GGA} = \frac{1}{2}\int d^3r n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r}), |\nabla n|) \tag{2.26}$$

Incorporation of this gradient allows some degree of inhomogeneity in the electron density to be considered. Thus, when compared to the local density approximation, the generalised gradient approximation usually provides a better approximation of the electronic structure for semiconductors and insulators[135]. The *Perdew-Burke-Ernzerhof*[136] (PBE) functional and the *Perdew-Burke-Ernzerhof functional revised for solids*[137] (PBEsol) are referred to in later chapters of this work. The PBE functional is one of the most commonly used GGA functionals for benchmarking DFT calculations,

primarily due to its relatively low computational cost and good accuracy for many material properties. The PBEsol functional is particularly accurate for the calculation of lattice parameters.

The Meta-Generalised Gradient Approximation One of the trade-offs with the generalised gradient approximation is that the introduction of the gradient breaks several exact constraints that are maintained for the LDA[99,138]. Meta-GGA functionals have since been developed to remediate this problem. Whilst their treatment of the exchange-correlation energy density can vary, most meta-GGA functionals are characterised by a second-order gradient expansion of the exchange-correlation energy density in addition to the first-order gradient expansion of the density and the density itself.

$$E_{xc}^{mGGA} = \frac{1}{2}\int d^3r n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r}), |\nabla n|, |\nabla^2 n|) \tag{2.27}$$

The inclusion of the second-order term is expected to better represent the exact constraints that are broken by the generalised gradient approximation in a similar fashion to how a Taylor expansion more accurately represents a given function if higher-order terms are included. The Strongly Constrained and Appropriately Normed (SCAN) functional[138] is used in this work and claims to satisfy (or very nearly satisfy) all of the exact constraints that are currently known for the meta-GGA class of functionals.

The Hybrid Functional Approximation Hybrid functionals are so named because they incorporate a portion of the exact electron exchange, obtained via HF theory, into the DFT calculation[99].

$$E_{xc}^{Hybrid} = \alpha E_x^{HF} + \beta E_{xc}^{DFT} \tag{2.28}$$

where α and β are the respective fractions of HF exchange and DFT exchange-correlation. In HF theory, the electron exchange energy is calculated exactly, via the aforementioned Slater determinant, thereby accounting for Pauli exclusion. This effect is often described as the *"exchange hole"* which is a representation of the reduced probability of finding another particle of the same type in the region of space (the *"hole"*) around a given particle. Moreover, mathematically, the electron exchange energy exactly cancels the *self-interaction* energy, which represents the spurious effect of an electron interacting with itself. For these reasons, the HF exchange is described as being *"exact"*. As mentioned previously, HF theory provides no such exact treatment for electron correlation (i.e. a *"correlation hole"*), which would need to consider the interaction effects of all electrons, not merely those of the same spin[99].

In the LDA and GGA, the exchange-correlation energy density is treated as a homogeneous electron gas. This means that both the exchange and correlation effects are approximated when using DFT. Thus, the inclusion of exact electron exchange improves the estimation of several properties, most

notably semiconductor and insulator band gaps which are often severely underestimated with LDA and GGA functionals[135].

An inconvenience of the Hybrid functional approach is that there is a large increase in the computational overhead. However, this cost is still significantly lower than a Hartree-Fock calculation[j]. Furthermore, when using hybrid functionals, the GGA component is often separable from the Hartree-Fock component. Therefore, it is common practice to further reduce the computational expense by completing a GGA calculation first and then completing the hybrid functional calculation using the output of the GGA calculation[139,140]. This procedure also has the added benefit of observing the difference in performance between the GGA and hybrid functional for a given calculation. The *Heyd, Scuzeria, and Ernzerhof* (*HSE06*) functional[141] is used in this work which incorporates 25% HF exchange, 75% PBE exchange, and a screened Coulomb potential to further reduce the computational expense. This functional is notably accurate for the properties of many crystalline material systems.

DFT + U The Hohenberg Kohn Theorems and Kohn-Sham Equation provide no prescription for improving the exchange-correlation functional. Exchange and correlation effects are more pronounced for localised, strongly interacting electrons. Examples include transition metal oxides and rare earth elements. *Self Interaction Corrections* (*SICs*), *DFT+U*, and the aforementioned *hybrid* functionals have been suggested to alleviate this problem. SICs correct for the self-interaction effect caused by using an approximate exchange-correlation functional. As such, SICs may vary greatly in their mathematical form and application. DFT+U is DFT with an additional, orbital-dependent interaction term (U). Said interaction term is only included for orbitals that are expected to be localised (atom-like). The value of U is coupled to this interaction term and has the effect of shifting the localised orbital energies[99]. This allows the self-interaction effects to be accounted for. DFT+U is based on the Hubbard model[142], wherein the Hamiltonian operator includes the additional U parameter[143].

$$\hat{H} = \hat{H}_{DFT} + \hat{U} \tag{2.29}$$

2.7 Practical Considerations

Besides the aforementioned distinctions between different levels of theory, further practical considerations are often required for electronic structure calculations of semiconducting materials. These are briefly summarised below:

[j]The relationship between the different classes of functional and the computational expense can be summarised as LDA < GGA < meta-GGA « Hybrid « Hartree-Fock Theory.

2.7.1 Electronic Minimisation Algorithms

A computationally tangible algorithm is required to minimise the total energy iteratively and in accordance with the variational principle[128]. The selection of such an algorithm is usually a compromise between reliability and computational cost. Two such examples are the Blocked-Davidson Algorithm[144,145] (BDA) and the Conjugate Gradient Algorithm[146,147,148] (CGA).

For simple calculations, the Blocked-Davidson algorithm is often used. The Blocked-Davidson algorithm is based on the Lanczos algorithm[149] but is adapted to be more efficient for Hamiltonian matrices in which the diagonal part of the matrix dominates the behaviour of the system. This is considered typical for plane wave basis sets[99]. According to the Blocked-Davidson algorithm, the eigenvectors and eigenvalues near the top and bottom of the energy spectrum can be calculated accurately without calculating the entire tridiagonal matrix of values. Since states in the middle of the energy spectrum represent the average energy, the top and bottom states can be estimated, as opposed to directly calculated. Thus, fewer iterative steps are needed to find the ground state, compared to processing each state individually.[150].

In some cases, the Blocked-Davidson algorithm leads to *metastable* (not the lowest energy) solutions or may exhibit non-convergent behaviour. In such situations, the conjugate gradient algorithm may be more effective. This algorithm uses the value and gradient of the energy, at the current iterative step and all previous steps, to calculate the optimal path to the energy minimum. Due to the retention of a large amount of energy and energy gradient data, the computational expense of the conjugate gradient algorithm is greater than the Blocked-Davidson algorithm. However, the conjugate gradient algorithm has the advantage of increased robustness and reliability.

Between each iterative step in the electronic minimisation algorithm, the calculated charge density is *mixed* with a suggested charge density to ensure self-consistency. It also ensures efficient convergence to the global minimum, as opposed to generating a completely new charge density, which could lead to excessively long optimisation pathways or non-convergent behaviour[151]. A popular mixing scheme[152,153,154], implemented in the *Vienna Ab Initio Software Package*[155,153,156,157] (VASP), is one such example. This mixing scheme considers an input charge density ρ_{in}, an output charge density ρ_{out}, and the difference between these two densities. By appropriate weighting and incorporating each of the successive input densities as a subspace,[k] a self-consistent solution can be obtained after a

[k] A subspace is a subset of a vector space that is itself another vector space. A vector space is a set of vectors associated with a field of scalars[100].

sufficient number of iterations[151].

2.7.2 Convergence Testing

Plane-wave basis sets are convenient for periodic solids as their mathematical form and simplicity complement the computational implementation. When using a plane wave basis set, the precision of the calculation can be tuned via a single parameter, the *plane wave cutoff energy*. This determines the number of plane wave basis functions that are included in the calculation[114]. The optimal value of the plane wave cutoff energy is usually determined by increasing it until the value of the property of interest is consistent. Higher values beyond this point are not required because the computational overhead increases with the value of the plane wave cutoff energy but with an insubstantial improvement in convergence. For instance, if a given property requires a level of precision of 0.01 eV, the optimal cutoff energy is the minimum value at which the calculated property is constant within 0.01 eV. The same is true for the Monkhorst-Pack grid density and other quantities.

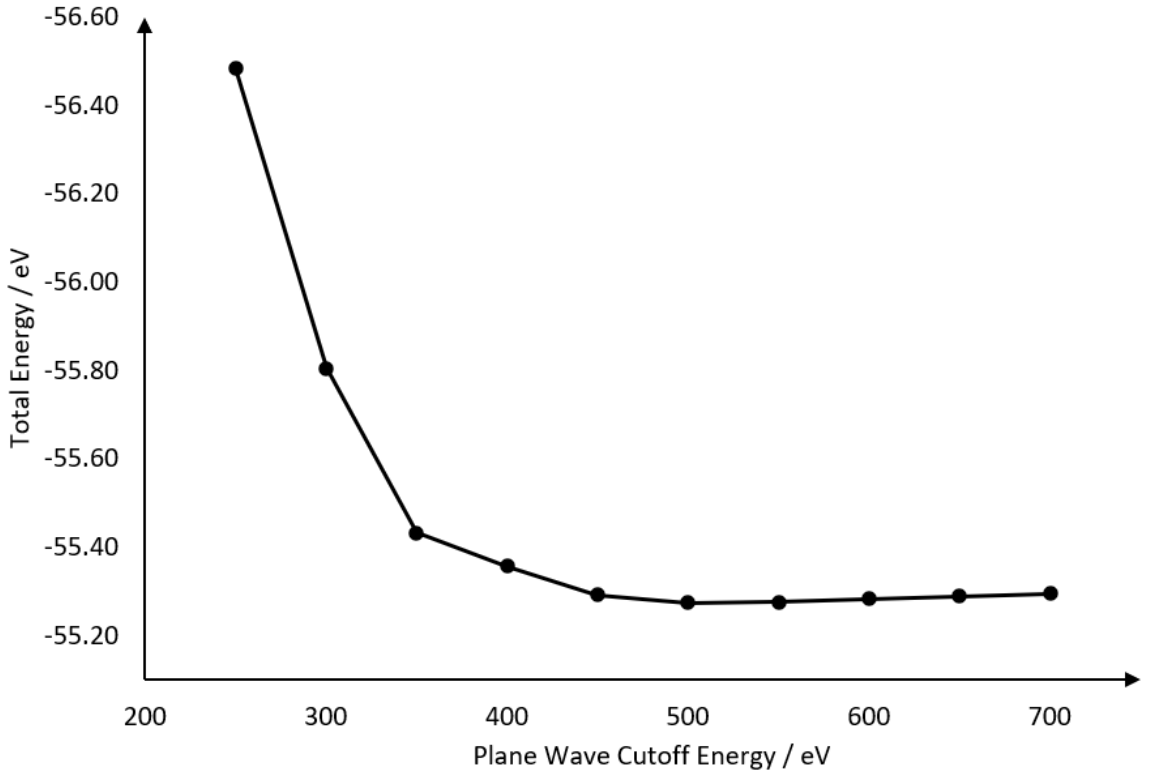


Figure 2.5: Convergence testing data of the plane wave cutoff energy for rutile titanium dioxide. The point of this figure is to illustrate how various quantities plateau after a certain value. In this case, there is no need to increase the plane wave cutoff energy beyond 500 eV. Different quantities (*total energy*, surface energy, ionisation potential, etc.) may plateau at different cutoff energies.

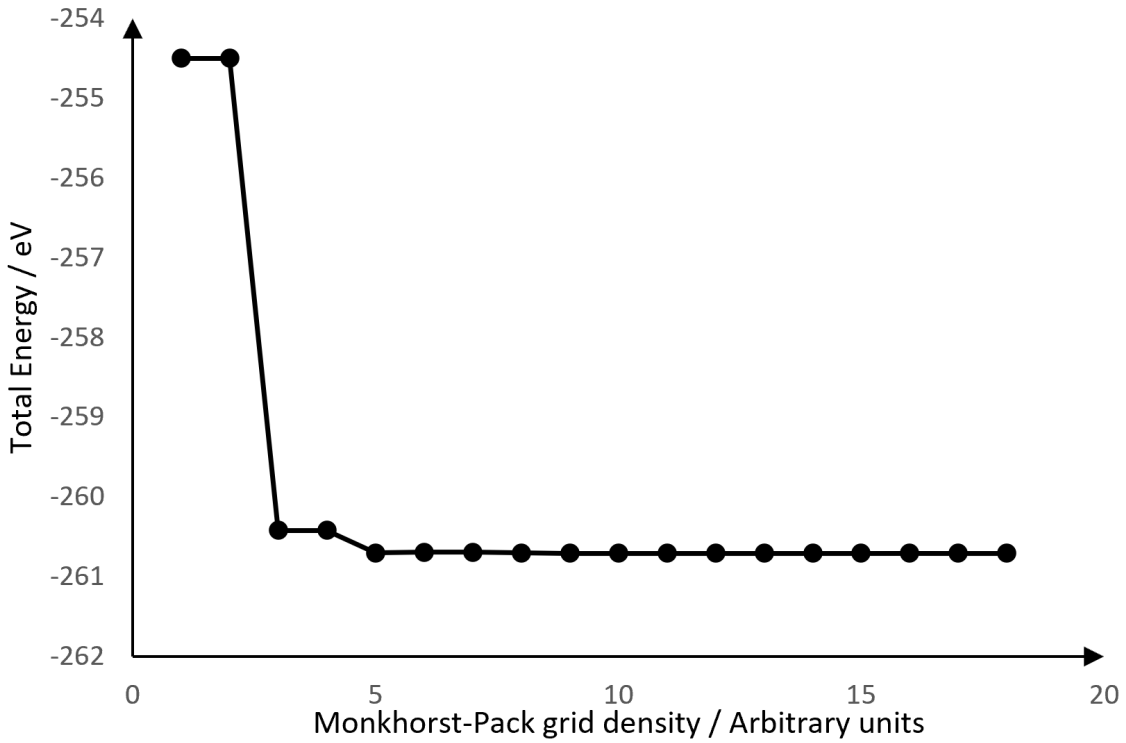


Figure 2.6: Convergence testing data of the Monkhorst-Pack grid density for a 50:50 germanium-lead alloy (See the *Catalytic Activity* chapter for a further discussion of such compounds). Similarly to Figure 2.5, there is no need to increase the Monkhorst-Pack grid density beyond 5 units, after which the total energy is consistent within an acceptable margin of error.

2.7.3 Geometry Optimisation Algorithms

Even with a suitable electronic minimisation algorithm and optimal convergence parameters, simply setting the lattice parameter to the experimentally determined lattice parameter can lead to inaccuracies in the calculated material properties. This is because the energetic ground state of the system varies between different functionals for the lattice parameter. Hence, once convergence testing is complete, it is common practice to optimise the volume, shape and atomic positions of the unit cell for the chosen functional. This is routinely performed by allowing the shape and atomic positions to relax for a set of fixed volumes across a reasonable range[158]. Assuming that the crystalline phase is stable, when the total energies of these systems are plotted as a function of the unit cell volume, the plot is roughly parabolic about the global minimum. The optimal unit cell volume can then be obtained by fitting an *Equation of State* (EoS) to the same plot[159].

Algorithms that allow the volume, shape and internal parameters to be optimised simultaneously are also available. However, these algorithms are often only accurate if the initial structure is very close to the global minimum[160]. In addition, to reduce the computational overhead of calculations that require the hybrid HSE06 functional, it is common practice to first optimise the geometry with the

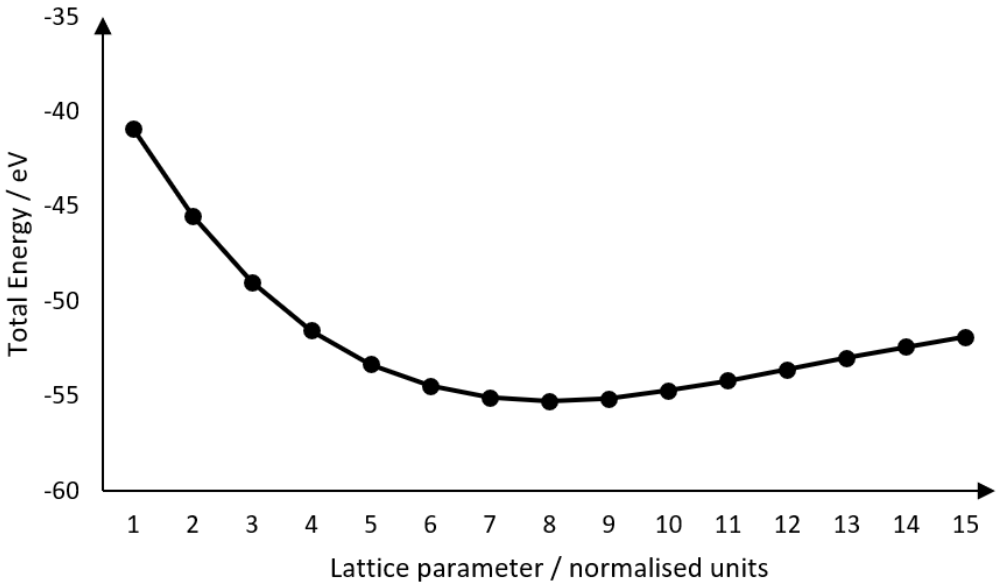


Figure 2.7: Plot of the variation in the total energy across various lattice parameters. The lowest energy point interpolated from this data, via an Equation of State, would be used as the lattice parameter for further calculations.

GGA PBEsol functional and to perform the electronic minimisation on the output with the HSE06 functional[161]. This is usually less computationally expensive than relaxing the lattice with the HSE06 functional and rarely produces spurious results[162], although ideally all calculations should be performed at the same level of theory. Much like the electronic minimisation procedure, different algorithms can be implemented to handle how the atomic positions are modified and optimised. The selection of which normally depends on a compromise between the accuracy/precision and the computational cost. In this work, all of the geometry optimisations are performed using the conjugate gradient algorithm[163].

2.7.4 Data Parallelisation

In the context of DFT. Parallelisation may be performed across electronic bands and/or across the number of **k**-points. Efficiency is usually maximised when both forms of parallelisation are used. Whilst parallelisation across bands is almost always recommended for multi-core or multi-node computers, parallelisation across **k**-points is often severely limited by the memory capabilities of the computer. This is because the orbital representations of the wavefunction are not parallelisable alongside each independent grouping of **k**-points. Hence the wavefunction data has to be copied multiple times according to the parallelisation scheme. Consequently, the parallelisation scheme is usually tailored according to the number of bands, **k**-points, computing nodes, and the amount of memory available per computing node[164,165].

2.8 Applications to Photoelectrochemical Cells

The above theoretical treatment of crystalline materials will now be brought into the context of photoelectrochemical cell electrodes by reviewing the previously mentioned photoelectrochemical cell criteria from a computational perspective.

Band gap The band gap is a routinely extracted parameter from a DFT calculation. Its extraction is amenable to a computational materials screening procedure because the positions of the electronic states as a function of the electron energy are among the most commonly produced outputs of an electronic structure calculation. As such, beyond the DFT calculation itself, there is very little additional effort required to calculate the band gap of a candidate semiconducting material. To date, the band gap is most accurately calculated via a hybrid functional. In particular, the HSE06 hybrid functional is known for its accuracy of crystalline semiconductor band gaps[135].

Band Alignment Despite the calculation of the band gap being a relatively trivial task, the absolute band energies are much more difficult to calculate. This is due to the periodic boundary conditions of the bulk calculation preventing the definition of an absolute reference potential. This problem is further elaborated on in the next chapter.

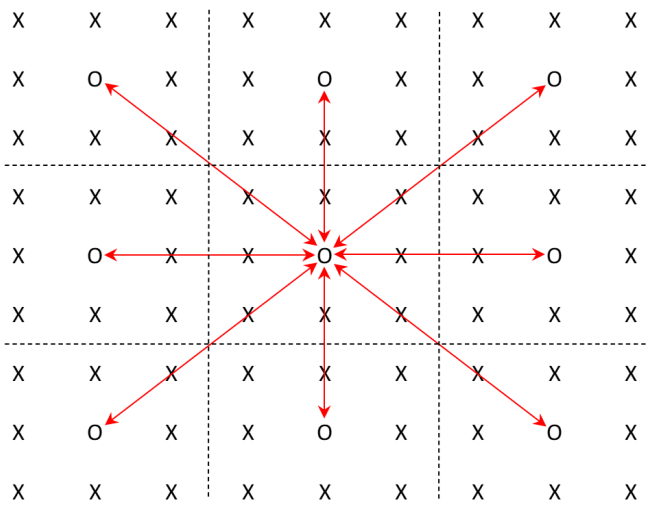


Figure 2.8: Illustration of the interaction between a central defect (O) and its periodic images within a supercell of a square lattice composed of lattice points (X). The red arrows represent this interaction. The dashed lines designate the arbitrarily defined bounds of the unit cell.

Photovoltage Both intrinsic and extrinsic defects, which affect the photovoltage through recombination mechanisms, may be modelled with DFT. In general, this involves creating a supercell expansion of the primitive unit cell. The size and dimensions of the supercell are often chosen to account for the defect itself in addition to minimising the extent of the interaction between said defects and their

periodic images (See Figure 2.8). Such interactions may also be mitigated by so-called *finite size corrections*[166]. Moreover, given the broad variety of hypothetical defects that may affect the photovoltage, as a first step in establishing a materials screening procedure, we only consider the intrinsic effects of *polarons* in this work, which do not require the removal or addition of atoms to the crystal. This is reviewed in more detail in the fourth chapter.

Corrosion Resistance, Catalytic Activity, and Catalytic Selectivity Assuming the various reaction mechanisms are known. These aspects of the photoelectrochemical cell may be calculated by computing the total energies of the reaction intermediates and then comparing these energies to other reaction mechanisms. This includes the OER, HER, corrosion mechanisms, and any side reactions. The catalytic activity of hypothetical corrosion-resistant alloys is the focus of the fifth chapter.

Band Alignment & Ionicity

3.1 Introduction

3.1.1 Band Alignment

The *Ionisation Potential* (IP) and *Electron Affinity* (EA) are important quantities for the selection of photoelectrochemical cell electrodes as they correspond to the absolute positions of the *band edge eigenvalues* for a semiconducting or insulating material[167]. Formally, the ionisation potential corresponds to the minimum energy change associated with the removal of an electron from the Valence Band Maximum (VBM) to the vacuum potential (ϕ). The electron affinity corresponds to the minimum energy change associated with the addition of an electron to the Conduction Band Minimum (CBM) from the vacuum potential. Together, the valence band maximum and conduction band minimum are referred to as the *band edge eigenvalues* (see Figure 3.1). The ionisation potential and electron affinity must be known to select candidate electrodes whose band edge eigenvalues straddle the respective hydrogen evolution reaction and oxygen evolution reaction redox potentials. The ionisation potential and electron affinity are commonly measured via photoemission techniques, such as photoelectron spectroscopy[24,19], or electrochemical techniques, such as cyclic voltammetry[24]. Since the ionisation potential and electron affinity are influenced by both the surface and bulk properties of the crystal[168,169], there is no singular ionisation potential or electron affinity for a given compound but rather multiple values. This is because each compound has multiple surface terminations according to its crystallographic planes, which each induce a different *surface/interface dipole*[99]. This is commonly referred to as *orientation dependence*[170]. Each of these surface terminations may be catalogued using the corresponding Miller index[16]. In addition to orientation dependence, it is challenging to

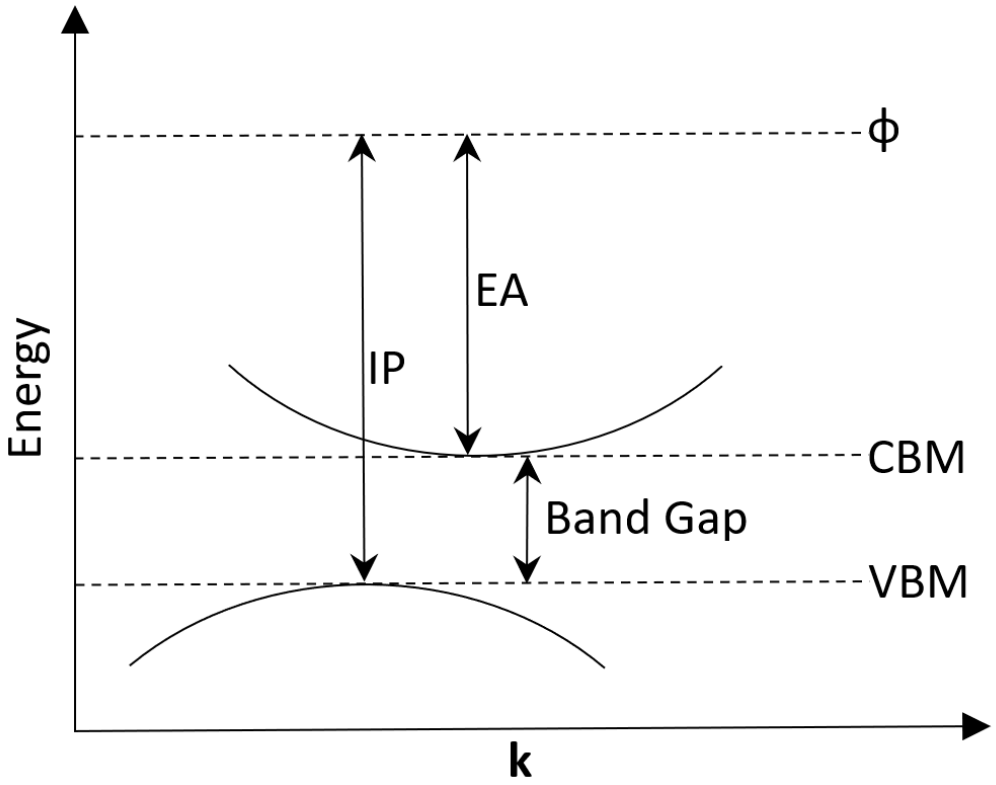


Figure 3.1: Illustration of how the band gap, ionisation potential (IP), electron affinity (EA), vacuum potential (ϕ), conduction band minimum (CBM), and valence band maximum (VBM) relate to one another.

ensure experimental consistency across different studies because sample qualities can vary. As such, there are many possible candidates for photoelectrochemical cell photoelectrodes, with multiple surface terminations and variable synthesisabilties. In light of this, the possibility of characterising such a large database of compounds is not economical through current experimental means. However, the computational screening of such a database may be a more achievable short-term goal.

3.1.2 Computational Band Alignment

One of the most common and intuitive methods of accurately estimating the ionisation potential and electron affinity is to create a slab model expansion of the primitive unit cell such that it contains a bulk component, the surface termination of interest, and a vacuum region[171] (see Figure 3.2). Beyond the creation of the surface, the purpose of the vacuum region is to calibrate the calculated eigenvalues according to an absolute reference potential[a]. Due to the periodic boundary conditions used in many crystalline density functional theory calculations, the eigenvaules from such calculations are arbitrarily defined[173]. Thus, the slab model method explicitly creates a vacuum region to properly align the

[a]For photoelectron spectroscopy[24], this reference potential is the electron energy at vacuum. For cyclic voltammetry, the reference potential is the Standard Hydrogen Electrode[172] (SHE).

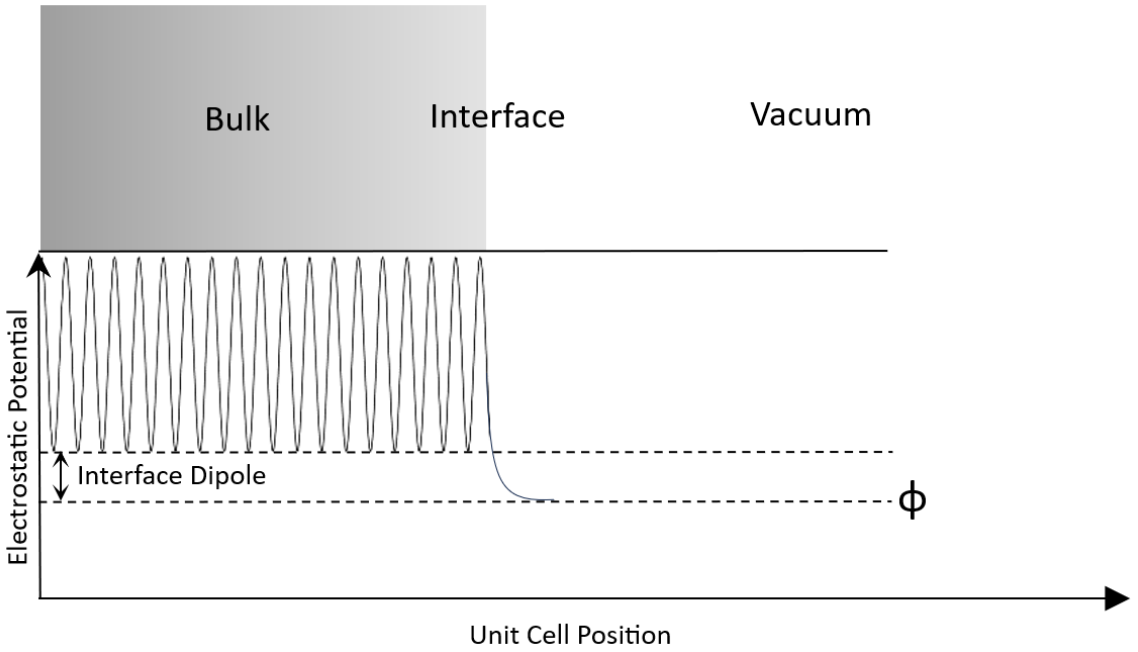


Figure 3.2: Illustration of how the electrostatic potential varies across the bulk, surface, and vacuum regions of a slab model. The interface dipole determines how the electrostatic potential varies across the boundary between the interface and vacuum regions. Thus, affecting the extent to which the vacuum potential differs from the bulk periodic potential.

eigenvalues to one another. This is referred to as *band alignment* and is performed by assuming Janak's theorem and deducting the calculated value of the electrostatic potential in the vacuum region (ϕ) from the calculated band edge eigenvalues (ϵ):

$$\epsilon_{absolute} = \epsilon - \phi \tag{3.1}$$

Whilst this method of computational band alignment is reliable and systematic, several challenges have prevented it from being applied to a large database of compounds[174]. These include the widths of the slab and vacuum in real space needing to be converged for the band edge eigenvalues, the atomic positions needing to be relaxed to account for structural differences in the bulk and surface layers of the slab, and hybrid functionals often producing the most accurate results[168].

With the computing resources that are available to date, the overhead per calculation remains uneconomical for a large database of compounds with multiple surface terminations. A common alternative to the slab model approach involves building an embedded cluster model[175]. However, this is also of comparable computational expense. This problem has incentivised the development of low-cost band alignment procedures. Past work has focused on the comparison of charge neutrality, impurity, core and defect transition levels[176,177,178,179,180], although none have yet to reach the desired compromise of computational efficiency and accuracy. However, a band alignment procedure initially introduced

by Frensley and Kroemer[25,26,27,28] ($F\&K$) may provide a useful representation of the bulk component of the ionisation potential and electron affinity. In this approach, the average interstitial electrostatic potential is used as an approximation of the vacuum reference potential.

Frensley and Kroemer used the *Zinc Blende* (ZB) crystal structure as a test case. The Zinc Blende lattice is composed of two oppositely charged, interpenetrating tetrahedral sublattices. Consequently, the two Zinc Blende interstices (Octahedral (O_h) and Tetrahedral (T_d)) have an equal and opposite electrostatic potential. These interstices are named as such due to their nearest neighbour coordination, where the tetrahedral interstice has four nearest neighbour atoms and the octahedral interstice has eight nearest neighbour atoms.

The equal and opposite relation in the electrostatic potential of O_h and T_d can be confirmed by calculating the *Madelung potential* at these sites via an *Ewald summation*[181]. In this formalism, the lattice is assumed to be composed of *point charges*. The Madelung potential is the electrostatic potential of an ion in the lattice within this point charge formalism (see Sections B.10 and B.9 for elaborations on the Ewald Summation and Madelung Potential respectively). A point charge is a representation of a charge that is restricted to a point in space. As such, it has no radius or volume. The Ewald summation is a method with which to efficiently calculate the Madelung potential by computing the short-range interactions in real space and the long-range interactions in reciprocal space. Sampling the interstices at this level of theory results in an equal and opposite potential (V) for O_h and T_d. Thus making their average (ϕ_{ZB}) equal to zero.

$$\phi_{ZB} = \frac{V^{O_h} + V^{T_d}}{2} = 0 \tag{3.2}$$

Equation 3.2 may be demonstrated numerically via the *General Utility Lattice Program*[182]. Examples of such calculations for the Zinc Blende, *Diamond*, *Rocksalt*, and *Perovskite* crystal structures are given in the appendices. Therefore, within this formalism, ϕ_{ZB} could serve as an effective vacuum reference potential. This would be advantageous for the computational screening of electrode band alignments because the only necessary density functional theory calculation would be a bulk primitive cell calculation. Such a calculation would carry a much lower computational overhead compared to the aforementioned slab and embedded cluster approaches. Despite this appeal, the validity of such a method remains debatable. Firstly, because there is no consideration of the surface in the calculation of ϕ_{ZB}, there is an implicit assumption that the bulk and surface components of the ionisation potential are separable. Where the Frensley and Kroemer method calculates only the bulk component, it is assumed that a suitable treatment of the surface component may be subsequently derived. As such,

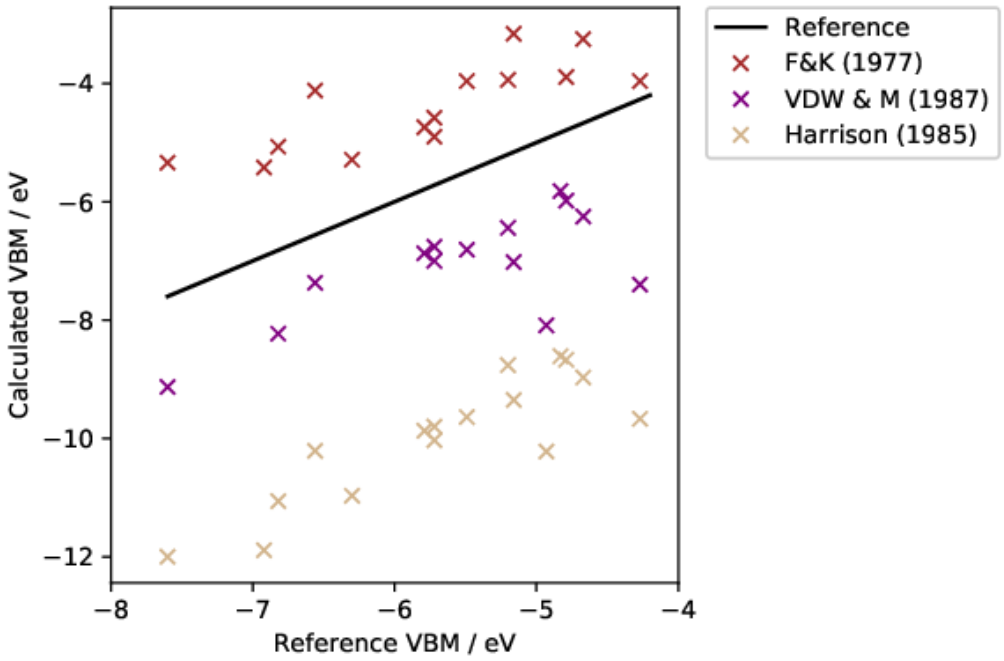


Figure 3.3: A plot of the data initially compiled by Tung[18]. The reference data was compiled by Pelatt et al[19]. Data points that are close to the *Reference* line indicate a stronger agreement with the reference data. It can be seen that the results of Frensley and Kroemer (*F&K*) appear to be no worse than those of Van de Walle and Martin (*VDW & M*) or Harrison. This figure was previously published by the author and their supervisor[20].

the separability of the ionisation potential in this way remains controversial. Secondly, it is not clear how reasonable ϕ_{ZB} is as an approximation to the vacuum potential, outside of a simplified point charge (Ewald) formalism. Real atoms in a crystalline system do indeed have radii and volumes. Moreover, the delocalisation of the electrons within the crystal is not accounted for in the Frensley and Kroemer method.

These concerns were highlighted in past literature, with Van de Walle and Martin[183] finding poor agreement between their theory of the ionisation potential and the Frensley and Kroemer method. This was attributed to there still being substantial charge density in the interstitial region, which offsets the accuracy of ϕ_{ZB}. Conversely, Yu et al[184] noted that the Frensley and Kroemer method was no better or worse than other band alignment procedures of the time, including the work of Van de Walle and Martin. Tung[18] later reviewed multiple ionisation potential calculations and found similar trends between the studies of Frensley and Kroemer[25,26,27,28], VDW & Martin[183], and Harrison[185]. The data from Tung's study is plotted in Figure 3.3, alongside a compilation of experimental measurements from Pelatt et al[19].

In recent years, variations of the Frensley and Kroemer method have also appeared independently. Choe et al[186] used a similar internal reference method for the alignment of metal-metal and metal-

semiconductor interfaces, whilst Butler et al[24] used the centres of *metal organic framework* pores[b] as an effective reference potential. Despite this recent interest, it is still unclear whether ϕ_{ZB} is a suitable reference potential. Benchmarking regarding the dependence of ϕ_{ZB} on the size of the interstices and the DFT exchange-correlation functional has not been determined, and could be quickly performed with modern computational resources. Given the renewed interest in low-cost alternatives to band alignment, it is desirable to investigate the accuracy and reliability of the Frensley and Kroemer method with modern computational tools and standards.

3.1.3 Ionicity

Alongside estimating the ionisation potential, Frensley and Kroemer used the interstitial potentials to calculate the relative *ionicity* of various Zinc Blende compounds[25,26,27,28]. Ionicity is the opposite of *covalency*. In a covalent crystal, such as pure silicon, the electrons are shared equally between the constituent atoms to form the chemical bonds that hold the crystal together. This is because the constituent atoms share the same *electronegativity*, which is the relative extent to which an atomic species attracts/repels electrons[16]. Conversely, in an ionic crystal, the atomic species differ in their electronegativities. As such, electrons from one species are transferred to the other species. The ionicity is a measure of the extent of this charge transfer for a given crystal structure.

Recent work by Xin et al[188] has indicated that doping with more electronegative species resulted in a substantial charge redistribution (an effective *"electronegativity difference"*) between adjacent lattice sites at the surface of the electrode. This further resulted in both enhanced photocatalytic activity and enhanced stability of the electrode material. Since the ionicity is a measure of the electronegativity difference, it could be a useful quantity for further consideration in this regard. Moreover, knowledge of the ionicity may, for example, help to rationalise subsequent investigations of the *lattice energy*. The lattice energy is the energy change associated with forming the crystal structure from its gaseous ions[16] and therefore relates to the synthesisabililty and solubility of the crystal structure.

Frensley and Kroemer Ionicity

In this regard, Frensley and Kroemer considered that the magnitude of the difference in the electrostatic potential of the octahedral and tetrahedral interstices (ΔV) is directly proportional to the magnitude

[b] Metal organic frameworks are highly porous crystal structures that are composed of metal atoms linked together by organic compounds[187].

of the ionicity.

$$\Delta V = V(O_h) - V(T_d) \tag{3.3}$$

Frensley and Kroemer further refined this metric by normalising ΔV for the lattice parameter of the compound (a) using a constant (8.287) derived from the Zinc Blende crystal topology.

$$\delta = \frac{\Delta V a}{8.287} \tag{3.4}$$

The *absolute ionicity* for a given ionicity metric, provided by Equation 3.4, is not considered to be practically useful on its own[189]. This is because there is no definite threshold for the maximum ionicity for a given metric. However, the *normalised ionicity* may be more useful as it allows for the direct comparison with other ionicity metrics:

$$I' = \frac{\delta - I_{min}}{I_{max} - I_{min}} \tag{3.5}$$

Where I' is the normalised ionicity and I_{max} and I_{min} are the maximal and minimal formal oxidation states (i.e. +1, +2, etc) of each compound respectively. The use of the formal oxidation states establishes a common threshold between different ionicity metrics when they are normalised according to Equation 3.5. Furthermore, since the lowest possible ionicity corresponds to a fully covalent structure, $I_{min} = 0$ in all cases. Thus, Equation 3.5 may be reduced to:

$$I' = \frac{\delta}{I_{max}} \tag{3.6}$$

As such, all further mentions of the ionicity in this work may be assumed as being normalised via Equation 3.6.

Bader Ionicity

An alternative way in which to consider the ionicity of a compound is via *Bader charge analysis*. Bader charge analysis[190,191] addresses the problem of partitioning the charge density from an electronic structure calculation into atoms. To partition the charge density, the method uses a *zero flux surface* which is a two-dimensional plane perpendicular to where the charge density between two atoms is a minimum. This means of partitioning the electron density provides an intuitive metric for comparing the ionicity between different compounds because the position of the zero flux surface is expected to

shift from the centre of a chemical bond according to the ionicity of the compound. For example, for a purely covalent interatomic bond, in which the density charge is shared exactly evenly between the two atoms, the zero flux surface would lie in exactly the centre of the bond and each atom would be calculated as having an identical *Bader charge*. Conversely, for an ionic system with an asymmetric charge distribution, the zero flux surface would lie closer to one atom than to another because the more electronegative atom would accumulate more charge and shift the position of the zero flux surface from the centre of the bond. In this case, the Bader charges for each atom would differ[192,193,194,195].

Philips Ionicity

Another alternative way in which to consider the ionicity of a compound is via the *Philips ionicity*. The Philips ionicity considers the *ionic character* (I_{ion}) and *covalent character* (I_{cov}) of a given charge to be separable quantities via:

$$Ionicity = \frac{I_{ion}^2}{I_{ion}^2 + I_{cov}^2} \tag{3.7}$$

Unlike the previous two ionicity metrics, the Philips ionicity metric considers the problem of ionicity from a spectroscopic perspective, with I_{ion} and I_{cov} being determined from spectroscopic data[196,197]. Given the wide variety of ionicity metrics and their varying conceptions of ionicity itself, it would be useful to see the similarities (if any) between the Frensley and Kroemer ionicity and other ionicity metrics.

3.2 Motivation and Materials of Interest

The remainder of this chapter aims to use the Frensley and Kroemer method to calculate the absolute band energies and ionicities for a range of compounds using modern computational tools. The accuracy of the Frensley and Kroemer method is reviewed accordingly and the aforementioned concerns from past literature are discussed. The Frensley and Kroemer ionicity is also compared to the aforementioned Bader and Philips ionicities to determine whether or not there are any common trends between these metrics. The focus in this chapter is not the direct determination of a suitable photoelectrochemical electrode, but rather the benchmarking of a possible means with which to find such an electrode. Consequently, little emphasis is put on a particular compound and its use within a photoelectrochemical cell, but more so on how the trends across compounds vary or correlate with established data. It is further noted that the results and figures within this chapter were previously published by the author

and their supervisor[20].

3.2.1 Zinc Blende Type Structures

Frensley and Kroemer's initial database of Zinc Blende compounds is retained and expanded in this chapter, for comparison and consistency. The Zinc Blende structure is illustrated in Figure 3.4.

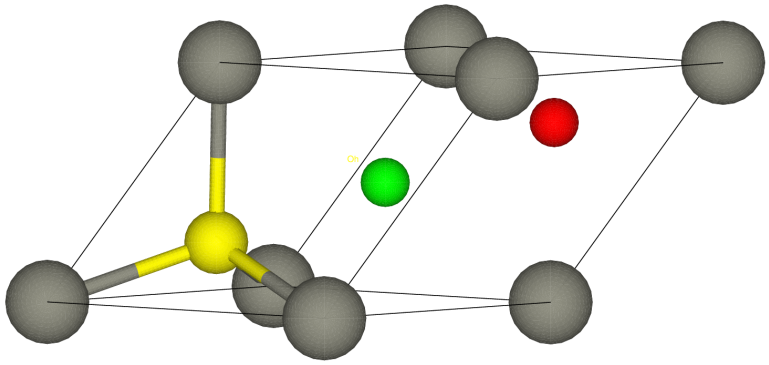


Figure 3.4: The $F\bar{4}3m$ Zinc Blende primitive unit cell. O_h and T_d are marked in green and red respectively. This image was produced using the VESTA program[21,22,23]. This figure was previously published by the author and their supervisor[20].

Silver iodide (AgI) is a semiconducting compound with a direct band gap[198] of 2.22 eV[1] and an electron mobility of 30 $cm^2V^{-1}s^{-1}$[1]. Despite having a relatively low solubility in water, it has seen little use in photoelectrochemical cells as the primary catalyst. Presumably due to its low electron mobility and its tendency for small amounts of silver iodide to dynamically dissolve in water and precipitate from the solution as a colloid suspension. Furthermore, this precipitated form of AgI is difficult to separate from the solution.[199]. Nevertheless, AgI still sees use in photoelectrochemical cells due to its high photocatalytic activity, in combination with higher mobility compounds[200].

Aluminium arsenide (AlAs) is a semiconducting material with an indirect band gap[201] of 2.16 eV[1] - 2.36 eV[19], an electron mobility of 1200 $cm^2V^{-1}s^{-1}$ and a hole mobility of 420 $cm^2V^{-1}s^{-1}$[1]. It is often alloyed or grown with gallium arsenide due to their similarity in lattice parameters. It may be useful as a photoelectrochemical cell electrode due to its insolubility in water[202] and enhanced charge carrier mobility when combined with gallium arsenide[203].

Aluminium phosphide (AlP) is a semiconducting material with an indirect band gap[204] of 2.42

eV[19] - 2.45 eV[1] and an electron mobility of 80 $cm^2V^{-1}s^{-1}$ [1]. It reacts with water to produce phosphine gas, which is flammable and toxic[205], likely explaining its absence from photoelectrochemical cells.

Aluminium antimonide (AlSb) is a semiconducting material with an indirect band gap[206] of 1.60 eV[19,1], an electron mobility of 200-400 $cm^2V^{-1}s^{-1}$ and a hole mobility of 550 $cm^2V^{-1}s^{-1}$ [1]. It is insoluble in water[207]. To date, only the honeycomb monolayer phase has received attention in photoelectrochemical cells[208].

Boron arsenide (BAs) is a semiconducting material with an indirect band gap[209] of 1.5 eV[1], electron mobility of 1400 $cm^2V^{-1}s^{-1}$, and hole mobility of 2100 $cm^2V^{-1}s^{-1}$ [210]. It is praised for its high electron and hole mobility and thermal conductivity, making it an ideal hypothetical material for electronics. It is also insoluble in water[211]. However, the material itself remains difficult to synthesise in large quantities. As such, substantial data for Boron Arsenide as a possible photoelectrochemical component is yet to emerge[212].

Boron nitride (BN) is an insulating material with an indirect band gap[213] of 4.6 eV[1] - 6.2 eV[19]. It is insoluble in water[214] and is not particularly electrically conductive. It is mainly used for its mechanical properties due to its material hardness, and chemical and thermal stability. A monolayer of the hexagonal phase has seen extensive use as a photoelectrochemical cell component, being isostructural with graphene[215].

Boron phosphide (BP) is a semiconducting material with an indirect band gap[216] of 2.1 eV, an electron mobility of 500 $cm^2V^{-1}s^{-1}$ and a hole mobility of 70 $cm^2V^{-1}s^{-1}$ [1]. It is praised for its chemical stability and high thermal conductivity, particularly concerning corrosion. As such, it has seen use as a photoelectrochemical cell electrode to facilitate hydrogen evolution[217].

Beryllium sulphide (BeS) is a semiconducting material with an indirect band gap[218] of 5.5 eV[219]. It has been considered for use in light-emitting diodes[220] but has not been considered for photoelectrochemical cell applications.

Beryllium selenide (BeSe) is a semiconductor with an indirect band gap[221] of 4.45eV[222]. It has been considered for use in optoelectronic nanodevices due to its ability to be alloyed with other compounds that lattice match with silicon[223].

Beryllium telluride (BeTe) is a semiconductor with an indirect band gap[224] of 3.09 eV[222] and an electron mobility of 20 $cm^2V^{-1}s^{-1}$ [1]. Literature concerning the use of this compound in photoelectrochemical cells, and its solubility, is sparse.

Cadmium sulphide (CdS) is a semiconductor with a direct band gap[225] of 2.42 eV[19] . It is insoluble in water[226] and has seen extensive research for photoelectrochemical cell applications, particularly for the hydrogen evolution reaction[227], due to its well-aligned bands. However, it suffers from corrosion and recombination processes during operation[228].

Cadmium selenide (CdSe) is a semiconductor with a direct band gap[229] of 1.74 eV[19]. It is insoluble in water[230] and has seen extensive research for photoelectrochemical cell applications due to its good photoabsorption and charge carrier separation capabilities[227].

Cadmium telluride (CdTe) is a semiconductor with a direct band gap[231] of 1.44 eV[1,19], electron mobility of 1200 $cm^2V^{-1}s^{-1}$, and hole mobility of 50 $cm^2V^{-1}s^{-1}$[1]. It has seen extensive use in photovoltaics (solar cells) due to its ideally sized band gap for maximal solar spectrum absorption efficiency. It is often combined with cadmium selenide to form the p-n junction of a solar cell. CdTe is insoluble in water and research is ongoing for their use in photoelectrochemical cells. However, corrosion remains an issue for this compound[228]. Furthermore, tellurium is a relatively rare element.

Copper bromide (CuBr) is a semiconductor with a direct band gap[232] of 2.91 eV[1]. It has seen little use as a photoelectrochemical cell component, possibly due to its solubility in water.[233]

Copper chloride (CuCl) is a semiconductor with a direct band gap[234] of 3.17 eV[1]. It has seen little use as a photoelectrochemical cell component, possibly due to its solubility in water. However, due to this solubility, it has seen extensive use as an intermediate in the *thermochemical* copper-chlorine process, which is an alternative means with which to facilitate water splitting using excess heat and recyclable catalysts.

Copper iodide (CuI) is a semiconductor with a direct band gap[235] of 2.95 eV[1]. It is insoluble in water[236] and has been suggested as a transparent p-type electrode for hole harvesting due to its high hole mobility and p-type conductivity.[237]

Gallium arsenide (GaAs) is semiconducting material with a direct band gap[70] of 1.35 eV[1] 1.42 eV[19], electron mobility of 8800 $cm^2V^{-1}s^{-1}$, and hole mobility of 400 $cm^2V^{-1}s^{-1}$[1]. It has exceptional photoabsorption properties[70] and is insoluble in water[238]. It is also known to perform well, in the short term, as a photoelectrochemical cell photoanode[239]. However, its long-term performance is severely limited by corrosion processes. Furthermore, GaAs is much less abundant and more costly to produce when compared to more established semiconductors such as silicon. GaAs is also used in this work as an example to demonstrate that the interstitial Madelung potential is zero for the Zinc Blende crystal structure (see appendices).

Gallium phosphide (GaP) is a semiconductor with an indirect band gap[240] of 2.24 eV[1] - 2.26 eV[19], an electron mobility of 300 $cm^2V^{-1}s^{-1}$ and hole mobility of 150 $cm^2V^{-1}s^{-1}$ [1]. It has seen extensive investigation as a photoelectrochemical cell photocathode due to its well-sized band gap and inherent corrosion resistance. However, it suffers from severe photovoltage losses[241].

Gallium antimonide (GaSb) is a semiconductor with a direct band gap[242] of 0.67 eV[1] - 0.726 eV[19], an electron mobility of 4000 $cm^2V^{-1}s^{-1}$, and a hole mobility of 1400 $cm^2V^{-1}s^{-1}$ [1]. It has a high electron mobility and has seen extensive research as a photoelectrochemical cell electrode as an alloy with gallium phosphide. This alloy can be tuned according to the antimony content to have a direct band gap and valence and conduction band edges that are well aligned for water splitting[243].

Indium arsenide (InAs) is a semiconductor with a direct band gap[244] of 0.36 eV[1,19], an electron mobility of 33000 $cm^2V^{-1}s^{-1}$ and a hole mobility of 460 $cm^2V^{-1}s^{-1}$ [1]. Despite its relatively high electron mobility, it has seen little use in photoelectrochemical cells thus far, possibly due to the narrow band gap.

Indium nitride (InN) is a semiconductor with a direct band gap[245] of 0.65 eV[19]. It has seen interest in photoelectrochemical cell development due to its absorption capabilities and high electron mobility[246].

Indium phosphide (InP) is a semiconductor with a direct band gap[247] of 1.27 eV[1] - 1.344 eV[19], an electron mobility of 4600 $cm^2V^{-1}s^{-1}$ and hole mobility of 150 $cm^2V^{-1}s^{-1}$ [1]. It has seen interest as a photoelectrochemical cell photocathode and is believed to suppress photovoltage losses in multi-layer photoelectrochemical cells. However, it suffers from corrosion processes after long-term operation.[248]

Indium antimonide (InSb) is a semiconductor with an direct band gap[249] of 0.163 eV[1] - 0.235[19], electron mobility of 78000 $cm^2V^{-1}s^{-1}$, and a hole mobility of 750 $cm^2V^{-1}s^{-1}$ [1]. It has a very high electron mobility but has mainly seen use in photodetectors. It is insoluble in water[250].

Silicon carbide (SiC) is a semiconductor with an indirect band gap[251] of 2.3 eV[1] - 2.36 eV[19] and an electron mobility of 4000 $cm^2V^{-1}s^{-1}$ [1]. As well as having high mechanical and chemical stability, its bands are well aligned for water splitting, However, the synthesis of acceptable quality crystals remains a challenge[252]. It is insoluble in water[253].

Zinc sulphide (ZnS) is a semiconductor with a direct band gap[254] of 3.54 eV[1] - 3.7 eV[19], and an electron mobility of 180 $cm^2V^{-1}s^{-1}$ and negligible hole mobility[1]. It is insoluble in water[255]. It has been considered for water splitting due to its chemical stability, suitable band gap for photoabsorption

and catalytic activity for hydrogen evolution. However, recombination remains a challenge for this material. Hence it is often used in combination with other materials that aid charge separation[256,257].

Zinc selenide (ZnSe) is a semiconducting material with a direct band gap[258] of 2.58 eV[1] - 2.82 eV[19], an electron mobility of 540 $cm^2V^{-1}s^{-1}$ and a hole mobility of 28 $cm^2V^{-1}s^{-1}$[1]. It is insoluble in water[259]. It has seen use in photoelectrochemical cell devices due to its chemical stability, well-aligned bands, and photoabsorption capabilities. However, similarly to ZnS, is often used in combination with other materials to enhance charge separation[260].

Zinc telluride (ZnTe) is a semiconducting material with a direct band gap[261] of 2.26 eV[1,19], an electron mobility of 340 $cm^2V^{-1}s^{-1}$ and hole mobility of 100 $cm^2V^{-1}s^{-1}$[1]. It has seen use in photoelectrochemical cell devices due to its chemical stability, well-aligned bands, and photoabsorption capabilities. It is insoluble in water[262]. Similarly to ZnS and ZnSe, it is often used in combination with other materials to enhance the charge separation capabilities[263].

3.2.2 Diamond Type Crystals

Diamond-type crystal structures are isostructural to Zinc Blende except that the two species that form the primitive cell are identical. As such, the same Frensley and Kroemer type analysis can be performed on the Diamond structure, as for Zinc Blende. In addition, since there is no difference in the electronegativities of both species, the interstitial potential difference is exactly zero. As such Diamond-type crystals serve as a useful reference point (i.e. fully covalent) for both the Frensley & Kroemer and Bader ionicity metrics. The two diamond-type crystals considered are:

Silicon (Si) Silicon is a semiconducting material with an indirect band gap[264] of 1.12 eV, and electron mobility of 1900 $cm^2V^{-1}s^{-1}$, a hole mobility of 500 $cm^2V^{-1}s^{-1}$[1]. It may have its uses within a photoelectrochemical cell due to its water insolubility[265], photoabsorbing properties, high abundance, good electron mobility, and existing manufacturing infrastructure. However, it has only seen use within a photoelectrochemical cell in combination with other materials because the silicon band gap is known to produce a poor photovoltage[266].

Germanium (Ge) Germanium is a semiconducting material with a direct band gap[267] of 0.67 eV, an electron mobility of 3800 $cm^2V^{-1}s^{-1}$, and a hole mobility of 1820 $cm^2V^{-1}s^{-1}$[1]. Its electron mobility is higher than that of silicon and it is also water-insoluble. However, it is much scarcer than silicon and exhibits poorer photoabsorption properties. As such, its use has been limited to a dopant in other materials to enhance their crystal stability[268].

3.2.3 Rocksalt Type Structures

Rocksalt is another type of cubic, binary crystal structure. Unlike Zinc Blende, the Rocksalt crystal structure involves two interpenetrating *cubic* lattices. Consequently, there is only one symmetry unique interstitial site in the Rocksalt crystal structure. This interstitial site may still be useful for the band alignment, but the ionicity cannot be calculated because there is no second interstitial site with which to calculate the electrostatic potential difference, as illustrated in Figure 3.5.

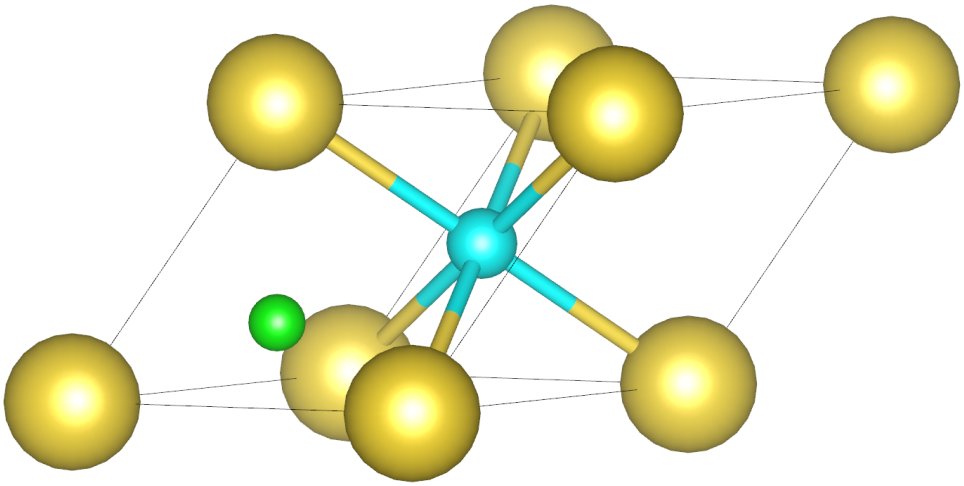


Figure 3.5: The $Fm\bar{3}m$ Rocksalt primitive unit cell. The interstice is marked in green. This image was produced using the VESTA program[21,22,23].

Sodium chloride (NaCl) is an insulating material with a direct band gap[269] of 8.75 eV[270]. It has not seen much use as a photoelectrochemical cell electrode, presumably due to its solubility in water[271]. It is merely used here as an example to demonstrate that the interstitial Madelung potential is zero for the Rocksalt crystal structure (see Section A.2).

Magnesium oxide (MgO) is an insulator with a direct band gap[272] of 7.8 eV[273]. It has seen use in photoelectrochemical cells due to its ability to enhance charge carrier separation[274,275]. However, its solubility in water[276] may raise questions about its long term stability.

3.2.4 Perovskite Type Structures

Perovskite is a cubic, ternary crystal structure. Unlike Zinc Blende and Rocksalt, the Perovskite crystal structure comprises three interpenetrating cubic lattices. Similarly to rocksalt, there is only one symmetry unique interstitial site in the Perovskite crystal structure, as illustrated in Figure 3.6.

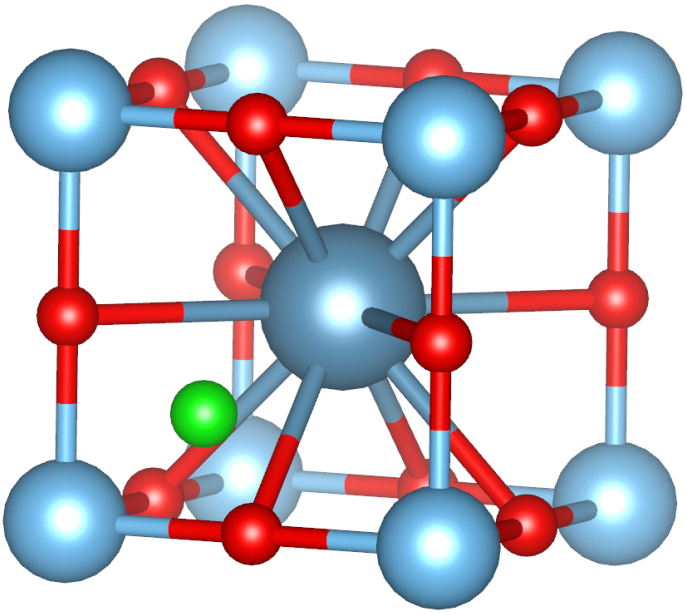


Figure 3.6: The $Pm\bar{3}m$ primitive unit cell. The interstice is marked in green. This image was produced using the VESTA program[21,22,23].

Calcium titanate ($CaTiO_3$) is a semiconducting material with a band gap of 3.4 eV - 3.97 eV[277]. It has seen use as a water splitting photocatalyst when combined with other, more catalytically active materials. It is used here merely as an example to demonstrate that the interstitial Madelung potential is zero for the Perovskite crystal structure (see Section A.2).

Caesium lead iodide ($CsPbI_3$) is a semiconductor with a band gap of 1.73 eV. It, and its derivatives, are among the most highly investigated compounds for solar cells, due to its relatively high photoabsorption efficiency. However, it is known to degrade in the long term under solar cell operating conditions[278]. For water splitting, degradation of the material may occur in the presence of water[279].

3.3 Methods

Each of the Zinc Blende compounds, silicon, germanium, caesium lead iodide, and magnesium oxide were tested using the Frensley and Kroemer method. For each compound, the primitive cell dimensions were optimised by varying the primitive cell volume over an appropriate range, then relaxing the shape and internal coordinates using the PBEsol functional. The total energy across this range of volumes was fitted to the *Murnaghan*[280] equation of state (See Section B.8) to deduce the optimal primitive cell volume.

Following this, a DFT electronic minimisation calculation was performed on each relaxed primitive cell using the LDA, PBEsol, SCAN and HSE06 functionals. A 13x13x13 **k**-point Γ-centred Monkhorst-Pack mesh, a plane wave cutoff energy of 500 eV, and a 170x170x170 Fast Fourier Transform grid were used in all of the above calculations. The tetrahedron method with Blöchl corrections was used to sample the Brillouin zone.

Upon completion of the DFT calculations, the *Macrodensity* code[281,20] was used to extract the value of the electrostatic potential at the interstices (O_h and T_d), which was then used to obtain ϕ_{ZB} via Equation 3.2. The absolute band edges were obtained using the band edge eigenvalues (as ϵ) and ϕ_{ZB} (as ϕ) via Equation 3.1.

The calculated values of O_h and T_d from Equation 3.2 were also used to calculate the Frensley and Kroemer ionicity via Equation 3.6. The Bader effective ionic charges were also extracted from the calculated charge densities using the *Henkleman* code[192] and further normalised via Equation 3.6.

3.4 Results and Discussion

3.4.1 Band Alignment

To determine the accuracy of the Frensley and Kroemer method with modern computational tools, the dependency of $\epsilon_{absolute}$ was investigated for the real-space resolution of the electrostatic potential and the interstitial volume. For clarity, a distinction is made here between the terms *sampling volume* and *Fast Fourier Transform (FFT) grid* in the text. The FFT grid defines the real space resolution of the electrostatic potential and charge density within the unit cell. This was converged similarly to the **k**-point grid density and plane wave cutoff energy for the calculated ionisation potential.

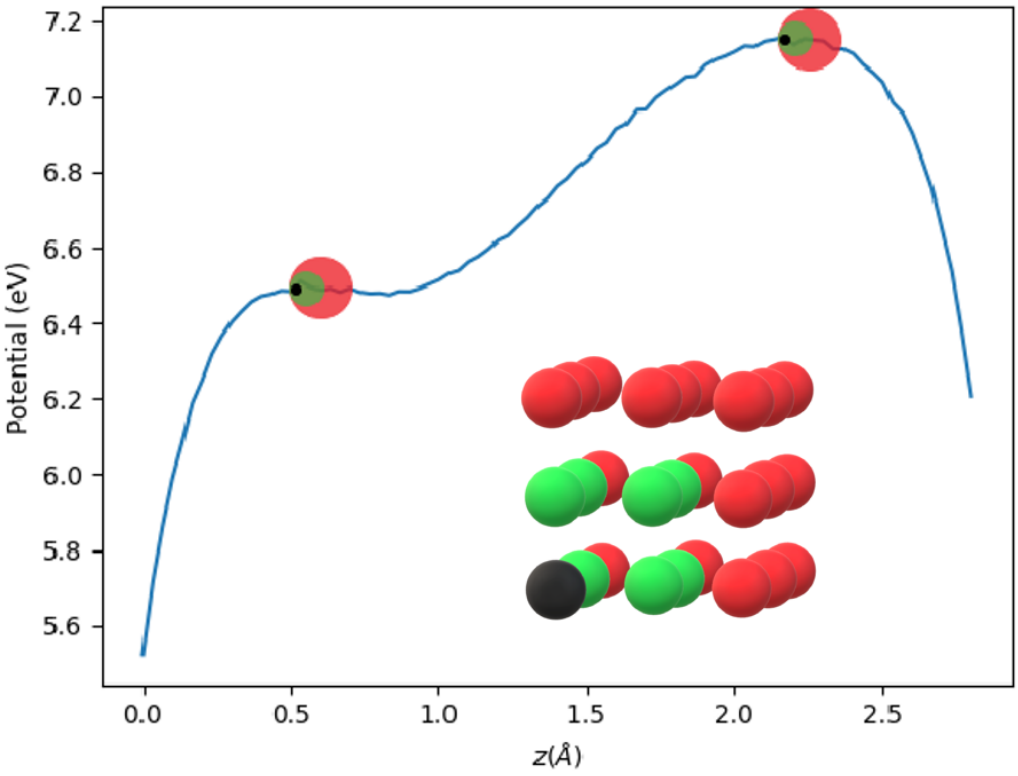


Figure 3.7: Illustration of the FFT grid sampling using the MacroDensity package[24,20]. Plot of the electrostatic potential across the [111] vector in the Zinc Blende lattice, which spans the O_h and T_d interstices. The interstices are positioned at the crests of the peaks in the potential. The array of coloured points in the lower portion illustrates the relationship of points in the sampling volume. The black points represent the sampling volume origin with 2x2x2 and 3x3x3 sampling volumes being represented in green and red respectively. This figure was adapted from work previously published by the author and their supervisor[20].

The default real-space FFT grid for the Zinc Blende system, set by VASP, is a 40x40x40 grid of points. However, an enhanced resolution of 100x100x100 points was necessary to ensure convergence of ϕ. Furthermore, a density of 170x170x170 was required to converge the charge density output for the Bader charge analysis. Unless otherwise stated, an FFT grid density of 170x170x170 was

used throughout this work. The *sampling volume* is the portion of FFT grid points extracted about each interstice via the *Macrodensity* package[24,20]. A 1x1x1 sampling volume corresponds to pointwise sampling, whereas a 10x10x10 sampling volume corresponds to a 10x10x10 grid of FFT grid points. This is illustrated in Figure 3.7.

For the calculation of ϕ, there is no need to increase the sampling volume beyond pointwise sampling. However, to appropriately represent the vacuum potential, the electrostatic potential around the interstice should be constant because the *true* vacuum potential is single-valued. Thus, systematically varying the sampling volume for ϕ allows for an evaluation of the local curvature in the interstitial region. This was used to investigate the appropriateness of the Frensley and Kroemer method for each of the compounds, as shown in Figure 3.8.

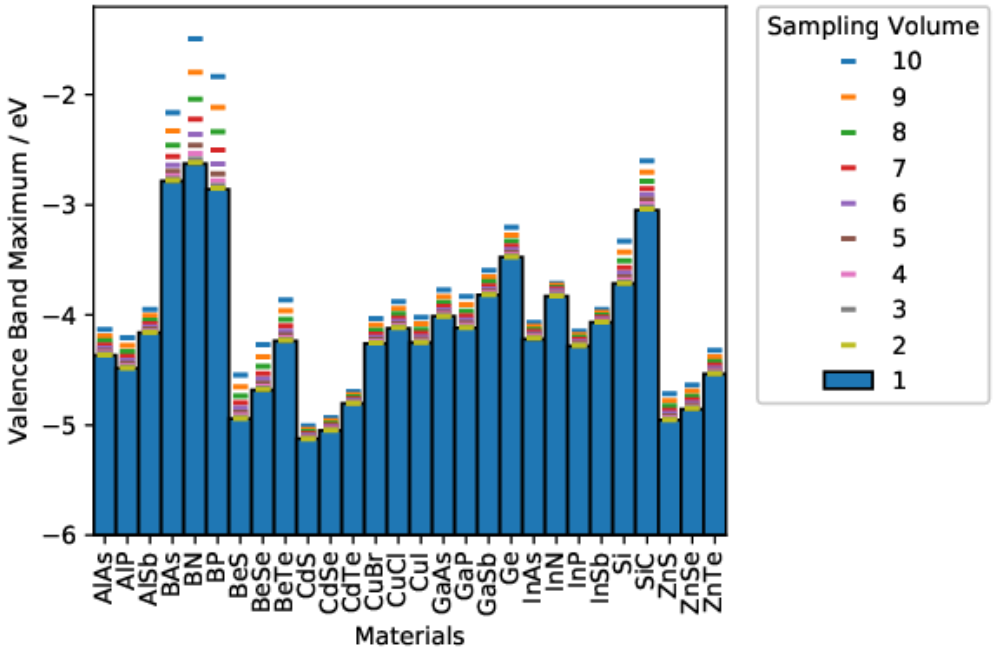


Figure 3.8: Variation in the VBM according to the sampling volume. Each number (n) in the legend represents an n x n x n grid of FFT grid points. The closer each grid is to pointwise sampling, the less local variation there must be in the electrostatic potential about the interstices. The PBEsol functional was used in this case. However, the same qualitative trends were observed for each of the other functionals and more coarse FFT grids. This figure was previously published by the author and their supervisor[20].

Figure 3.8 shows that the boron, beryllium and silicon compounds exhibit the greatest deviation from pointwise sampling, indicating a greater degree of curvature about the interstice. As such, it was expected that these compounds would produce the least accurate value of ϕ. The data in Figure 3.8 was produced using a 300x300x300 FFT grid to ensure that the interstitial region was sampled as precisely as possible. A common trait amongst the boron, beryllium and silicon compounds was their relatively small lattice parameters, which also implied a smaller interstitial volume. To investigate this

further, the deviation of the calculated VBM from reference data[19] and the degree of variation between large and small sampling volumes was plotted as a function of the lattice parameter in Figure 3.9.

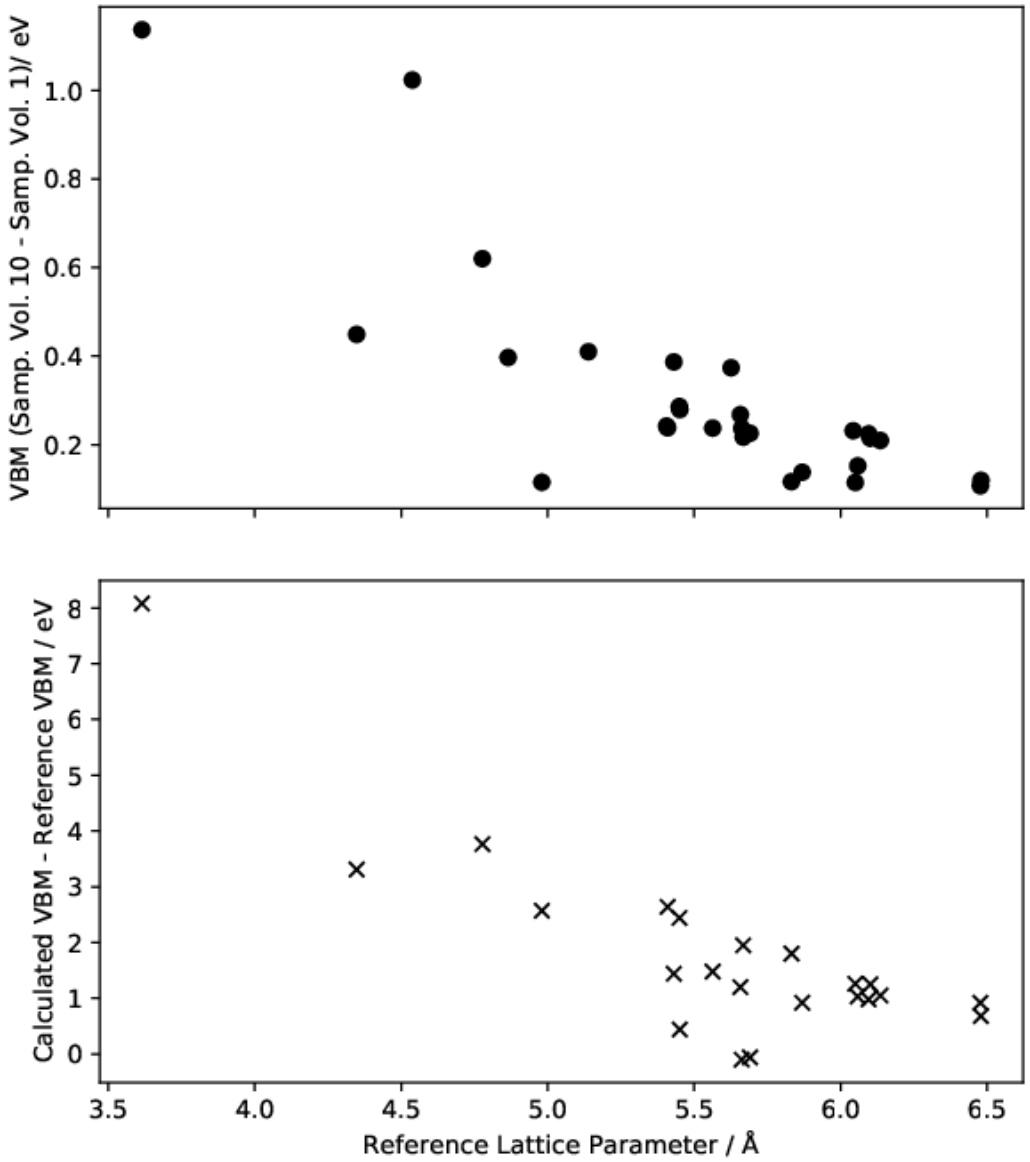
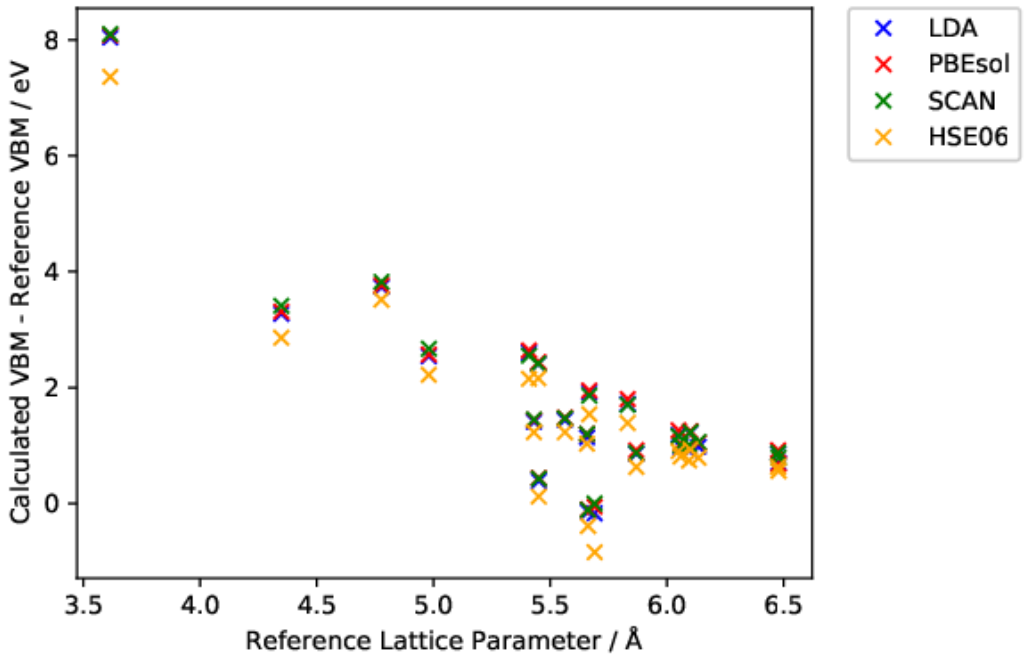


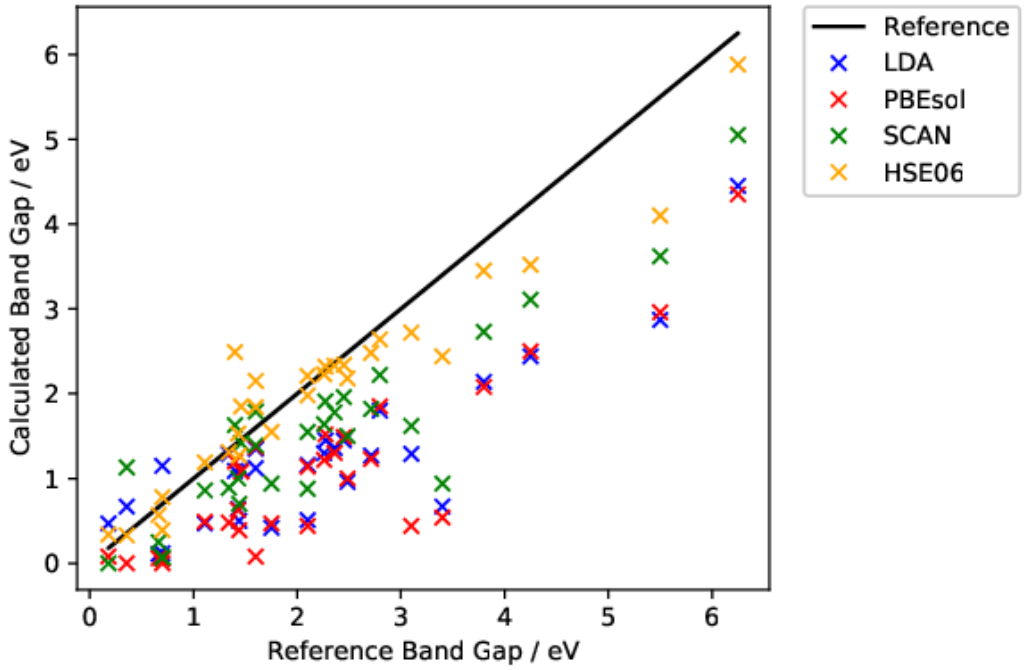
Figure 3.9: (Top) Difference in the VBM between a pointwise and 10x10x10 sampling volume versus the lattice parameter. (Bottom) Difference between the reference VBM[19] and the calculated VBM with single point sampling versus the lattice parameter. The data presented here was produced using the PBEsol functional but the same trend is observed for each of the other functionals. The data shows that the calculated VBM is less representative of the experimental result for smaller lattice parameter compounds. This figure was previously published by the author and their supervisor[20].

Figure 3.9 shows that there was indeed a correlation between the interstitial volume and the accuracy of the Frensley and Kroemer method of band alignment because the agreement between the calculated VBM and reference VBM[19] decreases rapidly for compounds with lattice parameters below 5Å.

This trend was maintained between both the upper and lower plots of Figure 3.9, indicating that adjusting the sampling volume was an adequate means to estimate the accuracy for a given compound. From Figure 3.10a, it was also clear that HSE06 had the lowest deviation from the reference data[19].



(a)



(b)

Figure 3.10: (a) Difference in the calculated and reference valence band maxima versus the lattice parameter of the compound. (b) Comparison of the calculated band gaps with their reference values from Pelatt[19]. Data points that are close to the black line indicate a strong agreement. The data shows that the calculated HSE06 band gaps and ionisation potentials are more representative of both the band gap and the ionisation potential. This figure was previously published by the author and their supervisor[20].

Since the CBM is calculated via $CBM = VBM + bandgap$, and because Figure 3.10b shows that HSE06 more accurately estimates experimental band gaps, HSE06 likely gives a better estimation of the absolute CBM merely due to its ability to predict the band gap more accurately.

To further verify that the accuracy of the Frensley and Kroemer method was not merely an artefact of the DFT calculations, the calculated valence band maxima were compared with Frensley and Kroemer's original work. The results are given in Figure 3.11 which shows that the two methods are in good agreement. Thus, the Frensley and Kroemer methodology appears to be largely insensitive to the choice of DFT exchange-correlation functional.

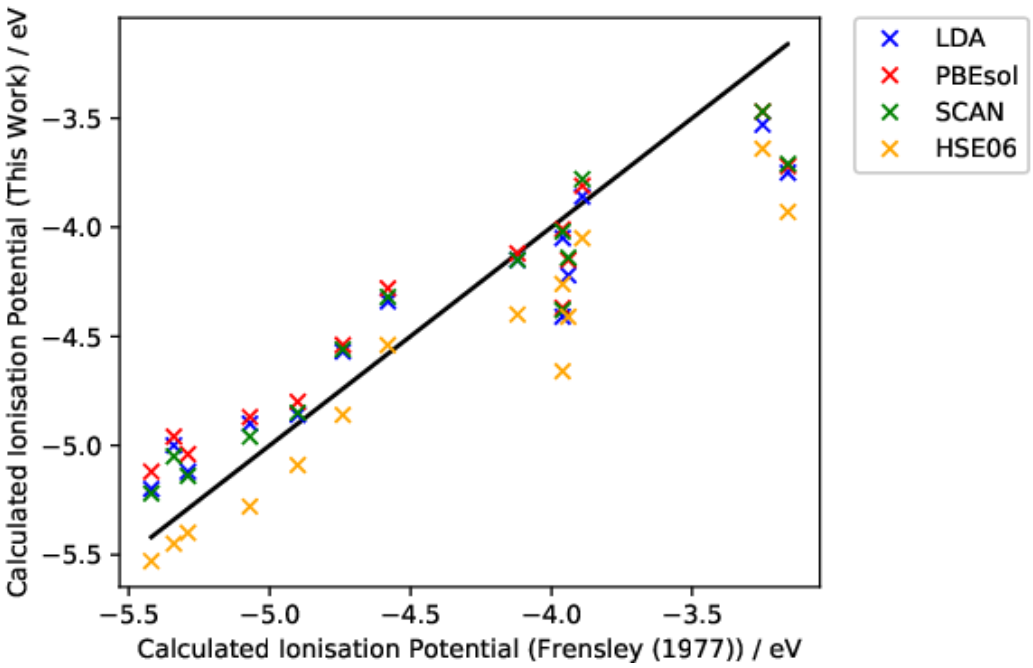


Figure 3.11: Comparison between this work and Frensley and Kroemer's original work[25,26,27,28]. Data points that are close to the black line indicate a strong agreement for that compound. This figure was previously published by the author and their supervisor[20].

Figure 3.12 shows the relative band edge alignment for all of the compounds investigated in this chapter. Boron nitride is a noteworthy outlier, having the only negative electron affinity. Whilst there have been reports of this in past literature[282], BN also has the smallest lattice parameter in the dataset. As such, the reason for the anomalous behaviour of boron nitride could not be conclusively determined based on the presently available data.

Figure 3.13 is a further plot of the HSE06 data, from Figure 3.12, but according to the Standard Hydrogen Electrode (SHE). The Standard Hydrogen Electrode is an alternative reference potential to the vacuum potential that is often more practical for electrochemical measurements. The conversion between the two reference potentials was provided by Trasatti[283] as being:

$$-\epsilon_{SHE} = \epsilon_{absolute} + 4.44 \tag{3.8}$$

where ϵ_{SHE} is the band energy according to the Standard Hydrogen Electrode and is conventionally

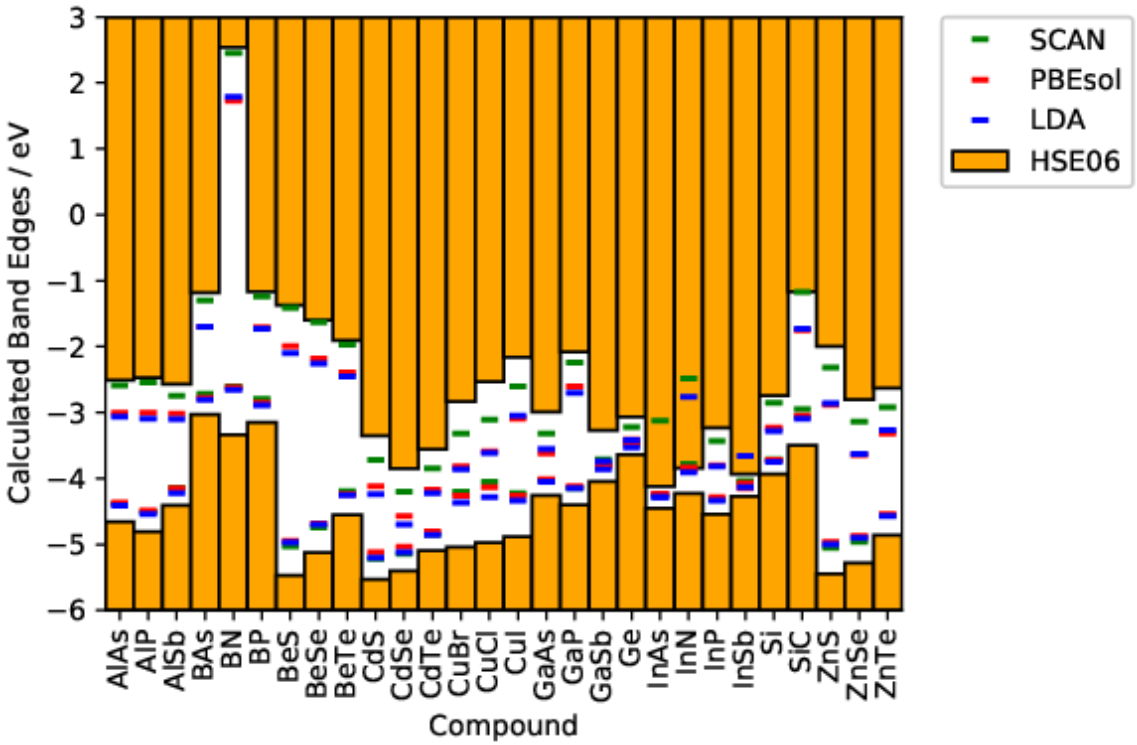


Figure 3.12: Plot of the band alignments for the compounds investigated in this chapter. The lower bars of each colour represent the VBM whilst the upper bars represent the CBM. This figure was previously published by the author and their supervisor[20].

plotted on an inverted axis, hence the negative factor. Upon comparison of Figure 3.13 with, for example, a compilation of band alignments from the work of Li et al[29] (See Figure 3[29]), it can be seen that the alignment of the band edges with the water splitting redox potentials is in good agreement for CdSe and CdS, with the band edges closely straddling the HER/OER redox potentials. However, despite having relatively large interstitial volumes, GaP and GaAs do not straddle the redox potentials as closely. Despite this, the *qualitative* band alignments (i.e. the relative ordering) between CdS, CdSe, GaP, and GaAs are in agreement with Figure 3 from Li et al's work[29]. This further indicates a trade-off with the Frensley and Kroemer method, wherein some quantitative precision is lost in exchange for greater computational speed. The precision of these alignments could, perhaps, be further improved with a consideration of the surface contribution to the ionisation potential. Si and SiC are also far from their expected alignments. This is likely due to their relatively small interstitial volumes.

To further understand the scope of the Frensley and Kroemer method, the same analysis was applied to the rocksalt compound MgO and the cubic perovskite $CsPbI_3$. Contour plots of the electrostatic potential of the primitive unit cells are given in Figure 3.14 which demonstrates a flattening of the potential in the interstitial regions. Using the *General Utility Lattice Program*[182] (GULP), the average electrostatic potential at the interstices was found to be 0V, via an Ewald summation. Using the HSE06 functional, a 500 eV plane wave cutoff energy, a 100x100x100 FFT grid and a Γ-centred 9x9x9

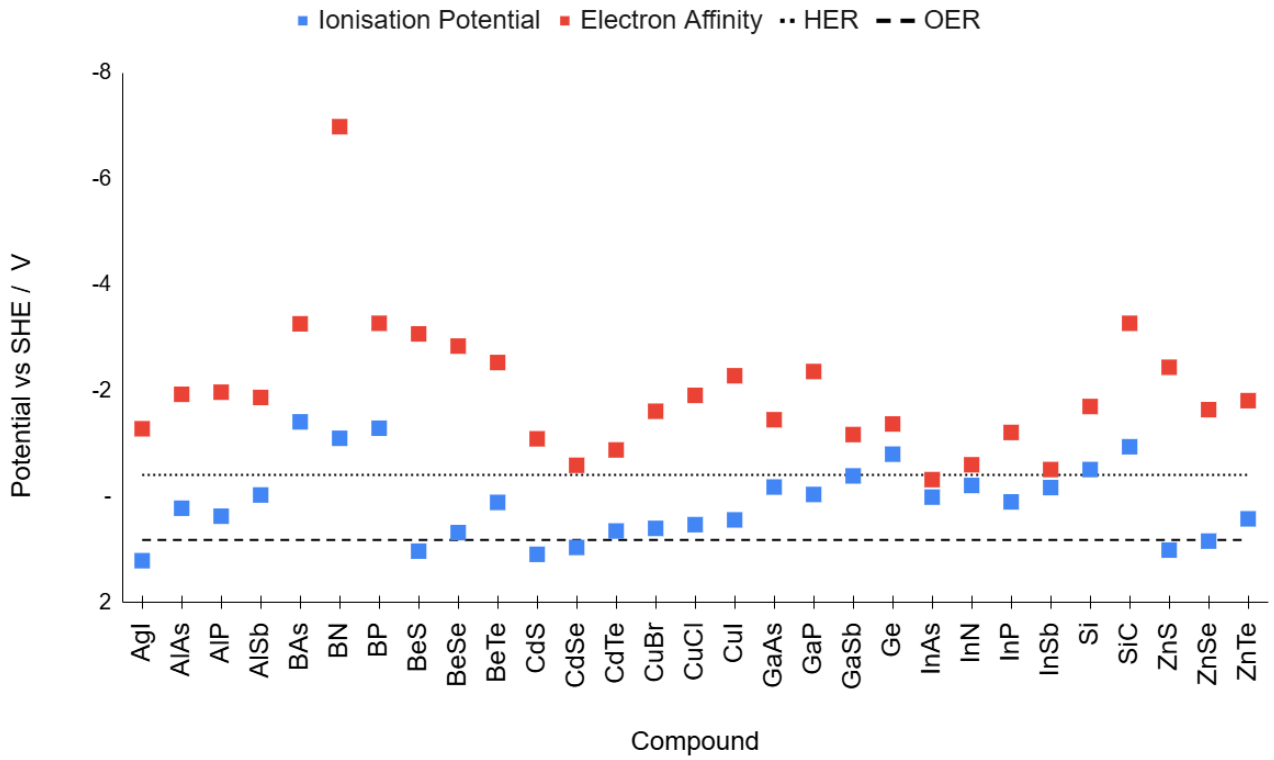
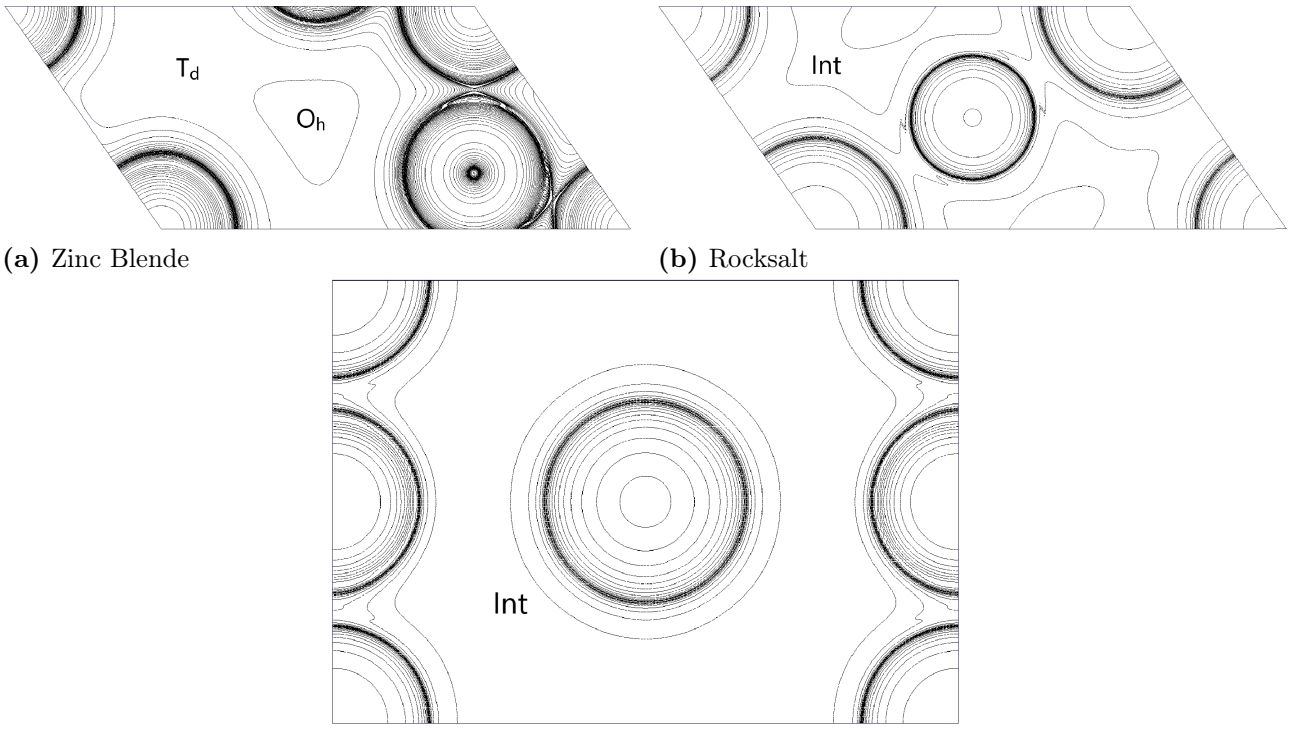


Figure 3.13: Plot of the band alignments for the compounds investigated in this chapter using the Standard Hydrogen Electrode as the reference potential. The red and blue markers represent the electron affinity and ionisation potential of each compound respectively. The dotted and dashed lines represent the HER and OER redox potentials according to the standard hydrogen electrode (-0.41 V and 0.82 V respectively[29]). Only the HSE06 values are included here, for simplicity and because they provide the best overall agreement with experimental data.

k-point grid density, an absolute VBM of -4.85 eV was calculated for MgO. This was within the range of reported literature values of 4.8 - 7 eV[168]. However, -4.85 eV is closer to reported PBE values within this range, despite being a HSE06 calculation. Moreover, it is well below the reported experimental value of 7.16 eV[168]. This could be due to the relatively small lattice parameter of MgO (4.219Å[273]) resulting in similar issues to the boron Zinc Blende compounds. $CsPbI_3$ had the same convergence parameters as above except a 4x4x4 **k**-point grid density was sufficient. In this case, an absolute VBM of -5.13 eV was calculated which was also in line with past literature values[284,285,286].

3.4.2 Ionicity

A comparison of the Frensley and Kroemer ionicity with the Bader and Philips ionicities is given in Figure 3.15 which demonstrates that there is little quantitative agreement between any of these ionicity metrics. However, there is a noticeable qualitative correlation between periodic groups. For instance, in Fig 3.15 (Top), the cadmium compounds follow a similar trend as for both the Bader and Frensley & Kroemer ionicities, with a near-constant offset from the black line. A similar situation can



(a) Zinc Blende

(b) Rocksalt

(c) Perovskite

Figure 3.14: Electrostatic (Hartree) potential contour plots for the AlAs (1-01) plane, the MgO (1-01) plane, and the $CsPbI_3$ (101) plane. The interstices are marked using O_h, T_d and *Int.* These images were produced using the VESTA software[21,22,23]. This figure was previously published by the author and their supervisor[20].

be seen for the indium compounds, but with a different offset. All of the ionicity metrics agreed that silicon and germanium were covalent. When compared with the Bader ionicity, the Philips ionicity[197] agreed poorly with the Frensley and Kroemer ionicity. This may be because the Bader and Frensley & Kroemer ionicity metrics were similar in their interpretation of the ionicity, via effective ionic charges. Conversely, the Phillips ionicity was based on a spectroscopic theory.

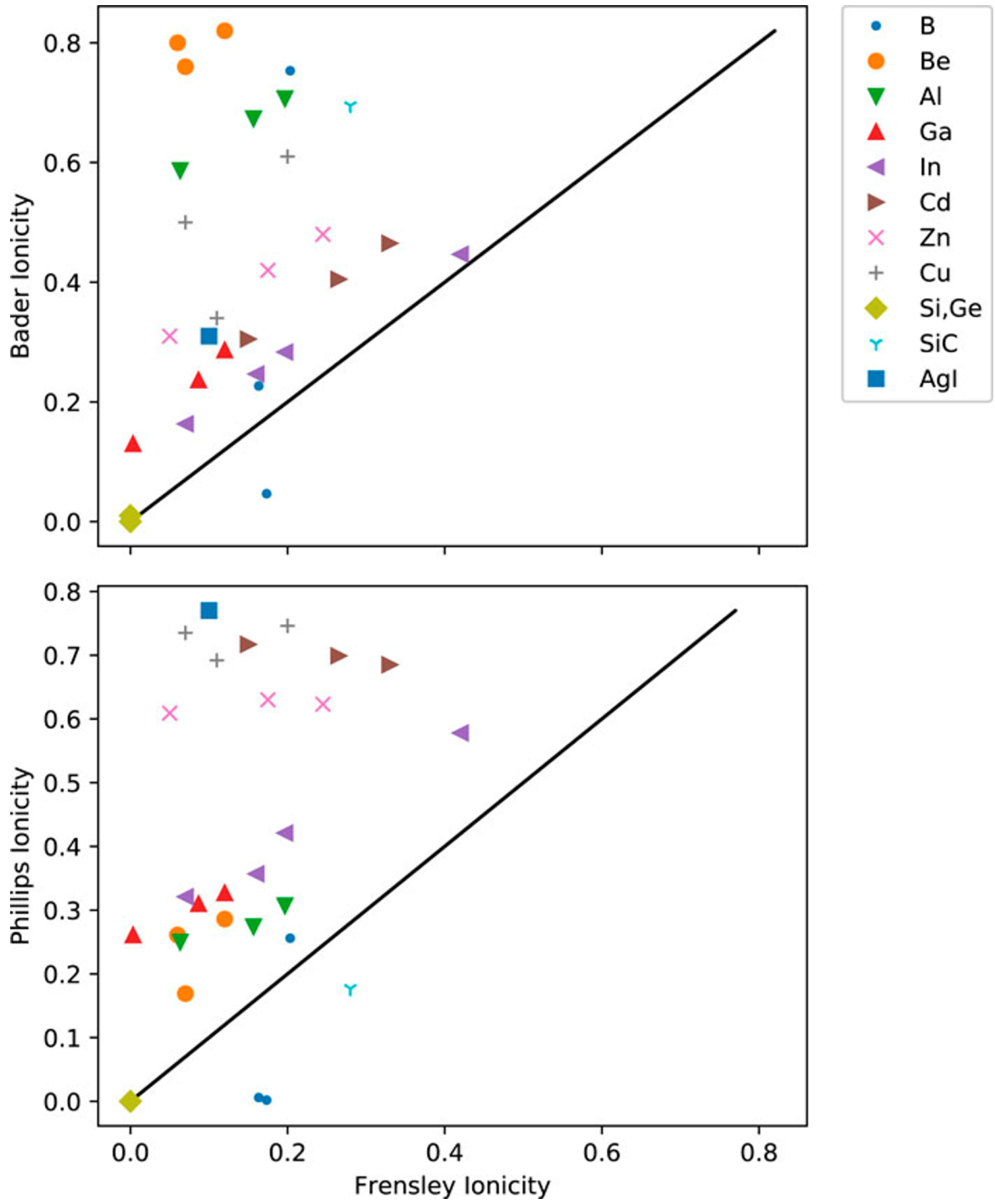


Figure 3.15: (Top) Comparison of the Frensley-Kroemer and Bader charge ionicities. The ionicity metrics are normalised such that 0.0 implies a fully covalent compound, whilst 1.0 implies a fully ionic compound. (Bottom) Comparison of the Frensley and Kroemer and Phillips ionicity. For both the top and bottom plots, the points that are close to the black line indicate a strong agreement. This figure was previously published by the author and their supervisor [20].

3.5 Conclusion

In this chapter, an approach that was first introduced by Frensley and Kroemer was used to rapidly screen the absolute band energies and ionicities of Zinc Blende, Diamond, Rocksalt and Perovskite crystals. Qualitative agreement with experimental data and strong agreement with Frensley and Kroemer's original data was observed. It was further determined that the hybrid functional produced the most accurate results due to its enhanced band gap accuracy. It was also observed that the quality of the estimation worsened as the interstitial volume decreased in size, particularly below 5Å. In this regard, a noteworthy outlier was the compound boron nitride, which was attributed to either its small interstitial volume or a chemical anomaly of the compound itself. The calculated ionicity appeared to follow many of the periodic trends of the Bader ionicity metric, as well as converging more quickly. However, it was found to be more limited in its application due to a need for at least two interstitial sites in the primitive unit cell. Given the knowledge of these limitations, the low computational overhead, and the simplicity of the Frensley and Kroemer method, the calculated ionicities and band alignments may prove useful as a basis for high-throughput screening of the bulk component of the ionisation potential. Further work and improvements for the current procedure may include:

A consideration of orientation dependence Since, the ionisation potential depends on the surface properties of the material, combining the Frensley and Kroemer method with a model of orientation dependence, or interface dipole corrections[287], may improve the estimation of the ionisation potential at a much lower cost than conventional methods. An interface dipole may be modelled as a parallel plate capacitor. This is because the interface dipole may itself be modelled as a parallel arrangement of dipoles, with the electrostatic potential varying smoothly from the surface to a constant value at vacuum. The question that remains is how various surface terminations affect this smooth transition. If this orientation dependence effect could be predicted accurately, it could be combined with the bulk band alignment procedure of Frensley and Kroemer to obtain both the bulk and surface components of the ionisation potential. Thereby obtaining more accurate results.

Application to a wider database of materials As demonstrated in this chapter, the Frensley and Kroemer method is readily applicable to other crystal structure types, such as Rocksalt and Perovskite. Therefore, it would be desirable to expand this methodology to a much wider database of compounds, to gather more data and verify the above conclusions.

Application to a wider variety of pseudopotentials Given that the pseudopotential approximation affects how the electrostatic potential varies from the atomic nucleus, it is conceivable that the choice of pseudopotential approximation affects the accuracy and precision of the Frensley and Kroemer method. Moreover, it would be useful to determine how, and if, the calculated absolute band energies would change when abandoning pseudopotential approximations entirely and favouring an *"all-electron"*[288] methodology for comparison. To this end, separately recalculating the present set of results using a different pseudopotential approximation and an all-electron calculation package would be instructive to better understand the limitations of the presently suggested method.

Reconsideration of the geometry optimisation procedure Fitting a set of optimised volumes to an equation of state is usually a reliable means to ensure accurate structural relaxations. However, Zinc Blende is a sufficiently simple system that is likely to produce accurate results even if the shape, volume, and internal parameters are modified simultaneously. Proceeding in this way would significantly reduce the benchmarking and postprocessing time for these calculations. Moreover, for a given material, if the lattice parameter calculated with PBEsol is different to the lattice parameter calculated with other functionals, the mixing of different levels of theory may have introduced stresses in the crystal that could have affected the resultant electrostatic potential. Therefore, whilst PBEsol is a cheap and accurate means to obtain accurate lattice parameters for many systems, it would be more consistent and systematic to use the same functional for both the structural optimisation and the final electronic minimisation step.

Self Trapping

4.1 Introduction

When a charge passes through a crystal lattice, the electrostatic interaction between the charge and the surrounding ions can be sufficient to distort the local crystal structure. This has the effect of influencing the movement of the charge itself. The magnitude of this effect differs between materials and is therefore an important consideration when understanding charge transport in a material. This local lattice polarisation is known as a *polaron*[289]. In the context of a photoelectrochemical cell, polarons may act as *trap states* (or *"self-trapped"* states) which may impede the efficiency of charge extraction.

Formally speaking, trap states capture and then release charge carriers following sufficient thermal excitation. Hence, if sufficiently close to their respective valence/conduction bands, they have the effect of slowing down charge transport but do not necessarily remove charge carriers from the system. On the other hand, a *recombination centre* tends to lie deeper within the band gap, and as such, is much more likely to induce recombination. This is because more energy is required to thermally excite the charge carrier from the recombination centre, as well as said recombination centre lying closer in energy to oppositely charged bands[70].

Recombination involves two opposing photogenerated charges neutralising each other before either charge can reach the external circuit. Therefore some of the energy absorbed during photoexcitation is lost[290]. However, in some situations, recombination may be desirable. For example, radiative recombination (or spontaneous emission)[70], wherein the charge carrier energy is released as a photon, is the key mechanism behind the function of light-emitting diodes. Another form of recombination

is *Auger*, wherein the kinetic energy of a first carrier is absorbed by a second charge carrier of the same charge. Correspondingly, the first charge carrier promptly recombines, whilst the second is promoted to a higher energy state. A third form of recombination, and the one most detrimental to photoelectrochemical devices, is *non-radiative recombination* (or *Shockley-Read-Hall recombination*)[70], wherein the energy released upon recombination is in the form of a phonon, which manifests as heat loss. These recombination mechanisms are illustrated in Figure 4.1.

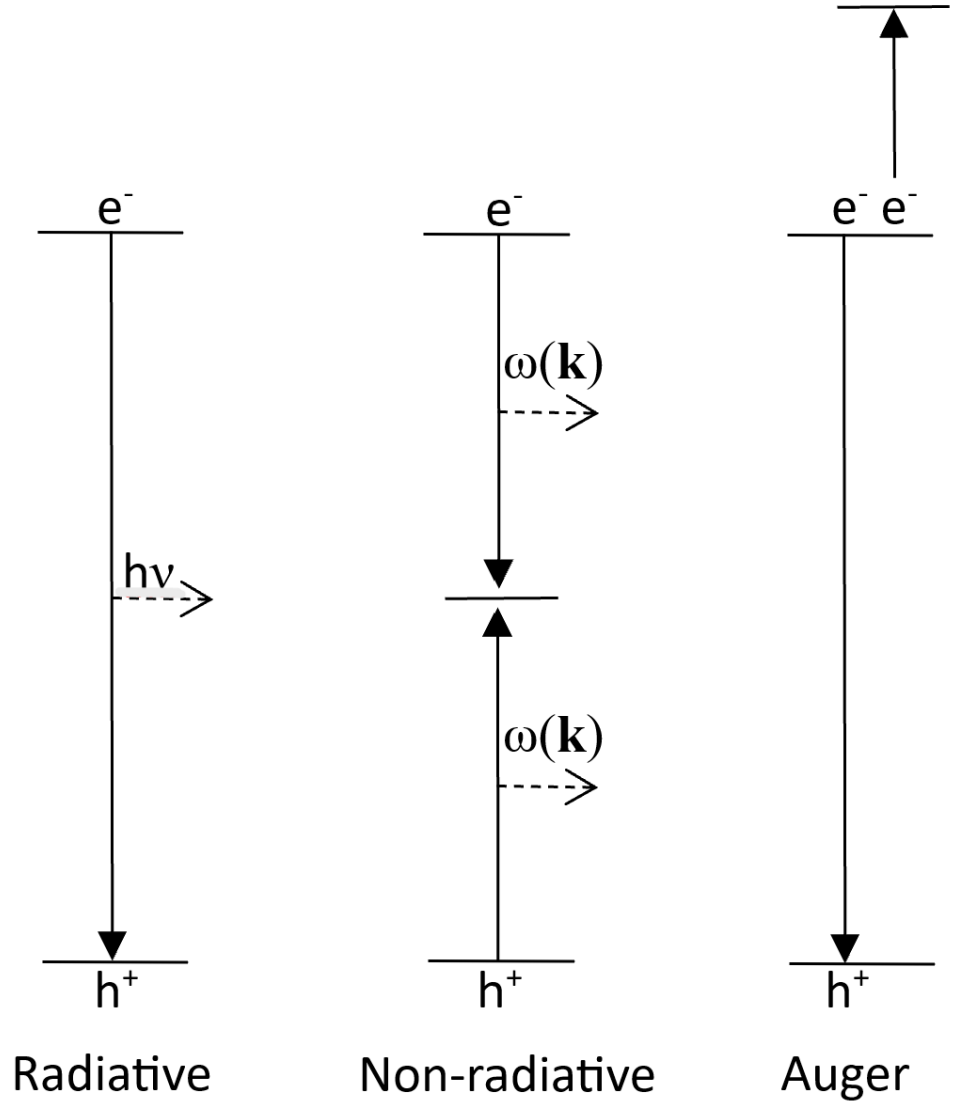


Figure 4.1: Simplified illustration of Radiative, Non-radiative, and Auger recombination for an electron in the conduction band and a hole in the valence band. $h\nu$ represents the emission of a photon. $\omega(\mathbf{k})$ represents the transfer of electron and hole energy to a phonon. The state in the middle of the Non-radiative diagram represents the recombination of the electron and hole via recombination centres.

Trapping and recombination may also occur via intrinsic and extrinsic defects (impurities, vacancies, interstitials, etc.)[291], which also create electronic states within the band gap. However, since the formation of a self-trapped polaron is an intrinsic quality of the material, understanding its formation

in isolation from other defects is an important first step in estimating a material's effectiveness as a photoelectrochemical cell electrode.

4.1.1 Types of Polarons

Polarons can be detected experimentally via Transient Absorption Spectroscopy[292] or Ultraviolet Photoelectron Spectroscopy[293] and can be loosely distinguished as[294] *small* polarons or *large* polarons. Small polarons are so named because they tend to influence a small radius around the charge, usually on the order of a shell of nearest neighbour ions or within one crystal unit cell. Large polarons tend to influence a wider radius, and a greater number of atoms, around a central charge[295]. Small and large polarons also differ in their transport mechanisms. Large polarons pass through the lattice more slowly than a free charge and their transport is further impeded with an increase in temperature. In contrast, small polarons bind the charge at a particular lattice site and require activation energy to move from lattice site to lattice site[296]. This is often facilitated through an increase in the temperature and the process has been more broadly referred to as *thermally activated hopping*[297]. Consequently, the transport mechanisms of small and large polarons have an inverse relationship to the temperature. Small polarons and large polarons can further be distinguished through their optical properties[298].

Beyond this distinction, a wide variety of other types of polarons or polaron-like particles may also be defined. These are summarised in a review by Franchini et al[295], but for the sake of keeping the scope of this chapter manageable, small polarons are the focus hereon.

4.1.2 Small Polaron Transport

Given that their transport mechanism involves hopping from one site in the lattice to another, polaron transport is most commonly modelled via *Marcus Theory*. Marcus theory[299,300,301,302] articulates the transition between lattice sites as the transition from a charge donor species, to a charge acceptor species, without the breaking of any chemical bonds. In this regard, Marcus theory is similar to transition state theory[303], except that in transition state theory chemical bonds are being broken and a *transition state* is formed. Marcus theory differs in that the coupling between the donor and acceptor is relatively weak. Correspondingly, no transition state is formed. As such, Marcus theory is well suited to representing charge transport between molecules, between lattice sites, and through a solvent. Diagrammatically, Marcus theory represents the transition of a polaron from one site to another using a *configuration coordinate diagram*[304], as illustrated in Figure 4.2.

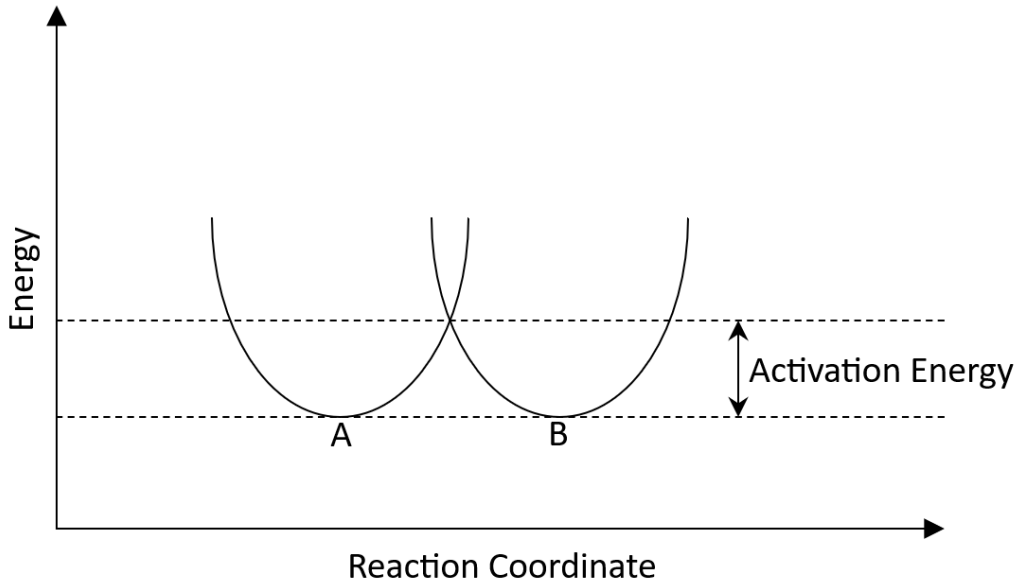


Figure 4.2: Illustration of small polaron transport based on a configuration coordinate diagram. Here, the polaron moves (or *"hops"*) from the donor site (A) to the acceptor site (B) upon receiving enough thermal energy to overcome the activation energy barrier. The parabolic shape of each site is due to a harmonic approximation of the available vibrational states at each site.

4.1.3 Small Polaron Formation

The configuration coordinate diagram also provides a convenient framework with which to understand polaron formation as well as transport. In the case of formation, the donor state is the *band state*, which assumes that the charge is fully delocalised throughout the lattice. The acceptor state is the polaron state. If the polaron state is lower in energy, then it is likely that a polaron will form (see Figure 4.3). If the band state is lower in energy, then polaron formation is unlikely and the electron will remain delocalised. A noteworthy quantity that can be extracted from such a framework is the self-trapping energy which is defined by Equation 4.1.

$$\Delta E_{ST} = E_{Polaron} - E_{Band} \tag{4.1}$$

Where ΔE_{ST} is the self-trapping energy, $E_{Polaron}$ is the total energy of the localised polaron state, and E_{Band} is the total energy of the delocalised band state.

4.1.4 Modelling Small Polarons with DFT

Whilst it is possible to model small polaron formation with DFT, there is not yet a method for simulating the properties of small polarons that is as reliable or generalised as for bulk electronic

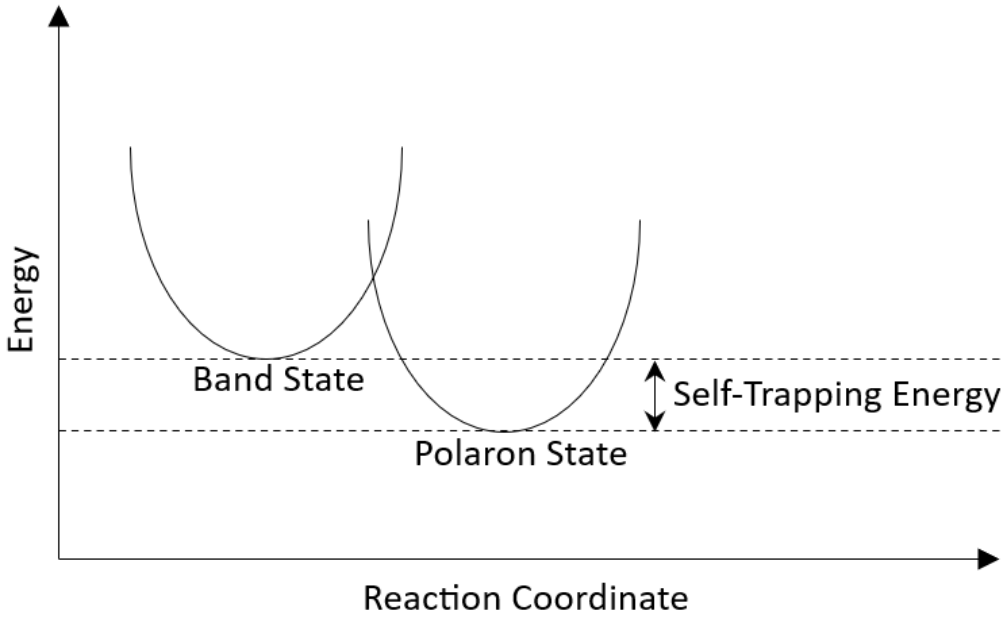


Figure 4.3: Illustration of small polaron formation based on Marcus theory. In this case, the polaron state is lower in energy than the band state, implying that self-trapping is likely. However, there may still be a thermal activation barrier for the self-trapping mechanism, which is represented by the crossing point of each parabola, as for Figure 4.2. In some cases, there may be no such barrier to self-trapping. These systems are described as being *"barrierless"*[30].

states. A similar statement could be made for all point defect calculations, but polaron calculations have specific challenges associated with their implementation. Standard DFT functionals (i.e. LDA or GGA) are insufficient for generating localised states within a bulk solid due to their insufficient treatment of electron correlation[305]. This necessitates the use of costly Hybrid or DFT + U functionals at both the benchmarking and production stages of a project[306].

DFT+U vs Hybrid Functionals

The DFT + U method uses the Hubbard model[307] to selectively force the electron occupancy of d and f orbitals to induce localisation[308]. The selection of an appropriate on-site Coulomb interaction term (U) is the benchmarking bottleneck for the DFT+U method because a suitable fitting protocol needs to be established. For instance, it is common practice to find a value of U that accurately represents the band gap and then use this value to calculate the electronic structure of the polaron[295].

For the Hybrid method, the fraction of the exact HF exchange that is incorporated into the Hamiltonian operator may be adjusted for the localisation of the electronic states, including polarons. However, similarly to DFT+U, there is a degree of ambiguity regarding the selection of a particular amount of HF-exchange[295,306]. Therefore, an appropriate fitting protocol is often required. Consequently, a high-precision DFT+U (or Hybrid) functional investigation may have to consider multiple fitting

protocols for each separate material, requiring many more calculations than simply using a given hybrid functional. It is respectfully submitted that this is a common reason why polaron studies are often restricted to one topical material or a small set of closely related materials. There is simply too much variation in the optimal calculation parameters between different materials and the benchmarking process is incredibly expensive.

Therefore, from a materials screening perspective, it may be more expedient and instructive to experiment first with a standard amount of HF exchange (e.g. 25% for the HSE06 functional), apply it to a wide range of compounds within a high-efficiency workflow, and then consider systematic adjustments to a further set of calculations based on any observed trends. Using this approach, such adjustments could be refined and implemented more reliably using a machine learning algorithm. This is because more variables would be controlled and consistently recorded across the database which would make pattern recognition based on the input parameters more reliable. In this regard, the first challenge would be how to establish such a *"high-efficiency workflow"*.

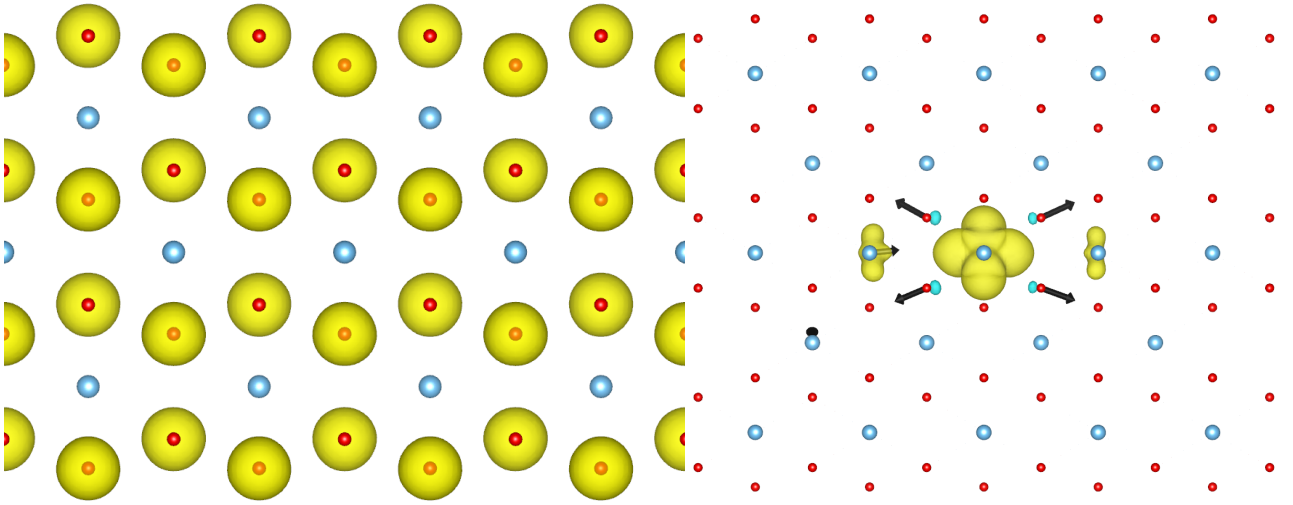
Structural and Electronic Relaxation of Polaron States

Even when hybrid functionals are used and additional charges are introduced into the bulk unit cell, if there is no symmetry-breaking defect[a] present in the unit cell to influence the localisation of added charges, the added charge often remains delocalised. To avoid this, several methods have been suggested to induce localisation, all of which involve breaking the symmetry of the bulk unit cell before a structural relaxation:

- Occupation matrix control[309]: Altering the occupation matrix of the cell such that the excess electron is localised to a specific site at the start of the calculation.
- Substitution[310]: A bulk atom is replaced with a foreign atom that accounts for the added charge. The structural relaxation is then performed, resulting in the localisation of the charge around the foreign atom. The foreign atom is then replaced with the original species and the relaxation is performed once more.
- Bond distortion[306]: The bonds around the desired site for localisation are stretched/compressed according to the expected shape of the polaron. The magnetic moment of the desired site (on which the polaron is expected to localise) may be modified, to better encourage localisation.

[a]Vacancy, Substitution, Interstitial, etc.

The results of such polaron calculations are often analysed by inspecting the *isosurface* of the charge density and by calculating the self-trapping energy from the respective relaxed polaron and band states. An example of these isosurface differences is given in Figure 4.4.



(a) Band State **(b)** Polaron State

Figure 4.4: Example of the electron band and polaron states for titanium dioxide using the bond distortion method. For the band state, it can be seen that the charge is shared evenly across the unit cell, indicating delocalisation. For the polaron state, it can be seen that the charge is highly localised on a d(xy) orbital of a (blue) titanium atom, with some participation of the d(z^2) orbitals of the nearest neighbour titanium atoms. The black arrows indicate the extent to which the attached atoms have moved during relaxation. As such, it can be seen that the nearest neighbour (red) oxygen atoms move outwards from the central titanium atom, likely due to electrostatic repulsion as a result of the excess charge density.

4.2 Motivation and Materials of Interest

Despite the abundance of methods to induce polaron formation, comparability between each of these methods, and within the same method, is difficult to establish. When comparing between methods, the issue is attributing the final polaron state to either the method itself or to a physical phenomenon. When comparing within the same method, the issue is attributing the polaron state to merely the specific amount of HF exchange or the value of the U parameter. For instance, Pham et al[306] observed that the localisation could be induced by altering the amount of HF-exchange of a hybrid functional calculation in much the same way as the U parameter is optimised in a DFT+U investigation.

Consequently, many studies focus on a single polaron property calculation for a single system, and using a single polaron generation methodology with specific benchmarks and standards. This approach is useful for obtaining good agreement with experimental results, and thereby obtaining a good understanding of a specific system, but can seldom be generalised across a large database of materials for

materials screening purposes. This is because the calculation parameters are carefully tailored to the system of interest and both the benchmarking and final calculations incur significant computational costs. Indeed, past work by Sezen et al[311] has stressed the need for a wider database of polaronic states.

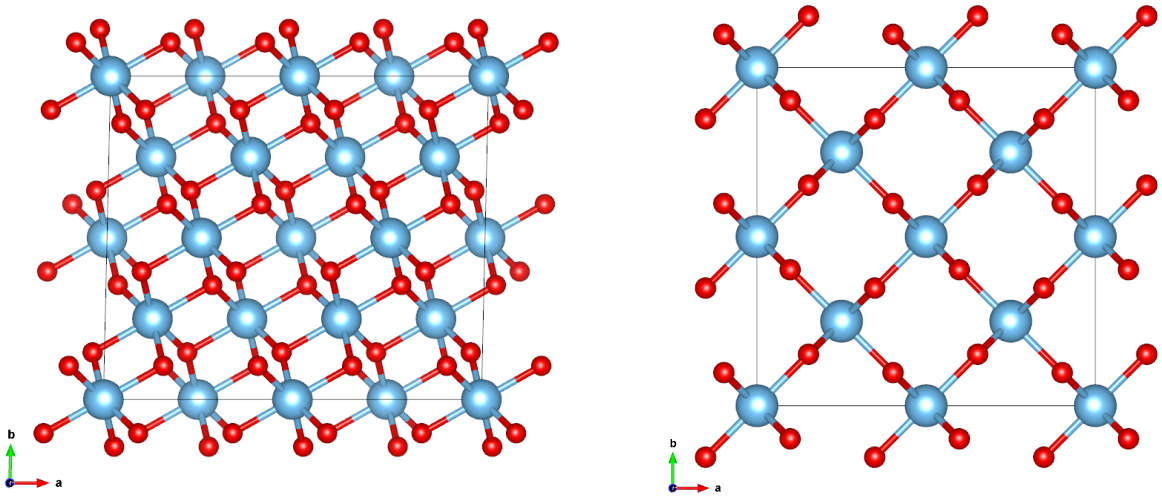
Therefore, to accelerate materials discovery and screening, it would be beneficial to develop a more flexible and predictable workflow for polaron formation. Specifically, to have an understanding of how a particular method or level of theory affects multiple compounds. Such an understanding may be achieved in combination with machine learning, provided that a large enough database of compounds could be generated efficiently. As mentioned, a key challenge is the large computational cost incurred by the polaron calculations. In particular, the memory requirements.

The remainder of this chapter attempts to address this problem by investigating the bond distortion method for a series of transition metal compounds and by attempting to make the polaron formation calculations as low-cost as possible. In this regard, a large cost-saving exercise may involve coarsening the **k**-point grid to a single **k**-point at the Γ-point, which corresponds to (0,0,0) in reciprocal space. Similar protocols have been used to cheaply exclude unlikely candidates in point defect screening tools wherein a large number of defect states are possible[312,313,314]. Such a procedure would not produce precise results and should not be considered as *quantitatively* representative of the real system or of experimental data. However, if sufficiently accurate, it could be used to efficiently and *qualitatively* determine the conditions necessary to induce polaron formation for further study. For example, the optimal degree of bond distortion may be determined using Γ-point sampling. Additionally, the optimised output of such tests may be used as the input charge density and wavefunction for further calculations at optimal Monkhorst pack grid densities.

Such calculations would be expected to reach completion faster than their fully converged counterparts because most of the structural relaxation steps would have been accounted for during the Γ-point calculation. This was suggested in a point defect screening study by Mosquera-Lois et al[314] who recommended further investigation of polarons. Moreover, some long-range behaviour of the system would still be accounted for, merely due to the use of a supercell. It would therefore be desirable to determine whether or not such coarse **k**-point sampling is sufficiently accurate and whether or not it can be used as the basis for more refined calculations. The bond distortion method was chosen specifically for its ease of implementation, requiring simple manipulations of the crystal structure and a minimal number of calculations. In a similar sense, the hybrid HSE06 functional was considered because of its ease of implementation, as opposed to the benchmarking of an optimal U parameter.

The materials investigated in this chapter traverse the transition metal oxide series and have seen extensive use in photoelectrochemical cells. This provides a diverse collection of crystal, electronic and magnetic structures across which to test the above hypothesis. In no particular order, the materials of interest for this chapter are as follows:

Titanium dioxide (TiO_2) is an insoluble[315] semiconductor for which there are two notable phases, *rutile* ($P4_2/mnm$) and *anatase* ($I4_1/amd$). In this context, the word *"phase"* refers to an alternate crystal structure with an equivalent composition. For example, the rutile and anatase structures have the same chemical formula (TiO_2) but different crystal structures, as illustrated by Figure 4.5. The rutile phase has a direct band gap of 3.0 eV[316]. The anatase phase has an indirect band gap of 3.2 eV[316].



(a) Anatase **(b)** Rutile

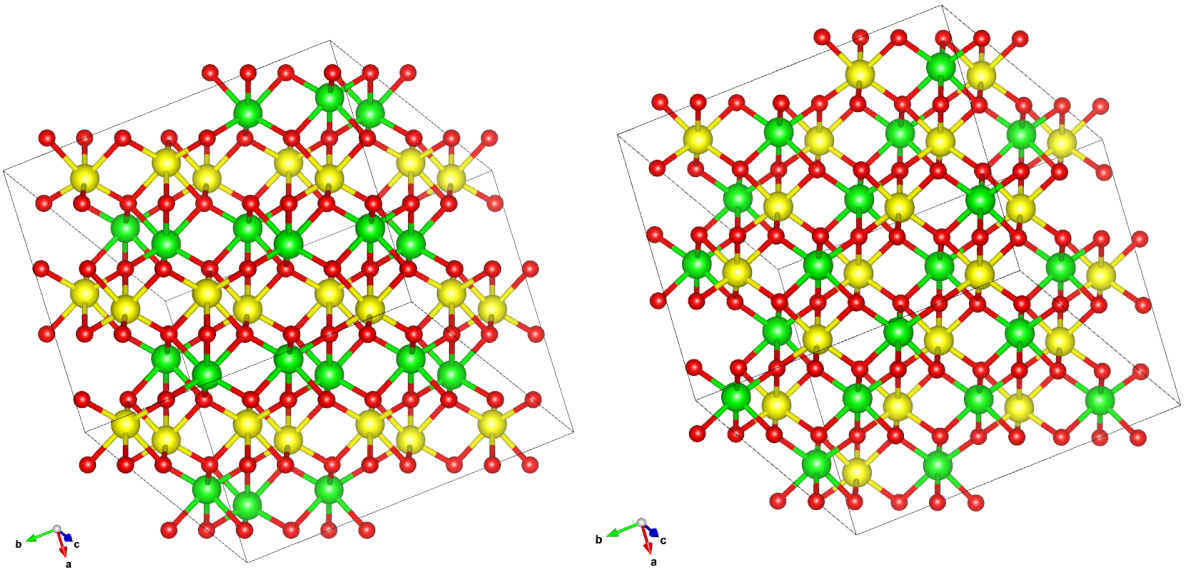
Figure 4.5: Illustration of the anatase and rutile TiO_2 crystal phases. The titanium atoms are given in blue. These images were produced using the Vesta[21,22,23] software.

The rutile phase of TiO_2 was used as the first photoelectrochemical cell photoanode and has since seen extensive use due to its low cost, chemical stability, corrosion resistance, non-toxicity, abundance, lack of expense and reasonable band alignment capabilities[317,29]. Furthermore, its extensive use throughout the industry has made it a common starting material with which to conduct experimental and theoretical investigations because the polaron morphologies have been well studied[318,319,306,320].

The anatase phase has also attracted attention in nanoparticle-type photoelectrochemical cells due to the stability of the crystal structure at this scale[321]. It has also been considered as a photoelectrochemical cell electrode for its acidic corrosion resistance[322,323]. Moreover, a mixed phase of both

anatase and rutile is known to be more photocatalytically active than either pure phase[316]. However, in both cases, the wide band gap inhibits this compound's ability to absorb light in the visible range. Its absorption capabilities are limited to to the UV range.[324,29].

Iron oxide (Fe_2O_3) is a semiconductor with a indirect[325] band gap of 2.2 eV[326] and corundum ($R\bar{3}c$) crystal structure[327]. It further has an antiferromagnetic magnetic structure within which unpaired spins of the iron atoms are arranged in alternating planes throughout the crystal, as illustrated in Figure 4.6a. It has been considered as a photoelectrochemical cell photoelectrode due to its low cost, light absorption capabilities, chemical stability, low toxicity, high abundance[328,325], and water insolubility[329]. However, it suffers from low charge carrier lifetimes which inhibits charge transport during operation[325]. A study by Ahart et al[326] has considered that this may be caused, at least in part, by the slow transport of hole polarons through the haematite lattice.



(a) Fe_2O_3 **(b)** Cr_2O_3

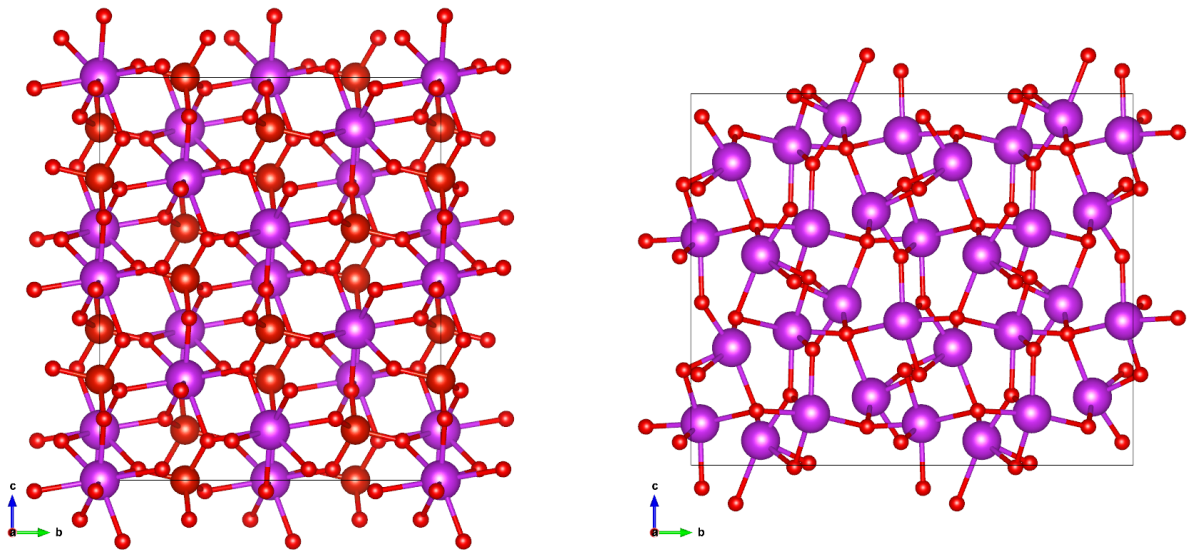
Figure 4.6: Illustration of the Fe_2O_3 and Cr_2O_3 crystal structures. The green and yellow atoms represent the alternate spin states of the iron and chromium atoms that determine the magnetic structures. As such, it can be seen that the Fe_2O_3 and Cr_2O_3 magnetic structures differ in that the spin states are distributed across different Miller planes throughout each crystal. This image was produced using the Vesta[21,22,23] software.

According to the calculations performed by Ahart et al[326], the Fe_2O_3 electron polaron is centred across two nearest neighbour iron atoms, whilst the hole polaron in centred on a single iron atom. However, similar calculations performed by Pastor et al[330] indicate that the Fe_2O_3 electron polaron is centred around a single iron atom. Both studies were similar in their approach to polaron formation, utilising a form of the bond distortion method to break the crystal symmetry. However, they each differed in their initial guess of the iron atom magnetic moments, the type of initial distortion applied to the system,

and the constraints applied to the DFT Hamiltonian operator. This difference further emphasises the desirability of a systematic investigation using a consistent level of theory for multiple materials.

Chromium oxide (Cr_2O_3) is a semiconductor with an indirect band gap[331] of 3 eV[332]. Cr_2O_3 has the same corundum structure as Fe_2O_3 but not the same antiferromagnetic magnetic structure, with the alternating planes of unpaired spins being distributed across a different axis within the crystal, as illustrated in Figure 4.6b. Hence, it is an interesting material to compare with Fe_2O_3 due to the differences in magnetic ordering.

Furthermore, in addition to its insolubility in water[333], the conduction band minimum of this material is known to be well aligned for water splitting applications[334]. As such, Cr_2O_3 has been considered as a candidate photocathode[335,334]. It is thought to suffer from poor photoabsorption properties due to its wide band gap[335] and the rapid recombination of charge carriers via defect states, of which polarons could be one such state. It has further been noted that such photoabsorption properties may be enhanced through nickel doping[336]. Past work by Dehn et al[337] has indicated that the electron polaron localises on a central chromium atom as part of a muon-polaron complex. However, literature concerning the electron and hole polarons is sparse. Hence, it would be useful to gather new polaron data on this material.



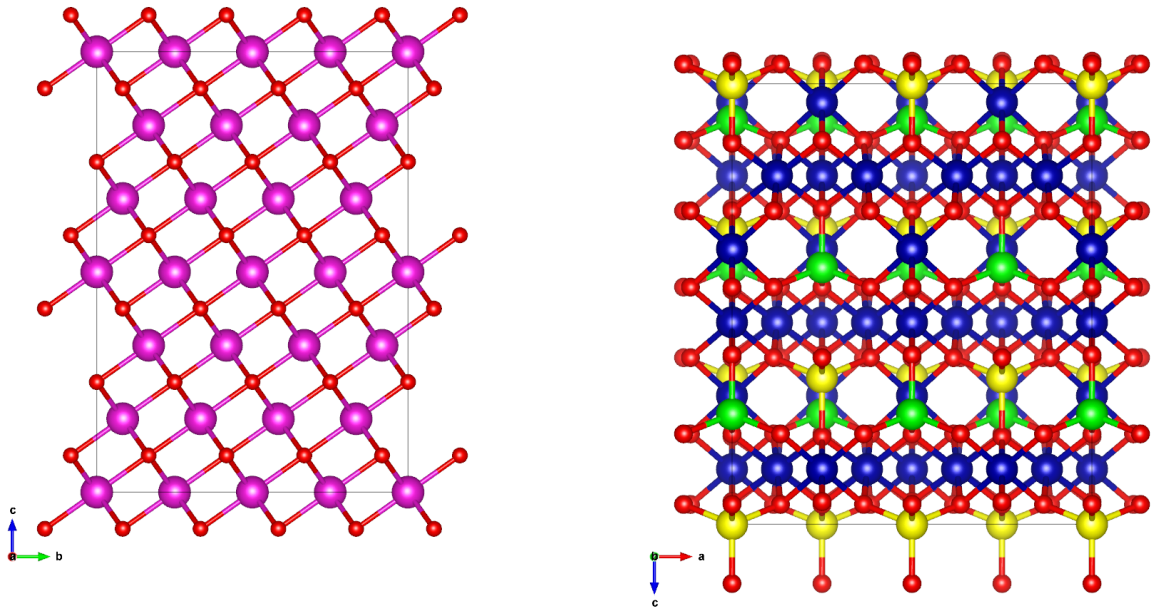
(a) $BiVO_4$ **(b)** Bi_2O_3

Figure 4.7: Illustration of the $BiVO_4$ and Bi_2O_3 crystal structures. The bismuth atoms are given in purple. The vanadium atoms are the larger red atoms. This image was produced using the Vesta[21,22,23] software.

Bismuth vanadate ($BiVO_4$) is an semiconductor whose monoclinic (C2/c) phase (see Figure 4.7a) has an indirect band gap[338] of 2.4-2.5 eV[339]. This particular phase has received a great deal of

attention, due to its well-sized band gap for solar absorption[335], low production costs, low toxicity, photostability, corrosion resistance,[339], and water insolubility[340]. However, its catalytic activity is limited by rapid recombination mechanisms[339]. According to calculations performed by Wiktor et al[341], the recombination of electron and hole polarons is considered to be a significant contributor to this effect.

Bismuth oxide (Bi_2O_3) is a water-insoluble[342] semiconductor whose tetragonal phase has received much attention due to its well-sized band gap for photoabsorption and charge carrier mobilities[343,344]. There are 3 other known phases of Bi_2O_3, of which, the monoclinic ($P2_1/c$ - band gap 2.5 eV[345]) phase is the most stable. Despite this apparent advantage, the monoclinic phase has received far less attention than the tetragonal phase, possibly due to the lack of an ideally sized band gap. In either case, literature regarding the polaron morphology of small polarons in this compound is sparse. However, charge transport mechanisms in nanostructured mixed compounds containing Bi_2O_3 have been attributed to small polaron hopping[346]. It is also established that the electrical conductivity can vary greatly between each of the four phases[347]. Hence it would be useful to gather such data for this material. The monoclinic phase (see Figure 4.7b) is considered here as a starting point.



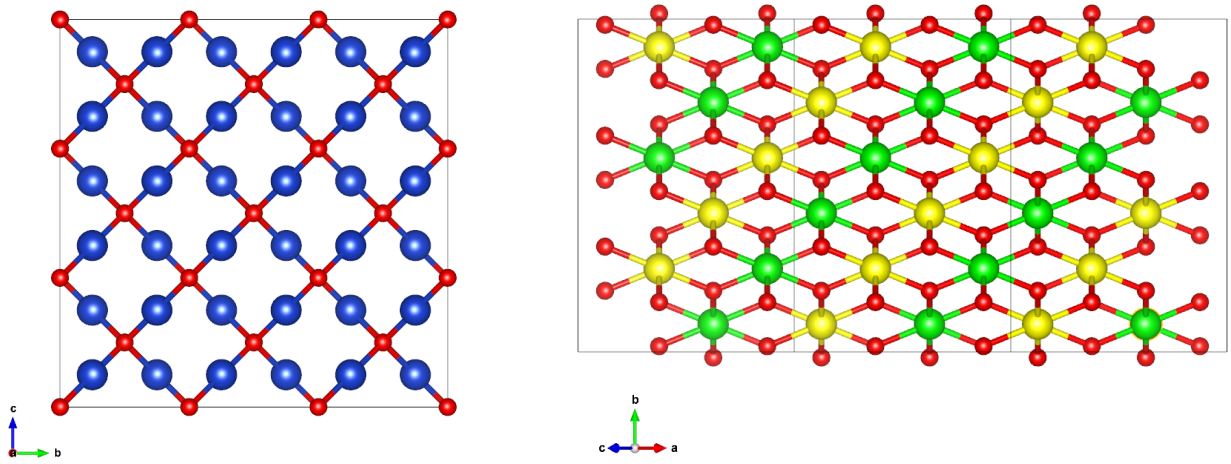
(a) CdO **(b)** Co_3O_4

Figure 4.8: Illustration of the CdO and Co_3O_4 crystal structures. The cadmium atoms are given in purple. The green and yellow atoms represent the alternate spin states of the cobalt atoms that define the magnetic structure. Spin inactive cobalt atoms are given in blue. This image was produced using the Vesta[21,22,23] software.

Cadmium oxide (CdO) is an insoluble[348] semiconductor with an indirect band gap of 0.84 eV eV[349]. It has the rocksalt ($Fm\bar{3}m$) structure (see Figure 4.8a) and has received attention for its well-sized band

gap for photoabsorption[350]. Literature concerning the polaron morphologies in CdO is sparse, though work by Karmakar et al[351] has indicated the formation of small polarons in nanocomposite structures at high temperatures. Hence, it will be useful to see if CdO presents any small polaron formation.

Cobalt oxide (Co_3O_4) is an insoluble[352] semiconductor with a band gap of 0.8 eV to 2 eV[353]. It has a spinel ($Fd\bar{3}m$) crystal structure (see Figure 4.8b) and has received attention due to its well-sized band gap for photoabsorption. However, one of its greatest limitations has been the efficient separation of photogenerated charges[354]. The reason for the large range of band gaps is due to the uncertainty of the cobalt oxide electronic structure. Past work by Smart et al[31] has calculated a band gap of 1.6 eV, attributing lower reported values to hole polaron states deep within the band gap. However, work by Singh et al[353] considered the band gap to be closer to 0.8 eV, but acknowledged that the resultant band gap and electronic structure were highly sensitive to the amount of HF exchange and value of U. Hence, in addition to testing the robustness of the presently suggested polaron screening procedure, further investigation into the polaron states of this particular material may be helpful. In this regard, Smart et al[31] found several hole polaron states within the band gap of cobalt oxide. Similar observations were made for the electron polaron, which can be found in the supplementary material of Smart's work[31].



(a) Cu_2O **(b)** CuO

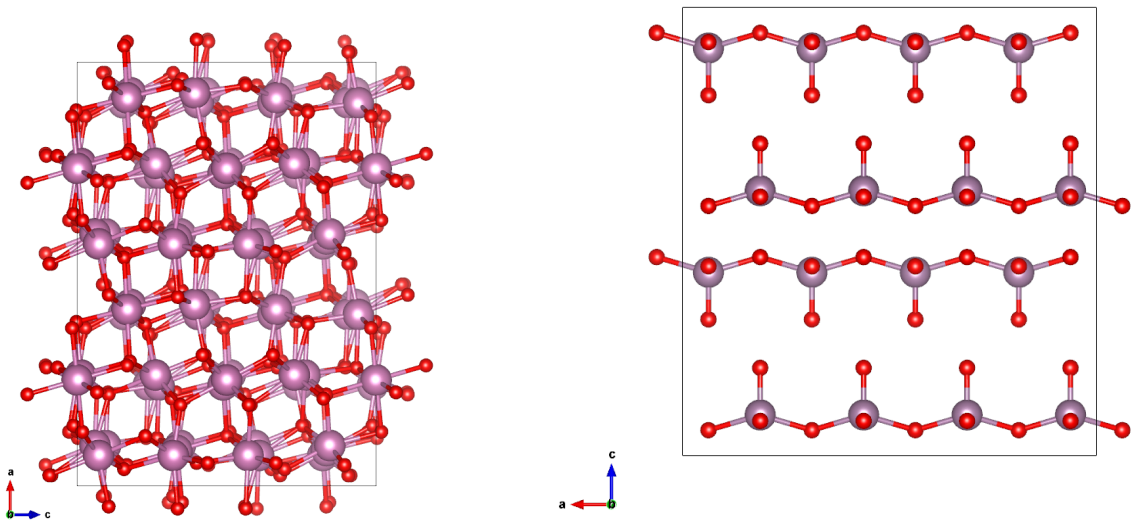
Figure 4.9: Illustration of the Cu_2O and CuO crystal structures. The spin inactive copper atoms are given in blue. The green and yellow atoms represent the spin states of the copper atoms that define the magnetic structure. This image was produced using the Vesta[21,22,23] software.

Copper (I) oxide (Cu_2O) is an insoluble[355] semiconductor with a direct band gap[356] of 2.1 eV[357]. It has a cubic ($Pn\bar{3}m$) crystal structure (see Figure 4.9a) and has received attention as a light-absorbing

photocathode, due to its well-sized band gap[358]. Literature concerning the polaron morphologies is sparse. However, it has been established that GGA +U has proven insufficient to describe certain defect states in this material[359]. A hybrid functional may provide useful data in this regard.

Copper (II) oxide (CuO) is a water-insoluble[360] semiconductor with a direct band gap[361] of 1.4 eV[357]. It has a monoclinic (C2/c) crystal structure and an antiferromagnetic magnetic structure wherein the unpaired spins are arranged in alternating layers[362] (see Figure 4.9b). It has received attention as a photocathode due to its well-sized and well-aligned band gap for water splitting, as well as its abundance[363] and high theoretical photocurrent density[364]. However, the material is known to suffer from corrosion, photo-induced degradation, and a low charge carrier mobility[361,364]. Literature concerning the polaron morphologies is sparse. However, the poor charge carrier conductivity has been attributed to the formation of polarons that are coupled to the magnetic structure[365].

Indium oxide (In_2O_3) is a water-insoluble[366] semiconductor with a direct band gap of 3 eV[367]. It has a cubic ($Ia\bar{3}$) crystal structure (see Figure 4.10a) and has received attention as a photoanode for its photoabsorption capabilities, electrical conductivity, and corrosion resistance[368]. Literature concerning the polaron morphologies is sparse, but their existence has been considered in past defect models for this system[369].



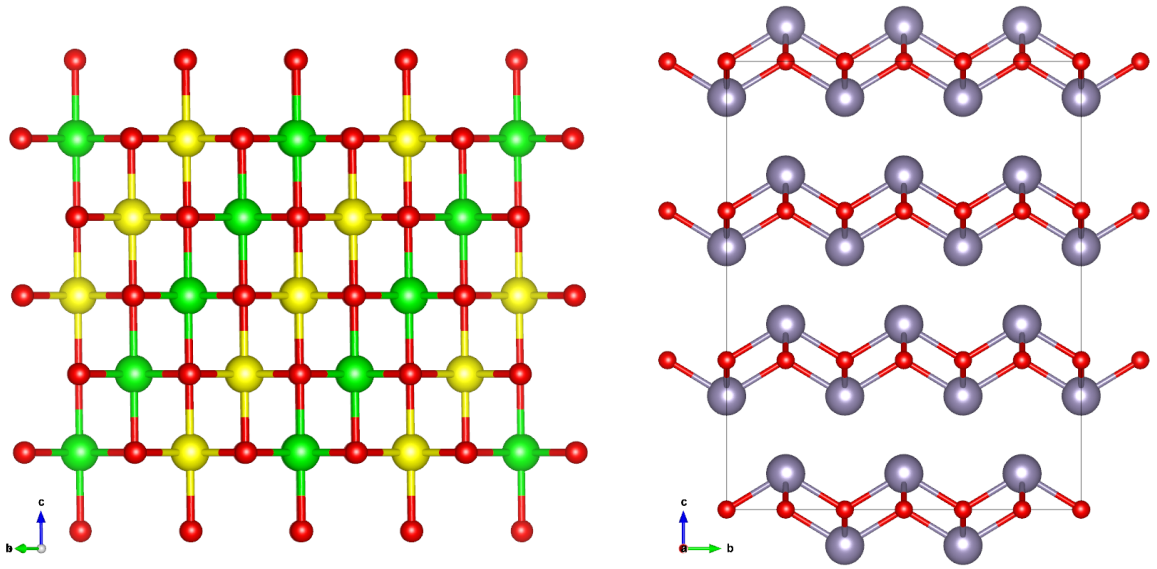
(a) In_2O_3 **(b)** MoO_3

Figure 4.10: Illustration of the In_2O_3 and MoO_3 crystal structures. The indium atoms are given in pink. The molybdenum atoms are given in grey. This image was produced using the Vesta[21,22,23] software.

Molybdenum Oxide (MoO_3) is a semiconductor with an indirect band gap of 3 eV[370]. It has an

orthorhombic (Pnma) crystal structure (see Figure 4.10b) and has received attention due to its ability to enhance the properties of water splitting electrodes when used in combination with other transition metal oxides[371,372]. It has received little attention as a stand-alone electrode, possibly due to its slight solubility in water[373]. Electron and hole polaron morphologies have been introduced in this material using the substitution method[374,375,376], wherein, the excess electron was found to localise on a central molybdenum *d* orbital.

Nickel oxide (NiO) is a water-insoluble[377] insulator with an indirect band gap[378] of 3.6 - 4 eV[379,380,381] However, intrinsic nickel vacancies result in an effective p-type (hole conducting) conductivity[381]. It has a Rocksalt crystal structure with an antiferromagnetic magnetic structure that is arranged in alternating layers of unpaired spins[382] (see Figure 4.11a). It has received attention in RuO_2-based photoelectrochemical cell electrodes as a promoter for both the HER and OER[383]. A nickel-nickel oxide heterojunction has also been found to have enhanced catalytic activity for the hydrogen evolution reaction[384]. It is believed that NiO suffers from low hole polaron mobility. However, past work by Yu et al[385], has had difficulty in simulating hole polaron formation with DFT + U. Yu et al did, however, succeed in generating an electron polaron that was centred on a nickel atom. Hence, it would be interesting to see whether or not the presently suggested method can reproduce these results.



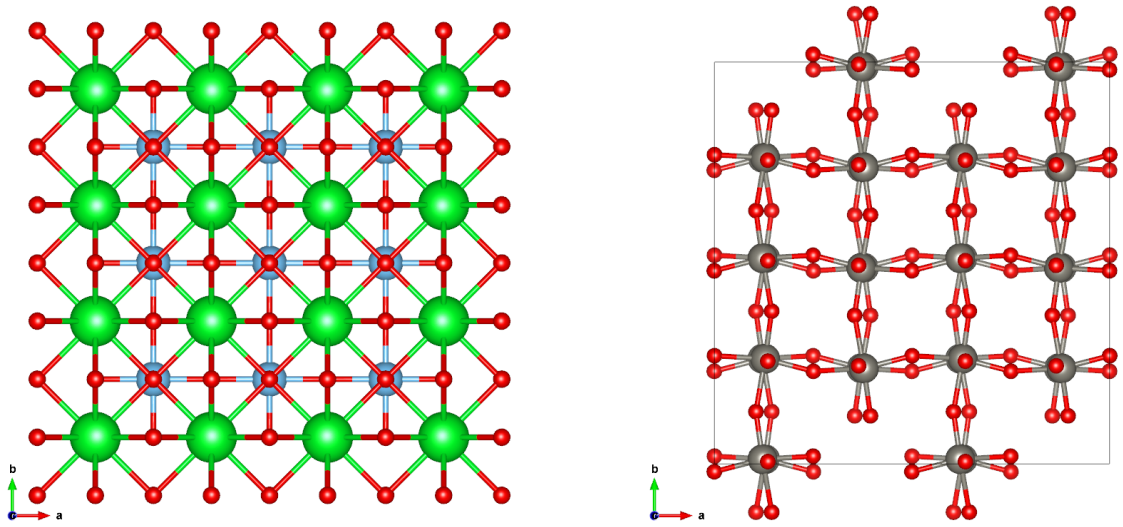
(a) NiO **(b)** SnO

Figure 4.11: Illustration of the NiO and SnO crystal structures. The green and yellow atoms represent the alternate spin states of the nickel atoms that define the magnetic structure. The tin atoms are given in grey. This image was produced using the Vesta[21,22,23] software.

Tin oxide (SnO) is an insoluble[386] semiconductor with a indirect band gap of 0.7 eV[387]. It has a tetragonal (P4/nmm) crystal structure (see Figure 4.11b) and has received attention for its well-

positioned band gap for visible-light photoabsorption[388] and high mobility[389]. However, it oxidises readily and irreversibly to SnO_2 in the presence of oxygen[389]. As such, further measures to limit degradation would be required. Literature concerning the polaron morphologies is sparse, most likely because the material is believed to exhibit band-like transport (i.e. not via polaron hopping)[389].

Strontium titanate ($SrTiO_3$) is a water-insoluble[390] semiconductor with an indirect band gap[391] of 3.25 ev[392]. It has the Perovskite ($Pm\bar{3}m$) crystal structure (see Figure 4.12a) and has received attention as a photocatalyst for its low cost, chemical stability, and optimal band alignment for water splitting. However, it is usually combined with other materials to compensate for its wide band gap and rapid recombination kinetics[393]. Work by Janotti et al[394] has indicated that the hole polaron localises on a central oxygen atom, with some participation from the nearest neighbour oxygen atoms, whilst the electron polaron localises on a central titanium atom.



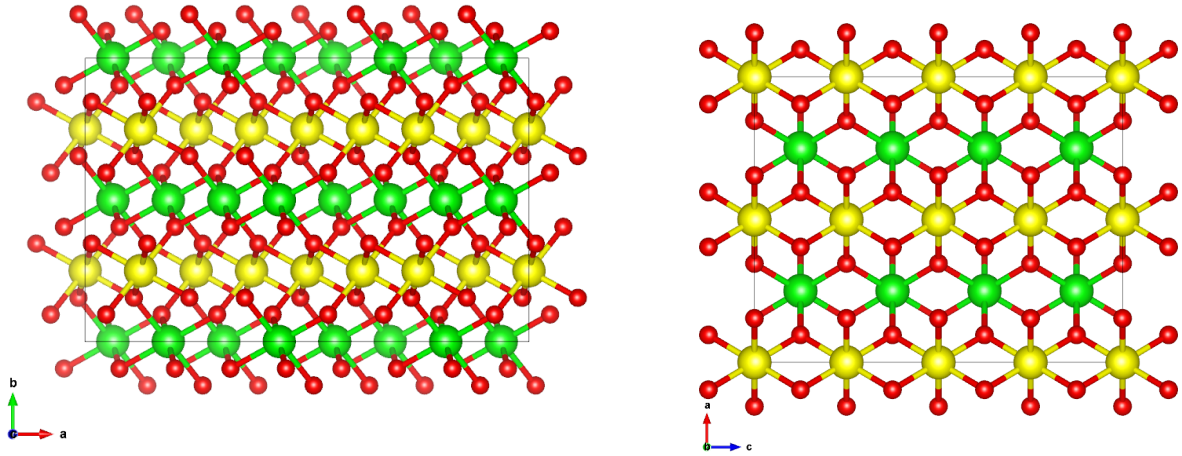
(a) $SrTiO_3$ **(b)** WO_3

Figure 4.12: Illustration of the $SrTiO_3$ crystal structure and WO_3 orthorhombic crystal structure. The strontium atoms are given in green and the titanium atoms are given in blue. The tungsten atoms are given in grey. This image was produced using the Vesta[21,22,23] software.

Tungsten oxide (WO_3) is a water-insoluble[395] semiconductor material with an indirect band gap of 2.6 - 2.9 eV[396]. It has received attention as a photoanode for its well-positioned (visible range) band gap and long hole diffusion length, which reduces the bulk recombination[397]. There are 10 known phases of WO_3, depending on the temperature, ranging from 0 K to 1170 K[398]. Each of these phases has the same octahedral coordination around the tungsten atom as shown below. However, they vary in their extent and arrangement of the octahedral tilting about each vertex.

Of these phases, the room temperature monoclinic ($P2_1/c$) phase has been extensively studied by Tao

and Liu[399]. The electron polaron localising on a central tungsten atom with significant contribution from the next two neighbouring shells of tungsten atoms. The hole polaron was found to localise on a central oxygen atom. However, work by Zhuiykov and Kats[396] has shown that the orthorhombic (Pbcn) phase (estimated indirect band gap 2.93[398]) can be obtained as ultra-thin nanoflakes via an appropriate sintering process[b]. Such nanoflakes and nanostructuring processes are expected to enhance the catalytic activity of WO_3 for the HER[396]. Given the absence of polaron data for such a system, and given the high structural similarity between the various WO_3 phases, it would be interesting to compare the calculated polaron morphologies of the orthorhombic phase (see Figure 4.12b) with the more established polaron morphologies of the monoclinic phase.



(a) VO_2 monoclinic phase **(b)** VO_2 rutile phase

Figure 4.13: Illustration of the monoclinic and rutile VO_2 crystal phases. The green and yellow atoms represent the alternate spin states of the vanadium atoms that define the magnetic structures. These images were produced using the Vesta[21,22,23] software.

Vanadium oxide (VO_2) is a water-insoluble compound[400] that has received attention due to its ability to transition from a semiconducting monoclinic ($P2_1/c$) phase (see Figure 4.13a) with an indirect band gap[401] of 0.6 eV to a metallic rutile phase at 340K[402] (see Figure 4.13b). This feature is expected to improve the catalytic activity of a photoelectrochemical cell as it is expected to result in an overall increased electrical conductivity[403,404]. The results of Liu et al[405] indicate the formation of an electron polaron centred around a single vanadium atom in the semiconducting phase, and that the formation of such polarons is crucial in facilitating the change in electrical properties between each phase. Literature

[b]Sintering involves compressing and heating a more diffuse material into a solid mass. The microstructure, and therefore phase, of the material is dependent on the sintering conditions.

regarding the hole polaron or polarons within the metallic phase is otherwise sparse. Hence it would be desirable to gain a deeper understanding of the polarons within this material.

Zinc oxide (ZnO) is a semiconductor with a direct band gap of 3.37 eV. It has a Wurtzite ($P6_3mc$) crystal structure (see Figure 4.14) and has received attention due to its low cost and high abundance. However, its wide band gap inhibits visible light absorption and recombination losses appear to be substantial[406]. A further issue may be its very slight solubility in water[407]. There is evidence for hole polaron formation in ZnO[311] though the literature regarding such polaron morphologies is sparse.

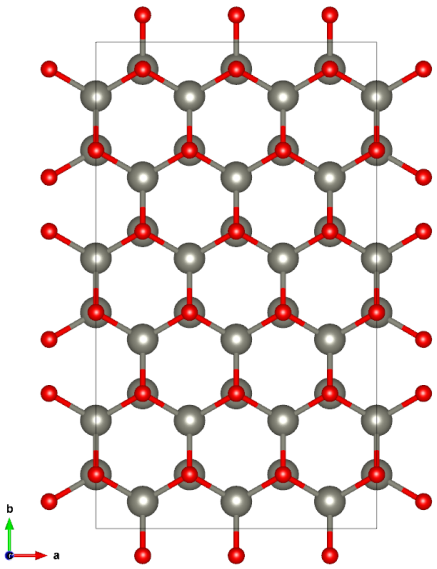


Figure 4.14: Illustration of the ZnO crystal structure. The zinc atoms are given in grey. This image was produced using the Vesta[21,22,23] software.

4.3 Methods

For each compound, the DFT calculations used a 550 eV plane wave cutoff energy and their Monkhorst Pack **k**-point grids were converged to 0.01 eV. The PBEsol functional was used to optimise the primitive unit cells at different volumes, before fitting the total energies to the Murnaghan Equation of State[280]. For each primitive cell, the *Atomic Simulation Environment*[408] (ASE) supercell optimiser[409] was used to create supercells that were as close to cubic as possible whilst keeping the total number of atoms in the system to a minimum. The supercells were selected such that there were at least 2 nearest neighbour shells of atoms around a central (target) atom.

To determine the polaron state, the magnetic moments in the system were adjusted[410] such that a *target atom(s)*, close to the centre of the unit cell, was reduced/oxidised relative to the bulk magnetic structure. The bond distortions and rattling of the atomic positions were created using a combination of the Deskins code[411,306] and ASE rattle function[412]. For atom-centred distortions, the bond lengths between the target atom and its nearest neighbours were altered by -20% to + 20% in 10% increments for separate unit cells. The same degree of rattling (within a standard deviation of 0.001) was used in all cases and an electron was added or removed by adjusting the electron count for each system[413]. For each incrementally distorted unit cell, the internal parameters were relaxed with a single **k**-point at the Γ-point. Out of these incrementally distorted unit cells, the system with the lowest total energy was inspected. If localisation was observed, the output charge density, wavefunction and structure were used as the inputs for further relaxations, at successively more dense **k**-point grids until the calculation was converged to within 0.01 eV of the total energy (these are provided in Section A.7). Non-localising systems and systems that struggled to achieve convergence were discarded from the pool of candidates.

To obtain the band state, an electron was either removed or added to the unit cell, as for the polaron state, but without distorting the initial structure. The electronic structure was preconverged using the PBE functional and the same **k**-point grid density as the converged polaron state calculations. A further electronic relaxation was then performed using the HSE06 functional and the Conjugate Gradient Algorithm[414]. Gaussian Smearing[415] of the partial occupancies with a 0.01 eV smearing width[416] was used for all bands. For enhanced performance, the coarsest possible Hartree-Fock FFT grid (determined automatically by VASP)[417], the rotation matrix between the occupied and unoccupied states was optimised during the electronic minimisation[418], the projection operators were evaluated in real-space[419], and parallelisation across bands and **k**-points was used throughout[420]. Using the band and polaron state total energies, the self-trapping energy was calculated via Equation 4.1.

4.4 Results and Discussion

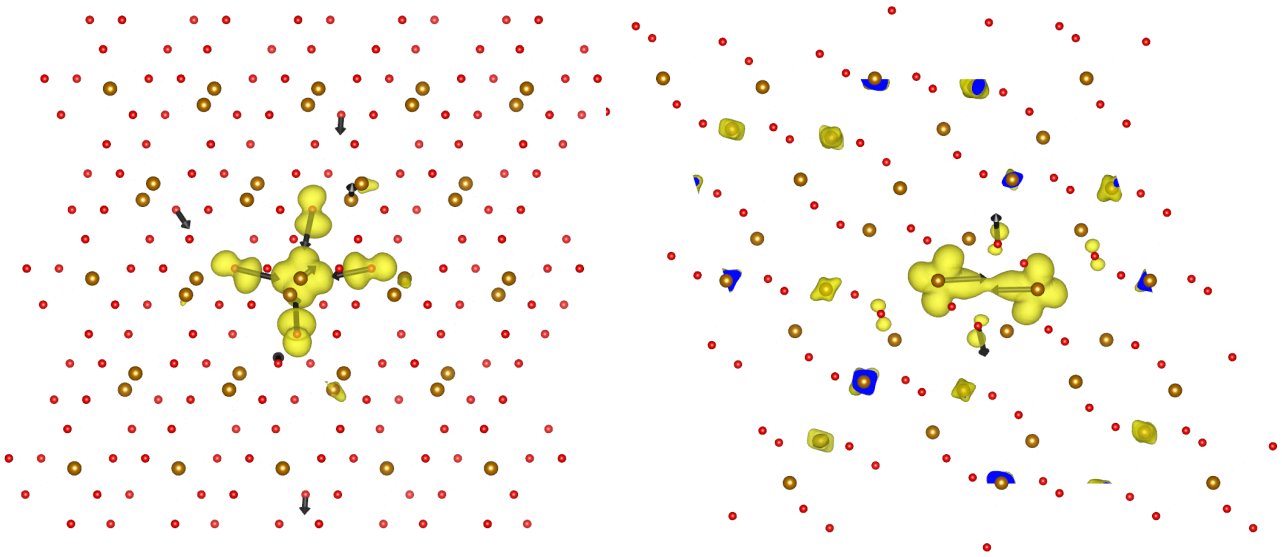
4.4.1 Converged Polaron Morphologies

In the following section, the polaron morphologies for each of the materials of interest are discussed. All such figures were produced using the VESTA program[21,22,23]. The *vesta_vectors* code[421] was used to re-centre the unit cell around the polaron as well as to add the black arrows to each figure. The black arrows indicate the relative magnitude of the displacements for each atom from their positions in the bulk unit cell. Displacements below 0.05 Angstroms were not included. In most cases, the magnitude of the displacements that were caused by polaron formation was very small. Consequently, smaller arrows can be seen far from the polaron centre, which are expected to be artefacts of the ASE rattle function[412].

To visualise the orbitals involved in polaron formation, the isosurface of the charge density is superimposed on each figure. Where possible, the isosurfaces shown correspond to the magnetisation density (spin-separated charge density) of the unit cell. However, for the open-shell systems, the magnetic structure often obscured the localisation of the polaron. In these cases, the band decomposed charge density of the polaron band, or a 2D slice of the spin density, was used. The isosurface level was set to 0.05 eV $Å^{-1}$in all cases.

Due to large variability in the computational overhead between different materials, not all of the materials could be fully converged for their respective Monkhorst-Pack **k**-point grids. These are discussed in their appropriate sections. Compounds that failed to show any localisation are also discussed here but were excluded from further study. The materials are listed in order of their contribution to the benchmarking process.

Fe_2O_3 The polaron isosurfaces for Fe_2O_3 are given in Figure 4.15. The isosurfaces appear to be in line with the results of Ahart et al[326], but not those of Pastor et al,[330] who observed localisation on a single iron atom. Presumably, these differences are due to the differences in the calculation parameters between the two studies. In particular, the type of initial distortion applied to the system before the structural relaxation. This is because the Fe_2O_3 relaxation was found to be very sensitive to the type of initial distortion and whether or not the ASE rattle function was used. Indeed, depending on the initial conditions, a Fe-centred state could also be converged upon using the presently proposed method.



(a) Hole **(b)** Electron

Figure 4.15: Fe_2O_3 polaron isosurfaces. It can be seen that the hole polaron is centred around an iron atom, with a substantial contribution from the iron d($x^2 - y^2$) orbital and nearest neighbour oxygen p orbitals in the d(xy) plane. There is little to no contribution from the oxygen p along the d(z^2) axis. The contributing oxygen atoms are shown to move inwards towards the metal centre, presumably to compensate for the positive charge. The electron polaron takes the form of a Fe-Fe dimer, where the iron d(xy) orbitals on nearest neighbour iron atoms share the added charge. This results in a shortening of the Fe-Fe distance in the dimer and a lengthening of the distance between oxygen atoms that bridge the two iron atoms.

Figure 4.16 summarises these observations from the benchmarking stage of the Fe_2O_3 calculations. For the electron polaron, the lowest energy states were consistently in the *Dimer* form. However, localisation on a single iron atom could also be observed if the distortion was initiated as such and if the ASE rattle function was not used. Indeed, using the ASE rattle function consistently resulted in lower energy states for both the electron and hole polarons. Thus, for consistency, the ASE rattle function was used in all further cases. The results given here do not necessarily mean that the results of Ahart et al[326] are more accurate than those of Pastor et al[330], but merely that the dimer state polaron appears to represent a *deeper* (lower energy) self-trapped state. Given that DFT always seeks to minimise the total energy of a given system, the method of polaron calculation suggested in this work should therefore be considered as a means to find the *deepest* self-trapped states, not necessarily those that dominate charge transport in the material.

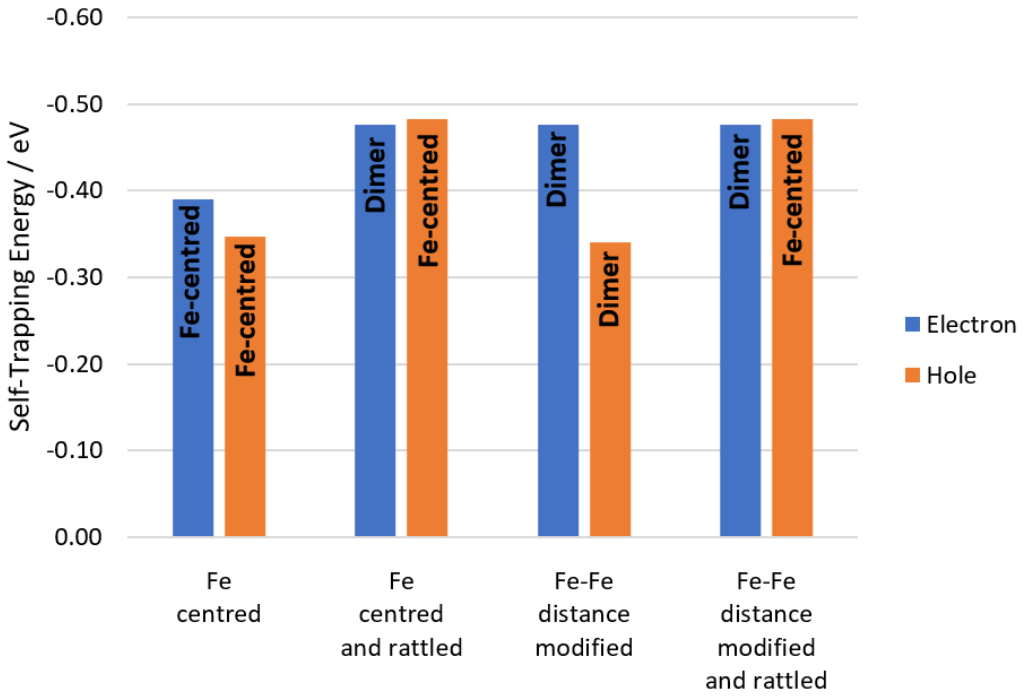
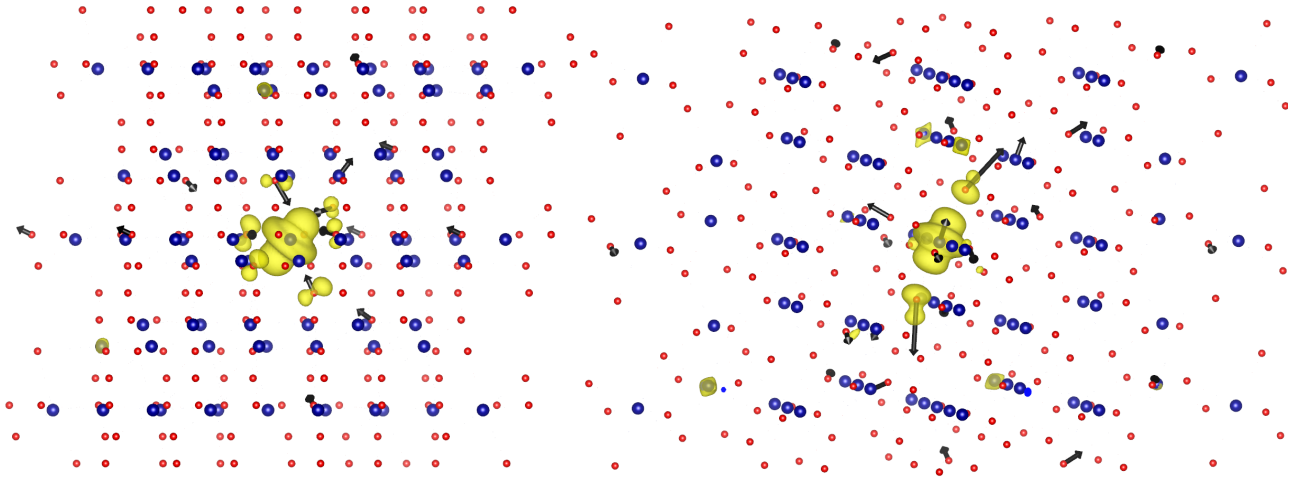


Figure 4.16: Binding energies of Fe_2O_3 using Γ-point calculations. The horizontal axis represents the initial configuration of the system, whilst the description on each bar denotes the resulting configuration following a structural relaxation. *"Fe-centred"* means that the resulting polaron was centred around a single iron atom and/or that the calculation was initiated as such. *"Dimer"* means that the resulting polaron was centred around two iron atoms. *"Fe-Fe distance modified"* means that the localisation was induced by moving two nearest neighbour iron atoms closer together or further apart. *"Rattled"* means that the ASE rattle function was used before the structural relaxation. It should be noted that these calculations were performed during the benchmarking stage of this investigation, using Γ-point sampling and a smaller Fe_2O_3 unit cell. Thus, the precise values of the binding energies differ from those given later in this chapter. Nevertheless, the same lowest energy morphology was reproduced for the larger unit cell.

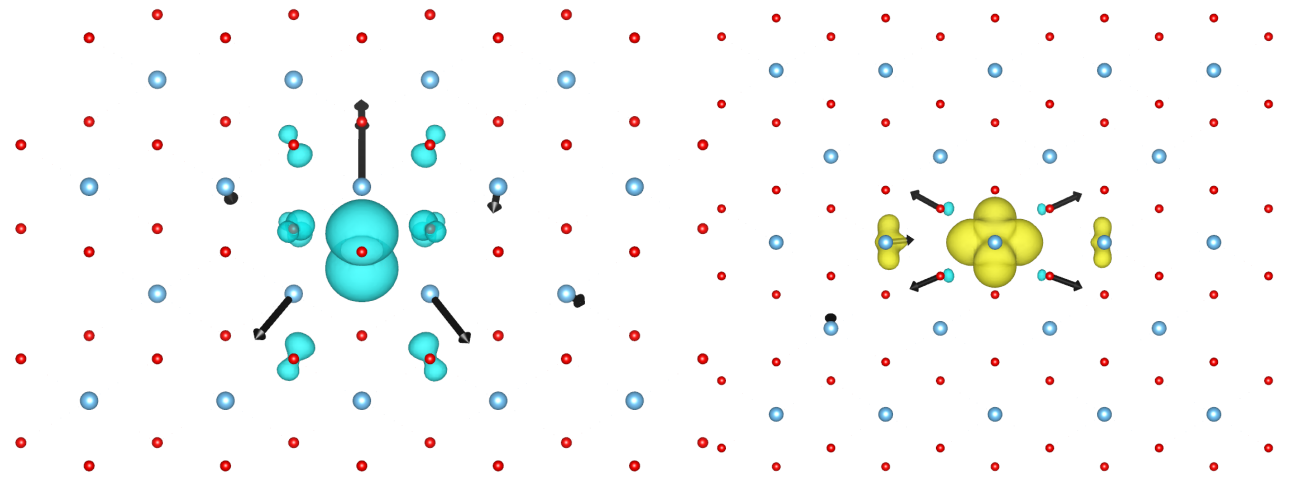
Cr_2O_3 The polaron isosurfaces for Cr_2O_3 are given in Figure 4.17. These morphologies support the observations of Dehn et al[337], albeit without the muon stabilisation. It is noted that no Cr-Cr dimer electron polarons were observed for this compound, unlike the isostructural Fe_2O_3. This may be due to the difference in the magnetic structures of Fe_2O_3 and Cr_2O_3. Whilst both systems involve alternating *up* and *down* spin layers of charge throughout the unit cell, these layers are orientated across different planes of the unit cell. This is such that the nearest neighbour atoms of Fe_2O_3 all share the same spin parity, whereas the opposite is true for Cr_2O_3. Since the Fe-Fe dimer polaron in Fe_2O_3 is formed between nearest neighbour iron atoms of the same spin, dimer polaron formation does not disturb the bulk magnetic structure of Fe_2O_3. In Cr_2O_3, this would not be the case. Thus, merely as a hypothesis for future consideration, there may be an energetic cost associated with Cr-Cr dimer polaron formation that is not present for Fe_2O_3.



(a) Hole **(b)** Electron

Figure 4.17: Cr_2O_3 polaron isosurfaces. It can be seen that both the electron and hole polarons are centred around a chromium atom. In both cases, the chromium d(z^2) orbital is shown to be involved in the binding of the unpaired spin. In the hole case, all six nearest neighbour oxygen p orbitals are involved, whereas in the electron case, the oxygen p orbitals along the chromium d(z^2) axis contribute noticeably more. The atomic displacements reflect this distinction with the hole polaron nearest neighbour oxygen atoms moving towards the central chromium atom in approximately equal measure, presumably to compensate for the positive charge. For the electron polaron, the d(z^2) axial nearest neighbours move away from the charge more strongly than the oxygen atoms in the d(xy) plane of the central chromium, presumably to accommodate the added negative charge.

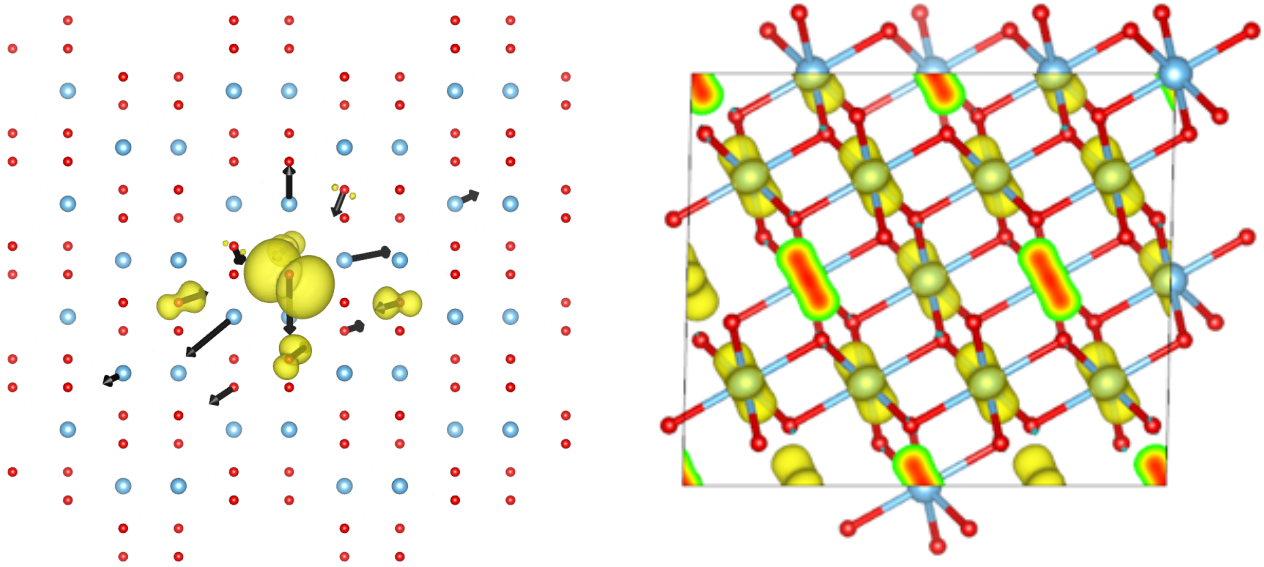
TiO_2 The polaron isosurfaces for the rutile phase of TiO_2 are given in Figure 4.18. The electron polaron morphology is in agreement with those of Elmaslmane et al[318], Deak et al[319], and Pham et al[306]. The hole polaron morphology is in agreement with those of Pham and Wang[320].



(a) Hole **(b)** Electron

Figure 4.18: TiO_2 (rutile) polaron isosurfaces. It can be seen that the hole polaron is centred around an oxygen atom, with the central oxygen p orbital and its nearest neighbours contributing the most to the binding of the unpaired spin. The eight nearest neighbour oxygen p orbitals contribute evenly to the polaron. The nearest neighbour titanium atoms also move outwards from the polaron centre. The electron polaron is centred around a titanium atom. The titanium d(xy) orbital contributes most strongly to its stabilisation, with the four nearest neighbour oxygen atoms in the d(xy) plane moving outwards from the polaron, likely due to the electrostatic repulsion.

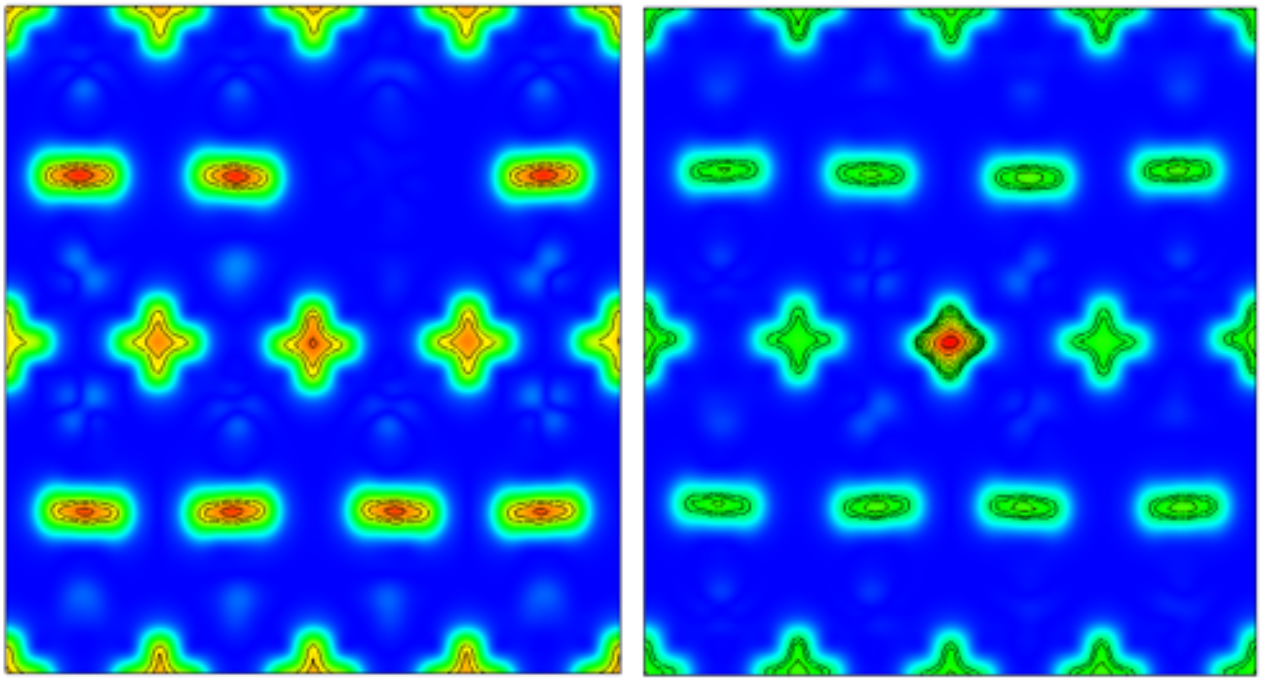
The polaron isosurfaces for the anatase phase are given in Figure 4.19. The hole polaron morphology agrees with the results of Elmaslmane et al[318] and Carey et al[422]. No electron polaron formation was observed which is in line with the observations of Elmaslmane et al[318].



(a) Hole **(b)** Electron

Figure 4.19: TiO_2 (anatase) polaron isosurfaces. It can be seen that the hole polaron is centred around an oxygen atom with the stabilising orbital contribution coming primarily from the central oxygen p and four nearest neighbour oxygen p orbitals. The nearest neighbour titanium atoms move outwards from the polaron site whereas the nearest neighbour oxygen atoms move inwards, presumably to compensate for the removed electron. The central oxygen atom moved downwards, closer to the plane of the contributing nearest neighbours. This may imply a stabilisation interaction through oxygen p-oxygen p orbital overlap. No polaron formation could be observed for the excess electron, given the even distribution of charge density across the unit cell.

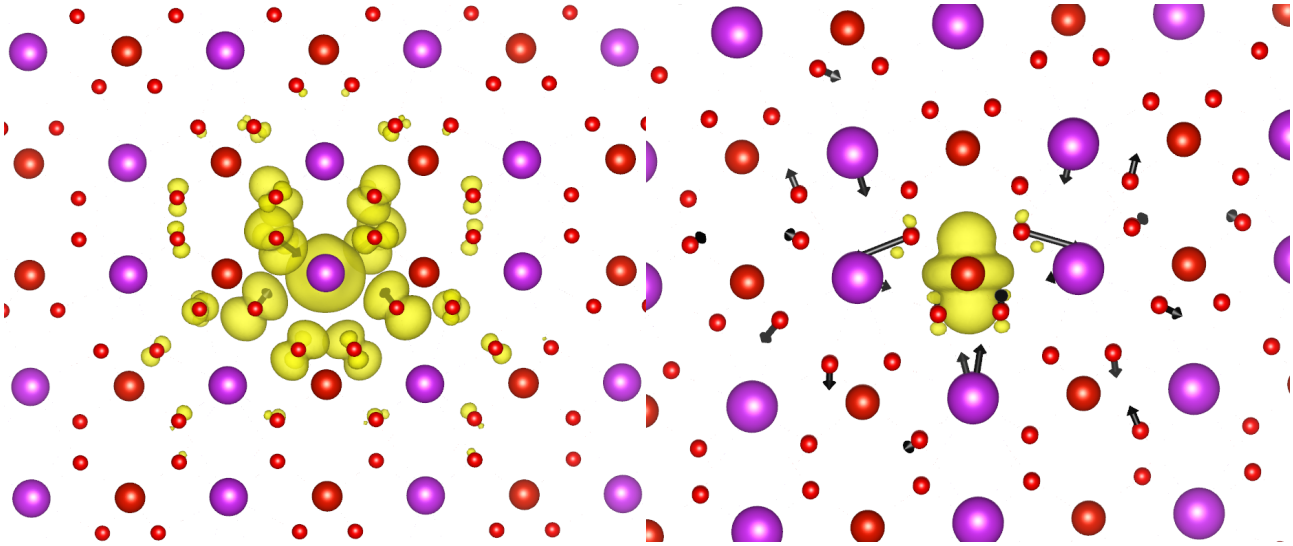
VO_2 The polaron isosurfaces for the VO_2 rutile phase are given in Figure 4.20. The isosurfaces for VO_2 were inconvenient for visualising the polarons as the surrounding (bulk) atom isosurfaces obscured the polaron isosurface. This may be due to the metallic nature of the VO_2 rutile phase. Therefore, a 2D slice of the charge density isosurface is given in Figure 4.20. Unfortunately, only the rutile phase could be studied because the calculation of the monoclinic phase proved to be prohibitively expensive. This may be alleviated in future work by choosing a more appropriate unit cell and abandoning the desire for a near-cubic unit cell.



(a) Hole **(b)** Electron

Figure 4.20: Two dimensional slice of the VO_2 (rutile) polaron charge density. It can be seen that both the electron and hole polarons are centred around a single vanadium atom. For the electron polaron, the d(xy) orbital appeared to contribute the most to the stabilisation of the polaron. For the hole polaron, there appears to be an absence of charge density relative to the surrounding atoms.

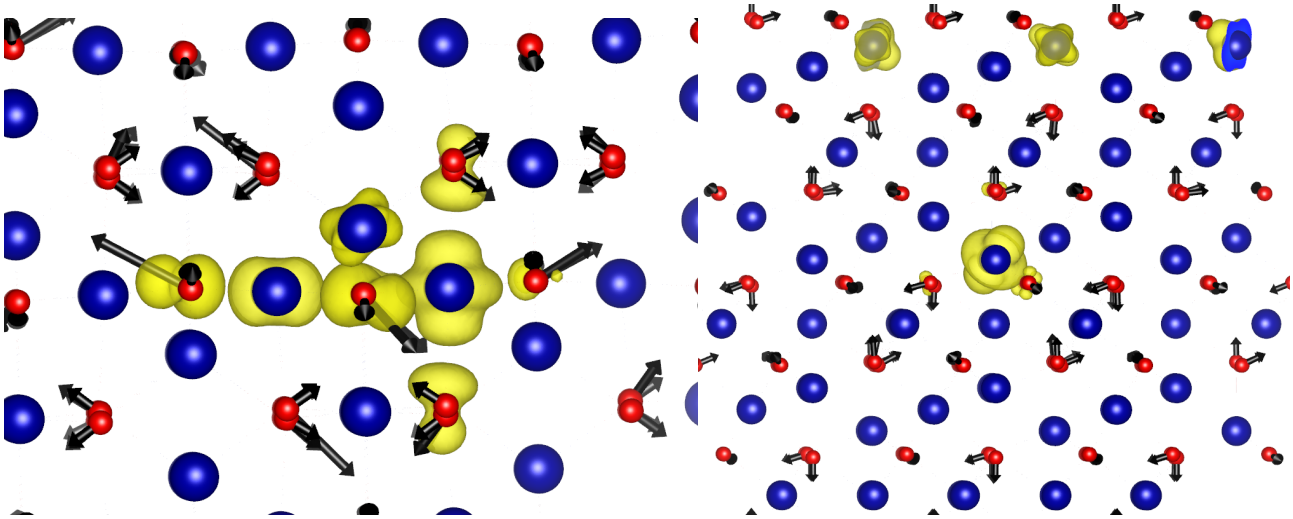
$BiVO_4$ The polaron isosurfaces for $BiVO_4$ are given in Figure 4.21. The electron polaron morphology is in line with those of Wu et al[423] and Pham et al[306]. The hole polaron morphology supports that of Wiktor et al[341].



(a) Hole **(b)** Electron

Figure 4.21: $BiVO_4$ polaron isosurfaces. It can be seen that the $BiVO_4$ electron polaron was centred around a single vanadium atom. The d(z^2) orbital of the central vanadium atom appears to accept the added charge. The hole polaron localises around a central bismuth atom with a substantial contribution from the neighbouring oxygen p orbitals, presumably to stabilise the added charge.

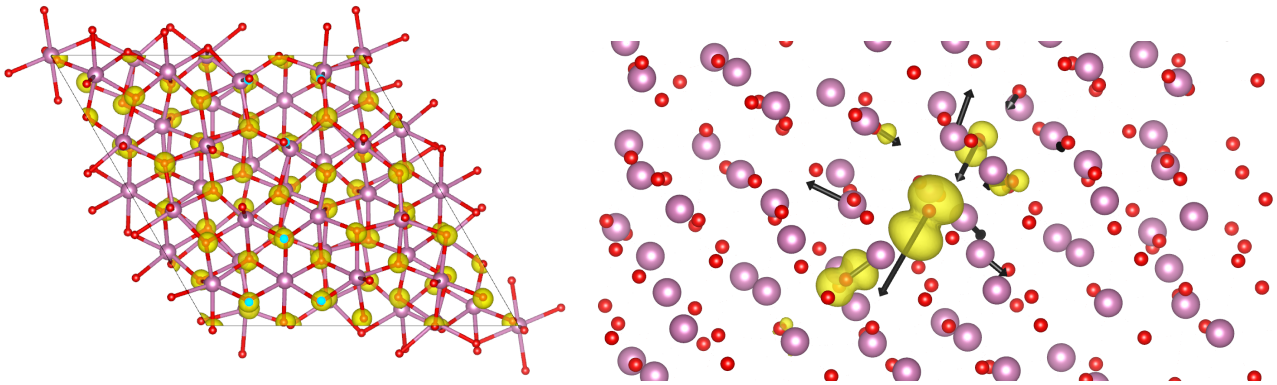
Co_3O_4 The polaron isosurfaces for Co_3O_4 are given in Figure 4.22. In both cases, substantial charge density was observed at the boundaries of the unit cell. This may indicate an inaccuracy in the results for Co_3O_4 as Ahart et al[326] mentioned that such an observation implies an insufficiently large supercell.



(a) Electron **(b)** Hole

Figure 4.22: Co_3O_4 polaron isosurfaces. It can be seen that the Co_3O_4 electron polaron appeared as a dimer polaron. However, unlike Fe_2O_3, the two nearest neighbour atoms used the d(xy) and d(z^2) orbitals to bind the charge. Neighbouring oxygen p orbitals also appeared to contribute to the stabilisation of the electron polaron. The hole polaron appeared to be stabilised by the d(z^2) orbital of a central cobalt atom.

In_2O_3 The polaron isosurfaces for In_2O_3 are given in Figure 4.23. As mentioned, literature concerning polaron formation in this material is sparse.

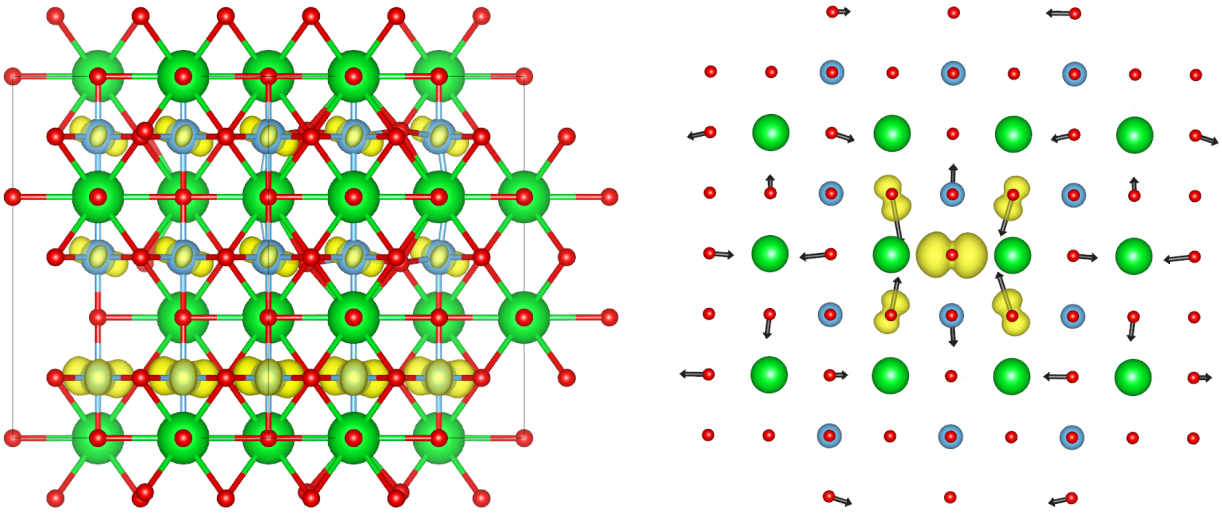


(a) Electron **(b)** Hole

Figure 4.23: In_2O_3 polaron isosurfaces. It can be seen that the In_2O_3 hole polaron was centred around a single oxygen atom and was stabilised through neighbouring oxygen p orbitals. No substantial localisation was observed for the electron polaron.

$SrTiO_3$ The polaron isosurfaces for $SrTiO_3$ are given in Figure 4.24. The hole polaron results are in line with the observations of Janotti et al[394]. However, the electron polaron results contradict the observations of Janotti et al. This may be due to different calculation parameters being used by Janotti

et al.



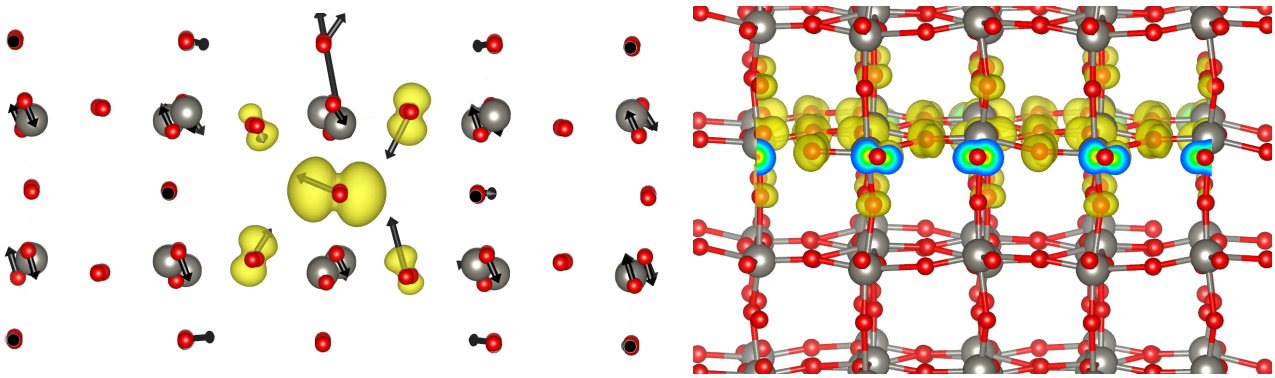
(a) Electron **(b)** Hole

Figure 4.24: $SrTiO_3$ polaron isosurfaces. The $SrTiO_3$ hole polaron was centred around a single oxygen atom and was stabilised through four neighbouring oxygen *p* orbitals in the same plane. These oxygen atoms moved towards the central atom accordingly. No substantial localisation was observed for the electron polaron, though there appears to be some asymmetry in the charge distribution. Perhaps indicating localisation under the right conditions.

4.4.2 Non-Converged Polaron Calculations

Due to the computational expense incurred, the localised states for WO_3, NiO, CuO, and Bi_2O_3 were not fully converged with the **k**-point grid density beyond Γ-point sampling. This was due to either the electronic minimisation steps being too costly, the structural relaxation of the polaron state requiring too many ionic steps, or the chosen unit cell being insufficient for polaron formation. Their preliminary results at the Γ-point are given below.

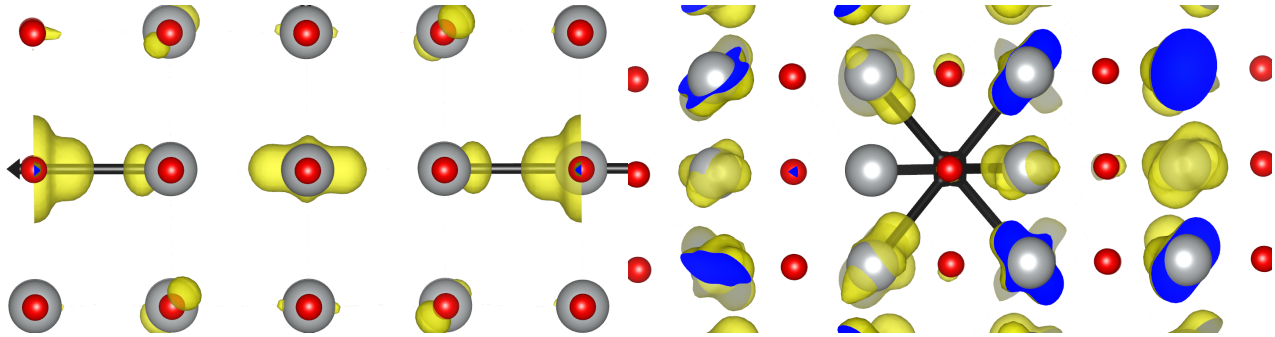
WO_3 The non-converged isosurfaces for WO_3 are given in Figure 4.25. These morphologies are in line with the observations of Tao and Liu[399] for the monoclinic phase. However, due to the limited size of the unit cell, it was not clear whether or not the electron polaron isosurface would extend further across the plane. A larger supercell would be required to confirm this. The calculation could not achieve full **k**-point convergence because the structural relaxation beyond the Γ-point relaxation proved to be incredibly slow, with an excessive number of structural relaxation steps. This may have been due to the orthorhombic phase relaxing into another, lower energy, phase of WO_3 whilst simultaneously stabilising the polaron state.



(a) Hole **(b)** Electron

Figure 4.25: WO_3 polaron isosurfaces. It can be seen that the hole polaron is centred around an oxygen atom, stabilised by nearest neighbour oxygen p orbitals. The electron polaron appeared to localise across a single plane of the WO_3 unit cell, but confirmation of this is limited by the bounds of the unit cell.

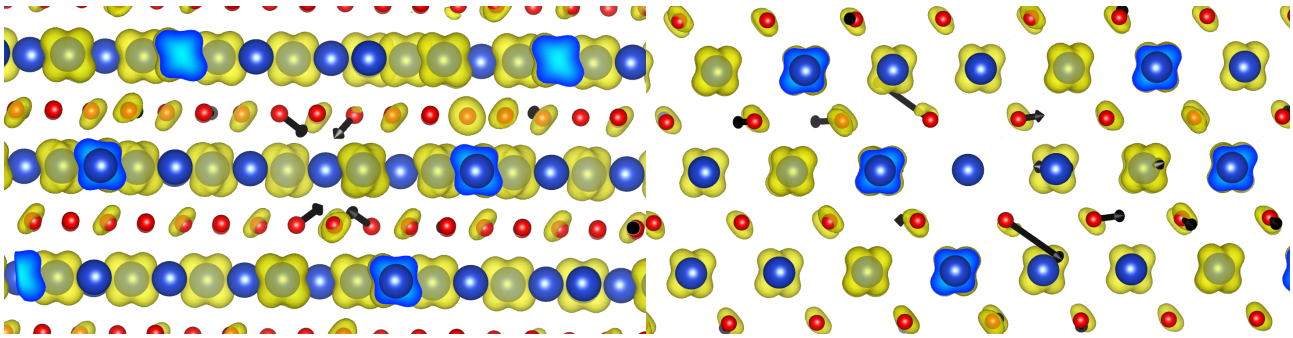
NiO The non-converged isosurfaces for NiO are given in Figure 4.26. It can be seen that the NiO supercell was of insufficient size to contain the hole and electron polarons. These morphologies contradict the work of Yu et al[385]. Better agreement may be obtained by using a larger supercell.



(a) Hole **(b)** Electron

Figure 4.26: NiO polaron isosurfaces. The hole polaron appears to be centred around a nickel atom. The electron polaron appears to be centred around an oxygen atom and is stabilised via neighbouring nickel d(xy) orbitals.

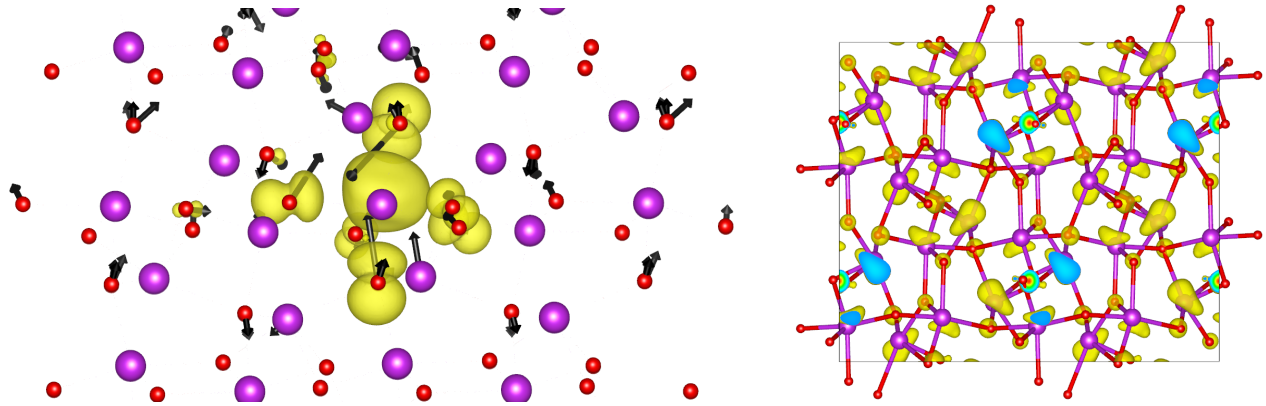
CuO The non-converged isosurfaces for CuO are given in Figure 4.27. In both cases, the spin density isosurface shows an absence of charge. This may indicate agreement with Smart et al regarding a *"flipped spin"*[365], and therefore inverted spin density, of the localised state. However, the calculation struggled to efficiently converge past the Γ-point sampled pre-convergence step.



(a) Hole **(b)** Electron

Figure 4.27: CuO polaron isosurfaces. The hole polaron appears to be centred around a copper atom, with neighbouring oxygen atoms moving towards the central copper atom. A similar observation is made for the electron polaron, except that the nearest neighbour oxygen atoms move outwards from the central copper atom.

Bi_2O_3 The non-converged isosurfaces for Bi_2O_3 are given in Figure 4.28. As mentioned, literature concerning the polaron isosurfaces for this material is sparse.



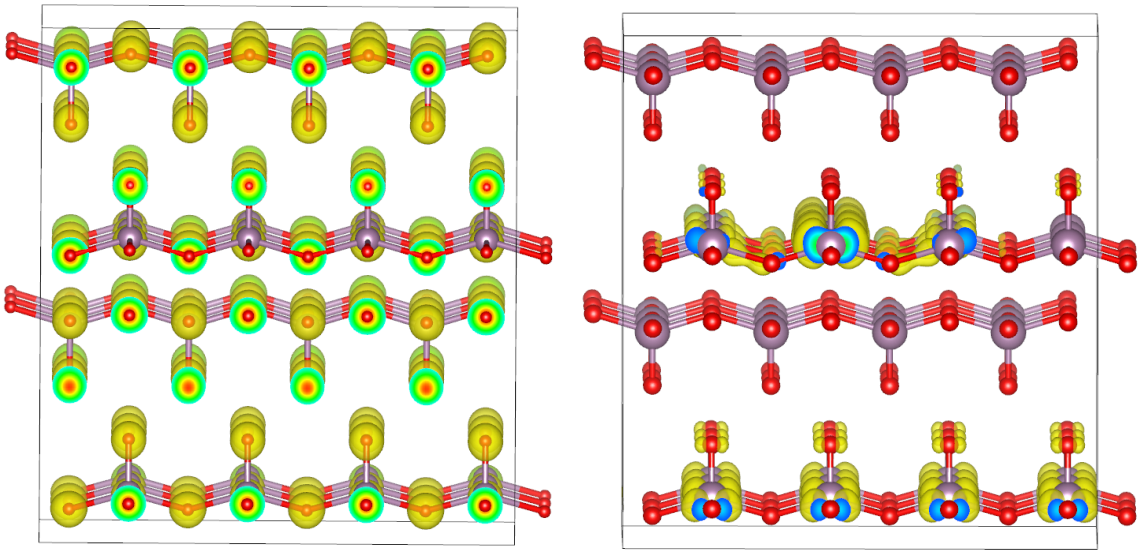
(a) Hole **(b)** Electron

Figure 4.28: Bi_2O_3 polaron isosurfaces. It can be seen that the hole polaron is centred around a single bismuth atom, with surrounding oxygen p orbitals stabilising the added charge. No substantial localisation was observed for the electron polaron, as shown by the diffuse and even distribution of the charge density isosurface.

4.4.3 Non-Localising Polaron Calculations

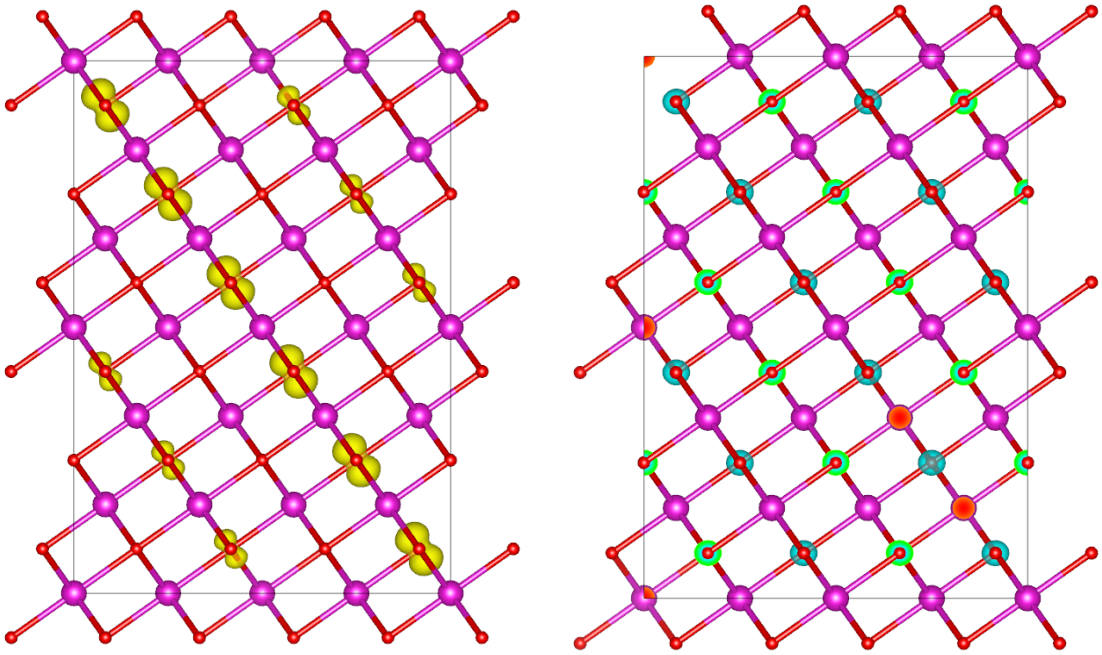
MoO_3, CdO, SnO, Cu_2O and ZnO each failed to localise any excess charge with Γ-point sampling. This may explain the sparseness of literature regarding CdO, SnO, Cu_2O and ZnO polaron morphologies, which may be difficult or unlikely to form. However, it does not account for the polaron MoO_3 morphologies observed in previous studies[374,375,376]. This may be due to the use of the bond distortion method (as opposed to the substitution method), the **k**-point sampling may have been too coarse, the wrong distortion may have been used, or the HSE06 Hamiltonian operator may have been insufficient. In any case, further investigation will be required. Merely for completeness, the preliminary

morphologies of each delocalised solution, using Γ-point sampling, are given below.



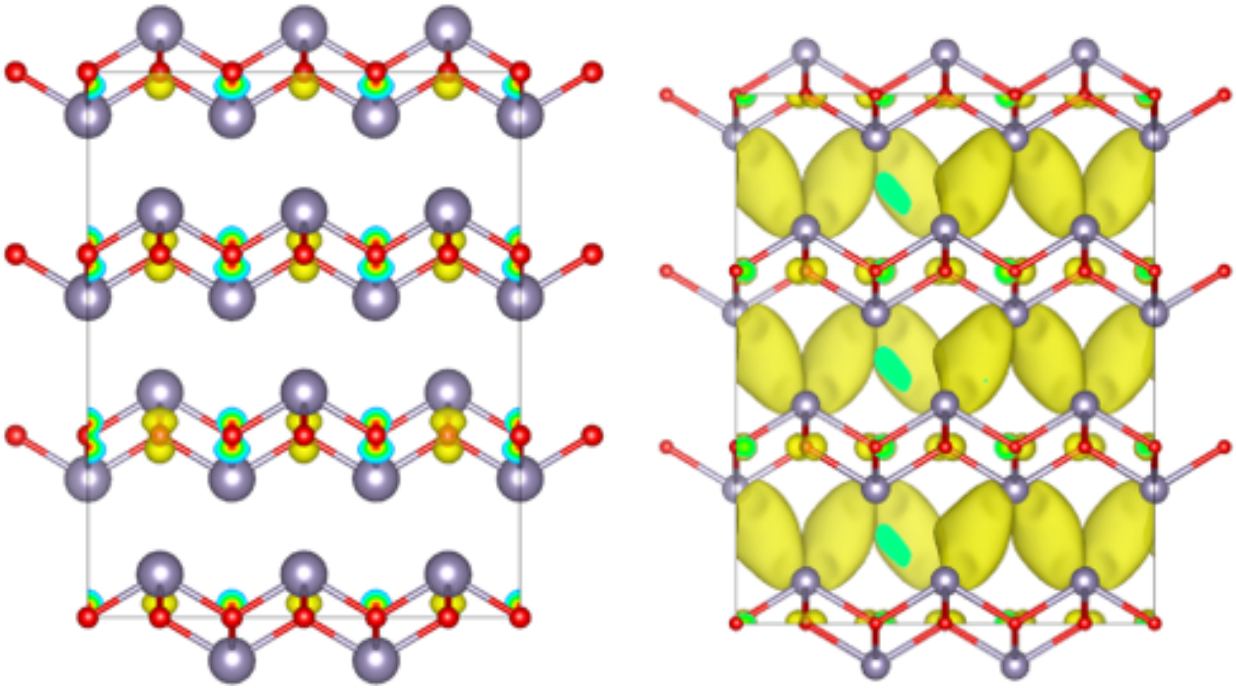
(a) Hole **(b)** Electron

Figure 4.29: MoO_3 delocalised isosurfaces. For the excess hole, no substantial localisation was observed by virtue of the even distribution of charge density across the atoms in the unit cell. For the excess electron, there is a noticeable asymmetry in the charge density distribution. However, this does not correspond to a clearly definable small polaron morphology.



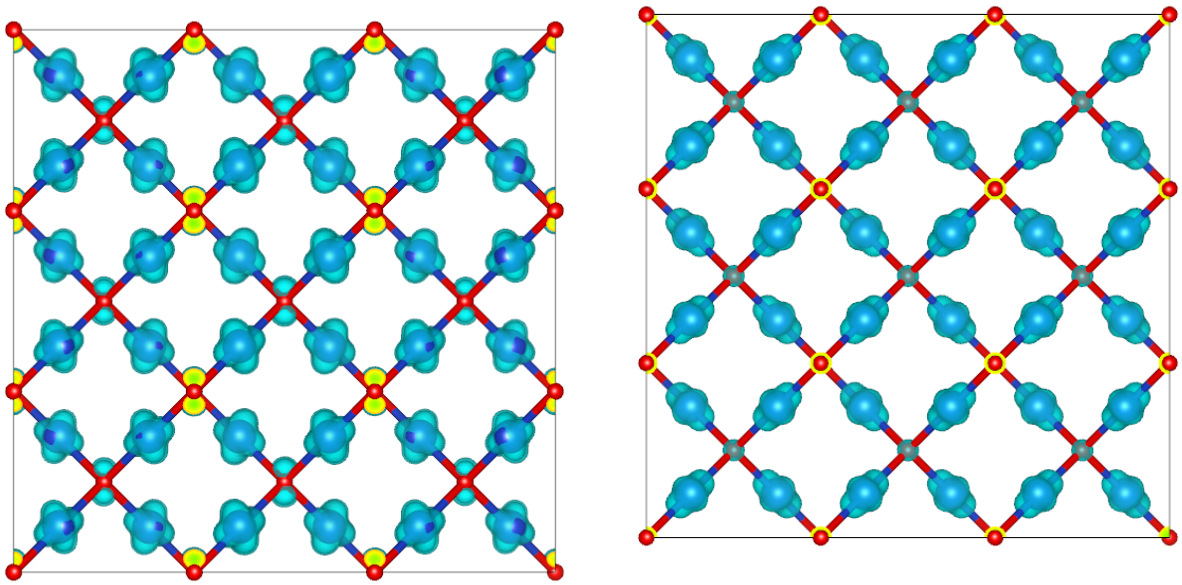
(a) Hole **(b)** Electron

Figure 4.30: CdO delocalised isosurfaces. In the case of an excess hole, there is a noticeable asymmetry in the charge density distribution. However, this does not correspond to a definable small polaron morphology. The charge is otherwise evenly distributed across the unit cell. In the case of an excess electron, the charge density is evenly distributed across the unit cell, indicating no substantial localisation.



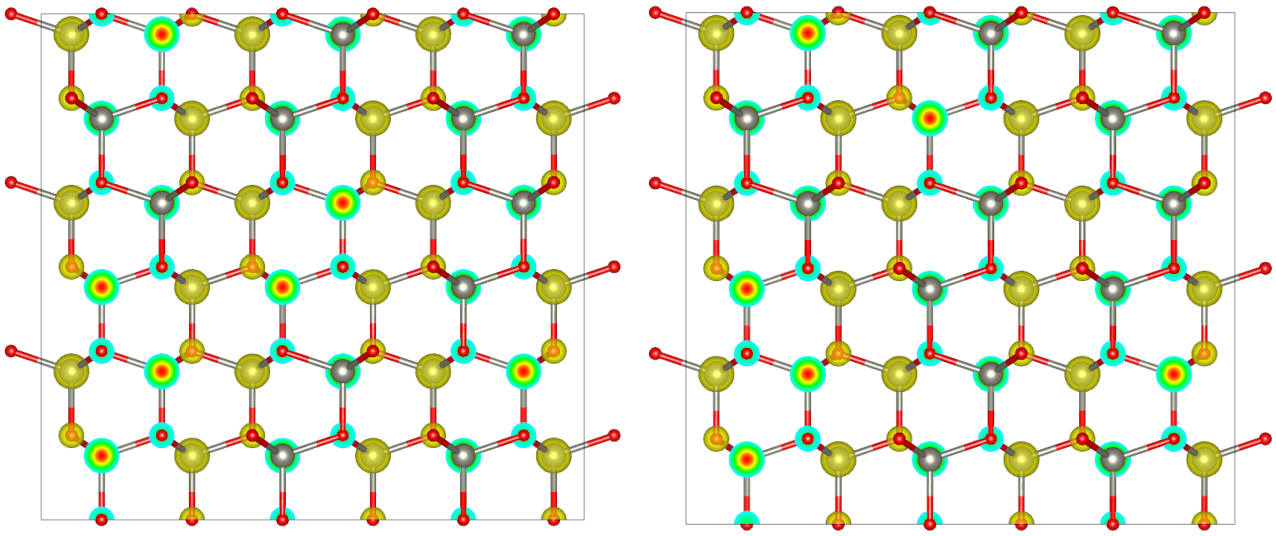
(a) Hole **(b)** Electron

Figure 4.31: SnO delocalised isosurfaces. In both the electron and hole cases, the excess charge density is spread evenly and symmetrically throughout the unit cell, implying delocalisation.



(a) Hole **(b)** Electron

Figure 4.32: Cu_2O delocalised isosurfaces. In both the electron and hole cases, the excess charge density is spread evenly and symmetrically throughout the unit cell, implying delocalisation.



(a) Hole **(b)** Electron

Figure 4.33: ZnO delocalised isosurfaces. In both the electron and hole cases, the excess charge density is spread evenly and symmetrically throughout the unit cell, implying delocalisation.

4.4.4 Converged Self-Trapping Energies

The self-trapping energies of the compounds that were able to achieve full **k**-point convergence were calculated to determine whether or not the presently suggested method produces reasonable results for the self-trapping energy. In Figure 4.34 the compounds are placed in order of their d-orbital occupancy. Other than the open shell oxides seemingly having greater self-trapping energies, implying more deeply trapped polarons than their closed shell counterparts, there appeared to be no trend between the d-orbital occupancy and the self-trapping energy. Co_3O_4 had a highly spurious energy, with the self-trapping energy being greater than the band gap.

To determine if the pre-convergence with a single **k**-point was causing this issue, a polaron relaxation of Co_3O_4 was performed without pre-converging the calculation with a coarse **k**-point grid. The resulting polaron morphology was completely different from the pre-converged case. The band decomposed charge density isosurfaces of the five partially occupied states in this fully converged Co_3O_4 calculation are given in Figure 4.35. The localisation appears to centre around an octahedral (d^6) cobalt atom for band 611, two neighbouring cobalt atoms for bands 614 & 615, and was less obvious for bands 612 & 613. The presence of multiple partially occupied states within the band gap supports the findings of Smart et al[31] concerning the presence of mid-gap states. However, the morphologies presented here do not agree with those of Smart et al. This may be due to the enforcement of different constraints on the Hamiltonian operator in their work.

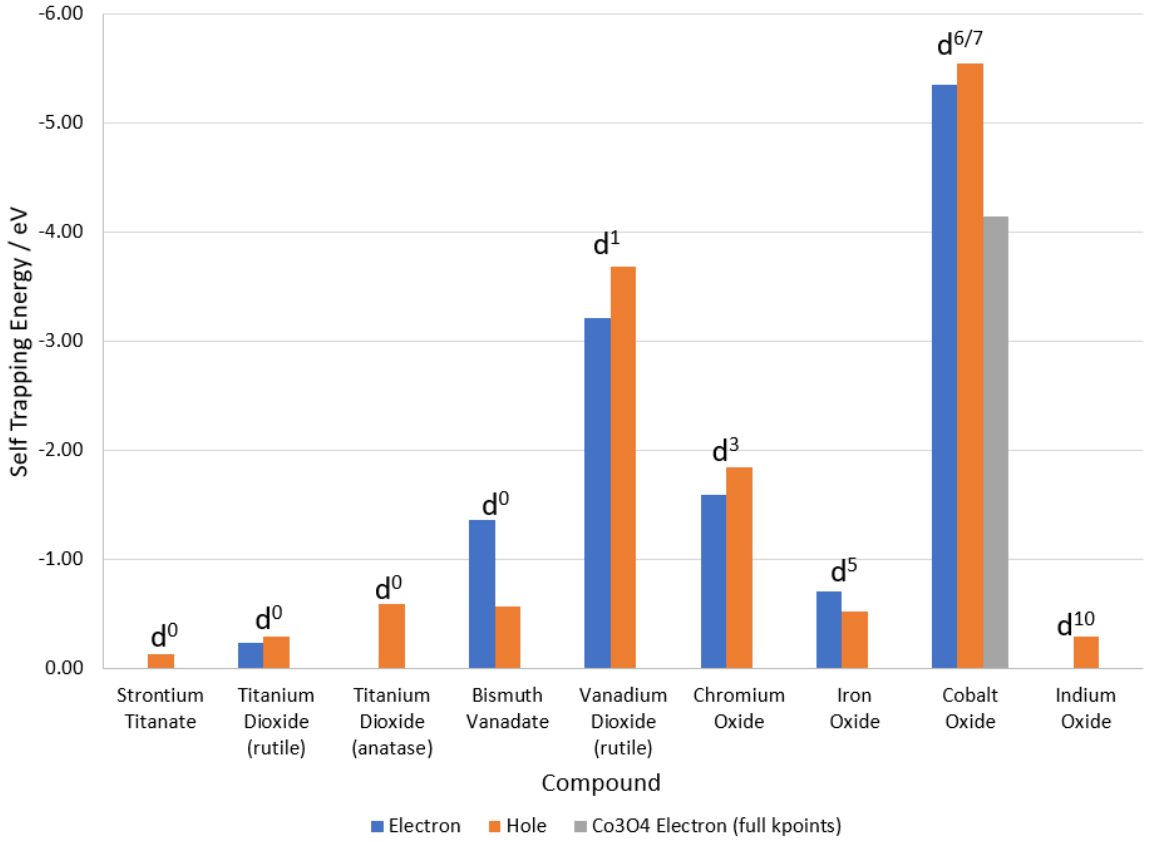
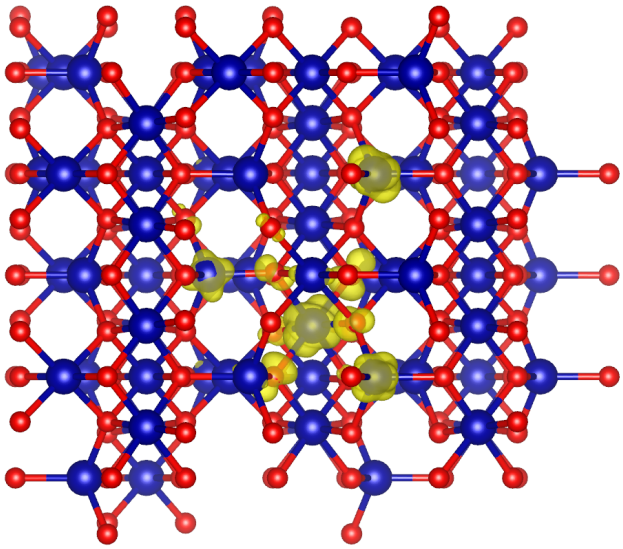
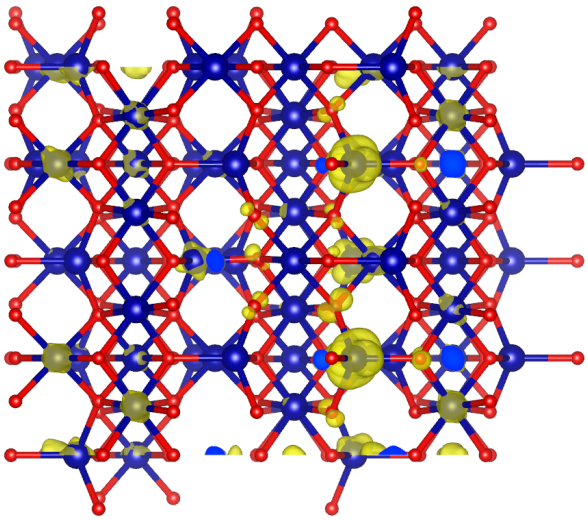


Figure 4.34: Fully converged electron and hole self-trapping energies. The d-orbital occupancy is given at the top of each bar. The grey bar corresponds to the Co_3O_4 electron polaron calculation that was not pre-converged with a coarse **k**-point grid.

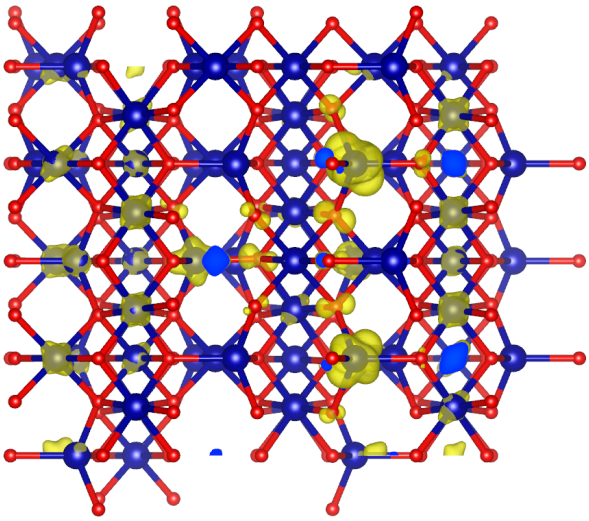
Furthermore, whilst this fully converged calculation reduced the self-trapping energy, it remained greater than the Co_3O_4 band gap. This may imply a general issue with simulating the polaron formation of Co_3O_4, aside from any inaccuracies that may introduced by preconverging with a coarse **k**-point grid. Singh et al[353] mentioned that the Co_3O_4 electronic structure would only be accurately represented when 5% of HF (as opposed to the usual 25%) exchange was used. Conversely, Smart et al[31] stressed the need to use the *generalised Koopmans' condition*, to maintain accuracy. As such, Co_3O_4 presents a challenging compound for the high throughput screening of polarons, requiring at least some degree of fine-tuning of the Hartree-Fock exchange to produce non-spurious results. The metallic VO_2 (rutile) phase also had comparably large self-trapping energies and, upon inspection of the eigenvalues, it also comprised three partially occupied states. However, since there is no band gap for this phase, it is difficult to tell how reasonable this self-trapping energy value is. It is respectfully submitted that further investigation of the electronic structure of these compounds is required.



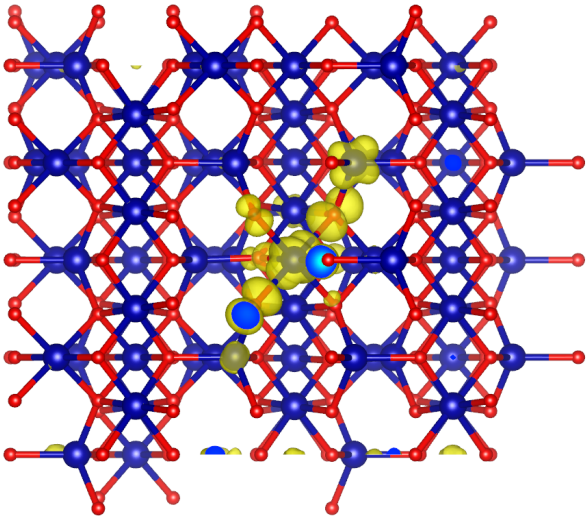
(a) Electron band 611



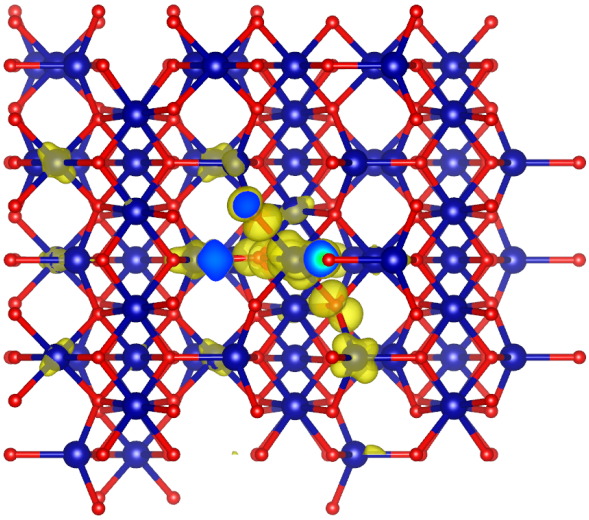
(b) Electron band 612



(c) Electron band 613



(d) Electron band 614



(e) Electron band 615

Figure 4.35: Band decomposed charge density of the highest energy, partially occupied states for the fully converged Co_3O_4 calculation. The cobalt orbitals involved in these states appear to be mixtures of d(xy) and d(z^2) orbitals. None of these morphologies correspond to those observed for the Γ-point calculations, nor those reported in past literature[31].

4.4.5 Reducing the Computational Cost

Despite these issues with Co_3O_4, the greatest cost-saving exercise was achieved through the use of a single **k**-point (at Γ) in reciprocal space. In general, it was observed that this level of precision was sufficient to determine whether or not localisation was successful, as well as the location of the polaron. Testing in this way allowed for the rapid pre-screening of multiple symmetry-breaking distortions. Doing so for a full **k**-point grid would have been, and was often found to be, prohibitively expensive for such a large dataset. However, as a result of the coarse **k**-point sampling, the total energies were expected to be inaccurate. Therefore, following the relaxation of these multiple starting geometries, the output state with the lowest total energy was selected and checked for localisation. It is noted that even

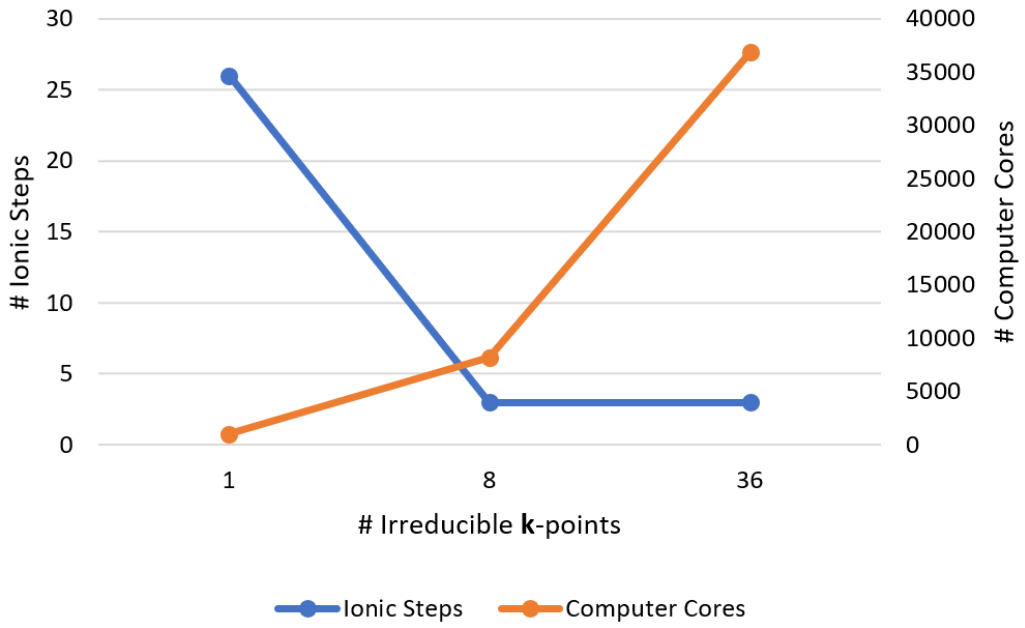


Figure 4.36: (Left axis - blue line) Number of geometry optimisation (ionic) steps in the structural relaxation of the Fe_2O_3 electron polaron for the number of **k**-points. (Right axis - orange line) Number of computer cores, required to complete the calculation for a given number of **k**-points within a 24-hour interval. This figure shows that the number of required computer cores increased rapidly with the number of **k**-points. However, the majority of the ionic steps were accounted for with a single **k**-point.

though the energies were quantitatively inaccurate, the qualitative relation between the different states (i.e. which state was lowest in energy) appeared to still be preserved. If localisation was successful, calculations with finer **k**-point grids were performed on the output until convergence was achieved. It was also found that it was more efficient to stagger the **k**-point convergence to ensure that the majority of the ionic steps for the structural relaxation were performed using the coarsest (cheapest) **k**-point grids. This is illustrated in Figure 4.36 where the highest number of ionic steps were accounted for with a single **k**-point. As mentioned in the methods section of this chapter, further performance (speed)

enhancing measures were also used in the calculations. These are standard implementations of the VASP code and are briefly elaborated on below:

PRECFOCK = FAST A more coarse FFT grid is used for the exact exchange (Hartree Fock) portion of the calculation. It is believed that the exact exchange component is relatively insensitive to the FFT grid density[417]. Hence a more coarse grid FFT is permissible, increasing the speed of the calculations whilst maintaining their accuracy.

LSUBROT = .TRUE. This measure uses an algorithm that optimises rotation matrices[c] between occupied and unoccupied states to calculate an optimal charge density matrix. It is believed to speed up hybrid functional calculations for narrow band gap semiconductors and metals[418].

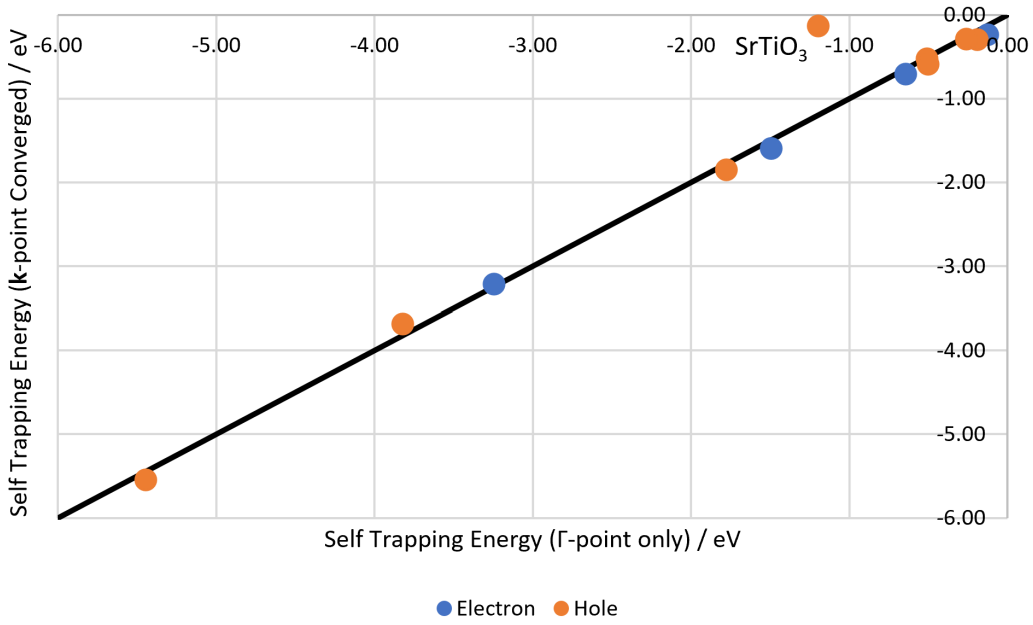
LREAL = Auto Density functional theory calculations may be performed in either real space or reciprocal space. As mentioned in previous sections, it is often more efficient to work in reciprocal space for bulk property calculations in which a primitive unit cell calculation is necessary. In the case of large supercells, such as those used for polaron calculations, it is often more efficient to evaluate the PAW projection operators in real space. Thus, VASP allows for the PAW projection operators to be evaluated in either real or reciprocal space[419].

4.4.6 Further Reducing the Computational Cost

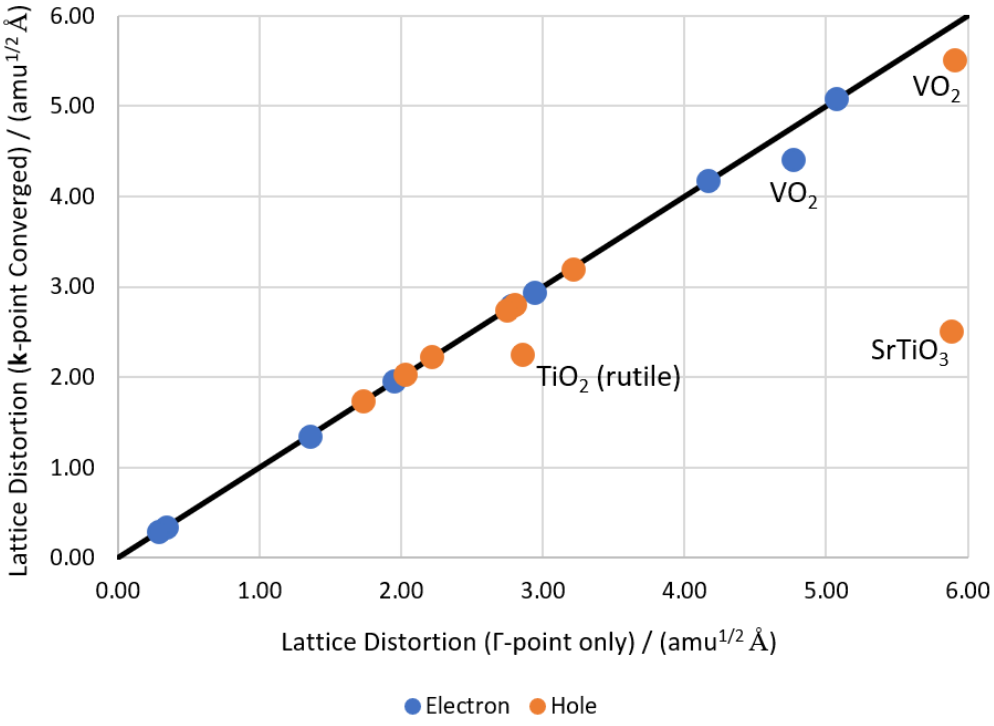
Given the computationally costly and labour-intensive process of refining the polaron calculations with finer **k**-point grids, and visually inspecting the orbitals involved, it is tempting to consider if useful information can be gleaned directly from the Γ-point calculations, without having to consider refined calculations. In this regard, further analyses of the differences between the Γ-point and fully converged calculations were performed.

These are provided in Figures 4.37, 4.38 and 4.39. From Figure 4.37 (a) it can be seen that the self-trapping energy was largely unaffected following the Γ-point pre-convergence, the only outlier being $SrTiO_3$. However, given that the results for Co_3O_4 changed drastically depending on whether Γ-point pre-convergence was used, it may be the case that pre-converging the calculation at the Γ-point may restrict the system to a particular morphology. If the morphology is accurate, such as for TiO_2, Fe_2O_3 and $BiVO_4$, a great deal of computational expense may be saved by only using Γ-point data. However, for compounds like Co_3O_4, it appears to have a substantial effect on the final result.

[c] A rotation matrix is a transformation matrix that has the effect of rotating a vector (i.e. in degrees or radians) when operating on the vector.



(a)



(b)

Figure 4.37: (a) Comparison of the self-trapping energies using a single **k**-point with the fully converged **k**-point grids. The distance from the black line indicates the degree to which the two values are in agreement. (b) Comparison of the overall lattice distortions using a single **k**-point with the fully converged **k**-point grids. The *Lattice Distortion* quantity was obtained using the *getdelQ.py* code of the *carrier capture* utility[32]. The *getdelQ.py* code obtains the mass-weighted coordinate difference between two structural input files and is thereby a means to represent the overall distortion within a unit cell. The distance from the black line indicates the degree to which the two values are in agreement.

If this is the case, the procedure suggested here may remain useful as a screening tool, albeit with further considerations for more complicated candidates. For instance, a comparison of the self-trapping energy with the band gap, as performed for Co_3O_4, could be used as a brief check of the reliability of the results, flagging the compound as a likely outlier that requires further consideration. Alternatively, increasing the size of the supercell whilst using a single **k**-point may be a more robust method of ensuring that the reciprocal/real space convergence and supercell size requirements are met. However, using this method may prove to be as expensive as a full **k**-point converged calculation.

From Figure 4.37 (b), it can be seen that the extent of the overall lattice distortion was largely unaffected by the number of **k**-points included in the calculation. Indeed, and in most cases, neither the crystal structure nor the isosurface varied considerably following an increase in the **k**-point grid density. Notable exceptions include the TiO_2 rutile hole polaron, both VO_2 polarons and the $SrTiO_3$ hole polaron. However, in each of these exceptional cases, the isosurfaces for the Γ-point and fully converged calculations were almost identical. Therefore, these exceptions are attributed to small and collective changes in the cell shape and internal coordinates. As such, the overall effect on the polaron morphology was inconsequential.

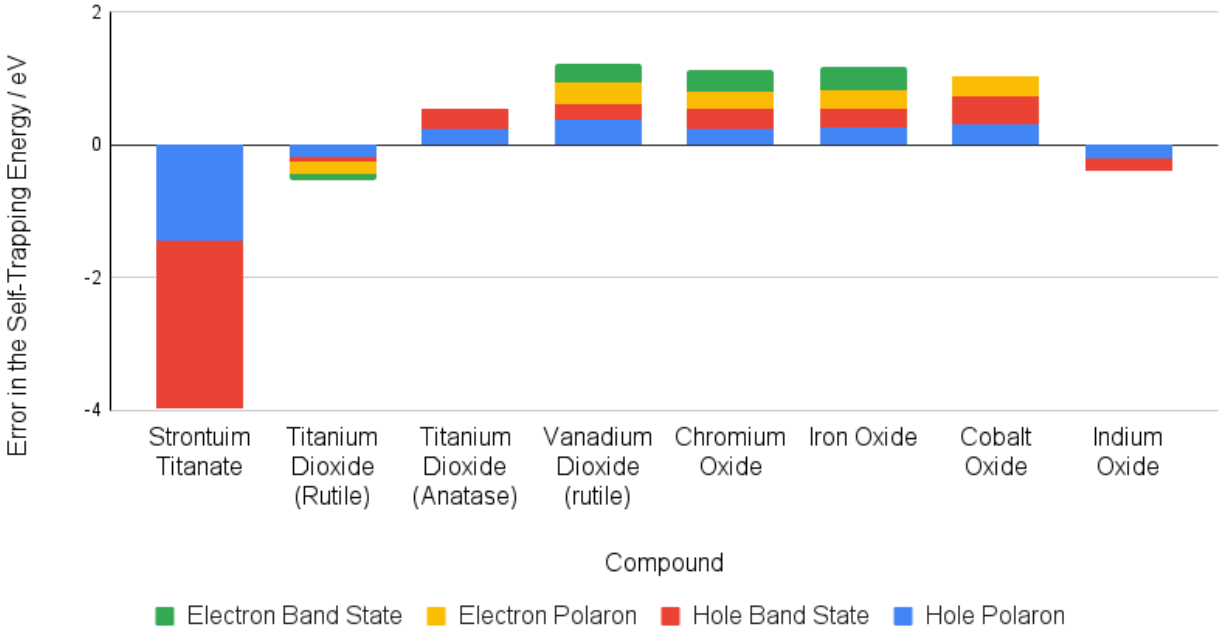


Figure 4.38: Comparison of the fully converged polaron and band state total energies with their respective Γ-point energies. The *"Errors"* displayed in this figure are the differences between the fully converged total energies and their corresponding Γ-point total energies.

To further investigate how the Γ-point and fully converged energies differed, the total energies of the polaron and band states were compared directly, as opposed to only comparing the resultant self-trapping energies. The results are shown in Figure 4.38. From this figure, it can be seen that, since the

calculated errors are all of the same sign and that the corresponding polaron and band state energies are similar in magnitude, the error in the polaron energy is substantially cancelled by the error in the band state. In other words, the band and polaron state energies shift by approximately the same amount as the **k**-point density is increased. Thus, when the self-trapping energy calculation is performed, the overall trend between different compounds is preserved. However, since the individual errors are of the order of 0.2 eV - 0.3 eV (i.e. the same order as many self-trapping energies), the *qualitative* but not *quantitative* accuracy of the Γ-point data must be emphasised. For example, the Fe_2O_3 hole polaron and band state errors are 0.27 eV and 0.28 eV respectively (see Table 4.2), leading to a difference in the calculated Γ-point and fully converged self-trapping energies of 0.01 eV, a very small quantitative error. However, the same quantitative error for $SrTiO_3$ is 1.06 eV. As mentioned, $SrTiO_3$ is a notable exception despite appearing to retain the same polaron morphology at higher **k**-point densities. Since the aforementioned error cancellation trend is still observed for $SrTiO_3$, albeit with unusually large quantitative errors for both the band and polaron states, it is clear that Γ-point sampling poorly represents the total energy of this particular material.

To reduce the computational cost even further, one may consider whether or not there is a relationship between the self-trapping energy and the extent of the structural distortion. Using the same data from Figure 4.37, Figure 4.39 shows that there appears to be a weak but noticeable correlation between the extent of the overall lattice distortion and the self-trapping energy. In addition, Co_3O_4 continues to be a noteworthy outlier. Therefore, a simple analysis of the extent of lattice distortion may be used as an even cheaper means to estimate the self-trapping energy, requiring only a Γ-point relaxation without a corresponding band state calculation.

Merely for the reader's reference, the tabulated values for the above figures are given below, to make clear which data are included in the above figures.

		Electron Energy / eV		**Hole Energy / eV**		**Self Trapping Energy / eV**	
Compound	d^x	**Polaron**	**Band State**	**Polaron**	**Band State**	**Electron**	**Hole**
$SrTiO_3$	0	-1290.66		-1298.54	-1298.41		-0.13
TiO_2(rutile)	0	-1036.09	-1035.86	-1040.76	-1040.47	-0.23	-0.29
TiO_2(anatase)	0	-1039.17		-1042.51	-1041.92		-0.59
$BiVO_4$	0	-1744.71	-1743.35	-1752.66	-1752.08	-1.36	-0.57
VO_2(rutile)	1	-1025.20	-1021.99	-1033.60	-1029.91	-3.21	-3.68
Cr_2O_3	3	-1844.78	-1843.19	-1856.74	-1854.89	-1.59	-1.85
Fe_2O_3	5	-1676.76	-1676.05	-1687.37	-1686.85	-0.71	-0.52
Co_3O_4	6 & 7	-1572.91	-1567.56	-1588.30	-1582.76	-5.35	-5.54
In_2O_3	10	-1682.48		-1689.26	-1688.97		-0.29

Table 4.1: Tabulated data for Figure 4.34.

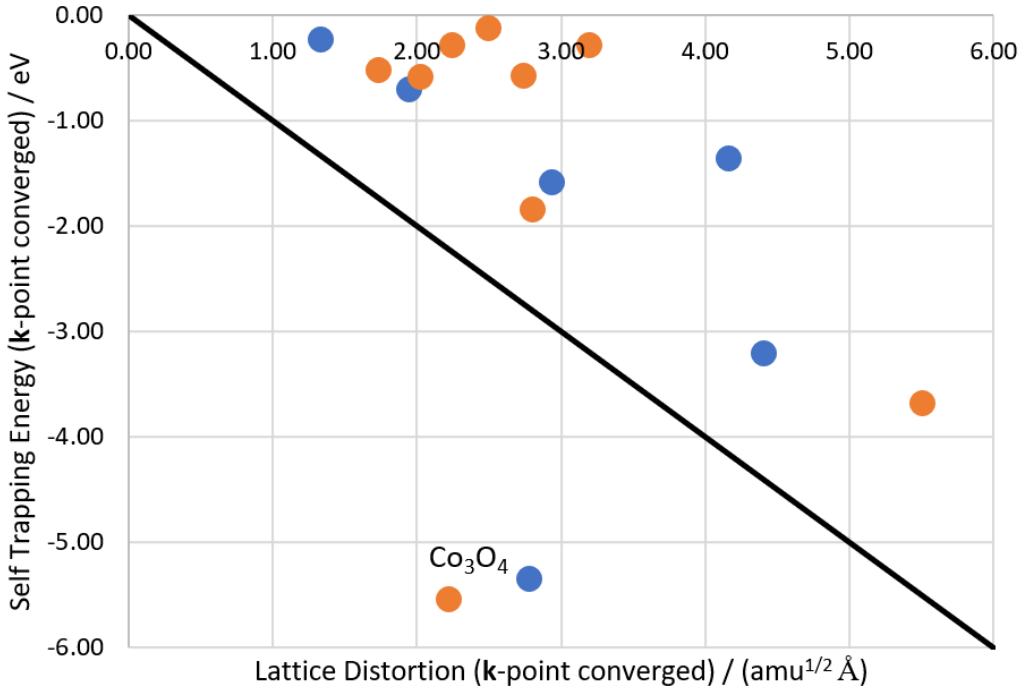


Figure 4.39: Comparison of the fully converged self-trapping energies with the overall lattice distortion. The distance from the black line indicates the degree to which the two values are in agreement.

	Lattice Distortion				**Total Energy / -eV**				**Self-Trapping Energy / -eV**	
	Polaron				**Polaron**		**Band State**			
Comp.	e^- Γ	h^+ Γ	e^- C	h^+ C	e^- Γ	h^+ Γ	e^- Γ	h^+ Γ	e^- Γ	h^+ Γ
$SrTiO_3$	5.08	5.89	5.08	2.50		1297.09		1295.90		1.19
TiO_2 (rutile)	1.36	2.86	1.33	2.24	1035.90	1040.58	1035.77	1040.39	0.13	0.19
TiO_2 (anatase)	0.35	2.03	0.34	2.03		1042.74		1042.24		0.50
$BiVO_4$	4.17	2.75	4.17	2.74						
VO_2 (rutile)	4.77	5.91	4.41	5.51	1025.53	1033.97	1022.29	-1030.14	3.24	3.82
Cr_2O_3	2.94	2.80	2.94	2.80	1845.02	1856.98	1843.52	1855.20	1.49	1.78
Fe_2O_3	1.96	1.73	1.95	1.74	1677.04	1687.64	1676.39	1687.13	0.64	0.51
Co_3O_4	2.78	2.22	2.78	2.22	1573.22	1588.61		1583.17		5.44
In_2O_3	0.29	3.22	0.29	3.20		1689.05		1688.79		0.26

Table 4.2: Tabulated data for Figures 4.37, 4.38, and 4.39. *C* denotes the fully converged data whilst Γ denotes the Γ-point data. The $BiVO_4$ electron, $BiVO_4$ hole, and Co_3O_4 electron self-trapping energies at the Γ-point are omitted due to the band state energies not being fully converged with respect to the total energy. These values could be obtained in future calculations by increasing the maximum number of electronic minimisation steps beyond the default values provided by VASP.

4.5 Conclusions and Future Work

In this chapter, an approach to rapidly screen the self-trapping behaviour of a range of topical transition metal oxides was assessed. Each calculation was preconverged using a single **k**-point calculation, to limit the computational cost. This confirmed or excluded the possibility of localisation for further study at fully converged **k**-point grids. The polaron morphologies for TiO_2, Cr_2O_3, Fe_2O_3, $BiVO_4$, CuO and $SrTiO_3$ (hole polaron only) appeared to agree with past studies and the self-trapping energies were within the band gap for most of the compounds. In addition, new polaron morphologies were found for the orthorhombic phase of WO_3, the rutile phase of VO_2, In_2O_3 and Bi_2O_3. A notable exception was Co_3O_4, whose self-trapping energy and polaron morphologies were affected by the pre-convergence step. However, the results remained spurious even without this pre-convergence step. Therefore, further investigation would be required to determine if the issues were caused solely by the deficiencies of the suggested method or also by the representation of the Co_3O_4 electronic structure. A final analysis of the differences between the Γ-point calculations and their fully converged counterparts showed that the self-trapping energies and structural distortions changed little between successive **k**-point grids. As such, the computational cost may be further reduced by performing a single Γ-point calculation, and if localised, using the extent of the structural distortion as a proxy for the self-trapping energy. This may be a less precise but more efficient means with which to screen candidates.

At the very least, the methods suggested here allow for localisation to be qualitatively determined for a large database of simple compounds at a comparatively low computational cost. Moreover, a better understanding of the method's limitations for more complicated compounds has been established. In particular, a robust and reliable consideration of the size of the supercell will be required to accurately capture the behaviour of systems such as NiO, WO_3, and the monoclinic phase of VO_2. With the above points in mind, a set of improvements or further work using the suggested method may include:

A revised approach to materials selection From a reduced human intervention perspective, it would be more efficient to consider benchmarking each crystal class, using an archetypal compound, before proceeding with a screening investigation across a database of compounds with that structure. Proceeding in this way, similarly to the preceding chapter and forthcoming chapter, has proved to be more efficient from both a benchmarking and production perspective. Furthermore, comparison between different levels of theory (i.e. different DFT functionals) is more reliably and systematically achieved in this way.

Comparison with fully converged calculations To determine conclusively that using coarse **k**-points has a substantial effect on the accuracy of the self-trapping energies. It may also be helpful to repeat the calculation for an entire series of compounds, in a similar fashion to the fully converged Co_3O_4 calculation, and then analyse the overall trend between the two datasets. To this end, it would be wise to select a crystal system that is relatively cheap to produce such converged calculations. From this work, the rutile system would be an ideal candidate, due to its relative simplicity, low computational expense, and a wide variety of candidate photoelectrochemical cell compounds, as will be seen in the next chapter.

Bader charges for verifying charge localisation Human intervention may further be reduced by incorporating Bader[424] charge analyses to partition the charge density following polaron formation and to detect inhomogeneities in the charge density. For instance, if a polaron state is formed, this inhomogeneity may be detected in the Bader charge analysis and subsequently retained for further study. If not, the candidate may be discarded. If properly incorporated into a computational workflow, delocalised candidates could be excluded without requiring visual inspection. Bader charge analyses have already been used in the context of polaron formation to determine the oxidation state of each atom following localisation of the excess charge[306,374].

Finite size corrections As mentioned in the third chapter, finite size corrections are often used to compensate for the self-interaction between a given defect and its periodic images. If facile to implement, such finite size corrections are recommended as they are believed to reduce the required size of the supercell and reduce the dependence of the results on the DFT functional[425].

Satisfying the Generalised Koopmans' Theorem The degree of HF exact exchange mixing for a hybrid functional (α) is often derived from empirical data. When considering other criteria with which to select the value of this parameter, it has been suggested[426,318] that the *Generalised Koopmans' Theorem* should be satisfied, wherein:

$$-I(N+1) = E(N+1) - E(N) = \varepsilon_{N+1}(N+1) \tag{4.2}$$

where I is the ionisation energy and E is the total energy for a system of N electrons, and ε_i is the i^{th} eigenvalue of the N electron system. As such, if considering other values of α for further study, it would be useful to use the suggested method to compute a set of *Koopmans-compliant* polaron states and compare their results with standard functional values of α.

Comparing the eigenvalues with bulk calculations To further understand how the self-trapping

process changes the material's electronic structure, it would be useful to consider the depth of the trap state relative to the bulk valence and conduction bands. Such data could be used to generate a diagram, similar to the Shockley-Read-Hall recombination diagram in Figure 4.1 for a database of candidate materials. This would be useful for visualising the relative depths of the trap states for many materials. However, this would require additional bulk calculations and postprocessing measures, alongside the band state and polaron state calculations, which would increase the computational cost.

Catalytic Activity

5.1 Introduction

As mentioned in the first chapter, another challenge concerning photoelectrochemical cell electrodes is their long-term corrosion resistance, catalytic activity, and catalytic selectivity. These three challenges may be considered to be interdependent because such corrosion reactions may be considered to be an unwanted side reaction that further reduces the overall catalytic activity.

An example of this issue is most pertinent for gallium arsenide photoelectrodes that facilitate the oxygen evolution reaction. The GaAs surface dissolves following the photogeneration of holes on the material's surface. However, a recent study[239] was able to demonstrate that the corrosion process may be mitigated by using a corrosion-resistant material as a passivation layer, TiO_2, on top of the more catalytically active material. Similar studies for the cadmium chalcogenides (CdS, CdSe, CdTe) have also been reported using an amorphous layer of TiO_2[427]. Whilst this may present a future solution to the corrosion problem, it remains desirable to create an electrode material that is corrosion-resistant itself or at least has a more catalytically active passivation layer.

In this regard, the reason TiO_2 serves as an effective passivation layer is due to its intrinsic corrosion resistance[428]. Bulk titanium metal is not corrosion resistant, but its surface is corrosion resistant due to the naturally formed TiO_2 oxide layer that forms on the surface of titanium metal when exposed to air. Moreover, the rutile phase is believed to be more corrosion-resistant than the anatase phase[429].

5.1.1 Rutile Alloys

Inconveniently, the highest-performing rutile materials for OER catalysis also happen to be based on rare minerals. IrO_2 and RuO_2 have received particular attention for their catalytic activity[59]. As such, it is desirable to find alternative materials that have comparable catalytic activity with substantially lower concentrations of IrO_2 and RuO_2. One means by which this has been proposed is to use a homogeneous *alloy* of more abundant rutile-forming compounds[59]. An alloy may be described as a mixture of two or more metals. Hence, strictly speaking, an alloy does not possess a crystal structure because there is no long-range order to the species occupancy of the lattice sites. As such, a homogeneous alloy is one in which the mixed species are evenly dispersed throughout the material in a randomised fashion. In this regard, an ongoing challenge is the selection of a particular set of species to mix at a particular composition. Such alloys may be synthesised, for example, via *sputter deposition* and their catalytic activities may be assessed via cyclic voltammetry[430]. The homogeneity of each mixture may be determined via inductively coupled plasma mass spectrometry[431]. To accelerate the selection of particular compositions, it would be desirable to screen candidate compositions according to their likelihood of homogeneity and their catalytic activity.

5.1.2 Simulating Alloy Structures

Given their disordered nature, classifying alloy materials within the framework of a crystal structure is difficult to standardise. Nevertheless, this can be achieved using modern computational tools and to an acceptable degree of accuracy. When simulating alloys, it is well understood that as the size of the unit cell increases, and the site occupancies are suitably randomised, the thermodynamic properties of the *true* alloy are more closely represented[432]. This convergence procedure is similar to that of the plane wave cutoff energy or **k**-point grid density. Indeed, a robust method of calculating alloy properties is by systematically increasing the size of a supercell, varying the site occupancies in multiple permutations, and calculating the thermodynamic properties until they are converged within a desired tolerance. Whilst this approach is automatable within a computational framework, the issue with such a method is that the computational overhead is relatively large because many atoms are required for each unit cell and many permutations of the system need to be calculated. Furthermore, since the supercell is still periodic, local correlations between nearby lattice sites and their periodic images can cause further deviations from the properties of the *truly random* alloy[433,432]. The overall effect is that the convergence of the randomised supercell size is very slow for the thermodynamic properties.

However, *Special Quasi-random Structures* (SQSs) are a notably accurate and efficient alternative for an estimation of the thermodynamic properties of a homogeneous alloy[432].

5.1.3 Special Quasi-random Structures

The SQS method[432] circumvents the local correlation problem by increasing the unit cell size via a non-uniform cluster expansion. Instead of merely adding more unit cells to the supercell to increase the accuracy, configurable clusters of lattice sites (doublets, triplets, quadruplets, etc) are selectively added and occupied with a particular species such that the local randomness per lattice site is maximised. The key assumption of the SQS method is that the local correlations between nearby lattice sites have a greater effect on the thermodynamic properties than those at distant lattice sites. Thus, by optimising the local *randomness* around each lattice site, a more accurate estimation of the thermodynamic properties of a truly random alloy can be reached with a much smaller unit cell (See Section B.11 for the mathematical arguments behind the SQS methodology). A popular means with which to efficiently assign site occupancies is via a *Monte-Carlo algorithm* according to Equation[434,433] 5.1.

$$Q = -\omega L + \sum_{\alpha} |\rho_{\alpha} - \rho_{\alpha}^{target}| \tag{5.1}$$

where Q is the *Objective Function* to be minimised during the Monte-Carlo algorithm, ω is a user-specified weighting for the class of cluster (doublets, triplets, quadruplets, etc), α is the cluster index that denotes the cluster in question, ρ_{α} represents the correlations of α, ρ_{α}^{target} represents the *target* correlations of the *truly random* configuration, and L is a user-specified radius such that all clusters within L are optimised ($|\rho_{\alpha} - \rho_{\alpha}^{target}| = 0$). A Monte Carlo algorithm is a probabilistic way to evaluate integrals across complicated boundaries. As opposed to evaluating said integral across an evenly spaced grid of points, points for evaluation are selected randomly. If a randomly selected point lies inside the boundary, it is retained. If it lies outside of the boundary, it is discarded[16]. As such, ω and L must be carefully chosen to optimise the cluster selection criteria for accurate results.

Conveniently, a suitably optimised SQS calculation at a particular composition is insensitive to the chemical species and thereby provides a template crystal structure for that particular composition. Furthermore, a database of SQS template structures[435] already exists for a 50:50 binary mixture of rutile compounds. Further composition ratios require the calculation of a separate SQS.

5.1.4 Estimating the Thermodynamic Properties with Density Functional Theory

Upon generation, geometry optimisation, and electronic minimisation of an SQS with DFT, the *Gibbs Free Energy* (ΔG_{mix}) and *Critical Temperature* (Tc_{mix}) *of Mixing* can be approximated by considering the ideal *Configurational Entropy of Mixing*, ΔS_{mix}, of two ideal gases at a constant temperature and pressure:

$$\Delta S_{mix} = -nR[x_1 ln(x_1) + x_2 ln(x_2)] \tag{5.2}$$

Where n is the *amount of substance*, R is the ideal gas constant, and x_i represents the ratio of each mixture component (i.e. $x_1 + x_2 = 1$). Since for a 50:50 binary mixture $x_1 = x_2 = 0.5$, the above equation can be reduced to:

$$\Delta S_{mix} = nRln(2) = Nk_B ln(2) \tag{5.3}$$

where N is the number of AO_2 formula units in the SQS unit cell and k_B is the Boltzmann constant. ΔG_{mix} can then be calculated via the standard relation:

$$\Delta G_{mix} = \Delta H_{mix} - T\Delta S_{mix} \tag{5.4}$$

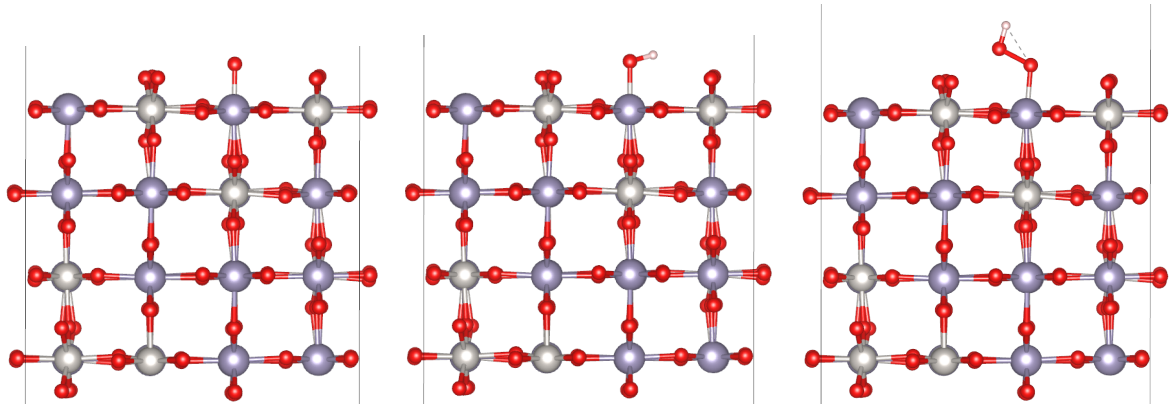
Where ΔH_{mix} is the *Enthalpy of Mixing* and T is the temperature. ΔH_{mix} is calculated according to Hess' law[436], by combining the total energies from DFT calculations of the SQS system and its corresponding pure systems. Correspondingly, Tc_{mix} is estimated via Equation 5.5:

$$Tc_{mix} \approx \frac{\Delta H}{\Delta S} \tag{5.5}$$

5.1.5 Estimating the Catalytic Activity with Density Functional Theory

The major efficiency limiting step to the catalytic activity of photoelectrochemical water splitting is believed to be the oxygen evolution reaction. In particular, the kinetics that occur at the OER electrode[59]. To estimate the catalytic activity with DFT and further understand the reaction kinetics, a

slab model of the desired surface needs to be constructed and optimised. The adsorbates (intermediates) involved in the OER are then added to the slab in the appropriate binding configuration, and the corresponding energy change upon adsorption provides the adsorption energy. An illustration of such OER slab models is given in Figure 5.1:



(a) Oxygen adsorbed. **(b)** Hydroxyl adsorbed. **(c)** Hydroperoxyl adsorbed.

Figure 5.1: Examples of the structural output for each intermediate for the OER adsorption calculation.

The adsorption energies can then be used to estimate the catalytic activity via a *volcano plot*. The volcano plot is so named because of its triangular shape and increasing catalytic activity as the *peak* is approached. Such plots were first introduced by Balandin[437,438], and rely principally on the *Sabatier Principle*[439]. Accordingly, compounds that fall on the right of the volcano plot bind the adsorbate too strongly, thereby making the desorption of the adsorbate rate limiting. Compounds that fall on the left-hand side of the volcano plot bind the adsorbate too weakly, thereby making the adsorption of the adsorbate rate limiting. Compounds at the peak of the volcano have the expected *ideal* binding strength with the surface and are expected to maximise the reaction rate. Therefore, the volcano plot provides a convenient means to compare the effectiveness of potential photoelectrochemical cell electrodes.

The Universal Scaling Relation

The most convenient general volcano plot form for the OER was introduced by Man et al[58]. Here, a particular step of the OER reaction pathway was found to be sufficient to describe the overall catalytic activity for a particular compound.

Step 2: $\mathrm{M(OH)} \longrightarrow \mathrm{M(O)} + \mathrm{H^+} + \mathrm{e^-}$

The corresponding thermodynamic relation is given by the difference in the Gibbs free energy in the oxygen adsorption step (ΔG^O) and the hydroxyl adsorption step (ΔG^{OH}). Furthermore, it was found that plotting the adsorption energy of the hydroperoxyl adsorption step (ΔE^{OOH}) and the hydroxyl adsorption step (ΔE^{OH}) produced a *universal scaling relation* for the OER:

$$\Delta E^{OOH} = \Delta E^{OH} + 3.2eV \tag{5.6}$$

This was because it was observed empirically that the difference in the hydroxyl and hydroperoxyl adsorption energies was almost constant (3.2 eV) and independent of the overall binding strength to a particular surface. The rationale for this observation, provided by Man et al[58], was that both the hydroxyl and hydroperoxyl intermediates bind to the surface via a single oxygen atom. Thus, from the perspective of the surface, hydroxyl, and hydroperoxyl intermediates undergo near-identical chemical interactions. Moreover, it was further observed that the adsorption energies of all three OER intermediates changed in a correlated manner with the applied potential (i.e. $\Delta G^{OH} < \Delta G^O < \Delta G^{OOH}$), regardless of the binding strength of the particular surface. This understanding of the correlations between adsorbate binding energies allowed the individual steps in the water splitting reaction to be reasonably represented by a single quantity:

$$\Delta G^O - \Delta G^{OH} = E^O - E^{OH} + 0.5E^{H_2} + 0.41 \tag{5.7}$$

Where E^{H_2} is the total energy of a hydrogen molecule, calculated via DFT, and 0.41 is a zero point energy correction, as suggested in Man et al's work[58] (see Section B.12 for a full derivation of the above equation). The volcano plot can then be constructed via the reaction overpotential (η) according to:

$$\eta = max[\Delta G^O - \Delta G^{OH}, \Delta G^{OOH} - \Delta G^O]/e - 1.23V \tag{5.8}$$

Which is further approximated using the universal scaling relation of Equation 5.6:

$$\eta \approx max[\Delta G^O - \Delta E^{OH}, 3.2 - \Delta G^O - \Delta E^{OH}]/e - 1.23V \tag{5.9}$$

The $''max''$ term indicates a selection of either $\Delta G^O - \Delta G^{OH}$ or $3.2 - \Delta G^O - \Delta G^{OH}$, whichever is larger. This distinction determines which side of the volcano the compound will fall upon when plotted; with $\Delta G^O - \Delta G^{OH}$ representing the right hand (overbinding) side and $3.2 - \Delta G^O - \Delta G^{OH}$ representing

the left hand (underbinding) side. The volcano is considered in terms of ΔG^O and ΔG^{OH} because the difference between ΔE^{OH} and ΔE^{OOH} is assumed to be constant by virtue of the aforementioned universal scaling relation. Consequently, the oxygen adsorption energy (ΔE^O) determines the variation in the required overpotentials for different surfaces. Therefore, it is only the difference in the binding energy between ΔG^O and ΔG^{OH} that needs to be calculated to obtain a complete (approximate) understanding of the binding characteristics for a given system. In other words, ΔE^{OOH} can always be estimated from ΔE^{OH}, via Equation 5.6, and it can be assumed that $\Delta G^{OH} < \Delta G^O < \Delta G^{OOH}$. Thus, only ΔG^O and ΔG^{OH} need to be calculated to compare the OER across different surfaces.

5.2 Motivation and Materials of Interest

With this in mind, computational materials screening could be used as a means for rapid materials discovery for the calculation of the thermodynamic and adsorption properties of multiple candidate compounds. Specifically, the feasibility of the homogeneity and the catalytic activity could be estimated and used to promote or exclude each candidate. The SQS method could be used to generate a bulk unit cell, followed by the rapid slab generation using existing computational tools, the thermodynamic properties and catalytic activity could then be investigated for several different compounds using the same SQS template structure. However, literature using SQS unit cells for surface slab models is sparse. This may be due to the difficulty in transforming the highly disordered SQS unit cell into a near-cubic slab model. However, this has been achieved by Moses et al[440] and Wu et al.[441] for band alignment. In addition, a more recent study by Svane et al[59], has stressed that the catalytic activity of rutile alloys requires further study, though did not use the SQS method to generate surfaces for their catalysis investigations. They instead used the traditional randomised occupancy generation approach, as described previously. It is noted that this was not a deficiency of Svane et al's[59] work because the emphasis was on the particular effects of specific surface site occupations for the OER. As such, a greater number of possible configurations were required for their study, as opposed to a reasonable average that would be offered by an SQS model.

For the remainder of this chapter, it is demonstrated that reasonably dimensioned slab models may be rapidly generated from SQS unit cells for rutile mixtures via an appropriate unit cell transformation. This allows the same cell to be used to investigate both bulk and surface properties. If accurate in both respects, this would result in a large reduction in the computational overhead and implementation because only a single unit cell would be needed to understand both the bulk and surface properties.

However, a remaining question is whether or not the SQS-generated surface slab models produce reasonable estimations of the catalytic activity of a given compound. In this regard, the topical and hypothetically corrosion-resistant, rutile alloys are a good starting point for such an investigation.

The following rutile-type materials were considered in this work, all of which have the presumed corrosion-resistant rutile ($P4_2/mnm$) structure. Furthermore, when selecting appropriate candidates, abundance is a key consideration. In this regard, the below rutile compounds are listed in ascending order of their abundance within the Earth's crust (in mg kg^{-1}):[1] Ir (0.001), Ru (0.001), Pt (0.005), Ge (1.5) , Sn (2.3), Pb (14), Mn (950), Ti (5650). These abundance values do not consider the feasibility of mining and processability. Therefore, a more in-depth techno-economic analysis would be required in due course. Nevertheless, given the large difference in abundance between platinum and germanium, it would be desirable to find a competitive alloy of germanium, tin, lead, manganese, or titanium. Ruthenium, iridium and platinum are included in this chapter as control samples. Other rutile-forming materials such as Osmium (0.0015 mg kg^{-1}) and Rhodium (0.001 mg kg^{-1}) have also been considered in past literature[59] but are not considered here due to their comparable abundance to Ru, Ir, and Pt[1] and would therefore merely serve as additional control compounds.

Iridium dioxide (IrO_2) is a metallic compound with an antiferromagnetic magnetic structure[442]. It has received a great deal of attention as a photoelectrochemical cell electrode for its exceptional chemical stability and catalytic activity for the OER. The most important issue concerning this compound is the low abundance of iridium. It is insoluble in water[443].

Ruthenium dioxide (RuO_2) is a metallic compound[444,445] or narrow, direct band gap semiconductor whose band gap requires further investigation[446,447]. It is believed to have an antiferromagnetic structure[448]. It has a high catalytic activity for the OER. However, over time, the ruthenium ions are further oxidised and dissolve into the solution. As such, the long-term stability of ruthenium dioxide is not sustainable under photoelectrochemical cell operating conditions[449]. The low abundance of ruthenium further compounds this issue. It is otherwise insoluble in water[450].

Platinum dioxide is a semiconductor with an indirect band gap[331] of 1.30 eV - 1.47 eV[451]and has a slightly distorted (Pnnm) rutile crystal structure[452,331]. It has a high OER catalytic activity, both alone[58] and in combination with other materials[453]. However, like ruthenium and iridium, platinum is still relatively rare. It is insoluble in water[454].

Germanium dioxide is a semiconductor with a direct[455] band gap of 4.68 eV[456]. It has been considered as a photoelectrochemical cell electrode material, due to its high charge carrier mobilities and

high charge separation rate.[455] However, it may encounter challenges due to its slight solubility in water[457]. It is much more abundant than the aforementioned compounds.

Manganese dioxide is a semiconductor with a direct[458] band gap of 1.08 eV[459]. It is believed to have a helical magnetic structure with antiferromagnetic ordering[459]. However, the helical magnetic structure is often disputed in other publications, in favour of a more conventional magnetic structure that is similar to IrO_2 or RuO_2[460]. Manganese dioxide benefits from being in high abundance. However, the rutile phase has received little attention due to its relatively low OER catalytic activity. However, the layered ($R\bar{3}m$) structure is believed to have a much higher intrinsic catalytic activity[461]. It is insoluble in water[462].

Lead dioxide is a semiconductor with a direct band gap of 0.61 eV[463]. However, it has a metal-like conductivity which is believed to be due to the natural deficiency of oxygen within the real crystal[93]. As such, it has attracted considerable attention as a highly conductive, low cost and chemically stable water splitting OER catalyst and as a catalyst for the generation of ozone. A great deal of research is now focused on elucidating the ozone and oxygen generation mechanisms[75]. Once understood, further steps could be taken to improve the catalytic selectivity of this material for the OER. A foreseeable issue with using this material would be the known toxicity of lead[464]. It is insoluble in water[465].

Tin dioxide is a semiconductor with a direct band gap of 3.6 eV[463,466]. It has been considered a photoelectrochemical cell photocatalyst for its photocatalytic activity, chemical stability, and UV absorption capabilities[467,468] A foreseeable issue with using this material would be the ongoing debate regarding the toxicity of tin compounds[469]. It is insoluble in water[470].

Titanium dioxide was introduced in the previous chapter.

5.3 Methods

5.3.1 SQS Structure Optimisation

An SQS template for the rutile structure was sourced from the *Alloy Theoretic Automated Toolkit* (ATAT) package[471,472,473,474]. The ATAT *sqs2tdb* function was used to generate an SQS unit cell for each compound.

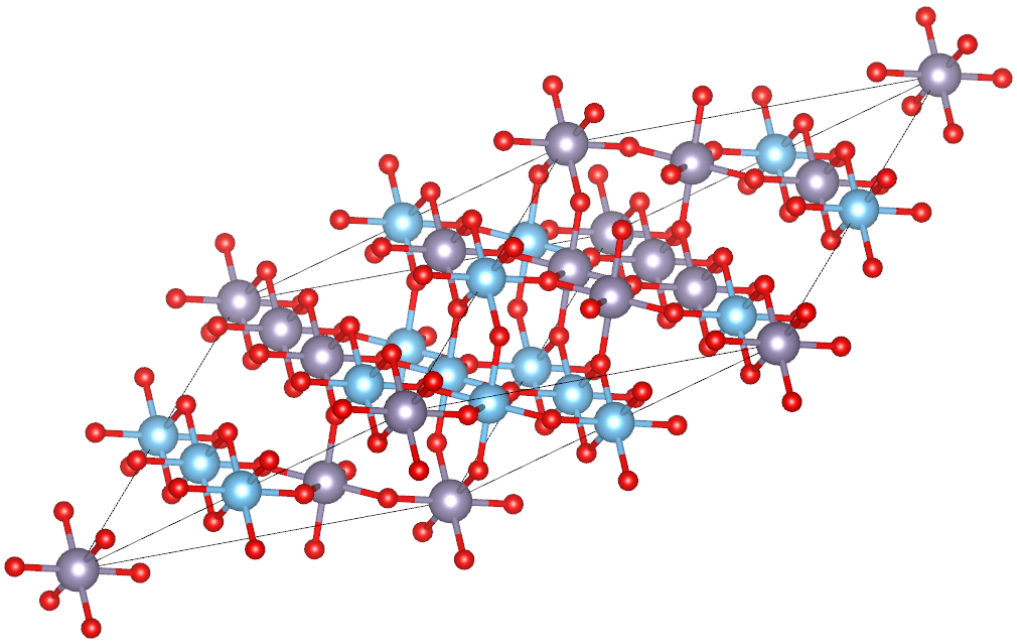


Figure 5.2: 50:50 Rutile SQS template unit cell. This unit cell contains sixteen AO_2 formula units.

Each SQS structure and pure AO_2 compound was relaxed with the PBEsol functional, a 500 eV plane wave cutoff energy, and a 13x9x9 Γ-centred **k**-point grid. In contrast to previous chapters, the unit cell shape and internal parameters were kept fixed across a range of volumes, before fitting to the Murnaghan Equation of State[280]. The reason for keeping the cell shape and internal parameters constant for a series of fixed volumes was that the SQS unit cells would often undergo substantial changes in the cell shape and internal parameters at certain cell volumes. This would cause the total energies to deviate substantially from the expected parabolic shape, making the equation of state fitting sub-optimal for finding the equilibrium volume. Conversely, fixing the unit cell shape and internal parameters resulted in more reasonable Equation of State fitting[a]. Subsequent relaxation of the volume, shape and internal parameters yielded structures that were equal to or lower in energy than the standard geometry optimisation procedure shown in previous chapters. This is demonstrated for $Ti_{0.5}Pt_{0.5}O_2$ and $Sn_{0.5}Pt_{0.5}O_2$ in Figure 5.3 and is supported by the findings of Wu et al[441]. The

[a]Determined by the R^2 of the Equation of State fit.

optimal volume was then used as a starting point for a simultaneous relaxation of the volume, shape and internal parameters of the unit cell.

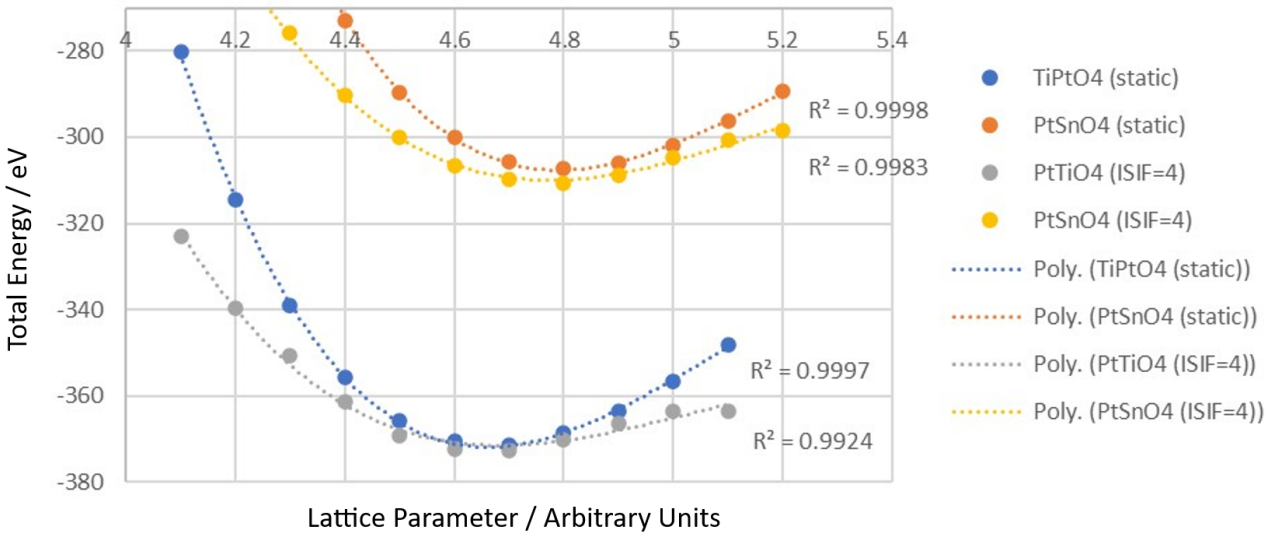


Figure 5.3: Plot of the volume scans for $Ti_{0.5}Pt_{0.5}O_2$ and $Sn_{0.5}Pt_{0.5}O_2$, wherein the internal parameters are held constant (static) and where they are allowed to relax at each volume (ISIF = 4 in the VASP syntax). Polynomial fits are also included to distinguish the 4 plots.

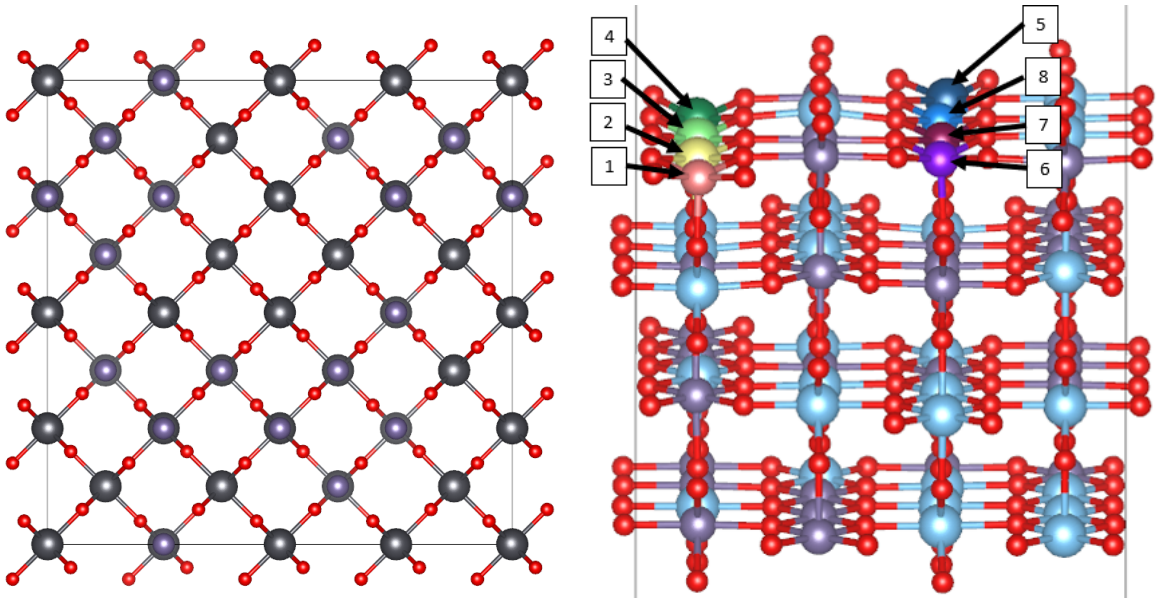
5.3.2 Thermodynamic Property Calculations

Following this, a PBE electronic minimisation was performed on the output structure to obtain ΔH_{mix}. The SQS template contained sixteen AO_2 formula units. Thus, via Equation 5.3, $S_{mix} = 0.956 meV K^{-1}$ and, an assumed room temperature of $300K$, $T\Delta S_{mix} = 287 meV$. These quantities were used to calculate ΔG_{mix} and T_c via Equations 5.4 and 5.5 respectively.

5.3.3 SQS Slab Generation

Given the non-cubic and non-symmetric shape of the SQS template, converting the structure into a slab was not trivial. In particular, the initial SQS structure that was used for the mixing energies would often produce spurious slabs that were substantially slanted across the unit cell and converged very slowly. To make the SQS unit cell more workable with automated slab generation software, expanding the SQS supercell with the ASE supercell optimisation tool[409], and then using the Pymatgen slab model generator[475,476] on the output, proved to generate more reasonable slab models. A corresponding slab model that was generated via this method is given in Figure 5.4b. It is noted that convergence of the adsorption energy to 0.01 eV with the slab width was much more difficult when compared to the pure systems. This may have been due to the lack of symmetry and long-range order of the mixed

system, causing asymmetries in the electrostatic potential, and thus limiting convergence. Whilst it is conceivable that convergence may be achievable to 0.01 eV, such slab models were expected to be impractically wide for materials screening. Indeed, the SQS method specifically aims to limit any such long-range order. As such, a lower degree of precision was deemed acceptable. An adsorption site numbering scheme was also used to track the variation across different adsorption sites. This is given in the below figure.



(a) **(b)**

Figure 5.4: (Left) Transformed 8x supercell using the ASE supercell optimisation tool. (Right) Illustration of the slab adsorption site numbering.

Following this, the *Pymatgen* slab model generator[475,476] was used to produce slabs of the non-cleaved rutile oxide 110 surface. This is believed to be the most chemically and structurally stable surface for rutile catalysis[477]. Each of these slabs was four cation layers thick with a 50 Å vacuum layer. These slab dimensions were per the work of Svane et al[59]. The calculations also included standard dipole corrections[478,479] which ensured that convergence of the pure system adsorption energies was within 0.01 eV. Correspondingly sized mixed slab systems achieved convergence between 0.1 - 0.05 eV. The dipole corrections were based on the work of Makov et al[480] and Neugebauer et al[481], which cancel out undesired dipole interactions between each surface and its periodic images, such that convergence is reached for narrower slab widths. The slab was relaxed using PBEsol, a 500 eV plane wave cut-off energy, and a 13x9x9 Γ-centred **k**-point grid, keeping the bottom two layers fixed in place. In separate calculations, an adsorbate was positioned above each of the 8 adsorption sites (see Figure 5.4b) at approximately the bulk metal-oxygen atom bond length. The structural relaxation was then repeated with the PBEsol functional, and the total energy was extracted following a PBE electronic minimisation.

For the $Ru_xTi_{(x-1)}O_2$ 25:75 mixture compositions, the *Integrated Cluster Expansion Toolkit*[33] (ICET) package was used to generate a corresponding SQS[b]. The same inputs for the 50:50 ATAT SQS (ω, L, etc.) were applied to the 25:75 ICET calculations. A 50:50 ICET SQS was further generated for comparison. The thermodynamic properties and catalytic activities of the ICET models were calculated as above.

[b]ICET was chosen for its ease of integration with postprocessing tools (mainly based on Python code).

5.4 Results and Discussion

5.4.1 Thermodynamic Properties

For conciseness, a shorthand is utilised in this chapter to describe each of the mixtures discussed in the following sections. Where a compound AB is mentioned, this refers to an $A_xB_{(x-1)}O_2$ mixture. The calculated thermodynamic properties for each of the 50:50 rutile mixtures are given in Figure 5.5. Via the colour scale, it can be seen that each of the quantities (ΔH_{mix}, ΔG_{mix}, and Tc_{mix}) are strongly correlated. Moreover, the lower the value of each quantity, the more facile the expected synthesis of the alloy. This is because each of these quantities is associated with an energetic barrier to homogeneous mixing. Thus, homogeneous mixtures of IrRu, MnRu and PtRu appear to be the most synthesisable, whilst MnSn, GePb and RuSn are among the least synthesisable. Despite having the lowest Gibbs free energy of mixing, MnPb had to be excluded from further study because the relaxed structure was distorted such that it no longer corresponded to the rutile crystal structure. Consequently, rutile 110 surface slabs could not be generated for MnPb. This was not an expected result and was only identified by visual inspection which is a clear limitation of the suggested screening procedure. This limitation may be alleviated by incorporating a symmetry analysis step following the structural relaxation of the SQS unit cell. Such a symmetry analysis could be performed automatically by *spglib* [482], which would confirm that the relaxed structure still corresponds to the rutile crystal structure. Another observation was that mixtures containing Pt also had consistently low energies, regardless of the species Pt was mixed with. However, given the relative scarcity of Pt, it is not an ideal candidate for a low-cost photoelectrochemical cell electrode. The Tc_{mix} values below 0K are attributed to Tc_{mix} being merely an approximation, not a robust and absolute metric.

ΔH(mix) / meV

	Ir	Mn	Pb	Pt	Ru	Sn	Ti
Ge	97	126	233	47	110	107	52
Ir		6	77	42	-31	140	95
Mn			-109	49	-26	345	163
Pb				30	115	45	164
Pt					-8	61	19
Ru						191	133
Sn							84
Ti							

ΔG(mix) / meV

	Ir	Mn	Pb	Pt	Ru	Sn	Ti
Ge	-190	-161	-54	-240	-177	-180	-235
Ir		-281	-210	-245	-318	-147	-192
Mn			-396	-238	-313	58	-124
Pb				-257	-172	-242	-123
Pt					-295	-226	-268
Ru						-96	-154
Sn							-203
Ti							

Tc(mix) / K

	Ir	Mn	Pb	Pt	Ru	Sn	Ti
Ge	101	132	244	49	115	112	55
Ir		6	80	44	-33	146	100
Mn			-114	51	-27	361	171
Pb				31	120	47	171
Pt					-8	64	20
Ru						200	139
Sn							88
Ti							

Figure 5.5: Thermodynamic quantities for each of the 50:50 rutile mixtures. The row and column for a particular cell give the corresponding chemical species. The colour scale indicates the relative magnitude of the quantity in each cell which reflects the relative likelihood of each composition to mix homogeneously; green being a high likelihood of mixing and red being a low likelihood of mixing.

5.4.2 Adsorption Energies

As can be seen from Figure 5.6, the majority of the data adheres to the universal scaling relation of Equation 5.6. The strongest outliers are attributed to inconsistencies during the structural relaxation. Such inconsistencies included an alternative binding of the adsorbate, non-binding of the adsorbate, or dissociation of part of the adsorbate after binding. In all cases, different bonds were being broken and formed compared to the single bond formed in Figure 5.1. Moreover, those that did not could be explained as being aberrations from the usual adsorption configuration (see Figure 5.7 for an example).

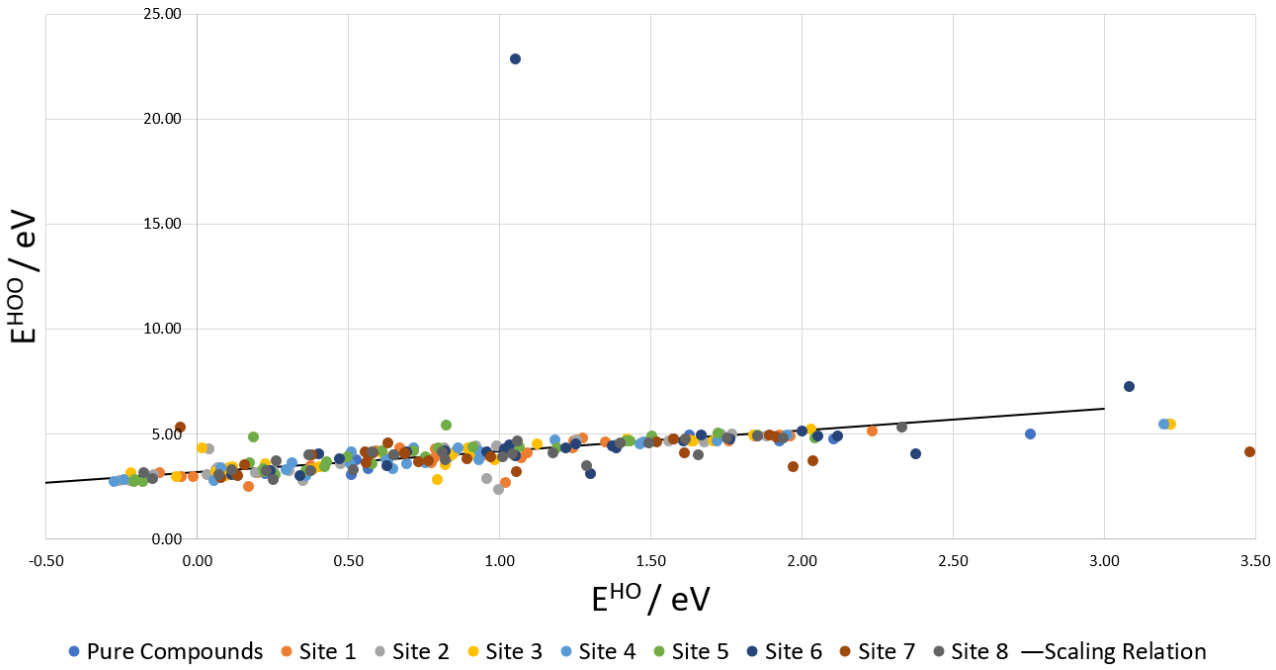
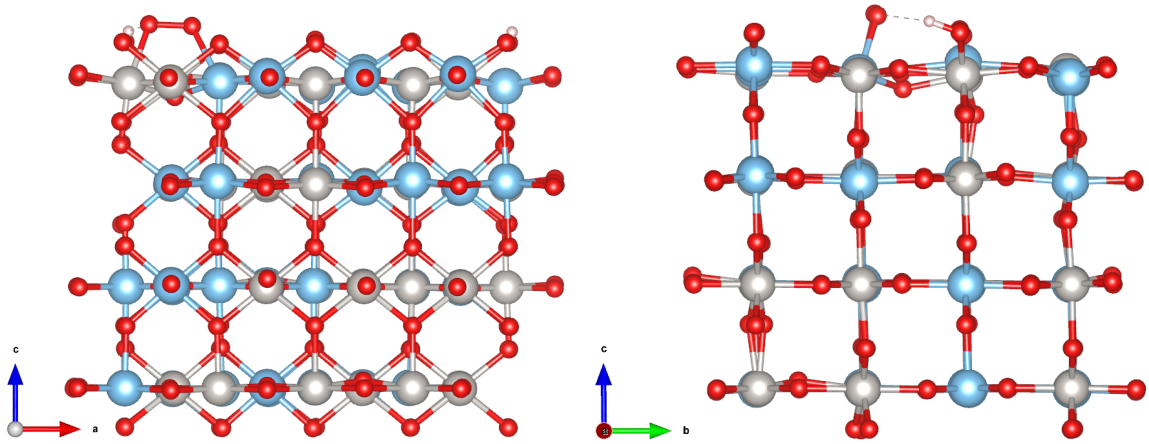


Figure 5.6: Plot of ΔE^{OOH} versus ΔE^{OH} for each of the 50:50 rutile mixtures. The black line represents the universal scaling relation of Equation 5.6. The *"Site"* numbering corresponds to the numbering of each adsorption site as presented in Figure 5.4b. This figure shows that the majority of the data from this work is in accordance with the aforementioned universal scaling relation.

This would have altered the resultant total energy of the adsorption process, creating the outlier energy. Despite these inconsistencies, the outliers themselves may not be unrepresentative of the OER process. On the contrary, they may be representative of an alternative OER reaction mechanism. However, Man et al's universal scaling relation was not established with such mechanisms in mind. Furthermore, structural consistency would make the screening procedure more expedient and easier to analyse. For future work, a possible means to prevent such outliers from occurring would be to restrict (freeze) the adsorbate ions, similarly to the frozen slab layers, along all dimensions except the c-axis (see Figure 5.7). Thus, during the structural relaxation of the system, the atoms of the adsorbate may only be lowered or raised from the slab, making them much less likely to bind to adjacent atoms on the slab. This may also have the added benefit of further reducing the calculation time because there would be fewer degrees of freedom available during such a calculation.



(a) Angle 1 **(b)** Angle 2

Figure 5.7: The corresponding structure for the outlier at approximately 23 eV on the E^{HOO} axis in Figure 5.6, presented at two different angles. It can be seen that the oxygen atoms of the hydroperoxyl adsorbate have each bound to a separate metal ion as opposed to a single oxygen atom binding to a single metal atom. Moreover, the hydrogen atom appears to have dissociated from the adsorbate and has bound to an adjacent bridging oxygen atom of the slab model.

Figure 5.8 gives the volcano plot data of Figure 5.6, including the outliers, using Equation 5.9. From left to right of Figure 5.8, the ordering of the pure compounds agrees well with that of Man et al[58]: (MnO_2, IrO_2, RuO_2, PtO_2, TiO_2 and SnO_2). The rearrangement of IrO_2 and RuO_2 may be due to the differences in the calculation parameters. In particular, Man et al. used the *Revised Perdew Burke Ernzerhof*[483] (RPBE) functional, whereas the *Perdew Burke Ernzerhof* (PBE) functional was used here. Indeed, the supplementary material of Svane et al's work demonstrated that a similar rearrangement of the activity of IrO_2 and RuO_2 could be achieved by using a different functional[59]. This may also explain why RuO_2 and PtO_2 are closer to the peak of the volcano in Man et al's data[58].

Regarding the mixed systems, each of the mixtures is situated on the right-hand side of the volcano, indicating an overbinding of the adsorbate. However, a notable exception was GePb. Depending on the particular site of adsorption, GePb sat on either the left or right-hand side of the volcano and this is illustrated via the error bars for GePb in Figure 5.8. Descending from the peak of the volcano, the most catalytically active mixtures are: GePb, MnRu, IrRu, GeIr, PbRu and MnTi. Ascending from the bottom of the volcano, the least catalytically active mixtures are GeSn, PbSn, SnTi, GeTi, PbTi and IrMn. Besides GePb, mixtures containing Ir and Ru were expected to be the most effective catalysts due to their superior performance as pure compounds. Using the same reasoning, compounds containing Sn and Ti were among the least effective catalysts.

Compounds containing Ge and Pb were among both the best and worst performing compounds depending on the species that they were combined with. This may indicate an ability of Ge and Pb to

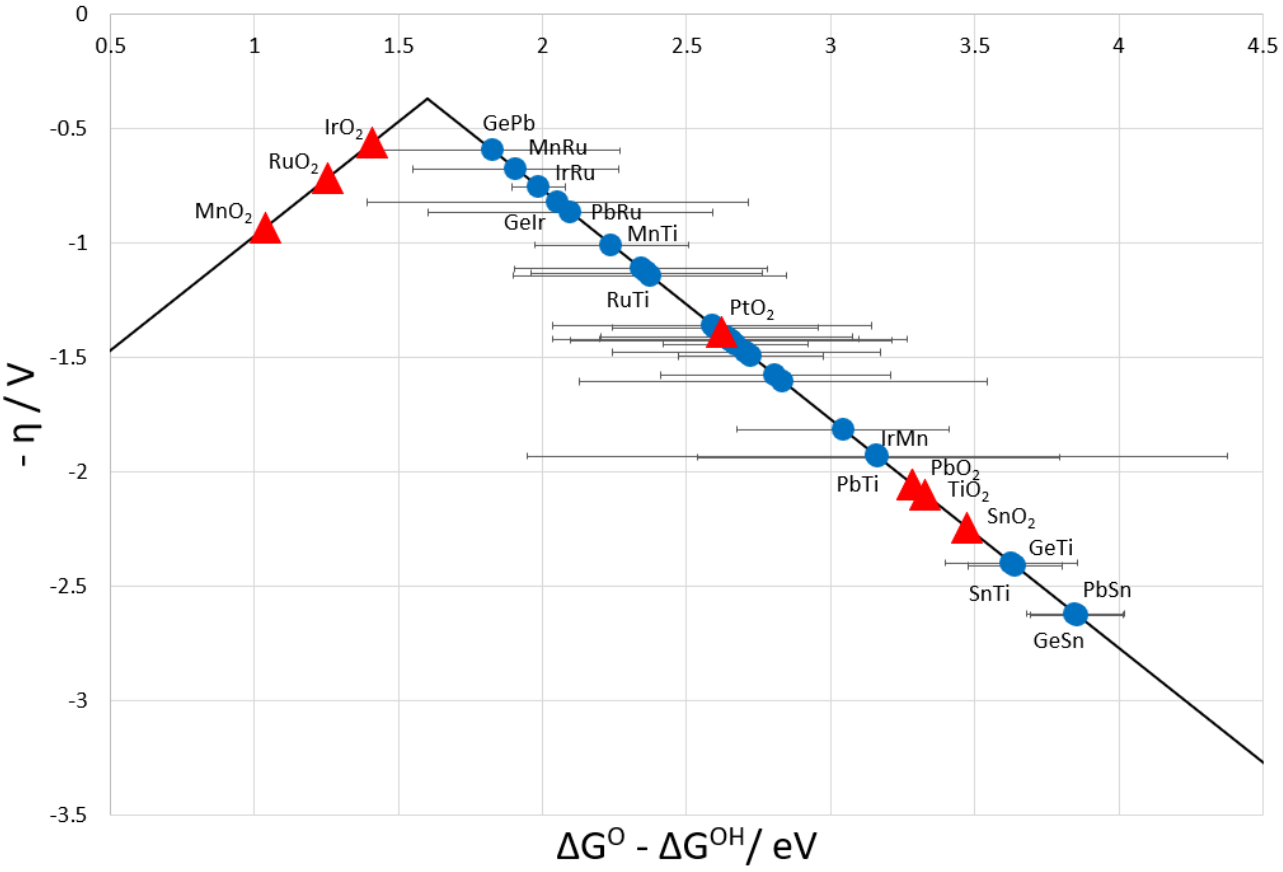


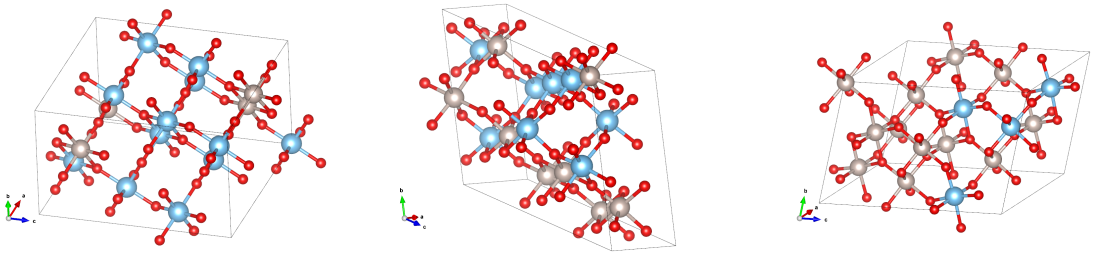
Figure 5.8: Volcano plot of the adsorption energies from this work. The black line is a linear plot of η to define the volcano. The red triangles correspond to the pure compounds and the blue circles correspond to the average adsorption energies of sites 1-8 for each of the mixtures (including the aforementioned outliers of Figure 5.6). The horizontal error bars correspond to the standard deviation of the mean for each mixture's adsorption energy.

passively maintain the catalytic activity. Therefore, they may be good candidates for an inert scaffold for more active compounds, such as RuO_2 or IrO_2. Given its position on the volcano and its abundance, GePb may seem to be a near-ideal candidate for further study. However, its high enthalpy of mixing (see Figure 5.5) may make a homogeneous synthesis prohibitive. Consequently, MnTi may prove a more promising compromise: being the sixth mixture furthest from the peak of the volcano, containing relatively abundant elements, and having a substantially lower mixing energy than that of GePb. Moreover, a recent study by Zhou et al[430] has highlighted the enhanced activity of Mn and Ti when used together in mixtures of four and five-component mixtures of Mn, Sb, Sn, Ti and O.

5.4.3 Intermediate Compositions

Whilst relying on the existing ATAT database, the above method was limited to 50:50 compositions. As such, it would be desirable to have a method in which intermediate compositions can be generated and tested. It is further noted from Figure 5.8 that, in many cases, the error bars demonstrated a

greater degree of variation, per adsorption site, than the 0.1 eV upper limit to the precision of the mixed system adsorption energies. As such, data across different compositions and an appropriately wide range of catalytic activities may still provide qualitative information, despite the aforementioned limit to the precision.



(a) 25 % Ru **(b)** 50 % Ru **(c)** 75 % Ru

Figure 5.9: SQS structures for the RuTi intermediate compositions generated via ICET[33].

To this end, the ICET package was used to generate SQS structures of the RuTi mixture with 25% 50% and 75% Ru. RuTi was chosen for this investigation because Ru and Ti are well-studied within the field and the 50:50 composition data presented minimal outliers. Furthermore, Ru and Ti each represent opposite extremes in catalytic activity. Thus, providing the maximal possible range within which to distinguish the effects of each component. If components of similar reactivity (i.e. Ru and Ir) were to be used, it may have been difficult to distinguish the performance of particular compositions from the error margin across different adsorption sites. For these calculations, the same benchmarks as the ATAT database structures were used. However, exhaustive enumeration of the SQS structures was used, as opposed to Monte-Carlo algorithm generation. This was because the computational expense was manageable using the available computational resources. This produced the structures in Figure 5.9. Their thermodynamic properties are given in Table 5.1.

Mixture	$\Delta H(mix)$	$\Delta G(mix)$	$Tc(mix)$
RuTi(ATAT)	132.62	-154.38	138.73
RuTi(25perc)	106.68	-466.32	55.82
RuTi(50perc)	132.41	-154.59	138.50
RuTi(75perc)	103.79	-469.21	54.31

Table 5.1: Thermodynamic properties of the RuTi SQS unit cells.

It can be seen that the ATAT and ICET 50:50 compositions are in good agreement. The 25% and 75% Ru compositions are of comparable synthesisability and are both expected to be more synthesisable than the 50:50 composition. The corresponding volcano plot is given in Figure 5.10. Due to the compositional differences, the binding sites did not consistently correspond to the same atomic species (Ru or Ti), nor were an equal number of species available for each binding site. This presents a general

challenge when studying non-50:50 compositions, as a simple averaging of all of the adsorption site binding energies (as for Figure 5.8) would skew the averaged result. For instance, a 75% composition of Ru may only perform better than a 25% composition due to a higher number of Ru sites that are available for adsorption.

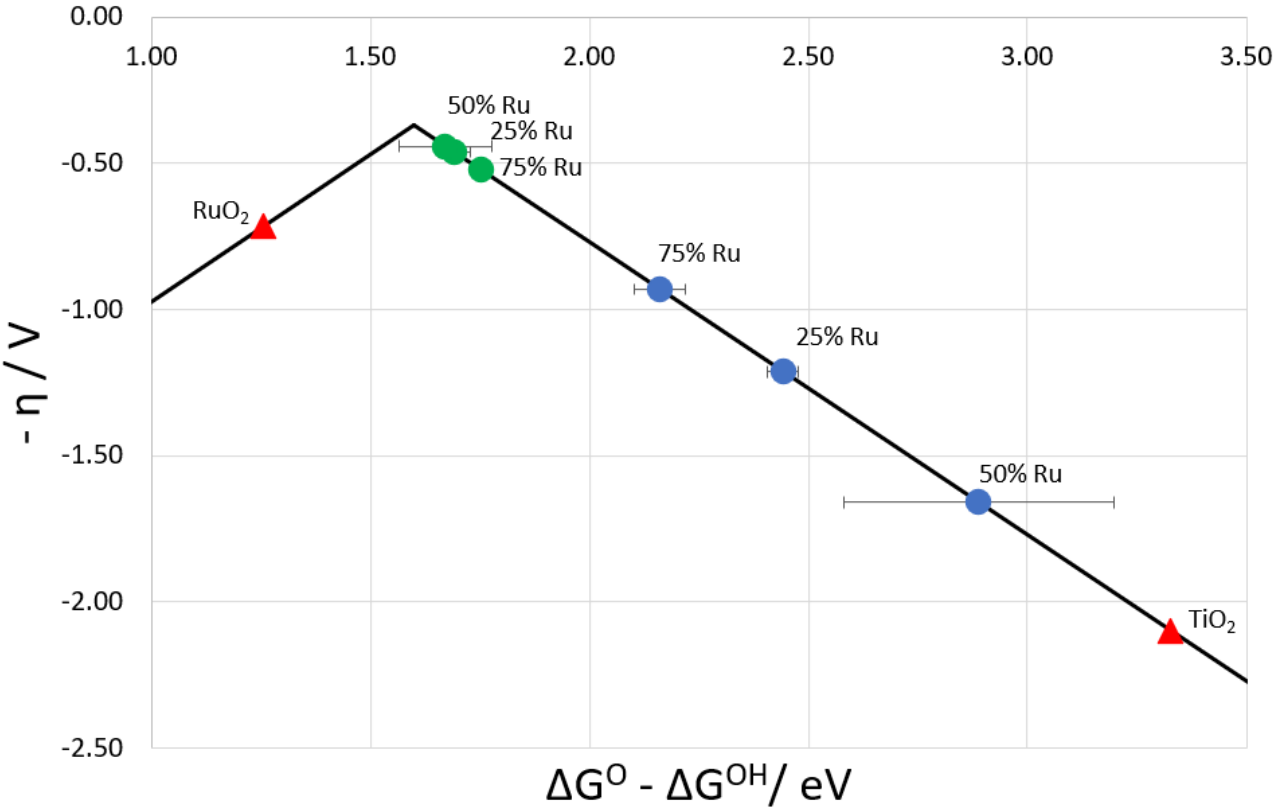


Figure 5.10: Volcano plot of the adsorption energies. The black line plots the ideal scaling relations that form the volcano. The red triangles correspond to the pure compounds. The blue circles correspond to the Ti adsorbate binding sites, whilst the green circles correspond to the Ru adsorbate binding sites. The error bars are the standard deviation between the two binding sites for each species (ruthenium or titanium).

Fortunately, each of the compositions shared the same atomic species for Ti binding sites 1 and 4 and Ru binding sites 5 and 6. The Ru binding sites may be closest to the peak of the volcano due to the superior catalytic activity of Ru. Unexpectedly, an overall higher Ru content did not necessarily imply a greater catalytic activity for both Ru and Ti binding sites. Another unexpected observation is that the 50% Ru composition performs more poorly than the 25% Ru composition when Ti is the binding site. However, the spread of the 50% Ru composition for Ti binding sites was much greater than that of Ru. This may be rationalised by considering the mere presence of Ru as the driving force in this step of the OER reaction, with the average adsorption energy being relatively insensitive to the local environment. Conversely, Ti as a binding site may be incredibly sensitive to the local environment. This would explain the wider spread in catalytic activity for the Ti binding site and the apparent lack of correlation with the bulk concentration of Ru.

5.5 Conclusions and Future Work

In this chapter, SQS models were used to rapidly screen a large number of binary rutile oxide mixtures for their thermodynamic properties and catalytic activities. The ordering of the catalytic activities of the pure compounds was found to agree with past studies[58,59]. Mixture slab generation was rapid and has the potential to be automatable with appropriate slab generation[475,476] and symmetry analysis[484,409] tools. Said mixed slabs suffered from an adsorption energy precision to within 0.1 - 0.05 eV. This was attributed to the inherent lack of long-range order that accompanied the mixed systems. However, this level of precision appeared to be sufficient to derive qualitative trends across a sufficiently wide range of catalytic activities. It is therefore recommended that the presently suggested method be used as such unless the absolute convergence of the mixed system adsorption can be improved. In this regard, future improvements to the presently suggested method include:

Improved dipole corrections Given the reduced degree of confidence in the quantitative precision of the mixed slab adsorption energies, caused by poor convergence of the slab total energy, it would be desirable to develop alternative dipole corrections, as opposed to the protocol used here, to compensate for the loss of long-range order within mixed slab systems.

Consistency of the slab model dimensions It is acknowledged that the generation of slab models for more complex (i.e. non-cubic) crystalline compounds may be more difficult to achieve. To address this problem, fixing the shape of the SQS supercell during an ICET SQS generation and using a larger (ideally near-cubic) supercell is recommended. This would substantially reduce the required amount of benchmarking and human intervention, at the cost of a slight increase in the computational cost per calculation due to the extra number of clusters that would be required to maintain the unit cell shape.

Spin polarisation IrO_2, MnO_2, and RuO_2 are each open shell oxides, each with their magnetic structures[485,442,486]. A consideration of this for the adsorption energy calculations may improve the agreement with future experimental results. However, this would inevitably increase the computational expense of the screening procedure because each calculation would be more expensive. Moreover, the complexity of the screening procedure would be increased for such mixtures as an initial *guess* magnetic structure would have to be specified and optimised in the electronic minimisation. A small-scale investigation into the differences between the closed-shell and open-shell cases would be helpful in this regard.

Choice of functional Since only the PBE functional was used in this work, it is unclear how impor-

tant the choice of functional is in the estimation of the catalytic activity. Man et al claimed that the RPBE functional[483] performed well for the compounds studied in their work[58] whilst Svane et al's work also stressed the sensitivity of the catalytic activity to the DFT functional[59]. For comparison, repeating this work with the RPBE functional may be useful. Additionally, obtaining the adsorption energy with HSE06, using the output of this work as a pre-converged basis, may also yield useful comparative data for hybrid functionals.

Electrochemical interface effects Similar to the discussion of interface dipoles in Chapter 3. A more sophisticated consideration of the electrochemical interface is expected to further refine the accuracy of the adsorption calculations from this work. For example, due to the mere application of a potential (voltage/bias) during operation, charges and dipoles from the electrolyte arrange themselves close to the surface of the electrode via electrostatic attraction. These ordered layers of electrolyte molecules have different properties to the *"bulk"* electrolyte, screening the electric field of the electrode, and causing fluctuations in the electrostatic potential moving from the electrode surface towards the bulk of the electrolyte. This effect is described as the *electrical double layer*. Moreover, the microstructure (and therefore the behaviour of) the electrical double layer is dependent on the applied potential[487]. As such, the adsorption energies from this work may be corrected in light of a quantitative model of the electrical double-layer (i.e. interface dipole correction) for a given system.

Summary

In summary, the first two chapters have outlined the challenges relating to photoelectrochemical water splitting and how such challenges may be addressed using density functional theory. The subsequent three chapters took three such density functional theory implementations and adapted them into more efficient workflows for a materials screening or inverse materials design procedure. In this regard, the results from each of these chapters are summarised in terms of the presently suggested ideal materials screening criteria below.

Inexpensive *"The calculations require minimal computing power. This would allow the procedure to analyse a large number of candidates in a short period of time."* The Frensley and Kroemer band alignment method merely required the geometry to be optimised and an electronic minimisation of the bulk primitive unit cell to be performed using an enhanced FFT grid. This was far less expensive than the traditional slab model generation method of band alignment. Regarding the polaron calculations, it was observed that reducing the number of **k**-points for a polaron calculation greatly reduced the computational overhead when compared to a fully converged counterpart. It was also suggested that the computational expense may be further reduced by using the structural distortion as a proxy for the self-trapping energy. Regarding the corrosion resistance chapter, an SQS model provided a single representative alloy structure as the basis for thermodynamic and catalytic activity calculations. When compared to the generation and enumeration of multiple structures for the calculation of each property, this greatly reduced the required benchmarking and verification time of such thermodynamic and catalytic activity calculations.

Automatable *"The implementation and analysis of the calculations require minimal human intervention. This would accelerate the output of data and reduce the chances of human error influencing*

the result." The Frensley and Kroemer method can be easily automated since automated geometry optimisation and electronic minimisation packages are already available. The only additional required step is the extraction of the interstitial potentials, which may be performed using the Macrodensity package[281,20]. Regarding the polaron calculations, a convenient means with which to automate polaron and defect calculations, using the bond distortion method, has recently been developed by Mosquera-Lois et al[314]. Therefore, a reduction in the **k**-point grid density would be easy to implement using their suggested package, *shakenbreak*[313]. Lastly, SQS models for any mixed crystal system may be rapidly generated using the ATAT or ICET packages. For facile slab model generation, said SQS models may be generated as being as close to cubic as possible, or made as such via supercell transformations.

Accurate *"The calculations produce results that, at minimum, qualitatively reflect the trends in the properties of known systems."* It has been demonstrated, both here and by Tung[18], that the Frensley and Kroemer method produces band alignment results that are at least on par with other models of the time. The polaron morphologies were also shown to be accurate for several simple, well-studied systems and had reasonable self-trapping energies once fully converged. Lastly, SQS models are known to accurately represent the thermodynamic properties of alloy systems and it has been demonstrated here that catalytic activity calculations produce results that are consistent with the scaling relations derived by Man et al[58].

Understood in terms of its limitations *"The limitations of the screening procedure are known and easily identifiable, to avoid misleading results."* It has been a general theme of this work that materials screening procedures must forego some degree of quantitative precision in exchange for computational speed. Notably, the Frensley and Kroemer method lacks a consideration of the interface dipole and its accuracy appears to decline with the magnitude of the lattice parameter. Regarding the polaron calculations, reducing the **k**-point density appears to directly affect the polaron morphology and self-trapping energy for some systems, notably Co_3O_4. Moreover, supercell size selection is difficult to establish *a priori*. Regarding the corrosion resistance chapter, the number of surface sites available for adsorption during an SQS-based slab calculation is dependent on the composition. As such, adsorption data for a particular mixture component, at low concentrations of said component, may be difficult to determine. Additional measures have been suggested to reduce the number of outlier adsorptions and to exclude highly distorted candidates.

Modular *"The components of the screening procedure function independently. This would enable convenient experimentation with different approximations, standards, and machine learning implementations. Practically, this also makes the development of the procedure, including error correction, more*

manageable." The Frensley and Kroemer method merely requires a reading of the spatially resolved electrostatic potential from an electronic structure calculation. Hence, it may be added to, or removed from, a given screening procedure without requiring substantial reconsideration. Regarding the polaron calculations, using a reduced **k**-point sampling for the polaron states may be implemented by merely altering the standard input file, whilst the structural distortions may be applied, adapted, or removed using the standard functions and packages that were mentioned in this work. Lastly, the SQS method may be used with/without other cluster expansion techniques, and the site occupancies of a given SQS model may be easily adjusted and rearranged using the ICET package[33,434]. Overall, this provides a modular and flexible base structure for further experimentation.

It is hoped that the methods and data presented in this work will inform a reader that is interested in the rapid screening of photoelectrochemical cell electrodes. It is further hoped that the procedures suggested here will promote the discovery and synthesis of more efficient and abundant compounds for photoelectrochemical cell device development. These procedures may be implemented as cost-saving measures for direct property calculations or as low-cost feature calculators for use in a machine-learning algorithm.

Bibliography

[1] W. M. Haynes, D. R. Lide and T. J. Bruno, *CRC handbook of chemistry and physics : a ready-reference book of chemical and physical data.*, CRC Press, Boca Raton, 94th edn., 2013 - 2013.

[2] L. J. Brillson, *Semiconductor Properties*, John Wiley and Sons, Ltd, 2012, pp. 549–550.

[3] O. Madelung, *Semiconductors: data handbook*, Springer Science & Business Media, 2004.

[4] R. H. Victoria, *Phys. Rev. B*, 1997, **56**, 4417–4421.

[5] R. Matzdorf, J. Skonieczny, J. Westhof, H. Engelhard and A. Goldmann, *J. Phys.: Cond. Mat.*, 1993, **5**, 3827–3836.

[6] K. Bushick, K. Mengle, N. Sanders and E. Kioupakis, *Appl. Phys. Lett.*, 2019, **114**, 022101.

[7] N. Doran and A. Woolley, *J. Phys. C: Solid State Phys.*, 1979, **12**, L321.

[8] C. Yang, M. Kneiß, F.-L. Schein, M. Lorenz and M. Grundmann, *Sci. Rep.*, 2016, **6**, 1–8.

[9] W. Yim, J. Dismukes, E. Stofko and R. Paff, *J. Phys. Chem. Solids*, 1972, **33**, 501–505.

[10] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder and K. A. Persson, *APL Materials*, 2013, **1**, 011002.

[11] *MAGNDATA: A Collection of magnetic structures with portable cif-type files*, `https://www.cryst.ehu.es/magndata/`, [Accessed 2024-03-14].

[12] S. V. Gallego, J. M. Perez-Mato, L. Elcoro, E. S. Tasci, R. M. Hanson, K. Momma, M. I. Aroyo and G. Madariaga, *Journal of Applied Crystallography*, 2016, **49**, 1750–1776.

[13] S. V. Gallego, J. M. Perez-Mato, L. Elcoro, E. S. Tasci, R. M. Hanson, M. I. Aroyo and G. Madariaga, *Journal of Applied Crystallography*, 2016, **49**, 1941–1956.

[14] J. Perez-Mato, S. Gallego, E. Tasci, L. Elcoro, G. De La Flor and M. Aroyo, *Annual Review of Materials Research*, 2015, **45**, 217–248.

[15] M. Farah Khaleda, B. Vengadaesvaran and N. Rahim, in *Energy Materials*, ed. S. Dhoble, N. Kalyani, B. Vengadaesvaran and A. Kariem Arof, Elsevier, 2021, pp. 525–566.

[16] M. Clugston, *The Penguin Dictionary of Science*, Penguin, 4th edn., 2014.

[17] A. M Ganose, A. J Jackson and D. O Scanlon, *Journal of Open Source Software*, 2018, **3**, 717.

[18] R. T. Tung, *Materials Science and Engineering: R: Reports*, 2001, **35**, 1–138.

[19] B. D. Pelatt, R. Ravichandran, J. F. Wager and D. A. Keszler, *Journal of the American Chemical Society*, 2011, **133**, 16852–16860.

[20] L. Harnett-Caulfield and A. Walsh, *The Journal of Chemical Physics*, 2021, **155**, 024113.

[21] F. Izumi and K. Momma, *Solid State Phenomena*, 2007, **130**, 15–20.

[22] K. Momma and F. Izumi, *Journal of Applied Crystallography*, 2008, **41**, 653–658.

[23] K. Momma and F. Izumi, *Journal of Applied Crystallography*, 2011, **44**, 1272–1276.

[24] K. T. Butler, C. H. Hendon and A. Walsh, *Journal of the American Chemical Society*, 2014, **136**, 2703–2706.

[25] H. Kroemer, *C R C Critical Reviews in Solid State Sciences*, 1975, **5**, 555–564.

[26] W. R. Frensley and H. Kroemer, *Journal of Vacuum Science and Technology*, 1976, **13**, 810–815.

[27] W. R. Frensley and H. Kroemer, *Applied Physics Letters*, 1977, **31**, 48–50.

[28] W. R. Frensley and H. Kroemer, *Physical Review B*, 1977, **16**, 2642–2652.

[29] X. Li, J. Yu, J. Low, Y. Fang, J. Xiao and X. Chen, *J. Mater. Chem. A*, 2015, **3**, 2485–2534.

[30] Y. Zhang, C. Zhang, X. Huang, Z. Yang, K. H. L. Zhang and Y. Yang, *The Journal of Physical Chemistry Letters*, 2021, **12**, 12033–12039.

[31] T. J. Smart, T. A. Pham, Y. Ping and T. Ogitsu, *Physical Review Materials*, 2019, **3**, 102401.

[32] WMD-group, *CarrierCapture.jl*, `https://github.com/WMD-group/CarrierCapture.jl/blob/master/script/get_del_Q.py`, [Accessed: 2024-03-14].

[33] M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod and P. Erhart, *Advanced Theory and Simulations*, 2019, **2**, 1900015.

[34] GOV.UK, *Energy consumption in the UK 2021*, 2022, https://www.gov.uk/government/statistics/energy-consumption-in-the-uk-2021, [Accessed: 2024-01-23].

[35] GOV.UK, *Energy and emissions projections: 2021 to 2040*, 2023, https://www.gov.uk/government/publications/energy-and-emissions-projections-2021-to-2040, [Accessed: 2024-01-23].

[36] Global Energy Monitor Wiki, *Air pollution from coal-fired power plants*, 2021, https://www.gem.wiki/Air_pollution_from_coal-fired_power_plants, [Accessed: 2024-01-31].

[37] US EPA, *Basic Information about Oil and Natural Gas Air Pollution Standards*, 2016, https://www.epa.gov/controlling-air-pollution-oil-and-natural-gas-operations/basic-information-about-oil-and-natural, [Accessed: 2024-01-31].

[38] MetGroup, *Gas vs. oil: differences between natural gas and crude oil*, https://group.met.com/en/media/energy-insight/natural-gas-vs-oil, [Accessed: 2024-01-23].

[39] IAEA, *What is Nuclear Energy? The Science of Nuclear Power*, 2022, https://www.iaea.org/newscenter/news/what-is-nuclear-energy-the-science-of-nuclear-power, [Accessed: 2024-01-31].

[40] N. E. Selin and C. Lehman, *Biofuel - Encyclopedia Britannica*, 2023, https://www.britannica.com/technology/biofuel, [Accessed: 2024-01-31].

[41] The Editors of Encyclopaedia Britannica, *Hydroelectric power - Encyclopedia Britannica*, 2023, https://www.britannica.com/science/hydroelectric-power, [Accessed: 2024-01-31].

[42] N. E. Selin, *Wind power - Encyclopedia Britannica*, 2024, https://www.britannica.com/science/wind-power, [Accessed: 2024-01-31].

[43] J. W. Lund, *Geothermal energy - Encyclopedia Britannica*, 2024, https://www.britannica.com/science/geothermal-energy, [Accessed: 2024-01-31].

[44] EnergyData.info, *Global Solar Atlas*, https://globalsolaratlas.info/map, [Accessed: 2024-01-23].

[45] A. Bernardes, D. Espinosa and J. Tenório, *Journal of Power Sources*, 2004, **130**, 291–298.

[46] K. Nwaigwe, P. Mutabilwa and E. Dintwa, *Materials Science for Energy Technologies*, 2019, **2**, 629–633.

[47] M. Winter and R. J. Brodd, *Chemical Reviews*, 2004, **104**, 4245–4270.

[48] Y. K. Patanwar, H.-M. Kim, D. Deb and Y. K. Gujjala, *International Journal of Hydrogen Energy*, 2024, **50**, 116–133.

[49] M. Masoudi, A. Hassanpouryouzband, H. Hellevang and R. S. Haszeldine, *Journal of Energy Storage*, 2024, **84**, 110927.

[50] I. Dincer and C. Acar, *International Journal of Hydrogen Energy*, 2015, **40**, 11094–11111.

[51] B. D. James, G. N. Baum, J. Perez and K. N. Baum, *Technoeconomic Analysis of Photoelectrochemical (PEC) Hydrogen Production*, 1218403, 2009.

[52] M. Kumar, B. Meena, P. Subramanyam, D. Suryakala and C. Subrahmanyam, *NPG Asia Materials*, 2022, **14**, 1–21.

[53] V. Kumaravel, J. Bartlett and S. C. Pillai, *ACS Energy Lett.*, 2020, **5**, 486–519.

[54] IEA, *Direct Air Capture 2022 - Analysis and key findings. A report by the International Energy Agency*, 2022, `https://www.iea.org/reports/direct-air-capture-2022`, [Accessed: 2023-05-17].

[55] Z. Chen, H. N. Dinh and E. Miller, *Photoelectrochemical Water Splitting: Standards, Experimental Methods, and Protocols*, Springer New York, New York, NY, 2013.

[56] A. Fujishima and K. Honda, *Nature*, 1972, **238**, 37–38.

[57] T. Naito, T. Shinagawa, T. Nishimoto and K. Takanabe, *Inorganic Chemistry Frontiers*, 2021, **8**, 2900–2917.

[58] I. C. Man, H. Su, F. Calle-Vallejo, H. A. Hansen, J. I. Martínez, N. G. Inoglu, J. Kitchin, T. F. Jaramillo, J. K. Nørskov and J. Rossmeisl, *ChemCatChem*, 2011, **3**, 1159–1165.

[59] K. L. Svane and J. Rossmeisl, *Angewandte Chemie International Edition*, 2022, **61**, e202201146.

[60] A. Lasia, *International Journal of Hydrogen Energy*, 2019, **44**, 19484–19518.

[61] A. J. Bard, *Journal of Photochemistry*, 1979, **10**, 59–75.

[62] A. J. Bard, *Science*, 1980, **207**, 139–144.

[63] A. J. Bard, *The Journal of Physical Chemistry*, 1982, **86**, 172–177.

[64] Merriam-Webster Dictionary, *Photovoltage*, `https://www.merriam-webster.com/dictionary/photovoltage`, [Accessed: 2024-01-23].

[65] E. A. Schmitt, M. Guidat, M. Nusshör, A.-L. Renz, K. Möller, M. Flieg, D. Lörch, M. Kölbach and M. M. May, *Cell Reports Physical Science*, 2023, **4**, 101606.

[66] A. Iqbal and K. H. Bevan, *The Journal of Physical Chemistry C*, 2018, **122**, 30–43.

[67] D. M. Smyth, *The defect chemistry of metal oxides*, Oxford University Press, New York, 2000.

[68] A. R. Zanatta, *Scientific Reports*, 2019, **9**, 11225.

[69] Y. H. Lee, *Science*, 2024, **383**, eado4308.

[70] J. Nelson, *The physics of solar cells*, Imperial College Press ; Distributed by World Scientific Pub. Co, London : River Edge, NJ, 2003.

[71] M. A. Khan, T. Al-Attas, S. Roy, M. M. Rahman, N. Ghaffour, V. Thangadurai, S. Larter, J. Hu, P. M. Ajayan and M. G. Kibria, *Energy & Environmental Science*, 2021, **14**, 4831–4839.

[72] Y. Kuang, M. J. Kenney, Y. Meng, W.-H. Hung, Y. Liu, J. E. Huang, R. Prasanna, P. Li, Y. Li, L. Wang, M.-C. Lin, M. D. McGehee, X. Sun and H. Dai, *Proceedings of the National Academy of Sciences of the United States of America*, 2019, **116**, 6624–6629.

[73] M. Li, C. S. Cucinotta and A. P. Horsfield, *Journal of Physics and Chemistry of Solids*, 2022, **170**, 110936.

[74] D. Eaves, G. Williams and H. McMurray, *Electrochimica Acta*, 2012, **79**, 1–7.

[75] G. Gibson, A. Morgan, P. Hu and W.-F. Lin, *Chemical Physics Letters*, 2016, **654**, 46–51.

[76] A. Chaves, J. G. Azadani, H. Alsalman, D. R. da Costa, R. Frisenda, A. J. Chaves, S. H. Song, Y. D. Kim, D. He, J. Zhou, A. Castellanos-Gomez, F. M. Peeters, Z. Liu, C. L. Hinkle, S.-H. Oh, P. D. Ye, S. J. Koester, Y. H. Lee, P. Avouris, X. Wang and T. Low, *npj 2D Materials and Applications*, 2020, **4**, 1–21.

[77] A. Jahan, M. Ismail, S. Sapuan and F. Mustapha, *Materials & Design*, 2010, **31**, 696–705.

[78] A. Zunger, *Nature Reviews Chemistry*, 2018, **2**, 1–16.

[79] Cambridge Dictionary, *Clock Speed*, `https://dictionary.cambridge.org/dictionary/english/clock-speed`, [Accessed: 2024-01-31].

[80] S. Dasgupta, *Computer Science: A Very Short Introduction*, Oxford University Press, 2016.

[81] G. Boole, *An Investigation of the Laws of Thought: On Which Are Founded the Mathematical Theories of Logic and Probabilities*, Cambridge University Press, Cambridge, 2009.

[82] D. Kahneman and A. Tversky, *Econometrica*, 1979, **47**, 263–291.

[83] A. L. Samuel, *IBM Journal of Research and Development*, 1959, **3**, 210–229.

[84] B. Oral, E. Can and R. Yildirim, *International Journal of Hydrogen Energy*, 2022, **47**, 19633–19654.

[85] Z. Wang, J. Cai, Q. Wang, S. Wu and J. Li, *npj Computational Materials*, 2021, **7**, 128.

[86] S. Gow, M. Niranjan, S. Kanza and J. G. Frey, *Digital Discovery*, 2022, **1**, 551–567.

[87] D. B. Farmer, R. Golizadeh-Mojarad, V. Perebeinos, Y.-M. Lin, G. S. Tulevski, J. C. Tsang and P. Avouris, *Nano Letters*, 2009, **9**, 388–392.

[88] The Editors of Encyclopaedia Britannica, *Resistivity - Encyclopaedia Britannica*, 2024, `https://www.britannica.com/science/resistivity`, [Accessed: 2024-01-31].

[89] The Editors of Encyclopaedia Britannica, *Semiconductor - Encyclopaedia Britannica*, 2024, `https://www.britannica.com/science/semiconductor`, [Accessed: 2024-01-31].

[90] J. W. Allen, *Nature*, 1960, **187**, 403–405.

[91] *Concise Science Dictionary*, ed. Oxford University Press, Oxford University Press, Oxford ; New York, 1984.

[92] M. S. Islam, J. Lee, S. Ganguli and A. K. Roy, *Scientific Reports*, 2023, **13**, 16656.

[93] D. O. Scanlon, A. B. Kehoe, G. W. Watson, M. O. Jones, W. I. F. David, D. J. Payne, R. G. Egdell, P. P. Edwards and A. Walsh, *Physical Review Letters*, 2011, **107**, 246402.

[94] W. Gerlach and O. Stern, *Zeitschrift für Physik*, 1922, **9**, 349–352.

[95] W. Pauli, *Wolfgang Pauli – Nobel Lecture, NobelPrize.org. Nobel Prize Outreach AB 2022*, 1946, `https://www.nobelprize.org/prizes/physics/1945/pauli/lecture/`, [Accessed: 2024-03-14].

[96] R. D. Knight, *Physics for scientists and engineers: a strategic approach; with modern physics. Textbook*, Pearson/Addison-Wesley, San Francisco, 2nd edn., 2009.

[97] The Editors of Encyclopaedia Britannica, *Wave function - Encyclopaedia Britannica*, 2023, https://www.britannica.com/science/wave-function, [Accessed: 2023-05-17].

[98] E. Schrödinger, *Physical Review*, 1926, **28**, 1049–1070.

[99] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press, 2004.

[100] *The Penguin dictionary of mathematics*, ed. R. D. Nelson and D. Nelson, Penguin Books, London, 4th edn., 2008.

[101] J. M. Herbert, in *J. M. Herbert*, 2023, pp. 69–118.

[102] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*, Academic Press, San Diego, 2nd edn., 2002.

[103] F. Bloch, *Zeitschrift f r Physik*, 1929, **52**, 555–600.

[104] R. Hoffmann, *Angewandte Chemie International Edition in English*, 1987, **26**, 846–878.

[105] Vaspwiki, *K-point integration*, https://www.vasp.at/wiki/index.php/K-point_integration, [Accessed: 2023-05-17].

[106] P. E. Blöchl, O. Jepsen and O. K. Andersen, *Phys. Rev. B*, 1994, **49**, 16223–16233.

[107] Vaspwiki, *ISMEAR*, https://www.vasp.at/wiki/index.php/ISMEAR, [Accessed: 2023-05-17].

[108] M. Y. Toriyama, A. M. Ganose, M. Dylla, S. Anand, J. Park, M. K. Brod, J. Munro, K. A. Persson, A. Jain and G. J. Snyder, *Comparison of the Tetrahedron Method to Smearing Methods for the Electronic Density of States*, 2021.

[109] Vaspwiki, *Number of k points and method for smearing*, https://www.vasp.at/wiki/index.php/Number_of_k_points_and_method_for_smearing, [Accessed: 2023-05-17].

[110] J. D. Pack and H. J. Monkhorst, *Physical Review B*, 1977, **16**, 1748–1749.

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

[112] S. Lehtola, *International Journal of Quantum Chemistry*, 2019, **119**, e25968.

[113] R. Ditchfield, W. J. Hehre and J. A. Pople, *The Journal of Chemical Physics*, 1971, **54**, 724–728.

[114] CASTEP, *Plane wave basis set - Theory in CASTEP*, https://www.tcm.phy.cam.ac.uk/castep/documentation/WebHelp/content/modules/castep/thcastepplanebasis.htm, [Accessed: 2023-05-17].

[115] P. Schwerdtfeger, *ChemPhysChem*, 2011, **12**, 3143–3155.

[116] P. E. Blöchl, *Physical Review B*, 1994, **50**, 17953–17979.

[117] T. Koopmans, *Physica*, 1934, **1**, 104–113.

[118] J. F. Janak, *Physical Review B*, 1978, **18**, 7165–7168.

[119] D. R. Hartree, *Mathematical Proceedings of the Cambridge Philosophical Society*, 1928, **24**, 89–110.

[120] D. R. Hartree, *Mathematical Proceedings of the Cambridge Philosophical Society*, 1928, **24**, 111–132.

[121] D. R. Hartree, *Mathematical Proceedings of the Cambridge Philosophical Society*, 1928, **24**, 426–437.

[122] E. Wigner, *Physical Review*, 1934, **46**, 1002–1011.

[123] L. P. Kadanoff, *Journal of Statistical Physics*, 2009, **137**, 777.

[124] J. Zaanen, G. A. Sawatzky and J. W. Allen, *Physical Review Letters*, 1985, **55**, 418–421.

[125] J. C. Slater, *Phys. Rev.*, 1929, **34**, 1293–1322.

[126] A. Zannoni, *On the Quantization of the Monoatomic Ideal Gas*, 1999.

[127] *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 1926, **112**, 661–677.

[128] C. Lanczos, *The Variational Principles of Mechanics 4th ed.*, University of Toronto Press, 1970.

[129] D. L. Strout and G. E. Scuseria, *The Journal of Chemical Physics*, 1995, **102**, 8448–8452.

[130] T. van Mourik, M. Bühl and M.-P. Gaigeot, *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 2014, **372**, 20120488.

[131] P. Hohenberg and W. Kohn, *Physical Review*, 1964, **136**, B864–B871.

[132] W. Kohn and L. J. Sham, *Physical Review*, 1965, **140**, A1133–A1138.

[133] D. Bagayoko, *AIP Advances*, 2014, **4**, 127104.

[134] J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1996, **77**, 3865–3868.

[135] P. Borlido, J. Schmidt, A. W. Huran, F. Tran, M. A. L. Marques and S. Botti, *npj Computational Materials*, 2020, **6**, 1–17.

[136] J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1997, **78**, 1396–1396.

[137] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, *Physical review letters*, 2008, **100**, 136406.

[138] J. Sun, A. Ruzsinszky and J. Perdew, *Phys. Rev. Lett.*, 2015, **115**, 036402.

[139] Vaspwiki, *Bandgap of Si using different DFT+HF methods*, 2022, `https://www.vasp.at/wiki/index.php/Bandgap_of_Si_using_different_DFT%2BHF_methods`, [Accessed: 2024-03-14].

[140] Vaspwiki, *Fcc Ni DOS with hybrid functional*, 2022, `https://www.vasp.at/wiki/index.php/Fcc_Ni_DOS_with_hybrid_functional`, [Accessed: 2024-03-14].

[141] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov and G. E. Scuseria, *The Journal of Chemical Physics*, 2006, **125**, 224106.

[142] *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1963, **276**, 238–257.

[143] *Nature Physics*, 2013, **9**, 523–523.

[144] Vaspwiki, *Davidson iteration scheme*, `https://www.vasp.at/wiki/index.php/Davidson_iteration_scheme`, [Accessed: 2023-05-17].

[145] D. ER, *Journal of Computational Physics*, 1975, **17**, 87–94.

[146] Vaspwiki, *Conjugate gradient optimization*, 2022, `https://www.vasp.at/wiki/index.php/Conjugate_gradient_optimization`, [Accessed: 2024-03-14].

[147] M. Hestenes and E. Stiefel, *Journal of Research of the National Bureau of Standards*, 1952, **49**, 409.

[148] T. A. Straeter, *On the Extension of the Davidon-Broyden Class of Rank One, Quasi-Newton Minimization Methods to an Infinite Dimensional Hilbert Space with Applications to Optimal Control Problems*, 1971, `https://ntrs.nasa.gov/citations/19710026200`, [Accessed: 2023-05-18].

[149] C. Lanczos, *Journal of Research of the National Bureau of Standards*, 1950, **45**, 255.

[150] M. Sadkane and R. B. Sidje, *Numerical Algorithms*, 1999, **20**, 217–240.

[151] Vaspwiki, *Density mixing*, 2022, https://www.vasp.at/wiki/index.php/Category:Density_mixing, [Accessed: 2024-03-14].

[152] D. D. Johnson, *Physical Review B*, 1988, **38**, 12807–12813.

[153] G. Kresse and J. Furthmüller, *Computational Materials Science*, 1996, **6**, 15–50.

[154] G. P. Kerker, *Physical Review B*, 1981, **23**, 3082–3084.

[155] G. Kresse and J. Hafner, *Physical Review B*, 1993, **47**, 558–561.

[156] G. Kresse and J. Furthmüller, *Physical Review B*, 1996, **54**, 11169–11186.

[157] G. Kresse and D. Joubert, *Physical Review B*, 1999, **59**, 1758–1775.

[158] Vaspwiki, *Cd Si volume relaxation*, https://www.vasp.at/wiki/index.php/Cd_Si_volume_relaxation, [Accessed: 2023-05-17].

[159] Vaspwiki, *Fcc Si*, https://www.vasp.at/wiki/index.php/Fcc_Si, [Accessed: 2023-05-17].

[160] Vaspwiki, *Energy vs volume Volume relaxations and Pulay stress*, https://www.vasp.at/wiki/index.php/Energy_vs_volume_Volume_relaxations_and_Pulay_stress, [Accessed: 2023-05-17].

[161] A. J. Jackson, J. M. Skelton, C. H. Hendon, K. T. Butler and A. Walsh, *The Journal of Chemical Physics*, 2015, **143**, 184101.

[162] G. M. Dongho Nguimdo and D. P. Joubert, *The European Physical Journal B*, 2015, **88**, 113.

[163] Vaspwiki, *IBRION*, https://www.vasp.at/wiki/index.php/IBRION, [Accessed: 2023-05-17].

[164] Vaspwiki, *Not enough memory*, https://www.vasp.at/wiki/index.php/Not_enough_memory, [Accessed: 2023-05-17].

[165] Vaspwiki, *Category:Parallelization*, https://www.vasp.at/wiki/index.php/Category:Parallelization, [Accessed: 2023-05-17].

[166] C. G. Rocha, A. R. Rocha, P. Venezuela, J. H. Garcia and M. S. Ferreira, *Scientific Reports*, 2018, **8**, 9348.

[167] A. Kahn, *Materials Horizons*, 2015, **3**, 7–10.

[168] A. J. Logsdail, D. O. Scanlon, C. R. A. Catlow and A. A. Sokol, *Physical Review B*, 2014, **90**, 155106.

[169] Y. Kumagai, K. T. Butler, A. Walsh and F. Oba, *Physical Review B*, 2017, **95**, 125309.

[170] L. Pielsticker, I. Zegkinoglou, Z.-K. Han, J. J. Navarro, S. Kunze, O. Karslıoğlu, S. V. Levchenko and B. Roldan Cuenya, *The Journal of Physical Chemistry Letters*, 2021, **12**, 2570–2575.

[171] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, *Reviews of Modern Physics*, 1992, **64**, 1045–1097.

[172] International Union of Pure and Applied Chemistry (IUPAC), *Standard hydrogen electrode*, 2019, `https://doi.org/10.1351/goldbook.S05917`, [Accessed: 2023-05-18].

[173] A. Zunger and M. L. Cohen, *Physical Review B*, 1978, **18**, 5449–5472.

[174] K. Brlec, D. Davies and D. Scanlon, *Journal of Open Source Software*, 2021, **6**, 3171.

[175] D. O. Scanlon, C. W. Dunnill, J. Buckeridge, S. A. Shevlin, A. J. Logsdail, S. M. Woodley, C. R. A. Catlow, M. J. Powell, R. G. Palgrave, I. P. Parkin, G. W. Watson, T. W. Keal, P. Sherwood, A. Walsh and A. A. Sokol, *Nature Materials*, 2013, **12**, 798–801.

[176] M. Cardona and N. E. Christensen, *Physical Review B*, 1987, **35**, 6182–6194.

[177] B. Höffling, A. Schleife, F. Fuchs, C. Rödl and F. Bechstedt, *Applied Physics Letters*, 2010, **97**, 032116.

[178] J. M. Langer, C. Delerue, M. Lannoo and H. Heinrich, *Physical Review B*, 1988, **38**, 7723–7739.

[179] S. Wei and A. Zunger, *Journal of Applied Physics*, 1995, **78**, 3846–3856.

[180] C. G. Van de Walle and J. Neugebauer, *Nature*, 2003, **423**, 626–628.

[181] P. P. Ewald, *Annalen der Physik*, 1921, **369**, 253–287.

[182] J. D. Gale, *J. Chem. Soc., Faraday Trans.*, 1997, **93**, 629–637.

[183] C. G. Van de Walle and R. M. Martin, *Physical Review B*, 1987, **35**, 8154–8165.

[184] E. T. Yu, J. O. McCaldin and T. C. McGill, in *Solid State Physics*, ed. H. Ehrenreich and D. Turnbull, Academic Press, 1992, vol. 46, pp. 1–146.

[185] W. A. Harrison, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, 1986, **4**, 1068.

[186] D.-H. Choe, D. West and S. Zhang, *Physical Review Letters*, 2018, **121**, 196802.

[187] V. F. Yusuf, N. I. Malek and S. K. Kailasa, *ACS Omega*, 2022, **7**, 44507–44531.

[188] X. Xin, Y. Li, Y. Zhang, Y. Wang, X. Chi, Y. Wei, C. Diao, J. Su, R. Wang, P. Guo, J. Yu, J. Zhang, A. J. Sobrido, M.-M. Titirici and X. Li, *Nature Communications*, 2024, **15**, 337.

[189] C. R. A. Catlow and A. M. Stoneham, *Journal of Physics C: Solid State Physics*, 1983, **16**, 4321.

[190] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, 1990.

[191] R. F. W. Bader, *Phys. Rev. B*, 1994, **49**, 13348–13356.

[192] W. Tang, E. Sanville and G. Henkelman, *Journal of Physics: Condensed Matter*, 2009, **21**, 084204.

[193] E. Sanville, S. D. Kenny, R. Smith and G. Henkelman, *Journal of Computational Chemistry*, 2007, **28**, 899–908.

[194] G. Henkelman, A. Arnaldsson and H. Jónsson, *Computational Materials Science*, 2006, **36**, 354–360.

[195] M. Yu and D. R. Trinkle, *The Journal of Chemical Physics*, 2011, **134**, 064111.

[196] C. Li, J. C. Li and Q. Jiang, *Solid State Communications*, 2010, **150**, 1818–1821.

[197] J. C. Phillips, *Science*, 1970, **169**, 1035–1042.

[198] B. Hönerlage, in *New Data and Updates for IV-IV, III-V, II-VI and I-VII Compounds, their Mixed Crystals and Diluted Magnetic Semiconductors*, ed. W. Martienssen and U. Rössler, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, vol. 44D, p. 42.

[199] Z. Tauanov and V. Inglezakis, *Science of The Total Environment*, 2019, **682**, 259–270.

[200] C. Zeng, Y. Hu, Y. Guo, T. Zhang, F. Dong, X. Du, Y. Zhang and H. Huang, *Applied Catalysis B: Environmental*, 2016, **194**, 62–73.

[201] The Materials Project, *mp-2172: AlAs (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-2172/`, [Accessed: 2024-03-13].

[202] ThermoFisher Scientific, *Safety Data Sheet - Aluminium Arsenide*, April 2022, Version 3, ALFAA14522.

[203] L. Balaghi, S. Shan, I. Fotev, F. Moebus, R. Rana, T. Venanzi, R. Hübner, T. Mikolajick, H. Schneider, M. Helm, A. Pashkin and E. Dimakis, *Nature Communications*, 2021, **12**, 6642.

[204] The Materials Project, *mp-1550: AlP (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-1550/`, [Accessed: 2024-03-13].

[205] GOV.UK, *Aluminium phosphide: general information*, `https://www.gov.uk/government/publications/aluminium-phosphide-health-effects-and-incident-management/aluminium-phosphide-general-information`, [Accessed: 2024-03-13].

[206] The Materials Project, *mp-2624: AlSb (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-2624/`, [Accessed: 2024-03-13].

[207] ThermoFisher Scientific, *Safety Data Sheet - Aluminium Antimonide*, February 2021, Version 2, ALFAA43455.

[208] M. A. Mohebpour and M. B. Tagani, *Scientific Reports*, 2023, **13**, 9925.

[209] The Materials Project, *mp-10044: BAs (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-10044/`, [Accessed: 2024-03-13].

[210] J. Shin, G. A. Gamage, Z. Ding, K. Chen, F. Tian, X. Qian, J. Zhou, H. Lee, J. Zhou, L. Shi, T. Nguyen, F. Han, M. Li, D. Broido, A. Schmidt, Z. Ren and G. Chen, *Science*, 2022, **377**, 437–440.

[211] American Elements, *Safety Data Sheet - Boron Arsenide*, January 2022, Version 1, 12005-69-5.

[212] J. Shin, G. A. Gamage, Z. Ding, K. Chen, F. Tian, X. Qian, J. Zhou, H. Lee, J. Zhou, L. Shi, T. Nguyen, F. Han, M. Li, D. Broido, A. Schmidt, Z. Ren and G. Chen, *Science*, 2022, **377**, 437–440.

[213] A. Tararan, S. Di Sabatino, M. Gatti, T. Taniguchi, K. Watanabe, L. Reining, L. H. G. Tizei, M. Kociak and A. Zobelli, *Physical Review B*, 2018, **98**, 094106.

[214] Sigma-Aldrich, *Safety Data Sheet - Boron Nitride*, January 2022, Version 8.4, 10043-11-5.

[215] S. Saha, *International Journal of Hydrogen Energy*, 2024, **51**, 503–514.

[216] The Materials Project, *mp-1479: BP (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-1479/, [Accessed: 2024-03-13].

[217] T. Suzuki, *Applied Surface Science*, 2022, **598**, 153844.

[218] The Materials Project, *mp-422: BeS (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-422/, [Accessed: 2024-03-13].

[219] B. A. Ayirizia, J. S. Brumfield, Y. Malozovsky and D. Bagayoko, *Materials*, 2021, **14**, year.

[220] B. A. Ayirizia, J. S. Brumfield, Y. Malozovsky and D. Bagayoko, *Materials*, 2021, **14**, 6128.

[221] The Materials Project, *mp-1541: BeSe (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-1541/, [Accessed: 2024-03-13].

[222] S. Laref and A. Laref, *Journal of Materials Science*, 2013, **48**, 5499–5508.

[223] T. Larbi, I. Toumi, K. Doll and M. Amlouk, *Surfaces and Interfaces*, 2021, **24**, 101087.

[224] The Materials Project, *Materials Data on BeTe by Materials Project*, 2020, https://www.osti.gov/servlets/purl/1200563/, [Accessed: 2024-01-15].

[225] The Materials Project, *Materials Data on CdS by Materials Project*, 2020, https://www.osti.gov/servlets/purl/1200195/, [Accessed: 2024-01-15].

[226] ThermoFisher Scientific, *Safety Data Sheet - Cadmium Sulphide*, February 2022, Version 3, AL-FAAA14544.

[227] A. B. Isaev, N. S. Shabanov, D. Sobola, K. Kaviyarasu, A. M. Ismailov and G. M. Omarov, in *Nanomaterials for Energy Conversion, Biomedical and Environmental Applications*, ed. K. Kasinathan, M. S. Elshikh and D. A.-A. Al Farraj, Springer Nature, Singapore, 2022, pp. 3–35.

[228] U. Abdullah, M. Ali and E. Pervaiz, *Molecular Catalysis*, 2021, **508**, 111575.

[229] The Materials Project, *mp-2691: CdSe (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-2691/, [Accessed: 2024-03-13].

[230] ThermoFisher Scientific, *Safety Data Sheet - Cadmium Selenide*, February 2021, Version 2, AL-FAA36323.

[231] The Materials Project, *mp-406: CdTe (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-406/, [Accessed: 2024-03-13].

[232] The Materials Project, *mp-22913: CuBr (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-22913/, [Accessed: 2024-03-13].

[233] ThermoFisher Scientific, *Safety Data Sheet - Copper(I) Bromide*, December 2021, Version 4, 7787-70-4.

[234] The Materials Project, *mp-22914: CuCl (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-22914/, [Accessed: 2024-03-13].

[235] The Materials Project, *mp-22895: CuI (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-22895/, [Accessed: 2024-03-13].

[236] ThermoFisher Scientific, *Safety Data Sheet - Copper(I) Iodide*, December 2021, Version 7, 7681-65-4.

[237] A. Liu, H. Zhu, M. Kim, J. Kim and Y. Noh, *Advanced Science*, 2021, **8**, 2100546.

[238] ThermoFisher Scientific, *Safety Data Sheet - Gallium Arsenide*, February 2021, Version 2, ALFAA88458.

[239] M. Arunachalam, R. S. Kanase, K. Zhu and S. H. Kang, *Nature Communications*, 2023, **14**, 5429.

[240] The Materials Project, *mp-2490: GaP (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-2490/, [Accessed: 2024-03-13].

[241] J. Ziegler, D. Fertig, B. Kaiser, W. Jaegermann, M. Blug, S. Hoch and J. Busse, *Energy Procedia*, 2012, **22**, 108–113.

[242] S. Adachi, in *Optical Constants of Crystalline and Amorphous Semiconductors: Numerical Data and Graphical Information*, ed. S. Adachi, Springer US, Boston, MA, 1999, pp. 227–237.

[243] H. B. Russell, A. N. Andriotis, M. Menon, J. B. Jasinski, A. Martinez-Garcia and M. K. Sunkara, *Scientific Reports*, 2016, **6**, 20822.

[244] F. Mezrag, N. Bouarissa and M. Boucenna, *Optik*, 2016, **127**, 1167–1170.

[245] S. Saib and N. Bouarissa, *Journal of Electronic Materials*, 2022, **51**, 3758–3765.

[246] M. Wu, H. Wu, S. Xie, W. Wang and G. Li, *CrystEngComm*, 2023, **25**, 6650–6659.

[247] The Materials Project, *mp-20351: InP (cubic, F-43m, 216)*, https://legacy.materialsproject.org/materials/mp-20351/, [Accessed: 2024-03-13].

[248] M. Löw, M. Guidat, J. Kim and M. M. May, *RSC Advances*, 2022, **12**, 32756–32764.

[249] E. O. Kane, *Journal of Physics and Chemistry of Solids*, 1957, **1**, 249–261.

[250] American Elements, *Safety Data Sheet - Indium Antimonide*, `https://www.americanelements.com/indium-antimonide-1312-41-0`, [Accessed: 2024-01-25].

[251] The Materials Project, *mp-8062: SiC (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-8062/`, [Accessed: 2024-03-13].

[252] J. Jian and J. Sun, *Solar RRL*, 2020, **4**, 2000111.

[253] *Safety Data Sheet - Silicon Carbide*, 2021, Version 3.

[254] The Materials Project, *mp-10695: ZnS (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-10695/`, [Accessed: 2024-03-13].

[255] ThermoFisher Scientific, *Safety Data Sheet - Zinc Sulphide*, February 2020, Version 2, 1314-98-3.

[256] A. Hassan, R. Liaquat, N. Iqbal, G. Ali, X. Fan, Z. Hu, M. Anwar and A. Ahmad, *Journal of Electroanalytical Chemistry*, 2021, **889**, 115223.

[257] G.-J. Lee, S. Anandan, S. J. Masten and J. J. Wu, *Renewable Energy*, 2016, **89**, 18–26.

[258] The Materials Project, *mp-1190: ZnSe (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-1190/`, [Accessed: 2024-03-13].

[259] ThermoFisher Scientific, *Safety Data Sheet - Zinc Selenide*, December 2021, Version 4, 1315-09-9.

[260] E. Sitara, H. Nasir, A. Mumtaz, M. F. Ehsan, M. Sohail, S. Iram, S. A. B. Bukhari, S. Ullah, T. Akhtar and A. Iqbal, *International Journal of Hydrogen Energy*, 2021, **46**, 25424–25435.

[261] The Materials Project, *mp-2176: ZnTe (cubic, F-43m, 216)*, `https://legacy.materialsproject.org/materials/mp-2176/`, [Accessed: 2024-03-13].

[262] *Safety Data Sheet - Zinc Telluride*, 2020, Version 2.

[263] M. Dong, W. Li, C. Ren, X. Ma, X. Li, S. Liu, Y. Li and S. Chen, *International Journal of Hydrogen Energy*, 2022, **47**, 11500–11510.

[264] The Materials Project, *mp-149: Si (cubic, Fd-3m, 227)*, `https://legacy.materialsproject.org/materials/mp-149/`, [Accessed: 2024-03-13].

[265] *Safety Data Sheet - Silicon*, 2021, Version 4.

[266] T. S. Teitsworth, D. J. Hill, S. R. Litvin, E. T. Ritchie, J.-S. Park, J. P. Custer, A. D. Taggart, S. R. Bottum, S. E. Morley, S. Kim, J. R. McBride, J. M. Atkin and J. F. Cahoon, *Nature*, 2023, **614**, 270–274.

[267] C. Ekuma, M. Jarrell, J. Moreno and D. Bagayoko, *Physics Letters A*, 2013, **377**, 2172–2176.

[268] K.-Y. Yoon, J. Park, M. Jung, S.-G. Ji, H. Lee, J. H. Seo, M.-J. Kwak, S. Il Seok, J. H. Lee and J.-H. Jang, *Nature Communications*, 2021, **12**, 4309.

[269] The Materials Project, *mp-22862: NaCl (cubic, Fm-3m, 225)*, `https://legacy.materialsproject.org/materials/mp-22862/`, [Accessed: 2024-03-13].

[270] I. S. Messaoudi, A. Zaoui and M. Ferhat, *physica status solidi (b)*, 2015, **252**, 490–495.

[271] ThermoFisher Scientific, *Safety Data Sheet - Sodium Chloride*, October 2023, Version 8, 7647-14-5.

[272] The Materials Project, *mp-1265: MgO (cubic, Fm-3m, 225)*, `https://legacy.materialsproject.org/materials/mp-1265/`, [Accessed: 2024-03-13].

[273] J. Mangaiyarkkarasi, S. Sasikumar, V. Meenakumari and J. S. Revathy, *MRS Advances*, 2023.

[274] Y. Li, Y.-K. Peng, L. Hu, J. Zheng, D. Prabhakaran, S. Wu, T. J. Puchtler, M. Li, K.-Y. Wong, R. A. Taylor and S. C. E. Tsang, *Nature Communications*, 2019, **10**, 4421.

[275] N. Mao and J.-X. Jiang, *Applied Surface Science*, 2019, **476**, 144–150.

[276] ThermoFisher Scientific, *Safety Data Sheet - Magnesium oxide*, December 2021, Version 4, 1309-48-4.

[277] M. Rizwan, Z. Usman, M. Shakil, S. S. A. Gillani, S. Azeem, H. B. Jin, C. B. Cao, R. F. Mehmood, G. Nabi and M. A. Asghar, *Materials Research Express*, 2020, **7**, 015920.

[278] K. Wang, Z. Jin, L. Liang, H. Bian, D. Bai, H. Wang, J. Zhang, Q. Wang and S. Liu, *Nature Communications*, 2018, **9**, 4544.

[279] S. Cheng and H. Zhong, *The Journal of Physical Chemistry Letters*, 2022, **13**, 2281–2290.

[280] F. D. Murnaghan, *Proceedings of the National Academy of Sciences*, 1944, **30**, 244–247.

[281] WMD-group, *Macrodensity*, `https://github.com/WMD-group/MacroDensity`, [Accessed: 2024-03-14].

[282] M. J. Powers, M. C. Benjamin, L. M. Porter, R. J. Nemanich, R. F. Davis, J. J. Cuomo, G. L. Doll and S. J. Harris, *Applied Physics Letters*, 1995, **67**, 3912–3914.

[283] *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 1986, **209**, 417–428.

[284] F. Zu, C. M. Wolff, M. Ralaiarisoa, P. Amsalem, D. Neher and N. Koch, *ACS Applied Materials & Interfaces*, 2019, **11**, 21578–21583.

[285] T. W. Kasel and C. H. Hendon, *ACS Applied Energy Materials*, 2020, **3**, 10328–10332.

[286] K. T. Butler, J. M. Frost and A. Walsh, *Materials Horizons*, 2015, **2**, 228–231.

[287] E. Welch and L. Scolfaro, *Computational Condensed Matter*, 2022, **30**, e00653.

[288] F. Gygi, *Journal of Chemical Theory and Computation*, 2023, **19**, 1300–1309.

[289] M. I. Dykman and E. I. Rashba, *Physics Today*, 2015, **68**, 10–11.

[290] G. P. Agrawal and N. K. Dutta, in *Semiconductor Lasers*, ed. G. P. Agrawal and N. K. Dutta, Springer US, Boston, MA, 1993, pp. 74–146.

[291] L. Li and E. A. Carter, *Journal of the American Chemical Society*, 2019, **141**, 10451–10461.

[292] J. M. Marin-Beloqui, K. J. Fallon, H. Bronstein and T. M. Clarke, *The Journal of Physical Chemistry Letters*, 2019, **10**, 3813–3819.

[293] Y. Chen, S. Ghosh, X. Liu, I. V. Zozoulenko, M. Fahlman and S. Braun, *The Journal of Physical Chemistry C*, 2021, **125**, 937–945.

[294] D. Emin, *Physics Today*, 1982, **35**, 34–40.

[295] C. Franchini, M. Reticcioli, M. Setvin and U. Diebold, *Nature Reviews Materials*, 2021, **6**, 560–586.

[296] P. Cox and F. Cox, *Transition Metal Oxides: An Introduction to Their Electronic Structure and Properties*, Clarendon Press, 1992.

[297] J. H. Bombile, S. Shetty, M. J. Janik and S. T. Milner, *Physical Chemistry Chemical Physics*, 2020, **22**, 4032–4042.

[298] D. Emin, *Physical Review B*, 1993, **48**, 13691–13702.

[299] R. A. Marcus, *The Journal of Chemical Physics*, 1956, **24**, 966–978.

[300] R. A. Marcus, *The Journal of Chemical Physics*, 1957, **26**, 867–871.

[301] R. A. Marcus, *The Journal of Chemical Physics*, 1957, **26**, 872–877.

[302] R. A. Marcus, *The Journal of Chemical Physics*, 1965, **43**, 679–701.

[303] K. J. Laidler and M. C. King, *The Journal of Physical Chemistry*, 1983, **87**, 2657–2664.

[304] A. Alkauskas, M. D. McCluskey and C. G. Van De Walle, *Journal of Applied Physics*, 2016, **119**, 181101.

[305] Z. Wang, C. Brock, A. Matt and K. H. Bevan, *Physical Review B*, 2017, **96**, 125150.

[306] T. D. Pham and N. A. Deskins, *Journal of Chemical Theory and Computation*, 2020, **16**, 5264–5278.

[307] J. Hubbard and B. H. Flowers, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1963, **276**, 238–257.

[308] M. Capdevila-Cortada, Z. Łodziana and N. López, *ACS Catalysis*, 2016, **6**, 8370–8379.

[309] J. P. Allen and G. W. Watson, *Physical Chemistry Chemical Physics*, 2014, **16**, 21016–21031.

[310] N. A. Deskins, R. Rousseau and M. Dupuis, *The Journal of Physical Chemistry C*, 2014, **118**, 13326–13327.

[311] H. Sezen, H. Shang, F. Bebensee, C. Yang, M. Buchholz, A. Nefedov, S. Heissler, C. Carbogno, M. Scheffler, P. Rinke and C. Wöll, *Nature Communications*, 2015, **6**, 6901.

[312] Doped, *Tips & Tricks*, `https://doped.readthedocs.io/en/latest/Tips.html`, [Accessed: 2024-01-17].

[313] I. Mosquera-Lois, S. R. Kavanagh, A. Walsh and D. O. Scanlon, *Journal of Open Source Software*, 2022, **7**, 4817.

[314] I. Mosquera-Lois, S. R. Kavanagh, A. Walsh and D. O. Scanlon, *npj Computational Materials*, 2023, **9**, 25.

[315] ThermoFisher Scientific, *Safety Data Sheet - Titanium(IV) oxide*, October 2023, Version 11, 13463-67-7.

[316] Y. Mi and Y. Weng, *Scientific Reports*, 2015, **5**, 11482.

[317] X. Zhang, H. Cui, M. Humayun, Y. Qu, N. Fan, X. Sun and L. Jing, *Scientific Reports*, 2016, **6**, 21430.

[318] A. R. Elmaslmane, M. B. Watkins and K. P. McKenna, *Journal of Chemical Theory and Computation*, 2018, **14**, 3740–3751.

[319] P. Deák, B. Aradi and T. Frauenheim, *Physical Review B*, 2012, **86**, 195206.

[320] H. H. Pham and L.-W. Wang, *Physical Chemistry Chemical Physics*, 2015, **17**, 541–550.

[321] F. Fasulo, G. Piccini, A. B. Muñoz-García, M. Pavone and M. Parrinello, *The Journal of Physical Chemistry C*, 2022, **126**, 15752–15758.

[322] R. Nebel, K. M. Macounová, H. Tarábková, L. Kavan and P. Krtil, *The Journal of Physical Chemistry C*, 2019, **123**, 10857–10867.

[323] N. Liu, S. Mohajernia, N. T. Nguyen, S. Hejazi, F. Plass, A. Kahnt, T. Yokosawa, A. Osvet, E. Spiecker, D. M. Guldi and P. Schmuki, *ChemSusChem*, 2020, **13**, 4937–4944.

[324] N. S. Azhar, M. F. M. Taib, O. H. Hassan, M. Z. A. Yahya and A. M. M. Ali, *Materials Research Express*, 2017, **4**, 034002.

[325] C. Xia, Y. Jia, M. Tao and Q. Zhang, *Physics Letters A*, 2013, **377**, 1943–1947.

[326] C. S. Ahart, K. M. Rosso and J. Blumberger, *Journal of the American Chemical Society*, 2022, **144**, 4623–4632.

[327] E. N. Maslen, V. A. Streltsov, N. R. Streltsova and N. Ishizawa, *Acta Crystallographica Section B Structural Science*, 1994, **50**, 435–441.

[328] X. Zong and C. Li, in *Metal Oxides in Heterogeneous Catalysis*, Elsevier, 2018, pp. 355–399.

[329] ThermoFisher Scientific, *Safety Data Sheet - Iron(III) oxide*, Decemeber 2021, Version 4, 116-500.

[330] E. Pastor, J.-S. Park, L. Steier, S. Kim, M. Grätzel, J. R. Durrant, A. Walsh and A. A. Bakulin, *Nature Communications*, 2019, **10**, 3962.

[331] The Materials Project, *Materials Data on PtO2 by Materials Project*, 2020, `https://www.osti.gov/servlets/purl/1189224/`, [Accessed: 2024-01-17].

[332] M. M. Abdullah, F. M. Rajab and S. M. Al-Abbas, *AIP Advances*, 2014, **4**, 027121.

[333] ThermoFisher Scientific, *Safety Data Sheet - Chromium(III) oxide*, February 2020, Version 2, 1308-38-9.

[334] K. Sekizawa, K. Oh-ishi and T. Morikawa, *Dalton Transactions*, 2020, **49**, 659–666.

[335] S. N. F. Mohd Nasir and M. A. Mat-Teridi, *Journal of Sol-Gel Science and Technology*, 2020, **93**, 1–5.

[336] Z. Fan, M. Zhu, S. Pan, J. Ge and L. Hu, *Ceramics International*, 2021, **47**, 13655–13659.

[337] M. Dehn, J. Shenton, S. Holenstein, Q. Meier, D. Arseneau, D. Cortie, B. Hitti, A. Fang, W. MacFarlane, R. McFadden, G. Morris, Z. Salman, H. Luetkens, N. Spaldin, M. Fechner and R. Kiefl, *Physical Review X*, 2020, **10**, 011036.

[338] The Materials Project, *mp-23506: VBiO4 (monoclinic, C2/c, 15)*, `https://legacy.materialsproject.org/materials/mp-504878/`, [Accessed: 2024-03-14].

[339] M. A., M. J., M. Ashokkumar and P. Arunachalam, *Applied Catalysis A: General*, 2018, **555**, 47–74.

[340] ThermoFisher Scientific, *Safety Data Sheet - Bismuth vanadium oxide*, February 2021, Version 2, ALFAA39220.

[341] J. Wiktor, F. Ambrosio and A. Pasquarello, *ACS Energy Letters*, 2018, **3**, 1693–1697.

[342] ThermoFisher Scientific, *Safety Data Sheet - Bismuth(III) oxide*, February 2020, Version 2, 1304-76-3.

[343] S. H. Hsieh, G. J. Lee, C. Y. Chen, J. H. Chen, S. H. Ma, T. L. Horng, K. H. Chen and J. J. Wu, *Journal of Nanoscience and Nanotechnology*, 2012, **12**, 5930–5936.

[344] M. Karimi Sahnesarayi, H. Sarpoolaky and S. Rastegari, *International Journal of Hydrogen Energy*, 2022, **47**, 7214–7227.

[345] V. P. Zhukov, V. M. Zhukovskii, V. M. Zainullina and N. I. Medvedeva, *Journal of Structural Chemistry*, 1999, **40**, 831–837.

[346] M. M. El-Desoky, N. A. Gazouly, A. E. Hannor and H. A. Yousef, *Journal of Materials Science: Materials in Electronics*, 2023, **34**, 829.

[347] T. Gujar, V. Shinde, C. Lokhande, R. Mane and S.-H. Han, *Applied Surface Science*, 2005, **250**, 161–167.

[348] ThermoFisher Scientific, *Safety Data Sheet - Cadmium oxide*, December 2021, Version 7, 1306-19-0.

[349] M. Burbano, D. O. Scanlon and G. W. Watson, *Journal of the American Chemical Society*, 2011, **133**, 15065–15072.

[350] I. Ahmad, I. Shahid, A. Ali, S. Zeb, L. Gao and J. Cai, *New Journal of Chemistry*, 2021, **45**, 17699–17708.

[351] S. Karmakar, H. S. Mohanty and D. Behera, *The European Physical Journal Plus*, 2021, **136**, 1038.

[352] ThermoFisher Scientific, *Safety Data Sheet - Cobalt(II,III) oxide*, February 2020, Version 3, 1308-06-1.

[353] V. Singh, M. Kosa, K. Majhi and D. T. Major, *Journal of Chemical Theory and Computation*, 2015, **11**, 64–72.

[354] T. Hong, Z. Liu, X. Zheng, J. Zhang and L. Yan, *Applied Catalysis B: Environmental*, 2017, **202**, 454–459.

[355] ThermoFisher Scientific, *Safety Data Sheet - Copper(I) oxide*, December 2021, Version 5, 1317-39-1.

[356] The Materials Project, *Materials Data on Cu2O by Materials Project*, 2020, `https://www.osti.gov/servlets/purl/1207131/`, [Accessed: 2024-01-19].

[357] B. K. Meyer, A. Polity, D. Reppin, M. Becker, P. Hering, B. Kramm, P. J. Klar, T. Sander, C. Reindl, C. Heiliger, M. Heinemann, C. Müller and C. Ronning, in *Semiconductors and Semimetals*, ed. B. G. Svensson, S. J. Pearton and C. Jagadish, Elsevier, 2013, vol. 88 of Oxide Semiconductors, pp. 201–226.

[358] L. Pan, Y. Liu, L. Yao, Dan Ren, K. Sivula, M. Grätzel and A. Hagfeldt, *Nature Communications*, 2020, **11**, 318.

[359] D. O. Scanlon, B. J. Morgan and G. W. Watson, *The Journal of Chemical Physics*, 2009, **131**, 124703.

[360] ThermoFisher Scientific, *Safety Data Sheet - Copper(II) oxide*, December 2021, Version 5, 1317-38-0.

[361] Q. Zhang, K. Zhang, D. Xu, G. Yang, H. Huang, F. Nie, C. Liu and S. Yang, *Progress in Materials Science*, 2014, **60**, 208–337.

[362] J. B. Forsyth, P. J. Brown and B. M. Wanklyn, *Journal of Physics C: Solid State Physics*, 1988, **21**, 2917–2929.

[363] N. H. Lam, N. T. N. Truong, N. Le, K.-S. Ahn, Y. Jo, C.-D. Kim and J. H. Jung, *Scientific Reports*, 2023, **13**, 5776.

[364] N. Hoang Lam, N. Tam Nguyen Truong, K.-S. Ahn, Y. Jo, S. Beom Kang, N. Huu Hieu, S. F. Shaikh, C.-D. Kim, M. Lee and J. Hak Jung, *FlatChem*, 2024, **43**, 100600.

[365] T. J. Smart, A. C. Cardiel, F. Wu, K.-S. Choi and Y. Ping, *npj Computational Materials*, 2018, **4**, 1–8.

[366] ThermoFisher Scientific, *Safety Data Sheet - Indium(III) oxide*, February 2020, Version 2, 1312-43-2.

[367] P. D. C. King, T. D. Veal, F. Fuchs, C. Y. Wang, D. J. Payne, A. Bourlange, H. Zhang, G. R. Bell, V. Cimalla, O. Ambacher, R. G. Egdell, F. Bechstedt and C. F. McConville, *Physical Review B*, 2009, **79**, 205211.

[368] H. Xu, S. Chen, K. Wang and X. Wang, *International Journal of Hydrogen Energy*, 2022, **47**, 38219–38228.

[369] Q. Hou, J. Buckeridge, T. Lazauskas, D. Mora-Fonz, A. A. Sokol, S. M. Woodley and C. R. A. Catlow, *Journal of Materials Chemistry C*, 2018, **6**, 12386–12395.

[370] Q. Qu, W.-B. Zhang, K. Huang and H.-M. Chen, *Computational Materials Science*, 2017, **130**, 242–248.

[371] S. Swathi, G. Ravi, R. Yuvakkumar, S. I. Hong, E. S. Babu, D. Velauthapillai and P. Kumar, *Ceramics International*, 2020, **46**, 23218–23222.

[372] J. V. Silveira, J. A. Batista, G. D. Saraiva, J. Mendes Filho, A. G. Souza Filho, S. Hu and X. Wang, *Vibrational Spectroscopy*, 2010, **54**, 179–183.

[373] ThermoFisher Scientific, *Safety Data Sheet - Molybdenum(VI) oxide*, December 2021, Version 6, 1313-27-5.

[374] Y. Natanzon, A. Azulay and Y. Amouyal, *Israel Journal of Chemistry*, 2020, **60**, 768–786.

[375] H. Ding, H. Lin, B. Sadigh, F. Zhou, V. Ozoliņš and M. Asta, *The Journal of Physical Chemistry C*, 2014, **118**, 15565–15572.

[376] H. A. Tahini, X. Tan, S. N. Lou, J. Scott, R. Amal, Y. H. Ng and S. C. Smith, *ACS Applied Materials & Interfaces*, 2016, **8**, 10911–10917.

[377] ThermoFisher Scientific, *Safety Data Sheet - Nickel(II) oxide*, December 2021, Version 4, 1313-99-1.

[378] The Materials Project, *Materials Data on NiO by Materials Project*, 2017, `https://www.osti.gov/servlets/purl/1193796/`, [Accessed: 2024-01-19].

[379] C. Díaz-Guerra, A. Remón, J. A. García and J. Piqueras, *physica status solidi (a)*, 1997, **163**, 497–503.

[380] N. M. Hosny, *Polyhedron*, 2011, **30**, 470–476.

[381] O. Seo, A. Tayal, J. Kim, C. Song, Y. Chen, S. Hiroi, Y. Katsuya, T. Ina, O. Sakata, Y. Ikeya, S. Takano, A. Matsuda and M. Yoshimoto, *Scientific Reports*, 2019, **9**, 4304.

[382] E. Ressouche, N. Kernavanois, L.-P. Regnault and J.-Y. Henry, *Physica B: Condensed Matter*, 2006, **385-386**, 394–397.

[383] J. Liu, Y. Zheng, Y. Jiao, Z. Wang, Z. Lu, A. Vasileff and S. Qiao, *Small*, 2018, **14**, 1704073.

[384] M. Gong, W. Zhou, M.-C. Tsai, J. Zhou, M. Guan, M.-C. Lin, B. Zhang, Y. Hu, D.-Y. Wang, J. Yang, S. J. Pennycook, B.-J. Hwang and H. Dai, *Nature Communications*, 2014, **5**, 4695.

[385] J. Yu, K. M. Rosso and S. M. Bruemmer, *The Journal of Physical Chemistry C*, 2012, **116**, 1948–1954.

[386] ThermoFisher Scientific, *Safety Data Sheet - Tin(II) oxide*, February 2020, Version 2, 21651-19-4.

[387] W. Zhou and N. Umezawa, *Physical Chemistry Chemical Physics*, 2015, **17**, 17816–17820.

[388] Q. Zheng, X. Zeng, W. Xia, Q. Lu, J. Sun and M. Luo, *Materials Research Express*, 2020, **7**, 065501.

[389] J. Raj, S. Biswas and V. Kamble, *Applied Surface Science*, 2023, **624**, 157078.

[390] ThermoFisher Scientific, *Safety Data Sheet - Strontium Titanate*, January 2022, Version 1, 1260-59-2.

[391] The Materials Project, *Materials Data on SrTiO3 by Materials Project*, 2020, `https://www.osti.gov/servlets/purl/1263154/`, [Accessed: 2024-01-19].

[392] C. Wang, H. Qiu, T. Inoue and Q. Yao, *International Journal of Hydrogen Energy*, 2014, **39**, 12507–12514.

[393] Y. Fan, Y. Liu, H. Cui, W. Wang, Q. Shang, X. Shi, G. Cui and B. Tang, *Nanomaterials*, 2020, **10**, 2572.

[394] A. Janotti, J. B. Varley, M. Choi and C. G. Van De Walle, *Physical Review B*, 2014, **90**, 085202.

[395] ThermoFisher Scientific, *Safety Data Sheet - Tungsten(VI) oxide*, December 2021, Version 5, 1314-35-8.

[396] S. Zhuiykov and E. Kats, *Nanoscale Research Letters*, 2014, **9**, 401.

[397] A. Jelinska, K. Bienkowski, M. Jadwiszczak, M. Pisarek, M. Strawski, D. Kurzydlowski, R. Solarska and J. Augustynski, *ACS Catalysis*, 2018, **8**, 10573–10580.

[398] J. C. Alvarez-Quiceno, G. M. Dalpian and J. M. Osorio-Guillén, *physica status solidi (b)*, 2015, **252**, 2290–2295.

[399] J. Tao and T. Liu, *The Journal of Physical Chemistry C*, 2023, **127**, 16204–16210.

[400] ThermoFisher Scientific, *Safety Data Sheet - Vanadium (IV) oxide*, February 2020, Version 3, 12036-21-4.

[401] M. Zayed, A. Elabbar and O. Yassin, *Physica B: Condensed Matter*, 2020, **582**, 411887.

[402] T. Bredow and B. Stahl, *Chemical Physics Letters*, 2018, **695**, 28–33.

[403] K.-H. Kim, K.-H. Kim, W. Choi, Y.-M. Kim, S.-H. Hong and Y.-H. Choi, *Nanoscale*, 2022, **14**, 8281–8290.

[404] S. Singh, T. A. Abtew, G. Horrocks, C. Kilcoyne, P. M. Marley, A. A. Stabile, S. Banerjee, P. Zhang and G. Sambandamurthy, *Physical Review B*, 2016, **93**, 125132.

[405] X. Liu, T. Yang, J. Wu, M. Sun, M. Chen, C. S. Tang, K. Han, D. Zhou, S. Zeng, S. Sun, S. Li, M. Yang, M. B. H. Breese, C. Cai, T. Venkatesan, A. T. S. Wee and X. Yin, *Optical detection of small polarons in vanadium dioxide and their critical role in mediating metal-insulator transition*, 2023, `http://arxiv.org/abs/2312.17419`, arXiv:2312.17419 [cond-mat, physics:physics].

[406] J. Kegel, I. M. Povey and M. E. Pemble, *Nano Energy*, 2018, **54**, 409–428.

[407] ThermoFisher Scientific, *Safety Data Sheet - Zinc oxide*, December 2021, Version 6, 1314-13-2.

[408] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, *Journal of Physics: Condensed Matter*, 2017, **29**, 273002.

[409] ASE documentation, *Tools for defect calculations*, `https://wiki.fysik.dtu.dk/ase/tutorials/defects/defects.html#id7`, [Accessed: 2024-03-14].

[410] Vaspwiki, *MAGMOM*, `https://www.vasp.at/wiki/index.php/MAGMOM`, [Accessed: 2023-06-02].

[411] Deskins-group, *Polaron-Strategies*, `https://github.com/Deskins-group/Other-Files`, [Accessed: 2023-06-02].

[412] ASE documentation, *The Atoms object - rattle*, `https://wiki.fysik.dtu.dk/ase/ase/atoms.html#ase.Atoms.rattle`, [Accessed: 2023-06-02].

[413] Vaspwiki, *NELECT*, `https://www.vasp.at/wiki/index.php/NELECT`, [Accessed: 2023-06-02].

[414] Vaspwiki, *ALGO*, `https://www.vasp.at/wiki/index.php/ALGO`, [Accessed: 2023-06-02].

[415] Vaspwiki, *ISMEAR*, `https://www.vasp.at/wiki/index.php/ISMEAR`, [Accessed: 2023-06-02].

[416] Vaspwiki, *SIGMA*, `https://www.vasp.at/wiki/index.php/SIGMA`, [Accessed: 2023-06-02].

[417] Vaspwiki, *PRECFOCK*, `https://www.vasp.at/wiki/index.php/PRECFOCK`, [Accessed: 2023-06-02].

[418] Vaspwiki, *LSUBROT*, `https://www.vasp.at/wiki/index.php/LSUBROT`, [Accessed: 2023-06-02].

[419] Vaspwiki, *LREAL*, `https://www.vasp.at/wiki/index.php/LREAL`, [Accessed: 2023-06-02].

[420] Vaspwiki, *KPAR*, `https://www.vasp.at/wiki/index.php/KPAR`, [Accessed: 2023-06-02].

[421] lucydot, *vesta_vectors*, `https://github.com/lucydot/vesta_vectors`, [Accessed: 2024-03-14].

[422] J. J. Carey, J. A. Quirk and K. P. McKenna, *The Journal of Physical Chemistry C*, 2021, **125**, 12441–12450.

[423] F. Wu and Y. Ping, *Journal of Materials Chemistry A*, 2018, **6**, 20025–20036.

[424] R. F. W. Bader, *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 2001, **105**, 276–283.

[425] S. Kokott, S. V. Levchenko, P. Rinke and M. Scheffler, *New Journal of Physics*, 2018, **20**, 033023.

[426] A. R. Elmaslmane, J. Wetherell, M. J. P. Hodgson, K. P. McKenna and R. W. Godby, *Physical Review Materials*, 2018, **2**, 040801.

[427] M. F. Lichterman, A. I. Carim, M. T. McDowell, S. Hu, H. B. Gray, B. S. Brunschwig and N. S. Lewis, *Energy & Environmental Science*, 2014, **7**, 3334–3337.

[428] N. G. Krishna, R. P. George and J. Philip, *Corrosion Science*, 2020, **172**, 108678.

[429] B. Li, L. Zhang, Y. Li, H. Li, L. Zhou, C. Liang and H. Wang, *Chemistry Letters*, 2019, **48**, 1355–1357.

[430] L. Zhou, A. Shinde, M.-C. Chang, R. Bruce Van Dover, M. O. Thompson and J. M. Gregoire, *Journal of Materials Chemistry A*, 2023, **11**, 25262–25267.

[431] A. Held, P. D. P. Taylor and C. Ingelbrecht, *Fresenius' Journal of Analytical Chemistry*, 1999, **364**, 437–439.

[432] A. Zunger, S.-H. Wei, L. G. Ferreira and J. E. Bernard, *Physical Review Letters*, 1990, **65**, 353–356.

[433] A. Van De Walle, P. Tiwary, M. De Jong, D. Olmsted, M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen and Z.-K. Liu, *Calphad*, 2013, **42**, 13–18.

[434] ICET, *Special quasirandom structures*, `https://icet.materialsmodeling.org/advanced_topics/sqs_generation.html`, [Accessed: 2023-06-02].

[435] A. Van De Walle, R. Sun, Q.-J. Hong and S. Kadkhodaei, *Calphad*, 2017, **58**, 70–81.

[436] *Chemistry3: introducing inorganic, organic and physical chemistry*, ed. A. Burrows, Oxford University Press, Oxford ; New York, 2009.

[437] A. Balandin, in *Advances in Catalysis*, Elsevier, 1969, vol. 19, pp. 1–210.

[438] H. Knözinger and K. Kochloefl, in *Ullmann's Encyclopedia of Industrial Chemistry*, ed. Wiley-VCH Verlag GmbH & Co. KGaA, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2003, p. a05_313.

[439] T. A. Batchelor, J. K. Pedersen, S. H. Winther, I. E. Castelli, K. W. Jacobsen and J. Rossmeisl, *Joule*, 2019, **3**, 834–845.

[440] P. G. Moses and C. G. Van De Walle, *Applied Physics Letters*, 2010, **96**, 021908.

[441] J.-C. Wu, J. Zheng, C. L. Zacherl, P. Wu, Z.-K. Liu and R. Xu, *The Journal of Physical Chemistry C*, 2011, **115**, 19741–19748.

[442] Y. Ping, G. Galli and W. A. Goddard, *The Journal of Physical Chemistry C*, 2015, **119**, 11570–11577.

[443] ThermoFisher Scientific, *Safety Data Sheet - Iridium (IV) Oxide*, February 2020, Version 2, 12030-49-8.

[444] Y. Z. Abdullahi, *Computational Condensed Matter*, 2021, **29**, e00614.

[445] N. Mehtougui, D. Rached, R. Khenata, H. Rached, M. Rabah and S. Bin-Omran, *Materials Science in Semiconductor Processing*, 2012, **15**, 331–339.

[446] J. Xu, T. Jarlborg and A. Freeman, *Physical Review B*, 1989, **40**, 7939–7947.

[447] J. S. De Almeida and R. Ahuja, *Physical Review B*, 2006, **73**, 165102.

[448] S. W. Lovesey, D. D. Khalyavin and G. van der Laan, *Physical Review B*, 2022, **105**, 014403.

[449] K. Du, L. Zhang, J. Shan, J. Guo, J. Mao, C.-C. Yang, C.-H. Wang, Z. Hu and T. Ling, *Nature Communications*, 2022, **13**, 5448.

[450] ThermoFisher Scientific, *Safety Data Sheet - Ruthenium (IV) Oxide*, December 2021, Version 6, 12036-10-1.

[451] H. A. H. Mohammed, G. M. Dongho-Nguimdo and D. P. Joubert, *Physica E: Low-dimensional Systems and Nanostructures*, 2021, **127**, 114514.

[452] R. D. Shannon, *Solid State Communications*, 1968, **6**, 139–143.

[453] A. Valizadeh, P. Aleshkevych and M. M. Najafpour, *Inorganic Chemistry*, 2022, **61**, 613–621.

[454] ThermoFisher Scientific, *Safety Data Sheet - Platinum (IV) Oxide*, December 2021, Version 5, 1314-15-4.

[455] R. Wang, Q. Li, W. Li, P. Jiang, R. Cong and T. Yang, *Chemical Communications*, 2021, **57**, 536–539.

[456] S. Chae, J. Lee, K. A. Mengle, J. T. Heron and E. Kioupakis, *Applied Physics Letters*, 2019, **114**, 102104.

[457] ThermoFisher Scientific, *Safety Data Sheet - Germanium (IV) Oxide*, February 2020, Version 2, 1310-53-8.

[458] R. Devi, V. Kumar, S. Kumar, M. Bulla, S. Sharma and A. Sharma, *Applied Sciences*, 2023, **13**, 7907.

[459] C. Zhou, J. Wang, X. Liu, F. Chen, Y. Di, S. Gao and Q. Shi, *New Journal of Chemistry*, 2018, **42**, 8400–8407.

[460] E. Cockayne and L. Li, *Chemical Physics Letters*, 2012, **544**, 53–58.

[461] Y.-F. Li and Z.-P. Liu, *Journal of the American Chemical Society*, 2018, **140**, 1783–1792.

[462] ThermoFisher Scientific, *Safety Data Sheet - Manganese (IV) Oxide*, December 2021, Version 5, 1313-13-9.

[463] G. S. L. Fabris, D. H. M. Azevedo, A. C. Alves, C. A. Paskocimas, J. R. Sambrano and J. M. M. Cordeiro, *Physica E: Low-dimensional Systems and Nanostructures*, 2022, **136**, 115037.

[464] G. Flora, D. Gupta and A. Tiwari, *Interdisciplinary Toxicology*, 2012, **5**, 47–58.

[465] ThermoFisher Scientific, *Safety Data Sheet - Lead (IV) Oxide*, December 2021, Version 5, 1309-60-0.

[466] M. Karmaoui, A. B. Jorge, P. F. McMillan, A. E. Aliev, R. C. Pullar, J. A. Labrincha and D. M. Tobaldi, *ACS Omega*, 2018, **3**, 13227–13238.

[467] M. Sun, Z. Gong, Z. Zhang, Y. Li, D. Xie, F. Wu and R. Li, *Thin Solid Films*, 2021, **732**, 138794.

[468] S. K. Jain, M. Fazil, N. A. Pandit, S. A. Ali, F. Naaz, H. Khan, A. Mehtab, J. Ahmed and T. Ahmad, *ACS Omega*, 2022, **7**, 14138–14147.

[469] K. A. Winship, *Adverse Drug Reactions and Acute Poisoning Reviews*, 1988, **7**, 19–38.

[470] CDC, *NIOSH Pocket Guide to Chemical Hazards - Tin(IV) oxide (as Sn)*, `https://www.cdc.gov/niosh/npg/npgd0616.html`, [Accessed: 2024-01-17].

[471] A. Van De Walle, *Calphad*, 2009, **33**, 266–278.

[472] A. Van De Walle, M. Asta and G. Ceder, *Calphad*, 2002, **26**, 539–553.

[473] A. Walle and G. Ceder, *Journal of Phase Equilibria*, 2002, **23**, 348–359.

[474] ATAT, *Alloy Theoretic Automated Toolkit - Home Page*, http://alum.mit.edu/www/avdw/atat/, [Accessed: 2024-03-14].

[475] pymatgen 2023.5.10 documentation, *pymatgen.core.surface module*, https://pymatgen.org/pymatgen.core.surface.html, [Accessed: 2023-06-02].

[476] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson and G. Ceder, *Computational Materials Science*, 2013, **68**, 314–319.

[477] H. A. Zakaryan, A. G. Kvashnin and A. R. Oganov, *Scientific Reports*, 2017, **7**, 10357.

[478] Vaspwiki, *IDIPOL*, https://www.vasp.at/wiki/index.php/IDIPOL, [Accessed: 2024-03-14].

[479] Vaspwiki, *Monopole Dipole and Quadrupole Corrections*, https://www.vasp.at/wiki/index.php/Monopole_Dipole_and_Quadrupole_Corrections, [Accessed: 2024-03-14].

[480] G. Makov and M. C. Payne, *Physical Review B*, 1995, **51**, 4014–4022.

[481] J. Neugebauer and M. Scheffler, *Physical Review B*, 1992, **46**, 16067–16080.

[482] Spglib, *Spglib v.2.1.0rc1*, https://spglib.readthedocs.io/en/latest/index.html, [Accessed: 2023-06-02].

[483] B. Hammer, L. B. Hansen and J. K. Nørskov, *Phys. Rev. B*, 1999, **59**, 7413–7421.

[484] A. Togo and I. Tanaka, *Spglib: a software library for crystal symmetry search*, 2018, http://arxiv.org/abs/1808.01590, [Accessed: 2024-01-31].

[485] H. Sato, K. Wakiya, T. Enoki, T. Kiyama, Y. Wakabayashi, H. Nakao and Y. Murakami, *Journal of the Physical Society of Japan*, 2001, **70**, 37–40.

[486] T. Berlijn, P. Snijders, O. Delaire, H.-D. Zhou, T. Maier, H.-B. Cao, S.-X. Chi, M. Matsuda, Y. Wang, M. Koehler, P. Kent and H. Weitering, *Physical Review Letters*, 2017, **118**, 077201.

[487] M. T. Darby and C. S. Cucinotta, *Current Opinion in Electrochemistry*, 2022, **36**, 101118.

[488] S. Biswas, *Fermi-Dirac Statistics*, 2015, arXiv:1402.3370 [physics].

[489] R. P. Feynman, *Physical Review*, 1939, **56**, 340–343.

[490] ASE documentation, *Equation of state (EOS)*, https://wiki.fysik.dtu.dk/ase/tutorials/eos/eos.html, [Accessed: 2024-03-13].

[491] P. W. Atkins and J. De Paula, *Atkins' Physical chemistry*, Oxford University Press, Oxford ; New York, 9th edn., 2010.

[492] R. Santamaria, in *Molecular Dynamics*, ed. R. Santamaria, Springer Nature Switzerland, Cham, 2023, pp. 327–356.

[493] *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 1913, **88**, 428–438.

[494] C. G. Shull, *Reviews of Modern Physics*, 1995, **67**, 753–757.

[495] The Editors of Encyclopedia Britannica, *Electron diffraction - Encyclopedia Britannica*, `https://www.britannica.com/science/electron-diffraction`, [Accessed: 2024-03-13].

[496] R. Berera, R. Van Grondelle and J. T. M. Kennis, *Photosynthesis Research*, 2009, **101**, 105–118.

[497] N. Elgrishi, K. J. Rountree, B. D. McCarthy, E. S. Rountree, T. T. Eisenhart and J. L. Dempsey, *Journal of Chemical Education*, 2018, **95**, 197–206.

[498] I. Fischer and S. T. Pratt, *Physical Chemistry Chemical Physics*, 2022, **24**, 1944–1959.

[499] J. E. Whitten, *Applied Surface Science Advances*, 2023, **13**, 100384.

[500] D. Depla, S. Mahieu and J. Greene, in *Handbook of Deposition Technologies for Films and Coatings*, Elsevier, 2010, pp. 253–296.

[501] L. Sun, G. Yuan, L. Gao, J. Yang, M. Chhowalla, M. H. Gharahcheshmeh, K. K. Gleason, Y. S. Choi, B. H. Hong and Z. Liu, *Nature Reviews Methods Primers*, 2021, **1**, 5.

[502] T. Van Acker, S. Theiner, E. Bolea-Fernandez, F. Vanhaecke and G. Koellensperger, *Nature Reviews Methods Primers*, 2023, **3**, 1–18.

[503] M. Namakka, M. R. Rahman, K. A. M. B. Said, M. Abdul Mannan and A. M. Patwary, *Environmental Nanotechnology, Monitoring & Management*, 2023, **20**, 100900.

[504] E. Kreyszig, *Advanced engineering mathematics*, Wiley, Hoboken, NJ, 9th edn., 2006.

[505] L. Brillouin, *Journal de Physique et le Radium*, 1930, **1**, 377–400.

[506] J. J. Jorgensen, J. E. Christensen, T. J. Jarvis and G. L. W. Hart, *Communications in Computational Physics*, 2022, **31**, 495–515.

Data tables

The following appendix lists the data used in the most pertinent figures presented in the preceding chapters. This is merely for the reader's reference should they require precise values for further study.

A.1 Band Alignment

Compound	Ionisation Potential / eV			
	Reference	**Frensley(1977)**	**VDWM(1987)**	**Harrison(1985)**
BeSe	NaN	NaN	NaN	NaN
CuBr	-4.200	NaN	NaN	NaN
InSb	-4.830	NaN	-5.82	-8.61
AlAs	-4.270	-3.96	-7.40	-9.67
AgI	NaN	NaN	NaN	NaN
BeTe	NaN	NaN	NaN	NaN
GaP	-6.560	-4.12	-7.37	-10.21
CdSe	-6.300	-5.29	NaN	-10.97
SiC	-6.360	NaN	NaN	NaN
BAs	-6.544	NaN	NaN	NaN
CuCl	NaN	NaN	NaN	NaN
GaSb	-4.790	-3.89	-5.98	-8.67
CdTe	-5.720	-4.90	-6.76	-9.80
ZnSe	-6.820	-5.07	-8.23	-11.06
BN	NaN	NaN	NaN	NaN
ZnTe	-5.790	-4.74	-6.87	-9.87
Ge	-4.670	-3.25	-6.25	-8.97
CuI	NaN	NaN	NaN	NaN
BeS	NaN	NaN	NaN	NaN
AlP	-4.930	NaN	-8.09	-10.22
InN	-6.450	NaN	NaN	NaN
GaAs	-5.490	-3.96	-6.81	-9.64
Si	-5.160	-3.16	-7.02	-9.35
AlSb	-5.200	-3.94	-6.44	-8.76
CdS	-6.920	-5.42	NaN	-11.89
InAs	-5.260	NaN	NaN	NaN
BP	NaN	NaN	NaN	NaN
InP	-5.720	-4.58	-7.00	-10.03
ZnS	-7.600	-5.34	-9.13	-12.00

Table A.1: Tabulated data for Figures 3.3, 3.10a, 3.11.

Compound	Lattice Parameter / Å		
	Reference	**LDA**	**PBEsol**
BeSe	5.13900	5.086	5.135
CuBr	5.69050	5.509	5.591
InSb	6.47877	6.470	6.543
AlAs	5.66220	5.638	5.683
AgI	6.50200	6.358	6.453
BeTe	5.62600	5.564	5.614
GaP	5.44950	5.426	5.476
CdSe	6.05000	6.025	6.097
SiC	4.34800	4.331	4.360
BAs	4.77700	4.742	4.779
CuCl	5.40570	5.219	5.301
GaSb	6.09540	6.066	6.130
CdTe	6.47700	6.421	6.499
ZnSe	5.66760	5.575	5.642
BN	3.61500	3.585	3.610
ZnTe	6.10100	6.002	6.072
Ge	5.65754	5.647	5.704
CuI	6.04270	5.866	5.946
BeS	4.86500	4.796	4.840
AlP	5.45100	5.435	5.472
InN	4.98000	5.009	5.058
GaAs	5.56315	5.627	5.685
Si	5.43072	5.404	5.438
AlSb	6.13550	6.122	6.171
CdS	5.83200	5.775	5.839
InAs	6.05838	6.060	6.124
BP	4.53800	4.492	4.523
InP	5.86875	5.879	5.934
ZnS	5.40930	5.304	5.362

Table A.2: Tabulated calculated data and reference[1,2,3] lattice parameter data for Figures 3.9 and 3.10a.

	Sampling Volume / $n \times n \times n$									
Compound	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
AlAs	-4.37	-4.36	-4.35	-4.34	-4.32	-4.30	-4.27	-4.24	-4.19	-4.13
AlP	-4.48	-4.48	-4.47	-4.46	-4.44	-4.41	-4.38	-4.33	-4.28	-4.21
AlSb	-4.16	-4.16	-4.15	-4.14	-4.13	-4.11	-4.08	-4.05	-4.00	-3.95
BAs	-2.78	-2.78	-2.76	-2.74	-2.70	-2.64	-2.56	-2.46	-2.33	-2.16
BN	-2.63	-2.61	-2.58	-2.53	-2.46	-2.36	-2.22	-2.04	-1.79	-1.49
BP	-2.86	-2.85	-2.83	-2.78	-2.72	-2.63	-2.50	-2.34	-2.12	-1.83
BeS	-4.94	-4.94	-4.93	-4.91	-4.88	-4.85	-4.80	-4.74	-4.65	-4.55
BeSe	-4.68	-4.68	-4.66	-4.65	-4.62	-4.58	-4.53	-4.47	-4.38	-4.27
BeTe	-4.23	-4.23	-4.22	-4.21	-4.18	-4.15	-4.10	-4.04	-3.96	-3.86
CdS	-5.12	-5.12	-5.12	-5.11	-5.10	-5.09	-5.08	-5.06	-5.04	-5.01
CdSe	-5.05	-5.05	-5.04	-5.04	-5.03	-5.02	-5.00	-4.98	-4.96	-4.93
CdTe	-4.80	-4.80	-4.80	-4.79	-4.78	-4.77	-4.76	-4.74	-4.72	-4.70
CuBr	-4.26	-4.25	-4.25	-4.24	-4.22	-4.20	-4.17	-4.14	-4.09	-4.03
CuCl	-4.12	-4.12	-4.11	-4.10	-4.08	-4.06	-4.03	-3.99	-3.94	-3.88
CuI	-4.25	-4.25	-4.24	-4.23	-4.22	-4.20	-4.17	-4.13	-4.08	-4.02
GaAs	-4.01	-4.01	-4.00	-3.99	-3.98	-3.95	-3.92	-3.88	-3.84	-3.77
GaP	-4.12	-4.11	-4.11	-4.09	-4.07	-4.04	-4.01	-3.96	-3.91	-3.83
GaSb	-3.82	-3.82	-3.81	-3.80	-3.78	-3.76	-3.73	-3.70	-3.65	-3.59
Ge	-3.47	-3.47	-3.46	-3.45	-3.43	-3.41	-3.37	-3.33	-3.27	-3.21
InAs	-4.22	-4.21	-4.20	-4.19	-4.18	-4.17	-4.15	-4.13	-4.10	-4.07
InN	-3.83	-3.83	-3.82	-3.81	-3.79	-3.78	-3.76	-3.75	-3.73	-3.71
InP	-4.28	-4.28	-4.27	-4.26	-4.25	-4.24	-4.22	-4.20	-4.18	-4.15
InSb	-4.07	-4.07	-4.06	-4.06	-4.05	-4.04	-4.02	-4.00	-3.98	-3.95
Si	-3.72	-3.71	-3.70	-3.68	-3.65	-3.62	-3.57	-3.51	-3.43	-3.33
SiC	-3.05	-3.04	-3.02	-2.99	-2.95	-2.91	-2.85	-2.79	-2.70	-2.60
ZnS	-4.95	-4.95	-4.94	-4.93	-4.91	-4.89	-4.86	-4.83	-4.78	-4.71
ZnSe	-4.85	-4.85	-4.84	-4.83	-4.82	-4.80	-4.77	-4.74	-4.69	-4.64
ZnTe	-4.54	-4.53	-4.52	-4.52	-4.50	-4.48	-4.46	-4.42	-4.38	-4.32

Table A.3: Tabulated data for Figures 3.8 and 3.9.

	Ionisation Potential / eV				Electron Affinity / eV			
Compound	**HSE06**	**SCAN**	**PBEsol**	**LDA**	**HSE06**	**SCAN**	**PBEsol**	**LDA**
BeSe	-5.12	-4.74	-4.68	-4.70	-1.60	-1.63	-2.18	-2.26
CuBr	-5.04	-4.20	-4.26	-4.37	-2.83	-3.32	-3.82	-3.86
InSb	-4.27	-4.04	-4.15	-4.13	-3.93	-4.04	-4.07	-3.66
AlAs	-4.66	-4.38	-4.37	-4.41	-2.51	-2.59	-3.00	-3.06
AgI	-5.65	-5.23	-5.15	-5.19	-3.16	-3.60	-3.95	-4.10
BeTe	-4.55	-4.19	-4.24	-4.25	-1.91	-1.97	-2.39	-2.45
GaP	-4.40	-4.15	-4.12	-4.15	-2.08	-2.24	-2.60	-2.70
CdSe	-5.40	-5.14	-5.04	-5.12	-3.85	-4.20	-4.57	-4.70
SiC	-3.50	-2.95	-3.05	-3.09	-1.17	-1.17	-1.75	-1.73
BAs	-3.03	-2.72	-2.78	-2.81	-1.18	-1.30	-1.70	-1.70
CuCl	-4.97	-4.05	-4.13	-4.28	-2.53	-3.11	-3.59	-3.61
GaSb	-4.05	-3.78	-3.81	-3.86	-3.27	-3.71	-3.81	-3.74
CdTe	-5.09	-4.85	-4.80	-4.86	-3.56	-3.85	-4.17	-4.22
ZnSe	-5.28	-4.96	-4.87	-4.90	-2.80	-3.14	-3.64	-3.63
BN	-3.34	-2.60	-2.62	-2.66	2.54	2.45	1.73	1.79
ZnTe	-4.86	-4.56	-4.54	-4.57	-2.63	-2.92	-3.32	-3.27
Ge	-3.64	-3.47	-3.47	-3.53	-3.07	-3.22	-3.41	-3.42
CuI	-4.88	-4.22	-4.26	-4.34	-2.16	-2.60	-3.09	-3.05
BeS	-5.47	-5.03	-4.95	-4.97	-1.37	-1.41	-1.99	-2.10
AlP	-4.81	-4.50	-4.49	-4.54	-2.47	-2.54	-3.00	-3.09
InN	-4.23	-3.78	-3.88	-3.91	-3.84	-2.48	-3.83	-2.76
GaAs	-4.26	-4.02	-4.01	-4.05	-2.99	-3.32	-3.62	-3.55
Si	-3.93	-3.71	-3.72	-3.75	-2.74	-2.85	-3.23	-3.28
AlSb	-4.41	-4.14	-4.15	-4.22	-2.57	-2.75	-3.02	-3.10
CdS	-5.53	-5.22	-5.12	-5.20	-3.35	-3.72	-4.12	-4.24
InAs	-4.45	-4.25	-4.23	-4.28	-4.12	-3.12	-4.23	-4.28
BP	-3.15	-2.79	-2.85	-2.89	-1.17	-1.24	-1.71	-1.73
InP	-4.54	-4.32	-4.28	-4.34	-3.23	-3.43	-3.80	-3.82
ZnS	-5.45	-5.05	-4.96	-5.00	-2.00	-2.32	-2.88	-2.86

Table A.4: Tabulated data for Figures 3.11 and 3.11.

	Ionicity		
Compound	**Phillips**	**Frensley**	**Bader**
BeSe	0.261	0.12	1.60
CuBr	0.735	0.07	0.50
InSb	0.321	0.21	0.49
AlAs	0.274	0.47	2.02
AgI	0.770	0.10	0.31
BeTe	0.169	0.14	1.52
GaP	0.327	0.36	0.86
CdSe	0.699	0.53	0.81
SiC	0.177	1.12	2.78
BAs	0.002	0.52	0.14
CuCl	0.746	0.20	0.61
GaSb	0.261	0.01	0.39
CdTe	0.717	0.30	0.61
ZnSe	0.630	0.35	0.84
BN	0.256	0.61	2.26
ZnTe	0.609	0.10	0.62
Ge	0.000	0.00	0.00
CuI	0.692	0.11	0.34
BeS	0.286	0.24	1.64
AlP	0.307	0.59	2.12
InN	0.578	1.26	1.34
GaAs	0.310	0.26	0.71
Si	0.000	0.00	0.01
AlSb	0.250	0.19	1.76
CdS	0.685	0.66	0.93
InAs	0.357	0.48	0.74
BP	0.006	0.49	0.68
InP	0.421	0.59	0.85
ZnS	0.623	0.49	0.96

Table A.5: Tabulated data for Figure 3.15.

	Band Gap / eV				
Compound	**Reference**	**HSE06**	**SCAN**	**PBEsol**	**LDA**
BeSe	4-4.5	3.52	3.11	2.50	2.44
CuBr	2.1	2.21	0.88	0.44	0.51
InSb	0.235	0.34	0.00	0.08	0.47
AlAs	2.36	2.15	1.79	1.37	1.35
AgI	1.4	2.49	1.63	1.20	1.09
BeTe	2.7	2.64	2.22	1.85	1.80
GaP	2.26	2.32	1.91	1.52	1.45
CdSe	1.74	1.55	0.94	0.47	0.42
SiC	2.36	2.33	1.78	1.30	1.36
BAs	1.46	1.85	1.42	1.08	1.11
CuCl	3.1	2.44	0.94	0.54	0.67
GaSb	0.726	0.78	0.07	0.00	0.12
CdTe	1.44	1.53	1.00	0.63	0.64
ZnSe	2.82	2.48	1.82	1.23	1.27
BN	6.2	5.88	5.05	4.35	4.45
ZnTe	2.26	2.23	1.64	1.22	1.30
Ge	0.67	0.57	0.25	0.06	0.11
CuI	3.1	2.72	1.62	1.17	1.29
BeS	5.5	4.10	3.62	2.96	2.87
AlP	2.42	2.34	1.96	1.49	1.45
InN	0.65	0.39	1.30	0.05	1.15
GaAs	1.42	1.27	0.70	0.39	0.50
Si	1.11	1.19	0.86	0.49	0.47
AlSb	1.6	1.84	1.39	1.13	1.12
CdS	2.42	2.18	1.50	1.00	0.96
InAs	0.36	0.33	1.13	0.00	0.00
BP	2.1	1.98	1.55	1.14	1.16
InP	1.344	1.31	0.89	0.48	0.52
ZnS	3.7	3.45	2.73	2.08	2.14

Table A.6: Tabulated reference[1,2,4,5,6,7,8,9,3] and calculated band gap data for Figure 3.10.

A.2 GULP Output Data

The following section contains snippets of the GULP input and output files of Ewald summation calculations that were referred to in the third chapter. The output files show that the interstitial potentials (*I1* and *I2*) average to zero.

A.2.1 Zinc Blende

Input File

```
dist single pot
cell
   5.6 5.6 5.6  90.000000  90.000000  90.000000
space
216
fractional
Ga   core  0.000000   0.000000   0.000000
As   core  0.250000   0.250000   0.250000
I1   core  0.500000   0.500000   0.500000
I2   core  0.750000   0.750000   0.750000
species
Ga core 1.0
As core -1.0
I1 core 0
I2 core 0
```

Output File

Space group (noncentrosymmetric) : F -4 3 M

Fractional coordinates of asymmetric unit :

No.	Atomic Label		x (Frac)	y (Frac)	z (Frac)	Charge (e)	Occupancy (Frac)
1	Ga	c	0.000000	0.000000	0.000000	1.00000	1.000000
2	As	c	0.250000	0.250000	0.250000	-1.00000	1.000000
3	I1	c	0.500000	0.500000	0.500000	0.00000	1.000000
4	I2	c	0.750000	0.750000	0.750000	0.00000	1.000000

Lattice summation method = Ewald (3-D)

Electrostatic potential at atomic positions :

Site No.	Atomic Label		Potential (V)	Derivatives (V/Angs) x	y	z
1	Ga	c	-9.727361	-0.000000	-0.000000	-0.000000
2	As	c	9.727361	0.000000	0.000000	0.000000
3	I1	c	-0.740037	-0.000000	-0.000000	-0.000000
4	I2	c	0.740037	0.000000	-0.000000	-0.000000

A.2.2 Rocksalt

Input File

```
dist single pot
cell
   5.64 5.64 5.64  90.000000  90.000000  90.000000
space
225
fractional
Na   core  0.000000   0.000000   0.000000
Cl   core  0.500000   0.000000   0.000000
I1   core  0.250000   0.250000   0.250000
species
Na core 1.0
Cl core -1.0
I1 core 0
```

Output File

```
  Space group (centrosymmetric)    :  F M 3 M

  Fractional coordinates of asymmetric unit :

-------------------------------------------------------------------------------
  No.  Atomic       x           y           z         Charge      Occupancy
       Label      (Frac)      (Frac)      (Frac)        (e)         (Frac)
-------------------------------------------------------------------------------
     1 Na    c    0.000000    0.000000    0.000000     1.00000    1.000000
     2 Cl    c    0.500000    0.000000    0.000000    -1.00000    1.000000
     3 I1    c    0.250000    0.250000    0.250000     0.00000    1.000000
-------------------------------------------------------------------------------

  Lattice summation method              =    Ewald          (3-D)

  Electrostatic potential at atomic positions :

-------------------------------------------------------------------------------
 Site  Atomic      Potential               Derivatives (V/Angs)
 No.   Label          (V)                 x           y           z
-------------------------------------------------------------------------------
     1 Na    c     -8.923584           0.000000    0.000000    0.000000
     2 Cl    c      8.923584          -0.000000   -0.000000   -0.000000
     3 I1    c      0.000000           0.000000    0.000000    0.000000
-------------------------------------------------------------------------------
```

A.2.3 Perovskite

Input File

```
dist single pot
cell
7.6438 7.6438 7.6438  90.000000  90.000000  90.000000
space
221
fractional
Ca   core  0.000000   0.000000   0.000000
O    core  0.500000   0.500000   0.000000
Ti   core  0.500000   0.500000   0.500000
I1   core  0.250000   0.250000   0.250000
species
Ca   core 2.0
O    core -2.0
Ti   core 4.0
I1   core 0
```

Output File

Space group (centrosymmetric) : P M 3 M

Fractional coordinates of asymmetric unit :

No.	Atomic Label		x (Frac)	y (Frac)	z (Frac)	Charge (e)	Occupancy (Frac)
1	Ca	c	0.000000	0.000000	0.000000	2.00000	1.000000
2	O	c	0.500000	0.500000	0.000000	-2.00000	1.000000
3	Ti	c	0.500000	0.500000	0.500000	4.00000	1.000000
4	I1	c	0.250000	0.250000	0.250000	0.00000	1.000000

Lattice summation method = Ewald (3-D)

Electrostatic potential at atomic positions :

Site No.	Atomic Label		Potential (V)	Derivatives (V/Angs) x	y	z
1	Ca	c	-10.148685	-0.000000	-0.000000	-0.000000
2	O	c	12.161952	-0.000000	-0.000000	0.000000
3	Ti	c	-23.317270	-0.000000	-0.000000	-0.000000
4	I1	c	-0.000000	0.000000	0.000000	0.000000

A.3 Self Trapping

Compound	# Primitive cells	MP Grid Density
Bi_2O_3	16	Nan
$BiVO_4$	16	3×4×3
CdO	12	Nan
Co_3O_4	12	4×5×3
Cr_2O_3	16	4×4×4
Cu_2O	27	Nan
CuO	6	Nan
Fe_2O_3	16	4×4×4
In_2O_3	6	3×4×4
MoO_3	12	Nan
NiO	1	Nan
SnO	27	Nan
$SrTiO_3$	27	4×4×4
WO_3	8	Nan
ZnO	27	Nan
TiO_2 (Rutile)	16	5×5×4
TiO_2 (Anatase)	16	4×5×4
VO_2 (Rutile)	16	5×5×4
VO_2 (Monoclinic)	16	Nan

Table A.7: Supercell parameters that were used for self-trapping calculations displayed in Section 4.2. Listed are the number of primitive unit cells included in the supercells and their converged Monkhorst-Pack grid densities. Note that for magnetic systems, such as NiO and CuO, the primitive cells correspond to the *magnetic* primitive cell and not the structural primitive cell which ignores any magnetic ordering of the lattice. Non-magnetic primitive cell structure files were sourced from the *Materials Project* database[10]. Magnetic primitive cell structure files were sourced from the *MAGNDATA* database[11,12,13,14]. Note also that the non-localising or non-converging polaron calculations do not have converged **k**-point data because they were discarded from further convergence testing.

A.4 Corrosion Resistance

Adsorbate	Energies / eV							
	TiO_2	**GeO_2**	**SnO_2**	**IrO_2**	**RuO_2**	**PbO_2**	**MnO_2**	**PtO_2**
None	-1672.28	-1204.27	-1174.03	-1321.24	-1400.59	-1009.87	-1464.80	-1096.08
O	-1675.25	-1206.81	-1176.56	-1327.26	-1406.70	-1011.76	-1471.47	-1100.87
OH	-1681.52	-1213.03	-1182.98	-1331.60	-1410.89	-1017.99	-1475.45	-1106.43
OOH	-1685.68	-1217.89	-1187.73	-1336.54	-1415.63	-1023.22	-1480.04	-1110.71

Table A.8: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV					
		PtRu	**IrMn**	**IrPt**	**PbSn**	**GeTi**	**PtSn**
None	None	-1248.65	-1392.35	-1208.51	-1089.99	-1435.32	-1130.62
O	1	-1254.52	-1398.22	-1214.10	-1092.86	-1437.82	-1135.27
O	2	-1253.11	-1393.23	-1214.22	-1092.30	-1438.62	-1134.99
O	3	-1252.94	-1394.54	-1214.07	-1092.33	-1438.59	-1133.79
O	4	-1253.01	-1393.91	-1212.75	-1092.37	-1438.81	-1134.58
O	5	-1253.21	-1395.40	-1212.83	-1091.95	-1438.49	-1135.65
O	6	-1254.62	-1396.36	-1212.80	-1091.87	-1437.83	-1135.48
O	7	-1254.74	-1393.72	-1212.85	-1092.06	-1437.76	-1136.34
O	8	-1254.66	-1398.36	-1214.13	-1092.03	-1437.74	-1135.64
O	AVG	-1253.85	-1395.47	-1213.47	-1092.22	-1438.21	-1135.22
O	MIN	-1254.74	-1398.36	-1214.22	-1092.86	-1438.81	-1136.34
O	STD	0.79	1.87	0.66	0.30	0.43	0.73
OH	1	-1259.15	-1403.28	-1219.15	-1099.11	-1444.92	-1141.32
OH	2	-1259.05	-1400.01	-1219.35	-1099.19	-1444.94	-1140.64
OH	3	-1258.91	-1400.00	-1219.28	-1099.22	-1444.77	-1140.60
OH	4	-1259.02	-1400.02	-1218.63	-1099.15	-1445.01	-1140.56
OH	5	-1259.10	-1402.40	-1218.81	-1098.82	-1444.77	-1140.78
OH	6	-1259.28	-1400.14	-1218.76	-1098.75	-1444.58	-1141.16
OH	7	-1259.39	-1403.28	-1218.82	-1098.90	-1444.62	-1141.42
OH	8	-1259.30	-1403.40	-1219.27	-1098.93	-1444.58	-1141.24
OH	AVG	-1259.15	-1401.57	-1219.01	-1099.01	-1444.77	-1140.96
OH	MIN	-1259.39	-1403.40	-1219.35	-1099.22	-1445.01	-1141.42
OH	STD	0.15	1.55	0.26	0.17	0.16	0.33
OOH	1	-1263.55	-1407.77	-1223.62	-1103.68	-1448.90	-1146.52
OOH	2	-1263.46	-1405.26	-1223.82	-1103.75	-1448.99	-1144.83
OOH	3	-1262.98	-1405.28	-1223.48	-1103.72	-1448.93	-1144.72
OOH	4	-1263.45	-1405.26	-1223.25	-1103.68	-1448.97	-1144.80
OOH	5	-1263.61	-1405.29	-1223.31	-1103.58	-1449.02	-1144.81
OOH	6	-1263.78	-1403.46	-1223.39	-1103.45	-1449.03	-1145.95
OOH	7	-1264.01	-1405.38	-1223.28	-1104.91	-1448.95	-1146.09
OOH	8	-1263.77	-1407.59	-1223.61	-1103.56	-1448.92	-1146.17
OOH	AVG	-1263.58	-1405.66	-1223.47	-1103.79	-1448.96	-1145.49
OOH	MIN	-1264.01	-1407.77	-1223.82	-1104.91	-1449.03	-1146.52
OOH	STD	0.29	1.31	0.19	0.43	0.04	0.71

Table A.9: Continued: Tabulated data for Figure 5.6. *AVG*, *MIN* and *STD* give the average, minimum and standard deviation respectively of all 8 adsorption site energies.

Adsorbate	Binding Site	Energies / eV					
		SnTi	**IrTi**	**GeIr**	**GeSn**	**GeRu**	**GeMn**
None	None	-1417.85	-1490.09	-1257.41	-1183.44	-1295.46	-1327.80
O	1	-1420.50	-1494.90	-1261.21	-1185.64	-1299.12	-1332.59
O	2	-1420.18	-1496.09	-1263.45	-1185.82	-1301.48	-1331.58
O	3	-1420.10	-1496.01	-1263.51	-1185.76	-1301.44	-1333.38
O	4	-1420.37	-1496.04	-1263.35	-1186.05	-1301.26	-1333.63
O	5	-1420.39	-1495.98	-1263.32	-1186.27	-1301.54	-1331.33
O	6	-1420.88	-1493.70	-1261.24	-1185.82	-1299.10	-1332.37
O	7	-1420.95	-1495.36	-1261.72	-1186.43	-1299.41	-1331.76
O	8	-1420.67	-1495.32	-1261.78	-1186.09	-1299.49	-1331.05
O	AVG	-1420.51	-1495.42	-1262.45	-1185.98	-1300.35	-1332.21
O	MIN	-1420.95	-1496.09	-1263.51	-1186.43	-1301.54	-1333.63
O	STD	0.29	0.77	0.98	0.25	1.08	0.89
OH	1	-1426.80	-1500.18	-1267.61	-1192.48	-1305.55	-1337.90
OH	2	-1427.16	-1501.19	-1267.32	-1192.84	-1306.08	-1337.86
OH	3	-1427.01	-1501.03	-1267.49	-1192.68	-1306.24	-1337.77
OH	4	-1427.26	-1501.13	-1268.53	-1192.93	-1306.04	-1337.98
OH	5	-1427.29	-1501.14	-1268.49	-1193.13	-1306.07	-1337.61
OH	6	-1427.05	-1500.15	-1265.91	-1192.56	-1305.64	-1337.62
OH	7	-1427.15	-1500.38	-1266.25	-1192.80	-1305.77	-1338.04
OH	8	-1426.97	-1500.45	-1267.92	-1192.82	-1305.68	-1337.62
OH	AVG	-1427.09	-1500.71	-1267.44	-1192.78	-1305.88	-1337.80
OH	MIN	-1427.29	-1501.19	-1268.53	-1193.13	-1306.24	-1338.04
OH	STD	0.15	0.43	0.89	0.19	0.24	0.16
OOH	1	-1431.27	-1504.16	-1271.47	-1196.91	-1309.95	-1342.53
OOH	2	-1431.57	-1505.66	-1272.91	-1197.22	-1310.60	-1341.83
OOH	3	-1431.56	-1505.50	-1272.96	-1197.15	-1310.80	-1342.17
OOH	4	-1431.72	-1505.54	-1272.94	-1197.35	-1310.56	-1342.59
OOH	5	-1431.58	-1505.71	-1273.04	-1197.46	-1310.75	-1341.81
OOH	6	-1431.29	-1504.25	-1271.71	-1197.06	-1309.67	-1342.21
OOH	7	-1431.44	-1504.36	-1272.08	-1197.22	-1309.89	-1341.60
OOH	8	-1431.40	-1505.18	-1271.76	-1197.23	-1309.85	-1341.50
OOH	AVG	-1431.48	-1505.05	-1272.36	-1197.20	-1310.26	-1342.03
OOH	MIN	-1431.72	-1505.71	-1273.04	-1197.46	-1310.80	-1342.59
OOH	STD	0.14	0.63	0.62	0.16	0.43	0.38

Table A.10: Continued: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV					
		RuTi	**IrSn**	**MnTi**	**PbTi**	**RuSn**	**MnSn**
None	None	-1528.56	-1240.59	-1558.87	-1330.91	-1276.73	-1300.66
O	1	-1532.43	-1246.51	-1563.20	-1333.48	-1281.26	-1306.22
O	2	-1534.80	-1246.55	-1563.46	-1333.64	-1282.66	-1303.98
O	3	-1534.56	-1246.49	-1563.51	-1335.54	-1282.74	-1303.90
O	4	-1533.34	-1244.65	-1563.52	-1336.07	-1282.51	-1304.32
O	5	-1532.37	-1244.48	-1563.47	-1332.87	-1282.62	-1304.73
O	6	-1533.39	-1245.47	-1564.32	-1332.28	-1279.76	-1305.78
O	7	-1533.51	-1244.72	-1564.32	-1333.06	-1280.59	-1304.80
O	8	-1534.49	-1246.49	-1562.66	-1333.27	-1280.62	-1305.84
O	AVG	-1533.61	-1245.67	-1563.56	-1333.78	-1281.60	-1304.95
O	MIN	-1534.80	-1246.55	-1564.32	-1336.07	-1282.74	-1306.22
O	STD	0.88	0.88	0.52	1.24	1.11	0.84
OH	1	-1539.23	-1251.58	-1568.67	-1339.82	-1287.01	-1310.29
OH	2	-1539.37	-1251.60	-1568.82	-1339.90	-1287.30	-1310.55
OH	3	-1539.37	-1251.68	-1568.85	-1339.95	-1287.48	-1310.41
OH	4	-1538.50	-1250.95	-1568.80	-1340.31	-1287.29	-1310.82
OH	5	-1538.64	-1250.96	-1568.83	-1339.93	-1287.24	-1310.92
OH	6	-1538.47	-1251.06	-1568.49	-1339.79	-1287.13	-1310.15
OH	7	-1538.75	-1251.08	-1568.85	-1339.89	-1287.05	-1310.57
OH	8	-1539.06	-1251.61	-1568.57	-1339.45	-1287.02	-1310.49
OH	AVG	-1538.92	-1251.31	-1568.74	-1339.88	-1287.19	-1310.53
OH	MIN	-1539.37	-1251.68	-1568.85	-1340.31	-1287.48	-1310.92
OH	STD	0.35	0.31	0.13	0.22	0.16	0.24
OOH	1	-1543.77	-1255.78	-1573.39	-1344.38	-1290.91	-1314.72
OOH	2	-1543.66	-1255.87	-1572.80	-1344.39	-1291.85	-1314.59
OOH	3	-1543.72	-1255.97	-1572.90	-1344.32	-1291.65	-1314.49
OOH	4	-1543.17	-1254.83	-1573.18	-1344.69	-1291.47	-1314.69
OOH	5	-1542.59	-1255.02	-1572.88	-1344.32	-1291.87	-1314.86
OOH	6	-1542.80	-1254.89	-1572.73	-1344.16	-1291.27	-1314.70
OOH	7	-1542.90	-1254.93	-1573.42	-1344.33	-1290.97	-1315.14
OOH	8	-1543.70	-1256.09	-1573.17	-1343.94	-1290.93	-1314.99
OOH	AVG	-1543.29	-1255.42	-1573.06	-1344.32	-1291.36	-1314.77
OOH	MIN	-1543.77	-1256.09	-1573.42	-1344.69	-1291.87	-1315.14
OOH	STD	0.45	0.51	0.25	0.20	0.38	0.20

Table A.11: Continued: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV					
		IrRu	**MnRu**	**GePb**	**GePt**	**PtTi**	**PbPt**
None	None	-1362.02	-1433.43	-1094.89	-1146.81	-1380.59	-1054.17
O	1	-1368.24	-1437.81	-1098.73	-1150.46	-1384.96	-1058.71
O	2	-1368.03	-1439.22	-1097.14	-1151.51	-1385.96	-1057.02
O	3	-1367.98	-1439.23	-1098.80	-1151.42	-1384.92	-1059.76
O	4	-1368.06	-1439.59	-1099.52	-1151.21	-1385.09	-1059.75
O	5	-1367.93	-1439.45	-1100.18	-1152.13	-1385.27	-1056.62
O	6	-1368.04	-1438.96	-1099.18	-1150.12	-1384.82	-1058.35
O	7	-1368.00	-1439.04	-1099.66	-1150.64	-1385.30	-1057.16
O	8	-1368.09	-1437.94	-1099.51	-1151.44	-1385.20	-1058.60
O	AVG	-1368.05	-1438.90	-1099.09	-1151.12	-1385.19	-1058.25
O	MIN	-1368.24	-1439.59	-1100.18	-1152.13	-1385.96	-1059.76
O	STD	0.09	0.62	0.86	0.62	0.33	1.13
OH	1	-1372.90	-1443.36	-1103.53	-1156.44	-1390.44	-1064.24
OH	2	-1373.15	-1444.11	-1104.03	-1157.61	-1391.11	-1063.28
OH	3	-1373.11	-1443.91	-1103.74	-1157.46	-1390.48	-1064.20
OH	4	-1373.16	-1444.22	-1103.81	-1157.32	-1390.60	-1064.42
OH	5	-1373.09	-1444.08	-1104.26	-1157.46	-1391.29	-1063.32
OH	6	-1372.77	-1443.28	-1103.71	-1156.65	-1390.41	-1064.39
OH	7	-1372.76	-1443.57	-1103.85	-1157.00	-1391.30	-1063.44
OH	8	-1372.82	-1443.48	-1103.91	-1156.87	-1391.20	-1063.39
OH	AVG	-1372.97	-1443.75	-1103.86	-1157.10	-1390.85	-1063.83
OH	MIN	-1373.16	-1444.22	-1104.26	-1157.61	-1391.30	-1064.42
OH	STD	0.16	0.35	0.21	0.40	0.38	0.49
OOH	1	-1377.42	-1447.84	-1108.13	-1160.51	-1396.27	-1068.43
OOH	2	-1377.62	-1448.66	-1108.26	-1161.80	-1396.20	-1067.53
OOH	3	-1377.23	-1448.44	-1108.04	-1161.58	-1395.19	-1068.55
OOH	4	-1377.63	-1448.40	-1108.30	-1162.17	-1394.63	-1068.73
OOH	5	-1377.54	-1448.58	-1108.36	-1161.84	-1395.34	-1067.49
OOH	6	-1377.33	-1447.51	-1108.35	-1160.73	-1376.09	-1068.54
OOH	7	-1377.32	-1448.11	-1108.40	-1161.08	-1395.42	-1068.46
OOH	8	-1377.35	-1448.02	-1108.36	-1161.01	-1395.24	-1068.55
OOH	AVG	-1377.43	-1448.20	-1108.27	-1161.34	-1393.05	-1068.28
OOH	MIN	-1377.63	-1448.66	-1108.40	-1162.17	-1396.27	-1068.73
OOH	STD	0.14	0.37	0.12	0.55	6.43	0.45

Table A.12: Continued: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV IrPb	MnPt	PbRu
None	None	-1167.66	-1277.78	-1205.94
O	1	-1169.33	-1281.98	-1210.22
O	2	-1173.31	-1282.05	-1209.94
O	3	-1173.54	-1281.86	-1210.37
O	4	-1173.36	-1282.14	-1209.98
O	5	-1171.46	-1282.37	-1211.31
O	6	-1171.95	-1281.70	-1210.51
O	7	-1171.98	-1283.25	-1210.59
O	8	-1171.87	-1282.21	-1210.60
O	AVG	-1172.10	-1282.19	-1210.44
O	MIN	-1173.54	-1283.25	-1211.31
O	STD	1.29	0.44	0.40
OH	1	-1177.19	-1287.56	-1215.18
OH	2	-1178.50	-1287.83	-1215.81
OH	3	-1178.52	-1287.84	-1216.40
OH	4	-1178.48	-1288.01	-1216.18
OH	5	-1178.35	-1288.23	-1216.05
OH	6	-1177.24	-1287.44	-1215.44
OH	7	-1177.48	-1287.89	-1213.33
OH	8	-1177.25	-1287.65	-1215.41
OH	AVG	-1177.87	-1287.81	-1215.47
OH	MIN	-1178.52	-1288.23	-1216.40
OH	STD	0.59	0.24	0.90
OOH	1	-1181.43	-1292.04	-1219.61
OOH	2	-1181.76	-1292.31	-1221.96
OOH	3	-1181.71	-1292.60	-1220.85
OOH	4	-1183.24	-1292.80	-1220.61
OOH	5	-1181.17	-1292.50	-1220.38
OOH	6	-1182.92	-1291.82	-1219.89
OOH	7	-1182.85	-1292.45	-1220.17
OOH	8	-1182.56	-1292.24	-1219.74
OOH	AVG	-1182.20	-1292.34	-1220.40
OOH	MIN	-1183.24	-1292.80	-1221.96
OOH	STD	0.73	0.29	0.71

Table A.13: Continued: Tabulated data for Figure 5.6.

	Energies / eV						
Adsorbate	**TiO_2**	**GeO_2**	**SnO_2**	**IrO_2**	**RuO_2**	**PbO_2**	**MnO_2**
O	4.55	4.98	4.99	1.51	1.41	5.63	0.85
OH	1.63	2.11	1.92	0.51	0.57	2.75	0.23
OOH	4.99	4.77	4.69	3.10	3.35	5.04	3.15

Table A.14: Continued: Tabulated data for Figure 5.6.

None	**Binding**	**Energies / eV**					
Adsorbate	**Site**	**PtRu**	**IrMn**	**IrPt**	**PbSn**	**GeTi**	**PtSn**
O	1	1.65	1.65	1.94	4.66	5.03	2.88
O	2	3.07	6.64	1.82	5.22	4.22	3.15
O	3	3.23	5.33	1.96	5.19	4.25	4.36
O	4	3.16	5.97	3.29	5.15	4.03	3.57
O	5	2.96	4.47	3.20	5.57	4.35	2.49
O	6	1.56	3.51	3.23	5.64	5.02	2.66
O	7	1.43	6.15	3.18	5.46	5.08	1.80
O	8	1.51	1.52	1.90	5.49	5.10	2.50
OH	1	0.38	-0.05	0.23	1.76	1.28	0.17
OH	2	0.47	3.21	0.03	1.68	1.25	0.85
OH	3	0.61	3.22	0.10	1.64	1.42	0.89
OH	4	0.51	3.20	0.75	1.72	1.18	0.93
OH	5	0.42	0.82	0.58	2.04	1.42	0.72
OH	6	0.24	3.08	0.63	2.12	1.61	0.34
OH	7	0.13	-0.05	0.56	1.97	1.57	0.08
OH	8	0.23	-0.18	0.11	1.94	1.61	0.25
OOH	1	3.49	2.97	3.28	4.71	4.82	2.50
OOH	2	3.59	5.49	3.09	4.64	4.73	4.18
OOH	3	4.07	5.47	3.43	4.67	4.79	4.30
OOH	4	3.60	5.49	3.66	4.71	4.75	4.22
OOH	5	3.44	5.46	3.60	4.81	4.69	4.20
OOH	6	3.26	7.28	3.52	4.94	4.69	3.06
OOH	7	3.03	5.36	3.63	3.48	4.76	2.92
OOH	8	3.28	3.16	3.30	4.83	4.80	2.84

Table A.15: Continued: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV					
		SnTi	**IrTi**	**GeIr**	**GeSn**	**GeRu**	**GeMn**
O	1	4.87	2.72	3.73	5.32	3.87	2.73
O	2	5.19	1.53	1.48	5.15	1.50	3.75
O	3	5.27	1.61	1.42	5.21	1.54	1.94
O	4	5.00	1.57	1.58	4.92	1.72	1.69
O	5	4.99	1.63	1.62	4.70	1.45	4.00
O	6	4.49	3.91	3.69	5.15	3.88	2.96
O	7	4.42	2.25	3.21	4.54	3.58	3.57
O	8	4.70	2.29	3.15	4.88	3.49	4.27
OH	1	1.92	0.79	0.67	1.84	0.78	0.77
OH	2	1.56	-0.22	0.96	1.47	0.25	0.82
OH	3	1.71	-0.07	0.79	1.63	0.10	0.91
OH	4	1.47	-0.17	-0.24	1.39	0.29	0.69
OH	5	1.43	-0.18	-0.21	1.19	0.26	1.06
OH	6	1.67	0.82	2.38	1.76	0.69	1.05
OH	7	1.57	0.58	2.03	1.52	0.56	0.63
OH	8	1.75	0.52	0.37	1.49	0.65	1.06
OOH	1	4.97	4.33	4.34	4.93	3.90	3.67
OOH	2	4.68	2.83	2.89	4.62	3.26	4.37
OOH	3	4.68	2.99	2.85	4.69	3.06	4.03
OOH	4	4.53	2.95	2.87	4.50	3.30	3.61
OOH	5	4.67	2.78	2.76	4.38	3.11	4.39
OOH	6	4.95	4.24	4.10	4.78	4.18	3.99
OOH	7	4.80	4.13	3.73	4.63	3.97	4.60
OOH	8	4.84	3.30	4.05	4.62	4.01	4.70

Table A.16: Continued: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV					
		RuTi	**IrSn**	**MnTi**	**PbTi**	**RuSn**	**MnSn**
O	1	3.65	1.60	3.19	4.95	2.99	1.96
O	2	1.29	1.56	2.94	4.80	1.59	4.21
O	3	1.52	1.62	2.89	2.90	1.51	4.29
O	4	2.74	3.46	2.88	2.37	1.74	3.86
O	5	3.72	3.63	2.93	5.57	1.63	3.46
O	6	2.69	2.64	2.07	6.16	4.49	2.40
O	7	2.57	3.39	2.07	5.37	3.66	3.38
O	8	1.59	1.62	3.73	5.17	3.63	2.35
OH	1	0.20	-0.13	1.07	1.96	0.59	1.24
OH	2	0.06	-0.14	0.92	1.88	0.30	0.99
OH	3	0.06	-0.22	0.90	1.84	0.12	1.12
OH	4	0.93	0.51	0.94	1.48	0.31	0.71
OH	5	0.80	0.50	0.91	1.85	0.36	0.61
OH	6	0.96	0.40	1.25	2.00	0.47	1.39
OH	7	0.69	0.38	0.89	1.89	0.55	0.97
OH	8	0.37	-0.15	1.18	2.33	0.58	1.04
OOH	1	3.18	3.20	3.87	4.93	4.21	4.34
OOH	2	3.30	3.11	4.47	4.92	3.28	4.46
OOH	3	3.23	3.01	4.36	4.99	3.48	4.57
OOH	4	3.79	4.15	4.08	4.62	3.65	4.37
OOH	5	4.37	3.96	4.38	4.99	3.26	4.20
OOH	6	4.15	4.09	4.54	5.14	3.86	4.36
OOH	7	4.06	4.05	3.85	4.97	4.16	3.92
OOH	8	3.26	2.89	4.10	5.37	4.19	4.07

Table A.17: Continued: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV					
		IrRu	**MnRu**	**GePb**	**GePt**	**PtTi**	**PbPt**
O	1	1.30	3.14	3.69	3.87	3.15	2.98
O	2	1.51	1.72	5.27	2.82	2.15	4.68
O	3	1.55	1.72	3.61	2.91	3.19	1.94
O	4	1.48	1.36	2.89	3.12	3.02	1.94
O	5	1.61	1.50	2.24	2.20	2.84	5.08
O	6	1.50	1.99	3.23	4.21	3.29	3.35
O	7	1.54	1.91	2.75	3.69	2.82	4.53
O	8	1.45	3.01	2.90	2.89	2.91	3.09
OH	1	-0.01	0.94	2.23	1.25	1.02	0.81
OH	2	-0.26	0.19	1.73	0.07	0.35	1.77
OH	3	-0.22	0.38	2.03	0.23	0.98	0.84
OH	4	-0.28	0.08	1.95	0.36	0.86	0.62
OH	5	-0.21	0.22	1.50	0.22	0.17	1.72
OH	6	0.11	1.02	2.05	1.03	1.05	0.65
OH	7	0.13	0.73	1.91	0.69	0.16	1.61
OH	8	0.07	0.82	1.85	0.81	0.26	1.66
OOH	1	2.99	3.98	5.16	4.69	2.71	4.14
OOH	2	2.80	3.16	5.03	3.40	2.78	5.04
OOH	3	3.18	3.38	5.25	3.63	3.79	4.02
OOH	4	2.78	3.43	4.99	3.04	4.35	3.84
OOH	5	2.87	3.24	4.93	3.36	3.64	5.08
OOH	6	3.08	4.31	4.94	4.48	22.89	4.03
OOH	7	3.09	3.71	4.89	4.13	3.57	4.10
OOH	8	3.06	3.80	4.92	4.19	3.74	4.01

Table A.18: Continued: Tabulated data for Figure 5.6.

Adsorbate	Binding Site	Energies / eV		
		IrPb	MnPt	PbRu
O	1	5.85	3.32	3.24
O	2	1.88	3.26	3.51
O	3	1.65	3.45	3.09
O	4	1.83	3.16	3.48
O	5	3.73	2.94	2.15
O	6	3.23	3.61	2.94
O	7	3.20	2.06	2.87
O	8	3.31	3.10	2.86
OH	1	1.35	1.09	1.63
OH	2	0.04	0.83	1.00
OH	3	0.01	0.82	0.41
OH	4	0.06	0.65	0.63
OH	5	0.19	0.43	0.75
OH	6	1.30	1.22	1.37
OH	7	1.05	0.77	3.48
OH	8	1.29	1.01	1.40
OOH	1	4.62	4.14	4.72
OOH	2	4.30	3.87	2.37
OOH	3	4.35	3.58	3.48
OOH	4	2.82	3.38	3.73
OOH	5	4.89	3.68	3.95
OOH	6	3.14	4.36	4.44
OOH	7	3.21	3.73	4.16
OOH	8	3.50	3.94	4.59

Table A.19: Continued: Tabulated data for Figure 5.6.

	$\Delta G^O - \Delta G^{HO}$ / eV	RHS value	LHS value	- η / V	Error bars
TiO_2	3.33	-0.13	3.33	-2.10	None
GeO_2	3.28	-0.08	3.28	-2.05	None
SnO_2	3.47	-0.27	3.47	-2.24	None
IrO_2	1.41	1.79	1.41	-0.56	None
RuO_2	1.25	1.95	1.25	-0.72	None
PbO_2	3.28	-0.08	3.28	-2.05	None
MnO_2	1.04	2.16	1.04	-0.93	None
PtO_2	2.62	0.58	2.62	-1.39	None
PtRu	2.36	0.84	2.36	-1.13	0.40
IrMn	3.16	0.04	3.16	-1.93	1.22
IrPt	2.60	0.60	2.60	-1.37	0.36
PbSn	3.85	-0.65	3.85	-2.62	0.17
GeTi	3.63	-0.43	3.63	-2.40	0.23
PtSn	2.81	0.39	2.81	-1.58	0.40
SnTi	3.64	-0.44	3.64	-2.41	0.16
IrTi	2.34	0.86	2.34	-1.11	0.44
GeIr	2.05	1.15	2.05	-0.82	0.66
GeSn	3.86	-0.66	3.86	-2.63	0.16
GeRu	2.59	0.61	2.59	-1.36	0.55
GeMn	2.65	0.55	2.65	-1.42	0.45
RuTi	2.37	0.83	2.37	-1.14	0.47
IrSn	2.71	0.49	2.71	-1.48	0.47
MnTi	2.24	0.96	2.24	-1.01	0.27
PbTi	3.17	0.03	3.17	-1.94	0.63
RuSn	2.66	0.54	2.66	-1.43	0.56
MnSn	2.64	0.56	2.64	-1.41	0.44
IrRu	1.98	1.22	1.98	-0.75	0.09
MnRu	1.91	1.29	1.91	-0.68	0.36
GePb	1.83	1.37	1.83	-0.60	0.44
GePt	3.04	0.16	3.04	-1.81	0.37
PtTi	2.72	0.48	2.72	-1.49	0.25
PbPt	2.65	0.55	2.65	-1.42	0.62
IrPb	2.83	0.37	2.83	-1.60	0.71
MnPt	2.67	0.53	2.67	-1.44	0.25
PbRu	2.10	1.10	2.10	-0.87	0.49

Table A.20: Continued: Tabulated data for Figure 5.8. *RHS* and *LHS* denote the right-hand-side and left-hand-side values respectively for the volcano plot.

Adsorbate	Binding Site (Species)	Energies / eV 25%	50%	75%
None	None	-1597.20	-1528.39	-1463.96
O	1 (Ti)	-1601.91	-1533.27	-1468.99
O	2 (Ti)	-1601.76	-1534.51	-1469.03
O	3 (Ru)	-1603.35	-1532.15	-1470.02
O	4 (Ti)	-1601.78	-1532.04	-1469.14
O	5 (Ru)	-1603.29	-1534.43	-1469.90
O	6 (Ru)	-1603.44	-1534.62	-1469.84
O	7 (Ti)	-1601.90	-1534.54	-1469.82
O	8 (Ti)	-1601.89	-1533.14	-1469.93
O	AVG-Ti	-1601.84	-1532.65	-1469.07
O	AVG-Ru	-1603.36	-1534.52	-1469.87
O	STD-Ti	0.06	0.62	0.07
O	STD-Ru	0.07	0.09	0.03
OH	1 (Ti)	-1607.26	-1538.52	-1474.08
OH	2 (Ru)	-1607.00	-1539.03	-1474.06
OH	3 (Ti)	-1607.77	-1538.52	-1474.53
OH	4 (Ti)	-1607.19	-1538.44	-1474.26
OH	5 (Ru)	-1607.98	-1538.95	-1474.59
OH	6 (Ru)	-1608.00	-1539.32	-1474.53
OH	7 (Ru)	-1607.16	-1539.12	-1474.60
OH	8 (Ti)	-1607.20	-1538.33	-1474.62
OH	AVG-Ti	-1607.23	-1538.48	-1474.17
OH	AVG-Ru	-1607.99	-1539.13	-1474.56
OH	STD-Ti	0.03	0.04	0.09
OH	STD-Ru	0.01	0.19	0.03
OOH	1 (Ti)	-1611.33	-1542.65	-1478.09
OOH	2 (Ti)	-1611.14	-1543.67	-1478.64
OOH	3 (Ru)	-1612.43	-1542.57	-1478.90
OOH	4 (Ti)	-1611.27	-1542.59	-1478.30
OOH	5 (Ru)	-1612.21	-1543.54	-1479.12
OOH	6 (Ru)	-1613.63	-1543.87	-1479.07
OOH	7 (Ru)	-1611.36	-1543.57	-1479.10
OOH	8 (Ru)	-1611.40	-1539.06	-1479.11

Table A.21: Raw DFT Energies for Figure 5.10.

Adsorbate	Binding Site (Species)	Energies / eV 25%	50%	75%
O	1 (Ti)	2.82	2.65	2.49
O	2 (Ti)	2.97	1.41	2.46
O	3 (Ru)	1.38	3.77	1.47
O	4 (Ti)	2.95	3.88	2.34
O	5 (Ru)	1.44	1.49	1.58
O	6 (Ru)	1.29	1.30	1.64
O	7 (Ti)	2.83	1.38	1.67
O	8 (Ti)	2.83	2.78	1.55
OH	1 (Ti)	0.82	0.74	0.76
OH	2 (Ru)	1.08	0.24	0.78
OH	3 (Ti)	0.30	0.75	0.30
OH	4 (Ti)	0.88	0.83	0.58
OH	5 (Ru)	0.10	0.32	0.24
OH	6 (Ru)	0.07	-0.05	0.30
OH	7 (Ru)	0.91	0.14	0.24
OH	8 (Ti)	0.88	0.94	0.22
OOH	1 (Ti)	4.27	4.14	4.27
OOH	2 (Ti)	4.46	3.13	3.72
OOH	3 (Ru)	3.17	4.22	3.46
OOH	4 (Ti)	4.33	4.20	4.06
OOH	5 (Ru)	3.40	3.25	3.24
OOH	6 (Ru)	1.97	2.92	3.28
OOH	7 (Ru)	4.24	3.23	3.26
OOH	8 (Ru)	4.20	7.74	3.25

Table A.22: Adsorption Energies for Figure 5.10.

	$\Delta G^O - \Delta G^{HO}$ / eV	RHS value	LHS value	- η / V	Error bars
25%Ru (Ti)	2.44	0.76	2.44	-1.21	0.03
50%Ru (Ti)	2.89	0.31	2.89	-1.66	0.31
75%Ru (Ti)	2.16	1.04	2.16	-0.93	0.06
25%Ru (Ru)	1.69	1.51	1.69	-0.46	0.04
50%Ru (Ru)	1.67	1.53	1.67	-0.44	0.11
75%Ru (Ru)	1.75	1.45	1.75	-0.52	0.02

Table A.23: Tabulated data for Figure 5.10. *RHS* and *LHS* denote the right-hand-side and left-hand-side values respectively for the volcano plot.

Clarifying Background Matter

The following appendix contains clarifying background matter that may aid the reader's understanding of some of the concepts mentioned in the preceding chapters.

B.1 The Fermi-Dirac Distribution

Adapted from the work of Biswas[488]

Consider an ideal gas of indistinguishable particles that obey the Pauli exclusion principle. The system is also connected with a thermal reservoir and particle reservoir such that the temperature (T) and chemical potential (μ) are fixed[a]. The grand potential for this system (Φ_G) is therefore minimised when the system is in equilibrium with μ and T. Since the particles are indistinguishable, it may be assumed that their single particles all have Hamiltonian operators of the same form and can be represented together as a linear sum of said operators. An example of such a grand potential can therefore be:

$$\Phi_G = -k_B T ln(\sum_{[n_j]} e^{-\frac{\sum_j n_j(\epsilon_j - \mu)}{k_B T}}) \tag{B.1}$$

where k_B is the Boltzmann constant, $[n_j]$ represents the microstate (quantum state[16]) for a given particle (n_j), ϵ_j is the single particle eigenvalue (energy) of the particle. Next, consider that there is no restriction on the number of particles for the grand canonical ensemble and that Pauli exclusion restricts the set of allowable n_j states to 0 or 1. Thus, the summation in the above equation can be reparameterised in terms of j, and not $[n_j]$, whilst substituting in $n_j = 1$ and 0 gives:

$$\Phi_G = -k_B T \sum_j ln(1 + e^{-\frac{(\epsilon_j - \mu)}{k_B T}}) \tag{B.2}$$

Next, consider the Boltzmann (canonical) distribution which considers the probability (p_j) of the occupation of energy level j within the grand canonical ensemble[16].

$$p_j = \frac{e^{-\frac{(\epsilon_j - \mu)}{k_B T}}}{\sum_i e^{-\frac{(\epsilon_i - \mu)}{k_B T}}} \tag{B.3}$$

Applying this to the equation to Φ_G allows the average number of particles ($\bar{n_j}$) at energy level j within the system:

$$\bar{n_j} = \sum_{[n_i]} n_j p_i = \frac{1}{1 + e^{-\frac{(\epsilon_j - \mu)}{k_B T}}} \tag{B.4}$$

[a] This is known as the grand canonical ensemble.

B.2 The Plausibility of the Schrödinger Equation

Adapted from pages 1263 - 1264 of Physics for Scientists and Engineers[96].

The Schrödinger equation cannot be derived mathematically by conventional means. Rather, its validity has been proven with time in a heuristic fashion. Nevertheless, the validity of the Schrödinger equation can be articulated via *plausibility arguments*. Consider the de Broglie relationship in terms of the electron kinetic energy:

$$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{h}{\sqrt{2m(\frac{1}{2}mv^2)}} = \frac{h}{\sqrt{2mE_K}} \tag{B.5}$$

Note that this implies that the de Broglie wavelength is inversely proportional to the kinetic energy of the electron. Now consider the second derivative of any wave-like eigenfunction ($\psi(x) = \psi_0 sin(\frac{2\pi x}{\lambda})$) of the differential operator ($\frac{d^2}{dx^2}$):

$$\frac{d^2\psi}{dx^2} = -\frac{4\pi^2}{\lambda^2}\psi_0 sin(\frac{2\pi x}{\lambda}) = -\frac{4\pi^2}{\lambda^2}\psi(x) \tag{B.6}$$

Assuming $\lambda = \lambda$, substituting the above de Broglie relationship into the above relationship and simplifying the resulting expression gives:

$$\frac{d^2}{dx^2}\psi(x) = -\frac{2m}{\hbar^2}E_K\psi(x) \tag{B.7}$$

Assuming that the total energy (E) of the system comprises a combination of kinetic (E_K) and potential (V) energy which are in constant exchange according to $E = E_K + V$. Incorporating this expression into the above equation gives:

$$\frac{d^2}{dx^2}\psi(x) = -\frac{2m}{\hbar^2}[E - V]\psi(x) \tag{B.8}$$

Rearranging for $E\psi(x)$ gives:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + V\psi(x) = E\psi(x) \tag{B.9}$$

Extending to 3 dimensions via the Laplacian operator (∇^2), gives the Time Independent Schrödinger Equation:

$$[-\frac{\hbar^2}{2m}\nabla^2 + V]\psi(x) = E\psi(x) \tag{B.10}$$

In general, the kinetic component of the Schrödinger equation is unchanged across different Hamiltonian operators. It is normally the *potential* energy component of the Hamiltonian operator that is the point of contention. In this regard, the wavefunction must be parameterised with tangible and physically meaningful quantities. To do this, boundary conditions are imposed on the wavefunction to solve the Schrödinger equation. These not only provide a mathematical form for the wavefunction but also define the quantisation of the electronic states by restricting the number of acceptable solutions for E to discrete values. The general boundary conditions imposed on a wavefunction are:

1. the wavefunction must be a continuous function,
2. the wavefunction = 0 outside of acceptable bounds (i.e. outside the crystal),
3. the wavefunction tends to 0 as the size of the acceptable volume for the electron tends to infinity which ensures that the above normalisation condition converges,
4. and the wavefunction is normalised.

B.3 Born-Von Karman Periodic Boundary Conditions

Adapted from pages 81-82 of Electronic Structure[99].

Consider a function ($f(\mathbf{r})$) which represents the density of the electrons in a perfect crystal. It follows that such a function would coincide with the periodicity of the lattice according to:

$$f(\mathbf{r}) = f(\mathbf{r} + \mathbf{T}) = f(\mathbf{r} + \sum_j N_j \mathbf{a}_j) \tag{B.11}$$

where $\mathbf{T}$ is any translation in terms of primitive lattice vectors ($\mathbf{a}_j$), $\mathbf{r}$ is the position in space, and N_j is an integer that represents the number of unit cells in the crystal. Thus, $f(\mathbf{r})$ is a finite and periodic function that can, like any such function, be represented by a Fourier transform defined in terms of a set of discrete Fourier components for N_j. An example of such a function is:

$$e^{i\mathbf{q}\cdot N_1\mathbf{a}_1} = e^{i\mathbf{q}\cdot N_2\mathbf{a}_2} = e^{i\mathbf{q}\cdot N_3\mathbf{a}_3} = 1 \tag{B.12}$$

where $\mathbf{q}$ is a wavevector for the system and i is the imaginary unit. A corresponding Fourier transform would therefore be:

$$f(\mathbf{q}) = \frac{1}{\Omega}\int_\Omega d\mathbf{r} f(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \tag{B.13}$$

Where Ω represents the volume of the crystal. The *boundary condition* is now invoked which is that $\mathbf{q}\cdot\mathbf{a}_j$ may only be chosen such that:

$$\mathbf{q}\cdot\mathbf{a}_j = 2\pi\frac{n}{N_j} \tag{B.14}$$

where n is an integer. This restricts the selection of Fourier components such that they only represent the lattice vectors.

Given the $2\pi x$ form of the above relation, Born-Von Karman boundary conditions are often referred to as *cyclic* boundary conditions. This is because, in each dimension (j), the unit cells form a closed loop, where the *first* unit cell is simultaneously the *final* unit cell. As such, Born-Von Karman boundary conditions are an effective and efficient approximation for representing the periodicity of a *continuous* and *infinite* crystal. However, this approximation breaks down for small values of N_j and crystal discontinuities. For example, surfaces and grain boundaries are not efficiently represented by these boundary conditions. However, in principle, they could be modelled with sufficiently large unit cells.

B.4 The Bloch Theorem

Adapted from pages 85-88 of Electronic Structure[99].

Recall the Time Independent Schrodinger Equation:

$$\hat{H}\psi(\mathbf{r}) = \left[-\frac{\hbar^2}{2m_e}\nabla^2 + V(\mathbf{r}) \right](\mathbf{r}) = \varepsilon_i(\mathbf{r}) \tag{B.15}$$

It is noted that $\hat{H}\psi(\mathbf{r})$ is translation invariant because both ∇^2 and $V(\mathbf{r})$ are translation invariant. Thus, we can define a *translation operator* $\hat{T}_{\mathbf{n}}\psi(\mathbf{r})$, where:

$$\hat{T}_{\mathbf{n}}\psi(\mathbf{r}) = \psi[\mathbf{r} + \mathbf{T}(\mathbf{n})] = \psi(\mathbf{r} + n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3) \tag{B.16}$$

Since $\hat{H}\psi(\mathbf{r})$ is invariant to any translation $\mathbf{T}_{\mathbf{n}}\psi(\mathbf{r})$, the Hamiltonian operator must *commute* with the translation operator $\hat{T}_{\mathbf{n}}$. Therefore:

$$\hat{H}\hat{T}_{\mathbf{n}} = \hat{T}_{\mathbf{n}}\hat{H} \tag{B.17}$$

Consequently, the eigenstates $\psi(\mathbf{r})$ of the Hamiltonian operator $\hat{H}$ can also be eigenstates of all $\hat{T}_{\mathbf{n}}$. Moreover, all eigenfunctions of $\hat{H}$ (solutions of the Schrodinger equation) can be articulated in terms of $\hat{T}_{\mathbf{n}}$:

$$\hat{T}_{\mathbf{n}}\psi(\mathbf{r}) = t_{\mathbf{n}}\psi(\mathbf{r}) \equiv \hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{B.18}$$

This relation enables a more efficient means to solve the Schrodinger Equation, for a crystal lattice because all translation operators $\hat{T}_{\mathbf{n}}$ obey the relation:

$$\hat{T}_{n_1}\hat{T}_{n_2} = \hat{T}_{n_1+n_2} \tag{B.19}$$

In other words, if a particle is translated at a distance n_1, then n_2, the overall translation is equal to n_1+n_2. Incorporating the above relation into the above eigenvalue equation further allows a translation

operator, corresponding to a Hamiltonian operator, to be expanded in terms of each of its primitive vectors $t(\mathbf{a}_i)$:

$$\hat{T}_{\mathbf{n}_1}\hat{T}_{\mathbf{n}_2}\psi(\mathbf{r}) = t_{\mathbf{n}_1+\mathbf{n}_2}\psi(\mathbf{r}) = t_{\mathbf{n}_1}t_{\mathbf{n}_2}\psi(\mathbf{r}) \tag{B.20}$$

wherein:

$$t_{\mathbf{n}} = [t(\mathbf{a}_1)]^{n_1}[t(\mathbf{a}_2)]^{n_2}[t(\mathbf{a}_3)]^{n_3} \tag{B.21}$$

Consequently, a Hamiltonian operator articulated in terms of translation operators can be more efficiently *block diagonalised*.

B.4.1 Imposing Born-Von Karman Boundary Conditions

Incorporating the Born-Von Karman periodic boundary conditions into Equation B.21 ensures that the wavefunction is bound to the periodicity of the Bravais Lattice. This results in:

$$t(\mathbf{a}_i) = e^{\frac{i2\pi}{N_i}} \tag{B.22}$$

where $t(\mathbf{a}_i)^{N_i} = 1$ and N_i represents the number of unit cells in the crystal.

B.4.2 Imposing Reciprocal Lattice Conditions

Correspondingly, the *reciprocal lattice* can be generated for said system using a periodic form of Equation B.13:

$$f(\mathbf{q}) = \frac{1}{\Omega}\sum_{n_i}\int_{\Omega} d\mathbf{r} f(\mathbf{r}) e^{(i\mathbf{q}\cdot\mathbf{r}+\mathbf{T}(n_i))} \tag{B.23}$$

this can be further separated into unit cell and periodic terms as:

$$f(\mathbf{q}) = \frac{1}{N_i}\sum_{n_i} e^{i\mathbf{q}\cdot\mathbf{T}(n_i)} \times \frac{1}{\Omega}\sum_{n_i}\int_{\Omega} d\mathbf{r} f(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} \tag{B.24}$$

Recalling the Born-Von Karman boundary conditions, a similar expression to Equation B.14 may be expressed:

$$\mathbf{q} \cdot \mathbf{T}(n_i) = 2\pi \times integer \tag{B.25}$$

Consider further restricting the set of allowable solutions to primitive lattice translations ($\mathbf{a}_i$) :

$$\mathbf{q} \cdot \mathbf{a}_j = 2\pi \times integer \tag{B.26}$$

Consider further a set of reciprocal lattice translation vectors ($\mathbf{b}_i$) that are defined such that $a_j \cdot b_i = 2\pi\delta_{ij}$ where δ_{ij} is an integer. This ensures that the only non-zero Fourier components that are selected correspond to lattice point vectors in reciprocal space. Thus, the real space and reciprocal space lattices may be readily inter-converted using Fourier transforms. Furthermore, the above equation for $t_{\mathbf{n}}$ can be articulated as:

$$t_{\mathbf{n}} = e^{i\mathbf{k}\cdot\mathbf{T_n}} \qquad \text{where} \qquad \mathbf{k} = \frac{n_1}{N_1}\mathbf{b}_1 + \frac{n_2}{N_2}\mathbf{b}_2 + \frac{n_3}{N_3}\mathbf{b}_3 \tag{B.27}$$

Substituting the above into Equations B.16 and B.18 gives:

$$\hat{T}_{\mathbf{n}}\psi(\mathbf{r}) = \psi[\mathbf{r} + \mathbf{T}(\mathbf{n})] = e^{i\mathbf{k}\cdot\mathbf{T_n}}\psi(\mathbf{r}) \equiv \hat{H}\psi(\mathbf{r}) - E\psi(\mathbf{r}) \tag{B.28}$$

B.5 The Born Oppenheimer Approximation

Consider an electronic and nuclear form of the Hamiltonian operator (H) of the Schrödinger equation:

$$H = T_{nuc} + T_{ele} + V_{nuc-nuc} + V_{ele-ele} + V_{ele-nuc} \tag{B.29}$$

where T and V are respective kinetic and electrostatic potential components. Consider electron motion as being unaffected by nuclear motion (neglect T_{nuc}), but is bounded by static electron-nuclear attraction (retain $V_{ele-nuc}$). The corresponding wavefunction (Ψ) can therefore be articulated as:

$$\Psi(\mathbf{R}, \mathbf{r}) = \Psi_{nuc}(\mathbf{R})\Psi_{ele}(\mathbf{R}, \mathbf{r}) \tag{B.30}$$

Consider that $\Psi_{ele}(\mathbf{R}, \mathbf{r})$ can be inserted into the Schrodinger equation to obtain solutions relating only to the electrons:

$$H_{ele} = T_{ele} + V_{nuc-nuc} + V_{ele-ele} + V_{ele-nuc} \tag{B.31}$$

Consider that, if said Schrodinger equation for $\Psi_{ele}(\mathbf{R}, \mathbf{r})$ is solved incrementally for various static values of $\mathbf{R}$ (i.e. $\mathbf{R}$ is treated as a parameter and not a variable), a *potential energy surface* ($E_{ele}(\mathbf{R})$) is obtained that is only a function of $\mathbf{R}$. Now, consider the Hamiltonian operator for $\Psi_{nuc}(\mathbf{R})$:

$$H_{nuc} = T_{nuc} + V_{nuc-nuc} + V_{ele-ele} + V_{ele-nuc} \tag{B.32}$$

Further consider the Hellman Feynman Theorem[489] which implies that if the spatial distribution of the electrons is determined, the forces of the entire system may be derived using classical electrostatic theory. Since $E_{ele}(\mathbf{R})$ represents the spatial distribution of solutions to the electronic Schrodinger equation, all of the nuclear components of the nuclear Hamiltonian operator may be replaced with $E_{ele}(\mathbf{R})$:

$$H_{nuc} = T_{nuc} + E_{ele}(\mathbf{R}) \tag{B.33}$$

Thus, the electronic component of the Schrödinger equation may be solved separately from the nuclear component and then said electronic solutions may be used to infer the solutions for the nuclear component.

B.5.1 The Hellmann-Feynman Theorem

Adapted from Martin's Electronic Structure[99]

The Hellmann-Feynman Theorem considers that force on the nuclei can be expressed solely in terms of the charge density.

Consider the force conjugate $\mathbf{F}$ determining the position of a nucleus $\mathbf{R}$ is given as:

$$\mathbf{F} = -\frac{\partial E}{\partial \mathbf{R}} \qquad \text{via the standard relation:} \qquad Work = \int \mathbf{F} d\mathbf{R} \tag{B.34}$$

Recall the general expression for the total energy:

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle \hat{H} \rangle = \langle \hat{T} \rangle + \langle \hat{V}_{e-e} \rangle + \int d^3 r V_{e-n}(\mathbf{r}) n(r) + E_{n-n} \tag{B.35}$$

Where $\hat{T}$ is the kinetic energy operator, $\hat{V}_{e-e}$ is the electron-electron interaction potential operator, V_{e-n} is the electron-nuclear (external) potential, and E_{n-n} is the nuclear-nuclear interaction energy. Using first-order perturbation theory, and assuming $\langle \Psi | \Psi \rangle = 1$, the above expressions may be combined to give:

$$\mathbf{F} = -\frac{\partial E}{\partial \mathbf{R}} = -\langle \Psi | \frac{\partial \hat{H}}{\partial \mathbf{R}} | \Psi \rangle - \langle \frac{\partial \Psi}{\partial \mathbf{R}} | \hat{H} | \Psi \rangle - \langle \Psi | \hat{H} | \frac{\partial \Psi}{\partial \mathbf{R}} \rangle - \frac{\partial E_{n-n}}{\partial \mathbf{R}} \tag{B.36}$$

Notice that the middle two terms vanish if it is further assumed that the exact ground state wavefunction is an extremity for all variations of the wavefunction. Also, consider that the expectation value of the external potential of the Hamiltonian operator is merely an integral over the charge density. Thus:

$$\mathbf{F} = -\frac{\partial E}{\partial \mathbf{R}} = -\int d^3 r \frac{\partial V_{e-n}(\mathbf{r})}{\partial \mathbf{R}} n(r) - \frac{\partial E_{n-n}}{\partial \mathbf{R}} \tag{B.37}$$

Perturbation Theory

Adapted from Martin's Electronic Structure[99]*. Definitions were provided by Clugston*[16]*.*

To clarify the above, *Perturbation Theory* is a means to solve complex differential equations by assuming a simple, known solution and then applying a small change (a perturbation) to that system. Upon

solving the corresponding equations, the derived solutions effectively act as corrections to the simplified model. An example would be using the exactly calculable wavefunction for a hydrogen atom as the basis for an approximate solution for a hydrogen atom acted on by an electric field. Mathematically, perturbation theory is usually represented as a power series of appropriate order. Solutions may be obtained via the Schrodinger equation:

$$\hat{H} = \hat{H}^0 + \lambda\Delta H + \lambda^2\Delta H... \tag{B.38}$$

B.6 Density Functional Theory

B.6.1 The First Hohenberg-Kohn Theorem

The *First Hohenberg-Kohn Theorem* argues that if the spatially resolved external (nuclear) potential (V_{ext}) in the ground state uniquely determines the spatial distribution of the electron density ($\rho_0(\mathbf{r})$), so too can said external potential be uniquely described by that electron density. Consequently, all of the properties of the system may be determined via the ground state electron density.

Proof

Adapted from pages 123-124 of Electronic Structure[99].

Consider two external potentials $V^1_{ext}(\mathbf{r})$ and $V^2_{ext}(\mathbf{r})$ which correspond to the same ground state density ($n_0(\mathbf{r})$), but differ by more than a constant[b]. Each external potential must have a corresponding Hamiltonian operator, $\hat{H}^1$ and $\hat{H}^2$, and ground state wavefunction, Ψ^1 and Ψ^2. Since, Ψ^2 cannot correspond to the ground state wavefunction of $\hat{H}^1$, therefore:

$$\langle \Psi^1 | \hat{H}^1 | \Psi^1 \rangle < \langle \Psi^2 | \hat{H}^1 | \Psi^2 \rangle \tag{B.39}$$

Next, consider the right-hand side of the above inequality in terms of $\hat{H}^1$ and $\hat{H}^2$:

$$\langle \Psi^1 | \hat{H}^1 | \Psi^1 \rangle < \langle \Psi^2 | \hat{H}^2 | \Psi^2 \rangle + \langle \Psi^1 | \hat{H}^1 - \hat{H}^2 | \Psi^1 \rangle \tag{B.40}$$

Incorporating the relevant external potentials and ground state density into the above equation, and solving the simple expressions for the energy (E), gives:

$$E^1 < E^2 + \int d^3r [V^1_{ext}(\mathbf{r}) - V^2_{ext}(\mathbf{r})] n_0(\mathbf{r}) \tag{B.41}$$

Consider also that the above inequality would also occur for E^2:

[b] Note that the superscripts in $V^1_{ext}(\mathbf{r})$ and $V^2_{ext}(\mathbf{r})$ are merely identifiers and not power operations

$$E^2 < E^1 + \int d^3r[V^2_{ext}(\mathbf{r}) - V^1_{ext}(\mathbf{r})]n_0(\mathbf{r}) \tag{B.42}$$

Therefore adding our above equations for E^1 and E^2 leads to the following contradiction:

$$E^1 + E^2 < E^1 + E^2 \tag{B.43}$$

Thus, there cannot be 2 inequivalent external potentials that result in the same ground state charge density. Therefore, V_{ext} is unique for each $n_0(\mathbf{r})$.

B.6.2 The Second Hohenberg-Kohn Theorem

The *Second Hohenberg-Kohn Theorem* argues that the energy and wavefunction, of such a system, are calculated exactly via a *universal functional* of the electron density ($E[n]$). In particular, the global minimum of $E[n]$ corresponds to the exact ground state energy (E_0) and density ($n_0(\mathbf{r})$)[131].

Proof

Adapted from pages 124-125 of Electronic Structure[99].

Consider that a universal functional for the total energy ($E[n]$) in terms of the electron density (n) has an external potential ($V_{ext}(\mathbf{r})$) defined by the nuclei.

$$E[n] = T_{ele}[n] + V_{ele-ele}[n] + V_{nuc-nuc} + \int d^3r V_{ext}(\mathbf{r})n(\mathbf{r}) \tag{B.44}$$

Define $F_{HK}[n]$ as the *internal energy* (kinetic + potential) of the interacting electronic system.

$$E[n] = F_{HK}[n] + V_{nuc-nuc} + \int d^3r V_{ext}(\mathbf{r})n(\mathbf{r}) \tag{B.45}$$

where $V_{nuc-nuc}$ is the interaction energy of the nuclei. Therefore:

$$F_{HK} = T[n] + E_{int}[n] \tag{B.46}$$

which must be universal, as T and E_{int} are only functionals of the density n. In this context, *"universal"* means the same for all systems regardless of V_{ext}. Consider a ground state density $n^1(\mathbf{r})$ with corresponding $V^1_{ext}(\mathbf{r})$, the energy is given by:

$$E^1 = E_{HK}[n^1] = \langle \Psi^1 | \hat{H}^1 | \Psi^1 \rangle \tag{B.47}$$

consider another density $n^2(\mathbf{r})$. Thus:

$$E^1 = \langle \Psi^1 | \hat{H}^1 | \Psi^1 \rangle < \langle \Psi^2 | \hat{H}^1 | \Psi^2 \rangle = E^2 \tag{B.48}$$

If $F_{HK}[n]$ were known, minimising E for the density ($n(\mathbf{r})$) would give the exact ground state density. Note that this does not help with excited states. A different treatment is required.

B.6.3 The Kohn-Sham Equation

Adapted from pages 135-138 of Electronic Structure[99].

DFT is an *ansatz*, meaning that it assumes the mathematical form of a problem. It replaces the many-body problem, which limits the application of HF theory, with an *auxiliary* system that can be calculated more cheaply. As such, DFT does not explicitly solve or approximate the Schrödinger equation. Rather, it solves a Schrodinger-like equation, known as the Kohn-Sham Equation. The Kohn Sham equation relies on two assumptions. The first is that the ground state density can be exactly represented by an auxiliary system of non-interacting particles. The second is that said auxiliary system may be represented via the kinetic energy operator, as for HF theory, and an effective *local potential* (V_{eff}) that acts on an electron of spin (σ) at a point ($\mathbf{r}$):

$$\hat{H}^\sigma_{DFT} = -\frac{1}{2}\nabla^2 + V^\sigma_{eff}(\mathbf{r}) \tag{B.49}$$

Note that $V^\sigma_{eff}(\mathbf{r})$ has no functional form yet. The electron density may be given as the sum of squares of the orbitals per spin:

$$n(\mathbf{r}) = \sum_\sigma n(\mathbf{r}, \sigma) = \sum_\sigma \sum_{i=1}^{N^\sigma} |\Psi^\sigma_i(\mathbf{r})|^2 \tag{B.50}$$

Correspondingly, the kinetic energy operator (T) is given as:

$$T = \frac{1}{2}\sum_{\sigma}\sum_{i=1}^{N^{\sigma}}\int d^3r|\nabla\Psi_i^{\sigma}(\mathbf{r})|^2 \tag{B.51}$$

Consider the classical coulomb interaction energy (Hartree energy) of the electron density interacting with itself.

$$E_{Hartree} = \frac{1}{2}\int d^3r d^3r' \frac{n(\mathbf{r})n(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|} \tag{B.52}$$

Combining the above expression into a separable Hamiltonian operator gives:

$$E_{KS}[n] = T[n] + \int d\mathbf{r}V_{ext}(\mathbf{r})n(\mathbf{r}) + E_{Hartree}[n] + E_{nuc-nuc} + E_{xc}[n] \tag{B.53}$$

In this way, all of the many-body effects of exchange are contained within the exchange-correlation energy term ($E_{xc}[n]$). Comparing Equation B.45 into Equation B.53, one can notice that:

$$E_{xc}[n] = F_{HK}[n] - T[n] - E_{Hartree}[n] \tag{B.54}$$

Therefore. $E_{xc}[n]$ is a functional of the electron density that is merely the difference in the kinetic and internal interaction energies of the system. If this were exactly calculable, then the exact ground state energy and density could be obtained. This relation is useful because the independent particle kinetic energy and long-range Hartree terms are separated from the exchange and correlation energy, which can be approximated as a *local* function of the density:

$$E_{xc}[n] = \int d\mathbf{r}n(\mathbf{r})\epsilon_{xc}([n],\mathbf{r}) \tag{B.55}$$

where $\epsilon_{xc}([n],\mathbf{r})$ is the energy per electron at $\mathbf{r}$ that is a functional of the *local* density around that point.

B.7 The Projector Augmented Wave Method

Adapted from pages 225-226 of Electronic Structure[99].

In the *Projector Augmented Wave* (*PAW*) formalism, the wavefunction is expressed as a sum of valance and core functions. This is similar to the *"muffin-tin"* approach wherein the core functions (close to the nucleus) are modelled similarly to an atom, but the interstitial (valence) regions are modelled via plane waves. Computationally, this separable problem is much less expensive.

Projectors ($\tilde{p}$) are functions that operate upon the wavefunction and are localised in space such that they are only non-zero within a given pseudopotential cutoff radius. The advantage of using projectors is that the pseudopotential operator is separable in terms of both electronic states and their assignments as core/valence states. This allows matrix calculations to be performed via products of projection operations. This is more efficient than a direct model-ion approach that involves radial integrations between pairs of functions. For example, consider the expansion valence wavefunction for a system $\psi(\mathbf{r})$ and a linear transformation (T) that relates the set of all-electron valence functions to the smooth functions within a sphere (T_0) around a given nucleus. Outside this sphere, it may be assumed that the result of the transformation is equal to 1.

$$\psi = T\tilde{\psi} \qquad \text{where} \qquad T = 1 + T_0 \tag{B.56}$$

Consider the expansion of $\psi(\mathbf{r})$ in terms of m partial waves within each sphere:

$$|\tilde{\psi}\rangle = \sum_m c_m |\tilde{\psi}\rangle \tag{B.57}$$

Transform according to T:

$$|\psi\rangle = T|\tilde{\psi}\rangle = \sum_m c_m |\psi\rangle \tag{B.58}$$

Combining the above relations, we may write the full wavefunction as:

$$|\psi\rangle = |\tilde{\psi}\rangle + \sum_m c_m[|\psi\rangle - |\tilde{\psi}\rangle] \tag{B.59}$$

Since the transformation is linear, we may represent the coefficients as a projection within each sphere.

$$c_m = \langle \tilde{p}_m | \tilde{\psi} \rangle \tag{B.60}$$

Selecting only solutions that satisfy the orthogonality conditions ($\langle \tilde{p}_m | \tilde{\psi}_{m'} \rangle = \delta_{mm'}$), the above may be further combined to obtain:

$$|\psi\rangle = 1 + \sum_m [|\psi\rangle - |\tilde{\psi}\rangle] \langle \tilde{p}_m| \tag{B.61}$$

The above expression may be applied separately to core and valence states by adjusting the precise form of $\tilde{p}$.

B.8 The Murnaghan Equation of State

Consider Hooke's Law:

$$F = k\Delta x \tag{B.62}$$

where F is the applied extensive force, k is the elastic constant for the material, and Δx is the length of extension of the material. In built into Hooke's Law is the assumption that the elastic constant is invariant under applied stress. However, it is known that the compressibility of a material varies under different applied pressures. Thus, in reality, k is not a constant and depends on the initial pressure acting on the system before extension. Murnaghan addressed this problem by considering infinitesimally small variations in stress for Hooke's Law. By considering Hooke's Law, in combination with the conservation of mass, and assuming that the elastic constant varied as a linear function of pressure, Murnaghan arrived at the following equation:

$$\frac{V_0 - V}{V_0} = 1 - (1 + kP)^{-\frac{1}{ck}} \tag{B.63}$$

where V_0 is the volume with no applied pressure, V is the volume of the material (crystal unit cell), P is the applied pressure, and c is a coefficient that represents the linear variation of k with P[280]. Modern Murnaghan equation of state fitting tools typically reparameterise the Murnaghan equation of state in terms of the total energy, unit cell volume, and bulk modulus[490] [c]. Thus, the equilibrium volume (V_0), lattice parameters and bulk modulus may be estimated by fitting the Murnaghan Equation of state across a range of unit cell volumes and their corresponding total energies.

[c] The bulk modulus is the material dependent change in volume under applied pressure[16]. Thus it serves a similar purpose as k in Hooke's Law.

B.9 The Madelung Potential

Adapted from Atkins' Physical Chemistry[491]

Consider an infinite one-dimensional chain of alternating positive and negative ions, separated by a constant distance (d). Assume that the overall system is charge neutral, thus $z_1 = -z_2$ where z_i is the charge of ion i in the chain. Consider an arbitrarily chosen central cation and its potential interaction with all particles to its right. This can be given as:

$$V_{cat} = -\frac{z^2e^2}{4\pi\varepsilon_0 d}(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4}...) = -\frac{z^2e^2}{4\pi\varepsilon_0 d}ln(2) \tag{B.64}$$

where ε_0 is the free space permittivity and e^2 is the charge of an electron. Multiplying this equation by two accounts for the ions to the left. Additionally, the same expression may be used to represent an arbitrarily chosen central anion (V_{ani}). Thus, the total potential per mole felt by an arbitrary pair of ions, divided by two to account for double counting, is therefore:

$$V = \frac{1}{2}N_A(V_{cat} + V_{ani}) = -2N_Aln(2)\frac{z^2e^2}{4\pi\varepsilon_0 d} \tag{B.65}$$

where N_A is Avogadro's constant[d]. This may be extended to three dimensions, following the particular lattice geometry, to get:

$$V = -M\frac{|z_Az_B|e^2}{4\pi\varepsilon_0 d} \tag{B.66}$$

where M is the dimensionless Madelung constant for that lattice geometry.

[d] Avogadro's constant is the number of species per mole of substance; a mole being the exact number of carbon atoms 12 grams of carbon 12. The mole is used to normalise chemical quantities to make assessing the stoichiometry of a chemical reaction more simple

B.10 The Ewald Summation

Adapted from pages 327-356 of Molecular Dynamics by Santamaria[492]

The *Ewald Summation* technique aims to efficiently calculate the electrostatic energies of a periodic system. It does this by separating the short-range and long-range interactions within the lattice and handling these components separately. The reasoning behind this is that the interaction energies between nearby atoms (short-range) have a greater dependency on the interatomic distance than distant (long-range) atoms. In particular, short-range interactions are more efficiently calculated in real space, while long-range interactions are more efficiently calculated in reciprocal space. Additionally, Gaussian functions are used to represent the atomic species as their mathematical form complements the transformations required to maximise efficiency. To outline the basics of the theory, consider an arbitrarily selected *central* unit cell within an infinitely periodic lattice. The central cell contains N atoms and is charge neutral such that:

$$\sum_{i}^{N} q_i = 0 \tag{B.67}$$

where q_i is the charge of atom i. For a given charge q_i at $\mathbf{r}_i$ within the central cell, its electrostatic interaction energy E_{el} with the rest of the lattice is given as:

$$E_{el} = \frac{1}{2}\frac{1}{4\pi\varepsilon_0}\sum_{n_1,n_2,n_3=-\infty}^{\infty}\sum_{i=1}^{N}\sum_{j=1}^{N'}\frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}L|} \tag{B.68}$$

where $\mathbf{n} = n_1, n_2, n_3$ represents the unit cell parameters and L is the unit cell length. Note that the central cell corresponds to $\mathbf{n} = 0$ in this formalism. The above equation is too expensive to compute directly. Thus, consider an alternative formalism where a distribution function $\rho_i(\mathbf{r})$ is used:

$$\rho_i(\mathbf{r}) = q_i\delta(\mathbf{r} - \mathbf{r}_i) \tag{B.69}$$

Reparameterising E_{el} accordingly gives:

$$E_{el} = \frac{1}{2}\frac{1}{4\pi\varepsilon_0}\sum_{\mathbf{n}}\sum_{i,j=1}^{N'}\int\int\frac{\rho_i(\mathbf{r})\rho_j(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'} + \mathbf{n}L|}d^3r d^3r' \tag{B.70}$$

Next, Consider the electrostatic potential v_i at point $\mathbf{r}$ excluding particle i:

$$v_i(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \sum_{\mathbf{n}} \sum_{j=1}^{N'} \int \frac{\rho_j(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'} + \mathbf{n}L|} d^3r' \tag{B.71}$$

$$v_i(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} [\sum_{\mathbf{n}} \sum_{j=1}^{N} \int \frac{\rho_j(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'} + \mathbf{n}L|} d^3r' - \int \frac{\rho_i(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} d^3r'] \tag{B.72}$$

Correspondingly:

$$E_{el} = \frac{1}{2} \sum_{i=1}^{N} \int \rho_i(\mathbf{r}) v_i(\mathbf{r}) d^3r \tag{B.73}$$

Next, consider separable short-range ($\rho_i^s(\mathbf{r})$) and long-range ($\rho_i^l(\mathbf{r})$) distribution functions:

$$\rho_i(\mathbf{r}) = \rho_i^s(\mathbf{r}) + \rho_i^l(\mathbf{r}) \tag{B.74}$$

where:

$$\rho_i^s(\mathbf{r}) = q_i \delta(\mathbf{r} - \mathbf{r}_i) - q_i g(\mathbf{r} - \mathbf{r}_i) \qquad \text{and} \qquad \rho_i^l(\mathbf{r}) = q_i g(\mathbf{r} - \mathbf{r}_i) \tag{B.75}$$

$g(\mathbf{r})$ is a Gaussian function that may be defined as:

$$g(\mathbf{r}) = \alpha e^{-\beta r^2} \qquad \text{where} \qquad \alpha = \frac{1}{(2\pi\sigma^2)^{\frac{3}{2}}} \qquad \text{and} \qquad \beta = \frac{1}{2\sigma^2} \tag{B.76}$$

The Gaussian function is used to screen the point charges that are represented by the Delta functions of the opposite sign. For reference, the electrostatic potential of the Gaussian function is given by:

$$v(r) = \frac{\alpha}{4\beta\varepsilon_0 r} \sqrt{\frac{\pi}{\beta}} erf(\sqrt{\beta} r) \tag{B.77}$$

erf is the error function that accounts for the inability to compute the integral of the Gaussian function directly. By combining the above equations and simplifying the result, the short-range electrostatic potential and energy may be derived:

$$E^s = \frac{1}{2}\sum_{i=1}^{N} q_i v_i^s(\mathbf{r}_i) \tag{B.78}$$

where:

$$v_i(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\sum_{\mathbf{n}}\sum_{j=1}^{N'}\frac{q_j(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}+\mathbf{n}L|}erfC(\sqrt{\beta}|\mathbf{r}-\mathbf{r'}+\mathbf{n}L|) \qquad \text{and} \qquad erfC(x) = 1 - erf(x) \tag{B.79}$$

E^s is short-ranged because $erfC(x)$ decays rapidly as a function of distance within the lattice. This quantity may therefore be efficiently computed in real space. To account for the self-interaction, the Poisson equation is solved for a Gaussian-type function in the limit of $erf(x \to 0)$. This results in:

$$E_{self} = \frac{1}{4\pi\varepsilon_0}\frac{1}{\sqrt{2\pi\sigma^2}}\sum_{i=1}^{N} q^2 \tag{B.80}$$

which is a constant. In the limit of large interatomic distance, $erf(x \to \infty) = 1$. This convergence is slow in real space but rapid in reciprocal space. Thus, for long-range interactions, it is preferable to work in reciprocal space. In particular, the Fourier transform of the long-range distribution function does not require summations across $\mathbf{n}$. This is accounted for by the usual reciprocal space translation arguments and incorporation of a Gaussian function:

$$\rho_i^l(\mathbf{r}) = \sum_{\mathbf{n}}\sum_{i=1}^{N}\rho_i^l(\mathbf{r}+\mathbf{n}L) \qquad , \qquad \tilde{\rho}_i^l(\mathbf{k}) = \frac{1}{V}\int_V \rho_i^l e^{-i\mathbf{kr}}d^3r = \frac{1}{V}\sum_{j=1}^{N} q_j \int \alpha e^{-\beta(\mathbf{r}-\mathbf{r}_j)^2}e^{-i\mathbf{kr}}d^3r \tag{B.81}$$

The corresponding long-range electrostatic potential is given by:

$$\tilde{v}^l(\mathbf{k}) = \frac{1}{V\varepsilon_0}\sum_{j=1}^{N}\frac{q_j}{k^2}e^{-i\mathbf{kr}_j}e^{-\frac{\sigma^2k^2}{2}} \tag{B.82}$$

Thus:

$$E^l = \frac{1}{2V\varepsilon_0}\sum_{k\neq 0}\frac{|S(\mathbf{k})|}{k^2}e^{-\frac{\sigma^2k^2}{2}} \qquad \text{where} \qquad |S(\mathbf{k})|^2 = |\sum_{j=1}^{N} q_j cos(\mathbf{kr}_j)|^2 + |\sum_{j=1}^{N} q_j sin(\mathbf{kr}_j)|^2 \tag{B.83}$$

Recombining the long, short, and self-interaction terms according to:

$$E_{el} = E^s + E^l - E_{self} \tag{B.84}$$

is overall more efficient than a direct summation.

B.11 Special Quasi-Random Structures

Adapted from the work of Zunger et al:[432]

For an AB-type mixed lattice, consider a *configuration* (σ) of atoms within a lattice that can be deconstructed into a series of geometric functions (f)

$$f = (k, m) \tag{B.85}$$

where k corresponds to the number of *vertices* (atoms) for a given f and m corresponds to the distances between said vertices. Further, consider a boolean occupation variable $\hat{S}_i$ that determines the occupation of a given lattice (i) as being either A or B (i.e. +1 or -1). The product (Π_f) of $\hat{S}_i$ for each i within a given f can be defined as:

$$\Pi_f(l, \sigma) \tag{B.86}$$

where l corresponds to the location and orientation of the given f within the lattice. Moreover, if there are D_f geometric functions per site (i), a lattice average ($\bar{\Pi}_f(\sigma)$) of Π_f across all l and symmetry-related f is defined as:

$$\bar{\Pi}_f(\sigma) = \frac{1}{N D_f} \sum_i \Pi_f(l, \sigma) \tag{B.87}$$

where N corresponds to the number of lattice sites. Further, consider the ensemble average[e] of a physical property (P) over all σ:

$$\langle P \rangle = \sum_{k,m} D_{k,m} \langle \bar{\Pi}_{k,m} \rangle p_{k,m} \tag{B.88}$$

where $p_{k,m}$ are the interaction parameters and $\langle \bar{\Pi}_{k,m} \rangle$ are the correlation functions. Consider that for a perfectly random binary system, the correlation function is given as:

$$\langle \bar{\Pi}_{k,m} \rangle_{Random} = (2x - 1)^k \tag{B.89}$$

since when the composition parameter (x) is equal to $\frac{1}{2}$, all correlations vanish. Thus, instead of

[e] An *ensemble* being the collection of possible configurations that may be possible in a dynamic system [16], the *ensemble average* being the average of a quantity across these configurations.

computing the ensemble average by experimenting with a large number of permutations, structures (S) may be selected such that the resulting correlation function is as close to $\langle \bar{\Pi}_{k,m} \rangle_{Random}$ as possible. This may be expressed as:

$$\langle P \rangle_{Random} - P(S) = \sum_{k \neq 0, k \neq 1, k, m} D_{k,m}[(2x-1)^k - \bar{\Pi}_{k,m}(S)]p_{k,m} \tag{B.90}$$

B.11.1 Monte Carlo SQS

Adapted from the work of Van de Walle et al:[433]

The above method arrives at a final structure via exhaustive enumeration. Whilst reliable, the overhead still increases rapidly as the cell size or complexity increases. The Monte Carlo SQS (mcsqs) method generates SQS models *stochastically*, such that the computational overhead is further decreased:

$$Q = -\omega L + \sum_{\alpha} |\rho_\alpha - \rho_\alpha^{target}| \tag{B.91}$$

where Q is the *Objective Function* to be minimised during the Monte-Carlo algorithm; ω is a user-specified weighting for the class of cluster (doublets, triplets, quadruplets, etc); α is the cluster index, denoting the cluster in question; ρ_α represents the correlations of α; ρ_α^{target} represents the *target* correlations of the *truly random* configuration; and L is a user-specified radius, such that all clusters within L are optimised ($|\rho_\alpha - \rho_\alpha^{target}| = 0$).

B.12 Derivation of the OER Scaling Relation

Adapted from the work of Man et al[58].

B.12.1 Chemical Potential Equations in Terms of DFT Energies

Consider the chemical potential (μ) which may be defined as the partial derivative (∂) of the Gibbs free energy (G) of a system of particles (N)[16] :

$$\mu_i = \left(\frac{\partial G}{\partial N_i}\right)_{T,P,N_{j\neq i}} \tag{B.92}$$

Where i and j are individual particles of the N particles, T is the temperature, and P is the pressure. Two systems in equilibrium have the same chemical potential. As such, consider a system of solvated protons and electrons that are in equilibrium with gaseous hydrogen.

$$2H^+(aq) + 2e^- \leftrightharpoons H_2(g) \tag{B.93}$$

The corresponding expression, in terms of chemical potentials, is therefore given as:

$$\mu_{H^+} + \mu_{e^-} = \frac{1}{2}\mu_{H_2(g)} \tag{B.94}$$

The chemical potential is further related to the activity (a_i) and the partial pressure (p_i) via:

$$\mu_i = \mu_i^0 + k_B T ln(a_i) \qquad \text{and} \qquad a_i = \frac{p_i}{p^0} \tag{B.95}$$

Where μ_i^0 is the chemical potential at standard conditions ($p^0 = 1$ bar, for example), k_B is the Boltzmann constant, and T is the temperature. Applying this to gaseous hydrogen, solvated protons, and solvated electrons gives:

$$\mu_{H_2(g)} = \mu_{H_2}^0 + k_B T ln(p_{H_2}) \tag{B.96}$$

$$\mu_{H^+} = \mu_{H^+}^0 + k_B T ln(a_{H^+}) \tag{B.97}$$

$$\mu_{e^-} = \mu_{e^-}^0 - eU \tag{B.98}$$

where eU represents the shift in electron energy under an applied bias[f]. As such, assuming a system in equilibrium at standard conditions:

$$\mu^0_{H^+} + \mu^0_{e^-} = \frac{1}{2}\mu^0_{H_2(g)} \tag{B.99}$$

Recall that the Gibbs free energy G is related to the enthalpy (H), temperature (T) and entropy (S) via:

$$G(T,p) = H(T,p) - TS(T,p) \tag{B.100}$$

Consider that the energy of a DFT calculation may equal the enthalpy in combination with a Zero Point Energy (ZPE) correction. Thus, the standard chemical potential of hydrogen ($\mu^0_{H_2(g)}$), in terms of calculable DFT quantities, can be expressed as:

$$\mu^0_{H_2(g)} = E^{H_2(g)}_{DFT} + ZPE_{H_2(g)} - TS^0_{H_2(g)} \tag{B.101}$$

Where $E^{H_2}_{DFT}$ is the total energy of a hydrogen molecule, $ZPE_{H_2(g)}$ is the zero point energy correction computed via DFT vibrational frequency calculations performed by Man et al[58], and $S^0_{H_2}$ is the standard entropy of hydrogen obtained from a standard data table cited by Man et al[58].

B.12.2 The Hydrogen Bonding Approximation

One aspect of the system that DFT cannot properly account for is the hydrogen bonds of liquid water. To avoid this problem entirely, an approximation is introduced in which the gaseous chemical potential is equal to the liquid chemical potential.

$$\mu_{H_2O(l)} = \mu_{H_2O(g)} \tag{B.102}$$

This only happens to be true at 0.035 bar and 298.15 K. As such:

$$\mu_{H_2O(g)} = E^{H_2O(g)}_{DFT} + ZPE - TS^0(0.035\text{bars}) \tag{B.103}$$

[f] A *bias* is simply an applied voltage[70].

B.12.3 The Oxygen Molecule Approximation

A further approximation is introduced to overcome the difficulty in calculating the energy of the oxygen molecule with DFT. In this regard, for the water splitting reaction:

$$2H_2O(l) \leftrightharpoons O_2(g) + 2H_2(g) \tag{B.104}$$

the corresponding experimental standard free energy change (2.46 eV) is used to consider the above relation in terms of the Gibbs free energy:

$$2G^0_{H_2O(l)} - G^0_{O_2(g)} - 2G^0_{H_2(g)} = 2.46 \text{ eV} \tag{B.105}$$

Substituting the above equations for hydrogen and water into the above equation and rearranging for $G^0_{O_2(g)}$ indirectly provides the value for oxygen.

$$\left(E^{O_2(g)}_{DFT} + ZPE_{O_2(g)} - TS^0_{O_2(g)}\right) = 4.92 + 2\left(E^{H_2O}_{DFT} + ZPE_{H_2O(g)} - TS^0_{H_2O}\right) \\ -2\left(E^{H_2(g)}_{DFT} + ZPE_{H_2(g)} - TS^0_{H_2(g)}\right) \tag{B.106}$$

B.12.4 Step One of the Oxygen Evolution Reaction

With each of the *reagents* of the OER described in terms of DFT energies, the *intermediates* of the OER can further be described in terms of DFT energies. Recall the first step of the OER:

$$H_2O(l) + M \leftrightharpoons M(OH) + H^+ + e^- \tag{B.107}$$

A corresponding equation in terms of the Gibbs free energy can be written:

$$\Delta G_1 = G_{M(OH)} + \mu_{H^+} + \mu_{e^-} - G_M - \mu_{H_2O(l)} \tag{B.108}$$

where $G_{M(OH)}$ and G_M are the energies of the surface with and without the hydroxyl adsorbate. Correspondingly, these quantities can be expressed in terms of DFT energies:

$$G_{M(OH)} = E^{M(OH)}_{DFT} + ZPE_{OH} - TS^0_{OH} \qquad \text{and} \qquad G_M = E^M_{DFT} \tag{B.109}$$

Moreover, by combining the preceding two equations, we arrive at:

$$\Delta G_1 = G_{M(OH)} + \mu^0_{H^+} + k_B T ln(a_{H^+}) + \mu^0_{e^-} - eU = G_M - \mu_{H_2O(g)} \tag{B.110}$$

Further substituting Equations B.101 and B.103 into the above equation, we arrive at:

$$\Delta G_1 = E^{M(OH)}_{DFT} - E^{M}_{DFT} - E^{H_2O(g)}_{DFT} - \left(E^{H_2O(g)}_{DFT} - \frac{1}{2} E^{H_2(g)}_{DFT} \right) + (\Delta ZPE - T\Delta S^0) + k_B T ln(a_{H^+}) - eU \tag{B.111}$$

B.12.5 Step Two of the Oxygen Evolution Reaction

Recall the second step of the OER:

$$M(OH) \leftrightharpoons M(O) + H^+ e^- \tag{B.112}$$

A corresponding equation in terms of the Gibbs free energy can be written:

$$\Delta G_2 = G_{M(O)} - G_{M(OH)} + \mu_{H^+} + \mu_{e^-} \tag{B.113}$$

The same method of substitution, as for the first step, can be applied to the above equation for the second step to obtain an equation in terms of DFT energies:

$$\Delta G_2 = E^{M(O)}_{DFT} - E^{M(OH)}_{DFT} + \frac{1}{2} E^{H_2(g)}_{DFT} + (\Delta ZPE - TS^0) + k_B T ln(a_{H^+}) - eU \tag{B.114}$$

B.12.6 Step Three of the Oxygen Evolution Reaction

Recall the third step of the OER:

$$M(O) + H_2O(l) \leftrightharpoons M(OOH) + H^+ + e^- \tag{B.115}$$

A corresponding equation in terms of the Gibbs free energy can be written:

$$\Delta G_3 = G_{M(OOH)} + \mu_{H^+} + \mu_{e^-} - G_{M(O)} - \mu_{H_2O(l)} \tag{B.116}$$

The same method of substitution, as for the previous two steps of the OER, can be applied to the above equation for the third step to obtain an equation in terms of DFT energies:

$$\Delta G_3 = E_{DFT}^{HOO*} - E_{DFT}^{O*} - \left(E_{DFT}^{H_2O(g)} - \frac{1}{2} E^{H_2}(g) \right) + (\Delta ZPE - T\Delta S^0) + k_B T ln(a_{H^+}) - eU \quad \text{(B.117)}$$

B.12.7 The Final Step of the Oxygen Evolution Reaction

Recall the third step of the OER:

$$M(OOH) \leftrightharpoons M + O_2(g) + H^+ + e^- \quad \text{(B.118)}$$

A corresponding equation in terms of the Gibbs free energy can be written:

$$\Delta G_4 = G_M + G_{O_2(g)} + \mu_{H^+} + \mu_{e^-} - G_{M(OOH)} \quad \text{(B.119)}$$

The same method of substitution, as for the previous three steps of the OER, can be applied to the above equation for the final step to obtain an equation in terms of DFT energies:

$$\Delta G_4 = E_{DFT}^{*} - E_{DFT}^{HOO*} + \left(2E_{DFT}^{H_2O(g)} - \frac{1}{2} E_{DFT}^{H_2(g)} \right) + 4.92 + (\Delta ZPE - T\Delta S^0) + k_B T ln(a_{H^+}) - eU \quad \text{(B.120)}$$

B.13 Experimental Methods

For the reader's reference, the following section contains background material regarding the various experimental methods that were mentioned in previous chapters. This section merely outlines the general ideas behind each technique and not necessarily the practical intricacies. Such information may be obtained from the citations provided in this section.

B.13.1 X-Ray Diffraction

A sample of the crystal is probed with a range of monochromatic X-rays at various orientations. As the X-ray passes through the crystal the beam is scattered by the electrons of the structure. The scattered beams usually destructively interfere, with the exceptions of specific characteristic angles at which they constructively interfere. These constructively interfering beams produce a diffraction pattern. Application of Bragg's Law[493] and a Fourier Transform to these diffraction patterns allows the crystal structure to be deduced.

$$n\lambda = 2dsin\theta \tag{B.121}$$

where n is the diffraction order, λ is the wavelength of the X-ray radiation, d is the distance between successive planes of atoms, and θ is the angle is incident radiation normal to the plane of the material's surface. Alternatives include neutron and electron diffraction which broaden the range of applications for the technique[494,495].

B.13.2 Stern-Gerlach Experiment

In the Stern Gerlach experiment[94], silver atoms were passed through an inhomogeneous (spatially varying) magnetic field to a detector. Particles with non-zero magnetic moments (i.e. electrons) would therefore be deflected before reaching the detector.

The magnetic field had to be inhomogeneous to ensure that the electrons were deflected. Assuming, that the electron (a charged particle) indeed possesses an angular momentum, then that angular momentum must produce a magnetic dipole because a moving charge is known to produce a magnetic field. However, if the magnetic dipole passes through a spatially homogeneous magnetic field, the

forces acting upon the magnetic dipole, due to the magnetic field, would cancel each other out and the particle would pass through undisturbed. However, a spatially inhomogeneous magnetic field would not produce this effect and there would be a net force acting on the magnetic dipole. Thus, altering its trajectory. Silver atoms were chosen due to their charge neutrality, ensuring that the deflection path was not too great and that the magnetic effects of the electrons dominated the observed result.

It was observed that, due to the quantum nature of spin, the deflected particles accumulated in one of two specific locations on the detector. Thus, supporting the existence of up and down spin states. If the electrons' angular momenta were insignificant, or not quantum in nature, then a continuous distribution of particles would have accumulated at the detector.

B.13.3 Transient Absorption Spectroscopy

A *pump* pulse is used to excite a fraction of electrons in the material to an excited state. Following a delay time (τ), a second *probe* pulse of the same wavelength (λ) but of lower intensity than the pump pulse, is used to excite the material further. The difference (ΔA) between the pump absorption (A_{pump}) and probe absorption (A_{probe}) can then be considered as a function ($\Delta A(\lambda, \tau)$) of the delay time and the wavelength.

$$\Delta A(\lambda, \tau) = A_{probe} - A_{pump} \tag{B.122}$$

By experimenting with various wavelengths and delay times, the spectroscopic data can be used to characterise the dynamic processes that may occur within the system. For example, a transient absorption spectrum presents a characteristic negative ΔA peak at the ground state absorption wavelength. This is known as the *ground state bleach* peak, so named because the pump pulse would have already promoted a substantial fraction of the electrons at this wavelength. Thus causing the probe pulse to excite fewer electrons than the probe pulse. Therefore, the ground state bleach peak is always negative[496]. Polarons can be detected using transient absorption spectroscopy by considering the absorption spectrum as a function of time and wavelength in light of the band gap[29,295]. For instance, excitations within the band gap and characteristic decay times of certain signals may indicate polaron formation.

B.13.4 Cyclic Voltammetry

Cyclic voltammetry uses an electrochemical cell to probe the electronic structure of an analyte, alongside an associated chemical process. The electrochemical cell comprises a working electrode, a reference electrode, a counter electrode and the analyte in question. By applying an electrical current to the system and observing how the system behaves as a function of the applied current and voltage, a voltammogram can be obtained. From this voltammogram, the various potentials (i.e. the ionisation potential and electron affinity) associated with the analyte can be deduced. This is because, as the potential applied to the working electrode is increased, the energy of the electrons in the working electrode increases accordingly. At a specific potential, the electrons are of sufficient energy to be transferred to the analyte. A redox process then occurs between the analyte, the working electrode, and the counter electrode which corresponds to identifiable signatures within the voltammogram. The reference electrode is a well-known reference potential that calibrates the overall reaction such that the band energies deduced for the analyte are absolute[497].

B.13.5 Photoelectron Spectroscopy

Photoelectron spectroscopy exploits the photoelectric effect which is the observation that an electron can only be removed from a material (i.e. ionised) via electromagnetic radiation once the radiation is of a sufficient threshold frequency. The intensity of the radiation does not affect this relation[16]. The threshold frequency is characteristic of the material in question.

As such, probing the material with various frequencies, and observing the photoelectrons produced, may be used to characterise the material. Moreover, the emitted photoelectrons provide useful information concerning the electronic structure. For instance, assuming that Koopmans' Theorem is valid, the energy bands of the material correspond to the characteristic threshold frequencies associated with the material. This includes the determination of the ionisation potential[498]. Generally speaking, valence electronic states are probed via ultraviolet radiation, whilst core electronic states are probed via X-rays. X-ray photoelectron spectroscopy may also be used to determine the oxidation state of an atom[499], thereby providing an estimation of the ionicity.

B.13.6 Sputter Deposition

Sputter deposition is a means to generate a thin film of material on a substrate. In this technique, a source of the material is bombarded with ionised particles such that fragments of the material are ejected from the surface. Under the correct conditions, these ejected atoms may land on the substrate, gradually forming a thin film. Sputter deposition is an example of a physical deposition process, wherein the chemistry of the materials involved is not exploited in the deposition process[500]. Mixed systems (alloys) may be created by using multiple sputtering sources targeted at a single substrate.

B.13.7 Chemical Vapour Deposition

Converse to physical vapour deposition, chemical vapour deposition is an alternative deposition process that exploits the chemistry of the materials involved. A precursor compound is vaporised and either reacts or decomposes, on the substrate surface to create a thin film[501].

B.13.8 Inductively Coupled Plasma Mass Spectrometry

In mass spectrometry, a material sample is ionised, causing charged fragments of the sample to break free from the bulk. The mass/charge ratio of these fragments can be used to characterise the sample according to the known masses of the constituent elements. For inductively coupled plasma mass spectrometry, the ionised fragments are in the form of an inductively coupled plasma that allows discrete samples of the material to be ionised in a time-resolved fashion. By statistically analysing each of these discrete samples across time, the homogeneity of the sample can be deduced[431,502].

B.13.9 Optical Absorption Spectroscopy

Optical absorption spectroscopy exploits the physical phenomenon that materials exhibit wherein the specific electromagnetic frequencies that are absorbed (and not transmitted, scattered or reflected) by a material are characteristic and unique to the material[16]. This is caused by, and is evidence of, the the quantisation of electronic states within the material.

B.13.10 Tauc Analysis

The band gap may be determined experimentally via a *Tauc analysis* following *optical absorption spectroscopy* wherein[29]:

$$(ah\nu)^{n/2} \qquad \text{is plotted against} \qquad h\nu \tag{B.123}$$

where h is Planck's constant, ν is the frequency of light, E_g is the band gap, a is the absorption coefficient of the material, $n = 1$ for a direct band gap and $n = 4$ for an indirect band gap. The resulting plot usually includes a linear portion the intercept of which corresponds to the band gap of the material.

B.13.11 Nanoparticle Synthesis

Nanoparticles may be synthesised via *"top-down"* methods, where the bulk material is ground down mechanically, ablated via a laser, or sonicated. Alternatively, they may also be synthesised in a *"bottom-up"* approach, where the nanoparticles are grown via a chemical synthesis. A comprehensive review of top-down and bottom-up processes is given by Namakka et al[503].

Mathematical Background

The purpose of this chapter is to introduce key mathematical concepts that may be required to understand the previous chapters. It is assumed that the reader has a basic knowledge of arithmetic, algebra and calculus.

C.1 Domains, Ranges, Functions and Functionals

A *function* (f) is a rule that assigns elements from a set (a collection of objects) (X) to another set (Y) such that such that each element of X corresponds to an element of Y. The *domain* of a function is the set of values that can be assumed by the independent variable (x) of the function. The *range* of a function is the set of elements that can be assumed by the dependent variable (y) of the function[100]. Thus:

$$f : X \to Y \qquad \text{and} \qquad y = f(x) \tag{C.1}$$

A *functional* is a function that has a domain that is a set of functions and a range that belongs to another set of functions or numbers[100]. An example of a functional is the *differential operator* ($\frac{d}{dx}$), which is a functional of the function $f(x)$. Generally speaking, the terms *"operator"* and *"function"* are often used interchangeably throughout literature. For this work, *operators* refer to well-known and used functions, such as the aforementioned differential operator.

Laplacian Operator

The Laplacian operator is a second-order differential operator applied across Euclidean space. As such, it provides a convenient shorthand for representing spatially resolved transformations.

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}, \frac{\partial^2}{\partial z^2} \right) \tag{C.2}$$

C.2 Matrices

A *matrix* is a rectangular array of numbers or functions. They are particularly useful for computers for storing interpreting and manipulating data:

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{C.3}$$

For example, spatially resolved charge density data can be stored in matrix form. Like scalar numbers, matrices can be added, subtracted, and multiplied. They may also be *transposed* to swap the rows and columns:

$$M^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \tag{C.4}$$

The *determinant* of a matrix, of order n, is a scalar associated with an $n \times n$ (*square*) matrix that describes the solutions of a linear system of equations when written in matrix form:

$$det(M) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{C.5}$$

The *inverse* of a matrix is defined as:

$$M^{-1} = \frac{1}{det(M)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{C.6}$$

The inverse of a matrix acts as the reciprocal of its counterpart, M. As such, when applied to another matrix or vector, it can function as division by M. *Diagonal matrices* are square matrices that have only non-zero terms across the diagonal.

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \tag{C.7}$$

Scalar matrices are diagonal matrices in which each of the diagonal elements is equal. As such, the matrix functions in the same way as a scalar number when multiplied.

$$\begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix} \tag{C.8}$$

Unit matrices are scalar matrices in which all of the diagonal values are equal to 1.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{C.9}$$

Tridiagonal matrices are square matrices whose elements are non-zero on the principal diagonal, the first *superdiagonal*, and the first *subdiagonal*:

$$\begin{bmatrix} a & h & 0 & 0 \\ e & b & i & 0 \\ 0 & f & c & j \\ 0 & 0 & g & d \end{bmatrix} \tag{C.10}$$

C.2.1 Diagonalisation

Diagonalisation transforms a non-diagonal matrix into a diagonal matrix. The overall computational power required for performing calculations using the diagonalised matrix is lower because common matrix manipulations may be simplified via the large number of zero terms within the diagonalised matrix. For example, the determinant of a diagonalised matrix is merely the product of its diagonal entries. Diagonalisation is performed via a *similarity transformation*, wherein an $n \times n$ diagonal matrix M^{Diag} is *similar* to another $n \times n$ matrix M if[504]:

$$M^{Diag} = P^{-1}MP \tag{C.11}$$

where P is also an $n \times n$ matrix. *Block Diagonalisation* may also be used in certain circumstances to further reduce the computational expense of diagonalisation. If the elements of a matrix are conveniently partitionable into subsets such that adjacent elements within each subset may be grouped into

their own smaller matrices, then said matrix is called a *block matrix*. Or more simply, a matrix whose elements are also matrices. Much like a standard matrix, block matrices may also be diagonalised into diagonal block matrices. Therefore, depending on the situation, it may be more computationally efficient to block diagonalise a given matrix.

C.3 Eigenvalues, Eigenvectors and Eigenfunctions

An *eigenvector* (ν) is a non-zero vector associated with a linear operator ($\hat{O}$) or matrix (M) that satisfies the relationship:

$$\hat{O}\nu = \lambda\nu \qquad \text{or} \qquad M\nu = \lambda\nu \tag{C.12}$$

where λ is a scalar quantity called the *eigenvalue*[16]. An *eigenfunction* is a type of eigenvector that is also a function. For example, the function:

$$f(x) = e^{\lambda x} \tag{C.13}$$

is an eigenfunction of the differential operator ($\frac{d^2}{dx^2}$) because $f(x)$ is a function that also satisfies Equation C.12 :

$$\frac{d^2}{dx^2}e^{\lambda x} = \lambda e^{\lambda x} \tag{C.14}$$

C.3.1 Eigenbases and Matrices

The concept of eigenvectors, eigenvalues, and eigenfunctions also applies to matrices and is particularly useful for diagonalisation. If an $n \times n$ matrix M has n distinct eigenvalues, then M is considered to have an *eigenbasis* (a basis of eigenvectors), $\nu_1, ..., \nu_n$[504]. This relation is useful because it allows any arbitrary vector, $\mathbf{r}$, to be decomposed into a linear combination of the eigenvectors of M:

$$\mathbf{r} = c_1\nu_1 +, ..., +c_n\nu_n \tag{C.15}$$

Moreover, upon diagonalisation of M via Equation C.11, the eigenvalues of M are preserved and expressed along the principal diagonal of M^{Diag}. Furthermore, the eigenvectors of M correspond to the matrix columns (*column vectors*) of P[504].

C.4 Taylor Series

Taylor's Theorem submits that for any smooth function ($f(\theta)$), there exists a single and infinite *Taylor series* that perfectly represents the function:

$$f(\theta)_{Taylor} = \sum_{n=0}^{\infty} \frac{f^n(\alpha)}{n!}(\theta - \alpha)^n = f(\alpha) + \frac{f'(\alpha)}{1!}(\theta - \alpha) + \frac{f''(\alpha)}{2!}(\theta - \alpha)^2 ... \tag{C.16}$$

where α corresponds to a location along $f(\theta)$. The utility of the Taylor series is in its ability to be truncated into a simpler and finite function to obtain an acceptable approximation at a point (α) along $f(\theta)$. For example, a complex function that is close to parabolic about a certain point of interest (α) may be reasonably approximated by a Taylor series expansion to the order θ^2 about that point. Thus, the only terms to be kept from Equation C.16 would be:

$$f(\theta) \approx f(\alpha) + \frac{f'(\alpha)}{1!}(\theta - \alpha) + \frac{f''(\alpha)}{2!}(\theta - \alpha)^2 \tag{C.17}$$

Such a finite function is called a *Taylor polynomial*.

C.4.1 Maclaurin Series

The *Maclaurin series* expansion of a function ($f(\theta)$) is the Taylor series at which $\alpha = 0$.

$$f(\theta)_{Maclaurin} = \sum_{n=0}^{\infty} \frac{f^n(0)}{n!}(\theta)^n = f(0) + \frac{f'(0)}{1!}\theta + \frac{f''(0)}{2!}\theta^2 ... \tag{C.18}$$

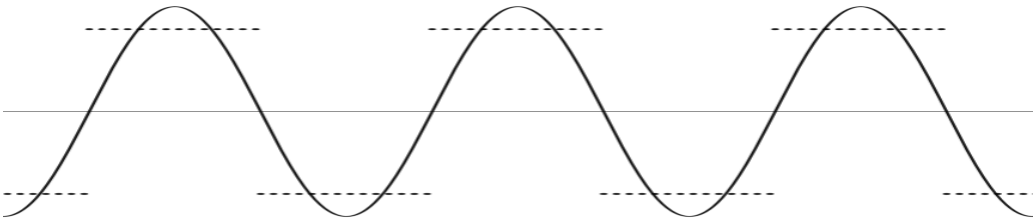
C.5 Fourier Series and Fourier Transforms

A Fourier Series is an infinite mathematical series that can be used to represent any function that is both infinite and periodic in terms of separable sine and cosine functions. Such a representation is often simpler to classify and process than the function itself, especially for a traditional computing system. The general equation for a Fourier series of a given function $f(x)$:

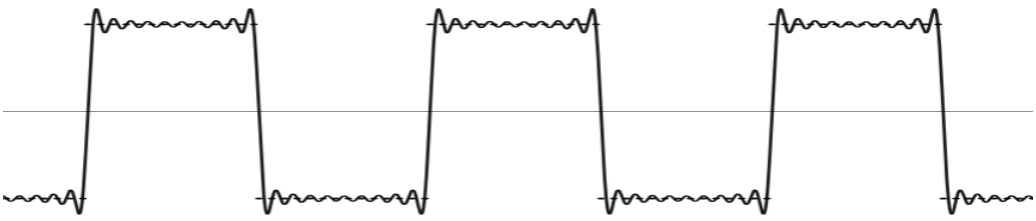
$$f(x) = a_0 + \sum_{n=1}^{\infty} \left[a_n cos\left(\frac{n\pi}{L}x\right) + b_n sin\left(\frac{n\pi}{L}x\right) \right] \tag{C.19}$$

$$\text{where} \qquad a_0 = \frac{1}{2L}\int_{-L}^{L} f(x)dx, \qquad a_n = \frac{1}{L}\int_{-L}^{L} f(x)cos(\frac{n\pi x}{L})dx, \qquad b_n = \frac{1}{L}\int_{-L}^{L} f(x)sin(\frac{n\pi x}{L})dx \tag{C.20}$$

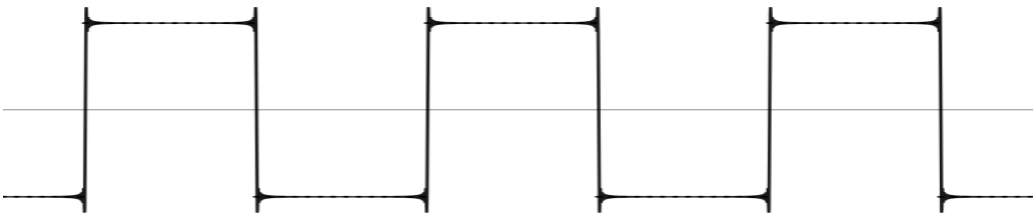
Similar to the Taylor series, if the computational overhead is a concern, the number of terms, n, within a Fourier series can be chosen to sufficiently represent $f(x)$:



(a) n=1



(b) n=10



(c) n=100

Figure C.1: An example of the Fourier series applied to a square wave function, represented by the dashed line, for 3 different values of n.

C.5.1 Fourier Integral

Fourier integrals are an extension of the Fourier Series to *non-periodic* functions. In general, the Fourier Integral for a non-periodic function is given as[504]:

$$f(x) = \int_0^\infty [A(w)cos(wx) + B(w)sin(wx)]dw \tag{C.21}$$

C.5.2 Fourier Transforms

A *Fourier transform* is a specific use of the Fourier integral, wherein a given function ($f(x)$) depending on parameter x is *transformed* into a new function ($f(w)$) which depends on parameter w[504]. For example, reciprocal space is the Fourier transform of Euclidean space.

$$\hat{f}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-iwx}dx \tag{C.22}$$

C.5.3 Inverse Fourier Transforms

Each Fourier transform also has an *inverse* transform. Thus allowing the two domains (e.g. reciprocal and Euclidean space) to be reversibly traversed and analysed[504].

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(w)e^{iwx}dw \tag{C.23}$$

C.5.4 Discrete and Fast Fourier Transforms

A *Discrete* Fourier Transform is an extension of the Fourier Transform to discretely sampled, evenly spaced data across a range, as opposed to a continuous function. The Discrete Fourier transform is particularly useful for Fourier analysis using a computer because data is stored in a computer's memory in discrete units and not across a continuous range[504].

$$\hat{f}_n = Nc_n = \sum_{k=0}^{N-1} f_k e^{-inx_k} \tag{C.24}$$

An issue with the Discrete Fourier Transform is it has a relatively large computational overhead due

to the large number of mathematical operations that often need to be performed. The *Fast* Fourier Transform reinterprets the Discrete Fourier Transform for more efficient computation by a computer. The result is a set of transformations that scale according to $log_2 N$, as opposed to N^2, where N are the number of data points for consideration[504].

C.6 Crystal Structures and Lattices

Many photoelectrochemical cell electrode materials are crystalline, meaning that they are composed of an ordered and periodic arrangement of atoms[16]. This regular arrangement of the atoms may be represented mathematically via a *lattice*[16]. Assuming three spatial dimensions, a lattice is a collection of all *lattice points* that satisfy the relation:

$$a_x v_x + a_y v_y + a_z v_z \tag{C.25}$$

where v_i are linearly independent vectors or points in Euclidean space and a_i are integers.[100]. Consequently, each lattice has at least one characteristic *lattice constant* that spans the distance between lattice points. The degrees of freedom for a lattice include the relative size of each of the lattice constants as well as the angles between them. This mathematical formalism provides a convenient means to categorise different crystal structures because the ordering of the atomic positions in a crystal structure can be mapped directly onto an ordering of lattice points within a lattice.

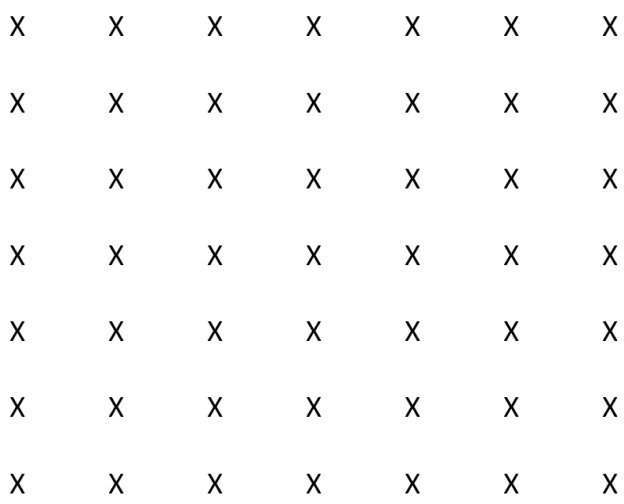


Figure C.2: A 2-dimensional representation of a square lattice. Each X may represent a lattice point and/or an atomic position within the crystal structure.

Further lattice geometries can be categorised according to 7 lattice systems[16]. An illustrative example in three dimensions is the cubic lattice wherein each of the lattice constants, (a, b, and, c) are equal, and the angles between each lattice vector is 90°[100].

Lattice system	Lattice constant	Angles
Cubic	$a = b = c$	$\alpha = \beta = \gamma = 90°$
Tetragonal	$a = b \neq c$	$\alpha = \beta = \gamma = 90°$
Orthorhombic	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90°$
Rhombohedral	$a = b = c$	$\alpha = \beta = \gamma \neq 90°$
Hexagonal	$a = b \neq c$	$\alpha = \beta = 90°, \gamma = 120°$
Monoclinic	$a \neq b \neq c$	$\alpha = \gamma \neq \beta$
Triclinic	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma \neq 90°$

Table C.1: List of 7 lattice systems. When drawn from a common lattice point, a, b and c correspond to lattice constants in Euclidean space. α, β, and γ correspond to the angles between b and c, a and c, and a and b respectively.

Each of the lattice systems is further specified into primitive, base-centred, body-centred and face-centred *Bravais lattices*. A Bravais lattice is a lattice arranged such that the infinite array of neighbouring lattice points appears to be identical when viewed in the same direction from any lattice point[16]. All Bravais lattices obey the following relation wherein a set of vectors, R, join each of the lattice points and span the lattice without overlapping or leaving any gaps:

$$R = n_1 a_1 + n_2 a_2 + n_3 a_3 \tag{C.26}$$

where n_i are integers and a_i are primitive translation vectors. In three dimensions, the relation between a Euclidean lattice vector ($\mathbf{a}_1$) and its Reciprocal counterpart ($\mathbf{b}_1$) is given as[99]:

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{|\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|} \tag{C.27}$$

The representation of the Euclidean primitive cell in reciprocal space is known as the *First Brillouin Zone*[505] (*FBZ*). This representation can be further reduced, by symmetry, to the *Irreducible Brillouin Zone*[506] (*IBZ*).

Within this volume of reciprocal space, the long-range, real space properties of the crystal structure can be calculated exactly by integrating over the irreducible Brillouin zone. Within this mathematical framework, the structure of any crystalline solid may be determined and categorised according to its space group[16]. The crystal structure is usually verified experimentally via X-ray diffraction (see Section B.13).

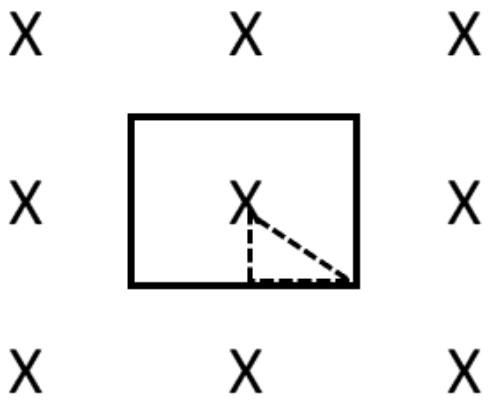


Figure C.3: The first and irreducible Brillouin zone for the square lattice of Figure C.2. The space enclosed by the solid box represents the first Brillouin zone, whilst the space enclosed by the dashed line represents the irreducible Brillouin zone. For reference, Γ corresponds to the coordinates (0,0,0) in reciprocal space.

C.7 Special Functions and Operators

C.7.1 Complex Numbers

A problem often encountered is the need to solve $\sqrt{-1}$, which has no real solution. To traverse this problem the *imaginary unit* (i) is defined as[100]:

$$i = \sqrt{-1} \tag{C.28}$$

Correspondingly, quantities that involve $\sqrt{-1}$ are deemed *imaginary*, whilst quantities that do not are considered *real*. Quantities that contain both real and imaginary components are deemed *complex* and have the form:

$$a + ib \tag{C.29}$$

where a and b are real numbers. Practically speaking, it is usually desirable to group the imaginary components, so that they may be treated separately from the real parts. In some cases, this may conveniently result in the imaginary parts cancelling each other out. For example, via $i^2 = -1$. The complex conjugate of a complex quantity is given as:

$$a - ib \tag{C.30}$$

C.7.2 Euler's Formula

Euler's formula is a particularly important relation because it provides a simple mathematical form as the basis for crystalline wavefunctions. Furthermore, Euler's formula simultaneously provides a means with which to deconstruct such wavefunctions into functions of sine and cosine for Fourier analysis. Thus, providing a much more convenient framework for computation.

$$e^{i\theta} = cos(\theta) + isin(\theta) \tag{C.31}$$

Derivation of Euler's Formula

Adapted from page 58 of Advanced Engineering Mathematics[504].

Consider the function $f(\theta)$ wherein:

$$f(\theta) = e^{i\theta} \tag{C.32}$$

Now consider the Maclaurin series expansion of $f(\theta)$:

$$f(\theta)_{Maclaurin} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!}... \tag{C.33}$$

Separating the real and imaginary components gives:

$$f(\theta)_{Maclaurin} = \left[1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!}...\right] + i\left[\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!}...\right] \tag{C.34}$$

Now consider the functions $g(\theta)$ and $h(\theta)$, wherein:

$$g(\theta) = cos(\theta) \qquad \text{and} \qquad h(\theta) = sin(\theta) \tag{C.35}$$

Now consider the respective Maclaurin expansions of $g(\theta)$ and $h(\theta)$:

$$g(\theta)_{Maclaurin} = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!}... \qquad \text{and} \qquad h(\theta)_{Maclaurin} = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!}... \tag{C.36}$$

Upon comparison of the above equations, it is clear that:

$$f(\theta) = g(\theta) + i[h(\theta)] \tag{C.37}$$

Thus:

$$e^{i\theta} = cos(\theta) + isin(\theta) \tag{C.38}$$

C.7.3 The Delta Function

The Delta function ($\delta(x)$) is a function with a value of zero for all points of x except for a specific point (x_0)[100]. Furthermore, the integral across all points of x is equal to 1. Thus:

$$\delta(x) = 0 \qquad \text{where} \qquad x \neq x_0 \qquad \text{but} \qquad \int_{-\infty}^{\infty} \delta(x)dx = 1 \tag{C.39}$$

The Delta function is a *generalised* function, meaning that it may map onto a continuous function ($f(x)$) as a linear functional. Given these properties, the Delta function is commonly referred to as a *sifting* function, because it may be applied to $f(x)$ at a given value of x_0 to obtain the value of $f(x)$

at x_0:

$$\int_{-\infty}^{\infty} \delta(x - x_0) f(x) dx = f(x_0) \qquad \text{or} \qquad f \xrightarrow{\delta_{x_0}(x)} f(x_0) \tag{C.40}$$

As such, the Delta function and its Fourier Series are commonly used to represent discrete quantities, such as impulses or point charges.

C.7.4 The Gaussian Function

The *Gaussian Function* is a function ($f(x)$) that represents the probability distribution with regards to a random variable (x)[100]:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{C.41}$$

where μ is the mean and σ is the standard deviation of a dataset of values relating to x. Beyond being a useful probability distribution, the Gaussian function may also be used as an adjustable approximation of the delta function and be used to smooth out (reduce the noise) a probabilistic or random dataset. This is known as *Gaussian smearing*.