Localised electronic states in model systems and semiconductors

Abdul Razak El-maslmane

A Thesis presented for the degree of
Doctor of Philosophy in Physics

Physics

November 2019

Dedicated to my mother and father

إهداء لأمي وأبي، بشرى و فتاح

Localised electronic states in model systems and semiconductors

A. R. El-maslmane

Submitted for the degree of Doctor of Philosophy in Physics
November 2019

Abstract

Accurately modelling charge trapping phenomena is a vital part of understanding and improving the behaviour of semiconductors in both current and future devices. While charge traps are frequently observed in doped crystals, the formation of so-called self-trapped polarons means that charges can trap even in defect-free bulk crystals. This in turn reduces charge carrier mobility in materials, sometimes to the detriment of the underlying device. More effective methods of modelling self-trapping typically introduce a few free parameters, each of which can significantly influence results obtained by the model. The focus of this thesis is on developing parameter-free, computationally inexpensive and accurate approaches to modelling charge trapping, with a focus on titanium dioxide, a material used in promising new photovoltaics. Through comparison to both experimental data and solutions to the exact many-electron Schrödinger equation, this thesis demonstrates that use of the generalised Koopmans' theorem in conjunction with hybrid functionals can yield strikingly accurate results. This technique is subsequently used to predict the formation of self-trapped charges in a number of titania phases, including the well-studied rutile and anatase, and less-known brookite, TiO_2(H), TiO_2(R) and TiO_2(B). Intrinsic point defects are also investigated in rutile and anatase, where it is found that several interesting and unique electronic phenomena occur in their vicinity.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

Formal acknowledgements

I would like to thank my supervisor Keith McKenna for his guidance, support and insight on research presented in this thesis. I would also like to thank academic members of staff Rex Godby, Phil Hasnip and Matt Probert for many fruitful discussions and suggestions throughout my enrollment at the University of York. I would also like to thank my collaborator (from Chapter 5) Matt Watkins for fruitful discussions and guidance in running `CP2K` calculations. Special thanks to Jack Wetherell, Matt Hodgson and Rex Godby on an engaging collaborative project presented in Chapter 4 of this thesis. I would also like to thank all members of Keith's group for helpful discussions, in no particular order: Claudio, Adam, James, Steven, Chuanjia, Jon and John.

Informal acknowledgements

As I come to the end of my PhD, a question I'm often asked, sometimes by myself, is whether pursuing a PhD is worthwhile. Without a doubt, the answer is yes. A friend of mine was told by his parents that it is not often you can do something you like and get paid to do it – this statement really sums up my last four years. It is the people I encountered on a near-daily basis that made the journey worthwhile. The academic staff in condensed matter, particularly Phil, Rex and Matt, are incredible researchers and educators. The conversations I've had with them over the years are always interesting, entertaining and inspire curiosity. Members of Keith's group – Claudio, Steven, Adam, James, Chuanjia, Jon and John – always made working on gloomy days like this one so much easier. I will miss sellotaping Adam and James' pens, mice and coffee cups to their desks. Members of Rex's group – Jack, Matt and Mike – have also been incredible people to work with. This trio are some of the most interesting characters I've ever met – they're probably mad, but they always made my day. Visits to the SLB were always entertaining with them. The countless other people I encountered in the department daily – Jenny, Manuel, Leo, Neville, Jacob and many others – played no small part in making my PhD so much more enjoyable. Last but not least, I'd like to thank Keith for both all his efforts with me over the years, and for keeping me off the streets. I could not have asked for a nicer, more helpful and more welcoming PhD supervisor. Without him, this thesis certainly would not exist and I sincerely thank him for his efforts.

Outside the physics department, I'd like to thank my family for their support throughout my enrollment. The number of times they have helped me over the years is beyond counting and I only hope I can someday return the favour. A hugely special thanks to Ellie Phillpotts, who has been the most supportive, caring and loving girlfriend I could have ever asked for. This year marks our seventh year together and she deserves more medals than a general for putting up with me. Hearing her proof read this thesis has really made me realise how much I've learned since January 2016.

DECLARATION

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 4 represents collaborative work that has been published: A. R. Elmaslmane, J. Wetherell, M. J. P. Hodgson, K. P. McKenna, R. W. Godby, 'Accuracy of electron densities obtained via Koopmans-compliant hybrid functionals', *Physical Review Materials (Rapid Communications) 2, 040801 (2018)*. J. Wetherell and I both participated equally in the development of code, execution of calculations and analysis under the guidance of R. W. Godby and K. P. McKenna. I prepared the first draft of the publication. Subsequent drafts were written collaboratively with all listed authors.

Chapter 5 represents collaborative work that has been published: A. R. Elmaslmane, M. B. Watkins, K. P. McKenna, 'First principles modelling of polaron formation in TiO_2 polymorphs' *Journal of Chemical Theory and Computation 14, 3740-3751 (2018)*. I performed all calculations and analysis under the guidance of K. P. McKenna. M. W. Watkins provided training on using `CP2K` and assistance on difficult calculations. All drafts of the publication were prepared by myself and reviewed by K. P. McKenna and M. W. Watkins.

PUBLICATIONS

The following are two publications that contribute to this thesis. Chapters which use wording from these publications have the relevant publication mentioned at the beginning of the chapter.

1. A. R. Elmaslmane, J. Wetherell, M. J. P. Hodgson, K. P. McKenna, R. W. Godby, 'Accuracy of electron densities obtained via Koopmans-compliant hybrid functionals', *Phys. Rev. Materials (Rapid Comms) 2, 040801 (2018)*

2. A. R. Elmaslmane, M. B. Watkins, K. P. McKenna, 'First principles modelling of polaron formation in TiO_2 polymorphs' *J. Chem. Theory Comput. 14, 3740-3751 (2018)*

The following is a study used to write Chapter 6 and is being prepared for publication.

3. A. R. Elmaslmane, K. P. McKenna, 'Intrinsic point defects in TiO_2 rutile and anatase'

The following study (in preparation) uses approaches developed in this thesis but has not been used to contribute to this thesis.

4. G. Schusteritsch, R. Ishikawa, T. Saito, A. R. Elmaslmane, K. Inoue, K. P. McKenna, Y. Ikuhara, C. J. Pickard, 'Localized anatase grain boundary phase embedded in rutile titanium dioxide'

CHAPTER 1

INTRODUCTION

Charge trapping in semiconductors is a widely discussed topic in condensed matter physics, materials chemistry and engineering, and is relevant to a vast range of materials. Many of these materials are used in a number of devices and applications such as batteries, photovoltaics, photocatalysis, self-cleaning glass and even nuclear reactors.[1–8] Certain applications may benefit from additional charge trapping, such as self-cleaning glass where trapped charges can interact with molecular impurities adsorbed onto the glass coating.[9] In other applications, such as photovoltaics, charge localisation leads to undesired effects which reduce device efficiency, such as non-radiative electron-hole recombination and a reduction in charge mobility.[10] It follows that developing accurate and predictive models of this phenomenon can provide a better understanding of key bottlenecks hindering applications. While a number of theoretical approaches have been used to model this phenomenon, many often require the input of experimental data, making them somewhat semi-empirical. Furthermore, this can cast doubt into whether these methods can be used predictively to model or improve the underlying material. The work presented in this thesis builds upon previous theoretical efforts and provides an inexpensive, accurate and non-empirical approach to modelling charge trapping. In this chapter, the motivations for performing studies shown later in this thesis are presented and discussed in the context of charge trapping more generally.

1.1 Motivation

An ideal semiconductor is one which has extended electronic states and a forbidden energy gap, known as the band gap, which separates occupied and unoccupied states.[11] Electrons in such systems occupy dispersive states, well-described by Bloch's theorem. The presence of localised, and strongly bound, electrons and holes near the Fermi energy can give rise to nearly-dispersionless states that lie within the band gap.[8] In some materials these states can only be induced via doping, whilst in others these can form intrinsically through charge carriers breaking the underlying crystal symmetry. The latter of these results in so-called *self-trapped* charges, or polarons, whose presence can hinder the underlying material's conductivity.[8] These form in materials where electron-phonon coupling is sufficiently strong, enabling charge carriers to distort local atomic configurations and occupy bound states. This behaviour has been observed in numerous materials used in new cutting-edge devices which may significantly impact the way we generate and store energy. Through use of theoretical models, charge trapping in materials can be studied to both understand and remedy key bottle-necks present in these devices. However, commonly employed models often rely on experimental data to be parametrised, reducing their predictive power. The series of studies presented in this thesis lay out an inexpensive, accurate and parameter-free approach to modelling localised electronic states in materials.

While theoretical techniques play a central part in this thesis, the material under study, titanium dioxide (TiO_2), is itself technologically significant with a number of valuable applications. As a strongly correlated material, TiO_2 is an example of a material where electronic interactions are difficult to accurately describe using conventional theoretical techniques. The material finds applications in a new generation of photovoltaics, called mesoscopic perovskite solar cells.[12] In the last five years, these have dramatically risen in energy efficiency, going from 14% in 2013 to around 25% this year, rivalling the 27% seen in more traditional silicon-based devices.[13] The primary advantage perovskite solar cells offer is low construction costs, which comes from the nanoporous nature of the devices.[12] A few factors hinder the commercial use of perovskite cells, the main one being the instability of the solar absorber material methyl-ammonium lead iodide ($CH_3NH_3PbI_3$).[14] Another factor, hindering efficiency, is electron-hole recombination, a process in which photo-generated electrons and holes non-radiatively recombine, emitting a phonon into the device.[10] As such

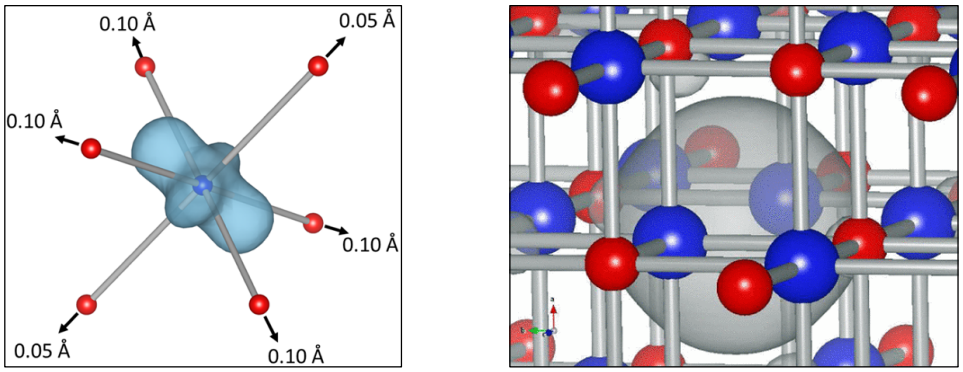

Figure 1.1: *(Left)* Isosurface of a self-trapped polaron in TiO_2 rutile taken from Ref. 15 (Wallace and McKenna, 2015), with titanium and oxygen atoms shown as grey and red spheres. Published by the American Chemical Society under CC BY 4.0. *(Right)* Isosurface of an F-centre defect state surrounding an oxygen vacancy defect in MgO taken from Ref 16 (Rinke *et al*, 2012). Here magnesium and oxygen ions are represented as blue and red spheres. Published by the American Physical Society under CC BY 3.0. DOI: 10.1103/PhysRevLett.108.126404. Article title, journal citation and full list of authors provided in the bibliography.

events reduce energy available to the device, they should ideally be minimised to achieve optimal device efficiency. Charge trapping in particular has been shown to be a leading cause of electron-hole recombination. Trapping has other unfortunate consequences, such as reducing conductivity, which can further reduce device efficiency. The combination of increasing electron-hole recombination and reducing conductivity essentially means that charge trapping is a compounded disadvantage that needs to be remedied.

1.2 Charge trapping in materials

In a large number of semiconductors, the presence of a charge trap will give rise to defect states that lie between the valence band minimum (VBM) and conduction band maximum (CBM). When probed further, these states are usually found to be localised either on an existing atom in the crystal or inside a defect site. The oxygen vacancy defect in MgO (shown in Fig. 1.1) is one such case, where zero, one or two electrons can be bound inside an oxygen vacancy site, resulting in a configuration called an *F-centre* defect.[16] Extrinsic defects, such as dopants or impurities, can also create trap states in a material. The case of Al-doped α-quartz (SiO_2) is one example, where the presence of these impurities results in holes trapping on oxygen ions in their vicinity;. This is in contrast to the pure crystal, where neither electrons nor holes can trap.[8] As previously mentioned, trap states need not occur only at defect sites; in some materials charges

can localise in the defect-free crystal by forming *self-trapped* polarons. A key difference between intrinsically self-trapped charges (such as polarons) and extrinsic trap states arising from impurities is the positioning and mobility of localised charges. Trapped charges from intrinsic point defects or dopants may only be stable in the vicinity of the defect, while crystals that can host self-trapped polarons may in principle form them anywhere. As a result of this, polarons can be considered somewhat more mobile, as they can migrate from site to site via a hopping mechanism. In spite of this apparent mobility, self-trapping charges may still not be a desirable result as overall charge mobility will be reduced.

The self-trapped charges that are a key topic of this thesis are often referred to in literature as small (or Holstein) polarons.[17] They are termed small as the localised charge carriers typically reside on the scale of one to three atoms, all separated by a short distance. Another species of polaron is also often discussed in literature, called a large (or Frölich) polaron, which is much more diffuse and can localise on a larger number of atoms than a small polaron.[17] The scale on which large polarons reside is often prohibitively large to be probed using atomistic techniques. Instead, many opt to use model Hamiltonians to describe them, with the key quantities obtained being an effective polaron mass and radius. These models are idealised and do not account for complex band structures, defects or include realistic electron-phonon interactions.[18] In fact, these models often describe phonons as being dispersionless. While these models may be idealistic, they can provide better physical intuition into the behaviour of polarons, something atomistic techniques may not provide. There are therefore two, somewhat separate, bodies of work that are used to describe polarons in materials: atomistic techniques and model Hamiltonians. This thesis fits into the former category, with its main focus being understanding and applying atomistic techniques. However, recent efforts have attempted to draw both these bodies of work together, using a mixture of model Hamiltonians and atomistic techniques to probe the behaviour of polarons.[18]

1.3 Self-trapping in materials

In its equilibrium structure, a defect-free crystal is assumed to be in its lowest energy configuration (disregarding the effects of temperature); distorting ions in a crystal will have some energy requirement. In materials that can host self-trapped charges, this distortion energy penalty E_{strain} is offset by the energy gained from localising the elec-

tron in a defect-state below the conduction band minimum. The so-called trapping energy E_t can therefore be thought of as the difference in two sets of these energies,

$$E_t^{\mathrm{e}} = \overbrace{(E_{\mathrm{bulk}} - E_{\mathrm{distorted}})}^{\text{Strain energy}} - \overbrace{(\varepsilon_{\mathrm{polaron}} - \varepsilon_{\mathrm{CBM}})}^{\text{Defect state energy}}, . \tag{1.1}$$

The first set of energies $E_{\mathrm{bulk}} - E_{\mathrm{distorted}}$ compare the energies of the pure bulk crystal E_{bulk} and the distorted bulk crystal $E_{\mathrm{distorted}}$, providing a strain energy. The second set of energies compares the eigenvalue of a delocalised CBM state and a localised polaron state, $\varepsilon_{\mathrm{CBM}}$ and $\varepsilon_{\mathrm{polaron}}$. A similar equation can be constructed for hole polarons,

$$E_t^{\mathrm{h}} = \overbrace{(E_{\mathrm{bulk}} - E_{\mathrm{distorted}})}^{\text{Strain energy}} - \overbrace{(\varepsilon_{\mathrm{VBM}} - \varepsilon_{\mathrm{polaron}})}^{\text{Defect state energy}}, \tag{1.2}$$

where the eigenvalues compared in this case are of a delocalised hole at the valence band maximum $\varepsilon_{\mathrm{VBM}}$ and a localise hole polaron state. By distorting the crystal, a defect state whose energy depth is $\varepsilon_{\mathrm{polaron}} - \varepsilon_{\mathrm{CBM}}$ is created. If the energy gained by occupying this defect state is larger than the strain energy required to create it, the localised charge configuration can be considered a stable one. The above (Eqn. 1.1) is useful for understanding the energetics of localising a state, however the trapping energy is more simply calculated as

$$E_t = E_{\mathrm{polaron}} - E_{\mathrm{deloc}} \tag{1.3}$$

where E_{polaron} denotes the energy of a localised polaron configuration and E_{deloc} is the energy of a delocalised electron in the defect-free, unperturbed bulk. [1]

The sheer number of materials where small polarons can form, as well as the substantial number of studies on self-trapping, is highlighted by Table 1.1. While the approaches developed in this thesis are only applied to TiO_2, they can be straightforwardly and inexpensively applied to other materials. Several materials in Table 1.1 could have a significant impact on the modern world if their complications are sufficiently dealt with; one particular material, LiO_2, has been shown to have a theoretical energy density exceeding 10,000 Wh/Kg, closely following gasoline's (approximately)

[1] It should be noted that Eqn. 1.1 is only true within particular frameworks, whereas Eqn. 1.3 is always true; the discussion which takes place in Chapter 4 elucidates this. While it may seem that Eqn. 1.1 does not consider electronic relaxation effects due to the presence of an additional charge, this information is contained within the eigenvalues.

13,000 Wh/Kg.[19] In short, this material is the best battery material currently known. It has a few properties prohibiting its large scale use, such as a large voltage gap between charge and discharge cycles (1000 mV), also called the overpotential.[20] Use of heavier alkali metals, such as NaO_2 and KO_2, causes a reduction in overpotentials to around 100 mV and 50 mV, but unfortunately also reduces theoretical energy capacities.[20] Conductivity in these materials is primarily governed by the movement of hole polarons, also called V_k centres, and metal ion conduction.[21] Our understanding of these battery materials is therefore dependent on employing appropriate and accurate models for localised charges in materials.

1.4 Simulating charge trapping

Density functional theory (DFT), an atomistic technique that is central to this thesis, is frequently used to model real materials from first principles (as seen in Table 1.1). Indeed, use of this approach has expanded rapidly, with the number of publications developing, or using, the technique growing from around 100 per year in 1985 to over 15,000 per year in 2015.[22] Owing to this, two DFT publications, Ref. 23 (Becke, 1993) and Ref. 24 (Lee, Yang and Parr, 1988), are in the Web of Science top ten most cited scientific papers of all time. A detailed discussion of DFT itself takes place in Chapter 2, however the above information should highlight its utility for many researchers as a modelling tool. Despite its successes, the application of DFT has systematic errors that render some of its predictions, particularly those surrounding the localisation of charge, as incorrect if left untreated.

Piecewise linearity of the total energy

Previous studies have established that the exact DFT functional yields a total energy that is piecewise linear with respect to fractional number of electrons[25] and have a slope corresponding to the frontier orbital eigenvalue (see Fig. 1.2).[26,27] Commonly used DFT approximations, such as the local density approximation (LDA) and generalised gradient approximation (GGA), give total energies which are convex with respect to fractional numbers of electrons and cause the spurious delocalisation of charge.[28–30] The Hartree-Fock (HF) approximation on the other hand exhibits concave behaviour and tends to over-localise charge. Hybrid functionals mix these two approaches together, introducing an additional parameter α which determines the ad-

mixture of HF and DFT. Another method, DFT+U, applies an occupation dependent correction to the energy.[31,32] The parameter introduced in this case is the Hubbard-U, which determines the size of the energy correction applied. Both DFT+U and hybrid functional approaches successfully alter charge localisation properties in a calculation, with the hope that more realistic localisation properties are obtained. By choosing an appropriate set of parameters, both approaches can provide energy curvatures closer to the correct piecewise linear behaviour.

The generalised Koopmans' theorem (GKT),[25,116–118] which must be satisfied for an exact Kohn-Sham (KS) or generalised Kohn-Sham (GKS) functional, provides a convenient way to formulate the requirement of piecewise linearity,

$$-I(N+1) \equiv E(N+1) - E(N) = \varepsilon_{N+1}(N+1). \tag{1.4}$$

where $I(N)$ and $E(N)$ are the total energy and ionisation energy of an N electron system and $\varepsilon_i(N)$ denotes the i^{th} eigenvalue of an N electron system. This theorem only makes reference to the highest occupied molecular orbital (HOMO) and so only addresses piecewise linearity between N and $N+1$ electrons. Strictly speaking, this is a necessary but not sufficient condition, however previous work has shown linearity is greatly improved when the condition is satisfied.[119–121] While application of the GKT to localised states (such as polarons) is straightforward using the above equation, its application to delocalised states has had less success.[122] This condition, used in subsequent chapters, enables an essentially parameter-free approach to obtaining the localisation properties of charge using DFT.

1.5 Experimental techniques

A number of experimental techniques are used to probe charge trapping, each of which generate data by exploiting particular features of trap states. It is often the case that ambiguities remain even when experimental data is available. To counter this, studies may use than one experimental technique with several samples, each constructed using a different approach. Below, a non-exhaustive list of experimental techniques relevant to charge trapping is shown, describing key experimental approaches. Other techniques, such as deep-level transient spectroscopy (DLTS), X-ray photoelectron spectroscopy (XPS) and electron energy loss spectroscopy (EELS), have not been described here in as much detail, although relevant information can be found

Chemical	Group	Carrier(s)	Theoretical technique(s)	Experiment(s)	E_g (eV)
$BiVO_4$	$C2/c$	e, h	Hybrid[33,34] DFT+U[35]	EPR[36]	2.4[37]
UO_2	$Fm\bar{3}m$	e, h	DFT+U*[,38]	Conductivity[39]	2.0[40]
Ga_2O_3	$C2/m$	h	Hybrid[41,42]	DLTS[43]	4.8[43]
CeO_2	$Fm\bar{3}m$	e	Hybrid[44] DFT+U[44]	Conductivity[45]	3.3[44]
TiO_2	$P4_2/mnn$	e, h	Hybrid[46] DFT+U[47] RPA[48]	EPR[49]	3.0[50]
TiO_2	$I4_1/amd$	e, h	Hybrid[46] DFT+U[47] RPA[48]	EPR	3.2[51]
HfO_2	$P2_1/c$	h	DFT+U[52] GW*[,53]	EPR[54]	5.6–5.7[55,56]
ZrO_2	$P2_1/c$	h	DFT+U[52] GW*[,53]	EPR[54]	5.6–5.7[57,58]
Fe_2O_3	$R\bar{3}c$	e	Hybrid[59] DFT+U[59–61]	EPR[62] Optical[63,64]	2.0[65]
Li_2O_2	$P6_3/mmc$	e, h	Hybrid[66,67] DFT+U[68] GW[69]	Optical[70] Conduct.[71]	4.9–7.8[69,72†]
LiO_2	$Pnnm$	e, h	Hybrid[73]	–	–
Na_2O_2	$P\bar{6}2m$	e, h	Hybrid[74,75] GW[74,75]	–	4.8–6.7[74,75†]
NaO_2	$Pnnm$	e, h	Hybrid[75] GW[75]	–	5.3[75†]
K_2O_2	$Cmce$	–	–	–	–
KO_2	$I4/mmm$	e, h	Hybrid‡, GW‡	Conductivity[21]	5.0‡
AgCl	$Fm\bar{3}m$	h	DFT+U[76]	EPR[77]	3.3[78]
KBr	$Fm\bar{3}m$	h	–	EPR[79]	8.0[80]
KCl	$Fm\bar{3}m$	h	–	EPR[79]	8.7[81]
$BaSnO_3$	$Pm\bar{3}m$	h	DFT+U[82]	EPR[83]	3.1[84]
$BaCeO_3$	$Pm\bar{3}m$	e, h	Hybrid[85]	Optical[86]	4.1–4.2[86,87]
$BaTiO_3$	$Pm\bar{3}m$	e	Hybrid[88] DFT+U[88,89]	Conductivity[90,91]	3.3–3.4[92,93]
$SrTiO_3$	$Pm\bar{3}m$	h	Hybrid[94] DFT+U[89] GW[95]	EPR[96] Optical[97]	3.3[98]
Cs_2HfCl_6	$Fm\bar{3}m$	e, h	Hybrid[99]	EPR[100,101] Optical[101]	6.0[99†]
Cs_4PbBr_6	$R\bar{3}c$	e, h	Hybrid[102]	Optical[103,104]	2.3[103–105]
$CsPbBr_3$	$Pm\bar{3}m$	e, h	Hybrid[102]	Optical[104]	2.4[103,104]
Cs_2ZrCl_6	$Fm\bar{3}m$	e, h	–	Optical[106]	–
Cs_2LiYCl	$Fm\bar{3}m$	e, h	Hybrid[107]	Optical[108,109] EPR[110]	7.5[109]
$MnWO_4$	$P2/c$	e, h	Hybrid[111]	Conductivity[112]	2.4–3.0[113,114]
$FeWO_4$	$P2/c$	e, h	Hybrid[111]	Conductivity[115]	2.0[114]

* Relevant, however polarons not discussed
– No data found
† No experimental data, theoretical gap(s) used
‡ Unpublished work

Table 1.1: A non-exhaustive list of theoretical and experimental studies on polaronic properties of materials. For each case, carriers e and/or h which are either theoretically predicted, or experimentally observed, to be trapped are shown, along with the shortest optical gap of the material. In the case of theoretical data, quasiparticle gaps were used. A study on KO_2, a potential alkali-metal-air battery material, was taken from the recent PhD thesis of S. Yang (2018),[20] a student under the supervision of Prof. Siegel from Refs 69 and 75.

in Refs 123, 124 and 125.

Electron paramagnetic resonance

Electron paramagnetic resonance (EPR) is used to probe unpaired spins in crystals through use of a magnetic field and photon source.[126] The magnetic field splits elec-

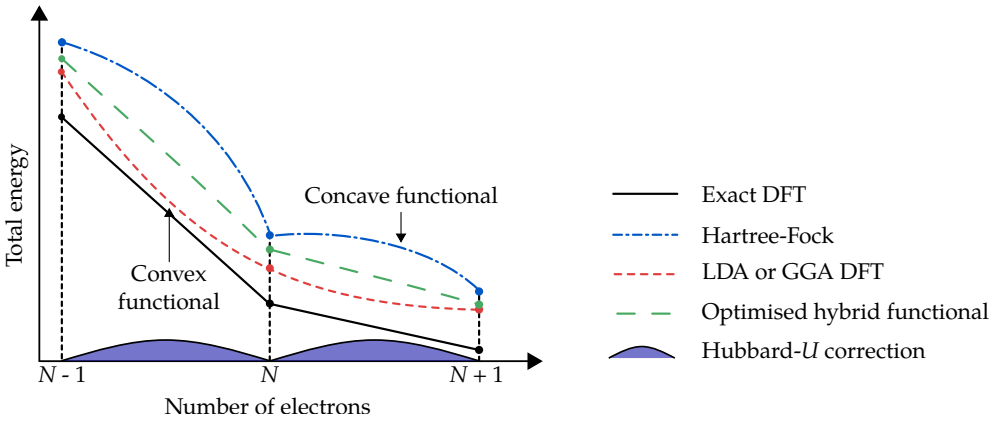

Figure 1.2: A schematic showing the variation of total energy for exact, LDA/GGA and HF with the number of electrons is shown. It can be seen that the exact case is piecewise-linear, LDA/GGA DFT has convex behaviour and HF is concave. An optimised hybrid functional would ideally follow the exact line closely. It should be noted that correcting piecewise linearity does not necessarily correct the total energy, but only its curvature between integer numbers of electrons.

tronic energy levels into spin-parallel and spin-antiparallel configurations (via the Zeeman effect), with the magnitude of the splitting determined by field strength. EPR experiments can be conducted by either applying light of a fixed-frequency and varying magnetic field strength, or through the reverse, and observing (typically microwave) photon absorbance. Usually, photon frequencies are kept fixed whilst the magnetic field is varied. The photon absorbance spectrum is monitored to obtain a g-factor, which depends on the chemical environment an electron is in. The hyperfine structure of atoms can be anisotropic with magnetic field direction, providing further information on the particular atomic orbitals occupied. While this technique can provide information on specific atomic orbitals an electron occupies, it can fail to differentiate between unpaired spins on interstitial and substitutional sites in crystals where symmetries of the two sites are similar.

Conductivity measurements

The first experimental observation of polarons was through conductivity measurements on oxidised uranium dioxide, an intrinsic polaronic semiconductor.[8,127] In the temperature range relevant to nuclear reactors, key thermal properties such as thermal conductivity are determined by the behaviour of small polarons. It was found that hole mobility μ increased with temperature, as described by the formula

$$\mu \propto \frac{1}{T} \exp\left(-\frac{\Delta E}{k_B T}\right) \tag{1.5}$$

which has a Boltzmann factor containing the polaron hopping activation energy ΔE. This increase in mobility with temperature is in contrast with systems where transport is governed by large polarons, which typically decrease with T due to the increased number of phonon scattering events. Measurements of the Seebeck coefficient in CeO_2 are independent of temperature, suggesting that the number of charge carriers does not change with temperature.[128] This implies that conductivity is governed mainly by hole polarons instead of dispersive states as in band theory.[127,128]

Optical measurements

A number of optical techniques can be used to probe polaronic states in semiconductors, such as photoluminescence (PL) or ultraviolet/visible light spectroscopy (UV/Vis). In PL experiments, high-frequency light is directed towards samples to enable the photoexcitation of charges. The absorbed light is then emitted from samples and collected to form a spectra which can provide information on intra-gap states. UV/Vis measures absorbed and reflected photon frequencies on samples, also providing a spectrum to probe gap states. infrared spectroscopy (IR) is another technique that has been used to probe polarons.[129] Similar to other optical approaches, IR experiments generally measure relative absorption by different photon frequencies.

Scanning tunnelling microscopy

Scanning tunnelling microscopy (STM) has been used in a number of TiO_2 studies to both probe and manipulate the material's surfaces.[47,130] The technique, which won its creators a Nobel price in 1981, uses an atomically sharp tip to detect electron tunnelling currents. Surface planes are typically scanned with either constant height or constant current settings. In the latter, the probe adjusts its height at each scanning point to ensure the appropriate current is maintained. Forward and reverse biases can be used to probe occupied and unoccupied states on surfaces. In scanning tunnelling spectroscopy (STS), a range of voltages are explored at a fixed height, providing a local density of states. STS can be particularly helpful when studying differences in electronic structure between defective and defect-free surfaces.

1.6 Research presented

While the work in this thesis primarily focuses on TiO_2, the techniques discussed can be easily and straightforwardly applied to other materials. As previously mentioned, TiO_2 is a strongly correlated material with complex electron-electron interactions which many theoretical techniques often fail to accurately describe. This makes it an ideal material for testing new theoretical techniques, ensuring their suitability. In the first study presented (Chapter 4), the accuracy of the model employed in later chapters (Chapter 5 and Chapter 6) is assessed, and highlights that electron densities and other quantities of interest, such as ionisation energies, are well-described by using the GKT in conjunction with hybrid functionals. Chapter 5 applies the approach to polaronic states in TiO_2 alongside other cost-reducing methods, such as the auxiliary density matrix method, Coulomb interaction truncation and Cauchy-Schwarz integral screening. The combination of theoretical techniques employed in Chapter 5 produces an inexpensive and non-empirical hybrid functional which reproduces results from other similar theoretical studies at a fraction of the computational cost. Furthermore, properties of both the crystal and polarons in TiO_2 are found to be in very good agreement with experimental data. In Chapter 6 the approach is applied once more to intrinsic defects in titania. Literature reviewed in this chapter clearly indicates a lack of consensus, from both experimental and theoretical studies, on the behaviour of point defects in the TiO_2. In this chapter, a number of intriguing findings are discussed, such as the multiple interesting characteristics of oxygen interstitial defects. In both the rutile and anatase phases of TiO_2, the interstitial displayed the strongest trapping behaviour, causing both electrons and holes to localise in its surroundings. For a few of the defects in Chapter 6, findings could explain certain experimentally observed differences between the two phases, such as the substantially higher conductivity of anatase over rutile. Trends could also be drawn by combining the findings of Chapter 6 and Chapter 5, which indicate that electrons are generally much more diffuse in bulk anatase than rutile.

CHAPTER 2

ELECTRONIC STRUCTURE THEORY

Many-body problems in physics are consistently included among the most computationally demanding problems in the world. The complexity of such problems becomes apparent when data from the UK national supercomputer ARCHER is considered, showing that on a typical month the most-used applications are those which specialise in many-body problems (see Fig. 2.1). Treating large numbers of particles without interactions is somewhat trivial; the complexities in many-body problems arise from accurately describing particle interactions. As negatively charged particles, electrons in a crystal exhibit a repulsion between one another which is dampened by the presence of other electrons. This effect is often called 'screening' and cannot be practically described without approximation. If particle interactions are treated exactly, the resulting complexity of the interacting electron problem scales poorly with numbers of electrons, making practical calculations infeasible; even if an exact solution could be obtained, the amount of memory required to store the solution would be enormous. In this chapter, a number of approaches are described that neither treat particle interactions exactly nor ignore them. By taking a happy medium between these two extreme cases, solving the many-electron problem becomes more feasible, practical and inexpensive.

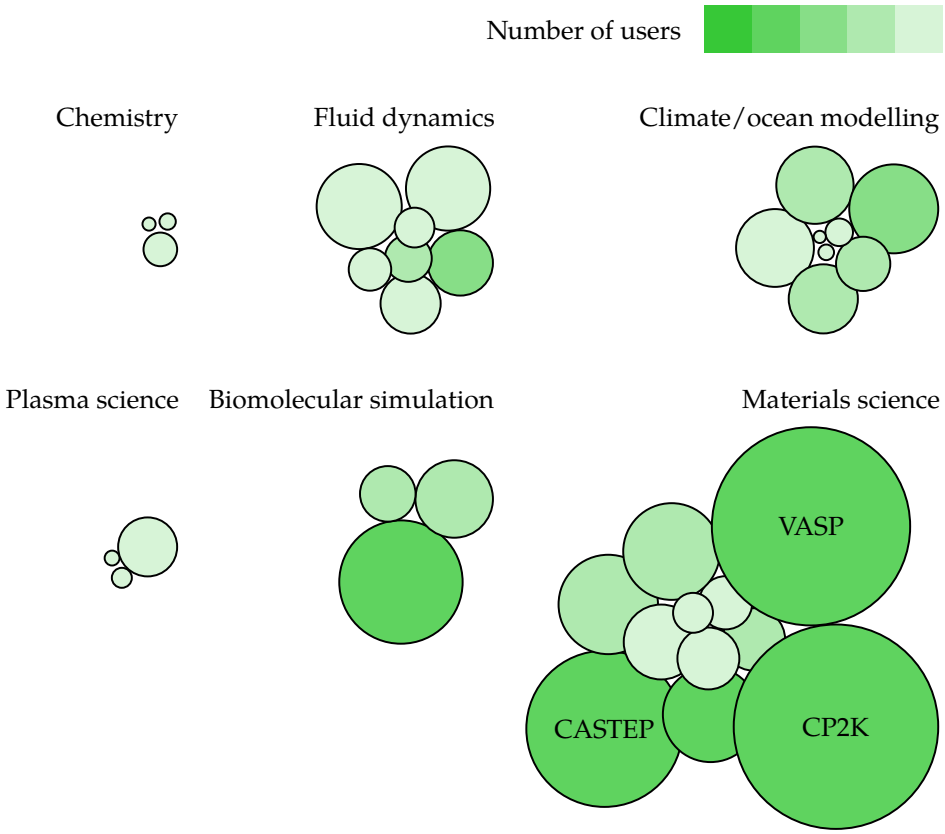

Figure 2.1: UK national supercomputer (ARCHER) application usage by research area for December 2018. The size of each circle is proportional to computational cost in node hours. `CP2K`, `VASP` and `CASTEP`, all of which are electronic structure theory codes, use the largest number of node hours across all applications.

2.1 Many-body quantum mechanics

The systems we aim to describe require simulations of hundreds of nuclei and a few thousand electrons. The wavefunctions Ψ_i, describing all electrons and nuclei in the system, and the Hamiltonian $\hat{\mathcal{H}}$ are both contained within the time independent Schrödinger equation,

$$\hat{\mathcal{H}}\Psi_i = E_i\Psi_i, \tag{2.1}$$

where E_i denotes the energy of each eigenfunction Ψ_i. In principle, solving this equation would allow access to all physical observables via the wavefunctions Ψ_i. These wavefunctions depend on all electron and nuclear positions (denoted $\mathbf{r}_i$ and $\mathbf{R}_i$),

$$\Psi_i = \Psi_i(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e}, \mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_n}). \tag{2.2}$$

In general, an analytic solution to the many-body Schrödinger equation is not possible, therefore the problem needs to be solved numerically. However, the problem can be simplified via a series of approximations to make numerical simulations more feasible.

2.1.1 The low temperature limit

While the effects of temperature in a crystal are occasionally non-negligible, the computational cost associated with taking these effects into account is substantial. In order to correctly assess these effects, a statistical picture, built up over a large number of calculations, is needed. When temperature effects are included, the kinetic energy of atoms in a crystal increases, meaning atoms exhibit more movement around their equilibrium positions. In addition, electrons in a crystal start to occupy higher energy levels as described by the Fermi-Dirac distribution. A commonly prescribed approximation when modelling materials is that calculations are done at the 0 Kelvin limit. At this temperature, the wavefunction describing a crystal only exists in the lowest possible energy, called the ground state wavefunction Ψ_0. While atoms at 0 K do exhibit zero-point motion, this is often disregarded in calculations, and only the equillibrium positions of atoms are considered. These approximations are removed in particular cases when a correct description of atomic motion becomes necessary. Many experimental techniques allow low temperatures to be probed, allowing comparison to theoretical data. The first approximation made is to primarily be concerned with the ground state wavefunction Ψ_0, and disregard zero-point motion.

2.1.2 The variational principle

The variational principle is a method which can be used to aid finding the ground state wavefunction of a system. By definition, a normalised ground state wavefunction Ψ_0 (with energy E_0) is related to any other arbitrary trial wavefunction Ψ (with energy E) as shown below:

$$\begin{aligned} E_0 &\leq E; \\ \langle \Psi_0 | \, \hat{\mathcal{H}} \, | \Psi_0 \rangle &\leq \langle \Psi | \, \hat{\mathcal{H}} \, | \Psi \rangle \, . \end{aligned} \tag{2.3}$$

By using this technique, it is possible to assess whether one wavefunction is closer to the ground state than another via the energy. The ground state energy may also be

degenerate, meaning more than one ground state wavefunction can exist.

2.1.3 The many-body Schrödinger equation

The Hamiltonian describing the electronic system, along with the nuclei, is given by

$$\hat{\mathcal{H}} = \hat{\mathcal{T}}_n + \hat{\mathcal{T}}_e + \hat{\mathcal{V}}_{n-n} + \hat{\mathcal{V}}_{e-n} + \hat{\mathcal{V}}_{e-e}, \tag{2.4}$$

where $\hat{\mathcal{T}}$ and $\hat{\mathcal{V}}$ denote kinetic energy and interaction terms. The subscripts n and e denote nuclear and electronic components of the Kinetic energy (i.e. $\hat{\mathcal{T}}_n$ and $\hat{\mathcal{T}}_e$). The interaction term subscripts $n-n$, $e-n$ and $e-e$ denote interactions between nuclei and electrons in all possible unique combinations. Using the standard definition of the Laplacian operator, the terms are fully expressed below (in Hartree atomic units):

$$\hat{\mathcal{T}}_n = -\frac{1}{2m_n}\sum_i^{N_n}\nabla_i^2, \qquad \hat{\mathcal{V}}_{n-n} = \sum_i^{N_n}\sum_{j>i}^{N_n}\frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|},$$

$$\hat{\mathcal{T}}_e = -\frac{1}{2}\sum_i^{N_e}\nabla_i^2, \qquad \hat{\mathcal{V}}_{e-e} = \sum_i^{N_e}\sum_{j>i}^{N_e}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|},$$

$$\hat{\mathcal{V}}_{e-n} = \sum_i^{N_e}\sum_{j}^{N_n}\frac{-Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}.$$

where charged particles interact via the Coulomb interaction. The quantity m_n is the nuclear atomic mass in Hartree units. In principle, solving the Schrödinger equation (Eqn. 2.1) with the above Hamiltonian will provide access to any physical observables. However, the equation still has no analytic solutions and is extremely difficult to solve computationally.

2.1.4 The Born-Oppenheimer approximation

The Born-Oppenheimer approximation is commonly used in electronic structure theory, and arises from the fact that the nuclei are substantially more massive than an electron, or in the case of crystals the reduced electron mass. This has two key consequences: first, the kinetic energy of the nuclei $\hat{\mathcal{T}}_n$ can be assumed to be very small. Secondly, the electronic and nuclear problems can be separated:

$$\Psi_0 = \Psi_n \Psi_e \tag{2.5}$$

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_e + \hat{\mathcal{H}}_n. \tag{2.6}$$

which decouples the electronic and nuclear wavefunctions. The corresponding Hamiltonian can therefore also be separated into a nuclear $\hat{\mathcal{H}}_n$ and an electronic $\hat{\mathcal{H}}_e$ part, thus simplifying the problem. While it may seem that the equation cannot be separated because of the $\hat{\mathcal{V}}_{e-n}$ term, which depends on both $\mathbf{r}_i$ and $\mathbf{R}_j$, the dependence of this interaction can be made parametric. Later sections discuss how to optimise the positions of nuclei, however for now they will be considered as fixed in space, meaning the term $\hat{\mathcal{V}}_{n-n}$ can be safely neglected. The resulting set of equations essentially reduces the problem to only solving the electronic part of the Hamiltonian,

$$\hat{\mathcal{H}} = \hat{\mathcal{T}}_e + \hat{\mathcal{V}}_{e-e} + v_{\text{ext}}, \tag{2.7}$$

where electrons sit inside a fixed external potential generated by the atomic nuclei. For the remaining sections, the (atomic) potential in which all electrons reside will be called $v_{\text{ext}}(\mathbf{r})$, the external potential, which is analogous to $\hat{\mathcal{V}}_{e-n}$ when all nuclei are fixed. The presence of an external electric field (i.e. not generated by any charge within a system) changes the form of the external potential, however such electric fields will not be considered.

2.2 Hartree and Hartree-Fock theories

While the Born-Oppenheimer approximation simplifies the problem posed by the many-body Schrödinger equation (Eqn. 2.4), the equation still has no analytic solutions. The Hartree approximation is one of the first approximations used to solve the many-electron problem. The Hartree ansatz proposed for the ground state wavefunction is given by

$$\Psi_0^{\text{H}}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\ldots\psi_N(\mathbf{r}_N), \tag{2.8}$$

which is a product state of many single-electron orbitals ψ_i. By using of the many-electron Schrödinger equation (Eqn. 2.7) and variational principle (Eqn. 2.3), the Hartree approximation can be found as

$$\left(-\frac{1}{2}\nabla_i^2 + \overbrace{\sum_{j\neq i}^{\text{occ}} \int \frac{|\psi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j}^{\text{Hartree potential}} + v_{\text{ext}}(\mathbf{r}_i)\right)\psi_i(\mathbf{r}_i) = \varepsilon_i\psi_i(\mathbf{r}_i), \tag{2.9}$$

where one essentially finds a single-particle picture of interacting electrons. In order to simplify notation, the orbital spin degree of freedom is suppressed. The Hartree interaction potential (labelled in Eqn. 2.9) essentially corresponds to an electrostatic potential generated by an electron density (of all other electrons ψ_j) that each orbital ψ_i resides in.

For the Hartree equations, the total energy is not simply a sum of single-particle eigenvalues. Rather, it is given by

$$E_{\text{tot}}^{\text{H}} = \sum_{i}^{\text{occ}} \varepsilon_i \overbrace{- \frac{1}{2} \sum_{i}^{\text{occ}} \sum_{j \neq i}^{\text{occ}} \iint \frac{|\psi_i(\mathbf{r}_i)|^2 |\psi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_i d\mathbf{r}_j}^{\text{Hartree interaction energy, } -E^{\text{H}}}, \tag{2.10}$$

where the additional term removes double-counting errors associated with electron interactions contained in the sum of the eigenvalues. The external potential energy and kinetic energy contributions from $\langle \psi_i | \, v_{\text{ext}} \, | \psi_i \rangle$ and $\langle \psi_i | \, -\frac{1}{2} \nabla_i^2 \, | \psi_i \rangle$ are contained within the eigenvalue sum once only, therefore do not need d.c. corrections.

While in principle calculations using Hartree theory can be calculated in reasonable time, the theory does not explicitly incorporate anti-symmetry of electrons, specifically

$$\Psi_0(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = -\Psi_0(\mathbf{r}_2, \mathbf{r}_1, \ldots, \mathbf{r}_N). \tag{2.11}$$

HF theory remedies this by using a different form for the many-electron wavefunction ansatz, given by

$$\Psi_0^{\text{HF}}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \ldots & \psi_N(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \ldots & \psi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{r}_N) & \psi_2(\mathbf{r}_N) & \ldots & \psi_N(\mathbf{r}_N) \end{vmatrix}, \tag{2.12}$$

known as a Slater determinant. It is simple to show for an N-electron case that wavefunction anti-symmetry is obeyed by using this form for the many-electron wavefunction.

Using the Slater-determinant wavefunction ansatz, the many-electron Schrödinger equation (Eqn. 2.7) and variational principle (Eqn. 2.3) produces the following expres-

sions, known as the HF equations.

$$\left(-\frac{1}{2}\nabla_i^2 + \overbrace{\sum_j^{\text{occ}} \int \frac{|\psi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j}^{\text{Hartree potential}} + v_{\text{ext}}(\mathbf{r}_i) - \hat{\mathcal{F}} \right) \psi_i(\mathbf{r}_i) = \varepsilon_i \psi_i(\mathbf{r}_i) \tag{2.13a}$$

$$\hat{\mathcal{F}}\psi_i(\mathbf{r}_i) = \int F(\mathbf{r}_i, \mathbf{r}_j)\psi_i(\mathbf{r}_j) d\mathbf{r}_j \tag{2.13b}$$

$$F(\mathbf{r}_i, \mathbf{r}_j) = \sum_j^{\text{occ}} \frac{\psi_j^\star(\mathbf{r}_i)\psi_j(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2.13c}$$

The form of these equations is quite different to the Hartree equations as they contain an additional operator, known as the (non-local) Fock exchange operator ($\hat{\mathcal{F}}$). It is called a non-local operator as it includes information into its potential that is non-local in space (i.e. $v(\mathbf{r}_i, \mathbf{r}_j)$ rather than $v(\mathbf{r}_i)$). It is only when the many-body wavefunction is given the appropriate degrees of freedom contained within the Slater determinant that this term arises. A consequence of non-local potentials is each orbital can see a different interaction potential to each other orbital.

The total energy given by the HF equations also contains an additional term corresponding to the Fock operator, given by [1]

$$E_{\text{tot}}^{\text{HF}} = \sum_i^{\text{occ}} \varepsilon_i - E_{\text{H}} + \underbrace{\frac{1}{2}\sum_i^{\text{occ}}\sum_j^{\text{occ}} \iint \frac{\psi_j^\star(\mathbf{r}_i)\psi_i^\star(\mathbf{r}_j)\psi_i(\mathbf{r}_i)\psi_j(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_i d\mathbf{r}_j}_{\text{Hartree-Fock exchange energy, } -E_{\text{x}}^{\text{HF}}}. \tag{2.14}$$

As the spin degree of freedom has been suppressed in this chapter, it should be mentioned that the exchange potential only affects electrons of like spin; for a two-electron helium atom with one electron in each spin channel, the HF equations are reduced to just the Hartree equations.

2.3 Correlated electrons

While HF accurately describes some systems, it behaves very poorly in others. A well-known failure relates to metals, where the HF predicts many of them to be insulators. Thus, this theory is surely lacking some vital component(s) required to ac-

[1] In the Hartree equations, the $i = j$ summation term is excluded from the Hartree energy. The HF equations are usually shown with this term present in both the Hartree and exchange potentials. Regardless of how the equations are presented, the $i = j$ terms cancel due to the Fock operator.

curately describe electrons in crystals – these effects are termed 'correlation' effects. Systems where exchange is the (main) ingredient required to accurately describe electrons are dubbed 'exchange-dominated'. In contrast, systems where correlation effects dominate are called 'strongly correlated'. The term correlation refers specifically to the movement of electrons. In strongly correlated systems, the motion of one electron would result in a significant response of the surrounding electrons, meaning the movement of charges becomes strongly intertwined. Unfortunately, many interesting systems we would like to model require at least some correlation effects included in order to be accurately described. This includes a material studied in this thesis (TiO_2). Thus, the relation between the correlation energy $E_{\text{correlation}}$, exchange energy $E_{\text{tot}}^{\text{HF}}$ and exact total energy E_{exact} of a system can be defined as

$$E_{\text{correlation}} = E_{\text{exact}} - E_{\text{tot}}^{\text{HF}}. \tag{2.15}$$

Despite improving upon Hartree theory, HF still lacks the degrees of freedom which the real ground state many-electron wavefunction has. The true wavefunction is known to require a large sum of Slater determinants. However, finding a set of analytic equations (like the HF equations) is not possible. Instead the problem must be solved numerically (at a high computational cost). Notable methods which methodically include correlation include configuration interaction, Møller-Plesset perturbation theory and the GW approximation. These more sophisticated theories essentially use HF as the 0^{th} order and expand their approach by including correlation effects systematically; in limiting cases, these theories essentially revert to becoming HF.

2.4 Kohn-Sham density functional theory

The KS approach makes use of density functionals (i.e. functionals of the density) and introduces an auxiliary system of non-interacting electrons that should, in principle, reproduce all effects of ground-state many-body systems. This approach introduces some correlation effects and is capable of being an exact theory. In practise, the KS approach often requires approximations in order to be useful. By comparing to more exact techniques, such as quantum Monte-Carlo (QMC), approximate KS approaches are assessed on their ability to reproduce ground-state energies and densities. The key advantage of KS DFT over other techniques lies in its relatively low computational cost and reasonable accuracy.

2.4.1 Hohenberg-Kohn theorems

Pierre Hohenberg and Walter Kohn proved two key theorems, work which later led to the development of the well-known DFT.[131] Walter Kohn later won the Nobel prize in chemistry in 1998 for his contributions to the field, highlighting the significance the KS approach had.

For a fixed number of electrons N_e, the external potential $v_{\text{ext}}(\mathbf{r})$ of a system determines its ground state electronic energy and wavefunction, therefore also the density. The ground state of a system may be degenerate, however degenerate ground states result in the same density and energy unless the degeneracy is lifted. Hohenberg and Kohn proved that this mapping, $\{v_{\text{ext}}\} \rightarrow \{\rho\}$, may also be reversed.[131]

The external potential is a unique functional of the ground state electronic density (to within an additive constant). The (ground state) total energy is therefore also uniquely determined by the same density.

In other words, there is a one-to-one mapping from the ground state density $\rho(\mathbf{r})$ to external potential in many-electron systems. As a result of this theorem, the ground state density can, in principle, uniquely determine all other properties of the system. Such properties include the many-body wavefunction, or energy

$$E[\rho] = F_{\text{HK}}[\rho] + \int \rho(\mathbf{r}) v_{\text{ext}}(\mathbf{r}) d\mathbf{r} \tag{2.16}$$

where $F_{\text{HK}}[\rho]$ is some unknown functional. However, the theorem does not elaborate on how to find the functional $F_{\text{HK}}[\rho]$. Further, it should be noted that this *only* applies to ground state densities; for an arbitrary density (ground state or otherwise), the theorem does not hold as the ground state density of one system may be the excited state of another.

The second Hohenberg-Kohn theorem simply states:

Only the true ground state density minimises the functional which delivers the ground state density.

If the ground state density is denoted ρ_0, the theorem simply states that[131]

$$\left.\frac{\delta E}{\delta \rho}\right|_{\rho=\rho_0} = 0. \tag{2.17}$$

2.4.2 Kohn-Sham equations

KS DFT makes use of the Hohenberg-Kohn theorems and provides a means in which to perform electronic structure calculations. The total energy functional in KS DFT can be decomposed as

$$E[\rho] = T_s[\{\psi_i[\rho]\}] + E_{\mathrm{H}}[\rho] + E_{\mathrm{ext}}[\rho] + E_{\mathrm{xc}}[\rho], \tag{2.18}$$

where T_s, E_H, E_{ext} and E_{xc} are the total single-particle kinetic, Hartree, external potential and so-called exchange-correlation (xc) energies.[132] Of these terms, only the first three have well-defined analytic expressions:

$$T_s[\{\psi_i[\rho]\}] = -\frac{1}{2}\sum_i^{\mathrm{occ}} \int \psi_i^\star(\mathbf{r})\nabla^2\psi_i(\mathbf{r})d\mathbf{r} \tag{2.19a}$$

$$E_{\mathrm{H}}[\rho] = \frac{1}{2}\iint \frac{\rho(\mathbf{r}_i)\rho(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|}d\mathbf{r}_i d\mathbf{r}_j \tag{2.19b}$$

$$E_{\mathrm{ext}}[\rho] = \int \rho(\mathbf{r})v_{\mathrm{ext}}(\mathbf{r})d\mathbf{r}. \tag{2.19c}$$

While analytic expressions for the (total) single-particle kinetic and Hartree energies are known, they are the KS approximations to their corresponding many-electron energies; the real many-electron kinetic and interaction energies are not equal to the single-particle kinetic and Hartree interaction energies. This means that the remaining term, $E_{\mathrm{xc}}[\rho]$, must account for both of these discrepancies for KS DFT to be exact.[132]

It is useful to note that the Hartree energy from the KS approximation *is not* equal to the Hartree energy from the Hartree approximation. If the reader looks closely, Eqn. 2.9 does not contain the so-called 'self-interaction' term; in the Hartree equations, the $j = i$ summation term is omitted from the interaction potential, meaning each orbital does not see its own density in the interaction potential. The HF approach also does not have the self-interaction term due to cancellation. As DFT does not omit this term in the interaction potential, the xc energy must also account for this.[132]

The potentials which correspond to the energies given in Eqn. 2.19b and Eqn. 2.19c can be obtained by performing a functional derivative with respect to the density,

$$\frac{\delta E_{\mathrm{H}}}{\delta\rho} = v_{\mathrm{H}}(\mathbf{r}) = \int \frac{\rho(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|}d\mathbf{r}_j \tag{2.20a}$$

$$\frac{\delta E_{\mathrm{ext}}}{\delta\rho} = v_{\mathrm{ext}}(\mathbf{r}). \tag{2.20b}$$

However, the potential corresponding to E_{xc}, $v_{\mathrm{xc}}(\mathbf{r})$, is still not known, as is the form of E_{xc} itself. The renowned KS equations are

$$\left(-\frac{1}{2}\nabla_i^2 + \overbrace{v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{xc}}(\mathbf{r}) + v_{\mathrm{ext}}(\mathbf{r})}^{\text{Effective potential, } v_{\mathrm{eff}}(\mathbf{r})} \right) \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}), \tag{2.21}$$

where all single-particle orbitals $\{\psi_i(\mathbf{r})\}$ see the same effective potential $v_{\mathrm{eff}}(\mathbf{r})$.

2.4.3 Exchange-correlation approximations

In essence, only the v_{xc} term needs to be found in order to solve these equations exactly, resulting in quantities otherwise obtained via solving the exact many-electron Schrödinger equation. Currently no exact form exists and the potential is known to be non-analytic. This behaviour is discussed in Section 2.4.5. In other words, finding the exact form of this potential will prove to be very challenging, if not impossible, for most systems. In this section, we detail a list of commonly used approximations to xc which render practical calculations possible.

The exchange correlation energy E_{xc} in common density functional approximations can be manipulated as follows.

$$E_{\mathrm{xc}}[\rho] = \int \rho\,(\mathbf{r})\,\varepsilon_{\mathrm{xc}}\,d\mathbf{r} = \underbrace{\int \rho\,(\mathbf{r})\,\varepsilon_{\mathrm{x}}\,d\mathbf{r}}_{E_{\mathrm{x}}[\rho]} + \underbrace{\int \rho\,(\mathbf{r})\,\varepsilon_{\mathrm{c}}\,d\mathbf{r}}_{E_{\mathrm{c}}[\rho]}, \tag{2.22}$$

where $\varepsilon_{\mathrm{xc}}$ is xc energy per electron for a homogeneous electron gas with density $\rho(\mathbf{r})$. The components ε_{x} and ε_{c} are the exchange and correlation parts of $\varepsilon_{\mathrm{xc}}$ respectively.

While an analytic expression for (the exact) exchange energy E_{x} can be found via the Fock operator (Eqn. 2.13b and Eqn. 2.13c), a similar expression for E_{c} is not known. One key point is that the HF potential is *non-local*, whilst KS DFT is primarily concerned with *local* potentials. Even when DFT is extended to include non-local potentials, so-called GKS DFT, using full HF exchange with an approximation to correlation generally yields poor results. Using insight from many-body perturbation theory (via the GW method), part of correlation essentially cancels out HF exchange (through use of the screened Coulomb interaction W for exchange interactions rather than the bare Coulomb one). This means that full HF exchange with the small portion of correlation contained within DFT is not a viable method for modelling materials.

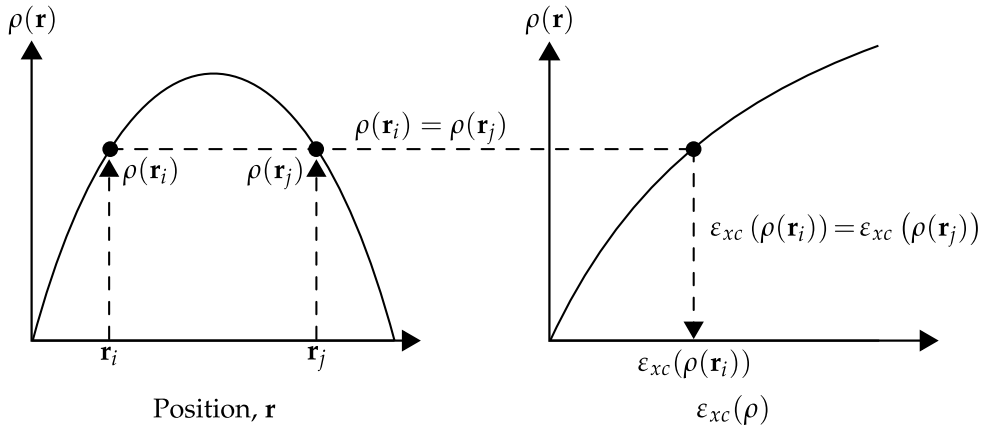

Figure 2.2: A schematic (corresponding to an infinite square well) which illustrates how the LDA calculates $\varepsilon_{\text{xc}}^{\text{LDA}}(\rho)$, thus the energy and potential. In this case, the densities at points $\mathbf{r}_i$ and $\mathbf{r}_j$ are the same, meaning both points in space would have the same value of $\varepsilon_{\text{xc}}^{\text{LDA}}$.

Local density approximation

The LDA was one of the first approximations to xc and has the form

$$\varepsilon_{\text{x}}^{\text{LDA}}(\rho(\mathbf{r})) = -k \cdot \rho(\mathbf{r})^{1/3} \tag{2.23}$$

where k is a constant found analytically that derives from the energy density of the homogeneous electron gas.[133] The form this constant takes depends on a number of factors, including the way electron spins are treated. There exist a variety of ways to fit the correlation part of the energy functional E_{c}, although most rely on parametrisation from QMC data. The most elegant form is from relatively recent work by Chachiyo *et al* (2016) and is given by

$$\varepsilon_{\text{c}}^{\text{LDA}}(\rho(\mathbf{r})) = a \ln\left(1 + \frac{b}{r_s} + \frac{b}{r_s^2}\right) \tag{2.24}$$

where a and b are constants.[134] The parameter r_s is a dimensionless quantity that is related to the density as $r_s^3(\mathbf{r}) = 3/4\pi\rho(\mathbf{r})$. In this simple case, the constants can either be fitted to QMC data or by using various constraints. Other notable approaches for finding $\varepsilon_c^{\text{LDA}}$ include the Ceperley-Adler, Perdrew-Wang and Vosko-Wilk-Nusair function fits, however all fits behave in a similar fashion to the fit provided above.[135–137]

For meaningful calculations to take place, the potential $v_{\text{xc}}^{\text{LDA}}$ and total energy $E_{\text{tot}}^{\text{LDA}}$ must also be calculated. Using the definitions given above, the xc potential can be

found as

$$v_{\mathrm{xc}}^{\mathrm{LDA}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}^{\mathrm{LDA}}}{\delta \rho(\mathbf{r})} = \varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho(\mathbf{r})) + \rho(\mathbf{r}) \frac{\partial \varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho(\mathbf{r}))}{\partial \rho(\mathbf{r})}, \tag{2.25}$$

and the total LDA energy is given by

$$E_{\mathrm{tot}}^{\mathrm{LDA}} = \sum_{i}^{\mathrm{occ}} \varepsilon_i - E_{\mathrm{H}} - \overbrace{\int \rho(\mathbf{r}) v_{\mathrm{xc}}^{\mathrm{LDA}}(\mathbf{r}) d\mathbf{r}}^{\text{Incorrect xc energy}} + \underbrace{\int \varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}\left(\rho(\mathbf{r})\right) \rho(\mathbf{r}) d\mathbf{r}}_{\text{Correct xc energy}}. \tag{2.26}$$

In the above (total energy) equation, the LDA xc potential includes a variationally incorrect contribution. This quantity, contained within the eigenvalue sum, is removed and replaced by the variationally correct form, as seen in Eqn. 2.26.

As shown in this section, the quantity $\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho(\mathbf{r}))$ is merely a fitted function of the density. With it being known, all quantities in the KS equations (Eqn. 2.21) are also known, and only $v_{\mathrm{xc}}(\mathbf{r})$ is approximated. In practical calculations, the LDA works by calculating potentials (and energies) point-wise, i.e. the LDA xc potential at point $\mathbf{r}_i$ with density $\rho(\mathbf{r}_i)$ is given by looking what value $\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}\left(\rho(\mathbf{r}_i)\right)$ should be at such a density. This is illustrated more clearly in Fig. 2.2. Indeed, the schematic highlights why the LDA carries the name *local* density approximation, as it only assesses the value of $\varepsilon_{\mathrm{xc}}$ by observing the density at one point in space. The *real* many body system is not quite so simple, as density contributions from other regions in space may also affect the xc potential.

Generalised gradient approximation

The GGA takes the LDA a step further; by considering not only the value of the density, but also the gradient of the density at a point in space, a more sophisticated approximation to xc can be found. This can be thought of as similar to a Taylor expansion to a function, with the first order contribution being only the density, and second order being the first derivative of the density. Higher order gradients can also be included (so-called meta-GGAs), potentially leading to more accurate $\varepsilon_{\mathrm{xc}}$ approximations, however currently these are known to have issues with stability.[138] Nonetheless, the form of GGA functionals takes the form

$$E_{\mathrm{xc}}[\rho] = \int \rho\left(\mathbf{r}\right) \varepsilon_{\mathrm{xc}}(\rho, \nabla\rho) d\mathbf{r}. \tag{2.27}$$

The most popular GGA functionals include the Perdew-Wang 1991 (PW91) and Perdew-Burke-Ernzerhof (PBE) parametrisations.[136,139]

Spin density approximations

For systems with spin, extending the LDA and GGA approaches merely separates the densities of both spin channels (α and β). For example, the LDA and GGA would take the form

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho_\alpha, \rho_\beta] = \int \rho\left(\mathbf{r}\right) \varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho_\alpha, \rho_\beta) d\mathbf{r} \tag{2.28a}$$

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[\rho_\alpha, \rho_\beta] = \int \rho\left(\mathbf{r}\right) \varepsilon_{\mathrm{xc}}^{\mathrm{GGA}}(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta) d\mathbf{r} \tag{2.28b}$$

where the quantities $\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}$ and $\varepsilon_{\mathrm{xc}}^{\mathrm{GGA}}$, as described for the spinless LDA case, are merely functions of the density. The densities of α and β spin channels would be varied independently, potentially leading to different spatial descriptions of the two densities (ρ_α and ρ_β). Consider the simple example of two well-separated hydrogen atoms and two electrons, with one electron in each spin channel. The ground state of such a system should be one full electron with each proton. If the density of both spin channels is not varied independently, it would not at all be possible to obtain the correct solution. The spin restricted case (also called the closed-shell solution) would better apply to a helium atom, where the spatial distribution of the two electrons is the same.

2.4.4 Meaning of orbital eigenvalues

In principle, solving the KS equations exactly (once only) can easily yield three quantities which are identical to the many-electron ones: the ground state density $\rho_0(\mathbf{r})$, energy E_0 and ionisation potential $I(N)$. The third quantity, $I(N)$, corresponds to the energy released by removing the N^{th} lowest energy electron from the system. In exact DFT, this is accessed from the HOMO $\varepsilon_{\mathrm{HOMO}}$.[140] This quantity arises somewhat by coincidence; the ionisation potential is dependent on the decay of the density, which exact DFT should replicate, towards the edge of a system.[25]

The unoccupied orbitals in DFT do not carry any information on excited states. As DFT is primarily a ground-state theory, it need not obtain excited state properties to model the ground state (in theory). The subsequent subsection discusses this further. Moreover, apart from the HOMO, all other occupied orbitals, in theory, also have no physical meaning. However, empirical evidence shows that there is large agreement

between occupied orbital band structures and those obtained from experiment.

2.4.5 Derivative discontinuity

The (eigenvalue) energy gap between occupied and unoccupied orbitals *should not* correspond to the experimental band gap in DFT. Exact DFT requires an energy contribution added to unoccupied states, called the derivative discontinuity term Δ_{xc}, that restores correct energy gaps.[25,140] This constant, discussed further in Chapter 4, arises from a constant shift in the xc potential around integer numbers of electrons,

$$v_{xc}^{N-\delta}(\mathbf{r}) + \Delta_{xc} = v_{xc}^{N+\delta}(\mathbf{r}), \tag{2.29}$$

where δ is an infinitesimal charge added to and removed from an integer number of electrons N.[140] In other words, the entire xc potential is shifted by a constant Δ_{xc} after adding a small amount of charge. This concept may seem slightly perplexing, as a constant shift in the potential would reproduce the same density. The key change occurs in the orbitals; a constant shift in the potential causes the orbitals to have different energies, which reproduce the same energy and density when inserted into the KS equations. In other words, unoccupied orbitals are subject to a constant shift in their energies when they become occupied. This can be written as [2]

$$\varepsilon_{N+1}(N-\delta) + \Delta_{xc} = \varepsilon_{N+1}(N+\delta). \tag{2.30}$$

In turn, this means that the energy gap between occupied and unoccupied orbitals E_g in exact functionals is actually given by[140]

$$E_g = [\varepsilon_{N+1}(N) + \Delta_{xc}] - \varepsilon_N(N). \tag{2.31}$$

While not being available from the eigenvalues, the quantity E_g can instead be obtained by performing the calculation

$$E_g = I(N) - A(N) \tag{2.32}$$

where $A(N)$ is the electron affinity of an N electron system. Consequently, the energy E_g can only be accessed via two separate sets of orbitals in the exact DFT case [3].

[2] $\varepsilon_N(M)$ is the N^{th} lowest energy orbital from an M electron system.
[3] The ionisation potential and electron affinity are defined as $-I(N) \equiv E(N) - E(N-1)$ and

An unfortunate consequence of Δ_{xc} means that the correct form of the xc potential must include a discontinuity. The implication is that the exact functional is not smooth with respect to the density and is therefore non-analytic:

$$\Delta_{xc} = v_{xc}^{N+\delta}(\mathbf{r}) - v_{xc}^{N-\delta}(\mathbf{r}) = \left.\frac{\delta E_{xc}}{\delta\rho(\mathbf{r})}\right|_{N+\delta} - \left.\frac{\delta E_{xc}}{\delta\rho(\mathbf{r})}\right|_{N-\delta}. \tag{2.33}$$

The LDA and PBE functionals are created using smooth functions of the density, meaning they do not have this feature.[140]

2.5 Hybrid functionals

As discussed in Chapter 1, both DFT and HF calculations tend to be systematically erroneous, but in opposing ways. Hybrid functionals exploit this by mixing these two approaches together. While this may initially seem haphazard from a pragmatic point of view, the approach can, for some systems, yield substantially improved results when comparing to experimental data. Hybrid functionals exist within a framework known as GKS DFT. In this, the xc potential is allowed to be both local *or* non-local. Use of hybrid functionals does come at a cost, as calculating HF exchange is an expensive procedure

2.5.1 Linearly-mixed functionals

Some commonly used hybrid approximations to xc are shown below. The simplest is given by

$$E_{xc}^{\mathrm{PBE}\alpha}(\alpha) = \alpha E_x^{\mathrm{HF}} + (1-\alpha)E_x^{\mathrm{PBE}} + E_c^{\mathrm{PBE}}, \tag{2.34}$$

which simply mixes PBE and HF exchange, with α being the fraction of HF exchange. The PBE0 functional is defined as the case when $\alpha = 1/4$.[141] For molecular (and aperiodic) systems B3LYP,

$$E_{xc}^{\mathrm{B3LYP}}(a_0, a_x, a_c) = E_x^{\mathrm{LDA}} + a_x\left(E_x^{\mathrm{GGA}} - E_x^{\mathrm{LDA}}\right) + a_0\left(E_x^{\mathrm{HF}} - E_x^{\mathrm{LDA}}\right)$$

$-A(N) \equiv E(N+1) - E(N)$, where $E(N)$ is the energy of an N electron system. While $I(N)$ needs both $E(N)$ and $E(N-1)$, in the exact DFT case it can instead be evaluated from the HOMO of an N electron system, $\varepsilon_N(N) = -I(N)$. The second quantity $A(N)$ must be obtained from $E(N+1) - E(N)$, which requires an additional calculation for the energy $E(N+1)$. Alternatively, the orbital $\varepsilon_{N+1}(N+1) = -I(N+1) \equiv A(N)$ can be used, although this also requires an additional calculation. Three sets of orbitals are required for inexact or non-Koopmans-compliant functionals as $\varepsilon_N(N) \neq -I(N)$.

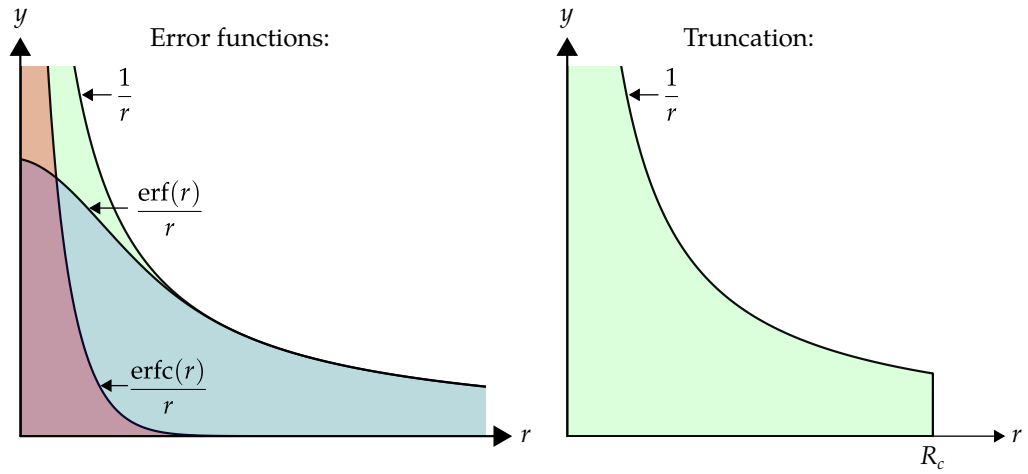

Figure 2.3: A schematic illustrating how range-separation can be introduced. For the error functions case, it can be seen that the terms $1/r \approx \mathrm{erf}(r)/r$ and $\mathrm{erfc}(r)/r \approx 0$ for large r. For the truncation case, large values of the truncation radius mean $1/R_c \approx 0$.

$$+E_{\mathrm{c}}^{\mathrm{LDA}} + a_{\mathrm{c}}\left(E_{\mathrm{c}}^{\mathrm{GGA}} - E_{\mathrm{c}}^{\mathrm{LDA}}\right), \tag{2.35}$$

is commonly used, where a_{x}, a_{c} and a_0 are fixed constants.[24] Its performance for condensed phase systems varies; it usually performs poorly for systems with small and medium energy gaps E_g, but better with large gap systems.

2.5.2 Range-separated functionals

So far, the only electron-electron interaction $u(\mathbf{r}_i, \mathbf{r}_j)$ discussed has been of the form

$$u(\mathbf{r}_i, \mathbf{r}_j) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} = \frac{1}{r}. \tag{2.36}$$

In order to reduce cost associated with HF exchange, it is useful to determine its short- and long-range components. The short-range components, when r is small contribute more to the exchange energy than the long-range components, when r is large, due to the inverse dependence on r (see exchange energy in Eqn. 2.14). In this section two approaches are detailed that alter the interaction term u to induce range separation. The first approach uses the error and complimentary error functions ($\mathrm{erf}(x)$ and $\mathrm{erfc}(x)$) to induce this. The second approach simply truncates the long-range HF component and adds a long-range correction term given by a local functional.

Error function interaction

A very successful hybrid functional for condensed phase systems by Heyd, Scuseria and Ernzerhof, called the HSE functional, works by changing the form of the Coulomb interaction.[142] The interaction u can be rewritten as

$$u(\mathbf{r}_i, \mathbf{r}_j) = \frac{1}{r} = \frac{\operatorname{erfc}(\omega \cdot r)}{r} + \frac{\operatorname{erf}(\omega \cdot r)}{r}. \tag{2.37}$$

Using the above, the HF exchange energy can be manipulated as [4].

$$E_{\mathrm{x}}^{\mathrm{HF}} = -\frac{1}{2}\sum_{i,j}\langle\psi_j\psi_i|\,\frac{1}{r}\,|\psi_i\psi_j\rangle \tag{2.38}$$

$$E_{\mathrm{x}}^{\mathrm{HF}} = \underbrace{-\frac{1}{2}\sum_{i,j}\langle\psi_j\psi_i|\,\frac{\operatorname{erfc}(\omega \cdot r)}{r}\,|\psi_i\psi_j\rangle}_{\text{Short-range exchange, } E_{\mathrm{x}}^{\mathrm{HF,\,SR}}} \underbrace{-\frac{1}{2}\sum_{i,j}\langle\psi_j\psi_i|\,\frac{\operatorname{erf}(\omega \cdot r)}{r}\,|\psi_i\psi_j\rangle}_{\text{Long-range exchange, } E_{\mathrm{x}}^{\mathrm{HF,\,LR}}}. \tag{2.39}$$

This separates long and short range components of HF exchange by introducing a range-separation parameter ω. The (semi-local) PBE functional can also be separated into similar terms. If the exchange parts of the PBE0 functional are expressed in terms of short- and long-range components, the following expression can be found.

$$E_{\mathrm{x}}^{\mathrm{HSE}}(\alpha, \omega) = \alpha E_{\mathrm{x}}^{\mathrm{HF,\,SR}} + \alpha E_{\mathrm{x}}^{\mathrm{HF,\,LR}} + (1-\alpha)E_{\mathrm{x}}^{\mathrm{PBE,\,SR}} + (1-\alpha)E_{\mathrm{x}}^{\mathrm{PBE,\,LR}} \tag{2.40}$$

The long-range PBE0 and HF components are very small, and tend to cancel each other out when large (i.e. $\alpha E_{\mathrm{x}}^{\mathrm{HF,\,LR}} \approx \alpha E_{\mathrm{x}}^{\mathrm{PBE,\,LR}}$) for sensible ω values. The resulting expression

$$E_{\mathrm{x}}^{\mathrm{HSE}}(\alpha, \omega) = \alpha E_{\mathrm{x}}^{\mathrm{HF,\,SR}} + (1-\alpha)E_{\mathrm{x}}^{\mathrm{PBE,\,SR}} + E_{\mathrm{x}}^{\mathrm{PBE,\,LR}} \tag{2.41}$$

essentially reduces the problem to only calculating short-range HF exchange. This simple re-expression reduces computational cost significantly; as r grows large, the short-range component of the Coulomb interaction essentially falls to zero, as seen in Fig. 2.3. In other words, when $E_{\mathrm{x}}^{\mathrm{HF,\,LR}} \approx E_{\mathrm{x}}^{\mathrm{HF}}$, which usually occurs at the characteristic length $2/\omega$, HF exchange need not be calculated at all as it cancels with $E_{\mathrm{x}}^{\mathrm{PBE,\,LR}}$.[143] The HSE06 functional is defined as when $\alpha = 1/4$ and $\omega = 0.2$ Å^{-1}.

[4] The $i = j$ summation term in the Hartree and HF exchange energies (Eqn. 2.14) is included in the context of hybrid functionals.

Truncated Coulomb interaction

An alternative hybrid functional, called the Coulomb-truncated PBEα (tr-PBEα), has a simpler means of inducing range-separation. The form of the Coulomb operator used in this case is given by

$$u_{\mathrm{tr}}(\mathbf{r}_i, \mathbf{r}_j) = \begin{cases} \frac{1}{r}, & \text{for } r \leq R_c \\ 0, & \text{for } r > R_c \end{cases} \tag{2.42}$$

where R_c is some specified cut-off radius. This truncation is illustrated in Fig. 2.3. In other words, the exchange energy for sufficiently large r values (i.e. the long range term) is not calculated. For appropriately chosen values of R_c, the long-range terms contribute little to the energy, meaning this should not be a problem. While the long-range effects of exchange have been omitted, this functional attempts to capture them using a long-range correction term via a local functional. This correction is calculated from the PBE functional and is termed $E^{\mathrm{PBE,\,LRC}}$. The form of this correction is compared with that of HSE in the subsequent section (Section 2.5.2), however a detailed discussion can be found in reference 144. The final exchange energy is given by,

$$E_{\mathrm{x}}^{\text{tr-PBE}\alpha}(\alpha, R_c) = \alpha E_{\mathrm{x}}^{\text{tr-HF}} + \alpha E_{\mathrm{x}}^{\mathrm{PBE,\,LRC}} + (1-\alpha)\, E_{\mathrm{x}}^{\mathrm{PBE}} \tag{2.43}$$

where

$$E_{\mathrm{x}}^{\text{tr-HF}} = -\frac{1}{2}\sum_{i,j} \langle \psi_j \psi_i |\, u_{\mathrm{tr}}\, | \psi_i \psi_j \rangle \,. \tag{2.44}$$

This functional is also called PBE0-TC-LRC and PBE0-TC, with and without the long-range correction.

Comparison of range-separation

PBE functionals incorporate gradient approximations into a functional by

$$\varepsilon_{\mathrm{x}}(\rho, \nabla\rho) = \varepsilon_{\mathrm{x}}^{\mathrm{LDA}}(\rho) \cdot F_{\mathrm{x}}^{\mathrm{PBE}}(\rho, \nabla\rho) \tag{2.45}$$

where $F_{\mathrm{x}}^{\mathrm{PBE}}(\rho, \nabla\rho)$ is known as the enhancement factor. This is usually expressed in terms of a fitted function $J_{\mathrm{x}}^{\mathrm{PBE}}(s, r)$ known as the spherically averaged exchange hole,

$$F_{\mathrm{x}}^{\mathrm{PBE}}(\rho, \nabla\rho) = -\frac{8}{9}\int_0^\infty k_F J_{\mathrm{x}}^{\mathrm{PBE}}(s, r)dr, \tag{2.46}$$

where $k_F = (3\pi^2\rho)^{\frac{1}{3}}$ and $s = |\nabla\rho|/2k_F\rho$. The ranged PBE parts in the HSE and tr-PBEα functionals are calculated as

$$F_{\mathrm{x}}^{\mathrm{PBE,\,SR}} = -\frac{8}{9}\int_0^\infty k_F J_{\mathrm{x}}^{\mathrm{PBE}}(s, r)\ \mathrm{erfc}\,(\omega \cdot r)\,dr \tag{2.47a}$$

$$F_{\mathrm{x}}^{\mathrm{PBE,\,LR}} = -\frac{8}{9}\int_0^\infty k_F J_{\mathrm{x}}^{\mathrm{PBE}}(s, r)\ \mathrm{erf}\,(\omega \cdot r)\,dr \tag{2.47b}$$

$$F_{\mathrm{x}}^{\mathrm{PBE,\,LRC}} = -\frac{8}{9}\int_{R_c}^\infty k_F J_{\mathrm{x}}^{\mathrm{PBE}}(s, r)dr. \tag{2.47c}$$

In other words, the HSE functional has a smooth PBE energy contribution whilst tr-PBEα does not.[144] As detailed in previous subsections, the same can be said of exchange energy contributions for these functionals.

2.6 Solid state calculations

At the start of this chapter, it was mentioned that the systems we aim to describe require a few hundred nuclei and a few thousand electrons. This begs the question: given real crystals contain $\sim 10^{23}$ atoms, how can one describe them using such a small system? A simulation containing this number of atoms is seemingly impossible. To this end, many simulations of solid state matter make use of periodic boundary conditions. While both naturally occurring and synthetic crystals usually contain defects, very pure crystals often have, at least in some small region, a repeating structure of atoms. In that region, particular quantities surrounding the electronic structure are also repeating. For example, the density follows the rule $\rho(\mathbf{r}) = \rho(\mathbf{r} + \mathbf{R})$, where $\mathbf{R}$ is a vector that is commensurate with the periodicity of the crystal.[11]

The periodic symmetry of a crystal is not necessarily mirrored by the symmetry of the single-particle orbitals, meaning $\psi(\mathbf{r})$ need not equal $\psi(\mathbf{r} + \mathbf{R})$. In general the symmetry of a system does not necessarily reflect in its solution(s) to the underlying differential equation. One example would be an empty (and boundless) box with a constant potential and one electron. In such a box, the potential is constant with respect to all translations, whereas solutions are of planewave form.

2.6.1 Periodic systems

Consider a 3D system with a periodic potential and density. The smallest repeating unit of such a crystal is called a unit cell and (can in general) span across three directions, given by $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$. These vectors, called primitive lattice vectors, have a magnitude that equals the length of each repeating unit. If a point $\mathbf{r}$ in one unit cell is translated by $\mathbf{R}$,

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \tag{2.48}$$

where n_1, n_2 and n_3 are integers, an equivalent point would be found in another repetition of the unit cell. Using this, an additional basis which corresponds to the atoms within each primitive cell can also defined. For example, atoms A_1 and A_2 could be given position co-ordinates $\mathbf{R}_1$ and $\mathbf{R}_2$ as

$$\mathbf{R}_1 = (1/2\,|\mathbf{a_1}|,\ 1/2\,|\mathbf{a_2}|,\ 1/4\,|\mathbf{a_3}|) \tag{2.49}$$

$$\mathbf{R}_2 = (1/2\,|\mathbf{a_1}|,\ 1/2\,|\mathbf{a_2}|,\ 3/4\,|\mathbf{a_3}|)\,. \tag{2.50}$$

With the above, our atomic nuclei have positions in real space that are well-defined, as is their periodicity via the lattice vectors. Another set of vectors whose relevance will be discussed below, called reciprocal lattice vectors, can also be defined:

$$\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \tag{2.51a}$$

$$\mathbf{b}_2 = 2\pi\frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_2 \cdot (\mathbf{a}_3 \times \mathbf{a}_1)} \tag{2.51b}$$

$$\mathbf{b}_3 = 2\pi\frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2)}. \tag{2.51c}$$

2.6.2 Bloch's theorem

Bloch theorem states that the wavefunction of a periodic crystal contains is described by

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r}) \cdot e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.52a}$$

$$\psi_{n,\mathbf{k}}(\mathbf{r}+\mathbf{R}) = \psi_{n,\mathbf{k}}(\mathbf{r}) \cdot e^{i\mathbf{k}\cdot\mathbf{R}}, \tag{2.52b}$$

where $\mathbf{k}$ is described by the reciprocal lattice vectors $\mathbf{k} = (\theta_1\mathbf{b}_1,\ \theta_2\mathbf{b}_2,\ \theta_3\mathbf{b}_3)$. The term $u_{n,\mathbf{k}}(\mathbf{r})$ is a function that is periodic with the crystal lattice vectors $\mathbf{a}_i$. Conversely, the term describing the wavefunction $\psi_{n,\mathbf{k}}(\mathbf{r})$ is not necessarily periodic. Essentially, the wavefunction describing the crystal $\psi_{n,\mathbf{k}}(\mathbf{r})$ has a periodic component that is modulated by an exponential envelope function $e^{i\mathbf{k}\cdot\mathbf{r}}$ that has the periodicity of the lattice. This equation contains the band index n, which is analogous to the single-particle eigenvalue indices, and an additional index corresponding to the vector $\mathbf{k}$.

If Bloch's equation (Eqn. 2.52) is substituted into the KS equations (Eqn. 2.21), a differential equation describing a periodic crystal can be found,

$$\left[-\frac{1}{2}\left(\nabla + i\mathbf{k}\right)^2 + v_{\text{eff}}(\mathbf{r})\right] u_{n,\mathbf{k}}(\mathbf{r}) = \varepsilon_{n,\mathbf{k}} u_{n,\mathbf{k}}(\mathbf{r}), \tag{2.53}$$

where $u_{n,\mathbf{k}}$ and $\psi_{n,\mathbf{k}}$ are both eigenvectors of the same Hamiltonian (at a given $\mathbf{k}$ value).[11] As the eigenfunction $u_{n,\mathbf{k}}$ is now periodic, solving these equations inside one unit cell would be the equivalent of solving the KS equations for an infinitely repeating crystal. It can be seen that the vector $\mathbf{k}$ needs to be specified in order to solve these equations. At a particular value of $\mathbf{k}$, called the Γ-point [$\mathbf{\Gamma} = (0,0,0)$] (i.e. the origin of the reciprocal lattice), the equation takes a similar form to the normal KS equations. In this case, solutions to the equation can be taken to be real, reducing computational cost.

2.6.3 k-point convergence

The reciprocal lattice also has its own version of the (primitive) unit cell, called the Brillouin zone. In addition to being periodic in real space, the function $u_{n,\mathbf{k}}(\mathbf{r})$ is also periodic with the reciprocal lattice vectors. This means that any $\mathbf{k}$ vector outside the Brillouin zone can be mapped to a symmetrically equivalent one within the Brillouin zone.

Many quantities in a periodic (DFT) calculation require an integral over the Brillouin zone. The density is one such example and is described by

$$\rho(\mathbf{r}) = \frac{\Omega}{(2\pi)^3}\int_{\text{BZ}} \rho_{\mathbf{k}}(\mathbf{r}) d\mathbf{k} \tag{2.54a}$$

$$\rho_{\mathbf{k}}(\mathbf{r}) = \sum_{i}^{\text{occ}} \psi^{\star}_{n,\mathbf{k}}(\mathbf{r})\psi_{n,\mathbf{k}}(\mathbf{r}), \tag{2.54b}$$

where Ω is the volume of the unit cell. Similar integrals exist for other variables,

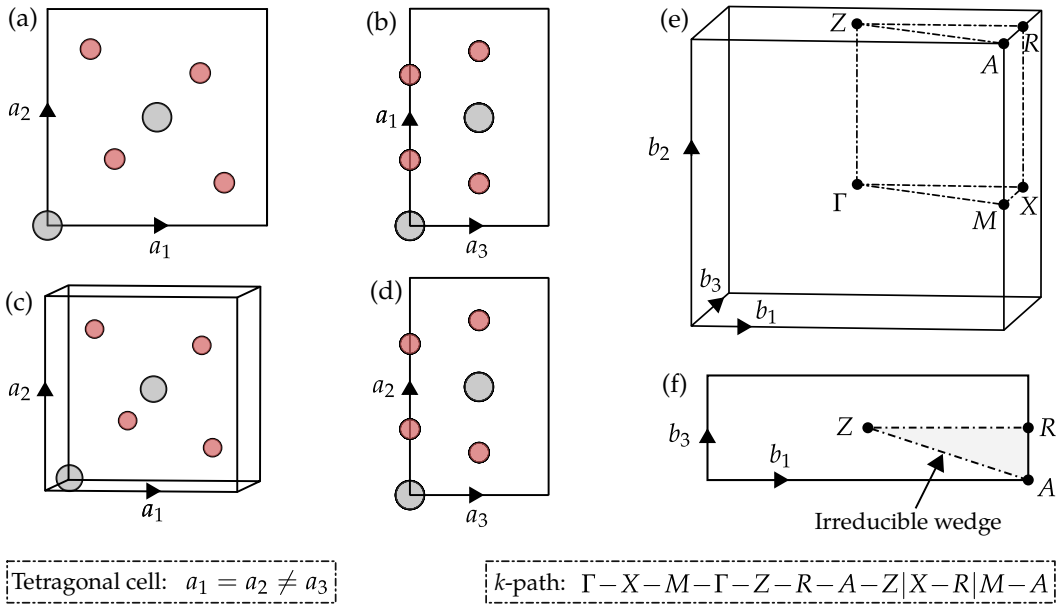

Figure 2.4: (a)-(d): A schematic illustrating the (tetragonal) TiO_2 rutile primitive cell. (e)-(f): The corresponding Brillouin zone is shown, with high symmetry points marked. As with the real space cell, the Brillouin zone contains many points which are symmetrically equivalent. An irreducible wedge can be constructed where all non-equivalent k-points are present. Band structure calculations for a tetragonal cell follow the high-symmetry k-point path indicated.

such as the energy. Luckily, electronic wavefunctions in a small enough region over **k**-space (i.e. the space of **k** vectors) are very similar, meaning **k**-space can be discretised. Convergence with number of **k**-point is therefore required for these quantities to be meaningful. Computational cost can be reduced by making use of symmetry within the Brillouin zone. As shown in Fig. 2.4 for a tetragonal cell, this means the volume of **k**-space that needs to be explored is reduced to a much smaller region, called an irreducible wedge.

The magnitude of reciprocal lattice vectors takes the form $2\pi/|\mathbf{a}_i|$ where $\mathbf{a}_i$ is a real-space lattice vector. As the size of a real-space cell increases, the Brillouin zone shrinks, meaning each **k**-point sampled can be mapped on more reciprocal lattice sites. In other words, if the Brillouin zone is sufficiently small, the integrals above (Eqn. 2.54a and Eqn. 2.54b) can be performed at only one value of **k**. This means that, rather than sampling more **k**-points, an alternative approach to converging the above integrals is by using a large enough real-space cell. In practise, this would mean duplicating a primitive cell in all directions creating a 'supercell'.

2.6.4 Optimising atomic configurations

Up until this section, atomic positions have been fixed and included into the electronic Hamiltonian via an external potential. In practise, it is not always known exactly where atomic positions lie in a crystal. It is therefore necessary to have a means of optimising these positions. A variety of algorithms exist to optimise atomic positions, however they all rely on calculation of forces from electronic structure calculations. The forces on each ion I at position $\mathbf{R}_I$ are calculated using the Hellman-Feynman theorem,

$$\mathbf{F}_I = -\frac{\partial E}{\partial \mathbf{R}_I} = -\langle \Psi_0 | \frac{\partial \hat{\mathcal{H}}}{\partial \mathbf{R}_I} | \Psi_0 \rangle \tag{2.55}$$

which essentially treats nuclei as classical particles. In order for the above equation to hold true, various corrections need to be considered. These include corrections for spurious forces that arise from use of finite basis sets, called Pulay forces. As shown in Fig. 2.5, the ground state orbitals $\{\psi_i\}$ are often found using density mixing schemes. As this is a not a variational approach, further corrections may also be considered to account for this.

While the forces on atoms in a primitive cell can be reduced to equal zero, the primitive cell itself may also be incorrect. The cell can be relaxed calculating the stress tensor $\sigma_{\alpha\beta}$,

$$\sigma_{\alpha\beta} = \frac{1}{\Omega}\frac{\partial E}{\partial \varepsilon_{\alpha\beta}}, \tag{2.56}$$

where $\varepsilon_{\alpha\beta}$ is an applied strain to the crystal. The components of the strain tensor correspond to strains in different directions. Once the stress tensor is minimised, the cell can be assumed to be in its equilibrium shape. By employing symmetry conditions, the number of tensor components that need to be calculated can be reduced.

2.7 Solving single-particle equations

In previous sections, procedures for finding the ground state electronic and atomic configurations have been discussed. Here, a discussion is presented on how typical electronic structure codes use all the above to find ground state configurations. Fig. 2.5 shows, how an initial guess for the ground state can be used to find the actual ground state. For the electronic convergence, the relevant single-particle equations are ap-

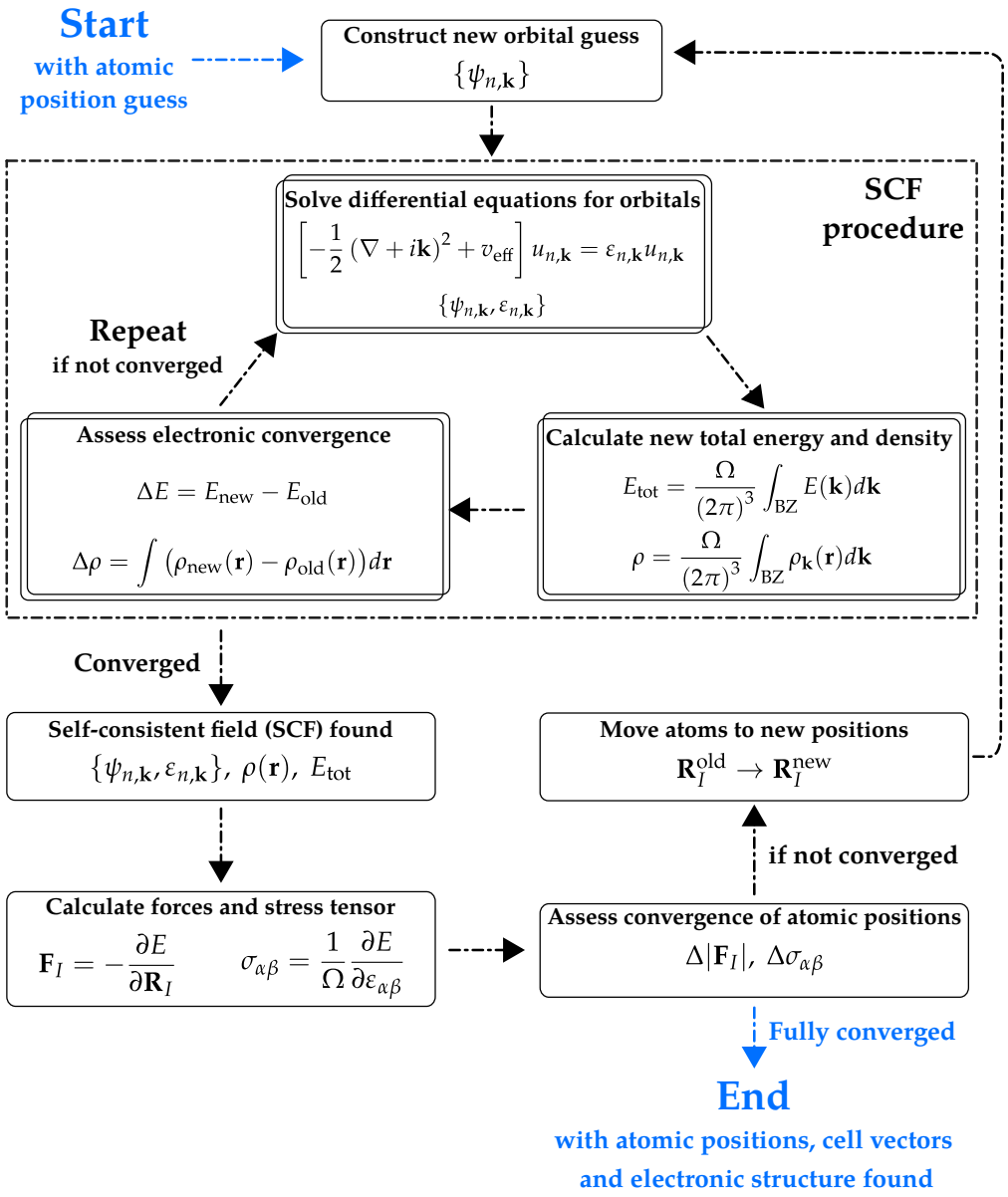

Figure 2.5: A schematic illustrating the procedure that happens when attempting to find the ground state electronic and atomic configurations. Starting from cell vector, atomic position and orbital guesses, $\{\mathbf{a}_i\}$, $\{\mathbf{R}_I\}$ and $\{\psi_{n,\mathbf{k}}\}$, this procedure is applied until convergence is reached with respect to energy, density, forces and stresses. In cases where atomic forces and/or cell vectors do not need to be optimised, only the SCF procedure is applied iteratively.

plied iteratively until the input and output orbitals do not change. The criteria for convergence is usually defined as when either the energy ΔE or density $\Delta\rho$ changes are below some specified cutoff. Once the electronic ground state has been reached, a self-consistent field (SCF) is found, and the atomic forces and stresses are assessed. If these are not converged, atomic positions are improved via the forces and lattice vectors via the stress tensor. The SCF procedure starts again with the new positions and the process entire process starts again. This process repeats until convergence is

reached (i.e. the input and output orbitals do not change by more than a specified measure, as seen in Fig. 2.5).

For the electronic part, there are numerous approaches to changing the density (hence orbitals) between each iteration, many of which use a history of the density from previous iterations. These approaches are dubbed 'mixing schemes' and essentially attempt to use an educated guess for what the ground state electronic configuration should be. Notable examples include Broyden and Pulay mixing. For simple systems, a pure mix of the previous and current densities would suffice for convergence, however this simple approach often leads to slower convergence. Usually the performance of these algorithms depends on the system under investigation, meaning the user must optimise their choice by testing.

Numerous algorithms exist to aid atomic configuration convergence as well. Many of these only converge on local minima, meaning the global minimisation must be done by other means, of which a variety exist. When a structure is near its local minima, approaches such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm or the conjugate gradients algorithm can be employed.[145] As with the electronic convergence, some algorithms may be more appropriate than others depending on the system under study.

2.8 Calculating defect properties

A number of defects types are considered in this thesis, such as purely electronic defects (e.g. polarons) and intrinsic point defects (e.g. vacancies and interstitials). As many bulk crystals are inherently stable (i.e. energy is released on their formation), forming defects in a perfect bulk crystal will usually carry an energy cost, called the defect formation energy E_f. This cost will vary depending on chemical conditions; forming excess oxygen defects in oxygen rich conditions may require less energy than forming an oxygen vacancy in the same conditions. Excess electrons or holes may bind to these defects, negating any p- or n-type doping. The energy released by binding excess charges may result in a more stable configuration. This means that different charge states q must also be explored for each possible defect. For each charge state a number of configurations may be found, however the most stable one is usually used to calculate the defect formation energy E_f at a given Fermi energy E_F. The quantity

E_f for a defect X in charge state q is given by

$$E_f[X^q] = E[X^q] - E[H^{q=0}] - \sum_i (n_i - m_i)\mu_i + q(\varepsilon_v + E_F). \tag{2.57}$$

$E[H^{q=0}]$ denotes the total energy for the defect-free host crystal (i.e. bulk) and μ_i the chemical potential for adding or removing a particular atomic species. [5] The quantities n_i and m_i are the number of times an atomic species is added and removed. ε_v denotes the valence band minimum and E_F the Fermi energy (relative to VBM). While an actual calculation will have the Fermi energy fixed, the above definition of E_f can be considered as varying the energy of a reservoir of electrons that interacts with electrons in the crystal. When E_F is varied, E_f can be thought of as the ease in which a defect will form when the crystal is $n-$ or $p-$type doped. Changes in μ_i reflect the ease in which a defect will form under particular chemical conditions. For example, an oxygen-rich environment will reduce the cost of an oxygen interstitial defect as there is an abundance of oxygen.

2.9 Discussion

Despite the large success of the single-particle approaches introduced above, there is not one universal approach that can be used for all types of systems. This may perhaps lead one to think that more sophisticated approximations must be employed, however their use also currently has severe limitations. In the following sections, the limitations of both single-particle and many-body approaches are briefly discussed.

2.9.1 Limitations of Hartree-Fock

As briefly discussed in Section 2.2, HF is an extremely poor choice of method to investigate materials that do not have a band gap. Specifically, this theory predicts many, if not all, metals to be insulators; even materials with a band gap (i.e. semiconductors and insulators) are predicted to have much larger gaps than experimental data suggests. This failure stems from the tendency of HF to over-localise electrons, a topic discussed in Chapter 1. An over-localised electron is also a very energetically bound one. By predicting this behaviour, electrons cannot delocalise as they do in metals. As the valence band states (i.e. the occupied ones) are more energetically bound, they

[5] The total energies in 2.57 are assumed to be free from finite size errors. Finite size errors, and the relevant corrections, discussed in more detail in Chapter 6.

lie much lower in energy than they should be. In turn, a gap is induced between the occupied and unoccupied states. Thus, by virtue of electron over-localisation, this theory predicts materials to have much larger band gaps than they actually do. For the same reason, the theory also fails to predict correct ionization energies for materials.

2.9.2 Limitations of Kohn-Sham density functional theory

In contrast to HF, KS DFT predicts metallic behaviour very well. However, this is not construction, but by a convenient cancellation of errors. As mentioned in Chapter 1, DFT has a tendency to under-localise electrons by virtue of the self-interaction error (SIE). As electrons in metals are naturally more delocalised, and DFT tends toward delocalisation anyway, the correct behaviour is found. In fact, DFT tends to reduce gaps in semiconductors, often predicting them to be metallic. This error is in direct contrast with HF. The approach therefore tends to predict much smaller band gaps due to electrons being weakly bound.

2.9.3 Limitations of hybrid functionals

The remaining chapters of this thesis discuss and address many of the limitations associated with hybrids, of which there are mainly three: computational cost, introduction of free parameters and accuracy. The computational cost limitation is addressed in Chapter 3, where various techniques are employed to reduce this. These techniques reduce cost to roughly two to three times the cost of (semi-)local DFT calculations. Regarding free parameters, hybrids also introduce at least one additional free parameter, the fraction of HF exchange α. Other functionals, such as HSE and tr-PBEα (discussed in Section 2.5), have an additional range-separation parameter that also needs to be tuned. The GKT (discussed in Chapter 1), introduces a means to remove the degrees of freedom associated with hybrid functionals. In Chapter 4 and Chapter 5 the condition is examined and used in model systems and semiconductors. The final concern, accuracy, is addressed in Chapter 4 by testing hybrid functionals on model systems where exact solutions can be found.

2.9.4 Beyond the single-particle picture

Employing higher level approximations to the time-independent Schrödinger equation may seem like a sensible choice given some of the limitations discussed above, however these carry too large a computational cost. Use of the so-called 'post-HF'

techniques, such as Møller-Plesset perturbation theory or configuration interaction, usually increases computational costs due to poor scaling with number of electrons. Other approaches, such as QMC, have even worse scaling. Consequently, obtaining convergence with respect to atomic positions is not feasible, as each of the electronic steps is far too costly. To this end, many employ these techniques on fixed geometries obtained from DFT, which can be a reasonable approach.

The electronic structure of a defective crystal is often affected by the concentration of its defects. The chosen size of a supercell determines defect concentration. For example, one vacancy per primitive cell corresponds to a much larger defect concentration than one vacancy in a $2 \times 2 \times 2$ supercell. Ideally, the chosen supercell would reflect defect concentrations seen in experimental conditions. In DFT simulations, defect concentrations tend to be larger than those in experiments due to computational costs. While defect concentrations can be made to match experimental conditions (at a larger computational cost), the larger cells tend to be not feasible for post-HF methods. Subsequently, post-HF methods are usually only really applied to high defect concentrations, meaning their description of defective crystals is often unphysical.

2.10 Summary

The relevant underlying approaches used in many electronic structure simulations of materials have been summarised in this chapter. In Section 2.2, the HF approximation is discussed, followed by KS DFT in Section 2.4 then a marriage of these two approaches, specifically hybrid functionals, in Section 2.5. The approaches used to enable periodic simulations, such as Bloch's theorem, are discussed in Section 2.6. Bloch's theorem can be used in conjunction with one of the single-particle approximations outlined in Section 2.2, Section 2.4 or Section 2.5, such as hybrid DFT, to model periodic systems. Finally, the approaches were critiqued in the previous section.

CHAPTER 3

IMPLEMENTATION OF THEORY

In Chapter 2 a set of commonly used approaches for simulating materials were outlined, all of which require a differential equation of some form to be solved. As none of these differential equations have general analytic solutions, they must be solved numerically. The performance of calculations, measured in terms of both accuracy and computational cost, is heavily influenced by particular implementation choices. For example, plane wave basis sets are slower compared to atomic ones for calculations that involve Hartree-Fock as a larger number of basis functions is required; conversely, atomic basis sets are susceptible to inaccurate orbital descriptions via the basis set superposition error. The particular choices made therefore heavily depend on intended use. In this chapter, the key implementations discussed are those of `CP2K`, `VASP` and `iDEA`, where each code serves a slightly different purpose. The main aim of the `iDEA` code is to benchmark new electronic structure approximations against exact solutions in model 1D systems. In contrast, `CP2K` and `VASP` apply existing approximations to real 3D materials and molecules, which in turn can be directly compared to experiment. In the subsequent sections, the key implementation details of the above three codes are discussed.

3.1 General techniques

There are a few general techniques that are common to some of the electronic structure implementations explored in this chapter, such as basis sets and pseudopotentials. These are discussed first, then a more detailed discussion of each code's implementations is presented.

3.1.1 Basis sets

In general an orbital $\psi_i(\mathbf{r})$ with index i may be represented in terms of a set of basis functions $\phi_\nu(\mathbf{r})$,

$$\psi_i(\mathbf{r}) = \sum_\nu C_{i\nu}\,\phi_\nu(\mathbf{r}) \tag{3.1}$$

where $C_{i\nu}$ are their corresponding coefficients. In order for a calculation to complete in a reasonable amount of time, the number of functions considered must be truncated at some point. While this does mean that the basis set is effectively incomplete, care can be taken to ensure that the number of basis functions included sufficiently describes an orbital.

Plane waves are a popular choice of basis for periodic electronic structure codes. The periodic component of a Bloch wavefunction $u_{n,\mathbf{k}}(\mathbf{r})$ (Eqn. 2.52) can be expressed in terms of periodic plane waves,

$$u_{n,\mathbf{k}}(\mathbf{r}) = \frac{1}{\Omega^{1/2}} \sum_{|\mathbf{G}|<G_C} C_{n,\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} \tag{3.2}$$

where the $C_{n,\mathbf{k}}(\mathbf{G})$ are the basis coefficients and Ω is the volume of the cell. Each of the plane waves has the periodicity of the lattice (i.e. $\mathbf{G}$ is an integer multiple of the reciprocal lattice vectors). The series is truncated at a specified cutoff G_C, with a corresponding energy $|\mathbf{G}|^2/2$ being a convenient measure of basis set completeness. Plane waves can also be used to represent the density,

$$\rho(\mathbf{r}) = \frac{1}{\Omega} \sum_{|\mathbf{G}|<G_C} \rho(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}. \tag{3.3}$$

The Fourier transformed (total) real-space density $\rho(\mathbf{G})$ evaluated at a point $\mathbf{G}$ is the coefficient of its corresponding plane wave $e^{i\mathbf{G}\cdot\mathbf{r}}$.

3.1.2 Pseudopotentials

Changes in core electron levels of a particular atomic species between materials have little impact on electronic structure near the Fermi energy. For example, the difference in bonding between metallic titanium and an oxide of titanium has little to do with changes in the $1s$ electron levels of titanium. If these core levels were somehow eliminated (or frozen) then the chemical properties of a material would not be substantially different. In addition, plane wave basis sets require a large number of basis functions to sufficiently describe these core electrons, meaning computational cost can be greatly reduced if core levels were eliminated. To this end, many electronic structure codes use a pseudopotential or effective core potential to alleviate some of the computational cost associated with these electrons. The pseudopotential essentially replaces the bare atom with one that has fewer electrons. The resulting pseudopotential enforces the condition that its remaining electrons have the properties

$$\varepsilon_i' = \varepsilon_i \tag{3.4}$$

$$\psi_i'(\mathbf{r})\big|_{|\mathbf{r}|>r_c} = \psi_i(\mathbf{r})\big|_{|\mathbf{r}|>r_c}, \tag{3.5}$$

where the primed quantities represent results from using a pseudopotential. The value r_c defines a spherical region, inside of which an orbital is allowed to differ from its all-electron counterpart. Outside this, orbitals must equal the all-electron orbitals. Norm-conserving and ultra-soft pseudopotentials are two commonly used flavours in electronic structure codes.

3.2 The iDEA code

The `iDEA` code is used to benchmark approximations against exact solutions to the many-electron Schrödinger equation. The many-electron problem becomes difficult to solve for systems in 3D with large numbers of particles; the code avoids this problem by operating in 1D with a few electrons. Currently the `iDEA` code can only solve this for three electrons. The Schrödinger equation is solved using a real-space finite differences approach for electrons without spin. Electrons interact via the softened Coulomb interaction

$$u(x_i, x_j) = \frac{1}{|x_n - x_m| + k} \tag{3.6}$$

where $k = 1$. An interaction of this form avoids a singularity at zero separation. The parameter k is often set to allow 1D data to be mapped onto a real 3D problem, usually exploiting a symmetry that the 3D system has. Written in C and Python, the code currently runs using parallelised linear algebra libraries on a single node. On a single high-end computing node, the three electron problem can be solved exactly within 24 hours.

3.2.1 Finite differences approach

The code solves differential equations using finite differences on a real-space grid. In this approach, the position variable is discretised, as are all functions of the position variable [1].

$$x \to [x_1, x_2, \ldots, x_n] \tag{3.7}$$

$$f(x) \to [f_1, f_2, \ldots, f_n]. \tag{3.8}$$

The single-particle approaches discussed in the previous chapter require the calculation of a second derivative. This can be found for each f_n by using a Taylor expansion:

$$f_{n+1} = f_n + \frac{df_n}{dx}\Delta x + \frac{d^2 f_n}{dx^2}\frac{(\Delta x)^2}{2} + \mathcal{O}(x^3) \tag{3.9}$$

$$f_{n-1} = f_n - \frac{df_n}{dx}\Delta x + \frac{d^2 f_n}{dx^2}\frac{(\Delta x)^2}{2} + \mathcal{O}(x^3) \tag{3.10}$$

$$f_{n-1} + f_{n+1} \approx 2f_n + \frac{d^2 f_n}{dx^2}(\Delta x)^2 \tag{3.11}$$

$$\frac{d^2 f_n}{dx^2} \approx \frac{f_{n-1} - 2f_n + f_{n+1}}{(\Delta x)^2}. \tag{3.12}$$

Using this, the second derivative operator can be turned into a matrix to speed up the calculation. This is given by:

$$\left(\frac{d^2}{dx^2}\right)\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_n \end{bmatrix} \approx \left(\frac{1}{\Delta x^2}\begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & 1 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}\right)\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} f_1'' \\ f_2'' \\ f_3'' \\ \vdots \\ f_n'' \end{bmatrix}. \tag{3.13}$$

[1] The grid is uniform in `iDEA`, $x_{n+1} - x_n = \Delta x \; \forall \; n$. To simplify notation, $f_n = f(n\Delta x)$

The above operator is only applied to orbitals when kinetic energy components are required. The systems investigated in `iDEA` are contained within an infinite potential barrier. The boundary condition imposed in `iDEA` enforces that orbitals decay to zero at the boundary of the system. The above matrix holds even for the edge cases, f_1 and f_n, as f_{-1} and f_{n+1} (which correspond to points outside the system) both equal zero.

3.2.2 Exact many-electron solutions

Solutions to the exact time-independent Schrödinger equation can be used to form a basis. An arbitrary many-body wavefunction Ψ may therefore be expressed using a linear sum of these solutions. In order to ensure consistency with previous sections in this chapter, the basis comprising of many-body Schrödinger equation solutions will be denoted ϕ_n. The time evolution of this arbitrary wavefunction is given by

$$\Psi(x_1, x_2, \ldots, x_n, t) = \sum_n C_n \phi_n e^{-iE_n t}. \tag{3.14}$$

If instead the system is propagated through negative imaginary time, $t \to -i\tau$, the above equation becomes an exponential decay, with higher energy basis functions decaying faster than others. As the ground-state wavefunction has the lowest energy, it will be the slowest decaying one. This means that the coefficients C_n will slowly vary as

$$\lim_{\tau\to\infty} \Psi(x_1, x_2, \ldots, x_n, \tau) = \Psi_0(x_1, x_2, \ldots, x_n). \tag{3.15}$$

Normalisation of Ψ must also be ensured during this process. This discussion is a proof of principle that propagation of an arbitrary wavefunction through negative imaginary time yields the exact solution to the many-electron Schrödinger equation. Excited states can also be found using this approach, however only ground-state solutions will be discussed in this work.

The implementation of wavefunction propagation uses the Crank-Nicholson algorithm with the equation

$$\left(1 + \frac{\Delta\tau}{2}\mathcal{H}\right)\Psi(x_1, x_2, \ldots, x_n, \tau + \Delta\tau) = \left(1 - \frac{\Delta\tau}{2}\mathcal{H}\right)\Psi(x_1, x_2, \ldots, x_n, \tau) \tag{3.16}$$

where $\Delta\tau$ is the time step, typically set at 10^{-3} (a.u.). The many-body Hamiltonian matrix $\mathcal{H}$ is constructed using a position basis and has $G^N \times G^N$ elements, where G and N

are number of real-space grid points and number of electrons. Solving an eigenvalue equation has a similar time complexity to matrix multiplication (i.e. scaling with the cube of matrix size), meaning calculations at each time step have a time complexity of around $\mathcal{O}(G^{3N})$. Typical calculations have around 300 grid points, meaning that each additional electron makes calculations around 10^8 times slower. Currently, calculations involving more than three electrons would be far too costly to perform. Use of additional dimensions causes the calculation to scale as $\mathcal{O}(G^{3ND})$ where D is the number of dimensions, making it clear why these are also not possible.

The many-body matrix expressed using a position basis in 1D is given by

$$\mathcal{H} = \left\langle x^{1'}, x^{2'}, \ldots, x^{N'} \right| \hat{\mathcal{H}} \left| x^1, x^2, \ldots, x^N \right\rangle \tag{3.17}$$

where the Hamiltonian operator is given by

$$\hat{\mathcal{H}} = \sum_i K_i + \sum_i v_{\text{ext}}(x_i) + \sum_i \sum_{j>i} u(x_i, x_j). \tag{3.18}$$

The term K_i is the kinetic energy operator acting on the position co-ordinate of electron i. In the two electron case, the many-body Hamiltonian matrix has the form

$$\left\langle x^{1'}, x^{2'} \right| \hat{\mathcal{H}} \left| x^1, x^2 \right\rangle =$$

$$\begin{array}{c} \\ (x_1^{1'}, x_1^{2'}) \\ (x_1^{1'}, x_2^{2'}) \\ \vdots \\ (x_1^{1'}, x_G^{2'}) \\ (x_2^{1'}, x_1^{2'}) \\ (x_2^{1'}, x_2^{2'}) \\ \vdots \\ (x_2^{1'}, x_G^{2'}) \\ \vdots \\ (x_G^{1'}, x_G^{2'}) \end{array} \begin{array}{c} \begin{array}{cccccccc} (x_1^1, x_1^2) & \cdots & (x_1^1, x_G^2) & (x_2^1, x_1^2) & \cdots & (x_2^1, x_G^2) & \cdots & (x_G^1, x_G^2) \end{array} \\ \begin{bmatrix} \mathcal{H}_{(1,1)} & \cdots & \mathcal{H}_{(1,G)} & \mathcal{H}_{(1,G+1)} & \cdots & \mathcal{H}_{(1,2G)} & \cdots & \mathcal{H}_{(1,G^2)} \\ \mathcal{H}_{(2,1)} & \cdots & \mathcal{H}_{(2,G)} & \mathcal{H}_{(2,G+1)} & \cdots & \mathcal{H}_{(2,2G)} & \cdots & \mathcal{H}_{(2,G^2)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathcal{H}_{(G,1)} & \cdots & \mathcal{H}_{(G,G)} & \mathcal{H}_{(G,G+1)} & \cdots & \mathcal{H}_{(G,2G)} & \cdots & \mathcal{H}_{(G,G^2)} \\ \mathcal{H}_{(G+1,1)} & \cdots & \mathcal{H}_{(G+1,G)} & \mathcal{H}_{(G+1,G+1)} & \cdots & \mathcal{H}_{(G+1,2G)} & \cdots & \mathcal{H}_{(G+1,G^2)} \\ \mathcal{H}_{(G+2,1)} & \cdots & \mathcal{H}_{(G+2,G)} & \mathcal{H}_{(G+2,G+1)} & \cdots & \mathcal{H}_{(G+2,2G)} & \cdots & \mathcal{H}_{(G+2,G^2)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathcal{H}_{(2G,1)} & \cdots & \mathcal{H}_{(2G,G)} & \mathcal{H}_{(2G,G+1)} & \cdots & \mathcal{H}_{(2G,2G)} & \cdots & \mathcal{H}_{(2G,G^2)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathcal{H}_{(G^2,1)} & \cdots & \mathcal{H}_{(G^2,G)} & \mathcal{H}_{(G^2,G+1)} & \cdots & \mathcal{H}_{(G^2,2G)} & \cdots & \mathcal{H}_{(G^2,G^2)} \end{bmatrix} \end{array}, \tag{3.19}$$

where each matrix element is given by a set of co-ordinates (x_k^1, x_l^2) and $(x_i^{1'}, x_j^{2'})$, illustrated above and on the left side of the matrix. The subscript refers to the grid co-ordinate index, while the superscript refers to the electron number. The matrix

elements are given by

$$\hat{\mathcal{H}}_{ab} = \left\langle x_i^{1'}, x_j^{2'} \right| K_1 \left| x_k^1, x_l^2 \right\rangle + \left\langle x_i^{1'}, x_j^{2'} \right| K_2 \left| x_k^1, x_l^2 \right\rangle + \left\langle x_i^{1'}, x_j^{2'} \right| u(x_1, x_2) \left| x_k^1, x_l^2 \right\rangle + \left\langle x_i^{1'}, x_j^{2'} \right| v_{\text{ext}}(x_1) \left| x_k^1, x_l^2 \right\rangle + \left\langle x_i^{1'}, x_j^{2'} \right| v_{\text{ext}}(x_2) \left| x_k^1, x_l^2 \right\rangle \tag{3.20}$$

where each i, j and k, l translate to a indices a and b within the many-body matrix. The elements can be simplified to

$$\hat{\mathcal{H}}_{ab} = \left\langle x_i^{1'} \right| K_1 \left| x_k^1 \right\rangle \delta(x_j^{2'}, x_l^2) + \left\langle x_j^{2'} \right| K_2 \left| x_l^2 \right\rangle \delta(x_i^{1'}, x_k^1) + \Big[u(x_1, x_2) + v_{\text{ext}}(x_1) + v_{\text{ext}}(x_2) \Big] \delta(x_i^{1'}, x_k^1) \delta(x_j^{2'}, x_l^2). \tag{3.21}$$

The size of the above matrix can be truly massive, meaning a large amount of memory is required. This problem is worsened in 3D; a calculation involving 14 electrons (i.e. a silicon atom) in 3D with 300 grid points requires $300^{(14\times3\times2)}$ floating point numbers to be stored, which is around 10^{94} terabytes of data. The Kronecker delta functions above should highlight that the many-body matrix is quite sparse, with most of the elements being zero. Despite this, the matrix is still quite large, highlighting why working in small numbers of dimensions is a necessity. In `iDEA`, this matrix is stored using sparse matrix techniques, thus avoiding memory issues.

3.2.3 Kohn-Sham DFT implementation

The KS DFT implementation in `iDEA` uses a non-standard form for the LDA which is parametrised from finite slabs in 1D systems.[146] These have been shown to be identical to more traditional approximations based on the homogeneous electron gas (HEG). The Hartree potential for DFT is calculated as

$$v_H(x) = \overbrace{\int u(x_i, x_j)\rho(x_j)dx_j}^{\text{Continuous}} \approx \overbrace{\sum_{x_j} u(x_i, x_j)\rho(x_j)\Delta x}^{\text{Discretised}}$$

$$\begin{bmatrix} v_H(x_1) \\ v_H(x_2) \\ \vdots \\ v_H(x_n) \end{bmatrix} = \begin{bmatrix} u(x_1,x_1) & u(x_2,x_1) & \cdots & u(x_n,x_1) \\ u(x_1,x_2) & u(x_2,x_2) & \cdots & u(x_n,x_2) \\ \vdots & \vdots & \ddots & \vdots \\ u(x_1,x_n) & u(x_2,x_n) & \cdots & u(x_n,x_n) \end{bmatrix} \begin{bmatrix} \rho(x_1) \\ \rho(x_2) \\ \vdots \\ \rho(x_n) \end{bmatrix} \Delta x \tag{3.22}$$

where the Coulomb interaction matrix has elements $u(x_i, x_j)$. Like the Hartree potential, local DFT approximations (e.g. the LDA) also give rise to a vector of discretised

values. The resulting Hamiltonian matrix,

$$\hat{\mathcal{H}}^{\text{LDA}} = \frac{1}{\Delta x^2}\begin{bmatrix} 1 & -2 & 0 & 0 \\ -2 & 1 & \ddots & 0 \\ 0 & \ddots & \ddots & -2 \\ 0 & 0 & -2 & 1 \end{bmatrix} + \begin{bmatrix} v_{H1} & 0 & 0 & 0 \\ 0 & v_{H2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & v_{Hn} \end{bmatrix} + \begin{bmatrix} v_{\text{ext1}} & 0 & 0 & 0 \\ 0 & v_{\text{ext2}} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & v_{\text{extn}} \end{bmatrix} + \begin{bmatrix} v_{\text{xc1}} & 0 & 0 & 0 \\ 0 & v_{\text{xc2}} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & v_{\text{xcn}} \end{bmatrix}, \tag{3.23}$$

is then used in an eigensolver. The lowest energy eigenvectors are then used as the occupied orbitals, giving access to the ground state density. This process is then iterated until convergence. Convergence is set as when the density residual,

$$\Delta\rho = \sum_i |\rho^{\text{new}}(x_i) - \rho^{\text{old}}(x_i)| \tag{3.24}$$

between new and old density iterations is less than $10^{-8}\ a_0^{-1}$. This typically gives an energy convergence of 10^{-6} (a.u.).

3.2.4 Hartree-Fock implementation

The Fock operator can be manipulated as follows on a real-space grid.

$$\begin{aligned} \hat{\mathcal{F}}\psi_k(x_j) &= \overbrace{\int F(x_i, x_j)\psi_k(x_j)dx_j}^{\text{Continuous}} \approx \overbrace{\sum_{x_j} F(x_i, x_j)\psi_k(x_j)\Delta x}^{\text{Discretised}} \\ &= \Delta x \underbrace{\begin{bmatrix} F(x_1,x_1) & F(x_1,x_2) & \cdots & F(x_1,x_n) \\ F(x_2,x_1) & F(x_2,x_2) & \cdots & F(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ F(x_n,x_1) & F(x_n,x_2) & \cdots & F(x_n,x_n) \end{bmatrix}}_{\text{Fock matrix } \mathbf{F}} \begin{bmatrix} \psi_k(x_1) \\ \psi_k(x_2) \\ \vdots \\ \psi_k(x_n) \end{bmatrix} \end{aligned} \tag{3.25}$$

In the above, the order of multiplication by Δx does not matter, meaning knowledge of the orbitals is no longer required before using an eigensolver. The Fock matrix elements $F(x_i, x_j)$ are calculated using a tensor product $\otimes$ of the discretised orbitals,

then a Hadamard product $\circ$ with the Coulomb matrix [2].

$$\mathbf{F} = \sum_{k}^{\text{occ}} \left(\begin{bmatrix} \psi_k^\star(x_1) \\ \psi_k^\star(x_2) \\ \vdots \\ \psi_k^\star(x_n) \end{bmatrix} \otimes \begin{bmatrix} \psi_k(x_1) \\ \psi_k(x_2) \\ \vdots \\ \psi_k(x_n) \end{bmatrix} \right) \circ \begin{bmatrix} u(x_1,x_1) & u(x_2,x_1) & \cdots & u(x_n,x_1) \\ u(x_1,x_2) & u(x_2,x_2) & \cdots & u(x_1,x_2) \\ \vdots & \vdots & \ddots & \vdots \\ u(x_1,x_n) & u(x_2,x_n) & \cdots & u(x_n,x_n) \end{bmatrix} \tag{3.26}$$

3.2.5 Hybrid functional implementation

The following functionality was implemented into the `iDEA` code for a part of this thesis. The work was collaborative with members of Rex Godby's group. The HF code implemented was vectorised to increase computational efficiency. The full implementation can be found on the `iDEA` GitHub page.[147] The hybrid Hamiltonian is given by

$$\begin{aligned} \hat{\mathcal{H}}^{\text{HYB}} = &-\frac{1}{2}\frac{d^2}{dx^2} + \begin{bmatrix} v_{H1} & 0 & 0 & 0 \\ 0 & v_{H2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & v_{Hn} \end{bmatrix} + \begin{bmatrix} v_{\text{ext1}} & 0 & 0 & 0 \\ 0 & v_{\text{ext2}} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & v_{\text{extn}} \end{bmatrix} \\ &+(1-\alpha) \begin{bmatrix} v_{\text{xc1}} & 0 & 0 & 0 \\ 0 & v_{\text{xc2}} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & v_{\text{xcn}} \end{bmatrix} + \alpha \begin{bmatrix} F(x_1,x_1) & F(x_1,x_2) & \cdots & F(x_1,x_n) \\ F(x_2,x_1) & F(x_2,x_2) & \cdots & F(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ F(x_n,x_1) & F(x_n,x_2) & \cdots & F(x_n,x_n) \end{bmatrix} \Delta x. \end{aligned} \tag{3.27}$$

3.3 The CP2K code

`CP2K` is one of the most popular electronic structure codes in the UK (see Fig. 2.1). Its implementation is efficient and scalable for calculations that involve HF exchange. Compared to other codes, `CP2K` is unique in its choice of basis sets, as it uses both augmented plane waves and atom-centred basis sets to describe its valence electrons. One drawback of `CP2K` is that it operates using the Γ-point only, with limited functionality for larger numbers of **k**-points. As discussed in Section 2.6.3, this means **k**-point convergence must instead be done by increasing system size.

---

[2] The matrix elements F_{ij} (which make the matrix $\mathbf{F}$) are given by $F_{ij} = \sum_k \phi_k^\star(x_i)\phi_k(x_j)u(x_i,x_j)\Delta x$. The tensor and Hadamard products are instead used to speed up the calculation. An alternative expression for the tensor product $\mathbf{v} \otimes \mathbf{w}$ is simply $\mathbf{v}\mathbf{w}^T$, where $\mathbf{v}$ and $\mathbf{w}$ are vectors. The Hadamard product is simply an element-wise multiplication of matrix elements.

3.3.1 Representation of the orbitals

CP2K uses contracted Gaussian-type orbitals (GTOs) for its basis functions,

$$\phi_\nu(\mathbf{r}-\mathbf{R}) = \sum_k^L c_k (x-R_x)^l (y-R_y)^m (z-R_z)^n e^{-\alpha_k(\mathbf{r}-\mathbf{R})^2} \tag{3.28}$$

where c_k are called the contraction coefficients and L is the contraction length. R_x, R_y and R_z denote the Cartesian coordinates of an atom. The quantities n, l and m denote the principle, azimuthal and magnetic quantum numbers. Each of the basis functions consists of a sum of primitive Gaussian functions whose width is dependent on the exponent α_k. If the basis functions ϕ_ν were only made of primitive Gaussians (i.e. the sum over k is removed in the above equation), a large number of basis functions ϕ_ν would be needed to accurately describe each orbital ψ_i. Instead, the approach can be made more efficient by having each basis function be comprised of a fixed combination of many primitive Gaussians. The proportion of each primitive function is described by the (fixed) contraction coefficients c_k. Compared to other choices of basis set, such as Slater-type orbitals (STOs), GTOs allow for a faster evaluation of four-centre integrals. This arises from the fact that a product of two GTOs results in another GTO. As a result, GTOs can be around five orders of magnitude faster than STOs. These basis functions are not necessarily orthogonal, meaning their overlap matrix may not equal the identity matrix ($S_{ij} = \langle \phi_i | \phi_j \rangle \neq \delta_{ij}$).

3.3.2 Representation of the density

In some electronic structure codes, orbitals and densities can be described by different basis sets. CP2K is one such case, where a primary basis set is used to represent both orbitals and densities, and an additional auxiliary basis set is used for representing densities. By having a dual representation of the density, parts of a calculation can be evaluated faster, such as the Hartree potential. CP2K has a representation of the density in terms of GTOs,

$$\begin{aligned}\rho(\mathbf{r}) = \sum_i^{\text{occ}} |\psi_i|^2 = \sum_i^{\text{occ}} \left(\sum_\mu C_{i\mu}^\star \phi_\mu^\star(\mathbf{r}) \right) \left(\sum_\nu C_{i\nu} \phi_\nu(\mathbf{r}) \right) &= \sum_{\mu\nu} \underbrace{\sum_i^{\text{occ}} C_{i\mu}^\star C_{i\nu}}_{P_{\mu\nu}} \phi_\mu^\star(\mathbf{r}) \phi_\nu(\mathbf{r}) \\ &= \sum_{\mu\nu} P_{\mu\nu} \phi_\mu^\star(\mathbf{r}) \phi_\nu(\mathbf{r}). \end{aligned} \tag{3.29}$$

where $P_{\mu\nu}$ are elements of the density matrix $\mathbf{P}$.

The Gaussian and plane waves approach (GPW) approach in CP2K uses GTOs as the primary basis set and plane waves as an auxiliary representation of the density. Using this, the KS potential to be evaluated at a reduced computational cost.

3.3.3 Hybrid functional implementation

In this section a series of approximations are used to reduce the computational cost associated with HF exchange integrals. While the techniques used do constitute an approximation, they essentially only filter out exchange integrals which do not contribute much to the total energy. The discussion begins with the Roothan equations, which provide a means to solve the HF equations when using basis sets.

Roothan equations

The HF equations (Eqn. 2.13a, Eqn. 2.13b and Eqn. 2.13c) can be expanded in terms of a basis set and manipulated as follows. [3]

$$\hat{\mathcal{H}}_i^{HF} \sum_\nu C_{i\nu}\,\phi_\nu(\mathbf{r}) = \varepsilon_i \sum_k C_{i\nu}\,\phi_\nu(\mathbf{r}) \tag{3.30a}$$

$$\int \phi_\mu^\star(\mathbf{r}) \left(\hat{\mathcal{H}}_i^{HF} \sum_\nu C_{i\nu}\phi_\nu(\mathbf{r}) \right) d\mathbf{r} = \int \phi_\mu^\star(\mathbf{r}) \left(\varepsilon_i \sum_\nu C_{i\nu}\phi_\nu(\mathbf{r}) \right) d\mathbf{r} \tag{3.30b}$$

$$\sum_\nu C_{i\nu} \underbrace{\left(\int \phi_\mu^\star(\mathbf{r}) \hat{\mathcal{H}}_i^{HF} \phi_\nu(\mathbf{r}) d\mathbf{r} \right)}_{\text{Fock matrix elements } F_{\mu\nu}} = \sum_\nu C_{i\nu} \underbrace{\left(\int \phi_\mu^\star(\mathbf{r}) \phi_\nu(\mathbf{r}) d\mathbf{r} \right)}_{\text{Overlap elements } S_{\mu\nu}} \varepsilon_i \tag{3.30c}$$

$$\sum_\nu F_{\mu\nu} C_{i\nu} = \sum_\nu S_{\mu\nu} C_{i\nu} \varepsilon_i \tag{3.30d}$$

The final set of equations are simply a matrix multiplication and can be expressed concisely as

$$\mathbf{FC} = \mathbf{SC}\varepsilon. \tag{3.31}$$

These equations, known as the Roothaan equations, are an eigenvalue problem that can be solved numerically. While the above resembles a generalised eigenvalue problem, a normal eigenvalue problem can be recovered by solving $\mathbf{S}^{-1}\mathbf{FC} = \mathbf{C}\varepsilon$ instead. However, using a generalised eigenvalue problem solver (e.g. one from LAPACK) is usually faster than inverting the $\mathbf{S}$ matrix and using a regular eigensolver. In the case

[3] Standard calculations include the self-interaction term in the Hartree potential and its correction in HF. In this section, references to the HF equations (Eqn. 2.13a, Eqn. 2.13b and Eqn. 2.13c) and hybrid functionals refer to those equations but with all $i = j$ terms included.

where basis functions are orthogonal, $\mathbf{S} = I$ where I is the identity matrix. In general, the matrix $\mathbf{C}$ can be complex, however they are real when only the Γ point is used in periodic calculations. The above Roothan equations are for closed-shell systems; for open-shell systems a pair of Roothan equations would be solved and are known as the Pople-Nesbet-Berthier equations. A similar equation to the Roothaan equations can be made for the KS equations

$$\mathbf{KC} = \mathbf{SC}\varepsilon, \tag{3.32}$$

where $\mathbf{K}$ is called the KS matrix.

Four-centre integrals

If the wavefunctions in the HF exchange energy (Eqn. 2.14) are expanded in terms of basis functions ϕ_ν, the following expression can be found:

$$E_{\mathrm{x}}^{\mathrm{HF}} = -\frac{1}{2}\sum_{\nu\mu\sigma\lambda}^{\mathrm{bf}} \underbrace{\left(\sum_{i}^{\mathrm{occ}} C_{i\nu}^{\star} C_{i\lambda}\right)}_{P_{\nu\lambda}} \underbrace{\left(\sum_{j}^{\mathrm{occ}} C_{j\mu}^{\star} C_{j\sigma}\right)}_{P_{\mu\sigma}} \underbrace{\iint \frac{\phi_\nu^\star(\mathbf{r}_i)\phi_\mu^\star(\mathbf{r}_i)\phi_\sigma(\mathbf{r}_j)\phi_\lambda(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_i d\mathbf{r}_j}_{\text{Shorthand } (\nu\mu|\sigma\lambda)} \tag{3.33}$$

$$E_{\mathrm{x}}^{\mathrm{HF}}[\mathbf{P}] = -\frac{1}{2}\sum_{\nu\mu\sigma\lambda}^{\mathrm{bf}} P_{\nu\lambda} P_{\mu\sigma} \, (\nu\mu|\sigma\lambda) \tag{3.34}$$

where the elements $P_{\mu\nu}$ are again components of the density matrix $\mathbf{P}$. These equations highlight very clearly the sheer number of exchange integrals, known as four-centre integrals, that need to be evaluated in a HF calculation. If these integrals are calculated as is, the scaling with both number of basis functions and system size is $\mathcal{O}(N^4)$.

Permutation symmetry

The above integrals have eight-fold symmetry that can be exploited,

$$\begin{aligned}(\nu\mu|\sigma\lambda) &= (\mu\nu|\sigma\lambda) = (\nu\mu|\lambda\sigma) = (\mu\nu|\lambda\sigma) \\ &= (\lambda\sigma|\mu\nu) = (\sigma\lambda|\mu\nu) = (\lambda\sigma|\nu\mu) = (\sigma\lambda|\nu\mu)\,,\end{aligned} \tag{3.35}$$

resulting in an eight-fold reduction in the number of integrals performed. By looking at the expanded form of the exchange energy above (Eqn. 3.33), it is clear that these symmetries arise from the the interchangeable order of basis function multiplication. While reducing the pre-factor associated with calculation scaling, this symmetry does

not change scaling with either number of basis functions or system size.

Integral screening

Near-field screening techniques can be used to reduce the computational cost of HF exchange integrals. A Cauchy-Schwarz inequality

$$(\nu\mu|\sigma\lambda) \leq (\mu\nu|\mu\nu)^{\frac{1}{2}} (\sigma\lambda|\sigma\lambda)^{\frac{1}{2}} \tag{3.36}$$

can be used to provide an upper bound on an exchange integral's value. If the right-hand side of the above inequality is below a specified threshold ε_{cs}, the left-hand integral can be discarded. The upper bound integrals that need to be performed are only two-centre integrals. Instead of computing and storing these two-centre integrals, they can instead be parametrised into a function of inter-atomic distances $|\mathbf{R}|_{\mu\nu}$ between each of the two basis functions (ϕ_μ and ϕ_ν) in the two centre integrals:

$$\log\left(\mu\nu|\mu\nu\right)\left(|\mathbf{R}_{\mu\nu}|\right) \lessapprox a_1|\mathbf{R}|_{\mu\nu} + a_0 \tag{3.37}$$

where a_1 and a_0 are fitted coefficients. In order to ensure accuracy, different fits are done at the start of each calculation for each different type of basis function, shell and atomic nucleus. This function is approximately an upper-bound to the (upper-bound of the) Cauchy-Schwarz inequality and reduces scaling with system size from $\mathcal{O}(N^4)$ to $\mathcal{O}(N^2)$. When both near- and far-field screening techniques are applied, system size scaling reduces to $\mathcal{O}(N)$. An illustration of this is shown in Fig. 3.1

The quantity ε_{cs} is usually set to around 10^{-6} to 10^{-10}. The error in the total energy from using this approach is approximated to be around ε_{cs} per electron. Calculations performed in Chapter 5 and Chapter 6 contain around 2000 electrons, amounting to an error of around 3 meV in the total energy. This error will likely be smaller when considering total energy differences.

Auxiliary density matrix method

Despite improvements in integral screening techniques, scaling with basis set quality remains $\mathcal{O}(N^4)$. In addition to the previously mentioned auxiliary basis `CP2K` uses for the density, the code also employs an additional basis for HF exchange integrals. This approach is known as the auxiliary density matrix method (ADMM) and can reduce the computational cost of HF calculations by a factor of around 20. In this section, a hat

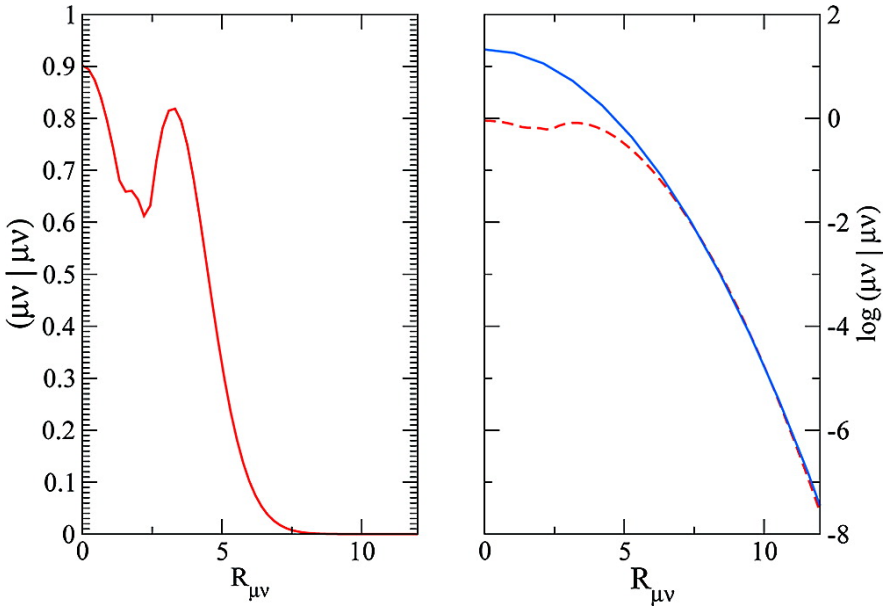

Figure 3.1: (Left) The two electron integral $(\mu\nu|\mu\nu)$ between lithium p- and hydrogen d-functions as a function of interatomic distance $|\mathbf{R}_{\mu\nu}|$. (Right) The dashed line shows the logarithm of the left hand panel data, while the solid (blue) line shows a fitted upper-bound function to the data. Reprinted with permission from *J. Chem. Theory Comput. 2009, 5, 11, 3010-3021*. Copyright 2009 American Chemical Society.

or a tilde is used to denote variables expressed in an auxiliary basis. The expression relating an orbital $\hat{\psi}_i$ and its auxiliary basis $\hat{\phi}_i$ is

$$\hat{\psi}_i(\mathbf{r}) = \sum_{\nu} \hat{C}_{\nu i} \hat{\phi}_\nu(\mathbf{r}), \tag{3.38}$$

where the new auxiliary basis set is either smaller in size or described by more rapidly decaying Gaussian primitive functions. The exact HF energy can be expressed in terms of the auxiliary density matrix $\hat{\mathbf{P}}$ as

$$E_{\mathrm{x}}^{\mathrm{HF}}[\mathbf{P}] = E_{\mathrm{x}}^{\mathrm{HF}}[\hat{\mathbf{P}}] + (E_{\mathrm{x}}^{\mathrm{HF}}[\mathbf{P}] - E_{\mathrm{x}}^{\mathrm{HF}}[\hat{\mathbf{P}}]). \tag{3.39}$$

While the above still requires use of the non-auxiliary density matrix for HF, the term can instead be approximated as

$$E_{\mathrm{x}}^{\mathrm{HF}}[\mathbf{P}] \approx E_{\mathrm{x}}^{\mathrm{HF}}[\hat{\mathbf{P}}] + (E_{\mathrm{x}}^{\mathrm{DFT}}[\mathbf{P}] - E_{\mathrm{x}}^{\mathrm{DFT}}[\hat{\mathbf{P}}]) \tag{3.40}$$

which reduces computational costs. Numerical tests have shown that the exchange energy difference is well-captured by DFT, as shown in reference 148.

The auxiliary basis coefficients can be chosen as the those which minimise the func-

tional

$$\Delta_1\left[\hat{\psi}_i\right] = \sum_i \int \left(\psi_i(\mathbf{r}) - \hat{\psi}_i(\mathbf{r})\right)^2 d\mathbf{r} \tag{3.41}$$

with respect to the auxiliary basis coefficients $\hat{C}_{i\mu}$. The non-auxiliary orbitals ψ_i are assumed to be orthonormal (i.e. $\langle\psi_i|\psi_j\rangle = \delta_{ij}$). The above functional is expanded in terms of a basis set and manipulated as

$$\Delta_1\left[\hat{\psi}_i\right] = \sum_i \int \left(\psi_i(\mathbf{r})\psi_i(\mathbf{r}) - \psi_i(\mathbf{r})\hat{\psi}_i(\mathbf{r}) + \hat{\psi}_i(\mathbf{r})\hat{\psi}_i(\mathbf{r})\right) d\mathbf{r} \tag{3.42}$$

$$\Delta_1\left[\hat{C}_{i\star}\right] = \sum_i \int \left(\sum_{nn'} C_{in}C_{in'}\phi_n\phi_{n'} - 2\sum_{nm} C_{in}\hat{C}_{im}\phi_n\hat{\phi}_m + \sum_{mm'} \hat{C}_{im}\hat{C}_{im'}\hat{\phi}_m\hat{\phi}_{m'}\right) d\mathbf{r} \tag{3.43}$$

$$\frac{\delta\Delta_1\left[\hat{C}_{i\star}\right]}{\delta\hat{C}_{i\star}} = \sum_i \int \left(-2\sum_{mn} C_{in}\phi_n\hat{\phi}_m + 2\sum_{mm'} \hat{C}_{im}\hat{\phi}_m\hat{\phi}_{m'}\right) d\mathbf{r}. \tag{3.44}$$

If the functional is minimised (i.e. $\delta\Delta/\delta\hat{C}_{i\star} = 0$), the following expression can be found:

$$\sum_{im}\hat{C}_{im}\sum_{mm'}\underbrace{\int \hat{\phi}_m(\mathbf{r})\hat{\phi}_{m'}(\mathbf{r})d\mathbf{r}}_{\text{Overlap elem. } \hat{S}_{mm'}} = \sum_{in} C_{in}\sum_{mn}\underbrace{\int \phi_n(\mathbf{r})\hat{\phi}_m(\mathbf{r})d\mathbf{r}}_{\text{Overlap elem. } Q_{nm}}. \tag{3.45}$$

In matrix notation, this would be

$$\hat{\mathbf{C}} = \hat{\mathbf{S}}^{-1}\mathbf{Q}\mathbf{C} \tag{3.46}$$

$$\hat{\mathbf{P}} = \hat{\mathbf{C}}\hat{\mathbf{C}}^T, \tag{3.47}$$

providing a means to project orbitals expressed in basis set onto another basis set using the projector matrix $\hat{\mathbf{S}}^{-1}\mathbf{Q}$.

The orbitals $\hat{\psi}(\mathbf{r})$ expressed in the auxiliary basis are not orthonormal using the above approach. It would benefit calculation scaling properties if the orbitals had such a property. This can be achieved by minimising an alternative functional,

$$\Delta_2\left[\tilde{\psi}_i\right] = \sum_i \int \left(\psi_i(\mathbf{r}) - \tilde{\psi}_i(\mathbf{r})\right)^2 d\mathbf{r} + \sum_{kl}\int \Lambda_{kl}\left(\tilde{\psi}_k\tilde{\psi}_l - \delta_{kl}\right) d\mathbf{r} \tag{3.48}$$

where orthogonality is be enforced via Lagrange multipliers. Following a similar procedure to the above, the projection between the primary and auxiliary basis is given

by [4]

$$\tilde{\mathbf{C}} = \hat{\mathbf{C}}\mathbf{\Lambda}^{-1/2} \tag{3.49}$$

$$\mathbf{\Lambda} = \hat{\mathbf{C}}^T\hat{\mathbf{S}}\hat{\mathbf{C}} \tag{3.50}$$

$$\tilde{\mathbf{P}} = \tilde{\mathbf{C}}\tilde{\mathbf{C}}^T \tag{3.51}$$

The use of this second scheme removes certain properties that a pure density matrix should have. For example, a pure matrix satisfies the condition

$$\mathrm{tr}(\mathbf{PS}) = \int \left(\sum_i^{\mathrm{occ}} C_{i\mu}\phi_\mu\right)\left(\sum_i^{\mathrm{occ}} C_{i\sigma}\phi_\sigma\right) d\mathbf{r} = \int \rho(\mathbf{r}) d\mathbf{r} = N_e \tag{3.52}$$

which ensures that the number of electrons is correct. When using the orthonormal orbitals $\tilde{\psi}$ this constraint is not met. Matrix purification is a technique that can be used to ensure that these constraints are met. If purified, the orthonormal density matrix $\tilde{\mathbf{P}}$ reverts to becoming the same as $\hat{\mathbf{P}}$. Further, the eigenvalues corresponding to $\tilde{\psi}$ cannot be used once purification has taken place. Using either the non-purified $\tilde{\psi}$ or non-orthogonal $\hat{\psi}$ introduces similar errors into a calculation as shown by Guidon *et al*,[148] so we opt for the (faster) non-purified approach (i.e. the $\tilde{\psi}$ given by Δ_2). Calculations performed in Chapter 5 assess the impact of different ADMM basis sets on relevant semiconductor properties. A more detailed discussion on the ADMM approach can be found in Reference 148.

3.3.4 Pseudopotentials

CP2K uses Goedecker-Teter-Hutter (GTH) pseudopotentials in its calculations, which describe pseudo-atoms by a local and non-local potential. GTH pseudopotentials are also norm-conserving, meaning they also satisfy the condition

$$\int_{|\mathbf{r}|<r_c} \psi_i'(\mathbf{r})\psi_j'(\mathbf{r})\, d\mathbf{r} = \int_{|\mathbf{r}|<r_c} \psi_i(\mathbf{r})\psi_j(\mathbf{r})\, d\mathbf{r}. \tag{3.53}$$

[4] The overlap matrix elements Q_{nm} for the orthogonal ADMM approach are defined in the same way as the previous case, i.e. $Q_{nm} = \int \phi_n(\mathbf{r})\tilde{\phi}_m(\mathbf{r})d\mathbf{r}$. Similarly, $\tilde{S}_{nm} = \int \tilde{\phi}_n(\mathbf{r})\tilde{\phi}_m(\mathbf{r})d\mathbf{r}$.

3.4 The VASP code

In contrast to CP2K, the code VASP expresses orbitals in terms of plane waves (see Eqn. 3.3), meaning the cutoff E_C is an exact measure of basis set completeness. Using plane waves reduces the cost of Fourier transforms and construction of the KS matrix. Pulay forces are also absent with this basis set choice. Generally, plane wave electronic structure codes perform quite well for calculations that do not require HF exchange. A large number of plane waves is usually required to accurately describe an orbital, meaning the $\mathcal{O}(N^4)$ for HF exchange with basis set size results in large computational cost. In addition, integral screening techniques (e.g. Cauchy-Schwarz) cannot be applied with plane waves.

Plane wave DFT calculations contain a series of steps which are rate limiting under different circumstances. Scaling of DFT calculations is $\mathcal{O}(N_e^3)$ due to the wavefunction orthonormalisation step required.[149] However, for smaller numbers of atoms the diagonalisation step has a small prefactor and computational cost is dominated by other operations. For many calculations, the cost of DFT calculations is dominated by fast Fourier transforms, which have a cost of $\mathcal{O}(N_e N_G \ln N_G)$ where N_G is the number of plane waves.[149] The so-called 'nearsightedness' or 'locality' of quantum mechanics can be exploited to achieve linear scaling DFT calculations through use of localised basis sets.[132]

3.4.1 Projector augmented wave method

The projector and augmented plane wave method used in VASP bears some resemblance to the approach used in CP2K . In this, space is divided into a series of augmentation (or core-like) regions and interstitial regions. The all electron and pseudopotential orbitals are denoted ψ_i and ψ_i'. In the interstitial regions both of these are equal. In the augmentation region, the two wavefunctions can differ. Within the augmentation region the two wavefunctions are related via

$$\psi_i(\mathbf{r}) = \psi_i'(\mathbf{r}) + \sum_\nu \left[\phi_\nu(\mathbf{r}) - \phi_\nu'(\mathbf{r})\right] \left\langle p_\nu' \middle| \psi_i' \right\rangle \tag{3.54}$$

which requires three key quantities to be defined – an all-electron basis $\{\phi_\nu\}$, a pseudo wavefunction basis $\{\phi_\nu'\}$ and a set of projectors p_ν'. The projectors must satisfy both

the above equation and the following conditions:

$$\sum_{\nu} |\phi'_{\nu}\rangle \langle p'_{\nu}| = 1 \tag{3.55}$$

$$\langle p'_{\nu}|\phi'_{\nu'}\rangle = \delta_{\nu,\nu'}. \tag{3.56}$$

Critically, Eqn. 3.54 only holds if the two basis sets have the property

$$\psi'_i(\mathbf{r}) = \sum_{\nu} \langle p'_{\nu}|\psi'_i\rangle \, \phi'_{\nu}(\mathbf{r}) = \sum_{\nu} C_{i\nu}\phi'_{\nu}(\mathbf{r}) \tag{3.57}$$

$$\psi_i(\mathbf{r}) = \sum_{\nu} \langle p_{\nu}|\psi_i\rangle \, \phi_{\nu}(\mathbf{r}) = \sum_{\nu} C_{i\nu}\phi_{\nu}(\mathbf{r}), \tag{3.58}$$

where p_ν in this case is a projector onto the all-electron basis. The above simply states that expansion coefficients of the pseudo and all-electron orbitals are the same in their respective basis expansions. The basis functions are defined using solutions of the Schrödinger equation for the isolated atom (which the pseudopotential is for). The augmentation region can be defined using a radial cutoff, outside of which both the pseudo wavefunction and all-electron wavefunction are equal. The pseudo wavefunction in `VASP` is expressed using a plane wave basis.

3.4.2 The Hubbard correction

DFT+U is another common approach to overcome the delocalisation errors present in DFT. The so-called Hubbard-U correction is implemented using

$$E_{\text{tot}}^{\text{DFT}+U}[\mathbf{n}^{I\sigma}] = \sum_{I\sigma} \frac{U^I}{2} \text{Tr}\left[\mathbf{n}^{I\sigma}\left(1 - \mathbf{n}^{I\sigma}\right)\right] \tag{3.59}$$

where $\mathbf{n}^{I\sigma}$ is the occupation matrix of an atom I for spin σ electrons. There exist a few ways to implement this correction, however the above form is the simplest and has nicer properties, such as rotational invariance. This is calculated using

$$n^{I\sigma}_{m,m'} = \sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle \psi_{n\mathbf{k}}| P^I_{m,m'} |\psi_{n\mathbf{k}}\rangle \tag{3.60}$$

where $P^I_{m,m'}$ is a projector onto a predefined localised state [5]. In `VASP` this state is defined using spherical harmonics, as shown in reference 150. The above occupation matrix elements are simply projections of the KS orbitals onto a localised basis. Using

[5] $f_{n\mathbf{k}}$ is the occupation of a KS state with indices n and $\mathbf{k}$. This number is between 0 and 1 for electrons with spin. In the closed-shell case, the number is between 0 and 2.

the above energy correction, the potential becomes repulsive for electrons that occupy an electronic state by less than half and attractive for the remainder. This behaviour forces electrons to localise and adds an energy penalty to delocalised states. The size of the repulsiveness is governed by the size of the Hubbard-U energy correction U^I.

A key difference between the DFT+U and hybrid approaches is that DFT+U provides additional parameters; while hybrid functionals have one key parameter (α), DFT+U introduces one parameter per orbital. In the case of TiO_2, a correction is needed for oxygen p (U^{O}_{2p}) and titanium $3d$ (U^{Ti}_{3d}) orbitals as polarons can localise on either orbital. In principle the GKT can be used to fit these parameters however the fitted values are interdependent. In turn, these values need to be either optimised self-consistently (through iterative fittings) or through other means. Furthermore, cases where defects are present can introduce different chemical environments on surrounding orbitals of the same kind, all of which also require further fitting. As a result, the DFT+U approach is much more time consuming and will likely provide only a qualitatively correct picture.

3.5 Summary

The techniques this chapter introduces three DFT implementations which use the theory provided in Chapter 2. While some additional electronic structure techniques are discussed, such as the DFT+U approach, this chapter largely discusses numerical techniques for solving single-particle equations. In subsequent chapters we both use and assess these techniques on relevant systems.

CHAPTER 4

ACCURACY OF KOOPMANS-COMPLIANT HYBRID FUNCTIONALS [1]

The work presented in this chapter compares results obtained from hybrid functionals to exact solutions of the many-electron Schrödinger equation. By performing such a comparison, it becomes possible to identify the accuracy of hybrid functionals overall, and also whether the Koopmans parameterisation yields accurate densities and energies. As mentioned in previous chapters, exact solutions come at a substantial computational cost, therefore restricting the numbers of electrons, even in 1D. The implementation details of this work are described in Chapter 3, alongside the approach used to obtain exact many-electron solutions. The results indicate that Koopmans' compliant hybrid functionals yield strikingly accurate electron densities and energies for both exchange dominated and correlated systems.

[1] This chapter represents collaborative work that has been published: A. R. Elmaslmane, J. Wetherell, M. J. P. Hodgson, K. P. McKenna, and R. W. Godby. 'Accuracy of electron densities obtained via Koopmans-compliant hybrid functionals'. *Phys. Rev. Materials* 2, 040801(R). The remaining sections of this chapter have been adapted from the publication.

4.1 Introduction

A key measure of success for any electronic-structure theory is its ability to yield accurate electron densities and energies for many-electron systems. For example, KS DFT[131,151] is in principle exact, but the use of an approximate xc potential, such as the LDA[152] or GGA,[139] is associated with a self-interaction error which can cause the spurious delocalisation of localised charge[153] and incorrect dissociation behavior for molecules.[154] Recently, hybrid functionals that mix HF exchange with a (semi-)local approximation (such as the LDA or GGA)[141] have become popular as an alternative approach to xc. However, hybrids introduce at least one additional parameter, the mixing parameter α. This is often determined empirically, e.g., via experimental data, or through the adiabatic connection.[155] In this chapter, α is determined using a group of more physically justified constraints, including the GKT.[31,119,156–158] While it has been shown that this constrained hybrid approach results in ionisation energies and band gaps close to experimental values,[119,157] to date the electron density of this approach has not been directly compared to the exact density.

4.2 Background

As Medvedev *et al* argue,[159] progress in the accuracy of electronic structure calculations requires improvements in both energies *and* densities. Srebro *et al* indirectly assessed densities obtained via hybrid functionals using the electric field gradient at the nucleus.[120] Reference 160 obtained densities from popular empirical hybrid functional parameterisations and found sensitivity to the value of the various mixing parameters. Good agreement between hybrid and coupled cluster densities (singles and doubles) has been found for the CO molecule.[161]

In order to address the density more directly, a set of model systems are considered where the many-body problem can be solved exactly for a small number of electrons, allowing for a direct comparison of densities, energy gaps and ionisation potentials (IPs) obtained from the constrained hybrid approach to the exact values. It is shown that an *ab initio* determination of α results in hybrid functionals yielding extremely accurate densities and gaps.

The exact total energy E (of a many-electron system) is piecewise linear with respect to the number of electrons, N.[25,116] In exact KS DFT, the slope of each straight-line segment $\partial E/\partial N$ is shown by Janak's theorem to equal the highest (partly) occu-

pied molecular orbital eigenvalue.[26] The usual approximate density functionals (LDA and GGA), and HF, exhibit nonzero curvature $\partial^2 E/\partial N^2$, which can lead to qualitatively wrong physical behavior.[28–30] The curvatures are of opposite signs which means that hybrid approximations benefit from a partial cancellation of these errors.[119,156]

The exact total energy difference $E(N-1) - E(N)$ is both the ionisation energy of the N–electron system, $I(N)$, and the electron affinity of the $(N-1)$–electron system, $A(N-1)$. In HF, the equivalent of Janak's theorem[140] shows that the slope $(\partial E/\partial N)_{N-\delta}$ is equal to the HOMO eigenvalue, and $(\partial E/\partial N)_{N+\delta}$ to the lowest unoccupied molecular orbital (LUMO) eigenvalue. In exact KS DFT, the LUMO eigenvalue differs from the negative electron affinity $-A$ by a discontinuity, Δ, in the xc potential.[25] Thus all three quantities $\varepsilon_N(N-1) + \Delta$, $\varepsilon_N(N)$ and $E(N) - E(N-1)$ should, in principle, be equal, where Δ is non-zero for exact DFT methods. But for approximate methods such as hybrids where exchange and correlation are explicitly analytical functionals of the single-particle orbitals and therefore exhibit zero derivative discontinuity Δ, the first quantity becomes $\varepsilon_N(N-1)$.[140,162,163] Three requirements may therefore be identified,

(A) $\varepsilon_N(N-1) = -A(N-1) \equiv E(N) - E(N-1)$,

(B) $\varepsilon_N(N-1) = \varepsilon_N(N)$,

(C) $\varepsilon_N(N) = -I(N) \equiv E(N) - E(N-1)$,

which may be used to constrain a hybrid calculation by enforcing internal consistency. In practice, the parameter α of the basic hybrid approach provides a single degree of freedom and so can be used to impose (A) the LUMO-A condition *or* (B) the LUMO-HOMO condition *or* (C) the HOMO-I condition, or GKT. The merits, as regards electron energies, of satisfying the last two conditions using a more elaborate hybrid form has been investigated[164–166]

A key point regarding the hybrid approach is that the derivative discontinuity Δ in the xc potential not only *is* zero, but also *should be* zero, when viewed from the perspective of many-body perturbation theory. This is most clearly seen by noting that the description of exchange and correlation in the hybrid approach includes a reduced-strength Fock operator, essentially mimicking the screened exchange operator that is at the heart of the well-known GW approximation to the self-energy operator,[167–170] plus LDA exchange and correlation reduced in strength. This identification of the hybrid approach's "self-energy" as a screened-exchange approximation to the exact self-

energy Σ_{xc}, as noted by other authors,[171,172] means that Σ_{xc} would yield exact electron addition and removal energies through its one-electron eigenvalues that then acquire the significance of quasiparticle energies. Hence in both the N and $(N-1)$–particle systems both the HOMO and LUMO energies may be regarded as fairly sophisticated approximations to the ionisation potential and electron affinity, and therefore require no Δ correction.

4.3 Methods

The hybrid functional that is tested straightforwardly mixes HF with an LDA xc potential:

$$v_{xc}^{HYB}(\alpha) = \alpha v_x^{HF} + (1-\alpha) v_{xc}^{LDA}, \tag{4.1}$$

where V_{xc}^{HYB}, V_{xc}^{LDA} and V_x^{HF} denote the hybrid and LDA xc potentials and the non-local HF exchange potential, respectively. This has the advantage of focusing more on the variational power of HF for exchange-dominated systems and accommodating better the cross-over between exchange and correlation when the LDA is applied to inhomogeneous systems.

The LDA used in this work is parameterised from finite slabs.[146] Testing has shown these give indistinguishable results compared to HEG-based LDAs. The retention of the full LDA correlation potential is also explored, mixing only the exchange terms,

$$v_{xc}^{HYB}(\alpha) = \alpha v_x^{HF} + (1-\alpha) v_x^{LDA} + v_c^{LDA}. \tag{4.2}$$

Hybrid functionals are assessed both in systems where correlation is relatively unimportant ("exchange-dominated") and systems in which correlation is more significant. The exact many-body wavefunction (used to compute the exact density) is obtained by direct solution of the many-body Schrödinger equation using the `iDEA` code.[173] The electrons interact via the softened Coulomb interaction

$$u(x,x') = \frac{1}{|x-x'|+1} \tag{4.3}$$

and are treated as spinless in order to model more closely the richness of correlation found in systems containing a large number of electrons.

In order to verify that the curvature $\partial^2 E/\partial N^2$ in the functionals tested is indeed

better using the constrained hybrid approach, the derivative of energy with respect to number of electrons $\partial E/\partial N$ is calculated. This means calculating energies for fractional numbers of electrons, which can accessed by setting the density and kinetic energy contributions as

$$\rho(x) = \sum_i f_i |\psi_i|^2 \tag{4.4}$$

$$T_s[\{\psi_i[\rho]\}] = \sum_i f_i \langle \psi_i | -\frac{1}{2}\nabla^2 | \psi_i \rangle = \sum_i f_i t_i. \tag{4.5}$$

This is in line with Reference 26. In the above t_i is the kinetic energy contribution of an orbital with index i to the total kinetic energy. The occupation number f_i satisfies $0 \leq f_i \leq 1$. While in principle each orbital has its own occupation, only the occupation of the HOMO may be fractional. By expressing the single-particle kinetic energy term as

$$t_i = \varepsilon_i - \langle \psi_i | \left(v_H + v_{\text{xc}}\right) | \psi_i \rangle \tag{4.6}$$

the total energy can be found as

$$E_{\text{tot}} = \sum_i f_i t_i + E_H + E_{xc} \tag{4.7}$$

$$E_{\text{tot}} = \sum_i f_i \varepsilon_i - \underbrace{\sum_i \left[\int \psi_i^\star \left(v_H + v_{\text{xc}}\right) \psi_i dx \right]}_{2E_H + 2E_{\text{xc}}} + E_H + E_{xc} \tag{4.8}$$

$$E_{\text{tot}} = \sum_i f_i \varepsilon_i - E_H - E_{xc}. \tag{4.9}$$

If the correct form of E_{xc} is substituted into the above, taking into account the incorrect variational form of the LDA xc energy, the following can be found

$$E_{\text{tot}}^{\text{HYB}} = \sum_i f_i \varepsilon_i - E_H + \alpha E_{\text{x}} - (1-\alpha) \left[\int \rho(\mathbf{r}) v_{\text{xc}}^{\text{LDA}}(\mathbf{r}) d\mathbf{r} - \int \varepsilon_{\text{xc}}^{\text{LDA}} \rho(\mathbf{r}) d\mathbf{r} \right] \tag{4.10}$$

where the density in this equation is given by Eqn. 4.4. Therefore, only the energy and density calculations need to be varied in order to access fractional numbers of electrons.

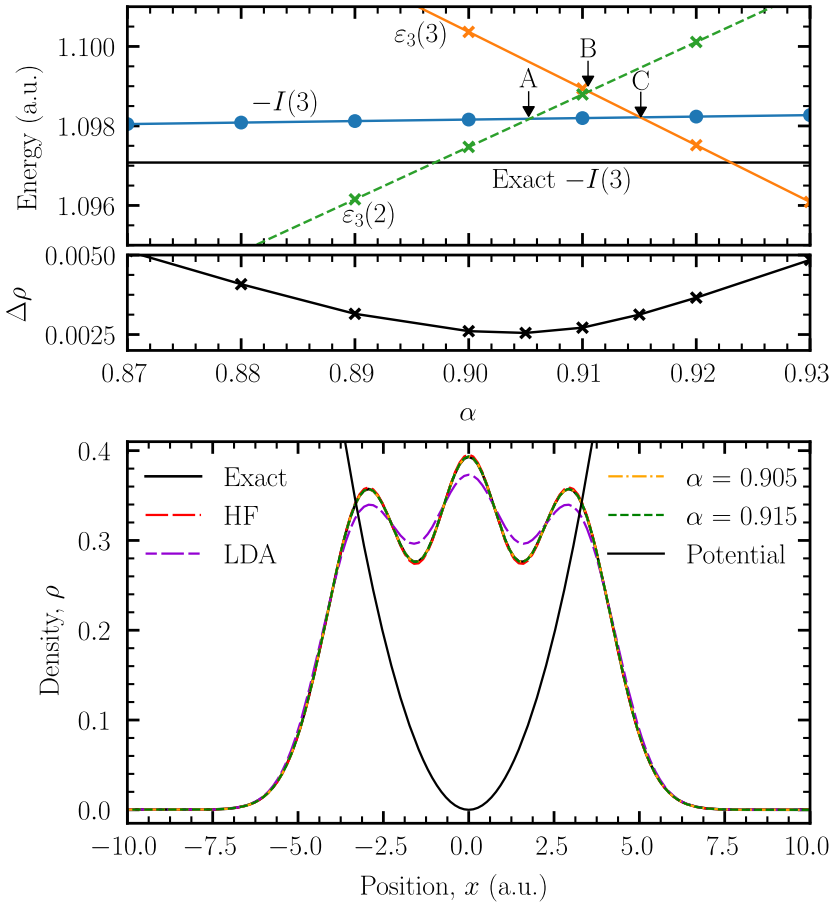

Figure 4.1: (Upper) The variation in hybrid ionisation energy $I(3)(= A(2))$, exact $I(3)(= A(2))$, $\varepsilon_3(3)$ and $\varepsilon_3(2)$ with α are illustrated for three electrons in an harmonic oscillator with $\omega = 0.25$, an exchange-dominated system. Energies are in Hartree atomic units. There are three 'crossing points': (A) A-LUMO, (B) HOMO-LUMO and (C) I-HOMO. (Center) The integrated absolute error in the density $\Delta\rho$ is shown for each value of α. This is defined as $\int |\rho^{\mathrm{EXT}}(x) - \rho^{\mathrm{HYB}}(x)| dx$ where the ρ^{EXT} and ρ^{HYB} correspond to the exact and hybrid densities. (Lower) The densities for crossings (A) and (C) are benchmarked against the exact, LDA and HF cases; the hybrid, HF and exact curves lie close together.

4.4 Exchange-dominated systems

Fig. 4.1 demonstrates for the harmonic well with angular frequency $\omega = 0.25$ (an exchange-dominated system) that application of any of the conditions (A)–(C) yields an α very close to pure HF, i.e., $\alpha \approx 1$, as expected. Other exchange-dominated systems tested yield similarly good results from the constrained hybrid.

Conditions (A)–(C) correspond to three 'crossing points', as shown in Fig. 4.1. Using the argument laid out previously, the self-energy should satisfy all three of these conditions. Generally, (A)–(C) correspond to different conditions that specify where the HOMO, LUMO and IP of a system lie with respect to one another. Although it clear from Fig. 4.1 that the three conditions cannot be exactly satisfied, the three cross-

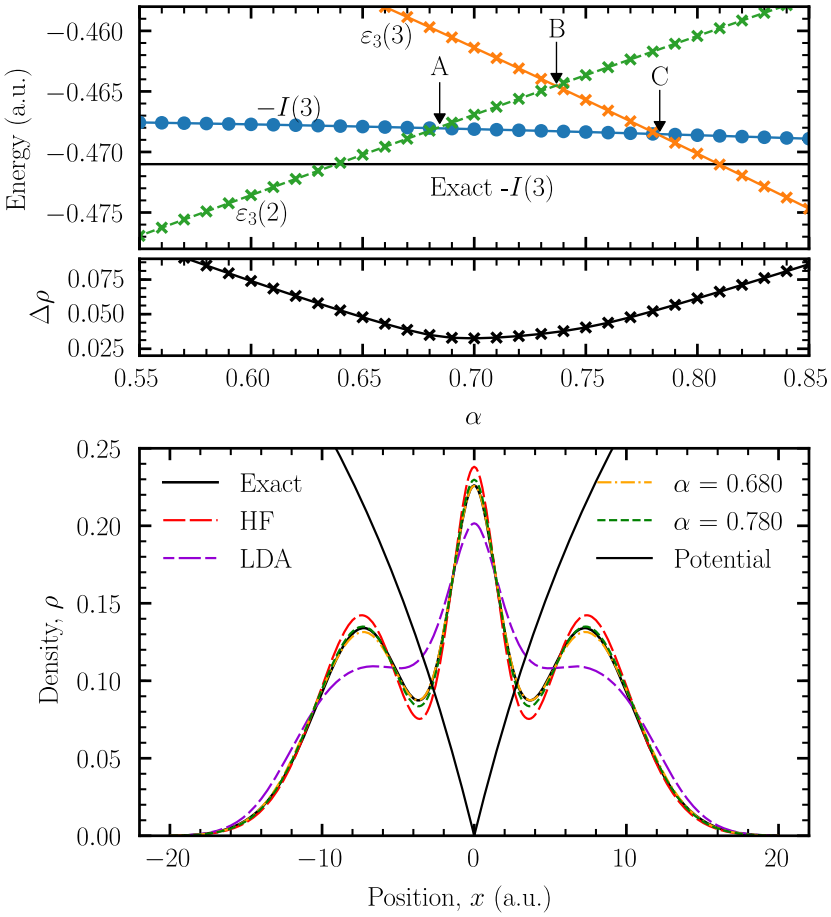

Figure 4.2: As Fig. 4.1, for three electrons in an atom-like external potential ($V_{\text{ext}}(x) = -1/(0.05|x|+1)$). The system is correlated as hf fails to predict the exact density and energy.

ing points lie pleasingly close together, and the density error $\Delta\rho$ (see figure caption) is small in their vicinity. Generally, it is found that densities obtained from α values lying between crossing points (A) and (C) are in excellent agreement with the exact case.

4.5 Correlated systems

Given that both of the underlying functionals usually fail to produce a near-exact density in these systems, one can ask: is a hybrid functional capable of reproducing a near-exact density for any value of α? Results are shown in Fig. 4.2 for three electrons in an atom-like potential. Once again, all three conditions (A)–(C) produce values of α that yield strikingly accurate densities.

Although in the exchange-dominated case crossing points (A) and (C) correspond to an α differing by only one percent, in the correlated system they differ more ($\sim$ 10%). Crucially, however, the density error $\Delta\rho$ corresponding to condition (A) and (C)

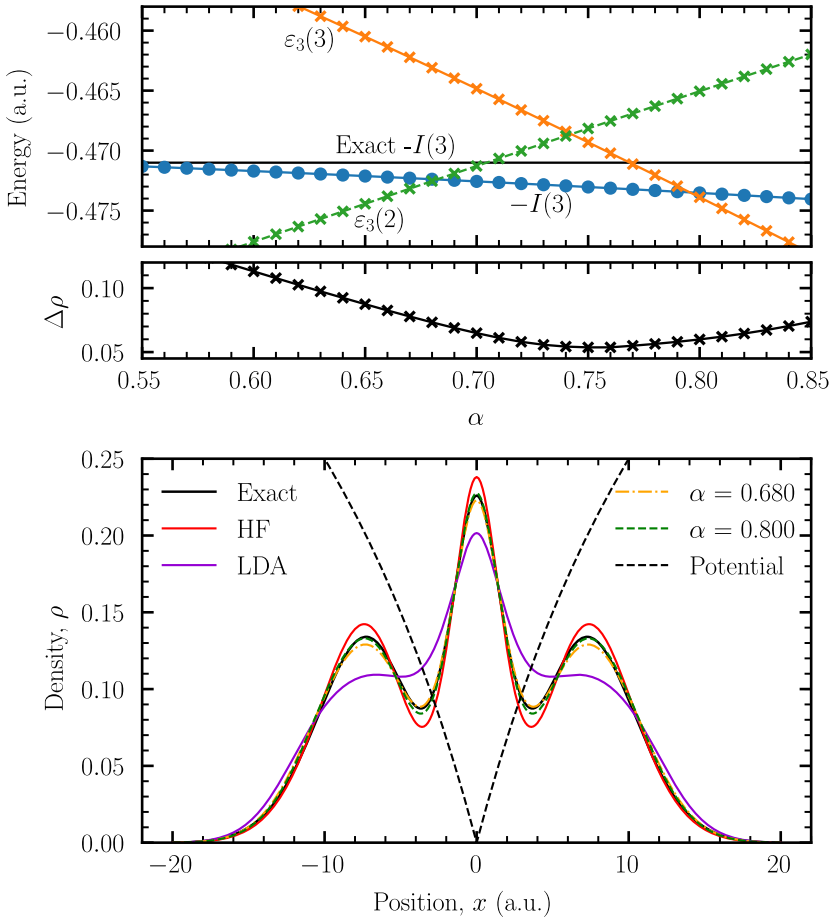

Figure 4.3: As Fig. 4.2, however using the more common exchange-only mixing.

is better than 0.03. Hence, as before, each density corresponding to these conditions is in excellent agreement with the exact. Of note is that condition (A) corresponds to a *slightly* better density than (C), the GKT, for both this correlated system and the exchange-dominated system. The alternative hybrid strategy of mixing only the exchange potentials yields accurate, but slightly inferior, densities, as seen in Fig. 4.3. This shows that the traditional approach of mixing only the exchange potentials and including full LDA correlation may not be the most accurate. Comparing to Fig. 4.2 shows that the density error using full xc mixing is twice as small.

Fractional numbers of electrons

Data for fractional numbers of electrons is shown in Fig. 4.4, where it can be seen that the HF case is exact for values leading up to one electron, however curvature is present for anything larger. This is as expected, as the HF energy and density are exact for one electron systems. Unlike HF, the LDA is inexact for all numbers of electrons. The α values corresponding to conditions (A) and (C) in the atom-like potential follow the

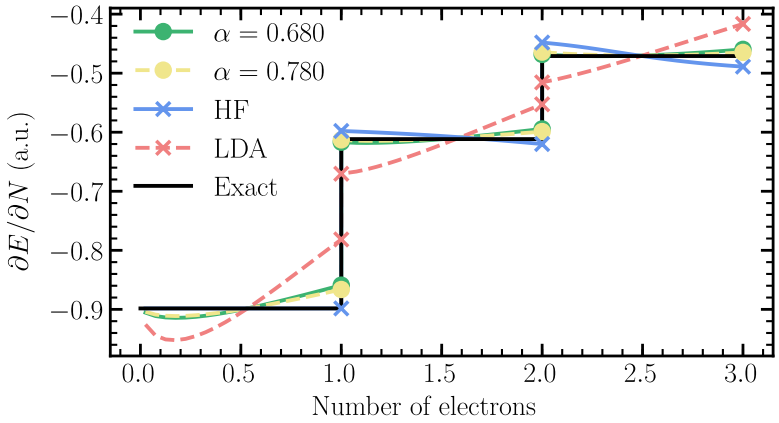

Figure 4.4: The derivative of energy with respect to total number of electrons N, $\partial E/\partial N$, for a number of approximations. The external potential and α values chosen are the same as that of Fig. 4.2. It was verified that the $\partial E/\partial N$ curve lies exactly on that of the HOMO eigenvalue within each approach. Each node at integer numbers of electrons corresponds to the HOMO and LUMO, with the lower energy value being the HOMO.

exact line much more closely than the LDA and HF between two and three electrons, the region where conditions (A)–(C) have been imposed. This suggests that the curvature has indeed been reduced. Comparing the curvature for conditions (A) and (C), it can be seen that the two are comparable to one another.

4.6 Fractional dissociation problem

It is now demonstrated that hybrids are capable of rectifying the fractional charge problem common to many xc approximations for molecular dissociation. Specifically, a system with two separated wells is tested where the usual DFT approximations inaccurately predict the amount of charge present in each well. Fig. 4.5 demonstrates that, when compared with the exact case, the constrained hybrid approach and HF yield near-exact densities. In addition, it can be seen that even for a small fraction of exact exchange ($\alpha = 0.200$), the correct charge in each well is obtained, and hence a large range of values of α yield accurate densities. However, the density has an incorrect shape within each well when an α not corresponding to conditions (A)–(C) is used.

4.7 Errors in the energy

It is shown in Table 4.1 that the accuracy of hybrid functionals for densities is not at the expense of energies. Of particular interest is the quasiparticle energy gap $(I - A)$, which the LDA and HF usually under- and over-estimate, respectively, as well as the values of I and A individually. This establishes contact with the performance of

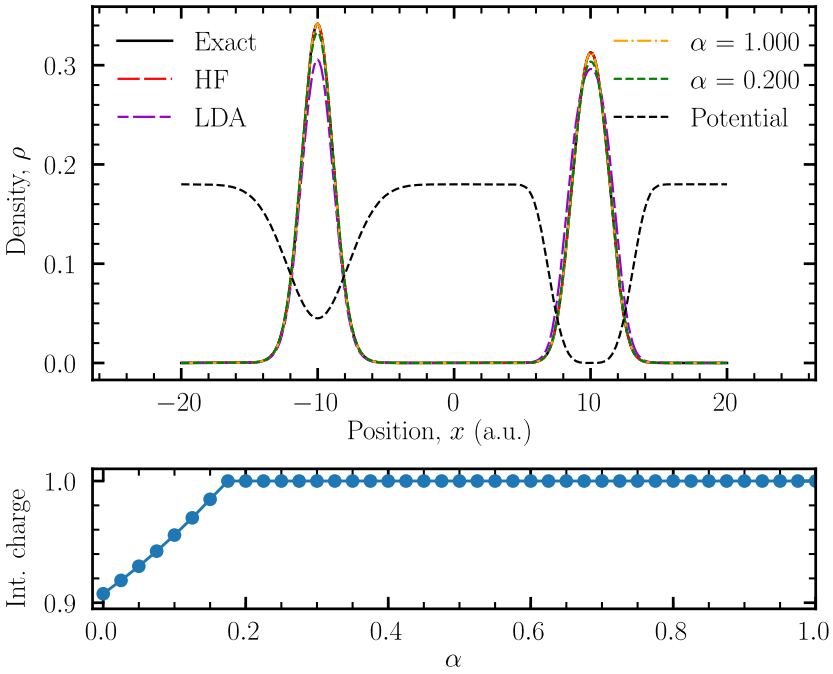

Figure 4.5: (Upper) Densities for various approximations are shown for an exchange-dominated asymmetric double-well potential. The dashed line, illustrating the potential (scaled by 0.15), shows that the two wells are asymmetric. The HF case follows the exact one, placing one electron in each well of a strongly localised system. The LDA predicts that an additional 0.1 electrons are present in the deeper well. The GKT yields $\alpha \approx 1$, effectively HF. The density for $\alpha = 0.2$ is shown, which places the correct charge in each well, but has an incorrect density shape. (Lower) The integrated charge of the left (shallower) well is shown for a range of α values.

Koopmans-compliant hybrids in 3D systems[119,165,169] and suggests that useful quasi-particle energies can be extracted from functionals which also produce an accurate density. The tendency of constrained hybrids to reduce these energy gaps from HF to near-exact levels further supports the idea that this approach is similar to a screened-exchange method.

	Quasiparticle gaps (a.u.)				Ionisation potentials (a.u.)			
	LDA	HF	Hybrid	Exact	LDA	HF	Hybrid	Exact
Harmonic	0.222	0.491	0.472	0.469	-0.761	-0.620	-0.629	-0.628
% error	53%	5%	1%	–	21.2%	1.3%	0.2%	–
Atom-like	0.037	0.172	0.152	0.141	0.551	0.620	0.608	0.612
% error	74%	22%	8%	–	9.9%	1.4%	0.5%	–

Table 4.1: The quasiparticle gap of two-electron systems as extracted from the LDA, HF and hybrid HOMO-LUMO eigenvalue differences, compared to the exact gap calculated from many-body total energies. Gaps are compared for the exchange-dominated (harmonic) and correlated (atom-like) systems. The two-electron IPs are shown for the same systems. These results are constrained using condition (C), though (A) and (B) yield similar results.

	E_{ext}	E_H	E_K	E_{xc}	Sum
Hartree-Fock	0.036	0.132	0.044	0.021	0.233
Hybrid	0.014	0.042	0.019	0.083	0.158
LDA	0.090	0.250	0.024	0.296	0.661

Table 4.2: The absolute error (in eV) in each component of the total energy for Hartree-Fock, hybrid and LDA approaches calculated for two electrons in an atom-like potential. Here condition (C) is used, however conditions (A) and (B) yeild similar results. In each case, these errors are broken down into the individual Kohn-Sham energy components.

In addition to the error in ionisation and eigenvalue gap energies, the error in the total energy is evaluated in Table 4.2. It can be seen that for all components bar E_{xc} the hybrid approach parametrised using condition (C) yields errors in total energy components that are reduced compared to other approaches. The component where hybrid functionals outperform the other approaches most is the Hartree energy, which can be attributed to the improved electron density obtained using this approach. This is made clearer when considering the form of the Hartree energy, which only depends on the electron density. Surprisingly, the error in total energy calculated using the LDA yields the lowest error. This arises from a cancellation of the errors in E_H and E_{xc}, which are roughly equal and opposite in sign.

4.8 Conclusion

Through direct comparison of solutions to the exact many-electron Schrödinger equation, it has been shown that hybrid functionals yield accurate densities *and* quasiparticle energy gaps in both exchange-dominated and correlated systems, if the fraction of exact exchange, α, is chosen using physically justified constraints, such as the generalised Koopmans' theorem. Particularly accurate densities are obtained from a hybrid strategy that mixes LDA correlation, as well as LDA exchange. The three studied constraints are all in close agreement with one another and all yield accurate densities and gaps. In double-well systems, hybrid functionals are found to perform well and are free from the fractional dissociation problem for a large α range. A key perspective is the interpretation of a hybrid method as a simple screened-exchange approximation within many-body perturbation theory.

CHAPTER 5

SELF-TRAPPING POLARONS IN TITANIA POLYMORPHS [1]

The previous chapter illustrated that quantities of interest, such as electron densities and IPs, were well-described in model systems by combining hybrid functionals with the GKT. The focus of this chapter shifts away from model systems and towards 3D periodic crystals, where a similar set of techniques are applied. The set of approaches described in Chapter 3, including ADMM, Coulomb interaction truncation and Cauchy-Schwarz screening, are also used to reduce the computational cost associated with HF exchange, providing an inexpensive hybrid functional. The self-trapped charges introduced in Chapter 1 are discussed in much more detail in this chapter, focusing primarily on six phases of TiO_2: rutile, anatase, brookite, TiO_2(B), TiO_2(H) and TiO_2(R). While a large body of literature relevant to rutile and anatase exists, the remaining phases are not well understood. The approach shown in this chapter is first compared to previous studies on rutile and anatase, and subsequently used to provide a set of predictions for the polaronic properties of the remaining TiO_2 phases.

[1] This chapter represents work that has been published: A. R. Elmaslmane, M. B. Watkins, and K. P. McKenna. 'First-Principles Modeling of Polaron Formation in TiO2 Polymorphs'. *J. Chem. Theory Comput.* 14, 3740–3751. The remaining sections of this chapter have been adapted from the publication.

5.1 Introduction

Charge trapping in semiconductors and insulators has widespread interest in the fields of physics, chemistry and materials science and controls performance for important applications, such as photocatalysis, superconductivity, tribocharging, magnetism and microelectronics.[1–7] The trapping of electrons or holes may take place at pre-existing defects (such as vacancies or impurities) or even in the perfect lattice in some materials (so-called polaronic self-trapping).[8,15,46,94,174,175] Predictively modelling these effects is vital to obtaining a deeper theoretical understanding of charge trapping and to guide materials optimisation for applications. Variants of KS DFT[131,151] are often used to model charge trapping.[15,46] Although DFT is exact in principle, practical calculations require an approximation to the xc potential. The SIE which plagues many xc approximations, such as the local density approximation[152] or (semi-local) GGA,[136,139] results in the spurious delocalisation of localised charges.[28–30,153,176,177] Other methods, such as DFT+U[178] and hybrid DFT,[141] aim to correct this behaviour, however they introduce additional parameters that are often fit to experimental data[142] limiting its predictive power. More sophisticated many-body methods that are in principle predictive, such as Møller-Plesset perturbation theory or the GW approximation, are extremely computationally demanding, especially for modelling polaronic charge trapping where full self-consistency for electrons and ions is essential. For these reasons, finding an inexpensive, predictive and transferable method to model charge trapping in materials is urgently needed.

5.2 Background

DFT+U has been applied successfully to model polarons in a range of metal oxides by introducing a Hubbard-U correction on cations and, optionally, anions. Both the GKT and comparison to experimental data have been used to determine these Hubbard-U values.[179,180] One shortcoming of these approaches is, although a qualitatively correct description of the polarons is obtained, quantitative results are often incorrect. For example, the anatase band gap given by DFT+U (for reasonable parameterisations) is significantly underestimated from the experimental band gap of 3.2 eV.[181] A similar situation is found in other crystals, such as zinc oxide, ceria and rutile-TiO_2.[181] This underestimation naturally yields incorrect polaron charge transition levels (CTLs), hindering comparison to experimental data. The DFT+U approach also predicts a

low (total) dielectric constant in titania,[182,183] meaning electron and ionic screening properties are underestimated. As finite size corrections are inversely proportional to the total dielectric constant ε,[184] an underestimated ε means that they become more of an issue.

Unlike DFT+U, hybrid functionals such as PBE0 or HSE06 can generally yield better-behaved crystal properties. These functionals usually have at least one parameter, the fraction of exact exchange α, which determines the proportion of HF exchange used. Some hybrid functionals make a distinction between short- and long-range exchange (e.g. HSE06), resulting in an additional range-separation parameter ω. The parameterisation of hybrid functionals usually comes in two varieties. A 'one-size-fits-all' approach, where parameters are fixed by comparison to a large set of training data and then used for a variety of different systems. The performance of such hybrids usually varies depending on the choice of system. An alternative approach is to fit hybrids for a specific system, either by using the GKT or by fitting to experimental data.[119,121,156] Fitting parameters to experimental properties such as the band gap may give reasonable results, however the approach is somewhat semi-empirical. Furthermore, the predictive power of such an experimental fitting can also be questioned. In contrast, methods using the GKT to parameterise the hybrid functional have been shown to yield extremely accurate ionisation and band gap energies.[119] However, the degree of transferability for these parameters remains somewhat unknown. A common issue regarding the use of hybrid functionals is the computational cost associated with them, usually around ten times the cost of a standard DFT calculation. To compensate for this, one is forced into using smaller supercells in periodic calculations. This can hinder the quality of a calculation as artificial interactions between periodically repeated defects tend to grow larger with smaller system sizes.[184] Therefore, while hybrid functionals can be superior to DFT+U, they come at a higher computational cost.

Previous experimental and theoretical results for charge trapping in TiO_2 are now discussed. The stability of electron polarons in rutile and anatase is a widely discussed question, with many differing reports from both experimental and theoretical work. While only self-trapped electron and hole polarons are considered in this chapter, numerous studies highlight the effects of reduced cation and oxidised anion species due to dopants, such as Nb, Li and H, and vacancy defects.[47,130,185–189] EPR,[49,190] IR[191] and STM[47] data suggest that small (Holstein) electron polarons can form in rutile localiz-

ing on titanium 3*d* orbitals. IR studies suggest that the polaron CTL lies roughly 0.3 eV below the CBM.[191] Other data from deep level transient spectroscopy, STM and optical absorption places the polaron CTL 0.5 eV,[192] 0.7 eV[47] and 0.8 eV[188] below the CBM, giving the CTL a range of 0.3–0.8 eV below the CBM. Of note is that oxygen vacancies give rise to defects known to lie across a similar, albeit slightly deeper, range of CTLs below the CBM. Yang *et al* determined the experimental thermal activation energy associated with polarons in rutile to be around 20-30 meV, although the mechanism for this was not found.[49] The two main mechanisms for this activation are polaron hopping or excitation into a conduction band-like state.

In rutile, the HSE06 hybrid functional has been shown to satisfy the GKT to within 0.1 eV, making it a suitable choice.[46,193,194] It predicts the electron polaron CTL is around 0.5 eV below the CBM, in good agreement with the 0.3–0.8 eV range (determined through optical measurements), and a quasiparticle gap of 3.4 eV, only a slight overestimate with respect to the optical gap value of 3.0 eV. An alternative parameterisation of the HSE functional (with a smaller range separation parameter, ω = 0.1 Å^{-1}) predicts the polaron CTL is 0.8 eV below the CBM,[195] however the linearity of this functional was not assessed. It is expected that such a parametrisation widens the quasiparticle gap and polaron CTL. Another hybrid functional, B3LYP, significantly overestimates the band gap of rutile[196] resulting in much deeper defect levels than HSE06, and is known to not satisfy the GKT.[197] The polaron trapping energy E_t, the difference between localised and delocalised electron solutions, provides a measure of polaron stability. This energy difference is predicted to be small with HSE06 and DFT+U ($<$ 0.1 eV),[47] and substantially larger with the random phase approximation (RPA) (0.6 eV),[48] although use of a PBE0 geometric configuration from the RPA study should be noted. Polaron migration barriers calculated using HSE (with ω = 0.1 Å^{-1}), the RPA and DFT+U (with U = 10.0 eV) yield 0.3 eV, 0.04–0.14 eV and 0.3 eV.[48,195,198]

In defect-free anatase, electronic states are likely more delocalised than that of rutile as suggested by EPR data.[190,199] This information leads us to believe that, unlike rutile, excess electrons in bulk anatase do not form self-trapped (small) electron polarons, but are much more diffuse. Further confirmation is obtained from STM experiments which find similar results in Nb doped samples.[47] This is also highlighted by conductivity measurements, which find Nb doped samples of anatase exhibit three-to-four orders of magnitude higher conductivity compared to rutile.[200,201] HSE06,[46] the RPA[48] and DFT+U (with U parameterised through the RPA)[47] all indicate that

small electron polarons are unstable in anatase, further indicating why anatase has a higher electron mobility than rutile. On the other hand, B3LYP predicts that electron polarons in anatase are stable.[9,196,202–204] RPA results indicate that a large (Fröhlich) polaron forms when using large system sizes,[48] however experiments indicate that this only occurs in the vicinity of a Nb defect.[47]

The behaviour of hole polarons in rutile and anatase is the opposite of the electron polaron case. Specifically, hole polarons can form in anatase, but not rutile. HSE06 results predict that excess holes in anatase form self-trapped hole polarons (on oxygen $2p$ orbitals), however hole polarons are not predicted to form in rutile (i.e. they are delocalised). HSE06 predictions suggests that hole polaron levels in anatase are a deep acceptor state, lying roughly 1.3 eV above the VBM with a trapping energy of 0.2 eV. EPR experiments suggest trapped holes can form in anatase.[205] The concentrations of trapped holes (O^- species) were higher than those of trapped electrons (Ti^{3+} species), primarily because most electrons were delocalised in the conduction band, and hence EPR silent. In contrast, rutile samples showed larger EPR signals attributed to electron trapping than hole trapping,[206] in agreement with HSE06 results. Other EPR experiments somewhat confuse the picture as it is suggested that hole trapping in rutile is observed.[207] There have been very few studies of polaron formation in phases of TiO_2 other than rutile and anatase. This includes an experimental investigation into polaron self-trapping in brookite,[190] which finds that anatase and brookite have similar trapping properties. There has also been a theoretical study of polaron formation using DFT+U which finds charge strongly localised on Ti sites in Li-doped TiO_2(B).[208]

5.3 Methods

The hybrid DFT implementation laid out in previous chapters (Chapter 2 and Chapter 3) is used to perform all calculations.[144,148,209–211] Gaussian type orbitals are mapped onto CP2K's multi-grid solver with a relative cutoff of 60 Ry. Five multi-grids, with the finest levels of grid having a cutoff of 600 Ry, are used. The primary oxygen and titanium basis sets used are of triple-ζ quality with valence and polarisation exponents.[212] Norm-conserving GTH pseudopotentials available with CP2K[213] are used. As CP2K only samples the Γ-point in reciprocal space, crystal properties must be converged with respect to system size as opposed to number of k-points. For all phases,

it is found that supercell dimensions are well-converged at around 18 Å. [2] Of note is that using larger system sizes results in band folding, which increases the number of bands sampled at each k-point in the Brillouin zone. This is equivalent to increasing the number of k-points sampled, meaning that the calculation remains accurate despite only sampling the Γ-point. Furthermore, using, using larger system sizes instead of a greater number of k-points reduces the size of the interaction between charged supercell images.

The exchange part of the hybrid functional employed (tr-PBEα) is described in Chapter 2 (Section 2.5.2). In both the short range ($r \leq R_c$) and long range ($r > R_c$) regions the full PBE correlation energy is used. This truncation is performed to reduce computational costs associated with HF exchange. However, predicted properties (e.g. lattice constants, band gaps, polaron trapping energies and charge transition levels) converge rapidly with respect to increasing R_c. It is found that crystal properties, such as lattice constants and the band gap, are well-converged at R_c = 6.0 Å in all phases. In addition to HF exchange truncation in the functional, `CP2K` has other tools to alleviate computational cost. One such method is the ADMM[148] (described in Chapter 3), which approximates exchange integrals by mapping orbitals onto smaller, more localised, basis sets. These significantly improve the speed of HF exchange calculations, meaning the approach can be used to model larger systems with reasonable computational cost. In later sections, variations in properties induced by changes in ADMM basis sets are considered to establish any error these may introduce. `CP2K` also takes advantage of other exchange integral screening methods which reduce computational cost, all of which are described in Chapter 3.

The procedure for determining the proportion of HF exchange for electron or hole polarons is as follows. Cell vectors are initially optimised for a large defect-free supercell. As hybrid functional `CP2K` calculations use the Γ-point only, this must be done using the supercell rather than the unit cell. Following this a distortion is applied to the geometry, pushing ions within 2.2 Å of a chosen ion radially outward by 0.1 Å. This distortion creates a precursor potential well for localisation of the polaron on a given ion. An additional electron (or hole) is then added to the supercell and atoms are relaxed until the magnitude of forces on atoms is less than 0.01 eV/Å. The final calculation (vertically) removes the added electron (or hole) from the system, keeping ions in the polaronic configuration. The above calculations yield all the necessary en-

[2] Graphs depicting this are illustrated in Appendix B.

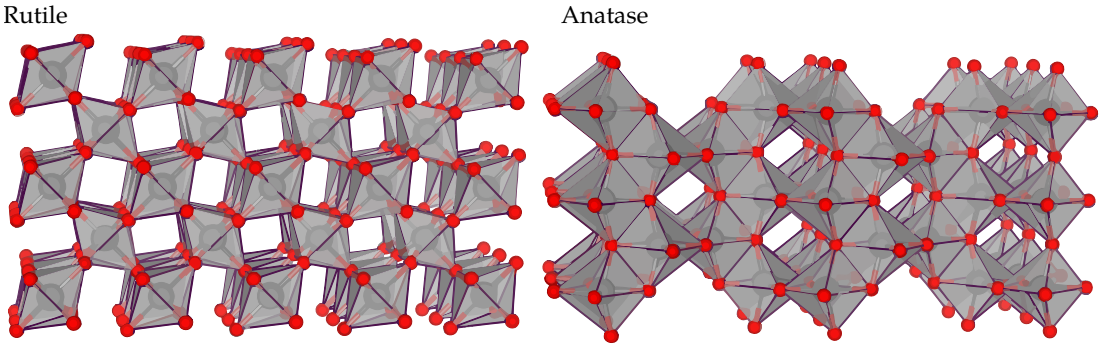

Figure 5.1: The crystal structures of rutile and anatase TiO_2. All sites of the same species in each phase are equivalent by symmetry. Oxygen atoms are shown in red whilst titanium atoms are shown in grey. Figure taken from *J. Chem. Theory Comput.* 14, 3740–3751.

ergies to test the GKT (Eqn. 1.4). With half-percent granularity, the range of α values which satisfy the condition to within 0.05 eV are identified. The degree to which the GKT is satisfied, also called the nonlinearity, is denoted ξ:

$$\xi = \begin{cases} E(N+1) - E(N) - \varepsilon_{N+1}(N+1), & \text{for electrons} \\ E(N) - E(N-1) - \varepsilon_N(N), & \text{for holes.} \end{cases} \tag{5.1}$$

A limit of $|\xi| \leq 0.05$ eV is set as physical properties of the crystal can change quickly with small changes in nonlinearity. Operating within this limit causes changes in the GKS gap of around 0.2 eV over the range $-0.05 \leq \xi \leq 0.05$ eV.

5.4 Anatase and rutile

In this section the approach described above is applied to model electron and hole polarons in anatase and rutile. The structures of the two crystals are shown in Fig. 5.1. Initially titanium FIT9 and oxygen pFIT3 (FIT9/pFIT3) ADMM basis sets are used, although alternatives are tested below to demonstrate basis set quality. $5 \times 5 \times 2$ and $4 \times 4 \times 6$ expansions of the anatase and rutile unit cells are used, with 600 and 576 atoms in the supercell.

In Fig. 5.2 α is tuned to satisfy the GKT for electron and hole polarons in rutile and anatase. It can be seen that the anatase electron polaron case does not satisfy the GKT in the range $0.25 \leq \alpha \leq 0.40$. Attempting to reduce α to a lower value ($\alpha = 0.225$) results in a delocalised conduction band state, suggesting that the electron polaron is not stable at the Koopmans-compliant value of α. Probing the point where the electron delocalises closely, $\xi = 0.29$ eV at $\alpha = 0.230$. A small move to $\alpha = 0.225$ results in a

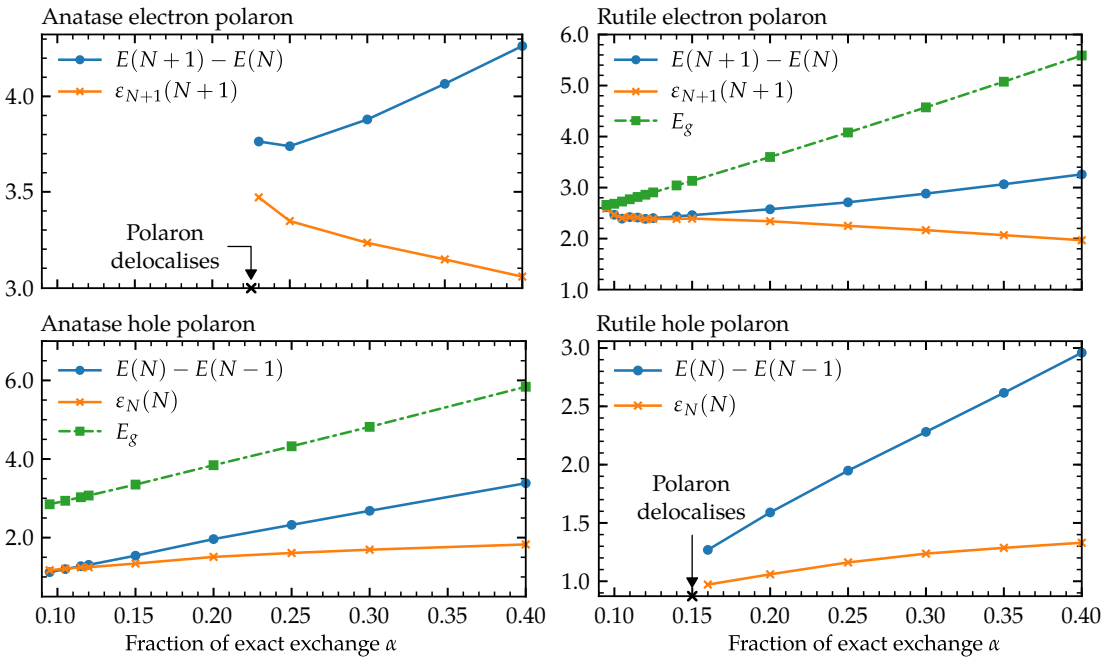

Figure 5.2: The variation in charge transition levels, (defect-free) band gap E_g and highest occupied eigenvalue with α for the anatase (left) and rutile (right) electron (upper) and hole (lower) polarons. With half-percent granularity, it is found that Koopmans theorem is satisfied at $\alpha = 0.105$ and $\alpha = 0.115$ for the anatase hole and rutile electron cases. Figure taken from *J. Chem. Theory Comput.* 14, 3740–3751.

completely delocalised electron, meaning that this method of constraining α predicts that even a metastable electron polaron configuration is not possible. Of note is that fully self-consistent optimisation of the geometry for a given α is essential for accurate prediction of polaron properties. For example, if one fixes the polaron geometry to that obtained with α = 0.25, the electron polaron remains localised even when α is reduced to 0.10, contrary to the behavior seen when the geometry is self-consistently optimised.

Unlike the electron polaron case, the anatase hole polaron does satisfy the GKT at α = 0.105 with $\xi \approx 0$ eV. The hole polaron spin density is illustrated in Fig. 5.3, where it can be seen that the hole polaron is largely localised on a single oxygen ion. The trapping energy E_t is defined as the energy difference between localised and delocalised states, which for the hole polaron is -0.21 eV. While it would be wrong to advocate use of a non-self-consistent geometry for prediction of polaron properties, for deep traps, such as the hole polaron in anatase, there is little difference in polaron geometry between α = 0.25 and α = 0.105.

In rutile, electron polarons satisfy the GKT whilst hole polarons do not, as shown in Fig. 5.2. As mentioned in Section 5.2, this result is also in agreement with findings us-

ing HSE06 by Deák *et al.*[46] However, unlike the anatase case, a large number of α values satisfy the GKT. As the value of α decreases across range where $-0.05 \leq \xi \leq 0.05$ eV, the electron polaron trapping energy increases from a value suggesting the polaron is mildly stable ($E_t = -0.07$ eV at $\alpha = 0.14$) to metastable ($E_t = 0.01$ eV at $\alpha = 0.10$). The localised solution cannot be found at $\alpha = 0.095$. Ultimately, as the polaronic state is capable of satisfying the GKT it may be that the polaron is stable. A low trapping energy suggests there is little difference energetically between localised and delocalised states, in agreement with HSE06 and DFT+U results.[46,47] The small trapping energy may explain why these polarons rapidly delocalise in EPR experiments when temperature is raised.[206] The spin density for the electron polaron is illustrated in Fig. 5.3, where it is clear that, unlike the hole polaron in anatase, the electron polaron is much more diffuse.

In the region $0.100 < \alpha < 0.140$ the GKT is satisfied to less than 0.05 eV for the rutile electron polaron. Across this range, the polaron CTL moves from 0.22 eV to 0.65 eV below the CBM. The largest value of α satisfying $|\xi| < 0.01$ eV ($\alpha = 0.115$) is selected as the optimal value for evaluation of polaron properties. In this case the electron polaron CTL is predicted to lie 0.40 eV below the CBM. Experimental work puts the CTL at 0.3–0.8 eV below CBM as previously mentioned, which is in good agreement with the predicted result. The prediction is also in good agreement with previous calculations performed using the HSE06 functional which places the polaron CTL at 0.5 eV below the CBM.[46]

The discussion in Section 5.2 highlighted that TiO_2 band gaps from B3LYP and DFT+U are generally over- and under-estimated, by around 1 eV.[181,196] In contrast, the HSE06 functional overestimates the gap only slightly by 0.2 eV.[46] Surprisingly, the outlined procedure yields defect-free band gaps in both rutile and anatase that deviate only by around 6% from the experimental values of 3.0 eV and 3.2 eV (with $\alpha = 0.115$ and $\alpha = 0.105$) respectively. In both cases, the gap is underestimated by around 0.2 eV, as seen in Table 5.1. It should be stressed that results have not been optimised for the band gap. However, it seems that applying the GKT to these localised states results in a band gap very close to the experimental value. If the upper limit of $|\xi|$ tolerance is used, errors in the band gaps of rutile and anatase shrink further. Lattice constants are also in good agreement with experiment, with errors of 0.3% (vector a) and 1.1% (vector c) in anatase and 0.7% (vector a) and 0.1% (vector c) in rutile. This is somewhat reassuring, as it suggests that the method is able to accurately describe

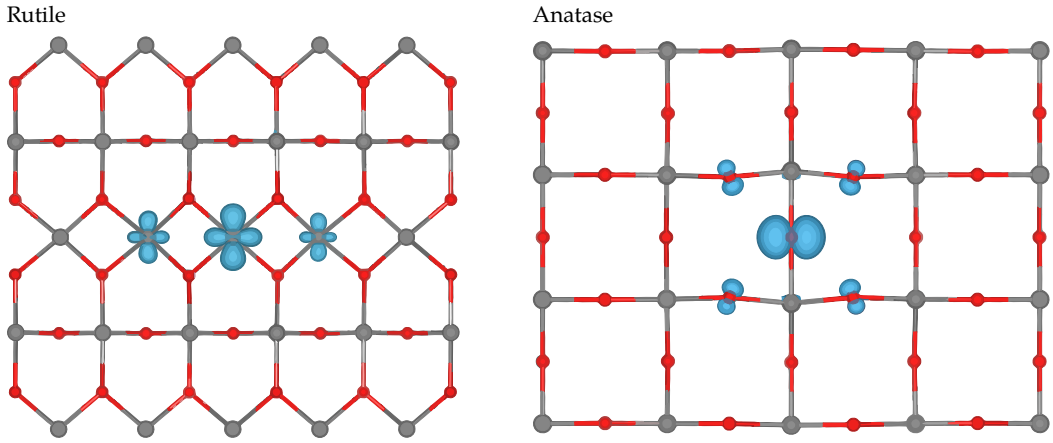

Figure 5.3: Spin density for a hole polaron in anatase and an electron polaron in rutile (isosurface value 0.01 a_0^{-3}). As the polarons lie largely on a single crystallographic plane, only a slice of the crystal is shown. Titanium and oxygen ions are represented by red and grey spheres respectively. The electron polaron, localised on a titanium 3*d* orbital, has a much more diffuse nature than the hole polaron, localised on an oxygen 2*p* orbital. In rutile, the visible isosurface extends around 5.8 Å, while in anatase the distance is 5.3 Å. Bader spin-density analysis shows that the central atom in rutile contains 50% of an electron, while the central atom in anatase contains 75% of a hole. Figure taken from *J. Chem. Theory Comput.* 14, 3740–3751.

both the polaronic and bulk properties of these crystals using only the GKT. This is also in line with results from finite systems, which report similar findings regarding gaps between the HOMO and the LUMO.[119]

The optimised values of α which satisfy the GKT for anatase and rutile, 0.105 and 0.115, are very close together. In rutile, one could satisfy the GKT with $|\xi| < 0.05$ eV using $0.100 \leq \alpha \leq 0.140$ and for anatase one can use $0.095 \leq \alpha \leq 0.115$, meaning there is a large overlap of α values. This means that one could feasibly choose the same value of α to model both electron and hole polarons in these phases. This transferability in α values is somewhat unexpected, however it allows for more than one phase to be described by the same functional. In later sections, it is shown that this transferability extends to other phases as well.

As the primary basis sets used are all of good quality, it is expected that the ADMM basis will likely be a key source of inaccuracy in subsequent calculations. To investigate the error introduced, changes in bulk and polaron properties are explored with differing titanium and oxygen basis sets. Table 5.1 illustrates that the choice of ADMM basis set has little effect on both the trapping energy, local magnetic moment and charge transition energy for both electron and hole polarons across both phases. Magnetic moments are calculated via Bader analysis of the spin density.[214–216] The choice of basis set has no effect on the optimised value of α or the nonlinearity. Similarly, bulk quantities change very little with increasing ADMM basis set quality. Using con-

Ti ADMM basis set / **O ADMM basis set**	FIT9/ cFIT3	FIT9/ FIT3	FIT9/ cpFIT3	FIT9/ pFIT3	FIT10/ pFIT3	FIT11/ pFIT3	Exp.
Anatase hole							
Lattice vector a	3.784	3.783	3.794	3.793	3.794	3.793	3.782[a]
Lattice vector c	9.708	9.701	9.623	9.617	9.605	9.605	9.502[a]
GKS gap E_g	3.03	3.00	2.94	2.94	2.93	2.93	3.23[b]
Nonlinearity ξ	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	-
Trapping energy E_t	-0.21	-0.21	-0.21	-0.21	-0.20	-0.20	-
Bader magnetisation	0.76	0.76	0.75	0.75	0.75	0.75	-
CTL - VBM	1.23	1.24	1.19	1.20	1.21	1.18	-
Rutile electron							
Lattice vector a	4.618	4.617	4.620	4.618	4.618	4.618	4.587[a]
Lattice vector c	2.961	2.959	2.960	2.958	2.958	2.957	2.954[a]
GKS gap E_g	2.82	2.79	2.84	2.82	2.82	2.81	3.03[c]
Nonlinearity ξ	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	-
Trapping energy E_t	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-
Bader magnetisation	0.48	0.49	0.49	0.48	0.48	0.49	-
CTL - CBM	-0.41	-0.38	-0.44	-0.40	-0.41	-0.42	-

All energies and lengths shown are in electron-volts and Angstroms.
[a] Ref 217 , [b] Ref 50 , [c] Ref 51

Table 5.1: Bulk and polaron properties are compared using different auxiliary oxygen basis sets in anatase and rutile. As only anatase holes and rutile electrons are predicted to be stable, only these two polarons are shown. The titanium basis FIT9 has 9 basis functions (FIT9), comprising of three *s*-, *p*- and *d*-functions. Titanium basis sets FIT10 and FIT11 have 10 and 11 basis functions, whilst the oxygen basis sets cFIT3, FIT3, cpFIT3 and pFIT3 have 4, 6, 5 and 7 basis functions. For each combination of the basis sets, the value of α which minimises nonlinearity ξ is found. It can be seen that the choice of basis set does not produce a difference in the optimised α. For anatase and rutile, the optimised α values are 0.105 and 0.115.

tracted basis sets, labeled with a lower case letter c, does not affect polaron or bulk properties. To this end, using FIT9/pFIT3 provides the best balance between accuracy and computational efficiency.

In the preceding calculations R_c was converged with respect to the predicted bulk lattice constant and band gap, resulting in the value $R_c = 6$ Å. This parameterisation is now compared with lower values of R_c. By varying R_c, the optimal value of α that satisfies the GKT (for the selected value of R_c) changes. As R_c increases, the amount of exchange energy increases, resulting in a slightly more concave functional (see Fig. 1.2). It was previously shown for rutile electron polarons that the using $R_c = 6$ Å with $\alpha = 0.115$ satisfies the GKT with nonlinearity $\xi < 0.05$ eV. For the values $R_c = 4$ Å and $R_c = 2$ Å it is found that $\alpha = 0.115$ and $\alpha = 0.2$ satisfy the GKT to within 0.05 eV. The α parameter effectively remains the same when going from $R_c = 6$ Å to $R_c = 4$ Å, suggesting that restricting exchange interactions over 4 Å yields similar results to 6 Å.

However, for R_c = 2 Å predicted polaron properties are very different. Taking R_c = 6 Å yields a band gap of 2.8 eV, with the electron polaron CTL being 0.4 eV below the CBM. In contrast, R_c = 2 Å yields a band gap of 3.1 eV and a polaron CTL that is roughly 0.9 eV below the CBM, much higher than that reported experimentally. This suggests that functionals with lower R_c values have a tendency to predict more stable polaronic states. A similar conclusion can be drawn from the polaron trapping energy, where using R_c = 2 Å and R_c = 4 Å results in $E_t = -0.14$ eV and $E_t = -0.02$ eV respectively. Comparing the range-separation used in this work and the popular HSE functional,[142] HSE more smoothly weighs short- and long-range Coulomb interactions (using the complimentary error and error functions). Short-range Coulomb interactions typically become negligible a length of $2/\omega$. Values of ω are usually around 0.2 Å^{-1},[143] meaning short-range exchange is negligible beyond 10 Å.

The nonlinearity ξ for polaronic states also depends on the supercell size. This was studied for the following supercell sizes (number of atoms in brackets): $3 \times 3 \times 5$ (270), $4 \times 4 \times 6$ (576) and $5 \times 5 \times 8$ (1200). Using the α optimised for the 576 atom system, nonlinearity was calculated for 270 and 1200 atom systems. It is found that ξ changed by 0.04 eV and 0.01 eV when moving from 270 to 576 and 576 to 1200 atom systems respectively. The next reasonable (cubic) supercell size to test for nonlinearity is far too large (1944 atoms). However, as system size (576 atoms) has a nonlinearity only different by 0.01 eV to much larger systems (1200 atoms), it provides a good balance between computational time and accuracy.

5.5 Brookite, TiO_2(H), TiO_2(R) and TiO_2(B)

Having benchmarked calculations on rutile and anatase, the model is now applied to model polaron formation in four other phases of titanium dioxide: brookite, TiO_2(H) (hollandite), TiO_2(R) (ramsdellite) and TiO_2(B). Each of these possess a number of lattice sites that are inequivalent by symmetry (see Fig. 5.4). Lattice sites are denoted using shorthand notation: e.g. O_a^{Br} and O_b^{Br} label oxygen sites a and b in the brookite crystal. To avoid confusion, rutile and anatase are abbreviated to Ru and An, while TiO_2(H), TiO_2(R) and TiO_2(B) are abbreviated as H, R and B.

For all localised polarons, the range of α values which satisfy $|\xi| < 0.05$ eV is determined (shown in Fig. 5.5). Across a large number of sites, there is significant overlap in the range of acceptable α values. This means that it is feasible to model all phases and sites using the same α whilst still satisfying the GKT. It can be seen

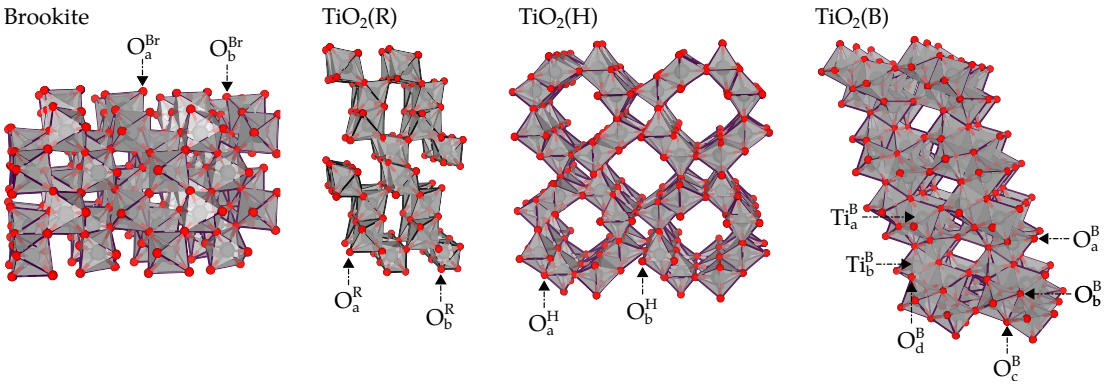

Figure 5.4: Crystal structure of brookite, TiO_2(H), TiO_2(R) and TiO_2(B). Titanium and oxygen ions are represented by red and grey spheres respectively and Ti-centered polyhedra are highlighted. Lattice sites which are inequivalent by symmetry are labeled. Figure taken from *J. Chem. Theory Comput.* 14, 3740–3751.

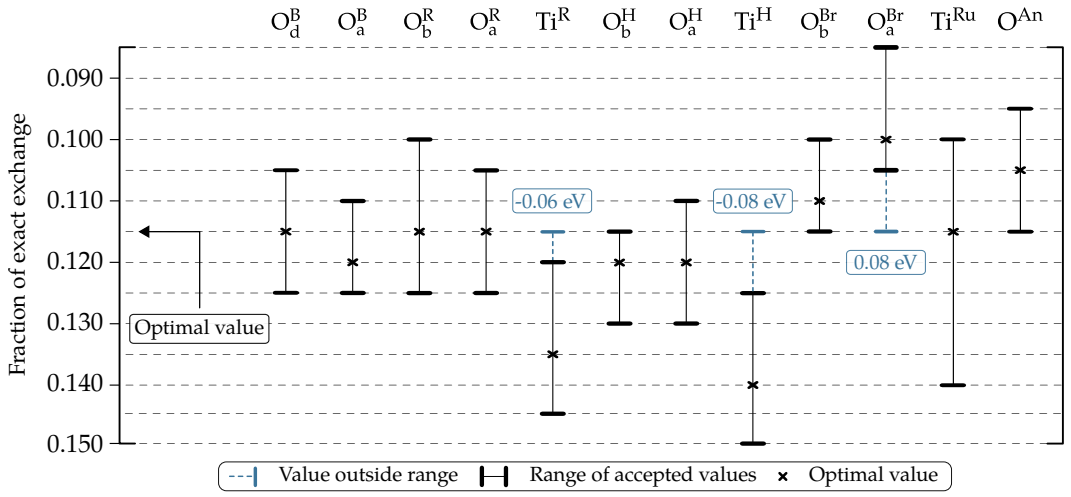

Figure 5.5: The range of α values which satisfy the generalised Koopmans' theorem to within 0.05 eV for polarons localised on different lattice sites. For each site, the point which has the smallest ζ is marked with a cross. For $\alpha = 0.115$, $|\zeta| < 0.08$ eV across all sites. Figure taken from *J. Chem. Theory Comput.* 14, 3740–3751.

$\alpha = 0.115$ ensures $|\zeta| < 0.08$ eV across all phases and sites. In the following subsections the predicted properties of polarons in each phase are describe in detail. The polaron properties of each phase are evaluated using an α that minimises nonlinearity for that phase. This allows us to compare CTLs of different polaron sites using the same functional (and hence the same lattice constant and band gap). These α values are shown in Table 5.2 alongside the predicted bulk properties. If instead, one uses $\alpha = 0.115$ to model all phases, good quantitative agreement is seen. In Table 5.2 the stability of each of the phases relative to anatase is shown, where it can be seen that the energy differences between phases are small. For phases where data is available (rutile, anatase and brookite), the energy differences are in line with previous (local)

Phase	Anatase	Rutile	Brookite	TiO_2(H)	TiO_2(R)	TiO_2(B)
α	0.105	0.115	0.105	0.130	0.125	0.120
E_g (eV)	2.94	2.82	3.20	3.76	3.65	3.85
Δ (eV)	0.000	0.031	0.017	0.054	0.054	-0.001
a (Å)	3.794	4.618	9.228	10.236	4.950	12.236
b (Å)	-	-	5.479	-	9.425	3.746
c (Å)	9.617	2.958	5.151	2.968	2.968	6.575
β (°)	90	90	90	90	90	107.0

Phase	Anatase[a]	Rutile[a]	Brookite[b]	TiO_2(H)[c]	TiO_2(R)[d]	TiO_2(B)[e]
a (Å)	3.782	4.587	9.184	10.164	4.902	12.197
b (Å)	-	-	5.447	-	9.459	3.754
c (Å)	9.502	2.954	5.145	2.963	2.956	6.654
β (°)	90	90	90	90	90	107.16

[a] Ref 217 [b] Ref 222 [c] Ref 223 [d] Ref 224 [e] Ref 225

Table 5.2: *(Upper)* Optimised α value for each TiO_2 phase with corresponding bulk lattice constants and band gaps. The internal energy Δ (per TiO_2 formula unit) is given for each of the phases, relative to anatase. *(lower)* Experimentally determined lattice parameters.

DFT calculations.[218,219] Diffusion QMC results yield, ignoring the effects of lattice dynamics, a similar energetic ordering.[220,221]

5.5.1 Polarons in brookite

It is found that brookite is in many ways similar to anatase. Specifically, the GKT predicts the formation of hole polarons, but not electron polarons. This is in good agreement with EPR data, which also suggests that the polaronic properties of the two crystals are quite similar.[190] The two inequivalent O sites in brookite, differing mainly by bond angles to neighboring Ti, are both capable of forming hole polarons. Both sites have (+/0) CTLs differing only by 0.02 eV at the optimal α for brookite (α = 0.105), as seen in Table 5.3. However, their trapping energies differ by around 0.1 eV, making O_a^{Br} the most stable trapping site in brookite. The predicted band gap of brookite is 3.2 eV, consistent with the experimentally determined optical values of 3.1-3.4 eV.[226]

5.5.2 Polarons in TiO_2(H) and TiO_2(R)

The phases TiO_2(H) and TiO_2(R) are now discussed together, as predictions suggest they exhibit similar charge trapping properties. TiO_2(H) contains two unique oxygen sites and one unique titanium site. It is found that the best α for all sites in this phase

is α = 0.130, as seen in Table 5.2. A small hole polaron, similar to that formed in anatase, is predicted on one oxygen site (O_b^H). On the second oxygen site (O_a^H), a multi-site small polaron, or molecular polaron, is predicted to form, previously unseen in titania. The multi-site polaron, illustrated in Fig. 5.6, is predicted to localise over two ions, with almost equal charge on both ions. Comparing the single- and multi-site hole polarons, it is found the (+/0) CTL from the multi-site polaron lies deeper in the band gap (with respect to the VBM) by 0.1 eV. Furthermore, the multi-site polaron is also more energetically stable, with a trapping energy 0.1 eV lower than the single-site polaron. On the only unique titanium site in TiO_2(H), a multi-site electron polaron can be found (illustrated in Fig. 5.6). The (-/0) CTL of this polaron is twice as deep as the electron polaron found in rutile, lying 0.8 eV below the CBM.

The two oxygen (O_a^H) ions on which the multi-site hole polaron forms (see Fig. 5.6) move closer together by 0.3 Å compared with their equilibrium position. In contrast, two nearby titanium ions (shared by the two oxygen ions on which the multi-site polaron resides) move apart by 0.3 Å. The single-site hole polaron (O_b^H) results in an asymmetric distortion of nearby titanium ions, with bond lengths increasing by 0.1–0.3 Å from their defect-free positions. With geometries compared, it is clear that the nature of polarons these two sites is different. Despite large changes in the geometry compared to the equilibrium positions, the multi-site polaron remains more energetically stable than the single-site one. A similar behaviour is found regarding multi-site electron polarons, where two titanium ions involved to move together by 0.2 Å, whilst two surrounding oxygen ions are pushed apart by 0.2 Å.

Single- and multi-site polarons are predicted to form in TiO_2(R), on sites O_a^R and O_b^R respectively. The geometric distortions of both polarons are also quite similar to the TiO_2(H) case. Of note is that in both phases multi-site polarons are more stable than single-site polarons, as seen in Table 5.3.

One distinguishing factor between single-site and multi-site polaron sites in the TiO_2(H) and TiO_2(R) phases is the bonding. The multi-site electron polarons form over two cations ions that bond with two of the same anions, shown in Fig. 5.6. Anions which display the same double-sharing behaviour also form multi-site hole polarons. The multi-site polarons reside on sites of the same symmetry.

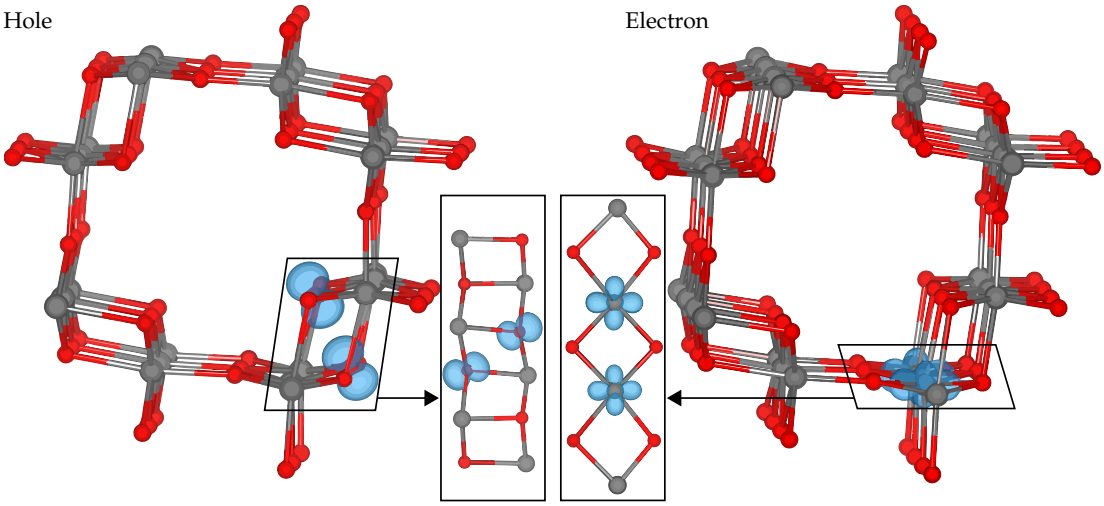

Figure 5.6: Molecular (two-site) electron and hole polarons in TiO_2(H). The polaron charge are shared over two equivalent sites, roughly 0.5e on each site. It is found that molecular polarons require the double-sharing of titanium ions by oxygen ions (or vice versa) to form. The double sharing behaviour for both cases is shown in the boxed figures. A spin density isosurface of 0.02 a_0^{-3} was used to generate the above figures. Figure taken from *J. Chem. Theory Comput.* 14, 3740–3751.

5.5.3 Polarons in TiO_2(B)

Of the four unique oxygen sites in TiO_2(B), only polarons on two sites (O_a^B and O_d^B) remain stable across the range $-0.05 < \xi < 0.05$, with one site (O_d^B) being the most stable trapping site predicted in this work. The site O_a^B is under-coordinated with two bonds to neighboring titanium ions, whilst O_d^B has the same coordination (three) as other phases. Although the GKT allows for polaron formation on both sites, site O_d^B is 0.1 eV more stable than site O_a^B. The site O_c^B is nearly capable of satisfying the GKT with $\xi = 0.06$ eV at $\alpha = 0.135$, however the site O_d^B is 0.47 eV more stable than O_c^B, implying that polarons are less likely to trap on O_c^B than O_d^B. The final site, O_b^B is four-coordinated and the GKT predicts a localised hole polaron is not possible on this site.

The electron polaron case in TiO_2(B) is slightly more complicated than the hole polaron case. While the sites Ti_a^B and Ti_b^B are both capable of forming small polarons that satisfy the GKT to within 0.06 eV, the polarons begin to delocalise over the range $0 < \xi \leq 0.05$ eV. Small polarons on Ti_a^B and Ti_b^B delocalise at $\alpha = 0.145$ and $\alpha = 0.140$ respectively, with the more delocalised states having similar total energies to the small polarons. At $\alpha = 0.145$, the small electron polaron on site Ti_b^B has a (-/0) CTL of 0.40 eV and a trapping energy -0.04 eV. As electron polarons are capable of satisfying the GKT over a small range, it plausible that they form in TiO_2(B). However, the small trapping

Site	Polaron type	CTL (eV)	Magnetisation	E_t (eV)
Ti^{Ru}	Single-site	-0.40	0.45	-0.02
Ti^{H}	Multi-site	-0.82	0.44, 0.37	-0.14
Ti^{R}	Multi-site	-0.82	0.45, 0.37	-0.35
O^{An}	Single-site	1.20	0.75	-0.20
O_a^{Br}	Single-site	1.11	0.72	-0.25
O_b^{Br}	Single-site	1.09	0.73	-0.16
O_a^{H}	Multi-site	1.71	0.41, 0.47	-0.42
O_b^{H}	Single-site	1.59	0.80	-0.34
O_a^{R}	Single-site	1.32	0.81	-0.31
O_b^{R}	Multi-site	1.48	0.42, 0.45	-0.43
O_a^{B}	Single-site	1.62	0.70	-0.42
O_d^{B}	Single-site	1.60	0.74	-0.52

Table 5.3: Charge transition levels, local magnetic moment (from Hirshfeld spin density analysis) and polaron trapping energy for all polarons studied in this work. Each of the systems shown above is in a spin-doublet formation. For electron polarons (on Ti sites) the (-/0) transition is shown in reference to the CBM, whilst hole polarons (on O sites) have (+/0) charge transitions that are shown in reference to the VBM. For multi-site polarons, two magnetisation numbers are shown, corresponding to the two sites over which the multi-site polaron resides. The four multi-site polarons break magnetisation symmetry, which may be an artefact of the Hirshfeld spin density analysis.

energies indicate that they may be less likely to form.

5.6 Comparison of polarons across phases

The trapping energies, CTLs and Hirshfeld spin density analysis of each polaron is shown in Table 5.3. These values are calculated using the ideal value of α for each phase (shown in Table 5.2). It can be seen (from the trapping energy) that site O_d^B is the most stable site of all sites explored. The multi-site hole polarons formed in TiO_2(H) (O_a^H) and TiO_2(R) (O_b^R) are both much more stable than their single-site polaron counterparts. In addition, multi-site electron polarons (Ti^H and Ti^R) are much more stable than the single-site electron polaron found in rutile. The deepest (-/0) and (+/0) CTLs correspond to multi-site polarons in TiO_2(H) and TiO_2(R). Sites O_a^{Br} and O_b^{Br} both have differing trapping energies, yet the CTLs are quite similar. The same can be said of sites O_a^B and O_d^B. In contrast, CTLs from multi-site polarons are distinctly split from those of small polarons, as seen in phases TiO_2(H) and TiO_2(R). As previously mentioned, the transferable and phase-specific α values are in good quantitative agreement.

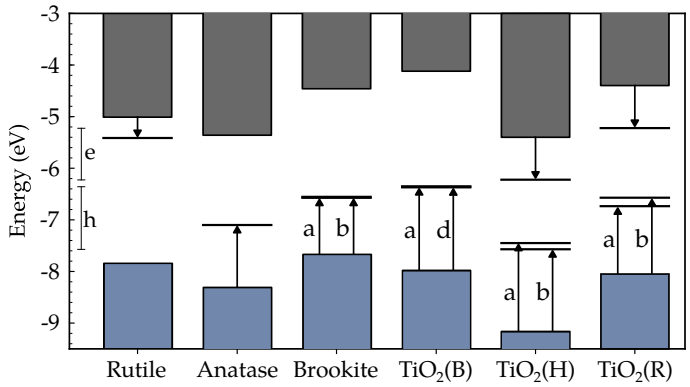

Figure 5.7: Alignment of polaron CTLs with respect to a common vacuum energy. The polaron CTLs and band gaps are computed using phase-specific optimised α values as shown in Table 5.2. For clarity, only the subscript letter index for each of the polaron sites is shown. Electron and hole polaron energy ranges with respect to vacuum are labeled as e and h. Figure taken from *J. Chem. Theory Comput.* 14, 3740–3751.

To compare polaron CTLs across different phases, the band alignment of the phases with respect to each other is required. Here the VBM of all phases is aligned using the ionisation energies obtained in a previous theoretical study,[227] which calculates these energies through a hybrid quantum mechanical/molecular mechanical embedded cluster approach. This allows us to compare the calculated band gaps and polaron CTLs for all phases (Table 5.3) against a common vacuum level reference (Fig. 5.7). The deepest electron and hole states from the vacuum level are the Ti^{H} and O_b^{H} polarons. In contrast, the shallowest electron and hole polaron sites are Ti^{R} and O_a^{B}. It can be seen that, across phases, electrons and holes have an energy range which is separated by roughly 0.2 eV when aligned to a common vacuum level. Of the phases explored in this work rutile is the only phase of TiO_2 which cannot form hole polarons.

5.7 Discussion

Some factors which may influence the accuracy of predictions shown are now discussed. As noted in Section 5.3 the use of the ADMM basis set allows for a vast speedup in calculation of exact exchange but also may introduce inaccuracies. However, bulk and polaronic properties of rutile and anatase were compared with a variety of titanium and oxygen basis sets and found that the choice of ADMM basis set has little influence. The choice of pFIT3/FIT9 yields similar results to pFIT3/FIT11, suggesting that including polarisation (f-)functions for titanium are largely unnecessary in these materials. Similarly, pFIT3/FIT9 gives similar results to FIT3/FIT9, also

suggesting that polarisation (*d*-)functions for oxygen are not required. These tests suggest that the choice of pFIT3/FIT9 is sufficiently accurate for the prediction of polaron properties.

Another potential source of error is the truncation of exact exchange in the hybrid functional beyond a cut-off radius R_c. The approach was to converge predicted bulk properties with respect to R_c to mimic the PBE0 functional but with reduced computational cost. This approach yielded R_c = 6 Å. However, the effects of using smaller cut-off values were also tested. While results obtained using R_c = 4 Å were not significantly different, reduction of the cut-off to 2 Å led to over-stabilisation polarons, absence of transferability between phases and predicted polaron properties inconsistent with experiment. It is notable that empirical attempts to tune the amount of HF exchange in the PBE0 functional to give a reasonable description of the band structure of anatase TiO_2 arrived at $\alpha = 0.125$.[228] This is very consistent with the value obtained in this chapter from application of the GKT providing further support for the approach.

The previous discussion (in Section 5.2) highlights that (small) electron polarons are stable in rutile, but not in anatase. In addition, some evidence suggests hole polarons are more stable in anatase compared to rutile. This in agreement with both experiment[47,190,199] and other theory.[46,48,193,194] It was shown that the predicted band gaps of 2.82 eV and 2.94 eV for rutile and anatase are in very good agreement with the experimentally observed values of 3.0 eV and 3.2 eV. The rutile (-/0) electron polaron CTL is found to be 0.4 eV below the CBM in this chapter, which is in agreement with experimental data placing the level at 0.3 eV, 0.5 eV and 0.7 eV.[47,191,192] It was highlighted that the rutile and anatase CTL data is also consistent with HSE06 findings.[46] In addition, the trapping energies of -0.02 eV and -0.20 eV for rutile and anatase are found to be in good agreement with HSE06. It was found that anatase and brookite have very similar charge trapping properties, behaviour which has been seen experimentally in Ref 190. Specifically, the two phases have much more delocalised electron polarons compared to the rutile case. Experimentally, little information exists on polaronic trapping in the phases TiO_2(H), TiO_2(R) and TiO_2(B). Therefore, the predictions shown are extremely valuable with regards to guiding future experimental work.

5.8 Conclusion

It has been demonstrated that hybrid functionals coupled with the GKT gives predictive capabilities with regards to localised charges in TiO_2. The stability and properties of small electron and hole polarons were assessed in various TiO_2 phases. Furthermore, by fitting the fraction of exact exchange to reduce nonlinear behaviour of energy with respect to occupation number of these polaronic states, a band gap differing from experimental results only around 6% was obtained. This suggests that this approach is not only capable of predicting polaron localisation properties, but also results in well-behaved bulk crystal properties. In addition to this, it was shown that the fraction of exact exchange α is transferable between phases. Using $\alpha = 0.115$ along with the tr-PBEα functional satisfies the GKT to within 0.08 eV for all phases. Parameters that influence the key results, such as the ADMM basis set, were explored ensuring that they are of sufficient quality. Work from this chapter suggests that the approach used is of reasonable quality, inexpensive, transferable and should be easily extendable to other metal-oxides capable of polaron formation.

Using this approach on the brookite, TiO_2(H), TiO_2(R) and TiO_2(B) phases has shown that intrinsic polaron formation is not unique to rutile and anatase. Unlike the remaining phases, TiO_2(H) and TiO_2(R) display multi-site electron and hole polarons, arising from to the unique nature of bonds between anions and cations. TiO_2(B) exhibits (small) hole polarons, but has shallow electron polaron states. While theoretical exploration of brookite is scarce, EPR data suggests the phase capable of forming hole polarons but not electron polarons, in agreement with findings shown here. Somewhat surprisingly, it is found that rutile is the only phase of titania which is incapable of forming hole polarons. The techniques employed in this work have been shown to be consistent with experimental data where available and transferable across all of the TiO_2 phases considered. Furthermore, the computational efficiency of the approach means it can be applied to more complex systems, such as surfaces, interfaces and doped systems, to provide greater insight into the nature of charge trapping in materials.

CHAPTER 6

INTRINSIC DEFECTS IN TITANIA

Purely-electronic defects, such as small electron and hole polarons, were discussed in the previous chapter (Chapter 5), where it was shown that the nature of these defects is sensitive to both the level of theory and its parametrisation. In this chapter, the focus shifts to intrinsic point defects, such as vacancies and interstitials. Oxygen and titanium defects are explored in both rutile and anatase, each with their thermodynamically feasible range of charge states. The GKT is also assessed for each charge state where possible. For most point defects, the tr-PBEα functional parametrisation used here yields similar results to previous studies. However, several stable charge traps relating to oxygen defects have been identified in this work, with many having binding energies substantially larger than the self-trapped polaron cases.

6.1 Introduction

The numerous applications of TiO_2, most notably its rutile and anatase phases, which include photocatalysis, superconductivity, tribocharging, magnetism and microelectronics have been discussed in previous chapters.[1–7] The change in electronic structure caused by point defects, such as oxygen vacancies, can give rise to trap states or scattering sites, possibly hindering the movement of charges. Applications such as photocatalysis rely on materials having particular properties to facilitate and enhance their use. Such properties may include the absence of charge trapping or adequate carrier concentrations. Intrinsic defects can also be used to dope and enhance a material; oxygen vacancies in oxides such as TiO_2 can release electrons into the conduction band, increasing *n*-type carrier concentrations. In contrast, other intrinsic defects, such as cation vacancies in metal oxides, may cause a reduction in *n*-type carrier concentrations. Studying these defects, and their impact on the electronic structure of the underlying material, would provide a better understanding of a defective material's properties and how its applications will be affected.

6.2 Background

Oxygen vacancies in rutile

The behaviour of oxygen vacancies in rutile has been investigated in a number of studies using DFT, DFT+U and hybrid functionals. The transition levels reported in previous literature for this defect are shown in Table 6.1. There is broad consensus that the V_O^{2+} charge state is stable across a wide range of Fermi energies. In this charge state, the vacancy causes a distortion of the surrounding crystal, pushing titanium ions away, and oxygen ions toward, the vacancy site. A number of configurations for other charge states have been identified in literature, including a closed-shell configuration where excess charge localises in the vacancy and a lower-energy, open-shell, configuration where excess charges form small polarons around the vacancy.[229–231] An LDA study (using a large supercell) finds delocalised charges (presenting as shallow gap states) for V_O^0, with a portion of charge located near the vacancy site.[232] Interestingly, if the experimental band gap is used instead of the calculated one, good qualitative agreement can be seen between GGA and screened exchange hybrid data regarding transition levels.[233,234] The range of charge states which are most thermodynamically

stable varies even between similar methods (see Fig. 6.1, Table 6.1).[231,234,235] Of note is that the behaviour of oxygen vacancies in rutile is highly-dependent on defect concentration (and therefore supercell size).[46] This may explain the discrepancy between HSE data in Table 6.1.

Oxygen deficient (110) surfaces of TiO_2 present an optical level peak at around 0.75 eV below the Fermi energy.[236] Another study suggests this peak lies around 1.1 eV below the band edge,[237] giving some indication of the optical depth of this level. A third study suggests that such a feature should instead be attributed to titanium interstitials as it remains present even when vacancies are healed.[238] Nevertheless, DFT eigenstates cannot be directly attributed to experimental optical measurements as electrons can excite into resonant or bound states above and below the conduction band edge. EPR data has identified the V_O^0 and V_O^{1+} charge states in rutile, which contains Ti^{3+} centres adjacent to the oxygen vacancy site along the c direction.[239,240] For V_O^0 the two centres were found to lie in a triplet ($S = 1$) configuration when illuminated with 442 nm (2.81 eV) laser light. The first of the two bound electrons was released at around 26 K ($\sim$ 2 meV), highlighting that one electron is only loosely bound. This low thermal activation energy suggests the defect essentially dopes the crystal n-type with one electron per vacancy.

It should be noted that defects in materials tend to self-regulate, preventing excessive n- or p-type doping through Fermi level pinning.[241] The Fermi energy is generally pinned at the intersection of donor and acceptor defects; under O-rich conditions the Fermi energy is pinned at around 1.2 eV, as seen in Fig. 6.1. Under O-poor conditions, this would correspond to an energy above the CBM, meaning both phases would display n-type characteristics typical of TiO_2 samples. As a consequence, excessively negative formation energies far from the pinned E_F have little meaning.

Oxygen vacancies in anatase

Like rutile, the V_O^{2+} charge state in anatase causes a distortion in the underlying crystal with no excess charge present. Information provided in Chapter 5, as well as other theoretical and experimental studies, indicates that small electron polarons do not form in bulk anatase.[199,235,249] It follows that the V_O^{1+} and V_O^0 solutions in anatase may differ to those found in rutile as excess charges will not localise on sites far from the vacancy. DFT+U reports suggest excess charges localise around titanium atoms nearest the oxygen vacancy,[46,229,250] while hybrid functionals predict charges localises inside the

Ref	Approach	Parameters	Size	Gap	States	2+/1+	1+/0	2+/0
Rutile								
231	HSE	{0.20, 0.20}	72	3.1	2+	-	-	-
242	HSE	{0.25, 0.20}	576	-	2+	-	-	-
243	LDA	-	72	1.7	2+	-	-	-
233	GGA	-	270	-	2+/0	-	-	2.1
234	sX	-	72	3.1	2+/0	-	-	2.3
244	GGA+U	{4.20, 5.25}	216	2.4	2+/0	-	-	1.8
245	G_0W_0	-	72	3.1	2+/0	-	-	2.8
235	HSE	{0.25, 0.20}	192	3.4	2+/1+/0	2.6	3.0	-
246	GGA+U	{3.25, -}	96	2.0	2+/1+	1.6	2.1	-
246	GGA+U	{3.25, 5.00}	96	2.7	2+/1+/0	2.4	2.7	-
Anatase								
247	HSE	{0.25, 0.20}	108	3.7	2+	-	-	-
248	LDA	-	108	2.2	2+	2.7	2.9	-
246	GGA+U	{3.23, -}	108	2.5	2+/1+	2.4	2.5	-
235	HSE	{0.25, 0.20}	96	3.6	2+/1+/0	3.2	3.5	-
246	GGA+U	{3.23, 5.00}	108	3.3	2+/1+/0	3.1	3.3	-
180	GGA+U	{4.20, 5.25}	108	2.6	2+/1+/0	2.2	2.3	-

All energies are given in electron-volts

Table 6.1: Supercell sizes, calculated band gaps and defect transition levels (with respect to the VBM) for a number of TiO_2 oxygen vacancy studies. The parametrisation used for HSE and DFT+U studies is provided as $\{U_{Ti}, U_O\}$ and $\{\alpha, \omega\}$ for studies. The behaviour of the Hubbard-U term strongly depends on the pseudopotential used, which can be found in the appropriate reference. The screened exchange study used a Thomas Fermi screening length of 2.15 $Å^{-1}$. The most thermodynamically stable charge states for each study are also noted. *(Rutile only)* Ref. 231 and Ref. 243 found only the (2+) charge state to be thermodynamically stable within the band edges, however the open-shell localised solution was not used to calculate the V_O^0 formation energy. Ref. 242 used a relatively large supercell, however the electron localisation configuration was not reported. Ref. 233 did not find a localised solution, however data extrapolated beyond the calculated band edges is in good quantitative agreement with a screened-exchange (sX) hybrid functional study. *(Anatase only)* Ref. 246 shows two sets of results for different DFT+U parametrisations. Ref. 247 found only the V_O^{2+} charge state to be stable.

vacancy site.[46] Both approaches also find that the V_O^0 configuration exists in a singlet state. A solution where one electron localises on a titanium ion near the vacancy and another localises on a next-nearest titanium site is reported to be similar in energy to the solution where two electrons localise on the nearest titanium sites.[180] Studies performed using the LDA do not find localised states, however this approach is known to favour such solutions.[248] Some studies did not report electronic configurations for this defect, however all reported transition levels for this defect are shown in Table 6.1.

Oxygen vacancies on anatase (101) surfaces bind excess electrons to the defect site as seen in STM data.[47] The presence of this defect introduces a gap state roughly 1 eV

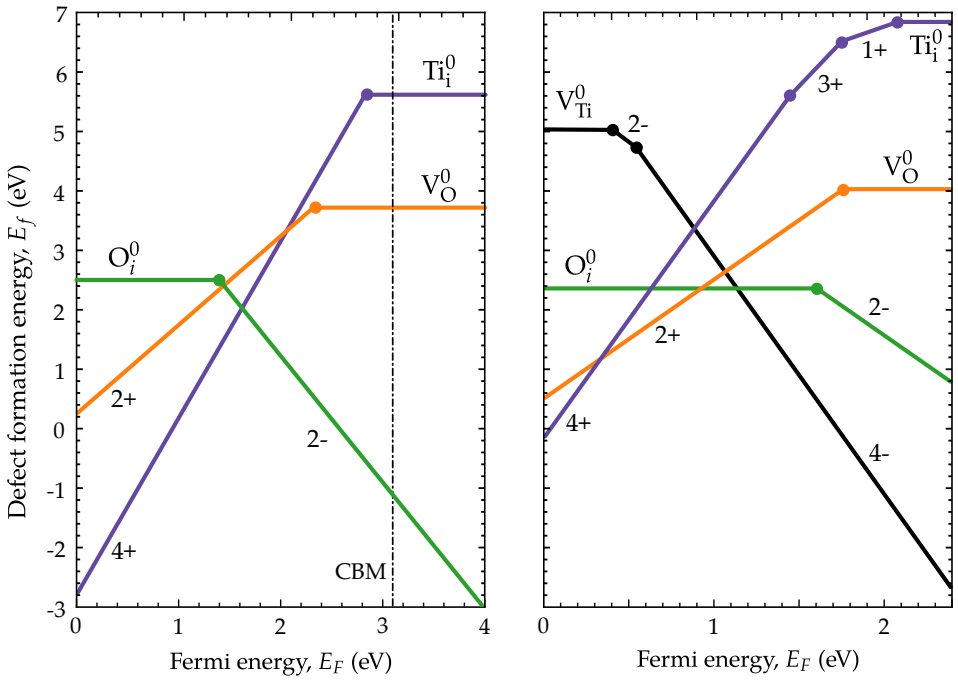

Figure 6.1: Defect formation energies for rutile reproduced from a screened-exchange hybrid functional study (left, Ref. 234) and a DFT+U study (right, Ref. 244). All energies were calculated using oxygen-rich chemical potentials (i.e. $\mu_O = \frac{1}{2}E[O_2]$).

below the CBM.[251] Unlike rutile, surface oxygen vacancies in anatase are less stable than bulk ones, allowing for migration of vacancies into the bulk.[252] The migration barrier for surface to subsurface migration is larger (0.6-1.2 eV) than bulk to bulk vacancy migration (0.2 eV).[253] Livraghi *et al* find two key EPR signals in anatase microcrystals, one attributed to diffuse electrons in bulk anatase ($g_{\parallel} = 1.962$, $g_{\perp} = 1.992$) and another, present in both chemically and thermally reduced samples, assigned to reductive processes involving surface layers ($g = 1.93$).[199] This information suggests that a larger number of vacancies in anatase are located on surfaces, and those in the bulk may be of a more diffuse nature. This picture is consistent with, and perhaps explains, the n-type behaviour often seen in TiO_2 anatase samples. More specifically, this may explain why anatase has higher charge mobility than rutile.[200,254]

Oxygen interstitials in rutile

DFT, DFT+U and hybrid functional studies of this defect in rutile suggest the interstitial forms a bond with a lattice oxygen, with the two atoms in a planar configuration, and causes a distortion in the surrounding crystal.[234,255,256] As oxygen ions in TiO_2 have a formal charge of O^{2-}, the addition of a neutral oxygen results in an O_2^{2-} molecule. A DFT+U study reports that a rotated configuration is lower in energy than

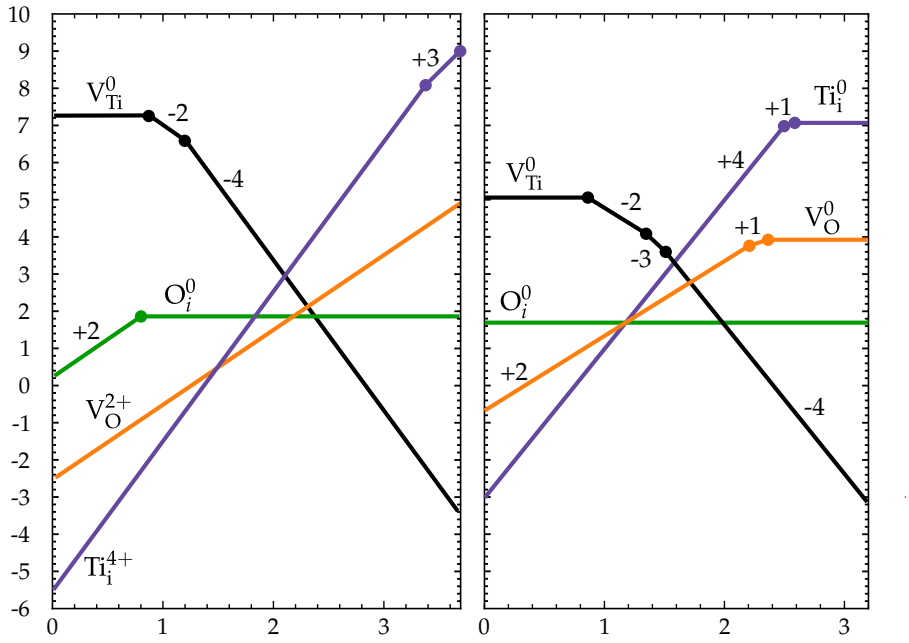

Figure 6.2: Defect formation energies for anatase reproduced from a HSE06 hybrid functional study (left, Ref. 247) and a DFT+U study (right, Ref. 180). All energies were calculated using oxygen-rich chemical potentials (i.e. $\mu_O = \frac{1}{2}E[O_2]$).

a planar one.[244] The most thermodynamically stable charge states for this defect are reported to be the O_i^0 and O_i^{2-}, with the O_i^{1-} charge state having a higher formation energy. There is good quantitative agreement between local DFT, hybrid functional and DFT+U studies regarding defect formation energies (of the neutral case), with two examples shown in Fig. 6.1.[234,244,255,256] Reported transition levels for this defect are shown in Table 6.2. There is little mention of positive charge states for this defect in literature.

Unfortunately experimental studies on this defect are few in number, however two relevant studies have been identified. The first suggests that interstitial oxygen may be the defect which causes a transformation from an n- to a p-type semiconductor when oxygen pressure is increased near 1000°.[257] The second reports that improved photocatalytic performance of TiO_2 can be achieved by using a preparation that allows for excess oxygen, however both anatase and rutile phases are present in these samples.[258]

Oxygen interstitials in anatase

A DFT+U study reports that the range of thermodynamically stable charge states for this defect are similar to the rutile case, with both forming O_i^0 and O_i^{2-}.[255] Of note is that this study did not incorporate a Hubbard-U value on oxygen atoms, which

Ref	Approach	Parameters	Size	Band gap	States	2+/0	0/2-
Rutile							
234, 259	sX	-	72	3.1	0/2-	-	1.4
256	GGA	-	216	-	0/2-	-	0.9
255	GGA+U	{3.00, - }	24	-	0/2-	-	1.3
244	GGA+U	{4.20, 5.25}	216	2.4	0/2-	-	1.6
Anatase							
247	HSE06	{0.25, 0.20}	108	3.7	2+/0	0.8	-
260	LDA	-	108	2.3	0	-	-
180	GGA+U	{4.20, 5.25}	108	2.6	0	-	-
255	GGA+U	{3.00, - }	24	-	0/2-	-	1.5

All energies are given in electron-volts.

Table 6.2: System sizes, calculated band gaps and defect transition levels for a number of TiO_2 oxygen interstitial studies. The most thermodynamically stable charge states, approach and parametrisation of the approach are also reported. Ref. 255 and Ref. 256 did not report a calculated band gap. Many studies found the O_i^{1-} charge state was less thermodynamically stable than O_i^{2-} at all Fermi energy values. (*Anatase only*) Ref. 180 and Ref. 260 find that only the O_i^0 charge state is stable in anatase.

may have prohibited an investigation into localised positive charge states. The study suggests the defect forms more easily in anatase than rutile and have similar transition levels.[255] Two studies conducted on this defect, using the LDA and DFT+U, found only the O_i^0 charge state to be stable, as shown in Table 6.2.[180,260] Only one study, using HSE, was found reporting positive charge states for this defect in anatase, however no transitions to negative charge states were reported.[247]

It has been observed through PL that oxygen-rich anatase has a reduced band gap compared to control samples.[261] In addition, O-rich anatase has been reported to remain as anatase even when annealed at 900°, a temperature that would usually trigger an irreversible phase transformation to rutile.[261] Core level spectroscopy data suggests that O_2 annealing of Nb-doped anatase compensates for the additional carriers provided by doping by introduction of oxygen interstitials.[262]

Titanium vacancies in rutile

This defect in TiO_2 is a p-type one that can introduce a number of holes into the crystal. The neutral defect introduces four holes, each of which get filled as electrons are added until none remain in the V_{Ti}^{4-} charge state. A DFT+U study reports the optimum configuration for the neutral defect to be four holes immediately surrounding the defect site.[244] One other study using the LDA indicates that oxygen ions move away from the vacancy, while titanium ions move towards it.[243] Further studies on

this defect in rutile are scarce, however those available have their transition levels shown in Table 6.3.

Prolonged oxidation of undoped TiO_2 single crystals at high temperatures (75kPa $P(O_2)$ at 1323 K for over 6000 hours) causes a reduction in conductivity over time as expected.[263,264] As the (irreversible) rutile to anatase phase transition occurs at around 600°, the remaining crystals are likely mainly composed of rutile. Strangely, performing the same experiment at lower temperatures (1123 K) causes an increase in conductivity of the material. This has been explained by increasing oxygen pressure required to undergo an *n*–*p* mobility transition.[265] Many of these effects have been attributed to titanium vacancy formation propagating from the surface to the bulk over prolonged periods of time.[263] However, this is in contrast with data from Fig. 6.1, which implies that *p*-type TiO_2 forms oxygen interstitials much more easily than titanium vacancies. A large body of work suggests that TiO_2 becomes a majority *p*-type conductor when in a strongly oxidised regime.[264–267]

Titanium interstitials in rutile

This *n*-type defect donates four electrons into the crystal in its neutral charge state. The reported transition levels for this defect are fairly wide-ranging, as seen in Table 6.4. In addition, the most thermodynamically stable charge states also vary even between similar approaches. A DFT+U study suggests excess charges do not localise on the defect itself but on nearby titanium ions, behaving like a weakly bound polaron.[244] Another DFT+U study found a similar result,[229] albeit with a small amount of charge on the interstitial. A HSE06 study reports the contrary, indicating that the Ti_i^{3+} and Ti_i^{2+} charge states result in defect-bound electrons that sit in a (closed-shell) singlet configuration.[235] Two hybrid studies, using screened exchange and B3LYP, suggest the interstitial ion only traps one electron, with the remaining three being donated into the crystal or around the defect site.[234,268] Another DFT+U study reports similar findings to Refs 268 and 234.[246]

Experiments using photoelectron spectroscopy (PES) and STM have indicated that a gap state 0.85 eV below the Fermi energy is mostly likely associated with interstitial titanium atoms.[238] This feature, which has previously been associated with bridging oxygen vacancies, was shown to be present even when the so-called bridging oxygen vacancies were healed. Interestingly, it is found that these interstitial atoms diffuse to the surface when a sufficient amount of oxygen is present. The re-emergence of this

Ref	Approach	Parameters	Size	Gap	States	4-/3-	4-/2-	3-/2-	2-/0
Rutile									
243	LDA	-	72	1.7	4-	-	-	-	-
244	GGA+U	{4.20, 5.25}	216	2.4	4-/2-/0	-	0.6	-	0.4
Anatase									
260	LDA	-	108	2.3	4-	-	-	-	-
247	HSE	{0.25, 0.20}	108	3.7	4-/2-/0	-	1.2	-	0.9
180	GGA+U	{4.20, 5.25}	108	2.6	4-/3-/2-/0	1.5	-	1.4	0.8

All energies are given in electron-volts.

Table 6.3: Supercell sizes, calculated band gaps E_g and defect transition levels for a number of studies on the TiO_2 titanium vacancy.

feature beneath the Fermi energy occurs after annealing at 400 K. It has been suggested that this occurs due to Ti atoms reacting with surface adsorbed oxygen, causing deeper Ti atoms to migrate into interstitial sites closer to the surface. This occurrence has been noted by other authors previously.[269–271] EPR data on this defect in rutile finds that in particularly non-stoichiometric samples a signal can be found ($g_1 = 1.9780$, $g_2 = 1.9746$, $g_3 = 1.9414$),[272,273] indicating a paramagnetic species. Other EPR data (obtained at 77 K) assigns two signals to surface and near-surface sites (dehydrated surface: $g_1 = 1.970$, $g_2 = 1.961$, $g_3 = 1.948$ and hydrated surface: $g_1 = 1.973$, $g_2 = 1.961$, $g_3 = 1.948$), and a negligible amount of signal associated with interstitial Ti^{3+} ($g_1 = 1.978$, $g_2 = 1.975$, $g_3 = 1.942$).[190] It has been argued by other authors that this final signal is actually associated with regular lattice sites and not interstitials,[274] raising a question as to whether bulk titanium interstitials can be detected through EPR. As indicated by HSE data (see previous paragraph), it may be that this state is EPR silent.

Titanium vacancies in anatase

Like the rutile case, there are very few studies on this defect, as seen in Table 6.3. A GGA study using a 48 atom supercells reports that the neutral form of this defect has a magnetic structure with total spin $4\mu_B$. Work using DFT+U suggests that excess holes generated by this defect sit on oxygen ions near the vacancy site, similar to a DFT+U report in rutile.[180,244]

An experimental study reports Ti deficient anatase can lead to improved photocatalytic activity and high charge mobility through p-type conduction.[275] p-type conduction was verified in this case through Hall effect measurements, which show a positive

Ref	Approach	Size	Gap	Stable	4+/3+	4+/0	3+/2+	3+/1+	2+/1+	2+/0	1+/0
Rutile											
233	GGA	270	-	4+	2.0	-	2.0	-	2.2	-	2.2
234	sX	72	3.1	4+/0	-	2.8	-	-	-	-	-
246	GGA+U	96	2.7	4+/3+/2+	2.3	-	2.6	-	2.7	-	2.8
246	GGA+U	96	2.0	4+/3+/2+/1+	1.8	-	1.8	-	2.0	-	2.0
244	GGA+U	216	2.4	4+/3+/1+/0	1.5	-	-	1.8	-	-	2.1
235	HSE	192	3.4	4+/3+/2+/1+/0	1.9	-	2.3	-	2.8	-	3.2
Anatase											
260	LDA	108	2.3	4+	-	-	-	-	-	-	-
246	GGA+U	96	2.5	4+/3+	2.3	-	2.5	-	2.5	-	2.7
246	GGA+U	96	3.3	4+/3+/2+	2.9	-	3.2	-	3.3	-	3.6
247	HSE	108	3.7	4+/3+/2+	3.4	-	3.7	-	3.8	-	4.0
235	HSE	96	3.6	4+/3+/2+/0	2.3	-	2.9	-	-	3.3	-

All energies are given in electron-volts.

Table 6.4: Supercell sizes, calculated band gaps E_g and defect transition levels for a number of TiO_2 titanium interstitial studies. (*Anatase only*) One further study DFT+U study (Ref. 180) with transition levels (4+/1+) = 2.5 eV and (1+/0) = 2.6 eV was not included in the table due to space restrictions. The transition levels can instead be seen in Fig. 6.2.

Hall coefficient. The experimental data is in contrast with HSE, which predicts that the Fermi energy is still 2.3 eV above the valence band even under extremely O-rich conditions.[247] The experimental study also reports room temperature ferromagnetism and an EPR signal at $g = 1.998$.[275] The absence of oxygen interstitials and presence of titanium vacancies was argued through comparison of X-ray diffraction (XRD) data, which shows a change in lattice constants, with DFT relaxed lattice constants of defective supercells. Use of positron annihilation indicates a longer lifetime of positrons compared to the defect-free bulk, which can be attributed to a reduced electron density and larger number of vacancies.

Titanium interstitials in anatase

As with the rutile case, this defect in anatase has a wide range of reported transition levels. In anatase, they even differ between studies using the same functional, as seen in Table 6.4. HSE06 data implies this defect localises two electrons near the defect site and delocalises the remaining two into the conduction band.[235] Two DFT+U studies give a different picture, with one electron localised around the interstitial site (forming Ti^{3+}) and the remaining three localising around the interstitial.[180,246] Calculations using the LDA approach predict that the four excess electrons are donated into the conduction band,[260] however the LDA is known to favour delocalisation of states.

One EPR study on this defect found Ti^{3+} centres in bulk anatase, however it was not possible to distinguish this defect from substitutional Ti^{3+}.[276] More recent studies have noted that an interstitial Ti^{3+} centre has yet to be unambiguously identified using EPR.[199] As mentioned previously, the oxygen vacancy in anatase may also give rise to Ti^{3+} centres near surfaces, which may make assigning EPR signals to this defect directly difficult.

All defects

Aggregating the studies referenced above, a clouded picture of charge localisation behaviour in TiO_2 emerges. The optimal configuration of charges surrounding point defects varies wildly across the theoretical techniques surveyed above, with most reasonable configurations identified in at least one study. Using oxygen vacancies in anatase as an example, the neutral charge state is found to have charges that either reside within the vacancy site, on neighbouring titanium atoms to the vacancy, or its next-nearest neighbours. Supercells used in most studies tend to be small, and may suffer from image charge interactions. Furthermore, using smaller cells may prohibit finding the optimal configuration of charges surrounding a defect, artificially raising defect formation energies. Experimental techniques in some cases cannot directly assign particular measurements to certain defects, confusing the picture further. Worse still, certain measurements may be misassigned to defects, as seen in some studies above. This highlights the urgent need for predictive theoretical techniques to be employed, with the aim of guiding future experimental and theoretical work.

As noted by previous authors, the *n*-type behaviour of TiO_2 has previously been attributed to titanium interstitial defects,[235] however recently it has emerged that these play more of a role in heavily reduced crystals, while oxygen vacancies dominate the mildly reduced regimes. However, this picture is at odds with a large number of theoretical approaches, most of which predict excess charge binds near defect sites. As self-trapped electron polarons have been shown to be only mildly stable at zero temperature in bulk rutile and anatase (both experimentally and theoretically), the behaviour of point defects may play a part in explaining why anatase has enhanced *n*-type behaviour over rutile. In particular, the four orders of magnitude larger conductivity seen in Nb-doped anatase compared to Nb-doped rutile may be explained, at least in part, by the behaviour of intrinsic point defects.

Defect	Symbol	Charge states				
Oxygen interstitial	O_i	O_i^{2+}	O_i^{1+}	O_i^{0}	O_i^{1-}	O_i^{2-}
Oxygen vacancy	V_O	V_O^{0}	V_O^{1+}	V_O^{2+}		
Titanium interstitial	Ti_i	Ti_i^{0}	Ti_i^{1+}	Ti_i^{2+}	Ti_i^{3+}	Ti_i^{4+}
Titanium vacancy	V_{Ti}	V_{Ti}^{0}	V_{Ti}^{1-}	V_{Ti}^{2-}	V_{Ti}^{3-}	V_{Ti}^{4-}

Table 6.5: The symbol and explored charge states for all intrinsic defects in rutile and anatase are shown.

6.3 Methods

All calculations in this section were performed using the hybrid functional implementation in CP2K along with GTH pseudopotentials.[144,148,209–211,213] As CP2K only samples the Γ-point in reciprocal space, crystal properties were converged with respect to system size instead of k-point mesh size. The tr-PBEα hybrid functional, discussed in previous chapters, was used with $\alpha = 0.115$ and $R_c = 6.0$ Å. The primary basis sets used were of triple-ζ quality and included valence and polarisation exponents.[212] The FIT3 and pFIT3 ADMM basis sets, provided in Ref. 277, were used for titanium and oxygen. In Chapter 5, testing of the above parameterisation was done to ensure that the bulk and polaron properties of rutile and anatase were well-converged. For rutile and anatase, $4 \times 4 \times 6$ and $5 \times 5 \times 2$ expansions of the conventional unit cells were used to construct the supercells used. These contain 576 and 600 atoms, respectively.

6.3.1 Transition levels

The defects explored in subsequent sections, and their possible charge states, are shown in Table 6.5. The trapping energy E_t for a charge localised around an intrinsic defect X^q can be calculated as

$$E_t[X^q] = E[X^q] - E[H^0] + \begin{cases} E[H^{-1}] - E[X^{q-1}], & \text{for } n\text{-type defects} \\ E[H^{+1}] - E[X^{q+1}], & \text{for } p\text{-type defects} \end{cases} \tag{6.1}$$

which provides a measure of trap stability by comparing the energy of a trapped charge near a defect $E[X^q]$ to one delocalised in the bulk $E[H^{\pm 1}]$ [1]. Using n-type defects as an example, it can be shown that the trapping energy E_t is equivalent to the charge transition levels with respect to the band edges between different charge states.

[1] $E[H^0]$ is the total energy of the host (i.e. defect free bulk) in a neutral cell, as defined in Chapter 2.

At the transition level between defects X^q and X^{q-1}, the two defect formation energies are equal,

$$\Big[E_f[X^q] - E_f[X^{q-1}]\Big]_{E_F=E_F^\star} = 0, \tag{6.2}$$

and have the same Fermi energy $E_F^\star$. The transition between the pristine bulk and a delocalised electron in the pristine bulk, which occurs near the conduction band edge E_F^{CBM}, also equals zero

$$\Big[E_f[H^{-1}] - E_f[H^0]\Big]_{E_F=E_F^{CBM}} = 0. \tag{6.3}$$

Both energy components of the above (Eqn. 6.3) are individually zero as well; the defect formation energy of the defect-free bulk is zero by definition, as is the defect formation energy of a delocalised bulk, band-like, electron when $E_F = E_F^{CBM}$. The sum of Eqn. 6.2 and Eqn. 6.3 is therefore

$$\Big[E_f[X^q] - E_f[X^{q-1}]\Big]_{E_F=E_F^\star} + \Big[E_f[H^{-1}] - E_f[H^0]\Big]_{E_F=E_F^{CBM}} = 0. \tag{6.4}$$

Expanding the formation energies (using Eqn. 2.57) gives

$$\underbrace{\Big(E[X^q] + E[H^{-1}]\Big)}_{\text{Reactants}} - \underbrace{\Big(E[X^{q-1}] + E[H^0]\Big)}_{\text{Reagents}} = E_F^{CBM} - E_F^\star = E_t[X^q] \tag{6.5}$$

where it has been assumed that the chemical potentials of atoms are the same for all formation energies. The LHS of Eqn. 6.5 can be thought of as the energy released or gained through the reaction

$$E[X^q] + E[H^{-1}] \rightleftharpoons E[X^{q-1}] + E[H^0] \tag{6.6}$$

which is equivalent to a trapping energy for systems containing crystallographic defects.

Charge state	$\|q\| = 1$	$\|q\| = 2$	$\|q\| = 3$	$\|q\| = 4$
Rutile correction	0.003	0.011	0.024	0.043
Anatase correction	0.015	0.059	0.133	0.236

Table 6.6: The size of image charge corrections in eV for rutile and anatase using the Lany-Zunger scheme.

Defect	O interstitial	O vacancy	Ti interstitial	Ti vacancy
Rutile correction	0.016	-0.005	0.064	-0.052
Anatase correction	0.012	-0.007	0.054	-0.047

Table 6.7: Potential alignment corrections (in eV) for each defect. These were calculated by comparing the bulk electrostatic potential to the geometrically unrelaxed and uncharged defect, as described in Ref. 278.

6.3.2 Chemical potentials

All defect formation energies are calculated under oxygen-rich chemical potentials, meaning the chemical potential of oxygen μ_O and titanium μ_{Ti} atoms are given by

$$\mu_O = \frac{1}{2}E[O_2] \tag{6.7}$$

$$\mu_{Ti} = E[TiO_2] - E[O_2] \tag{6.8}$$

where $E[TiO_2]$ in this case is the formation enthalpy per formula unit of TiO_2. For oxygen-poor conditions, chemical potentials are chosen at the TiO_2/Ti_2O_3 equilibrium,[234,260]

$$\mu_{Ti} + 2\mu_O = E[TiO_2] \tag{6.9}$$

$$2\mu_{Ti} + 3\mu_O = E[Ti_2O_3]. \tag{6.10}$$

Going from O-rich to Ti-rich conditions, oxygen and titanium chemical potentials shift by $\Delta\mu_O = 3.80$ eV and $\Delta\mu_{Ti} = -7.61$ eV for rutile. In anatase, the shifts are $\Delta\mu_O = 3.97$ eV and $\Delta\mu_{Ti} = -7.93$ eV.

6.3.3 Image charge correction

As charge states relating to the titanium defects can be high ($|q| = 4$), corrections to the total energy must be made to account for periodic image interactions. The Lany-Zunger implementation of the Makov-Payne corrections straight-forwardly provides

a third order correction.[184,278,279] The scheme provides an energy correction to charged systems ΔE_{ic} that is given by

$$\Delta E_{\text{ic}} = \frac{q^2 \alpha_M}{2\varepsilon L} \cdot f^{\text{LZ}} \tag{6.11}$$

$$f^{\text{LZ}} = 1 - c_{\text{sh}} \left(1 - \frac{1}{\varepsilon}\right) \tag{6.12}$$

where α_M is the Madelung constant, which only depends on the shape of the cell.[184,278] The dielectric constant ε should equal the high-frequency one ε^{∞} when only electronic relaxations are considered. If ionic relaxations take place, the static constant should instead be used instead. In the above correction, L is the distance between defects and c_{sh} is a shape factor that depends on the size of the cell.

As dielectric constants cannot be directly obtained from `CP2K` at present, these quantities were instead calculated in `VASP` for a PBEα hybrid functional with $\alpha = 0.115$. This will naturally cause an error, as the dielectric response of electrons and ions in these two DFT implementations will differ. However, the error in ε will not significantly change the image charge correction. To illustrate this, image charge corrections were calculated using LDA and HSE06 dielectric constants. For the $|q| = 4$ case, the largest charge considered, the difference in E_{ic} was found to be < 0.01 eV in anatase. This same difference was also found to be 0.03 eV in rutile. This highlights that use of a dielectric constant from `VASP` will not significantly impact defect formation energies. For rutile, the calculated dielectric constants (ε_a^{∞}, ε_c^{∞}, ε_a, ε_c) were found to be (6.90, 8.41, 224.6, 328.2). In anatase, these were (6.48, 6.11, 49.3, 28.9). The subscripts a and c correspond to the two inequivalent lattice vector directions. The associated image interaction corrections for are shown in Table 6.6.

6.3.4 Eigenvalue correction

Calculations involving a charged supercell include an error contained within eigenvalues that is proportional to the total energy error. With increasing system sizes, the change in eigenvalue energies corresponding to localised states can be attributed to two factors. First, an increase in the degrees of freedom, allowing the localised charge to extend further into the crystal. Second, an electrostatic interaction between localised states in periodic cells. While the first change cannot be accounted for straightforwardly using a *post-hoc* correction, the second can be approximated reasonably well. In Appendix C, we show for the case of an oxygen molecule in a large periodic box

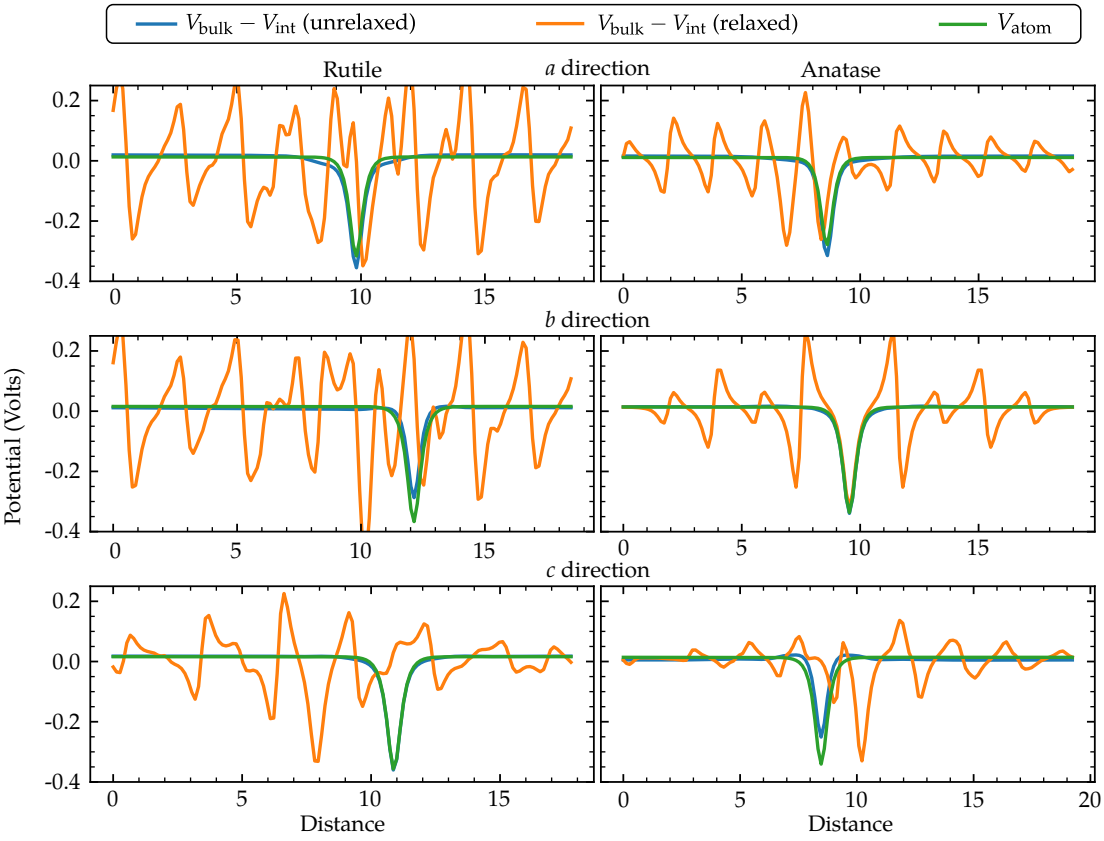

Figure 6.3: Electrostatic potentials for the oxygen interstitial defect in rutile (left) and anatase (right). The average electrostatic potential across cell slices is shown along the three crystallographic directions. For each direction and phase, three electrostatic potentials are compared: the difference between bulk and geometrically relaxed interstitial defect (orange), bulk and geometrically unrelaxed defect cell (blue) and the potential for only an oxygen atom in the same sized box (green). The smoother (green and blue) lines correspond to potentials used in the correction scheme proposed by Durrant *et al* (Ref 278). The Lany-Zunger potential alignment correction (Ref 279) often relies on a less smooth potential (yellow lines above).

that the correction (in hybrid DFT) is of the form

$$\varepsilon_{ic} = \varepsilon - \frac{2}{q}\Delta E_{ic} \tag{6.13}$$

where ε_{ic} and ε are the corrected and uncorrected eigenvalues. This correction matches the form used in Ref 280, but without a potential alignment correction included. Applying the same correction to the V_O^{1+} defect in MgO, an F-centre defect, results in ideal eigenvalue scaling with system size (also shown in Appendix C).

6.3.5 Potential alignment correction

In periodic DFT calculations, the average electrostatic potential is usually set to zero. When calculating defect formation energies, this can cause issues as the bulk and de-

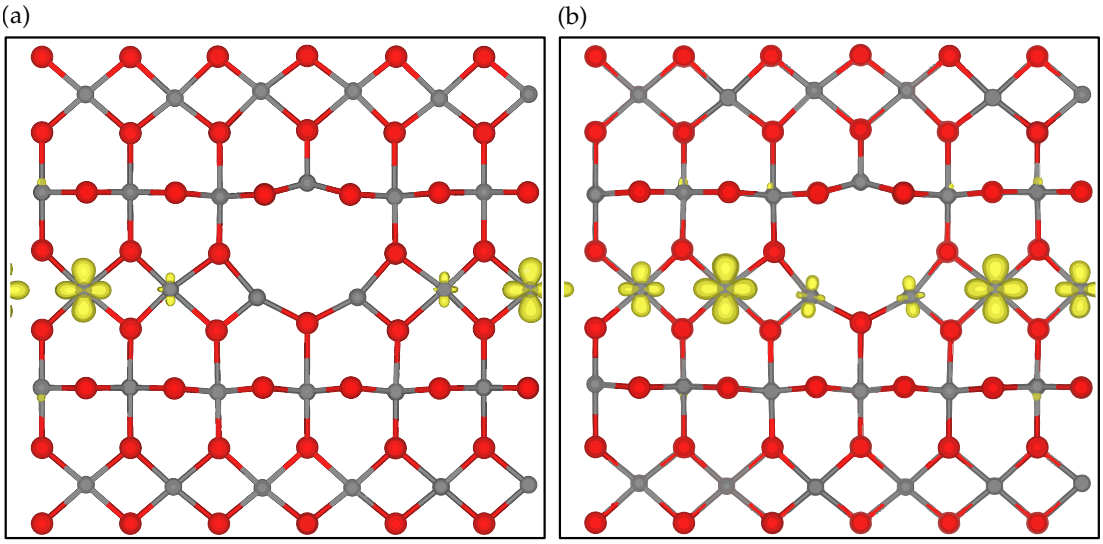

Figure 6.4: Illustrated is a portion of the rutile TiO_2 crystal with an oxygen vacancy present. Spin density isosurfaces for non-paired spins are shown for V_O^{1+} (a) and V_O^0 (b) charge states. Spin density isosurfaces are set at 0.01 a_0^{-1}.

fective cells' energies may have different references. The Lany-Zunger potential alignment procedure calculates a total energy correction to account for this using

$$V_{q/b}(\mathbf{r}) = V_q(\mathbf{r}) - V_b(\mathbf{r})\big|_{\mathbf{r}_{\text{far}}} \tag{6.14}$$

which compares the bulk electrostatic potential $V_b(\mathbf{r})$ to the defect potential $V_q(\mathbf{r})$. For the defect potential, a region far from the defect is used, allowing the defect cell's bulk-like region to align to the true bulk. As a result, the energy correction can vary depending on how and where the bulk-like region is defined. Durrant *et al* argue that use of Eqn. 6.14 corrects both the potential alignment *and* image charge interaction, meaning when used in conjunction with Eqn. 6.12 the image interaction correction is double-counted.[278] Instead, they propose that the potential alignment correction should be performed by comparing the defective, geometrically unrelaxed, and uncharged electrostatic potential to the defect-free bulk one. In their paper, they argue a key component of potential alignment is adding and removing atoms to periodic cells. As seen in Fig. 6.3, this approach removes much of the variation in the electrostatic potential due to ionic relaxation. Durrant *et al* also find that this component of potential alignment can be approximated by the electrostatic potential of a cell with just the atom relating to the defect in the simulation cell, as shown in Fig. 6.3.[278] The size of the potential alignment corrections used is given in Table 6.7.

Additional information can be calculated from the outputs of *ab initio* calculations,

such as optical gaps or shallow defect corrections, using the `sumo` package.[281] This is currently only available for calculations using `VASP`, although there are plans to include other *ab initio* codes to this package.

6.4 Oxygen defects

Vacancies in rutile

The presence of an V_O^{2+} defect in rutile causes a distortion of the lattice, moving nearby oxygen atoms towards the vacancy site and titanium atoms away from it. This is illustrated in Fig. 6.4. As self-trapping polarons can form in rutile, the excess electron present in the V_O^{1+} charge state can localise in a number of locations in the cell. The most stable localised configuration was found to be one where the electron is in the vicinity of the oxygen vacancy, as shown in Fig. 6.4. A fully delocalised configuration is found to be 0.1 eV higher in energy than the most stable localised one. The V_O^{2+}/V_O^{1+} transition level for this defect is found to be 0.12 eV below the CBM, suggesting that oxygen vacancies are a trapping site for excess electrons. For the V_O^0 charge state, the most stable configuration is also one where two electrons are localised in the vicinity of the defect, shown in Fig. 6.4. A change in the spin configuration of the two charges does not cause a significant change in energy, with the singlet and triplet configurations differing by only a few meV. The V_O^{1+}/V_O^0 transition level is 0.03 eV below the CBM, indicating that the first localised charge is much more strongly bound than the second. An V_O^{2+} configuration with two delocalised electrons is found to be 0.13 eV higher in energy than the localised V_O^0 configuration, confirming that a delocalised configuration is unfavourable. These energies are much larger than the polaron binding energy (found in Chapter 5), which was -0.02 eV. Fig. 6.5 shows *p*-type conditions would favour the V_O^{2+} charge state, while *n*-type conditions will produce mainly the V_O^0 charge state.

Interstitials in rutile

Previous studies have identified the optimal configuration for O_i^0, O_i^{1-} and O_i^{2-} charge states in rutile, however the O_i^{1+} charge state has not been reported yet in literature. Calculated charge isosurfaces and atomic configurations are shown are shown in Fig. 6.6 for the calculated most stable configurations of all charge states. The neutral O_i^0 defect displaces, and forms a bond with, a lattice oxygen, causing an O_2^{2-} pair to

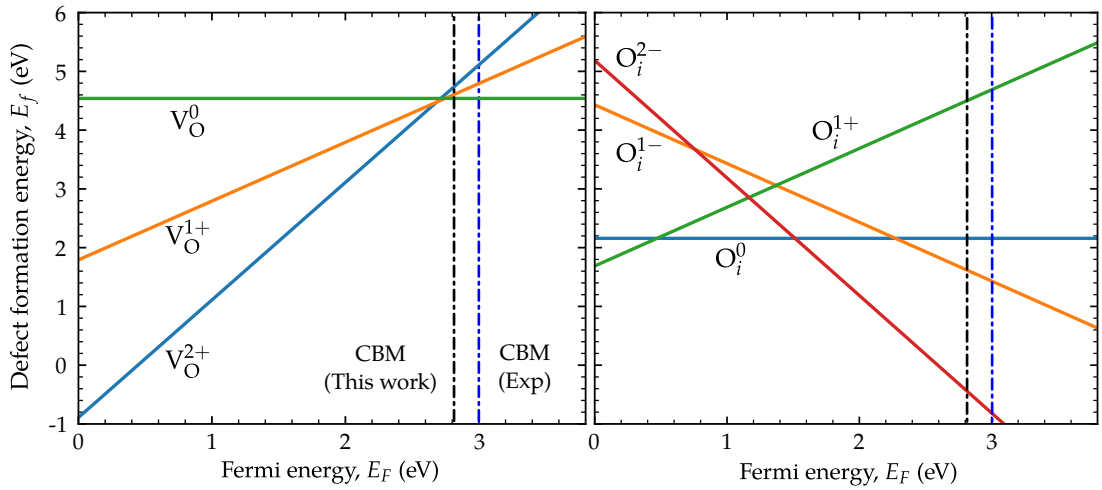

Figure 6.5: The variation in defect formation energy with Fermi energy for the lowest energy configurations of each charge state is shown for the oxygen vacancy (left) and oxygen interstitial (right) defects in rutile. The dashed vertical black and blue lines show the calculated and experimental CBM.

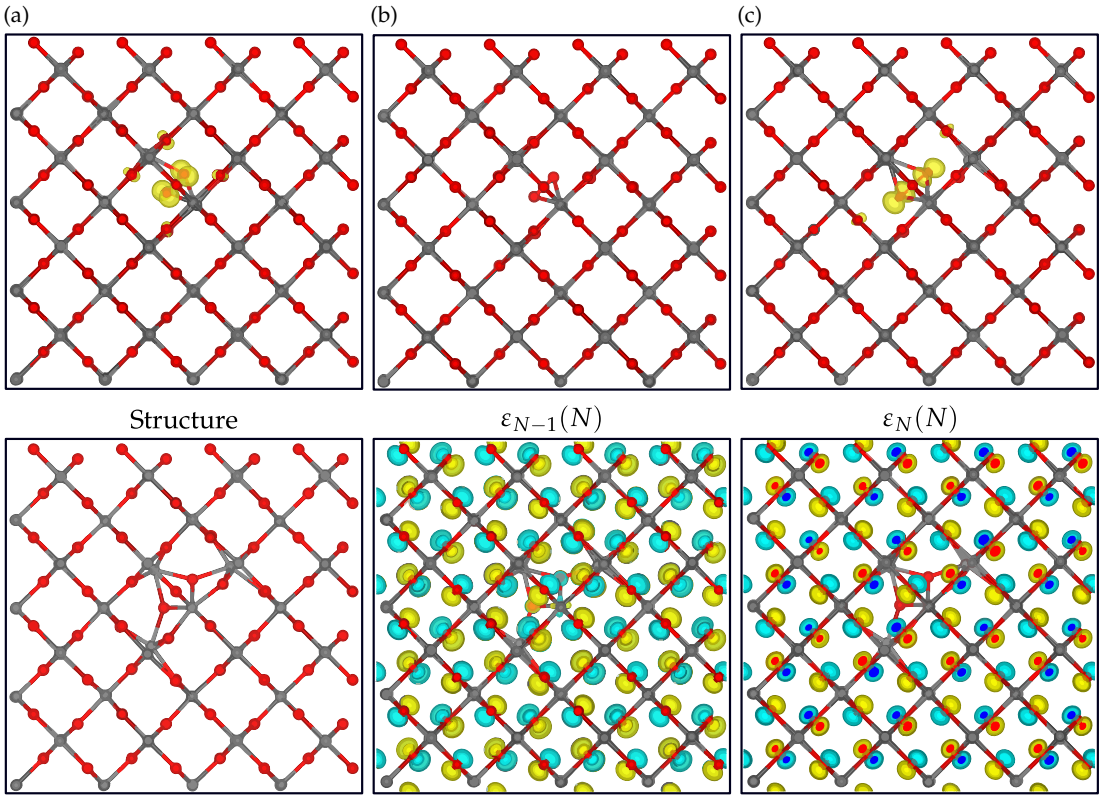

Figure 6.6: *(Upper)* Structure and spin density isosurfaces (where unpaired-spins are present) for O_i^{1+} (a), O_i^0 (b) and O_i^{1-} (c) charge states of the oxygen interstitial defect in rutile. *(Lower)* The Crystal structure and isosurfaces for the $\varepsilon_N(N)$ and $\varepsilon_{N-1}(N)$ orbitals are shown for the O_i^{2-} charge state in rutile. The HOMO has very few contributions which lie spatially on the interstitial defect site, while the following occupied orbital does.

form. The midpoint of the O_2^{2-} bond lies near the defect-free lattice oxygen site. The atoms surrounding the defect are also distorted as a consequence. The O_i^{1+} and O_i^{1-}

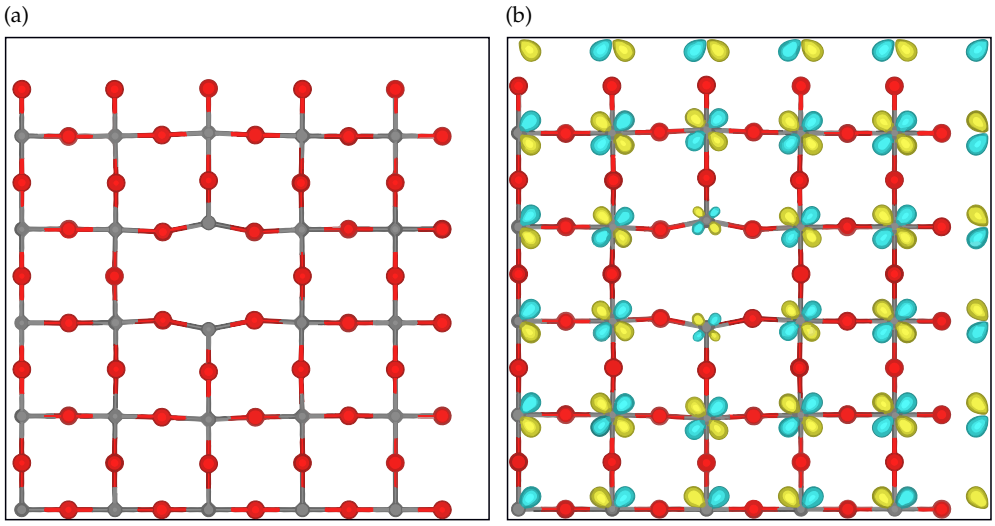

Figure 6.7: Isosurfaces and structures for the anatase V_O^{2+} (a) and V_O^{1+} (b) oxygen vacancy defect.

charge states are found to have excess charge localised to the defect site. The addition of an electron (to O_i^0) pushes the two oxygen atoms further apart by 0.43 Å, while the addition of a hole moves them closer together by 0.09 Å. The trapping energies for O_i^{1-} and O_i^{1+} are found to be -0.54 eV and -0.46 eV, indicating that these are stable traps. The O_i^{2-} charge state does not form localised solutions, however it can form spontaneously under *n*-type (and O-rich) conditions in rutile. This defect pushes the oxygen pair further apart, making the bond length 0.68 Å longer than the V_O^0 configuration. Interestingly, this defect produces no gap states as its two highest-energy orbital eigenvalues lie beneath the VBM, as shown in Fig. 6.6. With reference to the CBM, it can be seen that the O_i^{1-}/O_i^{2-} transition is much larger than the O_i^0/O_i^{1-} transition, making this a negative-U defect.[282] As seen in Fig. 6.5, the O_i^{1-} charge state is not thermodynamically stable compared to other defects. The previously unreported O_i^{1+} configuration is found to be the most thermodynamically stable defect under *p*-type conditions.

Vacancies in anatase

The positively charged V_O^{2+} defect in anatase, whose structure is shown in Fig. 6.7, causes a distortion in the underlying crystal as in the rutile case. While a localised solution can be found for the V_O^{1+} charge state, a delocalised one, shown in Fig. 6.7, is 0.52 eV lower in energy. The lowest energy V_O^0 configuration explored was one with two charges delocalised. This solution was found to be 0.47 eV lower in energy than

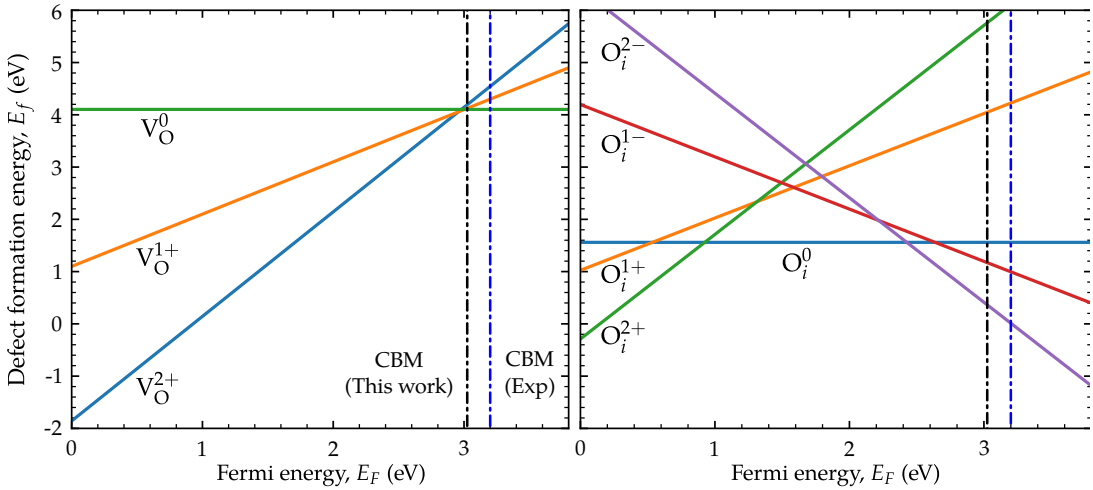

Figure 6.8: Variation of defect formation energy E_f against Fermi energy E_F in anatase under O-rich conditions for the oxygen vacancy defect (*left*) and oxygen interstitial defect (*right*).

one with a single charge localised and 0.95 eV lower in energy than a localised singlet configuration in the vicinity of the defect. This suggests that the behaviour of oxygen vacancies in anatase is fundamentally different to the rutile case, as oxygen vacancies (in the bulk) do not appear to be a prominent source of electron trapping. Fig. 6.8 shows that all transition levels for this defect lie close to the CBM. For most Fermi energy values, the V_O^{2+} defect is most thermodynamically stable.

Interstitials in anatase

The range of stable charge states for this defect goes from -2 to +2, as illustrated in Fig. 6.8. The calculated lowest energy structures and molecular orbital isosurfaces are shown in Fig. 6.9 for each charge state. With the exception of the neutral charge state, all oxygen interstitial defects are made more stable by localising excess charges near the defect. Similar to the rutile case, the added oxygen atom displaces a lattice oxygen, leaving a bonded O_2^{2-} unit. Adding an electron to the crystal increases the oxygen-oxygen bond length to 1.91 Å, making it 0.46 Å longer than the neutral configuration. Adding a second electron increases this distance to 2.16 Å, which is comparable to the titanium-oxygen bond distance in bulk TiO_2. Adding one hole to the O_2^{2-} pair reduces the bond length by 0.08 Å. The addition of a second hole causes the atomic configuration to change entirely. The resulting structure is one where the interstitial oxygen is bonded to two lattice oxygens, forming O_3^{2-} as seen in Fig. 6.9. Unlike the rutile case, all charged oxygen interstitial defects leave a gap state present. Under *p*-type (and O-rich) conditions, the O_i^{2+} defect can form spontaneously, while *n*-type condi-

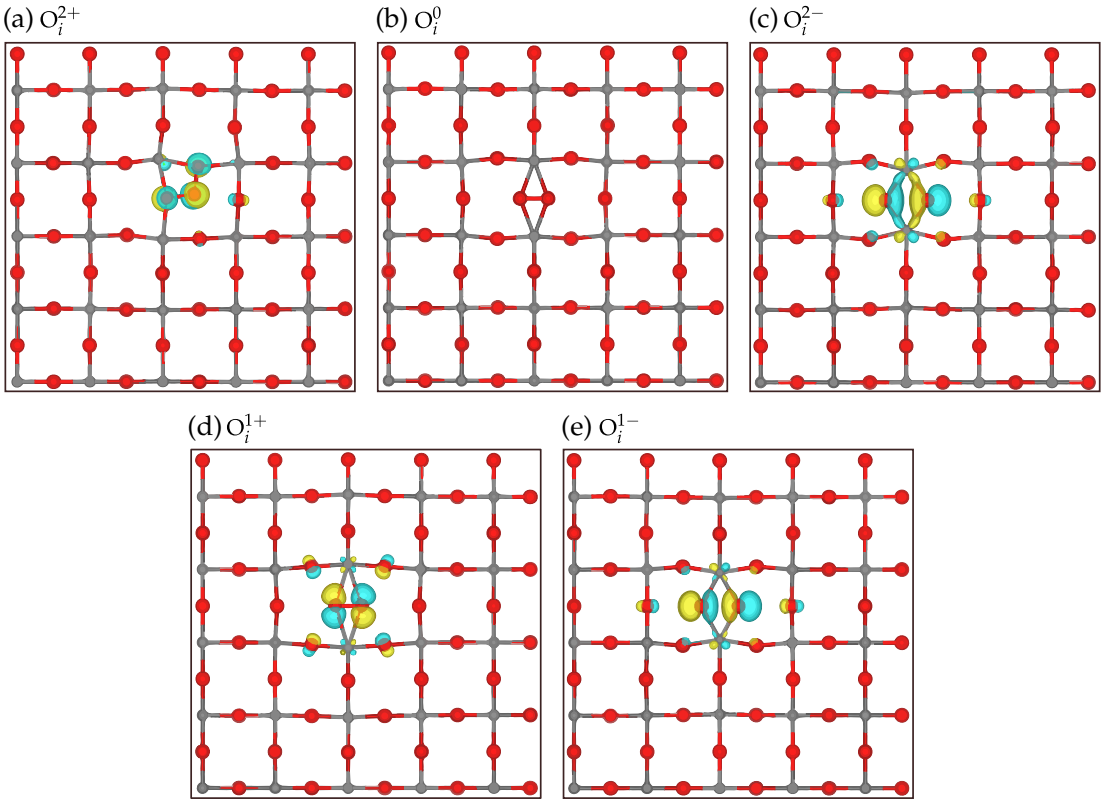

Figure 6.9: Structures and molecular orbital isosurfaces are shown for the thermodynamically stable range of charge states for the oxygen interstitial defect in anatase. Isosurfaces are constructed for the 0.05 a_0^{-1} value. The orbitals used to construct these isosurfaces are the LUMO (a), HOMO (c), LUMO (d) and HOMO (e). The HOMO up- and down-spin isosurfaces for the O_i^{2-} charge state only change in phase, so only the up-spin is shown here. The same is true for the LUMO up- and down-spin isosurfaces in O_i^{2+}.

tions can produce O_i^{2-}. For charge states with an even number of electrons, a spin singlet is formed, while odd numbered systems form a spin doublet. The transition levels for the charges surrounding the interstitial defect are much larger than polaron self-trapping energies calculated for anatase in Chapter 5.

6.5 Titanium defects

Vacancies in rutile

Neutral titanium vacancies in rutile are a *p*-type defect, donating four holes into the crystal. The V_{Ti}^{4-} charge state of this defect contains four excess electrons that fill the holes created by the vacancy, resulting in a vacancy and an accompanied distortion of nearby ions in the nearby crystal. This is illustrated in Fig. 6.10. The addition of a hole to this configuration results in V_{Ti}^{3-}, whose calculated lowest energy configuration places a hole localised in the vicinity of this defect. The addition of one and two more

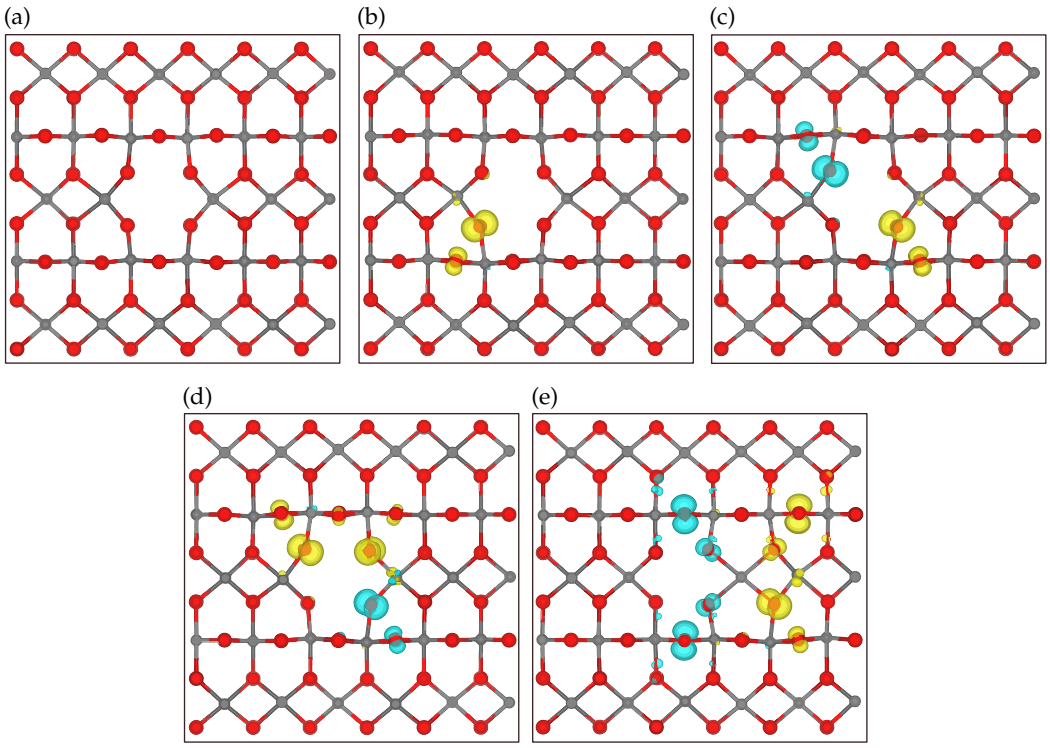

Figure 6.10: Structures and molecular orbital isosurfaces are shown for the thermodynamically stable range of charge states for the titanium vacancy defect in rutile. Isosurfaces are constructed for the 0.05 a_0^{-1} value. The charge states of the above isosurfaces are -4 (a), -3 (b), -2 (c), -1 (d) and 0 (e).

holes results in a similar configuration, with all three holes localised near the defect as shown in Fig. 6.10. In the V_{Ti}^{3-}, V_{Ti}^{2-} and V_{Ti}^{1-} cases all holes are mainly localised near on oxygen ions directly near the vacancy site. The addition of a fourth hole results in a change in configuration, with holes localising on oxygen ions near, but further along from the defect site, as illustrated in Fig. 6.10. The transition levels for these defects are shallow, but still sizable, as seen in Fig. 6.12. The V_{Ti}^{4-}/V_{Ti}^{3-} transition lies 0.16 eV above the VBM, and the V_{Ti}^{3-}/V_{Ti}^{2-} transition is 0.1 eV above the VBM. No other charge states are found to be stable above the valence band edge. Other tested V_O^0 configurations were higher in energy. One configuration placed charges mainly on two sites and was 0.08 eV higher in energy. A configuration where charges were localised on oxygen ions immediately near the titanium vacancy site was 0.62 eV higher in energy. Of note is that under *p*-type O-rich conditions, the formation energy of this defect is substantially higher than any oxygen defect.

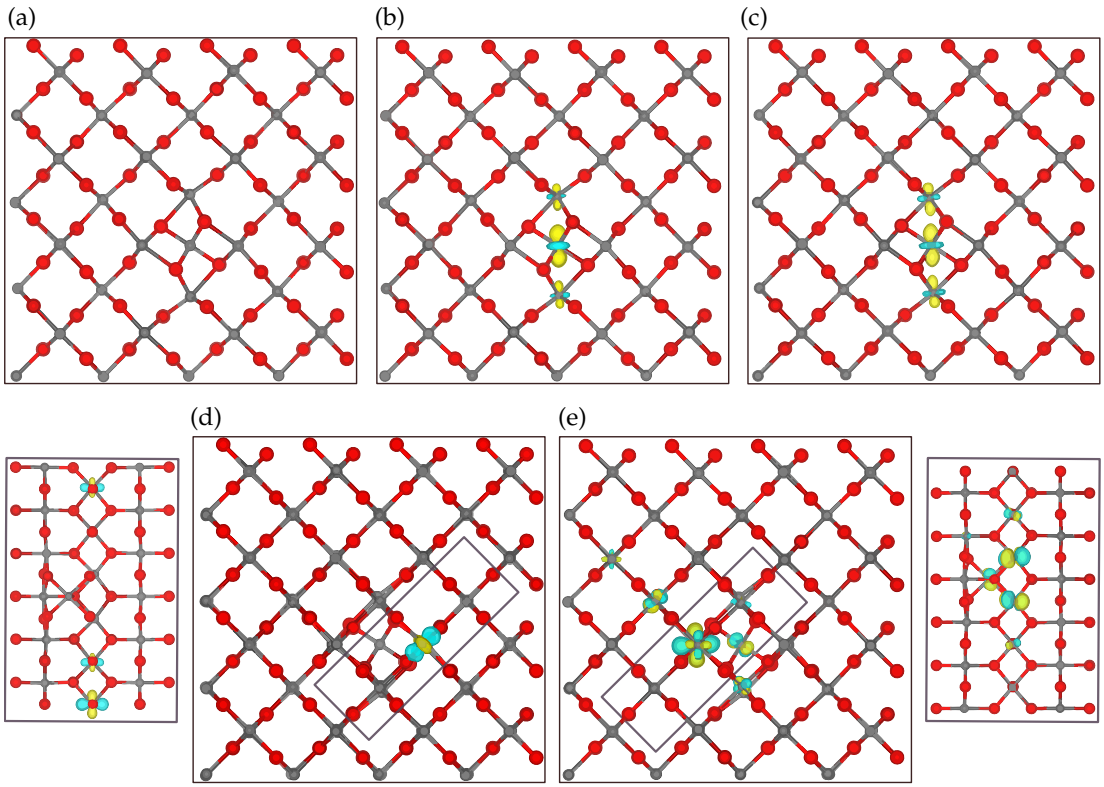

Figure 6.11: Structures and molecular orbital isosurfaces are shown for the thermodynamically stable range of charge states for the titanium interstitial defect in rutile. Isosurfaces are constructed for the 0.05 a_0^{-1} value. The charge states of the above isosurfaces are +4 (a), +3 (b), +2 (c), +1 (d) and 0 (e).

Interstitials in rutile

The neutral titanium interstitial Ti_i^0 is an *n*-type defect, donating four electrons into the crystal. In its Ti_i^{4+} charge state the defect causes a distortion in the surrounding crystal as seen in Fig. 6.11. The most favourable Ti_i^{3+} and Ti_i^{2+} charge state configurations have one and two electrons localise directly on the defect site, as seen in Fig. 6.11. In the Ti_i^{2+} charge state, the additional electrons sit in a singlet configuration. The interstitial defect sits in an octahedral configuration, bonding to six nearby oxygen ions. Two of each of the surrounding oxygen ions have the same bond length to the interstitial, lying 1.86 Å, 1.93 Å and 2.04 Å away from the ion. The additional electron in the Ti_i^{3+} charge state changes these same lengths change to 1.89 Å, 2.02 Å and 2.03 Å. In the Ti_i^{2+} charge state changes these lengths to 1.91 Å, 2.03 Å and 2.04 Å. The transition levels for these defects are quite shallow, lying 0.16 eV $Ti_i^{4+}/Ti_i^{3+}$and 0.10 eV Ti_i^{3+}/Ti_i^{2+}above the valence band edge. From 0.1 eV to the calculated valence band edge, only the Ti_i^{2+} charge state is stable.

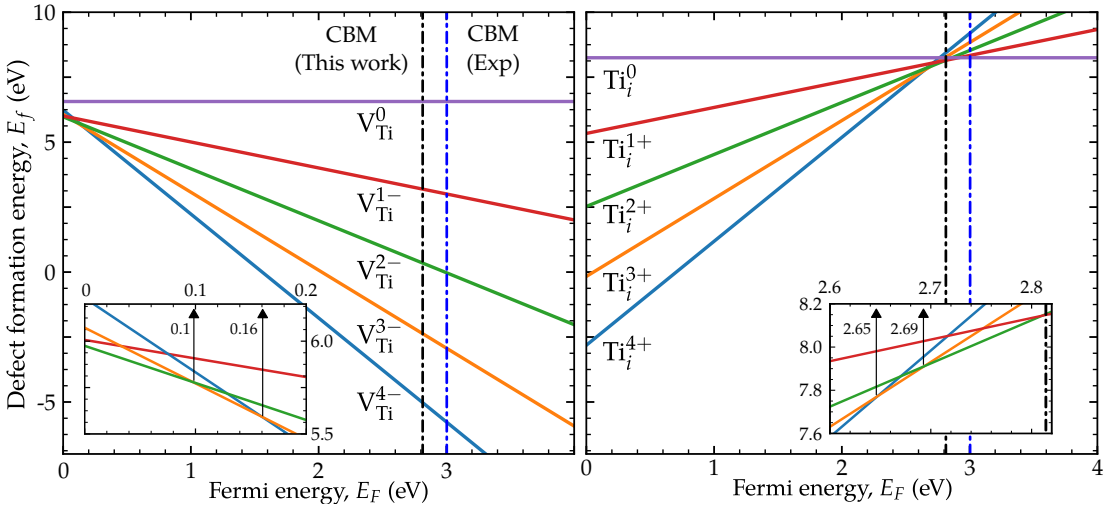

Figure 6.12: Calculated formation energies for the titanium vacancy (left) and titanium interstitial (right) defects in rutile for a number of charge states. (*Left, inset*) A magnified version of the large left figure near the valence band edge. The V_{Ti}^{4-}/V_{Ti}^{3-} and V_{Ti}^{3-}/V_{Ti}^{2-} transitions are 0.16 eV and 0.10 eV above the valence band edge. (*Right, inset*) A magnified version of the large right hand figure, focusing near the calculated conduction band edge (2.81 eV). The Ti_i^{4+}/Ti_i^{3+}, Ti_i^{3+}/Ti_i^{2+} and Ti_i^{2+}/Ti_i^{1+} transitions are at 2.65 eV, 2.69 eV and 2.81 eV above the valence band edge.

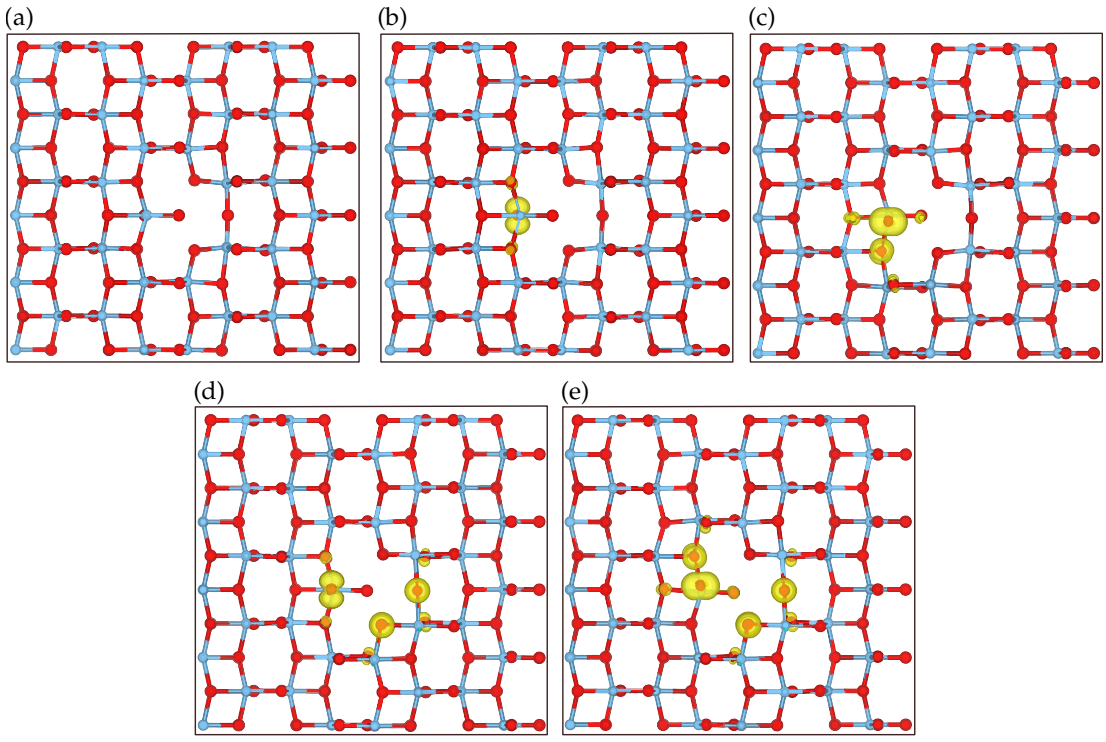

Figure 6.13: Structures and spin density isosurfaces are shown for the thermodynamically stable range of charge states of the titanium vacancy defect in anatase. Isosurfaces are constructed for the 0.05 a_0^{-1} value. The charge states of the above isosurfaces are -4 (a), -3 (b), -2 (c), -1 (d) and 0 (e).

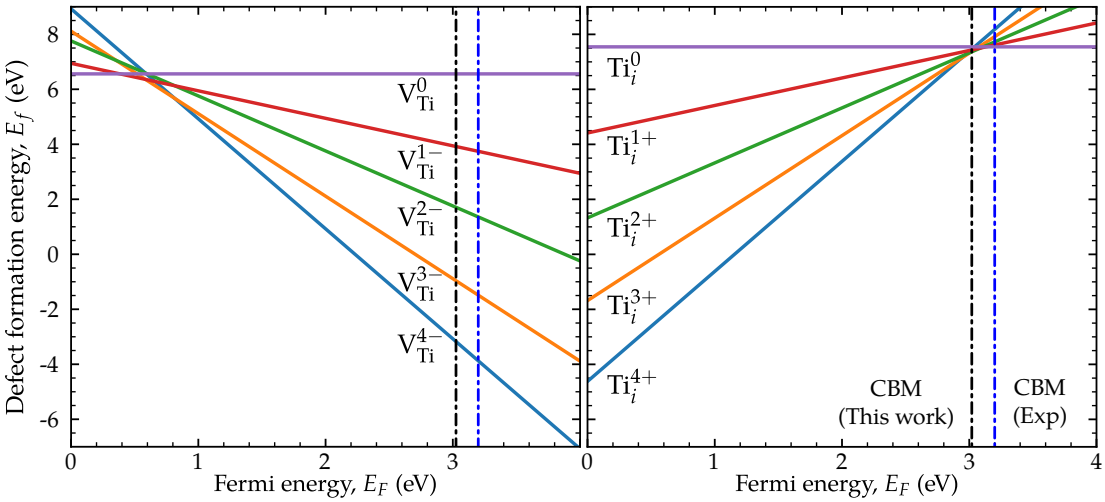

Figure 6.14: Calculated formation energies for the titanium vacancy (left) and titanium interstitial (right) defects in anatase for a number of charge states.

Vacancies in anatase

The neutral titanium vacancy in anatase introduces four holes into the crystal, similar to the rutile case, which get filled as electrons are added into the crystal. As self-trapped holes can form in anatase, it is anticipated that these charges may behave differently to the rutile case. In anatase, these holes are found to localise around the titanium defect site on nearby oxygen atoms, as shown in Fig. 6.13. All charge states with excess charge are found to be much more stable with spins aligned in parallel. In all cases where there is more than one hole present, maximising the total spin of localised charges gives solutions which are 0.11 eV to 0.12 eV more stable than configurations which minimise spin. This may partially explain experimental data which shows room temperature ferromagnetism in titanium deficient anatase.[275] It can be seen from Fig. 6.13 that the transition levels associated with this defect are much deeper than in the rutile case. This difference in behaviour of the two phases shows a trend similar to the self-trapping polaron picture, which is expected as holes localise much more easily in anatase than rutile.

Interstitials in anatase

This defect in anatase behaves differently to the rutile case, with delocalised electron configurations being more favourable than localised ones. In the Ti_i^{3+} case, which contains one excess electron, the localised configuration is found to be 0.16 eV less stable than the delocalised one. Localised charge configurations could not be found

for all other charge states other than Ti_i^{3+}. As seen in Fig. 6.14, all charge states have shallow transition levels, lying very close to the CBM.

6.6 Assessment of Koopmans' theorem

Compliance of Koopmans' condition is now assessed for charges localised around the defects discussed in this chapter. In Chapter 5, deviation from linearity was found to be 0.08 eV at most for small electron and hole polarons across several TiO_2 phases. It can be seen from Table 6.8 that the condition is fulfilled to this degree in 12 of 21 cases. Generally, the GKT is better fulfilled in oxygen defects compared to titanium ones. It can also be seen that different charge states of the same defect have quite different nonlinearities. This implies that different values of α should be used for different charge states. While such an approach may be used to assess the localisation properties of defects, it would not be possible to directly compare calculated defect formation energies. In Table 6.8, a positive non-linearity indicates that lower values of α would satisfy Koopmans' condition, while negative values suggest larger α values. For certain defects, such as the titanium vacancy, it may be the case that charges delocalise before the GKT is satisfied as seen in Chapter 5.

For the oxygen vacancy in anatase, optimising α for each charge state will likely yield localised defect states that are still metastable. Earlier it was found that localised solutions for the V_O^0 defect are 0.95 eV less favourable than delocalised solutions. The V_O^0 charge state would require lower values of α, which in turn would reduce the favourability of the localised solution even more. For the V_O^{1+} charge state, a higher value of α is required. However, data from Chapter 5 suggests that increasing α by 2% increases nonlinearity ξ by around 0.09 eV and increasing trapping energies by a similar amount. As delocalised solutions were more favourable by 0.47 eV (with α=0.115) for V_O^{1+}, this indicates that the localised solution in anatase will be around 0.3 eV higher in energy than a delocalised solution. The implication of this data is that (for reasonable values of α) localised electrons around oxygen vacancies in bulk anatase are much more diffuse than in bulk rutile.

6.7 Comparison of defect formation energies

Fig. 6.15 shows defect formation energies under both titanium and O-rich conditions for the previously discussed defects. It can be seen that the formation energies of V_O

Defect	q/q'	Rutile	Anatase	Defect	q/q'	Rutile	Anatase
V_O	1+/2+	-0.02	-0.12	Ti_i	3+/4+	-0.06	-0.02
V_O	0/1+	+0.07	+0.16	Ti_i	2+/3+	-0.12	–
O_i	2+/1+	-	+0.02	V_{Ti}	3-/4-	+0.30	+0.08
O_i	1+/0	+0.04	-0.05	V_{Ti}	2-/3-	+0.23	+0.12
O_i	1-/0	+0.01	-0.13	V_{Ti}	1-/2-	+0.34	+0.05
O_i	2-/1-	-	-0.06	V_{Ti}	0/1-	+0.05	+0.22

Table 6.8: Koopmans compliance for rutile and anatase defects discussed in this chapter. The nonlinearity ξ (see Eqn. 5.1) shown for each case for the fraction of exact exchange $\alpha = 0.115$. The charge states of $E(N)$ and $E(N-1)$ are presented as q/q'. This is done for all localised defects, even those where a localised solution was found not to be energetically stable.

and Ti_i are similar in both phases under *n*-type conditions. These defects are also the lowest energy ones in both *n*- and *p*-type Ti-rich environments. These defects only form transition levels below the conduction band edge in rutile, supporting the argument that electrons in bulk anatase tend to be of a much more diffuse nature. For hole states the reverse is true, which are more shallow in rutile than anatase. The deepest transition levels in both phases are caused by the oxygen interstitial. Fig. 6.15 highlights that the lowest energy *p*-type defect under *p*-type conditions is the oxygen interstitial, with at least a 2 eV lower formation energy than the titanium vacancy in all cases. Of note is that formation energies for titanium vacancies are extremely large under *p*-type Ti-rich conditions and have a formation energy of over 10 eV in both phases.

Killer defects in materials are often discussed as those which either prevent doping of specific charge carriers, or defects which cause catastrophic failure of the device a material is used for. Under O-rich conditions, the titanium vacancy and titanium interstitial are the main killer defects in both materials, preventing *n*- and *p*-type doping. For example, doping TiO_2 *n*-type would reduce titanium vacancy formation energies, increasing the defect's concentration. This defect would in turn absorb any excess electrons, negating the effect of doping. It can be seen (from Fig. 6.15) that Ti-rich conditions shift this equillibrium, making formation energies of defects preventing *n*-type doping very high. Under O-rich conditions the oxygen interstitial in anatase may prevent both *n*- and *p*-type doping, making this a more severe, or perhaps even *serial* killer defect.

Under *n*-type conditions, the relative stability of titanium interstitials to oxygen vacancies changes sign when going from O-rich to Ti-rich conditions, as seen in Fig. 6.15.

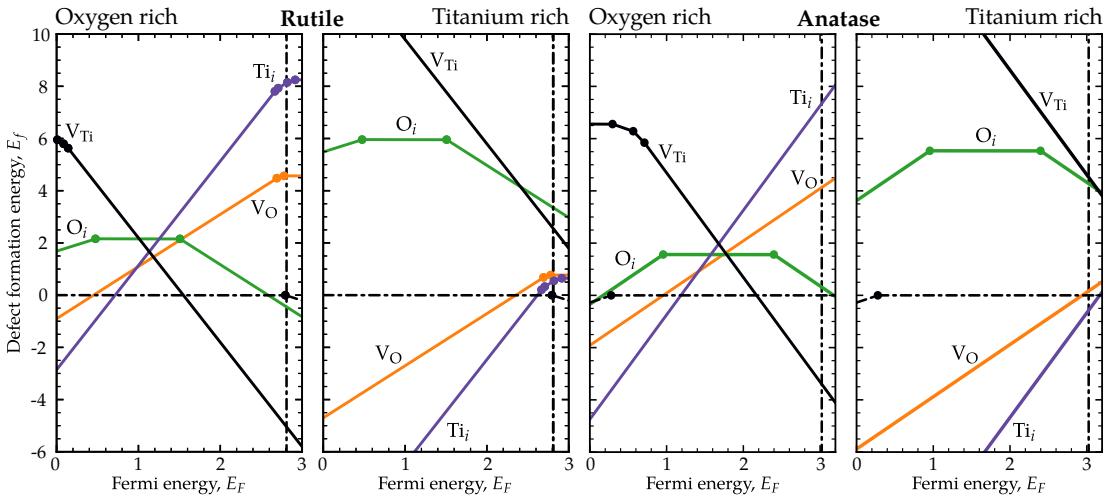

Figure 6.15: Aggregated defect formation energies for the most thermodynamically stable defects in rutile (*left two panels*) and anatase (*right two panels*). For both phases, these energies are shown for O-rich and Ti-rich conditions, corresponding to the energy of an isolated oxygen molecule and the the TiO_2/Ti_2O_3 equillibrium. (*Horizontal dashed line*) The transition from defect free bulk to a self-trapped polaronic configuration. (*Vertical dashed line*) the defect free bulk conduction band minimum. Defects which produce shallow delocalised states have been omitted.

This finding is in line with previous arguments suggesting that titanium interstitials play a larger role in heavily reduced samples, whereas oxygen vacancies will be more dominant in lightly reduced samples.[235] As Ti-rich chemical potentials correspond to those where the formation energies of TiO_2 and Ti_2O_3 are equal, they likely correspond to slightly extreme Ti-rich conditions. For this reason, true experimental chemical conditions are likely to correspond to chemical potential values that lie between the two limits shown in Fig. 6.15.

6.8 Single-particle defect levels

Fig. 6.16 indicates that the Kohn-Sham eigenvalues for the O_i^{1+} defect in anatase produce a gap state that is 1.4 eV above the VBM, and two very deep hole gap states 0.05 eV below the CBM for the O_i^{2+} charge state. In contrast to the positive charge state, the O_i^{2-} charge state in anatase produces gap states that are less deep, but still sizeable, lying 1.1 eV above the VBM. The O_i^{1-} defect in rutile produces an extremely deep electron state, lying 0.02 eV above the VBM. In the opposing spin channel (not shown in Fig. 6.16), an unoccupied state 2.4 eV above the VBM appears. The additional electron introduced by the O_i^{2-} charge state results in two electrons lying beneath the VBM.

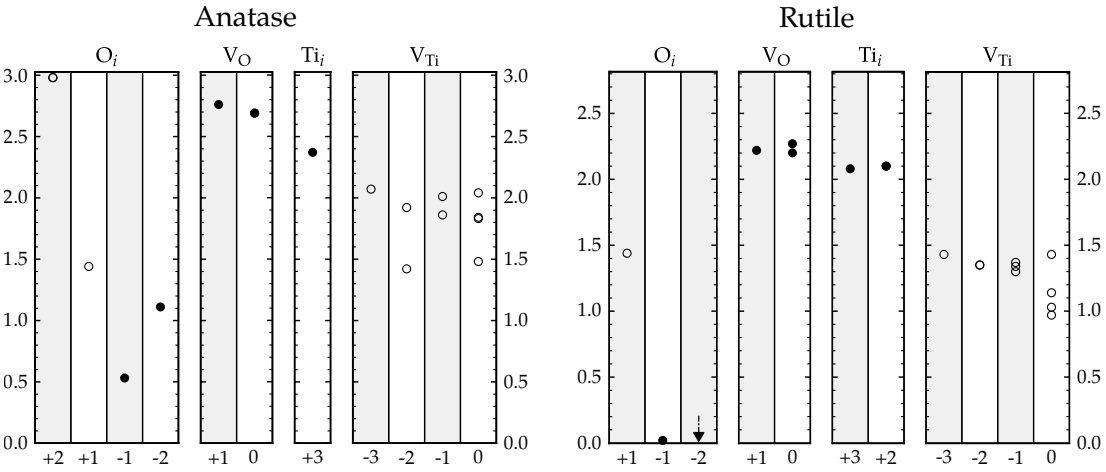

Figure 6.16: Single-particle eigenvalues of charges surrounding the defects discussed in this chapter shown between the VBM (at zero) and calculated CBM. Occupied levels are shown as filled circles, while unoccupied ones are unfilled. Localised anatase V_O and Ti_i levels are also shown, although they are found to be less stable than delocalised solutions. The O_i^{2-} defect in rutile has levels beneath the valence band edge, indicated with an arrow.

The behaviour of single-particle levels relating to the oxygen interstitial defect is unexpected, and has recently been noted by other authors[283] with other materials. These include cation vacancies in MgO, Li-ion battery materials and rare-earth nickel oxides. Dubbed *antidoping*, this effect occurs when unoccupied conduction band levels merge with the valence band on occupation. The presence of this effect would in turn cause a reduction in conductivity. As shown in Fig. 6.16, this behaviour occurs with O_i^{2-} in rutile and O_i^{2+} in anatase, with the effect being more pronounced in rutile. This indicates that oxygen interstitials would significantly reduce electron mobility in rutile and hole mobility in anatase. This information suggests that oxygen interstitials in rutile and anatase display a variety of characteristics, being negative-U, killer, and anti-doping defects.

6.9 Discussion

EPR measurements have shown that the oxygen vacancy defect in rutile produces two spin-parallel localised states, with the first state having a low thermal activation energy ($\sim$ 2 meV). This finding is similar to those found in our theoretical predictions, where the V_O^0/V_O^{1+} and V_O^{2+}/V_O^{1+} transition levels were found to be around 0.02 eV and 0.12 eV. These findings indicate that only one of the two electrons made from an oxygen vacancy is easily donated to the crystal. This behaviour is in direct contrast with the anatase phase, where it is found that delocalised solutions are favoured by

more than 0.4 eV. As oxygen vacancy defects are more dominant than titanium interstitial in the mildly reduced limit, the difference in conductivity between rutile and anatase samples may be explained, at least in part, by the behaviour of the oxygen vacancy defects in these phases.

The oxygen interstitial defect is perhaps the most interesting one discussed, having both positive and negative stable charge states in both phases. As mentioned earlier in this chapter, experimental studies directly probing this defect in rutile are scarce. When considered in light of the findings shown in this chapter, this may seem less puzzling. Specifically, the most stable negative charge state for this defect is found to induce delocalised states beneath the valence band minimum, which will be EPR silent. Moreover, as no occupied single-particle states are expected to form far within the band edges (for either O_i^{1-} and O_i^{2-}), states associated with this defect may be difficult to probe using PES. A proposed method of detection for this defect would be conductivity measurements, which should be reduced under *n*-type conditions when this defect is present. Under *p*-type conditions this defect would be easier to probe in rutile, as a deeply trapped localised hole with a gap state (around 1.6 eV above the VBM) is expected. Finally, as this defect is expected to have a formation energy around three times lower than the titanium vacancy under *p*-type O-rich conditions, it is anticipated that this will be the dominant *p*-type defect.

In anatase, it is expected that the oxygen interstitial displays both electron and hole trapping behaviour, with holes being more strongly trapped. Holes trapped by this defect will be localised but lie nearer to the conduction band edge, making them slightly more difficult to detect. In contrast, trapped electrons are anticipated to induce gap levels around 1.1 eV above the VBM. One experimental study finds that Nb-doped anatase with excess oxygen has a metal-to-insulator transition when annealed at an oxygen partial pressure of 4.6 kPa.[262] The interstitial was found to significantly reduce conductivity, in line with our findings. Another experimental study found a reduced gap in excess oxygen samples due to a rising VBM.[261] While such an effect was not observed in this study, it may be due to differences in defect concentration.

Prolonged exposure to high oxygen partial pressures have been shown to cause an *n*- to *p*-type conduction transition in rutile. Nowotny *et al* proposed that this effect is largely due to titanium vacancy diffusion into the bulk crystal.[263] While such a mechanism is no doubt feasible, no concrete evidence for this is provided. As shown previously, titanium vacancies in rutile have a formation energy around 4 eV higher

than oxygen interstitials under *p*-type O-rich conditions. The formation of a Frenkel defect may be more likely. The findings in this chapter, as well as other theoretical studies, indicate that titanium vacancies may play a larger role in reducing *n*-type conductivity than they do inducing *p*-type conductivity. As mentioned by other authors, oxygen interstitials may play a larger role in this transition.[257] The findings in this chapter, specifically those surrounding the O_i^{2-} charge state of the oxygen interstitial, straightforwardly explain reduced conductivity under high partial oxygen pressures.

Previously discussed experiments shows that under O-rich conditions anatase can turn into a *p*-type semiconductor, verified through a positive Hall co-efficient measurement. The simplest point of comparison to this work is through the observed ferromagnetism, where spin parallel solutions are found to be 0.11-0.12 eV more stable than spin paired solutions for all possible charge states. This behaviour is quite distinct from that of rutile, where the lowest energy configurations are those where holes are spin paired.

Previous theoretical studies on titanium interstitials in rutile have shown that configurations localising between one and two electrons are most stable, with electrons on either the titanium interstitials themselves or on nearby sites. The findings shown from this chapter suggest that up to that two electrons trap directly on the defect site. In the two electron case, one is much more strongly bound than the second, with transition levels for both of 0.10 eV and 0.16 eV. Previously discussed EPR data relating to this defect in rutile suggests that only a negligible signal amount can be attributed to bulk interstitials.[190] Other authors argue that this signal should in fact be attributed to substitutional Ti, not interstitials.[274] Data from this chapter suggests that this defect will not always produce an EPR signal in the bulk due to the spin-pairing of signals. The single-particle states for this defect are found to lie 0.7 eV below the CBM (in this work), indicating deeper levels than those in the oxygen vacancy. Continuing the trend seen with other defects, anatase shows a reduced ability to localise electrons on both the titanium interstitial and its surrounding atoms; an electron localised on the interstitial has an energy 0.16 eV higher than a delocalised solution. Results shown in this chapter indicate electrons will not trap on this defect, perhaps explaining why EPR signals have not yet been unambiguously attributed to this defect.[199,276]

6.10 Conclusion

The work shown in this chapter supports the argument that electrons tend to more localised in rutile than anatase. As TiO_2 samples generally tend to be oxygen deficient, the oxygen vacancy and titanium interstitial defects are perhaps the most significant; both of these defects result in electrons trapping more strongly in rutile compared to anatase. Delocalised electrons were found to be 0.52 eV, 0.95 eV and 0.16 eV more stable than localised ones for V_O^{1+}, V_O^0 and Ti_i^{3+} in anatase. The findings may at least partially why Nb-doped anatase has a conductivity four orders of magnitude larger than Nb-doped rutile.[47] The reverse is found to be true with holes, which are more deeply trapped in anatase than rutile. The main *n*- and *p*-type trapping defect found was the oxygen interstitial, which traps both electrons and holes more strongly than any other defect in both phases. This defect also displays interesting characteristics, being a negative-U, killer and antidoping defect. Positive charge states for this defect have been explored little in literature, with only one hybrid functional study on anatase. For applications which require high charge mobility, avoidance of this defect should be a key priority.

The results shown in this chapter illustrate that the transferability of Koopmans condition seen in Chapter 5 extends to some intrinsic defects but not others. Generally oxygen defects are better compliant with the GKT than titanium ones. The main defect found not to be Koopmans compliant was the titanium vacancy, with all charge states requiring a lower value of α to be used. It was also found that for some defects, such as the titanium vacancy, different charge states would require different values of α.

CHAPTER 7

CONCLUSIONS

Developing accurate models of charge trapping phenomena in semiconductors is vital to both understanding and developing technologies used in many new cutting-edge devices, such as the high-capacity batteries and nanoporous photovoltaics described in Chapter 1. While several approaches to tackle this problem exist, as demonstrated throughout this thesis, many are either prohibitively expensive or lack the predictive power needed to accurately model charge trapping. As a result, progress in the advancement of technologies can be hampered and made more expensive through the reliance of theoretical approaches on experimental data.

In this thesis an inexpensive, accurate and predictive approach to modelling charge trapping was identified and thoroughly tested on a set of challenging systems. By comparing to both exact theoretical models, as in Chapter 4, and experimental data on TiO_2, the use of physically justified constraints, such as the generalised Koopmans' theorem, with hybrid functionals was shown to yield extremely accurate results. As a result, the models developed in this thesis were subsequently used to investigate less-understood polymorphs of TiO_2 in Chapter 5, and intrinsic defects in Chapter 6 where there is a lack of consensus of the behaviour of excess charges. Outside this thesis, the approach outlined here has already been used to model charge trapping in more complex extended defects such as surfaces and grain boundaries.[284–286]

The primary focus of Chapter 4 was to assess the accuracy of electron densities and energies of hybrid functionals overall, and whether the constraint provided by the generalised Koopmans' theorem sufficiently captures exchange and correlation effects present in model systems. It was demonstrated, by comparing to the exact solution to the many-electron Schrödinger equation, that the approach is indeed sensitive to correlation effects and provides extremely accurate electron densities and energies. In addition to the generalised Koopmans' theorem, two further physically justified

constraints were also suggested and tested. An alternative mixing strategy in hybrid functionals, yielding superior densities, was identified and compared with more conventional schemes. The Koopmans' approach was found to be free from the fractional dissociation problem common in many DFT approximations, thus providing the correct charge localisation properties. The physical intuition behind this approach was found, when viewed from the perspective of many-body perturbation theory, to bear resemblance to a screened-exchange approximation, similar to the GW method.

Chapter 5 took the approach assessed in Chapter 4 further by using theoretical techniques to reduce the cost HF exchange, such as ADMM, Coulomb interaction truncation and Cauchy-Schwarz screening. The errors introduced by these techniques were assessed and found to be sufficiently small, enabling accurate calculations at low computational cost. The coulomb-truncated hybrid functional was parameterised using the Koopmans' condition for every localised polaronic state, where a number of significant findings were noted. First, the accuracy of single-particle band gaps was within six percent of experimental data. Second, parameters obtained from Koopmans' condition have a degree of transferability which can be beneficial in extended defects, such as surfaces and grain boundaries, where chemical environments may vary more dramatically across a simulation cell. Finally, the polaronic properties of all TiO_2 phases were compared, providing useful information to guide future experimental investigations.

It is evident from information provided in Chapter 6 that there is a lack of consensus, both experimentally and theoretically, on the behaviour of intrinsic point defects in TiO_2 rutile and anatase. The functional parameterised in Chapter 5 was used to probe these defects, referencing previous data when possible, with the aim of gaining a better understanding of excess charge behaviour. The substantially higher conductivity of mildly reduced anatase over rutile could be partially explained by the behaviour of oxygen vacancies, where excess electrons strongly preferred to be delocalised in anatase. Titanium interstitials may also explain this behaviour, where a one-electron localised solution was found to be metastable in anatase. Interstitial defects in both phases displayed interesting characteristics, being killer, antidoping and negative-U defects. Interestingly, oxygen interstitial traps may bind pairs of conduction band electrons into states beneath the valence band minimum, meaning these defects may be difficult to defect using EPR or optical techniques. A useful technique to correct eigenvalues was also discussed, which may prove helpful in applying Koop-

mans' theorem to systems where dielectric constants are small.

7.1 Future work

While the work presented in this thesis provides a route to using physically justified parametrisations of inexpensive hybrid functionals, it has not been extensively tested on materials other than TiO_2. A number of materials identified in Chapter 1 could be used to further validate the Koopmans' approach. Alkali-metal oxides can form a variety of polarons and have a relatively low dielectric constant, making them ideal for further testing with eigenvalue corrections. The alternative hybrid xc mixing scheme, shown in Chapter 4, could also be validated on real materials. This scheme was demonstrated to have reduced errors in the density, potentially increasing the accuracy of hybrid functional calculations. Finally, extending the approaches assessed in Chapter 4 to both molecules in 3D and extended states would be of great use to materials modelling.

Alternative approaches to fitting the fraction of exact exchange α can also be further investigated. One notable approach, briefly mentioned in Chapter 5, ensures that $\alpha = \varepsilon_\infty^{-1}$ where ε_∞ is the electronic contribution to the dielectric constant. Previous work has shown that application of such an approach to rutile yields $\alpha = 0.15$, which lies pleasingly close to the $\alpha = 0.115$ determined in Chapter 5.[287] Investigating differences in predictions made by the tr-PBEα, HSE and PBEα functionals available in `CP2K` and `VASP` would also be useful. Finally, it would be helpful to determine how HSE predictions change using various fixed values of ω, and α tuned to the GKT.

Appendix A

Slab vs HEG based LDAs

In order to verify that LDAs constructed from slabs (i.e. finite systems) perform in a similar fashion to HEG-based LDAs, the following two figures were made using slab- and HEG-based LDAs.

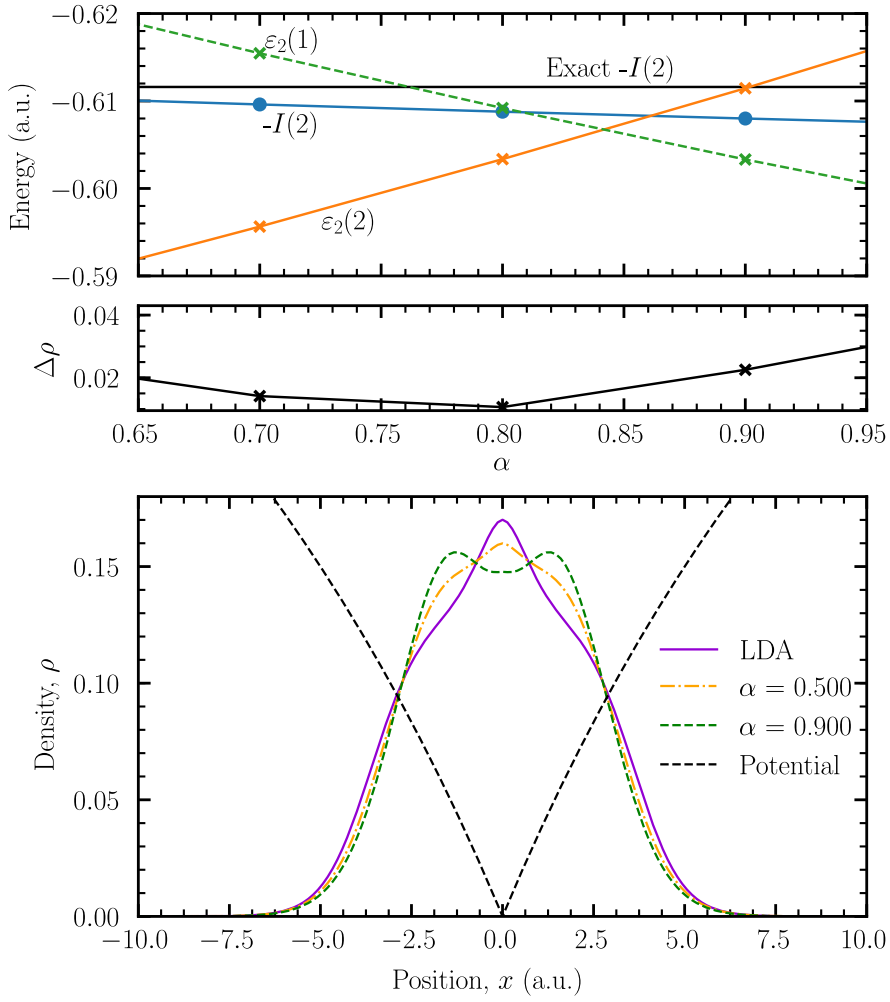

Figure A.1: Energies and densities calculated from a slab-based LDA for a two-electron system.

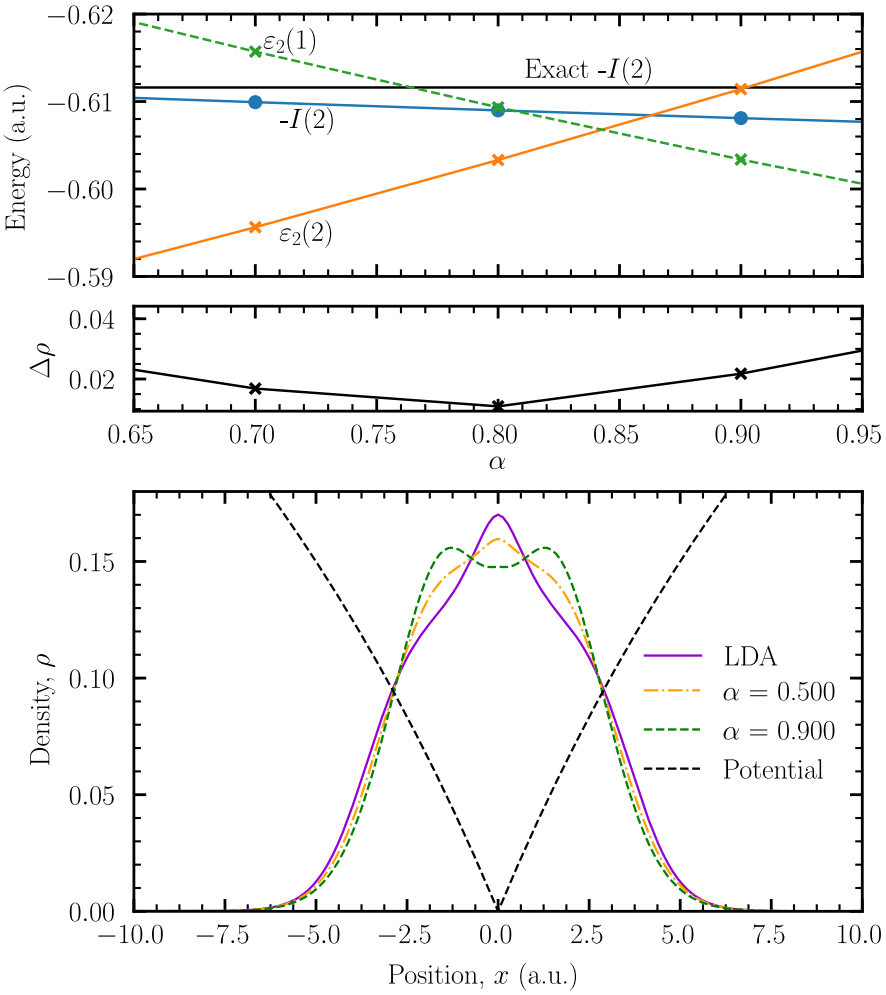

Figure A.2: Energies and densities calculated from a HEG-based LDA for a two-electron system.

APPENDIX B

POLARONS IN TITANIA POLYMORPHS: ADDITIONAL INFORMATION

This appendix contains information relevant to Chapter 5 that a reader may be interested to see, but is not central to the main text.

Cell size convergence

	Supercell A (Å)					Supercell B (Å)				
	Size	N	a_x	b_y	c_z	Size	N	a_x	b_y	c_z
Rutile	4×4×6	576	18.347	18.347	17.702	5×5×8	1200	22.933	22.933	23.600
Anatase	5×5×4	600	18.877	18.877	19.140	6×6×5	1080	22.658	22.658	23.910
TiO_2(B)	5×3×3	540	18.658	18.289	18.738	6×4×4	1152	22.394	24.381	24.985

Table B.1: Shown supercells used for size convergence. While the main text uses conventional cells for all phases, the irreducible cells for TiO_2(B) and anatase are used in this section. This gives access to a greater range of supercell sizes.

	Lattice vector a_x (Å)		Lattice vector b_y (Å)		Lattice vector c_z (Å)		Band gap E_g (eV)	
Cell	A	B	A	B	A	B	A	B
Rutile	4.587	4.587	-	-	2.950	2.950	4.08	4.08
Anatase	3.775	3.776	-	-	4.785	4.782	4.30	4.33
TiO_2(B)	3.732	3.732	6.096	6.095	6.426	6.426	5.14	5.10

Table B.2: Cell size convergence data is shown for the rutile, anatase and TiO_2(B) primitive cells. Of note is that the main text uses conventional anatase and TiO_2(B) cells.

Range-seperation convergence

	Lattice vector a (Å)				Lattice vector b (Å)			
R_c (Å)	2.0	4.0	6.0	8.0	2.0	4.0	6.0	8.0
Rutile	4.570	4.586	4.587	4.587	-	-	-	-
Anatase	3.760	3.771	3.772	3.772	-	-	-	-
TiO_2(B)	12.150	12.184	12.189	12.190	3.726	3.733	3.734	3.734

	Lattice vector c (Å)				Band gap E_g (eV)			
R_c (Å)	2.0	4.0	6.0	8.0	2.0	4.0	6.0	8.0
Rutile	2.947	2.949	2.950	2.951	3.45	4.05	4.08	4.09
Anatase	9.573	9.585	9.589	9.588	3.79	4.23	4.30	4.31
TiO_2(B)	6.229	6.245	6.246	6.246	4.60	5.06	5.13	5.15

Table B.3: HF truncation radius R_c convergence data is shown for the rutile, anatase and TiO_2(B) phases. Here the supercells are constructed from the conventional unit cells.

Supercell information

Phase	Atoms	Supercell size	Crystal symmetry	Space group	Inequiv. Ti atoms	Inequiv. O atoms
Anatase	600	5×5×2	Tetragonal	$I4_1/amd$	1	1
Rutile	576	4×4×6	Tetragonal	$P4_2/mnm$	1	1
Brookite	576	2×3×4	Orthorhombic	$Pbca$	1	2
TiO_2(H)	576	2×2×6	Tetragonal	$I4/m$	1	2
TiO_2(R)	576	4×2×6	Orthorhombic	$Pbnm$	1	2
TiO_2(B)	720	2×5×3	Monoclinic	$C2/m$	2	4

Table B.4: Crystal symmetries, simulation cell sizes and space groups are shown for each phase studied in this work. For each crystal, we show the number of symmetrically inequivalent titanium and oxygen sites.

Comparing phase-specific and transferable α values

Quantity	Phase α	Transferable α
Alpha value	0.105	0.115
O^{An} E_t (eV)	-0.21	-0.25
O^{An} CTL (eV)	1.20	1.26
Band gap (eV)	2.94	3.03
Lattice vector a (Å)	3.793	3.792
Lattice vector c (Å)	9.617	9.614

Brookite			TiO_2(B)		
Quantity	Phase α	Transferable α	Quantity	Phase α	Transferable α
Alpha value	0.105	0.115	Alpha value	0.120	0.115
O_a^{Br} E_t (eV)	-0.25	-0.24	O_a^{B} E_t (eV)	-0.42	-0.40
O_b^{Br} E_t (eV)	-0.16	-0.21	O_d^{B} E_t (eV)	-0.54	-0.52
O_a^{Br} CTL (eV)	1.11	1.20	O_a^{B} CTL (eV)	1.62	1.52
O_b^{Br} CTL (eV)	1.09	1.15	O_d^{B} CTL (eV)	1.60	1.51
Band gap (eV)	3.20	3.28	Band gap (eV)	3.89	3.80
Vector a (Å)	9.228	9.226	Vector a (Å)	12.236	12.240
Vector b (Å)	5.479	5.476	Vector b (Å)	3.746	3.746
Vector c (Å)	5.151	5.148	Vector c (Å)	6.287	6.289

TiO_2(H)			TiO_2(R)		
Quantity	Phase α	Transferable α	Quantity	Phase α	Transferable α
Alpha value	0.130	0.115	Alpha value	0.125	0.115
Ti^{H} E_t (eV)	-0.14	-0.10	Ti^{R} E_t (eV)	-0.35	-0.31
O_a^{H} E_t (eV)	-0.42	-0.35	O_a^{R} E_t (eV)	-0.31	-0.25
O_b^{H} E_t (eV)	-0.34	-0.26	O_b^{R} E_t (eV)	-0.43	-0.38
Ti^{H} CTL (eV)	-0.82	-0.67	Ti^{R} CTL (eV)	-0.82	-0.71
O_a^{H} CTL (eV)	1.71	1.53	O_a^{R} CTL (eV)	1.32	1.18
O_b^{H} CTL (eV)	1.59	1.39	O_b^{R} CTL (eV)	1.48	1.39
Band gap (eV)	3.76	3.63	Vector a (Å)	4.950	4.952
Vector a (Å)	10.236	10.241	Vector b (Å)	9.423	9.428
Vector c (Å)	2.968	2.969	Vector c (Å)	2.967	2.968

Table B.5: Compared are the predicted bulk crystal and polaron properties, such as trapping energies E_t and CTLs, given by the phase-specific and transferable ($\alpha = 0.115$) value. It can be seen that the difference between predicted properties is generally small. For electron polarons (on titanium sites) the CTL is given with respect to the CBM, while hole polarons (on oxygen sites) are given with respect to the VBM.

Polaron property range

	α value range		E_t (eV)		CTL (eV)	
	Min	Max	Min	Max	Min	Max
Ti^{Ru}	0.100	0.140	0.01	-0.07	-0.21	-0.66
Ti^{H}	0.125	0.150	-0.13	-0.20	-0.80	-0.92
Ti^{R}	0.120	0.145	-0.33	-0.45	-0.79	-0.93
O^{An}	0.095	0.115	-0.17	-0.25	1.12	1.28
O_a^{Br}	0.085	0.105	-0.15	-0.25	0.96	1.11
O_b^{Br}	0.100	0.115	-0.14	-0.21	1.05	1.17
O_a^{H}	0.110	0.130	-0.33	-0.42	1.54	1.71
O_b^{H}	0.115	0.130	-0.26	-0.34	1.47	1.59
O_a^{R}	0.105	0.125	-0.19	-0.31	1.15	1.32
O_b^{R}	0.100	0.125	-0.30	-0.43	1.31	1.48
O_a^{B}	0.110	0.125	-0.37	-0.45	1.53	1.66
O_d^{B}	0.105	0.125	-0.47	-0.57	1.49	1.64

Table B.6: For the range of α where $|\xi| < 0.05$ eV, the minimum and maximum values for trapping energy E_t and polaron CTL are shown.

Phase	E_g min (eV)	E_g max (eV)
Rutile	2.68	3.04
Anatase	2.85	3.03
Brookite	3.02	3.28
TiO_2(H)	3.58	3.76
TiO_2(R)	3.43	3.61
TiO_2(B)	3.71	3.90

Table B.7: Shown is the minimum and maximum band gap E_g values within the determined α range for each phase.

Variation in Coulomb-truncation radius

R_c	α	a_x	c_z	E_g	CTL	ξ	E_t	q
2.0	0.20	4.583	2.950	3.084	-0.924	0.013	0.774	-0.136
2.0	0.25	4.570	2.947	3.454	-1.245	0.225	0.825	-0.362

All energies and lengths are shown in eV and Å

Table B.8: Optimisation of α to the GKT in rutile with all the same parameters as the main text, however with R_c reduced to 2.0 Å. Using this approach yields non-transferable values of α between electrons and holes in rutile and anatase. The stability of polarons is also drastically changed compared to the main text. The reader should bear in mind that a titanium-oxygen bond is typically around 2.1 Å, which means using such low values of R_c could restrict the effects of exchange locally on each atom. In essence, this would be similar to a Lany-Zunger self-interaction correction. This illustrates that functionals with such low R_c predict somewhat unphysical properties.

Appendix C

Finite-size correction for eigenvalues

The form of the eigenvalue correction considered in Chapter 6 is:

$$\varepsilon_{\text{ic}} = \varepsilon - \frac{2}{q}\Delta E_{\text{ic}}. \tag{C.1}$$

Most DFT codes, including `CP2K`, set the average electrostatic potential to 0 in a calculation. For charged systems, this means both the total energy and eigenvalues may be offset by an arbitrary amount, as discussed in Chapter 6. Furthermore, electrostatic interactions between neighbouring cells can artificially change energies. While total energy corrections are fairly common in literature, corrections to eigenvalues are rarely used. Table C.1 shows application of the above equation to an oxygen molecule in periodic and aperiodic systems, where it can be seen that the aperiodic result is quickly recovered with the eigenvalue correction. In order to further verify the approach also performs well on extended systems, a q = +1 charge state of an unrelaxed F-centre defect in MgO was considered, shown in Fig. C.1. In order to verify the $2/q$ form of the correction is indeed correct, two holes were added to the oxygen molecule. The behaviour of the corrected values was similar to the q = +1 case, yielding similar energies to the aperiodic case.

Finite system testing

Length L	$E(N+1)$	$E(N)$	HOMO	ξ	$E(N+1)^\dagger$	HOMO†	$\xi^\dagger$
10*	-868.453	-867.667	2.258	-3.044	-	-	-
20*	-868.453	-867.667	2.258	-3.044	-	-	-
30*	-868.453	-867.667	2.258	-3.044	-	-	-
10	-870.251	-867.669	-1.500	-1.083	-868.209	2.586	-3.125
15	-869.742	-867.667	-0.369	-1.706	-868.380	2.355	-3.068
20	-869.444	-867.667	0.256	-2.033	-868.422	2.299	-3.054
25	-869.254	-867.667	0.645	-2.232	-868.437	2.279	-3.049
30	-869.125	-867.667	0.908	-2.366	-868.444	2.270	-3.047
35	-869.031	-867.667	1.098	-2.462	-868.447	2.266	-3.045
40	-868.960	-867.667	1.242	-2.534	-868.449	2.263	-3.045

† With finite size correction * Aperiodic reference system

Table C.1: Application of the eigenvalue correction proposed in Chapter 6 to an oxygen molecule inside periodic and aperiodic simulation cells. For the aperiodic case, no correction is needed as values are independent of system size, as shown above. Aperiodic cells also provide a reference value for the correction. It can be seen that the appropriate value of nonlinearity ξ is recovered for reasonably sized cells. This represents an extreme case of the correction as the dielectric constant of vacuum is unity.

Length L	$E(N-1)$	$E(N)$	LUMO	ξ	$E(N-1)^\dagger$	LUMO†	$\xi^\dagger$
20*	-855.678	-867.762	-16.161	-4.077	-	-	-
10	-857.843	-867.763	-11.999	-2.079	-855.800	-16.085	-4.122
15	-857.076	-867.762	-13.415	-2.729	-855.714	-16.139	-4.090
20	-856.715	-867.762	-14.109	-3.061	-855.693	-16.152	-4.083
25	-856.503	-867.762	-14.522	-3.263	-855.685	-16.157	-4.080
30	-856.363	-867.762	-14.797	-3.398	-855.682	-16.159	-4.079
35	-856.264	-867.762	-14.992	-3.495	-855.681	-16.160	-4.078
40	-856.191	-867.762	-15.139	-3.567	-855.680	-16.160	-4.078

† With finite size correction * Aperiodic reference system

Table C.2: As Table C.1, however for hole states on a charge deficient oxygen molecule ($q = +1$). It can be seen here that the correction works well for holes as well as electrons.

Extended system testing

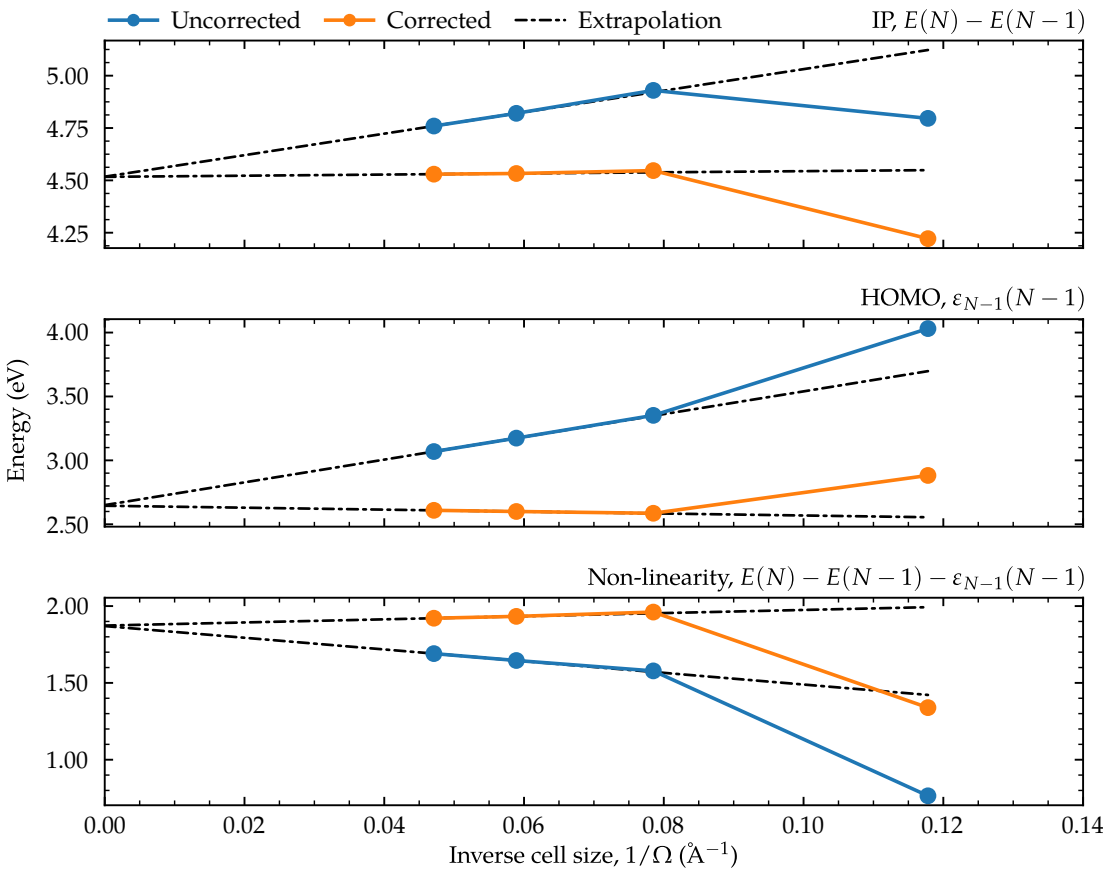

Figure C.1: Scaling behaviour of eigenvalues for an MgO F-centre defect. For each system size, one oxygen atom and one electron were removed and the electronic energy was converged. Only the electronic component of the dielectric constant was used for the image-charge correction as ions were not relaxed.

TiO_2 scaling tests

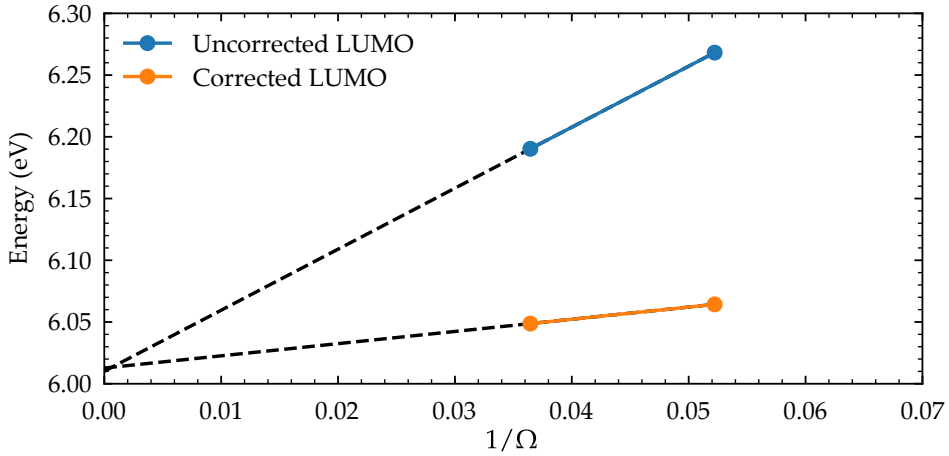

Figure C.2: Scaling behaviour of eigenvalues for a self-trapped hole polaron configuration in anatase. The polaron was created by radially (outward) shifting atoms surrounding one oxygen atom by 0.3 Å. These calculations were performed using DFT to allow larger system sizes to be investigated. Geometric relaxation would cause the hole to delocalise, therefore geometries were not relaxed.

ABBREVIATIONS

tr-PBEα Coulomb-truncated PBEα.

ADMM auxiliary density matrix method.

B3LYP Becke, 3-parameter, Lee-Yang-Parr.

BFGS Broyden-Fletcher-Goldfarb-Shanno.

CBM conduction band maximum.

CTL charge transition level.

DFT density functional theory.

DFT+U DFT with a Hubbard-U correction.

DLTS deep-level transient spectroscopy.

EELS electron energy loss spectroscopy.

EPR electron paramagnetic resonance.

GGA generalised gradient approximation.

GKS generalised Kohn-Sham.

GKT generalised Koopmans' theorem.

GPW Gaussian and plane waves approach.

GTH Goedecker-Teter-Hutter.

GTO Gaussian-type orbital.

GW An approximation to the self-energy in Hedin's equations.

HEG homogeneous electron gas.

HF Hartree-Fock.

HOMO highest occupied molecular orbital.

HSE Heyd-Scuseria-Ernzerhof.

HSE06 HSE functional with a particular parametrisation.

IP ionisation potential.

IR infrared spectroscopy.

KS Kohn-Sham.

LAPACK linear algebra package.

LDA local density approximation.

LUMO lowest unoccupied molecular orbital.

PBE Perdew-Burke-Ernzerhof.

PBE0 Perdew-Burke-Ernzerhof hybrid functional with α=0.25.

PBEα Like PBE0 hybrid functional but with other values of α.

PES photoelectron spectroscopy.

PL photoluminescence.

PW91 Perdew-Wang 1991.

QMC quantum Monte-Carlo.

RPA random phase approximation.

SCF self-consistent field.

SIE self-interaction error.

STM scanning tunnelling microscopy.

STO Slater-type orbital.

STS scanning tunnelling spectroscopy.

UV/Vis ultraviolet/visible light spectroscopy.

VBM valence band minimum.

xc exchange-correlation.

XPS X-ray photoelectron spectroscopy.

XRD X-ray diffraction.

BIBLIOGRAPHY

[1] M. A. Fox and M. T. Dulay, 'Heterogeneous photocatalysis,' *Chem. Rev.* **93**, 341 (1993)

[2] R. Asahi, T. Morikawa, T. Ohwaki, K. Aoki and Y. Taga, 'Visible-light photocatalysis in nitrogen-doped titanium oxides,' *Science* **293**, 269 (2001)

[3] A. Zecchina, D. Scarano, S. Bordiga, G. Spoto and C. Lamberti, 'Surface structures of oxides and halides and their relationships to catalytic properties,' *Adv. Catal.* **46**, 265 (2001)

[4] P. Kuiper, G. Kruizinga, J. Ghijsen, G. A. Sawatzky and H. Verweij, 'Character of holes in $Li_xNi_{1-x}O$ and their magnetic behavior,' *Phys. Rev. Lett.* **62**, 221 (1989)

[5] J. Osorio-Guillén, S. Lany, S. V. Barabash and A. Zunger, 'Nonstoichiometry as a source of magnetism in otherwise nonmagnetic oxides: magnetically interacting cation vacancies and their percolation,' *Phys. Rev. B* **75**, 184421 (2007)

[6] N. Mannella, W. L. Yang, K. Tanaka, X. J. Zhou, H. Zheng, J. F. Mitchell, J. Zaanen, T. P. Devereaux, N. Nagaosa, Z. Hussain and Z. X. Shen, 'Polaron coherence condensation as the mechanism for colossal magnetoresistance in layered manganites,' *Phys. Rev. B* **76**, 233102 (2007)

[7] T. Miyasaka and T. N. Murakami, 'The photocapacitor: an efficient self-charging capacitor for direct storage of solar energy,' *Appl. Phys. Lett.* **85**, 3932 (2004)

[8] A. M. Stoneham, J. Gavartin, A. L. Shluger, A. V. Kimmel, D. M. Ramo, H. M. Rønnow, G. Aeppli and C. Renner, 'Trapping, self-trapping and the polaron family,' *J. Phys. Condens. Matter* **19**, 255208 (2007)

[9] C. Sotelo-Vazquez, N. Noor, A. Kafizas, R. Quesada-Cabrera, D. O. Scanlon, A. Taylor, J. R. Durrant and I. P. Parkin, 'Multifunctional P-doped TiO_2 films:

a new approach to self-cleaning, transparent conducting oxide materials,' *Chem. Mater.* **27**, 3234 (2015)

[10] T. Kirchartz, L. Krckemeier and E. L. Unger, 'Research update: Recombination and open-circuit voltage in lead-halide perovskites,' *APL Mater.* **6**, 100702 (2018)

[11] N. Ashcroft and N. Mermin, *Solid State Physics*, (Saunders College 1976)

[12] Y. Rong, Z. Ku, A. Mei, T. Liu, M. Xu, S. Ko, X. Li and H. Han, 'Hole-conductor-free mesoscopic $TiO_2/CH_3NH_3PbI_3$ heterojunction solar cells based on anatase nanosheets and carbon counter electrodes,' *J. Phys. Chem. Lett.* **5**, 2160 (2014)

[13] *National Renewable Energy Laboratory (NREL) website, www.nrel.gov*

[14] R. Wang, M. Mujahid, Y. Duan, Z.-K. Wang, J. Xue and Y. Yang, 'A review of perovskites solar cell stability,' *Adv. Funct. Mater.* 1808843 (2019)

[15] S. K. Wallace and K. P. Mckenna, 'Facet-dependent electron trapping in TiO_2 nanocrystals,' *J. Phys. Chem. C* **119**, 1913 (2015)

[16] P. Rinke, A. Schleife, E. Kioupakis, A. Janotti, C. Rödl, F. Bechstedt, M. Scheffler and C. G. Van de Walle, 'First-principles optical spectra for f centers in MgO,' *Phys. Rev. Lett.* **108**, 126404 (2012)

[17] J. T. Devreese, *Polarons*, (American Cancer Society 2003)

[18] W. H. Sio, C. Verdi, S. Poncé and F. Giustino, 'Polarons from first principles, without supercells,' *Phys. Rev. Lett.* **122**, 246403 (2019)

[19] G. Girishkumar, B. McCloskey, A. C. Luntz, S. Swanson and W. Wilcke, 'Lithium-air battery: promise and challenges,' *J. Phys. Chem. Lett.* **1**, 2193 (2010)

[20] S. Yang, *Computational Prediction of Transport Properties in Battery Materials*, (University of Michigan 2018)

[21] O. Gerbig, R. Merkle and J. Maier, 'Electrical transport and oxygen exchange in the superoxides of potassium, rubidium, and cesium,' *Adv. Funct. Mater.* **25**, 2552 (2015)

[22] R. O. Jones, 'Density functional theory: its origins, rise to prominence, and future,' *Rev. Mod. Phys.* **87**, 897 (2015)

[23] C. Lee, W. Yang and R. G. Parr, 'Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density,' *Phys. Rev. B* **37**, 785 (1988)

[24] A. D. Becke, 'Density-functional thermochemistry. III. The role of exact exchange,' *J. Chem. Phys.* **98**, 5648 (1993)

[25] J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz, 'Density-functional theory for fractional particle number: derivative discontinuities of the energy,' *Phys. Rev. Lett.* **49**, 1691 (1982)

[26] J. F. Janak, 'Proof that $\frac{\partial e}{\partial n_i} = \epsilon$ in density-functional theory,' *Phys. Rev. B* **18**, 7165 (1978)

[27] M. M. Valiev and G. W. Fernando, 'Occupation numbers in density-functional calculations,' *Phys. Rev. B* **52**, 10697 (1995)

[28] P. Mori-Sánchez, A. J. Cohen and W. Yang, 'Localization and delocalization errors in density functional theory and implications for band-gap prediction,' *Phys. Rev. Lett.* **100**, 146401 (2008)

[29] A. J. Cohen, P. Mori-Sánchez and W. Yang, 'Insights into current limitations of density functional theory,' *Science* **321**, 792 (2008)

[30] A. J. Cohen, P. Mori-Snchez and W. Yang, 'Challenges for density functional theory,' *Chem. Rev.* **112**, 289 (2012)

[31] S. Lany and A. Zunger, 'Polaronic hole localization and multiple hole binding of acceptors in oxide wide-gap semiconductors,' *Phys. Rev. B* **80**, 085202 (2009)

[32] M. Cococcioni and S. de Gironcoli, 'Linear response approach to the calculation of the effective interaction parameters in the LDA + U method,' *Phys. Rev. B* **71**, 035105 (2005)

[33] J. Wiktor, F. Ambrosio and A. Pasquarello, 'Role of polarons in water splitting: the case of $BiVO_4$,' *ACS Energy Lett.* **3**, 1693 (2018)

[34] K. E. Kweon, G. S. Hwang, J. Kim, S. Kim and S. Kim, 'Electron small polarons and their transport in bismuth vanadate: a first principles study,' *Phys. Chem. Chem. Phys.* **17**, 256 (2015)

[35] A. Chaudhuri, L. Mandal, X. Chi, M. Yang, M. C. Scott, M. Motapothula, X. J. Yu, P. Yang, Y. Shao-Horn, T. Venkatesan, A. T. S. Wee and A. Rusydi, 'Direct observation of anisotropic small-hole polarons in an orthorhombic structure of $BiVO_4$ films,' *Phys. Rev. B* **97**, 195150 (2018)

[36] J. K. Cooper, S. B. Scott, Y. Ling, J. Yang, S. Hao, Y. Li, F. M. Toma, M. Stutzmann, K. V. Lakshmi and I. D. Sharp, 'Role of hydrogen in defining the n-type character of $BiVO_4$ photoanodes,' *Chem. Mater.* **28**, 5761 (2016)

[37] A. Kudo, K. Omori and H. Kato, 'A novel aqueous process for preparation of crystal form-controlled and highly crystalline $BiVO_4$ powder from layered vanadates at room temperature and its photocatalytic and photophysical properties,' *J. Am. Chem. Soc.* **121**, 11459 (1999)

[38] H. He, D. A. Andersson, D. D. Allred and K. D. Rector, 'Determination of the insulation gap of uranium oxides by spectroscopic ellipsometry and density functional theory,' *J. Phys. Chem. C* **117**, 16540 (2013)

[39] J. Devreese, R. De Coninck and H. Pollak, 'On the conduction mechanism in uranium dioxide,' *Phys. Status Solidi B* **17**, 825 (1966)

[40] H. Idriss, 'Surface reactions of uranium oxide powder, thin films and single crystals,' *Surf. Sci. Rep.* **65**, 67 (2010)

[41] P. Deák, Q. Duy Ho, F. Seemann, B. Aradi, M. Lorke and T. Frauenheim, 'Choosing the correct hybrid for defect calculations: a case study on intrinsic carrier trapping in β-Ga_2O_3,' *Phys. Rev. B* **95**, 075208 (2017)

[42] J. L. Lyons, 'A survey of acceptor dopants for β-Ga_2O_3,' *Semicond. Sci. Technol.* **33**, 05LT02 (2018)

[43] Z. Zhang, E. Farzana, A. R. Arehart and S. A. Ringel, 'Deep level defects throughout the bandgap of (010) β-Ga_2O_3 detected by optically and thermally stimulated defect spectroscopy,' *Appl. Phys. Lett.* **108**, 052105 (2016)

[44] L. Sun, X. Huang, L. Wang and A. Janotti, 'Disentangling the role of small polarons and oxygen vacancies in CeO_2,' *Phys. Rev. B* **95**, 245101 (2017)

[45] I. Naik and T. Tien, 'Small-polaron mobility in nonstoichiometric cerium dioxide,' *J. Phys. Chem. Solids* **39**, 311 (1978)

[46] P. Deák, B. Aradi and T. Frauenheim, 'Quantitative theory of the oxygen vacancy and carrier self-trapping in bulk TiO_2,' *Phys. Rev. B* **86**, 195206 (2012)

[47] M. Setvin, C. Franchini, X. Hao, M. Schmid, A. Janotti, M. Kaltak, C. G. Van de Walle, G. Kresse and U. Diebold, 'Direct view at excess electrons in TiO_2 rutile and anatase,' *Phys. Rev. Lett.* **113**, 086402 (2014)

[48] C. Spreafico and J. VandeVondele, 'The nature of excess electrons in anatase and rutile from hybrid DFT and RPA,' *Phys. Chem. Chem. Phys.* **16**, 26144 (2014)

[49] S. Yang, A. T. Brant, N. C. Giles and L. E. Halliburton, 'Intrinsic small polarons in rutile TiO_2,' *Phys. Rev. B* **87**, 125201 (2013)

[50] A. Amtout and R. Leonelli, 'Optical properties of rutile near its fundamental band gap,' *Phys. Rev. B* **51**, 6842 (1995)

[51] B. Kraeutler and A. J. Bard, 'Heterogeneous photocatalytic decomposition of saturated carboxylic acids on titanium dioxide powder. decarboxylative route to alkanes,' *J. Am. Chem. Soc.* **100**, 5985 (1978)

[52] K. P. McKenna, M. J. Wolf, A. L. Shluger, S. Lany and A. Zunger, 'Two-dimensional polaronic behavior in the binary oxides *m*-HfO_2 and *m*-ZrO_2,' *Phys. Rev. Lett.* **108**, 116403 (2012)

[53] H. Jiang, R. I. Gomez-Abal, P. Rinke and M. Scheffler, 'Electronic band structure of zirconia and hafnia polymorphs from the *GW* perspective,' *Phys. Rev. B* **81**, 085119 (2010)

[54] S. Wright and R. C. Barklie, 'Electron paramagnetic resonance characterization of defects in monoclinic HfO_2 and ZrO_2 powders,' *J. Appl. Phys* **106**, 103917 (2009)

[55] J. Aarik, H. Mndar, M. Kirm and L. Pung, 'Optical characterization of HfO_2 thin films grown by atomic layer deposition,' *Thin Solid Films* **466**, 41 (2004)

[56] M. Balog, M. Schieber, M. Michman and S. Patai, 'Chemical vapor deposition and characterization of HfO_2 films from organo-hafnium compounds,' *Thin Solid Films* **41**, 247 (1977)

[57] H. Nohira, W. Tsai, W. Besling, E. Young, J. Petry, T. Conard, W. Vandervorst, S. D. Gendt, M. Heyns, J. Maes and M. Tuominen, 'Characterization of ALCVD-Al_2O_3 and ZrO_2 layer using X-ray photoelectron spectroscopy,' *J. Non-Cryst. Solids* **303**, 83 (2002)

[58] R. Puthenkovilakam and J. P. Chang, 'Valence band structure and band alignment at the ZrO_2/Si interface,' *Appl. Phys. Lett.* **84**, 1353 (2004)

[59] Z. Wang, C. Brock, A. Matt and K. H. Bevan, 'Implications of the DFT+U method on polaron properties in energy materials,' *Phys. Rev. B* **96**, 125150 (2017)

[60] N. Adelstein, J. B. Neaton, M. Asta and L. C. De Jonghe, 'Density functional theory based calculation of small-polaron mobility in hematite,' *Phys. Rev. B* **89**, 245115 (2014)

[61] T. J. Smart and Y. Ping, 'Effect of defects on the small polaron formation and transport properties of hematite from first-principles calculations,' *J. Phys. Condens. Matter* **29**, 394006 (2017)

[62] A. Jahagirdar, N. Dhananjaya, D. Monika, C. Kesavulu, H. Nagabhushana, S. Sharma, B. Nagabhushana, C. Shivakumara, J. Rao and R. Chakradhar, 'Structural, EPR, optical and magnetic properties of α-Fe_2O_3 nanoparticles,' *Spectrochim. Acta A* **104**, 512 (2013)

[63] A. J. E. Rettie, W. D. Chemelewski, B. R. Wygant, J. Lindemuth, J.-F. Lin, D. Eisenberg, C. S. Brauer, T. J. Johnson, T. N. Beiswenger, R. D. Ash, X. Li, J. Zhou and C. B. Mullins, 'Synthesis, electronic transport and optical properties of Si:α-Fe_2O_3 single crystals,' *J. Mater. Chem. C* **4**, 559 (2016)

[64] L. M. Carneiro, S. K. Cushing, C. Liu, Y. Su, P. Yang, A. P. Alivisatos and S. R. Leone, 'Excitation-wavelength-dependent small polaron trapping of photoexcited carriers in α-Fe_2O_3,' *Nat. Mater.* **16**, 819 (2017)

[65] M. Al-Kuhaili, M. Saleem and S. Durrani, 'Optical properties of iron oxide ($\alpha - Fe_2O_3$) thin films deposited by the reactive evaporation of iron,' *J. Alloys Compd.* **521**, 178 (2012)

[66] J. Kang, Y. S. Jung, S.-H. Wei and A. C. Dillon, 'Implications of the formation of small polarons in Li_2O_2 for Li-air batteries,' *Phys. Rev. B* **85**, 035210 (2012)

[67] S. P. Ong, Y. Mo and G. Ceder, 'Low hole polaron migration barrier in lithium peroxide,' *Phys. Rev. B* **85**, 081105 (2012)

[68] J. M. Garcia-Lastra, J. S. G. Myrdal, R. Christensen, K. S. Thygesen and T. Vegge, 'DFT+U study of polaronic conduction in Li_2O_2 and Li_2CO_3: implications for Li-air batteries,' *J. Phys. Chem. C* **117**, 5568 (2013)

[69] M. D. Radin and D. J. Siegel, 'Charge transport in lithium peroxide: relevance for rechargeable metal-air batteries,' *Energy Environ. Sci.* **6**, 2370 (2013)

[70] J. M. Garcia-Lastra, J. D. Bass and K. S. Thygesen, 'Communication: Strong excitonic and vibronic effects determine the optical properties of Li_2O_2,' *J. Chem. Phys.* **135**, 121101 (2011)

[71] O. Gerbig, R. Merkle and J. Maier, 'Electron and ion transport in Li_2O_2,' *Adv. Mater.* **25**, 3129 (2013)

[72] J. S. Hummelshøj, J. Blomqvist, S. Datta, T. Vegge, J. Rossmeisl, K. S. Thygesen, A. C. Luntz, K. W. Jacobsen and J. K. Nrskov, 'Communications: elementary oxygen electrode reactions in the aprotic Li-air battery,' *J. Chem. Phys.* **132**, 071101 (2010)

[73] S. Li, J. Liu and B. Liu, 'First-principles study of the charge transport mechanisms in lithium superoxide,' *Chem. Mater.* **29**, 2202 (2017)

[74] R. B. Araujo, S. Chakraborty and R. Ahuja, 'Unveiling the charge migration mechanism in Na_2O_2: implications for sodium-air batteries,' *Phys. Chem. Chem. Phys.* **17**, 8203 (2015)

[75] S. Yang and D. J. Siegel, 'Intrinsic conductivity in sodium-air battery discharge phases: Sodium superoxide vs sodium peroxide,' *Chem. Mater.* **27**, 3852 (2015)

[76] S. Loftager, P. Garca-Fernndez, J. A. Aramburu, M. Moreno and J. M. Garcia-Lastra, 'Stability and polaronic motion of self-trapped holes in silver halides: Insight through DFT+U calculations,' *J. Phys. Chem. C* **120**, 8509 (2016)

[77] P. G. Baranov, N. G. Romanov, O. G. Poluektov and J. Schmidt, 'Self-trapped excitons in ionic-covalent silver halide crystals and nanostructures: High-frequency EPR, ESE, ENDOR and ODMR studies,' *Appl. Magn. Reson.* **39**, 453 (2010)

[78] H. Kanzaki and S. Sakuragi, 'Optical absorption and luminescence of excitons in silver halides containing isoelectronic impurities. Part II. AgBr:Cl^- and AgBr,' *J. Phys. Soc. Jpn.* **29**, 924 (1970)

[79] T. G. Castner and W. Känzig, 'The electronic structure of V-centers,' *J. Phys. Chem. Solids* **3**, 178 (1957)

[80] L. J. Page and E. H. Hygh, 'Calculation of energy bands in alkali halides,' *Phys. Rev. B* **1**, 3472 (1970)

[81] D. M. Roessler and W. C. Walker, 'Electronic spectra of crystalline NaCl and KCl,' *Phys. Rev.* **166**, 599 (1968)

[82] G. Geneste, B. Amadon, M. Torrent and G. Dezanneau, 'DFT+U study of self-trapping, trapping, and mobility of oxygen-type hole polarons in barium stannate,' *Phys. Rev. B* **96**, 134123 (2017)

[83] H. Mizoguchi, P. Chen, P. Boolchand, V. Ksenofontov, C. Felser, P. W. Barnes and P. M. Woodward, 'Electrical and optical properties of Sb-doped $BaSnO_3$,' *Chem. Mater.* **25**, 3858 (2013)

[84] H. Mizoguchi, H. W. Eng and P. M. Woodward, 'Probing the electronic structures of ternary perovskite and pyrochlore oxides containing Sn^{4+} or Sb^{5+},' *Inorg. Chem.* **43**, 1667 (2004)

[85] M. Swift, A. Janotti and C. G. Van de Walle, 'Small polarons and point defects in barium cerate,' *Phys. Rev. B* **92**, 214114 (2015)

[86] M. E. Kompan, Y. M. Baikov, B. A. T. Melekh and A. V. Yakubovich, 'Luminescence spectra of nominally pure $BaCeO_3$ perovskite crystals,' *Phys. Solid State* **44**, 1263 (2002)

[87] T. He, P. Ehrhart and P. Meuffels, 'Optical band gap and urbach tail in Y-doped $BaCeO_3$,' *J. Appl. Phys.* **79**, 3219 (1996)

[88] E. J. Granhed, A. Lindman, C. Eklf-sterberg, M. Karlsson, S. F. Parker and G. Wahnstrm, 'Band vs. polaron: vibrational motion and chemical expansion of hydride ions as signatures for the electronic character in oxyhydride barium titanate,' *J. Mater. Chem. A* **7**, 16211 (2019)

[89] P. Erhart, A. Klein, D. Åberg and B. Sadigh, 'Efficacy of the DFT+U formalism for modeling hole polarons in perovskite oxides,' *Phys. Rev. B* **90**, 035204 (2014)

[90] H. Ihrig, 'On the polaron nature of the charge transport in $BaTiO_3$,' *J. Phys. C* **9**, 3469 (1976)

[91] J. P. Boyeaux and F. M. Michel-Calendini, 'Small polaron interpretation of $BaTiO_3$: transport properties from drift mobility measurements,' *J. Phys. C* **12**, 545 (1979)

[92] S. H. Wemple, 'Polarization fluctuations and the optical-absorption edge in $BaTiO_3$,' *Phys. Rev. B* **2**, 2679 (1970)

[93] F. M. Michel-Calendini and G. Mesnard, 'Band structure and optical properties of tetragonal BaTiO3,' *J. Phys. C* **6**, 1709 (1973)

[94] A. Janotti, J. B. Varley, M. Choi and C. G. Van de Walle, 'Vacancies and small polarons in $SrTiO_3$,' *Phys. Rev. B* **90**, 085202 (2014)

[95] C. Bhandari, M. van Schilgaarde, T. Kotani and W. R. L. Lambrecht, 'All-electron quasiparticle self-consistent *GW* band structures for $srtio_3$ including lattice polarization corrections in different phases,' *Phys. Rev. Materials* **2**, 013807 (2018)

[96] H. Tan, Z. Zhao, W.-b. Zhu, E. N. Coker, B. Li, M. Zheng, W. Yu, H. Fan and Z. Sun, 'Oxygen vacancy enhanced photocatalytic activity of pervoskite $SrTiO_3$,' *ACS Appl. Mater. Interfaces* **6**, 19184 (2014)

[97] R. Leonelli and J. L. Brebner, 'Time-resolved spectroscopy of the visible emission band in strontium titanate,' *Phys. Rev. B* **33**, 8649 (1986)

[98] K. van Benthem, C. Elssser and R. H. French, 'Bulk electronic structure of $SrTiO_3$: Experiment and theory,' *J. Appl. Phys.* **90**, 6156 (2001)

[99] B. Kang and K. Biswas, 'Carrier self-trapping and luminescence in intrinsically activated scintillator: Cesium hafnium chloride (Cs_2HfCl_6),' *J. Phys. Chem. C* **120**, 12187 (2016)

[100] M. Buryi, R. Krl, V. Babin, J. Pterek, V. Vanek, P. Veverka, M. Kohoutkov, V. Laguta, M. Fasoli, I. Villa, F. Cova, A. Vedda and M. Nikl, 'Trapping and recombination centers in cesium hafnium chloride single crystals: EPR and TSL study,' *J. Phys. Chem. C* **123**, 19402 (2019)

[101] R. Krl, V. Babin, E. Mihkov, M. Buryi, V. V. Laguta, K. Nitsch and M. Nikl, 'Luminescence and charge trapping in Cs_2HfCl_6 single crystals: Optical and magnetic resonance spectroscopy study,' *J. Phys. Chem. C* **121**, 12375 (2017)

[102] B. Kang and K. Biswas, 'Exploring polaronic, excitonic structures and luminescence in $Cs_4PbBr_6/CsPbBr_3$,' *J. Phys. Chem. Lett.* **9**, 830 (2018)

[103] H. Zhang, Q. Liao, Y. Wu, J. Chen, Q. Gao and H. Fu, 'Pure zero-dimensional Cs_4PbBr_6 single crystal rhombohedral microdisks with high luminescence and stability,' *Phys. Chem. Chem. Phys.* **19**, 29092 (2017)

[104] J.-H. Cha, J. H. Han, W. Yin, C. Park, Y. Park, T. K. Ahn, J. H. Cho and D.-Y. Jung, 'Photoresponse of $CsPbBr_3$ and Cs_4PbBr_6 perovskite single crystals,' *J. Phys. Chem. Lett.* **8**, 565 (2017)

[105] Y. Zhang, M. I. Saidaminov, I. Dursun, H. Yang, B. Murali, E. Alarousu, E. Yengel, B. A. Alshankiti, O. M. Bakr and O. F. Mohammed, 'Zero-dimensional Cs_4PbBr_6 perovskite nanocrystals,' *J. Phys. Chem. Lett.* **8**, 961 (2017)

[106] K. Saeki, Y. Fujimoto, M. Koshimizu, D. Nakauchi, H. Tanaka, T. Yanagida and K. Asai, 'Luminescence and scintillation properties of Cs_2HfBr_6 and Cs_2ZrBr_6 crystals,' *Jpn. J. Appl. Phys.* **57**, 030310 (2018)

[107] K. Biswas and M.-H. Du, 'Energy transport and scintillation of cerium-doped elpasolite Cs_2LiYCl_6: hybrid density functional calculations,' *Phys. Rev. B* **86**, 014102 (2012)

[108] E. V. D. van Loef, P. Dorenbos, C. W. E. van Eijk, K. W. K. mer and H. U. G. del, 'Scintillation and spectroscopy of the pure and Ce^{3+}-doped elpasolites: Cs_2LiYX_6(X $\sim$ Cl, Br),' *J. Phys. Condens. Matter* **14**, 8481 (2002)

[109] A. Bessière, P. Dorenbos, C. W. E. van Eijk, L. Pidol, K. W. Krmer and H. U. Gdel, 'Spectroscopy and anomalous emission of Ce doped elpasolite Cs_2LiYCl_6,' *J. Phys. Condens. Matter* **16**, 1887 (2004)

[110] T. Pawlik and J.-M. Spaeth, 'Electron and hole centres in the x-irradiated elpasolite crystal studied by means of electron paramagnetic resonance and electron nuclear double resonance,' *J. Phys. Condens. Matter* **9**, 8737 (1997)

[111] K. Hoang, 'Polaron formation, native defects, and electronic conduction in metal tungstates,' *Phys. Rev. Materials* **1**, 024603 (2017)

[112] M. Dissanayake, O. Ileperuma and P. Dharmasena, 'A.C. conductivity of $MnWO_4$,' *J. Phys. Chem. Solids* **50**, 359 (1989)

[113] J. Ruiz-Fuertes, S. López-Moreno, J. López-Solano, D. Errandonea, A. Segura, R. Lacomba-Perales, A. Muñoz, S. Radescu, P. Rodríguez-Hernández, M. Gospodinov, L. L. Nagornaya and C. Y. Tu, 'Pressure effects on the electronic and optical properties of AWO_4 wolframites ($A =$ Cd, Mg, Mn, and Zn): the distinctive behavior of multiferroic $MnWO_4$,' *Phys. Rev. B* **86**, 125202 (2012)

[114] T. Ejima, T. Banse, H. Takatsuka, Y. Kondo, M. Ishino, N. Kimura, M. Watanabe and I. Matsubara, 'Microscopic optical and photoelectron measurements of MWO_4 (M=Mn, Fe, and Ni),' *J. Lumin.* **119-120**, 59 (2006)

[115] E. Schmidbauer, U. Schanz and F. J. Yu, 'Electrical transport properties of mono- and polycrystalline FeWO4,' *J. Phys. Condens. Matter* **3**, 5341 (1991)

[116] M. Levy, J. P. Perdew and V. Sahni, 'Exact differential equation for the density and ionization energy of a many-particle system,' *Phys. Rev. A* **30**, 2745 (1984)

[117] C. O. Almbladh and U. von Barth, 'Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues,' *Phys. Rev. B* **31**, 3231 (1985)

[118] J. P. Perdew and M. Levy, 'Comment on "significance of the highest occupied Kohn-Sham eigenvalue",' *Phys. Rev. B* **56**, 16021 (1997)

[119] L. Kronik, T. Stein, S. Refaely-Abramson and R. Baer, 'Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals,' *J. Chem. Theory Comput.* **8**, 1515 (2012)

[120] M. Srebro and J. Autschbach, 'Tuned range-separated time-dependent density functional theory applied to optical rotation,' *J. Chem. Theory Comput.* **8**, 245 (2012)

[121] M. Srebro and J. Autschbach, 'Does a molecule-specific density functional give an accurate electron density? The challenging case of the CuCl electric field gradient,' *J. Phys. Chem. Lett.* **3**, 576 (2012)

[122] V. Vlek, H. R. Eisenberg, G. Steinle-Neumann, L. Kronik and R. Baer, 'Deviations from piecewise linearity in the solid-state limit with approximate density functionals,' *J. Chem. Phys.* **142**, 034107 (2015)

[123] D. V. Lang, 'Deeplevel transient spectroscopy: a new method to characterize traps in semiconductors,' *J. Appl. Phys.* **45**, 3023 (1974)

[124] S. Hüfner, *Photoelectron Spectroscopy: principles and applications*, (Springer 2003)

[125] F. Hofer, F. P. Schmidt, W. Grogger and G. Kothleitner, 'Fundamentals of electron energy-loss spectroscopy,' *IOP Conference Series: Materials Science and Engineering* **109**, 012007 (2016)

[126] J. A. Weil and J. R. Bolton, *Basic Principles of Paramagnetic Resonance*, (John Wiley & Sons 2006)

[127] M. Reticcioli, U. Diebold, G. Kresse and C. Franchini, *Small Polarons in Transition Metal Oxides*, (Springer International Publishing 2019), 1–39

[128] H. Tuller and A. Nowick, 'Small polaron electron transport in reduced CeO_2 single crystals,' *Journal of Physics and Chemistry of Solids* **38(8)**, 859 (1977)

[129] H. Sezen, M. Buchholz, A. Nefedov, C. Natzeck, S. Heissler, C. Di Valentin and C. Wöll, 'Probing electrons in TiO_2 polaronic trap states by ir-absorption: evidence for the existence of hydrogenic states,' *Sci. Rep.* **4**, 3808 EP (2014)

[130] C. M. Yim, M. B. Watkins, M. J. Wolf, C. L. Pang, K. Hermansson and G. Thornton, 'Engineering polarons at a metal oxide surface,' *Phys. Rev. Lett.* **117**, 116402 (2016)

[131] P. Hohenberg and W. Kohn, 'Inhomogeneous electron gas,' *Phys. Rev.* **136**, B864 (1964)

[132] R. M. Martin, *Electronic Structure: basic Theory and Practical Methods*, (Cambridge University Press 2004)

[133] R. G. Parr, *Horizons of Quantum Chemistry: density Functional Theory of Atoms and Molecules*, (Springer Netherlands 1980)

[134] T. Chachiyo, 'Communication: simple and accurate uniform electron gas correlation energy for the full range of densities,' *J. Chem. Phys.* **145**, 021101 (2016)

[135] D. M. Ceperley and B. J. Alder, 'Ground state of the electron gas by a stochastic method,' *Phys. Rev. Lett.* **45**, 566 (1980)

[136] J. P. Perdew and Y. Wang, 'Accurate and simple analytic representation of the electron-gas correlation energy,' *Phys. Rev. B* **45**, 13244 (1992)

[137] S. H. Vosko, L. Wilk and M. Nusair, 'Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis,' *Can. J. Phys.* **58**, 1200 (1980)

[138] J. Gräfenstein, D. Izotov and D. Cremer, 'Avoiding singularity problems associated with meta-GGA (generalized gradient approximation) exchange and correlation functionals containing the kinetic energy density,' *J. Chem. Phys.* **127**, 214103 (2007)

[139] J. P. Perdew, K. Burke and M. Ernzerhof, 'Generalized gradient approximation made simple,' *Phys. Rev. Lett.* **77**, 3865 (1996)

[140] W. Yang, A. J. Cohen and P. Mori-Sánchez, 'Derivative discontinuity, bandgap and lowest unoccupied molecular orbital in density functional theory,' *J. Chem. Phys.* **136**, 204111 (2012)

[141] J. P. Perdew, M. Ernzerhof and K. Burke, 'Rationale for mixing exact exchange with density functional approximations,' *J. Chem. Phys.* **105**, 9982 (1996)

[142] J. Heyd, G. E. Scuseria and M. Ernzerhof, 'Hybrid functionals based on a screened coulomb potential,' *J. Chem. Phys.* **118**, 8207 (2003)

[143] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber and J. G. Angyán, 'Screened hybrid density functionals applied to solids,' *J. Chem. Phys.* **124**, 154709 (2006)

[144] M. Guidon, J. Hutter and J. VandeVondele, 'Robust periodic Hartree-Fock exchange for large-scale simulations using Gaussian basis sets,' *J. Chem. Theory Comput.* **5**, 3010 (2009)

[145] J. Nocedal and S. Wright, *Numerical optimization*, (Springer 2006)

[146] M. T. Entwistle, M. J. P. Hodgson, J. Wetherell, B. Longstaff, J. D. Ramsden and R. W. Godby, 'Local density approximations from finite systems,' *Phys. Rev. B* **94**, 205134 (2016)

[147] S. Adamson, J. Chapman, T. Durrant, A. R. Elmaslmane, M. Entwistle, R. Godby, M. Hodgson, P. Lillystone, A. Long, R. Oliver, J. Ramsden, E. Richardson, M. Smith, L. Talirz and J. Wetherell, *github.com/godby-group/idea-public*

[148] M. Guidon, J. Hutter and J. VandeVondele, 'Auxiliary density matrix methods for Hartree-Fock exchange calculations,' *J. Chem. Theory Comput.* **6**, 2348 (2010)

[149] G. Kresse and J. Furthmller, 'Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set,' *Comput. Mater. Sci.* **6**, 15 (1996)

[150] O. Bengone, M. Alouani, P. Blöchl and J. Hugel, 'Implementation of the projector augmented-wave LDA+U method: application to the electronic structure of NiO,' *Phys. Rev. B* **62**, 16392 (2000)

[151] W. Kohn and L. J. Sham, 'Self-consistent equations including exchange and correlation effects,' *Phys. Rev.* **140**, A1133 (1965)

[152] A. D. Becke, 'Density-functional exchange-energy approximation with correct asymptotic behavior,' *Phys. Rev. A* **38**, 3098 (1988)

[153] J. P. Perdew and A. Zunger, 'Self-interaction correction to density-functional approximations for many-electron systems,' *Phys. Rev. B* **23**, 5048 (1981)

[154] E. J. Baerends, 'Exact exchange-correlation treatment of dissociated H_2 in density functional theory,' *Phys. Rev. Lett.* **87**, 133004 (2001)

[155] A. D. Becke, 'A new mixing of Hartree-Fock and local density-functional theories,' *J. Chem. Phys.* **98**, 1372 (1993)

[156] V. Atalla, I. Y. Zhang, O. T. Hofmann, X. Ren, P. Rinke and M. Scheffler, 'Enforcing the linear behavior of the total energy with hybrid functionals: implications for charge transfer, interaction energies, and the random-phase approximation,' *Phys. Rev. B* **94**, 035140 (2016)

[157] S. Refaely-Abramson, S. Sharifzadeh, M. Jain, R. Baer, J. B. Neaton and L. Kronik, 'Gap renormalization of molecular crystals from density-functional theory,' *Phys. Rev. B* **88**, 081204 (2013)

[158] I. Dabo, A. Ferretti, N. Poilvert, Y. Li, N. Marzari and M. Cococcioni, 'Koopmans' condition for density-functional theory,' *Phys. Rev. B* **82**, 115121 (2010)

[159] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew and K. A. Lyssenko, 'Density functional theory is straying from the path toward the exact functional,' *Science* **355**, 49 (2017)

[160] A. D. Bochevarov and R. A. Friesner, 'The densities produced by the density functional theory: comparison to full configuration interaction,' *J. Chem. Phys.* **128**, 034102 (2008)

[161] Y. He, J. Gräfenstein, E. Kraka and D. Cremer, 'What correlation effects are covered by density functional theory?' *Mol. Phys.* **98**, 1639 (2000)

[162] J. P. Perdew, W. Yang, K. Burke, Z. Yang, E. K. U. Gross, M. Scheffler, G. E. Scuseria, T. M. Henderson, I. Y. Zhang, A. Ruzsinszky, H. Peng, J. Sun, E. Trushin and A. Görling, 'Understanding band gaps of solids in generalized Kohn-Sham theory,' *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2801 (2017)

[163] A. Seidl, A. Görling, P. Vogl, J. A. Majewski and M. Levy, 'Generalized Kohn-Sham schemes and the band-gap problem,' *Phys. Rev. B* **53**, 3764 (1996)

[164] T. Stein, J. Autschbach, N. Govind, L. Kronik and R. Baer, 'Curvature and frontier orbital energies in density functional theory,' *J. Phys. Chem. Lett.* **3**, 3740 (2012)

[165] S. Refaely-Abramson, R. Baer and L. Kronik, 'Fundamental and excitation gaps in molecules of relevance for organic photovoltaics from an optimally tuned range-separated hybrid functional,' *Phys. Rev. B* **84**, 075144 (2011)

[166] T. Stein, L. Kronik and R. Baer, 'Reliable prediction of charge transfer excitations in molecular complexes using time-dependent density functional theory,' *J. Am. Chem. Soc.* **131**, 2818 (2009)

[167] L. Hedin, 'New method for calculating the one-particle Green's function with application to the electron-gas problem,' *Phys. Rev.* **139**, A796 (1965)

[168] F. Aryasetiawan and O. Gunnarsson, 'The *GW* method,' *Rep. Prog. Phys.* **61**, 237 (1998)

[169] T. Stein, H. Eisenberg, L. Kronik and R. Baer, 'Fundamental gaps in finite systems from eigenvalues of a generalized Kohn-Sham method,' *Phys. Rev. Lett.* **105**, 266802 (2010)

[170] V. Atalla, M. Yoon, F. Caruso, P. Rinke and M. Scheffler, 'Hybrid density functional theory meets quasiparticle calculations: a consistent electronic structure approach,' *Phys. Rev. B* **88**, 165122 (2013)

[171] M. A. L. Marques, J. Vidal, M. J. T. Oliveira, L. Reining and S. Botti, 'Density-based mixing parameter for hybrid functionals,' *Phys. Rev. B* **83**, 035119 (2011)

[172] N. P. Brawand, M. Vörös, M. Govoni and G. Galli, 'Generalization of dielectric-dependent hybrid functionals to finite systems,' *Phys. Rev. X* **6**, 041002 (2016)

[173] M. J. P. Hodgson, J. D. Ramsden, J. B. J. Chapman, P. Lillystone and R. W. Godby, 'Exact time-dependent density-functional potentials for strongly correlated tunneling electrons,' *Phys. Rev. B* **88**, 241102 (2013)

[174] B. J. Morgan and G. W. Watson, 'A DFT+U description of oxygen vacancies at the TiO_2 rutile (110) surface,' *Surf. Sci.* **601**, 5034 (2007)

[175] A. L. Shluger, K. P. McKenna, P. V. Sushko, D. M. Ramo and A. V. Kimmel, 'Modelling of electron and hole trapping in oxides,' *Model. Simul. Mater. Sci. Eng.* **17**, 084004 (2009)

[176] P. Mori-Sánchez, A. J. Cohen and W. Yang, 'Many-electron self-interaction error in approximate density functionals,' *J. Chem. Phys.* **125**, 201102 (2006)

[177] R. Haunschild, T. M. Henderson, C. A. Jiménez-Hoyos and G. E. Scuseria, 'Many-electron self-interaction and spin polarization errors in local hybrid density functionals,' *J. Chem. Phys.* **133**, 134116 (2010)

[178] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys and A. P. Sutton, 'Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA+U study,' *Phys. Rev. B* **57**, 1505 (1998)

[179] Z. Hu and H. Metiu, 'Choice of U for DFT+U calculations for titanium oxides,' *J. Phys. Chem. C* **115**, 5841 (2011)

[180] B. J. Morgan and G. W. Watson, 'Polaronic trapping of electrons and holes by native defects in anatase TiO_2,' *Phys. Rev. B* **80**, 233102 (2009)

[181] M. E. A. de Dompablo, A. Morales-Garca and M. Taravillo, 'DFT+U calculations of crystal lattice, electronic structure, and phase stability under pressure of TiO_2 polymorphs,' *J. Chem. Phys.* **135**, 054503 (2011)

[182] B. Lee, C. ki Lee, C. S. Hwang and S. Han, 'Influence of exchange-correlation functionals on dielectric properties of rutile TiO_2,' *Curr. Appl. Phys.* **11**, S293 (2011)

[183] E. Shojaee, M. Abbasnejad, M. Saeedian and M. R. Mohammadizadeh, 'First-principles study of lattice dynamics of TiO_2 in brookite and cotunnite structures,' *Phys. Rev. B* **83**, 174302 (2011)

[184] S. Lany and A. Zunger, 'Accurate prediction of defect properties in density functional supercell calculations,' *Model. Simul. Mater. Sci. Eng.* **17**, 084002 (2009)

[185] C. Di Valentin, G. Pacchioni, A. Selloni, S. Livraghi and E. Giamello, 'Characterization of paramagnetic species in N-doped TiO_2 powders by EPR spectroscopy and DFT calculations,' *J. Phys. Chem. B* **109**, 11414 (2005)

[186] C. Di Valentin, G. Pacchioni and A. Selloni, 'Reduced and n-type doped TiO_2: nature of Ti^{3+} species,' *J. Phys. Chem. C* **113**, 20543 (2009)

[187] B. J. Morgan, D. O. Scanlon and G. W. Watson, 'Small polarons in Nb- and Ta-doped rutile and anatase TiO_2,' *J. Mater. Chem.* **19**, 5175 (2009)

[188] V. N. Bogomolov and D. N. Mirlin, 'Optical absorption by polarons in rutile (TiO_2) single crystals,' *Phys. Status Solidi B* **27**, 443 (1968)

[189] N. A. Deskins, R. Rousseau and M. Dupuis, 'Distribution of Ti^{3+} surface sites in reduced TiO_2,' *J. Phys. Chem. C* **115**, 7562 (2011)

[190] S. Livraghi, M. Rolando, S. Maurelli, M. Chiesa, M. C. Paganini and E. Giamello, 'Nature of reduced states in titanium dioxide as monitored by electron paramagnetic resonance. II: rutile and brookite cases,' *J. Phys. Chem. C* **118**, 22141 (2014)

[191] D. A. Panayotov, S. P. Burrows and J. R. Morris, 'Infrared spectroscopic studies of conduction band and trapped electrons in UV-photoexcited, H-atom n-doped, and thermally reduced TiO_2,' *J. Phys. Chem. C* **116**, 4535 (2012)

[192] C. N. Duckworth, A. W. Brinkman and J. Woods, 'Deep level transient spectroscopy in TiO_2:Nb,' *Phys. Status Solidi A* **75**, K99 (1983)

[193] P. Deák, B. Aradi and T. Frauenheim, 'Polaronic effects in TiO_2 calculated by the HSE06 hybrid functional: dopant passivation by carrier self-trapping,' *Phys. Rev. B* **83**, 155207 (2011)

[194] P. Deák, B. Aradi, T. Frauenheim, E. Janzén and A. Gali, 'Accurate defect levels obtained from the HSE06 range-separated hybrid functional,' *Phys. Rev. B* **81**, 153203 (2010)

[195] A. Janotti, C. Franchini, J. B. Varley, G. Kresse and C. G. Van de Walle, 'Dual behavior of excess electrons in rutile TiO_2,' *Phys. Status Solidi RRL* **7**, 199 (2013)

[196] F. Labat, P. Baranek and C. Adamo, 'Structural and electronic properties of selected rutile and anatase TiO_2 surfaces: an *ab initio* investigation,' *J. Chem. Theory Comput.* **4**, 341 (2008)

[197] F. Bruneval, '*GW* approximation of the many-body problem and changes in the particle number,' *Phys. Rev. Lett.* **103**, 176403 (2009)

[198] N. A. Deskins and M. Dupuis, 'Electron transport via polaron hopping in bulk TiO_2: a density functional theory characterization,' *Phys. Rev. B* **75**, 195212 (2007)

[199] S. Livraghi, M. Chiesa, M. C. Paganini and E. Giamello, 'On the nature of reduced states in titanium dioxide as monitored by electron paramagnetic resonance. I: the anatase case,' *J. Phys. Chem. C* **115**, 25413 (2011)

[200] S. X. Zhang, D. C. Kundaliya, W. Yu, S. Dhar, S. Y. Young, L. G. Salamanca-Riba, S. B. Ogale, R. D. Vispute and T. Venkatesan, 'Niobium doped TiO_2: intrinsic transparent metallic anatase versus highly resistive rutile phase,' *J. Appl. Phys* **102**, 013701 (2007)

[201] D. Morris, Y. Dou, J. Rebane, C. E. J. Mitchell, R. G. Egdell, D. S. L. Law, A. Vittadini and M. Casarin, 'Photoemission and stm study of the electronic structure of Nb-doped TiO_2,' *Phys. Rev. B* **61**, 13445 (2000)

[202] C. Di Valentin and A. Selloni, 'Bulk and surface polarons in photoexcited anatase TiO_2,' *J. Phys. Chem. Lett.* **2**, 2223 (2011)

[203] F. De Angelis, C. Di Valentin, S. Fantacci, A. Vittadini and A. Selloni, 'Theoretical studies on anatase and less common TiO_2 phases: bulk, surfaces, and nanomaterials,' *Chem. Rev.* **114**, 9708 (2014)

[204] E. Finazzi, C. D. Valentin, G. Pacchioni and A. Selloni, 'Excess electron states in reduced bulk anatase TiO_2: comparison of standard GGA, GGA+U, and hybrid DFT calculations,' *J. Chem. Phys.* **129**, 154113 (2008)

[205] T. Berger, M. Sterrer, O. Diwald, E. Knzinger, D. Panayotov, T. L. Thompson and J. T. Yates, 'Light-induced charge separation in anatase TiO_2 particles,' *J. Phys. Chem. B* **109**, 6061 (2005)

[206] I. R. Macdonald, R. F. Howe, X. Zhang and W. Zhou, '*In situ* EPR studies of electron trapping in a nanocrystalline rutile,' *J. Photochem. Photobiol. A* **216**, 238 (2010)

[207] S. Yang, A. T. Brant and L. E. Halliburton, 'Photoinduced self-trapped hole center in TiO_2 crystals,' *Phys. Rev. B* **82**, 035209 (2010)

[208] B. J. Morgan and P. A. Madden, 'Lithium intercalation into TiO_2(B): a comparison of LDA, GGA, and GGA+U density functional calculations,' *Phys. Rev. B* **86**, 035147 (2012)

[209] J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, 'CP2K: atomistic simulations of condensed matter systems,' *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 15 (2014)

[210] M. Guidon, F. Schiffmann, J. Hutter and J. VandeVondele, '*Ab initio* molecular dynamics using hybrid density functionals,' *J. Chem. Phys.* **128**, 214104 (2008)

[211] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, 'Quickstep: fast and accurate density functional calculations using a mixed Gaussian and plane waves approach,' *Comput. Phys. Commun.* **167**, 103 (2005)

[212] J. VandeVondele and J. Hutter, 'Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases,' *J. Chem. Phys.* **127**, 114105 (2007)

[213] M. Krack, 'Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals,' *Theor. Chem. Acc.* **114**, 145 (2005)

[214] W. Tang, E. Sanville and G. Henkelman, 'A grid-based Bader analysis algorithm without lattice bias,' *J. Phys. Condens. Matter* **21**, 084204 (2009)

[215] M. Yu and D. R. Trinkle, 'Accurate and efficient algorithm for Bader charge integration,' *J. Chem. Phys.* **134**, 064111 (2011)

[216] E. Sanville, S. D. Kenny, R. Smith and G. Henkelman, 'Improved grid-based algorithm for Bader charge allocation,' *J. Comput. Chem.* **28**, 899 (2007)

[217] J. K. Burdett, T. Hughbanks, G. J. Miller, J. W. Richardson and J. V. Smith, 'Structural-electronic relationships in inorganic solids: powder neutron diffraction studies of the rutile and anatase polymorphs of titanium dioxide at 15 and 295 K,' *J. Am. Chem. Soc.* **109**, 3639 (1987)

[218] J. Muscat, V. Swamy and N. M. Harrison, 'First-principles calculations of the phase stability of TiO_2,' *Phys. Rev. B* **65**, 224112 (2002)

[219] T. Zhu and S.-P. Gao, 'The stability, electronic structure, and optical property of TiO_2 polymorphs,' *J. Phys. Chem. C* **118**, 11385 (2014)

[220] J. Trail, B. Monserrat, P. López Ríos, R. Maezono and R. J. Needs, 'Quantum Monte Carlo study of the energetics of the rutile, anatase, brookite, and columbite TiO_2 polymorphs,' *Phys. Rev. B* **95**, 121108 (2017)

[221] Y. Luo, A. Benali, L. Shulenburger, J. T. Krogel, O. Heinonen and P. R. C. Kent, 'Phase stability of TiO_2 polymorphs from diffusion quantum Monte Carlo,' *New J. Phys.* **18**, 113049 (2016)

[222] W. H. Baur, 'Atomabstnde und bindungswinkel im brookit, TiO_2,' *Acta Crystallogr.* **14**, 214 (1961)

[223] T. Sasaki, M. Watanabe and Y. Fujiki, 'Structure of $K_{1.0}Ti_8O_{16}$ and $K_{0.0}Ti_8O_{16}$,' *Acta Crystallogr. B* **49**, 838 (1993)

[224] J. Akimoto, Y. Gotoh, Y. Oosawa, N. Nonose, T. Kumagai, K. Aoki and H. Takei, 'Topotactic oxidation of ramsdellite-type $Li_{0.5}TiO_2$, a new polymorph of titanium dioxide: TiO_2(R),' *J. Solid State Chem.* **113**, 27 (1994)

[225] M. Ben Yahia, F. Lemoigno, T. Beuvier, J.-S. Filhol, M. Richard-Plouet, L. Brohan and M.-L. Doublet, 'Updated references for the structural, electronic, and vibrational properties of TiO_2(b) bulk using first-principles density functional theory calculations,' *J. Chem. Phys.* **130**, 204501 (2009)

[226] A. Di Paola, M. Bellardita and L. Palmisano, 'Brookite, the least known TiO_2 photocatalyst,' *Catalysts* **3**, 36 (2013)

[227] J. Buckeridge, K. T. Butler, C. R. A. Catlow, A. J. Logsdail, D. O. Scanlon, S. A. Shevlin, S. M. Woodley, A. A. Sokol and A. Walsh, 'Polymorph engineering of TiO_2: demonstrating how absolute reference potentials are determined by local coordination,' *Chem. Mater.* **27**, 3844 (2015)

[228] O. Lamiel-Garcia, K. C. Ko, J. Y. Lee, S. T. Bromley and F. Illas, 'When anatase nanoparticles become bulklike: properties of realistic TiO_2 nanoparticles in the 16 nm size range from all electron relativistic density functional theory based calculations,' *J. Chem. Theory Comput.* **13**, 1785 (2017)

[229] B. J. Morgan and G. W. Watson, 'Intrinsic n-type defect formation in TiO_2: a comparison of rutile and anatase from GGA+U calculations,' *J. Phys. Chem. C* **114**, 2321 (2010)

[230] A. Janotti, C. Franchini, J. B. Varley, G. Kresse and C. G. Van de Walle, 'Dual behavior of excess electrons in rutile TiO_2,' *Phys. Status Solidi RRL* **7**, 199 (2013)

[231] A. Janotti, J. B. Varley, P. Rinke, N. Umezawa, G. Kresse and C. G. Van de Walle, 'Hybrid functional studies of the oxygen vacancy in TiO_2,' *Phys. Rev. B* **81**, 085212 (2010)

[232] E. Cho, S. Han, H.-S. Ahn, K.-R. Lee, S. K. Kim and C. S. Hwang, 'First-principles study of point defects in rutile TiO_{2-x},' *Phys. Rev. B* **73**, 193202 (2006)

[233] H. Iddir, S. Öğüt, P. Zapol and N. D. Browning, 'Diffusion mechanisms of native point defects in rutile TiO_2: *ab initio* total-energy calculations,' *Phys. Rev. B* **75**, 073203 (2007)

[234] H.-Y. Lee, S. J. Clark and J. Robertson, 'Calculation of point defects in rutile TiO_2 by the screened-exchange hybrid functional,' *Phys. Rev. B* **86**, 075209 (2012)

[235] P. Deák, B. Aradi and T. Frauenheim, 'Oxygen deficiency in TiO_2: similarities and differences between the Ti self-interstitial and the O vacancy in bulk rutile and anatase,' *Phys. Rev. B* **92**, 045204 (2015)

[236] U. Diebold, 'The surface science of titanium dioxide,' *Surf. Sci. Rep.* **48**, 53 (2003)

[237] A. K. Ghosh, F. G. Wakim and R. R. Addiss, 'Photoelectronic processes in rutile,' *Phys. Rev.* **184**, 979 (1969)

[238] S. Wendt, P. T. Sprunger, E. Lira, G. K. H. Madsen, Z. Li, J. Ø. Hansen, J. Matthiesen, A. Blekinge-Rasmussen, E. Lægsgaard, B. Hammer and F. Besenbacher, 'The role of interstitial sites in the Ti 3d defect state in the band gap of titania,' *Science* **320**, 1755 (2008)

[239] A. T. Brant, N. C. Giles, S. . Yang, M. A. R. Sarker, S. Watauchi, M. Nagao, I. Tanaka, D. A. Tryk, A. Manivannan and L. E. Halliburton, 'Ground state of the singly ionized oxygen vacancy in rutile TiO_2,' *J. Appl. Phys.* **114**, 113702 (2013)

[240] S. Yang, L. E. Halliburton, A. Manivannan, P. H. Bunton, D. B. Baker, M. Klemm, S. Horn and A. Fujishima, 'Photoinduced electron paramagnetic resonance study of electron traps in TiO_2 crystals: oxygen vacancies and Ti^{3+} ions,' *Appl. Phys. Lett.* **94**, 162114 (2009)

[241] J.-H. Yang, W.-J. Yin, J.-S. Park and S.-H. Wei, 'Self-regulation of charged defect compensation and formation energy pinning in semiconductors,' *Scientific Reports* **5**, 16977 (2015)

[242] A. Janotti and C. G. Van de Walle, 'LDA+U and hybrid functional calculations for defects in ZnO, SnO_2, and TiO_2,' *Phys. Status Solidi B* **248**, 799 (2011)

[243] H. Peng, 'First-principles study of native defects in rutile TiO_2,' *Phys. Lett. A* **372**, 1527 (2008)

[244] S. Abdelouahed and K. P. McKenna, 'Relevance of non-equilibrium defect generation processes to resistive switching in TiO_2,' *J. Appl. Phys.* **118**, 134103 (2015)

[245] A. Malashevich, M. Jain and S. G. Louie, 'First-principles DFT+GW study of oxygen vacancies in rutile TiO_2,' *Phys. Rev. B* **89**, 075205 (2014)

[246] G. Mattioli, P. Alippi, F. Filippone, R. Caminiti and A. Amore Bonapasta, 'Deep versus shallow behavior of intrinsic defects in rutile and anatase TiO_2 polymorphs,' *J. Phys. Chem. C* **114**, 21694 (2010)

[247] A. Boonchun, P. Reunchan and N. Umezawa, 'Energetics of native defects in anatase TiO_2: a hybrid density functional study,' *Phys. Chem. Chem. Phys.* **18**, 30040 (2016)

[248] J. M. Sullivan and S. C. Erwin, 'Theory of dopants and defects in Co-doped TiO_2 anatase,' *Phys. Rev. B* **67**, 144415 (2003)

[249] S. Livraghi, S. Maurelli, M. C. Paganini, M. Chiesa and E. Giamello, 'Probing the local environment of Ti^{3+} ions in TiO_2 (rutile) by 17O HYSCORE,' *Angew. Chem.* **50**, 8038 (2011)

[250] K. Yang, Y. Dai, B. Huang and Y. P. Feng, 'Density-functional characterization of antiferromagnetism in oxygen-deficient anatase and rutile TiO_2,' *Phys. Rev. B* **81**, 033202 (2010)

[251] A. G. Thomas, W. R. Flavell, A. R. Kumarasinghe, A. K. Mallick, D. Tsoutsou, G. C. Smith, R. Stockbauer, S. Patel, M. Grätzel and R. Hengerer, 'Resonant photoemission of anatase TiO_2 (101) and (001) single crystals,' *Phys. Rev. B* **67**, 035110 (2003)

[252] P. Scheiber, M. Fidler, O. Dulub, M. Schmid, U. Diebold, W. Hou, U. Aschauer and A. Selloni, '(Sub)surface mobility of oxygen vacancies at the TiO_2 anatase (101) surface,' *Phys. Rev. Lett.* **109**, 136103 (2012)

[253] M. Setvín, U. Aschauer, P. Scheiber, Y.-F. Li, W. Hou, M. Schmid, A. Selloni and U. Diebold, 'Reaction of O_2 with subsurface oxygen vacancies on TiO_2 anatase (101),' *Science* **341**, 988 (2013)

[254] T. Dittrich, J. Weidmann, F. Koch, I. Uhlendorf and I. Lauermann, 'Temperature- and oxygen partial pressure-dependent electrical conductivity in nanoporous rutile and anatase,' *Appl. Phys. Lett.* **75**, 3980 (1999)

[255] H. Kamisaka and K. Yamashita, 'Theoretical study of the interstitial oxygen atom in anatase and rutile TiO_2: electron trapping and elongation of the r(O-O) bond,' *J. Phys. Chem. C* **115**, 8265 (2011)

[256] H. Jeong, E. G. Seebauer and E. Ertekin, 'First-principles description of oxygen self-diffusion in rutile TiO_2: assessment of uncertainties due to enthalpy and entropy contributions,' *Phys. Chem. Chem. Phys.* **20**, 17448 (2018)

[257] A. G. Hollister, P. Gorai and E. G. Seebauer, 'Surface-based manipulation of point defects in rutile TiO_2,' *Appl. Phys. Lett.* **102**, 231601 (2013)

[258] L.-L. Tan, W.-J. Ong, S.-P. Chai and A. R. Mohamed, 'Band gap engineered, oxygen-rich TiO_2 for visible light induced photocatalytic reduction of CO_2,' *Chem. Commun.* **50**, 6923 (2014)

[259] H.-Y. Lee and J. Robertson, 'Doping and compensation in Nb-doped anatase and rutile TiO_2,' *J. Appl. Phys.* **113**, 213706 (2013)

[260] S. Na-Phattalung, M. F. Smith, K. Kim, M.-H. Du, S.-H. Wei, S. B. Zhang and S. Limpijumnong, 'First-principles study of native defects in anatase TiO_2,' *Phys. Rev. B* **73**, 125205 (2006)

[261] V. Etacheri, M. K. Seery, S. J. Hinder and S. C. Pillai, 'Oxygen rich titania: a dopant free, high temperature stable, and visible-light active anatase photocatalyst,' *Adv. Funct. Mater.* **21**, 3744 (2011)

[262] H. Nogawa, T. Hitosugi, A. Chikamatsu, S. Nakao, Y. Hirose, T. Shimada, H. Kumigashira, M. Oshima and T. Hasegawa, 'Carrier compensation by excess oxygen atoms in anatase $Ti_{0.94}Nb_{0.06}O_{2+\delta}$ epitaxial thin films,' *Jpn. J. Appl. Phys.* **49**, 041102 (2010)

[263] M. K. Nowotny, T. Bak, J. Nowotny and C. C. Sorrell, 'Titanium vacancies in nonstoichiometric TiO_2 single crystal,' *Phys. Status Solidi B* **242**, R88 (2005)

[264] M. K. Nowotny, T. Bak and J. Nowotny, 'Electrical properties and defect chemistry of TiO_2 single crystal. IV. Prolonged oxidation kinetics and chemical diffusion,' *J. Phys. Chem. B* **110**, 16302 (2006)

[265] M. K. Nowotny, T. Bak and J. Nowotny, 'Electrical properties and defect chemistry of TiO_2 single crystal. II. Thermoelectric power,' *J. Phys. Chem. B* **110**, 16283 (2006)

[266] M. K. Nowotny, T. Bak and J. Nowotny, 'Electrical properties and defect chemistry of TiO_2 single crystal. I. Electrical conductivity,' *J. Phys. Chem. B* **110**, 16270 (2006)

[267] M. K. Nowotny, T. Bak and J. Nowotny, 'Electrical properties and defect chemistry of TiO_2 single crystal. III. Equilibration kinetics and chemical diffusion,' *J. Phys. Chem. B* **110**, 16292 (2006)

[268] E. Finazzi, C. Di Valentin and G. Pacchioni, 'Nature of Ti interstitials in reduced bulk anatase and rutile TiO_2,' *J. Phys. Chem. C* **113**, 3382 (2009)

[269] M. A. Henderson, S. Otero-Tapia and M. E. Castro, 'The chemistry of methanol on the TiO_2(110) surface: the influence of vacancies and coadsorbed species,' *Faraday Discuss.* **114**, 313 (1999)

[270] M. Li, W. Hebenstreit and U. Diebold, 'Oxygen-induced restructuring of the rutile TiO_2(110)(1 $\times$ 1) surface,' *Surf. Sci.* **414**, L951 (1998)

[271] M. Li, W. Hebenstreit, U. Diebold, M. A. Henderson and D. R. Jennison, 'Oxygen-induced restructuring of rutile TiO_2(110): formation mechanism, atomic models, and influence on surface chemistry,' *Faraday Discuss.* **114**, 245 (1999)

[272] M. Aono and R. R. Hasiguti, 'Interaction and ordering of lattice defects in oxygen-deficient rutile TiO_{2-x},' *Phys. Rev. B* **48**, 12406 (1993)

[273] M. Chiesa, M. C. Paganini, S. Livraghi and E. Giamello, 'Charge trapping in TiO_2 polymorphs as seen by electron paramagnetic resonance spectroscopy,' *Phys. Chem. Chem. Phys.* **15**, 9435 (2013)

[274] A. T. Brant, S. . Yang, N. C. Giles and L. E. Halliburton, 'Hydrogen donors and Ti^{3+} ions in reduced TiO_2 crystals,' *J. Appl. Phys.* **110**, 053714 (2011)

[275] S. Wang, L. Pan, J.-J. Song, W. Mi, J.-J. Zou, L. Wang and X. Zhang, 'Titanium-defected undoped anatase TiO_2 with p-type conductivity, room-temperature ferromagnetism, and remarkable photocatalytic performance,' *J. Am. Chem. Soc.* **137**, 2975 (2015)

[276] R. F. Howe and M. Gratzel, 'EPR study of hydrated anatase under UV irradiation,' *J. Phys. Chem.* **91**, 3906 (1987)

[277] A. R. Elmaslmane, M. B. Watkins and K. P. McKenna, 'First-principles modeling of polaron formation in TiO_2 polymorphs,' *J. Chem. Theory Comput.* **14**, 3740 (2018)

[278] T. R. Durrant, S. T. Murphy, M. B. Watkins and A. L. Shluger, 'Relation between image charge and potential alignment corrections for charged defects in periodic boundary conditions,' *J. Chem. Phys.* **149**, 024103 (2018)

[279] G. Makov and M. C. Payne, 'Periodic boundary conditions in ab initio calculations,' *Phys. Rev. B* **51**, 4014 (1995)

[280] W. Chen and A. Pasquarello, 'Correspondence of defect energy levels in hybrid density functional theory and many-body perturbation theory,' *Phys. Rev. B* **88**, 115104 (2013)

[281] A. Ganose, A. Jackson and D. Scanlon, 'sumo: Command-line tools for plotting and analysis of periodic *ab initio* calculations,' *Journal of Open Source Software* **3(28)**, 717 (2018)

[282] G. D. Watkins, *Negative-U properties for defects in solids*, (Springer 1984)

[283] Q. Liu, G. M. Dalpian and A. Zunger, 'Antidoping in insulators and semiconductors having intermediate bands with trapped carriers,' *Phys. Rev. Lett.* **122**, 106403 (2019)

[284] J. J. Carey and K. P. McKenna, 'Does polaronic self-trapping occur at anatase TiO_2 surfaces?' *J. Phys. Chem. C* **122(48)**, 27540 (2018)

[285] J. A. Quirk, V. K. Lazarov and K. P. McKenna, 'Electronic properties of 112 and 110 twin boundaries in anatase TiO_2,' *Adv. Theory Simul.* **2**, 1900157 (2019)

[286] J. J. Carey and K. P. McKenna, 'Screening doping strategies to mitigate electron trapping at anatase TiO_2 surfaces,' *J. Phys. Chem. C* **123**, 22358 (2019)

[287] J. H. Skone, M. Govoni and G. Galli, 'Self-consistent hybrid functional for condensed systems,' *Phys. Rev. B* **89**, 195112 (2014)