University of Southampton

Faculty of Natural and Environmental Sciences

Chemistry

Density Functional Theory Applied to Metallic Nanoparticles

Author:
Jolyon Aarons

Supervisor:
Prof. Chris-Kriton Skylaris

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Computational Chemistry Group
SCHOOL OF CHEMISTRY

January 2018

"Other than the laws of physics, rules have never really worked out for me."

Craig Ferguson

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Chemistry

Doctor of Philosophy

DENSITY FUNCTIONAL THEORY APPLIED TO METALLIC NANOPARTICLES

by Jolyon Aarons

This thesis will focus on DFT for calculations of large metallic nanoparticles. It will show new algorithms that were developed for reduced scaling DFT methods for metals; the testing, verification and design of new descriptors for predicting the catalytic activity of metallic nanoparticles; application of large-scale DFT calculations to model nanoparticle sequences to show size and oxygen adsorption coverage trends, and finally the application of these techniques and knowledge to perform a study of oxygen adsorption on real-world, experimentally determined platinum nanoparticles in collaboration with the Nellist group at Oxford materials.

We explore the binding of atomic oxygen to cuboctahedral platinum nanoparticles of up to ∼1000 atoms using DFT calculations in ONETEP. We demonstrate convergence to the infinite slab limit for single oxygen adsorption in chapter 4 and correlate adsorption strength against popular descriptors for catalytic activity, such as the d-band centre approach. This approach is possible because of work which will be described in chapter 3 to implement angular momentum projected density of states calculations in ONETEP. The effects of oxygen coverage on the Pt_{55} and Pt_{147} cuboctahedral nanoparticles will also be analysed, which serves to advance our simulations towards realistic conditions.

We show in our investigation into half monolayer, hemispherical oxygen coverage on platinum nanoparticles that oxygen tends to gravitate towards the edges and lower coordinated sites in the nanoparticle and away from the centres of facets. This effect correlates with the site specific, single oxygen adsorption energies on Pt_{309} and experimental platinum nanoparticles which is presented in chapter 5. We show that when subdividing the binding of monolayers of oxygen into only (111) and (100) facets that these have a lower adsorption strength per oxygen atom than combined (100) and (111) facets as well as lower binding strength than single oxygen adsoprtion.

In the next part of the study, which is discussed in chapter 5, we show large scale DFT calculations on real platinum nanoparticles, which were measured by the Nellist group at Oxford materials using advanced electron microscopy techniques. These DFT calculations provide the electronic structure of the experimentally measured nanoparticles, which allowed us to apply electron density based catalytic activity descriptors to the nanoparticles, such as the d-band centre approach, or our own electronic density based descriptor described in chapter 3. We find that surface roughness of the experimental nanoparticles contributes to more potential oxygen binding sites with low electron density, which correlatates with stronger oxygen adsorption strength in our model, when compared with the relative smoothness of cuboctahedral and truncated octahedral facets. In the analysis which is presented in chapter 5, the proportion of sites which lie within 0.2 eV of the oxygen binding strength required for optimum catalytic activity is predicted with high efficiency, based on our catalytic activity descriptor.

Finally, in chapter 6 we describe a new method for large scale DFT calculations on metallic systems which we call the AQuA-FOE method. We show how this method can have a computational cost which increases effectively linearly with the number of atoms. The AQuA-FOE method works by implicitly heating and quenching the electrons in the system to find the one-particle density matrix, while conserving the electron number. We show validation of this method inside the EDFT procedure by comparing numerically with the diagonalisation based EDFT that is already implemented in ONETEP showing agreement in the energies to better than 10^{-5} E_H per atom. We will also demonstrate the effectively linear-scaling computational cost of our method with calculation times on regular truncated octahedral Palladium nanoparticles ranging from 2,406 to 12,934 atoms.

Contents

List of Tables

List of Figures

Declaration of Authorship

I, Jolyon Aarons, declare that this thesis titled, 'Density Functional Theory Applied to Metallic Nanoparticles' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
- Parts of the work presented in this thesis have been published in academic journals [8, 11, 12] and my contribution has been clearly stated in those chapters.

Signed: _____

Date: _____

Acknowledgements

I would like to thank my supervisor, Chris for keeping the project on track and keeping a cool head, even when I lost it.
Thanks to my family: Mum, Dad, Robin and Toby, and Jingdi for all the support throughout my time in Southampton, I couldn't have got to this point without you.
Thanks to the EPSRC and to Johnson Matthey for funding my project.
Thanks to Valerio for being such a good friend and for keeping me (mostly) sane while in Southampton, we had a great time I guess. Thanks for acting as the celebrant at my wedding. Thanks also to Dr. Martina Stella for keeping Valerio (mostly) sane.
Thanks to Dr. Jacek Dziedzic for being an always approachable and helpful source of all knowledge. I have met few other people who I think could be considered a polymath, but Jacek is a true renaissance man.
In my first two years, Dr. Dave Case was another great inspiration for me. Being in different offices and research groups, we mainly spoke in the pub. Despite this, Dave told me things about chemistry which I never would have known otherwise and somewhat incredibly Dave is remarkably well read in the arts and humanities as well as science. A pub session never went by without an interesting fact about Afghanistan. I also want to express my deep gratitude to Dave for helping my wife when her cousin was severely ill.
Almost as soon as I joined the department in Southampton, I met Tom Gee (now Dr. Tom Gee). Always affable and always a good laugh Tom helped me get through several difficult weeks by cracking jokes and listening to my moaning on Friday evenings. Thanks!
Lucas Garcia Verga. Since joining the group Lucas' considerable talent in physical sciences became obvious. I have benefitted from his advice and deep insight on many occasions. Thanks to Lucas for that and also for his sound "internal politics" advice, as well as for continuing some of the work that I began.
Dr. Hairul Hisham Hamzah for reliably coming with us on free coffee and donut day and for being always jovial, even when extremely stressed. Thanks also for just being a good friend.
Tom Ellaby has worked tirelessly since arriving in the group to get up to speed with calculations, both for the continuation of this project and for the diabolically difficult alumina calculations. Thanks for that because it took a lot of stress off me, but also for having interesting anecdotes about most things and being a generally outgoing person.
Thanks to Dr. James Womack for teaching me that there's more to computational chemistry than just DFT! Jwack is knowledgeable about almost everything in the field and has been a great source of help and suggestions whenever I have had either scientific or computational issues.
Prof. Andrea Russell for being my academic advisor, her suggestions regarding our adsorption calculations were invaluable for steering the project in the direction of usefulness. I also want to express my gratitude to her for the support and encouragement she gave me during some difficult times in my project.
Many thanks to Dr. Misbah Sarwar. I am very sure that without her as my industrial supervisor at Johnson Matthey this project would not have got off the ground. Misbah provided important advice, suggestions and project course corrections from the very beginning, as well as putting us in contact with the Nellist group for our experimental collaboration. Thanks also to Misbah

for taking the time to carefully read through this thesis at the last minute! Your comments and suggestions were invaluable, I only hope I can implement everything before I submit.
Dr. Lewys Jones is one of the most enthusiatic and capable scientists I have ever met. No matter what we presented of recent calculations, Lewys could always find an interesting story in the results and some offer some insight I couldn't provide myself. When I spent some time in Oxford to finish the paper Lewys was accomodating and helpful as well as extremely hardworking. Thanks!
Thanks to Alex Gheorghiu for being a good laugh. Even though Alex only spent a year in the group, he had a lasting impact on morale. An extremely hard working student who has the right mixture of cynicism and optimism. "Here's my results... but it's just a bunch of numbers in a row".
Thanks to Steve Fox for helping me find my bearings in my first year in Southampton, with setting up my computer and for being an enthusiastic lunch companion with always a story to tell!
Karl Wilkinson although no longer with us, continues to inspire me by his example, and his dedication to his wife and career over the course of his life. I will never forget how despite being seriously ill, Karl somehow kept a sense of humour and managed to persevere with his work.
Max Phipps is just one of those guys. Loveable and frustrating in equal measure. I count Max as being one of the good friends I made in Southampton, I've never met anyone quite like him. I mean that in a good way. I think.
Hong-Tao Xue for helping me on my quest to learn Chinese, for interesting anecdotes and for the audacious, but profound truths he spoke.
Chris Sampson. The office wouldn't have been the same without Sampson. The culture of the office still reflects Sampson's personality. Thanks also to Sampson for giving me some good advice when I was facing a difficult situation of my own making and also for forgiving me for some of my stupid antics.
Phil Blowey for being a conscientious final year project student in my first year with excellent repartee (and badminton skills).
Thanks to Dr. Álvaro Ruiz-Serrano for doing the necessary work in ONETEP to make my project possible. Thanks also for discussing possible directions for the project with me on my first days in Southampton. That advice proved to be invaluable to me.
Gabriele Boschetto is one of the most easy-going and friendly people I've met and made core hours a pleasure (almost). Thanks for all the funny conversations over coffee.
Thanks to Frank Longford for finally making the group less isolated and helping us to socialise with other groups.
Thanks also to the rest of the office guys, you know who you are!
Thanks to Dave Inwood for being a approachable and helpful friend on the experimental side, as well as for interesting conversations.
Dr. Nick Hine for always being approachable about problems in ONETEP and helping with fixing some of the problems which I created! Thanks also for offering me a job in Warwick, I am looking forward to working with you!
Dr. Doğan Özkaya for some really great suggestions in our meetings in Harwell, as well as being so enthusiastic about the work. It was inspiring.

Thanks to Kevin Duff for bearing with me while I write this thesis to get back to you about the changes in EDFT.
Thanks to Prof. Matt Probert for lending a sympathetic ear on many occasions during this project.
Thanks to Dr. Phil Hasnip for invaluable advice and suggestions for making electronic optimisers faster for metallic systems.
Dr. Ondrej Hovorka and Veronika have been a constant in my life since we left York and arrived in Southampton at the same time. Thank you a lot for all the good times and the support and suggestions throughout the whole of my project.
Thanks to Prof. Rex Godby for giving me firm advice about PhD projects before I started.
Thanks to Dr. David Kinnison for teaching me the value of workplace safety and for letting me off with (stern) warnings on two occasions.
Prof. Peter Nellist for remaining enthusiastic about our collaboration and for driving the work forwards.
Dr. Syma Khalid for being sympathetic to me at a time when I hadn't earned any sympathy, also thanks for several interesting conversations in the pub / Christmas meals afterwards.
Dr. Edmond Lee for being an astonishingly tolerant office companion for the first 3 and a half years of my project.
Since I was in York, Dr. Gionni Marchetti has remained a good friend to me and my wife and despite being far apart these days. Thanks for the impromptu conversations every so often!
Thanks to the guys from Matt and Rex's group office in York such as Ed Higgins, Dr. Matthew Hodgson and Dr. James Ramsden for a friendly welcome every time I visited York during this project, as well as for coming to my wedding. In particular, thanks to Dr. Aaron Hopkinson who offered me some very helpful support on more than one occasion.
Thanks to the Essex group for occasionally inviting us to things.
Thanks to Andreas Friberg for supportive comments relating to the submission of this thesis.
Finally, thanks to Ann French for keeping me fueled with caffeine and biscuits throughout my time in Southampton.

Abbreviations

ONETEP	**O**rder-**N** **E**lectronic **T**otal **E**nergy **P**ackage
NGWF	**N**on-orthogonal **G**eneralised **W**annier **F**unction
(**E**)**DFT**	(**E**nsemble) **D**ensity **F**unctional **T**heory
LDA	**L**ocal **D**ensity **A**pproximation
GGA	**G**eneralised **G**radient **A**pproximation
(**p**)**DOS**	([angular momentum] **p**rojected) **D**ensity **O**f **S**tates
(**F**)**OE**	(**F**ermi) **O**perator **E**xpansion
XC	e**X**change **C**orrelation ([functional])
PEM(FC)	**P**roton **E**xchange **M**embrane (**F**uel **C**ell)
ORR	**O**xygen **R**eduction **R**eaction
MO	**M**olecular **O**rbital
PW	**P**lane **W**ave
SW	**S**pherical **W**ave
PAW	**P**rojector **A**ugmented **W**ave
CG	**C**onjugate **G**radient
PExSI	**P**ole **Ex**pansion with **S**elective **I**nversion
ADF-STEM	**A**nnular **D**ark **F**ield - **S**canning **T**ransmission **E**lectron **M**icroscopy
PaPaFOE	**Pa**dé **Pa**rtitioned **F**ermi **O**perator **E**xpansion
AQuAFOE	**A**nnealing and **Qu**enching **A**lgorithm for **F**ermi **O**perator **E**xpansion

For Jingdi.

Chapter 1

Introduction

1.1 Background

The physical laws governing the interactions of subatomic physical particles are known as quantum mechanics. Having been vindicated by countless experiments from the early 20^{th} century up to the present day, quantum mechanics has allowed some of the most accurate predictions of the physical world in all of science.
The principal equation in quantum mechanics is the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi, \tag{1.1}$$

where i is the unit imaginary number, $\hbar$ is the reduced Planck constant, $h/2\pi$, Ψ is the wavefunction which represents the state of the system and $\hat{H}$ is the Hamiltonian operator, which operates on the wavefunction to give the energy of the system. The Schrödinger equation is a wave equation which represents how the system evolves in time, but for stationary states, where the wave function is separable and remains stationary in space, provided the Hamiltonian operator is time independent, the time independent Schrödinger equation is written as an eigenvalue equation as

$$\hat{H}\Psi = E\Psi, \tag{1.2}$$

where E are the energy eigenvalues of the Hamiltonian. In practice, the Schrödinger equation is solvable only for very few systems analytically and becomes rapidly infeasible to solve numerically with increasing system size.

1.1.1 The Problem of the Many-body Wavefunction

The fundamental object in the original formulation of quantum mechanics is the wavefunction. The wavefunction holds all information about an isolated system and physical observables are obtained from it by applying operators. For a single particle, before

considering spin, it is expressed as

$$\Psi(\vec{r}, t), \tag{1.3}$$

where the arguments t and $\vec{r}$ index over time and the three dimensioned coordinate space of the particle. As particles are added to the system, the dimensionality of the wavefunction increases, for instance, for two particles:

$$\Psi(\vec{r_1}, \vec{r_2}, t), \tag{1.4}$$

the spatial dimensionality is now 6. This has fairly serious implications, in that when we consider the general N-body problem in quantum mechanics, the wavefunction will have $3N + 1$ dimensions;

$$\Psi(\vec{r_1}, \vec{r_2}, \cdots, \vec{r_N}, t). \tag{1.5}$$

In order to solve a quantum mechanical problem, such as this on a digital computer it would be necessary to discretise space and time, but even assuming a very coarse grid in space in each dimension, it is apparent that merely the storage of such an object becomes rapidly infeasible. This is because for each value of each dimension, every value of every other dimension must be considered. When moving from 1D to 2D, given a constant grid spacing along each dimension, the number of values to be evaluated is squared, and cubed when going to 3D; likewise, when considering a generalised Hilbert space of $3N$ dimensions, the number of grid-points to be evaluated is n^{3N}, where n is the number of grid-points along each dimension. If a ridiculously coarse grid-spacing is chosen simply for the sake of the argument, for instance: 2 grid points in each dimension, then only 13 particles are needed before the wavefunction alone would fill the largest commonly available consumer hard disk drive twice over, as of 2014 (assuming four byte, double precision arithmetic):

$$D_{14,2} = ((2^{3*14}) * 4\text{B})/(1024 * 1024 * 1024 * 1024) = 16\text{TB}. \tag{1.6}$$

Now, given a slightly less extremely (but still extreme) coarse grid of 10 grid-points along each dimension,

$$D_{14,10} = ((10^{3*14}) * 4\text{B})/(1024 * 1024 * 1024 * 1024) = 3.6 * 10^{30}\text{TB}, \tag{1.7}$$

which is 9 billion times the number of transistors ever produced up until 2015, according to Intel[13]. Try to store the many-body wavefunction for 30 particles at this coarse sampling and more Terabytes would be needed than there are atoms in the observable universe. So quantum mechanics offers predictive power unrivaled in accuracy on the atomic and electronic scale, but this power can never be fully harnessed in the many-body wavefunction formulation.

1.2 Density Functional Theory

Density Functional Theory (DFT) offers a way to sidestep this problem. The main premise of DFT is that ground state observables (physical properties of the system at its lowest energy configuration) are available from the 3 dimensioned density of the system, rather than the $3N$ dimensioned wavefunction. The way to show why this is so, is through the Hohenberg-Kohn theorems which are presented in the next sub-section.

1.2.1 The Hohenberg-Kohn Theorems

Assuming first of all that there exist two external potentials, $V_1(\vec{r})$ and $V_2(\vec{r})$ which both give rise to the same charge density, n_0 which are associated with two Hamiltonians, $\hat{H}_1$ and $\hat{H}_2$ and wavefunctions, Ψ_1 and Ψ_2.
By the variational principle:

$$E_1 = \langle \Psi_1 | \hat{H}_1 | \Psi_1 \rangle < \langle \Psi_2 | \hat{H}_1 | \Psi_2 \rangle, \tag{1.8}$$

meaning that the Hamiltonian of a system has a lower minimum energy eigenvalue when applied to the wavefunction of the same system compared with when applied to a different wavefunction. Where $\langle \Psi_2 | \hat{H}_1 | \Psi_2 \rangle$ can be written as:

$$\langle \Psi_2 | \hat{H}_2 | \Psi_2 \rangle + \langle \Psi_2 | \hat{H}_1 - \hat{H}_2 | \Psi_2 \rangle \tag{1.9}$$

and because the number of electrons is constant in both systems, the kinetic and electronic part of $\hat{H}_1 - \hat{H}_2$ cancel, giving $\hat{V}_{ext1} - \hat{V}_{ext2}$, so:

$$E_1 < E_2 + \int [V_2(\vec{r}) - V_1(\vec{r})] n_0(\vec{r}) d\vec{r} \tag{1.10}$$

But, for $\hat{H}_2$:

$$E_2 < E_1 + \int [V_1(\vec{r}) - V_2(\vec{r})] n_0(\vec{r}) d\vec{r} \tag{1.11}$$

And when added together, the latter two equations give:

$$E_1 + E_2 < E_1 + E_2 \tag{1.12}$$

Proving by contradiction, therefore, that the external potential and hence, the ground state energy of a system of electrons is a unique functional of the ground state charge density.[14] This is the first Hohenberg-Kohn theorem.
The Hamiltonian can be written as

$$\hat{H} = \hat{T} + \hat{V_{ee}} + \hat{V_{ext}}, \tag{1.13}$$

where the potential term has been split into electron-electron and external parts. The wavefunction which minimises the expectation value of the Hamiltonian gives the ground-state energy and this wavefunction has been shown to be unique to one charge density. It follows that Ψ is a functional of density.

$$\Psi = \Psi[n(\vec{r})] \tag{1.14}$$

Consequently, the ground-state energy must also be a functional of density.

$$E[n(\vec{r})] = \langle \Psi | \hat{H} | \Psi \rangle \tag{1.15}$$

Since the energy contribution of an electro-static potential is known from electromagnetism (eq.1.16), the total ground-state energy of the system can be written as eq.1.17 for some external potential.

$$U = \frac{1}{2} \int V(\vec{r}) n(\vec{r}) d\vec{r} \tag{1.16}$$

$$E[n(\vec{r})] = F[n(\vec{r})] + \int V(\vec{r}) n(\vec{r}) d\vec{r} \tag{1.17}$$

Also, if the ground state density n_0 of a system corresponds uniquely to a particular Hamiltonian $\hat{H}_1$, then some other density n_1, which corresponds to a different Hamiltonian $\hat{H}_2$, the energy of the system is given by (eq.1.15) and by (eq.1.8) it is clear that there is some charge density which minimises the total energy of the system, which is consequently the ground-state energy. Applying the variational principle leads to the Euler-Lagrange minimisation problem (eq.1.18). This is the second Hohenberg-Kohn theorem.

$$\mu = \frac{\partial F[n(\vec{r})]}{\partial n(\vec{r})} + V(\vec{r}) \tag{1.18}$$

1.2.2 Kohn-Sham DFT

So far, the formulation of DFT has been conceptually elegant, but not particularly useful for actual computation, as it is known that there is a unique functional of the charge density which minimises the total energy of the system, but there is no clear way to know what this functional actually is.

Kohn-Sham DFT presents a way to evaluate at least part of the functional, through the reintroduction of wavefunctions. Rather than the many-body beast introduced previously, however, this is a wavefunction of a system of non-interacting particles and hence, its Hilbert space after discretisation spans Nn^3 values rather than n^{3N} in the interacting case.

Solving a system of non-interacting particles in an effective potential designed to give the same density (and hence energy) of the interacting problem is equivalent to solving

the eigenvalue equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\text{eff}}(\vec{r})\right)\phi_i(\vec{r}) = \varepsilon_i\phi_i(\vec{r}), \tag{1.19}$$

where the density of the system is written as

$$n(\vec{r}) = \sum_i^N |\phi_i(\vec{r})|^2. \tag{1.20}$$

In order to determine the effective potential v_{eff} in Kohn-Sham DFT, it is necessary to compute all components of the energy functional,

$$E[n] = T[n] + \int d\vec{r}\; v_{\text{ext}}(\vec{r})n(\vec{r}) + V_H[n] + E_{\text{XC}}[n], \tag{1.21}$$

including both those which are expressed in terms of Kohn-Sham orbitals and those expressed in terms of the electronic density. In the above expression $T[n]$ is the Kohn-Sham Kinetic energy, V_H is the Hartree potential, or the electron-electron Coulombic potential v_{ext} is the external potential, which is usually the electron-nucleus potential and E_{XC} is the exchange correlation energy. The Kinetic energy is expressed in terms of the Kohn-Sham wave functions as

$$T[n] = \sum_{i=1}^N \int d\vec{r}\; \phi_i^*(\vec{r})\left(-\frac{\hbar^2}{2m}\nabla^2\right)\phi_i(\vec{r}), \tag{1.22}$$

All the other terms are written in terms of the density, which is calculated from the wave functions as above, where the Hartree term is given as:

$$V_H = \frac{e^2}{2}\int d\vec{r}\int d\vec{r}'\;\frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|}. \tag{1.23}$$

By inspection, the kinetic energy term is already present in the Kohn-Sham eigenvalue equation, so that the external potential is given from the energy functional as:

$$v_{\text{eff}}(\vec{r}) = v_{\text{ext}}(\vec{r}) + e^2\int\frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r}' + \frac{\delta E_{\text{xc}}[n]}{\delta n(\vec{r})} \tag{1.24}$$

where the last, exchange-correlation term,

$$v_{\text{xc}}(\vec{r}) \equiv \frac{\delta E_{\text{xc}}[n]}{\delta n(\vec{r})}, \tag{1.25}$$

has no known analytic form and must be approximated. The minimum ground state density can then be found (subject to a decent approximation for the XC functional) by variationally optimising with respect to the wave functions / orbitals. Because the orbitals give the density, the density gives the potential and the potential gives a new

set of orbitals, this can be solved to self-consistency and is often known as self-consistent field (SCF) Kohn-Sham DFT.

The idea is that if the charge density of a system uniquely determines its wavefunction and hence its physical observables in a one to one relationship, then it does not matter how this charge density is constructed. So, if a ficticious system of non-interacting particles is designed to give the same charge density as the physical system of interacting particles, then the Kinetic energy of the system could be calculated[1].

1.2.3 Exchange-Correlation Functionals

In Kohn-Sham DFT, all of the unknowns are bundled up into what is known as the Exchange-Correlation (XC) term. This is in effect what is not known of the mapping from the system of interacting particles on the the system of non-interacting Kohn-Sham particles. What is remarkable, however, is that despite not knowing a form for this term, a great many approximations have been proposed, some of which have been tremendously sucessful, in terms of accuracy and predictive power.

The exchange-correlation term may be further divided into two terms. The exchange energy, otherwise known as Pauli repulsion is the energy associated with the change in sign of the wavefunction upon exchange of indistinguishable Fermions[2], which serves to keep them separated from each other. The second term is the correlation energy, which represents any electron-electron interaction outside of the Coulombic interaction of Hartree's term. The XC term can then be written as the sum,

$$E_{XC}(n) = E_X(n) + E_C(n). \tag{1.26}$$

1.2.3.1 Local Density Approximation (LDA)

The LDA class of XC functionals are the simplest class which are useful in practical calculations. LDA functionals are most broadly defined as those which are functionals solely of the electronic density field and the most widely used and well-known of these are all based on the homogeneous electron gas model. This represents an approximation that maps electron density at each point in the simulation cell to a homogeneous electron gas of the same density, for which analytic expressions exist for the calculation of exchange and correlation terms;

$$E_{XC}^{\mathrm{LDA}}[n] = \int n(\mathbf{r})\epsilon_{XC}(n)\ \mathrm{d}\mathbf{r}, \tag{1.27}$$

[1] It is worth noting that this does not violate the first Hohenberg-Kohn theorem because HK-1 states only that there is a unique mapping between the charge density and many-body wavefunction for interacting particles, but says nothing against producing a matching charge density from different kinds of particle

[2] It has been assumed from this point onward, that particles are referred to that they are spin 1/2 integer Fermions unless otherwise stated.

where $\epsilon_{XC}(n)$ is an exchange-correlation energy distribution function for charge density n.

The exchange energy of a homogeneous electron gas is known exactly[15],

$$E_{\mathrm{X}}^{\mathrm{LDA}}[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int n(\mathbf{r})^{4/3}\, \mathrm{d}\mathbf{r} \tag{1.28}$$

which makes expressions for the XC energy in terms of the homogenous electron gas appealing. The correlation energy term is not known analytically, however, and provides most of the variation between the various LDA XC functionals. Limits of the correlation energy are known in the high[16] and low[17] density limits, together with computed intermediate values (typically computed with quantum Monte Carlo techniques), then functions which interpolate between the various pieces of evidence and match the behaviour in the limits are produced[18], with the particular qMC measurements determining the major differences between the various LDA schemes.

LDA schemes are expected to be a good approximation in the case where the density is slowly varying, because the sampling from the independent homogenous electron gas problems at each point in the simulation cell are discontinuous. If the density is slowly varying, the discontinuity will be minimised. Despite this, remarkable agreement from LDA results has been shown for metallic systems

1.2.3.2 Generalised Gradient Approximation (GGA)

As opposed to LDA functionals, GGA functionals take into account the gradient vector field of electron density as well as solely the electron density.

$$E_{XC}^{\mathrm{GGA}}[n] = \int n(\mathbf{r})\epsilon_{XC}(n, \Delta n)\, \mathrm{d}\mathbf{r}, \tag{1.29}$$

This is important because in many systems under study the electron density may vary more rapidly than is appropriate for modelling with an LDA functional based on the homogeneous electron gas, however, due to the many successes of LDA functionals, it would be good to base improvements upon such a model.

Most GGAs seek to reproduce the correct physics predicted by the LDA in some well defined limits as well as fitting computed datapoints in and respecting constraints on the electron density gradient. One of the first useful and reliable GGA functionals was the PW91.

The Perdew-Burke-Ernzerhof, or PBE functional was developed as a simplification to the methodology used to conglomerate all of the data used in the development of the PW91 functional, while retaining all of the relevant physics [19]. It is one of the most widely used XC functionals because of its versatility.

The PBE functional defines the exchange contribution as

$$E_{\mathrm{X}}^{\mathrm{PBE}}\left[n(\mathbf{r})\right] = \int n(\mathbf{r})\epsilon_{\mathrm{X}}^{\mathrm{LDA}}(n)F_{\mathrm{X}}(s)d\mathbf{r}, \tag{1.30}$$

where

$$F_{\mathrm{X}}(s) = 1 + \kappa - \frac{\kappa}{1 + bs^2/\kappa}, \tag{1.31}$$

and

$$s = \frac{|\nabla n|}{2k_F n}. \tag{1.32}$$

The parameter $k_F = (3\pi^2 r)^{1/3}$ and $\kappa = 0.84$ and $b = 0.21951$. These were derived by the authors from the Lieb-Oxford bound and by requiring that the gradient coefficient for correlation cancels the exchange coefficient.
As for the correlation contribution in the PBE functional, this is written as

$$E_{\mathrm{C}}^{\mathrm{PBE}}\left[n(\mathbf{r})\right] = \int n\left[\epsilon_{\mathrm{C}}^{\mathrm{LDA}}(n) + h(n,t)\right]d\mathbf{r}, \tag{1.33}$$

where

$$h(n,t) = \gamma \ln\left\{1 + \frac{\beta t^2}{\gamma}\left[\frac{1 + At^2}{1 + At^2 + A^2t^4}\right]\right\}, \tag{1.34}$$

$$A = \frac{\beta}{\gamma e^{-\epsilon_{\mathrm{C}}^{\mathrm{LDA}}/\gamma} - 1}, \tag{1.35}$$

and

$$t = \frac{|\nabla n|}{2k_s n}, \tag{1.36}$$

where the parameter $k_s = \sqrt{4k_F/\pi}$ and $\beta = 0.066725$ and $\gamma = 0.031091$. The β and γ parameters were derived by the authors from the high-density limit of the gradient expansion of the homogeneous electron gas correlation energy expression. The PBE functional is one of the most widely used of the generalised gradient exchange correlation options and we will make extensive use of its reparameterised cousin, rPBE in this work. rPBE is a re-parameterisation of PBE developed within the Nørskov group, which sought to improve adsorption energies, particularly in the context of catalysis, i.e. reaction barriers [20]. This XC functional has been shown to be good over a wide class of systems and observables [21], it is one of the most accurate of the GGA class of functionals.

1.2.3.3 Hybrid Functionals

It must be stated that every known XC functional is an approximation. Many of them are far more accurate than one might expect, considering the complexity of the problem being modelled, however, they will never give the correct values for every observable in every material. The XC functional used to perform predictive material and chemical

calculations must be carefully chosen to be appropriate based on prior studies and known system applicability.

Hybrid functionals are a class of empirical XC functionals based on the observation that DFT and Hartree-Fock methods tend to get properties wrong in opposite directions. A proportion of Hartree-Fock exact exchange,

$$E_{\mathrm{X}}^{\mathrm{HF}}[\{\psi(\mathbf{r})\}] = -\frac{1}{2}\sum_{i,j}\int\int \psi_i^*(\mathbf{r})\psi_j^*(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_i^*(\mathbf{r}')\psi_j^*(\mathbf{r}')\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}', \tag{1.37}$$

is added to the XC functional. For example, the PBE0 function is written as

$$E_{\mathrm{XC}}^{\mathrm{PBE0}}\left[n\left(\mathbf{r}\right),\{\psi\left(\mathbf{r}\right)\}\right] = \frac{1}{4}E_{\mathrm{X}}^{\mathrm{HF}}\left[\{\psi\left(\mathbf{r}\right)\}\right] + \frac{3}{4}E_{\mathrm{X}}^{\mathrm{PBE}}\left[n\left(\mathbf{r}\right)\right] + E_{\mathrm{C}}^{\mathrm{PBE}}\left[n\left(\mathbf{r}\right)\right], \tag{1.38}$$

in terms of the Hartree-Fock exact exchange and the PBE exchange and correlation. This emperical mixing is arbitrary, and many possible mixing schemes exist, each tuned in some way to cancel out error with respect to a test basket of systems for which experimental or more accurate results exist.

1.2.4 Basis Functions in Kohn-Sham DFT

An important adverse effect of Kohn-Sham DFT is the (re-)introduction of wavefunctions and hence, basis functions are required to be used in calculations. Two of the oldest and most widely established basis sets in DFT are the plane-waves and the various flavours of Gaussian basis sets.

Typically, the basis set employed in a calculation is decided based on the material or molecule under study. If the calculation is of an isolated molecule, then it is likely that a localised basis set comprised of, for instance Gaussian functions would be appropriate. For a bulk or surface calculation, where at least two of the spatial dimensions of the simulation cell are periodic, then a plane-wave basis set is probably a reasonable choice. The first step, irrespective of the choice of basis set is to expand the wavefunctions in terms of the basis functions:

$$\psi_i = \sum_{\alpha} c_\alpha \phi_\alpha, \tag{1.39}$$

where ψ_i are molecular orbitals or electron bands and ϕ_α are basis functions.

1.2.4.1 Plane-waves

If the system under study is inside a unit cell defined by vectors $(\vec{a_1}, \vec{a_2}, \vec{a_3})$, periodicaly repeating in space, then any point in the cell can be mapped to the equivalent point in any of the periodic image cells with a translation vector of the appropriate number of lattice vectors. For instance, if

$$\vec{r} \rightarrow \vec{r} + \vec{R}, \tag{1.40}$$

where $\vec{r}$ represents some point in the cell and $\vec{R}$ represents some linear combination of lattice vectors given by

$$\vec{R} = n_1\vec{a_1} + n_2\vec{a_2} + n_3\vec{a_3}, \tag{1.41}$$

if n_i are integers, then such a system is infinitely periodically repeating in all dimensions in space, but can be represented solely (i.e. all of the physics can be captured) by one cell and a vector $\vec{R}$. What has been described so far in simply periodic boundary conditions, ideally suited to a periodically repeating system. Plane waves share this symmetry, so are well matched in such cases. The plane wave expansion of the cell may be written

$$\psi_n(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} \tilde{\psi}(\vec{G}) e^{i\vec{G}\cdot\vec{r}}, \tag{1.42}$$

where Ω is the volume of the cell, $\vec{G}$ is a wave-vector and quantities over-written with a tilde ($\tilde{}$) have been Fourier transformed. Applying equation (1.40) to plane waves, the plane waves which are considered are only those which share the same periodicity as the lattice, or those for which:

$$e^{i\vec{G}\cdot(\vec{r}+\vec{R})} = e^{i\vec{G}\cdot\vec{r}}. \tag{1.43}$$

Then a grid of wave-vectors which satisfy this criterion can be written in terms of the reciprocal lattice vectors as:

$$\vec{G}_{j_1,j_2,j_3} = (j_1 - \frac{N_1}{2})\vec{b_1} + (j_2 - \frac{N_2}{2})\vec{b_2} + (j_3 - \frac{N_3}{2})\vec{b_3}, \tag{1.44}$$

and where the reciprocal lattice vectors are defined in the usual way as:

$$\vec{b_1} = 2\pi\frac{\vec{a_2}\times\vec{a_3}}{\Omega}, \tag{1.45}$$

$$\vec{b_2} = 2\pi\frac{\vec{a_1}\times\vec{a_3}}{\Omega}, \tag{1.46}$$

$$\vec{b_3} = 2\pi\frac{\vec{a_1}\times\vec{a_2}}{\Omega}. \tag{1.47}$$

Finally, the plane-wave expansion of the bands is

$$\psi_n(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \sum_{j_3=1}^{N_3} \tilde{\psi}_n(\vec{G}_{j_1,j_2,j_3}) e^{i\vec{G}_{j_1,j_2,j_3}\cdot\vec{r}}, \tag{1.48}$$

where the three summations go over the dimensions of the computational grid, so that a denser grid involves more Fourier components and hence can represent frequency components in the bands that are higher (more rapidly oscillating) than a coarser grid.
Due to the Planck relationship between frequency and energy, the density of the grid and maximum frequency component can be written as a single transferrable parameter which defines the accuracy of the basis set. The cut-off energy determines a sphere of

'radius' E_{cut} in reciprocal space and can be increased systematically until convergence with any given physical observable. The inequality for the cut-off energy of the electron bands is given by

$$E_{\text{cut}} > \frac{1}{2}|\vec{G}|^2. \tag{1.49}$$

The electronic density is also frequently represented in terms of plane-waves and contains 8 times more plane-waves as an equivalent density because the density is formed from the wave functions by a band-band inner product. This is given by

$$4E_{\text{cut}} > \frac{1}{2}|\vec{G}|^2. \tag{1.50}$$

Typically, plane-wave basis sets are used for periodic calculations, but instead open boundary conditions can be employed through the use of a cut-off Coulomb operator. [22] This method cuts off all long-range interaction in periodic cells and restricts interactions to those within a pre-defined radius in real space. The method is applied in reciprocal space.

1.2.5 Pseudopotential Approximation

Electrons in the core region tend to be concentrated close to the atomic nucleus, where the nuclear, Coulombic potential is large. This leads to rapidly varying core wave functions and densities, which in turn require many Fourier coefficients to accurately represent in a plane-wave basis. As most chemistry is only concerned with the valence electrons, in practice, relatively few electronic states are modelled in a modified "pseudo-potential", while the core states are frozen and interact in the calculation only as a combined effective potential together with the nucleus.

It is important in practice that the potential in the valence region is accurate. For this reason, a form is chosen which accurately matches the correct potential outside of a given radius around the nucleus r_c (See figure 1.1). Within this radius, the form is chosen to vary more smoothly than a Coulombic potential would and to have a finite limiting behaviour as the radius goes to zero. The larger the cut-off radius (in common vernacular, the larger this is, the "softer" the pseudo-potential is said to be), the fewer Fourier terms are needed, but the less accurate is the approximation.

In the pseudopotential approximation, the core and valance parts of the electronic problem are separated:

$$\hat{H}|\psi_c\rangle = \varepsilon_c|\psi_c\rangle \tag{1.51}$$

$$\hat{H}|\psi_v\rangle = \varepsilon_v|\psi_v\rangle, \tag{1.52}$$

where $|\psi_c\rangle$ are the core states and $|\psi_v\rangle$ are the valance states. A set of smooth valance

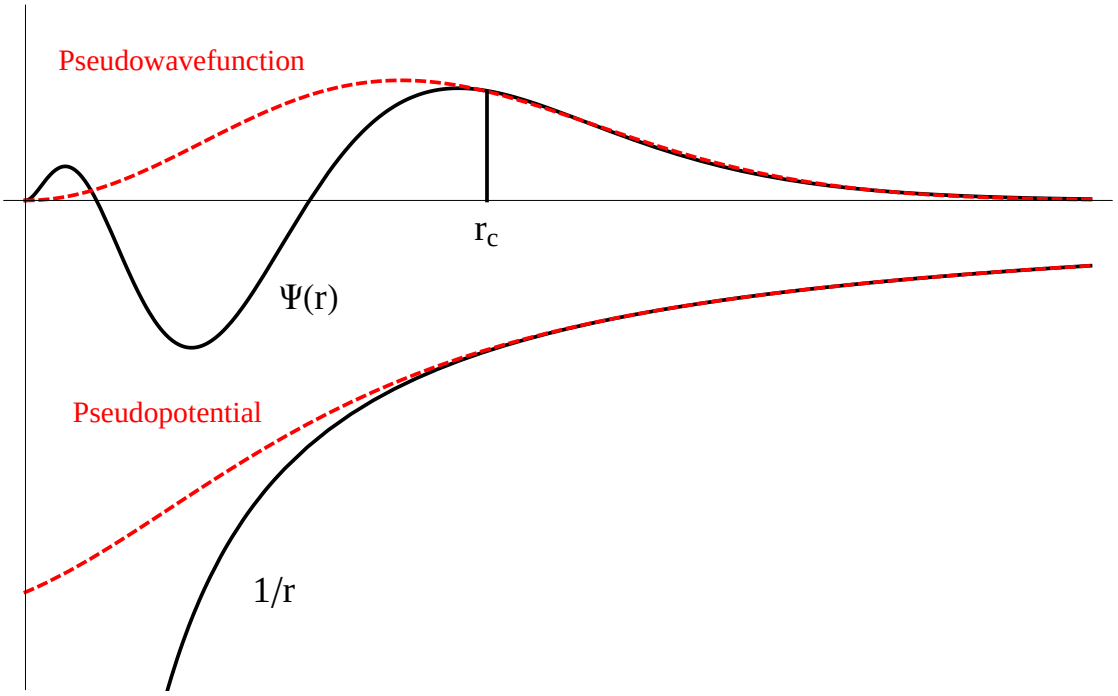


FIGURE 1.1: Pseudopotentials are configured to have a profile which exactly matches the all electron potential outside of a nuclear-centred sphere of radius r_c. Within this sphere, the potential is allowed to differ from the all electron version to reduce the number of plane waves required to represent it accurately. Figure courtesy of Valerio Vitale[2].

states can be constructed from these in the form suggested by Phillips and Kleinman[23], with a lower maximum frequency component required for an accurate plane wave basis representation. This can be expressed as

$$|\hat{\psi}_v\rangle = |\psi_v\rangle + \sum_c B_{cv}|\psi_c\rangle, \tag{1.53}$$

where B_{cv} are orthogonality coefficients given by

$$B_{cv} = \langle\psi_c|\hat{\psi}_v\rangle. \tag{1.54}$$

The pseudowavefunction also obeys a Schrödinger equation,

$$\hat{H}_{\mathrm{PS}}|\hat{\psi}_v\rangle = \varepsilon_v|\hat{\psi}_v\rangle, \tag{1.55}$$

with the Hamiltonian

$$\hat{H}_{\mathrm{PS}} = \hat{H} + \sum_c (\varepsilon_v - \varepsilon_c)\,|\psi_c\rangle\langle\psi_c|. \tag{1.56}$$

The term added onto the Hamiltonian is an energy-dependent repulsive potential, that cancels the nuclear Coulomb potential in the core region. The potential associated with

the new Hamiltonian is then

$$V_{\mathrm{PS}} = V + \sum_{c} (\varepsilon_v - \varepsilon_c) \, |\psi_c(\mathbf{r})\rangle\langle\psi_c(\mathbf{r}')|. \tag{1.57}$$

which is a sum of the unscreened nuclear potential and this non-local repulsive potential. In practice, the non-local part of the pseudopotential is usually given semi-local form, where it can reproduce the scattering behaviour of the nucleus by being formed from several contributions with resolved angular momentum channels.
The way semi-local pseudopotentials are usually set up is due to Kleinman and Bylander[24], which avoids the computational problem with the way semi-local pseudopotentials are constructed naively:

$$V_{\mathrm{PS}} = V_{\mathrm{loc}} + \sum_{l} \sum_{m=-1}^{l} |Y_{lm}\rangle V_l \langle Y_{\mathrm{lm}}|, \tag{1.58}$$

which scales quadratically with the number of basis functions. The Kleinman and Bylander form avoids this problem by constructing a separable set of matrix elements which are a sum of a ratio of two integrals involving only one basis function in each integral. This means that the scaling is linear and so this construction is more efficient in practice. The representation is given by

$$V_{\mathrm{PS}} = V_{\mathrm{loc}} + \sum_{l} \sum_{m=-1}^{l} \frac{|V_l \chi_{lm}\rangle\langle\chi_{\mathrm{lm}} V_l|}{\langle\chi_{lm}|V_l|\chi_{lm}\rangle}, \tag{1.59}$$

where the eigenstates of the pseudo Hamiltonian are given by

$$\chi_{lm}(\mathbf{r}) = R_{(\mathrm{PS})\,l}(\mathbf{r}) Y_{\mathrm{lm}}(\theta, \phi). \tag{1.60}$$

One way to determine the optimum cut-off radius is to compare the integral of the amplitude-squared of the all-electron potential and pseudo-potential and to choose the value which gives the same value for each. Pseudo-potentials of this form are known as norm-conserving, where they have the same charge within the core region, but different shape, but Gauss's law ensures that they are indistinguishable from the all electron equivalent, electrostatically.
Further reduction of the number of plane waves which are necessary for an accurate representation, known as softening in the pseudo-potentials is possible, via either a Vanderbilt ultrasoft or projector augmented wave approach.

1.2.6 The Projector Augmented Wave (PAW) Method

A similar approach to the Vanderbilt ultrasoft pseudopotential method is the Projector Augmented Wave or PAW method. PAW offers the possibility of performing accurate

all-electron calculations in a computationally efficient manner like plane-wave pseudopotential calculations. This is done by defining a linear operator between all electron and pseudo wavefunctions, which maps from one to the other in spherical core regions centred on atoms. In so doing, the pseudo wavefunctions which are defined to vary smoothly on a grid of a given sampling without aliasing can be optimised *in situ* and the all electron wavefunctions can be represented in terms of another basis set and effectively "put back" into the wavefunction when needed to apply operators. This means that for quantities that require accurate wavefunctions and densities close to the nucleii, such as Nuclear Magnetic Resonance (NMR) calculations which are impossible with conventional pseudopotentials, PAW can be used.

$$\left|\psi_n\right\rangle = \left|\tilde{\psi}_n\right\rangle + \left(|\varphi_\nu\rangle - |\tilde{\varphi}_\nu\rangle\right)\left\langle\tilde{p}^\nu\middle|\tilde{\psi}_n\right\rangle, \tag{1.61}$$

where the projectors $\langle\tilde{p}^\nu|$ project onto the pseudo partial waves, $\langle\tilde{\varphi}_\nu|$. This then allows for the construction of the all electron wavefunction $\left|\psi_n\right\rangle$ as in equation 1.61 from the pseudo wavefunction $\left|\tilde{\psi}_n\right\rangle$ by projecting out the pseudo partial waves and replacing these spherical regions with the all electron partial waves, $|\varphi_\nu\rangle$.
In practice, this operation is not performed explicitly, as a basis which could accurately represent the all electron wavefunctions would be required, making PAW redundant if such a basis set were possible. Instead, operator matrix elements and expectation values are calculated by calculating contributions from the pseudo wavefunction and all electron partial wave terms separately and accumulating the result.

1.2.7 Gaussian Basis Functions

Gaussian basis functions were developed in 1950 by S. Boys after noting that Slater type orbitals (STOs), themselves based on eigenfunctions of the Hydrogenic Schrödinger equation could be modified to be more easily integrable.
Hydrogen orbitals are given by:

$$\phi(r,\theta,\phi) = R_{n,l}(r)Y_{l,m}(\theta,\phi), \tag{1.62}$$

where $R_{n,l}$ is the radial and $Y_{l,m}$ the angular contribution to the solutions. $Y_{l,m}$ are spherical harmonics, whereas $R_{n,l}$ are linear combinations of spherical Neumann and Bessel functions. To obtain STOs, this radial contribution is replaced with a smooth node-less exponential decay which matches the hydrogenic radical solutions in the short and long range limits, i.e. it has the exponential tail at infinite separation and the Coulombic cusp singularity at the nucleus.
The problem with STOs is that no analytic form exists for the overlap, so a four-index numerical integral must be performed, which becomes very expensive as the number of

electrons in a calculations increases. To avoid this issue, gaussian orbitals were introduced, which substitute the e^{-r} radial contribution for Gaussians e^{-r^2}. Of course, the integration is now trivial, but the limiting behaviour is lost, because the Gaussian function has zero gradient at the origin and non-exponential decay in the limit $r \to \infty$. This presents a number of problems, but in practice, over decades of numerical calculations, Gaussian functions have proved to be a useful tool, and sometimes a highly accurate basis set in computational Chemistry, when used in an appropriate linear combination.
Compared with plane waves, Gaussian functions have several advantages and several disadvantages. Gaussian orbitals do not possess the systematic improvement in basis set accuracy that plane-waves offer, whereas plane-waves need pseudo-potentials to remove the rapidly varying potential of the core-electrons, whereas Gaussian orbitals can be used adequately for all electrons in an atom. Plane-waves are orthogonal and can represent any potential or density given a high enough sampling (plane-wave cut-off energy), but Gaussians typically need far fewer functions to represent the same system with the same accuracy (provided a suitable arrangement of basis functions had been chosen).
Finally, Gaussians are localised, whereas plane-waves are delocalised. This is neither an advantage nor a disadvantage. Usually, the particular basis set is chosen based on the geometry of the system under study.

1.3 Density Matrix Formalism

Kohn-Sham DFT can also be written in terms of a density kernel, which is a generalisation of occupancy to non-orthogonal functions. Where, for orthogonal systems, the occupancy of the states of a Hamiltonian can be imposed by multiplying the Hamiltonian by a diagonal matrix with diagonal elements of the occupancy of a state, a non-orthogonal system can have state occupancy imposed upon it without diagonalisation by multiplication by a non-diagonal matrix. This matrix is defined to be the matrix with occupancy eigenvalues in the same basis as (same eigenvectors as) the Hamiltonian:

$$\rho(\vec{r}, \vec{r}') = \sum_n f_n \psi_n^*(\vec{r}) \psi_n(\vec{r}), \tag{1.63}$$

where in the case of Kohn-Sham DFT, the single-electron wave functions, ϕ_n are the eigenvectors of the Kohn-Sham Hamiltonian. In the case of systems at zero Kelvin, the density matrix is obviously idempotent, as the occupancy profile is a Heaviside function; one up to the chemical potential and zero at higher energy states. This means that:

$$\rho(\vec{r}, \vec{r}') = \rho^2(\vec{r}, \vec{r}'), \tag{1.64}$$

which has implications for the calculation of density matrix computationally. This property also implies that the trace of the density kernel is equal to the total number of electrons, which is true not only for zero Kelvin systems, but also finite temperature systems.
Two further important properties of the density matrix are (assuming spin degeneracy): firstly that the band-structure total energy of the system may be written in terms of the trace of the product of the Hamiltonian and density kernel as

$$E_{bs} = 2\mathrm{Tr}(\rho H_{KS}), \tag{1.65}$$

and the electronic density of the system may be written in terms of the density matrix as

$$n(\vec{r}) = 2\rho(\vec{r}, \vec{r}). \tag{1.66}$$

1.4 Linear Scaling DFT

In a seminal 1996 paper, Prof. Walter Kohn introduced the concept of "nearsightedness of electron matter" and shows that local electronic properties depend on the external electronic potential only in the local region of space. The effect of changing the electronic potential at some point in space on the electronic properties at some other point decays to zero rapidly as the distance between the two points is increased.

$$\rho(\vec{r}, \vec{r}') \to 0 \text{ as } |\vec{r} - \vec{r}'| \to \infty \tag{1.67}$$

For an insulating material, the decay function with distance is exponential and for a metal it is algebraic. In practice, this means that interactions occurring over (or equivalently elements of the density matrix corresponding to) a distance of greater than some arbitrary cut-off radius can be assumed to be zero and hence not calculated, but that the cut-off distance which will give acceptable accuracy in computed quantities will be further for a metal than an insulator.
This is important for computationally efficient DFT calculations because linear-algebraic operations such as matrix-vector and matrix-matrix products scale as $O(N^2)$ and $O(N^3)$ with matrix dimension (N) for dense matrices (where all of the elements are considered because they may be non-zero), but are$O(N)$ scaling when dealing with sparse matrices (where the number of non-zero elements, $N_{NZE} << N$) and the bandwidth of the matrix remains constant as the dimension is increased. Since the Hamiltonian can be made sparse through the use of a localised basis set, but the density matrix remains dense whatever the basis-set, the nearsightedness of electron matter must necessarily be exploited to perform all linear algebra on sparse matrices and hence achieve linear scaling

with system size and the ability to apply DFT to large systems which would be inhibited by cubic scaling.

Care must be taken when adjusting the cut-off distance, because if it is reduced too far, then the accuracy of the simulation will be impaired, however, while if it is made too large then linear scaling may not be achieved. It is another convergence parameter which calculations must be tested against to ensure decent accuracy.

1.4.1 ONETEP

As localised basis-functions are necessary for linear-scaling calculations, it might be expected that Gaussian orbitals would be the basis-set of choice. It is of course possible to create a linear scaling Gaussian code and this has been achieved [25], however, as with Gaussian basis codes in general, speed comes at the expense of systematic basis set improvement which can be achieved with plane-wave codes. A code with a plane-wave basis set, on the other hand cannot obviously be made linear scaling because basis-functions have no locality in space.

ONETEP is a code which achieves plane-wave systematic improvement of the basis-set with linear scaling. This is possible because of Wannier functions, defined as

$$\Phi(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\mathrm{BZ}} e^{-i\vec{k}\cdot\vec{R}} \psi_{\vec{k}}(\vec{r}), \tag{1.68}$$

where $\psi_{\vec{k}}$ are bands as before and the summation is over a discretised Brillouin zone. The set of Wannier functions are an orthonormal basis-set, but which are not necessarily localised. This comes from the freedom to multiply the bands by any phase factor and not affect the result, afforded by the bands being eigenvectors of the Hamiltonian. For this reason, the phase factor $e^{-i\vec{k}\cdot\vec{R}}$ could become any phase factor and the generalised Wannier functions would still be a valid orthonormal basis for the bands. It has been shown that there is always a choice of phase-factor which leads to localisation.[26]

Expressed slightly differently, this implies that the electron bands of a system which are plane-waves through Bloch's theorem can always be represented in terms of localised functions if the correct matrix of phase factors is known. This is useful in that it shows that a localised basis of rotated (in phase space) plane-waves is always possible, but is only helpful in this form if the optimised eigenvectors of the Hamiltonian are already known. Then the rotation with maximises locality under some norm can be found to give a generalised Wannier function basis.

ONETEP approaches this problem from a different perspective. A set of basis functions are assumed to be local to begin with, so are defined on local lattices centred on each atom. Because these functions must remain local, they are constructed from linear combinations of p-sinc functions (the 3 dimensional generalisation of the cardinal sine

function). This choice of function allows for the direct recovery of a plane-wave basis through a Fourier transform [27], hence reinforcing their equivalency, but also crucially allows for the systematic improvement of the basis-set by increasing the density of the real-space grid of p-sinc functions from which ONETEP's Wannier basis functions are built.

In ONETEP, to minimise to computational expense, the vast majority of the computation is done in the subspace of the Wannier functions, rather than the p-sinc support functions, as this is vastly reduced in dimension. In order to achieve this, the density matrix must be written in a separable form as

$$\rho(\vec{r},\vec{r}') = \sum_{\alpha,\beta}\sum_{\vec{R}} \phi_{\alpha,\vec{R}}(\vec{R}) K^{\alpha,\beta} \phi^{*}_{\beta,\vec{R}}(\vec{r}'), \tag{1.69}$$

where the Wannier functions have become non-orthogonal because of this construction and are known as "Non-orthogonal Generalised Wannier Functions", or NGWFs in ONETEP terminology. These are the basis functions in ONETEP. The non-orthogonality is expressed tensorially, where upper and lower indices have their standard tensor meaning (contra-variant and co-variant quantities, respectively). $K^{\alpha,\beta}$ is known as the density kernel and is a generalisation of state occupancy to non-orthogonal functions.

1.4.2 PAW in ONETEP

As the central idea of PAW is to define a transformation operator between all-electron and pseudised eigenstates, to use it in a density matrix based code such as ONETEP, where eigenstates are not available, the idea must be recast in terms of density matrices[28]. In ONETEP, the PAW density matrix can be considered in terms of the NGWFs as

$$\rho = \tilde{\rho} + \sum_{\nu\mu} \left(|\varphi_\nu\rangle \langle\varphi_\mu| - |\tilde{\varphi}_\nu\rangle \langle\tilde{\varphi}_\mu| \right) \langle \tilde{p}^\nu | \tilde{\rho} | \tilde{p}^\mu \rangle \,, \tag{1.70}$$

where the density matrix in conventional ONETEP becomes $\tilde{\rho}$

$$\tilde{\rho} = |\phi_\alpha\rangle \, \mathbf{K}^{\alpha\beta} \, \langle\phi_\beta| \,. \tag{1.71}$$

As in conventional, norm-conserving ONETEP, the density matrix is never explicitly computed, but its implicit representation through the NGWFs and density kernel in the NGWF basis is. In constrast with norm-conserving ONETEP where $\tilde{\rho}$ is made to conserve the number of electrons through $\mathrm{Tr}[\tilde{\rho}] = N_e$, in PAW mode, this conservation is must be through $\mathrm{Tr}[\rho] = N_e$ and is enforced via the density kernel:

$$\mathbf{K}^{\alpha\beta}\mathbf{S}_{\beta\alpha} = N_e. \tag{1.72}$$

The other important condition on the density matrix is that its eigenvalues equal the state occupancies. In zero Kelvin DFT, this is enforced via the density kernel as an idempotency condition:

$$\mathbf{K}^{\alpha\gamma}\mathbf{S}_{\gamma\delta}\mathbf{K}^{\delta\beta} = \mathbf{K}^{\alpha\beta}, \tag{1.73}$$

where the NGWF overlap matrix has its standard form: $\mathbf{S}_{\gamma\delta} = \langle\phi_\gamma|\phi_\delta\rangle$. So that through idempotency, the occupancy eigenvalues have only zeros and ones. In finite temperature ONETEP calculations, where the state occupancy is fractional, the effective generalised eigenvalue problem to be solved in order to impose the correct occupancy distribution is given by:

$$\mathbf{K}^{\alpha\gamma}\mathbf{S}_{\gamma\delta}\mathbf{M}^{\delta}{}_{i} = \mathbf{M}^{\alpha}{}_{i}\lambda_i, \tag{1.74}$$

where $\{\mathbf{M}\}$ are the eigenvectors of the Hamiltonian matrix and λ_i are the ocupancies of the molecular orbitals. In order to apply either of these options in ONETEP, with PAW, no changes are necessary beyond norm-conserving calculations if the overlap matrix is redefined for PAW to be

$$\mathbf{S}_{\alpha\beta} = \langle\phi_\alpha|\left[\mathbf{1} + |\tilde{p}^\nu\rangle\left(\langle\varphi_\nu|\varphi_\mu\rangle - \langle\tilde{\varphi}_\nu|\tilde{\varphi}_\mu\rangle\right)\langle\tilde{p}^\mu|\right]|\phi_\beta\rangle\,. \tag{1.75}$$

In ONETEP PAW, the quantities close to nucleii which oscillate rapidly and are not representable on the standard ONETEP uniform grids of p-sinc functions, such as the all electron partial waves $|\varphi_\mu\rangle$, are represented on atom centred radial grids, bespoke routines are available for performing integrals of these quantities with low frequency quantities such as NGWFs.
In PAW, the electronic density, which is needed to calculate the density dependent components of the Hamiltonian can be written down in terms of the PAW projectors

$$n(\vec{r}) = \tilde{n}(\vec{r}) + n^1(\vec{r}) - \tilde{n}^1(\vec{r}), \tag{1.76}$$

where $\tilde{n}(\vec{r})$ is just the electronic density from conventional ONETEP,

$$\tilde{n}(\vec{r}) = \phi_\alpha(\vec{r})\mathbf{K}^{\alpha\beta}\phi_\beta(\vec{r}), \tag{1.77}$$

and the core region all-electron and pseudo densities are given as

$$n^1(\vec{r}) = \varphi_\nu(\vec{r})\rho^{\nu\mu}\varphi_\mu(\vec{r}), \tag{1.78}$$

and

$$\tilde{n}^1(\vec{r}) = \tilde{\varphi}_\nu(\vec{r})\rho^{\nu\mu}\tilde{\varphi}_\mu(\vec{r}), \tag{1.79}$$

respectively. The projection of the density matrix within the PAW core spheres is calculated simply as

$$\rho^{\nu\mu} = \langle \tilde{p}^{\nu} | \phi_{\alpha} \rangle \, \mathbf{K}^{\alpha\beta} \, \langle \phi_{\beta} | \tilde{p}^{\mu} \rangle \,, \tag{1.80}$$

since it has only 'on-atom' contributions, it is block diagonal with only ν and μ blocks on the same atom having non-zero contribution.

For most calculations, such as standard singlepoint energies and geometry optimisations, PAW may be used without ever performing overlap integrals of the all electron core regions with quantities on the uniform grid. This is achieved by setting up the so-called compensation density

$$\hat{n}(\mathbf{r}) = \sum_{\mathrm{L,M}} \rho^{\nu\mu} \hat{Q}^{\mathrm{LM}}_{\nu\mu}. \tag{1.81}$$

In the compensation density the difference between the pseudo and all electron densities are fitted with low bandwidth (soft), angular momentum resolved functions, $\hat{Q}^{\mathrm{LM}}_{\nu\mu}$ which are designed to replicate the density difference. Using the compensation density, the total energy can then be written in three terms, without integrals of all electron core regions with uniform grid quantities:

$$E = \tilde{E} + E^1 + \tilde{E}^1, \tag{1.82}$$

where the individual terms are given by

$$\begin{aligned} \tilde{E} = K^{\alpha\beta} \langle \phi_{\beta} | -\frac{1}{2}\nabla^2 | \phi_{\alpha} \rangle \quad &+ \\ E_{\mathrm{XC}}[\tilde{n} + \tilde{n}_c] \quad &+ \quad E_{\mathrm{H}}[\tilde{n} + \hat{n}] + \int V_{\mathrm{H}}[n_{\mathrm{Z_c}}](\tilde{n} + \hat{n}) d\mathbf{r} + E_{\mathrm{II}} \end{aligned} \tag{1.83}$$

$$\begin{aligned} E^1 = \rho^{\nu\mu} \langle \varphi_{\mu} | -\frac{1}{2}\nabla^2 | \varphi_{\nu} \rangle \quad &+ \\ E_{\mathrm{XC}}[n^1 + n_c] \quad &+ \quad E_{\mathrm{H}}[n^1] + \int V_{\mathrm{H}}[n_{\mathrm{Z_c}}] n^1 d\mathbf{r} \end{aligned} \tag{1.84}$$

$$\begin{aligned} \tilde{E}^1 = \rho^{\nu\mu} \langle \tilde{\varphi}_{\mu} | -\frac{1}{2}\nabla^2 | \tilde{\varphi}_{\nu} \rangle \quad &+ \\ E_{\mathrm{XC}}[\tilde{n}^1 + \tilde{n}_c] \quad &+ \quad E_{\mathrm{H}}[\tilde{n}^1 + \hat{n}] + \int V_{\mathrm{H}}[\tilde{n}_{\mathrm{Z_c}}](\tilde{n}^1 + \hat{n}) d\mathbf{r}. \end{aligned} \tag{1.85}$$

1.5 Finding the Ground-State of Insulators

The total energy functional which is minimised in ONETEP with norm-conserving pseudopotentials is given by

$$\begin{aligned} E\left[\{\phi_{\alpha}(\mathbf{r})\}, K^{\alpha\beta}\right] \quad &= E_{\mathrm{K}}\left[\{\phi_{\alpha}(\mathbf{r})\}\right] + E_{\mathrm{NL}}\left[\{\phi_{\alpha}(\mathbf{r})\}\right] + E_{\mathrm{L}}\left[\{\phi_{\alpha}(\mathbf{r})\}, K^{\alpha\beta}\right] \\ &+ E_{\mathrm{H}}\left[\{\phi_{\alpha}(\mathbf{r})\}, K^{\alpha\beta}\right] + E_{\mathrm{XC}}\left[\{\phi_{\alpha}(\mathbf{r})\}, K^{\alpha\beta}\right] + E_{\mathrm{Ewald}}, \end{aligned} \tag{1.86}$$

where the Kinetic energy and non-local pseudopotential energies are "density independent", and depend solely on the NGWFs. The local pseudopotential, Hartree term and exchange-correlation terms depend on the density and hence require both the NGWFs and the density kernel for computation. The Ewald term comes from the Ewald summation[29], which is used to efficiently compute the electrostatic interations between nuclei, with periodic boundary conditions.

This total electronic energy functional is minimised self consistently to the ground state in a two loop approach, where the total energy is minimised with respect to the NGWF p-sinc expansion coefficients in the outer loop and with respect to the density kernel matrix elements in the inner loop. This approach can be summarised as finding the minimum of the Lagrangian

$$E_{\mathrm{GS}} = \min_{\{\phi_\alpha(\mathbf{r})\}} \min_{\{K^{\alpha\beta}\}} \mathcal{L}\left[\{\phi_\alpha(\mathbf{r})\}, K^{\alpha\beta}\right]. \tag{1.87}$$

For insulators, we know that the state occupancy distribution is a step function and so the density kernel is idempotent. The McWeeny purification[30]

$$\mathbf{K} \rightarrow 3\mathbf{KSK} - 2\mathbf{KSKSK} \tag{1.88}$$

takes an almost idempotent density kernel and makes it more idempotent by effectively applying the polynomial $3x^2 - 2x^3$ to the eigenvalues of $\mathbf{K}$. This polynomial is sigmoidal on the interval [0:1] and provided the original eigenvalues are in the interval $[\frac{1-\sqrt{3}}{2}:\frac{1+\sqrt{3}}{2}]$, the eigenvalues of the purified matrix will become closer to 0 for eigenvalues < 0.5 and 1 for eigenvalues > 0.5. This transform can be applied recursively to produce idempotent matrices.

The problem with this approach is that it is non self consistent and non variational. Li, Nunes and Vanderbilt developed a variational approach which is often referred to as just "LNV" [31, 32]. This technique defines a new functional in terms of an auxiliary density kernel, $L^{\alpha\beta}$. The LNV functional, which is minimised subject to the contraint that the number of occupied states is kept constant, is given by

$$E_{\mathrm{LNV}} = \mathrm{tr}\left[(3\mathbf{LSL} - 2\mathbf{LSLSL})\mathbf{H}\right]. \tag{1.89}$$

For LNV optimisation, provided the density kernel $\mathbf{K}$ is kept weakly idempotent, with eigenvalues in the interval [0:1] and the constraint on the number of occupied states is enforced, the derivative of the LNV energy functional with respect to auxiliary density kernel $\mathbf{L}$ points in the direction of lower energy, and hence, also leads to an idempotent density kernel. The density kernel is kept weakly idempotent through the use of an auxiliary density by $\mathbf{K} = 3\mathbf{LSL} - 2\mathbf{LSLSL}$. The constraint on number of occupied

states can be enforced by multiplying E_{LNV} by $N_{\mathrm{occ}}/\mathrm{tr}(3\mathbf{LSL} - 2\mathbf{LSLSL})$, where N_{occ} is the target number of occupied states.

A standard conjugate gradients (CG) non-linear optimisation procedure is used to minimise the LNV functional, making use of the functional derivative of E_{LNV}

$$\frac{\partial E_{\mathrm{LNV}}}{\partial L_{\alpha\beta}} = 6(\mathbf{SLH} + \mathbf{HLS})_{\alpha\beta} - 4(\mathbf{SLSLH} + \mathbf{SLHLS} + \mathbf{HLSLS})_{\alpha\beta}, \tag{1.90}$$

however, since the auxiliary density kernel is contravariant, a contravariant gradient is expected and the one which is used in the CG minimisation procedure, which is given as [33]

$$\frac{\partial E_{\mathrm{LNV}}}{\partial L^{\alpha\beta}} = \left[\mathbf{S}^{-1}\right]^{\alpha\gamma} \frac{\partial E_{\mathrm{LNV}}}{\partial L_{\gamma\delta}} \left[\mathbf{S}^{-1}\right]^{\delta\beta}. \tag{1.91}$$

The CG procedure is performed in ONETEP using a line search

$$\mathbf{L} \rightarrow \mathbf{L} + \gamma \left[\frac{\partial E_{\mathrm{LNV}}}{\partial L^{\alpha\beta}}\right], \tag{1.92}$$

where the parameter α is chosen so as to minimise the energy along the search direction. The LNV procedure may be started from an auxiliary density kernel generated with purification through equation 1.88, starting from a scaled Hamiltonian matrix as the initial density kernel

$$\mathbf{K}_0 = \left[\mathbf{H} - \mu\mathbf{I}\right]/w, \tag{1.93}$$

where w is an estimate for the spectral width of the Hamiltonian matrix, perhaps determined through a norm of the Hamiltonian matrix. This method has been shown to be an efficient way to start LNV [34].

As for the outer loop, where the energy is minimised with respect to the NGWF psinc expansion coefficients, a CG procedure is also used in a similar fashion, but with the appropriate search direction, this time formed from the LNV energy functional by taking the derivative with respect to the NGWFs

$$\frac{\partial E_{\mathrm{LNV}}}{\partial \phi^{\alpha\,\dagger}(\mathbf{r})} = \left[\hat{H}\phi_\beta(\mathbf{r})K^{\beta\alpha} + \phi_\beta(\mathbf{r})(\mathbf{QS})^\beta{}_\alpha\right], \tag{1.94}$$

where the $\mathbf{Q}$ matrix is given by

$$\mathbf{Q} = 3\mathbf{LHL} - 2\mathbf{LSLHL} - 2\mathbf{LHLSL}, \tag{1.95}$$

in the case of LNV. In the case of Ensemble-DFT optimisation, the $\mathbf{Q}$ matrix has a different form which will be discussed in chapter 2, section 2.3.2 (see equation 2.45 for more details).

As with the inner loop gradient, the tensorial characteristics must be taken into account. The NGWFs are covariant vectors, so the contravariant quantity in equation 1.94 must

be multiplied by the overlap matrix **S** to make it covariant

$$\frac{\partial E_{\mathrm{LNV}}}{\partial \phi_\alpha^\dagger(\mathbf{r})} = S_{\alpha\beta}\frac{\partial E_{\mathrm{LNV}}}{\partial \phi^{\beta\,\dagger}(\mathbf{r})}. \tag{1.96}$$

With the exception of the line search in the NGWFs, analagous to the line search in the inner loop,

$$\phi_\alpha(\mathbf{r}) \rightarrow \phi_\alpha(\mathbf{r}) + \lambda\frac{\partial E_{\mathrm{LNV}}}{\partial \phi_\alpha^\dagger(\mathbf{r})}, \tag{1.97}$$

this completes the LNV scheme as performed by ONETEP. The procedure is iteratively repeated with an inner loop within each outer loop iteration, until convergence is reached in both loops.

1.6 An introduction to DFT for Metals

The majority of DFT calculations are performed with the Kohn-Sham approach which describes the energy of the system as a functional of the electronic density $n(\mathbf{r})$ and molecular orbitals $\{\psi_i\}$ The Kohn-Sham equations are non-linear and in practice they need to be solved iteratively until the wavefunctions, occupancies and the density no longer change with respect to each other, which is what is termed a self-consistent solution to these equations. Typically this solution is obtained by a Self-Consistent-Field (SCF) process where $\hat{v}_{\mathrm{KS}}[n]$ is built from the current approximation to the density; then the Kohn-Sham equations are solved to obtain new wavefunctions and occupancies to build a new density; from that new density a new $\hat{v}_{\mathrm{KS}}[n]$ is constructed and these iterations continue until convergence.

This formulation of DFT is suitable for calculations on materials with a band gap (or HOMO-LUMO gap in molecules), and a wide range of algorithms have been developed for the efficient numerical solution of these equations. The absence of a gap at the Fermi level of metallic systems makes the application of DFT approaches for insulators unsuitable for metallic systems. The extension of DFT to finite electronic temperature by Mermin can overcome this limitation by providing a canonical ensemble statistical mechanics treatment of the electrons. In this approach, the existence of a universal functional $F_T[n]$ of the electronic density for the canonical ensemble electronic system at temperature T is shown, and the Helmholtz free energy of the electronic system is written as:

$$A[n] = F_T[n] + \int v_{\mathrm{ext}}(\mathbf{r})\, n(\mathbf{r})\, d\mathbf{r} \tag{1.98}$$

where $v_{\mathrm{ext}}(\mathbf{r})$ is the external potential and $F_T[n]$ contains the kinetic energy, the electron-electron interaction energy and the entropy of the electronic canonical ensemble.

A Kohn-Sham mapping of canonical ensemble DFT to a system of non-interacting electrons can be carried out by analogy with the derivation of standard Kohn-Sham DFT for

zero electronic temperature. In this description, the electronic system is represented by the single particle states (molecular orbitals) $\{\psi_i\}$ which are solutions of a Kohn-Sham eigenvalue equation. The electronic density is constructed from all the single particle states

$$n(\mathbf{r}) = \sum_i f_i \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}). \tag{1.99}$$

where the *fractional* occupancies f_i of the states follow the Fermi-Dirac distribution:

$$f_i^{(FD)}(\epsilon_i) = \left[1 + \exp\left(\frac{\epsilon_i - \mu}{\sigma}\right)\right]^{-1}, \tag{1.100}$$

where μ is the chemical potential and $\sigma = k_B T$, where T is the electronic temperature and k_B is the Boltzmann constant. For this distribution of occupancies, the electronic entropy is given by

$$S(f_i) = -k_B \sum_i f_i \ln(f_i) + (1 - f_i)\ln(1 - f_i)\ . \tag{1.101}$$

As in the zero temperature case, the non-interacting system is constructed to have the same density as that of the interacting system. The Helmholtz free energy on the interacting electronic system is expressed as:

$$\begin{aligned} A[T, \{\varepsilon_i\}, \{\psi_i\}] = \sum_i f_i \langle \psi_i | \hat{T} | \psi_i \rangle + \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \\ + E_H[n] + E_{xc}[n] - TS[\{f_i\}], \end{aligned} \tag{1.102}$$

and consists of the kinetic energy of the non-interacting electrons, and the known expressions for the external potential energy and Hartree energy of the electrons and the unknown exchange-correlation energy expression. Also, the entropic contribution to the electronic free energy $-TS[\{f_i\}]$ is included. In practice, this is a functional not only of the density but also of the molecular orbitals (as they are needed for the calculation of the non-interacting kinetic energy, and the orbital energies, which determine their fractional occupancies.

Another aspect which is particularly relevant for DFT calculations of metallic systems is Brillouin zone sampling. Because of the extremely complicated Fermi surface in some metallic systems, incredibly dense k-point sampling must be done to sample the Brillouin zone adequately, for instance by using a Monkhurst-Pack grid [35], or the VASP tetrahedron method [36]. As the systems become larger, even in metallic systems, the k-point sampling becomes less demanding as the bands flatten.

Solving these canonical ensemble Kohn-Sham equations is not trivial and presents more difficulties than working with the zero temperature Kohn-Sham equations[37]. One has now to determine, in principle, an infinite number of states (instead of just N states in

the zero temperature case) - although we can in practice neglect the states whose energy is higher than a threshold beyond which the occupancies are practically zero. Another complication is the fact that most exchange-correlation functionals that are used in practice have been developed for zero temperature, so their behaviour and accuracy in a finite temperature calculation is not well understood.

Due to operations such as diagonalization, the computational effort to perform DFT calculations, whether on insulators or metals, formally increases with the third power in the number of atoms. This scaling constitutes a bottleneck in efforts to apply DFT calculations to more than a few hundred atoms, as is typically the case in problems involving biomolecules and nanostructures. Walter Kohn showed the path for removing this limitation for insulators with his theory of the "nearsightedness of electronic matter"[38] which states that the 1-particle density matrix (or equivalently the Wannier functions) decay exponentially in a system with a band gap:

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_i f_i \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}') \propto e^{-\gamma|\mathbf{r}-\mathbf{r}'|} \quad . \tag{1.103}$$

Several linear-scaling DFT programs have been developed during the last couple of decades, based on reformulations of DFT in terms of the density matrix or localized (Wannier-like) functions[39][40][41][42][43] and a wide array of techniques have been formulated to allow linear scaling calculations[44][45]. Typically these methods take advantage of the exponential decay to construct highly sparse matrices and use operations such as sparse matrix multiplication and storage where the CPU and memory used scale only linearly with system size.

However, for metallic systems, there is not yet a linear-scaling reformulation. The density matrix for metals is known to have algebraic rather than exponential decay at zero temperature and while exponential decay is recovered at finite temperature, the exponent and the temperature at which it becomes useful for linear-scaling have not been explored adequately for practical use[46].

Due to the importance of calculations on metallic systems, methods have been developed with the aim of performing such calculations in a stable and efficient manner and these methods will be expored further and much more deeply in Chapter 2.

1.7 Catalysis and the Motivation for this Project

Electronic structure theory calculations, using the Density Functional Theory (DFT) approach are widely used to compute and understand the chemical and physical properties of molecules and materials. The study of metallic systems, in particular, is an important area for the employment of DFT simulations as there is a broad range of practical applications. These applications range from the study of bulk metals and surfaces to the

study of metallic nanoparticles, which is a rapidly growing area of research due to its technological relevance [47].

For example, metallic nanoparticles have optical properties that are tunable and entirely different from those of the bulk material as their interaction with light is determined by their quantization of energy levels and surface plasmons. Due to their tunable optical properties, metal nanoparticles have found numerous applications in biodiagnostics [48][49] as sensitive markers, such as for the detection of DNA by Au nanoparticle markers. Another very promising area of metallic nanoparticle usage concerns their magnetic properties which intricately depend not only on their size but also on their geometry[50], and can exhibit effects such as giant magnetoresistance[51][52]. However, by far the domain in which metallic nanoparticles have found most application so far is the area of heterogeneous catalysis. Catalytic cracking of hydrocarbons produces the fuels we use, the vehicles we drive contain catalysts to control the emissions released, and the production of certain foodstuffs also rely on catalytic processes, all of which can use metallic nanoparticles. Catalysis also has a significant role to play in Proton Exchange Membrane (PEM) fuel cells, for example, offering a promising source of clean energy, producing electricity by the electrochemical conversion of hydrogen and oxygen to water. Either monometallic or alloyed, metallic nanoparticles are used as catalysts in these processes, and anchored to a support such as an oxide or carbon. The size, shape and composition (e.g. core-shell, bulk alloy, segregated structure)[53] of these nanoparticles influence their chemical properties, and catalytically important sizes of nanoparticles (diameters of 2-10nm) can consist of hundreds to thousands of atoms.

Even though extended infinite surface slabs of one type of crystal plane are often used as models[54], which correspond to the limit of very large nanoparticles ($\gtrsim$ 5nm) we can see from Figure 4.1 that this limit is not reached before nanoparticles with thousands of atoms are considered. The slab model has been applied successfully in screening metal and metal alloys for a variety of reactions, providing guidance on how to improve current catalysts or identifying novel materials or compositions [55]. However, these types of models, while providing useful insight, do not capture the complexity of the nanoparticle[56], for example, the effect the particle size has on properties or edge effects between different crystal planes. The influence of the support can modify the electronic structure and geometry of the nanoparticle as well as cause "spillover effects" which may, in turn, affect catalytic activity[57]. Metallic nanoparticles are also dynamic, and interaction with adsorbates can induce a change in their shape and composition, which can modify their behaviour. A fundamental understanding of how catalytic reactions occur on the surfaces of these catalysts is crucial for improving their performance, and DFT simulations play a key role [58][59] [60][61][62].

The need to understand and control the rich and unique physical and chemical properties of metallic nanoparticles provides the motivation for the development of suitable DFT

methods for their study. Conventional DFT approaches used to model metallic slabs are unsuited to modelling metallic nanoparticles larger than hundreds of atoms as they are computationally very costly with an increasing number of atoms. Therefore, development of DFT methods for metals with reduced (ideally linear) scaling of computational effort with the number of atoms is essential to modelling nanoparticles of appropriate sizes for the applications mentioned above. Such methods also allow for the introduction of increased complexity into the models, such as the effect of the support or the influence of the environment such as the solvent.

In this project we were, therefore, interested in studying industrially important nanoparticles for catalysis and developing methods which allow for more realistic conditions to be simulated. The outcomes of this work will be discussed in chapters 4 and 5 and development of methods to perform the calculations is discussed in chapters 3 and 6. In the remainder of this section, we will discuss the background of catalysis and of the computational study of catalysis.

1.7.1 Computational Study of Catalysis

The appeal of using computational techniques in catalyst development is that under the related assumptions that such techniques can be made sufficiently accurate and fast enough to be useful, that the throughput of testing proposed catalysts could be greater, because producing the chemicals (arranging the atoms) and setting up the experiments (running the calculations) are much less involved on the computer screen. What's more, in a simulation, properties which may be difficult or impossible to measure experimentally, or at least that may take several different experiments to measure along with a set of other measurements can usually be calculated with built-in functionality or can be implemented reasonably straightforwardly after the derivation of an appropriate equation. Computational catalysis is likely to always be intimately linked with experiment, not least because the value in it is always to test and possibly predict the behaviour of catalysts in real-world applications, but also because the trade-off between expediency and accuracy of present-day techniques necessitates regular sanity checks and comparisons between computed quantities and those obtained through experiment.

The approach to improving catalysts in general is often tied closely with the analysis of trends. If the catalytic behaviour of a set of materials or variants of the same material is plotted against some property thereof, and a relationship between this property and a descriptor of catalytic performance, such as activity is established, then materials which maximise the activity (for instance) may be predicted.

It is in the calculation of properties of predicted catalysts based on experimentally derived

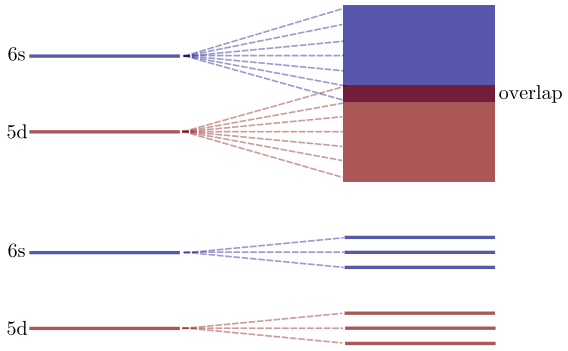


FIGURE 1.2: *Top*: In bulk metallic systems, the atomic energy levels split to form a continuum band structure, where the energy of the lowest conduction band and highest valence band overlap, so that electronic occupancy of the conduction band is energetically favourable. *Bottom*: In a nano-cluster, the energy levels may remain discrete, being too few in number to form a continuum. In a sufficiently diminutive nano- cluster, however, as insufficient splitting of energy levels has occurred to form continuous bands, and the energy level structure may be such that there is still a gap between the highest valence and lowest conduction energy levels, even when formed from atomic species which are bulk-metallic.

trends that computational catalyst study is presently employed. If a proven computational technique can be shown to reproduce the same trend as was measured experimentally, then it is a reasonable assumption that this technique could be used to assess the merit of predicted catalysts within this trend.

1.7.2 Proton Exchange Membrane Fuel Cells (PEMFCs)

As the world continues to develop, it will be necessary to reduce and eventually eliminate its reliance upon the diminishing supply of fossil fuels. In order to achieve this, a reliable and convenient means through which to store and transport energy must be fashioned. If hydrogen is to replace gasoline in this capacity, several challenges must be met, such as: sufficient mass production of hydrogen (the world currently consumes around 1.8×10^{20} J in automotive applications and this has been projected to increase by 38% by 2040 [63]), development of storage systems with enough specific energy and energy density for vehicular application and performance engines capable of replacing those in machines today (with the current best in class generation and drive systems, a specific energy of $\sim$ 10MJ/kg and energy density of $\sim$ 10MJ/L must be reached from current values of $\sim$ 5.7MJ/kg and $\sim$ 2.8MJ/L respectively [64]).

One particularly novel and promising possibility for the energy conversion *from* that in stored hydrogen is to use fuel cells to generate electricity. Proton exchange membrane (PEM) fuel cells can convert molecular hydrogen and oxygen to usable electric current without the use of a combustion-coupled heat engine, thereby allowing for the extraction of energy from the fuel without the limits imposed on efficiency by Carnot's theorem. Such a system was employed on NASA's Apollo program, from the eighth mission onward

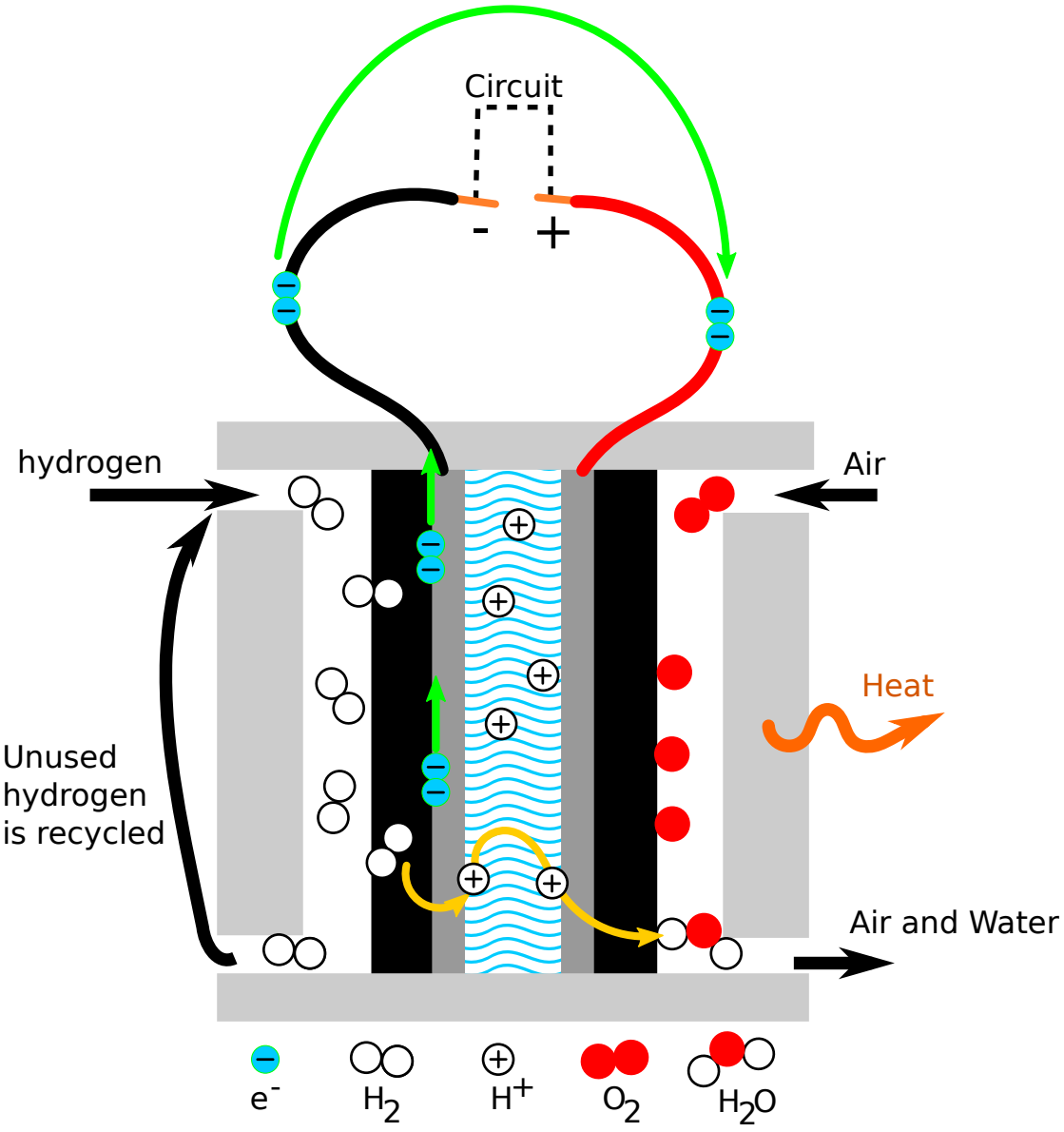


FIGURE 1.3: A hydrogen PEM fuel cell cartoon schematic showing how humidified hydrogen gas and oxygen from air is passed into the inlets on the anode and cathode sides of the fuel cell, respectively. The hydrogen gas flows past the anode catalyst and any which does not react is cycled past again. The hydrogen is catalytically split into hydrogen ions which pass through the proton exchange membrane and electrons which provide a current between the anode and cathode that can do work in an electrical circuit. The extraction of energy is up to 60% efficient [3]. The remaining energy is lost as heat. The fuel cell must be cooled to avoid drying the PEM or boiling the solution. The only product is water.

as it was capable of running on the hydrogen and oxygen fuel used for the rocket motors, while providing electricity in a package far lighter than equivalent battery and solar cell technologies of the day could provide, and having the added advantage that it could supply drinking water for the astronauts on board [65].

At the anode of the fuel cell, hydrogen is split into electrons and protons, while at the cathode, hydrogen ions are reacted with oxygen to form water. Both the anode and the

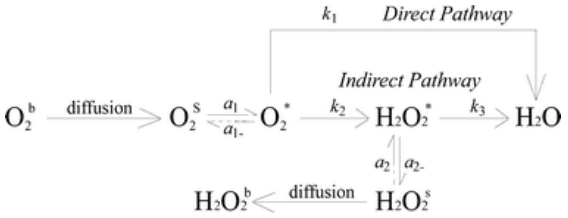


FIGURE 1.4: A reaction diagram showing the progression of the ORR on a hydrogen PEM-fuel cell cathode, where k are rate constants and a, equilibrium probabilities [4].

cathode require catalysts to split the molecular gases into ions which will be reacted in turn.

In a PEM fuel cell, the oxygen reduction reaction (ORR) occurs on the cathode, whereby molecular oxygen is reduced in the presence of protons, to form water. [66]. ORR can occur via a 'four-electron', direct pathway, where molecular oxygen is directly reduced to H_2O, or 'two-electron' indirect pathway where oxygen is serially reduced to H_2O_2, then H_2O, as shown in figure 1.4 [67].

The preferred pathway in ORR has been determined to be the direct pathway in acidic and alkaline electrolytic environments [68]. Although it has been suggested that both pathways may operate in parallel in platinum catalysts [69], and that the dominant reaction pathway may depend upon the surface concentration of oxygen [58], much is still unknown about the precise reaction kinetics.

Electrolytic membranes have been developed with high proton conductivity and hence, in PEM fuel cells, the ORR on the cathode is presently the main bottleneck in their performance; the slow kinetics of the reaction limit the efficiency of energy conversion, with respect to mass in particular. To reach a target energy conversion rate, the surface area of the device must thusly be altered. The most important consequence of this relationship is that with current cathode-catalyst technology, PEM fuel-cells may be too heavy or bulky for mobile applications. Furthermore, as platinum is used most commonly as the catalyst, increasing the mass of the catalyst is also expensive[70].

As improving the Pt mass activity is of such great concern to groups attempting to make PEM fuel cells viable in mobile and automotive applications, several schemes have been employed to improve it, both by increasing surface area and by improving the activity per site [71], such as depositing Pt on a substrate to form a mono-layer [72] forming platinum nano-particles [73] and depositing Pt on nano-particles [74][75].

Specifically, for automotive applications, the platinum-specific power density needs to reach $< 0.2 g_{Pt}/kW$, at cell voltage of $\geq 0.65V$ [71], because a reduction in the over-potential means that the cell is operating closer to the thermodynamically determined reaction potential, and inefficient endothermic intermediate reaction steps have been reduced [58].

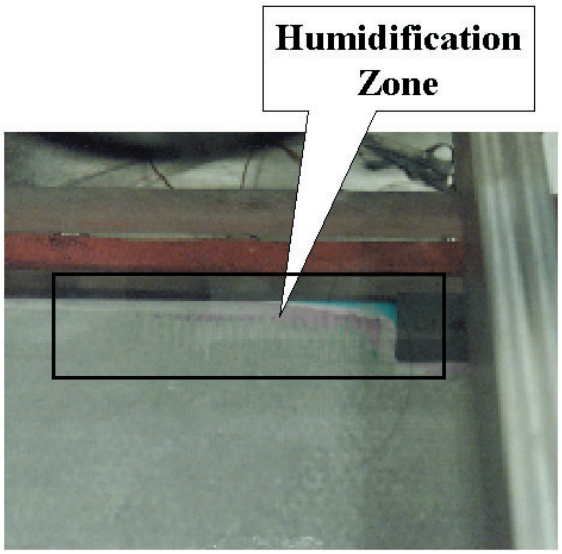


FIGURE 1.5: A photograph showing the depletion of water on a fuel-cell membrane, which will reduce its proton conductivity and the overall efficiency of the cell. Reproduced from Wheeler *et al* [5].

An increase of the Pt-specific power density to the required level represents an approximately five-fold increase in stable catalytic activity from the best in class (in 2009) PEM fuel-cell performance with a pure Pt(111) surface catalyst [76].

The most commonly used electrolyte membranes are perfluorosulphonated acids (PFSA), such as Nafion by Dupont, which must be wet in order to support proton transfer. The requirement of water to function limits the maximum temperature at which present-day fuel cells can operate to under 100 °C. Despite this, alternative membrane materials and modified PFSA membranes with non-water solvents are in development, which would allow for an efficiency gain associated with higher running-temperature, principally due to the faster kinetics on the cathode.

With regard to the membrane electrolyte assembly (MEA), further work must be done on water-management systems if water-solvent based PFSA electrolyte membranes are to be used in high-power, highly efficient PEMFCs. The effect of having an under-performing water-management system can be appreciated in figure 1.5 where the membrane has partially dried out, reducing the efficiency of the cell. While the membrane must be kept moist, flushing with liquid water is not an option, despite usually having a liquid water cooling-system, because efficiency can be reduced if ion-rich electrolyte solvent is flushed. The system, therefore, must balance the need for humidification with a need to minimise excess liquid water, finely. This is achieved in practice with careful temperature management, optimisation of the porosity of the separator between the coolant and electrolyte and careful tuning of the topology of the MEA[5].

Finer grained analyses have been performed on the reaction by several groups with density functional theory calculations and experimentally using the technique of Damjanovic

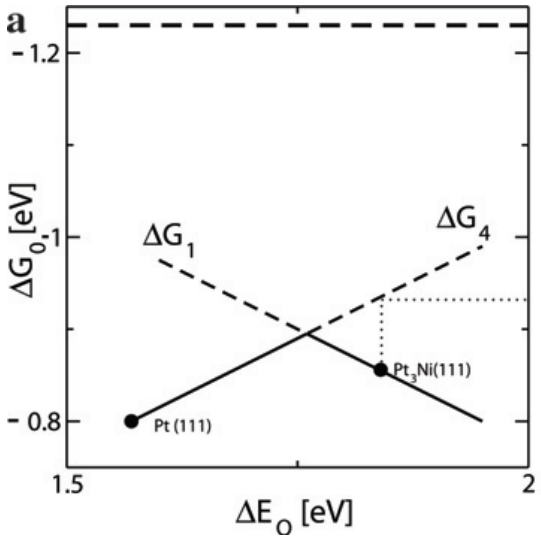


FIGURE 1.6: A 'volcano'-plot, showing how catalytic activity varies (through ΔG_0). The rate-limiting step is either that corresponding to ΔG_1 or ΔG_4. Reproduced from Rossmeisl *et al* [6].

et al [77], showing how the reaction mechanism is split into several intermediate-steps[73], as can be seen in figure 1.7; furthermore, topology can strongly affect the reaction mechanism, for example, it has been shown that the binding strength of intermediate species appears to decrease with decreasing surface lattice constant [78][69].

It has been empirically shown that there is a roughly linear relationship between the adsorption energies of O* and both OH* and HOO* [6]. This same linear relationship was said to be representative of both changes to the catalyst species and of adsorbate coverage allowing both to be investigated in the same analysis.

As a method has been developed for calculating the free-energy of intermediate reaction steps ΔG_i using density functional theory (DFT), by Nørskov *et al* [58], and the reaction rate is directly related to the overall free-energy change in the reaction, then these ΔG_i can be used as a relative measure of the activity, since the slowest intermediate rate corresponds to the upper-bound on the overall reaction rate.

Plotting all ΔG_i corresponding to endothermic intermediate reaction steps, as a function of the adsorption energy of oxygen leads to a graph like the one in figure 1.6, where it can be seen clearly which reaction intermediate step is limiting the rate of the reaction at a given data-point.

In the case of hydrogen bonding to a platinum surface, the binding energy is contributed from a filled, bonding molecular orbital (MO) formed from the hydrogen 1s orbital and the platinum 6s band (the anti-bonding MO is empty), and hybridisation of this σ orbital with the platinum d-band, resulting in a filled, bonding $(d-\sigma)$ orbital and a partially filled antibonding, $(d-\sigma)^*$ orbital. The σ bond varies little across the transition metal elements, however the level of filling of the $(d-\sigma)^*$ MO is determined by the surface electronic states of the metal and hence vary tremendously with altered surface geometry.

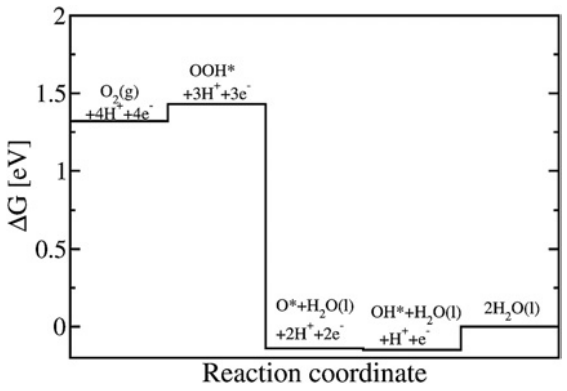
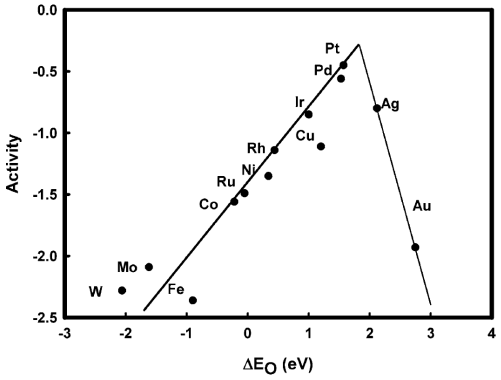
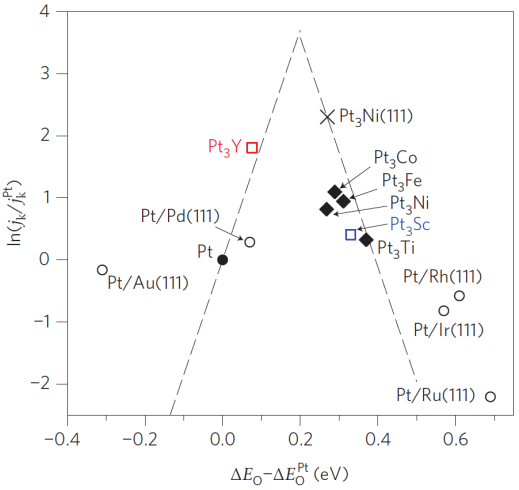


FIGURE 1.7: A reaction-path energy transition diagram, showing the identified intermediate steps in the ORR. Two of the intermediate reaction steps require energy; the first and last. Reproduced from Nørskov *et al* [6].



(A) A 'volcano'-plot, showing how catalytic activity varies across the elemental catalysts for the ORR reaction. Reproduced from Nørskov *et al* [58].



(B) A 'volcano'-plot, showing how catalytic activity varies across Pt(111) mono-layers on various substrate metals, versus pure Pt(111). Reproduced from Greeley *et al* [76].

FIGURE 1.8: Figures showing how activity varies with differing surface composition.

It has been observed that the filling of this anti-bonding MO is highly correlated with the energy of the d-band centre [79]. This observable can then be used experimentally or in simulation to find better catalysts, as filling of the $(d-\sigma)^*$ destabilises the hydrogen-catalyst bond, so that this can be used with the Sabatier principle to find a surface with just the right amout of binding.

The argument for using d-band centres to predict catalytic activity was made for hydrogen, but this was later extended to oxygen [80].

Chapter 2

Density Functional Theory for Metals

A version of this review was published in JCP in 2016:

Jolyon Aarons, Misbah Sarwar, David Thompsett, and Chris-Kriton Skylaris. Perspective: Methods for large-scale density functional calculations on metallic systems. *The Journal of Chemical Physics*, 145(22):220901, dec 2016. ISSN 0021-9606. doi: 10.1063/1.4972007[11].

Abstract

This chapter reviews the theory of DFT calculations on metallic systems with a focus on methods for large-scale calculations, as required for the study of metallic nanoparticles with thousands of atoms. Linear-scaling DFT methods for calculations on insulators, with thousands of atoms are now sufficiently efficient and the computational hardware is adequately powerful that running such large calculations is currently feasible. Such methods are not, however, applicable to metallic systems, where the continuum of states through the chemical potential and the partial occupation of those states must be correctly accounted for. Less mature and advanced, methods for calculations on metallic systems with large numbers of atoms, with correct treatment of the electrons have been developed, however, driven by current research challenges in areas such as energy and bioscience.

DFT calculations on large metallic nanoparticles, together with appropriate human analysis can provide insight from quantum mechanics into poorly understood aspects of problems in technologically important processes such as fuel-cell catalysis and magnetic materials. To perform these large calculations, approaches for electronic energy minimization

which are appropriate for metals need to be used. In this chapter, the (Pulay) density-mixing and EDFT approaches which were mentioned in chapter 1 will be presented in greater detail. Additionally, because diagonalisation of the Hamiltonian matrices of such large systems are either impractical or impossible, methodologies which are required to impose partial state occupancies from smeared occupancy distributions (such as the Fermi-Dirac distribution) without access to the electronic Hamiltonian eigenvalues, will be explored in depth. Particular attention will be paid to the classes of methods known as Fermi Operator Expansions and Integral Expansions.
Great progress has been made in the last decade in advancing the state of the art in partial-occupancy expansion methods towards larger systems and reducing prefactors and scaling. This chapter attempts to put these methods in context, present the methods' benefits as well as their shortcomings with respect to each other and finally to offer predictions for the possible future directions of the field.

2.1 Introduction

This chapter forms an in-depth discussion of methods for large-scale DFT simulations of metallic systems. It will focus on methods for metallic nanoparticle applications, as the motivation of this project has always been on metallic nanoparticles for industrially important catalytic reactions, however, the techniques are applicable widely to other classes of metallic systems.
The majority of this chapter forms a review of "operator expansion" methods. Operator expansion methods are a promising area in recent DFT metals methodology research which can serve to remove the cubic scaling bottleneck in very large metals calculations, by replacing the work of diagonalising the Hamiltonian matrix with a method which has a reduced scaling. We also discuss associated methodology and algorithms which are needed for some of these techniques, such as chemical potential search and matrix inversion, as well as matrix sparsity, although this is discussed in much greater depth in chapter 6. As an aside, there is also a brief section introducing KKR methods, as an alternative to DFT based techniques, which is frequently used by other groups for solid-state metals calculations.
Before introducing operator expansions, this chapter will begin with a general background to the methods for performing sucessful DFT calculations on metals, expanding on the brief introduction in chapter 1.

2.2 DFT for Metals

In DFT for insulators, it is not necessary to introduce the concept of state occupancy, because in ground state systems with a bandgap at the Fermi level, occupancy is determined by the Aufbau principle. That is, the electronic states are filled with spin-pairs of electrons or single electrons in ascending order of energy up to the Fermi energy.
In systems without a bandgap, the situation becomes more complicated. At zero Kelvin, the Fermi-Dirac distribution function is a Heaviside step function centred on the Fermi energy implying integer occupancy, and where the density operator is given by $\hat{\Gamma}_N = |\psi\rangle\langle\psi|$ and energy by $\langle H\rangle = \langle\psi|H|\psi\rangle = \mathrm{tr}(\hat{H}\hat{\Gamma}_N)$. Numerically, this is a problem however, as in the optimisation of the Kohn-Sham wavefunction, states may increase or decrease in energy with respect to each other. If this happens at the Fermi energy, then states will be discontinuously occupied or unoccupied, leading to discontinuous changes in the total energy of the system, where a small change in density leads to a large change in energy and resulting in a scenario in which a numerical optimisation scheme will struggle to proceed in a descent direction.
Knowing that at finite temperature, the Fermi-Dirac function becomes continuous through the chemical potential, a way to achieve convergent behaviour in the optimisation scheme is to increase the electronic temperature, in the way originally derived for Hohenberg-Kohn DFT by Mermin [81]. At finite (non-zero) temperature, pure states become mixed states, where occupancy cannot be represented by a vector, instead a density matrix is required, so that the density operator becomes

$$\hat{\Gamma}^0_N = \sum_i f_i|\psi_i\rangle\langle\psi_i|, \tag{2.1}$$

where f_i is the fraction in each state, $|\psi_i\rangle$.
This will, however, change the total energy of the system in that now the quantity being considered is a total electronic Helmholtz *free* energy($A = E - TS$), given by

$$A[\hat{\Gamma}_N] = \mathrm{tr}\left\{\hat{\Gamma}_N\left(\hat{H} + \frac{1}{\beta}\ln\left(\hat{\Gamma}_N\right)\right)\right\}, \tag{2.2}$$

where β is the thermodynamic quantity, $\beta = 1/k_BT$. The applied temperature should therefore be the minimum required to see convergent behaviour in the optimisation, in the hope that the effect on the total energy of the system will also be minimal and that relative energies will be comparable with the zero Kelvin equivalents.
Another side-effect of having a smoothly varying sigmoid through the chemical potential, rather than a step is that occupancy becomes real-valued on the closed interval $[0, 1]$. This is not a problem quantum mechanically speaking; it is perfectly admissible for states to be partially occupied, provided that each state is normalised to have one electron per

spin channel.

2.2.1 Smearing Functions

It has been addressed that physically, the reason that numerical optimisation of the total energy of metals is unstable at zero Kelvin is that the Fermi-Dirac distribution function

$$f(\epsilon_i, \mu, \beta) = \frac{1}{1 + e^{(\epsilon_i - \mu)\beta}} \tag{2.3}$$

is discontinuous through the Fermi level, and a way to remedy this is to apply a small electronic temperature to make the optimisation possible but not to significantly affect the energies and properties of the system. This temperature can be viewed as a numerical device to improve the convergence of the algorithm.

Therefore, there is no reason that any smoothly varying sigmoidal function should not be used in place of the Fermi-Dirac distribution function (1.100). For metallic systems, this choice of smearing function is a major consideration and there are several options to choose from. The major benefit of Fermi-Dirac occupation is that the electronic smearing corresponds to a physical thermal distribution at temperature T. If thermally distributed electrons are not of interest and the free energies obtained can be "corrected" to approximate zero Kelvin energies, then any other sigmoidal distribution which converges to a step function in some limit might be used. A significant downside to Fermi-Dirac smearing is that the function tails off very slowly, so a large number of very slightly occupied conduction bands must be used to capture all of the occupied states fully.

As a well known example of a non Fermi-Dirac occupation function, the error function as obtained by integrating a Gaussian bell-curve is smooth, sigmoidal and monotonic, although without a closed algebraic form. An advantage of "Gaussian smearing" as it is known is that the curve decays more rapidly to zero and one than the Fermi-Dirac distribution. This is advantageous because it reduces the number of empty states required above the chemical potential as fewer of those high energy states which may have some partial occupancy at some point during the optimisation process with Fermi-Dirac smearing will be partially occupied with Gaussian smearing.

Gaussian smearing solves the issue with the long tails of the distribution neatly. The smearing function is given by

$$f_i^{(G)}(\epsilon_i) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\epsilon_i - \mu}{\sigma}\right)\right], \tag{2.4}$$

where a shifted and scaled error function is used for state occupancy. Using Gaussian smearing, therefore, means that calculations will need relatively fewer partially occupied conduction states to be effective. In this approach, the "smearing width", σ no longer has a physical interpretation and the free energy functional to be minimized becomes an

analogue *generalized* free energy. Despite this, the approach has been used successfully for decades (it is the default smearing scheme in the CASTEP plane-wave DFT code, for instance) and the results can be effectively extrapolated to zero σ [37] by using

$$E_{\sigma=0} = \frac{1}{2}\left(A' + E\right) + O(\sigma^2), \tag{2.5}$$

where A' is the generalized free energy functional. This is a *post hoc* correction and hence is applied at the end of a calculation, while forces and stresses are not variational with respect to this unsmeared energy.

Another approach to recover non-smeared results for metals is to use a smearing function which knocks out the σ dependence of the generalized entropy. First order Methfessel-Paxton Hermite polynomial smearing [82],

$$f_i^{(MP)}(\epsilon_i) = \frac{1}{\sqrt{\pi}}\left(\frac{3}{2} - \left(\frac{\epsilon_i - \mu}{\sigma}\right)^2\right) e^{-\left(\frac{\epsilon_i-\mu}{\sigma}\right)^2}, \tag{2.6}$$

has only a quartic dependence on σ, so that results obtained using this approach need not be extrapolated back to zero σ, but may be used directly. A significant disadvantage of this method is that it yields non-physical negative occupancies. This can lead to difficulties in finding the particle number conserving chemical potential, as the thermal distribution has degeneracies, but may also lead to more serious concerns such as areas of negative electron density.

Marzari-Vanderbilt "cold smearing"[83] solves all of these issues by using a form

$$f_i^{(MV)}(x_i) = \frac{1}{\pi}\left(ax^3 - x^2 - \frac{3}{2}ax + \frac{3}{2}\right) e^{-x^2}, \tag{2.7}$$

where $x_i = (\epsilon_i - \mu)/\sigma$ and a is a free parameter for which the authors suggest a value of -0.5634.

With a non-Fermi-Dirac smearing function, the functional of energy which is minimised becomes a generalised free-energy functional in that the "entropy" contribution is not thermal entropy, but some generalisation of it with absolutely no physical meaning, which must be considered carefully.

This generalised entropy, $S[\{f_n\}]$ can be expressed in general, in terms of the integral of a broadening function (differential of the occupancy function) given by [84]

$$S[\{f_n\}] = \int_{-\infty}^{-\frac{\epsilon-\mu}{\sigma}} -tg(t)dt. \tag{2.8}$$

If the form of the broadening function is chosen to be

$$g(t) = \frac{2}{(e^{t/2} + e^{-t/2})^2}, \tag{2.9}$$

or the differential of the Fermi-Dirac distribution function, then the generalised entropy function recovers the standard Fermi-Dirac entropy,

$$S[\{f_n\}] = -2\sum_n (f_n/2)\log(f_n/2) + (1 - f_n/2)\log(1 - f_n/2). \tag{2.10}$$

Using this formalism allows either grand canonical DFT or any generalised smearing scheme which adheres to a set of conditions, such as that state occupancy lies on the open interval $(0, 1)$ or $(0, 2)$ for spin-polarised systems, and that the occupancy is sampled from a smooth and monotonic function of state energy.
Using a non-step function occupancy distribution will affect the total energy and forces obtained from the system, but with a low (cold) enough (generalised-) temperature, the relative energies and other observables should be very close to the zero-Kelvin limit. The Mermin-Hohenberg-Kohn free energy functional is given as

$$A_V[n] = F_T[n] + \int V_{\text{ext}}(\vec{r})n(\vec{r})\ \vec{dr}, \tag{2.11}$$

which is the equivalent of the Hohenberg-Kohn energy functional in zero temperature DFT. The same treatment may be used as in Kohn-Sham DFT to obtain an energy functional for an equivalent non-interacting quasi-particle system, so as to improve the evaluation of the kinetic energy term, but at the expense of introducing a set of orbitals. More generally, however, any smooth occupancy operator can be used rather than the Fermi-Dirac function, provided it offers an expression for the entropy. So for any distribution of electronic state occupancies, the generalised free energy is given by

$$\begin{aligned} A[\sigma; \{\phi_n\}, \{f_n\}] = \sum_n f_n \langle\phi_n|\hat{T}_c + \hat{V}_{nl}|\phi_n\rangle \quad &+ \quad E_{H,xc}[n] - \sigma S[\{f_n\}] + \mu(N - \sum_n f_n) \\ &+ \quad \sum_n f_n \epsilon_n(\langle\phi_n|\phi_n\rangle - 1), \end{aligned} \tag{2.12}$$

assuming orthogonal basis functions, where μ and $f_n\epsilon_n$ are Lagrange multipliers used to ensure charge conservation and orbital normalisation, respectively, and σ is some scalar representing the width of the distribution given by the derivative of the state occupancy distribution. This would be temperature if f_n were distributed according to the Fermi-Dirac distribution and the expression would recover Mermin DFT.
The energy functional has three arguments; (generalised) temperature, Hamiltonian and wave-functions. Temperature is purely used as a numerical device and is usually fixed for the duration of a calculation, leaving two variable arguments. This means that the functional must be minimised with respect to both, knowing that the problem will be self consistent at the ground state.

2.3 Methods for Solving the Mermin-Kohn-Sham Equations

With a fractional state occupancy, approaches like LNV (Discussed in Chapter 1) will not work, as they rely on the idempotency of the density matrix. With fractional occupancy, the density matrix is sometimes said to be "weakly idempotent"[85, p. 235] and techniques such as LNV or purification will lead to full idempotency and a step function (zero Kelvin) occupancy distribution.

With a step function occupancy distribution, any state crosses the Fermi-level at zero-Kelvin, for instance in electronic optimisation, then it is spontaneously (de-)occupied by the zero-Kelvin step function limit of the Fermi-Dirac distribution function. This is a significant impediment to performing calculations on metals, because the energy derivatives which are used to direct the system to the ground-state are discontinuous and hence useless (i.e. a small change in density will result in a large change in energy). To avoid this, an optimisation technique which is compatible with fractional state occupancy must be used. In this section, two such approaches will be discussed. These techniques (Density Mixing and Ensemble DFT) are mainly used for systems without bandgaps or with narrow bandgaps (where states lie very close to the Fermi-level) but will work also for insulating systems, despite being less efficient.

2.3.1 Density Mixing

One method for minimising the electronic energy in a DFT calculation is known as direct inversion in the iterative subspace (DIIS). First introduced by Peter Pulay for Hartree-Fock calculations [86] [87], this method was later adapted for DFT calculations by Kresse *et al* [88] and applied successfully to systems of up to 1000 metal atoms using the VASP code[89]. A similar technique has been implemented in the linear scaling DFT code CONQUEST[90], and while not linear scaling for metals, the authors show how such techniques may be applied successfully to density matrix based DFT approaches and perform some operations in a linear scaling way.

The central assumption in this method is that a good approximation to the solution can be constructed as a linear combination of the approximate solutions of the previous m iterations.

$$x^{i+1} = \sum_{j=0}^{m-1} \alpha_{i-j}\, x^{i-j}, \tag{2.13}$$

where x can represent any of the variables of the solution, such as the Hamiltonian matrix, the density or the one-particle wavefunctions. The DIIS method constructs a set of linear equations to solve which yield the expansion coefficients α_j.

Kresse *et al* discuss two uses for DIIS, RMM-DIIS, for the iterative diagonalization of a Hamiltonian and Pulay mixing of densities. RMM-DIIS, which is a form of iterative diagonalization allows for the first N eigenpairs of a Hamiltonian matrix to be found without performing a full diagonalization of the whole matrix - detailed information can be found in Kresse and Furthmüller[88]. The method for Pulay mixing of densities is described as follows.

Directly inputting the output density (from equation 1.99) into the next construction of the Hamiltonian can result in an unstable SCF procedure where large changes in the output density result from small changes in the input, known as charge-sloshing. Density mixing attempts to damp oscillations in the SCF procedure, by mixing densities from previous iterations with the density produced from the wavefunctions and occupancies at the current iteration (the output density).

The simplest approach is linear mixing of densities where the new density, $\mathrm{n}^{i+1}(\mathbf{r})$ is constructed from the density of the previous iteration as

$$\mathrm{n}_{\mathrm{in}}^{i+1}(\mathbf{r}) = \alpha\, \mathrm{n}_{\mathrm{out}}^{i}(\mathbf{r}) + (1-\alpha)\, \mathrm{n}_{\mathrm{in}}^{i}(\mathbf{r}), \tag{2.14}$$

where $\mathrm{n}_{\mathrm{in}}^{i}(\mathbf{r})$ and $\mathrm{n}_{\mathrm{out}}^{i}(\mathbf{r})$ are the input and output SCF solutions at the ith iteration. A better option regarding stability and efficiency is to use a mixing scheme based upon a history of previous densities, such as Pulay's DIIS procedure. Under the constraint of electron number conservation $\sum_i \alpha^i = 1$, the next input density is given as

$$\mathrm{n}_{in}^{i+1}(\mathbf{r}) = \sum_{j=0}^{m-1} \alpha_{i-j}\, \mathrm{n}_{in}^{i-j}(\mathbf{r}). \tag{2.15}$$

In DIIS, the mixing coefficients α_i are found by firstly considering the density residual,

$$R[n_{in}^{i}(\mathbf{r})] = n_{out}^{i}(\mathbf{r}) - n_{in}^{i}(\mathbf{r}), \tag{2.16}$$

which can incidentally be used to reformulate linear mixing as

$$n_{in}^{i+1}(\mathbf{r}) = n_{in}^{i}(\mathbf{r}) + \alpha R[n_{in}^{i}(\mathbf{r})]. \tag{2.17}$$

If the DIIS assumption is used that the residuals are linear in $n_{in}(\mathbf{r})$, then

$$R[n_{in}^{i+1}(\mathbf{r})] = \sum_{j=0}^{m-1} \alpha_{i-j}\, R[\mathrm{n}_{in}^{i-j}(\mathbf{r})], \tag{2.18}$$

the Pulay mixing coefficients, α_i which minimize the norm of the residual associated with the current iteration are given as

$$\alpha_i = \frac{\sum_j (\mathbf{A}^{-1})_{ji}}{\sum_{kl} (\mathbf{A}^{-1})_{kl}}, \tag{2.19}$$

where $\mathbf{A}$ is the matrix with elements a_{ij} the dot products of the residuals with index i and j.

To apply such a scheme for metals, a preconditioner is additionally required to ensure the convergence has an acceptable rate. The Kerker preconditioner damps long-range components in density changes more than short-range components[91], because, in metals, long-range changes in the density are often the cause of charge-sloshing effects,

$$G(k) = A\frac{k^2}{k^2 + k_0^2} \tag{2.20}$$

where A is a mixing weight and the parameter k_0 effectively defines a "length-scale" for what is meant by long-range. The k are the wave vectors via which the density is represented. Thus $G(k)$ is used to multiply the n_{in}^{i+1} density in reciprocal space and thus damp the "charge-sloshing" that can occur at long wavelengths. Furthermore, as in all approaches for metals, a smeared occupancy distribution must be used. Head-Gordon *et al* have explored expansions of the various smearing functions and present comparisons of convergence with the order of the expansions and the number of operations involved [92].

2.3.2 EDFT

Density mixing is non-variational in the sense that it can produce "converged" solutions below the minimum, but more than this it can take a long time to reach convergence, or it can even be unstable without a reliable mixing scheme and preconditioner. This is particularly the case in systems with a large number of degrees of freedom. Ideally, a variational approach, where every step is guaranteed to lower the energy towards the ground state energy would be preferred over a density mixing approach, if it could ensure that the ground state energy would always be reached through a stable progression. Marzari *et al* proposed such a scheme[93] in 1997, which in the literature has become familiar as Ensemble DFT (EDFT). EDFT should not be confused with Mermin's original finite temperature DFT formalism that is often referred to with the same name. A variational progression towards the ground state energy is achieved by decoupling the problems of optimising the 1-particle wavefunctions and of optimising the electronic occupancy of these wavefunctions with respect to the energy of the system. This is achieved by performing an occupancy optimization process with fixed wavefunctions at every step in the optimization of the wavefunctions themselves.

The Helmholtz free energy of equation 6.1 is expressed in an equivalent way in terms of the occupancies of the molecular orbitals $A[T;\{f_i\},\{\psi_i\}]$, due to the existence of a one-to-one mapping between the orbital energies and their occupancies . The method generalizes the occupancies to non-diagonal form $\{f'_{i,j}\}$ by working with molecular orbitals $\{\psi'_i\}$ which can be considered to be a unitary transformation of the orbitals in which the occupancies are diagonal. In practice, the following energy expression is used $A[T;\{f'_{i,j}\},\{\psi'_i\}]$. Working with this expression provides a stable direct energy minimization algorithm because the optimization of occupancies is not slowed down by nearly degenerate unitary rotations of the molecular orbitals. The optimization of the energy is done in two nested loops as follows: within the inner loop, occupancy contribution to the energy is minimized and in the outer loop the orbital contribution is minimized using the projected functional

$$A[T,\{\psi'_i\}] = \min_{\{f'_{ij}\}} A[\mathrm{T};\{\psi'_i\};\{f'_{ij}\}], \tag{2.21}$$

which allows an unconstrained optimization of both the orbitals and the occupancies.
To minimise the Mermin-Kohn-Sham equations, Marzari *et al* [93] propose a two-loop "Ensemble-DFT" approach, with wave-function minimisation as the outer loop and Hamiltonian or density matrix minimisation nested within each iteration. This means that upon each iteration of the inner loop, the Hamiltonian matrix corresponding to the minimum functional must be found. To do this, a trial Hamiltonian is constructed as a non-self-consistent guess at the argument of the minimum of the functional, where all density dependent terms are updated (wave-function dependent terms are constant as the wave-functions are constant during the inner loops) and used as a search direction toward the minimum, so that

$$H_{ij}^{n+1} = (1-\lambda)H_{ij}^{n} + \lambda\tilde{H}_{ij}^{n}. \tag{2.22}$$

where H and $\tilde{H}$ are the initial Hamiltonian and the trial Hamiltonian and λ is a scalar mixing parameter on the interval $(0,1)$. Applying a 1D numerical optimiser to λ will give the optimum Hamiltonian for a given set of wave functions along the search direction. In principle, this is a fast process, however in order to produce $\tilde{H}^m$, the Hamiltonian is reconstructed as

$$\tilde{H}^m = \langle\psi|\hat{T} + \hat{V}_{ext} + \hat{V}_H[n^m] + \hat{V}_{XC}[n^m]|\psi,\rangle \tag{2.23}$$

which requires a trial electronic density n^m be constructed from the occupancies and basis functions as

$$n^m(\mathbf{r}) = \sum_n f_i\psi_i(\mathbf{r})\psi_i^*(\mathbf{r}), \tag{2.24}$$

which in turn requires that the trial Hamiltonian be diagonalised to construct the density

matrix. This is the step in the algorithm which will limit the size of the system which can be studied, as diagonalisation is a cubically scaling operation. For an alternative approach which has the potential to reduce the scaling of this step, see chapter 3. The density matrix at each step is given in terms of the smearing function, $f(\epsilon)$ and the eigenvectors of the hamiltonian, Z as:

$$\rho^m = Z^m \mathrm{diag}(f(\epsilon), \mu) Z^{m*}, \tag{2.25}$$

where $f(\epsilon, \mu)$ has the correct number of electrons, determined by the chemical potential. As for the outer loop, this proceeds in an analogous way, defining a linear mixing function

$$|\psi^{l+1}\rangle = |\psi^l\rangle + \nu|\Xi^l\rangle, \tag{2.26}$$

where Ξ^l is a search direction in the wave-function space. Rather than defining it based on a trial wave-function, in the wave-function space, a derivative can be taken which is usually used to form a conjugate-gradients search direction, but in the simplest case can just be the steepest descent direction, with no pre-conditioning:

$$\frac{\partial A'^l}{|\psi^l\rangle} = \hat{H}^l |\psi^l\rangle \rho^l \tag{2.27}$$

Where A' is the functional projected onto the lower dimensional surface of A with constant Hamiltonian.

Despite the clear advantages of the Marzari method, one weakness is that the mapping from the occupancies to the unbounded range of orbital energies is very ill-conditioned, as typically there is a large number of occupancies close to zero that can map on to very different orbital energies. This issue can be allieviated by performing the inner-loop search in terms of the Hamiltonian matrix rather than the density matrix [1]. Alternatively, Freysoldt, *et al* [95] have attempted to address this issue by developing an equivalent scheme where one works directly with the molecular orbital energies instead of the occupancies. In the spirit of the Marzari approach, they employ a non-diagonal representation of orbital energies, which is the Hamiltonian matrix, and minimize the following functional

$$A[\mathrm{T}] = \min_{\{H'_{ij}\}\{\psi'_i\}} A[\mathrm{T}; \{\psi'_i\}; \{H'_{ij}\}]. \tag{2.28}$$

The combined minimization of $\{H'_{ij}\}$ and $\{\psi'_i\}$ is performed using line searches along an augmented search direction in a joint space.

[1] The search has been performed in terms of the Hamiltonian in Castep for many years, but anecdotally the energy surface in the line search is more complicated than the quadratic or cubic often needed to fit the occupancy search, possibly leading to poorer performance[94]

2.3.2.1 EDFT in ONETEP

Recently an EDFT method has been implemented within the ONETEP DFT package[7]. The implementation of EDFT in ONETEP follows the prescription outlined in the previous sections very closely, but differs slightly in order to handle the NGWF basis. ONETEP already optimizes the NGWFs variationally with respect to the energy in a manner similar to the outer loop of the EDFT method, minimising the energy functional along a descent direction in NGWF space projected onto the surface of constant Hamiltonian. The psinc basis set allows a systematic convergence to the complete basis set limit with a single plane-wave cutoff energy parameter much like plane waves do for periodic calculations. Techniques such as PAW, which has gained favour in many codes for the efficient representation of core electrons, may also be applied to the spatially localized NGWFs $\{\phi_\alpha\}$ in EDFT calculations.
The inner loop minimises the energy functional with respect to a descent direction in Hamiltonian space with constant NGWFs, subject to constraints on orthogonality and the number of electrons, which is determined entirely by the chemical potential.
In ONETEP, the density matrix is constructed as an expansion in the NGWFs as

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi^*_\beta(\mathbf{r}'), \tag{2.29}$$

where $K^{\alpha\beta}$ is the generalized occupancy of the NGWFs, i.e. its eigenvalues are $\{f_i\}$ and is known as the density kernel. EDFT is implemented [7] by taking the Hamiltonian eigenvalue approach of Freysoldt *et al*, but also taking advantage of the localized nature of the NGWFs which obviously also requires dealing with non-orthogonality and hence careful consideration of the tensorial nature of every quantity used. As usual, the density can be constructed from the density matrix as

$$n(\mathbf{r}) = \rho(\mathbf{r},\mathbf{r}). \tag{2.30}$$

In ONETEP EDFT, the free energy functional is optimized firstly with respect to the Hamiltonian matrix, $\{H_{\alpha,\beta}\}$ while keeping the NGWFs $\{\phi_\alpha\}$ fixed. The orbital contribution to the free energy can then be minimized by optimising a projected Helmholtz functional,

$$A'[\mathrm{T};\{\phi_\alpha\}] = \min_{\{H_{\alpha\beta}\}} A[\mathrm{T};\{H_{\alpha\beta}\};\{\phi_\alpha\}] \tag{2.31}$$

with respect to $\{\phi_\alpha\}$.
The steepest descent direction in the NGWF space is given by the negative of the gradient,

$$|\tilde{\Gamma}^{(l)}_\alpha = \hat{H}^{(l)}|\phi^{(l)}_\gamma\rangle K^{\gamma\beta} S^{(l)}_{\beta\alpha} - |\phi^{(l)}_\gamma\rangle K^{\gamma\delta} H^{(l)}_{\delta\alpha}, \tag{2.32}$$

which is then modified for occupancy and preconditioning, before having non-local components projected out. This descent direction can be used directly in a steepest descent optimisation, or conjugated in any of the standard ways to perform conjugate gradients optimisation. Either way, a line search is performed along the descent direction as

$$|\phi_\alpha^{(l+1)}\rangle = |\phi_\alpha^{(l)}\rangle + \eta|\chi_\alpha^{(l)}\rangle, \tag{2.33}$$

where $\chi_\alpha^{(l)}$ is the modified, local descent direction and η is a scalar mixing parameter, determined by line-search.

The eigenvectors of the Hamiltonian can then be updated for the change in basis using

$$\tilde{M}^\alpha_{\ i} = S^{\alpha\beta(l+1)}\langle\phi_\beta^{(l+1)}|\phi_\gamma^{(l)}\rangle M^\gamma_{\ i}, \tag{2.34}$$

allowing the density kernel to be reconstructed without diagonalisation as

$$\tilde{K}^{\alpha\beta} = \sum_i^{N_b} \tilde{M}^\alpha_{\ i} f_i \tilde{M}^{\dagger\alpha}_{\ \ i} \tag{2.35}$$

and subsequently the density as

$$n^{l+1}(\vec{r}) = \phi_\alpha^{(l+1)}(\vec{r})\tilde{K}^{\alpha\beta}\phi_\beta^{*(l+1)}(\vec{r}), \tag{2.36}$$

before updating the density dependent terms in the functional (the Hartree and XC terms) and the orbital dependent terms, and rotating the Hamiltonian into the current basis.

The inner loop is performed by diagonalising the Hamiltonian as

$$H^{(m)}_{\alpha\beta} M^{\beta(m)}_{\ \ i} = S_{\alpha\beta} M^{\beta(m)}_{\ \ i} \epsilon_i^{(m)}, \tag{2.37}$$

but in practice, the diagonalization of the Hamiltonian within the inner loop is done by orthogonalising, and then by solving a standard eigenvalue problem using parallel solvers. It can alternatively be done directly with a generalized parallel eigensolver, but to facilitate the expedient replacement of the eigensolver with an expansion method (discussed in the next section) orthogonalization is performed first. The orthogonalization step can be carried out in many ways; using Gram-Schmidt, Cholesky or Löwdin methods, or simply by taking the inverse of the overlap matrix and left-multiplying the Hamiltonian matrix by this to construct a new orthogonal Hamiltonian.

The Löwdin approach requires the overlap matrix to the power 1/2 and -1/2. Due to work of Jansík *et al* [96], it is possible to construct Löwdin factors in a linear-scaling way using a Newton-Shultz approach, without an expensive eigendecomposition.

The Hamiltonian matrix in the basis of Löwdin orthogonalized orbitals can be written as

$$H'_{ij} = (\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2})_{ij} \tag{2.38}$$

and subsequently, the Kohn-Sham Hamiltonian eigenvalues can be found by diagonalising this orthogonal matrix with a standard parallel eigenvalue solver.
The next step in this EDFT inner-loop method is to construct a non-self consistent density matrix from the Hamiltonian matrix. This density matrix can be represented as a function of the eigenvalues ε_i of the Hamiltonian as:

$$K^{\alpha\beta} = \sum_{i}^{N} M^{\alpha}{}_{i} f(\varepsilon_i) M_i^{\dagger\,\beta}, \tag{2.39}$$

where the eigenvalues (band occupancies) of the finite-temperature density kernel, $K^{\alpha\beta}$ are given in terms of a smearing function (1.100): The matrix M contains the eigenvectors of the Hamiltonian eigenproblem:

$$H_{\alpha\beta} M^{\beta}{}_{i} = S_{\alpha\beta} M^{\beta}{}_{i}\, \varepsilon_i. \tag{2.40}$$

Such eigenvalue based approaches will always scale as $O(N^3)$ as they employ matrix diagonalization algorithms. The construction of the Hamiltonian and orthogonalization procedures are however linear-scaling which significantly reduces the prefactor of this approach (See Fig. 2.1) versus a comparable EDFT implementation in conventional plane-wave codes.
the density kernel is constructed from the eigenvectors and the Fermi-Dirac (or alternative smearing function) applied to the eigenvalues as

$$K^{\alpha\beta(m)} = \sum_{i}^{N_b} M^{\alpha(m)}_{i} f_i^{(m)} M_i^{\dagger\beta(m)}, \tag{2.41}$$

where the chemical potential used to construct the occupancy is adjusted to conserve electron-number. Since the NGWFs do not change in the inner loop, only density dependent terms need to be evaluated for the energy functional. The density is evaluated as in the outer loop for new density kernel. The full Hamiltonian is given by

$$\tilde{H}^{(m)}_{\alpha\beta} = \langle \phi_\alpha | \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_H[n^{(m)}] + \hat{V}_{XC}[n^{(m)}] | \phi_\beta \rangle, \tag{2.42}$$

and minimization is performed in the space of Hamiltonians and a search direction is constructed as

$$\Delta^{(n)}_{\alpha\beta} = \tilde{H}^{(n)}_{\alpha\beta} - H^{(n)}_{\alpha\beta}, \tag{2.43}$$

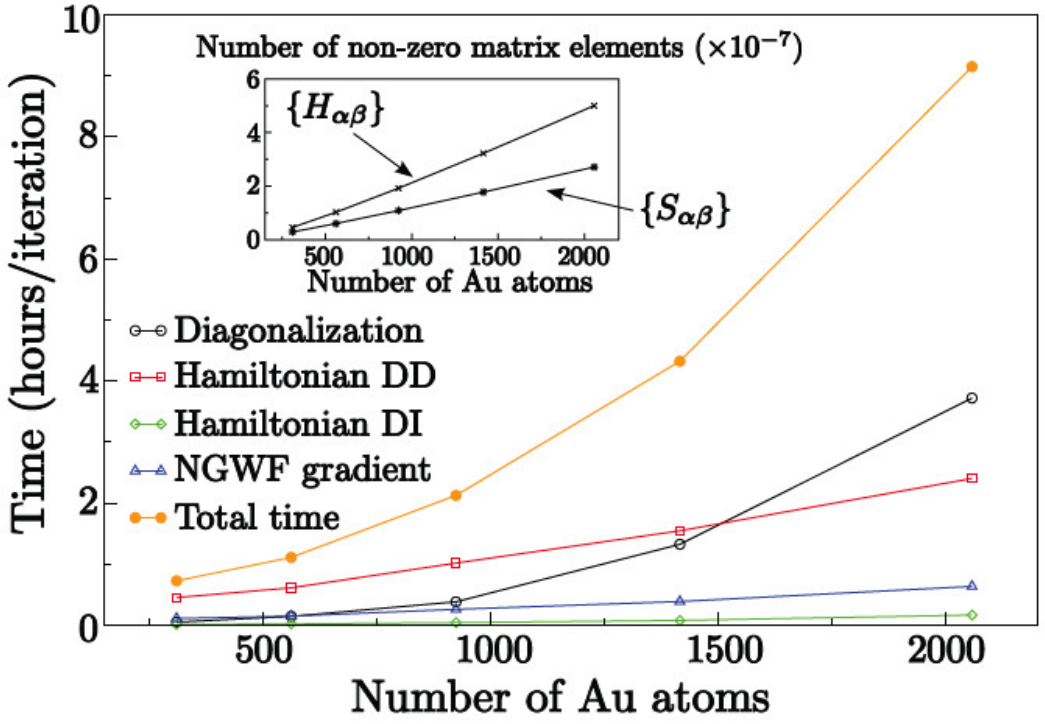


FIGURE 2.1: The scaling of the computational effort with the number of atoms in the EDFT metals method in ONETEP. The computationally demanding steps of the calculation such as the construction of the density dependent (DD) and density independent (DI) parts of the Hamiltonian and the NGWF gradient are linear-scaling and thus allow large numbers of atoms to be treated. However, there is a diagonalization step which is cubic-scaling and eventually dominates the calculation time. Reproduced from [7].

where a new Hamiltonian matrix, $\tilde{H}$ has been constructed from the updated density kernel. The Hamiltonian can then be updated as

$$H_{\alpha\beta}^{(n+1)} = H_{\alpha\beta}^{(n)} - \lambda\Delta_{\alpha\beta}^{(n)}, \tag{2.44}$$

where λ is a scalar line search parameter. Once the inner loop has converged, and the contribution to the energy from the Hamiltonian can no longer be reduced, the outer loop is resumed and this process continues to self-consistency. At self-consistency, i.e. when the calculation has converged, the Hamiltonian will commute with the density kernel. It should be noted that the outer loop proceeds in the same way as described in section 1.5 of chapter 1, except that the **Q** matrix in equation 1.95 is instead given by

$$\mathbf{Q} = -\mathbf{KHS}^{-1}, \tag{2.45}$$

because the derivative with respect to NGWFs needed for the search direction in the outer loop is in this case a derivative of the EDFT total energy functional, rather than the LNV version, so the derivative $\partial E_{\mathrm{LNV}}/\partial\phi_\alpha^\dagger(\mathbf{r})$ in the outer loop line search of equation 1.97 is replaced with the tensorially corrected

$$\frac{\partial E_{\mathrm{EDFT}}}{\partial\phi_\alpha^\dagger(\mathbf{r})} = S_{\alpha\beta}\frac{\partial E_{\mathrm{EDFT}}}{\partial\phi^{\beta\,\dagger}(\mathbf{r})}, \tag{2.46}$$

where $\partial E_{\mathrm{EDFT}}/\partial \phi^{\beta\,\dagger}(\mathbf{r})$ is given by

$$\frac{\partial E_{\mathrm{EDFT}}}{\partial \phi^{\alpha\,\dagger}(\mathbf{r})} = \left[\hat{H}\phi_{\beta}(\mathbf{r})K^{\beta\alpha} + \phi_{\beta}(\mathbf{r})(\mathbf{QS})^{\beta}{}_{\alpha}\right], \tag{2.47}$$

but where the $\mathbf{Q}$ matrix is defined as in equation 2.45 and the density kernel has been calculated using the EDFT inner loop procedure and has fractional occupancy eigenvalues.

2.4 Matrix Inversion

In all of the methods which follow in section 2.5, there is a need for matrix inversion or factorization. In fact, even if one needs an orthogonal representation of the Hamiltonian matrix in general, or for computing derivatives of molecular orbitals for optimizing these orbitals, an inverse or factorization of the overlap matrix is necessary.

The overlap matrix of strictly localized functions is itself a highly sparse matrix, since spatial separation of atoms in systems with many atoms, ensures that most matrix elements will be zero and hence $\mathbf{S}$ is localized. As for the inverse overlap sparsity of insulators, this can be shown to be exponentially localized by treating $\mathbf{S}$ as a Hamiltonian matrix and taking its Green's function at a shift of zero[32]. Since Green's functions are always exponentially localized at shifts outside of the eigenvalue range of the "Hamiltonian" (the overlap matrix is positive definite), the inverse overlap matrix too is exponentially localized. Nunes and Vanderbilt[32] state that the decay length of the exponential localization is dependent upon the ratio of maximum eigenvalue to minimum eigenvalue. So, systems involving overlap matrices with large ℓ^2 condition number will have a long decay length in the inverse overlap matrix.

In many cases it may be desirable to entirely avoid matrix inversion and solve matrix equations directly using a decomposition, but this entirely depends on whether a matrix decomposition can be efficiently computed for a sparse matrix in parallel[97].

If inversion is called for, as in several of the techniques in section 2.5, because the matrices involved are sparse, conventional inversion techniques cannot be employed efficiently, so more specialized techniques are used. The Newton-Shultz-Hotelling (NSH) inversion is a generalization of the application of the Newton-Raphson method for iterative inversion of scalars to matrices. As in the scalar case, the roots of an equation, in this case $\mathrm{f}(\mathbf{X}) = \mathbf{Q} - \mathbf{X}^{-1}$ are found iteratively with the Newton-Raphson approach. This can be performed simply by recursive application of the iteration:

$$\mathbf{X}_{n+1} = \mathbf{X}_n(2\mathbf{I} - \mathbf{QX_n}), \tag{2.48}$$

where $\mathbf{Q}$ is the matrix to be inverted and $\mathbf{X}_n$ converges quadratically to the inverse

matrix in the limit of $n \to \infty$, provided that $\mathbf{Q}$ is non-singular and $\mathbf{X}_0$ is initialized with a matrix which guarantees convergence of the iteration[98],

$$\mathbf{X}_0 = \alpha \mathbf{Q}^T, \tag{2.49}$$

where $\alpha = 1/(||\mathbf{Q}||_1 \, ||\mathbf{Q}||_\infty)$ is a good choice in general, according to Pan and Schreiber[99]. This is discussed in detail in Ozaki[100], along with several other possibilities for matrix inversion in linear-scaling electronic structure theory calculations. This method, however, has been picked up by the community and is used extensively in modern methods[101] and codes[102].

Another approach which has been developed for the inversion of matrices on the poles of a contour integral (see section 2.5.4) but which may be applicable more widely is Selected Inversion (SI). In effect this amounts to a Cholesky expansion (or LDL or LU decomposition) of the matrix in order to compute selected matrix elements, such as the diagonal, of the inverse matrix exactly [103].

Assuming that the matrix to be inverted is sparse then it can be decomposed as

$$\mathbf{L}\mathbf{D}\mathbf{L}^\dagger = \mathbf{P}\mathbf{Q}\mathbf{P}^T, \tag{2.50}$$

with sparse $\mathbf{L}$, which are lower triangular factor matrices and where $\mathbf{D}$ is a diagonal matrix and $\mathbf{P}$ is a permutation matrix chosen to reduce the amount of "fill-in" in the sparsity pattern of $\mathbf{L}$ with respect to $\mathbf{Q}$. If we write $\mathbf{P}\mathbf{Q}\mathbf{P}^T \to \mathbf{Q}$ for simplicity, then the inverse matrix,

$$\mathbf{Q}^{-1} = \mathbf{L}^{-\dagger}(\mathbf{L}\mathbf{D})^{-1}, \tag{2.51}$$

so, provided that the LDL factorization of the sparse $\mathbf{Q}$ matrix can be performed, the potentially sparse (but likely with high fill-in) $\mathbf{Q}^{-1}$ matrix may be calculated trivially by back substitution.

Rather than forming the inverse matrix in this fashion, however, SI may be used to calculated solely the selected elements of the inverse matrix on a pre-defined sparsity pattern. Firstly, the LDL factorization can be computed recursively, using the block factorization:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ 0 & \mathbf{I} \end{pmatrix}. \tag{2.52}$$

In the case that $\mathbf{A}$ is a scalar, a and the matrix to be factorized is block-symmetric, so that $\mathbf{C}$ is a vector $\mathbf{c} = \mathbf{b}^T$, and then

$$\begin{pmatrix} a & \mathbf{b} \\ \mathbf{b}^T & \mathbf{D} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mathbf{l} & \mathbf{I} \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \mathbf{D} - \mathbf{b}^T\mathbf{b}/a \end{pmatrix} \begin{pmatrix} 1 & \mathbf{l}^T \\ 0 & \mathbf{I} \end{pmatrix} \tag{2.53}$$

where the vector $\mathbf{l} = \mathbf{b}/a$. The Schur complement, $\mathbf{S} = \mathbf{D} - \mathbf{b}^T\mathbf{b}/a$ can then be factorized recursively and so the inverse can be written using the symmetric block matrix inverse formula, for scalar a, as

$$\mathbf{Q}^{-1} = \begin{pmatrix} a^{-1} + \mathbf{l}^T\mathbf{S}^{-1}\mathbf{l} & -\mathbf{l}^T\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{l} & \mathbf{S}^{-1} \end{pmatrix}. \tag{2.54}$$

In this way, the inverse can be computed recursively, along with the factorization, by descending through the recursion hierarchy until the Schur complement can be formed and inverted using scalar operations, and then the elements of each successive inverse are calculated from the inverse of the previous level, as the hierarchy is ascended. By computing the inverse recursively in this fashion, the authors[103] show that if the diagonal of the inverse matrix is required, only those elements of the inverse Schur complements which have an index equal to the index of a non-zero element of an $\mathbf{l}$ vector need to be calculated. This corresponds to calculating the inverse matrix for only those elements for which the $\mathbf{L}$ matrix has a non-zero element. The authors report computational scaling of $O(N^{3/2})$ for 1D systems.

Other methods exist for calculating selected elements of the inverse of a sparse matrix with similarly reduced scaling. The FIND algorithm[104], which works by permuting the original matrix to make a desired diagonal element obtainable as the trivial solution to the equation with one unknown after an LDL or LU decomposition.

This is then repeated for every diagonal element, but the computational cost is made manageable by performing partial decompositions and reusing information. Another option is based on Takahashi's equations[105], this has the advantage of not needing to invert the triangular factors.

An important point to note with these techniques for inversion is that a sparse matrix factorization must be used for them to be practical and beneficial. Routines for doing this are available in libraries such as SuperLU [97] and MUMPS [106]. Both of these libraries contain distributed-memory parallel implementations of matrix factorizations for sparse matrices, but it is well known that the parallel scaling of such approaches is somewhat limited [107].

If a general sparse matrix is factorized in such a way, however, then the factors may be dense, so the approach does not exploit the sparsity at all and so is not helpful in

making a fast algorithm. In order to circumvent such a situation, the rows and columns of the matrix to be factorized must be reordered prior to factorization with a fill-in reducing reordering. The ideal reordering in terms of number of non-zero elements in the triangular factor is unknown, as calculating it is an NP-complete problem, however good [108] algorithms exist for finding approximations. Unfortunately, the best of these are not well parallelizable. Despite this, libraries such as ParMETIS[109] and PT_SCOTCH[110] exist and do this operation with the best parallel algorithms currently known. It is perhaps worth noting that the best performing reordering methods differ, depending whether serial or parallel computers are used for the calculation.

2.5 Expansion Approaches

2.5.1 Introduction to expansion approaches

As in linear-scaling methods for insulators, the idea behind expansion and other reduced scaling approaches is to be able to perform an SCF calculation without using an (inherently) cubic-scaling diagonalization step. Thus these methods attempt to compute a converged density matrix for metallic systems directly from the Hamiltonian, using for example (potentially linear-scaling) matrix multiplications, as the Fermi operator expansion approach which is one of the early approaches of this kind. in this way, the cycle of having to diagonalize the Hamiltonian to obtain wavefunctions and energies from which to build the occupancies and the density is no longer needed.

The idea for Fermi operator expansions (FOE) was first proposed by Goedecker & Colombo in 1994 [111]. The density matrix with finite-temperature occupancies, which can be constructed as a function of the eigenvalues of the one-particle Hamiltonian, as in EDFT, is instead built as a matrix-polynomial expansion of the occupancy function. In most works this occupancy function is the matrix analogue of the Fermi-Dirac occupancy function, but can equally be a generalized occupancy function, such as Gaussian smearing, or either Methfessel-Paxton[112] or Marzari-Vanderbilt[93] cold-smearing. For simplicity, this work will only consider Fermi-Dirac style occupancy smearings, however. As an alternative to the eigenvalue based methods mentioned in the previous section, FOE methods begin instead by writing the occupancy formula in matrix form:

$$f(\mathbf{X}) = (\mathbf{I} + e^{\mathbf{X}})^{-1}, \tag{2.55}$$

where I is the identity matrix and

$$\mathbf{X} = (\mathbf{H} - \mu\mathbf{I})\beta. \tag{2.56}$$

where $\mathbf{H}$ is the Hamiltonian matrix expressed in an orthonormal basis, μ is the chemical potential and $\beta = 1/k_BT$. For instance, $\mathbf{H}$ can be obtained using a Löwdin orthogonalization, see equation 2.38, using for instance, the previously mentioned iterative refinement method of Jansík *et al*, or the with the combination of a recursive factorization of the overlap matrix and iterative refinement of the approximate result[113], as proposed by Rubensson *et al* [114]. This method for calculating the Löwdin factor (and inverse Löwdin factor) allows for rapid convergence in a Newton-Schultz style refinement, scheme, while providing heuristics for the requisite parameters, without reference to even extremal eigenvalues. As the title of the work suggests, for sparse matrices, this method can be implemented in a linear scaling way.
In practice, the matrix formula of equation 2.55 cannot be applied directly as the condition number of the matrix to be inverted will be too large in general. Instead, the operation in 2.55 is performed as an approximate matrix-expansion of this function.
In all of the following methods, matrix inversion plays an important role. The following section include discussion of three of the major flavours of expansion approaches; firstly the rational expansions, where the density matrix is made up from a sum involving inverses of functions of the Hamiltonian matrix, secondly the Chebyshev expansions, where the density matrix is formed from as a matrix Chebyshev polynomial approximating the Fermi-Dirac function of the eigenvalues, and thirdly the recursive approaches, where simple polynomial functions are applied recursively to the Hamiltonian matrix to produce the density matrix.

2.5.2 Chebyshev expansion approaches

A way to perform the operation of applying the Fermi-Dirac occupation function to the eigenvalues of a Hamiltonian matrix was proposed by Goedecker and Teter in 1995 as a Chebyshev expansion [115]. In order to do this, the Fermi function is written as a Chebyshev expansion

$$f(\mathbf{X}) = \sum_{i=0}^{N} a_i \mathbf{T}_i(\mathbf{X}), \tag{2.57}$$

where $\{\mathbf{T}_i\}$ are the Chebyshev matrices of the first kind, of degree i and $\{a_i\}$ are the expansion coefficients. The Chebyshev matrices are in the standard form:

$$\begin{aligned} \mathbf{T}_0(\mathbf{X}) &= \mathbf{I} \\ \mathbf{T}_1(\mathbf{X}) &= \mathbf{X} \\ \mathbf{T}_{n+1}(\mathbf{X}) &= 2\mathbf{X}\mathbf{T}_n(\mathbf{X}) - \mathbf{T}_{n-1}(\mathbf{X}). \end{aligned} \tag{2.58}$$

The coefficients can be developed by simply taking the Chebyshev expansion of the

scalar Fermi-Dirac function and applying these to compute a Chebyshev Fermi operator evaluation. As the Chebyshev polynomials can also be defined trigonometrically; $T_n(\cos(\omega)) = \cos(n\omega)$, then the coefficients can be found simply through a Discrete Cosine Transform:

$$\{a_i\} = \mathrm{DCT}\left(\frac{1}{1+e^{\cos(x)}}\right), \tag{2.59}$$

where $x_i = 2(((e_i - \mu)\beta) - e_0)/(e_N - e_0) - 1$ so that the range of equispaced e_i covers at least the range of the eigenrange of the Hamiltonian matrix (e_0:e_N) and the interval is scaled and shifted to cover the useful interpolative range of Chebyshev polynomials $(-1:1)$. This operation need only be performed once, the coefficients are stored for every subsequent evaluation and is effectively negligible in the complexity and timing of the algorithms based on such an expansion. This way of performing the expansion requires approximately N terms for a given accuracy, where N is a function of the smearing width, β, the required accuracy 10^{-D} and the width of the Hamiltonian eigenvalue spectrum ΔE. This implies that if matrix multiplication operations can be at best $O(N)$ scaling for tight-binding calculations such as those for which this technique was first proposed, then of the order of $M \approx D\beta\Delta E$ matrix multiplications would be required [116].

An improvement to the original Chebyshev series of Goedecker was proposed by Liang and Head-Gordon in 2003; this uses a divide and conquer approach to re-sum the terms of a truncated Chebyshev series[92], which is closely related to the divide and conquer approach for standard polynomials suggested in S. Paterson and L. J. Stockmeyer [117]. This approach factors the polynomial into a number of terms which share common subterms. The authors propose three algorithms, the simplest of which is recursive binary subdivision, where the polynomial terms are grouped into even and odd sub-polynomials, for instance

$$\sum_{i=0}^{N} a_i \mathbf{X}^i = \sum_{j=0}^{N_{\mathrm{even}}} a_{2j}\mathbf{X}^{2j} + \sum_{j=0}^{N_{\mathrm{odd}}} a_{2j+1}\mathbf{X}^{2j+1}, \tag{2.60}$$

then a factor of $\mathbf{X}$ can be taken out of the odd sum, to give

$$\sum_{i=0}^{N} a_i \mathbf{X}^i = \sum_{j=0}^{N_{\mathrm{even}}} a_{2j}(\mathbf{X}^2)^j + \mathbf{X}\sum_{j=0}^{N_{\mathrm{odd}}} a_{2j+1}(\mathbf{X}^2)^j, \tag{2.61}$$

If this process is performed recursively, almost a factor of two in matrix multiplications can be saved for each sub-division. If done carefully, the amount of work to produce these terms and combine them to construct the full polynomial is less than the amount of work to construct the polynomial directly, because of the unduplicated effort. To experience the most gain, this approach is repeated recursively, until the sub-division can no longer be performed (efficiently). Using such a divide and conquer scheme reduces the scaling of a Chebyshev decomposition of the Fermi operator to $\mathrm{O}(\sqrt{N})$ number of matrix multiplications.

Rather than decomposing the Fermi operator directly in terms of a truncated matrix-polynomial, Krajewski and Parrinello propose a construction based on an exact decomposition of the grand-canonical potential for a system of non-interacting Fermions[118]:

$$\Omega = -2\mathrm{Tr}\ \ln(\mathbf{I} + e^{-X}). \tag{2.62}$$

By decomposing the quantity in parentheses as

$$\mathbf{I} + e^{-X} = \prod_{l=1}^{P} \mathbf{M}_l \mathbf{M}_l^*, \tag{2.63}$$

where

$$\mathbf{M}_l = \mathbf{I} - e^{i(2l-1)\pi/2P} e^{-\mathbf{H}/2P}, \tag{2.64}$$

where P is an integer which Ceriotti at al. suggest should be in the range 500-1000 for optimal efficiency[119]. Krajewski *et al* show that expectation values of physical observables can be calculated using a Monte-Carlo, stochastic method leading to an $O(N)$ scaling approach, at the expense of noise on values of the calculated properties[120]. In a separate publication, the authors also show that the approach scales linearly without Monte-Carlo sampling in the case of 1D systems such as carbon nanotubes[121].
In later work, Ceriotti, Kühne and Parrinello extend the approach to allow better scaling in general. With this approach the usually expensive and difficult application of the matrix-exponential operator is reduced to a number of more tractable matrix exponential problems, in the sense that approximating them effectively through low-order truncated matrix-polynomial expansions is achievable. Through use of this formalism, the grand-canonical density matrix, otherwise known as the density matrix with Fermi-Dirac occupancy can be expressed exactly in terms of this expansion by taking the appropriate derivative of the grand potential,

$$\frac{\delta\Omega}{\delta\mathbf{H}} = (\mathbf{I} + e^{-X})^{-1} = \frac{2}{P}\sum_{l=1}^{P}[\mathbf{I} - \mathfrak{Re}(\mathbf{M}_l^{-1})]; \tag{2.65}$$

the authors go on to show that the majority of computational effort in such an approach is in the inversion of the $\mathbf{M}_l$ matrices and that the condition number of those matrices with low values of l is significantly larger than for those with higher l. Furthermore, the condition number drops off rapidly with l, approaching 1 after tens of terms. In practice, the authors have shown that to efficiently invert all $\mathbf{M}_l$ matrices in an expansion, the majority of low condition-number matrices can be handled through an expansion about an approximate diagonal matrix, with the significant advantage that these terms can all be formed from the same intermediate matrices, reducing the computational effort markedly. The remaining terms are handled via a Newton-Shultz-Hotelling method,

which is more expensive, but required only for approximately 10 matrices to achieve maximum efficiency.

The approach of Ceriotti *et al*, makes use of the fast-resumming improvements to the Chebyshev expansion for the decomposition in the high-l regime. It also uses initialization from previous matrices in the low-l regime. This work provides a useful framework for analysis of the problems facing such methods, but alone does not bring the scaling factor down from the previous state of the art. In order to achieve that, Richters and Kühne go on to improve the approach by computing only the real components of high condition number $\mathbf{M}_l^{-1}$ and computing the high-l expansion by approximating the inverse of $\mathbf{M}_l$ directly as a Chebyshev expansion[122]. By taking this approach, the scaling of the method is reduced to $O(\sqrt[3]{N})$ matrix multiplications[123].

2.5.3 Recursive methods

Another class of methods which are closely related to purification used in linear-scaling DFT techniques for insulators are the recursive operator expansions. Techniques such as these first appeared in the 1970s with the Haydock approach for tight-binding calculations, which can be performed in a linear scaling manner [124][125].

Niklasson [126] proposed taking a low order Padé approximant of the Fermi operator (2.55),

$$\mathbf{X}_n(\mathbf{X}_{n-1}) = \mathbf{X}_{n-1}^2 \left(\mathbf{X}_{n-1}^2 + (\mathbf{I} - \mathbf{X}_{n-1})^2\right)^{-1} \tag{2.66}$$

and applying it recursively starting with the initial

$$\mathbf{X}_0 = \frac{1}{2}\mathbf{I} - (\mathbf{H} - \mu_0\mathbf{I})\,\beta/2^{2+N}, \tag{2.67}$$

where N is the number of iterates in the expansion. With this method the Fermi operator can be approximated as

$$f(\mathbf{H}, \mu, \beta) = \mathbf{X}_N(\mathbf{X}_{N-1}(\cdots(\mathbf{X}_0)\cdots)). \tag{2.68}$$

Niklasson reports that the scheme is quadratically convergent and in practice the number of iterates can often be kept low ($N \approx 10$). Given that one inversion must be performed, or one linear equation solved per iteration which can be seeded from the the previous iterate, if using an iterative method, this technique is expected to be very quick in practice.

2.5.4 Rational expansion approaches

Goedecker was the first to introduce methods for the rational series expansion of the finite temperature density matrix [127]. The Fermi operator can be expressed as,

$$f(\mathbf{H},\mu,\beta) = \frac{1}{2k_BT}\int_{-\infty}^{\mu}\left(2\mathbf{I} + \frac{1}{2!}\left(\frac{\mathbf{H}-\mu'\mathbf{I}}{k_BT}\right)^2 + \frac{1}{4!}\left(\frac{\mathbf{H}-\mu'\mathbf{I}}{k_BT}\right)^4 + \cdots\right)^{-1} d\mu', \tag{2.69}$$

which still contains a very expensive and ill-conditioned inversion operation. Writing the expression in this form does however allow it to be further approximated by expressing the integrand as a partial fraction expansion truncated to order n:

$$f(\mathbf{H},\mu,\beta) = \int_{-\infty}^{\mu}\sum_{\nu=1}^{n}\frac{C_\nu}{(\mathbf{H}-\mu'\mathbf{I}) - k_BT\,(A_\nu + iB_\nu)}d\mu', \tag{2.70}$$

The coefficients $C_\nu = A_\nu + iB_\nu$, can be calculated, and the partial fraction decomposition is very quickly convergent [128]. The author suggests an n of 16, giving a compact expression:

$$f(\mathbf{H},\mu,\beta) = \sum_{\nu=1}^{n/4}\left[\int_{\Pi_\nu}\frac{2iB_\nu}{\mathbf{H}-z\mathbf{I}}dz + \int_{\Lambda_\nu^+}\frac{A_\nu - iB_\nu}{\mathbf{H}-z\mathbf{I}}dz + \int_{\Lambda_\nu^-}\frac{A_\nu + iB_\nu}{\mathbf{H}-z\mathbf{I}}\,dz\right], \tag{2.71}$$

where the z values are the complex value points along the path used to evaluate each integral by quadrature. Three paths in the complex plane, Π_ν, Λ_ν^+ and Λ_ν^- are used for quadrature. In essence, the approach works by re-expressing the occupancy formula (2.55) as in equation 2.71 which consists of contour integrals. These could be formally evaluated from their residues, but we do not have access to the poles of the function X. Thus, these integrals are computed by numerically integrating around a contour surrounding the eigenspectrum.

At every point z on the quadrature path, a Hermitian matrix, $\mathbf{H} - z\mathbf{I}$ must be inverted. Thus, rational expansion performed as in the contour integral expansion of Goedecker has scaling of O(ln(M)), but results in a large number of matrix multiplications if performed with an iterative inversion algorithm) [128].

In subsequent work, Lin Lin *et al* propose an alternative scheme based on contour integration of the matrix Fermi-Dirac function, so that:

$$f(\mathbf{X}) = \mathfrak{Im}\sum_{l=1}^{P}\frac{\omega_l}{\mathbf{H} - (z_l+\mu)\mathbf{S}}, \tag{2.72}$$

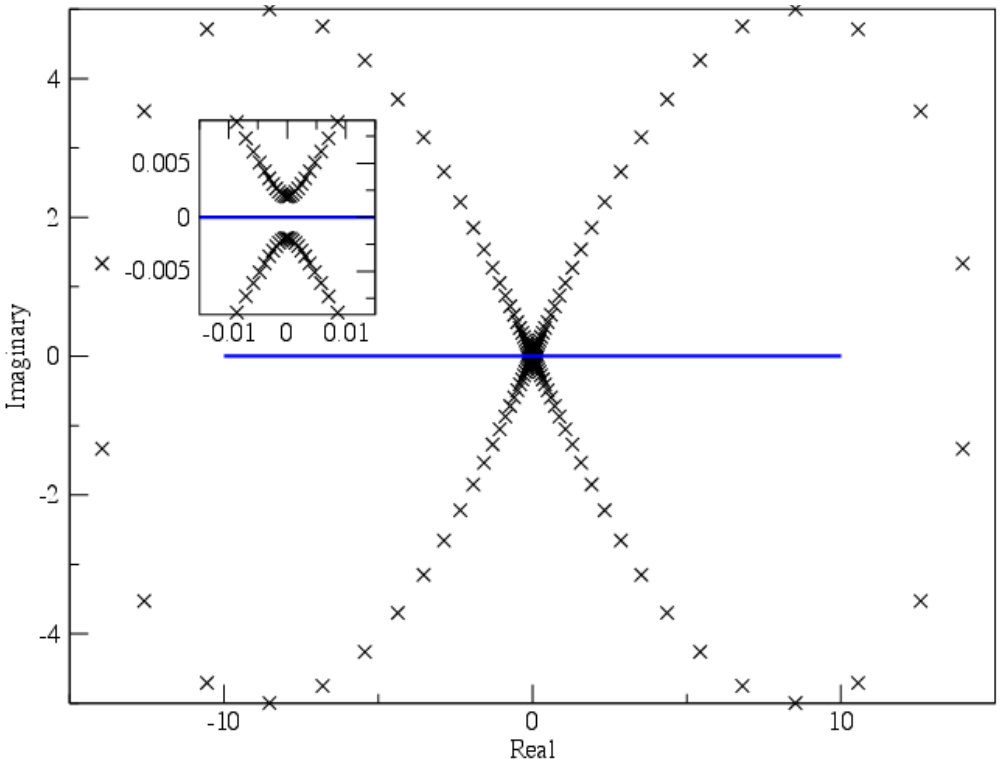


FIGURE 2.2: The typical eigenspectrum of a Hamiltonian matrix for a metallic system is shown on the real axis of this Argand diagram (blue line). The discretized contour integral of a matrix smearing function, as used in the PEXSI method, with poles on the eigenvalues is taken around the black contour, avoiding non-analytic points ($i\pi/\beta$ and $-i\pi/\beta$) on the imaginary axis.

where the complex shifts z_l and quadrature weights ω_l can be calculated from solely the chemical potential, the Hamiltonian eigenspectrum width and the number of points on the contour integral, P. A contour based on the complex shifts suggested by the authors is shown in figure 2.2. Similar techniques to these are used in KKR and related multi-scattering techniques, (see section 2.7).

In performing this integral, the authors are careful to avoid the non-analytic parts of Fermi-Dirac function in the complex plane on the imaginary axis greater than π/β and less than $-\pi/\beta$, by constructing a contour which encompasses the eigenspectrum of the Hamiltonian matrix, but "necks" sufficiently at zero on the real axis to pass though the analytic window while enveloping all of the real eigenvalues. A figure showing such a contour is given in Fig. 2.2.

Within Lin Lin *et al*, the authors show the poor convergence of the Matsubara expansion (expansion into even and odd imaginary frequency components) that Goedecker also reported, though they offer a solution to this in terms of fast multipole (FMP) methods. They do however show that the integral expansion based on a contour integral requires fewer inversions than this Matsubara based approach, even with the FMP method. These methods scale as O(ln(M)) where M is the number of matrix inversions [129].

Lin-Lin & Roberto Car, *et al* have proposed a multipole expansion based on Matsubara theory[130]. Given that

$$(\mathbf{I} + e^{X})^{-1} = (\mathbf{I} - \tanh(X/2))/2, \tag{2.73}$$

then the Matsubara representation can be written, using the pole-expansion of tanh as

$$(\mathbf{I} + e^{X})^{-1} = \mathbf{I} - 4\mathfrak{Re}\sum_{l=1}^{\infty}(\mathbf{X} - (2l-1)\pi i\mathbf{I})^{-1}, \tag{2.74}$$

then combining poles into multi-poles, the overall scaling of the method can be reduced to $O(\ln(N_I)$, where N_I represents the number of (complex-valued) matrices to be inverted. While this method may be considered almost linear scaling, one could argue that the prefactor is large (when using an inversion method such as NSH).

Such an approach necessarily involves a great deal of complex valued matrix inversions (for each point on the quadrature). This would be a significant problem if an iterative inversion method like NSH was employed, as for instance in a given system, if 50 quadrature points were required for accurate results and perhaps 10 iterations of the NSH algorithm were required to converge each inversion then as many as 1000 matrix multiplications would be required to perform this algorithm. It is necessary, therefore that a more appropriate matrix inversion method be used. This issue led to the development of the selective inversion (SI) method by the same authors (see section 2.4). Assuming that the Hamiltonian matrix is sparse and so are the matrices $(\mathbf{H} - (z_l + \mu)\mathbf{S})$ on each of the points on the quadrature z_l, then each of these matrices can be decomposed as

$$\mathbf{LDL}^{\dagger} = \mathbf{H} - (z_l + \mu)\mathbf{S} = \mathbf{Q}_l, \tag{2.75}$$

where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{D}$ is diagonal. and the inverse constructed as in section 2.4.

Together with the contour integral approach, or the pole-expansion (PEX) for calculating the density matrices, the authors have named this approach PEXSI. SI is useful because up until its development all of the contour integral / rational expansions needed a large number of expensive iterative inversions even though the best rational / contour integral expansions scale as $O(\ln(M))$.

It is also important to note that expansion methods become computationally advantageous when the matrices under consideration are sparse. This is clear in the case of insulators due to the short-sightedness of electronic matter, however the exponential decay of the density matrix is also recovered for metals at finite electronic temperature [46].

2.6 Chemical Potential Search

A significant issue when using an expansion approach in the canonical ensemble is to find the correct chemical potential. With standard cubic-scaling plane-wave DFT, this is not considered an issue, principally because the eigenvalues of the Hamiltonian are readily available. For this reason, it is not computationally expensive to perform a search in the eigenvalue space for the chemical potential which gives the correct electron number. This is not possible without a diagonal representation, as for each trial chemical potential, a new density matrix must be calculated, making the whole process many times more expensive.

Several methods have been proposed to improve the situation, including a finite difference representation of the change in number of electrons with respect to chemical potential [131].

Niklasson recommends[132] using the analytic derivative of the density matrix with respect to the chemical potential

$$\frac{\partial \mathbf{K}'}{\partial \mu} = \beta \mathbf{K}'(\mathbf{I} - \mathbf{K}'), \tag{2.76}$$

where $\mathbf{K}'$ is the density matrix represented in an orthogonal basis and $\beta = 1/k_B T$. Then when using a Newton-Raphson optimization process, the electron number can be corrected by altering the chemical potential as:

$$\mu_m = \mu_{m-1} + [N_{\text{occ}} - \text{Tr}(\mathbf{K}')]/\text{Tr}[\beta \mathbf{K}'(\mathbf{I} - \mathbf{K}')] \tag{2.77}$$

at each step in the optimization.

2.7 KKR and related approaches

The Korringa-Kohn-Rostoker method or KKR predates the Hohenberg-Kohn theorem and Kohn-Sham DFT, having been introduced in 1947 by Korringa[133] and 1954 by Kohn and Rostoker[134].

The main principle of the KKR method is that by reproducing the scattering behaviour of electrons and nuclei that the physics of the system will also be reproduced. So, the Lippmann-Schwinger integral, scattering equation is solved rather than the Schrödinger differential equation. The complete and orthogonal set of eigenvectors of the Hamiltonian matrix can be chosen as a basis to expand the Green's function. In extension to this approach, the Green's functions can instead be written in reciprocal space, and then the integrals become contour integrals over the eigenspectrum of the Green's function, where eigenvalues lie on poles[135]. Techniques such as the contour integral approaches

to expanding DFT density matrices have very similar analogues in KKR and were applied in such techniques earlier on[136].

KKR approaches remain important for large scale metallic systems, provided that LDA or LSDA quality results are a valid approximation, as they often are in purely metallic systems. KKR-type methods such as the locally self consistent multiple scattering (LSMS) technique have allowed for linear scaling (order N) calculations to be performed on metallic systems within the L(S)DA formalism since the 1990s[137], with the ability to study alloys coming slightly later [138]. Recent work has seen augmented-KKR approaches which combine the benefits of KKR with DFT calculations [139] and methods for linear scaling tight binding KKR on systems of tens of thousands of atoms[140].

It is also worth noting that KKR-type techniques are often applied in a muffin-tin approach and conventionally, the local scattering environments of the individual atoms are joined together through boundary conditions between the atomic environments. These multiple-scattering techniques are particularly suitable for parallelization, as the environments are mostly self-contained and independent, except at the boundary.

2.8 Conclusions and Outlook

Density Functional Theory calculations on metallic systems (i.e. systems with zero band gap) cannot be done with DFT approaches that have been developed for insulators as such approaches are not designed to cope with a continuum of energy levels at the chemical potential. Mermin's formulation of finite temperature DFT provided the theoretical basis for DFT calculations on metallic systems. A great deal of progress has been made over the last thirty years in studying metallic systems based on Mermin's DFT, starting with approaches such as density mixing and the more stable ensemble DFT. In the last decade rapid progress has been made on expansion methods, which however were introduced much earlier.

Advanced implementations of Mermin's DFT have allowed calculations with over 1000 atoms to be performed. For example Alfè *et al* have performed calculations on molten iron with over 1000 atoms which has provided new unique insights about processes taking place in the Earth's core[89], which can not be obtained by experimental means. Other examples can be found in the work of Nørskov *et al* who have studied small molecule adsorption on platinum nanoparticles of up to 1500 atoms[141] and in Skylaris and Ruiz-Serrano who have reported calculations on gold nanoparticles with up to 2000 atoms[7]. These three examples have used approaches which minimize the number of $O(N^3)$ operations, either the approach of Kresse *et al* as in VASP and GPAW, in the first two examples respectively or the approach of Skylaris *et al* in the latter example. This turns out to be key for calculations on systems up to the low thousands of atoms,

at which point the cubic diagonalization step dominates and a different approach with lower computational scaling is required.

This chapter reviewed the methods for calculations on metallic systems with emphasis on recent developments that promise calculations with larger numbers of atoms. Fractional occupancies of states are needed when dealing with metallic systems and conventionally these are computed after diagonalization of the Hamiltonian which is impractical for large systems. One way to avoid diagonalization is to use expansion methods which construct a finite temperature density matrix via matrix expansion of the Hamiltonian, without needing to access its eigenvalues. Expansion methods have typically been slow and so unsuitable for increasing the size of the system one can study or reducing the calculation time. Recent developments have improved this situation.

There is considerable choice, however in the particular expansion chosen to compute the density matrix. The variants fall roughly into two categories; those based on an expansion via Chebyshev polynomials and those based on a rational or contour integral expansion. At present, on balance, it seems that the *PEXSI* approach which is based on a contour integral combined with a novel matrix inversion approach is the most computationally efficient option. Time will tell whether the Chebyshev expansion methods can be developed further to have improved scaling with system size and become competitive to the best rational expansion methods, or if stochastic methods become usable.

In the future the ultimate target in methods for metallic systems would be to have computational CPU and memory cost which increase linearly with the number of atoms as is the case for insulators where a number of linear-scaling DFT programs currently exist. This is not at all trivial as the temperature and spatial cutoff at which exponential decay of the density matrix in a metal would be sufficient to have enough matrix sparsity for linear-scaling algorithms and acceptable accuracy is not clear and would need to be explored carefully. So far, the EDFT metals method in ONETEP, which works in the localized non-orthogonal generalized Wannier function framework, is an intermediate step towards the development of a linear-scaling method for metals as it benefits from the framework of a linear-scaling DFT approach (e.g. sparse matrices and algorithms) but it still requires a cubically-scaling diagonalisation step as it is based on the EDFT formalism. Two possible extensions of this framework based on expansion methods in ONETEP have been implemented as part of this project and are described in detail in Chapter 6. In particular, we show the the AQuA-FOE approach, which we have developed and describe therein to have a linear scaling computational cost.

Based on this review of the available literature discussing the application of expansion approaches to DFT calculations, the sizes (i.e. number of atoms) of metallic systems which have been studied with these methods is quite limited, in comparison to, for instance, the sizes of insulating systems studied with the same methods. There are several reasons for this. For example, these algorithms have higher prefactors than do

the algorithms involved in calculating idempotent density matrices, as used in linear-scaling DFT.

Production calculations of metallic systems with methods which purport to have a reduced or even linear-scaling computational cost with system size have only very recently begun being reported [142] for the large numbers of atoms that one would have expected, because the crossover point at which these methods become advantageous over diagonalization remains a very large number of atoms even on currently used High Performance Computing facilities.

Also, algorithms such as density mixing do not work as well for large systems, as has been seen in practical observations of the scaling of SCF iterations with the number of atoms in conventional DFT for metallic systems. If this is indeed a factor, then it is possible that an alternative, such as EDFT will be required.

Even with these caveats, the class of operator expansions applied to density matrix DFT methods appears to be the strongest contender for reducing the scaling of accurate DFT calculations on metallic systems in the future. If such methods are further developed, and actually start to be routinely applied as computational power increases, we expect that they will have a great impact on metallic nanostructure and bulk metal surface engineering for industrial applications in optics, magnetics, catalysis and other areas in which the unique properties of metallic systems can be exploited.

Chapter 3

Catalytic Descriptors

The method to generate (local) angular momentum projected density of states and d-band centres from ONETEP was used to compute descriptors in:

L. G. Verga, J. Aarons, M. Sarwar, D. Thompsett, A. E. Russell, and C.-K. Skylaris. Effect of graphene support on large pt nanoparticles. *Phys. Chem. Chem. Phys.*, 18: 32713–32722, 2016. doi: 10.1039/C6CP07334D. URL `http://dx.doi.org/10.1039/C6CP07334D`[12],

in which the contribution of J. Aarons was solely this method.

Abstract

DFT calculations are inherently expensive, even when using advanced techniques to reduce the scaling and overhead. When trying to characterise the performance of large numbers of materials and systems, with high throughput for rapid classification and design of new catalysts, it would be advantageous to use models parametrised from accurate DFT results, but to reduce the number of DFT calculations actually required. These kinds of models which give predictions of the activity of catalytic materials with reduced cost are known as descriptors. Several kinds of descriptors will be discussed in this chapter ranging from those which have already been used and proven by other groups, such as the d-band model and generalised coordination number to new ideas which we have used in this work and have published together with the Nellist group (see Chapter 5).
This chapter will also explore how decriptors can be combined and how existing descriptors such as d-band centres were implemented and used in the ONETEP code as a part of this project.

3.1 The d-band Model and Angular Momentum Projected Density of States (pDOS) in ONETEP

The d-band model, proposed by Nørskov *et al* [143] provides a qualitative interpretation of adsorption energies of ligands on transition metal surfaces, which can in turn, via Sabatier's principle, be used as a predictor of catalytic activity. Chemisorption of ligands takes place via bond formation between the ligand valence orbitals and the surface bands of the metallic surface. The d-band centre model is based on the general observation that the s energy levels of metallic surfaces form a continuum and hence delocalised bands which are wide and do not vary much between different surfaces, resulting in a single narrow band when mixed with an orbital of the ligand. On the other hand the d energy levels are comparatively localised and instead form distinct narrower peaks in the band structure of the surface. Thus the strength of the bond is ultimately determined by the narrower d bands of the transition metal surface and their position with respect to the Fermi level. The higher the d bands are in energy with respect to the Fermi level, the closer in energy they are to the orbitals of the ligand and thus they will mix more strongly with them to produce low energy occupied bonding and empty anti-bonding bands.

Therefore, Nørskov proposes that the "d-band centre", which is the energy weighted average of the d-band can be used as a predictor of ligand binding strength as it is well correlated with the resultant energy of the anti-bonding states. Since the splitting of the adsorbate state into bonding / anti-bonding states came about through interaction with the d-band, the resulting average energy of the anti-bonding states is determined mostly by the distribution of d-band energies in the raw metal. It is observed that as the energy of the d-band centre of a transition metal increases, so does the ligand adsorption affinity.

The density of states (DOS) is a function of the energy ε defined as:

$$\rho(\varepsilon) = \sum_i \langle \psi_i | \psi_i \rangle \delta(\varepsilon - \varepsilon_i), \tag{3.1}$$

where ε_i are the eigenvalues of the Hamiltonian, or band energies and ψ_i its eigenfunctions or bands. The d-band centre with respect to the Fermi level $\varepsilon_d - \varepsilon_f$, is defined as the average energy of the d-electrons

$$\varepsilon_d - \varepsilon_f = \frac{\int \varepsilon \rho_d(\varepsilon) f(\varepsilon)\ d\varepsilon}{\int \rho_d(\varepsilon) f(\varepsilon)\ d\varepsilon} \quad . \tag{3.2}$$

where $\rho_d(\varepsilon)$, is the DOS of the d electrons (either for all atoms or for a subset of the atoms as is the case when we are only interested on the atoms on a surface) and $f(\varepsilon)$ is the occupancy of the bands at energy ε.

In order to compute a projected density of states such as $\rho_d(\varepsilon)$ in ONETEP, the NGWFs need to be projected onto basis functions with angular momentum resolution and for this purpose we use sets of spherical waves with the same centres and localisation spheres as the NGWFs. Furthermore, in addition to projection to angular momentum we would like to partition each eigenvalue into contributions associated with a subset of atoms (local projected DOS), or both (local and angular momentum projected DOS). The electron bands, $\psi_i(\vec{r})$ can be written in terms of the NGWFs, $\phi_\alpha(\vec{r})$ using the eigenvalues of the one-particle Hamiltonian:

$$\psi_i(\vec{r}) = \phi_\alpha(\vec{r}) M^\alpha{}_i, \tag{3.3}$$

where we use the summation convention over repeated Greek letter indices, that correspond to non-orthogonal quantities. In terms of the NGWFs the normalisation condition of the bands becomes:

$$M^\dagger{}_i^\alpha \langle \phi_\alpha | \phi_\beta \rangle M^\beta{}_i = 1. \tag{3.4}$$

Now to give angular momentum and atomic resolution, we can insert a basis of truncated spherical waves,

$$\chi_{A,k_{n,l},l,m}(\vec{r}) = j_l(kr) Y_{l,m}(\Omega) H(a - r), \tag{3.5}$$

where A represents an atom, n,l and m are the quantum numbers, j_l are spherical bessel functions of the first kind, $Y_{l,m}$ are the spherical harmonics and Ω is the solid angle corresponding to $\vec{r}$, H is the Heaviside step function, and finally a the are radii of the spheres within which the basis functions reside. We can choose at this point either to use the full set of spherical waves on each atomic centre within a given psinc kinetic energy cut-off or to contract to a smaller set of functions. We call these spherical wave contracted functions (SWCFs):

$$\chi_{A,l,m_l}(\vec{r}) = \sum_k \omega_k \chi_{A,k_{n,l},l,m_l}(\vec{r}), \tag{3.6}$$

where ω_k are weights which in practice are unity. An identity operator in terms of the SWCFs can be constructed as follows:

$$\hat{1} = |\chi_{A,l,m}\rangle O^{A,l,m|B,l',m'} \langle \chi_{B,l',m'}| \tag{3.7}$$

where an implicit summation is assumed over repeated indices and $O^{A,l,m|B,l',m'}$ is the inverse overlap matrix of the SWCFs. By inserting (3.4) and (3.7) in equation 3.1 we

obtain an decomposition of the DOS in terms of atoms and angular momentum channels:

$$\rho(\varepsilon) = \sum_i M_i^{\dagger\alpha} \langle \phi_\alpha | \phi_\beta \rangle M^\beta{}_i \, \delta(\varepsilon - \varepsilon_i) \tag{3.8}$$

$$= \sum_i M_i^{\dagger\alpha} \langle \phi_\alpha | \chi_{A,l,m} \rangle O^{A,l,m|B,l',m'} \langle \chi_{B,l',m'} | \phi_\beta \rangle M^\beta{}_i \, \delta(\varepsilon - \varepsilon_i) \tag{3.9}$$

$$= \sum_i R_i^{B,l',m'} T_{B,l',m',i} \, \delta(\varepsilon - \varepsilon_i) \tag{3.10}$$

where the quantities $R_i^{B,l',m'}$ and $T_{B,l',m',i}$ are defined from the above equation. We can then define a set of weights as the Hadamard product of R and the transpose of T, so that finally the full DOS becomes

$$\rho(\varepsilon) = \sum_{B,l',m',i} W_{(B,l',m'),i} \, \delta(\varepsilon - \varepsilon_i) \tag{3.11}$$

The above expression allows us full flexibility to process the DOS according to our requirements by using the matrix of weights $W_{(B,l',m'),i}$ allows us to construct various types of DOS. For example, in order to construct the DOS for the d electrons we can add up all the weights for $l' = 2$ while if we want to restrict this DOS to a subset of atoms of our system (such as only the atoms on a surface $\mathcal{S}$) we can further restrict the summation to only these atoms:

$$\rho_{d,\mathcal{S}}(\varepsilon) = \sum_{B \in \mathcal{S},m',i} W_{(B,2,m'),i} \, \delta(\varepsilon - \varepsilon_i) \ . \tag{3.12}$$

This expression was derived for the case of norm-conserving pseudopotentials where ψ_i are the pseudo wavefunctions. In the case of PAW calculations [144], equation 3.1 becomes

$$\rho(\epsilon) = \sum_i \langle \tilde{\psi}_i | \hat{\tau}^\dagger \, \hat{\tau} | \tilde{\psi}_i \rangle \delta(\varepsilon - \varepsilon_i), \tag{3.13}$$

where $\tilde{\psi}_i$ are the smooth valence wavefunctions and $\hat{\tau}$ is the PAW operator that transforms these to the valence wavefunctions with all-electron shapes in the core. This results in replacing the integrals $\langle \chi_{B,l',m'} | \phi_\beta \rangle$ with $\langle \chi_{B,l',m'} | \hat{S} | \phi_\beta \rangle$ where $\hat{S}$ is the overlap operator (unity operator) in the PAW formalism [144]. These integrals are evaluated using the existing PAW machinery of ONETEP.

Finally, the number of electrons that occupy a particular DOS can be obtained as the area under the DOS multiplied by the occupancy, so for the example of equation 3.12 we would obtain the number of electrons as follows:

$$N_e(d, \mathcal{S}) = \int f(\varepsilon) \rho_{d,\mathcal{S}}(\varepsilon) \, d\varepsilon \tag{3.14}$$

provided of course that the weights have been normalised appropriately, so that

$$N_e = \sum_{A,l,m,i} W_{(A,l,m),i} f(\varepsilon_i) \tag{3.15}$$

is the total number of electrons.

3.1.1 Validation tests

Using this approach on various systems shows the correctness of the implementation and the effectiveness of the method. Firstly, on a single atom system, O, the pDOS partitions the two peaks into an s and p contribution as would be expected and sums together to give the total DOS. Moving on to larger systems, a Pt_{13} cuboctahedral nanoparticle

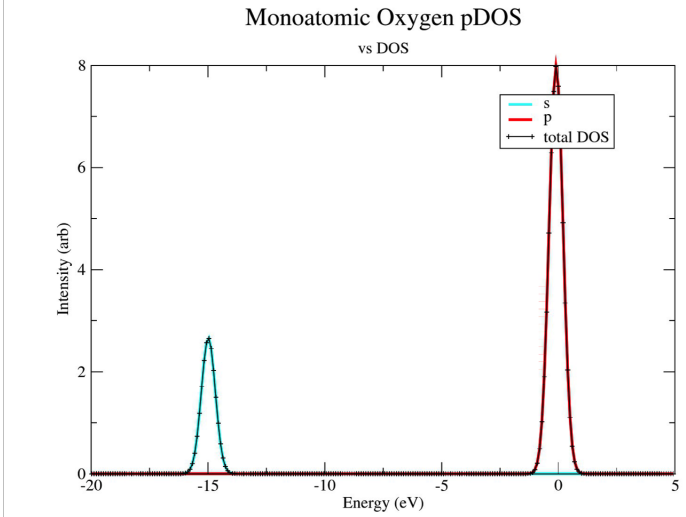


FIGURE 3.1: The angular momentum projected DOS for a single oxygen atom showing that the normalisation is correct and the partitioning is sensible.

which will be discussed in depth in the subsequent sections of this chapter was used as a test calculation. A single-point energy calculation was performed in both Castep and ONETEP with simulation parameters set to be as similar as possible. A plane-wave cut-off energy of 850eV was used in a 50Å cell with a Rappe GGA pseudo potential (see next section) and an rPBE XC functional. Both calculations were set up to output pDOS weights in OptaDOS format and then the OptaDOS program [145] was used on both sets of outputs with a 0.4eV adaptive Gaussian smearing. Beneath the chemical potential, the profiles are almost identical (see figure ??). Above the chemical potential, differences are expected. This is discussed in L.E. Ratcliff *et al* [146] and the techniques discussed therein could be used if accurate conduction states were necessary. In this work, however, they are not as in the integration of the d-band for the calculation of d-band centres, the

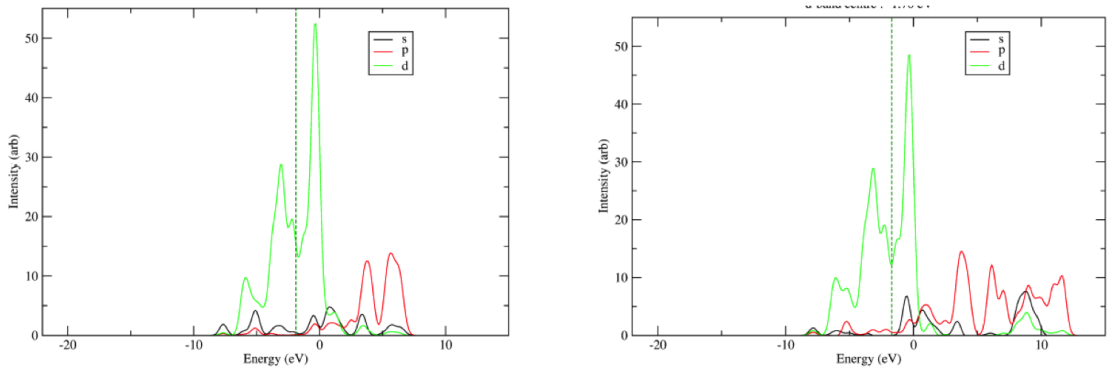


FIGURE 3.2: The angular momentum projected DOS for a Pt_{13} Cubooctahedral nanoparticle, as calculated with Castep and ONETEP showing good agreement up to the chemical potential.

states are weighted by the occupancy and hence, very little contribution comes from the conduction states.

3.2 Generalised Coordination Number

To describe generalised coordination number, we must first revisit the standard coordination number, which is defined for an atom with index i in the total set of atoms as the count of nearest neighbour atoms. As an example which will be important in the rest of this thesis, atoms of bulk FCC platinum have 12 nearest neighbour platinum atoms, so are hence 12 coordinated. Taking these 12 atoms together with the central reference atom out of the bulk and into isolation forms the first cuboctahedral nanoparticle, Pt_{13} which we will study in greater detail in chapter 4. This nanoparticle has a central atom which still has a coordination number of 12, but now that it has been cleaved from the bulk FCC platinum, the other 12 atoms are not 12 coordinated. Instead these "vertex" atoms have coordination number of 5.

If instead of only cleaving one atom plus its 12 nearest neighbours from the platinum FCC bulk, we chose to cleave it, its 12 nearest neighbours and *their* nearest neighbours, without any multiple counting, then we would have a cuboctahedron with 55 atoms. This cuboctahedron is comprised of a core atom, a shell of 12 atoms and a further shell of another 42 atoms. Now if we consider the coordination number of atoms in this nanoparticle, we find that the 13 "core" atoms are 12 coordinated, the "vertex" atoms still have coordination number of 5, but we now have atoms with intermediate coordination. This Pt_{55} nanoparticle has an "edge" atom between the vertex sites which has coordination of 7. There is also an atom at the centre of the square facets of this nanoparticle, which were cleaved along the {100} planes of the FCC lattice, with coordination number of 8.

We can keep doing this to get larger and larger cuboctahedral nanoparticles from bulk FCC platinum (or any FCC metal), and we will keep adding integers between 5 and 12, however the number of distinct coordination numbers saturates at Pt_{147} (the next cuboctahedral "magic number" after 55). To get finer grained coordination detail, we turn to the *generalised coordination number* of Calle-Vallejo *et al*, which takes into account the coordination of an atom's nearest neighbours' coordination as well as its own. This is important because if we want to segregate atoms into groups which may experience similar chemical environments then we can start by segregating them into groups by their local *geometric* environment. Having finer detail than just the integers between 1 and 12 inclusive, is crucial when we have nanoparticles with hundreds or even thousands of surface atoms as we do for real experimental nanoparticles involved in heterogeneous catalysis (see chapter 5).

The generalised coordination number is defined as

$$\bar{\mathrm{CN}}(i) = \sum_{j=1}^{n_i} \mathrm{cn}(j)/\mathrm{cn}_{\max}, \tag{3.16}$$

where n_i is the number of nearest neighbours of the i^{th} atom, $\mathrm{cn}(j)$ is the coordination number of the j^{th} atom and $\mathrm{cn}_{\max}$ is the maximum coordination number of the symmetry of the system (for example, for FCC this is 12).

The authors of [10] show clearly that this method has better correlation with oxygen binding strength than the mere coordination number and that it has much greater specificity. We will make use of the generalised coordination number extensively in chapter 5. It can be seen clearly from figure 5.10(b) in chapter 5 that this descriptor is linearly correllated with oxygen binding strength calculated explicitly with DFT in ONETEP.

3.3 Electronic Descriptor

We investigated the possibility of using catalytic descriptors to allow for cheaper prediction of catalytic activity than purely through relatively expensive (geometry relaxed) binding energy calculations with DFT. We also wanted to use the strengths of ONETEP to perform large-scale DFT calculations, such as the ability to compute electronic densities of relatively large metal systems in the model. This was partly the motivation for implementing the (l)pDOS method in ONETEP, described in section 3.1 for calculating d-band centres.

The descriptor described in this section is also based on the electronic density, like the d-band model, but we also wanted to incorporate the primary advantage of the generalised coordination number as we see it. We want a descriptor which is based on the density, but which does not average quantities over facets, instead allowing us to resolve the

descriptor on individual binding sites like generalised coordination. The advantage with respect to the generalised coordination number of this approach is that it is atomic species specific since it is based on the electronic density. Hence, core-shell nanoparticles and other bi-metallic nanoparticles should be resolved with such a descriptor unlike with generalised coordination number. This was the motivation for developing a new descriptor for catalytic activity, but we see a role for every descriptor described in this chapter in our application work.

The way we wanted to approach this problem was to find an isosurface of a quantity which encloses the nanoparticle and then to plot the values of the density onto this. With the right isosurface, the values of the descriptor at atomic sites (and between) could be read off by projecting the binding position of adsorbate atoms onto the isosurface and finding the value of density at this intersection on the isosurface. To do this in a high-throughput manner, as we would have to do with the many nanoparticles being produced in experiments, through our collaboration with the Nellist group (see 5), we would need a way to automatically determine where on the nanoparticles oxygen atoms would bind.

3.3.1 α-hulls and automatically determining oxygen binding sites

In order to automatically determine where on the nanoparticles the oxygen binding locations are, we need first to be able to determine a manifold connecting the surface atoms of the nanoparticle. The way we achieved this in this work was to take an α-shape, or α-hull of the nanoparticle atoms. An α-hull is a generalisation of a convex-hull which results from the operation of taking a ball of radius R_b and defining the manifold of every point that the centre of the ball can reside where its outside is in contact with at least one point in the cloud of points defined by the atomic centres and where no atomic centre is within R_b. This operation is effectively the same as if we could take a ball covered in paint and roll it over the nanoparticle, painting the spaces between atoms. Once the paint dried, we would have a shrink-wrapped manifold representing the outside of the nanoparticle, but with no holes with radius $> R_b$. This is shown in figure 3.3.

Unfortunately, this physical method for rolling a ball through space and monitoring collisions would be expensive to run. Instead we opted to use Edelsbrunner's algorithm [147]. In essence, this algorithm boils down to three steps:

1. Compute the Delaunay triangulation (gives tetrahedra in 3D) of the point-set defined by the atomic centres of the atoms in the nanoparticle

2. For every triangle face of each tetrahedron, compute the circumsphere and then determine whether any other points from the point-set lie within this circumsphere. If none do, then this triangle forms part of the manifold.

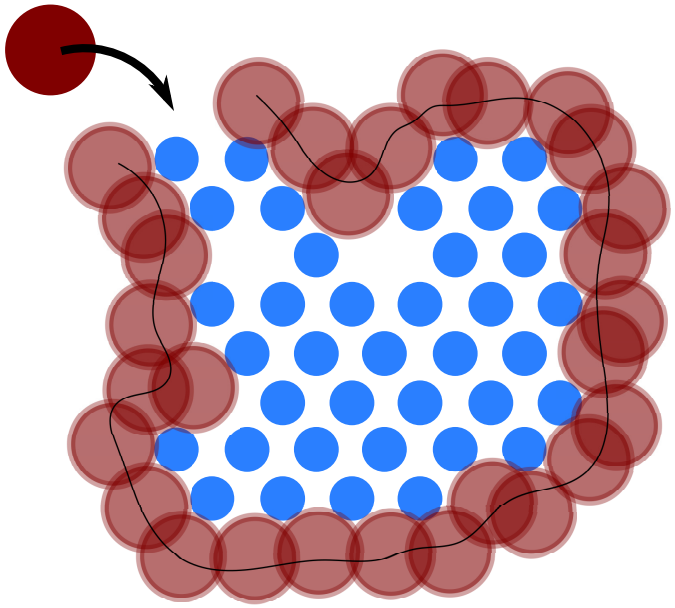


FIGURE 3.3: The effective procedure of generating an α-hull for the nanoparticles (atoms represented by the blue circles) is to take a sphere of radius R_b (in red) and roll it over the surface of the nanoparticle to generate manifold which represents the surface of the nanoparticle.

3. Decide whether parts of the manifold(s) are the outside of the α-hull or internal, disgard anything internal.

Firstly, the Delaunay triangulation[148] is defined as the triangulation (connection of points into triangles) of a set of point for which the circumcircles (circle with each point of a triangle on its circumference) do not contain any of the other points in the set. This leads to a triangulation which tends to avoid extremely acute triangles. This triangulation is easily generalised to 3D, where instead the circumspheres of the tetrahedra are not allowed to contain any of the other points in the set.

Several algorithms exist to compute the Delaunay triangulation of a set of points in 3D, but of these divide and conquer[149] has been shown to be the fastest[150]. In divide and conquer, the set of points is divided recursively in half until only three or four points exist in each partition. A Delaunay triangulation is calculated in any way on each of the many partitions. At this point, the triangulations are merged in pairs, recursively until the Delaunay triangulation of the whole set is obtained. In this work, we used the divide and conquer Delaunay method in the qhull library[151].

The next step in the α-hull algorithm of Edelsbrunner in 3D is to compute the circumcentre of a sphere of radius R_b on every triangle face of every tetrahedron and then the centres of the spheres. It is then simply a matter of finding those spheres which do not either circumscribe four points (one of the Delaunay spheres) or contain points. All the spheres in this set exist on the outside of the manifold, their circumscribed triangles form the manifold itself.

The α-hull must be optimised with respect to the radius of the test sphere, R_b. If the radius is made too big then the surface features of the nanoparticle will be averaged out and the hull will approach the convex hull. If the radius is made too small then the test sphere may pass between the outer points of the nanoparticle and produce unintended cavities in the hull. In the case of platinum nanoparticles, a radius of 2.3 Å was found to be a good compromise.

After finding the α-hull, the symmetries of the polygons of the hull must be found. The first step is to form a network data structure of the triangles on the manifold linking each triangle to its neighbours, simply by working out which triangles share an edge. Next, triangles are grouped into sets which are approximately coplanar. This is achieved by finding the equation of the plane of each triangle, then finding whether the scalar product of normal vectors is close to one within some tolerance.

Triangles with (111) symmetry may be found by finding all those which have a value of $\sqrt{(\alpha - 60.0)^2 + (\beta - 60.0)^2 + (\gamma - 60.0)^2}$ beneath some threshold, where α,β and γ are the three angles of the triangles. Triangles which form part of a site with (100) symmetry may be found by finding neighbouring, coplanar triangles which each have an angle of approximately 90 degrees and are joined along the hypotenuse. These triangles are merged to form (100) quads in the data-structure and labelled as such. With this information it is possible to find the projections of the hollow sites on the hull as the centroid of either the triangles for (111) sites or quads for (100) sites, the projections of bridge sites on the hull as the mid-points of the edges connecting any of the shapes and the projection of top sites as the points themselves. The binding sites are then calculated by moving off the hull along the normal vectors (or average of adjacent normal vectors for bridge and top sites) from the projected sites. The distance to travel along the normals must be known *a priori*.

The binding sites, projected on the α-hull were found, for bridge sites, by taking the midpoint of the line connecting two atoms; for 4-fold sites by taking the centroid of the quadrilateral formed by two polygons of the α-hull, if those polygons are connected along their hypotenuse, have an angle of approximately 90 degrees (85 to 95 degrees) and are co-planar; and for 3-fold sites, as the centroid of the α-hull polygons.

To find the position of the adsorbate binding site; on the bridge site, the average of the normal vectors of the adjoining polygons can be used to project off the α-hull. For 3- and 4-fold sites, the normal vector of the α-hull polygon itself can be used.

A program was written in Fortran to find the sites using the above procedure, and making use of the Qhull library (available from www.qhull.org). This program will be made freely available from the Utilities section of the ONETEP code website:
http://www.onetep.org/Main/Utilities

3.3.2 Electronic Descriptor Isosurfaces

The first option we considered was to use the electrostatic potential as has been used in biological descriptors for a similar purpose before[152, 153], to assess the binding affinity of ligands on biological macromolecules. These methods typically use solutions to the Poisson-Boltzmann equation to get an electrostatic potential isosurface[154], whereas we will use electrostatic potentials from DFT. In order to do this with ONETEP, we needed to implement a routine in ONETEP to compute and output an electrostatic potential which was compatible with PAW.

The electrostatic potential in DFT should be just the sum of the electronic Hartree term and nuclear Coulombic term,

$$V_{\mathrm{ES}}(\mathbf{r}) = V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{ext}}(\mathbf{r}), \tag{3.17}$$

however this is complicated by the use of pseudopotentials or the PAW approach in ONETEP. With norm-conserving pseudopotentials, V_{ext}, which is the local part of the pseudopotential can be calculated along with $V_{\mathrm{H}}(\mathbf{r})$ in one step by computing the Hartree potential of the valence *and* core density.

This procedure introduces a difficulty, however, in that in practice the interpolation between grids introduces ringing noise on the ONETEP fine potential grid after the reciprocal space application of the Hartree operator. This noise is very small in magnitude, too small to affect the results of our calculations and has high frequency on the fine potential grid, which we can filter out to a good degree using a simple odd-even filter.

To deal with the issue in general, as it affects our visualisation of electrostatic potentials, for norm conserving pseudopotentials we have opted to use the "smeared ion" approach. This had previously been implemented in ONETEP by Dziedzic *et al* [155] for performing implicitly solvated calculations with the "multigrid" approach to computing the Hartree operator in real space. By substituting the smeared atomic core regions, which are just gaussian functions with a width of the core sphere of each atom, into the total density,

$$n_{\mathrm{TOT}} = n_{\mathrm{ion}} + n_{\mathrm{elec}}, \tag{3.18}$$

where n_{ion} is the smeared core density and n_{elec} is the density of the valence electrons we are able to get smoothly varying electrostatic potential isosurfaces which are accurate outside of the core spheres.

In the case of PAW, we do not need to turn to anything outside of the PAW formalism. We use PAW because the "fine" grid in plane-wave codes is not fine enough to represent the core electrons, so we cannot output the PAW all electron potential on this uniform grid. Doing so results in large aliasing artifacts which swamp the potential and make plotting a smooth isosurface impossible. Instead we simply opt to output the Hartree

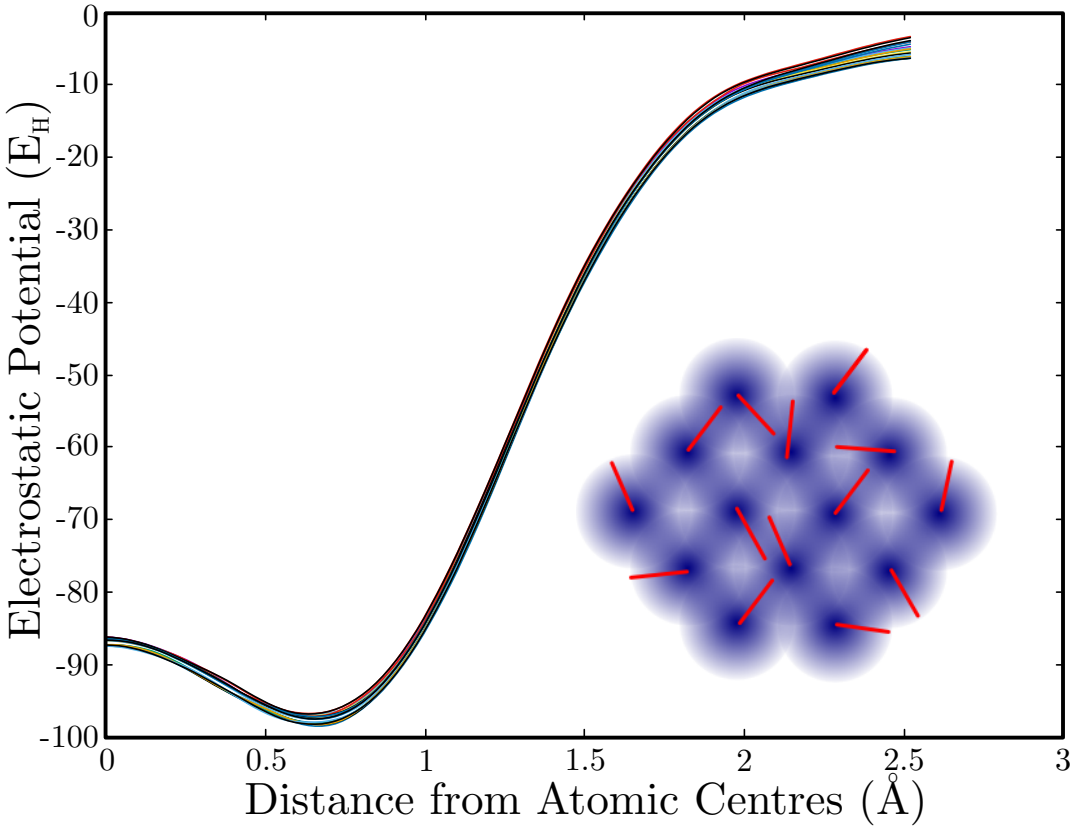


FIGURE 3.4: To find the minimum point of the electrostatic potential we find the potential along a set of equal length random vectors out of each atomic centre (inset). The plot of the potential along these vectors is shown for cuboctahedral Pt_{309}, showing the spread of the distribution.

potential of the pseudo-density. This is computed by summing the valence density with the PAW augmentation density

$$n_{\mathrm{TOT}} = n_{\mathrm{aug}} + n_{\mathrm{elec}}, \tag{3.19}$$

and applying the Hartree operator in reciprocal space. This is achieved by taking the Fourier transform of the density $n(\mathbf{r}) \underset{FFT}{\rightarrow} \tilde{n}(\mathbf{G})$ and feeding it into the Hartree operator,

$$\tilde{V}_{\mathrm{H}}(\mathbf{G}) = \frac{4\pi n_{\mathrm{TOT}}(\mathbf{G})}{\Omega_{\mathrm{SC}} G^2}, \tag{3.20}$$

where $\mathbf{G}$ are wavevectors (points in reciprocal space), G are the modulus of $\mathbf{G}$ and Ω_{SC} is the volume of the simulation cell. The electrostatic potential is then inverse Fourier transformed back to direct space $\tilde{V}_{\mathrm{H}}(\mathbf{G}) \underset{iFFT}{\rightarrow} V_{\mathrm{H}}(\mathbf{r})$, and outputted. If non-periodic boundary conditions are required and the dipole moment of the system is large enough that a sufficient vacuum gap is too expensive to use, then a Coulomb cutoff[156] or multigrid [155] (real space) approach to computing the Hartree potential may be used.

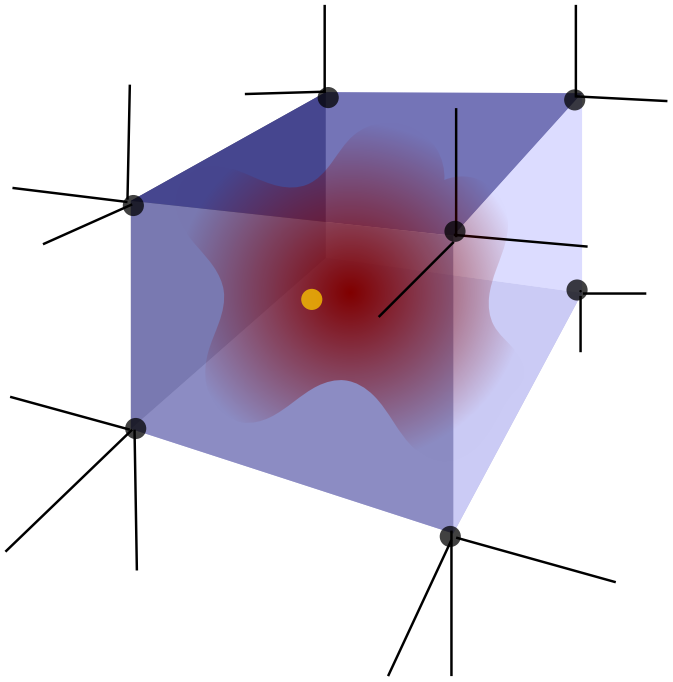


FIGURE 3.5: The tricubic interpolation defines and interpolant in each cube of gridpoints, based on the values and derivatives of the function at those 8 gridpoints. From this interpolant, the value of any point within the cube can be found.

Once we had a set of smooth electrostatic potentials, the next step was to ensure that the isosurfaces that we plotted from these electrostatic potentials had the same meaning on differing nanoparticle systems. To achieve this we needed a reference potential. Several options were considered including $\lim_{\mathbf{r}\to\infty} V_{\mathrm{ES}}$ which is difficult to approximate with the periodic boundary conditions of our calculations, and zero, which was problematic due to its arbitrary nature under a pseudopotential approximation. The reference we chose was the minimum point of the electrostatic potential, which we found by plotting the potential along vectors out of the atomic centres of each atom. We use a single random vector for each atom, where every vector in the set has the same length. This is shown pictorially in figure 3.4 along with examples of the potential along the set of vectors plotted on the main axes (for Pt_{309}).

The direction of the random vectors which we define in figure 3.4, outward from the atomic centres may not coincide with any points on the uniform grid of the electrostatic potential. To avoid this issue, we use tricubic interpolation on the grid. Tricubic is the 3D generalisation of cubic interpolation described in Lekien and Marsden [157].

The method described in Lekien and Marsden defines an interpolant in each cube of gridpoints (shown in figure 3.5), based on the values and derivatives of the function at those 8 gridpoints. The interpolated value of the function at any point within this cube can be found by using the following formula in terms of the interpolant values, a_{ijk} and

powers of the components of the coordinates, x^i, y^j and z^k as:

$$f(x,y,z) = \sum_{i,j,k=0}^{3} a_{ijk} x^i y^j z^k, \tag{3.21}$$

where the three dimensioned a_{ijk} can be flattened into a vector:

$$\bar{a}_{1+i+4j+16k} = a_{ijk} \qquad \forall \; i,j,k \in \{0,1,2,3\}, \tag{3.22}$$

so that if we construct a vector b_i of the function at the 8 vertices, label the vertices as $p_1 \ldots p_8$ in an arbitrary, but consistant fashion, and use directional derivatives upto cubic order in the following way:

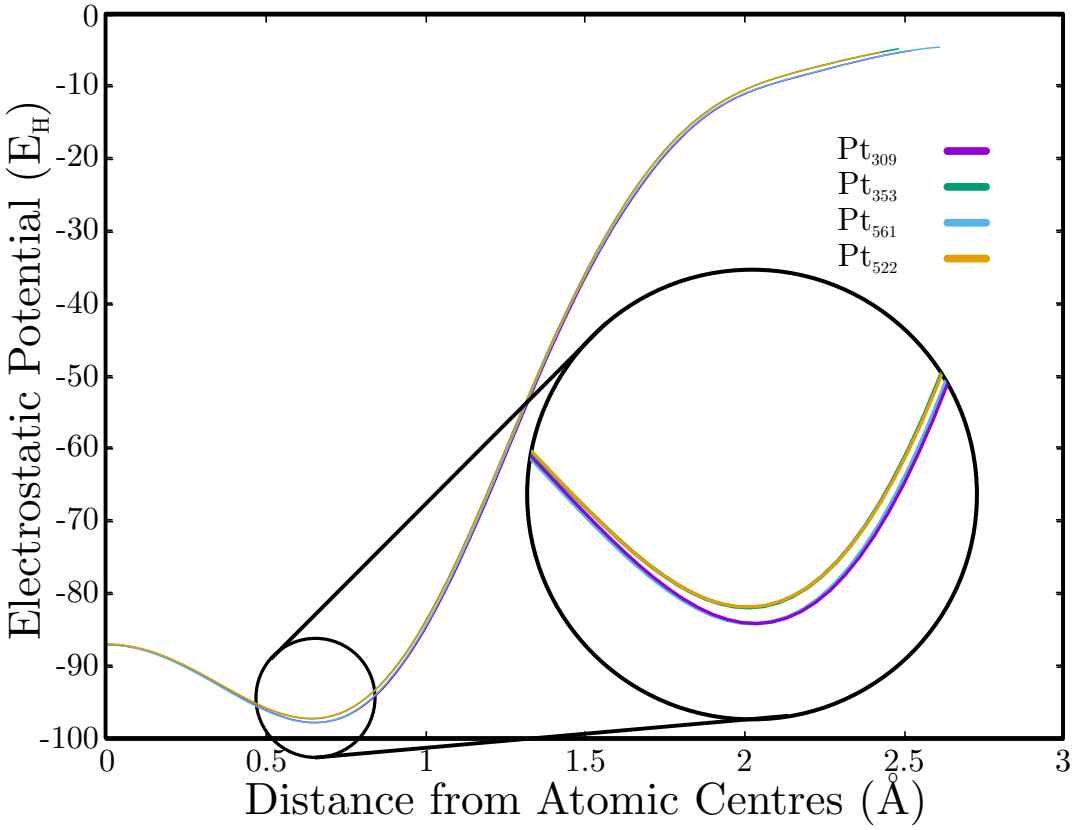


FIGURE 3.6: This figure shows the average of the potentials as found by the method described in figure 3.4 for 4 platinum nanoparticles. Cuboctahedral Pt_{309} and Pt_{561}, and experimental Pt_{353} and Pt_{522}. This allows us to find the minimum points of the various nanoparticles which are zoomed in the inset and shift them to the same reference potential.

$$b_i = \begin{cases} f(p_i), & \text{if } 1 \leq i \leq 8 \\ \frac{\partial f}{\partial x}(p_{i-8}), & \text{if } 9 \leq i \leq 16 \\ \frac{\partial f}{\partial y}(p_{i-16}), & \text{if } 17 \leq i \leq 24 \\ \frac{\partial f}{\partial z}(p_{i-24}), & \text{if } 25 \leq i \leq 32 \\ \frac{\partial^2 f}{\partial x \partial y}(p_{i-32}), & \text{if } 33 \leq i \leq 40 \\ \frac{\partial^2 f}{\partial x \partial z}(p_{i-40}), & \text{if } 41 \leq i \leq 48 \\ \frac{\partial^2 f}{\partial y \partial z}(p_{i-48}), & \text{if } 49 \leq i \leq 56 \\ \frac{\partial^3 f}{\partial x \partial y \partial z}(p_{i-56}), & \text{if } 57 \leq i \leq 64 \end{cases} \tag{3.23}$$

then the vector $\bar{a}_{1+i+4j+16k}$ and hence a_{ijk} via equation 3.22 can be found once and for all for each cube of gridpoints by constructing b_i from finite differences and computing

$$\bar{\mathbf{a}} = \mathbf{B}^{-1}\mathbf{b}, \tag{3.24}$$

where the matrix $\mathbf{B}^{-1}$ has elements which may be found in [158]. For the finite difference approximations to the derivatives in (3.23), we used a centred difference approach.
After using this interpolation method to take points from the electrostatic potential on the uniform fine grid, along the vectors of random direction out of each each atomic centre, the 1D electrostatic potential values were averaged over every centre, giving an electrostatic potential curve for each nanoparticle, shown in figure 3.6.
At this point we can find the minima of the averaged electrostatic potentials for each nanoparticle. The shifts to the value of each minima from a reference potential are calculated, before the negative of these shifts are added onto the function of electrostatic potential at every point on the uniform grid. This procedure serves to ensure that every potential has the same reference potential, and in practice, we use one nanoparticle's minimum in the average electrostatic potential as this reference.

3.3.3 Computing Values from the Electronic Descriptor

With electrostatic potentials from ONETEP which have been shifted to have an equivalent reference potential, as well as the electronic charge density from ONETEP on the same uniform grid, we can begin to compute values of our descriptor at sites corresponding to oxygen binding sites on the nanoparticle. A method for finding every possible oxygen binding site on FCC platinum nanoparticles was shown in section 3.3.1 and with this site data, we have all the information required to compute values from the electronic descriptor.

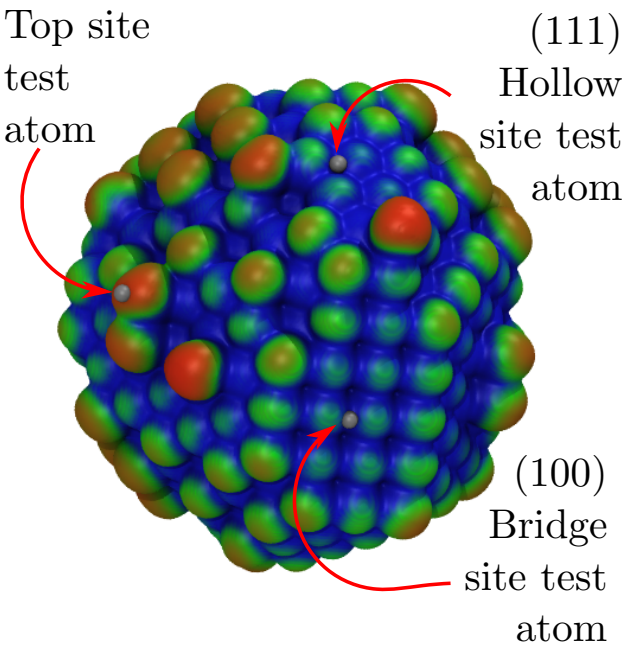


FIGURE 3.7: When computing the values of the electronic descriptor on sites, we use test atoms at the positions where adsorbates would bind to define vectors to intersect with the isosurface under the site and read off values.

The definition of our electronic descriptor is the value of the electronic charge density of a bare nanoparticle (no adsorbates) on an isosurface of the electrostatic potential of the same system, which is closest to the position of a test atom for which we would like the value of the electronic descriptor. This "test atom" does not exist either in the ONETEP calculation or in this descriptor post-processing. It is merely the coordinates of an adsorbate atom if *it were* adsorbed in the chosen binding site (this is shown pictorially in figure 3.7). Also, the particular isovalue of the electrostatic potential isosurface does not matter, apart from that the isosurface should be a closed manifold enclosing the nanoparticle atoms. It should also not be too far away from the nanoparticle that electrostatic potential differences cannot be resolved. In chapter 5 where we have used this approach, we found that an isovalue of $0\,E_H$ gave a reference potential corrected isosurface which was $\sim$2.9 Å away from platinum atomic centres on atop sites.

To automatically find the point on the isosurface which is closest to a test atom, we firstly compute a fine grid of coordinates on the isosurface, by using the tricubic interpolants which we described in the previous section in a root finding algorithm on the difference between the target isovalue and equation 3.21. We then find the closest point in this set on the isosurface to the test atom coordinates using a simple linear search algorithm [159] because the site of the isosurface set is small. We then find the value of the tricubic interpolant of the density at these coordinates and that is the value of the electrostatic potential. This procedure is repeated for every test atom.

It can be seen clearly from figure 5.10(a) in chapter 5 that this descriptor is linearly correllated with oxygen binding strength calculated explicitly with DFT in ONETEP. This fact leads us to the conclusion that this procedure can also be used sucessfully as a descriptor, by calculating the value of the descriptor at each site of interest and reading off the predicted binding strength.

3.4 Predicting Catalytic Activity with Descriptors

The values of the descriptors described in this chapter correlate with the binding strength of oxygen on metal nanoparticles. To be useful for determining the usefulness of a particular nanoparticle as a catalyst, for instance on the cathode of a PEM fuel cell, a descriptor must not only be able to predict binding strength of oxygen, but also to predict the catalytic activity of a nanoparticle.

In chapter 5 we do this in a conceptually simple way. Because the optimal oxygen adsorption strength has already been predicted by other groups [58, 76, 80], we use these predictions as our baseline.

We firstly parameterise a calibration curve by fitting a polynomial through data of oxygen binding strength on a set of binding sites on a model nanoparticle vs the value of the chosen site specific descriptor, which is also applied on each of these sites (either generalised coordination number or our electronic descriptor).

Using the α-hull method (see section 3.3.1), we find all of the binding sites on a nanoparticle(s) of interest. We apply the chosen site specific descriptor to these all of these sites and use the calibration curve to give us a set of predicted binding strengths.

We then find the proportion of this set of binding strengths which lie within some energy interval around the optimal adsorption strength. In the work which is described in chapter 5 we choose to use an interval of ± 0.2 eV. We argue that nanoparticles with a larger proportion of binding sites close to the optimal binding strength for catalysis should make better catalysts.

3.5 Conclusions

This chapter discussed three catalytic descriptors, the generalised coordination number which is a purely geometric descriptor that gives site specific values; the d-band centre which is based on local angular-momentum projected density of states and can give surface averages values, and our electronic descriptor which is based on the electronic density and electrostatic potential from DFT calculations of bare systems. All of these descriptors operate with coordinates of or DFT calculations of bare systems, but predict the binding strength of systems with adsorbates.

Requiring no calculations, but only coordinate data of the nanoparticle atoms, the generalised coordination number is the cheapest of these descriptors. It has the disadvantage of not being based on the electronic structure of the system, and hence cannot predict changes to activity based on changes of species, but is site specific. The d-band centre is not site specific, but it is based on electronic structure and sensitive to species and in particular to lattice strain and metal atom separation. Our electronic descriptor is site specific and based on the electronic structure, so sensitive to species changes. The latter two calculation

We have described in this chapter how each of these descriptors have been implemented: the generalised coordinate number as a standalone script, d-band centre as a routine of ONETEP which computes an angular momentum projected density of states calculation, as implemented in this project, and our electronic descriptor which has been implemented as a series of post-processing tools as well as a routine to compute and output electrostatic potentials from ONETEP after single point energy calculations.

We have used all of these descriptors in our collaboration with the Nellist group, where we apply them to experimentally measured nanoparticles, which can be seen in greater detail in chapter 5.

Chapter 4

Simulations of Oxygen Binding on Platinum Nanoparticles

Abstract

This chapter will explore the binding of oxygen to platinum nanoparticles using DFT calculations. We were interested in heterogeneous catalysis with metal nanoparticles in general, rather than a particular chemical reaction, however, the results presented in this chapter focus on oxygen adsorption on platinum.

Platinum is a good example because it is used as the catalyst in so many catalytic reactions, including the hydrogenation of liquid vegetable oils to margerines[160], in cracking large crude oil molecules to form smaller molecules such as those in petrol[160], in the catalytic decomposition of hydrogen peroxide[161], and of course on both the anodes and cathodes of fuel cells as explored in chapter 1 in section 1.7.2.

Calculations have been performed on cuboctahedral nanoparticles of up to 1000 atoms using the ONETEP package using the techniques described in chapter 2. Initially we performed a study of oxygen adsorption on single facets of the nanoparticles to assess binding strength trends with nanoparticle size and to determine at which size a slab model would be an adequate model for a facet of the nanoparticle. The second part of the study explores oxygen coverage of Pt_{147} to determine how binding strength is affected by fractions of a monolayer of oxygen coverage. We have also applied the techniques outlined in chapter 3 as descriptors of the systems explored in this chapter.

4.1 Cuboctahedral Nanoparticles

Initially, the choice was made to study the adsorption of small molecules on platinum nanoparticles. Platinum was chosen partly because it lies toward the peak of the Sabbatier curve for catalytic activity in the oxygen reduction reaction (ORR) figure 1.8a.

There were, however, a great many more options as far as the choice of nanoparticle composition was concerned[162–164].

Platinum exists in the bulk as a face-centred cubic (FCC)metal [165], so any shape which can be formed by enclosing a region of bulk within intersecting cleavage planes could be used as the geometry of the nanoparticle under study. However, the number of possible shapes is huge, for instance, tetrahedra, cubes, octahedra, truncated cubes, cuboctahedra and truncated octahedra are all possible regular geometries. Irregular geometries, as would be found in nature, increase the number of possible atomic arrangements to the point that it would be impossible to list all possible configurations of even relatively small nanoparticles, let alone to perform meaningful studies on such large datasets.

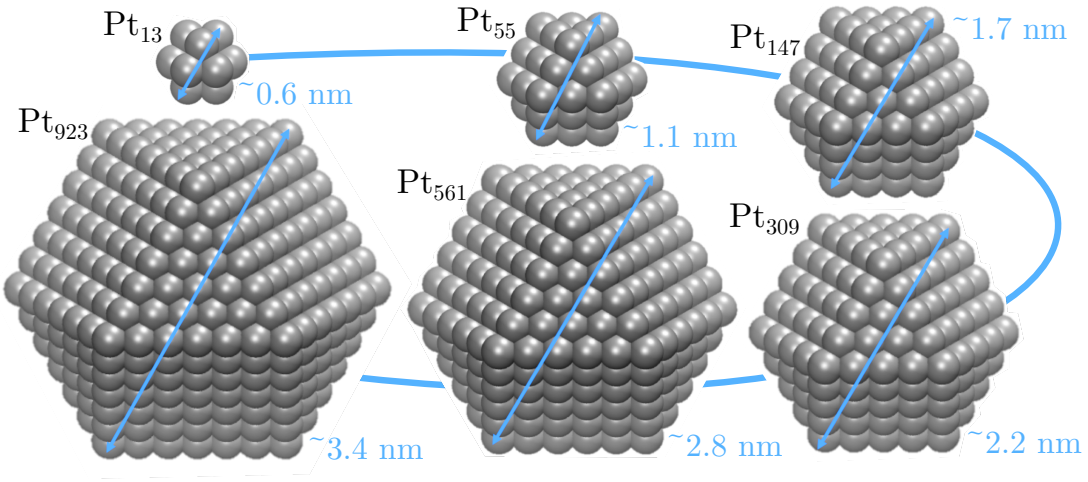


FIGURE 4.1: The first 6 "magic numbers" of platinum cuboctahedral nanoparticles showing how the number of atoms scales cubically with the diameter; 5 nm is not reached until the nanoparticle has 2869 atoms.

To avoid issues with the number of possible configurations, the cuboctahedra were chosen as a model system on which to perform adsorption calculations (see figure 4.1). This is partly because several experimental studies [166–168] have shown that approximately cuboctahedral nanoparticles can be produced experimentally. Also, cuboctahedral nanoparticles have a surface area split between (111) and (100) facets, increasing the number of potential interesting adsorption sites (see figure 4.2).

Using a model structure in this way it is possible to perform a study of how molecular adsorption is affected by increasing nanoparticle size, while minimising the effect of external variables such as composition and surface structure. Cuboctahedra of several "magic numbers", or atom counts as successive shells of atoms are added to the seed cuboctahedron remain similar in shape when cleaved from the bulk, retaining the same number of facets of hexagonal and square symmetries and the same relative areas.

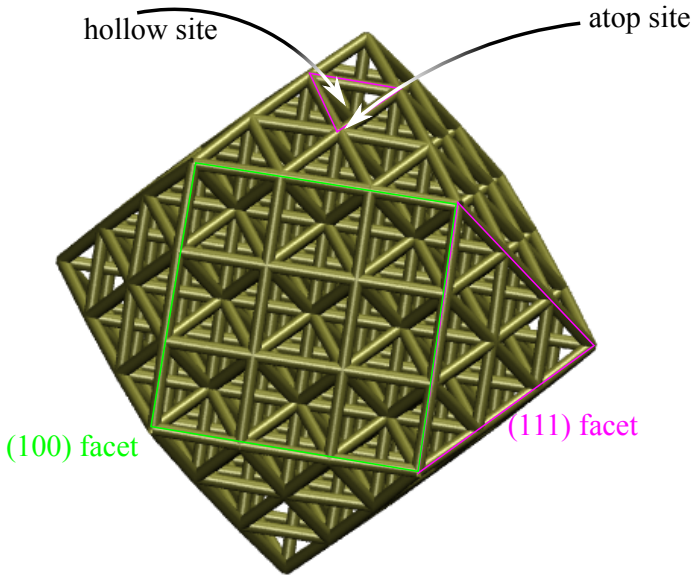


FIGURE 4.2: The facets of the Cuboctahedral nanoparticle, with sites on the (111) facet labelled.

The first 5 magic numbers of the cuboctahedral nanoparticles are (1, 13, 55, 147, 309) and are given in general as[169]:

$$C_n = (2n+1) \times \frac{(5n^2+5n+3)}{3}, \tag{4.1}$$

which clearly indicates a cubic order scaling and hence, calculations of many hundreds or even thousands of atoms will be extremely expensive or even impossible on current resources with conventional plane-wave DFT packages, though the 13 and 55 atom cases should present no difficulty and the 147 and 309 atoms systems ought to be possible on supercomputers. Our interest in these systems, however, was for chemically relevant systems, which are interesting for catalysis in the $\sim$ 2–10nm size range, involving hundreds to thousands of atoms. For these systems we anticipated that we would need to use ONETEP and its EDFT functionality in this study.

As EDFT is a relatively recent addition to ONETEP, it was imperative for us to test and validate the functionality against established, conventional methods, in software which are known to be reliable through years of published research output. We opted for Castep as the benchmark DFT package to validate against as it has had an implementation of EDFT since the late 1990s, it can use the same pseudopotential format as ONETEP (for norm conserving pseudopotentials) and it uses a plane-wave basis set which is formally

equivalent to the p-sinc basis in ONETEP. We chose the Pt_{55} system with a set of molecular adsorbates as the testbed on which EDFT functionality in ONETEP would be validated.

4.2 Validation Against Castep

Modern DFT packages offer hundreds of parameters and configuration options, as well as using inequivalent unit systems for input/output and differing conventions, for instance, using thermodynamic beta (β) rather than temperature. So rigorous comparison between DFT packages needs very careful matching of each input parameter, including units and conventions. The similarity between the Castep and ONETEP input files and compatibility of pseudopotential files goes some way to minimising the risk of an unfair test, however, a great deal of care was still taken to ensure that all remaining settings were equivalent; particularly those parameters which were set from default values rather than values in the input files. Also, Castep uses eV for energies and Åfor lengths by default, whereas ONETEP uses E_H for energies and a_0 for lengths by default.
Some aspects of the calculations will always remain inequivalent, and this is unavoidable due to the differing natures of the two methodologies. An example of this is in the representation of the valence electrons (delocalised electron bands vs. NGWFs). Another is the plane-wave kinetic energy cut-off corresponding to a sphere of wavevectors in reciprocal space in Castep, but the p-sinc kinetic energy cut-off corresponding to a cube in reciprocal space in ONETEP. This means that a calculation performed in Castep with a particular plane-wave energy cut-off is non-equivalent to an otherwise identical as possible calculation performed in ONETEP and there is *no* value which is equivalent. This is not a problem, however, because in the limit of high energy cut-off, where additional wave-vectors with further increased cut-off have a negligible contribution to the total energy of the system, the total sum of contributions from wave-vectors should be comparable between approaches.
For this reason, together with the usual arguments for ensuring a large enough basis-set size to accurately represent the Kohn-Sham wave-functions, the calculations first had to be converged *individually* with respect to plane-wave energy cut-off.
EDFT also requires extra, partially occupied basis functions above the chemical potential, we chose to have enough of these in our calculations that the occupancy of the highest energy state is negligible.
At finite temperature, states above the chemical potential are required because they are partially occupied in a metal and truncating this distribution will affect the total energy. A secondary reason for needing a large number of conduction states is that during the electronic energy minimisation process, the extra degrees of freedom allow the optimisation path toward the energy to temporarily move electron density from valence

to conduction states. For these reasons it is important to ensure that enough extra NGWFs in ONETEP and extra electron bands in Castep have been made available to the calculation. We do not know this number *a priori*, instead we monitor the calculations to ensure that at least one or two states are empty at every iteration of the calculation.

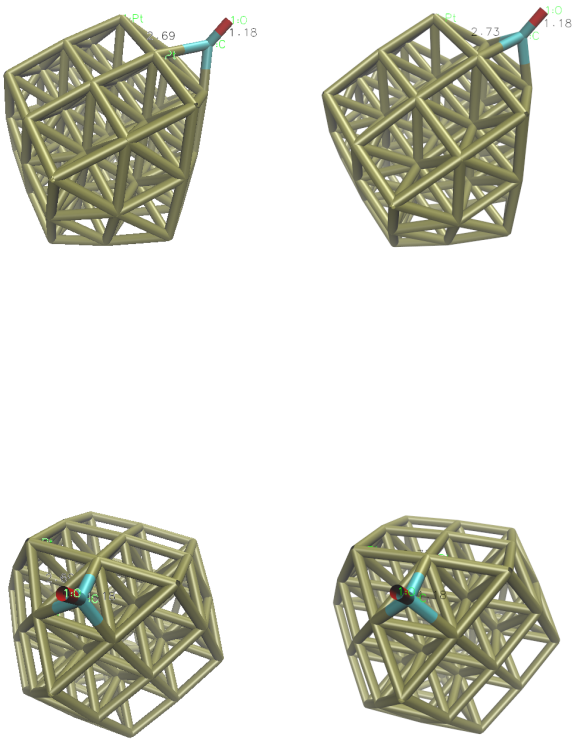


FIGURE 4.3: The orientation and proximity to the nanoparticle of the CO ligand after geometry optimisation is almost identical between Castep (left) and ONETEP (right).

After converging with respect to kinetic energy cutoff in both codes and number of NGWFs in ONETEP we found that we could find the same geometries (see figure 4.3) and absolute and relative energies table 4.1 in both codes using the same (Rappe-Bennett) set of norm conserving pseudopotentials [170] and spin polarised calculations.

We found that 850 eV kinetic energy cutoff was enough to converge the ONETEP calculations, so this value was used for both sets of calculations. We also used the Γ-point approximation in Castep because we are simulating isolated systems and because this most closely matches the ONETEP calculations which are Γ-point only.

We have compared the ONETEP EDFT approach on some small demonstrative calculations on Pt_{13} against Castep, which is a conventional plane-wave code with an alternative

TABLE 4.1: A table showing the remarkable agreement between Castep and ONETEP on converged calculations. The results show that not only the relative energies agree, but also the total energies.

(eV)	Castep	ONETEP
Pt55 Total Energy	-39220.418	-39220.413
O2 Total Energy	-858.217	-857.879
Combined Total Energy	-40079.377	-40079.105
Adsorption Energy	-0.741	-0.812
CO Total Energy	-584.501	-584.228
Combined Total Energy	-39806.478	-39806.198
Adsorption Energy	-1.559	-1.556

EDFT implementation.[171] The same norm-conserving pseudopotentials were used in both Castep and ONETEP.

We have computed the oxygen binding energy of a fully relaxed Pt_{13} nanoparticle according to equation 4.7, with both codes, while applying periodic boundary conditions in a cubic simulation cell with sides of 25 Å. The energies of the individual species, as well as the binding energy, as computed with each code, are shown in Table 4.2. We can observe that the binding energies agree to within 0.01 eV which is about 0.23 kcal/mol which is well within the level of precision required for chemical accuracy (which is of the order of 1 kcal/mol).

TABLE 4.2: Comparison of energies between Castep and ONETEP on Pt_{13}. All calculations performed with periodic boundary conditions in 25 Å, cubic simulation cells. All energies are in eV.

	Castep	ONETEP
Pt_{13}	-9302.94	-9302.47
$Pt_{13} - O_{(111)}$	-9734.48	-9733.94
$1/2O_2$	-429.11	-428.94
Binding Energy	-2.44	-2.53

4.3 DFT calculation set up

Our Density functional theory (DFT) calculations were performed with the ONETEP linear-scaling DFT program,[40] using its ensemble-DFT (EDFT) method,[7] which allows large-scale DFT calculations on metallic systems with thousands of atoms. ONETEP is based on the one-particle density matrix formulation of DFT and employs a two-resolution approach to describe the density matrix. At the coarse level it employs a

set of non-orthogonal generalised Wannier functions (NGWF) which are optimised *in situ*[172] during the calculation as they are expanded in a basis set of psinc functions which are equivalent to plane waves. The quality of the psinc basis set can be controlled with a single parameter, which is equivant to the kinetic energy cut-off in plane wave approaches. As the psinc functions are independent of atomic positions, ONETEP calculations do not suffer from basis set superposition error.[173] NGWFs are, strictly localised in spherical regions and so they are appropriate for the isolated nanoparticle systems under investigation here as vacuum is effectively free in terms of computational cost.

ONETEP can either use norm-conserving pseudopotentials or the Projector Augmented-Wave (PAW) approach [144] to represent the core electrons. When using the norm-conserving approach, we have used the norm-conserving PBE library of the Rappe group[170]. When performing PAW calculations, we have used the PBE data from the GBRV [174] library.

For all the calculations we employed the rPBE generalised gradient approximation exchange-correlation functional of Hammer and Nørskov,[20] which is known to be particularly good for calculations of binding energies on surfaces[21]. Since the highest psinc kinetic energy cut-off is determined by the atomic species with the hardest pseudopotential, which in out case is platinum, we investigated the convergence of the total energy with respect to the kinetic energy cut-off for a Pt_{13} nanoparticle and we have chosen a kinetic energy cut-off of 850eV for all our calculations as at this cut-off the total energy of Pt_{13} is already converged to 0.0001 E_h.

Our calculations were performed in simulation cells which obey periodic boundary conditions, but with large vacuum gaps and the spherical Coulomb-cutoff approach in order to eliminate interactions between periodic images.[156] The radius of the sphere used for the cut-off in the Coulomb operator was calculated using:

$$r_{\mathrm{cc}} = \max(d_{\mathrm{aa}}) + 2r_{\mathrm{NGWF}} + 5, \tag{4.2}$$

where d_{aa} is the inter-atomic separation, r_{NGWF} is the radius of the NGWF localisation spheres and the 5 is to account for Gibbs oscillations associated with the cut-off in the Coulomb operator, which is performed in reciprocal space. The radius of the NGWF spheres was chosen to be 4.8 based on platinum total energy convergence data. The Brillouin zone was sampled solely at the Γ point. The Fermi-Dirac occupancy smearing scheme was used, with 0.1 eV electronic temperature which corresponds to 1160 K and a plane-wave kinetic energy preconditioning parameter of 3.8$\mathrm{Å}^{-1}$.

For geometry optimisation calculations the BFGS quasi-Newton method was employed to relax ionic coordinates in cells of fixed dimension.[175] The geometry optimisation tolerances that we used were: $2.72\times10^{-5}eV$ in electronic free energy, 0.1eV/ in maximum

absolute force component and 0.003 in atomic displacement.

We have also performed calculations on Pt slabs using the CASTEP code. The slabs consisted of 5 layers of Pt cleaved along the (111) surface in the geometry of bulk Pt as previously optimised with CASTEP. The top 4 layers were geometry optimised with a vacuum gap of 15.0 Å. The cell consisted of 4 hollow sites so that placing an O on one of these sites corresponds to 0.25% monolayer coverage. A 5×5×1 Monkhorst-Pack grid of k-points was used and the calculations were performed with CASTEP's EDFT algorithm using Fermi-Dirac occupancy smearing with a width of 0.1 eV, as in the ONETEP calculations and the same plane wave kinetic energy cut-off (850 eV) as ONETEP was used.

TABLE 4.3: A table showing several of the most important parameters used in ONETEP, which would be necessary to reproduce the calculations that have been performed.

Parameter	Value
Plane-wave cut-off energy	850 eV
XC functional	rPBE
Fermi-Dirac smearing width in EDFT	0.1 eV
Coulomb cut-off radius	$\emptyset(nanoparticle) + 2r_{NGWF} + 5$Å
Plane-wave kinetic energy pre-conditioning	2.0 eV
Number of NGWFs	Pt : 9 H : 1 O : 4 C : 4
NGWF radius	9 Å

4.3.1 Computation of binding energies

Binding energies of O to Pt nanoparticles were calculated with reference to an isolated O_2 molecule and nanoparticles as:

$$E_{\mathrm{b}} = E_{\mathrm{p+O}} - E_{\mathrm{p}} - \frac{1}{2}E_{\mathrm{O_2}}, \tag{4.3}$$

where the nanoparticle energies were computed in cubic cells of dimension at least $2r_{\mathrm{cc}}$+5. The O_2 energy was computed in a cubic cell of 10 as a triplet electronic state, after the strucutre of O_2 was optimised.

We can differentiate our binding energy calculations into two cases by the type of structures we used. In the first case, the nanoparticles are generated as perfect cuboctahedrals (cleaved from bulk Pt with a Pt-Pt distance of 2.76 Å) and are kept constrained to their initial geometries throughout the calculation of the binding energy. So, in this case the binding energy is computed as

$$E_{\mathrm{b}}^{\mathrm{Con}} = E_{\mathrm{p+O}}^{\mathrm{Con}} - E_{\mathrm{p}}^{\mathrm{Con}} - \frac{1}{2}E_{\mathrm{O_2}}, \tag{4.4}$$

where the superscript Con indicates that in all these structures the Pt atoms are held fixed and constrained in the perfect cuboctahedral geometry in which the particle was constructed. Therefore $E_{\mathrm{p+O}}^{\mathrm{Con}}$ signifies that this is the energy of the nanoparticle with only the O atom coordinates relaxed.
In the second case, the binding energy is computed as

$$E_{\mathrm{b}}^{\mathrm{FRel}} = E_{\mathrm{p+O}}^{\mathrm{FRel}} - E_{\mathrm{p}}^{\mathrm{FRel}} - \frac{1}{2}E_{\mathrm{O_2}}, \tag{4.5}$$

where the superscript FRel indicates that the atomic positions in the structures used to calculate this binding energy have been fully relaxed. This is a more rigorous approach to calculate the binding energy of O on Pt nanoparticles. This procedure is especially time-consuming in the stage of structural relaxation of the nanoparticle plus O system (needed to obtain $E_{\mathrm{p+O}}^{\mathrm{FRel}}$) as it takes a large number of geometry optimisation steps due to the way most common geometry optimisers operate [176].
This posed a significant obstacle, especially for the larger nanoparticles, and in order to avoid it in future calculations, we have developed an alternative approach, which is summarised in the equation below.

$$E_{\mathrm{b}}^{\mathrm{Rel}} = E_{\mathrm{p+O}}^{\mathrm{Rel}} - E_{\mathrm{p}}^{\mathrm{FRel}} - \frac{1}{2}E_{\mathrm{O_2}}, \tag{4.6}$$

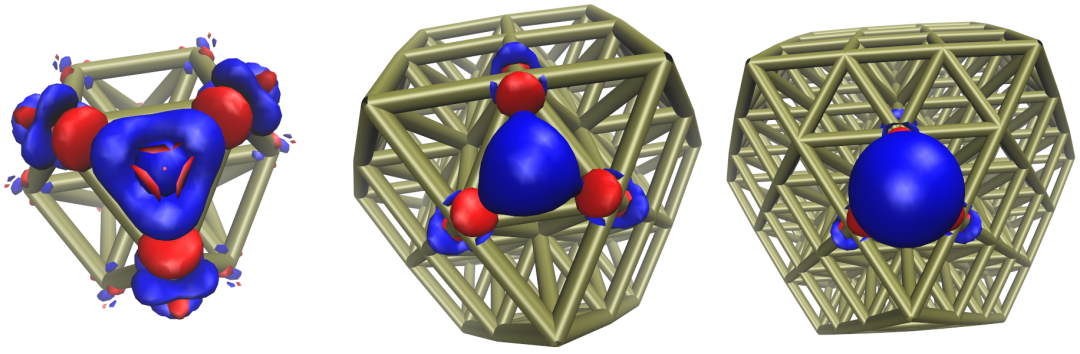


FIGURE 4.4: Hirshfeld defect densities of O on the (111) facets of cuboctahedral Pt nanoparticles from 13 to 147 atoms.

In this equation the superscript Rel denotes a partially relaxed nanoparticle+O system. To obtain this system we cleave a fragment from the fully relaxed Pt nanoparticle. The fragment is cleaved between the third and fourth layer of atoms below a (111) facet. Then the the O atom is added on the (111) facet, and the fragment+O is relaxed, while keeping the deepest layer of Pt atoms constrained to their original positions. After the relaxation has converged, the two layers of Pt atoms closest to the surface (which are the ones that were relaxed) replace the equivalent atoms of the full system, including

the position of the O atom. Then a single-point energy calculation is performed in the partially-relaxed nanoparticle+O system to obtain $E^{\mathrm{Rel}}_{\mathrm{p+O}}$. This idea is based on Hirshfeld charge defect density calculations, which indicate that the effect of adding a molecule to a (111) facet of a Pt nanoparticle extends to the third layer only (see figure 4.4).

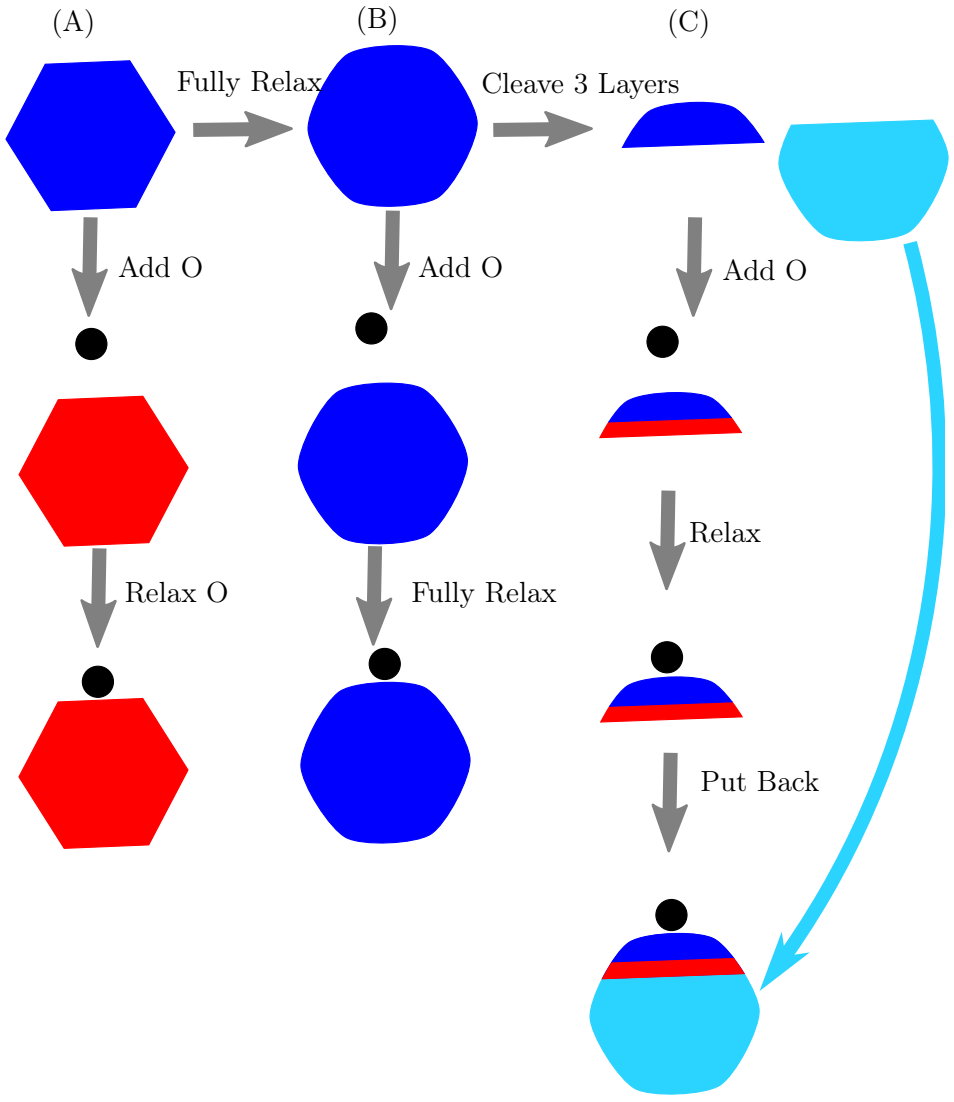


FIGURE 4.5: The three procedures which have been discussed are: A) To keep the nanoparticle constrained to bulk coordinates, while relaxing only the oxygen atom, B) To relax the nanoparticle atoms, before relaxing all of the atoms of the system of oxygen plus nanoparticle, C) To cleave a three-layer piece of the fully relaxed, bare nanoparticle and to relax the oxygen atom and the top two layers of nanoparticle atoms in this fragment system, with the bottom layer constrained, before calculating the energy with these relaxed atoms put back into the full system. Red indicates constrained atoms, light blue indicates that the coordinates of these atoms are unchanged in both systems.

The different procedures for calculating binding energies are summarised in the flowchart shown in Figure 4.5. We have not used the method in figure 4.5(c), but intend to use this approach for calculating binding energies of monolayers of oxygen on single facets.

In the slab calculation with CASTEP, the binding energies were computed by fully relaxing the Pt atoms, except for the bottom (fifth) layer, with and without the O ligand.

We have also performed calculations with higher than single atom coverage on nanoparticles, with one oxygen atom on each (111) facet, for a total of 8 oxygen atoms in total; and also half-monolayer coverage on a hemisphere of the nanoparticles. For these calculations we have used the PAW approach and kept the Pt atoms fixed at bulk-like positions. The oxygen atoms were then optimised on these constrained systems

$$E_{\mathrm{b}}^{8(\mathrm{O})} = E_{\mathrm{p+8(O)}}^{\mathrm{ORel}} - E_{\mathrm{p}}^{\mathrm{Const}} - \frac{1}{2}E_{\mathrm{O_2}}, \tag{4.7}$$

4.4 Results of Adsorption Calculations

Using this approach, we have attempted a study of how the binding energy of oxygen scales with the size of platinum nanoparticle and how these results compare with surface calculations performed in Castep. When comparing adsorption energies of the systems, we solely use procedure (B) from Figure 4.5 and the corresponding binding energy expression from Equation 4.7.

When performing the structural relaxation of the nanoparticle-oxygen systems, the initial position of the oxygen atom on the nanoparticles was chosen to be in the HCP hollow site, 2Åabove the plane formed by the three nearest-neighbour Pt atoms. The particular hollow site was chosen to be the most central one on the facet. On the nanoparticles with a hollow site at the centroid of the triangular (111) facet, this was chosen, whereas on those with a top-site at the centroid, a site with the top-site at the corner of the hollow-site triangle was chosen.

We firstly performed all our calculations with the Rappe-Bennett norm-conserving pseudopotentials. We used the rigid method (a) from figure 4.5 to give the dataset presented in red in figure 4.6. This was obviously unsatisfactory, so we attempted to fully optimise all of the atoms in the full set as in method (b) from figure 4.5. This gives the much more satisfactory dataset shown in black in figure 4.6, which appears to converge towards the Castep ultrasoft slab calculation, as we would expect for systems of increasing size.

At this stage, we also wanted to run PAW calculations on the same systems because we were concerned about the transferability of the norm conserving pseudopotentials we used in these calculations, i.e. a norm conserving pseudopotential optimised for the bulk may not be transferrable to the facets and other low coordinated sites of the nanoparticles. For this reason we opted to run the calcultions again, but to save on computation resources and time we chose to run with the nanoparticle geometries we had obtained for the norm conserving calculations and only optimise the oxygen atoms on these. Unfortunately, this resulted in an energy shift of about 0.5 eV, due to strain

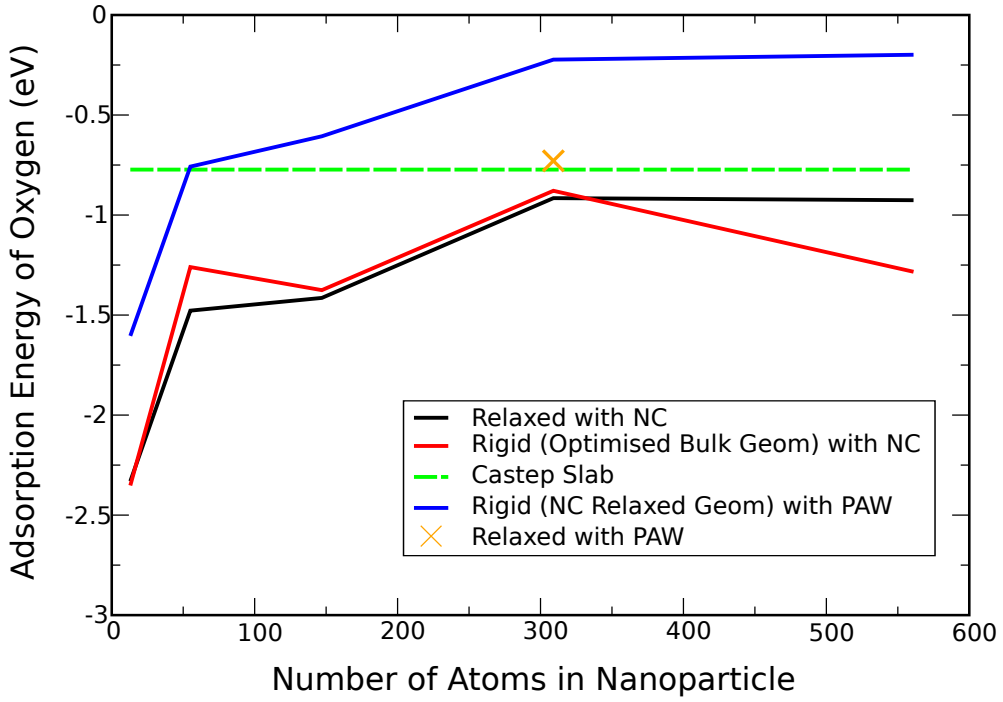


FIGURE 4.6: The binding energies of a single oxygen atom on the HCP sites of (111) facets on cuboctahedral, platinum nanoparticles of increasing size. We have computed three sets. An electronically optimised set of norm-conserving results with nanoparticles with Pt atoms in bulk lattice-like positions, where only the oxygen atom has been geometry optimised (red). A fully (structurally and electronically) optimised, norm-conserving set (black), where every atom has been optimised: it is clear that this is necessary for reliable adsorption energies if we compare with the Castep slab limit (green). Finally we have performed PAW calculations. Because of a lack of computer resources, we have computed these with the norm conserving geometries and only optimised the oxygen atom. This appears to be inadequate. If we fully optimise every atom, as we have done in the case of Pt_{309} (orange), we can see that PAW already almost matches the ultrasoft pseudopotential, slab calculation in Castep even at 309 atoms. We intend to do the full set with PAW and full geometry optimisation as further work.

on the nanoparticle as a result of the GBRV PAW and Rappe-Bennett pseudopotential datasets having different platinum nearest neighbour distances.

The full set of norm conserving results is shown in table 4.4 and in table 4.5 for PAW, which gives the binding energies of oxygen on cuboctahedral nano-particles of magic numbers 13 to 561, and the binding energy trend is shown in figure 4.6.

We have performed the Pt_{309} calculation with PAW and full geometry optimisation, so we can see that if we use the GBRV PAW dataset and optimise fully, this system lies almost exactly on the Castep slab result (see figure 4.7. If we take the data for the constrained PAW calculations and shift it so that the Pt_{309} result lies on the fully relaxed

TABLE 4.4: Properties of the O-nanoparticle systems which we have studied. The numbers include charges on atoms, as calculated using the DDEC/c3 method [1]. All nanoparticle calculations were performed using EDFT in ONETEP using the parameters explained in the text. The slab calculations were performed using Castep and equivalent parameters.

Property	Number of Atoms				
	13	55	147	309	561
Diameter (Å)	5.34	10.67	16.11	21.57	27.03
Relaxed					
O Binding Energy (eV)	-2.33	-1.48	-1.42	-0.92	-0.93
O-Pt nearest neighbour distance (Å)	2.01	2.10	2.10	2.06	2.05
Pt-Pt average distance on binding site (Å)	2.67	2.80	2.75	2.74	2.86
Nearest neighbour platinum average charge (e)	0.56	0.46	0.46	0.46	0.39
Charge on O Atom (e)	-0.20	-0.26	-0.26	-0.19	-0.21
Constrained					
O Binding Energy (eV)	-2.35	-1.26	-1.38	-0.88	-1.28
O-Pt nearest neighbour distance (Å)	2.00	2.03	2.05	2.04	2.01
Pt-Pt average distance on binding site (Å)	2.80	2.80	2.80	2.80	2.80
Nearest neighbour platinum average charge (e)	0.49	0.44	0.45	0.50	0.42
Charge on O Atom (e)	-0.18	-0.17	-0.18	-0.17	-0.16
Relaxed 5 layer, 4 site Slab					
O Binding Energy (eV)			-0.78		
O-Pt nearest neighbour distance (Å)			2.04		
Pt-Pt average distance on binding site (Å)			2.87		

result, we can see a trend the we would expect, flattening out after Pt_{309} towards Pt_{561}. We think that the argument that these constrained PAW calculations are under lattice strain is compelling and would like to run the full set with fully geometry optimised atomic coordinates. Unfortunately, due to a lack of time and computational resource constraints, we were unable to complete these calculations for the submission of this thesis, but we intend to run these calculations.

As might be expected, as the size of the nanoparticle increases, several properties of the nanoparticle-O systems begin to converge to the slab limit, including the oxygen binding energy and geometric properties of the Pt-O system. This is similar to the result reported by Li *et al* [141], but on the HCP rather than FCC site on the cuboctahedral (111) facets. Li *et al* show that the FCC sites converge to the slab limit after 147 atoms, whereas we find that for HCP sites this convergence does not happen until after 309 atoms. We chose to study the HCP rather than FCC site partly because of this pre-existing study by the Nørskov group, but also because the HCP sites are closer to the

TABLE 4.5: Properties of the O-nanoparticle systems which we have studied. The numbers include charges on atoms, as calculated using the DDEC/c3 method [1]. All calculations were performed using EDFT in ONETEP using the PAW approach on the (111) facets.

Property	Number of Atoms				
	13	55	147	309	561
Diameter (Å)	5.34	10.67	16.11	21.57	27.03
Relaxed O					
O Binding Energy (eV)	-1.60	-0.76	-0.61	-0.22	-0.20
Relaxed 8(O)					
O Binding Energy (eV)	-1.73	-0.77	-0.64	-0.24	-0.47
Relaxed 5 layer, 4 site Slab 5x5 Super-cell					
O Binding Energy (eV)			-0.62		
O-Pt nearest neighbour distance (Å)			2.04		
Pt-Pt average distance on binding site (Å)			2.87		

TABLE 4.6: Properties of the O-nanoparticle systems which we have studied. All calculations were performed using EDFT in ONETEP using the PAW approach on the (100) facets.

Property	Number of Atoms				
	55	147	309	561	923
Diameter (Å)	10.67	16.11	21.57	27.03	
Relaxed O					
O Binding Energy (eV)	-1.15	-0.82	-0.91	-1.04	-1.13

catalytic optimum binding strength of -0.52 eV (which is discussed in chapter 5).

We have also followed the pDOS procedure which was outlined in chapter 3 has been applied to the sequence of cuboctahedral, platinum nanoparticles. The results show that the systems are dominated by the d-band with relatively little but significant s-band character, but almost no p-band intensity, as one might except given the 9 5d electrons and 1 6s electron in the platinum valence states.

The systems also lose the single peaks of isolated systems and the DOSs become more and more band-like as the size of the nanoparticle increases (see figure 4.8). These pDOSs have also been projected onto the local set of angular momentum states *only* on the set of {111} facets, so that we can observe the trend in this (111) facet *local*-pDOS with single oxygen binding strength. The way we will draw a correlation in adsorption strength follows Nørskov *et al* closely [79], in that we calculate the d-band centre (energy weighted average of the d projected DOS) to correlate with. This is also plotted in figure 4.8 and shows convergence towards the slab limit (which was calculated using Castep).

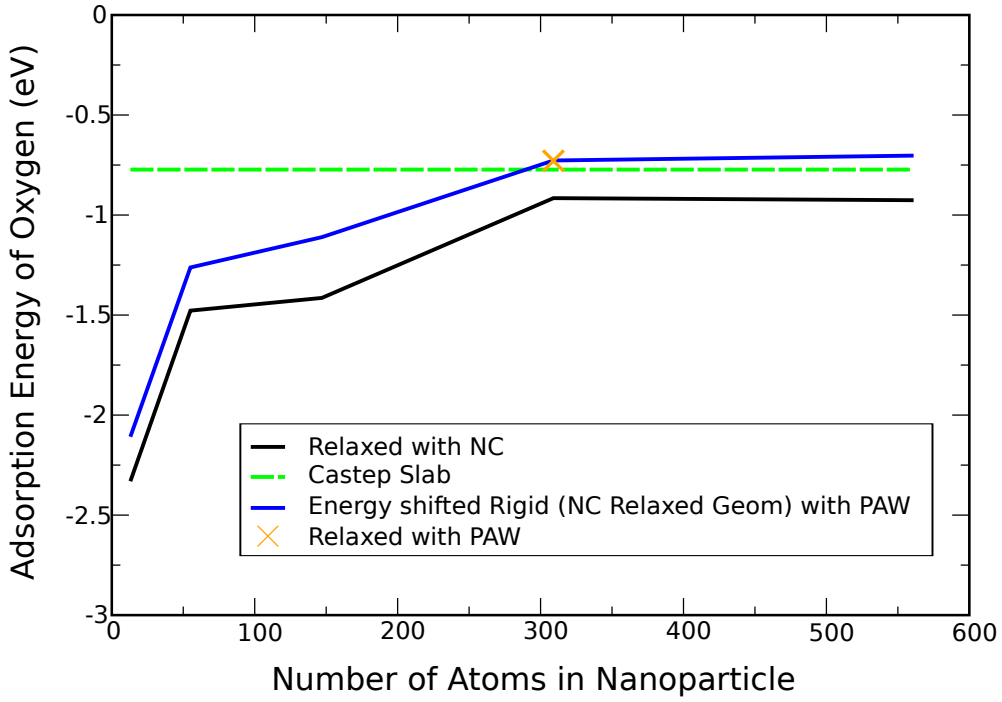


FIGURE 4.7: In this figure, we show the effect of shifting the PAW, Oxygen adsorption set shown in figure 4.6 so that the Pt_{309} result lies on the fully optimised Pt_{309} PAW result.

Plotting the local pDOS results against the oxygen adsorption strength on the (111) facets of the series of platinum cuboctahedral nanoparticles and examining the trend we can see the clear linear trend in (PAW) binding energy with d-band centre as we would expect (figure 4.9). This provides evidence (together with the literature) that d-band centre can be used as a good desciptor for binding strength as we have shown in chapter 5.

4.5 Oxygen Coverage Results

It can be seen from the results in the PAW dataset 4.5 having one oxygen atom per facet makes little difference to the binding energies with respect to having one oxygen atom per nanoparticle. This dataset also shows a jump in energy on Pt_{561} with eight oxygens. We also explored the effect of adding more than one oxygen atom per facet. For these calculations we opted to constrain platinum atoms, as the extra computational resource and time required to fully relax the degrees of freedom of tens of oxygen atoms together with the platinums was unfortunately well beyond what we had available to us. Because of this, we expect that an additive strain energy term will be present in our binding

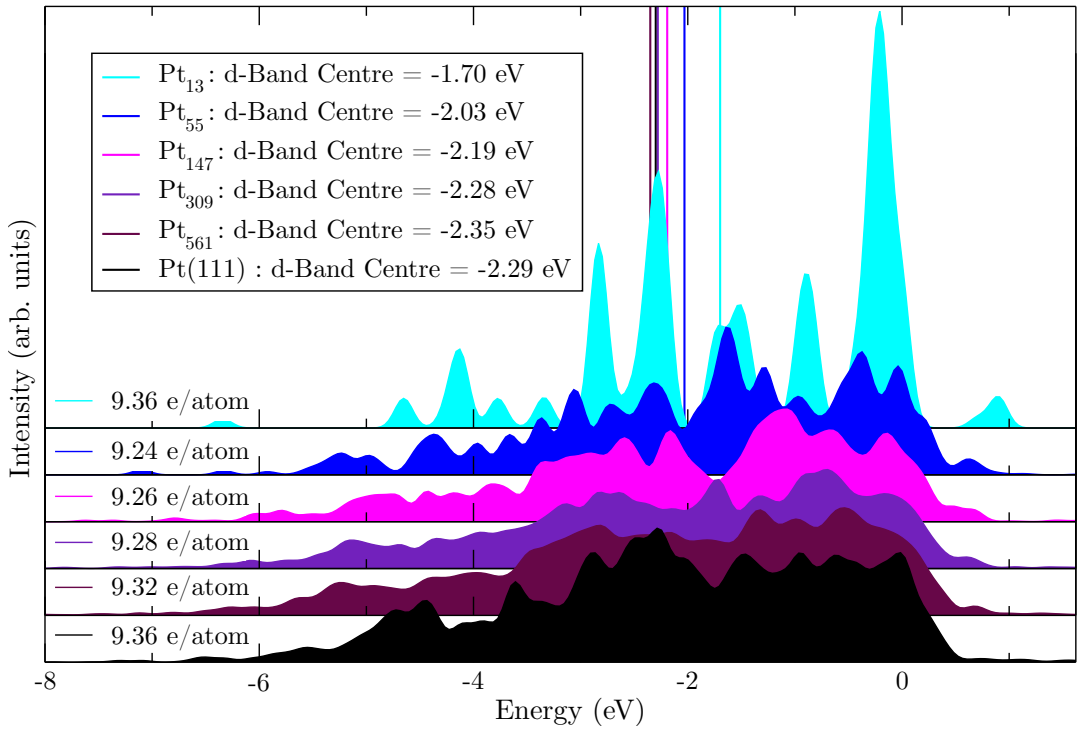


FIGURE 4.8: The angular momentum projected density of states on a (111) facet of cuboctahedral Pt_{13} to Pt_{561}, in the D-band shows that the states become more continuous ('band-like') with increasing size and from the D-band centres, that the D-band shifts away from the Fermi level. This corresponds to a increase in the filling of the (d-σ)* state upon adsorption of a ligand and should result in weaker binding. In these plots, the chemical potentials are at 0 eV.

energies similar to what we have already observed with the single oxygen atom adsorption study results seen in figure 4.6.

We have also chosen to add oxygen atoms to only one hemisphere of the nanoparticles. The justification for this is that in experimentally measured nanoparticle geometries, the nanoparticle can be embedded in the support. When embedded in such a way, the hemisphere of the nanoparticle which faces the support is not exposed and cannot take part in catalytic reactions. The initial coverage that was chosen for these calculations was a $\sim 1/2$ monolayer, where we covered (111) hollow sites and (100) bridge sites, and no other sites. We ran the calculations without spin polarisation, because we did not expect any oxygen molecules to form and the simulation cell contained an even number of electrons.

In practice, however, the calculations resulted in some of the oxygen atoms coming together and forming oxygen molecules. Any molecules that form must be in a singlet spin state because the calculations are not spin polarised. Despite this, we expect that if the spin state was unconstrained, then the molecules would still form, as the triplet state is lower in energy.

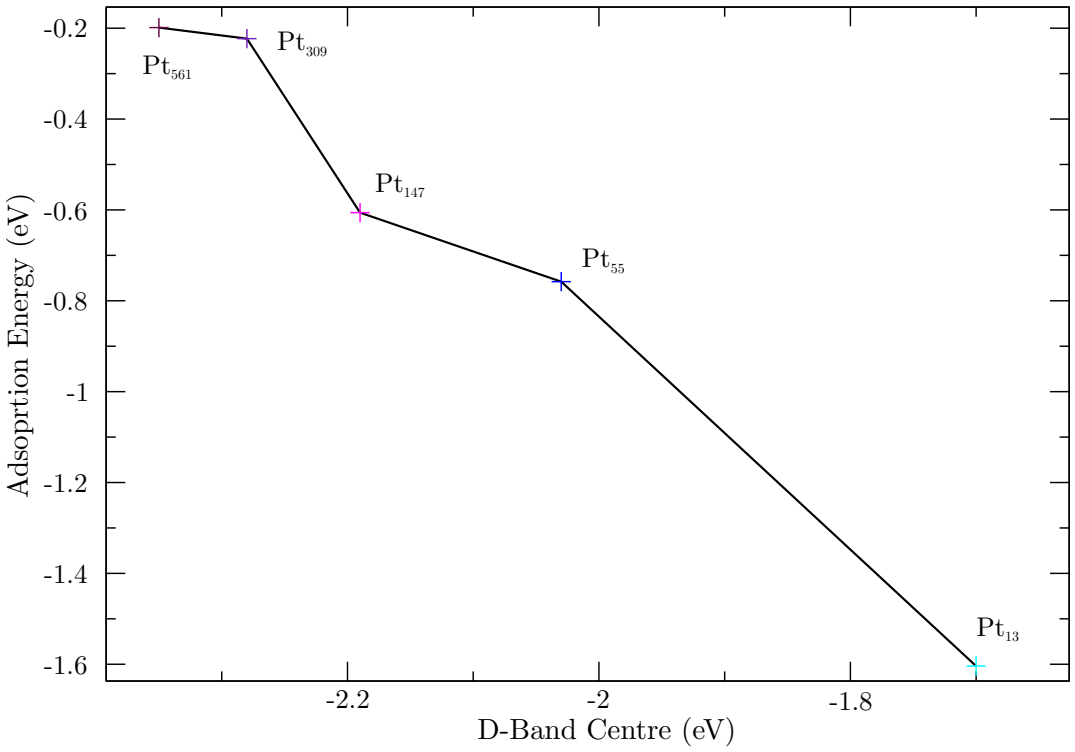


FIGURE 4.9: It is clear that with increasing nanoparticle size, that the adsorption strength of oxygen diminishes, and correlates well with the D-band centre, as expected from figure 4.8. Oxygen adsorption strength was calculated with respect to isolated, optimised, triplet O_2 and the D-band centres are all with respect to the chemical potential of the corresponding Pt nanoparticle system.

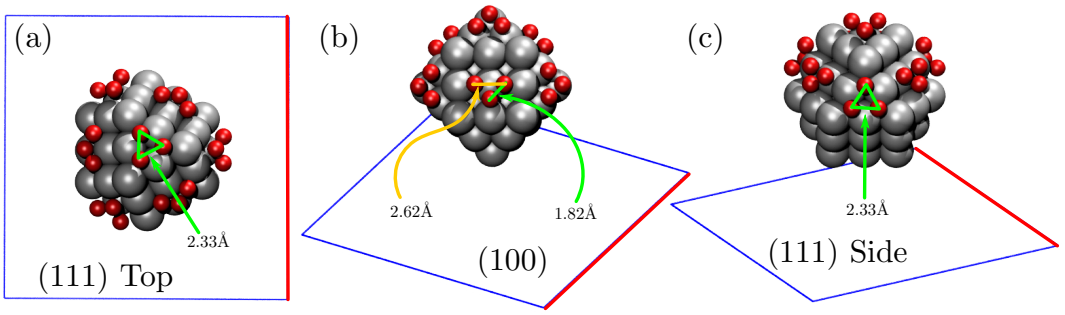


FIGURE 4.10: Three views of the initial geometry of the Pt_{55} nanoparticle with half monolayer, hemispherical oxygen coverage. Panel (a) shows the oxygen atoms on the HCP sites of the top (111) facet. Panel (b) shows the oxygens on three of the four bridge sites of the (100) facets and Panel (c) shows the top-like arragement of the other (111) facets.

In the initial configuration, the Pt_{55} cuboctahedral nanoparticle with half monolayer, hemispherical coverage has three oxygen atoms per facet. (111) facets have one oxygen atom per HCP site, (100) facets have an oxygen atom on three of the four central bridge sites. In figure 4.10(a), the top view shows that the oxygen atom separation distance is 2.33 Å. This is the same on each of the three other (111) facets, shown in figure 4.10(c).

Figure4.10(b) shows that the bridge sites which lie on opposite sides of the central (100) atom are spaced 2.62 Å apart, whereas, the bridge site which is furthest from the top facet is only 1.82 Å away from each of these oxygen atoms. The blue square with a red edge shows the face of the simulation cell which is parallel with, and furthest from the top facet. The edge that is coloured red is the same in each panel of figure 4.10 as well as those of figures 4.11 and 4.12.

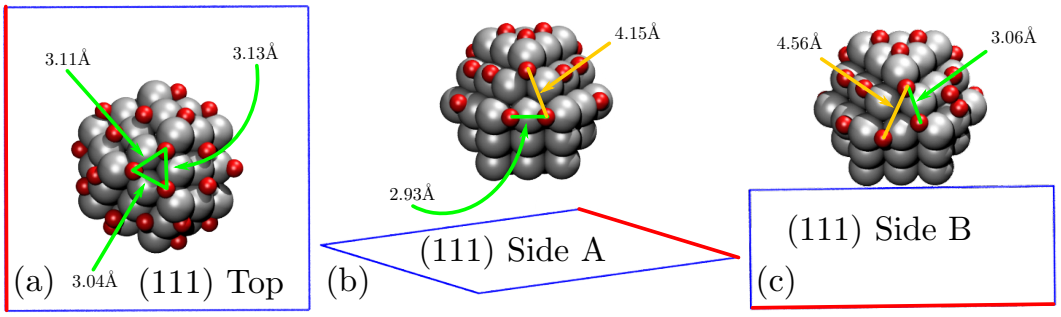


FIGURE 4.11: Three views of the relaxed geometry of the (111) facets of the Pt_{55} nanoparticle with half monolayer, hemispherical oxygen coverage. Panel (a) shows the oxygen atoms on the HCP sites of the top (111) facet. Panel (b) shows the oxygens on two of the other (111) facets, where two HCP oxygens have migrated onto edge bridge sites. Panel (c) shows the final configuration where only one oxygen atom migrated.

After optimisation of the Pt_{55} cuboctahedral nanoparticle with half monolayer, hemispherical coverage, the oxygen atoms on the top (111) facet have repelled slightly with respect to the initial configuration (see figure 4.10(a)). The oxygens remain in the HCP hollow sites, however, in a configuration just off hollow site centre, having moved approximately 0.7 Å away from the central atom of the facet towards the vertices (figure 4.11(a)). The two of the remaining three (111) facets have configuration as in figure4.11(b), where the atom closest to the top facet remains in the HCP hollow site. The remaining two atoms have moved onto the edge (111)–(100) bridge sites furthest away from the top facet and are separated from each other by 2.93 Å. None of the atoms on other facets are closer than this to any of the atoms on these facets. Finally, the fourth (111) facet also has two edge sites and one HCP hollow site, but instead, the remaining HCP hollow site is one of those furthest from the top and the oxygen which was on the site closest to the top has migrated onto the edge site.

In contrast with the (111) facets, the (100) facets show three distinct arrangements after optimisation of the Pt_{55} cuboctahedral nanoparticle with half monolayer, hemispherical coverage. Two facets (shown in figure 4.12 panels a and b) show two of the oxygen atoms remaining on (100) bridge sites. These are the central bridge sites which are on a line parallel with the top facet. These oxygens have moved slightly further apart than the initial structure (from 2.62 Å to 2.88 Å) and one of the oxygens on the third configuration remains in one of these sites. In figure 4.12(a), we see that the third atom has migrated

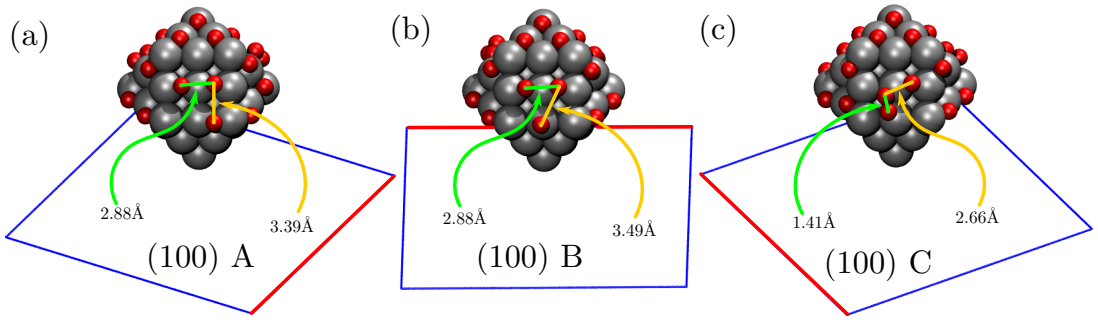


FIGURE 4.12: Three views of the relaxed geometry of the (100) facets of the Pt_{55} nanoparticle with half monolayer, hemispherical oxygen coverage. Panel (a) shows the case after relaxation where there are two occupied bridge sites and one edge-bridge site, panel (b) shows the case where we get two bridge sites and an atop site and panel (c) shows the case where there is one occuped bridge site and one singlet oxygen molecule straddling two bridge sites.

onto the edge site furthest away from the top facet, whereas in figure 4.12(b), the oxygen sits on the atop site of the bottom edge of the facet. Figure 4.12(c), on the other hand show that the two remaining oxygen atoms have formed a molecule, which must be in a singlet state because these calculations are not spin polarised. The molecule lies parallel to the facet and each atom sits just off each of the initial bridge sites. The molecule has a bond distance of 1.41 Å , which is slightly more than the 1.38Å we see with isolated oxygen molecules in the triplet spin state with the same pseudopotentials and simulation settings.

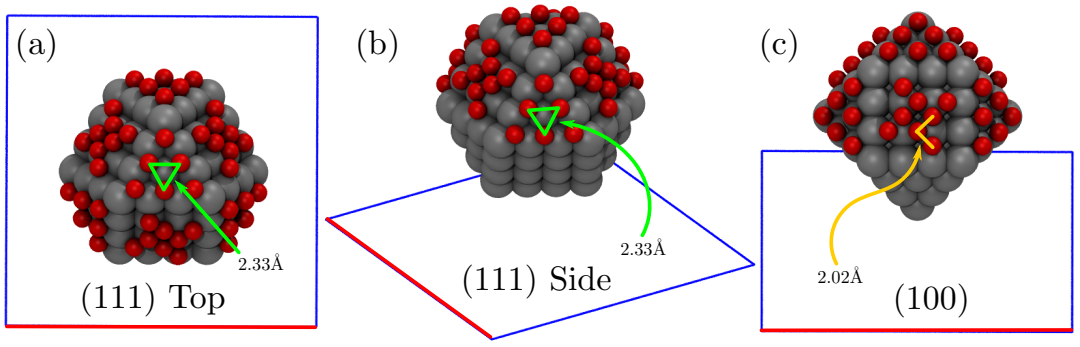


FIGURE 4.13: The initial geometry (before relaxation) of the Pt_{147} and hemispherical oxygen coverage. Each alike facet has exactly the same initial arragement of oxygen atoms, with 6 in the HCP hollow sites on the (111) facets shown in panels (a) and (b). The (100) facets, shown in panel (c) have 8 oxygens on the bridge sites.

Our first attempt to do half-monolayer, hemispherical oxygen packing on Pt_{147} resulted in 6 oxygen atoms per each of the four (111) facets, and 8 oxygen atoms per each of the three (100) facets. The oxygens on the (111) facets are positioned in the HCP hollow sites and the oxygens on the (100) facets are in the central bridge sites, avoiding bridge sites on the edges. This coverage is more than 50% depending on the counting method. On the (100) facets, there are 12 possible non-edge, bridge sites and we have occupied 8 of them.

On the (111) facets, there are 9 possible hollow sites and we have occupied 6 of them. This picture changes somewhat if edge bridge sites are included in the decision: this adds another 12 sites to the (100) total. So, we either have $(4\times6+3\times8)/(4\times9+3\times12) = 2/3$ or we have $(4\times6+3\times8)/(4\times9+3\times24) = 4/9$, which bracket 1/2 and so this seemed like a good starting point. We positioned the atoms at the centres of the HCP hollow sites, 2Å along the outward facing normals to the planes intersecting the atomic centres on the (111) facets, for oxygen neighbour-neighbour distances of 2.33Å on facet (figure 4.13 panels a and b). On the (100) facets, the oxygen atom nearest neighbour distance was 2.02Å(see figure 4.13(c)). The oxygens on the (100) facets were also configured to be 2Å along the outward facing normals to the planes intersecting the platinum atomic centres.

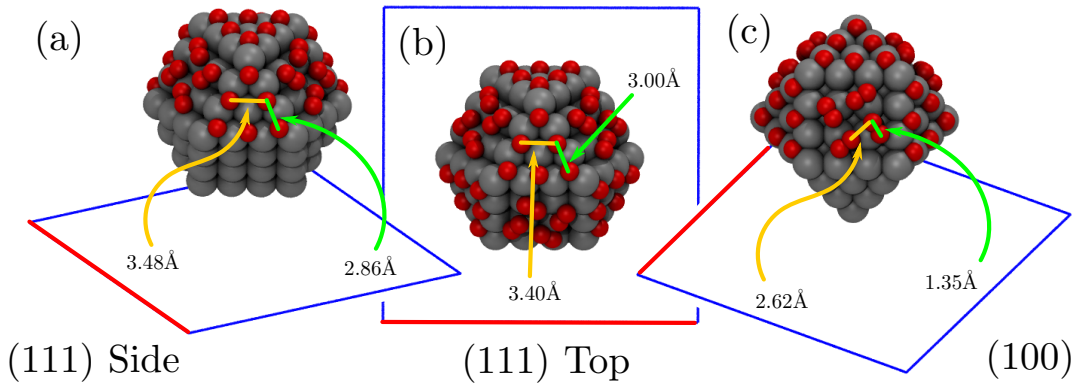


FIGURE 4.14: The relaxed oxygen geometry of the Pt_{147} and hemispherical oxygen coverage. Three of the 6 oxygen atoms migrate off the HCP and onto the edge sites on the (111) facets, see panels (a) and (b). Molecules form on the (100) facets (panel (c)), which project roughly on the the initial positions of the oxygens.

Unfortunately, upon geometry optimisation with ONETEP, we found that the oxygen atoms on the (100) facets formed molecules, with a bond length of 1.35 Å. The molecules remained over the bridge-bridge endpoints of the starting geometry, but one of the oxygen atoms of each molecule raised off the facet by 0.7Å, with the other remaining in its initial position. The nearest separation distance of oxygen molecules on the (100) facets was found to have become 2.62 Å, with the top-site energy barrier separating molecules (see figure 4.14(c)). As in the Pt_{55} case, these molecules must be in the singlet spin state, as we have not run spin-polarised calculations, due to our initial expectations. On the (111) facets the picture was closer to our expectation: the picture in both the top and three other (111) facets was almost identical, with the three oxygens closest to the nanoparticle vertices remaining in the HCP hollow sites and those on the sites close to edges migrating onto the edge (111)–(100) bridge sites. The separation between oxygen atoms on edge sites to HCP sites was found to be between 3Å on the top facets (see figure 4.14(b)) and 2.86Å on the other (111) facets (see figure 4.14(a)). As for the edge to edge separation,

these distances were approximately equal everywhere: 3.4Å on the top facet and 3.48Å on the others.

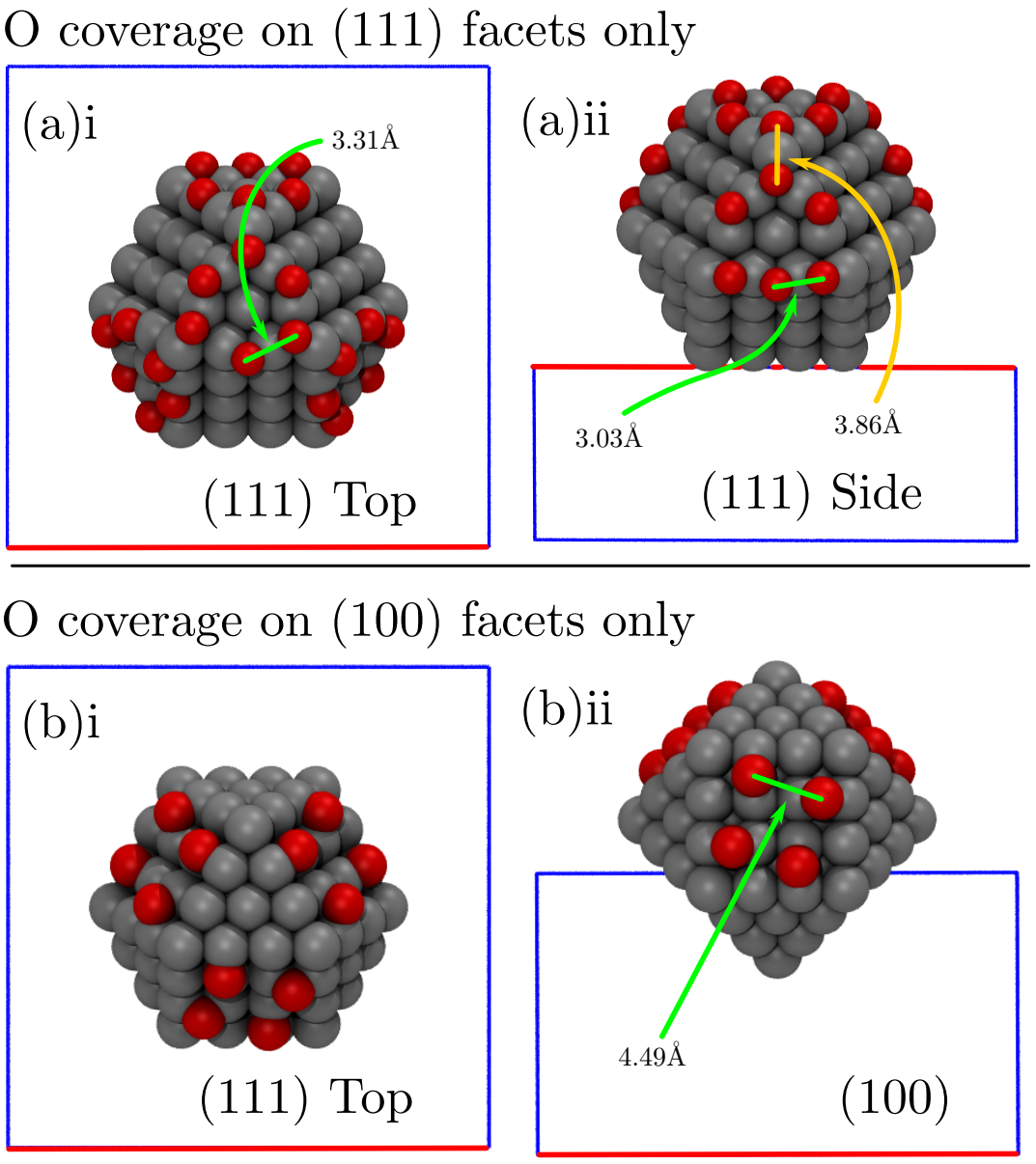


FIGURE 4.15: The two nanoparticles with coverage on only one type of facet. Shown above in the (a) panels is the (111) case and below, in the (b) panels is the (100) case. The initial oxygen positions were taken from the geometry in figure 4.14, with only the four oxygen atoms closest to the nanoparticle kept on the (100) facets, and the nanoparticles relaxed to give the structures shown here.

The challenge after noticing the over-packing on the (100) facets of the Pt_{147} nanoparticle with hemispherical approximately half-monolayer oxygen packing was to salvage something of these extremely expensive calculations. To optimise to convergence took 2 weeks on Archer. We decided to take the optimised geometry and separate the oxygen atoms onto two separate nanoparticles. One with only (111) facets covered with oxygen (shown

in figure 4.15(a)i and ii) and the other with only (100) facets covered with oxygen (shown in figure 4.15(b)i and ii). The nanoparticle with (111) oxygen packing had initial oxygen geometry and the same packing as the fully covered nanoparticle in figure 4.14 as no molecules formed. After optimisation, some of these oxygen atoms did move off the initial sites, on the non-top (111) facets. Two of the oxygens initially on HCP sites furthest from the top facet migrated onto edge sites, leaving only a single oxygen on the initial HCP site closest to the top site. This atom ended up 3.86Å away from its cousin on the top facet (figure 4.15(a)ii). On the (100) facets, we opted to keep the oxygen atoms that remained on bridge sites after the relaxation shown in figure 4.14 and remove the atoms at the other end of each molecule leaving a total of 4 atoms per (100) facet, which are separated by 4.49Å (see figure 4.15(b)ii) after relaxation of this configuration. If we add these two configurations together, then we either have $(4\times6+3\times4)/(4\times9+3\times12) = 1/2$ in the first possible counting scheme or $(4 \times 6 + 3 \times 4)/(4 \times 9 + 3 \times 24) = 1/3$, in the other. Giving exactly one half monolayer coverage in the first counting scheme.

Once we had a set of approximately half-monolayer, hemispherical oxygen coverage Pt_{147} calculations between the (111)–only set and (100)–only set, we opted to combine them for a further round of geometry relaxation. This did not result in the formation of molecules, most likely because the separation between oxygens on adjacent (111) and (100) facets remained large (2.78Å see figure 4.16(b)). The oxygen atoms on the (100) facets did not move appreciably, perhaps owing to the isolation of the chemical environment of each oxygen– leading to a good approximation to this combined system with the (100)-only system. Likewise, the (111) facets changed only very little on geometry optimisation. One more of the oxygen atoms on an HCP site on the top facet moved into an edge bridge site, breaking the three-fold symmetry of the system (See figure 4.16(a)). The only noticable effect of relaxation on the other (111) facets was for the oxygen-oxygen nearest neighbour distane to decrease slighly from 3Å to 2.9Å on the bottom edge sites of the facets.

At first glance, it may seem strange to have molecules of oxygen forming on platinum nanoparticles which are used in industry as catalysts to reduce oxygen molecules. Firstly we should frame these adsorption studies as models and outline the important approximations that we have made. Real platinum nanoparticles for ORR are supported and possibly solvated, we have fixed the coordinates of the platinum atoms in the structural relaxation of the oxygen atoms and we are studying these adsorption effects on a nanoparticle which is smaller than the ∼2-10 nm size range used in real platinum cathode catalysts.

These simplifications may all play a role in this phenomenon, but we can narrow down what is likely to be the main contribution. We have run solvated calculations on Pt_{55} and Pt_{13} with single oxygen atom adsorption and saw negligible effect on the binding energy in these test calculations, whereas with a hydroxyl adsorbate, we saw a large

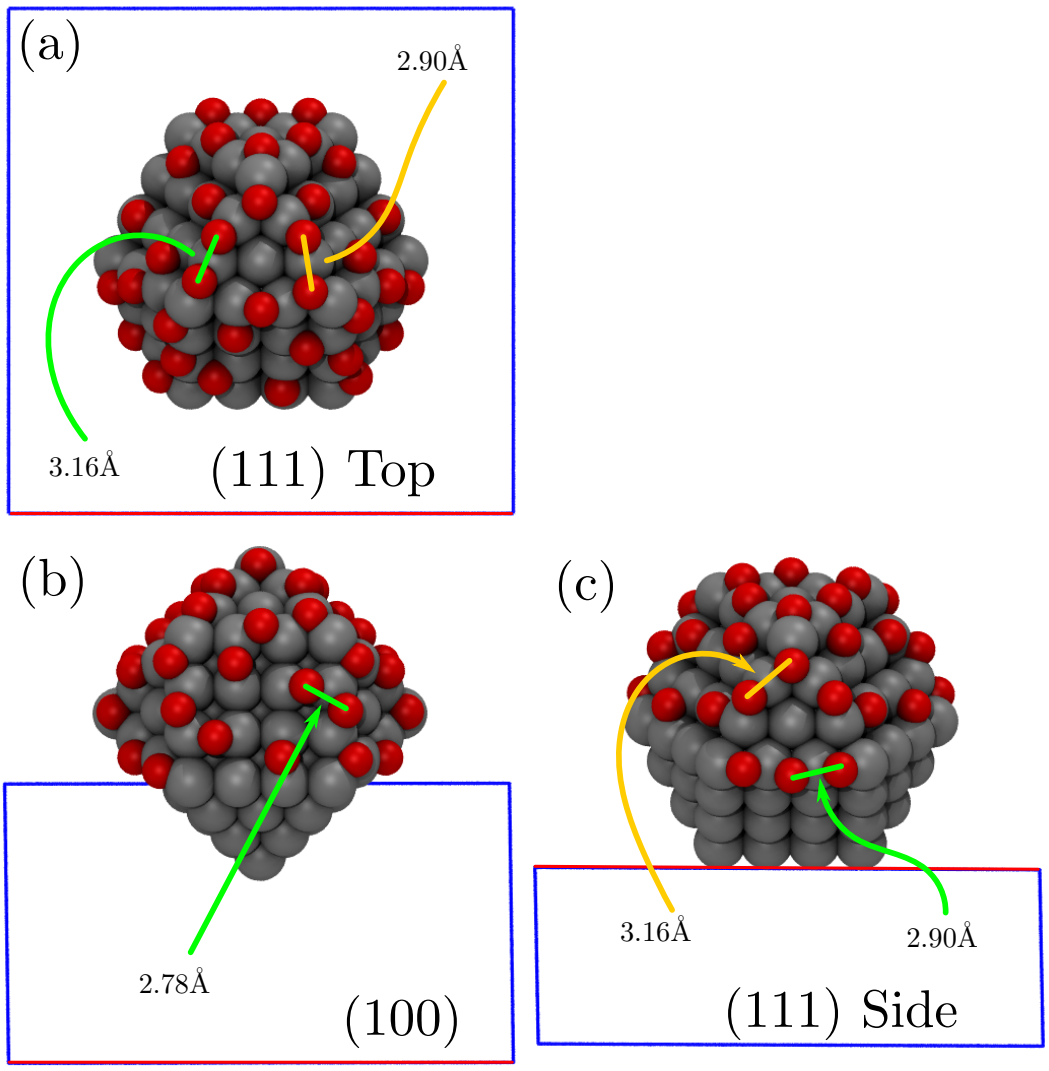


FIGURE 4.16: After combining the final, relaxed oxygen coordinates on the Pt_{147} nanoparticle from figure 4.15, the result was geometry optimised a further time to give the relaxed structure shown in this figure. There is little significant movement, apart from a symmetry breaking single HCP to edge bridge migration on the top (111) facet, shown in panel (a). There is a sligh tightening in the oxygen-oxygen neighbour distances on the other (111) facets (see panel (c)) and almost no change on the sparsely populated (100) facets shown in panel (b).

effect. Therefore it is likely that solvation plays only a very minor role for adsorption of ligands with zero or small dipole moment. A similar effect has been reported in the literature, particularly in the study by Rossmeisel *et al* [58].

We have also performed studies of adsorption of oxygen on graphene supported nanoparticles[12] and found that the electronic density difference between supported and non supported nanoparticles penetrate to a significant degree (> 0.01 eÅ^{-3}) to 3 or 4 layers. For Pt_{147} this is not deep enough to reach the point of adsorption of oxygen on these systems. Of course, with a hemispherically embedded nanoparticle, the situation could be different,

but we cannot speculate on that given that we have not yet done the calculations.
The smallness of the nanoparticle probably diminishes rather than enhances the chances of forming oxygen, as we have shown that oxygen binds more strongly with decreasing size. It is therefore less likely than with larger nanoparticles that the oxygen atoms will have the mobility to form oxygen molecules.
This leaves two contenders, constrained platinum positions in the relaxation, as mentioned, but also related to this: that the oxygen atoms were placed by hand in positions which may not be physical. To find the optimal (initial) packing of oxygen on the surface is a combinatorial problem and well beyond the scope of this project, so we had to make assumptions based on where single oxygen atoms would bind as to where fractions of monolayers would bind. It is known from cyclic voltammetry experiments that at high oxygen concentration oxygen tends to go subsurface in platinum [177]. Schnadt *et al* show that this place exchange mechanism can produce a complicated set of subsurface oxygen structures on silver [178] based on a combined experimental and DFT study. Gu *et al* show a similar picture, with many possible structures on platinum slabs in a purely DFT based study [179]. Because of this it is possible that by fixing the platinum through the geometry relaxation and hence not allowing the oxygen to go subsurface, that when we introduce unphysical concentrations at the surface that it is forced into forming molecules.

TABLE 4.7: The binding strength of oxygen per atom on the Pt_{147} nanoparticle with half monolayer hemispherical coverage.

Facets	(100)	(111)	(111) + (100)
Number of oxygen atoms	12	24	36
E_B (eV)	-0.860	-0.948	-0.759

For multiple oxygen coverage on the Pt_{147} nanoparticle, the binding energy per oxygen atom is stronger on the (111) facets than on the (100) facets. The combined (100) and (111) coverage calculations result in oxygen atoms that are bound even more weakly than this. We cannot directly compare these calculations to the fully relaxed Pt_{309} single-oxygen binding calculations in chapter5 because these coverage calculations have been performed with constrained, non-relaxed nanoparticle geometries, although the relative binding strengths are likely to be similar We can, however, directly compare with the results in tables 4.5 and 4.6, which show that single oxygen atoms on (100) bridge sites are bound more strongly than single oxygens on (111) HCP hollow sites.
We can see that the per-oxygen binding strength of oxygens on (100) bridge sites is only 0.04 eV different from that of single oxygen binding stength on (100) bridge sites. This is probably because the oxygen atoms are 4.49 Å apart (see figure 4.15) and so they experience a similar local chemical environment to the single oxygen case. As for the

half monolayer coverage on the (111) facets only, these oxygen atoms are separated by only 3 Å and so they are more likely to interact as this separation is only slightly more than double the O_2 bond distance we calculate with ONETEP. The increase in binding strength does not appear to be a result of this interaction, however, as the oxygen atoms are repelled from each other on the (111) facets (see figure 4.15). Instead, because the oxygen atoms move into edge-bridge sites, which have a much stronger relative binding strength than HCP (111) sites or on-facet (100) bridge sites, this is likely to be the reason for the 0.19 eV increase in per oxygen binding strength.

If we take this into account, although we do not have binding data of oxygen atoms on the edge sites of the Pt_{147} nanoparticle, we can compute relative quantities from data on Pt_{309} in table 5.2, to assess the plausability of this theory. On Pt_{309}, the binding strength of single oxygen on (100) bridge sites is -1.405 eV. On (111) HCP hollow sites it is -0.729 eV and on edge sites it is -1.763 eV. Using the fact that we have 18 edge sites occupied after relaxation of the oxygen positions and 6 remain in HCP hollow sites, we can compute a weighted average to predict the average binding strength from Pt_{309} numbers : $(18 \times (-1.763eV) + 6 \times (-0.729eV))/24 = -1.505eV$.

This number is not compariable with our calculations on Pt_{147} with hemispherical, half monolayer coverage, but ratios of energies on the same nanoparticle may be compariable with ratios of energies on a different nanoparticle. Taking the ratio of (100) bridge site binding with (111) HCP hollow binding, using numbers from Pt_{309}, we get $(-1.405eV)/(-0.729eV) = 1.93$, for example. If instead we use the weighted average binding, we get a relative binding of $(-1.405eV)/(-1.505eV) = 0.93$, with numbers from Pt_{309}. Taking ratio of the average (100) bridge site binding on Pt_{147} with the *calculated* binding strength of oxygen for the relaxed (111) HCP hollow system (where 18 oxygens have migrated to edges sites) we get $(-0.860eV)/(-0.948eV) = 0.91$. This implies that the effect of site migration of oxygen atoms has a much stronger effect on binding strength than interaction with other adsorbed oxygen atoms.

Interestingly, if we apply the same reasoning and analysis to the system with oxygens on (100) and (111) facets, the picture is different. The average binding strength based on Pt_{309} numbers comes out as $(12 \times (-1.405eV) + 19 \times (-1.763eV) + 5 \times (-0.729eV))/36 = -1.500eV$ and the strength relative to (100) with Pt_{309} numbers is $(-1.405eV)/(-1.500eV) = 0.94$. The strength relative to (100) with numbers from table 4.7, calculated on Pt_{147} is $(-0.860eV)/(-0.759eV) = 1.13$.

It is likely that this effect *is* related to the effect of coverage, because the simple weighted average does not give a similar ratio. Geometric evidence for this being the case can be seen in figure 4.16(b). The separation between oxygen atoms on the (100) bridge sites and on edge-bridge sites is only 2.78 Å, which is the closest we have witnessed after the relaxation of the geometries. Furthermore, these sites are both bridge sites with one common platinum atom on the (111)–(100) edge of the nanoparticle (see figure 4.17).

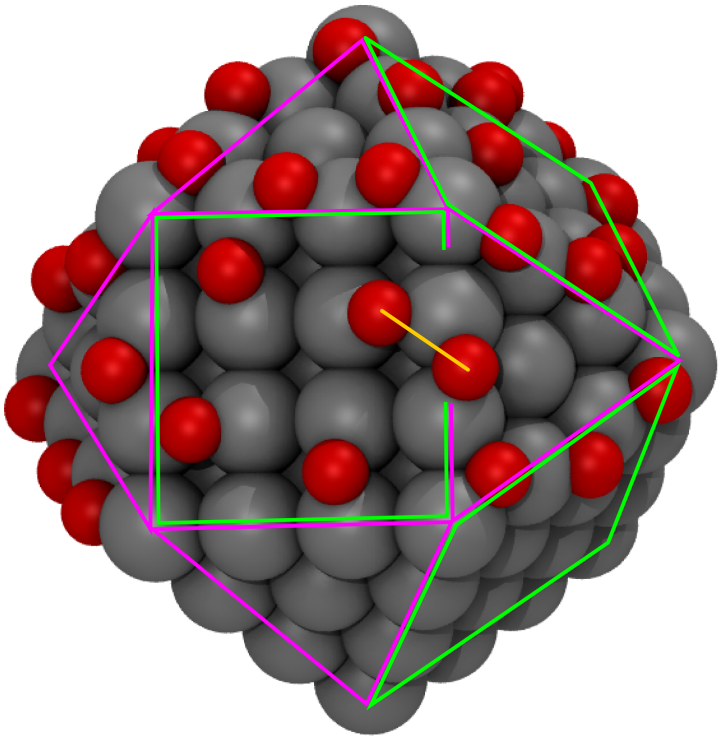


FIGURE 4.17: A closeup of panel (b) from figure 4.16, showing the bridge sites off the same edge atom occupied by the two oxygen atoms (highlighted with a yellow line), with respect to the facets highlighted in green (100) and pink (111).

In the original high coverage Pt_{147} nanoparticle in figure 4.14(c), the oxygen atoms which started out on bridge sites (albeit on facet bridge sites: see figure 4.13(c)) with similar geometry formed oxygen molecules. In this slightly different configuration, where one of these bridge sites is on the edge, with probably a higher oxygen adsorption strength (see table 5.2), and where there is a lower oxygen packing density, there may be a higher barrier to forming oxygen molecules. The oxygens, are however geometrically close enough to each other that they will interact in the calculation.

4.6 Conclusions

We have explored the binding of atomic oxygen to platinum nanoparticles of up to ∼1000 atoms using DFT calculations in ONETEP. We firstly explored the effect of binding single oxygen atoms per nanoparticle and increasing the size of the nanoparticle. This study was performed using the cuboctahedral sequence on nanoparticles cleaved from the platinum FCC lattice, so to increase in size we included another shell of platinum atoms on the cuboctahedron, advancing to the next cuboctahedral magic number. The binding strength vs cuboctahedral magic number was explored and we also investigated

the effect of popular descriptors for catalytic activity, such as the d-band centre approach (the implementation of this in ONETEP is described in chapter 3), which we correlate against the adsorption strength. We subsequently went on to explore the effects of oxygen coverage on the Pt_{55} and Pt_{147} cuboctahedral nanoparticles to attempt to advance further towards realistic conditions in our simulations.

Our study of oxygen adsorption strength with nanoparticle size indicated that as expected, the binding strength of oxygen decreases with nanoparticle size, converging on the limit of oxygen adsorption on the infinite platinum slab (as calculated with Castep in periodic boundary conditions). Due to the work we did on angular momentum projected density of states in ONETEP we were also able to correlate these adsorption strengths with d-band centre and obtained a linear trend as we expected based on the trends reported in the literature by Nørskov, *et al* [58, 79].

The investigation of oxygen coverage revealed that with the half monolayer, hemispherical oxygen coverage that we eventually settled on, oxygen tends to gravitate towards the edges and lower coordinated sites in the nanoparticle and away from the centres of facets. This effect correlates precisely with the site specific, single oxygen adsorption study on Pt_{309} and experimental platinum nanoparticles which has been documented in chapter 5.

The difficulty we faced in trying to find an oxygen coverage regime on the Pt_{147} nanoparticle without oxygen molecules forming turned out to give further insight into oxygen coverage mechanics. The O_2 molecules form due to the constraints we impose on the platinum atoms during the geometry relaxation, disallowing subsurface migration of the oxygen and too high an initial packing density of oxygen on the surface. The breaking up of the relaxation in two: firstly a nanoparticle with oxygen solely on the (111) facets, then the (100) facets, before combining the resulting oxygen coordinates onto a nanoparticle with half monolayer, hemispherical oxygen coverage allowed us to see the energetic effect of various binding types in finer resolution than we would have been able otherwise. The effect of oxygen site migration appears to have the greater effect on per oxygen binding strengths, with respect to oxygen-oxygen interaction on a nanoparticle, particularly on the (111) facets. The (100) coverage was found to be low enough that very little change was observed in average oxygen binding strength vs binding strength of single oxygens to (100) bridge sites.

As further work we intend to run single facet oxygen coverage calculations using the method described in figure 4.5(C), perhaps up to the Pt_{561} nanoparticle. Before this, however, we will repeat the calculations in the former part of the work and run the single oxygen adsorption study with PAW, but with full oxygen and platinum geometry relaxation.

Chapter 5

Predicting the oxygen binding properties of platinum nanoparticle ensembles by combining high-precision electron microscopy & DFT

This chapter describes DFT and analysis work we have done in collaboration with the Nellist electron microscopy group from Oxford materials. All of the experimental work presented in this chapter was performed principally by Lewys Jones and Aakash Varambhia, of the Nellist group. This work was published in Nanoletters in 2017:

Jolyon Aarons, Lewys Jones, Aakash Varambhia, Katherine E MacArthur, Dogan Ozkaya, Misbah Sarwar, Chris-Kriton Skylaris, and Peter D Nellist. Predicting the oxygen-binding properties of platinum nanoparticle ensembles by combining high-precision electron microscopy and density functional theory. *Nano Letters*, 17(7):4003–4012, 2017[8].

Abstract

5.1 Background and Pre-existing Work

In order to make systematic improvements to the activity of catalysts and to predict new, better catalysts, we need ways to characterise the function of existing catalysts. By making changes to their composition and atomic arrangement, the new candidates can be tested using the same methods and potentially better catalysts can be found. Ideally, this would also be a high-throughput technique so that many candidate structures

could be probed, to find a good option more quickly. We opted to combine atomic resolution microscopy of heterogeneous, platinum fuel-cell catalysts with large-scale DFT calculations. This allows us to probe chemical properties and reaction desciptors at the DFT level of theory for real, experimentally observed nanoparticle structures and relate how these structures affect the applicability of the nanoparticles for their intended purpose of catalysing reactions such as the oxygen reduction reaction at the cathode of a proton-exchange-membrane fuel cell. This combination is ideal because we want to run DFT on more than just model nanoparticles and we want to get more properties and predictions out of the electron microscopy than just structures.

Real-world heterogeneous catalysts exhibit a wide range of active morphologies. To understand how activity is related to structure, we can aim to first characterise an ensemble of real nanoparticles to explain how the activity of this subset is distributed.

In terms of taking experimental measurements, there are several possible options. X-Ray Adsorption Fine Structure (EXAFS) measurements are one possiblity which has been used in previous work to provide illumination of the chemical environments present in catalytic materials[180]. Such techniques provide inherently collective measurements, so while large temporally-resolved[181] and spatially-resolved[182] studies of platinum and platinum-based alloy fuel-cell cathode catalysts exist, they do not provide insight into how the morphology of individual nanoparticles affects catalytic activity.

To resolve the structures of individual nanoparticles in 3D, the Nellist group at Oxford University Materials performed high-throughput electron-microscopy measurements as the experimental part of our collaboration. Application of this technique provided an ensemble of real nanoparticles, measured from the real world which can be fed into metals DFT methods in ONETEP meant for large numbers of atoms, to calculate the electronic structures of the nanoparticle ensembles. Based on these DFT calculations we were then able to parameterise an accurate and widely applicable descriptor of oxygen binding-energy which can be applied to the full ensemble of nanoparticles which were obtained by experimental means. This means that oxygen binding energy can be predicted with less computational cost than if full oxygen binding calculations were performed with DFT on every new experimentally measured nanoparticle. With such large ensembles of nanoparticles and binding energy trends, we are then able to make further predictions of which morphologies of nanoparticle approaches the optimal binding strengths.

The particular type of electron microscopy which the Nellist group used as part of this collaboration is high-resolution Annular Dark-Field Scanning Transmission Electron Microscopy (ADF-STEM). The idea behind STEM is that a thin, electron-transparent sample is placed between the electron gun and the detector. In bright-field imaging, the Electron Energy Loss Spectrum (EELS) is measured and allows the experimentalist to map a picture of the sample as the electron beam is scanned over the material. In order to improve on the Z-contrast of bright-field imaging, Annular Dark Field (ADF) imaging

can be used, where the electrons are not Bragg scattered, so unaffected by abberations in the focussing apparatus[183]. This allows sharper, atomically resolved pictures to be formed from samples of the catalysts, as projections along crystalographic azamuths. To resolve fully 3D structures using the electron microscope, or to do so called electron tomography would require very high electron doses, which would likely melt or move the metallic nanoparticles unless they are infeasibly well anchored to the support and/or stable under temperature[184–187]. What the Nellist group have done with ADF-STEM is to avoid tomography by aligning a crystalographic azamuth of the nanoparticle samples with the beam direction. The columns of atoms are then resolved giving a 2D projection of each nanoparticle[188, 189]. As the nanoparticles contain only a single atomic species (platinum) the intensities of the columns corresponds to the number of atoms contained within[190], and hence the technique is effectively 2.5D, giving the exact positions of the columns and the number of atoms within each[9]. The approach is robust to abberations in the electron probe [190–192], and requires solely a model for placing the columns of atoms with respect to each other in the reconstructed 3D coordinates.

To determine where each column lies with respect to each other, some assumptions are first made: the nanoparticles are made to have atoms only on FCC Pt lattice sites and hence to be perfect crystalline samples. Further to this end, the columns are assumed to be contiguous, so that no internal FCC lattice sites are allowed to be empty. By using these simplifications, the columns can be relaxed in a simple pair-potential model, keeping the atoms of the columns rigidly fixed with respect to all atoms in the same column. The Nellist group use a Lennard-Jones potential for this purpose [9]. The accuracy of this model has been verified through a master weighing curve (see Figure 5.3). Doing this reconstruction in software gives a particular advantage in throughput, in that compared with tomography and statistical parameterisation approaches[193–195] there is no need for reorienting the sample for separate measurements along different crystalographic azimuths. This reorientation is slow as it is labour intensive on the experimentalist.

Throughput is also an issue with DFT calculations. Cubically scaling plane-wave calculations have been used for many years to make accurate and successful predictions of reactions on metallic catalyst surfaces [80, 196–199]. The main limitation in these calculations is size. When modelling nanoparticles with such techniques, models are required to reduce the size of the simulation cells and number of atoms. For instance, facets may be modelled using periodic slab models with similar surface symmetry to the facet under consideration. Alternatively, if the nanoparticles are produced as models by cleaving shells of regular polyhedra (such as cuboctahedra, icosohedra or truncated octahedra) from the bulk lattice then very small (sub-)nanoparticles may be produced which might be run in full with conventional DFT methods. These models do not however encompass the wide range of surface geometries which we have witnessed in the ensembles of

nanoparticles measured experimentally by the Nellist group. For this reason, cutting-edge reduced scaling techniques for DFT such as those available in the ONETEP package [40] become very appealing with the promise to allow calculations on full (experimentally measured) nanoparticles in the important size range for heterogeneous catalysis of 2-10 nm for the first time.

What we want to be able to do is to perform reduced scaling DFT calculations on metallic systems with thousands of atoms with the same high level of accuracy as conventional cubically scaling plane-wave DFT. Unfortunately, such methods do not yet exist for metals, although linear-scaling FOE methods are a promising area of current research (see chapter 3 and [11]), they do not presently have the accuracy to give the chemical precision which we require. Low prefactor metals methods are, however, available today in ONETEP, enabling the calculation of properties on metallic nanoparticles of up to thousands of atoms[7]. These methods still rely on a cubically scaling diagonalization, but the cost is greatly reduced to the point that it does not feature significantly in relative timings until the regime of thousands of atoms.

With such techniques we are able, for the first time to run full quantum mechanically accurate calculations of not only experimentally measured nanoparticles, but also of large idealised, polyhedral structures to compare and contrast how deviation from regular, faceted nanoparticles affects catalytic activity and adsorbate adsorption strength. This is important because while a wealth of previous work has focussed on idealised structures [200–202], we know that experimental structures have significant surface disorder which may contribute to the inaccuracy of predictions made with faceted nanoparticle models in DFT.

In this collaboration between Southampton University and Oxford University, we have chosen the ORR as a benchmark to compare results obtained with experimentally measured and idealised structures. The reason for this was partly that this project was funded by Johnson-Matthey, who are particularly interested in improving the ORR for application at the cathode of proton exchange membrane fuel cells. In starting the collaboration, we hoped to be able to address, to some extent the relative catalytic activities of various platinum nanoparticle morphologies, and assess this through the binding energies of ORR intermediates, such as atomic oxygen on sites on their surfaces.

Catalytic activity is affected through the Sabatier principle by the binding strength of ORR intermediates[80]. To maximise activity, the catalyst should bind O_2 sufficiently strongly to facilitate its catalytic bond breaking, but also the intermediates should be weakly bound enough that they can dissociate and not block sites after reacting. Shao *et al* use the binding strength of oxygen as a descriptor for activity[201], and we have chosen to take the same approach. Previous work based on extrapolations from platinum slab DFT suggest that the optimal binding strength of oxygen for maximum catalytic

activity should be between 0.2 eV [76, 80] and 0.25 eV [58] weaker than the binding energy of oxygen in the hollow site of a (111) platinum slab.

5.2 Experimental Measurements at Oxford Materials

The experimental work in this collaboration was performed by the Nellist group at Oxford University Materials using a JEOL 3000F STEM machine, fitted with a Fischione 3000 ADF detector. The samples were commerically produced ORR catalysts of platinum nanoparticles on carbon-black support, manufactured by Johnson-Matthey for PEM fuel-cells. The Nellist group prepared the commercial platinum nanoparticle sample by grinding, followed by sonication for 30min in 99.99% pure ethanol, before dropping it onto a holey carbon support grid. The loaded grid was baked for 12 hours at 80°C using a heating element fitted to a Fischione model 1020 plasma cleaner, where the unit was only used for heating, not plasma cleaning. The distribution of nanoparticle sizes can be seen in figure 5.3 or at low magnification in figure 5.1.

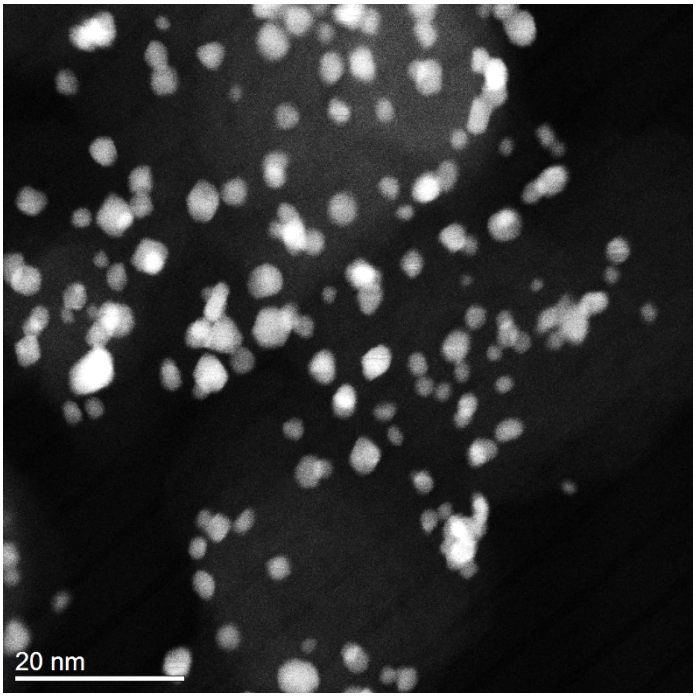


FIGURE 5.1: Example of the low-magnification images used for the sizing analysis. Reproduced from [8].

The resolution of the images in figure 5.1 is not sufficient to resolve individual atomic columns but is sufficient for nanoparticle sizing. Several tens of images were recorded and automated contrast thresholding was used to identify the nanoparticle sizes. A total of 1342 nanoparticles were sized for the population analysis. In addition to the size histogram, HR-STEM images were recorded over as wider size range as possible using a

JEOL ARM200CF fitted with a probe-aberration corrector. Before recording the high-resolution images, an efficiency scan was made of the ADF detector with fixed amplifier gain and offset conditions, which is shown in figure 5.2.

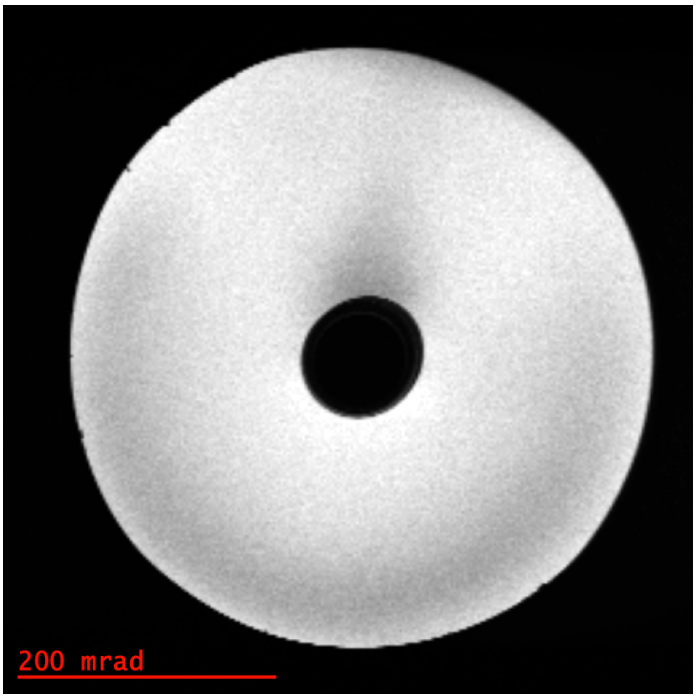


FIGURE 5.2: Example of ADF detector efficiency scan. Reproduced from [8].

The Nellist group determined that the mean nanoparticle size in the sample was 3.46 nm. In addition to statistics on the size distribution, atomically resolved STEM images were taken of the whole range of nanoparticles. Techniques pioneered by Lewys Jones in 2016 to minimise the sources of experimental error[203] from misalignment in the sample and detector were also employed by the group in taking the measurements for this collaborative work. The experiment was also optimised with respect to electron dose to find the minimum dose which could reliably achieve the atomic column counting sensitivity[204]. This dose was found to be $1.38 \times 10^4 e^- Å^{-2}$ per frame which, for a typical series of 20 frames, becomes $2.76 \times 10^5 e^- Å^{-2}$ after integration. A 12 cm long camera was used and rapid, successive image capture with a 90° scan-rotation between successive frames was employed to minimize scanning drift and distortion and sample damage[205–207].

Five samples of averaged frames from the time-series acquisition are shown in figure 5.4. By comparing sequential sets of averaged frames in this way, the Nellist group were able to determine when hopping events took place and average the maximum number of frames will imperceptible beam-damage. This method allows the maximum signal to noise ratio (SNR), stopping before introducing spurious artifacts into the averages. Maximising the number of frames in the averages in this way allowed the Nellist group

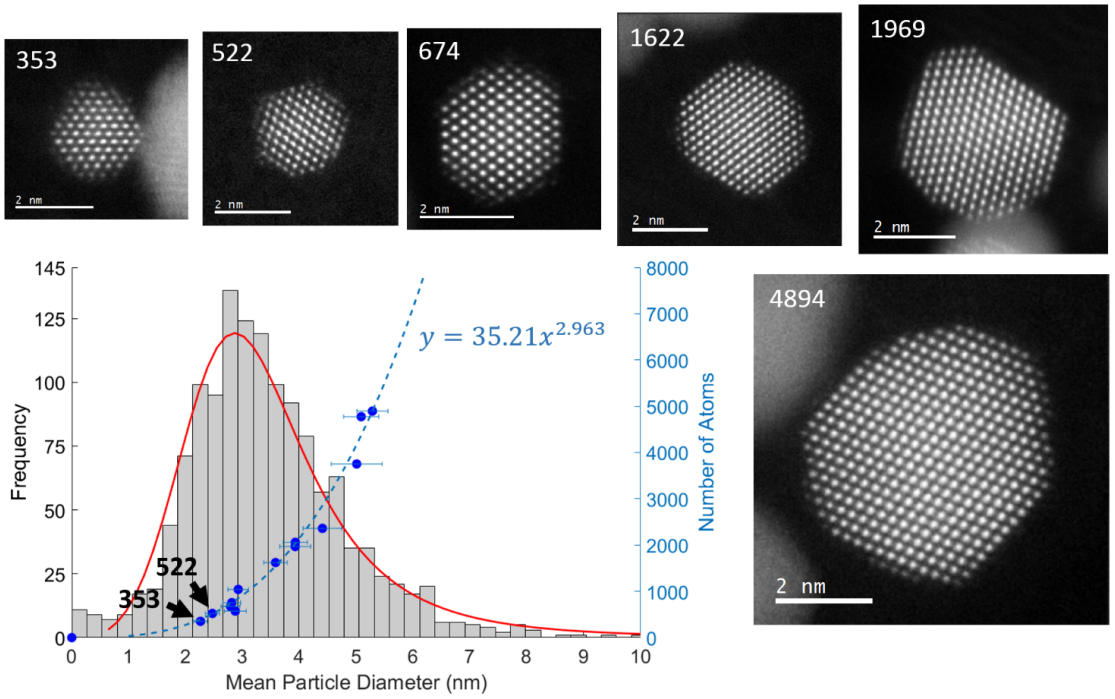


FIGURE 5.3: This histogram was produced using 1342 seperate nanoparticle measurements (mean diameter 3.46nm, median: 3.25nm) representative STEM images are inset with their number of atoms labelled. The overlaid scatter plot (blue points) shows the measured number of atoms (right axis) as a function of nanoparticle diameter. The nanoparticles with 353 and 522 atoms are shown and their relative positions in the population histogram are highlighted. Reproduced from [8].

to get atomic resolution pictures with the nanoparticles aligned with the beam along a crystalographic azamith down to 2 nm in diameter.

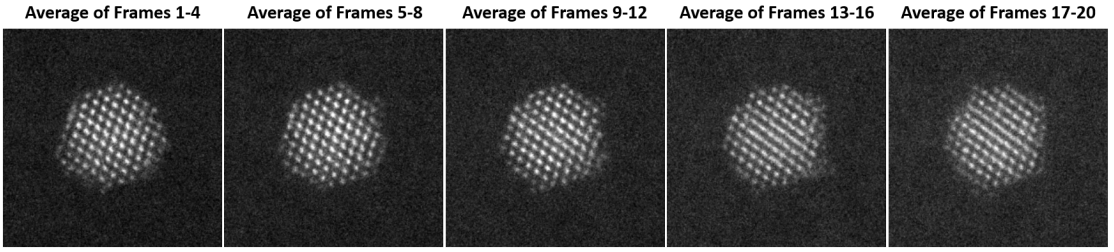


FIGURE 5.4: Integrated frame-sets through the series acquisition set to equal greyscales. These averages over four frames allow the STEM operators to determine how many frames to average in the production atom-counting images, since the averages should not go over hopping events. The averages should be over unchanged nanoparticles to minimise artifacts in the averages from the beam. At the same time, to maximise signal to noise ratio, the number of frames in the average should be maximised. Reproduced from [8].

In the later average frames shown in figure 5.4 a some nanoparticle tilt can just be observed in the thicker centre of the nanoparticle, hopping events have also occured in the bottom right of the average of frames 13-16 as well as on the top left of the sample in the average of frames 17-20. The production average in this case can be taken up to

about frame 12 without introducing many artifacts from beam damage or effects from the sample tilt.

For this work, the Nellist group managed to capture images of 13 nanoparticles with sufficient accuracy to allow accurate atom counts, for which the reconstructions are shown in figure 5.5. Many more were captured with lower accuracy, but sufficient to measure mean nanoparticle diameter, which can be seen in the histogram in figure 5.3 together with raw averaged pictures which went into build the reconstructions.

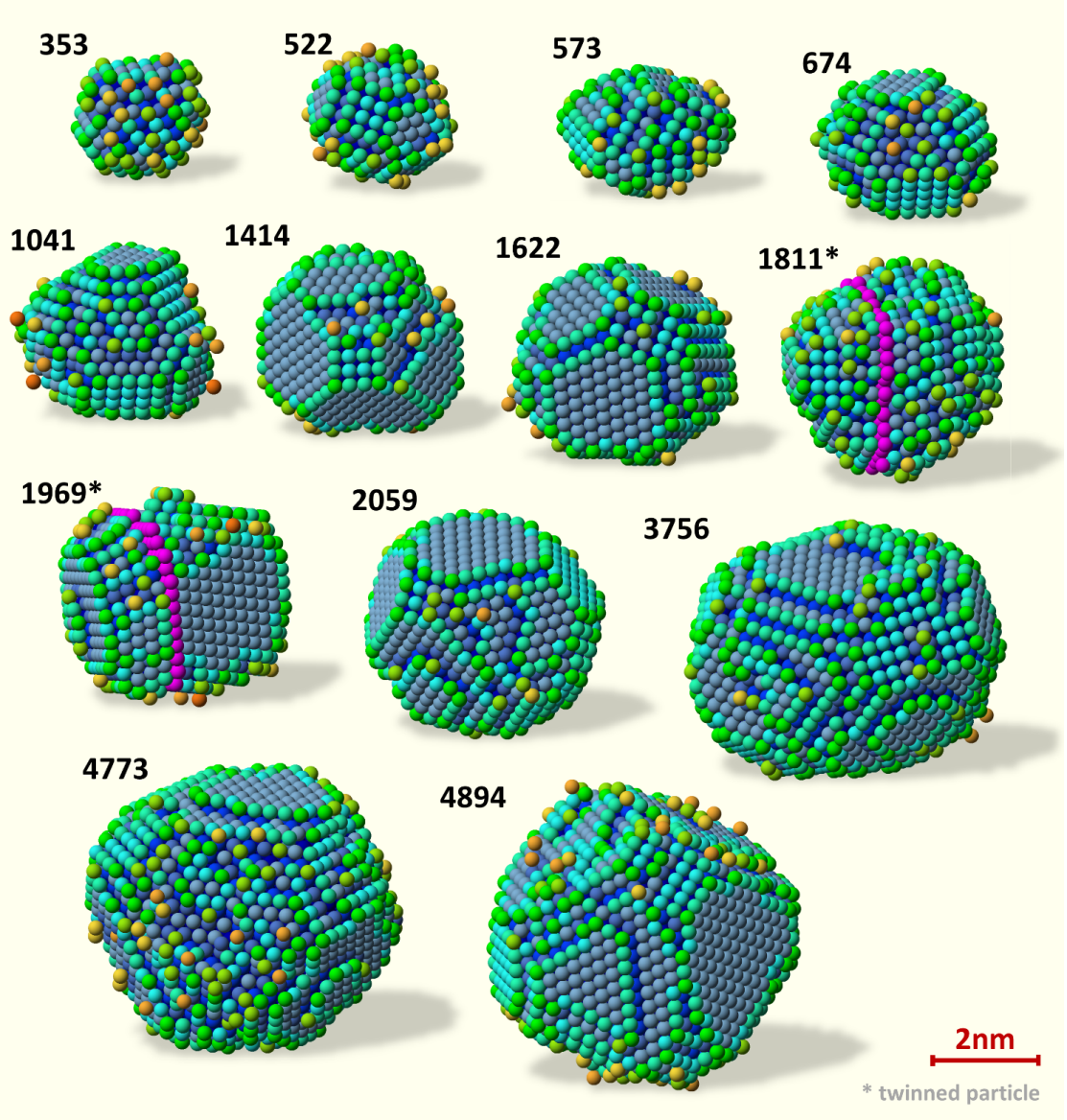


FIGURE 5.5: Montage of all the experimentally determined models. Nanoparticles labeled with an asterisk contain twin planes indicated in pink.

Atom-counts from the high-resolution frames are also plotted in 5.3 as a function of nanoparticle diameter, where the fit follows a cubic relation. The measured pre-factor of 35.21 is within a 0.2% agreement with the value of 35.25 expected for a bulk-like number density of Pt atoms[208]. This confirms that the overall 'nano-weighing' of each

nanoparticle is accurate, giving a good amount of confidence in the *accuracy* of the atom-count assignments for each atomic column in the HR-STEM images. The *precision* of these assignments is dictated by other factors including sample-tilt, scan-distortion[207] and Poisson noise[204], though every effort was made by the Nellist group to minimise these.

5.3 Nanoparticle Reconstructions

After the Nellist group captured the images in the STEM machine, image alignment was performed using Lewis Jones' Smart Align algorithm[207]. The data were then normalised using the ADF efficiency scan using the Absolute Integrator software[9] following the electron flux weighted (EFW) workflow approach [209]. A correction was made for the presence of the amorphous carbon support[204], and the calculated scattering cross-sections[190] were compared with image simulations performed using the µSTEM software. The ADF STEM image simulations were calculated using the quantum excitation of phonons (QEP) method[210] provided through the µSTEM software developed at the University of Melbourne. The software is optimized to provide fast parallel calculations on a graphical processing unit (GPU). An 8x8 Pt unit cell (viewed down either the <110> or <100> zone-axis) was sampled on a 2048x2048 reciprocal mesh, using microscopes settings to match those used experimentally. The thermal mean-square displacement of the atoms was calculated using the Gao and Peng parameterisation[211] and a temperature of 300K. An average was taken over 30 atomic configurations generated using these thermal displacement values. The scattering cross section was then calculated from these simulated images for different crystal thicknesses by integrating over an atomic column to create the simulation library.

The coordinates of the nanoparticle columns and atom counts along these columns were then used to reconstruct 3D coordinate model using the energy minimisation method described in [9] and [209]. As an initial configuration, the columns of atoms were positioned relative to each with their midpoints all on a plane. The positions of the columns relative to each other were then relaxed in the z (beam) direction. This relaxation was performed using a Lennard-Jones model parameterised for platinum. A genetic algorithm was used to search as wide as possible energy landscape; this is based on the work of Yu et al, but using the experimental atom-count data rather than computationally expensive full-field image matching[212]. Atoms were not allowed to move from one projected column to another, nor were the columns allowed to move in x or y, as both of these movements would be in violation of the experimental observations.

We do not claim that the reconstructed 3D geometries are *unique*. Other models fit the experimental data, such as the trivial case of the z-direction mirror image, and energetically indistinguishable structures that could have been found (such as an adatom

on an otherwise clean facet moving a short distance to an equivalent position, as seen in the DFT structure optimisation of Pt_{522}). Instead, these models provide us with experimentally determined, highly probable 3D structure coordinates to work with.

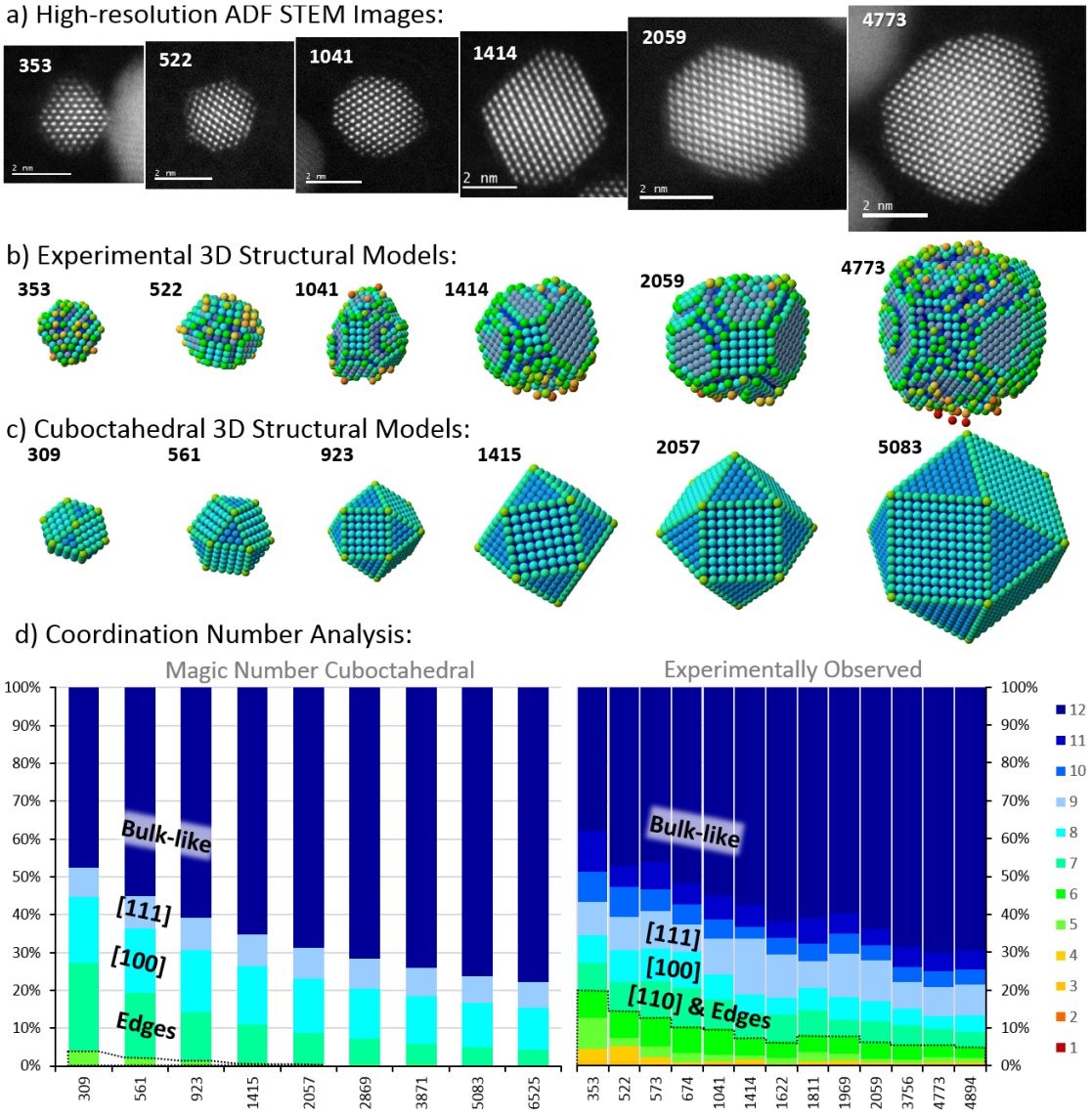


FIGURE 5.6: HR-STEM images (a) and accompanying hard-sphere models (b) for experimental nanoparticles observed with atom-counts near to magic-numbers; models are coloured by atom coordination numbers[9] and rotated to show their dominant facets. Magic number cuboctahedra are shown for comparison rotated to the same orientations (c). Population histograms (d) show the coordination number fractions as a function of nanoparticle atom-count for both magic (left) and experimentally observed nanoparticles (right). The 7, 8, and 9 coordination atoms corresponding loosely to both (110) facets and edges between facets, and (100) and (111) surfaces respectively are labelled, as well as the 'bulk-like' atoms of $\geq$10 coordination. The dashed area represents the fraction of low-coordination adatom, corner and step sites. Reproduced from [8].

The coordination number information from Figure 5.6 reveals some important results.

Firstly, as nanoparticle size increases the proportion of bulk-like coordinations (≥10 coordination) increases; a classic surface area to volume effect. Secondly, all the experimental nanoparticles observed showed some surface roughness with significantly more atoms of lower coordination than cuboctahedral nanoparticles of equivalent size. The dashed area in Figure 5.6 represents the low-coordination sites (6 and fewer) and this is 20% of the *total* atoms in the 353 atom nanoparticle, rising to more than 46% if we consider these as a proportion of the *surface* atoms. Of the nanoparticles observed experimentally, some contain a number of atoms close to a cuboctahedral 'magic numbers' (Figure 5.6).

5.4 Modelling with DFT

The energy minimisations, which were performed by Lewys Jones of the Nellist group, with Lennard-Jones and genetic algorithms provided a computationally efficient but somewhat simplified approach to determining the optimum structure for each nanoparticle; as a result, these were followed by a more computationally intensive and fully unconstrained geometry relaxation performed using the Sutton-Chen[213] force field within the DL_POLY4 software[214]. During this relaxation, a small relaxation of the atomic separations was observed around the periphery of the nanoparticles (RMS differences between the initial and final structure of 0.1 Å). The same 0.1 Å RMS difference was observed in the larger Pt_{522} structure but with the exception that an adatom on the surface migrates to its nearest neighbour site: this single platinum atom moves 2.6 Å in this relaxation. Due to these relatively small changes, we believe the experimental 3D models from the STEM experiments to be reliable.

ONETEP with the Coulomb cutoff approach was used in calculating all the electrostatic interactions in the energy functional.[156] The Coulomb cutoff approach allows us to apply a spherical truncation to the electrostatic potential and provided that the sphere covers all of the electronic density of the nanoparticle, the calculations are effectively isolated in space (rather than having periodic boundary conditions. Other more rapidly diminishing terms such as exchange-correlation are taken care of more simply though the use of a moderate 15 Å vacuum gap on each side of the simulation cell.

We converged rigorously with respect to the radius of the Coulomb cutoff sphere and found that the radius of the sphere should be

$$R = \max(r_{ij}) + 2r_{NGWF} + a \tag{5.1}$$

that is that the sphere is chosen to enclose the maximum separation of atomic centers r_{ij}, plus two NGWF radii r_{NGWF} and an extra constant distance a (set to 10.0 a_0) to avoid ringing as the operator is applied in momentum space.

The core electrons were described by the Projector Augmented Wave (PAW) approach with a PAW potential that treated the 5p, 5d, 6s, 6p electrons as valence, with a core charge of 16. The PAW data was taken from version 1.4 of the GBRV set[174]. Using NGWFs, which are local in space, leads to efficiency gains as compared with using a plane-wave basis. An *in situ* optimisation of the NGWFs ensures that far fewer functions are required for accurate calculations than atomic orbital basis set calculations and that near-complete basis set accuracy is achieved as a result of the psinc basis set, which is equivalent to plane waves. In addition, since the NGWFs are expanded in terms of basis functions that do not depend on atomic coordinates, ONETEP does not suffer from basis-set superposition errors, and can be systematically converged to the complete basis set limit with a single parameter (kinetic energy cut-of).

12 NGWFs were used per Pt atom. A psinc kinetic energy cut-off of 850 eV was used which ensured that the geometries were fully converged with respect to the basis set. The NGWFs exist within localisation spheres so that they have a fixed maximum extent in real space, the calculations were also converged with respect to this parameter and a radius of 9.0 a_0 was found to be sufficient.

A Fermi-Dirac occupancy smearing around the Fermi level, of 0.1 eV, was used. Using a smearing function for the occupancies of the electronic energy levels is a common technique in DFT calculations of metallic systems. This approach allows for smooth convergence towards the ground state energy minimum while affecting the system minimally[81]. Full geometry optimisations of the nanoparticles' atomic degrees of freedom were performed with the BFGS algorithm and forces were converged to a tolerance of 0.03 eV/Å. Calculations on Pt (111) slabs were also performed with the same protocol. In this case the slab consisted of 5 layers of Pt atoms, with 180 atoms of Pt in total. The Pt slab geometry was obtained from geometry optimization of the slab and had a lattice parameter of 3.913 Å. The calculations were done at the gamma point with a periodic supercell (6x6) in the slab plane and open boundary conditions in the vacuum direction. Owing to the computational demands of DFT, only the two smallest experimental models containing 353 and 522 atoms were used for DFT calculations, as well as the cuboctahedral nanoparticles with 309 and 561 atoms which are of similar size to these two experimental nanoparticles. These required significant supercomputer resources, and so nanoparticles larger than this ($>$ 2.8 nm diameter) were not studied here with DFT. Electronic calculations were performed with the ensemble-DFT method[7] for metallic systems in the ONETEP linear-scaling DFT program[40]. The Revised Perdew-Burke-Ernzerhof (RPBE) exchange correlation functional was used[20].

Although only two cuboctahedral, and two experimental structures were computed, our aim was to obtain descriptive information from the DFT calculations which can be used to predict chemical properties across the observed ensemble of nanoparticles. To achieve this, we use the approach described in chapter 3.3.2, where we first map the

local electronic-density calculated on an isosurface of the total electrostatic potential of the nanoparticle (0.0 E_h has been used throughout this work). Figure 5.7 shows these maps for the two smallest experimental nanoparticles and their closest cuboctahedral analogues.

We also performed (111) slab calculations to validate the use of fixed metal atom positions (relaxed oxygen positions) versus fully relaxed positions. These points are also shown on Figure 5.10 and lie within the spread of the oxygen-relaxed data.

TABLE 5.1: Comparison of oxygen-relaxed and fully-relaxed binding energies on a platinum slab model.

Site	***Gen. Coord. No.***	***Binding Energy***	
		Oxygen Relaxed (eV)	***Fully Relaxed (eV)***
Atop (111) on slab	*7.5*	*0.310*	*0.178*
Hollow (111) on slab FCC	*6.95*	*-0.751*	*-0.982*
Hollow (111) on slab HCP	*7.5*	*-0.472*	*-0.628*

By comparing the electronic charge density maps between the experimentally measured and cuboctahedral nanoparticles, it is immediately apparent that the experimental structures show a far greater number of sites of low electrondensity compared to their cuboctahedral counterparts. Atoms with lower coordination numbers exhibit lower electron density values on their surfaces (as sampled on the electrostatic isosurface) due to the well-known Smoluchowski effect[215].

The Smoluchowski effect results in a dipole moment, with negative side closer to the surface which has the effect of reducing the workfunction on the low-coordinated atoms. The dipole moment is caused by differences in bonding between the adatom and the surface atoms; the lowest energy state is reached by forming bonds with states close to the Fermi level and the substrate has more of these states than an adatom, through containing more atoms. Because of this, valence charge density flows from the adatom towards the substrate, causing a smeared, spread out electronic charge density profile above the adatom and a net dipole moment. This effect describes the observed[216–218] spreading out of electronic charge density at sharp, surface corrugations and defects resulting in loss of electron density above low coordinated sites (see Figure 5.13). This is particularly visible when contrasting the relatively high electronic charge density on facets of the cuboctahedral nanoparticles with the lower density values on the edges and

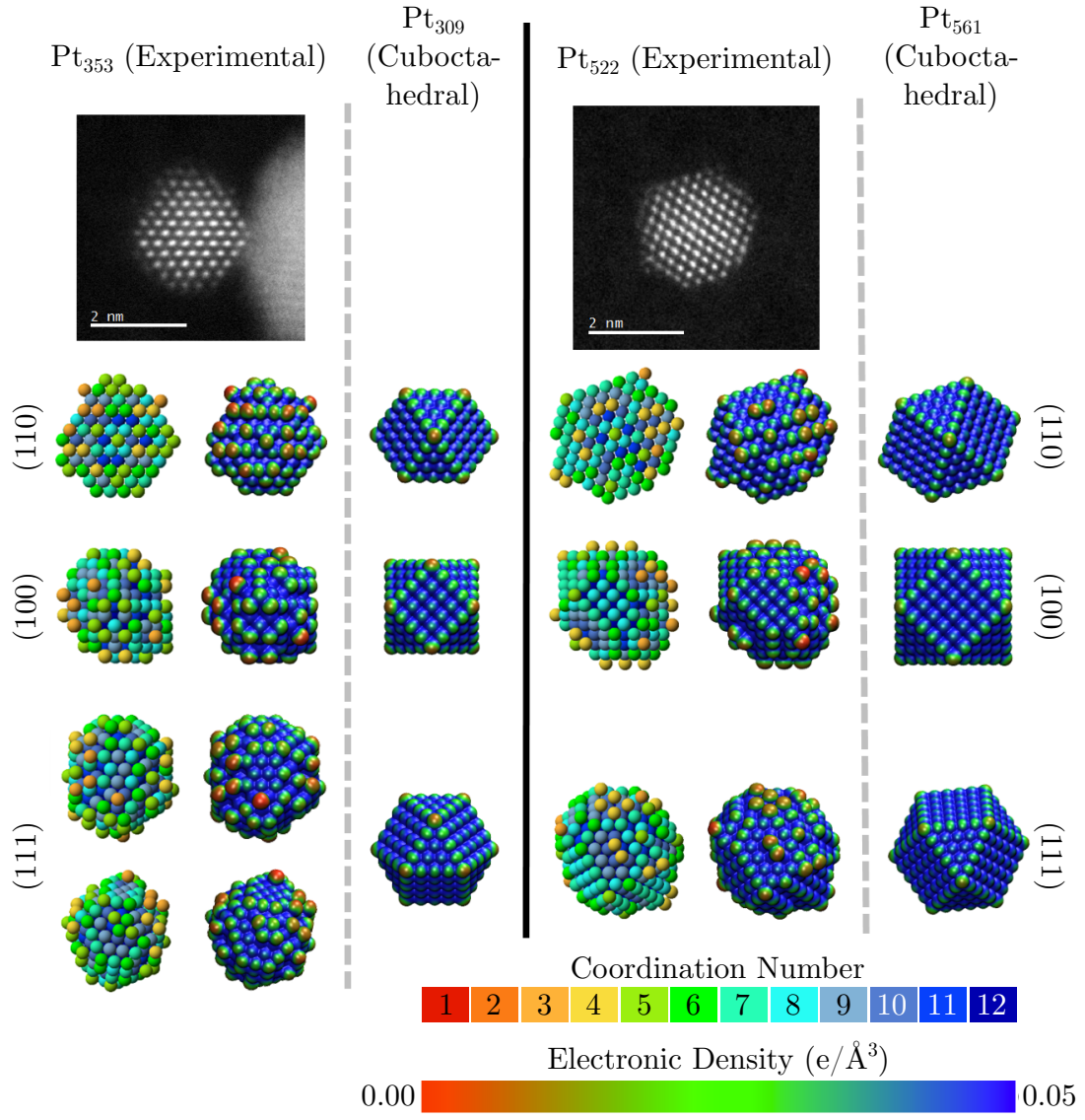


FIGURE 5.7: Coordination number and DFT analyses for the smaller (left) and larger (right) nanoparticles studied. Hard-sphere structural models were calculated from the ADF data and are coloured according to their coordination number; these are shown rotated to various viewing directions to highlight their morphology. Alongside the structural models, the DFT results showing isovalue surfaces of total electrostatic potential at 0.0 E_h colourised by the local electronic-density. Size analogue cuboctahedra DFT results are shown for comparison.

even more electron deficient vertices. Also, when examining the experimental nanoparticles, very high-density sites are visible under vacancy defects[219], where the atomic centre has high coordination number, and very electron deficient isosurface sites can be seen on adatoms with low coordination number.

The reduced work function for the low coordination atoms that results from the Smoluchowski effect can be considered together with the computed total charges of these atoms, which, rather paradoxically, show that the actual electronic charge that can be apportioned to the low-coordinated atoms is larger than the high-coordination atoms. Thus, the low-coordinated atoms are overall electron-rich with a lower work function, so they are stronger binding sites for electronegative ligands such as O. This can be verified from the results of Figure 5.10(a) where it can be seen that as the value of the electron density on the isosurface decreases in magnitude as the computed binding energy of O atoms on these sites increases.

The results from the electronic charge map descriptor are compared with the other descriptors we have investigated, namely oxygen binding-energies at various sites (which can be directly related to catalytic activity), the d-band centres[220] (see figure 5.8), and partitioning of electronic charge on atoms.

5.4.1 Calculated d-band Plots

The electronic density isosurfaces also highlight a great diversity of surface sites between the cuboctahedral and experimental models. To evaluate this systematically an algorithm was developed to identify all the possible atop, bridge, and hollow sites across the surface of the nanoparticles, described in chapter 3, and implemented as a stand-alone program (this will be deposited in the 'Utilities' section of www.onetep.org). The results of this are illustrated for the Pt_{309} model in figure 5.11. Using this code, the generalised coordination number (gCN) of every possible site for an oxygen adsorbate was calculated (using the formula from Calle-Vallejo et al.[10]) for every one of the experimental and magic-number atomistic models. The results for the three smallest nanoparticles are shown in Figure 5.9.

The histograms in Figure 5.9 reveal several insights. Firstly as expected, the Pt_{309} cuboctahedral model nanoparticle exhibits only a small number of unique gCN oxygen binding sites; a total of 25 unique-sites describe every possible symmetry-related adsorption site, and a systematic DFT study is relatively tractable. For the experimental Pt_{353} of similar size, the surface roughness leads to a far greater number of unique site geometries, 273 unique sites in total. The largest of the experimental models in this work exhibits ≈8,300 surface adsorption sites with 539 unique geometries.

Depending on future investigation that will be performed, many useful quantities can be extracted from these histograms. It is easy to extract, for instance, 'the fraction of

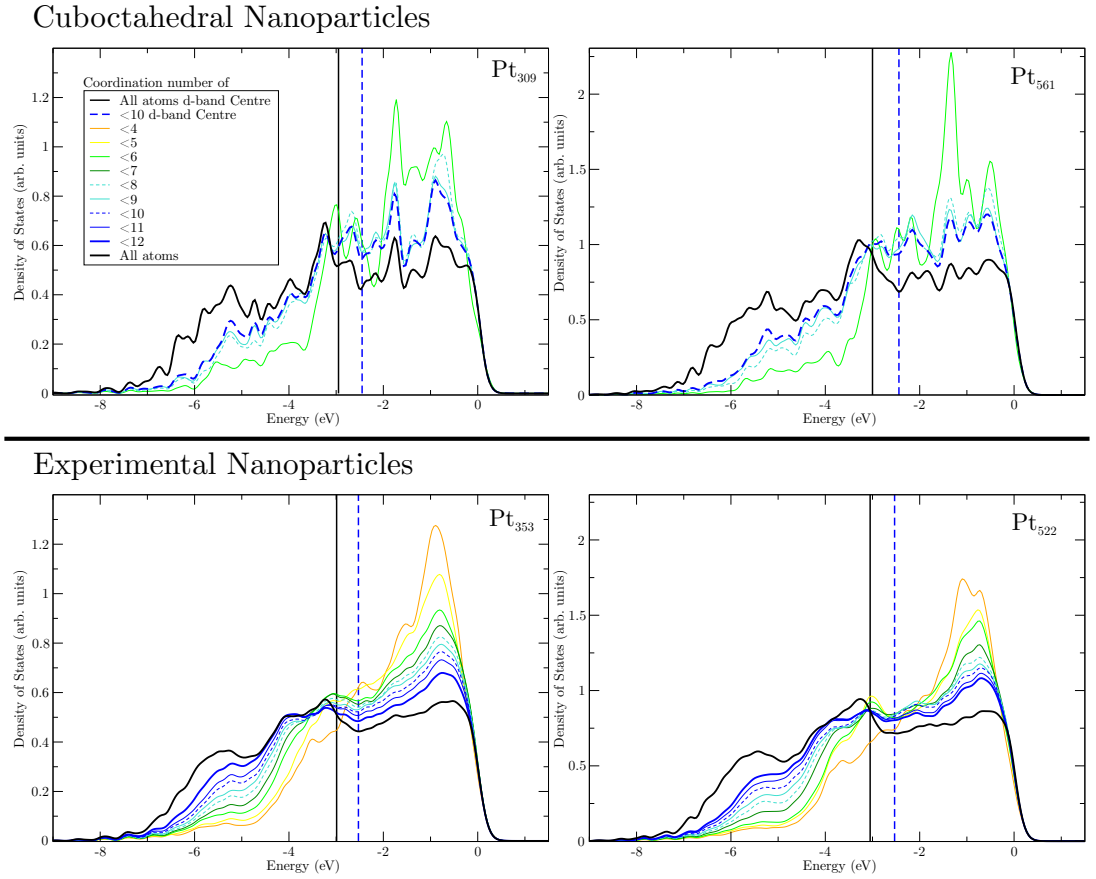


FIGURE 5.8: The angular momentum projected density of states (d-channel) for the experimental and cuboctahedral nanoparticles. The lines correspond to increasingly coordinated atoms being included in the LP-DOS sum, up to all atoms in the system (coloured as in Figure 5.7). The plots converge with increasing coordination contribution towards the surface, then when considering the 12-coordinated internal atoms, there is a jump towards a bulk-like density of states.

hollow-sites out of all sites' or 'the number of total oxygen-binding sites with a gCN greater than 8' against nanoparticle size. In this work, having access to a large collection of experimental nanoparticle models of increasing size and of adsorption site diversity, we do not seek to study each of the many thousand unique adsorption sites, but rather to extract some single parameter to compare nanoparticles across the ensemble.

While the colour coded isosurface representations provide a visual qualitative descriptor of the strength of O binding, quantitative values can also be extracted across the ensemble of nanoparticles measured, using the approach described in chapter 3. Figure 5.10 shows the oxygen binding-energies across a variety of sites from a Pt_{309} cuboctahedral nanoparticle and also platinum slab models (as computed with DFT calculations) both as a function of the electron density value on the potential isosurface, and as a function of their generalised coordination number[221]. Further models for the adatom variants used for the calculations leading to Figure 5.10 are shown in figure 5.13). Electronic-density values, for each binding site (whether atop, bridge, or hollow) are used as descriptors of

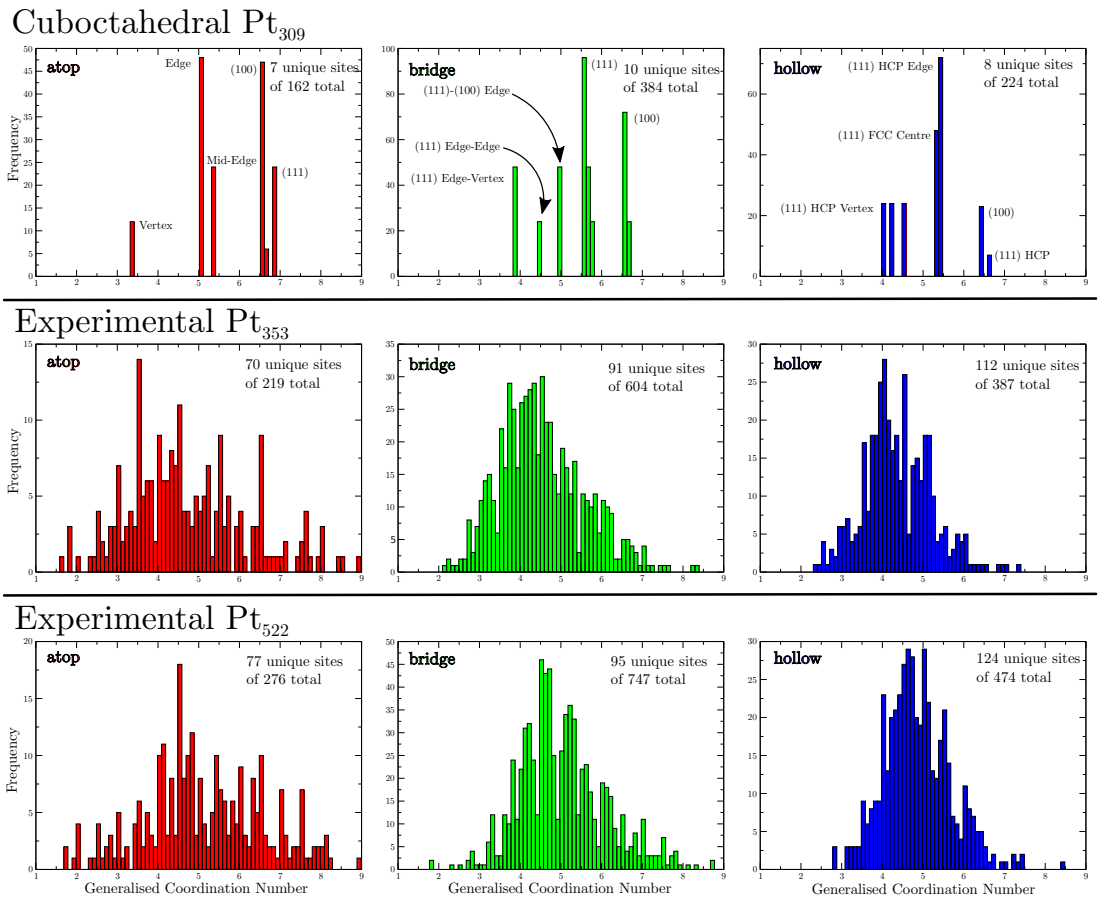


FIGURE 5.9: Histograms of the generalised coordination numbers of all possible surface atop, bridge and hollow oxygen binding sites, for the cuboctahedral Pt_{309} and the experimental Pt_{353} and Pt_{522} nanoparticles. For all plots, the number of total sites are indicated and the number of unique sites.

O binding strength, providing linear relationships for each binding-energy as a function of electron density value. Information about how these values are extracted, for each type of site (atop, bridge or hollow). One advantage of such electronic descriptors is that they are expected to apply also to alloyed nanoparticles, or particles with built in surface strains, which we intend to explore in future work.

The DFT O adsorption energies were computed according to the following equation

$$E_{bind} = E_{NP+O} - E_{NP} - \frac{1}{2}E_{O_2}, \tag{5.2}$$

where E_{NP} is the total energy of the bare, adsorbate-free nanoparticle in vacuum fully relaxed with DFT. E_{NP+O} is the total energy of the nanoparticle with a single oxygen atom adsorbed onto the site being probed, and E_{O_2} is the total energy of an oxygen molecule in isolation fully relaxed with DFT. To maximise the throughput of unique sites evaluated in a reasonable computation time, the Pt positions in the nanoparticle are kept fixed in their vacuum optimised geometries while each O atom is fully relaxed.

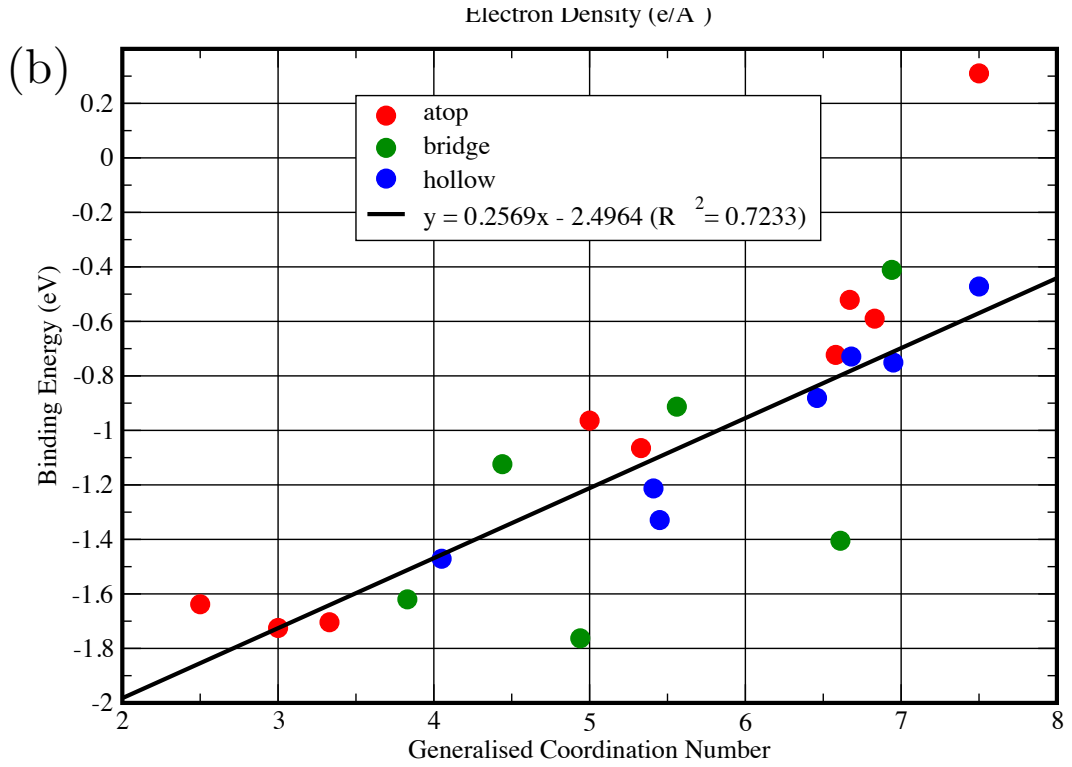


FIGURE 5.10: Calculated oxygen binding-energies for various oxygen-relaxed atop, bridge and hollow sites from both a cuboctahedral Pt_{309} nanoparticle, and from slab calculations, as a function of electron density (a) and gCN (b). The gCN values are calculated using the formula from Calle-Vallejo et al.[10]. Black diamonds represent the results of fully-relaxed (oxygen and platinum) calculations.

To validate this approximation, results were compared with slabs using full relaxation. In these calculations, the degrees of freedom of all atoms in the three surface layers were free to relax due to the presence of the oxygen, along with the oxygen degrees of freedom. The bottom two layers remained constrained to be bulk-like. The change in oxygen binding strength, observed when allowing the platinum atoms to relax due to the presence of the oxygen (Figure 5.10, black diamonds), is small relative to the energy difference between sites of differing character and well within the scatter of the data. Using the fixed platinum constraint is therefore justified and allows us to probe a far greater number of sites in a reasonable computational time frame. Results from the cuboctahedra and the slab calculations are shown together in Figure 5.10.

Binding-energy calculations for oxygen at various sites (Figure 5.10a) shows that oxygen binds most strongly to sites predicted to have low electron densities. Similarly, surface atom d-projected DOS, resolved according to the coordination number of surface atoms shows that reducing the coordination number shifts the d-band centre to higher (more positive) values. This is in accordance with the d-band model of Nørskov[197] which suggests enhancements of binding strength as the d-band centre becomes more positive. From our calculations the $Pt_{(111)}$ binding-energy (at a hollow FCC site on a slab) was

TABLE 5.2: Tabulated DFT Site Data.

Site Type	Short Code	Location	gCN	e- Density ($e/Å^3$)	Binding Energy (eV)
Atop	A1	(111) on slab	*7.50*	0.0430	0.310
	A2	(111) mid facet Pt_{309}	*6.83*	0.0312	-0.590
	A3	(100) mid facet Pt_{309}	*6.67*	0.0325	-0.521
	A4	Edge site Pt_{309}	*5.33*	0.0200	-1.065
	A5	Vertex site Pt_{309}	*3.33*	0.0056	-1.704
	A6	Adatom on (100) Pt_{309}	*3.00*	0.0050	-1.725
	A7	Adatom on (111) Pt_{309}	*2.50*	0.0024	-1.638
	A8	(100) facet Pt_{309}	*6.58*	0.0320	-0.723
	A9	Edge site Pt_{309}	*5.00*	0.0180	-0.964
Bridge	B1	(100) mid facet Pt_{309}	*6.61*	0.0590	-1.405
	B2	(111) mid facet Pt_{309}	*5.56*	0.0600	-0.913
	B3	(111) on slab	*6.94*	0.0920	-0.411
	B4	(100)-(111) Edge site Pt_{309}	*4.94*	0.0510	-1.763
	B5	(111) facet Pt_{309}	*4.44*	0.0500	-1.124
	B6	Edge site Pt_{309}	*3.83*	0.0350	-1.620
Hollow	H1	(100) Pt_{309}	*6.46*	0.0960	-0.881
	H2	(111) HCP Pt_{309}	*6.68*	0.0830	-0.729
	H3	(111) FCC Pt_{309}	*5.41*	0.0740	-1.213
	H4	(111) FCC on slab	*6.95*	0.1100	-0.751
	H5	(111) HCP on slab	*7.50*	0.1130	-0.472
	H6	(111) HCP edge Pt_{309}	*5.45*	0.0780	-1.329
	H7	(111) HCP vertex Pt_{309}	*4.05*	0.0630	-1.471

found to be -0.739eV; taking this, and the weaker binding predicted to be desirable from the literature[58, 76, 80], yields an optimum chemical activity with an oxygen binding energy of -0.52eV.

Oxygens are only found to bind strongly (more negative binding energy) to atop sites for gCNs lower than ≈3.5. At higher coordination, we predict that the bridge and hollow sites *do* have some strong binding sites. Unlike the simple geometric coordination numbers of the metal atoms shown in Figure 5.6 & Figure 5.7, the generalised coordination number of the oxygen-adsorbate sites (Figure 5.9 & Figure 5.10) capture a more subtle weighted average of the neighbouring atoms. The quality of the generalised coordination number as a descriptor of oxygen binding-energy has been shown to offer a powerful

tool for catalyst evaluation[10, 200, 219, 221], and has been exploited in the engineering of defective extended surfaces, but not so far for nanoparticles[219]. The approach has also already been reported to yield reliable results across different sizes of metallic nanoparticles[10, 219].

From the data in Figure 5.10b) we fit a linear trend which allows us to extrapolate from the individual, and computationally expensive, DFT calculations and onto the many thousand unique sites across the experimental observations, and predict a binding-energy for every surface site see section 3.4 of chapter 3 for more details. So-called "volcano plots" (see chapter 1 for more details) that follow from the Sabatier principle of catalyst activity, whether from different materials[80, 222] or for platinum surface sites of different coordinations[219], show that pure platinum surfaces always bind oxygen slightly too strongly (desorption limited). At present is it not possible for us to calculate computationally the chemical-activity of every site across the particle surface due to the sheer number of sites and a lack of computer resources. Here, as a useful metric of overall particle catalytic activity, we consider the fraction of surface sites within 0.2 eV of the optimum binding strength of -0.52 eV, and this fraction can be rapidly calculated using the linear fit in Figure 5.10b). While this may not explicitly yield a quantitative relationship to chemical-activity, it does offer a rapidly calculated metric for comparing any given particle with reference to the rest of the ensemble. This 'fraction of sites within 0.2eV' ($F_{\Delta<0.2eV}$) metric is more subtle than simply counting the fraction of {111} sites, is sensitive to particle-size, shape and roughness, and so can be used to evaluate the comparative catalytic performance of nanoparticles from a real catalyst sample.

To summarise; from high-quality ADF images and computationally efficient relaxations, we have the 3D locations of the metal atoms, the locations of all possible oxygen binding sites, and the generalised coordination numbers of all those sites. Combining this with the DFT master-trend from Figure 5.10(b), we are able to calculate a $F_{\Delta<0.2eV}$ for each observed particle. This site-identification and master-trend scaling is rapidly calculated, even for very large particles, and is shown in Figure 5.14 for both the experimental particles and for thermodynamically predicted Wulff shapes (truncated cuboctahedra), for cuboctahedra and quasi-spherical comparators.

Using the α-hull code described in Chapter 3 the positions and natures (atop, bridge or hollow) of all possible oxygen binding sites were identified.

From these data the generalised coordination numbers of all sites were calculated as shown in the histograms, see figure 5.9

Once the α-hull has been found, we can use it to find the values of electronic density on the electrostatic potential isosurface as described previously (after expressing all potential isosurfaces with respect to the same reference potential value) at the binding sites. This is done by taking the pre-computed position of the adsorbate binding site and finding the vector perpendicular to the α-hull which passes through this point. The intersection

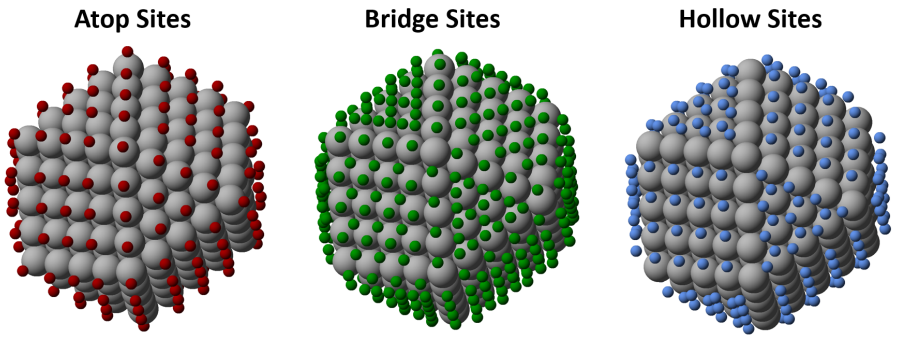


FIGURE 5.11: Illustrations of all possible atop, bridge and hollow oxygen binding sites.

point of this vector with the isosurface of the electrostatic potential at the chosen isovalue is then used to find the value of electronic density at this point on the electrostatic potential isosurface. This value of electronic density is the electronic descriptor value at this binding site.

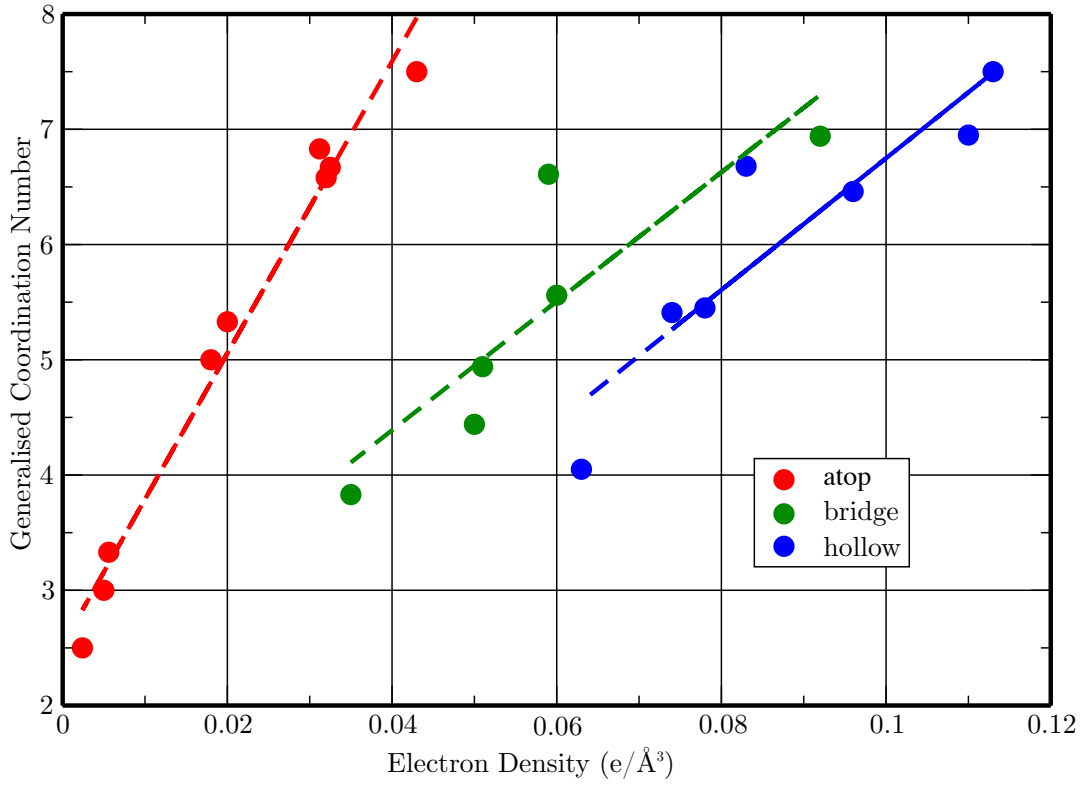


FIGURE 5.12: Plot of generalised coordination number as a function of electronic density for the atop, bridge and hollow sites.

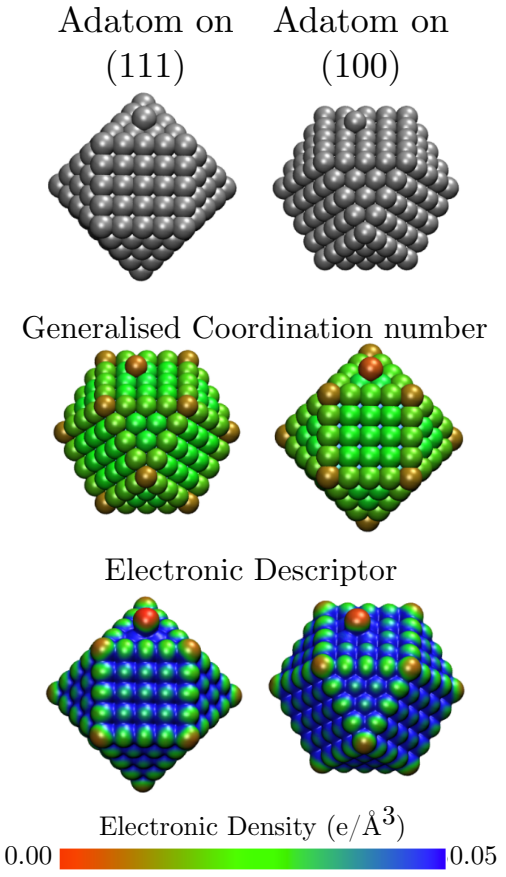


FIGURE 5.13: Structural ball model (left), the generalised coordination numbers (middle), and isovalue surfaces of total electrostatic potential colored by electronic density (right).

5.5 Results

For the Wulff-like truncated octahedral nanoparticles we see a monotonic curve dominated by simple size effects. Even though, by definition, all the models with truncated octahedral symmetry have similar {111}/{100} face area ratios, at small diameters the excess of low coordinated edges and vertices results in a smaller $F_{\Delta<0.2eV}$. This $F_{\Delta<0.2eV}$ falls to zero for truncated octahedra below 1.5 nm and is consistent with the data reported by Shao et al. for their nanoparticles smaller than 2 nm[201].

The Nellist group was unable to observe experimentally, nanoparticles smaller than 2 nm on-axis in the HR-STEM; in addition to these being only a small fraction of all present, even when found these are easily damaged under the electron beam. A similar trend is observed in the cuboctahedral series, though with the sharper tips of such models yielding very low-coordinated sites, the $F_{\Delta<0.2eV}$ falls to zero for all models below 2.0 nm.

The $F_{\Delta<0.2eV}$ for the experimentally observed nanoparticles was found to always be lower than for the truncted octahedral nanoparticles but higher than the cuboctahedra. That is to say that the experimental results are all bounded by one of these two geometric

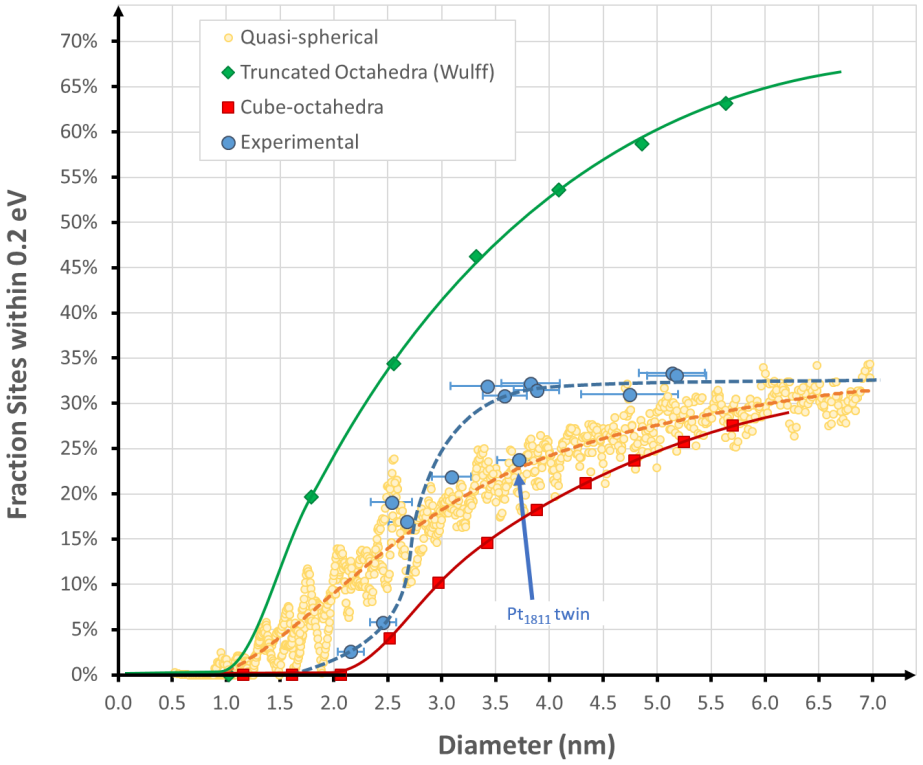


FIGURE 5.14: Fraction of surface sites within 0.2eV of optimal binding ($F_{\Delta<0.2eV}$) for the 13 experimentally observed particles, and truncated-octahedra, cuboctahedra and quasi-spherical reference particles. Reproduced from [8].

models. From this we might conclude, that while it is generally valid to use perfect-model geometric-series for simplified surface studies, the true experimental surface will always reveal more subtle details. From the structural models (Figure 5.6), and the gCN histogram analysis (Figure 5.9) we observe both an increase in roughness compared with Wulff shapes, and also a general lowering of the surface gCN. Because the lowest coordination sites are so strongly over-bound, rough adatom sites effectively poison the surface locally blocking otherwise high-coordination sites below. The downward shift of the experimental curve from the Wulff trend for the largest and smallest nanoparticles encompass the various topographic-effects, reducing the $F_{\Delta<0.2eV}$ for these models. In contrary, relative to the cuboctahedral models, we see the opposite effect where $F_{\Delta<0.2eV}$ is now improved in comparison to the perfect models (consistent with the increased fraction of {111} facets observed). Both these relative results indicate that caution must be exercised when comparing experimental chemical activity measurements to geometric simulations, something we hope to study in a future work.

As a further comparison, a series of 1600 randomly generated single-crystal quasi-spherical models were also generated and the $F_{\Delta<0.2eV}$ evaluated. These models were cleaved from a FCC platinum bulk lattice with only atoms within a radius threshold from the origin kept. This radius was made larger to create models with larger diameter and to form the

series. Because of the discrete atomic nature of the spherical models, the surfaces are not perfectly smooth but rather a mixture of small facets and rougher high-index sites. The $F_{\Delta<0.2eV}$ reaches most closely the truncated octahedral (Wulff) trend for nanoparticles in the size range of 2.7-3.5nm; with these models (with their increase fraction of {111} facets) sitting higher than the quasi-spherical trend but still falling short of the Wulff models. The location of this optimum is consistent with the industrially optimised synthesis process (determined from trial and improvement), and the specific-activity observations previously reported[201]. For the very largest experimentally determined nanoparticle structures, these appear more spherical, suggesting that they are further from their thermodynamic optimum shape and their shape is more kinetically limited, and indeed the $F_{\Delta<0.2eV}$ approaches the quasi-spherical trend. From inspection of Figure 5.5, it can be seen that the larger nanoparticles do not show a commensurate increase in {111} facet area and have become more spherical.

Two twinned nanoparticles were measured experimentally at 3.7 nm (1811 atoms) and at 3.8 nm (1969 atoms). While one of these was well below the wider trend, as only two of the thirteen nanoparticles in this study contained twins, a wider study of twinned nanoparticles would be required before making any specific conclusions about these. Finally, the cuboctahedral series of nanoparticles show the lowest $F_{\Delta<0.2eV}$ of any series, and underestimate the $F_{\Delta<0.2eV}$ at 3.5 nm compared to the experimental result by the same amount that Wullf-shapes overestimate this. Figure 5.14 then highlights two important points; firstly presumptive model geometries alone should not be relied upon when studying expected chemical behaviours, and secondly, that if they are then catalyst synthesis optimisation can be misinformed.

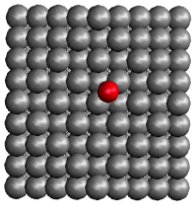
5.6 Conclusion

TIn conclusion, the three-dimensional structural models presented here, with up to 5000 atoms each, can each be calculated from experimental data in only a few hours on a standard desktop computer[9]. These structural models can be used as inputs to reduced-scaling DFT calculations to yield the electronic structure of the experimental nanoparticles. It is observed that increased roughness, compared to the equilibrium Wulff structures, lead to more sites with lower electron density and therefore higher O-binding reducing activity. Using a DFT-generated binding-energy master-trend calibration plot, an estimate for the 'fraction of sites within 0.2eV' ($F_{\Delta<0.2eV}$) can be calculated in only a few seconds for an ensemble of experimentally determined structures. This would be computationally unfeasible to do by DFT calculation of individual adsorption energies for each of the sites of each nanoparticle. By leveraging the accuracy of large-scale DFT calculations on small models, via the generalised coordination calibration relation, we can gain insight into a far wider ensemble of atomic models than would otherwise be possible.

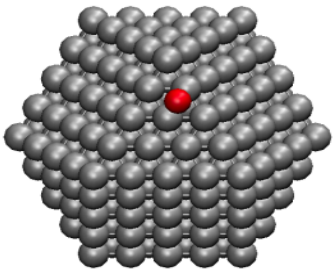
This high-throughput approach agrees well with other reported experimental[201, 222] and simulated[221] findings for this system and offers the potential to accelerate the development of better tuned catalysts for the ORR. Similar calibrations for other species have been reported and may lead to the optimisation of other systems[223]. In this work, the generalised coordination descriptors were evaluated for pure metal nanoparticles. While these may not be expected to be simply transferable to doped or alloyed cases, the electronic descriptors can be; these come at an increased computational cost with a single point energy calculation needed for each nanoparticle structure that needs to be studied but are nevertheless dramatically less demanding than having to do a DFT calculation of adsorption energy for each distinct site. In the future, the $F_{\Delta<0.2\mathrm{eV}}$ metric seems well suited to future extension from gCN to electronic-density, exploiting other experimental inputs, such as changes in binding-energy arising from composition[80] or atomic displacement (strains)[220]. The current rapid pace of progress in experimentally mapping composition and strain in the STEM merits revisiting this concept in the near future.

TABLE 5.3: Additional surface sites probed by DFT. Short-codes correspond to those in Table 5.2 below.

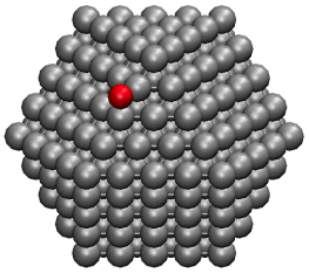
B1 – Bridge on {100} facet



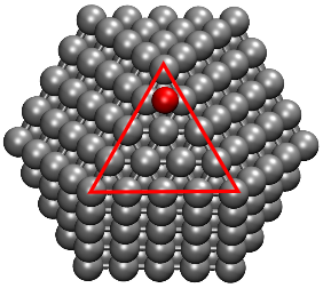
B2 – Bridge on {111} facet



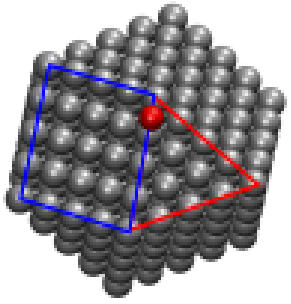
B4 – Bridge on {100}-{111} edge



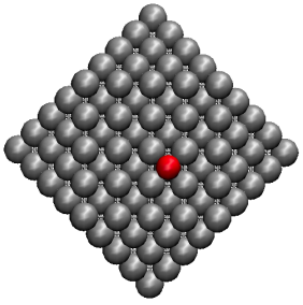
B5 – (111) Edge:Edge



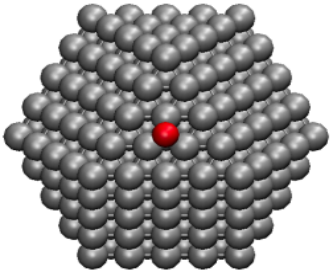
B6 – (111) Edge:Vertex



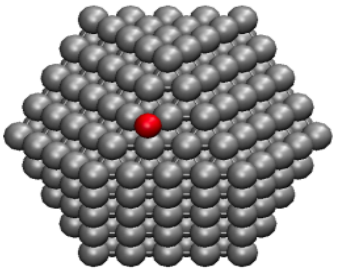
H1 – Hollow on {100} facet



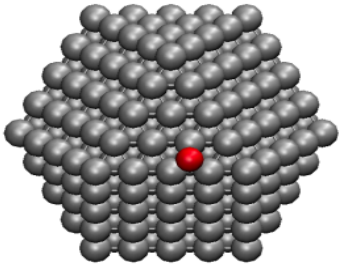
H2 – Hollow on {111} HCP site



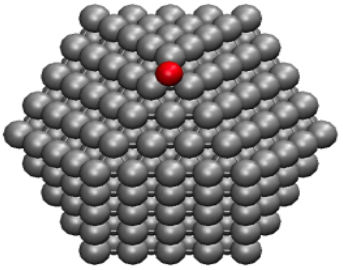
H3 – Hollow on {111} FCC site



H6 – {111} Hollow near Edge



H7 – {111} Hollow near vertex



Chapter 6

Electronic annealing Fermi Operator Expansion for DFT calculations on metallic systems

The work presented in this chapter, specifically the "AQuA-FOE" method has been described in a manuscript which we have submitted to JCP, which may be published at some point either before or after the examination of this thesis :
Jolyon Aarons and Chris-Kriton Skylaris. Electronic annealing fermi operator expansion for dft calculations on metallic systems. *The Journal of Chemical Physics*, Submitted 23-Aug-2017[224].

Abstract

Density Functional Theory (DFT) calculations with computational effort which increases linearly with the number of atoms (linear-scaling DFT) have been successfully developed for insulators, taking advantage of the exponential decay of the one-particle density matrix. For metallic systems, the density matrix is also expected to decay exponentially at finite electronic temperature and linear-scaling DFT methods should be possible by taking advantage of this decay. Here we present a method for DFT calculations at finite electronic temperature for metallic systems which is effectively linear-scaling ($O(N)$). Our method generates the elements of the one-particle density matrix and also finds the required chemical potential and electronic entropy using polynomial expansions. A fixed expansion length is always employed to generate the density matrix, without any loss in accuracy by the application of a high electronic temperature followed by successive steps of temperature reduction until the desired (low) temperature density matrix is obtained. We have implemented this method in the ONETEP linear-scaling (for insulators) DFT code which employs local orbitals that are optimised in situ. By making use of the

sparse matrix machinery of ONETEP, our method exploits the sparsity of Hamiltonian and density matrices to perform calculations on metallic systems with computational cost that increases asymptotically linearly with the number of atoms. We demonstrate the linear-scaling computational cost of our method with calculation times on Palladium nanoparticles with up to ∼13,000 atoms.

6.1 Introduction

To compute a density matrix in Density Functional Theory (DFT) calculations, an occupancy function is applied to the Hamiltonian matrix. In DFT calculations for metallic systems, the Mermin finite electronic temperature formulation of Kohn-Sham DFT is employed. In this formalism the density matrix can be obtained by diagonalizing the Hamiltonian matrix, applying a sigmoidal function, such as the Fermi-Dirac function to the energy eigenvalues and then using the Hamiltonian eigenvectors to transform these occupancy eigenvalues back into the original space of the Hamiltonian. This direct approach is, however, dependent on a cubically scaling diagonalization, so it becomes rapidly intractable as the number of basis states increases.

In the case where the Hamiltonian matrix is very large and sparse, many operations, including matrix products and even inversions may be performed in a much reduced computational complexity [40]. In such cases, it could be beneficial to avoid the eigendecomposition completely by forming a matrix function analogue of the scalar occupation function which constructs a density matrix, given a Hamiltonian matrix argument. The sigmoidal function is likely to be non-linear, so an approximation to it can be made based on linear operations (matrix products, for instance) which operates on a pre-defined domain. Such methods are often known as Fermi operator expansions (FOE) [128], when the Fermi-Dirac function is used as the occupancy function, though they are generalizable to any occupancy smearing function[92]. The functional form of the expansion is also flexible, with options including Taylor expansions, among various improved alternatives. Such methods[92, 119, 127–129, 225, 226], have been developed with the aim that the pre-factors involved in computing the resultant matrices are lower. The prefactors are still, however, large enough to make FOE methods impractical for systems involving fewer than tens of thousands of electronic states (bands or molecular orbitals) to solve for.

In this work we present a new method for DFT calculations on metallic systems where a fixed expansion length of the FOE is always employed, without any loss in accuracy. This is achieved by the application of a high electronic temperature followed by successive steps of temperature halving until the desired (low) temperature density matrix is obtained via a quenching approach. We call this method the Annealing and QUenching

Algorithm FOE or AQuA-FOE and we have implemented it in the ONETEP [40] linear-scaling (for insulators) DFT code which employs local orbitals that are optimised in situ (Non-orthogonal Generalised Wannier Functions, NGWFs) and provide large basis set accuracy in the calculations. Making use of the sparse matrix machinery of ONETEP [227–229], our method exploits the sparsity of Hamiltonian and density matrices (which is expected to decay exponentially for metallic systems at high temperatures [46]) to perform calculations on metallic systems with reduced computational cost. Our method is expected to be linear-scaling with the number of atoms, for large enough systems, and we attempt to reach this limit in this work, with timings for self-consistent iterations on metallic nanoparticles with up to ~13,000 atoms. An integral part of our method is a sparse matrix algorithm for finding the chemical potential and for calculating the electronic entropy, as these are also essential to finite temperature DFT.
In sections II-IV we describe the general theory behind our new method without the complexities of non-orthogonal bases and sparse matrices. These are introduced in sections VI and VII. In section VII we provide numerical validation by comparing calculations against the standard diagonalisation-based EDFT approach of ONETEP on metallic nanoparticles and then explore the timings of AQUA-FOE on Pd nanoparticles ranging from ~2,400 to ~13,000 atoms. We finish with some conclusions and thoughts for future applications of this method.

6.2 Kohn-Sham DFT in the Mermin Finite Temperature Formulation

For calculations of metallic systems, we must work with electrons with fractional occupancy distribution. This can be achieved by using the finite temperature Kohn-Sham equations [230] inspired by the Mermin formulation of DFT[81]. In this prescription, we minimize the Helmholtz free energy of the interacting electronic system, which is expressed as:

$$\begin{aligned} A[T,\{\varepsilon_i\},\{\psi_i\}] = \sum_i f_i \langle\psi_i|\hat{T}|\psi_i\rangle + \int \upsilon_{\text{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \\ + E_H[n] + E_{xc}[n] - TS[\{f_i\}], \end{aligned} \tag{6.1}$$

this consists of the kinetic energy of the non-interacting electrons, expressions for the external potential energy and Hartree energy of the electrons and the unknown exchange-correlation energy expression. The entropic contribution to the electronic free energy $-TS[\{f_i\}]$ is included.

To calculate the occupancies of the electronic states with finite temperature, the Fermi-Dirac (F-D) distribution can be used

$$f(\epsilon, \mu, \beta) = \frac{1}{1 + e^{(\epsilon-\mu)\beta}} \ . \tag{6.2}$$

Comparing this with the logistic function;

$$l(x) = \frac{1}{1 + e^{x}}, \tag{6.3}$$

it can be seen that the F-D function can be written in terms of the logistic function where $x \to (\epsilon - \mu)\beta$. This is useful because the logistic function can, in turn, be written in terms of the hyperbolic tangent. This representation allows the use of trigonometric identities to simplify computations, as we will see later.

$$l(x) = \frac{1}{2}\left(1 + \tanh\left(\frac{x}{2}\right)\right). \tag{6.4}$$

The finite temperature formulation of Kohn-Sham density functional theory results in occupancies of states which follow the F-D distribution. The energies of the orbitals are the eigenvalues of the Hamiltonian matrix which is often non-diagonal. Their occupancies, which follow the F-D distribution are the eigenvalues of the density matrix, which is also non-diagonal and it is essentially the matrix F-D function of the Hamiltonian.
If we have a Hamiltonian $\mathbf{H}$, then its eigenvalue expansion is:

$$\mathbf{H} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\dagger}, \tag{6.5}$$

where $\mathbf{Q}$ is a unitary matrix of eigenvectors, and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues. If we know the largest absolute eigenvalue ($|\epsilon|_{\max} = \alpha$), we can say that the eigenspectrum of $\mathbf{H}$ lies in the interval $[-\alpha, \alpha]$. As the occupancy function will be applied to all eigenvalues, any approximation to it must be accurate to some tolerance within this interval.
In this work, we show that in order to compute efficiently an FOE-type expansion, the range of energy eigenvalues can be scaled to reduce its spectral radius and then the resulting density matrix can be quenched to have the occupancy eigenvalues corresponding to the original energy eigenvalues. This is effectively an annealing and quenching procedure on the electronic temperature of the system and is done using solely matrix multiplication and without need for any matrix diagonalisation or inversion. So this approach, which we have named the Annealing and Quenching Algorithm for FOE, or AQuA-FOE could result in linear-scaling computational effort if the matrices involved have sufficient sparsity and sparse matrix storage and multiplication algorithms are employed.

Before getting to this we will describe another approach to compute an FOE-type expansion in an efficient way, where the range of energy eigenvalues can be partitioned into three intervals. These intervals will map into occupancies of 0, 1, and values between 0 and 1, which is where the FOE expansion will take place. We will then describe why this method which we call the Padé approximant and Partitioning FOE, or PaPa-FOE was put on hold for further work, and we have prioritised the advancement of the AQuA-FOE method instead.
Whichever FOE method is used to generate density matrices from Hamiltonian matrices, the expansion is valid only within a given energy eigenvalue interval. For accurate results, a large number of terms in the expansion must be used, such that the expansion is valid over the full range of energy eigenvalues of the system.

6.3 Eigenvalue subspace partitioning

One observation we can make of the eigenvalue spectrum of density kernels with Fermi-Dirac state occupancies is that after $\sim \pm 15$ smearing units in the argument of the distribution function, that the eigenvalues are less than 10^{-16} away from 1 and 0. This means that we only need an FOE which represents well the distribution between -15 and 15, outside of this range, we could find the matrices which have these eigenvalues using a conventional technqiue leading to idempotency such as purification or other linear scaling methods for insulators discussed in chapter 1.
To do this we would need to partition the eigenspectrum of Hamiltonian matrices without diagonalising. This can be achieved through the use of projection matrices, projecting out eigenvalues above and below threshold. The projection matrix can be created using a purification technique and then applied by multiplying by the Hamiltonian matrix to partition the Hamiltonian eigenvalue spectrum. The full method we have devised is presented in the following section.

6.3.1 Eigenvalue Subspace Partitioning with reduced prefactor

The Hamiltonian matrix is firstly shifted so that the trial chemical potential is at zero and scaled into units of smearing widths,

$$\mathbf{H}' = (\mathbf{H} - \mu\mathbf{I})\beta, \tag{6.6}$$

(see fig 6.1.a). The next step is to perform a step function of $\mathbf{H}'$

$$\mathbf{P}_\mu = \theta(-\mathbf{H}'), \tag{6.7}$$

(see fig 6.1.b) which can be computed using a recursive matrix multiplication based approach such as the stable Newton-Schulz sign function iteration of Chen and Chow[231]. The resulting $\mathbf{P}_\mu$ matrix is idempotent and has the correct eigenvalues (occupancies) of the zero temperature density matrix. The matrix $\mathbf{P}_\mu$ does not, in general, follow the F-D distribution around μ, as it is idempotent. However, this matrix has the correct eigenvalues of the finite temperature density matrix beyond the interval $[-c, c]$ of eigenvalues of $\mathbf{H}'$ for a certain cutoff threshold c. This is because outside this interval, the F-D function converges to 1 and 0 to machine precision.
The $\mathbf{P}_\mu$ matrix can be scaled and shifted to produce the signum matrix $\mathbf{G}_\mu$ with eigenvalues g_j,

$$\mathbf{G}_\mu = \mathrm{sgn}(\mathbf{H}') = \mathbf{I} - 2\mathbf{P}_\mu, \tag{6.8}$$

(see fig 6.1.c) which allows for the calculation of the absolute Hamiltonian (the matrix with eigenvalues as the absolute values of those of the Hamiltonian). This operation is achieved by multiplying the Hamiltonian from the left by the signum matrix (see fig 6.1.d)

$$\mathbf{H}_{\mathrm{abs}} = \mathbf{G}_\mu \mathbf{H}'. \tag{6.9}$$

A projector into the eigenvalue domain below $\mu + c$ is produced by performing a second purification on the absolute Hamiltonian at $\mu + c$ (see fig 6.1.e),

$$\mathbf{P}_{\mu+c} = \theta(c\,\beta\,\mathbf{I} - \mathbf{H}_{\mathrm{abs}}), \tag{6.10}$$

likewise, the projector into the eigenvalue domain above the $\mu + c$ threshold can be constructed trivially (see fig 6.1.f) as

$$\mathbf{R}_{\mu+c} = \mathbf{I} - \mathbf{P}_{\mu+c} \tag{6.11}$$

A matrix with eigenvalues equal to those of the absolute Hamiltonian matrix for those of its eigenvalues below the $\mu + c$ threshold and equal to zero for eigenvalues above this threshold can be calculated by multiplying by the former projector,

$$\mathbf{H}_{|\epsilon|<c} = \mathbf{H}_{\mathrm{abs}} \mathbf{P}_{\mu+c}, \tag{6.12}$$

(see fig 6.1.f). A series expansion of the electronic smearing distribution of any type (for instance Chebyshev, contour integral, recursive, etc.) is then applied to $\mathbf{H}_{|\epsilon|<c}$, which effectively applies the function to the matrix with absolute eigenvalues of $\mathbf{H}'$ in an interval of $2c$ about μ.

$$\mathbf{K}_{|\epsilon|<c} = \tanh(\mathbf{H}_{|\epsilon|<c}), \tag{6.13}$$

(see fig 6.1.g). By applying this procedure, the expansion of the smearing function need

only be computed in the positive argument domain. At this point, the matrix (6.13) has the absolute eigenvalues of the density matrix in the $[-c, c]$ energy range. To correct the sign of the eigenvalues within this interval, a multiplication by the signum matrix is performed (fig 1.h)

$$\mathbf{K}_{-c<\epsilon<c} = (\mathbf{I} + \mathbf{K}_{|\epsilon|<c}\mathbf{G}_{\mu})/2. \tag{6.14}$$

The contribution from the outer domains (those eigenvalues corresponding to energy eigenvalues outside of the interval $[\mu - c, \mu + c]$) can be included by forming

$$\mathbf{K} = \mathbf{K}_{-c<\epsilon<c} + (\mathbf{R}_{\mu+c}\mathbf{G}_{\mu} + \mathbf{I})/2. \tag{6.15}$$

(See fig 1.i) The advantages of this approach are that: (1) fewer terms are required in the series expansion of the smearing function because we are only considering the positive domain, and (2) there is a fixed upper bound on the number of terms at a given machine precision as determined by the chosen interval $[-c, c]$.

6.4 A Padé approximant based Fermi operator expansion

A Padé approximant is a rational approximation to a function $F(x)$ where the numerator and denominator are power series expansions truncated after m and n terms, respectively:

$$F(x) \approx [m/n]_F(x) = \frac{q^{(m)}(x)}{r^{(n)}(x)}, \tag{6.16}$$

in common notation. Padé approximants are desirable because they converge much more rapidly with additional expansion terms than other power series expansions [232]. In this work, we will be forming the Padé approximant of the Fermi operator as a matrix function.

Our Padé expansion of the Fermi function (which is equivalent to a tanh function) is centred at the chemical potential. We can find the numeric range away from the chemical potential that a matrix Padé approximant with certain m and n values can accurately represent $\tanh(x)$ on a particular computer system, by computing the error for the scalar functions as below:

$$\|[m/n]_{\tanh}(\alpha) - \tanh(\alpha)\| = \varepsilon^{[m/n]}(\alpha), \tag{6.17}$$

and noting that in 64bit (Intel x86_64) double precision floating point arithmetic, the coefficient (denoted $c_i^{[m/n]}$) of the highest term that is representable exactly is the 13th, then

$$\bar{\alpha} \rightarrow max(\alpha) \in \mathbb{R} \,|\, \varepsilon^{[13/13]}(\alpha) < t. \tag{6.18}$$

where t is some threshold, typically set to the square root of machine ε [233].

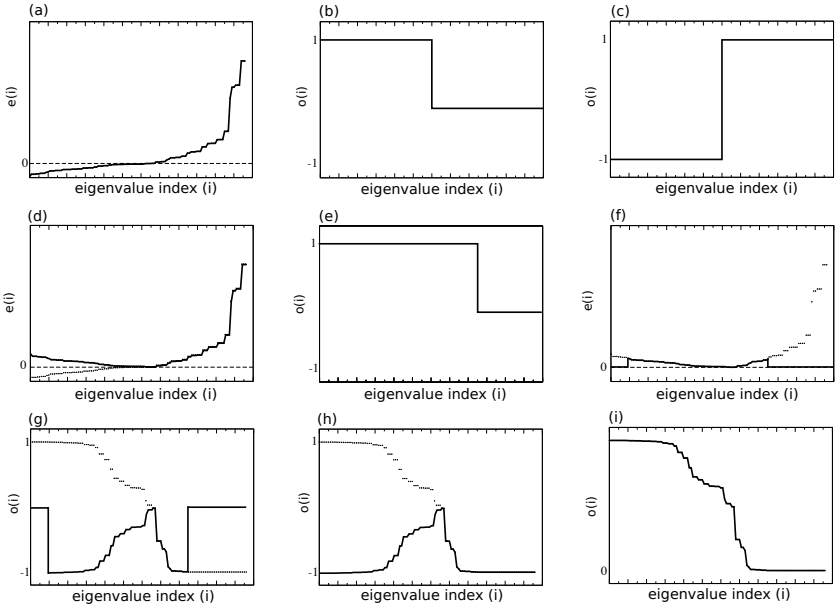


FIGURE 6.1: The effect on the eigenvalues of the Hamiltonian matrix of applying the steps of section 6.3.1. The sorted eigenvalue profile of a typical Hamiltonian matrix are shown in (a), with energy on the y axis and the chemical potential μ highlighted at the dotted line. In (b), the eigenspectrum of a projection matrix (in the same basis as the Hamiltonian matrix) with eigenvalues of 1 corresponding to those eigenvalues in the Hamiltonian matrix beneath the chemical potential. A signum matrix is produced from (b) by subtracting $\mathbf{I}/2$ and multiplying by -2; its eigenspectrum is shown in (c). This matrix can then be multiplyed by $\mathbf{H}$ to give the absolute Hamiltonian matrix, which is a matrix in the basis of the Hamiltonian, but with absolute values of the Hamiltonian matrix' eigenvalues (d). Another projection matrix is calculated from the Hamiltonian matrix, but this time at $\mu + c$ (e). This matrix is then used to send eigenvalues of the absolute Hamiltonian matrix to zero (project them out) if they are above $\mu + c$ in energy, shown in (g). A Fermi-Dirac distribution up to at least $\mu + c$ can then be applied to this matrix using the series expansion, giving the result in (h). Finally the resultant matrix can be summed with the additive inverse of the projection matrix in (e), which is shown in (f) and then multiplied by the signum matrix (c) to assign the correct sign to the eigenvalues.

In the case where $|\alpha| > \bar{\alpha}^{([m/n])}$, we need to scale the eigenspectrum of $\mathbf{H}$ by some factor, chosen to be a power of 2, so that we can recover the desired result with the hyperbolic double angle formula; the minimum scaling is given by:

$$\text{ceil}\left(\frac{|\alpha|}{\bar{\alpha}^{(d)}}\right) = n. \tag{6.19}$$

At this point, we have

$$(\mathbf{I} + e^{-\mathbf{H}'/2^{n-1}})^{-1} = \frac{1}{2}\left(\mathbf{I} + \tanh\left(\frac{\mathbf{H}'}{2^n}\right)\right). \tag{6.20}$$

To recover the desired result, we need to apply the hyperbolic double angle formula,

$$\tanh(2\mathbf{H}') = \frac{2\tanh(\mathbf{H}')}{\tanh^2(\mathbf{H}') + \mathbf{I}}, \tag{6.21}$$

$n-1$ times. This does not, however, involve computing multiple inverses, as multiple angle formulae can be constructed, for example, for $4\mathbf{H}$:

$$\tanh(4\mathbf{H}') = \frac{4\tanh^3(\mathbf{H}') + 4\tanh(\mathbf{H}')}{\tanh^4(\mathbf{H}') + 6\tanh^2(\mathbf{H}') + I}. \tag{6.22}$$

Obviously, when we have computed $\mathbf{X} = \tanh((\mathbf{H} - \mu\mathbf{I})\beta/2^n)$ using the Padé approximant, it would be greatly advantageous to minimise the number of further applications of the Padé approximant, inverses and matrix products, in this order. The Padé approximant, at the 13th order expansion, requires 12 matrix multiplications and one inverse and is thankfully only needed once per evaluation. The iterative inverse algorithm (Newton-Schulz-Hotelling) can be easily initialised with the result from the previous step, resulting in a great reduction in the number of matrix multiplications required.
For example, if $n = 8$ was required to keep the eigenspectrum of the Hamiltonian within the range that the Padé approximant can reproduce, the procedure would be as follows: Calculate Padé approximant of scaled matrix,

$$\mathbf{X} = \text{Padé}^{[13/13]}_{\tanh}((\mathbf{H} - \mu\mathbf{I})\beta/2^8), \tag{6.23}$$

use the $n = 8$ multiple angle formula to get the solution to the non-scaled problem $(\tanh(\mathbf{H} - \mu\mathbf{I}))$,

$$\tanh(\mathbf{H} - \mu\mathbf{I}) = \frac{8\mathbf{X}^7 + 56\mathbf{X}^5 + 56\mathbf{X}^3 + 8\mathbf{X}}{\mathbf{X}^8 + 28\mathbf{X}^6 + 70\mathbf{X}^4 + 28\mathbf{X}^2 + \mathbf{I}}, \tag{6.24}$$

and finally map back on to the logistic function (eqn.6.4).

6.4.1 Combining eigenvalue partitioning with Padé Fermi operator expansion

In practice, when calculating the finite-temperature density matrix, we use a combination of the Padé and the eigenvalue partitioning approaches. While the eigenvalue partitioning technique admits any expansion for the central eigenvalue domain, we find the Padé scheme to be efficient and comparable in terms of matrix multiplications to the Chebyshev expansion with divide and conquer [92].
When combining the Padé and eigenvalue partitioning in such a way, there is no need use a hyperbolic multiple angle formula except if one requires machine precision. In the case of 64bit precision, a 13 term Padé approximant combined with a triple angle formula are enough to give accurate eigenvalues up to the point of maximal and minimal occupancy in the interval with fractional occupancies. Outside of this interval, the (trivial) fully occupied and unoccupied eigenvalue intervals are handled separately, as they are created by the eigenvalue partitioning technique.

6.5 Issues with the Eigenvalue Partitioning approach and Electronic quenching

What we have found in practice, when applying our implementation of the PaPa-FOE technique described above is that our projection matrices are not sufficiently accurate to get good accuracy. When we multiply the Hamiltonian matrix by a projection matrix $\mathbf{P}$, and multiply the Hamiltonian matrix by $(\mathbf{I} - \mathbf{P})$, we would expect that the sum of the two resultant matrices has the eigenvalues of the original Hamiltonian matrix. Instead, we find that the eigenvalues are perturbed, particularly around the partition. This problem is excentuated when we apply functions to the eigenspectrum partitions. Since we are dealing with non-linear functions, when we apply the FOE to the central partition Hamiltonian matrix and sum this density matrix of the central partition with the density matrix for the partition with full occupancy, some eigenvalues go above one. The situation on the zero occupancy can be worse, leading to some eigenvalues going negative! For this reason we opted to create another scheme, which is similar in spirit, but does not suffer from the same issues with accuracy.

6.5.1 Electronic quenching

We propose that rather than increasing the number of terms for systems of increasing eigenvalue spectrum width (which is exacerbated in low electronic temperature), a fixed-length FOE is used which is valid on a preordained interval, but the electronic temperature is increased by a sufficient multiple of the target temperature so that the full spectrum of the hot-electron Hamiltonian lies within the interval of validity of the FOE. Then, the low (target) temperature density matrix can be recovered, as we will show in this section.
The Hamiltonian matrix is firstly shifted so that the trial chemical potential is at zero and scaled into units of smearing widths,

$$\mathbf{H}' = (\mathbf{H} - \mu\mathbf{I})\beta \ . \tag{6.25}$$

The scaled and shifted Hamiltonian matrix is then annealed by dividing by a sufficiently large power of 2:

$$\mathbf{H}_{\text{hot}} = \mathbf{H}'/2^n, \tag{6.26}$$

so that its spectrum lies on the desired interval as the temperature of this “hot" Hamiltonian goes from $\beta = 1/T$ to $\beta_{\text{hot}} = 1/(2^n T)$. The exponent, n can be determined as

$$n = \text{ceiling}\left(\log_2(\rho_{\mathbf{H}'}/c)\right), \tag{6.27}$$

where $\rho_{\mathbf{H}'}$ is the spectral radius (largest absolute eigenvalue) of $\mathbf{H}'$ and c is the domain of energy eigenvalues within which the FOE is valid. $\rho_{\mathbf{H}'}$ can be determined either by using the Gershgorin circle theorem or using a power series iteration, as we do in this work. The chosen FOE is then applied to $\mathbf{H}_{\text{hot}}$ to obtain $\mathbf{K}_{\text{hot}}$. Then $\mathbf{K}_{\text{hot}}$ is annealed n times to obtain the density matrix corresponding to the low (target) temperature Hamiltonian matrix at β. To do this, the density matrix is firstly transformed to have eigenvalues of the hyperbolic tangent function:

$$\mathbf{R}_{\text{hot}} = 2\mathbf{K}_{\text{hot}} - \mathbf{I} \ , \tag{6.28}$$

or in other words the range of the eigenvalues of the density matrix is scaled from [0,1] to [-1,1]. To this, the matrix analogue of the hyperbolic double angle formula can be applied:

$$\mathbf{R}_{\text{new}} = \frac{2\mathbf{R}_{\text{old}}}{(\mathbf{I} + \mathbf{R}_{\text{old}}^2)}. \tag{6.29}$$

This formula is first applied with $\mathbf{R}_{\text{old}} = \mathbf{R}_{\text{hot}}$ and repeated n times to give the desired density matrix. With each iteration of the above formula the temperature of the density matrix is halved. The target temperature density matrix can then be recovered as

$$\mathbf{K} = (\mathbf{R} + \mathbf{I})/2. \tag{6.30}$$

It is worth noting that to apply equation 6.29 requires an inversion or to compute the solution to a linear equation. To avoid this, we take a Chebyshev expansion of equation 6.29 which requires 37 terms to reach machine precision, although accuracy of 10^{-9} can be obtained with only 25 terms and this can be evaluated with as few as 12 matrix multiplications using the divide and conquer approach of Head-Gordon [92]. We should note here that the Chebyshev expansion of the hyperbolic double angle formula that we discuss here is different from the Chebyshev expansion of the Fermi-Dirac distribution which can be used to compute $\mathbf{K}_{\text{hot}}$. It is interesting to note that here we use a recursion in the temperature scaling of the FOE, which is itself computed using a Chebyshev expansion, while Niklasson *et al* proposed a recursion method directly for the FOE[226, 234]. Niklasson's recursion method could be used instead to compute the high temperature FOE.

For a given chemical potential, the scaling in terms of number of matrix products to apply this algorithm is given by:

$$N_{\text{MP}} = N_{\text{FOE}}(c) + nN_{\text{HI}}, \tag{6.31}$$

where N_{FOE} is the number of matrix products required to compute the FOE, N_{HI} is the number of matrix products required to compute an application of the hyperbolic double

angle formula as a Chebyshev expansion. The number of matrix products required to apply the FOE depends on c, but in practice this is constant and we always use a value of 15 smearing widths (β_{hot}) as there is no need to change this aside from optimization of the total number of matrix multiplications. This is a compromise between the FOE length and number of hyperbolic double angle formula evaluations. The hyperbolic double angle formula is always evaluated with a fixed number of terms in the Chebyshev expansion and hence, a fixed number of matrix multiplications. So, the only variable in the total number of matrix products is the spectral width of the Hamiltonian matrix, $\rho_{\mathbf{H}'}$.
As the number of atoms increases, if the material is homogeneous, we do not expect the spectral width of the Hamiltonian matrix to increase as it asymptotically reaches the bulk value. However, as an extreme upper bound, if the spectral width were to increase linearly with system size, then the amount of matrix multiplications required to apply this algorithm would only increase logarithmicaly, according to equation 6.27. In practice, this increase will be somewhere between zero and logarithmic. So, with a non-increasing, or logarithmically increasing number of matrix multiplications for an increase in system size (number of atoms), if the matrix multiplication can be made to be linear-scaling with dimension, as is the case for sparse matrices in the ONETEP linear-scaling DFT program for sufficient sparsity, then the whole algorithm may be viewed as having linear-scaling computational cost with system size.

6.6 Combining Hamiltonian Annealing and Quenching with a Chebyshev Fermi Operator Expansion

The Chebyshev FOE, as described originally by Goedecker and Teter [115] is achieved by taking a Chebyshev expansion of the 1D Fermi-Dirac function:

$$\{a_i\} = \text{DCT}\left(\frac{1}{1+e^{\cos(x_i)}}\right), \tag{6.32}$$

where $x_i = 2(((e_i - \mu)\beta) - e_0)/(e_P - e_0) - 1$ so that the range of equispaced e_i covers at least the interval of the negative to positive spectral radius of the Hamiltonian matrix (e_0:e_P). The interval is scaled and shifted to cover the useful interpolative range of Chebyshev polynomials ($-1:1$). DCT() refers to the Discrete Cosine Transform operation. $\{a_i\}$ are the Chebyshev expansion coefficients. The density matrix is formed by taking Chebyshev polynomials of the Hamiltonian matrix and summing using the weights, a_i:

$$\mathbf{K}(\mathbf{H}') = \sum_{i=0}^{P} a_i \mathbf{T}_i(\mathbf{H}'), \tag{6.33}$$

where $\{\mathbf{T}_i\}$ are the Chebyshev matrices of the first kind, of degree i. The Chebyshev matrices are in the standard form:

$$\begin{aligned} \mathbf{T}_0(\mathbf{H}') &= \mathbf{I} \\ \mathbf{T}_1(\mathbf{H}') &= \mathbf{H}' \\ \mathbf{T}_{n+1}(\mathbf{H}') &= 2\mathbf{H}'\mathbf{T}_n(\mathbf{H}') - \mathbf{T}_{n-1}(\mathbf{H}'). \end{aligned} \tag{6.34}$$

This assumes that either $\mathbf{H}'$ is in an orthogonal basis or it has been orthogonalised prior to application. If non-orthogonal basis functions are used, then $\mathbf{H}'_{\alpha\beta}$ is a covariant quantity and $\mathbf{K}^{\alpha\beta}$ should be contravariant. This can be achieved by raising one index of the Hamiltonian matrix, $\mathbf{H}'^{\alpha}{}_{\beta}$ either by multiplying by the inverse overlap matrix or solving $\mathbf{S}_{\gamma\alpha}\mathbf{H}'^{\alpha}{}_{\beta} = \mathbf{H}'_{\gamma\beta}$, to make it contra-covariant. In so doing, the Chebyshev products are all well defined and the resultant $\mathbf{K}^{\alpha}{}_{\beta}$ matrix is also contra-covariant. To make the density matrix fully contravariant in the form necessary for computing the electronic density, the column index can be raised by, for instance multiplying by the inverse overlap matrix on the right, or solving another linear matrix equation – this is explored in more detail in section 6.8.

This way of performing the expansion requires approximately P terms for a given accuracy, where P is a function of the smearing width, β, the required accuracy 10^{-D} and the Hamiltonian spectral width, ρ. We use the improvement to the original Chebyshev series of Goedecker, proposed by Liang and Head-Gordon in 2003; this uses a divide and conquer approach to re-sum the terms of a truncated Chebyshev series[92]. Using this scheme reduces the cost of a Chebyshev representation of the Fermi operator to $\mathrm{O}(\sqrt{P})$ number of matrix multiplications. In the best case scenario of a local orbital method where sufficient matrix sparsity leads to a cost of matrix multiplication being $O(N)$ this would lead to a cost per SCF iteration proportional to $O(N^{3/2})$.

In order to reduce this cost from $O(N^{3/2})$ to $O(N)$ we limit the width of the FOE to the constant interval $[-c, c]$ as we mention in section 6.5. We apply this FOE to a high temperature Hamiltonian, with a spectral radius less than c. In effect this means that we keep on raising the temperature with Hamiltonian spectral radius while keeping the FOE expansion length small and independent of $\rho_{\mathbf{H}'}$. Then the quenching formula of equation equation 6.29 is applied n times which is proportional to spectral width of the system by equation 6.27 so the overall scaling per SCF iteration is $O(N \log_2(\rho_{\mathbf{H}'}/c))$.

To avoid the inversion in equation 6.29, we also evaluate it using a Chebyshev expansion of equation 6.29, however this always has a constant number of terms (fixed length) for a given accuracy because the eigenvalues of tanh are already in the interval [-1:1].

6.7 Electronic Entropy

Since we have a finite electronic temperature, the energy which is minimised to find the ground state of a system is the Helmholtz free energy. Therefore, an electronic entropy term must be calculated.
The entropic contribution of the i*th* state is well known, as we are dealing with non-interacting, finite temperature Fermions:

$$s_i = f_i \ln(f_i) + [1 - f_i]\ln(1 - f_i), \tag{6.35}$$

on a state-by-state basis where f_i is the occupancy of the ith state. So, if the eigenvalues of the Hamiltonian matrix and hence the occupancies were easily accessible, then calculating the entropy would be simply a matter of using (6.35) for every occupancy, f_i and then summing them to get the contribution to the entropy:

$$\mathcal{S} = \sum_i s_i \ . \tag{6.36}$$

In the case where we are performing a Fermi-operator-expansion and eigenvalues are not available, another approach must be taken. In this case we aim to calculate an entropy matrix, in the same basis as the density matrix, with eigenvalues s_i.
This can be achieved in a number of ways but conceptually, using a matrix logarithm is the simplest method. Analogously to (6.35), the entropy matrix can be calculated as:

$$\mathcal{S} = \mathrm{tr}[\mathbf{K}\ \ln(\mathbf{K}) + [\mathbf{I} - \mathbf{K}]\ \ln(\mathbf{I} - \mathbf{K})] \ , \tag{6.37}$$

which concentrates all of the computation to the matrix logarithm function and the entropic contribution to the free energy can be written as

$$-T \sum_i s_i = -T\mathcal{S}. \tag{6.38}$$

The difficulty arises when attempting to compute the matrix logarithm, two of which must be performed to evaluate (6.37). Firstly, in order to calculate a matrix logarithm, the matrix argument must be positive definite, but $\mathbf{K}$ necessarily has eigenvalues close to zero by construction, which may be indistinguishable from zero on a finite precision computer.
In order to overcome this limitation, those eigenvalues closer to zero than some pre-defined threshold can be projected out, with the justification that the corresponding eigenvalue from the pre-multiplicative matrix, $\mathbf{K}$ or $\mathbf{I} - \mathbf{K}$, would have zeroed out this very large negative number.

Performing this operation adds a good deal of computation onto an already expensive calculation, as two extra step function projection operations must be performed for every electronic entropy calculation.

Computing the matrix logarithm is a very computationally demanding operation in itself. One way to calculate a matrix logarithm is by using the Mercator series

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \tag{6.39}$$

which can be applied to a general matrix $\mathbf{A}$ by linearity:

$$\ln(\mathbf{A}) = \ln(\mathbf{I}+\mathbf{X}) = \mathbf{X} - \frac{\mathbf{X}^2}{2}\cdots \tag{6.40}$$

Doing this in practice is ill-advised due to the poor convergence properties of the series. An alternative proposed by Kenney and Laub[235], which has been rigorously explored and revised in [236] is to use an inverse scaling and squaring approach. This is roughly the scaling and squaring algorithm for calculating matrix exponentials with the operations in reverse order.

The steps in the inverse scaling and squaring algorithm are firstly to calculate $\mathbf{B} = \mathbf{A}^{1/(2^m)}$, where m is sufficiently large that the result $\mathbf{B}$ is arbitrarily close to an identity matrix. $\ln(\mathbf{B})$ can then be approximated by an n^{th} order Padé approximant to $\ln(1+x)$ as

$$\ln(\mathbf{B}) \approx \mathrm{pa}^{[m/m]}(\mathbf{B}-\mathbf{I}), \tag{6.41}$$

so that the final result may be calculated as:

$$\ln(\mathbf{A}) = 2^m \ln(\mathbf{B}). \tag{6.42}$$

Applying this algorithm, presented by Higham[237] costs $(12 + 2m - 2/3)$ matrix multiplications, which may be a huge expense, depending on the cost of matrix products. An alternative to this approach is given in Hale, *et al*[238] which computes the logarithm in terms of inversions, via a contour integral. This approach may be less expensive depending on the cost of inversion.

An alternative to the direct approach of calculating (6.37) with matrix logarithms is to find a function which approximates (6.37) with reduced computational overhead.

One possible form, which we propose here, is:

$$s(x) \simeq ax^2 + \frac{b}{c + dx - dx^2} - e - ax = y(x), \tag{6.43}$$

where $a = 1.96056$, $b = 0.0286723$, $c = 0.114753$, $d = 1.98880$ and $e = 0.249860$.

This is an inverse quadratic with a quadratic fit to the error subtracted from it. This

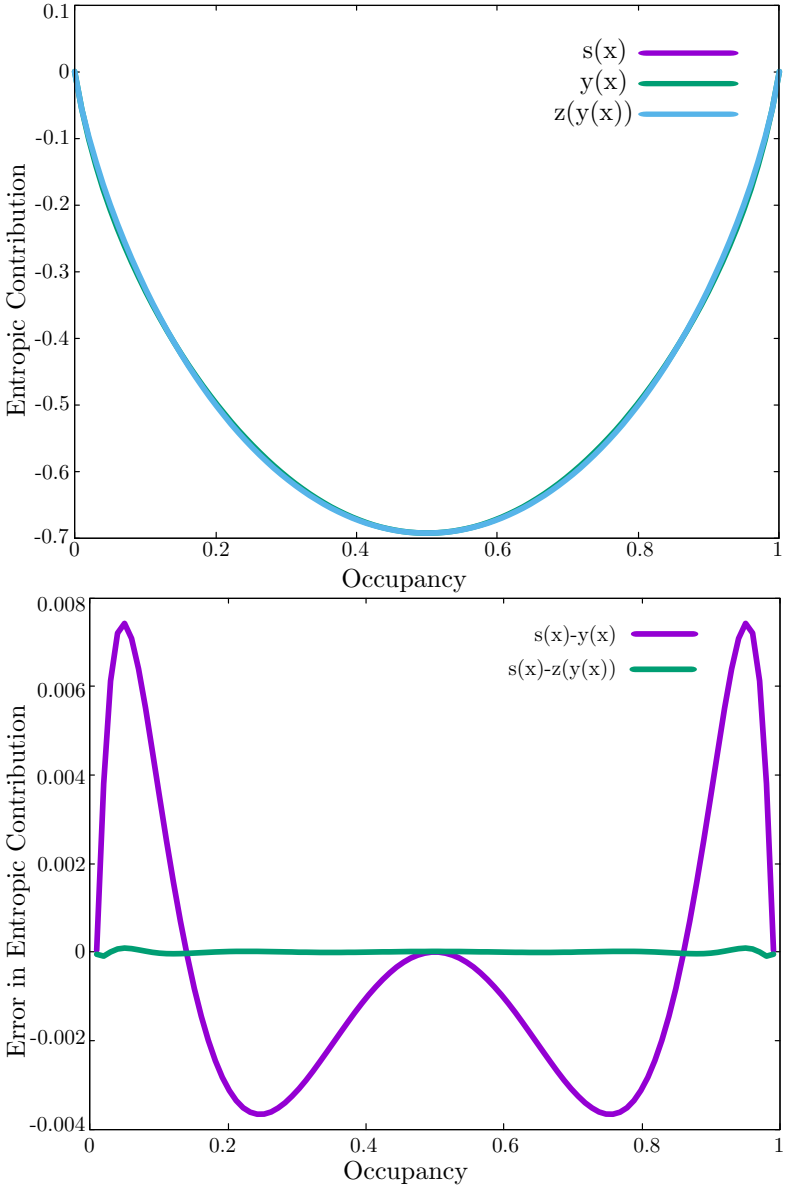


FIGURE 6.2: The two entropy approximations suggested in this work are overlayed in (a) and the relative errors shown in (b)

form approximates the entropy expression, and it has a mean squared error of 1.0×10^{-6} for each s_i approximated. A further refinement can be made to this by fitting to the error of this expression with another Padé approximant, such as

$$s(x) \simeq y + \frac{a'y + y^4 + b'y^3 + c'y^2 - d'}{e' - h'y - b'y^2 - k'y^3} = z(y), \tag{6.44}$$

where $a' = 0.01548792$, $b' = 1.1542349$, $c' = 0.3418894$, $d' = 4.964447 \times 10^{-6}$, $e' = 0.04446540$, $h' = 1.806607$, $k' = 5.0611405$ and where y are values calculated with (6.43). This form approximates the entropy expression, and it has a mean squared error of 1.0×10^{-8} for each s_i approximated.

6.8 Implementation in Terms of Non-Orthogonal Local Orbitals in the ONETEP code

A major problem in applying the techniques described in the previous sections is that the orthogonalization of the Hamiltonian matrix prior to application of the expansion method is costly and severely limits the sparsity of the orthogonalized Hamiltonian matrix. This has the knock-on effect that each of the matrix products in the expansion method is more expensive than using non-orthogonal matrices and the feasibility of the method is limited to systems where the full or almost full N^2 Hamiltonian matrix may be stored in memory.

We aim to implement our approach within the framework of the ONETEP [40] linear-scaling (up to now for insulators) DFT code which is based on a representation of the molecular orbitals in terms of a set of non-orthogonal strictly localised orbitals $\{\phi_\alpha\}$ which are optimised in situ [172]. These localised orbitals are called Non-orthogonal Generalised Wannier Functions (NGWFs) and are expressed in a basis set of psinc functions [239] which are equivalent to a plane wave basis set. In the conventional diagonalisation-based EDFT method implemented in ONETEP [7] the covariant Hamiltonian matrix $H_{\alpha\beta}$ in the representation of the NGWFs is optimised in the inner loop (the EDFT loop) of a two nested loop procedure, while the NGWFs are optimised in the outer loop using a conjugate gradients approach. ONETEP constructs the density matrix as an expansion in the NGWFs as:

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi^*_\beta(\mathbf{r}'), \tag{6.45}$$

where $K^{\alpha\beta}$ is the generalized occupancy of the NGWFs, i.e. its eigenvalues are $\{f_i\}$ and is known as the density kernel. In ONETEP EDFT, the free energy functional is optimized firstly with respect to the Hamiltonian matrix, $\{H_{\alpha\beta}\}$ while keeping the NGWFs $\{\phi_\alpha\}$ fixed. The orbital contribution to the free energy can then be minimized by optimising a projected Helmholtz functional,

$$A'[\mathrm{T};\{\phi_\alpha\}] = \min_{\{H_{\alpha\beta}\}} A[\mathrm{T};\{H_{\alpha\beta}\};\{\phi_\alpha\}] \tag{6.46}$$

with respect to $\{\phi_\alpha\}$. In conventional EDFT the density kernel can be generated explicitly as a function of the eigenvalues ε_i (obtained by diagonalisation) of the Hamiltonian as:

$$K^{\alpha\beta} = \sum_i^N M^\alpha{}_i f(\varepsilon_i) M_i^{\dagger\,\beta}, \tag{6.47}$$

where the eigenvalues (band occupancies) of the finite-temperature density kernel, $K^{\alpha\beta}$ are given in terms of the F-D smearing function and the matrix M contains the eigenvectors of the Hamiltonian eigenproblem:

$$H_{\alpha\beta}M^{\beta}_{\ i} = S_{\alpha\beta}M^{\beta}_{\ i}\,\varepsilon_i. \tag{6.48}$$

Such eigenvalue based approaches will always scale as $O(N^3)$ as they employ matrix diagonalization algorithms.

The aim of this work is to replace the diagonalisation, the building of the density kernel, the chemical potential search and the computation of the entropy in EDFT with our new AQuA-FOE method. Up to now ONETEP has used internally fully covariant and fully contravariant tensors but for our work here we will follow the contra-covariant approach suggested by Gibson, Haydock and LaFemina [240]. This involves working directly with the Hamiltonian in the natural representation, which is the contra-covariant Hamiltonian matrix, or $H^{\alpha}_{\ \beta}$.

$H^{\alpha}_{\ \beta}$ can be used in place of the H in any of the matrix products described and the result is tensorially correct, and as both matrices have the same eigenvalues, the resulting contra-covariant density matrix has the correct eigenvalues. The standard contravariant density matrix can then be formed by either multiplying on the right by the inverse overlap matrix to raise the index, or by solving

$$K^{\alpha}_{\ \beta} = K^{\alpha\gamma}S_{\gamma\beta}, \tag{6.49}$$

for $K^{\alpha\gamma}$.

The method of Gibson, *et al* works by assuming that the Hamiltonian is a local operator, and that when applied to a state, the value of the result at position $\mathbf{r}$ depends only upon the value in a small localization region surrounding the state at position $\mathbf{r}$. This must be the case because when applying the Hamiltonian operator, the contra-covariant form is the correct matrix represntation of the operator and this preserves the locality of the states. Hence, $H^{\alpha}_{\ \beta}$ is a sparse matrix, if $H_{\alpha\beta}$ is sparse even if $S^{\alpha\beta}$ is not. Using this logic, the authors suggest solving the matrix equation

$$S_{\gamma\alpha}H^{\alpha}_{\ \beta} = H_{\gamma\beta}, \tag{6.50}$$

for $H^{\alpha}_{\ \beta}$, without constructing the inverse overlap matrix, $S^{\gamma\alpha}$. This can be done in practice approximately using the same locality arguments as above by taking the sparsity pattern of the resultant contra-covariant Hamiltonian matrix as the subspace of local interactions: that is to construct small matrices for each column of the resulting matrix from the non-zero elements of $H_{\gamma\beta}$ on the mask:

$$M_i = \mathbf{v}_i\mathbf{v}_i^T, \tag{6.51}$$

where $\mathbf{v}_i$ is the column vector of nonzero elements from the ith column of $H_{\alpha\beta}$. This

"nonzero elements of the ith column squared matrix problem" can be solved to give the ith column of the resultant matrix:

$$ {}^{(i)}S_{\gamma\alpha}\,{}^{(i)}H^{\alpha}_{\ \beta} = {}^{(i)}H_{\gamma\beta}, \tag{6.52} $$

and then the elements are put back into the sparsity pattern to produce the approximate $H^{(A)\alpha}_{\ \ \beta}$.

The accuracy of the approximation can be checked trivially by using

$$ \Delta = S_{\gamma\alpha}H^{(A)\alpha}_{\ \ \beta} - H_{\gamma\beta}. \tag{6.53} $$

and the norm of Δ gives an estimate for the error. This can be minimized by increasing the size of the localization regions of the local functions.

In practice, we use the same sparsity pattern for $H^{\alpha}_{\ \beta}$ as we do for $H_{\alpha\beta}$, as was suggested in Gibson *et al.* In the calculations we have performed so far, this has proved adequate based on the error estimate given by taking $||\Delta||_F$, but if this error ever proved too large, we have the option to reduce the sparsity of $H^{\alpha}_{\ \beta}$ by multiplying the radii of the spherical localization regions by a value > 1.0 and constructing a more accurate sparsity pattern from these more delocalized regions.

The sparsity pattern of the contra-covariant Hamiltonian matrix is only one of the considerations for sparsity, however. We must also pay close attention to the sparsity pattern of the density matrices. In most linear-scaling DFT packages, the single particle density matrices are kept sparse through either a geometric cutoff between atomic centres, or a truncation by estimating the absolute magnitude of matrix elements and setting to zero, those elements which are smaller than the threshold.

In this work, we have opted to use an alternative method based on the nature of the FOE-type approaches. Since we perform a power series expansion, we opt to limit the sparsity of the contra-covariant density matrix to be no more dense than the $H^{\alpha}_{\ \beta}H^{\beta}_{\ \gamma}$ term. In so doing, we limit the accuracy of small systems, but calculations on these systems are likely to be possible with conventional methods, or with dense matrices so the compromise is worth it. In large systems, where resulting sparsity may be adequate for linear or near-linear scaling, care must be taken to ensure that an appropriate electronic smearing is used, as the error induced by the sparsity is reduced when raising the electronic temperature.

We find that for a truncated octahedral Al_{2406} nanoparticle that electronic temperature has a large effect on the achievable accuracy (fig. 6.3). At 2.5 eV smearing, we reach a maximum error of about 10^{-4} in the elements of the density kernel , whereas if we increase to 25 eV or 250 eV, we get maximum errors of 10^{-6} and 10^{-8}, respectively.

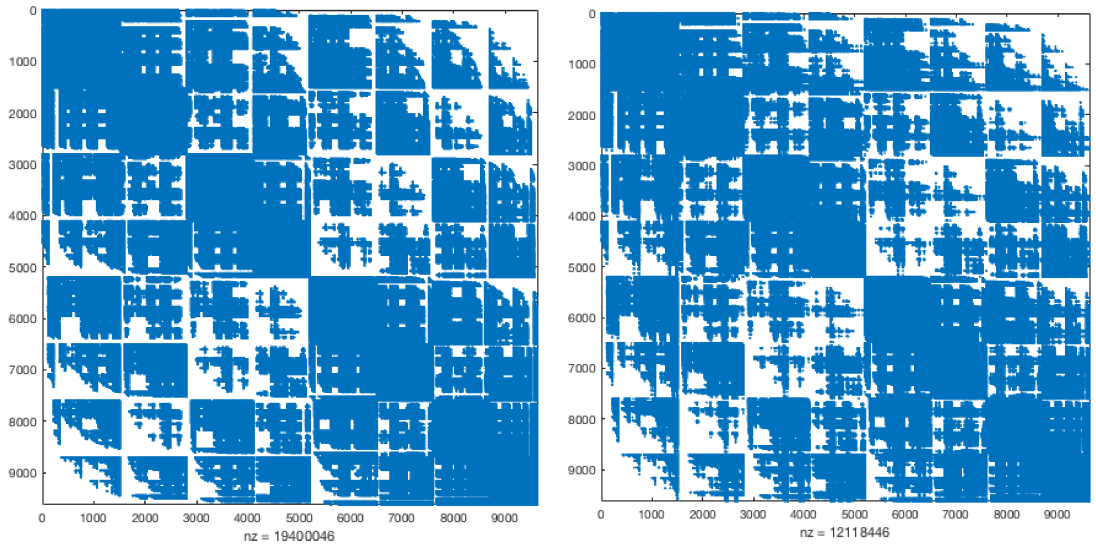


FIGURE 6.3: A comparison showing the sparsity pattern of the $H^{\alpha}{}_{\beta}H^{\beta}{}_{\gamma}$ matrix (above) and the sparsity pattern resulting from a 2.5 eV smeared density kernel calculated without truncation and then having a truncation applied to its elements which are smaller than a threshold of 10^{-4} (below). These are calculations on a truncated octahedral Al nanoparticle with 2406 atoms (4 NGWFs per atom). The axes labels in this figure represent matrix row and column indices and the number beneath each of them shows the number of non-zero elements in each.

6.9 Chemical potential search

When applying the Fermi-Dirac distribution at a particular electronic temperature to the eigenvalues of a Hamiltonian matrix, the chemical potential μ is found trivially by applying the Fermi function to all of the eigenvalues with a trial value of chemical potential μ_0, summing all of the results of these applications. The resulting scalar (number of electrons) is compared with the desired number of electrons. This process is repeated with a modification to the chemical potential until the difference between calculated and desired number of electrons is below some arbitrary threshold.

This approach is equivalent to using a root finding method on the following equation, with the chemical potential μ as the independent variable:

$$\Delta N_e = N_e - \sum_{i=1}^{N} \frac{1}{1 + e^{(\epsilon_i - \mu)\beta}}, \tag{6.54}$$

where N_e is the target number of electrons, N is the size of the eigenspace, ϵ_i are the eigenvalues of the Hamiltonian matrix, μ is the chemical potential and β is thermodynamic temperature $\beta = 1/kT$.

This direct approach is efficient if the eigenvalues are known. However, when performing a Fermi-operator-expansion (FOE) in the non-diagonal space, the eigenvalues are not known, but the number of particles can still be calculated as the trace of the density

matrix. A similar approach could then be used:

$$\Delta N_e = N_e - \text{trace}\left(\frac{\mathbf{I}}{\mathbf{I} + e^{(\mathbf{H}-\mu\mathbf{I})\beta}}\right), \tag{6.55}$$

where $\mathbf{H}$ is the Hamiltonian matrix (which is assumed to be orthogonal or be contra-covariant as described in section 6.8) and $\mathbf{I}$ is the identity matrix of dimension N. The problem with this approach is that each evaluation of the density matrix is a relatively expensive operation, which one would not want to perform many times per SCF iteration. A solution can be found by noticing that a density matrix calculated at a particular chemical potential can be modified to be at a different chemical potential at a lower cost than recalculation. This is possible because of the correspondence between the Fermi-Dirac function and the hyperbolic tangent function:

$$\frac{\mathbf{I}}{\mathbf{I} + e^{(\mathbf{H}-\mu\mathbf{I})\beta}} = \frac{1}{2}\left(\mathbf{I} + \tanh\left(\frac{(\mathbf{H}-\mu\mathbf{I})\,\beta}{2}\right)\right). \tag{6.56}$$

With this in mind, hyperbolic, trigonometric identities may be employed to compute useful quantities; for instance,

$$\tanh(x \pm y) = \frac{\tanh(x) \pm \tanh(y)}{1 \pm \tanh(x)\tanh(y)}, \tag{6.57}$$

or for matrices, where $\rho_{\mathbf{H},\mu,\beta}$ is the density matrix calculated for a Hamiltonian, $\mathbf{H}$, at a chemical potential μ, then

$$\tanh\left(\frac{((\mathbf{H}-\mu\mathbf{I})\beta)}{2} \pm \frac{\beta\Delta\mu\mathbf{I}}{2}\right) = \frac{\rho_{\mathbf{H},\mu,\beta} \pm \tanh(\frac{\beta\Delta\mu}{2})\mathbf{I}}{\mathbf{I} \pm \tanh(\frac{\beta\Delta\mu}{2})\rho_{\mathbf{H},\mu,\beta}}. \tag{6.58}$$

Provided that $\mathbf{I} \pm \tanh(\frac{\beta\Delta\mu}{2})\rho_{\mathbf{H},\mu,\beta}$ can be inverted efficiently, this method should involve far less computation than recalculation of the density matrix.

If the matrix is inverted with a Newton-Shulz-Hotelling algorithm, then each time the root-finding algorithm calls for a new trial chemical potential, then the inverse can be initialised with the inverse from the previous trial point saving significant calculation effort.

What we do in practice is to actually to use a further Chebyshev expansion of equation 6.58. The coefficients for this expansion have to be computed every time a new $\Delta\mu$ is used, but this is a relatively inexpensive operation. When computing the expansion we take the scalar form of equation 6.58 as:

$$q(x) = \frac{x + c}{1 + xc}, \tag{6.59}$$

on the domain [-1,1], where c is $\beta\Delta\mu/2$. We then evaluate the Chebyshev expansion of $\rho_{\mathbf{H},\mu,\beta}$ using these coefficients, again using the divide and conquer algorithm of Liang

and Head-Gordon, *et al.* to provide us with the density matrix at the updated chemical potential, $\rho_{\mathbf{H},\mu+\Delta\mu,\beta}$.

We can go further than this, however, and use derivatives with respect to the chemical potential to speed up the root search. The first derivative of the density matrix with respect to chemical potential, is:

$$\frac{\partial \rho}{\partial \mu} = -\frac{\beta}{4}\left(\mathbf{I} - \rho^2\right), \tag{6.60}$$

because

$$\frac{d(\tanh(x))}{dx} = \operatorname{sech}^2(x) = 1 - \tanh^2(x). \tag{6.61}$$

Furthermore, because the trace operator commutes with the differential operator, we can write:

$$\frac{\partial N_e}{\partial \mu} = -\frac{\beta}{4}\left(1 - \operatorname{trace}(\rho^2)\right). \tag{6.62}$$

As $N_e(\mu)$ is a monotonic function (when using the Fermi-Dirac distribution), the root finding ought to be simple with Newton's method. It is complicated slightly, however, because $N_e(\mu)$ has multiple stationary points which increase in breadth with Hamiltonian eigenvalue-spacing and decreasing temperature.

Using Newton's method at one of these points would send the next trial point off to ±infinity, so a safe-guarded version is used in practice, where a trust-region is defined and if Newton's method is going to send the next point outside of these bounds, then a bisection step is performed.

6.10 Validation tests

In ONETEP, we have implemented the Annealing and QUenching Algorithm - FOE (AQuA-FOE) approach that we present in this paper for calculating the density matrix from a given Hamiltonian matrix and also to calculate an electronic entropy matrix. This type of FOE is applied at every step of the ensemble-DFT (EDFT) approach, which is used as the electronic energy minimization technique.

To test the AQuA-FOE method we have carried out numerical comparisons against the conventional diagonalisation-based EDFT method which is available in ONETEP [7]. For this we use small cuboctahedral platinum nanoparticles. We performed all of these calculations in ONETEP with EDFT, at 500 eV kinetic energy cut-off, 9.0 Bohr radius localization spheres for the NGWFs, and with PAW[28], using the data from the GBRV pseudopotential dataset [174]. The platinum-platinum distance was set to the bulk value of 2.8 Å. We ran the AQuA-FOE scheme where the chemical potential search was configured to halt once the chemical potential had been found to within $10^{-6}\mathrm{E}_H$ - $10^{-8}\mathrm{E}_H$. We also used the refined approximation to the entropy which was described

in section 6.7 and an electronic smearing width of 0.1 eV. With this prescription for a cuboctahedral platinum nanoparticle with 55 atoms, the convergence in energy of the AQuA-FOE scheme with chemical potential search stopping criterion can be seen in Table 6.1. We observe rapid convergence of the energy to a value of -4914.74445 E_H while the

TABLE 6.1: The convergence in total energy of AQuA-FOE for cuboctahedral, platinum nanoparticles with respect to chemical potential search

E_H	Pt_{55}	Pt_{147}
10^{-6}	-4914.74211	-13137.32070
10^{-7}	-4914.74475	-13137.33179
10^{-8}	-4914.74445	-13137.33169

the same converged total energy with diagonalization is -4914.74442 E_H. With a 147-atom Pt nanoparticle we calculated an energy of -13137.33174 E_H with a diagonalization based technique and the converged AQuA-FOE value is -13137.33169 E_H. There was also no difference in the number of iterations (either optimization of the Hamiltonian in the inner loop or optimization of the NGWFs in the outer loop) between the all the AQuA-FOE and the diagonalisation calculations.

Next we examined the scaling of the AQuA-FOE method with system size using truncated octahedral Pd nanoparticles with 2406, 4033, 6266, 9201, 12934 atoms. We chose these as truly 3-dimensional examples as they are a much more stringent test of density matrix decay as compared to 1- and 2-dimensional systems which can be more straightforward to demonstrate reduced or linear-scaling performance with the number of atoms due to a lower crossover point with cubic-scaling methods. We chose to run with the Hamiltonian matrix squared sparsity for the 1-particle density matrix, as described in section 6.8 and an electronic smearing of 0.5 eV. We also used the refined approximation to the entropy which was described in section 6.7. A psinc basis set kinetic energy cut-off of 500 eV was used and 9 NGWFs per Pd atom were employed, with NGWF radii of 6.0 a_0.

We ran all of these calculations on Archer, the UK's national supercomputer using 2400 MPI processes with 2 OpenMP threads per process. The calculations were compared against the timings of ONETEPs diagonalization based EDFT method for each system with the same number of processes and threads. The eigendecomposition was performed by the Scalapack implementation in the Intel MKL library and all of the sparse matrix operations in the AQuA-FOE calculations were performed using the SPAM3 sparse matrix algebra library which is an integral part of the ONETEP code. We did not have sufficient computing resources to perform all of the calculations to convergence, running instead a fixed number of four inner and four outer loop iterations to assess average time per iteration. From previous experience with the EDFT approach, we expect a small increase in the number of outer loop iterations with system size [7].

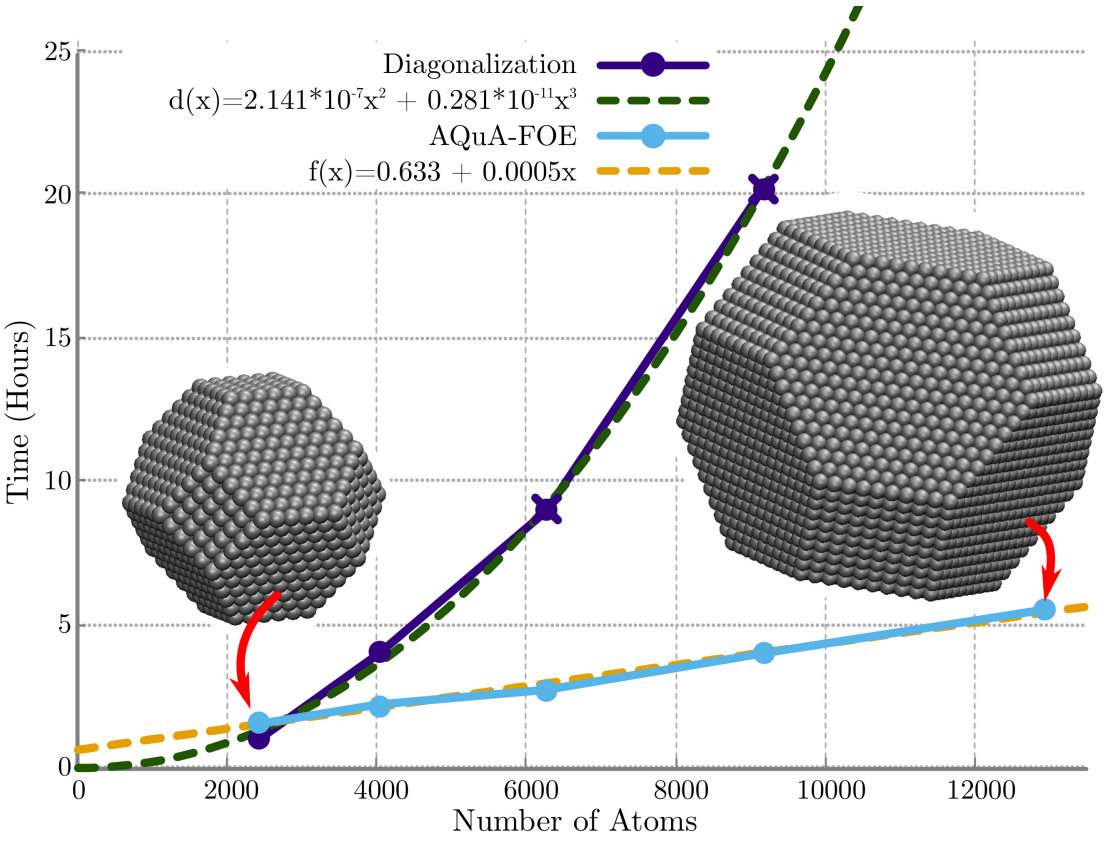


FIGURE 6.4: A comparison of EDFT in ONETEP using both Diagonalization with Scalapack (deep purple) and AQuA-FOE (light blue). We have performed calculations with 4 inner and 4 outer loop (EDFT) iterations for regular truncated octahedral nanoparticles of Palladium with 2406, 4033, 6266, 9201 and 12934 atoms. The run-times are presented on the y-axis, a cubic fit is shown through the diagonalization timings (deep green) and a linear fit is shown though the AQuA-FOE timings (orange). The geometries of the Pd_{2406} and Pd_{12934} nanoparticles are inset.

With ONETEP, the current implementation of EDFT with AQuA-FOE and Hamiltonian matrix squared sparsity has crossover point with ONETEPs highly optimized dense-matrix, diagonalization based EDFT scheme at ~3000 atoms (see figure 6.4). The observed linear-scaling ($O(N)$, to be precise) of the AQuA-FOE method shows that is it 5 times quicker when we reach Pd_{9201} (4 hours per EDFT inner loop iteration) than the cubically scaling diagonalization technique (20 hours per EDFT inner loop iteration). The 12934 atom system with diagonalization was excessively demanding in terms of computational expense and so we were not able to run it for comparison.

We expect that with the ever increasing performance of computational resources, and after further optimization of the AQuA-FOE method and code (perhaps through the use of the SelInv[103] algorithm rather than Chebyshev expansion to perform the quenches), that this method will become increasingly useful for calculations of industrially important large, metallic systems in technological applications in fields such as heterogeneous catalysis and biosensing.

6.11 Conclusions

We have presented the AQuA-FOE method for running DFT calculations on large metallic systems with a computational cost which increases effectively linearly with the number of atoms. The ONETEP code, which has already supported linear-scaling DFT calculations on insulators for many years provided the framework for developing our new method. Most of the machinery of ONETEP such as non-orthognal local orbitals optimised *in situ* (strictly localized orbitals, which we call NGWFs), and sparse, CPU-distributed matrices was used within our method. As in linear-scaling calculations on insulators, the exponential decay in the one-particle density matrix, which is also exhibited in finite temperature metallic systems was exploited to achieve reduced scaling.

The AQuA-FOE method works by running a Fermi Operator Expansion (FOE) on a Hamiltonian at several times the desired temperature. The resultant "hot" density matrix is then quenched repeatedly until the desired (lower) temperature is reached. This means that only a fixed, constant number of matrix multiplications is required to perform the FOE. Since each temperature-halving quenching operation also has a fixed cost, increase in the number of matrix products for different systems can only come from increase in the required number of quenches. However, for a given material, the Hamiltonian spectral width is asymptotically constant with increasing number of atoms. Hence, the number of quenching operations will also remain constant with the number of atoms and the full method is effectively linear-scaling with system size as the cost of each matrix product is linear-scaling with system size in the limit of sufficient matrix sparsity for large systems.

Our method finds the elements of the one-particle density matrix along with the electron number conserving chemical potential, and also electronic entropy using polynomial expansions which also have a constant number of terms with system size. We have shown validation calculations of AQuA-FOE inside the EDFT procedure by comparing numerically with the diagonalisation based EDFT that is already in ONETEP showing agreement in the energies to better than 10^{-5} E_H per atom. We have also demonstrated the effectively linear-scaling computational cost of our method with calculation times on Palladium nanoparticles ranging from 2406 to 12,934 atoms.

We expect that this method will become increasingly useful as supercomputing power becomes greater and more available. Complex metallic materials are possible to study with this approach, including large metallic nanoparticles which have a growing number of applications in important technological areas such as catalysis and biomolecular markers.

We have also shown our PaPa-FOE method which works by dividing the Hamiltonian eigenvalue spectrum into three. A partition which has full occupancy, a partition which has zero occupancy and a partition with intermediate, fractional occupancy. These partitions are Hamiltonian matrices where the other eigenvalues outside of the partition

are projected to zero. In so doing, we can apply the FOE to the central partition, only, with a fixed cost in terms of matrix products as its width is determined by the smearing function rather than the material. The outer partitions are given by the projection matrices themselves. The method currently has issues with accuracy on account of the projection matrices having limited accuracy or the sparsity patterns of the projection matrix – Hamiltonian matrix products being insufficient. As further work, it would be useful to make this method work accurately, because at least in theory it involves less computational effort (the prefactor is lower) than our AQuA-FOE scheme.

Chapter 7

Conclusions

The work presented in this thesis has focussed on DFT for calculations on large metallic nanoparticles. We have developed new algorithms for reduced scaling DFT methods for metals; we have tested, verified and devised new descriptors for predicting the catalytic activity of metallic nanoparticles; we have applied large-scale DFT calculations to model nanoparticle sequences to show size and oxygen adsorption packing trends, and finally we have applied these techniques and knowledge to perform a study of oxygen adsorption on real-world, experimentally determined platinum nanoparticles in collaboration with the Nellist group at Oxford materials.

We have explored the binding of atomic oxygen to platinum nanoparticles of up to ∼1000 atoms using DFT calculations in ONETEP. We firstly explored the effect of binding single oxygen atoms per nanoparticle and increasing the size of the nanoparticle. This study was performed using the cuboctahedral sequence of nanoparticles cleaved from the platinum FCC lattice. We investigated the binding strength vs cuboctahedral magic number, as well as the effect of popular descriptors for catalytic activity, such as the d-band centre approach (the implementation of this in ONETEP is described in chapter 3), which we correlate against the adsorption strength.

Our study of oxygen adsorption strength with nanoparticle size indicated that the binding strength of oxygen decreases with nanoparticle size, converging on the limit of oxygen adsorption on the infinite platinum slab (as calculated with Castep in periodic boundary conditions). Due to the work we did on angular momentum projected density of states in ONETEP we were also able to correlate these adsorption strengths with the d-band centre and obtained a linear trend, which is a new result for nanoparticles but in line with what has been reported in the literature by Nørskov, *et al*[197] for platinum slabs. We subsequently went on to explore the effects of oxygen coverage on the Pt_{55} and Pt_{147} cuboctahedral nanoparticles to attempt to advance further towards realistic conditions in our simulations. This part of the investigation revealed that with the half monolayer, hemispherical oxygen coverage that we eventually settled on, oxygen tends to gravitate

towards the edges and lower coordinated sites in the nanoparticle and away from the centres of facets. This effect correlates with the site specific, single oxygen adsorption energies on Pt_{309} and experimental platinum nanoparticles which has been documented in chapter 5. The breaking up of the relaxation in two: firstly a nanoparticle with oxygen solely on the (111) facets, then the (100) facets, before combining the resulting oxygen coordinates onto a nanoparticle with half monolayer, hemispherical oxygen coverage allowed us to see the energetic effect of various binding types in finer resolution than we would have been able otherwise. The effect of oxygen site migration appears to have the greater effect on per oxygen binding strengths, with respect to oxygen-oxygen interaction on a nanoparticle, particularly on the (111) facets. The (100) coverage was found to be low enough that very little change was observed in average oxygen binding strength vs binding strength of single oxygens to (100) bridge sites.

We then went on to perform large scale DFT calculations on real platinum nanoparticles, which were measured by the Nellist group at Oxford materials using ADF STEM. These DFT calculations gave us the electronic structure of the experimental nanoparticles which allowed us to apply electron density based catalytic activity descriptors to the nanoparticles, such as the d-band centre approach, or our own electronic density based descriptor described in chapter 3. We found that surface roughness of the experimental nanoparticles contributes to more potential oxygen binding sites with low electron density, which correlatates with stronger oxygen adsorption strength in our model, when compared with the relative smoothness of cuboctahedral and truncated octahedral facets. Using regression through our DFT data, we were able to generate calibration curves that we have used to predict the binding strength of sites which we have not directly run DFT calculations on, and furthermore, to predict the proportion of sites which lie with 0.2 eV of the oxygen binding strength required for optimum catalytic activity. This process takes a tiny fraction of the time required to do geometry relaxed DFT calculations on new experimental nanoparticles, and so we claim that this is a high-throughput computational technique to compliment new high-throughput experimental techniques, such as those performed by the experimental team during this project.

Because we have a highly sensitive electronic density based descriptor, we would also expect that this technique is applicable to multi-species nanoparticles in the future, in contrast to geometric descriptors.

Looking to the future of large scale DFT calculations on metallic systems and the future of this work to be undertaken by the group, we have presented the AQuA-FOE method for running DFT calculations on large metallic systems with a computational cost which increases asymptotically with the number of atoms. The AQuA-FOE method works by running a Fermi Operator Expansion (FOE) on a Hamiltonian at several times the target temperature. The resultant “hot” density matrix is then quenched repeatedly until the

target (lower) temperature is reached. This means that only a fixed, constant number of matrix multiplications is required to perform the FOE.
Our AQuA-FOE method finds the elements of the one-particle density matrix along with the electron number conserving chemical potential. We also devised a method for calculating electronic entropy using polynomial expansions with a constant number of terms. We have shown validation calculations of AQuA-FOE inside the EDFT procedure by numerical comparison with the diagonalisation based EDFT that is already in ONETEP showing agreement in the energies to better than 10^{-5} E_H per atom. We have also demonstrated the effectively linear-scaling computational cost of our method with calculation times on regular truncated octahedral Palladium nanoparticles ranging from 2,406 to 12,934 atoms. We expect that this method will become increasingly useful as supercomputing power becomes greater and more available. Complex metallic materials are possible to study with this approach, including large metallic nanoparticles which have a growing number of applications in important technological areas such as catalysis and biomolecular markers.

7.1 Further work

Further work on this project has already begun with the set of single oxygen adsorption calculations on cuboctahedral nanoparticles with fully geometry relaxed structure and PAW. This together with coverage calculations on Pt_{309} will complete our study of oxygen adsorption on platinum against nanoparticle size and coverage density of oxygen.
Other members of the group continue to advance these calculations to ever greater sizes and to more accurate conditions, including the effects of graphene and oxide supports, and solvation. These state of the art calculations promise to give great insight into the computational study of catalysis, which is hugely important in industry. I look forward, with great anticipation, to seeing the results of these projects.
Another aspect of this work which merits further study is the FOE techniques that we have pioneered during this project. There is great scope for improving the performance, functionality and applicability of these techniques, which will be an exciting area of research for years to come. One area which is already being developed for EDFT in ONETEP by a member of the ONETEP developers group is spin relaxation which it may be possible to combine with FOE techniques. Linear scaling methods for metals will be the future of this work.

Appendix A

Matlab Implementation of the new EDFT/FOE Ideas in Chapter 6

Although the ideas in chapter 6 have been successfully implemented in ONETEP in parallel with zero matrix diagonalisations or extremal eigenvalue evaluations, a (very much simplified) version implemented in Matlab is included here in case it might aid understanding of the technique.

A.1 Projection Operator based FOE

```
% An implementation of our projection based FOE method with the rational
% eigenvalue interval represented as a Chebyshev polynomial expansion.
function [T,bde,matmuls,mmvec] = jolyfoe4(Ain,M,mu,smear)
    matmuls=0;
    mmvec=0;
    Mlow=M^(-1/2);
    Horth=(Mlow'*Ain*Mlow);
    N=size(Ain,1);
    trace(Horth)
    A=-((Horth-eye(N)*mu)/smear);
    fprintf('\n ==>\n');
    trace(A)
    bde=max(abs(eig(A)))-min(abs(eig(A)));

    A=A/2;

    cutoff=15;

    min(abs(eig(A)));
    min(eig(A));
    max(eig(A));
```

```
norm(A,'fro');
norm(A-eye(size(A,1))*cutoff,'fro');

medium_T=zeros(N);

if(min(abs(eig(A)))>cutoff)
    [medium_T,mmsi]=signnsh(A,1e-13); %#ok<ASGLU>
    matmuls=matmuls+mmsi;
    low_T=zeros(N);
    high_T=zeros(N);

else
    if(max(eig(A))>cutoff)
        [sign_cutoff,mmsi]=signnsh((A-eye(size(A,1))*cutoff),1e-13);
        min(abs(eig(A)));
        high_T=(sign_cutoff+eye(size(sign_cutoff,1)))/2;

        matmuls=matmuls+mmsi;
        tmp_A=A*sign_cutoff;
        matmuls=matmuls+1;

        low_A=(A-tmp_A)/2;
        mmvec(1)=mmsi;

    else
        high_T=zeros(N);
        low_A=A;
        mmsi=0;
    end

    if(min(eig(A))<-cutoff)
        [sign_cutoff,mmsi]=signnsh((low_A+eye(size(A,1))*cutoff),1e-13);

        matmuls=matmuls+mmsi;
        mmvec(2)=mmsi;
        tmp_A=low_A*sign_cutoff;
        matmuls=matmuls+1;
        medium_A=(low_A+tmp_A)/2;

        low_T=(sign_cutoff-eye(size(sign_cutoff,1)))/2;
        max(abs(real(eig(medium_A))));
    else

        low_T=zeros(N);
        medium_A=low_A;

    end
end
```

```
if(min(abs(eig(A)))<cutoff)
    medium_A=medium_A/15;
    [medium_T,matmuls_cheby]=fastchebyfun(medium_A);
    matmuls=matmuls+matmuls_cheby;
end

    T=low_T+medium_T+high_T;

    T=(T+T')/2;

    trace(T)
    T= (eye(size(T,1))+T)/2;

end
```

A.2 Chemical Potential Search and Density Kernel Update Algorithm

```
% An implementation of our Chemical potential search method with
% hyperbolic gradients and density kernel update. This version applies
% a safeguarded Newton-Raphson root finder.

function [mu,total_matmuls,rho_out]=findfermi3(Ne,smearing,t,ham,overlap,lowb,highb)
  N=size(ham,1);
  machep = 2^-53;
  sa = lowb;
  sb = highb;
  total_matmuls=0;
  [rho,bde,matmuls]=jolyfoe4(ham,overlap,sa,smearing);
  total_matmuls=total_matmuls+matmuls;
  T=2*rho-eye(N);
  fa = trace(rho);
  Dmu=sb-sa;
  invdenom=false;
  [lowmu,T,matmuls,invdenomout]=changemu(T,sa,Dmu,smearing,ham,overlap);
  if(invdenomout~=false)
      invdenom=invdenomout;
  end
  total_matmuls=total_matmuls+matmuls;
  rho=0.5*((eye(N)+T));
  fb = trace(rho);

  c = sa;
```

```
fc = fa;
e = sb - sa;
d = e;

sold=sb;
while(true)
  if(abs(fc)<abs(fb))
    sa = sb;
    sb = c;
    c = sa;
    fa = fb;
    fb = fc;
    fc = fa;
  end

  tol = 2.0 * machep * abs ( sb ) + t;
  m = 0.5 * ( c - sb );

  if(abs(m)< tol || fb == 0.0)
    break
  end

  if(abs(e) < tol || abs(fa) <= abs(fb))
    e = m;
    d = e;
  else
    s = fb / fa;
    if(sa==c)
      p = 2.0 * m * s;
      q = 1.0 - s;
    else
      q = fa/fc;
      r = fb/fc;
      p = s*(2.0*m*q*(q-r)-(sb-sa)*(r-1.0));
      q = (q-1.0)*(r-1.0)*(s-1.0);
    end

    if(0.0<p)
      q = -q;
    else
      p = -p;
    end
    s = e;
    e = d;

    if(2.0*p < 3.0*m*q-abs(tol*q) && p<abs(0.5*s*q))
      d = p/q;
    else
```

```
        e = m;
        d = e;
      end
    end

    sa = sb;
    fa = fb;

    if(tol<abs(d))
      sb = sb+d;
    elseif(0.0<m)
      sb = sb+tol;
    else
      sb = sb-tol;
    end

    Dmu=sb-sold;
    if(invdenom~=false)
        [newmu,T,matmuls,invdenomout]=changemu(T,sold,Dmu,smearing,ham,overlap,invdenom);
        if(invdenomout~=false)
            invdenom=invdenomout;
        end
    else
        [newmu,T,matmuls,invdenomout]=changemu(T,sold,Dmu,smearing,ham,overlap);
        if(invdenomout~=false)
            invdenom=invdenomout;
        end
    end

    total_matmuls=total_matmuls+matmuls;

    fb=0.5*(trace(eye(N)+T))-Ne;
    sold=sb;

    if((0.0<fb && 0.0<fc) || (fb<=0.0 && fc<=0.0))
      c = sa;
      fc = fa;
      e = sb-sa;
      d = e;
    end
    end
  mu = sb;
  rho_out=0.5*((eye(N)+T));
  return
end


% Chemical potential derivative of the Tanh matrix
```

```
function df=deriv(T,smearing)
T2=T^2;

deriv1=-(1/(4))*(1-trace(T2));
df=deriv1;
end


% Change the chemical potential of a density kernel without reconstructing
% it (provided that the condition number is low enough that this is more
% efficient.
function [newmu,newT,matmuls,newinvinit]=changemu(Tinit,mu,Dmu,smearing,ham,overlap,invinit)
N=size(Tinit,1);

condi=cond(Tinit*tanh(Dmu/(2*smearing))+eye(N));
tanh(Dmu/(2*smearing))
matmuls_reinit=0;
matmuls_inv=0;

newmu=(Dmu) + mu;

if((abs(condi)>1e4)||(abs(Dmu/(2*smearing))>15))
    fprintf('reinit');
    [newrho,bde,matmuls]=jolyfoe4(ham,overlap,newmu,smearing); %fprintf('--> %d, %d\n',trace(newrh
    matmuls_reinit=matmuls_reinit+matmuls;
    newT=2*newrho-eye(N);
    newinvinit=false;

else
    denom=(Tinit*tanh(Dmu/(2*smearing))+eye(N));
    if (~exist('invinit', 'var'))
        invdenom=denom'*(1.0/(norm(denom,1)*norm(denom,inf)));
        fprintf('here');
    else
        invdenom=invinit;
    end
    D=9;
    [invdenom,matmuls] = nsh_inverse(denom,invdenom,D);
    newinvinit=invdenom;
    fprintf('set to:')
    newT=(Tinit+eye(N)*tanh(Dmu/(2*smearing)))*invdenom;
    matmuls_inv=matmuls_inv+matmuls+1;
end
matmuls_reinit
matmuls_inv
matmuls=matmuls_inv+matmuls_reinit;
end
```

A.3 Important Implementations of Algorithms from Literature used in the Above Methods

```
% An iterative implementation of the Lowdin decomposition as described in
% Jansik and Olsen.

function [Zk,Yk] = lowdin(overlap)

cond(overlap)

N=size(overlap,1);
Zk=eye(N);
Yk=overlap;

Seigs=eig(overlap);
epsmin=min(Seigs);
epsmax=max(Seigs);
lambda=2.0/(epsmax+epsmin);

finishearly=false;
i=0;
while not(finishearly)
    i=i+1;
    Xk=lambda*Yk*Zk;
    if(norm(Xk-eye(N))<10e-10)
        finishearly=true;
    end
    T2=0.5*(3*eye(N)-Xk);
    Zk=Zk*T2;
    Yk=T2*Yk;
    if(finishearly)
        break
    end
end
lambda
Zk=sqrt(lambda)*Zk;
Yk=sqrt(lambda)*Yk;
end
```

```
% An iterative implementation of the Matrix sign function as described in
% Chen and Chow.
function [X,matmuls,iter] = signnsh(A,tol)
    conv=0;
    matmuls=0;
```

```
    evs=abs(eig(A));
    %norm(A,'fro')
    emin=min(evs);
    emax=max(evs);
    %X0=A/emax;
    X0=A/norm(A,'inf');
    X2=X0^2;
    matmuls=matmuls+1;
    y=emin/emax
    y=1
    i=0;
    !iter=0;
    while true
        i=i+1;
        alpha=sqrt(3/(1+y+y^2));

        X0=0.5*alpha*X0*(3*eye(size(A,1))-(alpha^2)*X2);
        matmuls=matmuls+1;
        y=0.5*alpha*y*(3-(alpha^2)*y^2);

        X2=X0^2;
        matmuls=matmuls+1;
        iter(i)=norm(X2-eye(size(A,1)),'fro');
        if(i>1)
            conv(i)=abs(iter(i))/(abs(iter(i-1))^2);
        end

        if(iter(i)<tol)
            X=X0;
            %i
            break;
        end
    end
    matmuls
end
```

```
% An implementation of Liang & Head-Gordon divide and conquer Chebyshev
% polynomial resumming.

function [T,matmuls] = fastchebyfun(X)
%lbar=9 % round(P/10) %12%round(P/4)
ml=336;%round(((1.0/2.0)+(((max(eigs))*D*log(10.0))/(pi*(2*lbar-1)))));
%ml=150;
syms x
s = 'tanh(x*15)';
N=size(X,1);
```

```
tanhcheb=chebfun(s,ml+1);

dccoeffs = cell(1,1);

    c=tanhcheb.funs.coeffs(ml+1:-1:1);
    %c=tanhcheb.funs.coeffs;
    c(1:2:ml+1)=0;
    %c=tanhcheb.funs.coeffs;
    % Construct binary tree
    layers=floor(log2(ml))-3;
    jh=0;
    jl=0;
    jlold=1;
    jhold=1;
    num=0;
    num(1:2^(layers+1)-1)=0;
    num(1)=(ml+1)*2;
    parent=0;
    parent(1:2^(layers+1)-1)=0;
    start=0;
    start(1:2^(layers+1)-1)=0;
    start(1)=1;
    for i=0:layers
        jl=jl+2^(max(0,i-1));
        jh=jh+2^(i);
        for j=jl:jh
            parent(j)=jlold+(floor((j-jl+2)/2)-1);
            if(mod(j,2)==0)
                num(j)=floor(num(parent(j))/2);
            else
                num(j)=floor((num(parent(j))-1)/2);
            end
        end
        jlold=jl;
        jhold=jh;
    end

    for i=2:numel(num)
        start(i)=start(i-1)+num(i-1)+1;
    end
    coeffs=0;
    coeffs(1:(ml+1))=0;
    coeffs(1:ml+1)=c;

    jl=0;
    jh=0;
    jlold=1;
    jhold=ml+1;
```

```
    for i=2:numel(num)
        jl=jhold+1;
        jh=jl+num(i);
        if(mod(i,2)==0)
            for j=jh:-1:jl
                k=j-jl;
                coeffs(j)=coeffs(2*k+start(parent(i)));
            end
        else
            coeffs(jh)=2*coeffs(2*num(i)+1+start(parent(i)));
            for j=jh-1:-1:jl+1
                k=j-jl;
                coeffs(j)=2*coeffs(2*k+1+start(parent(i)))-coeffs(k+1+jl);
            end
            coeffs(jl)=coeffs(start(parent(i))+1)-(coeffs(jl+1)/2);
        end
        jlold=jl;
        jhold=jh;

    end
    dccoeffs{1} = coeffs(start(numel(num)-2^layers+1):numel(coeffs));
    dccoeffs{1}*dccoeffs{1}';
%end

calclayerstart=start(numel(num)-2^layers+1);
exes = cell(1,layers+1);
for i = 1:layers+1
    exes{i} = zeros(N);
end
matmuls_divconq=0;
exes{1}=X;
for i=2:layers+1
    matmuls_divconq=matmuls_divconq+1;
    exes{i}=(2*exes{i-1}*exes{i-1})-eye(N);
end


accumat = cell(1,2^layers);
for i=1:2^layers
    accumat{i} = zeros(N);
end
NLcheb_acc=cell(1,2^layers);
for i=1:2^layers
    NLcheb_acc{i}=zeros(N);
end
for i=0:max(num(numel(num)-2^layers+1:numel(num)))
    if(i==0)
        T0=eye(N);
```

```
        T=T0;
    elseif(i==1)
        T1=exes{layers+1};
        T=T1;
    else
        T=2*exes{layers+1}*T1 - T0;
        matmuls_divconq=matmuls_divconq+1;
        T0=T1;
        T1=T;
    end

    for j=(2^(layers-1))+1:2^(layers)
        if(i<=num(numel(num)-2^layers+j))

                accumat{j}=accumat{j}+(dccoeffs{1}(start(numel(num)-2^layers+j)-start(numel

        end
    end
end


k=0;
for i=1:layers-1
    nmats=2^i;
%     for j=1:2^(layers-i)
    for j=(2^(layers-i-1))+1:2^(layers-i)
        %accumat{nmats*(j-1)+1} = accumat{nmats*(j-1)+1} + exes{layers-i+1}*accumat{nmats*(
        k=k+1;
        norm(accumat{nmats*(j-1)+1+(2^(i-1))},'fro');
        %matmuls_divconq=matmuls_divconq+1;
        accumat{nmats*(j-1)+1} = accumat{nmats*(j-1)+1} + exes{layers-i+1}*accumat{nmats*(j
        matmuls_divconq=matmuls_divconq+1;
    end
end
accumat{1} = accumat{1} + exes{1}*accumat{1+(2^(layers-1))};
norm(accumat{1},'fro');
    matmuls_divconq=matmuls_divconq+1;
matmuls=matmuls_divconq;
%compcheb=accumat{round((2^layers)/2)+1};
T=accumat{1};
end
```

Appendix B

Other adsorbate calculations on Pt cuboctahedra

B.1 Hydroxide and Hydroxyl Radical

While both ORR mechanisms have a step involving a bound hydroxyl radical, the decision was made to perform calculations on bound hydroxide initially. This is based on the observation from the twisted 55 atom cubooctahedron that the convergence speed of hydroxide is much improved with respect to the hydroxyl radical system, partly perhaps because the hydroxyl radical system has an odd number of electrons and hence must be modelled spin-polarised. This at least doubles the cost of the calculation, but it was also observed in the twisted system that the number of iterations was also much increased. The adsorption energies of both systems are similar, but importantly the final geometries of the systems are very similar, so it made sense to run the hydroxide geometry optimisation first, as it is fast, and subsequently initialise the hydroxyl radical calculation with the initial configuration from the converged hydroxide system. This has the advantage that it gives more results for potentially less expense.

N_{atoms}	13	55	147
E_{ads}	-3.5512	-3.8423	-3.5502

TABLE B.1: The change in energy of adsorption of OH as the size of the nanoparticle is increased (in eV).

Examining the trend in energy of adsorption with increasing nanoparticle size for the hydroxyl ligand shows that the Pt_{13} and Pt_{147} systems have approximately the same binding strength of -3.55eV, while Pt_{55} appears to be more strongly bound. This was an unexpected result but it is made stranger by the comparison with the slab model in Castep, which shows significantly weaker adsorption strength.

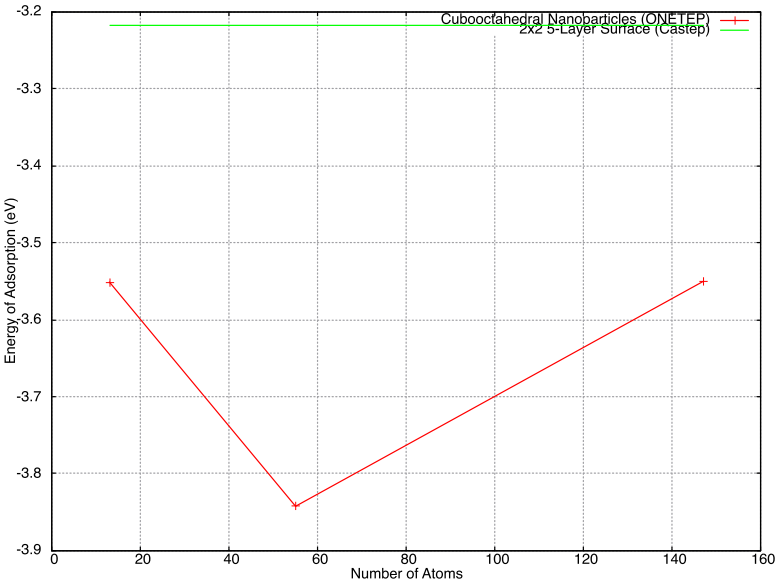


FIGURE B.1: The trend in adsorption strength of OH versus the size of cubooctahedral nanoparticle on a (111) site against the a slab model in Castep (25% packing) is inconclusive at present. It is clear that Pt_{55} appears to be more strongly bound than Pt_{13} and Pt_{147} and that the slab is 0.3eV more weakly bound than both.

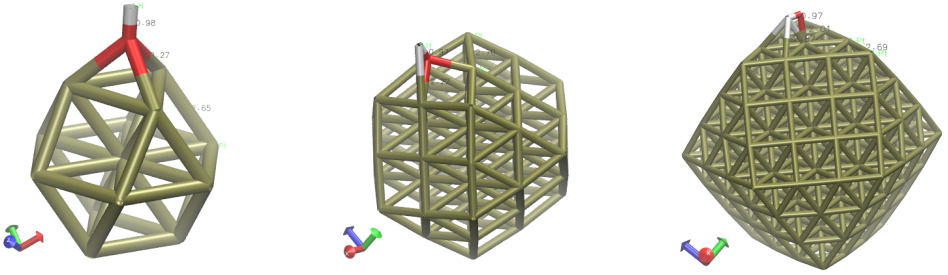


FIGURE B.2: The lowest energy configurations of OH on Pt_{13}, Pt_{55} and Pt_{147}.

As for the discrepancy between the Castep slab results and the ONETEP nanoparticle results, it would be very interesting to try the Pt_{309} to determine whether it moves closer in energy to the slab. These calculations are in progress. Another possibility is that the 25% packing of ligands onto the (111) surface in the Castep calculation is affecting the agreement with the single ligand nanoparticle calculations. This would seem fairly likely, given the excellent agreement between identical Pt_{55} systems in Castep and ONETEP. The quicker option to test this would be to recompute the slabs in Castep with lower packing, but it would also be interesting to try more packing on the nanoparticle systems in the future, as well.

B.2 Carbon Monoxide

The final set of calculations that were performed were not on reactants or intermediates from the ORR, but on carbon monoxide, an important poison of the platinum anode which slows the hydrogen oxidation reaction by preferentially adsorbing onto active catalytic sites and blocking these for hydrogen oxidation.
Carbon monoxide presents a difficulty as far as DFT with a PBE based XC functional is concerned, as it has been shown [241] many times to adsorb preferentially to the FCC hollow site over the atop site as measured experimentally. This problem has been studied in great depth by Nørskov [242] and Kresse [243], but no truly satisfactory resolution is reached other than that most conventional XC functionals are inadequate and that the PBE based functionals represent a trade-off between accurate binding energies and accurate bulk and surface physics (lattice constants, bulk moduli, etc.). So, an alternative GGA XC functional such as BLYP can be used to improve the adsorption energies to the point that the atop site is preferred over the FCC hollow, but with 10% error in the lattice constant and poor experimental agreement for most other observables.
It was decided that for this study, rPBE would be used for every calculation including CO, for consistency and the binding site order discrepancy put down to a limitation of the XC functional.

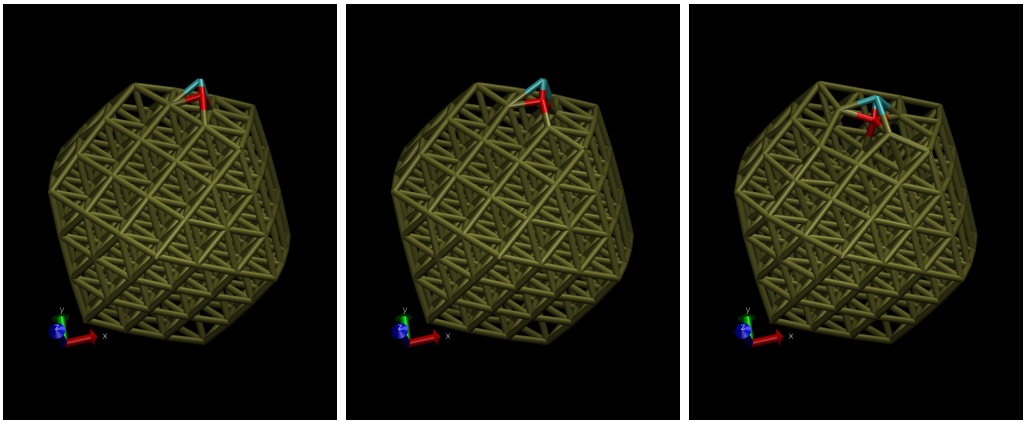


FIGURE B.3: The progression of a geometry optimisation of CO on a Pt_{147} nanoparticle, showing a failure mode for a pathological starting configuration.

The carbon monoxide geometry optimisation calculations presented several difficulties as far as convergence was concerned; the initial configuration that was chosen for the Pt_{147} system to begin with lead to a geometry optimised configuration where the carbon atom appeared to *absorbed* by the nanoparticle. Since the configuration had not been encountered in literature and the energetics appeared at odds with the results for the adsorption on the other sized nanoparticles (1eV energy of adsorption (absorption?) as opposed to 2eV), the calculation was restarted with different starting geometry.

The second attempt on Pt_{147} was much more successful and appeared to fall into an FCC 3-fold hollow site with the oxygen atom 30° degrees off the facet plane - carbon perpendicular, almost in a bridge site. The Pt_{55} system exhibits a similar configuration, whereas the Pt_13 system has a perpendicular molecule bond, through the nanoparticle.

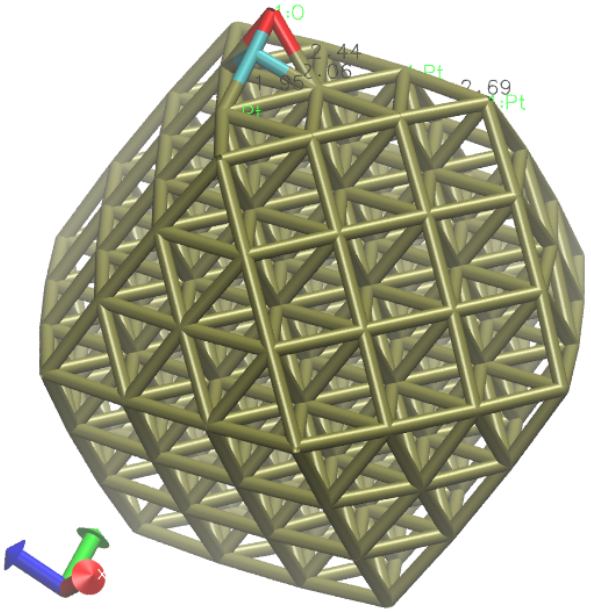


FIGURE B.4: The final configuration of a geometry optimisation calculation of CO on a Pt_{147} nanoparticle exhibits a $\mu_3 - \eta^2$ configuration.

The second issue which presented itself during the optimisation of the $Pt_{147} - CO$ systems remains unexplained; as the electronic minimisation progressed, at some configurations, the energy gradient would increase without bound, while the electronic energy was still reduced with each step. This is almost certainly a but in the code, probably in the EDFT line search and will be investigated further. The issue did not present a problem as far as running the calculation is concerned, only in that itsignificantly slowed the calculations down, as the maximum number of inner and outer loop iterations would be performed for these configurations. Such pathological behaviour has been seen only for CO.

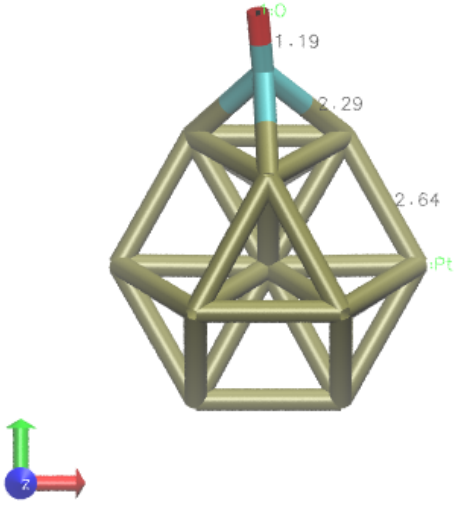


FIGURE B.5: The geometry optimised configuration of CO on Pt_{13} shows an almost perpendicular bond with the surface

N_{atoms}	13	55	147
E_{ads}	-2.1900	-1.9170	-1.8047

TABLE B.2: The change in energy of adsorption of CO as the size of the nanoparticle is increased (in eV).

B.3 Oxygen

Both molecular oxygen and atomic oxygen adsorbates were chosen for study. Molecular oxygen for the adsorption of oxygen onto the catalyst at the start of the reaction and atomic oxygen because it features in both the associative 2-electron and the dissociative 4-electron pathways. It is true that atomic oxygen is not be adsorbed onto the surface in the manner we are simulating during the reaction in reality, but the ground-state energy and coordinates of the atom on the facet will be in agreement with a point on the transition state path.

Molecular oxygen was modelled in a triplet state as it is found in its ground state in nature. This was achieved by performing a spin polarised calculation with a total spin of 2 in ONETEP, which corresponds to an overall spin of 1 in natural units. The atomic oxygen system, on the other hand, has an even electron number and was not expected

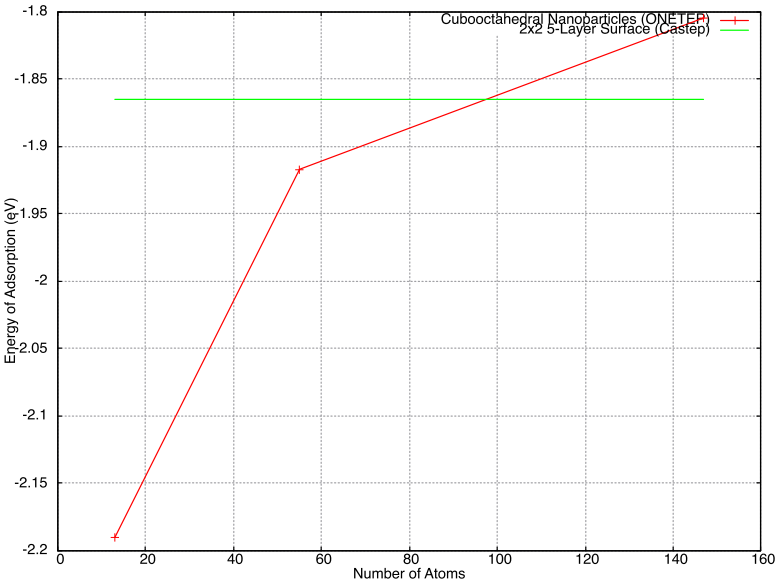


FIGURE B.6: The trend in adsorption strength of carbon monoxide versus the size of cubooctahedral nanoparticle on a (111) site against the a slab model in Castep (25% packing) shows that binding strength increases with decreasing nanoparticle size.

to be spin polarised, so was performed with spin of 0, with a single spin-channel. The adsorption energy of atomic oxygen was calculated as:

$$E_{\text{ads}} = E^{\text{opt}}_{\text{O on particle}} - E^{\text{opt}}_{\text{particle}} - \frac{1}{2}E^{\text{opt}}_{O_2}, \tag{B.1}$$

whereas molecular oxygen was calculated as:

$$E_{\text{ads}} = E^{\text{opt}}_{\text{O on particle}} - E^{\text{opt}}_{\text{particle}} - E^{\text{opt}}_{O_2}. \tag{B.2}$$

Across the three sizes of cubooctahedra, the molecular oxygen was found to bind in a similar minimum energy configuration. This is independent of whether solely the molecular coordinates were relaxed on a constrained nanoparticle, or all atomic coordinates were allowed to relax. The observed configuration is a $\mu_3 - \eta^2$ bridge configuration.

N_{atoms}	13	55	147
E_{OOads}	-1.4795	-0.7334	-0.2984
E_{Oads}	-2.2908	-1.4451	-0.7292

TABLE B.3: The change in energy of adsorption of molecular and atomic oxygen as the size of the nanoparticle is increased (in eV).

The fully geometry optimised calculations show that as the dimension of the nanoparticle is decreased, the strength of adsorption of oxygen is increased, both for the atomic and molecular oxygen.

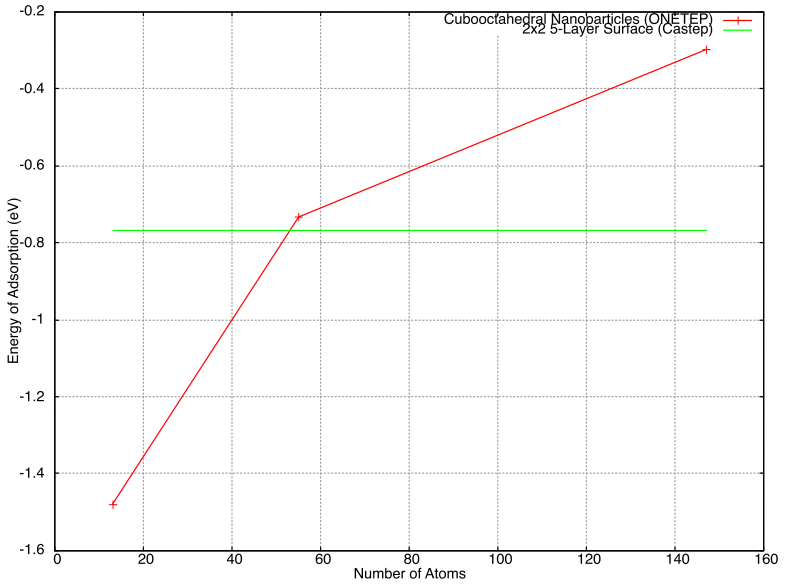


FIGURE B.7: The trend in adsorption strength of molecular oxygen versus the size of cubooctahedral nanoparticle on a (111) site against the a slab model in Castep (25% packing) increases with decreasing nanoparticle size, but that the slab model (at least at this packing) does not appear to be asymptotic with the nanoparticle model.

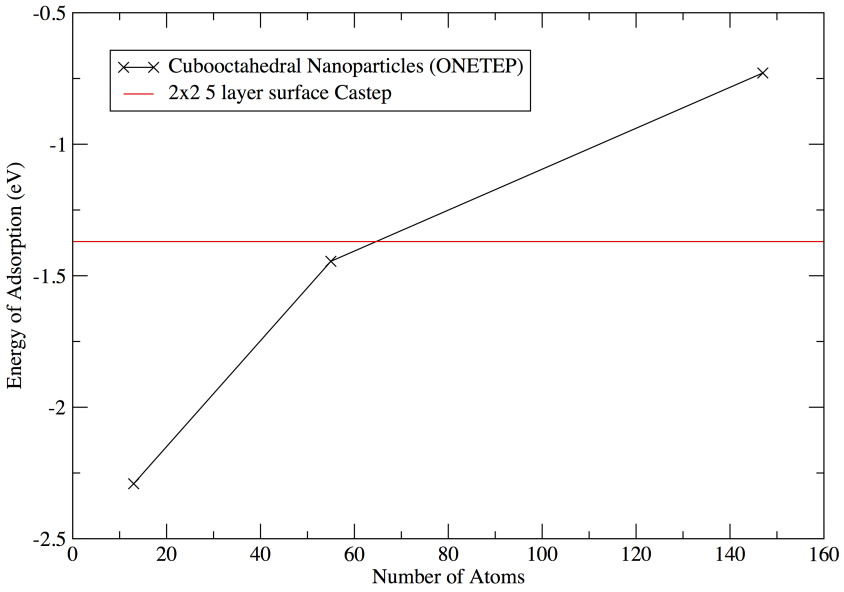


FIGURE B.8: The trend in adsorption strength of atomic oxygen versus the size of cubooctahedral nanoparticle on a (111) site against the a slab model in Castep (25% packing) shows much the same trend as for molecular oxygen.

B.4 Comparison of Adsorption Energies on Fully Relaxed and Constrained Nanoparticles

The calculations presented so far have been performed by geometry optimising the nanoparticle and ligand in isolation, before moving the ligand close the the surface of the nanoparticle and geometry optimising the full system. We believe that this is the most rigorous way to ensure the obtained relative energies are accurate and reliable, though it seems from prior studies in the literature [76] that the second step may be performed with the metal, nanoparticle atoms constrained and only the ligand atoms free to relax. The latter approach is likely to be faster than the former, as the geometry optimisation algorithm has to deal with fewer degrees of freedom. It was hypothesised, however, that the approach misses several important degrees of freedom (surface relaxations are impossible, for instance) which would lead to lower total energies. On the other hand, if this simplified approach could be used, a great deal of computer time could be saved.
In order to test the hypothesis, the calculations which were carried out using a full geometry optimisation were repeated with constrained metal atoms. The results are quite remarkable in places. For some nanoparticle-ligand combinations, very little difference is noted when comparing the two approaches, however, in most cases the full geometry optimisation gives lower total and adsorption energies than that with constrained metal atoms. In the case of molecular oxygen, the adsorption energy changes sign when the full geometry optimisation is performed, for the starting geometry we have used, only becoming stable with the full geometry optimisation.

TABLE B.4: On the 55 atom cubooctahedron, the adsorption energies of the ligands can vary by more than 1eV with and without full geometry optimisation. At least for the starting geometry used for this test, the molecular oxygen test was not stable without full optimisation

(eV)	Rigid	Relaxed
Pt55 (111) CO	-0.959	-1.917
Pt55 (111) OH-	-1.883	-3.843
Pt55 (111) O	-0.159	-1.445
Pt55 (111) OO	0.490	-0.733
Pt55 (111) OHrad	-5.083	-2.612

B.5 Classical Molecular Dynamics

The first work that was performed was to run the coordinates through a molecular dynamics (MD) geometry optimiser. While the (z) coordinates have already been optimised

TABLE B.5: On the 55 atom cubooctahedron, the adsorption energies of the ligands can vary by more than 1eV with and without full geometry optimisation. At least for the starting geometry used for this test, the molecular oxygen test was not stable without full optimisation

(eV)	Rigid	Relaxed
Pt147 (111) CO	-1.030	-1.805
Pt147 (111) OH-	-4.504	-3.550
Pt147 (111) O	-0.690	-0.729
Pt147 (111) OO	0.154	-0.298
Pt147 (111) Ohrad	-3.049	-3.127

to some extent, it was deemed important to perform an unconstrained geometry optimisation prior to any calculations and as reliable interatomic potential models for platinum are available in the DL POLY4 MD package, this would be a quick and simple test of the data received from Oxford.

Initially a Sutton-Chen potential was chosen with parameters from [244] as this had been used successfully in the group for gold nanoparticles. A damped MD geometry optimisation (known as a 'zero temperature' MD in the DL POLY vernacular) was performed on the system resulting in very slight atomic movement to the extent that no atom moved out of a lattice site.

As a further test of the reliability of the calculation, as the change to the coordinates had been so slight, a short MD run was performed on the system at a temperature of 300K, with a timestep of 0.25fs, for 1000ns. This yielded troubling results, in that the nanoparticle did not hold structural integrity for even the length of the calculation and instead flew apart. In order to address the issue, potential parameters were checked for dimensional equivalence and the time-step was reduced several times to no avail.

The situation was solved by using a different potential. The paper [244] also gives parameters for a modification to the Sutton-Chen potential known as the Gupta potential. Using a Gupta potential set up in the same way, the particle geometry relaxed slightly in the geometry optimisation, particularly in one dimension, as might be expected, but crucially the nanoparticle held structural integrity for the duration of the MD test irrespective of the time-step.

An interesting finding from the room temperature MD calculation is that the surface atoms are effectively in flux. This is in line with the findings of the Nellist group when imaging snapshots of the same particle in quick succession, though the temperature of the particle in the experimental setup is not known precisely because the beam is thought to impart a large quantity of energy on the nanoparticles. This result shows

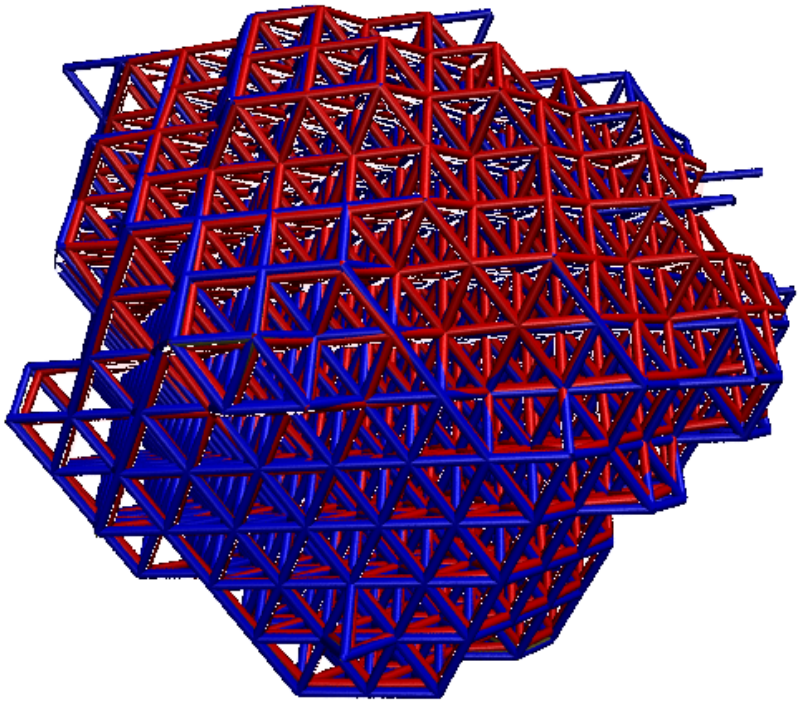


FIGURE B.9: The results of the damped MD optimisation of the experimental particle with a Gupta potential (blue) vs a Sutton-Chen potential (red). The Gupta coordinates move only slightly from the experimentally measured coordinates.

that although the third coordinates of the atoms in the nanoparticle under consideration are not directly experimentally accessible, the high mobility of atoms suggests that the exact nuclear coordinates of a snapshot are less important than having a generally representative coordination and level of disorder.

This data was then fed into ONETEP for DFT electronic minimisation. As the nanoparticle consists of 943 atoms, it was considered too large to perform a full DFT geometry optimisation on, particularly given that this was only to be a test system. There is nothing to preclude such calculations on similarly large nanoparticles in the future using L-BFGS in ONETEP except computational resource availability and time.
The calculation was run with parameters inspired by those from the cuboctahedral system calculations, in particular, the same pseudo-potentials and cut-off energies were used, with a Coulomb cut-off radius of 45 Åand a cubic cell of edge length 70 Å. The maximum atom-atom distance in the nanoparticle is 2.7 Å.
Convergence of the calculation was in line with prior expectation, taking 60 wavefunction optimisation iterations to converge to the agreed upon convergence criteria. So, the number of requisite iterations did increase with the jump in atom count, but it was still possible to run the single-point energy calculation on 1000 cores of the national supercomputer in less than one week including job queue time. As part of the calculation, electrostatic potential and electron density were saved to disk for analysis.

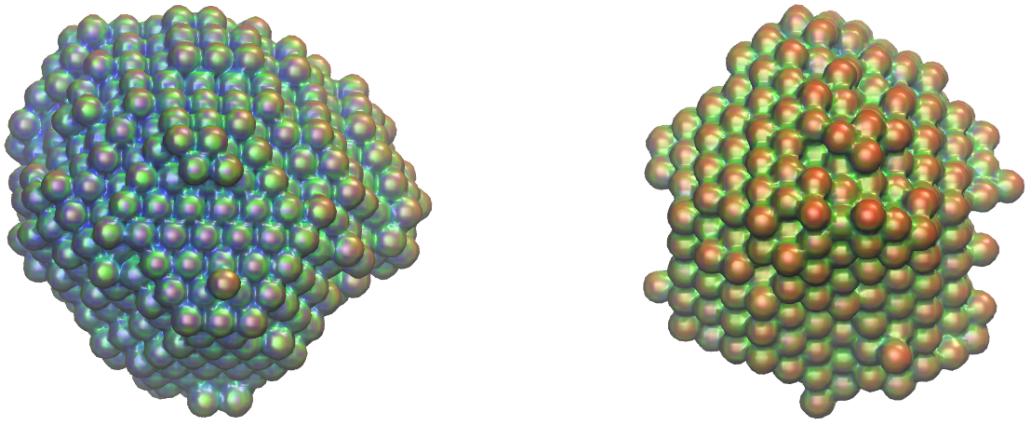


FIGURE B.10: A plot of the electrostatic potential at an isosurface of -2eV coloured by the electron density, showing areas of high to low electron density in red through green to blue. The left plot shows the particle from the top, looking toward the support, the right plot shows a profile view in the same orientation as that in fig. **??**.

Viewing the electrostatic density on an isosurface of -2eV coloured by the electron density shows which regions on the surface of the nanoparticle are relatively rich and poor in electron density (fig **??**). It was hypothesised that this information could be a useful indicator of potential adsorption sites and their relative strength, given prior knowledge of the electrophilic/phobic nature of the particular ligand of interest. In order to test the hypothesis, atomic oxygen was chosen as the ligand, as it has the lowest degrees of

freedom of the ligands which were tested on cuboctahedra, need not be performed spin polarised (calculation will be somewhat faster) and because it is known to be electrophilic.

Appendix C

A Hierarchical Matrix Inversion Method

Matrix inversion is often considered very bad practice[245] and methods to solve linear equations such as

$$\mathbf{Ax} = \mathbf{b}, \tag{C.1}$$

by factorising $\mathbf{A}$ rather than inverting it are well known, widely used and understood. Unfortunately, there are important cases in density matrix DFT where inverting the matrix is necessary. Such an example is if we want the elements of a density kernel through a rational representation of the matrix analogue of the Fermi-Dirac function. In other cases, such as for switching from a covariant representation of a quantity to a contravariant representation by multiplication by the inverse overlap matrix, are perfectly possible with a linear equation solver as we describe in chapter 6.

In this appendix, a method to invert a parallel sparse matrix is discussed in use the event that inversion is unavoidable. The method is based on the hierarchical application of the equations

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix}, \tag{C.2}$$

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}, \tag{C.3}$$

and the observation that the block columns of these two equivalent representations are independent of one another, such that

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}, \tag{C.4}$$

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix}, \tag{C.5}$$

are also correct. This is useful, because in a parallel computation, where the matrix is distributed over processors by column, the form in equation C.2 requires that the processors which have the latter block columns complete their operations, i.e. $(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}$ before the processors with the left block columns could start operations. This problem would be compounded with every step in the hierarchy, if the algorithm was made heirarchical.

The form in equation C.4 looks on first glance like it should be more costly to compute, but on a parallel, column distributed matrix, it offers better parallelisation oportunity. An algorithm will now be presented for heirarchically inverting a matrix using this form with only systolic (nearest-neighbour) explicit communications and matrix multiplications. The algorithm also requires *processor local* inversions of *small* matrix blocks at the lowest level of the heirarchy.

The first step is to define the form of the parallel, sparse matrices. They are column distributed over processors and as we decend the heirarchy they become more and more block diagonal,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{D} \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{A}'_1 & 0 & 0 & 0 \\ 0 & \mathbf{D}'_1 & 0 & 0 \\ 0 & 0 & \mathbf{A}'_2 & 0 \\ 0 & 0 & 0 & \mathbf{D}'_2 \end{bmatrix} \tag{C.6}$$

where the primed blocks in C.6 are half the dimension of the non-primed blocks, and share the sparsity pattern with their equivalent non-primed elements, but where off-diagonal non-zero elements become zero elements.

The first step in the algorithm is to set up four primitive matrices, which are block diagonal representations of the matrix at the next level up in the hierarchy

$$\alpha = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}, \tag{C.7}$$

$$\beta = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}, \tag{C.8}$$

$$\gamma = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}, \tag{C.9}$$

$$\delta = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}, \tag{C.10}$$

this is the picture at the first level of the hierarchy, at subsequent levels, matrix α is set up like in equation C.6, and the others transform in the same way with hierarchy level. α is generated trivially, by creating a sparsity pattern from the matrix at the previous level with zeroed **B** and **C** blocks and copying the (local) data into it. This is the same idea for γ, as the matrix is distributed by column, this requires no communications to set up.

The matrices β and δ *do* require communications to form them. Firstly, in the case of β the processors holding the sparsity pattern of the columns which make up **B** must send it to those processors which hold the sparsity pattern for **C** and *vice versa*. The sparsity pattern for the distributed matrix, β is then created and the matrix allocated. The data must then be sent and received in the same way as the sparsity patterns, into the newly allocated memory.

At this point, the α matrix must be inverted. This means that either the algorithm descends to the next level, or if the blocks are of sufficiently low dimension that they are local to processors and either dense or zero blocks (this shouldn't happen in theory, as it implies a singular matrix), then we use a standard Cholesky decomposition based inversion algorithm on them, or ignore them, respectively. This inversion is performed out of place into a matrix with the same sparsity pattern as the α matrix at this level, this matrix will be called

$$i \to \alpha^{-1}. \tag{C.11}$$

At this stage there is no further use for the α matrix and it can be deallocated.

Now a matrix is created which we will call *ib*, which is created with a sparsity pattern which is created as the product of the α^{-1} and β sparsity patterns. The product of α^{-1}

and β is then put into its elements

$$ib \to \alpha^{-1}\beta = \begin{bmatrix} \mathbf{A}^{-1}\mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1}\mathbf{C} \end{bmatrix}. \tag{C.12}$$

This matrix has row dimension of the α matrix and column dimension of the β matrix. At this stage we no longer require the β matrix and it can be deallocated.
The next step is to create a matrix which we will call *ci*, which is created with a sparsity pattern which is created as the product of the γ and α^{-1} sparsity patterns. The product of γ and α^{-1} is then put into its elements

$$ci \to \gamma\alpha^{-1} = \begin{bmatrix} \mathbf{C}\mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}\mathbf{D}^{-1} \end{bmatrix}. \tag{C.13}$$

This matrix has row dimension of the γ matrix and column dimension of the α matrix. Now we create a matrix which we will call *cib*, which is created with a sparsity pattern which is produced as the product of the *ci* and β sparsity patterns in union with the δ sparsity pattern (for the next step). The product of *ci* and β is then put into its elements

$$cib \to \gamma\alpha^{-1}\beta = \begin{bmatrix} \mathbf{C}\mathbf{A}^{-1}\mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}\mathbf{D}^{-1}\mathbf{C} \end{bmatrix}. \tag{C.14}$$

This matrix has row dimension of the *ci* matrix and column dimension of the β matrix. At this stage we have the peak number of block matrices allocated per level (6), which puts a hard limit on the memory usage of the algorithm. At this stage, the γ matrix is no longer required and may be deallocated.
The next step is to negate the matrix in *cib* and add γ to it, so that

$$dcib \to \delta - \gamma\alpha^{-1}\beta, \tag{C.15}$$

where the *cib* matrix is merely renamed *dcib* and all operations take place in place in its memory and sparsity pattern. After this δ can be deallocated.
At this point, the *dcib* matrix must be inverted. This means that either the algorithm descends to the next level again, or if the blocks are of sufficiently low dimension that they are local to processors and either dense or zero blocks (this shouldn't happen in theory, as it implies a singular matrix), then we use a standard Cholesky decomposition based inversion algorithm on them, or ignore them, respectively. This inversion is performed out of place into a matrix with the same sparsity pattern as the *dcib* matrix at this level,

this matrix will be called s (for Schur compliment):

$$s \to (\delta - \gamma\alpha^{-1}\beta)^{-1}. \tag{C.16}$$

At this stage there is no further use for the *dcib* matrix and it can be deallocated.
Now a matrix is created which we will call *sci*, which is created with a sparsity pattern which is created as the product of the s and ci sparsity patterns. The product of s and ci is then put into its elements

$$sci \to (\delta - \gamma\alpha^{-1}\beta)^{-1}\gamma\alpha^{-1} = \begin{bmatrix} (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}. \tag{C.17}$$

This matrix has row dimension of the s matrix and column dimension of the ci matrix. At this stage we no longer require the s matrix or the ci matrix and then can be deallocated.
After this step we need to create a matrix which we will call *ibsci*, which is created with a sparsity pattern which is created as the product of the *ib* and *sci* sparsity patterns. The product of *ib* and *sci* is then put into its elements

$$ibsci \to \alpha^{-1}\beta(\delta - \gamma\alpha^{-1}\beta)^{-1}\gamma\alpha^{-1} = \begin{bmatrix} \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}. \tag{C.18}$$

At this stage there is no further use for the *ib* matrix and it can be deallocated.
Next a block diagonal matrix is formed from the *sci* matrix, in a similar way to α

$$\mu = \begin{bmatrix} \mathbf{0} & \mathbf{B}_{\text{sci}} \\ \mathbf{C}_{\text{sci}} & \mathbf{0} \end{bmatrix}, \tag{C.19}$$

it is generated trivially, by creating a sparsity pattern from the *sci* matrix with zeroed $\mathbf{A}_{\text{sci}}$ and $\mathbf{D}_{\text{sci}}$ blocks and copying the (local) data into it. At this stage there is no further use for the *sci* matrix and it can be deallocated.
In the next step, the *ibsci* matrix is renamed as the *iibsci* matrix as the i matrix is added into it, in place

$$iibsci \to \alpha^{-1}\beta(\delta - \gamma\alpha^{-1}\beta)^{-1}\gamma\alpha^{-1} + \alpha^{-1}, \tag{C.20}$$

and at this stage there is no further use for the i matrix and it can be deallocated.
Finally, the inverse matrix is produced as the sum of the *iibsci* matrix and the μ matrix in a *iibsci* sparsity pattern and passed up the hierarchy. The remaining *iibsci* and μ are deallocated. This level returns control to the level above and this continues until we return back to the top of the hierarchy.

A simple serial implmentation of this idea was prepared in Fortran 2003 and it was found to work, but for large matrices, the errors introduced grew with the number of levels in the hierarchy. This could be solved simply by performing a single iteration of Newton-Schulz-Hotelling at each level, but this adds to the cost. This could be explored in greater depth, and/or implemented for true distributed, sparse matrices if someone had a need for it. An advantage is that this method can be run without any operations performed, as a dry run, only to determine a sparsity pattern of the inverse sparse matrix, and therein may lie the algorithms true value.

Bibliography

[1] Louis P. Lee, Nidia Gabaldon Limas, Daniel J. Cole, Mike C. Payne, Chris-Kriton Skylaris, and Thomas A. Manz. Expanding the scope of density derived electrostatic and chemical charge partitioning to thousands of atoms. *Journal of Chemical Theory and Computation*, 10(12):5377–5390, 2014. doi: 10.1021/ct500766v. URL http://dx.doi.org/10.1021/ct500766v. (Cited on pages ix, 95, and 96.)

[2] Valerio Vitale. *Computational methods for first-principles molecular dynamics with linear-scaling density functional theory*. PhD thesis, University of Southampton, 2017. (Cited on pages xi and 12.)

[3] David Jacobson. PEM Fuel Cells. http://physics.nist.gov/MajResFac/NIF/pemFuelCells.html, 2006. [Online; accessed 17-Sep-2017]. (Cited on pages xi and 29.)

[4] Andrzej Wieckowski. General discussion. *Faraday Discuss.*, 140:185–207, 2009. doi: 10.1039/B814700K. URL http://dx.doi.org/10.1039/B814700K. (Cited on pages xi and 30.)

[5] DJ Wheeler, JS Yi, R Fredley, D Yang, T Patterson, and L VanDine. Advancements in fuel cell stack technology at international fuel cells. *Journal of New Materials for Electrochemical Systems*, 4(4):233–238, 2001. (Cited on pages xi and 31.)

[6] J. Rossmeisl, A. Logadottir, and J.K. Nørskov. Electrolysis of water on (oxidized) metal surfaces. *Chemical Physics*, 319(1–3):178 – 184, 2005. ISSN 0301-0104. doi: http://dx.doi.org/10.1016/j.chemphys.2005.05.038. URL http://www.sciencedirect.com/science/article/pii/S0301010405002053. <ce:title>Molecular Charge Transfer in Condensed Media - from Physics and Chemistry to Biology and Nanoengineering in honour of Alexander M. Kuznetsov on his 65th birthday</ce:title>. (Cited on pages xi, 32, and 33.)

[7] Álvaro Ruiz-Serrano and Chris-Kriton Skylaris. A variational method for density functional theory calculations on metallic systems with thousands of atoms. *The Journal of Chemical Physics*, 139(5):054107, 2013. ISSN 00219606. doi: http://dx.doi.org/10.1063/1.4817001. URL http://scitation.aip.org/content/aip/

journal/jcp/139/5/10.1063/1.4817001. (Cited on pages xii, 46, 49, 62, 88, 114, 122, 152, 158, and 160.)

[8] Jolyon Aarons, Lewys Jones, Aakash Varambhia, Katherine E MacArthur, Dogan Ozkaya, Misbah Sarwar, Chris-Kriton Skylaris, and Peter D Nellist. Predicting the oxygen-binding properties of platinum nanoparticle ensembles by combining high-precision electron microscopy and density functional theory. *Nano Letters*, 17 (7):4003–4012, 2017. (Cited on pages xiv, xv, xvi, xix, 111, 115, 116, 117, 120, and 133.)

[9] Lewys Jones, Katherine E. MacArthur, Vidar T. Fauske, Antonius T. J. van Helvoort, and Peter D. Nellist. Rapid estimation of catalyst nanoparticle morphology and atomic-coordination by high-resolution z-contrast electron microscopy. *Nano Letters*, 14(11):6336–6341, nov 2014. ISSN 1530-6992. doi: 10.1021/nl502762m. URL http://dx.doi.org/10.1021/nl502762m. PMID: 25340541. (Cited on pages xv, 113, 119, 120, and 134.)

[10] Federico Calle-Vallejo, José I. Martínez, Juan M. García-Lastra, Philippe Sautet, and David Loffreda. Fast prediction of adsorption properties for platinum nanocatalysts with generalized coordination numbers. *Angewandte Chemie - International Edition*, 53(32):8316–8319, 2014. ISSN 15213773. doi: 10.1002/anie.201402958. (Cited on pages xvi, 71, 125, 128, and 130.)

[11] Jolyon Aarons, Misbah Sarwar, David Thompsett, and Chris-Kriton Skylaris. Perspective: Methods for large-scale density functional calculations on metallic systems. *The Journal of Chemical Physics*, 145(22):220901, dec 2016. ISSN 0021-9606. doi: 10.1063/1.4972007. (Cited on pages xix, 35, and 114.)

[12] L. G. Verga, J. Aarons, M. Sarwar, D. Thompsett, A. E. Russell, and C.-K. Skylaris. Effect of graphene support on large pt nanoparticles. *Phys. Chem. Chem. Phys.*, 18:32713–32722, 2016. doi: 10.1039/C6CP07334D. URL http://dx.doi.org/10.1039/C6CP07334D. (Cited on pages xix, 65, and 105.)

[13] Tech radar. Intel predicts 1,200 quintillion transistors in the world by 2015, September 2011. URL http://www.techradar.com/news/computing-components/processors/intel-predicts-1-200-quintillion-transistors-in-the-world-by-2015-1025851. (Cited on page 2.)

[14] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B): B864–B871, Nov 1964. doi: 10.1103/PhysRev.136.B864. (Cited on page 3.)

[15] Paul AM Dirac. A theory of electrons and protons. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 126, pages 360–365. The Royal Society, 1930. (Cited on page 7.)

[16] Murray Gell-Mann and Keith A Brueckner. Correlation energy of an electron gas at high density. *Physical Review*, 106(2):364, 1957. (Cited on page 7.)

[17] David M Ceperley and BJ Alder. Ground state of the electron gas by a stochastic method. *Physical Review Letters*, 45(7):566, 1980. (Cited on page 7.)

[18] John P Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23(10):5048, 1981. (Cited on page 7.)

[19] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, Oct 1996. doi: 10.1103/PhysRevLett.77.3865. URL http://link.aps.org/doi/10.1103/PhysRevLett.77.3865. (Cited on page 7.)

[20] B. Hammer, L. B. Hansen, and J. K. Nørskov. Improved adsorption energetics within density-functional theory using revised perdew-burke-ernzerhof functionals. *Phys. Rev. B*, 59(11):7413–7421, Mar 1999. ISSN 0163-1829. doi: 10.1103/PhysRevB.59.7413. URL http://link.aps.org/doi/10.1103/PhysRevB.59.7413. (Cited on pages 8, 89, and 122.)

[21] Ke Yang, Jingjing Zheng, Yan Zhao, and Donald G Truhlar. Tests of the rpbe, revpbe, τ-hcthhyb, ωb97x-d, and mohlyp density functional approximations and 29 others against representative databases for diverse bond energies and barrier heights in catalysis. *The Journal of chemical physics*, 132(16):164117, 2010. (Cited on pages 8 and 89.)

[22] Carlo A. Rozzi, Daniele Varsano, Andrea Marini, Eberhard K. U. Gross, and Angel Rubio. Exact coulomb cutoff technique for supercell calculations. *Phys. Rev. B*, 73:205119, May 2006. doi: 10.1103/PhysRevB.73.205119. URL http://link.aps.org/doi/10.1103/PhysRevB.73.205119. (Cited on page 11.)

[23] James C Phillips and Leonard Kleinman. New method for calculating wave functions in crystals and molecules. *Physical Review*, 116(2):287, 1959. (Cited on page 12.)

[24] Leonard Kleinman and DM Bylander. Efficacious form for model pseudopotentials. *Physical Review Letters*, 48(20):1425, 1982. (Cited on page 13.)

[25] Konstantin N Kudin and Gustavo E Scuseria. Linear-scaling density-functional theory with gaussian orbitals and periodic boundary conditions: Efficient evaluation of energy and forces via the fast multipole method. *Physical Review B*, 61 (24):16440, 2000. (Cited on page 17.)

[26] W. Kohn. Analytic properties of bloch waves and wannier functions. *Phys. Rev.*, 115:809–821, Aug 1959. doi: 10.1103/PhysRev.115.809. URL http://link.aps.org/doi/10.1103/PhysRev.115.809. (Cited on page 17.)

[27] Chris-Kriton Skylaris, Arash A. Mostofi, Peter D. Haynes, Oswaldo Diéguez, and Mike C. Payne. Nonorthogonal generalized wannier function pseudopotential plane-wave method. *Phys. Rev. B*, 66:035119, Jul 2002. doi: 10.1103/PhysRevB.66.035119. URL http://link.aps.org/doi/10.1103/PhysRevB.66.035119. (Cited on page 18.)

[28] Nicholas DM Hine. Linear-scaling density functional theory using the projector augmented wave method. *Journal of Physics: Condensed Matter*, 29(2):024001, 2016. (Cited on pages 18 and 159.)

[29] Paul P Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921. (Cited on page 21.)

[30] Rev McWeeny. Some recent advances in density matrix theory. *Reviews of Modern Physics*, 32(2):335, 1960. (Cited on page 21.)

[31] X-P Li, RW Nunes, and David Vanderbilt. Density-matrix electronic-structure method with linear system-size scaling. *Physical Review B*, 47(16):10891, 1993. (Cited on page 21.)

[32] RW Nunes and David Vanderbilt. Generalization of the density-matrix method to a nonorthogonal basis. *Physical Review B*, 50(23):17611, 1994. (Cited on pages 21 and 50.)

[33] John M Millam and Gustavo E Scuseria. Linear scaling conjugate gradient density matrix search as an alternative to diagonalization for first principles electronic structure calculations. *The Journal of chemical physics*, 106(13):5569–5577, 1997. (Cited on page 22.)

[34] DR Bowler and MJ Gillan. Density matrices in o (n) electronic structure calculations: theory and applications. *Computer physics communications*, 120(2-3): 95–108, 1999. (Cited on page 22.)

[35] Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976. doi: 10.1103/PhysRevB.13.

5188. URL http://link.aps.org/doi/10.1103/PhysRevB.13.5188. (Cited on page 24.)

[36] Peter E. Blöchl, O. Jepsen, and O. K. Andersen. Improved tetrahedron method for brillouin-zone integrations. *Phys. Rev. B*, 49:16223–16233, Jun 1994. doi: 10.1103/PhysRevB.49.16223. URL http://link.aps.org/doi/10.1103/PhysRevB.49.16223. (Cited on page 24.)

[37] M J Gillan. Calculation of the vacancy formation energy in aluminium. *Journal of Physics: Condensed Matter*, 1(4):689, 1989. URL http://stacks.iop.org/0953-8984/1/i=4/a=005. (Cited on pages 24 and 39.)

[38] W Kohn. Density functional and density matrix method scaling linearly with the number of atoms. *PHYSICAL REVIEW LETTERS*, 76(17):3168–3171, APR 22 1996. ISSN 0031-9007. doi: {10.1103/PhysRevLett.76.3168}. (Cited on page 25.)

[39] D. R. Bowler, R. Choudhury, M. J. Gillan, and T. Miyazaki. Recent progress with large-scale ab initio calculations: the CONQUEST code. *physica status solidi (b)*, 243(5):989–1000, 2006. ISSN 1521-3951. doi: 10.1002/pssb.200541386. URL http://dx.doi.org/10.1002/pssb.200541386. (Cited on page 25.)

[40] Chris-Kriton Skylaris, Peter D. Haynes, Arash A. Mostofi, and Mike C. Payne. Introducing ONETEP: Linear-scaling density functional simulations on parallel computers. *The Journal of Chemical Physics*, 122(8):084119, 2005. ISSN 00219606. doi: http://dx.doi.org/10.1063/1.1839852. URL http://scitation.aip.org/content/aip/journal/jcp/122/8/10.1063/1.1839852. (Cited on pages 25, 88, 114, 122, 138, 139, and 152.)

[41] Joost VandeVondele, Urban Borštnik, and Jürg Hutter. Linear scaling self-consistent field calculations with millions of atoms in the condensed phase. *Journal of Chemical Theory and Computation*, 8(10):3565–3573, 2012. doi: 10.1021/ct200897x. URL http://dx.doi.org/10.1021/ct200897x. PMID: 26593003. (Cited on page 25.)

[42] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. García, and J. M. Soler. Linear-scaling ab-initio calculations for large and complex systems. *physica status solidi (b)*, 215(1):809–817, 1999. ISSN 1521-3951. doi: 10.1002/(SICI)1521-3951(199909)215:1<809::AID-PSSB809>3.0.CO;2-0. URL http://dx.doi.org/10.1002/(SICI)1521-3951(199909)215:1<809::AID-PSSB809>3.0.CO;2-0. (Cited on page 25.)

[43] Luigi Genovese, Alexey Neelov, Stefan Goedecker, Thierry Deutsch, Seyed Alireza Ghasemi, Alexander Willand, Damien Caliste, Oded Zilberberg, Mark Rayson,

Anders Bergman, and Reinhold Schneider. Daubechies wavelets as a basis set for density functional pseudopotential calculations. *The Journal of Chemical Physics*, 129(1):014109, 2008. doi: http://dx.doi.org/10.1063/1.2949547. URL http://scitation.aip.org/content/aip/journal/jcp/129/1/10.1063/1.2949547. (Cited on page 25.)

[44] Stefan Goedecker. Linear scaling electronic structure methods. *Rev. Mod. Phys.*, 71:1085–1123, Jul 1999. doi: 10.1103/RevModPhys.71.1085. URL http://link.aps.org/doi/10.1103/RevModPhys.71.1085. (Cited on page 25.)

[45] D R Bowler and T Miyazaki. O(N) methods in electronic structure calculations. *Reports on Progress in Physics*, 75(3):036503, mar 2012. ISSN 0034-4885. doi: 10.1088/0034-4885/75/3/036503. URL http://stacks.iop.org/0034-4885/75/i=3/a=036503. (Cited on page 25.)

[46] S Goedecker. Decay properties of the finite-temperature density matrix in metals. *Physical Review B*, 58(7):3501, 1998. (Cited on pages 25, 60, and 139.)

[47] Riccardo Ferrando, Julius Jellinek, and Roy L Johnston. Nanoalloys: from theory to applications of alloy clusters and nanoparticles. *Chemical reviews*, 108(3):845–910, 2008. (Cited on page 26.)

[48] Prashant K. Jain, Wenyu Huang, and Mostafa A. El-Sayed. On the universal scaling behavior of the distance decay of plasmon coupling in metal nanoparticle pairs: A plasmon ruler equation. *Nano Letters*, 7(7):2080–2088, 2007. doi: 10.1021/nl071008a. URL http://dx.doi.org/10.1021/nl071008a. (Cited on page 26.)

[49] Nathaniel L. Rosi and Chad A. Mirkin. Nanostructures in biodiagnostics. *Chemical Reviews*, 105(4):1547–1562, 2005. doi: 10.1021/cr030067f. URL http://dx.doi.org/10.1021/cr030067f. PMID: 15826019. (Cited on page 26.)

[50] Cono Di Paola, Roberto D'Agosta, and Francesca Baletto. Geometrical effects on the magnetic properties of nanoparticles. *Nano letters*, 16(4):2885–2889, 2016. (Cited on page 26.)

[51] L. V. Lutsev, A. I. Stognij, and N. N. Novitskii. Giant magnetoresistance in semiconductor/granular film heterostructures with cobalt nanoparticles. *Phys. Rev. B*, 80:184423, Nov 2009. doi: 10.1103/PhysRevB.80.184423. URL http://link.aps.org/doi/10.1103/PhysRevB.80.184423. (Cited on page 26.)

[52] Wendong Wang, Fengwu Zhu, Jun Weng, Jimei Xiao, and Wuyan Lai. Nanoparticle morphology in a granular cu-co alloy with giant magnetoresistance. *Applied Physics Letters*, 72(9):1118–1120, 1998. doi: http://dx.doi.org/10.1063/1.120942.

URL http://scitation.aip.org/content/aip/journal/apl/72/9/10.1063/1.120942. (Cited on page 26.)

[53] Francesca Baletto and Riccardo Ferrando. Structural properties of nanoclusters: Energetic, thermodynamic, and kinetic effects. *Reviews of modern physics*, 77(1): 371, 2005. (Cited on page 26.)

[54] Yuguang Ma and Perla B Balbuena. Pt surface segregation in bimetallic pt 3 m alloys: a density functional theory study. *Surface Science*, 602(1):107–113, 2008. (Cited on page 26.)

[55] Gustavo E Ramirez-Caballero and Perla B Balbuena. Surface segregation of core atoms in core–shell structures. *Chemical Physics Letters*, 456(1):64–67, 2008. (Cited on page 26.)

[56] Ingmar Swart, Frank MF De Groot, Bert M Weckhuysen, Philipp Gruene, Gerard Meijer, and Andre Fielicke. H2 adsorption on 3d transition metal clusters: A combined infrared spectroscopy and density functional study. *The Journal of Physical Chemistry A*, 112(6):1139–1149, 2008. (Cited on page 26.)

[57] F. Tournus, K. Sato, T. Epicier, T. J. Konno, and V. Dupuis. Multi-$l1_0$ domain copt and fept nanoparticles revealed by electron microscopy. *Phys. Rev. Lett.*, 110: 055501, Jan 2013. doi: 10.1103/PhysRevLett.110.055501. URL http://link.aps.org/doi/10.1103/PhysRevLett.110.055501. (Cited on page 26.)

[58] Jens Kehlet Nørskov, Jan Rossmeisl, Ashildur Logadottir, LRKJ Lindqvist, John R Kitchin, Thomas Bligaard, and Hannes Jonsson. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *The Journal of Physical Chemistry B*, 108 (46):17886–17892, nov 2004. ISSN 1520-6106. doi: 10.1021/jp047349j. (Cited on pages 26, 30, 32, 33, 81, 105, 109, 115, and 129.)

[59] JR Kitchin, Jens Kehlet Nørskov, MA Barteau, and JG Chen. Modification of the surface electronic and chemical properties of pt (111) by subsurface 3d transition metals. *Journal of Chemical Physics*, 120(21):10240–10246, 2004. (Cited on page 26.)

[60] Jeff Greeley and Manos Mavrikakis. Alloy catalysts designed from first principles. *Nature materials*, 3(11):810–815, 2004. (Cited on page 26.)

[61] Yuguang Ma and Perla B Balbuena. Surface properties and dissolution trends of pt3m alloys in the presence of adsorbates. *The Journal of Physical Chemistry C*, 112(37):14520–14528, 2008. (Cited on page 26.)

[62] Anders Hellman, Andrea Resta, NM Martin, Johan Gustafson, Adriana Trinchero, P-A Carlsson, Olivier Balmes, Roberto Felici, Richard van Rijn, JWM Frenken, et al. The active phase of palladium during methane oxidation. *The journal of physical chemistry letters*, 3(6):678–682, 2012. (Cited on page 26.)

[63] Energy Information Administration (US). *International Energy Outlook, 2013*. Government Printing Office, 2013. (Cited on page 28.)

[64] Steven G Chalk and James F Miller. Key challenges and recent progress in batteries, fuel cells, and hydrogen storage for clean energy systems. *Journal of Power Sources*, 159(1):73–80, 2006. (Cited on page 28.)

[65] Francis T Bacon. Fuel cells, past, present and future. *Electrochimica Acta*, 14(7): 569–585, 1969. (Cited on page 29.)

[66] Chaojie Song and Jiujun Zhang. *Electrocatalytic Oxygen Reduction Reaction*. Springer London, 2008. ISBN 978-1-84800-935-6. doi: 10.1007/978-1-84800-936-3_2. URL http://dx.doi.org/10.1007/978-1-84800-936-3_2. (Cited on page 30.)

[67] Dong-Hee Lim and Jennifer Wilcox. Mechanisms of the oxygen reduction reaction on defective graphene-supported pt nanoparticles from first-principles. *The Journal of Physical Chemistry C*, 116(5):3653–3660, 2012. (Cited on page 30.)

[68] Ernest Yeager. Electrocatalysts for o_2 reduction. *Electrochimica Acta*, 29(11): 1527–1537, 1984. (Cited on page 30.)

[69] NM Markovic, TJ Schmidt, V Stamenkovic, and PN Ross. Oxygen reduction reaction on pt and pt bimetallic surfaces: A selective review. *FUEL CELLS-WEINHEIM-*, 1(2):105–116, 2001. (Cited on pages 30 and 32.)

[70] A Panchenko, MTM Koper, TE Shubina, SJ Mitchell, and E Roduner. Ab initio calculations of intermediates of oxygen reduction on low-index platinum surfaces. *Journal of the Electrochemical Society*, 151(12):A2016–A2027, 2004. (Cited on page 30.)

[71] Hubert A Gasteiger, Shyam S Kocha, Bhaskar Sompalli, and Frederick T Wagner. Activity benchmarks and requirements for pt, pt-alloy, and non-pt oxygen reduction catalysts for pemfcs. *Applied Catalysis B: Environmental*, 56(1):9–35, 2005. (Cited on page 30.)

[72] J Zhang, Y Mo, MB Vukmirovic, R Klie, K Sasaki, and RR Adzic. Platinum monolayer electrocatalysts for o2 reduction: Pt monolayer on pd (111) and on

carbon-supported pd nanoparticles. *The Journal of Physical Chemistry B*, 108 (30):10955–10964, 2004. (Cited on page 30.)

[73] Olivier Antoine, Yann Bultel, and Robert Durand. Oxygen reduction reaction kinetics and mechanism on platinum nanoparticles inside nafion< sup>®</sup>. *Journal of Electroanalytical Chemistry*, 499(1):85–94, 2001. (Cited on pages 30 and 32.)

[74] Minhua Shao, Kotaro Sasaki, Nebojsa S Marinkovic, Lihua Zhang, and Radoslav R Adzic. Synthesis and characterization of platinum monolayer oxygen-reduction electrocatalysts with co–pd core–shell nanoparticle supports. *Electrochemistry Communications*, 9(12):2848–2853, 2007. (Cited on page 30.)

[75] Jia X Wang, Hiromi Inada, Lijun Wu, Yimei Zhu, YongMan Choi, Ping Liu, Wei-Ping Zhou, and Radoslav R Adzic. Oxygen reduction on well-defined core- shell nanocatalysts: particle size, facet, and pt shell thickness effects. *Journal of the American Chemical Society*, 131(47):17298–17302, 2009. (Cited on page 30.)

[76] J Greeley, IEL Stephens, AS Bondarenko, Tobias Peter Johansson, Heine Anton Hansen, TF Jaramillo, Jan Rossmeisl, and Jens Kehlet Norskov. Alloys of platinum and early transition metals as oxygen reduction electrocatalysts. *Nature Chemistry*, 1(7):552–556, oct 2009. ISSN 1755-4330. doi: 10.1038/nchem.367. (Cited on pages 31, 33, 81, 115, 129, and 186.)

[77] A Damjanovic, MA Genshaw, and JO'M Bockris. Distinction between intermediates produced in main and side electrodic reactions. *The Journal of Chemical Physics*, 45:4057, 1966. (Cited on page 32.)

[78] Lars Grabow, Ye Xu, and Manos Mavrikakis. Lattice strain effects on co oxidation on pt (111). *Physical Chemistry Chemical Physics*, 8(29):3369–3374, 2006. (Cited on page 32.)

[79] B Hammer and JK Norskov. Why gold is the noblest of all the metals. *Nature*, 376(6537):238–240, 1995. (Cited on pages 33, 96, and 109.)

[80] Vojislav Stamenkovic, Bongjin Simon Mun, Karl JJ Mayrhofer, Philip N Ross, Nenad M Markovic, Jan Rossmeisl, Jeff Greeley, and Jens K Norskov. Changing the activity of electrocatalysts for oxygen reduction by tuning the surface electronic structure. *Angewandte Chemie*, 118(18):2963–2967, apr 2006. ISSN 1433-7851. doi: 10.1002/anie.200504386. (Cited on pages 33, 81, 113, 114, 115, 129, 130, and 135.)

[81] N. David Mermin. Thermal properties of the inhomogeneous electron gas. *Phys. Rev.*, 137(5A):A1441–A1443, Mar 1965. ISSN 0031899X. doi: 10.1103/

PhysRev.137.A1441. URL http://link.aps.org/doi/10.1103/PhysRev.137.A1441. (Cited on pages 37, 122, and 139.)

[82] M. Methfessel and A. T. Paxton. High-precision sampling for brillouin-zone integration in metals. *Phys. Rev. B*, 40:3616–3621, Aug 1989. doi: 10.1103/PhysRevB.40.3616. URL http://link.aps.org/doi/10.1103/PhysRevB.40.3616. (Cited on page 39.)

[83] Nicola Marzari, David Vanderbilt, Alessandro De Vita, and M. C. Payne. Thermal contraction and disordering of the al(110) surface. *Phys. Rev. Lett.*, 82:3296–3299, Apr 1999. doi: 10.1103/PhysRevLett.82.3296. URL http://link.aps.org/doi/10.1103/PhysRevLett.82.3296. (Cited on page 39.)

[84] Nicola Marzari. *Ab-initio molecular dynamics for metallic systems*. PhD thesis, University of Cambridge, 1996. URL http://theossrv1.epfl.ch/uploads/Group/Bibliography/Nicola%20Marzari_these_04-1996.pdf. accessed 17-Sep-2017. (Cited on page 39.)

[85] G. Galli. *Large-Scale Electronic Structure Calculations Using Linear Scaling Methods*. Wiley-VCH Verlag GmbH & Co. KGaA, 2005. ISBN 9783527603107. doi: 10.1002/3527603107.ch10. URL http://dx.doi.org/10.1002/3527603107.ch10. (Cited on page 41.)

[86] Peter Pulay. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters*, 73(2):393–398, 1980. (Cited on page 41.)

[87] Peter Pulay. Improved scf convergence acceleration. *Journal of Computational Chemistry*, 3(4):556–560, 1982. (Cited on page 41.)

[88] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16): 11169, 1996. (Cited on pages 41 and 42.)

[89] D Alfè, MJ Gillan, and GD Price. Ab initio chemical potentials of solid and liquid solutions and the chemistry of the earth's core. *The Journal of chemical physics*, 116(16):7127–7136, 2002. (Cited on pages 41 and 62.)

[90] Lianheng Tong. Metal CONQUEST, HECToR dCSE report. http://www.hector.ac.uk/cse/distributedcse/reports/conquest/conquest.pdf, March 2011. Online; accessed 17-Sep-2017. (Cited on page 41.)

[91] GP Kerker. Efficient iteration scheme for self-consistent pseudopotential calculations. *Physical Review B*, 23(6):3082, 1981. (Cited on page 43.)

[92] WanZhen Liang, Chandra Saravanan, Yihan Shao, Roi Baer, Alexis T Bell, and Martin Head-Gordon. Improved Fermi operator expansion methods for fast electronic structure calculations. *The Journal of chemical physics*, 119(8):4117–4125, 2003. (Cited on pages 43, 55, 138, 145, 147, and 149.)

[93] Nicola Marzari, David Vanderbilt, and Mike C Payne. Ensemble density-functional theory for ab initio molecular dynamics of metals and finite-temperature insulators. *Physical review letters*, 79(7):1337, 1997. (Cited on pages 43, 44, and 53.)

[94] Phil Hasnip. Private communication, Dec 2016. (Cited on page 45.)

[95] Christoph Freysoldt, Sixten Boeck, and Jörg Neugebauer. Direct minimization technique for metals in density functional theory. *Phys. Rev. B*, 79:241103, Jun 2009. doi: 10.1103/PhysRevB.79.241103. URL http://link.aps.org/doi/10.1103/PhysRevB.79.241103. (Cited on page 45.)

[96] Branislav Jansík, Stinne Høst, Poul Jørgensen, Jeppe Olsen, and Trygve Helgaker. Linear-scaling symmetric square-root decomposition of the overlap matrix. *The Journal of chemical physics*, 126(12):124104, 2007. (Cited on page 47.)

[97] Xiaoye S Li and James W Demmel. Superlu_dist: A scalable distributed-memory sparse direct solver for unsymmetric linear systems. *ACM Transactions on Mathematical Software (TOMS)*, 29(2):110–140, 2003. (Cited on pages 50 and 52.)

[98] William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical Recipes: The art of scientific computing (Cambridge*, volume 683. Cambridge Univ. Press, 1992. (Cited on page 51.)

[99] Victor Pan and Robert Schreiber. An improved newton iteration for the generalized inverse of a matrix, with applications. *SIAM Journal on Scientific and Statistical Computing*, 12(5):1109–1130, 1991. (Cited on page 51.)

[100] T Ozaki. Efficient recursion method for inverting an overlap matrix. *Physical Review B*, 64(19):195110, 2001. (Cited on page 51.)

[101] Adam HR Palser and David E Manolopoulos. Canonical purification of the density matrix in electronic-structure theory. *Physical Review B*, 58(19):12704, 1998. (Cited on page 51.)

[102] Peter D Haynes, Chris-Kriton Skylaris, Arash A Mostofi, and Mike C Payne. Density kernel optimization in the ONETEP code. *Journal of Physics: Condensed Matter*, 20(29):294207, 2008. (Cited on page 51.)

[103] Lin Lin, Chao Yang, Juan C Meza, Jianfeng Lu, Lexing Ying, et al. Selinv—an algorithm for selected inversion of a sparse symmetric matrix. *ACM Transactions on Mathematical Software (TOMS)*, 37(4):40, 2011. (Cited on pages 51, 52, and 161.)

[104] S. Li and E. Darve. Extension and optimization of the FIND algorithm: Computing Green's and less-than Green's functions. *Journal of Computational Physics*, 231(4):1121 – 1139, 2012. ISSN 0021-9991. doi: http://dx.doi.org/10.1016/j.jcp.2011.05.027. URL http://www.sciencedirect.com/science/article/pii/S002199911100338X. (Cited on page 52.)

[105] Franҫis Henry Rouet. *Thesis: Partial computation of the inverse of a large sparse matrix - application to astrophysics.* Institut national polytechnique de Toulouse, 2009. (Cited on page 52.)

[106] Patrick R Amestoy, Iain S Duff, Jean-Yves L'Excellent, and Jacko Koster. A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM Journal on Matrix Analysis and Applications*, 23(1):15–41, 2001. (Cited on page 52.)

[107] Nicholas IM Gould, Jennifer A Scott, and Yifan Hu. A numerical evaluation of sparse direct solvers for the solution of large sparse symmetric linear systems of equations. *ACM Transactions on Mathematical Software (TOMS)*, 33(2):10, 2007. (Cited on page 52.)

[108] George Karypis and Vipin Kumar. A parallel algorithm for multilevel graph partitioning and sparse matrix ordering. *Journal of Parallel and Distributed Computing*, 48(1):71–95, 1998. (Cited on page 53.)

[109] George Karypis, Kirk Schloegel, and Vipin Kumar. Parmetis: Parallel graph partitioning and sparse matrix ordering library. *Dept. of Computer Science, University of Minnesota*, 1997. (Cited on page 53.)

[110] Cédric Chevalier and François Pellegrini. Pt-scotch: A tool for efficient parallel graph ordering. *Parallel computing*, 34(6):318–331, 2008. (Cited on page 53.)

[111] Stefan Goedecker and L Colombo. Efficient linear scaling algorithm for tight-binding molecular dynamics. *Physical review letters*, 73(1):122, 1994. (Cited on page 53.)

[112] MPAT Methfessel and AT Paxton. High-precision sampling for brillouin-zone integration in metals. *Physical Review B*, 40(6):3616, 1989. (Cited on page 53.)

[113] Anders M. N. Niklasson. Iterative refinement method for the approximate factorization of a matrix inverse. *Phys. Rev. B*, 70:193102, Nov 2004. doi: 10.1103/

PhysRevB.70.193102. URL http://link.aps.org/doi/10.1103/PhysRevB.70.193102. (Cited on page 54.)

[114] Emanuel H. Rubensson, Nicolas Bock, Erik Holmström, and Anders M. N. Niklasson. Recursive inverse factorization. *The Journal of Chemical Physics*, 128(10):104105, 2008. doi: http://dx.doi.org/10.1063/1.2884921. URL http://scitation.aip.org/content/aip/journal/jcp/128/10/10.1063/1.2884921. (Cited on page 54.)

[115] Stefan Goedecker and M Teter. Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals. *Physical Review B*, 51 (15):9455, 1995. (Cited on pages 54 and 148.)

[116] Roi Baer and Martin Head-Gordon. Chebyshev expansion methods for electronic structure calculations on large molecular systems. *The Journal of chemical physics*, 107(23):10003–10013, 1997. (Cited on page 55.)

[117] Michael S Paterson and Larry J Stockmeyer. On the number of nonscalar multiplications necessary to evaluate polynomials. *SIAM Journal on Computing*, 2(1): 60–66, 1973. (Cited on page 55.)

[118] Florian R. Krajewski and Michele Parrinello. Stochastic linear scaling for metals and nonmetals. *Phys. Rev. B*, 71:233105, Jun 2005. doi: 10.1103/PhysRevB.71.233105. URL http://link.aps.org/doi/10.1103/PhysRevB.71.233105. (Cited on page 56.)

[119] Michele Ceriotti, Thomas D Kühne, and Michele Parrinello. An efficient and accurate decomposition of the Fermi operator. *The Journal of chemical physics*, 129 (2):024707, 2008. (Cited on pages 56 and 138.)

[120] Florian R. Krajewski and Michele Parrinello. Linear scaling electronic structure monte carlo method for metals. *Phys. Rev. B*, 75:235108, Jun 2007. doi: 10.1103/PhysRevB.75.235108. URL http://link.aps.org/doi/10.1103/PhysRevB.75.235108. (Cited on page 56.)

[121] Florian R. Krajewski and Michele Parrinello. Linear scaling for quasi-one-dimensional systems. *Phys. Rev. B*, 74:125107, Sep 2006. doi: 10.1103/PhysRevB.74.125107. URL http://link.aps.org/doi/10.1103/PhysRevB.74.125107. (Cited on page 56.)

[122] Dorothee Richters and Thomas D Kühne. Self-consistent field theory based molecular dynamics with linear system-size scaling. *The Journal of chemical physics*, 140(13):134109, 2014. (Cited on page 57.)

[123] Michele Ceriotti, Thomas D Kühne, and Michele Parrinello. A hybrid approach to fermi operator expansion. *arXiv preprint arXiv:0809.2232*, 2008. (Cited on page 57.)

[124] R Haydock, V Heine, and M J Kelly. Electronic structure based on the local atomic environment for tight-binding bands. ii. *Journal of Physics C: Solid State Physics*, 8 (16):2591, 1975. URL http://stacks.iop.org/0022-3719/8/i=16/a=011. (Cited on page 57.)

[125] R Haydock, V Heine, and M J Kelly. Electronic structure based on the local atomic environment for tight-binding bands. *Journal of Physics C: Solid State Physics*, 5 (20):2845, 1972. URL http://stacks.iop.org/0022-3719/5/i=20/a=004. (Cited on page 57.)

[126] Anders M. N. Niklasson. Implicit purification for temperature-dependent density matrices. *Phys. Rev. B*, 68:233104, Dec 2003. doi: 10.1103/PhysRevB.68.233104. URL http://link.aps.org/doi/10.1103/PhysRevB.68.233104. (Cited on page 57.)

[127] S Goedecker. Integral representation of the Fermi distribution and its applications in electronic-structure calculations. *Physical Review B*, 48(23):17573, 1993. (Cited on pages 58 and 138.)

[128] Stefan Goedecker. Low complexity algorithms for electronic structure calculations. *Journal of Computational Physics*, 118(2):261–268, 1995. ISSN 0021-9991. doi: http://dx.doi.org/10.1006/jcph.1995.1097. URL http://www.sciencedirect.com/science/article/pii/S0021999185710972. (Cited on pages 58 and 138.)

[129] Lin Lin, Jianfeng Lu, Lexing Ying, and E Weinan. Pole-based approximation of the Fermi-Dirac function. *Chinese Annals of Mathematics, Series B*, 30(6):729–742, 2009. (Cited on pages 59 and 138.)

[130] Lin Lin, Jianfeng Lu, Roberto Car, and E Weinan. Multipole representation of the Fermi operator with application to the electronic structure analysis of metallic systems. *Physical Review B*, 79(11):115133, 2009. (Cited on page 60.)

[131] Lin Lin, Mohan Chen, Chao Yang, and Lixin He. Accelerating atomic orbital-based electronic structure calculation via pole expansion and selected inversion. *Journal of Physics: Condensed Matter*, 25(29):295501, 2013. URL http://stacks.iop.org/0953-8984/25/i=29/a=295501. (Cited on page 61.)

[132] Anders M. N. Niklasson, Peter Steneteg, and Nicolas Bock. Extended lagrangian free energy molecular dynamics. *The Journal of Chemical Physics*, 135(16):164111,

2011. doi: http://dx.doi.org/10.1063/1.3656977. URL http://scitation.aip.org/content/aip/journal/jcp/135/16/10.1063/1.3656977. (Cited on page 61.)

[133] J Korringa. On the calculation of the energy of a Bloch wave in a metal. *Physica*, 13(6-7):392–400, 1947. (Cited on page 61.)

[134] W Kohn and N Rostoker. Solution of the schrödinger equation in periodic lattices with an application to metallic lithium. *Physical Review*, 94(5):1111, 1954. (Cited on page 61.)

[135] Duane D Johnson, FJ Pinski, and GM Stocks. Fast method for calculating the self-consistent electronic structure of random alloys. *Physical Review B*, 30(10): 5508, 1984. (Cited on page 61.)

[136] FJ Pinski and GM Stocks. Fast method for calculating the self-consistent electronic structure of random alloys. ii. optimal use of the complex plane. *Physical Review B*, 32(6):4204, 1985. (Cited on page 62.)

[137] Yang Wang, G. M. Stocks, W. A. Shelton, D. M. C. Nicholson, Z. Szotek, and W. M. Temmerman. Order- N multiple scattering approach to electronic structure calculations. *Phys. Rev. Lett.*, 75:2867–2870, Oct 1995. doi: 10.1103/PhysRevLett.75.2867. URL http://link.aps.org/doi/10.1103/PhysRevLett.75.2867. (Cited on page 62.)

[138] I. A. Abrikosov, A. M. N. Niklasson, S. I. Simak, B. Johansson, A. V. Ruban, and H. L. Skriver. Order- N Green's function technique for local environment effects in alloys. *Phys. Rev. Lett.*, 76:4203–4206, May 1996. doi: 10.1103/PhysRevLett.76.4203. URL http://link.aps.org/doi/10.1103/PhysRevLett.76.4203. (Cited on page 62.)

[139] Aftab Alam, Suffian N Khan, Andrei V Smirnov, DM Nicholson, and Duane D Johnson. Green's function multiple-scattering theory with a truncated basis set: An augmented-kkr formalism. *Physical Review B*, 90(20):205102, 2014. (Cited on page 62.)

[140] Rudolf Zeller. Towards a linear-scaling algorithm for electronic structure calculations with the tight-binding korringa–kohn–rostoker green function method. *Journal of Physics: Condensed Matter*, 20(29):294215, 2008. (Cited on page 62.)

[141] Lin Li, Ask H Larsen, Nichols A Romero, Vitali A Morozov, Christian Glinsvad, Frank Abild-Pedersen, Jeff Greeley, Karsten W Jacobsen, and Jens K Nørskov. Investigation of catalytic finite-size-effects of platinum metal clusters. *The journal of physical chemistry letters*, 4(1):222–226, 2012. (Cited on pages 62 and 95.)

[142] Phanish Suryanarayana. On nearsightedness in metallic systems for o (n) density functional theory calculations: A case study on aluminum. *Chemical Physics Letters*, 679:146–151, 2017. (Cited on page 64.)

[143] Jens K Nørskov, Frank Abild-Pedersen, Felix Studt, and Thomas Bligaard. Density functional theory in surface chemistry and catalysis. *Proceedings of the National Academy of Sciences of the United States of America*, 2011. (Cited on page 66.)

[144] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953–17979, Dec 1994. doi: 10.1103/PhysRevB.50.17953. URL http://link.aps.org/doi/10.1103/PhysRevB.50.17953. (Cited on pages 68 and 89.)

[145] Andrew J. Morris, Rebecca J. Nicholls, Chris J. Pickard, and Jonathan R. Yates. Optados: A tool for obtaining density of states, core-level and optical spectra from electronic structure codes. *Computer Physics Communications*, 185(5):1477 – 1485, 2014. ISSN 0010-4655. doi: http://dx.doi.org/10.1016/j.cpc.2014.02.013. URL http://www.sciencedirect.com/science/article/pii/S0010465514000460. (Cited on page 69.)

[146] Laura E. Ratcliff, Nicholas D. M. Hine, and Peter D. Haynes. Calculating optical absorption spectra for large systems using linear-scaling density functional theory. *Phys. Rev. B*, 84:165131, Oct 2011. doi: 10.1103/PhysRevB.84.165131. URL http://link.aps.org/doi/10.1103/PhysRevB.84.165131. (Cited on page 69.)

[147] Herbert Edelsbrunner and Ernst P. Mücke. Three-dimensional alpha shapes. *ACM Trans. Graph.*, 13(1):43–72, January 1994. ISSN 0730-0301. doi: 10.1145/174462.156635. URL http://doi.acm.org/10.1145/174462.156635. (Cited on page 72.)

[148] B.N. Delaunay. Sur la sphère vide. a la mémoire de georges voronoï. *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na*, 1934 (6):793–800, 1934. URL https://zbmath.org/?q=an:0010.41101|60.0946.06. (Cited on page 73.)

[149] Der-Tsai Lee and Bruce J Schachter. Two algorithms for constructing a delaunay triangulation. *International Journal of Computer & Information Sciences*, 9(3): 219–242, 1980. (Cited on page 73.)

[150] Jonathan Richard Shewchuk. Triangle: Engineering a 2d quality mesh generator and delaunay triangulator. In *Applied computational geometry towards geometric engineering*, pages 203–222. Springer, Berlin, Heidelberg, 1996. (Cited on page 73.)

[151] C Bradford Barber, David P Dobkin, and Hannu Huhdanpaa. The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22 (4):469–483, 1996. (Cited on page 73.)

[152] Michael K. Gilson, Kim A. Sharp, and Barry H. Honig. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *Journal of Computational Chemistry*, 9(4):327–335, 1988. ISSN 1096-987X. doi: 10.1002/jcc.540090407. URL http://dx.doi.org/10.1002/jcc.540090407. (Cited on page 75.)

[153] Walter Rocchia, Sundaram Sridharan, Anthony Nicholls, Emil Alexov, Alessandro Chiabrera, and Barry Honig. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *Journal of Computational Chemistry*, 23(1):128–137, 2002. ISSN 1096-987X. doi: 10.1002/jcc.1161. URL http://dx.doi.org/10.1002/jcc.1161. (Cited on page 75.)

[154] Michel F. Sanner, Arthur J. Olson, and Jean-Claude Spehner. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers*, 38(3):305–320, 1996. ISSN 1097-0282. doi: 10.1002/(SICI)1097-0282(199603)38:3<305::AID-BIP4>3.0.CO;2-Y. URL http://dx.doi.org/10.1002/(SICI)1097-0282(199603)38:3<305::AID-BIP4>3.0.CO;2-Y. (Cited on page 75.)

[155] Jacek Dziedzic, Stephen J Fox, Thomas Fox, Christofer S Tautermann, and Chris-Kriton Skylaris. Large-scale dft calculations in implicit solvent—a case study on the t4 lysozyme l99a/m102q protein. *International Journal of Quantum Chemistry*, 113(6):771–785, 2013. (Cited on pages 75 and 76.)

[156] Nicholas D. M. Hine, Jacek Dziedzic, Peter D. Haynes, and Chris-Kriton Skylaris. Electrostatic interactions in finite systems treated with periodic boundary conditions: Application to linear-scaling density functional theory. *The Journal of Chemical Physics*, 135(20):204103, 2011. ISSN 00219606. doi: 10.1063/1.3662863. URL http://arxiv.org/abs/1109.1970$\delimiter"026E30F$nhttp://scitation.aip.org/content/aip/journal/jcp/135/20/10.1063/1.3662863. (Cited on pages 76, 89, and 121.)

[157] Francois Lekien and J Marsden. Tricubic interpolation in three dimensions. *International Journal for Numerical Methods in Engineering*, 63(3):455–471, 2005. (Cited on pages 77 and 213.)

[158] Nicolas Bigaouette and Francois Lekien. Coefficients of the $\mathbf{B}^{-1}$ matrix in [157], December 2014. URL https://github.com/nbigaouette/libtricubic/blob/master/tricubic-1.0/src/libtricubic/coeff.h. accessed 17-Sep-2017. (Cited on page 79.)

[159] Donald E. Knuth. *The Art of Computer Programming, Volume 3: (2Nd Ed.) Sorting and Searching*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998. ISBN 0-201-89685-0. (Cited on page 80.)

[160] Jeffrey D. Bracken. The history and use of our earth's chemical elements: A reference guide (krebs, robert e.). *Journal of Chemical Education*, 76(4):475, 1999. doi: 10.1021/ed076p475. URL http://dx.doi.org/10.1021/ed076p475. (Cited on page 83.)

[161] DW McKee. Catalytic decomposition of hydrogen peroxide by metals and alloys of the platinum group. *Journal of Catalysis*, 14(4):355–364, 1969. (Cited on page 83.)

[162] Nuria Lopez, TVW Janssens, BS Clausen, Y Xu, Manos Mavrikakis, T Bligaard, and Jens Kehlet Nørskov. On the origin of the catalytic activity of gold nanoparticles for low-temperature co oxidation. *Journal of Catalysis*, 223(1):232–235, 2004. (Cited on page 84.)

[163] James Cookson. The preparation of palladium nanoparticles. *Platinum Metals Review*, 56(2):83–98, 2012. ISSN 0032-1400. doi: doi:10.1595/147106712X632415. URL http://www.ingentaconnect.com/content/matthey/pmr/2012/00000056/00000002/art00003. (Not cited.)

[164] Gledison S. Fonseca, Giovanna Machado, Sergio R. Teixeira, Gerhard H. Fecher, Jonder Morais, Maria C.M. Alves, and Jairton Dupont. Synthesis and characterization of catalytic iridium nanoparticles in imidazolium ionic liquids. *Journal of Colloid and Interface Science*, 301(1):193 – 204, 2006. ISSN 0021-9797. doi: https://doi.org/10.1016/j.jcis.2006.04.073. URL http://www.sciencedirect.com/science/article/pii/S0021979706003882. (Cited on page 84.)

[165] JW Arblaster. Crystallographic properties of platinum. *Platinum Metals Review*, 41(1):12–21, 1997. (Cited on page 84.)

[166] Ana Rodriguez, Catherine Amiens, Bruno Chaudret, Marie-Jose Casanove, Pierre Lecante, and John S Bradley. Synthesis and isolation of cuboctahedral and icosahedral platinum nanoparticles. ligand-dependent structures. *Chemistry of materials*, 8(8):1978–1986, 1996. (Cited on page 84.)

[167] Ruttala Devivaraprasad, Tathagata Kar, Arup Chakraborty, Ramesh Kumar Singh, and Manoj Neergat. Reconstruction and dissolution of shape-controlled pt nanoparticles in acidic electrolytes. *Phys. Chem. Chem. Phys.*, 18:11220–11232, 2016. doi: 10.1039/C5CP07832F. URL http://dx.doi.org/10.1039/C5CP07832F. (Not cited.)

[168] Tulio C. R. Rocha, Herbert Winnischofer, and Daniela Zanchet. *Structural Aspects of Anisotropic Metal Nanoparticle Growth: Experiment and Theory*. Wiley-VCH Verlag GmbH & Co. KGaA, 2012. ISBN 9783527652570. doi: 10.1002/9783527652570.ch6. URL http://dx.doi.org/10.1002/9783527652570.ch6. (Cited on page 84.)

[169] N. J. A. Sloane. Centered icosahedral (or cuboctahedral) numbers, also crystal ball sequence for f.c.c. lattice., October 2014. URL http://oeis.org/A005902. (Cited on page 85.)

[170] Joseph W Bennett. Discovery and design of functional materials: integration of database searching and first principles calculations. *Physics Procedia*, 34:14–23, 2012. (Cited on pages 87 and 89.)

[171] Stewart J Clark, Matthew D Segall, Chris J Pickard, Phil J Hasnip, Matt IJ Probert, Keith Refson, and Mike C Payne. First principles methods using castep. *Zeitschrift für Kristallographie*, 220(5/6/2005):567–570, 2005. (Cited on page 88.)

[172] Chris-Kriton Skylaris, Arash A. Mostofi, Peter D. Haynes, Oswaldo Diéguez, and Mike C. Payne. Nonorthogonal generalized wannier function pseudopotential plane-wave method. *Phys. Rev. B*, 66:035119, Jul 2002. doi: 10.1103/PhysRevB.66.035119. URL http://link.aps.org/doi/10.1103/PhysRevB.66.035119. (Cited on pages 89 and 152.)

[173] Chris-Kriton Skylaris and Peter D. Haynes. Achieving plane wave accuracy in linear-scaling density functional theory applied to periodic systems: A case study on crystalline silicon. *The Journal of Chemical Physics*, 127(16):164712, 2007. doi: http://dx.doi.org/10.1063/1.2796168. URL http://scitation.aip.org/content/aip/journal/jcp/127/16/10.1063/1.2796168. (Cited on page 89.)

[174] Kevin F. Garrity, Joseph W. Bennett, Karin M. Rabe, and David Vanderbilt. Pseudopotentials for high-throughput DFT calculations. *Computational Materials Science*, 81:446–452, 2014. ISSN 0927-0256. doi: 10.1016/j.commatsci.2013.08.053. URL http://www.sciencedirect.com/science/article/pii/S0927025613005077. (Cited on pages 89, 122, and 159.)

[175] Nicholas D. M. Hine, Mark Robinson, Peter D. Haynes, Chris-Kriton Skylaris, Mike C. Payne, and Arash A. Mostofi. Accurate ionic forces and geometry optimization in linear-scaling density-functional theory with local orbitals. *Phys. Rev. B*, 83:195102, May 2011. doi: 10.1103/PhysRevB.83.195102. URL http://link.aps.org/doi/10.1103/PhysRevB.83.195102. (Cited on page 89.)

[176] M W Finnis. A faster way to relax interfaces in supercells. *Journal of Physics: Condensed Matter*, 16(27):S2671, 2004. URL http://stacks.iop.org/0953-8984/16/i=27/a=006. (Cited on page 91.)

[177] Kotaro Sasaki, Nebojsa Marinkovic, Hugh S. Isaacs, and Radoslav R. Adzic. Synchrotron-based in situ characterization of carbon-supported platinum and platinum monolayer electrocatalysts. *ACS Catalysis*, 6(1):69–76, 2016. doi: 10.1021/acscatal.5b01862. URL http://dx.doi.org/10.1021/acscatal.5b01862. (Cited on page 106.)

[178] Joachim Schnadt, Jan Knudsen, Xiao Liang Hu, Angelos Michaelides, Ronnie T Vang, Karsten Reuter, Zheshen Li, Erik Lægsgaard, Matthias Scheffler, and Flemming Besenbacher. Experimental and theoretical study of oxygen adsorption structures on ag (111). *Physical Review B*, 80(7):075424, 2009. (Cited on page 106.)

[179] Zhihui Gu and Perla B. Balbuena. Absorption of atomic oxygen into subsurfaces of pt(100) and pt(111): Density functional theory study. *The Journal of Physical Chemistry C*, 111(27):9877–9883, 2007. doi: 10.1021/jp0711693. URL http://dx.doi.org/10.1021/jp0711693. (Cited on page 106.)

[180] Stephen W. T. Price, Nicholas Zonias, Chris Kriton Skylaris, Timothy I. Hyde, Bruce Ravel, and Andrea E. Russell. Fitting EXAFS data using molecular dynamics outputs and a histogram approach. *Physical Review B - Condensed Matter and Materials Physics*, 85(7):075439, feb 2012. ISSN 10980121. doi: 10.1103/PhysRevB.85.075439. (Cited on page 112.)

[181] Nozomu Ishiguro, Takahiro Saida, Tomoya Uruga, Shin-ichi Nagamatsu, Oki Sekizawa, Kiyofumi Nitta, Takashi Yamamoto, Shin-ichi Ohkoshi, Yasuhiro Iwasawa, Toshihiko Yokoyama, and Mizuki Tada. Operando Time-Resolved X-ray Absorption Fine Structure Study for Surface Events on a Pt 3 Co/C Cathode Catalyst in a Polymer Electrolyte Fuel Cell during Voltage-Operating Processes. *ACS Catalysis*, 2(7):1319–1330, jul 2012. ISSN 2155-5435. doi: 10.1021/cs300228p. (Cited on page 112.)

[182] Shinobu Takao, Oki Sekizawa, Shin-ichi Nagamatsu, Takuma Kaneko, Takashi Yamamoto, Gabor Samjeské, Kotaro Higashi, Kensaku Nagasawa, Takuya Tsuji, Motohiro Suzuki, Naomi Kawamura, Masaichiro Mizumaki, Tomoya Uruga, and Yasuhiro Iwasawa. Mapping Platinum Species in Polymer Electrolyte Fuel Cells by Spatially Resolved XAFS Techniques. *Angewandte Chemie International Edition*, 53(51):14110–14114, dec 2014. ISSN 14337851. doi: 10.1002/anie.201408845. (Cited on page 112.)

[183] P. D. Nellist, S. J. Pennycook, L. Spenadel, M. Boudart, S. F. Adler, J. J. Keavney, D. C. Koningsberger, B. C. Gates, A. S. Fung, P. A. Tooley, M. J. Kelley, D. C. Koningsberger, B. C. Gates, D. E. Jesson, S. J. Pennycook, M. S. Isaacson, D. Kopf, M. Ohtsuki, M. Utlaut, M. S. Isaacson, M. Ohtsuki, M. Utlaut, S. Iijima, M. M. J. Treacy, S. B. Rice, J. M. Cowley, E. Zeitler, M. G. R. Thomson, A. J. McGibbon, S. J. Pennycook, D. E. Jesson, D. C. Koningsberger, and M. Vaarkamp. Direct Imaging of the Atomic Configuration of Ultradispersed Catalysts. *Science*, 274 (5286):413–415, oct 1996. ISSN 0036-8075. doi: 10.1126/science.274.5286.413. (Cited on page 113.)

[184] Kees Joost Batenburg and Jan Sijbers. DART: A Practical Reconstruction Algorithm for Discrete Tomography. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 20(9):2542–2553, sep 2011. ISSN 1941-0042. doi: 10.1109/TIP.2011.2131661. (Cited on page 113.)

[185] Bart Goris, T Roelandts, Kees Joost Batenburg, H Heidari Mezerji, and Sara Bals. Advanced reconstruction algorithms for electron tomography: from comparison to combination. *Ultramicroscopy*, 127:40–7, apr 2013. ISSN 1879-2723. doi: 10.1016/j.ultramic.2012.07.003. (Not cited.)

[186] Sandra Van Aert, Kees Joost Batenburg, Marta D Rossell, Rolf Erni, and Gustaaf Van Tendeloo. Three-Dimensional Atomic Imaging of Crystalline Nanoparticles. *Nature*, 470(7334):374–377, feb 2011. ISSN 1476-4687. doi: 10.1038/nature09741. (Not cited.)

[187] Sara Bals, Marianna Casavola, Marijn A van Huis, Sandra Van Aert, Kees Joost Batenburg, Gustaaf Van Tendeloo, and Daniël Vanmaekelbergh. Three-Dimensional Atomic Imaging of Colloidal Core-Shell Nanocrystals. *Nano letters*, 11(8):3420–3424, aug 2011. ISSN 1530-6992. doi: 10.1021/nl201826e. (Cited on page 113.)

[188] James M. LeBeau, S.D. Findlay, L. J. Allen, and S. Stemmer. Quantitative Atomic Resolution Scanning Transmission Electron Microscopy. *Physical Review Letters*, 100(20):206101–, may 2008. ISSN 0031-9007. (Cited on page 113.)

[189] James M. LeBeau, Scott D. Findlay, Leslie J. Allen, and Susanne Stemmer. Standardless atom counting in scanning transmission electron microscopy. *Nano letters*, 10(11):4405–8, nov 2010. ISSN 1530-6992. doi: 10.1021/nl102025s. (Cited on page 113.)

[190] Haibo E, Katherine E MacArthur, Timothy J Pennycook, Eiji Okunishi, A J D'Alfonso, N. R. Lugg, Leslie J. Allen, and Peter D. Nellist. Probe integrated scattering cross sections in the analysis of atomic resolution HAADF

STEM images. *Ultramicroscopy*, 133:109–19, oct 2013. ISSN 1879-2723. doi: 10.1016/j.ultramic.2013.07.002. (Cited on pages 113 and 119.)

[191] Gerardo T. Martinez, Annick De Backer, Andreas Rosenauer, Jo Verbeeck, and Sandra Van Aert. The effect of probe inaccuracies on the quantitative model-based analysis of high angle annular dark field scanning transmission electron microscopy images. *Micron*, 63:57–63, dec 2014. ISSN 09684328. doi: 10.1016/j.micron.2013.12.009. (Not cited.)

[192] K.E. MacArthur, A.J. D'Alfonso, D. Ozkaya, L.J. Allen, and P.D. Nellist. Optimal ADF STEM imaging parameters for tilt-robust image quantification. *Ultramicroscopy*, 156:1–8, sep 2015. ISSN 03043991. doi: 10.1016/j.ultramic.2015.04.010. (Cited on page 113.)

[193] A J den Dekker, Sandra Van Aert, A van den Bos, and D Van Dyck. Maximum likelihood estimation of structure parameters from high resolution electron microscopy images. Part I: a theoretical framework. *Ultramicroscopy*, 104(2):83–106, sep 2005. ISSN 0304-3991. doi: 10.1016/j.ultramic.2005.03.001. (Cited on page 113.)

[194] Sandra Van Aert, Jo Verbeeck, Rolf Erni, Sara Bals, M Luysberg, D Van Dyck, and Gustaaf Van Tendeloo. Quantitative atomic resolution mapping using high-angle annular dark field scanning transmission electron microscopy. *Ultramicroscopy*, 109 (10):1236–44, sep 2009. ISSN 1879-2723. doi: 10.1016/j.ultramic.2009.05.010. (Not cited.)

[195] Annick De Backer, Gerardo T. Martinez, Andreas Rosenauer, and Sandra Van Aert. Atom counting in HAADF STEM using a statistical model-based approach: methodology, possibilities, and inherent limitations. *Ultramicroscopy*, 134:23–33, nov 2013. ISSN 1879-2723. doi: 10.1016/j.ultramic.2013.05.003. (Cited on page 113.)

[196] M. Mavrikakis, B. Hammer, and J. K. Nørskov. Effect of Strain on the Reactivity of Metal Surfaces. *Physical Review Letters*, 81(13):2819–2822, sep 1998. ISSN 0031-9007. doi: 10.1103/PhysRevLett.81.2819. (Cited on page 113.)

[197] J. K. Nørskov, Thomas Bligaard, Jan Rossmeisl, and Claus H. Christensen. Towards the computational design of solid catalysts. *Nature Chemistry*, 1(1):37–46, apr 2009. ISSN 1755-4330. doi: 10.1038/nchem.121. (Cited on pages 128 and 163.)

[198] Ifan E. L. Stephens, Alexander S. Bondarenko, Francisco J. Perez-Alonso, Federico Calle-Vallejo, Lone Bech, Tobias P. Johansson, Anders K. Jepsen, Rasmus Frydendal, Brian P. Knudsen, Jan Rossmeisl, and Ib Chorkendorff. Tuning the Activity of Pt(111) for Oxygen Electroreduction by Subsurface Alloying. *Journal of the*

American Chemical Society, 133(14):5485–5491, apr 2011. ISSN 0002-7863. doi: 10.1021/ja111690g. (Not cited.)

[199] Francesc Viñes, José R. B. Gomes, and Francesc Illas. Understanding the reactivity of metallic nanoparticles: beyond the extended surface model for catalysis. *Chemical Society Reviews*, 43(14):4922, 2014. ISSN 0306-0012. doi: 10.1039/c3cs60421g. (Cited on page 113.)

[200] B. C. Han, C. R. Miranda, and G. Ceder. Effect of particle size and surface structure on adsorption of O and OH on platinum nanoparticles: A first-principles study. *Physical Review B - Condensed Matter and Materials Physics*, 77(7):075410, feb 2008. ISSN 10980121. doi: 10.1103/PhysRevB.77.075410. (Cited on pages 114 and 130.)

[201] Minhua Shao, Amra Peles, and Krista Shoemaker. Electrocatalysis on Platinum Nanoparticles: Particle Size Effect on Oxygen Reduction Reaction Activity. *Nano Letters*, 11(9):3714–3719, sep 2011. ISSN 1530-6984. doi: 10.1021/nl2017459. (Cited on pages 114, 132, 134, and 135.)

[202] Sergey Dobrin. CO oxidation on Pt nanoclusters, size and coverage effects: a density functional theory study. *Physical Chemistry Chemical Physics*, 14(35): 12122, 2012. ISSN 1463-9076. doi: 10.1039/c2cp41286a. (Cited on page 114.)

[203] L. Jones. Quantitative ADF STEM: acquisition, analysis and interpretation. *IOP Conference Series: Materials Science and Engineering*, 109(1):012008, feb 2016. ISSN 1757-8981. doi: 10.1088/1757-899X/109/1/012008. (Cited on page 116.)

[204] A. De Backer, G.T. Martinez, K.E. MacArthur, Lewys Jones, A. Béché, P.D. Nellist, S. Van Aert, a. De Backer, and S. Van Aert. Dose limited reliability of quantitaive ADF STEM for nano-particle atom-counting. *Ultramicroscopy*, 151: 56–61, dec 2015. ISSN 03043991. doi: 10.1016/j.ultramic.2014.11.028. (Cited on pages 116 and 119.)

[205] A. De Backer, A. De wael, J. Gonnissen, and S. Van Aert. Optimal experimental design for nano-particle atom-counting from high-resolution STEM images. *Ultramicroscopy*, 151:46–55, nov 2015. ISSN 03043991. doi: 10.1016/j.ultramic.2014.10.015. (Cited on page 116.)

[206] Xiahan Sang and James M. LeBeau. Revolving scanning transmission electron microscopy: Correcting sample drift distortion without prior knowledge. *Ultramicroscopy*, 138C:28–35, dec 2013. ISSN 1879-2723. doi: 10.1016/j.ultramic.2013.12.004. (Not cited.)

[207] Lewys Jones, Hao Yang, Timothy J. Pennycook, Matthew S. J. Marshall, Sandra Van Aert, Nigel D. Browning, Martin R. Castell, and Peter D. Nellist. Smart Align—a new tool for robust non-rigid registration of scanning microscope data. *Advanced Structural and Chemical Imaging*, 1(1):8, dec 2015. ISSN 2198-0926. doi: 10.1186/s40679-015-0008-4. (Cited on pages 116 and 119.)

[208] C. Solliard and M. Flueli. Surface stress and size effect on the lattice parameter in small particles of gold and platinum. *Surface Science*, 156(PART 1):487–494, 1985. ISSN 00396028. doi: 10.1016/0039-6028(85)90610-7. (Cited on page 118.)

[209] G.T. Martinez, Lewys Jones, A. De Backer, A. Béché, J. Verbeeck, S. Van Aert, and P.D. Nellist. Quantitative STEM normalisation: The importance of the electron flux. *Ultramicroscopy*, 159:46–58, dec 2015. ISSN 03043991. doi: 10.1016/j.ultramic.2015.07.010. (Cited on page 119.)

[210] B. D. Forbes, A. V. Martin, Scott D. Findlay, Adrian J. D'Alfonso, and Leslie J. Allen. Quantum mechanical model for phonon excitation in electron diffraction and imaging using a Born-Oppenheimer approximation. *Physical Review B*, 82 (10):104103, sep 2010. ISSN 1098-0121. doi: 10.1103/PhysRevB.82.104103. (Cited on page 119.)

[211] H. X. Gao and L.-M. Peng. Parameterization of the Temperature Dependence of the Debye–Waller Factors. *Acta Crystallographica Section A Foundations of Crystallography*, 55(5):926–932, sep 1999. ISSN 0108-7673. doi: 10.1107/S0108767399005176. (Cited on page 119.)

[212] Min Yu, Andrew B. Yankovich, Amy Kaczmarowski, Dane Morgan, and Paul M. Voyles. Integrated Computational and Experimental Structure Refinement for Nanoparticles. *ACS Nano*, 10(4):4031–4038, apr 2016. ISSN 1936-0851. doi: 10.1021/acsnano.5b05722. (Cited on page 119.)

[213] A. P. Sutton and J. Chen. Long-range Finnis–Sinclair potentials. *Philosophical Magazine Letters*, 61(3):139–146, 1990. ISSN 0950-0839. doi: 10.1080/09500839008206493. URL http://www.tandfonline.com/doi/abs/10.1080/09500839008206493. (Cited on page 121.)

[214] Ilian T Todorov, William Smith, Kostya Trachenko, and Martin T Dove. Dl_poly_3: new dimensions in molecular dynamics simulations via massive parallelism. *Journal of Materials Chemistry*, 16(20):1911–1918, 2006. (Cited on page 121.)

[215] R. Smoluchowski. Anisotropy of the electronic work function of metals. *Physical Review*, 60(9):661–674, 1941. ISSN 0031899X. doi: 10.1103/PhysRev.60.661. (Cited on page 123.)

[216] M Vladimirova, M Stengel, a. De Vita, a Baldereschi, M Böhringer, K Morgenstern, R Berndt, and W.-D Schneider. Supramolecular self-assembly and selective step decoration on the Au(111) surface. *Europhysics Letters (EPL)*, 56(2):254–260, 2007. ISSN 0295-5075. doi: 10.1209/epl/i2001-00514-3. (Cited on page 123.)

[217] Harald Ibach. Vibration spectroscopy of water on stepped gold surfaces. *Surface Science*, 604(3-4):377–385, 2010. ISSN 00396028. doi: 10.1016/j.susc.2009.11.034. (Not cited.)

[218] Patrick Han and Paul S. Weiss. Electronic substrate-mediated interactions. *Surface Science Reports*, 67(2):19–81, 2012. ISSN 01675729. doi: 10.1016/j.surfrep.2011.11.001. (Cited on page 123.)

[219] Federico Calle-Vallejo, Jakub Tymoczko, Viktor Colic, Quang Huy Vu, Marcus D. Pohl, Karina Morgenstern, David Loffreda, Philippe Sautet, Wolfgang Schuhmann, and Aliaksandr S. Bandarenka. Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science*, 350(6257):185–189, oct 2015. ISSN 0036-8075. doi: 10.1126/science.aab3501. (Cited on pages 125 and 130.)

[220] Jacob Gavartin, Misbah Sarwar, Dimitrios Papageorgopoulos, David Gunn, Sonia Garcia, Alexander Perlov, Arek Krzystala, Dan L. Ormsby, David Thompsett, Gerhard Goldbeck-Wood, Amity Andersen, and Sam French. Exploring Fuel Cell Cathode Materials: A High Throughput Calculation Approach. *ECS Transactions*, 25(1):1335–1344, 2009. doi: 10.1149/1.3210689. (Cited on pages 125 and 135.)

[221] Federico Calle-Vallejo, David Loffreda, Marc T. M. Koper, and Philippe Sautet. Introducing structural sensitivity into adsorption–energy scaling relations by means of coordination numbers. *Nature Chemistry*, 7(5):403–410, apr 2015. ISSN 1755-4330. doi: 10.1038/nchem.2226. (Cited on pages 126, 130, and 135.)

[222] F. H B Lima, J. Zhang, M. H. Shao, K. Sasaki, M. B. Vukmirovic, E. A. Ticianelli, and R. R. Adzic. Catalytic Activity-d-Band Center Correlation for the O 2 Reduction Reaction on Platinum in Alkaline Solutions. *The Journal of Physical Chemistry C*, 111(1):404–410, jan 2007. ISSN 1932-7447. doi: 10.1021/jp065181r. (Cited on pages 130 and 135.)

[223] Federico Calle-Vallejo, Philippe Sautet, and David Loffreda. Understanding Adsorption-Induced Effects on Platinum Nanoparticles: An Energy-Decomposition

Analysis. *The Journal of Physical Chemistry Letters*, 5(18):3120–3124, sep 2014. ISSN 1948-7185. doi: 10.1021/jz501263e. (Cited on page 135.)

[224] Jolyon Aarons and Chris-Kriton Skylaris. Electronic annealing fermi operator expansion for dft calculations on metallic systems. *The Journal of Chemical Physics*, Submitted 23-Aug-2017. (Cited on page 137.)

[225] Phanish Suryanarayana. On spectral quadrature for linear-scaling density functional theory. *Chemical Physics Letters*, 584:182 – 187, 2013. ISSN 0009-2614. doi: http://dx.doi.org/10.1016/j.cplett.2013.08.035. URL http://www.sciencedirect.com/science/article/pii/S0009261413010452. (Cited on page 138.)

[226] Anders MN Niklasson, Marc J Cawkwell, Emanuel H Rubensson, and Elias Rudberg. Canonical density matrix perturbation theory. *Physical Review E*, 92(6): 063301, 2015. (Cited on pages 138 and 147.)

[227] Nick DM Hine, Peter D Haynes, Arash A Mostofi, C-K Skylaris, and Mike C Payne. Linear-scaling density-functional theory with tens of thousands of atoms: Expanding the scope and scale of calculations with onetep. *Computer Physics Communications*, 180(7):1041–1053, 2009. (Cited on page 139.)

[228] NDM Hine, PD Haynes, AA Mostofi, and MC Payne. Linear-scaling density-functional simulations of charged point defects in alo using hierarchical sparse matrix algebra. *The Journal of chemical physics*, 133:114111, 2010. (Not cited.)

[229] Karl A. Wilkinson, Nicholas D. M. Hine, and Chris-Kriton Skylaris. Hybrid MPI-OpenMP Parallelism in the ONETEP Linear-Scaling Electronic Structure Code: Application to the Delamination of Cellulose Nanofibrils. *Journal of Chemical Theory and Computation*, 10(11):4782–4794, 2014. (Cited on page 139.)

[230] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965. (Cited on page 139.)

[231] Jie Chen and Edmond Chow. A stable scaling of newton-schulz for improving the sign function computation of a hermitian matrix. 2014. (Cited on page 142.)

[232] Nicholas J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, 2008. ISBN 9780898717778. ISBN 978-0-898716-46-7. (Cited on page 143.)

[233] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, 2nd edition, 2006. (Cited on page 143.)

[234] Petros Souvatzis and Anders M. N. Niklasson. Extended lagrangian born-oppenheimer molecular dynamics in the limit of vanishing self-consistent field optimization. *The Journal of Chemical Physics*, 139(21):214102, 2013. doi: 10.1063/1.4834015. URL http://dx.doi.org/10.1063/1.4834015. (Cited on page 147.)

[235] Charles Kenney and Alan J. Laub. Condition estimates for matrix functions. *SIAM Journal on Matrix Analysis and Applications*, 10(2):191–209, 1989. doi: 10.1137/0610014. URL http://dx.doi.org/10.1137/0610014. (Cited on page 151.)

[236] Awad H. Al-Mohy and Nicholas J. Higham. Improved inverse scaling and squaring algorithms for the matrix logarithm. *SIAM Journal on Scientific Computing*, 34(4):C153–C169, 2012. doi: 10.1137/110852553. URL http://dx.doi.org/10.1137/110852553. (Cited on page 151.)

[237] Sheung Hun Cheng, Nicholas J Higham, Charles S Kenney, and Alan J Laub. Approximating the logarithm of a matrix to specified accuracy. *SIAM Journal on Matrix Analysis and Applications*, 22(4):1112–1125, 2001. (Cited on page 151.)

[238] Nicholas Hale, Nicholas J. Higham, and Lloyd N. Trefethen. Computing aα, log(a), and related matrix functions by contour integrals. *SIAM Journal on Numerical Analysis*, 46(5):2505–2523, 2008. ISSN 0036-1429. doi: 10.1137/070700607. (Cited on page 151.)

[239] Arash A Mostofi, Peter D Haynes, Chris-Kriton Skylaris, and Mike C Payne. Preconditioned iterative minimization for linear-scaling electronic structure calculations. *The Journal of chemical physics*, 119(17):8842–8848, 2003. (Cited on page 152.)

[240] Andrew Gibson, Roger Haydock, and John P LaFemina. Ab initio electronic-structure computations with the recursion method. *Physical Review B*, 47(15):9229, 1993. (Cited on page 154.)

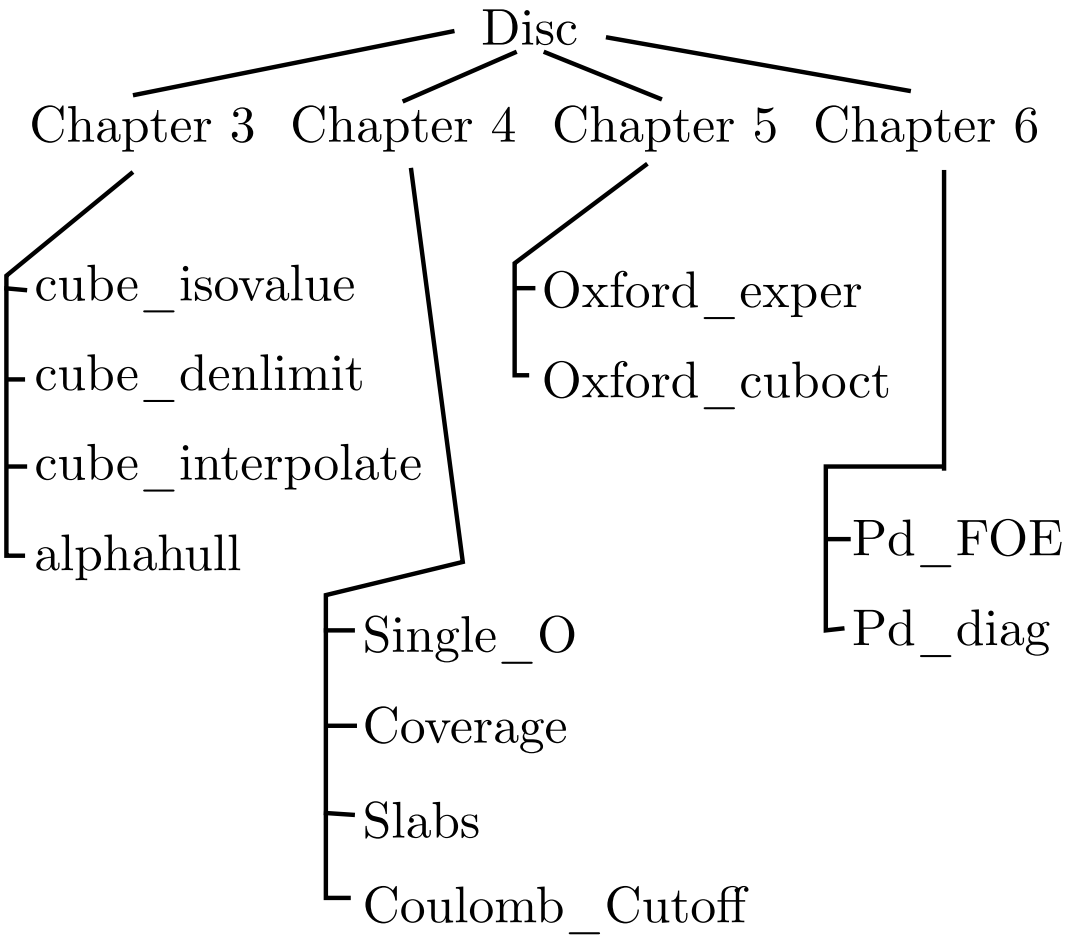
[241] Peter J Feibelman, Bjørk Hammer, Jens Kehlet Nørskov, F Wagner, M Scheffler, R Stumpf, R Watwe, and J Dumesic. The co/pt (111) puzzle. *The Journal of Physical Chemistry B*, 105(18):4018–4025, 2001. (Cited on page 181.)

[242] Ping Liu, Ashildur Logadottir, and Jens Kehlet Nørskov. Modeling the electro-oxidation of co and h< sub> 2</sub>/co on pt, ru, ptru and pt< sub> 3</sub> sn. *Electrochimica Acta*, 48(25):3731–3742, 2003. (Cited on page 181.)

[243] G Kresse, A Gil, and Ph Sautet. Significance of single-electron energies for the description of co on pt (111). *Physical Review B*, 68(7):073401, 2003. (Cited on page 181.)

[244] Fabrizio Cleri and Vittorio Rosato. Tight-binding potentials for transition metals and alloys. *Phys. Rev. B*, 48:22–33, Jul 1993. doi: 10.1103/PhysRevB.48.22. URL http://link.aps.org/doi/10.1103/PhysRevB.48.22. (Cited on page 187.)

[245] J.D. Cook. Centered icosahedral (or cuboctahedral) numbers, also crystal ball sequence for f.c.c. lattice., January 2010. URL https://www.johndcook.com/blog/2010/01/19/dont-invert-that-matrix/. accessed 17-Sep-2017. (Cited on page 191.)


Disc
Chapter 3
Chapter 4
Chapter 5
Chapter 6
cube_isovalue
cube_denlimit
cube_interpolate
alphahull
Oxford_exper
Oxford_cuboct
Pd_FOE
Pd_diag
Single_O
Coverage
Slabs
Coulomb_Cutoff