New Applications of Quantum Monte Carlo

Nicholas Hine

Thesis submitted for the degree of

Doctor of Philosophy of the University of London

and the

Diploma of Imperial College

Condensed Matter Theory Group

The Blackett Laboratory

Imperial College

London

December 2007

Abstract

This thesis comprises a trio of separate yet thematically-linked investigations within the field of electronic structure that address challenging problems to which quantum Monte Carlo, the most accurate electronic structure method capable of treating large extended systems, has been applied only rarely or indeed never before.

The first is the study of polarization and localization in many body systems. Density functional and variational quantum Monte Carlo methods are used to examine the behaviour of the many-electron localization length near band insulator to metal transitions in various one- and two-dimensional model systems. We examine when the localization length, known to be infinite in metals and finite in insulators and assumed to diverge at a metal-insulator transition, serves as a useful diagnostic of the transition. We also present a novel derivation of the formulae for polarization and localization in many body systems that expresses the connection of these two quantities to the concept of Many Body Wannier Functions.

The second is a study of the surface energy of the electron gas. The electron gas, or 'jellium', is an important benchmark system for electronic structure methods and serves as a simple model for metals. We apply diffusion Monte Carlo (DMC), and a new method for determining surface energies without an unreliable comparison of bulk and slab results, to the calculation of the surface energy. Our method improves on previous calculations in a number of ways and removes many sources of error that plagued previous calculations.

The final study applies DMC to the properties, specifically the formation energy, of point defects in Alumina. Alumina is an important technological material, and there are many important unanswered questions relating to its properties that stem from the behaviour and concentration of its point defects. We address many of the problems associated with the calculation of formation energies with electronic structure calculations, and present the first calculations of formation energies of charged point defects in DMC.

Acknowledgements

I am hugely indebted to a large number of people who have made the research represented in this thesis not only possible but also a great deal of fun. The Condensed Matter Theory Group at Imperial has been an awesome place to work and I would like to thank all my colleagues here who have contributed a great deal to keeping me happy and intellectually stimulated over the last three years. My colleagues in quantum Monte Carlo around the world, but particularly of the TCM group in Cambridge, have also been a source of enormous help and valuable discussions.

My supervisor Matthew Foulkes has been a source of unending enthusiasm, support, patience, and wisdom, and I will always be grateful for all that I have learnt from him.

Finally, I would like to thank my parents, all my friends and flatmates in London, and in particular Hattie, for supporting me and putting up with me over the last few years.

Contents

Chapter 1

Introduction

In recent decades, computational simulation has come to play an ever increasing role in advancing our understanding of physical processes. Bridging the gap as it does between pure theory and experiment, there is a wide spectrum of approaches one can take depending on the tractability of the problem at hand to analytic solution: the computational problem left once one has described a physical system as well as one can analytically might be anywhere from brute-force number crunching of the simplest underlying rules that describe a highly realistic system, to the realization of highly-abstracted rules that describe a model system. Often, the choice depends on the level of approximation one is willing to put up with, and the level of quantitative accuracy required of the results.

Computational electronic structure is a well-established field with a wide range of available theories that span almost this entire spectrum: from models where useful results can be extracted at the pen-and-paper level such as classical point charges and empirical tight-binding, through pair-potential methods and fully quantum-mechanical yet approximate theories including Hartree Fock (HF) and Density Functional Theory (DFT), to quantum chemistry methods capable of any level of precision one requires, at the cost of very great computational effort. Sitting very much at the 'accurate' end of that range is the set of methods known as quantum Monte Carlo, which, while not quite able in practice to reach the absolute heights of accuracy obtainable with quantum chemistry, turns out to be applicable to a large range of situations which are entirely out of reach of those techniques. The principal advantage of quantum Monte Carlo depends on what it is compared to: measured relative to post-Hartree-Fock quantum chemistry, it is its favourable scaling with system size that renders it feasible in simulating extended systems, whereas compared to Density Functional Theory, it is the greatly increased accuracy of its results, the result of its relative absence of arbitrary approximations. In many cases, most notably any sort of periodically-repeating solid crystal, it is the only technique capable of combining anything near the level referred to as "chemical" accuracy, usually thought of as 1 kcal/mol, or 0.043 eV/atom, with feasible computing time, favourable scaling with system size, and size-consistency[*].

The fundamental means of description of electrons in condensed matter is their wavefunction. DFT replaces this, in ways we will address in Chapter 2, with the electron density. DFT is, as will be shown, capable of being exact in principle for energies, and indeed for other operators as the Hohenberg-Kohn theorems show, but relies in practice on being able to write down in each case an unknown functional of the density. Many of the most significant breakthroughs in modern physics have come from an understanding of the characteristics of the wavefunctions of the elementary excitations of the system: superconductivity and superfluidity, the quantum hall effect and many others. It should not, therefore, come as a surprise that advances in the treatment of wavefunctions should also be capable of providing us with a more accurate way of calculating the properties of condensed matter required of electronic structure calculations.

In spite of the apparently well-founded belief that QMC is capable of providing answers of greater accuracy and perhaps, in view of its more fundamental level of description, greater insight, in reality this has not been widely demonstrated. There exists a need for the QMC community to demonstrate the use of QMC in new cir-

---

[*] By size consistency, we mean that for a periodic system, a sample of twice the size will have twice the total energy: this is not always exactly the case in QMC, which does suffer finite size effects, but they are in general quirks of the implementation rather than fundamental failings of the method.

cumstances to add weight to argument that it gives accurate results. Additionally, given the enormously widespread usage of DFT in physics and materials science, it is of great value to determine under what circumstances existing calculations are correct and when they are lacking, and to provide benchmark results to aid the development of improvements to the technical aspects of DFT calculations.

In this thesis, we present three such applications. To set the scene, Chapter 2 begins by introducing the many-body problem and its one-electron simplifications, and Chapter 3 describes the types of quantum Monte Carlo used in this research — variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC). The theory behind both algorithms is discussed, as well as some of the details of their implementation.

In Chapter 4 we discuss the mathematics behind quantities relating to polarization and localization in many-body systems. We present a new derivation of the many-body version of the 'Modern Theory' of Polarization and Localization which highlights the connection to the concept of many-body Wannier functions introduced by Souza, Wilkens and Martin [156]. We also present the results of a variety of calculations of such in model systems in DFT and VMC.

In Chapter 5, we discuss the calculation of the surface energy of the electron gas, an important system both as a benchmark of other methods, and as a demonstration of the accuracy and relative absence of approximation of QMC. Our results improve greatly on previous attempts to calculate surface energies within QMC by systematically identifying and minimizing the many sources of error that plague such calculations.

In Chapter 6 we present DMC calculations of the formation energy of point defects in Alumina. This is important not only as a proof of concept of DMC in large systems, but as a tool to address a wide range of uncertainties surrounding calculation of such quantities in DFT, such as bandgap-correction issues, DFT overbinding, and many others.

While calculations in QMC remain technically challenging and computationally demanding at present, they are nevertheless vital for providing a test and benchmark of the accuracy of other methods, principally DFT, and for resolving questions which DFT is unable to treat with sufficient accuracy. With time, increased computing power and the ever increasing maturity of available software codes, QMC may yet develop into a mainstay of the electronic structure community.

Chapter 2

Many-Electron Quantum Mechanics

2.1 The Schrödinger Equation and Many-body Wavefunctions

The central result of quantum mechanics is that the behaviour of all physical systems is the result of the properties of the wavefunction Ψ of the particles involved, which is governed by the Schrödinger equation. The defining element of the Schrödinger equation is the Hamiltonian $\hat{H}$ of the system, which incorporates all of the interactions between elements of the system and all the contributions to its energy. In a large system with a large number N of interacting particles, this becomes a fiendishly complex object, with of order N^2 or more terms which strongly couple the $3N$ spatial coordinates $\{\mathbf{r}_i\}$ and N spin coordinates $\{\sigma_i\}$, rendering the equation highly non-separable. Though they can be treated within the frameworks we are about to discuss, we will neglect for the purposes of this thesis the effect of magnetic fields and any direct contribution to the Hamiltonian of the spins of the electrons, as the only effect of spin in these applications is indirect, via the restrictions on the wavefunction imposed by exchange and antisymmetry. We will also discard relativity at this point, even though in heavier atoms it can have a significant effect on core electrons, as it will not be relevant to any of the studies presented here and makes the results considerably less neat.

The Schrödinger equation for a collection of N_e electrons at positions $\{\mathbf{r}_i\}$ with spins $\{\sigma_i\}$ interacting via the Coulomb interaction with N_n point nuclei of atomic number $\{Z_I\}$ and masses $\{M_I\}$ at positions $\{\mathbf{R}_I\}$ is

$$\hat{H}\Psi(\{\mathbf{r}_i,\sigma_i\},\{\mathbf{R}_I\},t) = i\hbar\frac{\partial\Psi}{\partial t}\ , \tag{2.1}$$

where

$$\begin{aligned}\hat{H} = &-\frac{\hbar^2}{2m_e}\sum_i^{N_e}\frac{\partial^2\Psi}{\partial\mathbf{r}_i^2} - \frac{\hbar^2}{2}\sum_I^{N_n}\frac{1}{M_I}\frac{\partial^2\Psi}{\partial\mathbf{R}_I^2} \\ &-\sum_i^{N_e}\sum_I^{N_n}\frac{Z_I e^2}{4\pi\epsilon_0|\mathbf{r}_i-\mathbf{R}_I|} + \frac{1}{2}\sum_I^{N_n}\sum_{J\neq I}^{N_n}\frac{Z_I Z_J e^2}{4\pi\epsilon_0|\mathbf{R}_I-\mathbf{R}_J|} \\ &+\frac{1}{2}\sum_i^{N_e}\sum_{j\neq i}^{N_e}\frac{e^2}{4\pi\epsilon_0|\mathbf{r}_i-\mathbf{r}_j|}\ .\end{aligned} \tag{2.2}$$

We will adopt the standard practice of using the Hartree system of atomic units throughout this thesis unless otherwise stated. In this system, length, time and mass are redefined such that $\hbar = e = m_e = 4\pi\epsilon_0 = 1$, in order both to simplify the equations and to render all the terms numerically of order ~ 1 to aid computational accuracy. Eq. 2.2 then becomes

$$\begin{aligned}\hat{H} = &-\frac{1}{2}\sum_i^{N_e}\frac{\partial^2\Psi}{\partial\mathbf{r}_i^2} - \frac{1}{2}\sum_I^{N_n}\frac{1}{M_I}\frac{\partial^2\Psi}{\partial\mathbf{R}_I^2} - \sum_i^{N_e}\sum_I^{N_n}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|} \\ &+\frac{1}{2}\sum_I^{N_n}\sum_{J\neq I}^{N_n}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|} + \frac{1}{2}\sum_i^{N_e}\sum_{j\neq i}^{N_e}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\end{aligned} \tag{2.3}$$

The first step towards solving this is to make what is known as the Born-Oppenheimer approximation. On the grounds that the timescale associated with the motion of whole atomic nuclei is expected to be much longer than that associated with the movement of individual electrons, when looking for stationary states of the Hamiltonian of energy E we make the following separation of variables:

$$\Psi(\{\mathbf{r}_i,\sigma_i\},\{\mathbf{R}_I\},t)=\psi(\{\mathbf{r}_i,\sigma_i\};\{\mathbf{R}_I\})\chi(\{\mathbf{R}_I\})\mathrm{e}^{-iEt}\ . \tag{2.4}$$

That is to say, the nuclear coordinates $\{\mathbf{R}_I\}$ enter into the electronic wavefunctions ψ solely as parameters, with respect to which ψ is smoothly varying. The Schrödinger equation then becomes

$$\left(-\frac{1}{2}\sum_i\nabla^2_{\mathbf{r}_i}+\frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}-\sum_{i,I}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|}\right)\psi_\alpha=E_\alpha(\{\mathbf{R}_I\})\psi_\alpha \tag{2.5}$$

for the electrons, and

$$\left(-\frac{1}{2}\sum_I\frac{1}{M_I}\nabla^2_{\mathbf{R}_I}+\frac{1}{2}\sum_{I\neq J}\frac{Z_IZ_J}{|\mathbf{R}_I-\mathbf{R}_J|}+E_\alpha(\{\mathbf{R}_I\})\right)\chi_{\alpha\beta}=E_{\alpha\beta}\chi_{\alpha\beta} \tag{2.6}$$

for the ions.

It can be seen that Eq. 2.4 is not actually an exact eigenstate of $\hat{H}$ by noting that matrix elements of the form $\langle\Psi_{\alpha'\beta'}|\hat{H}|\Psi_{\alpha\beta}\rangle$ are not simply $E_{\alpha\beta}\delta_{\alpha\alpha'}\delta_{\beta\beta'}$ but contain terms including $\dfrac{1}{M_I}\nabla^2_{\mathbf{R}_I}\psi_\alpha$. However, the large mass ratio m_e/m_p keeps these terms small enough that this approximation holds very well, especially for heavier ions. Standard practice at this point is to consider the ions as fixed, classical point charges, to discard Eq. 2.6 and consider only the electronic problem quantum-mechanically, with the ion-ion interaction energy added as a constant value E_{II}. We return to referring to the N-electron wavefunction as Ψ and we are now solving the marginally more tractable

$$\left(-\frac{1}{2}\sum_i\nabla^2_{\mathbf{r}_i}+\frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}-\sum_{i,I}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|}+E_{II}\right)\Psi(\{\mathbf{r}_i\})=E\Psi(\{\mathbf{r}_i\})\ , \tag{2.7}$$

and treating the ion coordinates $\{\mathbf{R}_I\}$ as fixed parameters.

In the rest of this chapter, we will make a series of approximations to reformulate this intractable N-electron problem as a collection of manageable one-electron problems. In the following chapter, we will discuss how to improve on this by returning to the form above, armed with the knowledge gained from the one-electron treatment, and trying to tackle it directly.

2.2 Single-Electron Methods

Single-electron, or 'independent-particle', approaches to solving Eq. 2.7 rely on mapping the many-electron problem onto a series of single-electron problems involving a Hamiltonian where the explicit electron-electron interactions are replaced by an effective potential $V^{\mathrm{eff}}(\mathbf{r})$:

$$\left[-\frac{1}{2}{\nabla_{\mathbf{r}}}^{2}+V_{i}^{\mathrm{eff}}(\mathbf{r})\right]\psi_{i}(\mathbf{r})=\epsilon_{i}\psi_{i}(\mathbf{r})\ . \tag{2.8}$$

The effective potential V_i^{eff} may be different for each orbital i, as in Hartree theory, or it may be a single value for all the orbitals as in Density Functional Theory (DFT). We will explain briefly how this effective potential is obtained in Hartree and Hartree-Fock theory before describing DFT in Section 2.3.

The simplest attempt at a solution for $\Psi(\{\mathbf{r}_i\})$ would be a product of one-electron orbitals $\psi_i(\mathbf{r})$:

$$\Psi(\{\mathbf{r}_i\})=\prod_i\psi_i(\mathbf{r}_i)\ , \tag{2.9}$$

whose optimal solution, by the variational principle, will obey

$$\frac{\delta}{\delta\psi_i^{*}(\mathbf{r}_i)}\left(\langle\Psi|\hat{H}|\Psi\rangle-\sum_j\lambda_j\langle\psi_j|\psi_j\rangle\right)=0\ . \tag{2.10}$$

The Lagrange multipliers λ_j enforce the normalization of the orbitals. Inserting the Hamiltonian of Eq. 2.7 gives us

$$\left(-\frac{1}{2}\nabla_{\mathbf{r}}^{2}+\sum_{j\neq i}\int\mathrm{d}\mathbf{r}'\,\frac{|\psi_j(\mathbf{r}')|^{2}}{|\mathbf{r}-\mathbf{r}'|}-\sum_I\frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}\right)\psi_i(\mathbf{r})=\lambda_i\psi_i(\mathbf{r})\ , \tag{2.11}$$

which is an effective Schrödinger equation for the single particle orbital ψ_i in the form of Eq. 2.8. However, Eq. 2.9 is not an acceptable wavefunction for electrons, as it is not antisymmetric under exchange of electron coordinates, which the wavefunctions of Fermions must be.

The antisymmetry comes in the form of a Slater determinant, which is the result of the application of an antisymmetrization operator $\hat{A}$ to the product wavefunction of Eq. 2.9:

$$\Psi_{HF}(\{\mathbf{r}_i,\sigma_i\}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1,\sigma_1) & \psi_1(\mathbf{r}_2,\sigma_2) & \dots & \psi_1(\mathbf{r}_N,\sigma_N) \\ \psi_2(\mathbf{r}_1,\sigma_1) & \psi_2(\mathbf{r}_2,\sigma_2) & \dots & \psi_2(\mathbf{r}_N,\sigma_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{r}_1,\sigma_1) & \psi_N(\mathbf{r}_2,\sigma_2) & \dots & \psi_N(\mathbf{r}_N,\sigma_N) \end{vmatrix} . \tag{2.12}$$

If we apply the Hamiltonian to this, we get

$$\langle\Psi_{HF}|\hat{H}|\Psi_{HF}\rangle = \sum_i H_i + \frac{1}{2}\sum_{ij}(J_{ij} - K_{ij}) \ , \tag{2.13}$$

where H_i are the single particle contributions

$$H_i = \int \mathrm{d}^3\mathbf{r}\,\psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla_\mathbf{r}^2 + \sum_I \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|} \right) \psi_i(\mathbf{r}) \tag{2.14}$$

and the J_{ij} are the Coulomb integrals as in Hartree theory:

$$J_{ij} = \int\int \mathrm{d}^3\mathbf{r}'\,\mathrm{d}^3\mathbf{r}\,\psi_i(\mathbf{r})\psi_i^*(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_j^*(\mathbf{r}')\psi_j(\mathbf{r}') \ . \tag{2.15}$$

The K_{ij} terms, however, are the exchange terms resulting from the antisymmetry:

$$K_{ij} = \int\int \mathrm{d}^3\mathbf{r}'\,\mathrm{d}^3\mathbf{r}\,\psi_i(\mathbf{r})\psi_j^*(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{r}')\psi_i^*(\mathbf{r}')\,\delta_{\sigma_i\sigma_j} \ . \tag{2.16}$$

It is not necessary to exclude $i = j$ terms in Eq. 2.13 because $K_{ii} = J_{ii}$. The same procedure as above gives us an effective-potential form to find the orbitals ψ_i. We write

$$\left(-\frac{1}{2}\nabla^2_{\mathbf{r}} + \hat{j} + \hat{k} - \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}\right)\psi_i(\mathbf{r}) = \lambda_i\psi_i(\mathbf{r}) \ , \tag{2.17}$$

where we have defined

$$\hat{j}\psi_i(\mathbf{r}) = \sum_j \psi_i(\mathbf{r}) \int \mathrm{d}^3\mathbf{r} \, \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \tag{2.18}$$

and

$$\hat{k}\psi_i(\mathbf{r}) = \sum_j \psi_j(\mathbf{r}) \int \mathrm{d}^3\mathbf{r} \, \frac{\psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \delta_{\sigma_i\sigma_j} \ . \tag{2.19}$$

We can now solve Eq. 2.17 for the orbitals $\psi_i(\mathbf{r})$ to obtain the Hartree-Fock energy through Eq. 2.13.

In Hartree-Fock, the effect of antisymmetry in keeping electrons apart is explicit, and can be expressed as the *exchange energy* in Eq. 2.16, the difference between the Hartree and Hartree-Fock energies. However, antisymmetry only affects electrons of like-spin: it does nothing to keep apart electrons of opposite spin, so the Hartree-Fock energy is always higher than the true ground state. The difference between the Hartree-Fock energy and the true ground state, the result of further correlation not included in exchange antisymmetry, is referred to as the *correlation energy*, which DFT attempts to approximately include, as well as an approximation to exchange, by taking a rather different approach to the problem.

2.3 Density Functional Theory

Density functional theory is founded on a pair of theorems by Hohenberg and Kohn [74], which state that for a given external potential $v_{\text{ext}}(\mathbf{r})$ acting on a many-electron system, it is possible to define a functional $E_{\text{HK}}[n]$ of the electron density $n(\mathbf{r})$ which is minimized and equal to the exact ground state energy when $n(\mathbf{r})$ is the ground state electron density, and that for a given number of interacting particles in an external potential $v_{\text{ext}}(\mathbf{r})$, this potential is uniquely determined by the ground state density $n(\mathbf{r})$. The theory was extended to spin polarized systems by von Barth and

Hedin [173], in terms of a unique functional of the spin density $E\left[n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})\right]$. Reviews of density functional theory focusing on different areas can be found by, among others: Jones and Gunnarsson, 1989 [84]; Dreizler and Gross, 1990 [39]; and Gross, Runge, and Heinonen, 1991 [67]. In this section we will give an extremely brief overview of the basic theory, without any details of the implementation of the method — which in many modern codes [34, 63, 152] does not closely resemble the route outlined here.

Because they are so central to the theory, and their proofs are so simple, we will outline a proof of the Hohenberg-Kohn theorems. Proof that $n(\mathbf{r})$ uniquely determines $v_{ext}(\mathbf{r})$ proceeds by contradiction: assume that there are two external potentials $v_{ext}^{(1)}(\mathbf{r})$ and $v_{ext}^{(2)}(\mathbf{r})$ which differ by more than a constant, producing different Hamiltonians $\hat{H}^{(1)}$ and $\hat{H}^{(2)}$, which thus have different ground state wavefunctions $\Psi^{(1)}$ and $\Psi^{(2)}$. Then, because $\Psi^{(2)}$ is not the ground state of $\hat{H}^{(1)}$, the variational principle demands that

$$E^{(1)} = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle < \langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle \ . \tag{2.20}$$

The assumption of strict inequality is discussed further in [74], and the result can be generalized to degenerate cases where this cannot be assumed. We write the last term of Eq. 2.20 as

$$\langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle = \langle \Psi^{(2)} | \hat{H}^{(2)} | \Psi^{(2)} \rangle - \langle \Psi^{(2)} | \hat{H}^{(2)} - \hat{H}^{(1)} | \Psi^{(2)} \rangle \ . \tag{2.21}$$

The interactions between particles are identical in both Hamiltonians, so only the interaction of the density with the external potential differs, giving

$$E^{(1)} < E^{(2)} + \int \mathrm{d}^3\mathbf{r} \ \left[v_{\mathrm{ext}}^{(1)}(\mathbf{r}) - v_{\mathrm{ext}}^{(2)}(\mathbf{r}) \right] n(\mathbf{r}) \tag{2.22}$$

Interchanging the subscripts also gives us

$$E^{(2)} < E^{(1)} + \int \mathrm{d}^3\mathbf{r} \ \left[v_{\mathrm{ext}}^{(2)}(\mathbf{r}) - v_{\mathrm{ext}}^{(1)}(\mathbf{r}) \right] n(\mathbf{r}) \tag{2.23}$$

Adding the two together gives the evidently contradictory result $E^{(1)} + E^{(2)} < E^{(1)} + E^{(2)}$, completing the proof. Crucially, this also ensures that the ground state wavefunction itself, and thus all the properties of the system, are uniquely determined by $n(\mathbf{r})$.

Proof of the second theorem, that there exists a universal functional for the energy $E[n(\mathbf{r})]$, proceeds by a similar approach, and a presentation can be found in, for example Ref. [116]. The original Hohenberg-Kohn proof only applied to a certain subset of possible densities: those that are the ground state density of the electron Hamiltonian with some external potential $v_{\text{ext}}(\mathbf{r})$, also known as "V-representable". It was later extended to the more general class of "N-representable" densities — those that are derivable from any valid N-electron wavefunction Ψ_N — by the approach of Levy and Lieb [108, 109, 112] involving a constrained search over the space of wavefunctions that produce a particular density.

The Hohenberg-Kohn theorem and its extensions are powerful arguments for the value of a density-based approach, but on their own they provide no recipe for calculating wavefunctions or other properties of the system from its density. The key to implementing a functional in practice is a substitution, known as the Kohn-Sham ansatz [94], which replaces the original many-body problem with an auxiliary independent-particle problem. The assumption is that the ground state density of the full interacting system, representing the minimization of the unknown Hohenberg-Kohn functional, is equal to the ground state density of a system of independent particles, which is found by minimizing a known functional where the explicit interactions are replaced by a local effective potential $V_{\text{eff}}(\mathbf{r})$. The original Hohenberg-Kohn theorem states that a unique functional $E_{\text{HK}}[n]$ exists for the energy of the full interacting system, which can be written as

$$E_{\text{HK}}[n] = T[n] + E_{\text{int}}[n] + E_{\text{ext}}[n] \equiv F_{\text{HK}}[n] + E_{\text{ext}}[n] \; , \tag{2.24}$$

where $F_{\text{HK}}[n]$ is the sum of the kinetic and interaction energies $T[n]$ and $E_{\text{int}}[n]$ of the particles, and

$$E_{\text{ext}}[n] = \int d\mathbf{r}\, V_{\text{ext}}(\mathbf{r})\, n(\mathbf{r}) \tag{2.25}$$

is the energy of the external potential (atoms, quantum wells, jellium slabs etc). The Kohn Sham approach replaces the total energy functional with

$$E_{\text{KS}}[n] = T_{\text{s}}[n] + E_{\text{H}}[n] + E_{\text{xc}}[n] + E_{\text{ext}}[n] \tag{2.26}$$

where

$$E_{\text{H}}[n] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}'\, \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.27}$$

is the Hartree Energy, $T_s[n]$ is the kinetic energy of the noninteracting system, and $E_{xc}[n]$ is the exchange-correlation energy, which attempts to group together all the many body effects of exchange and correlation:

$$E_{\text{xc}}[n] \;=\; F_{\text{HK}}[n] - (T_{\text{s}}[n] + E_{\text{H}}[n]) \;=\; T[n] - T_{\text{s}}[n] + E_{\text{int}}[n] - E_{\text{H}}[n] \tag{2.28}$$

We can write the density $n(\mathbf{r})$ and the non-interacting kinetic energy $T_{\text{s}}[n]$ in terms of the N lowest-energy Kohn-Sham orbitals $\psi_i(\mathbf{r})$ of the auxiliary Hamiltonian (to be defined shortly) as

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 \tag{2.29}$$

and

$$T_{\text{s}} = -\frac{1}{2}\sum_{i=1}^{N}\langle\psi_i|\nabla^2|\psi_i\rangle = \frac{1}{2}\sum_{i=1}^{N}\int |\nabla\psi_i(\mathbf{r})|^2 d\mathbf{r}\ . \tag{2.30}$$

The spin labels have been omitted here for clarity, but in spin polarized systems there will be separate spin up and spin down densities, and spin-dependent orbitals and eigenvalues.

The ground state of the Kohn-Sham auxiliary system, and thus the ground state density and energy of the full interacting system, is found by minimizing the total energy with respect to the orbitals, using the chain rule on the quantities expressed as functionals of the density,

$$\frac{\delta E_{\rm KS}}{\delta\psi_i^*(\mathbf{r})} = \frac{\delta T_{\rm s}}{\delta\psi_i^*(\mathbf{r})} + \left[\frac{\delta E_{\rm ext}}{\delta n(\mathbf{r})} + \frac{\delta E_{\rm H}}{n(\delta\mathbf{r})} + \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\right]\frac{\delta n(\mathbf{r})}{\delta\psi_i^*(\mathbf{r})} = 0\ , \tag{2.31}$$

subject to the constraints

$$\langle\psi_i|\psi_j\rangle = \delta_{i,j}\ . \tag{2.32}$$

Eqs. 2.29 and 2.30 give

$$\frac{\delta T_{\rm s}}{\delta\psi_i^*(\mathbf{r})} = -\frac{1}{2}\nabla^2\psi_i(\mathbf{r}); \qquad \frac{\delta n(\mathbf{r})}{\delta\psi_i^*(\mathbf{r})} = \psi_i(\mathbf{r})\ . \tag{2.33}$$

If we use the method of Lagrange Multipliers for the diagonal $(i = j)$ terms of the constraints, the Euler-Lagrange equation obtained is

$$\left(H_{\rm KS} - \epsilon_i\right)\psi_i(\mathbf{r}) = 0\ , \tag{2.34}$$

where H_{KS} is the effective Kohn-Sham Hamiltonian, given by

$$H_{\rm KS} = -\frac{1}{2}\nabla^2 + V_{\rm KS}(\mathbf{r})\ , \tag{2.35}$$

and the Kohn-Sham potential is given by

$$V_{\rm KS}(\mathbf{r}) = V_{\rm ext}(\mathbf{r}) + \frac{\delta E_{\rm H}}{\delta n(\mathbf{r})} + \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\ . \tag{2.36}$$

The solutions of Eq. 2.34 are found by matrix diagonalization, so they are all mutually orthogonal, ensuring that the off-diagonal constraints are imposed automatically.

An exact exchange-correlation functional $E_{\rm xc}[n]$, with which these equations would be exact and equivalent to the full Hohenberg-Kohn form, would have to express all the complications of the many-body correlation at long and short ranges, and is clearly far too complex an object for there to be any realistic hope of approximating it. Indeed it can be proved that the exact functional is not even analytic, due to the required behaviour of the functional as the number of electrons in the system is varied [130, 149]. Therefore, the most commonly used approach is to approximate E_{xc} by a local functional of the density, and optionally its gradient. The most familiar approximations of this type are the Local (Spin) Density Approximation (L(S)DA) and the Generalized Gradient Approximation (GGA). In the case of the LDA, the approximate exchange-correlation energy combines the exact exchange energy density for the homogeneous electron gas and a fit [133] to the correlation energy density based on diffusion Monte Carlo data for the Homogeneous Electron Gas (HEG) [31].

Clearly, in general, the eigenfunctions produced by solving Eq. 2.35 do not necessarily produce the same density as the $n(\mathbf{r})$ used to construct $V_{\text{KS}}(\mathbf{r})$ when Eq. 2.29 is used to recalculate the density. The energies and wavefunctions are only meaningful once they are *self-consistent*, that is to say, the output $n(\mathbf{r})$ from the wavefunctions that solve Eq. 2.35 is the same as that which produced $V_{\text{KS}}(\mathbf{r})$. Therefore, starting from an initial guess (which does not need to be very close to the solution) a procedure is used which mixes the input and output densities in such a way as to produce a self-consistent density where $n_{\text{out}}(\mathbf{r}) = n_{\text{in}}(\mathbf{r})$. Sometimes, linear mixing, where a certain fraction α of the new density is mixed with $(1-\alpha)$ of the old is sufficient, but in more difficult cases there are a large number of more complex mixing schemes which exist to ensure convergence, such as the Broyden method [20]. Alternatively, it is possible to directly minimise the interacting energy by using a scheme such as conjugate gradients directly upon the wavefunction coefficients in the basis used.

Density Functional Theory and its extensions can be used to produce energies and electron densities in very good agreement with experiment. The nature of the independent particle approach used to make the Kohn-Sham method feasible means that the specific eigenvalues and eigenvectors it produces are not in general physically meaningful as individual single-particle states, but the requirement that they produce the correct density means that they make a good starting point for explicit many-body calculations. In particular, they make good starting points for the single particle orbitals in a quantum Monte Carlo calculation, when combined with a Jastrow factor to include the effects of correlation.

2.4 Pseudopotentials

From consideration of the single-electron orbitals of individual atoms, it is clear that the lowest-energy eigenstates, those of the 'core' electrons surrounding a nucleus are very tightly bound to that nucleus, and do not extend far from it spatially. Consequently, the potential in which they move is dominated by that from the nucleus and they are comparatively unaffected by any environment in which the atom might be placed in a calculation. This is the basis of the pseudopotential approximation [134], which seeks to replace the strong Coulomb potential of the nucleus and the tightly bound, rapidly-varying orbitals of the core electrons with an effective ionic potential which acts on the remaining 'valence' electrons [35]. A pseudopotential is generated for a lone atom in a prior calculation, and is then used in a calculation of a molecule or solid to represent the nucleus and its core electrons, which are assumed to be 'frozen' in the same orbitals they occupy in the atom. The generation involves the calculation of the orbitals and the effective potential for the full all-electron atom within a given single-electron method — generally DFT or Hartree-Fock. The valence orbitals are then modified to define the pseudowavefunctions ψ_{ps}. The advantage of the pseudowavefunctions over the true valence wavefunctions is that they do not oscillate rapidly inside the core region (which would be necessary if they had to retain orthogonality with the core wavefunctions) and can thus be described much more easily with representations which rely on smoothness, such as plane waves. The Schrödinger equation is then inverted to find the pseudopotential V_{ps} which would have generated the pseudowavefunctions, under the constraint that the eigenvalues are the same as those of the original valence states.

A pseudopotential is constructed to match the true potential outside a specified radius r_c, the core radius, and the pseudowavefunctions that go with it match the true orbitals beyond this distance also. It is also desirable to reproduce the

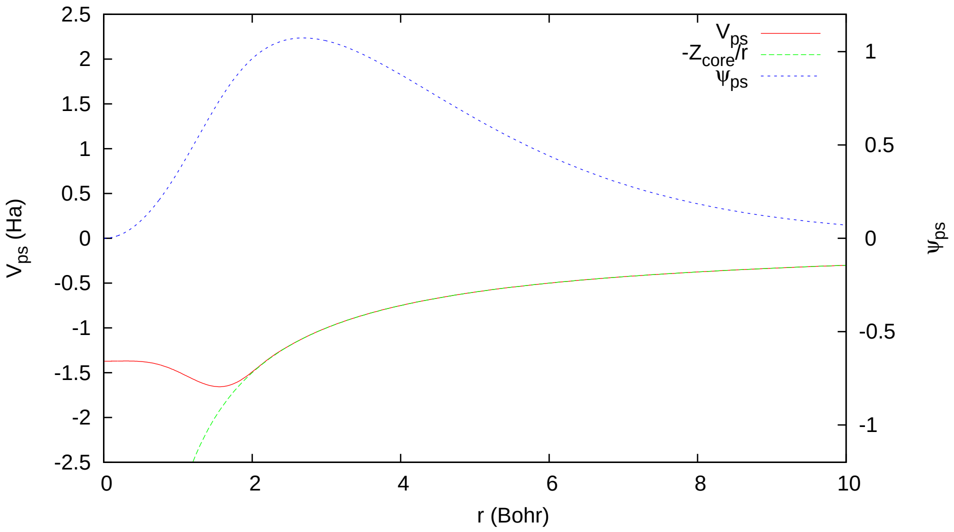

Figure 2.1: A Hartree-Fock generated pseudopotential for Al, with $Z_{\text{core}} = 3$, hence 3 valence electrons. Against the left scale, the red line shows $V_{\text{ps}}^{l=1}$, the pseudopotential for the p channel, compared to $-3/r$

phase shift upon scattering of a plane wave by the core. Because this phase shift is different for different angular momentum states, pseudopotentials must in general be non-local operators, to reproduce the scattering properties for all angular momentum states. The phase shift is always matched at the valence eigenvalue, but various techniques can be used to ensure this happens over a wider range of energies.

One condition that is often applied is that of norm-conservation, which ensures that the charge in the core region (corresponding to the integral $\int \mathrm{d}^3\mathbf{r}\, \psi_{\text{ps}}(\mathbf{r})$ up to $r = r_c$) is the same [68] as for the true wavefunctions. Norm conservation ensures that the energy derivative of the phase shift is correct at the valence eigenvalue, so that the real wavefunction and pseudowavefunction phase shifts agree (modulo 2π) over a wider range of energies. The norm-conservation condition is relaxed in other constructions such as the Vanderbilt 'ultrasoft' pseudopotentials [169] to enable other favourable properties to be enforced.

A nonlocal pseudopotential is often written in a 'semi-local' form, where one of the angular momentum channels is chosen as local and the rest are expressed as the difference relative to the local channel:

$$\hat{V}_{\mathrm{ps}} = \hat{V}_{\mathrm{loc}} + \sum_{lm} (V_l - V_{\mathrm{loc}}) \hat{P}_{lm} \tag{2.37}$$

where $\hat{V}_{\mathrm{loc}}$ is the pseudopotential of the local channel, V_l are the pseudopotentials for angular momentum channel l, and $\hat{P}_{lm}$ are projectors which project out the different angular momentum components lm of the electronic wavefunctions. It is sufficient to express $\hat{V}_{\mathrm{ps}}$ in terms of only a small number of angular momentum channels: usually only $l = 0, 1, 2$ are used. The choice of which channel to make the local channel is arbitrary, but is usually the highest l channel, as this saves the number of projectors that need to be calculated. A further improvement can be made to reduce the number of integrals required to calculate the projections, by writing the pseudopotential in Kleinman-Bylander form [90]:

$$\hat{V}_{\mathrm{KB}} = \hat{V}_{\mathrm{loc}} + \sum_{lm} \frac{|\delta \hat{V}_l \phi_{lm}\rangle \langle \phi_{lm} \delta \hat{V}_l|}{\langle \phi_{lm} | \delta \hat{V}_l | \phi_{lm} \rangle} \tag{2.38}$$

where the $|\phi_{lm}\rangle$'s are eigenstates of the pseudohamiltonian.

The quality of a pseudopotential is determined by its *transferability*, i.e. how accurate it remains in a variety of different chemical environments. A good pseudopotential will be able to reproduce properties such as binding energies, geometries and forces in a variety of different forms, with different types of bonding and different structures. The pseudopotentials we use in this work in Chapter 6 have been thoroughly tested in advance and found to be highly transferable.

Chapter 3

Quantum Monte Carlo

3.1 Background

Monte Carlo methods encompass a broad spectrum of stochastic techniques for simulation of physical and mathematical systems, whose only unifying characteristic is the use of random (or more usually pseudo-random) numbers. The strength of these methods lies in their ability to sample configuration spaces of systems with large numbers of coupled degrees of freedom in an efficient manner. In particular, Monte Carlo integration is a numerical quadrature method for approximate evaluation of multi-dimensional integrals by sampling random points within the configuration space and evaluating the integrand at each of these points. After enough samples have been accumulated, the sample mean asymptotically approaches the expectation value of the integrand over the whole space. The advantage of this method in terms of accuracy for a given computational effort becomes more and more pronounced as the dimensionality d of the integral increases: conventional mesh based methods scale increasingly poorly with system size as d increases. For example, using the d-dimensional generalization of Simpson's rule and a cubic mesh of M points, the error scales as $M^{-2/d}$, whereas a Monte Carlo sample of M points will have an error decaying according to the central limit theorem as $M^{-1/2}$ independent of d. For $d > 4$, therefore, Monte Carlo methods begin to win out over traditional mesh based quadrature. For other quadrature methods the exponents vary, but no matter how one divides up one's configuration space, with increasing dimensionality Monte Carlo eventually becomes the more efficient technique. It is this fact that suits it ideally to the integrals in enormous numbers of dimensions that characterize the expectation values of many-body quantum systems. In this section we briefly introduce the Monte Carlo technique, focussing on those aspects relevant to QMC.

We denote by a vector $\mathbf{R}$ a single position in a many-dimensional configuration space, and by $\mathcal{P}(\mathbf{R})$ an associated probability density, which obeys

$$\mathcal{P}(\mathbf{R}) \geq 0 \qquad \text{and} \qquad \int \mathrm{d}\mathbf{R}\, \mathcal{P}(\mathbf{R}) = 1 \; . \tag{3.1}$$

Many problems in physics amount to evaluating for some probability distribution estimates of the mean

$$\mu_f = \int \mathrm{d}\mathbf{R}\, f(\mathbf{R})\mathcal{P}(\mathbf{R}) \; , \tag{3.2}$$

and variance

$$\sigma_f = \int \mathrm{d}\mathbf{R}\, [f(\mathbf{R}) - \mu_f]^2\, \mathcal{P}(\mathbf{R}) \; . \tag{3.3}$$

of some function $f(\mathbf{R})$.

An estimate I of μ_f may be obtained by evaluating the mean of its value at a finite number M of points $\mathbf{R}_i$ sampled from $\mathcal{P}(\mathbf{R})$:

$$I = \frac{1}{M} \sum_{i=1}^{M} f(\mathbf{R}_i) \; . \tag{3.4}$$

Here the central limit theorem comes into play, in that not only does this mean tend to μ_f as $M \to \infty$, but also the error on the estimate will be given by

$$\sigma_I = \frac{\sigma_f}{\sqrt{M}} \; , \tag{3.5}$$

allowing us to obtain an arbitrarily accurate estimate of μ_f by increasing M until a desired σ_I is obtained.

More generally, this technique can be used to evaluate any integral of the form

$$I = \int_{\Omega} \mathrm{d}\mathbf{R}\, g(\mathbf{R}) \tag{3.6}$$

by splitting $g(\mathbf{R})$ into a 'score function' $f(\mathbf{R})$ and an 'importance function' $\mathcal{P}(\mathbf{R})$, such that $g(\mathbf{R}) = f(\mathbf{R})\mathcal{P}(\mathbf{R})$.

Quantum Monte Carlo refers to the application of Monte Carlo methods to simulation of quantum systems. There are again a wide range of techniques that fall under this name, but of these two are commonly used in application to finding ground state properties of extended fermion systems. The first, and simpler, of these is variational Monte Carlo, which applies the aforementioned Monte Carlo integration scheme to the evaluation of expectation values of quantum mechanical operators, primarily the Hamiltonian, on a trial wavefunction. The second, a more involved but potentially considerably more accurate and useful method, is diffusion Monte Carlo, which uses a Green function to propagate the Schrödinger equation in imaginary time so as to project out the ground state component of a starting trial function. These two methods account for the majority of applications of QMC to solids, but there are other variations as well. For example, path integral Monte Carlo [29] uses the path integral formulation of quantum mechanics to simulate systems at non-zero temperature. Stochastic sampling can also be applied to methods more traditionally associated with quantum chemistry such as Møller-Plesset theory by sampling the graphs corresponding to excitations in the space of Slater determinants [160]. In this chapter we discuss only VMC and DMC, which are the QMC techniques employed in this work.

3.2 Variational Monte Carlo

Variational Monte Carlo is the simplest means of applying the ideas of Monte Carlo methods to quantum mechanics. One would like to be able to make an estimate $\bar{O}$ of the expectation value of a Hermitian operator $\hat{O}$ representing some observable, which for a wavefunction Ψ_T of unknown normalization can be achieved by evaluating an integral of the form

$$\bar{O} = \frac{\int \mathrm{d}\mathbf{R}\, \Psi_T^*(\mathbf{R})\hat{O}\Psi_T(\mathbf{R})}{\int \mathrm{d}\mathbf{R}\, \Psi_T^*(\mathbf{R})\Psi_T(\mathbf{R})} \; . \tag{3.7}$$

This can be put in a form easily amenable to the method described above if it can be written as

$$\bar{O} = \frac{\int \mathrm{d}\mathbf{R} \left(\frac{\hat{O}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}\right) |\Psi_T(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R}\; |\Psi_T(\mathbf{R})|^2} \; . \tag{3.8}$$

The quantity in parentheses, $O_L(\mathbf{R}) = \left(\frac{\hat{O}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}\right)$ is known as the 'local' value of $\hat{O}$ as it depends only on the wavefunction at the specific point $\mathbf{R}$ in configuration space.

In essence, what this involves is 'guessing' a trial solution $\Psi_T(\mathbf{R})$ to the Schrödinger equation for the N-electron wavefunction. The normalized probability density $|\Psi_T(\mathbf{R})|^2/\int \mathrm{d}\mathbf{R}\, |\Psi_T(\mathbf{R})|^2$ acts as the importance function, and by using the local value of $O_L(\mathbf{R})$ as the score function and integrating over the $3N$-dimensional configuration space of the electron coordinates $\mathbf{R}$, one can estimate the expectation value of the operator $\hat{O}$ for this wavefunction. If $\hat{O}$ is the Hamiltonian $\hat{H}$, this gives a variational upper bound on the ground state energy, which can be improved by improving the quality of the trial wavefunction.

The fact that the estimate is variational, in that $\bar{E} \geq E_0$ where E_0 is the true ground state energy, can be seen by expanding in the basis of energy eigenstates for which $\hat{H}\Psi_i = E_i\Psi_i$ (we neglect normalization as Ψ_T was not itself assumed to be normalized) as

$$\Psi_T = \Psi_0 + \sum_{i>0} \alpha_i \Psi_i \; . \tag{3.9}$$

Inserting this into Eq. 3.7 for $\hat{O} = \hat{H}$ gives an energy of

$$E_T = E_0 + \frac{\sum_{i>0} |\alpha_i|^2 (E_i - E_0)}{\sum_{i>0} |\alpha_i|^2 + 1} \; , \tag{3.10}$$

so the error in the energy estimate $\Delta E = E_T - E_0$ is always positive and is of order α_i^2. Linear reduction of the error in the wavefunction will therefore give quadratic reduction of the error in the energy. A similar argument can be applied to Ψ_T close to one of the other eigenstates E_j, but in that case there is no assurance the energy found will be an upper bound on the true eigenvalue, as some values of $E_i - E_j$ will be negative.

The strict variational theorem applies as long as certain criteria are met for the trial wavefunction. Firstly, Ψ_T must obey the right symmetry properties and boundary conditions (antisymmetry under exchange of electron coordinates, and the right periodicity or decay to infinity). Secondly, $\hat{H}\Psi_T$ must be defined everywhere, so Ψ_T and $\nabla\Psi_T$ must be continuous everywhere the potential is finite. Finally, to enable the Monte Carlo method to work, the integrals $\int \mathrm{d}\mathbf{R}\, \Psi_T^2$, $\int \mathrm{d}\mathbf{R}\, \Psi_T^2 \hat{H}\Psi_T$, and $\int \mathrm{d}\mathbf{R}\, (\hat{H}\Psi_T)^2$, must all exist (the latter because otherwise the statistical error on the energy will be infinite).

To actually apply this method in practice, one needs: (a) a means of sampling a series of configurations $\mathbf{R}$ from a probability distribution given by $|\Psi(\mathbf{R})|^2$, (b) methods to efficiently evaluate the local value of the Hamiltonian or other operator, and (c) a method of generating and systematically improving upon a trial guess of the ground-state wavefunction $\Psi(\mathbf{R})$. The following subsections will examine how each of these are achieved.

3.2.1 The Metropolis Algorithm

The Metropolis Algorithm [120] lies at the heart of practical implementations of Monte Carlo methods: it allows highly efficient sampling of configurations from an arbitrary probability distribution by evolving the position of a walker (or many walkers) exploring the configuration space by taking a series of small steps, which are accepted or rejected according to the following rules:

1. Initialize the position $\mathbf{R}$ at random in the system.
2. Make a trial move to a new position $\mathbf{R}'$ chosen from some probability density function $T(\mathbf{R}' \leftarrow \mathbf{R})$.

3. Generate a uniform deviate ν from the interval [0, 1] and compare it to the probability $A(\mathbf{R}' \leftarrow \mathbf{R}) = \mathrm{Min}\left(1, \frac{T(\mathbf{R}'\leftarrow\mathbf{R})P(\mathbf{R}')}{T(\mathbf{R}\leftarrow\mathbf{R}')P(\mathbf{R})}\right)$ where $P(\mathbf{R})$ is the probability density we are trying to sample. If $\nu \leq A$, add the new position of the configuration $\mathbf{R}'$ to the *walk* comprising the series of recorded configurations. Otherwise, the configuration remains at $\mathbf{R}$ and another point on the walk is added here.
4. Return to step (2) and continue until sufficient points have been accumulated.

The only restriction on the probability distribution $T(\mathbf{R}' \leftarrow\mathbf{R})$ is that if $T(\mathbf{R}' \leftarrow\mathbf{R})$ is nonzero then $T(\mathbf{R}\leftarrow\mathbf{R}')$ must also be. The details of the choice, however, affect the efficiency of the algorithm by determining the number of steps over which successive configurations are serially correlated. If a typical $\mathbf{R} - \mathbf{R}'$ is large, the moves will be rejected too often and $\mathbf{R}$ will stay the same for many consecutive moves. If it is too small, $f(\mathbf{R}')$ will barely differ from $f(\mathbf{R})$. To maximize the efficiency one must balance one form of serial correlation against the other. In practice, one frequently chooses moves from a Gaussian distribution with a spread optimized such that around 50% of moves are accepted.

A detailed derivation of the working of this algorithm can be found in Feller [48], but some insight as to its operation can be seen by assuming that a very large population of walkers has reached an equilibrium distribution $n(\mathbf{R})$ which is constant in time and obeys detailed balance. This latter condition states that the number of walkers moving per step from some particular phase-space volume element $\mathrm{d}\mathbf{R}$ to any other $\mathrm{d}\mathbf{R}'$ is the same as the number moving from $\mathrm{d}\mathbf{R}'$ to $\mathrm{d}\mathbf{R}$. The probability of a walker moving from $\mathrm{d}\mathbf{R}$ to $\mathrm{d}\mathbf{R}'$ is the probability $T(\mathbf{R}' \leftarrow \mathbf{R})\mathrm{d}\mathbf{R}'$ of choosing $\mathrm{d}\mathbf{R}'$ to move to, multiplied by the probability $A(\mathbf{R}' \leftarrow \mathbf{R})$ of accepting such a move, so the rate of walkers doing so is

$$n(\mathbf{R})\mathrm{d}\mathbf{R}\, A(\mathbf{R}' \leftarrow \mathbf{R})T(\mathbf{R}' \leftarrow \mathbf{R})\mathrm{d}\mathbf{R}' \tag{3.11}$$

and the rate of walkers from $\mathrm{d}\mathbf{R}'$ to $\mathrm{d}\mathbf{R}$ is

$$n(\mathbf{R}')\mathrm{d}\mathbf{R}'\, A(\mathbf{R} \leftarrow \mathbf{R}')T(\mathbf{R} \leftarrow \mathbf{R}')\mathrm{d}\mathbf{R}\ . \tag{3.12}$$

Detailed balance ensures these are equal, giving

$$\frac{n(\mathbf{R})}{n(\mathbf{R}')} = \frac{A(\mathbf{R} \leftarrow \mathbf{R}')T(\mathbf{R} \leftarrow \mathbf{R}')}{A(\mathbf{R}' \leftarrow \mathbf{R})T(\mathbf{R}' \leftarrow \mathbf{R})}\ , \tag{3.13}$$

and from the definition of A in step 3 we know that

$$\frac{A(\mathbf{R} \leftarrow \mathbf{R}')}{A(\mathbf{R}' \leftarrow \mathbf{R})} = \frac{T(\mathbf{R}' \leftarrow \mathbf{R})P(\mathbf{R})}{T(\mathbf{R} \leftarrow \mathbf{R}')P(\mathbf{R}')}\ , \tag{3.14}$$

so the equilibrium density obeys:

$$\frac{n(\mathbf{R})}{n(\mathbf{R}')} = \frac{P(\mathbf{R})}{P(\mathbf{R}')}\ . \tag{3.15}$$

We have therefore shown that, after a sufficient equilibration period, the Metropolis algorithm is capable of delivering a distribution of walkers such that $n(\mathbf{R})$ is proportional to $P(\mathbf{R})$. Repeated use of the algorithm generates a chain of configurations $\mathbf{R}$, distributed as required but somewhat serially-correlated, in that $\mathbf{R}$ and $\mathbf{R}'$ are not independent because $\mathbf{R}$ has not moved very far in one step.

In practice, rather than choosing moves in which the whole vector $\mathbf{R}$ changes, it is more efficient to move electrons one-by-one. There are several benefits to this algorithm, such as the fact that the contribution to the averages of trial moves that were rejected can be included in the summations, and that an efficient two-level version of the algorithm [36] can be implemented which avoids computation of unnecessary quantities and considerably speeds up the calculation.

3.2.2 Trial Wavefunctions

We now consider how to construct a trial wavefunction and systematically improve upon it for a given system. Recent ideas suggest that analysis of the full unknown many-body wavefunction may allow one to predict a pattern to the hierarchy of importance of various possible terms in the mathematical form of the wavefunction as the level of approximation is increased [75]. This hierarchy begins with a set of independent one- or two-electron states and gradually includes higher and higher levels of correlation. Current calculations use only the first two or three levels of this hierarchy, which in their simplest forms represent a Slater determinant, combined with one-, two- and three-electron Jastrow factors (see Section 3.2.3) and the introduction of backflow (Section 3.2.4). In this and the following sections we will examine these forms.

In general (some recent calculations are moving beyond this) trial wavefunctions are built on a Slater determinant of single particle orbitals. This can be justified on the same grounds used to justify one electron-theories such as Hartree-Fock — namely that the dominant factor in the exchange-correlation energy is exchange rather than correlation. A Slater determinant, as previously introduced in Section 2.2, is the simplest means of including a form of wavefunction that is necessarily antisymmetric in all the electron coordinates and is thus capable of including the exchange energy correctly:

$$D(\mathbf{X}) = \hat{A}\prod_i \psi_i(\mathbf{r}_i,\sigma_i) = \begin{vmatrix} \psi_1(\mathbf{r}_1,\sigma_1) & \psi_1(\mathbf{r}_2,\sigma_2) & \dots & \psi_1(\mathbf{r}_N,\sigma_N) \\ \psi_2(\mathbf{r}_1,\sigma_1) & \psi_2(\mathbf{r}_2,\sigma_2) & \dots & \psi_2(\mathbf{r}_N,\sigma_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{r}_1,\sigma_1) & \psi_N(\mathbf{r}_2,\sigma_2) & \dots & \psi_N(\mathbf{r}_N,\sigma_N) \end{vmatrix} \tag{3.16}$$

For practical reasons, it is usually efficient to replace this full determinant by the product of spin-up and spin-down determinants

$$D(\mathbf{X}) \;\rightarrow\; D^{\uparrow}(\mathbf{r}_1,\dots,\mathbf{r}_{N_\uparrow})D^{\downarrow}(\mathbf{r}_{N_\uparrow+1},\dots,\mathbf{r}_N) \tag{3.17}$$

The orbitals in these two determinants are normally identical in a closed-shell calculation, but they may differ if the system is spin-polarized. Even if the orbitals are the same, this function is not strictly antisymmetric under exchange of electrons with opposite spins, but it can be shown that it gives the same expectation value as the full determinant for any spin-independent operator (see for example [54]).

The source of these orbitals will depend on the kind of lower-level calculation feasible and accurate for the system. In the case of single atoms, Hartree-Fock orbitals are generally preferred [65], for reasons to do with the accuracy of the nodal surface (see Section 3.3.4). In solids, where Hartree-Fock is not as accurate and is considerably more demanding, the Kohn-Sham orbitals of density functional theory are frequently used. Generally, even in cases where the treatment of exchange and correlation included in these methods is inadequate to capture some aspect of the relevant physics, they are an accurate enough starting point that the addition of correlation in ways we will describe shortly, via Jastrow factors and backflow, is sufficient to produce a wavefunction of high accuracy.

Several alternative starting points have been proposed recently for particular systems. Pairing wavefunctions, which introduce explicitly the kind of pairing of same-spin and opposite spin electrons that is required to adequately describe, for example, the Bardeen-Cooper-Schrieffer (BCS) wavefunction of a superconductor, can be used in place of the single-particle orbitals of the Slater determinant [26]. Pfaffians, a mathematical construction likened to a generalization of the idea of determinants, can be used in combination with pairing functions to generate even more accurate wavefunctions [10]. The 'Resonating Valence Bond' wavefunction [25] obtained by applying a Jastrow correlation term to an 'Antisymmetrized Geminal Power' [13] wavefunction, based upon singlet pairs of electrons, is particularly suited for describing the electronic structure of molecules. These forms will not be detailed here as they are not used in this work but it is important to note that the Slater determinant is not the only possible starting point.

3.2.3 Jastrow Factors

The Jastrow Factor $e^{J(\mathbf{R})}$ is an attempt to include the effect of short-ranged correlation by selectively reducing the amplitude of the wavefunction when pairs or larger groupings of electrons or other particles are close to each other, while not having any effect when they are far apart. An ideal Jastrow factor would in principle include the effects of general N-body correlations, but in practice it is common to discard higher than 3-body terms, and often only 2-body terms are included. The general form of a 2-electron Jastrow factor is:

$$J(\mathbf{R}) = -\frac{1}{2}\sum_{i}\sum_{j\neq i} u_{\sigma_i\sigma_j}(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i} \chi(\mathbf{r}_i) \ . \tag{3.18}$$

The two-body u-terms are usually simple functions of the electron-electron distances r_{ij}: one function $u_{\uparrow\uparrow}(r_{ij}) = u_{\downarrow\downarrow}(r_{ij})$ for two same-spin electrons and another function $u_{\uparrow\downarrow}(r_{ij}) = u_{\downarrow\uparrow}(r_{ij})$ for two different-spin electrons.

Some of the properties of these functions can be obtained by considering the dominant terms in the Hamiltonian as two electrons approach each other, i.e. as $r_{ij} \to 0$. Because the wavefunction must remain an eigenstate, the divergence in the Coulomb energy must be exactly cancelled by an equal and opposite divergence in the kinetic energy, which a single Slater determinant cannot provide. This approach gives rise to so-called ‘cusp conditions’ [86] which constrain the dependence of the many-electron wavefunction on the interparticle distance r_{ij} in the limit $r_{ij} \to 0$. For the approach of two same-spin electrons, it can be shown [86] that

$$\left.\frac{\partial u_{\uparrow\downarrow}(r_{ij})}{\partial r_{ij}}\right|_{r_{ij}=0} = \frac{1}{2} \ , \tag{3.19}$$

and for different-spins that

$$\left.\frac{\partial u_{\uparrow\uparrow}(r_{ij})}{\partial r_{ij}}\right|_{r_{ij}=0} = \frac{1}{4} \ . \tag{3.20}$$

In homogeneous systems, much can be determined about the required form of the Jastrow factor by analytical arguments. In an inhomogeneous system, however, it is more common to try to come up with a very general parametrized form obeying these cusp conditions and extending only to a certain cutoff radius, and then try to fit the best parameters by variance minimization, which is described in Section

3.2.6.

On their own, the u terms would tend to reduce the amplitude of the wavefunction in regions of high electron density, because there is a greater chance of multiple electrons being close to each other there. The effect of the one-body χ term is to counteract this and keep the density roughly what it was before. This is important because the single-particle orbitals are generally calculated within a self-consistent calculation like DFT or HF, and these schemes both yield rather accurate electron densities in most cases. The χ terms extend throughout the system but have peaks at positions where the electron density is high: in an atom, molecule or crystal this is near the nuclei; in a quantum dot, it will be at the dot centre. The χ function is often represented as a sum of radial functions which depend only on the distances r_{iI} from electron i to nucleus I.

A further complication caused by Jastrow factors relates to spin: a suitably chosen combination of Slater determinants of orbitals of particular spins $\hat{s}_z$ can be made to be an eigenfunction of the total spin operator $\hat{S}^2$, which is a quantity expected to be a good quantum number in most of the systems under study. However, inclusion of a Jastrow factor of the type shown above can easily be shown to introduce a certain amount of 'spin-contamination' such that the resulting Slater-Jastrow trial wavefunction is not an eigenfunction of $\hat{S}^2$. Fortunately, it has been shown [77] that the effect of this contamination is negligible in most cases and can safely be ignored.

3.2.4 Backflow

In Section 3.3.4 we will introduce the importance of the accuracy of the nodal surface in the determination of energies in DMC. Because a Jastrow factor $J(\mathbf{R})$ is always chosen to be real, $e^{J(\mathbf{R})}$ is nowhere negative, so it does not move the nodes of a wavefunction. As will be shown below, it therefore has no direct effect on the DMC energy except to lower its variance. One of the ways the nodal surface of the wavefunction can be improved is to introduce a different type of correlation known as backflow correlation. This was first introduced in the context of wavefunctions for liquid ^{4}He [49] and ^{3}He [50], and later applied to the Homogeneous Electron Gas (HEG) [100, 101], and most recently to atomic, molecular and even solid systems [113, 40, 19].

Conceptually, the idea behind backflow is to alter not the single-particle orbitals themselves but the coordinates $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ that go into them. The coordinate $\mathbf{r}_i$ is replaced by a new variable that depends both on $\mathbf{r}_i$ and on the positions of the other electrons and nuclei close to $\mathbf{r}_i$, allowing the electrons, in effect, to flow round each other so as to increase their average separation. Because this has a great deal in common with the classical fluid dynamical description of the fluid flow around an obstacle, it inherits much of its terminology from classical fluid dynamics. In practice, the introduction of backflow is capable not only of reducing the variance of both VMC and DMC, so lowering the error bars for a given run time, but also of retrieving a much larger fraction of the correlation energy in both methods, getting closer to the ground state. Backflow is also a very compact parameterization compared to other methods of improving the nodes such as large expansions in Slater determinant space, which makes it easier to optimize backflow wavefunctions. A comparison of QMC simulations with and without backflow is found in [113].

3.2.5 Operator Estimators in VMC

From the configurations produced by application of the Metropolis algorithm, one may wish to evaluate averages of a variety of operators. These may include the total energy, one of its components, other quantities such as Fourier components of the density, or as will be discussed later, operators relating to polarization and localization. Clearly, unless the wavefunction is an exact eigenstate of the operator being sampled, the recorded values will be subject to some statistical fluctuation. Without some idea of the bounds on the error, the mean itself is rather useless, but as the values will also be serially correlated it is not trivial to extract a meaningful error bar. One systematic way to identify and remove the effect of serial correlation is called reblocking [53].

Evaluation of the local value $O_L(\mathbf{R}_i)$ of operator $\hat{O}$ at a series of M configurations $\mathbf{R}_i$ gives a series of values O_i. Clearly the mean is

$$\bar{O} = \frac{1}{M}\sum_{i=1}^{M} O_i \ , \tag{3.21}$$

and the sample variance is

$$s_{M-1}^2 = \frac{1}{M-1}\sum_{i=1}^{M}(O_i - \bar{O})^2 \ . \tag{3.22}$$

If every sample was independent, this would be able to give a good estimate of the true error bounds ΔO on the mean, which would simply be the standard deviation of the estimator of the mean, $s_{N-1}/\sqrt{N}$. However, the aforementioned correlation means this is an underestimate of the true uncertainty.

Reblocking works by grouping the data points O_i into blocks of length L by taking the means $\bar{O}_j$ of the jth set of L points:

$$\bar{O}_j = \sum_{i=(j-1)L+1}^{jL} O_i \ . \tag{3.23}$$

The idea is that once L is comparable to the correlation time (in steps), then $\bar{O}_j$ will be uncorrelated with $\bar{O}_{j+1}$, and while the mean of all the values of $\bar{O}_j$ will be the same as $\bar{O}$, the variance will become

$$s_{M'-1}^2 = \frac{1}{M'-1}\sum_{j=1}^{M'}(\bar{O}_j - \bar{O})^2 \ , \tag{3.24}$$

with $M' = M/L$. This can be used to get a much more accurate estimate of the true statistical uncertainty of the sample mean, $\Delta O = \sqrt{s_{M'-1}^2/M'}$. With some trivial modifications, these formulae can be redefined to apply even where L is not an exact divisor of M. In practice, one must examine the trend of $s^2_{M'-1}/M'$ as a function of L to find the point where it levels out to a plateau: beyond this, L is greater than the correlation time and $s^2_{M'-1}/M'$ has converged and represents the true variance.

An alternative approach [81] involves trying to directly estimate the number of VMC steps in the correlation time, N_c, and simply multiplying up the raw error by $\sqrt{N_c}$ — this can be used for an automated estimate provided the error is relatively well-behaved.

The operator that is most commonly evaluated in QMC is, of course, the Hamiltonian, to give the total energy: this can be a costly and time-consuming process so it is worthy of special consideration here. There is an obvious split between kinetic energy and potential energy, and further between local and nonlocal parts of the potential. We thus write the local energy $E_L(\mathbf{R})$ as

$$E_L(\mathbf{R}) = \frac{\hat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} = \frac{\hat{T}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} + \frac{\hat{V}_{nl}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} + V_{loc}(\mathbf{R}) \; . \tag{3.25}$$

The kinetic energy can be decomposed into the contribution of each electron:

$$\frac{\hat{T}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} = \sum_i K_i(\mathbf{R}) = \sum_i \Psi^{-1}(\mathbf{R})(-\frac{1}{2}\nabla_i^2)\Psi(\mathbf{R}) \; , \tag{3.26}$$

which, for all useful choices of Slater determinant and Jastrow factor, can be evaluated analytically. There are several possible estimators for the kinetic energy, differing by integrations by parts (see [125]). They can be compared to each other to give a test that the calculation is working sensibly.

The electron-electron part of the potential energy can be evaluated either directly from the Coulomb interaction in finite systems, or via the Ewald [46] or Modified Periodic Coulomb (MPC) [177] interactions in periodic systems. The local part of the pseudopotential is also easily evaluated.

If our Hamiltonian contains nonlocal pseudopotentials, as is vital for simulation of realistic systems of anything other that first-row atoms, the Hamiltonian $\hat{H}$

contains a sum over the electrons i:

$$V_{nl} = \sum_i \Psi^{-1} \hat{V}_{nl,i} \Psi \ , \tag{3.27}$$

each of which require the evaluation of an integral. For a single atom at the origin, these integrals can be evaluated as a sum over the nonlocal channels l [47]

$$V_{nl,i} = \sum_l V^{ps}_{nl,l}(r_i) \frac{2l+1}{4\pi} \int \mathrm{d}^3\mathbf{r}_i \, P_l[cos(\theta'_i)] \frac{\Psi(\mathbf{r}_1, \ldots, \mathbf{r}'_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_N)}{\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_N)} \ , \tag{3.28}$$

where P_l is a Legendre polynomial. This integral is calculated numerically for $l = 0, 1, 2$ using a quadrature rule that integrates products of spherical harmonics, rotating the grid randomly for each integration to avoid systematic bias. The accuracy of this integration is controlled by a parameter l_{max} which controls the maximum spherical harmonic included. The treatment of nonlocal pseudopotentials thus represents a very costly part of the calculation. Moreover, although the integrals can be evaluated arbitrarily accurately in VMC there are further complications in DMC (see Section 3.3.6).

3.2.6 Wavefunction Optimization

Many of the ingredients used in the representation of the wavefunction contain variational parameters that need to be optimized to make the trial wavefunction Ψ_T as close as possible to the true ground state Ψ_0. In VMC, the accuracy of the wavefunction is the primary determining factor of the accuracy of the calculation, as the calculated energy is simply the expectation value of the Hamiltonian with this wavefunction. In DMC, the wavefunction plays a lesser but nevertheless vital role as a guide function and defines the nodal surface, as will be explained in the next section. Additionally, the DMC evaluation of quantities that do not commute with the Hamiltonian require use of both the DMC and VMC expectation values to calculate the “mixed” expectation value (see Section 3.3.5) and this relies heavily on the accuracy of the trial wavefunction. Consequently, no matter how one intends to do the calculation, there will be variational parameters that must be optimized. In many simple systems such as the homogeneous electron gas, good values of the parameters can often be obtained by comparison with analytical approaches such as calculations with the Random Phase Approximation — which, for example, can give the long-range behaviour of the two-body Jastrow factor. In more realistic systems, however, the parameters must be optimized as a preliminary step of the calculation itself.

Two principal methods exist for optimization of the wavefunction. Energy minimization [126, 161, 168] is perhaps the more intuitive route, involving minimization of the expectation of the energy with respect to the variational parameters, but in practice, despite recent advances in energy minimization, the older technique of variance minimization [166, 88] remains more common. Both methods rely on definition of a cost function and a minimization scheme, for which there are a wide range of choices. In this subsection we shall outline only the principles of the methods, not the details of their implementation.

Variance minimization relies on the fact that, for any eigenstate of the Hamiltonian, the variance of the local energy is zero, since $\hat{H}\Psi(\mathbf{R}) = E\Psi(\mathbf{R})$ everywhere. When the wavefunction is very close to the ground state, the dependence of both the energy and the variance on the accuracy of the wavefunction is at best quadratic, as both are almost minimized. Usually, however, the variance has a much stronger dependence on the quality of the trial wavefunction and is thus an easier quantity to minimize accurately. In terms of the local energy $E_L(\mathbf{R})$, the total energy for a real wavefunction $\Psi(\mathbf{R})$ parametrized by a set of parameters $\{\alpha\}$ could be written as

$$E = \frac{\int \mathrm{d}\mathbf{R}\, \Psi(\mathbf{R})^2 E_L(\mathbf{R})}{\int \mathrm{d}\mathbf{R}\, \Psi(\mathbf{R})^2} \,, \tag{3.29}$$

and the variance of this energy over the configuration space spanned by $\mathbf{R}$ is

$$\sigma^2 = \frac{\int \mathrm{d}\mathbf{R}\, \Psi(\mathbf{R})^2 (E_L(\mathbf{R}) - E)^2}{\int \mathrm{d}\mathbf{R}\, \Psi(\mathbf{R})^2} \,. \tag{3.30}$$

The standard variance minimization procedure generates a set of electron configurations $\{\mathbf{R}\}$ distributed according to the wavefunction $|\Psi^{\{\alpha_0\}}(\mathbf{R})|^2$ corresponding to some initial set of parameters α_0, and then uses nonlinear minimization algorithms to iteratively find the parameters $\{\alpha\}$ that minimize the variance. One usually sticks with a fixed set of configurations until the minimization process has reached a tolerable level of convergence. During this, the algorithm is referred to as either *reweighted* or *unreweighted* depending on whether the weights of the configurations are allowed to vary during the minimization procedure according to the change in wavefunction with the parameters, or whether they are kept constant. Once the procedure is completed, a new set of configurations is generated for the new wavefunction, and the process repeats until the wavefunction is as optimized as it can be for this parameterization.

There is also an alternative procedure that uses the fact that many of the parameters to be optimized appear as linear parameters in the Jastrow exponent, and at most quartic parameters in the variance. This allows direct analytic minimization techniques to be used [41].

Energy minimization in the presence of significant error bars on total energies and their derivatives with respect to the wavefunction parameters is very challenging, and a large number of tricks for dealing with such issues must be deployed for the method to succeed. As energy minimization is not used in this work, we will not discuss these tricks here, but the outline of a successful scheme can be found in [168].

3.3 Diffusion Monte Carlo

Diffusion Monte Carlo has a considerably more lofty goal than does variational Monte Carlo: for a given Hamiltonian, it can, in principle, project out the ground state in such a way as to give exact ground state energies for quantum systems. In this section, we show how the algorithm achieves this, concentrating more on the theoretical underpinnings than the technical details of the implementation, which are well explained in [125]. The specific techniques described here are not unique but represent the way the work of this thesis was carried out. For a wider discussion of DMC methods, presenting a range of alternative methodologies, see [54].

3.3.1 Imaginary Time Propagation

We begin from the Schrödinger equation, Eq. 2.1, and introduce the new variable $\tau = it$ to obtain

$$-\frac{\partial}{\partial\tau}\Phi(\mathbf{R},\tau) = (\hat{H} - E_T)\Phi(\mathbf{R},\tau)\ , \tag{3.31}$$

which is known as the imaginary-time Schrödinger equation. We have also included an arbitrary energy offset E_T for reasons which will become clear shortly. We can demonstrate the effect of evolving imaginary time τ, by writing $\Phi(\mathbf{R},t)$ in terms of the complete set of eigenvalues $\{E_i\}$ and eigenvectors ϕ_i of the Hamiltonian, which are of course stationary states whose time dependence is simply $\mathrm{e}^{-i(E_i - E_T)t}$:

$$\Phi(\mathbf{R},t) = \sum_i a_i\phi_i(\mathbf{R})\mathrm{e}^{-i(E_i - E_T)t}\ . \tag{3.32}$$

If $\tau = it$ this becomes

$$\Phi(\mathbf{R},\tau) = \sum_i a_i\phi_i(\mathbf{R})\mathrm{e}^{-(E_i - E_T)\tau}\ , \tag{3.33}$$

so in the limit of $\tau \rightarrow \infty$, and setting $E_T = E_0$ the ground state is projected out:

$$\lim_{\tau\rightarrow\infty}\Phi(\mathbf{R},\tau) = a_0\phi_0(\mathbf{R})\ . \tag{3.34}$$

This tells us that if we can simulate the imaginary time evolution of a wavefunction Φ, we can project out its ground state component ϕ_0 and thus evaluate the ground state energy E_0. The way to achieve this evolution is to express it in terms of the Green function for the Imaginary time Schrödinger equation (ITSE).

We begin by recasting Eq. 3.31 as an integral equation:

$$\Phi(\mathbf{R}; t + \tau) = \int G(\mathbf{R}', \mathbf{R}; \tau)\Phi(\mathbf{R}'; t)\mathrm{d}\mathbf{R}' \ , \tag{3.35}$$

where $G(\mathbf{R}', \mathbf{R}; \tau)$ is a Green function which is the solution to

$$-\frac{\partial}{\partial \tau} G(\mathbf{R}', \mathbf{R}; t) = (\hat{H} - E_T) G(\mathbf{R}', \mathbf{R}; t) \tag{3.36}$$

with the initial condition

$$G(\mathbf{R}', \mathbf{R}; 0) = \delta(\mathbf{R}' - \mathbf{R}) \ . \tag{3.37}$$

It can easily be seen that

$$G(\mathbf{R}', \mathbf{R}; \tau) = \langle \mathbf{R}' | \exp[-\tau(\hat{H} - E_T)] | \mathbf{R} \rangle \ . \tag{3.38}$$

Therefore, for any arbitrary wavefunction Ψ, we can apply the operator $\exp[-\tau(\hat{H} - E_T)]$ to propagate imaginary time. If we now expand in the basis of energy eigenfunctions, by inserting $\sum_i |\phi_i\rangle\langle\phi_i|$ into Eq. 3.38, we can show as above that this projects out the ground state:

$$\begin{aligned}
\lim_{\tau\to\infty} \exp[-\tau(\hat{H} - E_T)]\Psi(\mathbf{R}) &= \lim_{\tau\to\infty} \int G(\mathbf{R}', \mathbf{R}; \tau)\Psi(\mathbf{R}')\mathrm{d}\mathbf{R}' \\
&= \lim_{\tau\to\infty} \int \sum_i \mathrm{e}^{-\tau(E_i - E_T)} \phi_i^*(\mathbf{R}')\phi_i(\mathbf{R})\Psi(\mathbf{R}')\mathrm{d}\mathbf{R}' \\
&= \lim_{\tau\to\infty} \mathrm{e}^{-\tau(E_0 - E_T)} \langle \phi_0 | \Psi \rangle \phi_0(\mathbf{R}) \ .
\end{aligned}$$

We have thus turned the problem of finding the ground state into one of finding the Green function G, assuming we have some way of representing the wavefunction $\Psi(\mathbf{R})$ in such a way that we can apply G to it.

3.3.2 Walkers and Diffusion

The solution to the problem of representing the wavefunction becomes apparent if we consider the simplest possible Hamiltonian: that of N non-interacting particles with no potential energy. The Hamiltonian $\hat{H}$ is then simply the sum of the kinetic energy operators and the ITSE is

$$-\frac{\partial}{\partial t}\Phi(\mathbf{R},t) = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2\Phi(\mathbf{R},t) \ . \tag{3.39}$$

We see at once that this is equivalent to the well known diffusion equation in a $3N$ dimensional space, for which the Green function is a Gaussian peak spreading with time:

$$G(\mathbf{R},\mathbf{R}';\tau) = (2\pi\tau)^{-3N/2}\exp\left[\frac{-(\mathbf{R}-\mathbf{R}')^2}{2\tau}\right] \ . \tag{3.40}$$

In the classical diffusion of liquids or gases, the diffusing substance (be it liquid, gas, solid, heat or otherwise) can be equally well viewed either as a density field or as a large collection of individual particles. This latter way is significant here as it suggests a means of representing our wavefunction Φ as a population of "particles" each undergoing stochastic diffusion. Note that each of these particles, referred to as *walkers* and represented by a $3N$-dimensional position vector $\mathbf{R}_k$, represents an entire configuration of all the N electrons in the system, not just a single electron. Their diffusion is through the $3N$-dimensional configuration space, and one move may represent anything from a single electron moving in one dimension, to the whole set of N electrons moving at once.

These walkers are evolved in time using the Green function as a transition probability density. The application of the Green function describes this propagation as follows: a wavefunction $\Phi(\mathbf{R},t)$ is sampled by a set of M walkers at configuration-space positions $\mathbf{R}_k$:

$$\Phi(\mathbf{R},t) \rightarrow \sum_{k=1}^{M}\delta(\mathbf{R}-\mathbf{R}_k) \ , \tag{3.41}$$

where we use the symbol "→" to represent the correspondence between a collection of walker positions and a wavefunction. After time $\Delta\tau$, this becomes

$$\Phi(\mathbf{R}, t + \Delta\tau) \to \sum_{k=1}^{M} G(\mathbf{R}, \mathbf{R}_k; \Delta\tau) \; . \tag{3.42}$$

By comparison with Eq. 3.40, each of these can be seen to describe a set of N Gaussians of variance $3N\tau$. From these Gaussians, we re-sample a new set of discrete points $\mathbf{R}_k$. G is independent of the initial time t, depending only on the size of the timestep $\Delta\tau$, so we can iterate the same procedure repeatedly. After a sufficient number of iterations, the $\tau \to \infty$ limit will have been reached and the positions $\mathbf{R}_k$ sample the ground state in a statistical sense.

Unfortunately, a Hamiltonian containing only kinetic energy terms is of minimal use. To obtain a more useful Green function we must include a potential $U(\mathbf{R})$ which is a general function of all the particle coordinates, including interactions between the particles and interactions with an external potential. As there is no longer a simple expression for the Green function for general $U(\mathbf{R})$, an approximation to $\exp[-\Delta\tau(\hat{T} + \hat{U} - E_T)]$ for small timestep $\Delta\tau$ is the only way to proceed. We make use of the Trotter formula for exponentials of sums of operators:

$$\mathrm{e}^{-\Delta\tau(\hat{T}+\hat{U})} \simeq \mathrm{e}^{-\Delta\tau\hat{U}/2}\mathrm{e}^{-\Delta\tau\hat{T}}\mathrm{e}^{-\Delta\tau\hat{U}/2} \; , \tag{3.43}$$

and the Green function becomes, to order $(\Delta\tau)^3$,

$$G(\mathbf{R}, \mathbf{R}'; \Delta\tau) = C \exp\left[\frac{-(\mathbf{R} - \mathbf{R}')^2}{2\Delta\tau}\right] \exp\left[-\Delta\tau(U(\mathbf{R}) + U(\mathbf{R}') - 2E_T)/2\right] \; , \tag{3.44}$$

where C is a constant.

The trick that allows us to continue using “diffusing” walkers is to treat the new potential term as a reweighting of the Green function. As each walker now moves around the configuration space, it gains or loses weight w_k according to whether it is in a region of low or high potential energy, respectively: spending time $\Delta\tau$ at position $\mathbf{R}$ causes the walker’s weight to be multiplied by $\exp[-\Delta\tau U(\mathbf{R})]$.

The wavefunction Ψ is now no longer represented by the distribution of walkers alone, but by the distribution of total walker weight. The algorithm just described is not, however, stable in a real calculation: the ensemble of walkers rapidly ends up dominated by the single walker that has spent the longest time in favourable regions where $E_L < E_T$, and picked up an exponentially large weight in the process. This is clearly a desperately inefficient way to proceed.

The further ingredient that makes this approach feasible is to use a *branching* algorithm to incorporate proliferation and decay of the walkers as they explore the configuration space. The reweighting factor P, where

$$P = \exp\left[-\Delta\tau(U(\mathbf{R}) + U(\mathbf{R}') - 2E_T)/2\right] \tag{3.45}$$

is reinterpreted as specifying the mean number of (unreweighted) offspring of the current walker at the next time step:

i) If $P < 1$ the walker continues with probability P, otherwise it is removed.

ii) If $P > 1$ the walker continues and spawns, on average, $P - 1$ copies of itself at the same position in configuration space.

The following formula for the number of walkers M_{new} present after the end of a timestep incorporates both of these possibilities:

$$M_{new} = \mathrm{INT}(P + \eta)\ , \tag{3.46}$$

where η is a random number from a uniform distribution between 0 and 1. In this scheme, E_T plays a more significant role, as it is used to keep the population stable. By dynamically adjusting E_T the mean number of walkers is kept approximately constant around some predetermined value. If E_T was not adjusted, walkers either die out or proliferate catastrophically in regions of low potential energy until the available computational memory is exceeded.

One algorithm that can be used to adjust E_T is expressed in the formula

$$E_T(\tau) = E_{best}(\tau) - \frac{1}{g\tau} \ln \frac{M(\tau)}{M_0} \tag{3.47}$$

where $E_{best}(\tau)$ is the current best estimate of the ground-state energy, $g\tau$ is some multiple of the timestep, and $M(\tau)$ and M_0 are the current and target total weights or total populations.

Given that the walkers will spend equal amounts of time in regions where the energy is above and below the mean, E_T likewise spends equal time above and below E_0; it can thus be shown that the average value over a sufficiently long run will be equal to the ground state energy E_0. This 'growth estimator' can be used as a way of extracting the energy, but the so called *mixed estimator* that will be described in the next section turns out to be numerically better behaved.

3.3.3 Importance Sampling

As it stands, this algorithm is already capable in principle of evaluating the ground state energy for a given Hamiltonian. However, a little thought as to the behaviour of real wavefunctions shows that it very rapidly becomes impossibly inefficient as the number of electrons increases. The dimensionality of the configuration space to be explored is $3N$ and although the space of favourable configurations grows quickly with N, the space of unfavourable configurations grows vastly much faster. Left undirected, then, the diffusion will cause the walkers to wander off to regions where the amplitude of the wavefunction $\Phi(\mathbf{R})$ is very low, which contribute negligibly to the integral for the energy. We would like, therefore, to be able to guide the diffusion process towards areas of high amplitude of the wavefunction — effectively turning it into a drift-diffusion process. However, we do not know beforehand where these areas will be for the unknown ground state Φ_0.

We can overcome this by introducing a trial wavefunction Ψ_T to bias the process towards areas of high amplitude, at the cost of no longer sampling configurations from the true ground state wavefunction Φ_0. We first note the following identity:

$$\begin{aligned}
\frac{1}{2}\nabla^2(\Psi\Phi) &= \frac{1}{2}\Psi\nabla^2\Phi + \frac{1}{2}\Phi\nabla^2\Psi + \nabla\Phi.\nabla\Psi \\
&= \frac{1}{2}\Psi\nabla^2\Phi - \frac{1}{2}\Phi\nabla^2\Psi + \nabla.(\Phi\nabla\Psi) \\
\Rightarrow \frac{1}{2}\Psi\nabla^2\Phi &= \frac{1}{2}\nabla^2(\Psi\Phi) + \frac{1}{2}\Phi\Psi\frac{\nabla^2\Psi}{\Psi} - \nabla.\left(\Phi\Psi\frac{\nabla\Psi}{\Psi}\right)
\end{aligned}$$

which allows us to re-write the ITSE (Eq. 3.31) multiplied by a trial function Ψ_T

$$\Psi_T\frac{\partial\Phi}{\partial t} = -\Psi_T(-\frac{1}{2}\nabla^2\Phi + U(\mathbf{R})\Phi - E_T\Phi) \tag{3.48}$$

in terms of a new variable $f(\mathbf{R})$, where $f(\mathbf{R}) = \Psi_T(\mathbf{R})\Phi(\mathbf{R})$,

$$\begin{aligned}
\frac{\partial f}{\partial \tau} &= \frac{1}{2}\nabla^2 f - \nabla.(f\nabla\ln|\Psi_T|) + \frac{1}{2}f\frac{\nabla^2\Psi}{\Psi} - fU(\mathbf{R}) + E_T f \\
&= \frac{1}{2}\nabla^2 f - \nabla.(\mathbf{v}f) + (E_T - E_L)f \ , \qquad (3.49)
\end{aligned}$$

where $\mathbf{v}$ is the drift velocity, defined by

$$\mathbf{v} = \frac{1}{2}\nabla\ln(|\Psi_T(\mathbf{R})|^2) \ , \tag{3.50}$$

and $E_L(\mathbf{R}) = \hat{H}\Psi_T(\mathbf{R})/\Psi_T(\mathbf{R})$ is the local energy of the trial function at $\mathbf{R}$.

Eq. 3.49 corresponds to a new integral equation with a modified Green function:

$$f(\mathbf{R}', t+\tau) = \int G^*(\mathbf{R},\mathbf{R}';\tau)f(\mathbf{R},t)\mathrm{d}\mathbf{R} \ , \tag{3.51}$$

where $G^*(\mathbf{R},\mathbf{R}';\tau) = \Psi_T(\mathbf{R})G(\mathbf{R},\mathbf{R}';\tau)/\Psi_T(\mathbf{R}')$. Once again, we make a short timestep approximation to obtain, after considerable algebra,

$$\begin{aligned}
G^*(\mathbf{R},\mathbf{R}';\Delta\tau) &\simeq (2\pi\Delta\tau)^{-3N/2}\exp\left[-\frac{[\mathbf{R}-\mathbf{R}'-\Delta\tau\nabla\ln|\Psi_T|^2]^2}{2\Delta\tau}\right]\times \qquad (3.52) \\
&\quad \exp[-\Delta\tau(E_L(\mathbf{R}') + E_L(\mathbf{R}) - 2E_T)/2] \ .
\end{aligned}$$

There are several important differences to note about applying this equation to the construction of a DMC algorithm. Firstly, we are no longer simulating merely diffusion with varyingly weighted walkers: the process is now more akin to a *drift-diffusion* process, with $\mathbf{v}(\mathbf{R})$ as a drift velocity. The particles no longer diffuse around the configuration space at random but are guided by this drift velocity into regions of large Ψ_T. Secondly, without importance sampling we were treating Φ_0 as our distribution. Now, it is not $|\Phi_0|^2$ but $f = \Psi_T \Phi_0$ according to which the configurations are distributed in the limit of large τ. This has a number of implications for estimation of operators, but, significantly, not for those that commute with the Hamiltonian. For the total energy, for example, the true ground state energy can still be sampled by evaluating $E_L(\mathbf{R})$:

$$\begin{aligned} E_0 &= \frac{\left\langle \Psi_T | \hat{H} | \Phi_0 \right\rangle}{\langle \Psi_T | \Phi_0 \rangle} \\ &= \frac{\left\langle \Phi_0 | \hat{H} | \Psi_T \right\rangle}{\langle \Phi_0 | \Psi_T \rangle} \\ &= \frac{\int f(\mathbf{R})[\hat{H}\Psi_T/\Psi_T] \mathrm{d}\mathbf{R}}{\int f(\mathbf{R}) \mathrm{d}\mathbf{R}} \\ &= \lim_{M \to \infty} \frac{1}{M} \sum_m \frac{\hat{H}\Psi_T(\mathbf{R}_m)}{\Psi_T(\mathbf{R}_m)} \ . \end{aligned} \tag{3.53}$$

Finally, the function controlling population growth of the walkers is no longer just the potential energy $U(\mathbf{R})$ but is instead the full local energy $E_L(\mathbf{R})$, which varies much less strongly as a function of electron positions than the potential, especially with a good trial function Ψ_T which is close to an energy eigenstate.

In practical codes, a number of further modifications to the Green function are made, either to increase efficiency or to remove potential sources of bias. The scheme implemented in the code used in this research is that of Umrigar, Nightingale and Runge [167]. The most important further step is to introduce a Metropolis accept/reject step to ensure that the algorithm obeys detailed balance. This works by only accepting trial moves $\mathbf{R} \rightarrow \mathbf{R}'$ with a probability

$$A(\mathbf{R}' \leftarrow \mathbf{R}) = \mathrm{Min}\left(1, \frac{G^*(\mathbf{R}, \mathbf{R}'; \Delta\tau)|\Psi_T(\mathbf{R}')|^2}{G^*(\mathbf{R}', \mathbf{R}; \Delta\tau)|\Psi_T(\mathbf{R})|^2}\right)$$

or rejecting them otherwise. This only has an effect when $\Delta\tau$ is finite and serves to reduce the timestep error, and ensures that in the limit of perfect importance sampling, when $\Psi_T = \Phi_0$, the result can be exact for finite timestep.

There are also significant changes made to the Green function in the vicinity of a bare nucleus, designed to limit the drift vector to remove its divergence there, and to prevent electrons overshooting nuclei. Finally, one can make modifications to prevent bias due to population control, as the above algorithm is only exact in the limit of very large populations.

These improvements, combined with the much lower variance of the local energy and the greatly-increased efficiency of sampling in areas of greater significance to the energy, result in an increase in efficiency of many orders of magnitude. This is what makes importance-sampled DMC feasible even for large numbers of electrons $N \sim 1000$.

3.3.4 The Fixed Node Approximation

One very important assumption implicit in the use of diffusive walkers to represent a wavefunction has been neglected so far. While a density of walkers or walker weights can only be a positive quantity, the wavefunction of a system of fermions must, in order to obey antisymmetry, have both negative regions and positive regions. If left to evolve unconstrainedly, the algorithm described so far will cause the wavefunction to evolve towards the *bosonic* ground state, which is symmetric under exchange of electrons, even though the trial function is antisymmetric. The most obvious method to solve this would be to allow distinct positive and negative walkers diffusing around separately, with the value of the wavefunction being the difference in weighted densities, but in this case the positive and negative populations collapse individually to the bosonic ground state and the difference in weighted densities tends exponentially to zero and nothing has been achieved.

A number of methods exist for attempting to solve this problem, which is one instance of the wider 'Fermion Sign Problem' that plagues a variety of methods in quantum mechanics. The principal method is to perform release node calculations [31], in which, starting from the fixed node distribution, the behaviour of the walkers is changed such that they are allowed to cross the nodes of the wavefunction, but each time they do so they pick up a sign-change that means they contribute to statistical averages with the opposite sign. If the relaxation time from the fixed-node distribution to the bosonic ground state times the energy difference between the fermionic and bosonic ground states is less than unity, then a reliable fermionic ground state energy can be extracted. However, this is rarely the case in practice, and nevertheless the calculations are computationally extremely demanding, so the method is rarely used in anything except the homogeneous electron gas.

The most commonly used method, therefore, is to use what is known as the Fixed Node Approximation. Here, the function $f(\mathbf{R}) = \Psi_T(\mathbf{R})\Phi_0(\mathbf{R})$ is forced to remain non-negative everywhere by rejecting any moves in which a walker crosses a node of the trial function. This is easy to test for as if a node as been crossed, $\Psi_T(\mathbf{R}')$ will have opposite sign to $\Psi_T(\mathbf{R})$. When this restriction is enforced, $\Phi_0(\mathbf{R})$ automatically has the same nodes as $\Psi_T(\mathbf{R})$. We then refer to $\Phi_0(\mathbf{R})$ as the fixed-node ground state, sometimes written $\Phi_0^{FN}(\mathbf{R})$ to distinguish it from the true ground state (the label is usually dropped in cases where we have no means of obtaining or distinguishing this true ground state). The corresponding energy E_0^{FN} can be shown to be variational [123, 145], in that $E_0^{FN} \geq E_0$ where E_0 is the energy of the true ground state.

In this approximation, walkers are confined to the nodal pocket in which they start. If all the nodal pockets are of the same energy, this will not present a problem as the so-called "Tiling theorem" [30, 139] ensures that all the nodal pockets of the ground state of a Hamiltonian are all equivalent by permutation symmetry. Since we are always working with nodal surfaces from wavefunctions which are the ground states of some Hamiltonian (be it the Kohn-Sham or Hartree-Fock or whatever method we used for the generation of the orbitals), this always holds in practice.

3.3.5 Operator Estimators in DMC

In variational Monte Carlo, the distribution of the sampled points $\mathbf{R}_k$ in configuration space is a stochastic representation of the probability distribution of the trial wavefunction, $|\Psi_T(\mathbf{R})|^2$. Operator expectation values are therefore evaluated by taking the mean over values of the local operator at each step. In DMC, however, the sampled points are distributed according to $f(\mathbf{R}) = \Psi_T(\mathbf{R})\Phi_0(\mathbf{R})$, which complicates the evaluation of some expectation values.

For operators that commute with the Hamiltonian, there is no problem. By manipulations equivalent to Eq. 3.53 one can show that evaluating the so called "mixed" estimator

$$O_{mixed} = \sum_k O_L(\mathbf{R}_k) \tag{3.54}$$

over a set of configuration-space positions sampling $f(\mathbf{R})$ yields exactly the same expectation value as if we had sampled $|\Phi_0(\mathbf{R})|^2$. Mathematically, this follows because

$$O_{mixed} = \frac{\int \mathrm{d}\mathbf{R}\, \Phi_0(\mathbf{R})\hat{O}\Psi_T(\mathbf{R})}{\int \mathrm{d}\mathbf{R}\, \Phi_0(\mathbf{R})\Psi_T(\mathbf{R})} = \frac{\int \mathrm{d}\mathbf{R}\, \Phi_0\hat{O}\Phi_0}{\int \mathrm{d}\mathbf{R}\, \Phi_0\Phi_0} \,. \tag{3.55}$$

in these cases. Evaluating the mixed estimator is still useful even when the operator does not commute with the Hamiltonian, especially if the wavefunction is a good approximation to the true ground state. Expectation values of operators such as Fourier components of the density are evaluated in this way. However, since inevitably Ψ_T differs from Φ_0, the mixed estimator differs from the ground state value O_0, with an error linear in the error in the trial wavefunction.

In these cases, it is sometimes an improvement to use the so-called "extrapolated" estimator. Writing

$$\Psi_T = \Phi_0 + \Delta\Psi \tag{3.56}$$

and noting that the VMC expectation value O_{VMC} is

$$O_{VMC} = \frac{\int \mathrm{d}\mathbf{R}\, \Psi_T(\mathbf{R})\hat{O}\Psi_T(\mathbf{R})}{\int \mathrm{d}\mathbf{R}\, \Psi_T(\mathbf{R})\Psi_T(\mathbf{R})} \ , \tag{3.57}$$

we note that if we were to take

$$O_{ext} = 2O_{mixed} - O_{VMC} \ , \tag{3.58}$$

then the term linear in $\Delta\Psi$ cancels and we are left with

$$\begin{aligned} O_{ext} &= \frac{\int \mathrm{d}\mathbf{R}\, \Phi_0(\mathbf{R})\hat{O}\Phi_0(\mathbf{R})}{\int \mathrm{d}\mathbf{R}\, \Phi_0(\mathbf{R})\Phi_0(\mathbf{R})} + \mathcal{O}\left[(\Delta\Psi)^2\right] \\ &= O_0 + \mathcal{O}\left[(\Delta\Psi)^2\right] . \end{aligned}$$

This works reliably if the trial wavefunction is highly accurate, but is less useful otherwise as the error can fluctuate wildly.

Finally, a variety of schemes have been proposed to attempt to access the "pure" distribution $|\Phi_0(\mathbf{R})|^2$, and thus to evaluate "pure" estimators $\langle\Phi_0|\hat{O}|\Phi_0\rangle$. The most well established of these is called forward (or future) walking [14], which projects the walker through imaginary time, to find the set of weights $w(\mathbf{R}) = \Phi_0(\mathbf{R})/\Psi_T(\mathbf{R})$ required to change the walker density obtained in DMC, $\Psi_T(\mathbf{R})\Phi_0(\mathbf{R})$, into the exact ground state. Once these are known, the exact ground state expectation value can be obtained by simply multiplying the local operator $O_L(\mathbf{R})$ by these weights and normalizing. Future walking does, however, suffer from serious stability issues and scales very poorly with system size. The same is true of the pure distribution DMC method of Caffarel [21, 22].

The Hellman-Feynman theorem has also been applied to the DMC algorithm as a whole in an attempt to sample pure estimators [59]. It may be possible to think of this approach as an optimized approach to future walking without many of the drawbacks of the original formulation.

3.3.6 Pseudopotentials in DMC

One major problem remains in our method — that of evaluating energies of systems of atoms described using nonlocal pseudopotentials. In our discussion of the local energy thus far we have assumed that it could be expressed in a local form, as

$$\frac{\hat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})} = E_L(\mathbf{R}) \tag{3.59}$$

where $E_L(\mathbf{R})$ is simply a function of the electron coordinates $\mathbf{R}$. If there are nonlocal pseudopotentials present, however, this is no longer possible as the action of the the nonlocal projectors on the fixed node ground state, $\hat{V}_{nl} = \sum_i \frac{\hat{V}_{nl,i}\Phi_0}{\Phi_0}$, is unknown as we do not know an analytic form for Φ_0.

The most common scheme used to circumvent this problem is known as the pseudopotential localization approximation [121]. This replaces the true Hamiltonian with an effective Hamiltonian, using the trial wavefunction instead of the true ground state:

$$H_{eff} = K + V_{loc} + \frac{\int \mathrm{d}\mathbf{R}' \, \langle \mathbf{R} | \, \hat{V}_{nl} \, | \mathbf{R}' \rangle \, \Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} \tag{3.60}$$

where K is the kinetic energy operator and V_{loc} is the local part of the potential energy. Propagating this Hamiltonian through imaginary time yields the distribution $f_{eff}(\mathbf{R}) = \Psi_{eff}(\mathbf{R})\Psi_T(\mathbf{R})$ where $\Psi_{eff}(\mathbf{R})$, which has taken the role of $\Phi_0(\mathbf{R})$, is the fixed node ground state of the effective Hamiltonian. The last term in the effective Hamiltonian above is a local function of $\mathbf{R}$, but this has come at the cost of making the total energy no longer a variational estimate of the ground state energy. To see this, we note that for the trial wavefunction, $H_{eff}\Psi_T/\Psi_T = H\Psi_T/\Psi_T$, so

$$E_{eff} = \frac{\langle \Psi_{eff} | H_{eff} | \Psi_T \rangle}{\langle \Psi_{eff} | \Psi_T \rangle} = \frac{\langle \Psi_{eff} | H | \Psi_T \rangle}{\langle \Psi_{eff} | \Psi_T \rangle} \tag{3.61}$$

but this is not equal to the variational fixed node energy of the true Hamiltonian H

acting on Ψ_{eff}, which would be

$$E_{FN} = \frac{\langle \Psi_{eff} | H | \Psi_{eff} \rangle}{\langle \Psi_{eff} | \Psi_{eff} \rangle} \,, \tag{3.62}$$

because H and H_{eff} differ. We therefore no longer have a variational principle on E_{eff}, which may now be above or below the true ground state. In practice, this is found to be a adequate approximation as long as the trial wavefunction is good [121].

An alternative approach, the recently-proposed Casula scheme [24], modifies the Green function and the details of the algorithm in a variety of ways, with the result that the effective ground state energy E_{eff} is variational on the true ground state, so that $E_{eff} \geq E_0$. Indeed, in general, the ground state energy of the effective Hamiltonian *without* this scheme seems to come out usually slightly lower than with it. This improvement in accuracy comes at the cost of a potentially increased timestep-dependence, necessitating smaller timesteps and thus slower convergence of error bars.

3.3.7 Full Algorithm

We now have all the prerequisites to write down a practical DMC algorithm:

1. Generate a set of configurations $\mathbf{R}_k$ with $k = \{1, \ldots, N_{config}\}$ distributed according to $|\Psi_T|^2$ using variational Monte Carlo. These configurations are used as initial positions for the DMC walkers.

2. Propagate each walker by a suitably-chosen timestep $\Delta\tau$:

$$\mathbf{R}' = \mathbf{R} + \Delta\tau \boldsymbol{\chi} + \Delta\tau \nabla \ln |\Psi| \tag{3.63}$$

   where $\boldsymbol{\chi}$ is a normally-distributed $3N$-dimensional vector of unit variance.

3. Impose the node-crossing rejection step and the detailed balance rejection step.

4. Branching and decay of walkers is implemented by replacing a walker with

$$M_{new} = \mathrm{INT}(\exp[-\Delta\tau(E_L(\mathbf{R}) + E_L(\mathbf{R}') - 2E_T)/2] + \eta) \tag{3.64}$$

copies of itself, where η is a uniformly-distributed random number between 0 and 1.

5. The trial energy E_T is adjusted by $E_T(\tau) = E_{best}(\tau) - \frac{1}{g\Delta\tau} \ln \frac{M(\tau)}{M_0}$ to keep walker population roughly constant.

6. If a sufficient number of equilibration steps have already taken place, accumulate the current mean of the local energy over the walkers, $\bar{E_L}$. Also accumulate the local values of any other operators required, for later statistical processing.

7. Return to step 2 and iterate until sufficient data has been acquired.

This roughly describes the process in CASINO [125], the code used for all the QMC work in this thesis.

3.3.8 Sources of Error in Diffusion Monte Carlo

As previously stated, the goal of diffusion Monte Carlo is to find the exact ground state of a given Hamiltonian. However, in reality, because of approximations that must be made for the practical implementation of the algorithm or the evaluation of the action of Hamiltonian on the wavefunction, there are inevitable sources of error affecting its results beyond the mere statistical noise. While we have touched on many of these already in this section, and will return to some of them in the course of the chapters to come on specific applications, we will summarize the significant sources of error briefly now.

Timestep Error

Timestep error occurs because of the approximation made in Eq. 3.52 that the local energy and drift velocity terms $E_L(\mathbf{R})$ and $\mathbf{v}(\mathbf{R})$ in the Green function are a smoothly varying function of the electron positions, and thus that $E_L(\mathbf{R}) \simeq E_L(\mathbf{R}')$ and $\mathbf{v}(\mathbf{R}) \simeq \mathbf{v}(\mathbf{R}')$. It is possible to show that the error introduced by the timestep error is of order $(\Delta\tau)^2$ and is thus easily controllable. However, the error becomes notably worse in certain situations. When the wavefunction is a rapidly varying function of electron position, as is the case, for example, when an electron is close to a bare nucleus, then the prefactor increases sharply. Reducing the timestep reduces the error, but this must always be balanced against the increased serial correlation and thus the longer runtimes required to achieve a given error bar; it is therefore important to try and minimize such effects. For this reason, it is rare to attempt calculations in DMC with bare nuclei even if using a basis capable of accurately representing the orbitals of the core electrons. The pseudopotential approximation aids in this regard, although it introduces an error of its own due to pseudopotential localization. The control of timestep bias is discussed further in Section 6.3.3.

Fixed Node Error

The error resulting from the application of the fixed node approximation has already been discussed, as have the techniques for overcoming it. The variational inequality $E_0^{FN} \geq E_0$ ensures that in the absense of pseudopotentials we always find a higher energy than the true ground state, so in some situations the success of our simulation can be judged by the fraction of the total energy that is recovered. When highly accurate quantum chemistry calculations are available to give the true ground state energy, we can test the accuracy of the nodes by determining the fraction of the correlation energy recovered by fixed node DMC for a particular trial wavefuncion. In modern simulations with backflow and multideterminant trial wavefunctions, this can be upwards of 99%, even for difficult systems like full-core atoms beyond the first row.

The nodal surface is an $(Nd-1)$ dimensional object, but the only certain piece of information we have on it *a priori* is that the wavefunction is zero when two electrons coincide. This only defines a $(N-1)d$ dimensional system. For $d=1$, a one dimensional system, $Nd - 1$ and $(N - 1)d$ are the same, so the nodes are exact and highly accurate results can therefore be obtained (eg [42]). However, for all other systems, the coincidence points constrain the nodal surface somewhat but do not define it. The fixed node approximation then introduces what is generally thought of as an uncontrolled approximation. In Section 5.2.5 we present a novel way of estimating the magnitude of the error for a particular system.

Finite Size Effects

Finite size effects result from the use of the supercell approximation, in which one attempts to model an infinite (or at least, effectively infinite) system with underlying periodicity, by a supercell containing a finite number of repetitions of this periodic part. Finite size errors can take a number of forms, and indeed represent the major challenge to be overcome in several parts of this thesis, although in different applications they manifest themselves in different ways.

The simplest kind of such errors, the 'single-particle' finite size effects, are the result of replacing the smooth density of states of an infinite system by the discrete states of a finite system. Consider, for example, an electron gas, with single-electron orbitals which are simply plane waves $\mathrm{e}^{\imath \mathbf{k}.\mathbf{r}}$. In a finite system we are limited to **k**-points commensurate with the supercell we have chosen. For a cubic supercell of side length L_α, we require $k_\alpha = 2\pi/L_\alpha$. For the finite system, we sample the Brillouin zone with a finite sum $\sum_{\mathbf{k}}$ over these **k**-points. The sum approximates the integral $\int \mathrm{d}\mathbf{k}$ but inevitably, as the sampling is discrete, some sort of error is made. In the small supercells we are often limited to in QMC, this error can be large. It is often possible to improve the sampling merely by offsetting the entire grid from $\mathbf{k} = 0$ or by averaging over different offsets. This is discussed further in the discussion on boundary conditions for periodic many body systems in Section 4.2.

It is often adequate to assume that the errors due to the **k** quantization are roughly the same in DFT as they are in QMC. If this is the case, the QMC energy for an infinite system may be obtained from the QMC energy for a system of size L

by adding corrections obtained within DFT:

$$E_T^{QMC}(\infty) - E_T^{QMC}(L) = E_T^{DFT}(\infty) - E_T^{DFT}(L)\ .$$

We can thus correct the QMC energies by simply adding the difference between a very large system in DFT (or a truly infinite one where this is possible analytically) and one of the largest size we can manage in QMC.

Other types of finite size error have subtler effects on the energies, and are unique to truly many-body methods. These many-body errors can affect both the kinetic and the potential energies, and have only recently begun to be addressed theoretically. They result from a number of effects due to the constraints imposed by the size of the supercell. In the kinetic energy, the fact that electrons see, beyond a certain distance, only periodic replicas of themselves and their nearest neighbours, means that a certain contribution to the kinetic energy due to long-range correlation is lost. In the potential energy, in which the Ewald method is traditionally used to model the interaction of periodic copies of the electrons in the supercell, the fact that the correlation hole has to be ‘squashed’ into a single supercell affects the calculated exchange-correlation energy. Both effects can be dealt with using similar methods, as addressed in a recent paper by Chiesa et al. [33]. We discuss the application of these ideas to more complex systems in Section 5.2.6.

Chapter 4

Localization Lengths

4.1 Polarization and Localization

The properties that, on a macroscopic level, distinguish between metals and insulators are dc conductivity and static polarization: an insulator has vanishing conductivity at low frequency and low temperature, and sustains static polarization in an applied electric field. Two conditions on a system necessary for dc conductivity are that there must be excitations at vanishingly small energies above the ground state, and that these must represent charge distributions delocalized throughout the system.

The traditional band theory based approach, based on a Slater determinant of one-electron orbitals, explains metallic behaviour in terms of the accessibility of low-lying single particle excitations and the question of whether a particular wavefunction represents a metal or an insulator is easily answered: if the density of one-electron states is finite at the Fermi level, the system is a metal; if not, it is an insulator.

An alternative method is to consider whether the electrons in the system are free to move or if they are localized. A suitable unitary transformation of the determinant of Bloch functions of an insulator gives an equivalent determinant of exponentially localized one-electron Wannier functions [18]. The exponential localization of the Wannier functions implies a corresponding exponential localization of the one-electron density matrix $\rho(\mathbf{r}, \mathbf{r}')$ as a function of $\mathbf{r} - \mathbf{r}'$, and it has been shown [148] that any system for which the second moment of $\rho(\mathbf{r}, \mathbf{r}')$ is finite must be insulating.

In the one-electron theory of disordered systems, the energy eigenfunctions may be localized and a finite density of states at the Fermi level no longer implies conducting behaviour, but the more general idea that insulating behaviour results from wavefunction localization survives. The connection between the insulating state and localization of electron states was originally established by Kohn in a milestone paper in 1964 [92], where he showed that the ground state many-electron wavefunction of an insulator displays localization in a specific way, referred to as disconnectedness. The wavefunction of the extended system is referred to as *disconnected* if it can be expressed as a sum of functions $\Psi_{\mathbf{M}}$ which are individually localized in disconnected regions of the $3N$ dimensional hyperspace defined by the N electron coordinates, with an overlap $\langle \Psi_{\mathbf{M}} | \Psi_{\mathbf{M}'} \rangle$ which vanishes exponentially with system size. Kohn argued that *all* metal to insulator transitions, even those in strongly correlated interacting solids, are accompanied by this form of localization of the many-electron wavefunction.

What has become known as the 'modern theory' of polarization and localization has its origins in King-Smith and Vanderbilt's 1993 paper, where they proposed a method for the calculation of polarization changes in extended solids [85], which recast the problematic Kubo formula approach in terms of ground state properties only, and established the link to Wannier functions. Marzari and Vanderbilt later advanced the understanding of this link in a 1997 paper [117] introducing maximally localized Wannier functions. In 1999, Resta and Sorella extended the formalism to discuss localization in many-electron systems [144], and a further breakthrough was the rigorous mathematical description of polarization and localization in terms of probability generating functions by Souza, Wilkens and Martin in 2000 [156]. A review by Resta [141] summarizes the development of the theory up to 2002.

There have been a handful of calculations of localization lengths [148, 172], and many calculations of quantities depending on polarization [151, 12]. These have mostly been within single-particle approaches, principally density functional theory within the LDA, and little work exists on the potentially more interesting correlated-electron case. Among the exceptions are two QMC studies, one of phase transitions in the one-dimensional ionic Hubbard model [176], and one of the dielectric response of periodic systems [165]. In the latter, it was observed that the inclusion of many-body correlation effects dramatically affected the value of the polarizability and improved its convergence with system size by decreasing the localization length, a conclusion supported by calculations in which self-interaction corrections were included in calculations of Born effective charges [51].

In this chapter, we present a novel derivation of the known expressions for polarizations and localization lengths in terms of many body wavefunctions, discuss their evaluation within the frameworks of density functional theory and quantum Monte Carlo, and investigate their behaviour in model systems consisting of periodic arrays of quantum dots. These models, which combine inherent tunability with a simplified analogy to the behaviour of real atoms, should make an excellent testing ground for the application of these recent ideas within a fully many-body theory.

4.2 Boundary Conditions and Periodicity in Extended Systems

A crystal of a solid of macroscopic size contains of order 10^{24} atoms and has dimensions of order 10^8 times the lattice parameter defined by the primitive cell vectors $\mathbf{A}_i$ ($i = 1, \ldots, d$ where d is the spatial dimension). This scale difference underlies the simplifying assumption of taking a solid as being infinite in extent and using periodic boundary conditions. In the following, we briefly recount the consequences of this assumption in one electron theory, before exploring its implications for many-body wavefunctions of the sort used in quantum Monte Carlo calculations.

In the one-electron picture of a perfect crystal, Bloch's theorem provides us with a recipe for using this periodicity: the symmetry of the one-electron effective potential ensures that the single-particle eigenfunctions of the Hamiltonian are also eigenfunctions of lattice translations $\hat{T}_{\mathbf{R}}$ for all lattice vectors $\mathbf{R} = n_1\mathbf{A}_1 + \cdots + n_d\mathbf{A}_d$. The Hamiltonian obeys

$$\hat{H}(\mathbf{r}+\mathbf{R}) = \hat{H}(\mathbf{r}) \ , \tag{4.1}$$

and this eventually leads (see eg [16, 89]) to eigenstates of the Bloch form:

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \mathrm{e}^{i\mathbf{k}.\mathbf{r}} u_{i\mathbf{k}}(\mathbf{r}) \tag{4.2}$$

where $u_{i\mathbf{k}}$ is periodic in $\mathbf{r}$ such that $u_{i\mathbf{k}}(\mathbf{r}+\mathbf{R}) = u_{i\mathbf{k}}(\mathbf{r})$, and the Bloch wavevectors $\mathbf{k}$ may be chosen to lie in the first Brillouin zone. The Bloch functions obey $\psi_{i\mathbf{k}}(\mathbf{r}+\mathbf{R}) = \mathrm{e}^{i\mathbf{k}.\mathbf{R}}\psi_{i\mathbf{k}}(\mathbf{r})$, which is known as Bloch's theorem.

The individual Bloch functions are only determined by the Hamiltonian up to an arbitrary $\mathbf{k}$-dependent phase — the choice of phase as a function of $\mathbf{k}$ being referred to as a gauge. For simplicity we limit our consideration to gauges in which $\psi_{\mathbf{k}}$ is a differentiable function of $\mathbf{k}$ and $\psi_{i\mathbf{k}} = \psi_{i(\mathbf{k}+\mathbf{G})}$ for all reciprocal lattice vectors $\mathbf{G}$. This also ensures that $u_{i(\mathbf{k}+\mathbf{G})} = \mathrm{e}^{-i\mathbf{G}.\mathbf{r}} u_{i\mathbf{k}}$. Gauge transformations that preserve these properties take the form:

$$\psi_{\mathbf{k}} \rightarrow \mathrm{e}^{i\phi(\mathbf{k})}\psi_{\mathbf{k}} \tag{4.3}$$

where $\phi(\mathbf{k}) = \mathbf{k}.\mathbf{R} - \beta(\mathbf{k})$, with $\beta(\mathbf{k})$ being a periodic, differentiable function of $\mathbf{k}$.

Within one-electron theory, Bloch's theorem allows us to reduce the problem of solving the Schrodinger equation for an infinite periodic crystal to that of solving it for a single unit cell subject to "twisted" boundary conditions. By averaging over twists, which is achieved by integrating with respect to $\mathbf{k}$ over the first Brillouin zone, an exact result is obtained for the infinite periodic crystal.

In many electron theory, by contrast, the electron-electron interactions are treated explicitly rather than via an effective potential, and the coordinates of electrons in nearby cells are explicitly correlated. A wavefunction for the electrons in a single unit cell would be inadequate to describe a realistic system, so consider an N-electron simulation cell, preferably many times larger than the unit cell of the underlying crystal, with sides $\mathbf{A}_i, \ldots, \mathbf{A}_d$. Given any lattice vector $\mathbf{R}$ defined in terms of these simulation cell vectors, the Hamiltonian is invariant under translation of any single electron by this lattice vector:

$$\hat{H}(\mathbf{r}_1, \ldots, \mathbf{r}_i + \mathbf{R}, \ldots, \mathbf{r}_N) = \hat{H}(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N) \; . \tag{4.4}$$

This symmetry is not satisfied in the real system because the Coulomb interaction is not periodic, but under periodic boundary conditions the Coulomb interaction is replaced by the periodic Ewald interaction obtained by solving Poisson's equation subject to PBCs across the simulation cell and the above symmetry becomes exact. As the simulation cell is made larger, the finite-size errors resulting from this replacement diminish, and the properties of the artificially periodic system tend to those of a real crystal.

The effect of this symmetry is that the N-electron wavefunction Ψ can only change by a phase factor when any single electron is translated by a lattice vector $\mathbf{R}$. Moreover, because the electrons are all identical, the phase factor must be the same no matter which of the electron coordinates we choose to translate. This can be see by postulating that the phase factors are different for $\mathbf{r}_1$ and $\mathbf{r}_2$, giving two equations

$$\Psi(\mathbf{r}_1 + \mathbf{R}, \mathbf{r}_2, \ldots) = \mathrm{e}^{i\mathbf{k}_1 \cdot \mathbf{R}} \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) \tag{4.5}$$

and

$$\Psi(\mathbf{r}_1, \mathbf{r}_2 + \mathbf{R}, \ldots) = \mathrm{e}^{i\mathbf{k}_2 \cdot \mathbf{R}} \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) \; . \tag{4.6}$$

Applying the permutation operator to Eq. 4.5

$$\Psi(\mathbf{r}_2, \mathbf{r}_1 + \mathbf{R}, \ldots) = -\mathrm{e}^{i\mathbf{k}_1 \cdot \mathbf{R}} \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) \tag{4.7}$$

and then translating the second coordinate on the LHS by $-\mathbf{R}$ we get

$$\Psi(\mathbf{r}_2, \mathbf{r}_1, \dots) = -\mathrm{e}^{i(\mathbf{k}_1-\mathbf{k}_2).\mathbf{R}}\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots) \ . \tag{4.8}$$

Finally, permuting back gives

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots) = \mathrm{e}^{i(\mathbf{k}_1-\mathbf{k}_2).\mathbf{R}}\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots) \ , \tag{4.9}$$

which can only be fulfilled for $\mathbf{k}_1 - \mathbf{k}_2 \in \{\mathbf{G}\}$ where $\{\mathbf{G}\}$ is the set of reciprocal lattice vectors of the simulation cell. Because the phase factor $\mathrm{e}^{i\mathbf{k}.\mathbf{R}}$ is unchanged by translations of $\mathbf{k}$ by elements of $\{\mathbf{G}\}$, we can map $\mathbf{k}$ into the first Brillouin zone without any loss of generality, and we conclude that $\mathbf{k}_1 = \mathbf{k}_2 = \mathbf{k}$ for all the electrons. We thus have the Bloch condition on the many-body wavefunction $\Psi_{\mathbf{k}}$:

$$\Psi_{\mathbf{k}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathrm{e}^{i\mathbf{k}.\mathbf{X}}\Phi_{\mathbf{k}}(\mathbf{r}_1, \dots, \mathbf{r}_N) \tag{4.10}$$

where $\Phi_{\mathbf{k}}$ is periodic in all the electron coordinates over a simulation cell. Note the appearance of the so-called electronic dipole operator,

$$\mathbf{X} = \sum_{i=1}^{N} \mathbf{r}_i \ , \tag{4.11}$$

in the phase factor, playing the role of the electron coordinate. The same restrictions on gauge transformations $\phi(\mathbf{k})$ apply to $\Psi_{\mathbf{k}}$ as those described for Eq. 4.3.

Incidentally, there is also a second symmetry in a perfect crystal due to the periodic nature of the external potential. If one translates every electron coordinate simultaneously by a lattice vector of the *primitive* cell (not the simulation cell), the Hamiltonian again remains invariant. Although this symmetry imposes a second Bloch-like condition on the form of the wavefunction, it is not as useful as the above and will not be required here.

4.3 Localization Lengths in One-Electron Theories

A great deal of previous work has gone into developing the theory of polarization and localization in one-electron theories [85, 117, 148, 170]. As the focus of this work is on many-body methods, the detailed derivations of these results will not be reproduced here, especially as they are essentially a simplification of the many-body results of the next section. For simplicity we describe only the situation where $\mathbf{A}_i, \ldots, \mathbf{A}_d$ are orthogonal. The generalization to non-orthogonal simulation cells is straightforward but complicates the notation.

Restated in terminology closer to that used in the following section, the central result of King-Smith and Vanderbilt's paper [85] is that a change $\Delta\mathbf{P} = \mathbf{P}^{(1)} - \mathbf{P}^{(0)}$ in the polarization $\mathbf{P}^{\lambda}$ parameterized by some variable λ which varies between 0 and 1 can be expressed in terms of the periodic parts $u_{i\mathbf{k}}^{(\lambda)}$ of the n_b bands of one-electron Bloch functions as

$$P_\alpha^{(\lambda)} = \frac{iq_e}{(2\pi)^d} \sum_i^{n_b} \int \mathrm{d}\mathbf{k} \langle u_{i\mathbf{k}}^{(\lambda)} | \partial/\partial k_\alpha | u_{i\mathbf{k}}^{(\lambda)} \rangle \;, \tag{4.12}$$

where $\alpha = 1, \ldots, d$ is a Cartesian direction. For the set of Wannier functions defined by

$$w_i^{(\lambda)}(\mathbf{r} - \mathbf{R}) = \frac{1}{V_k} \int \mathrm{d}\mathbf{k}\, \mathrm{e}^{-i\mathbf{k}.\mathbf{R}}\, \psi_{i\mathbf{k}}^{(\lambda)}(\mathbf{r}) \;, \tag{4.13}$$

with $V_k = (2\pi)^d/V$ the volume of the Brillouin zone, the polarization $\mathbf{P}^{(\lambda)}$ is the first moment of the Wannier function centered on $\mathbf{R} = 0$:

$$\mathbf{P}^{(\lambda)} = \frac{q_e}{V} \sum_i^{n_b} \langle w_i^{(\lambda)} | \mathbf{r} | w_i^{(\lambda)} \rangle \;, \tag{4.14}$$

where q_e is the charge on an electron. It can be shown that gauge transformations among the occupied orbitals are only capable of changing this quantity by discrete jumps of whole quanta of polarization $q_e\mathbf{R}/V$.

The sum of the second moments of the occupied Wannier functions can be expressed in **k**-space in a very similar fashion. Marzari and Vanderbilt [117] showed that the quadratic spread Ω of the Wannier functions of the occupied Bloch orbitals

$$\Omega = \sum_i^{n_b} \left[\langle w_i | r^2 | w_i \rangle - |\langle w_i | \mathbf{r} | w_i \rangle|^2 \right] \tag{4.15}$$

can be decomposed into a gauge-dependent part $\tilde{\Omega}$ and a gauge-invariant part Ω_I, where

$$\begin{aligned} \Omega_I &= \frac{V}{(2\pi)^d} \sum_\alpha \int \mathrm{d}\mathbf{k} \Bigg(\sum_i^{n_b} \Big\langle \frac{\partial u_{i\mathbf{k}}}{\partial k_\alpha} \Big| \frac{\partial u_{i\mathbf{k}}}{\partial k_\alpha} \Big\rangle & (4.16) \\ & - \sum_i^{n_b} \sum_j^{n_b} \Big\langle \frac{\partial u_{i\mathbf{k}}}{\partial k_\alpha} \Big| u_{j\mathbf{k}} \Big\rangle \Big\langle u_{j\mathbf{k}} \Big| \frac{\partial u_{i\mathbf{k}}}{\partial k_\alpha} \Big\rangle \Bigg) \, , & (4.17) \end{aligned}$$

so the expectation value of the squared localization length, often written $\langle r^2 \rangle$ even though strictly it is not $\sum_i \langle \psi_i | r^2 | \psi_i \rangle$, can be obtained as $\langle r^2 \rangle = \Omega / N \geq \Omega_I / N$.

Finally, Resta [142] and later Resta and Sorella [144] showed that by considering the Slater determinant $\Psi_\mathbf{k}$ formed from the occupied one-electron orbitals, the polarization and localization can be expressed in terms of the quantity z_N^α, defined by

$$z_N^\alpha = \langle \Psi_{\mathbf{k}=0} | \mathrm{e}^{i(2\pi/L_\alpha) X_\alpha} | \Psi_{\mathbf{k}=0} \rangle \, , \tag{4.18}$$

as

$$P_\alpha = \lim_{N \to \infty} \frac{q_e}{2\pi} \mathrm{Im} \ln z_N^\alpha \tag{4.19}$$

and

$$\Omega_I = - \sum_\alpha \left(\frac{L_\alpha}{2\pi} \right)^2 \ln |z_N^\alpha|^2 \, . \tag{4.20}$$

These results have found use in a number of calculations, using both many-body methods ([165, 176]), and DFT ([148] and many others). In Section 4.4.4 we discuss how to evaluate P_α and Ω_I in DFT and present a minor improvement of our own, but first we will discuss the similar relationships in many-body systems, proceeding down a somewhat different track to previous work and giving a new derivation of the above relationships.

4.4 Polarization and Localization in Correlated Many-Electron Systems

4.4.1 Many-Body Wannier Functions

As we showed in Sec.4.3, imposition of twisted boundary conditions on the Hamiltonian ensures that its many-body eigenstates are of a Bloch-like form

$$\Psi_{\mathbf{k}}(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \mathrm{e}^{i\mathbf{k}.\mathbf{X}}\Phi_{\mathbf{k}}(\mathbf{r}_1, \ldots, \mathbf{r}_N) \ , \tag{4.21}$$

and it is natural to imagine defining a many-body Wannier function (MBWF) in analogy to the single-electron version:

$$W_{\mathbf{R}} = \frac{1}{V_k}\int \mathrm{d}\mathbf{k}\mathrm{e}^{-i\mathbf{k}.\mathbf{R}}\Psi_{\mathbf{k}} \ . \tag{4.22}$$

There is, however, a fundamental difference between MBWFs and normal WFs that it is important to point out. In a normal single-particle WF of an insulator, if the integral $\mathrm{d}\mathbf{k}$ is genuinely continuous, the resulting WF is exponentially localized in real space [18]. On the other hand, it is only in the d-dimensional coordinate $\mathbf{X} = \mathbf{r}_1 + \mathbf{r}_2 + \ldots$ that $W_{\mathbf{R}}$, which is defined over the whole dN dimensional configuration space of N particles, is localized. In the remaining $dN - d$ dimensions of the configuration space, the function remains extended.

If $W_{\mathbf{R}=\mathbf{0}}$ has a peak at some configuration-space point $(\mathbf{r}_1, \ldots, \mathbf{r}_2)$, then by shifting each electron coordinate $\mathbf{r}_i$ by a lattice vector $\mathbf{R}_i$, one arrives at the peak of a different function $W_{\mathbf{R}}$, with $\mathbf{R}$ such that $\sum_i \mathbf{R}_i = \mathbf{R}$. If one chooses, $\sum_i \mathbf{R}_i = \mathbf{0}$, every electron has been shifted to a different cell but the MBWF still has the same value. It cannot therefore be said to be localized in real space in any way, merely in some specific directions in configuration space.

One must therefore take care when defining the moments of these MBWFs, as expressions like $\langle W_0|\mathbf{X}|W_0\rangle$ are not necessarily normalized. In fact, as emphasized by Resta [142], $\hat{\mathbf{X}}$ is not even an admissible operator in a Hilbert space obeying PBCs, as the result of applying it to a function in the space is not periodic and thus does not occupy the same Hilbert space as the original system. However, one can still proceed if the system and the operators are carefully defined.

To avoid problems of normalization, we switch temporarily to considering a finite "supercell" of copies of the simulation cell. This supercell has sides of length $\tilde{A}_\alpha = L_\alpha A_\alpha$ for integer L_α. By demanding that the wavefunctions be truly periodic over this supercell, we have discretized the allowed k-points such that there are $N_k = L_1 L_2 \dots L_d$ of them, which can be written as $\mathbf{k} = (m_1\delta k_1, \dots, m_d\delta k_d)$ where $m_1, \dots, m_d$ are integers and $\delta k_\alpha = 2\pi/\tilde{A}_\alpha$. The Wannier function is then

$$W_{\mathbf{R}} = \frac{1}{\sqrt{N_k}} \sum_{\mathbf{k}} \mathrm{e}^{i\mathbf{k}.\mathbf{R}} \Psi_{\mathbf{k}} \; . \tag{4.23}$$

Several routes could now be followed to define moments: the easiest, producing the closest analogy to the single-particle versions, is to define a new operator $\hat{\mathbf{Q}}$ which replaces $\hat{\mathbf{X}}$, according to

$$\hat{Q}_\alpha \; = \; \frac{\sin(\delta k_\alpha \hat{\mathbf{X}}_\alpha)}{\delta k_\alpha} \; = \; \frac{\sin(\delta \mathbf{k}_\alpha \cdot \hat{\mathbf{X}})}{\delta k_\alpha} \; , \tag{4.24}$$

where $\delta\mathbf{k}_\alpha = (0, \dots, \delta k_\alpha, \dots, 0)$ is a vector of length δk_α in the α direction. For a finite supercell, $\hat{Q}_\alpha$ is periodic over $\tilde{A}_\alpha$, while in the limit $L_\alpha \to \infty$, it can be seen that $\hat{Q}_\alpha \to \hat{X}_\alpha$ as required. We can then evaluate $\hat{Q}_\alpha|W_{\mathbf{0}}\rangle$ as follows:

$$\begin{aligned}
\hat{Q}_\alpha |W_{\mathbf{0}}\rangle &= \frac{1}{\delta k_\alpha \sqrt{N_k}} \left(\frac{e^{i\delta \mathbf{k}_\alpha \cdot \hat{\mathbf{X}}} - e^{-i\delta \mathbf{k}_\alpha \cdot \hat{\mathbf{X}}}}{2i} \right) \sum_{\mathbf{k}} |\Psi_{\mathbf{k}}\rangle \\
&= \frac{-i}{2\delta k_\alpha \sqrt{N_k}} \left(e^{i\delta \mathbf{k}_\alpha \cdot \hat{\mathbf{X}}} \sum_{\mathbf{k}} |\Psi_{\mathbf{k}-\delta\mathbf{k}_\alpha}\rangle - e^{-i\delta \mathbf{k}_\alpha \cdot \hat{\mathbf{X}}} \sum_{\mathbf{k}} |\Psi_{\mathbf{k}+\delta\mathbf{k}_\alpha}\rangle \right) \\
&= \frac{i}{\sqrt{N_k}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\hat{\mathbf{X}}} \left(\frac{|\Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha}\rangle - |\Phi_{\mathbf{k}-\delta\mathbf{k}_\alpha}\rangle}{2\delta k_\alpha} \right) \\
&= \frac{i}{\sqrt{N_k}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\hat{\mathbf{X}}} \Delta_{k_\alpha} |\Phi_{\mathbf{k}}\rangle \ , \qquad (4.25)
\end{aligned}$$

where Δ_{k_α} is a discretized derivative with respect to k_α and we were able to replace $\sum_{\mathbf{k}} |\Psi_{\mathbf{k}}\rangle$ with $\sum_{\mathbf{k}} |\Psi_{\mathbf{k}\pm\delta\mathbf{k}_\alpha}\rangle$ because $|\Psi_{\mathbf{k}}\rangle$ is a periodic function of $\mathbf{k}$.

Evaluating the expectation value of $\hat{Q}_\alpha$ on $W_{\mathbf{0}}$ gives

$$\langle W_{\mathbf{0}}|\hat{Q}_\alpha|W_{\mathbf{0}}\rangle = \frac{i}{N_k} \sum_{\mathbf{k}'} \sum_{\mathbf{k}} \left\langle \Phi_{\mathbf{k}'} \middle| e^{-i\mathbf{k}'.\hat{\mathbf{X}}} e^{i\mathbf{k}.\hat{\mathbf{X}}} \middle| \Delta_{k_\alpha} \Phi_{\mathbf{k}} \right\rangle . \qquad (4.26)$$

It is clear that elements of this summation with $\mathbf{k} \neq \mathbf{k}'$ do not contribute: to prove this, one can note that both $\Phi_{\mathbf{k}'}$ and $\Delta_{k_\alpha}\Phi_{\mathbf{k}}$ are periodic over the simulation cell in every coordinate and so may be represented as dN-dimensional Fourier series involving simulation-cell reciprocal lattice vectors only. Then, since $\mathbf{k}$ and $\mathbf{k}'$ are within the simulation-cell Brillouin zone and cannot differ by a non-zero simulation-cell reciprocal lattice vector, it follows that the real-space integral gives zero unless $\mathbf{k} \neq \mathbf{k}'$, leaving

$$\langle W_{\mathbf{0}}|\hat{Q}_\alpha|W_{\mathbf{0}}\rangle = \frac{i}{N_k} \sum_{\mathbf{k}} \left\langle \Phi_{\mathbf{k}} \middle| \Delta_{k_\alpha} \Phi_{\mathbf{k}} \right\rangle . \qquad (4.27)$$

In the limit of very large supercells we can return to an integral over $\mathbf{k}$: assuming a choice of gauge for which $\Phi_{\mathbf{k}}$ is a smooth function of $\mathbf{k}$, this yields

$$\langle W_{\mathbf{0}}|\hat{Q}_\alpha|W_{\mathbf{0}}\rangle = \frac{i}{V_k} \int d\mathbf{k} \left\langle \Phi_{\mathbf{k}} \middle| \partial_{k_\alpha} \Phi_{\mathbf{k}} \right\rangle , \qquad (4.28)$$

The one-electron version of this result was originally derived by Blount [17]. The matrix elements in Eq. (4.28) can be evaluated by integrating over the region of configuration space corresponding to the volume of the supercell in each electron coordinate, with both $W_{\mathbf{0}}$ and $\Phi_{\mathbf{k}}$ normalized to unity over that region.

Now, by comparing Eq. (4.28) with Eq. (4.12), and the equivalent many-electron expressions of Ortiz and Martin [127], we see in analogy to the single-particle version, that the electronic polarization is given by $(P_{\mathrm{el}})_\alpha = q_e \langle W_{\mathbf{0}} | \hat{Q}_\alpha | W_{\mathbf{0}} \rangle / V$, so in the large supercell limit this is

$$(P_{el})_\alpha = \frac{i q_e}{(2\pi)^d} \int \mathrm{d}\mathbf{k} \; \langle \Phi_{\mathbf{k}} | \partial_{k_\alpha} \Phi_{\mathbf{k}} \rangle \tag{4.29}$$

This demonstrates that the change in polarization in response to an adiabatic change of Hamiltonian can be obtained from the change in the first moment of the many-body Wannier function $W_{\mathbf{0}}$. Periodic gauge transformations of the type defined in Eq. (4.3) shift the right-hand side of Eq. (4.28) by the lattice vector $\mathbf{R}$, corresponding to a shift in the polarization by an integer multiple of the quantum of polarization [85, 170].

A similar approach may be used to find the second moment of $W_{\mathbf{0}}$, giving the spread of $\hat{\mathbf{X}}$ about its mean value, from which we can calculate the localization length. Proceeding as above gives

$$\hat{Q}_\alpha \hat{Q}_\beta \, |W_{\mathbf{0}}\rangle = -\frac{1}{\sqrt{N_k}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{X}} \Delta_{k_\alpha} \Delta_{k_\beta} \, |\Phi_{\mathbf{k}}\rangle \;\; , \tag{4.30}$$

and hence the large supercell limit of the tensor spread functional,

$$\Omega_{\alpha\beta} = \langle W_{\mathbf{0}} | \, \hat{Q}_\alpha \hat{Q}_\beta \, | W_{\mathbf{0}} \rangle - \langle W_{\mathbf{0}} | \, \hat{Q}_\alpha \, | W_{\mathbf{0}} \rangle \, \langle W_{\mathbf{0}} | \, \hat{Q}_\beta \, | W_{\mathbf{0}} \rangle \;\; , \tag{4.31}$$

can be written as

$$\begin{aligned}\Omega_{\alpha\beta} &= -\Bigg(\int \frac{d\mathbf{k}}{V_k} \Big\langle \Phi_{\mathbf{k}} \Big| \partial_{k_\alpha} \partial_{k_\beta} \Phi_{\mathbf{k}} \Big\rangle \\ &\quad - \int \frac{d\mathbf{k}}{V_k} \int \frac{d\mathbf{k}'}{V_k} \Big\langle \Phi_{\mathbf{k}} \Big| \partial_{k_\alpha} \Phi_{\mathbf{k}} \Big\rangle \Big\langle \Phi_{\mathbf{k}'} \Big| \partial_{k'_\beta} \Phi_{\mathbf{k}'} \Big\rangle \Bigg) . \end{aligned} \tag{4.32}$$

Following [117], we split the spread functional into a gauge invariant part Ω_I and a gauge dependent part $\widetilde{\Omega}$, simply by choosing Ω_I to include the terms in which the phase $e^{i\phi(\mathbf{k})}$ cancels:

$$\Omega = \Omega_I + \widetilde{\Omega} , \tag{4.33}$$

where

$$(\Omega_I)_{\alpha\beta} = -\int \frac{d\mathbf{k}}{V_k} \Big(\langle \Phi_{\mathbf{k}} | \partial_{k_\alpha} \partial_{k_\beta} \Phi_{\mathbf{k}} \rangle - \langle \Phi_{\mathbf{k}} | \partial_{k_\alpha} \Phi_{\mathbf{k}} \rangle \langle \Phi_{\mathbf{k}} | \partial_{k_\beta} \Phi_{\mathbf{k}} \rangle \Big) \tag{4.34}$$

and

$$\widetilde{\Omega}_{\alpha\beta} = \int \frac{d\mathbf{k}}{V_k} \int \frac{d\mathbf{k}'}{V_k} \langle \Phi_{\mathbf{k}} | \partial_{k_\alpha} \Phi_{\mathbf{k}} \rangle \Big(\langle \Phi_{\mathbf{k}'} | \partial_{k'_\beta} \Phi_{\mathbf{k}'} \rangle - \langle \Phi_{\mathbf{k}} | \partial_{k_\beta} \Phi_{\mathbf{k}} \rangle \Big) . \tag{4.35}$$

Thus we see that $(\Omega_I)_{\alpha\alpha} = N \langle r_\alpha^2 \rangle_c$, where $\langle r_\alpha^2 \rangle_c$ is the squared localization length in the α direction (as defined by Resta [144] and in Souza, Wilkens, Martin [156]). Furthermore, by re-expressing the many-body Bloch functions in terms of the Wannier functions, the gauge-dependent part of the functional may be written as

$$\widetilde{\Omega}_{\alpha\beta} = \sum_{\mathbf{R} \neq \mathbf{0}} \Big\langle W_{\mathbf{R}} | \hat{Q}_\alpha | W_{\mathbf{0}} \Big\rangle \Big\langle W_{\mathbf{0}} | \hat{Q}_\beta | W_{\mathbf{R}} \Big\rangle . \tag{4.36}$$

If the many-body Wannier functions of an insulator are taken as "disconnected" Kohn's functions, it must be possible to choose a gauge in which those localized on different vectors $\mathbf{R}$ are non-overlapping in configuration space when the simulation cell is large enough. The gauge-dependent part of the spread therefore tends to zero with increasing simulation-cell size [156]. This allows us to limit our consideration to the gauge-invariant part Ω_I.

4.4.2 Discretization

The above definitions of $(\Omega_I)_{\alpha\beta}$ and P_α provide a description of the behaviour of both quantities in the thermodynamic limit, as $L \to \infty$. If we wish to calculate them in a simulation, however, in a necessarily finite system, we must return to a discretized form.

We define the function $f(\mathbf{q}) = \langle \Phi_\mathbf{k} | \Phi_{\mathbf{k}+\mathbf{q}} \rangle$ and make use of the following identities:

$$\partial_{q_\alpha} \ln f = \frac{1}{f} \partial_{q_\alpha} f \tag{4.37}$$

and

$$\partial^2_{q_\alpha} \ln f = \frac{1}{f} \partial^2_{q_\alpha} f - \frac{1}{f^2} (\partial_{q_\alpha} f)^2 \tag{4.38}$$

so that we can rewrite Eqs. 4.29 and 4.34 as

$$(P_{el})_\alpha = \frac{iq_e}{(2\pi)^d} \int \mathrm{d}\mathbf{k} \, \left[\partial_{q_\alpha} \ln \langle \Phi_\mathbf{k} | \Phi_{\mathbf{k}+\mathbf{q}} \rangle \right]_{\mathbf{q}=0} \tag{4.39}$$

and

$$(\Omega_I)_{\alpha\alpha} = -\frac{1}{V_k} \left[\int \mathrm{d}\mathbf{k} \, \partial^2_{q_\alpha} \ln \langle \Phi_\mathbf{k} | \Phi_{\mathbf{k}+\mathbf{q}} \rangle \right]_{\mathbf{q}=0} \tag{4.40}$$

We also need to discretize the derivatives and integrals, for which we use discretized forms of the derivative operators:

$$\begin{aligned} \partial_{q_\alpha} \ln f &\simeq (\ln f(\mathbf{q} + \delta \mathbf{q}_\alpha) - \ln f(\mathbf{q}))/\delta q_\alpha \\ \partial^2_{q_\alpha} \ln f &\simeq (\ln f(\mathbf{q} + \delta \mathbf{q}_\alpha) + \ln f(\mathbf{q} - \delta \mathbf{q}_\alpha) - 2 \ln f(\mathbf{q}))/\delta q^2_\alpha \end{aligned} \tag{4.41}$$

which gives us

$$(P_{el})_\alpha = \frac{iq_e}{N_k V} \sum_\mathbf{k} \left(\ln \langle \Phi_\mathbf{k} | \Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha} \rangle - \ln \langle \Phi_\mathbf{k} | \Phi_\mathbf{k} \rangle \right) / \delta k_\alpha \tag{4.42}$$

and

$$(\Omega_I)_{\alpha\alpha} = \quad -\tfrac{1}{N_k}\textstyle\sum_{\mathbf{k}} \tfrac{1}{\delta k_\alpha^2}\Big[\ln \langle \Phi_{\mathbf{k}} | \Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha} \rangle \\ + \ln \langle \Phi_{\mathbf{k}} | \Phi_{\mathbf{k}-\delta\mathbf{k}_\alpha} \rangle - 2 \ln \langle \Phi_{\mathbf{k}} | \Phi_{\mathbf{k}} \rangle \Big] \, , \tag{4.43}$$

By expanding $\langle \Phi_{\mathbf{k}} | \Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha} \rangle$ order by order in $\delta\mathbf{k}_\alpha$ one can show that the term inside the logarithm of the polarization expression is pure imaginary — as indeed it must be on physical grounds so that the observable P_{el} is real. The periodicity of the summand over the simulation-cell Brillouin zone allows us to make the substitution $\mathbf{k} \to \mathbf{k} + \delta\mathbf{k}_\alpha$ in the second term of Eq. 4.43, which changes it from $\langle \Phi_{\mathbf{k}} | \Phi_{\mathbf{k}-\delta\mathbf{k}_\alpha} \rangle$ to $\langle \Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha} | \Phi_{\mathbf{k}} \rangle$, so in all we have

$$(P_{el})_\alpha = \frac{-q_e}{N_k \delta k_\alpha V} \sum_{\mathbf{k}} \operatorname{Im} \ln \langle \Phi_{\mathbf{k}} | \Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha} \rangle \tag{4.44}$$

$$(\Omega_I)_{\alpha\alpha} = -\frac{1}{N_k} \sum_{\mathbf{k}} \frac{1}{\delta k_\alpha^2} \ln |\langle \Phi_{\mathbf{k}} | \Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha} \rangle|^2 \, . \tag{4.45}$$

4.4.3 Ansatz Wavefunction

Eqs. 4.44 and 4.45 would be usable expressions in a real system if we were able to calculate the many-electron wavefunction at every point on a grid of twist vectors and calculate overlaps between wavefunctions of different twists. However, it is usually more convenient to work with a single value of twist $\mathbf{k}$, and calculate expectation values of a single wavefunction. Moreover, in diffusion Monte Carlo, where we never know the true ground state wavefunction itself, the above overlaps could not be evaluated. We can put Eqs. 4.44 and 4.45 into forms suitable for evaluation as single expectation values by considering the wavefunction of a much larger system. This is a variant of the argument of [144, 156], and like them uses an ansatz about the form of the wavefunction for the larger system and relates the polarization and the localization length to a single expectation value of the "many-body phase operator" introduced by Resta [142].

In the course of discretizing, we introduced a supercell containing $N_k = L_1 L_2 \dots L_d$ simulation cells. The N-electron Wannier functions $W_{\mathbf{R}}(\mathbf{r}_1, \dots, \mathbf{r}_N)$, of this supercell have been suggested as having the properties of the "disconnected" functions introduced by Kohn [92] and are believed to be exponentially localized in insulators. More precisely, $W_{\mathbf{R}}(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is localized in that region of the dN-dimensional configuration space where the d-dimensional vector $\mathbf{X} = \mathbf{r}_1 + \dots + \mathbf{r}_N$ is close to $\mathbf{R}$. The exponential localization in phase space implies that the Hamiltonian matrix elements between neighbouring Wannier functions are exponentially small, and hence that the ground state energy of the simulation cell is exponentially insensitive to the choice of twist vector $\mathbf{k}$. This insensitivity to boundary conditions is a defining characteristic of the insulating state.

Previously, this supercell was introduced for mathematical convenience but was not viewed as a physical system; in particular, its wavefunctions had the same number of electrons, N, as the smaller simulation cell. Suppose, however, that we now consider a real supercell system containing N_k simulation cells and $\widetilde{N} = N_k N$ electrons subject to periodic boundary conditions. The wavefunction $\widetilde{\Psi}(\mathbf{r}_1, \dots, \mathbf{r}_{\widetilde{N}})$ of this physical supercell is not known and would be challenging to calculate, but the availability of N_k N-electron Bloch functions $\Psi_{\mathbf{k}}$ and the supposition that electron correlation in an insulator is a short-ranged effect suggests an appealing Hartree-Fock-like ansatz. Assigning each of the N_k $\mathbf{k}$-points an integer index $c = 1, \dots, N_k$, we write, as in [144, 156],

$$\widetilde{\Psi}(\mathbf{r}_1, \dots, \mathbf{r}_{\widetilde{N}}) = \hat{A} \prod_{\mathbf{k}}^{N_k} \Psi_{\mathbf{k}}(\mathbf{r}_{N(c-1)+1}, \dots \mathbf{r}_{N(c-1)+N}) \; , \tag{4.46}$$

where $\hat{A}$ is the antisymmetrization operator. The ansatz assumes that the $\widetilde{N} = N_k N$ electrons in the supercell correlate in groups of N at a time, so that $\widetilde{\Psi}$ can be written as an antisymmetrized product of N_k different N-electron wavefunctions. If the Hamiltonian of the physical supercell were a sum of terms involving each group of N electrons separately, as in a non-interacting system, $\widetilde{\Psi}$ would be the exact ground state. In an interacting system $\widetilde{\Psi}$ is clearly inexact, but since each N-electron wavefunction $\Psi_{\mathbf{k}}$ was originally an allowable periodic wavefunction for a smaller section of the solid, we are effectively requiring that beyond some range L, the difference between the interactions of genuinely distinguishable electrons and between periodic copies of the same electron becomes negligible. As the number of electrons N in the simulation cell and the cell volume V increase, the importance of correlations of longer range than the size of the simulation cell should decrease and $\widetilde{\Psi}$ should become more accurate.

As in [144, 156], we evaluate the expectation value

$$z_N = \langle\widetilde{\Psi}|e^{-i\delta\mathbf{k}_\alpha\cdot\widetilde{\mathbf{X}}}|\widetilde{\Psi}\rangle \tag{4.47}$$

of the "many-body phase operator" $e^{-i\delta\mathbf{k}_\alpha\cdot\widetilde{\mathbf{X}}}$ in this ansatz wavefunction, where $\widetilde{\mathbf{X}} = \mathbf{r}_1 + \ldots + \mathbf{r}_{N_kN}$. The overlap between two Slater determinants is the determinant of the matrix of overlaps between the "orbitals", which here are N-electron wavefunctions labelled by a Bloch wavevector $\mathbf{k}$. This remains true even in this more complex many body case because any term in the antisymmetrization in which the same electron coordinate appears with different boundary conditions in the bra and the ket will vanish upon integration. We can thus write the expectation value as:

$$\left\langle\widetilde{\Psi}|e^{-i\delta\mathbf{k}_\alpha\cdot\widetilde{\mathbf{X}}}|\widetilde{\Psi}\right\rangle = \prod_{\mathbf{k}}^{N_k}\langle\Phi_{\mathbf{k}}|\Phi_{\mathbf{k}+\delta\mathbf{k}_\alpha}\rangle\ . \tag{4.48}$$

Comparing this result with Eq. (4.45), we see that

$$(\Omega_I)_{\alpha\alpha} = \frac{-1}{N_k(\delta k_\alpha)^2}\ln\left|\langle\widetilde{\Psi}|e^{-i\delta\mathbf{k}_\alpha\cdot\widetilde{\mathbf{X}}}|\widetilde{\Psi}\rangle\right|^2 , \tag{4.49}$$

and hence that the localization length $\langle r_\alpha^2\rangle_c = (\Omega_I)_{\alpha\alpha}/N$ is given by

$$\langle r_\alpha^2\rangle_c = \frac{-1}{\widetilde{N}(\delta k_\alpha)^2}\ln\left|\left\langle\widetilde{\Psi}\middle|e^{-i\delta\mathbf{k}_\alpha\cdot\widetilde{\mathbf{X}}}\middle|\widetilde{\Psi}\right\rangle\right|^2 . \tag{4.50}$$

This is Resta and Sorella's [144] result for the square of the localization length, or quadratic spread, in the α direction. Henceforth, we shall treat the large system as the single system we are studying and drop the tildes.

Eq. 4.50 was originally obtained by supposing that the electron density could be decomposed into a sum of localized components, Fourier transforming, and applying the ansatz described above to produce a large enough system for convergence. This melds well with the proposed identification [156] of many-body Wannier functions with the "disconnected" parts of the many-body wavefunction hypothesized by Kohn [92]. Since the disconnected parts are non-overlapping in the configuration space of the electron coordinates, each can be considered as providing a separate contribution to the electron density in real space.

4.4.4 Evaluation of Localization Lengths in a Plane Wave Basis

In the work presented here, which uses DFT in 2D systems, localization lengths are calculated from the one-electron Bloch functions using an adaptation of the method described in [148]. In this approach, the many-body wavefunction Ψ in Eq. (4.50) is a Slater determinant of single-particle orbitals, and the expectation value $\left\langle\Psi\middle|e^{-i\delta\mathbf{k}_\alpha\cdot\mathbf{X}}\middle|\Psi\right\rangle$ is expressed as $\left\langle\Psi\middle|\Phi\right\rangle$, where Φ is a Slater determinant of the orbitals $\psi_{n\mathbf{k}}(\mathbf{r}_j)$ of Ψ, each multiplied by $e^{-i\delta\mathbf{k}_\alpha\cdot\mathbf{r}_j}$. The overlap of two Slater determinants is a determinant of the overlaps of the individual orbitals, and as shown in [148], the only terms $\langle\psi_{n\mathbf{k}}|e^{-i\delta\mathbf{k}_\alpha\cdot\mathbf{r}}|\psi_{n'\mathbf{k}'}\rangle$ of this determinant which survive the integration over the whole system are those for which $\mathbf{k}' = \mathbf{k}+\delta\mathbf{k}_\alpha+\mathbf{G}$, where $\mathbf{G}$ is a primitive reciprocal lattice vector (which may be zero). This makes the matrix very sparse indeed. The determinant can be factorized as

$$\left\langle\Psi\middle|e^{-i\delta\mathbf{k}_\alpha\cdot\mathbf{X}}\middle|\Psi\right\rangle = \prod_{\mathbf{k}} \det S^\alpha(\mathbf{k}) \; , \tag{4.51}$$

where $S_{nn'}(\mathbf{k})$ is a matrix of overlaps between the periodic parts of the Bloch functions at points $\mathbf{k}$ and $\mathbf{k}+\delta\mathbf{k}_\alpha+\mathbf{G}$. If both $\mathbf{k}$ and $\mathbf{k}+\delta\mathbf{k}_\alpha$ lie within the Brillouin zone (so that $\mathbf{G}=\mathbf{0}$), then

$$S^\alpha_{nn'}(\mathbf{k}) = \langle u_{n\mathbf{k}}|u_{n'\mathbf{k}+\delta\mathbf{k}_\alpha}\rangle \tag{4.52}$$

If $\mathbf{k}$ lies inside the Brillouin zone but $\mathbf{k}+\delta\mathbf{k}_\alpha$ lies outside, then

$$\begin{aligned} S^\alpha_{nn'}(\mathbf{k}) &= \langle\psi_{n\mathbf{k}}|\, e^{-i\delta\mathbf{k}_\alpha\cdot\mathbf{r}}\,|\psi_{n'\mathbf{k}+\delta\mathbf{k}_\alpha-\mathbf{G}}\rangle \\ &= \langle u_{n\mathbf{k}}|\, e^{-i\mathbf{G}\cdot\mathbf{r}}\,|u_{n'\mathbf{k}+\delta\mathbf{k}_\alpha-\mathbf{G}}\rangle \end{aligned}$$

where $\mathbf{G}$ is chosen such that $\mathbf{k}+\delta\mathbf{k}_\alpha-\mathbf{G}$ lies inside the Brillouin zone (because that was the state included in the original Slater determinant).

A further improvement we present in this work is to note that in a plane-wave basis, the matrix elements can be evaluated directly from the plane-wave coefficients $c^*_{n\mathbf{k}}$ of bands n and n' using Parseval's relation, as:

$$S^\alpha_{nn'}(\mathbf{k}) = \sum_{\mathbf{G}'} c^*_{n\mathbf{k}}(\mathbf{G}')c_{n'\mathbf{k}+\delta\mathbf{k}_\alpha-\mathbf{G}}(\mathbf{G}'+\mathbf{G}) \ , \tag{4.53}$$

where $\mathbf{G}$ is the reciprocal lattice vector that brings $\mathbf{k}+\delta\mathbf{k}_\alpha$ back into the first Brillouin zone if required. This expression allows very efficient calculation of z_N and thus the localization lengths, without the need to calculate overlaps in real space.

4.4.5 Evaluation of Localization Lengths in Quantum Monte Carlo

The form of Eq. 4.50 is ideally suited to evaluating the localization length within the framework of quantum Monte Carlo, and has been used in a small number of previous applications in situations such as calculation of dielectric properties [165, 176]. As with the evaluation of any expectation value in QMC, the result is subject to an associated statistical error and in this section we examine its behaviour. In cases where the localization length is large, and $|z_N|$ is correspondingly small, the error is somewhat poorly behaved and must be treated carefully.

In VMC, the algorithm provides us with a series of configurations $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ of electron positions distributed according to the probability distribution defined by the trial wavefunction. The expectation value of the many-body phase operator in Eq. 4.47 is evaluated by taking the sum of the electron coordinates $\mathbf{X}(\mathbf{R}) = \mathbf{r}_1 + \mathbf{r}_2 + \ldots + \mathbf{r}_N$ for each sampled configuration, calculating the quantity

$$z_N(\mathbf{R}) = e^{-i\delta\mathbf{k}_\alpha\cdot\mathbf{X}(\mathbf{R})} , \tag{4.54}$$

and averaging over a sufficient number M of statistically independent configurations to obtain:

$$\overline{z}_N = \frac{1}{M}\sum_{m=1}^{M} z_N(\mathbf{R}_m) . \tag{4.55}$$

As M tends to infinity, the sample mean $\overline{z}_N$ tends to z_N and the localization length can be obtained from the formula

$$\langle r_\alpha^2\rangle_c \;=\; -\frac{1}{N}\frac{A_\alpha^2}{(2\pi)^2}\ln|z_N|^2 \;\approx\; -\frac{1}{N}\frac{A_\alpha^2}{(2\pi)^2}\ln|\overline{z}_N|^2 . \tag{4.56}$$

The complication is that $z_N(\mathbf{R}_m)$ is a complex number whose modulus is always unity but whose mean z_N is very small if the localization length is large. To ensure the result is not swamped by the statistical error $\sigma_{|\overline{z}_N|}$ in $|\overline{z}_N|$, the number of configurations M must be large enough that $\sigma_{|\overline{z}_N|} \ll |z_N|$.

The error in the localization length can be estimated as follows. Given a complex random variable such as $\overline{z}_N = \overline{x}_N + i\,\overline{y}_N$, where $\overline{x}_N$ and $\overline{y}_N$ are statistically independent real random variables with population mean values x_N and y_N, and the uncertainties on the components obey $\sigma_{\overline{x}_N} \ll x_N$ and $\sigma_{\overline{y}_N} \ll y_N$, it is easy to show that

$$\sigma_{\ln(|\overline{z}_N|^2)} = \sigma_{\ln(\overline{x}_N^2+\overline{y}_N^2)} \approx \frac{\sqrt{4x_N^2\sigma_{\overline{x}_N}^2 + 4y_N^2\sigma_{\overline{y}_N}^2}}{x_N^2+y_N^2} . \tag{4.57}$$

Assuming that the M individual readings $x_N(\mathbf{R})$ and $y_N(\mathbf{R})$ used to calculate the sample mean values $\overline{x}_N$ and $\overline{y}_N$ are uncorrelated, $\sigma^2_{\overline{x}_N} = \sigma^2_{x_N}/M$ and $\sigma^2_{\overline{y}_N} = \sigma^2_{y_N}/M$, so

$$\sigma_{\ln(|\overline{z}_N|^2)} \approx \frac{\sqrt{4x_N^2\sigma^2_{x_N} + 4y_N^2\sigma^2_{y_N}}}{\sqrt{M}(x_N^2 + y_N^2)} , \tag{4.58}$$

where σ_{x_N} and σ_{y_N} are the errors in a single reading $x_N(\mathbf{R})$ or $y_N(\mathbf{R})$. The error in a single reading of z_N near a metal-insulator transition can be estimated by assuming its mean is in fact zero and that its possible values are distributed uniformly around the unit circle in the complex plane, so that

$$\sigma^2_{x_N} = \sigma^2_{y_N} = \frac{1}{2\pi}\int_0^{2\pi} \cos^2\theta \, \mathrm{d}\theta = \frac{1}{2} . \tag{4.59}$$

Finally, combining Eqs. 4.58 and 4.59 with Eq. 4.56, we obtain

$$\sigma_{\langle r^2_\alpha \rangle_c} = \frac{A^2_\alpha \sqrt{2}}{(2\pi)^2 N |z_N| \sqrt{M}} . \tag{4.60}$$

Since $\langle r^2_\alpha \rangle_c = -\kappa \ln |z_N|^2$, where κ is a constant, $|z_N|$ approaches zero like $e^{-\langle r^2_\alpha \rangle_c / 2\kappa}$ as $\langle r^2_\alpha \rangle_c$ tends to infinity, as we might expect at an appropriate type of insulator-to-metal transition. It then requires exponentially longer runs to achieve a given error bar on $\langle r^2_\alpha \rangle_c$. For metallic systems with a small number of electrons in the conduction band or holes in the valence band, the values of $z_N(\mathbf{R})$ from configuration to configuration are very strongly correlated and it takes many more than M VMC steps to accumulate M uncorrelated VMC samples. This occurs because the values of $z_N(\mathbf{R})$ depend strongly on the current positions of the "extra" or "missing" electrons in the cell and makes the determination of whether z_N is in fact 0 or is merely very close to it all but impossible.

4.5 Model Systems

In this section, we examine the behaviour of the localization lengths of the electrons of various 2D model systems. In two dimensions (and also in quasi-1D systems with periodicity in only one direction) most of the important effects of bandstructure on localization behaviour are already present, while the calculations remain relatively tractable, even in QMC. We study these systems primarily with density functional theory, although we use also variational Monte Carlo in certain cases.

Previous work [144] has indicated that the localization length correctly describes the various phases of a one-dimensional system of correlated electrons displaying a band insulator to Mott insulator transition with a metallic phase at the transition: a divergence is clearly observed in $\langle x^2 \rangle_c$ at sufficiently large system sizes. Given that the localization of Wannier functions is closely related to the energy gap, and that the localization length of a metal in band theory is always infinite, it might be supposed that the localization length would in general diverge as a metal-insulator transition is approached from the insulating side. As we shall show by a number of examples, however, this is not the case: the presence or absence of a divergence in one-electron theory depends on the nature of the bands that are crossing. The same behaviour is observed in many-electron quantum Monte Carlo simulations using a Slater-Jastrow trial function containing a Slater determinant of single-particle orbitals from band theory. To examine the behaviour of the localization length in a range of different situations, we study different arrangements of a two-dimensional system of potential wells we shall refer to as ‘quantum dots’.

The simulations in this section were performed with a new 2D DFT code written by the author, named DOTDFT, which uses a plane wave basis to represent the orbitals. DOTDFT uses the two-dimensional parameterization of the local density approximation from [8], based on the QMC calculations of the correlation energy of the 2D HEG by Tanatar and Ceperley [159]. The single-particle orbitals are expressed in a basis of two-dimensional plane-waves with coefficients $c_{i,\mathbf{k}+\mathbf{G}}$, such that

$$\psi_{i\mathbf{k}} = \sum_{\mathbf{G}} c_{i,\mathbf{k}+\mathbf{G}} \mathrm{e}^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \tag{4.61}$$

This basis is complete and capable of representing any suitable Bloch function if the sum over $\mathbf{G}$ includes the entire infinite set of reciprocal lattice vectors. In practice, the basis must be finite so the sum over $\mathbf{G}$ is truncated at some finite value of $|\mathbf{G}|$, such that $|\mathbf{G}| < G_{\mathrm{cut}}$. This is equivalent to setting a minimum length scale over which the Bloch function is able to vary rapidly, as it is the large $\mathbf{G}$ terms that will introduce short wavelengths and the corresponding high kinetic energies. It is important, therefore, to ensure that the value of G_{cut}, usually written in terms of the maximum plane-wave energy $E_{\mathrm{cut}} = \frac{1}{2}G_{\mathrm{cut}}^2$, is large enough that the energy is converged with respect to the size of the basis. This is tested beforehand for all the systems described here.

The code follows the approach outlined in Chapter 2 and achieves self-consistency by mixing the output and input densities using the Broyden method [20]. The Hamiltonian consists of the kinetic energy (which is diagonal in the plane wave representation), the Hartree potential, the exchange-correlation potential, and an analytically specified external potential as detailed below. It is diagonalized directly, using the linear algebra package LAPACK [7]. This approach is feasible in these systems, even though it would not be in real materials with very large basis sets, because the low dimensionality and use of smooth external potentials means the basis can be relatively small: a few hundred plane waves is usually sufficient.

4.5.1 Finite Size Effects

Previous studies [117, 144, 148] of electron localization in one-electron theories have noted the relatively slow and monotonically increasing convergence of the localization length as a function of system size. Sgiarovello, Peressi and Resta [148] examined the localization length $\langle x^2 \rangle_c$ in GaAs, and Marzari and Vanderbilt [117] the spread Ω of the maximally localized Wannier functions. Both came to similar conclusions.

The slow convergence can be explained by reference to Eq. 4.50 and consideration of the ansatz required to derive it. For the localization length to be well converged, the k-point sampling must be dense enough to sample the variation of the Bloch functions with $\mathbf{k}$; the finite difference approximation used in Eq. 4.45 is then a good approximation to the continuous k derivative in Eq. 4.40. Failure to converge with respect to $\delta\mathbf{k}_\alpha$ will appear as a variation of the localization length with the number of k points in the grid.

In one-electron theory, the convergence of Brillouin zone integrals may often be improved by shifting the entire grid of $\mathbf{k}$ vectors by a small amount $\widetilde{\mathbf{k}}$ relative to Γ. This corresponds to changing the strict periodic boundary conditions applied across the supercell to boundary conditions incorporating a twist vector $\widetilde{\mathbf{k}}$, which must be the same for all one-electron orbitals. A supercell twist $\widetilde{\mathbf{k}} = \delta\mathbf{k}_\alpha$ shifts every $\mathbf{k}$ vector to one of its neighbours and is thus equivalent to a zero twist, implying that $\widetilde{\mathbf{k}}$ may be chosen to lie within the supercell Brillouin zone. The variation of the localization length with $\widetilde{\mathbf{k}}$ has not been investigated previously.

To examine the convergence of the localization length, a two-dimensional array of potential wells capable of exhibiting a wide range of bandstructures was studied. The external potential consisted of a square array of Gaussian wells or “dots” of the form $V(\mathbf{r}) = -V_d \exp(-|\mathbf{r} - \mathbf{r}_c|^2/\rho^2)$, where $\mathbf{r}_c$ sets the center of the dot (in the center of the unit cell), ρ defines the width of the Gaussian, and V_d the dot depth. Neighbouring dots were separated by a distance a (the side of the unit cell, which is the same as the simulation cell in this case). The Kohn-Sham orbitals for this external potential were calculated on an equally spaced grid of $M \times M$ k-points within the Brillouin zone.

Figure 4.1 shows the localization length obtained by calculating z_N (see Eq. 4.47) for a weakly bound array of dots, each containing six electrons and thus three filled bands of doubly-occupied states, for a range of $M \times M$ k-point grids offset by $\widetilde{\mathbf{k}}$ from the Γ point.

The value of z_N was calculated from from the plane-wave representation of the orbitals as in 4.4.4. The parameters chosen were $\rho = 3$ and $a = 10$, with V_d tuned to a value only just yielding an insulating state with the first three bands fully occupied. If V_d is reduced further, the energy of the third band at $\mathbf{k} = (\frac{\pi}{a}, 0)$ becomes higher than that of the fourth band at $\mathbf{k} = (0, 0)$ and the system becomes metallic. The smallest energy gap is therefore indirect and the localization length

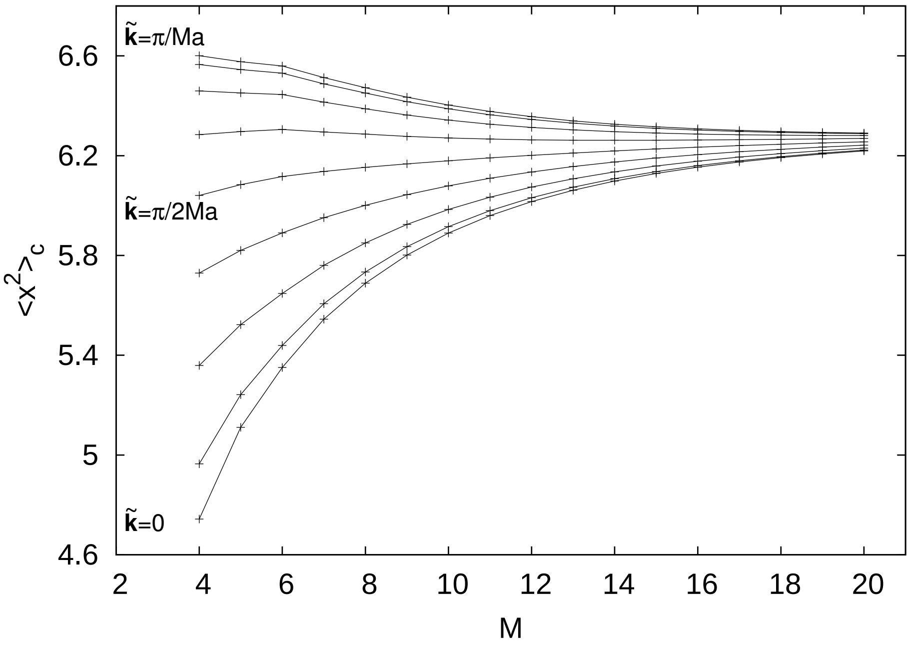

Figure 4.1: Convergence of the localization length as system size is increased. Different choices of offset of k-point grid strongly affect the speed of convergence.

might be expected to be only weakly dependent on the position of the grid, especially given that the total DFT energy is converged to five significant figures for $M \gtrsim 5$. However, as seen in Figure 4.1, in this situation chosen for its proximity to a transition, the convergence is slow and a 20 × 20 supercell is required to converge $\langle x^2 \rangle_c$ to 1%. A k-point grid centered on Γ is seen to yield relatively slow convergence, in agreement with previous studies that have shown the same behaviour for total energies [139]. In this particular system, the most rapid convergence is achieved when $\widetilde{\mathbf{k}} = \pi/(2Ma)$.

4.5.2 1D Chains

This is the simplest realizable model with interesting properties as regards metal-insulator transitions. In this set-up, a two-dimensional periodic array of dots is made quasi-one-dimensional by effectively isolating adjacent one-dimensional chains of dots. The unit cell is widened in the y direction and a strong barrier potential is introduced to separate adjacent unit cells in that direction. Every point on the k-space grid used for the Brillouin zone integration has $k_y = 0$, with equally spaced values of k_x. The simplicity of the one-dimensional bandstructure produced by this arrangement makes it easy to distinguish the effect of the symmetry properties of the Bloch functions on the behaviour of the localization tensor near the insulator-to-metal transition.

The lowest three eigenfunctions of an isolated dot that is not symmetric in the x and y directions can be classified as s-like, p_x-like and p_y-like: the s-like functions are nodeless, while the p_x- and p_y-like functions have nodes along the y- and x-axes, respectively, and change sign under reflections in those axes. The s- and p_x-like orbitals of a chain of dots oriented along the x axis mix to form hybrid bands, but the p_y-like orbitals mix only with each other and form a completely separate band. By varying the strengths of the confinement in the x- and y-directions independently, the p_y-like band can be shifted relative to the s and p_x bands. Starting from an insulating system with only the s and p_x bands occupied, this makes it possible to lower the energy of the p_y band until a band insulator to metal transition occurs.

We model this arrangement with a unit cell of size $a \times b$ and an asymmetric quartic potential of the form:

$$\begin{aligned} V(x,y) &= \frac{1}{2}\omega_x^2\left((x-a/2)^2 - \frac{2}{a^2}(x-a/2)^4\right) \\ &\quad + \frac{1}{2}\omega_y^2\left((y-b/2)^2 - \frac{2}{b^2}(y-b/2)^4\right) . \end{aligned} \tag{4.62}$$

By setting $\omega_y > \omega_x$ and $b > a$, we ensure there is an enormous barrier in the y direction, increasing the energy of the p_y band and creating a gap between the p_y band and the s and p_x bands. Then, filling up the s and p_x bands, the Kohn-Sham equations are solved self-consistently for gradually lower values of ω_y until the upper p_y band falls to meet the higher of the two symmetric bands at $\mathbf{k} = 0$ and the energy gap reduces to zero (see Fig.4.2(a)).

Because, by symmetry, there can be no matrix elements between states in the

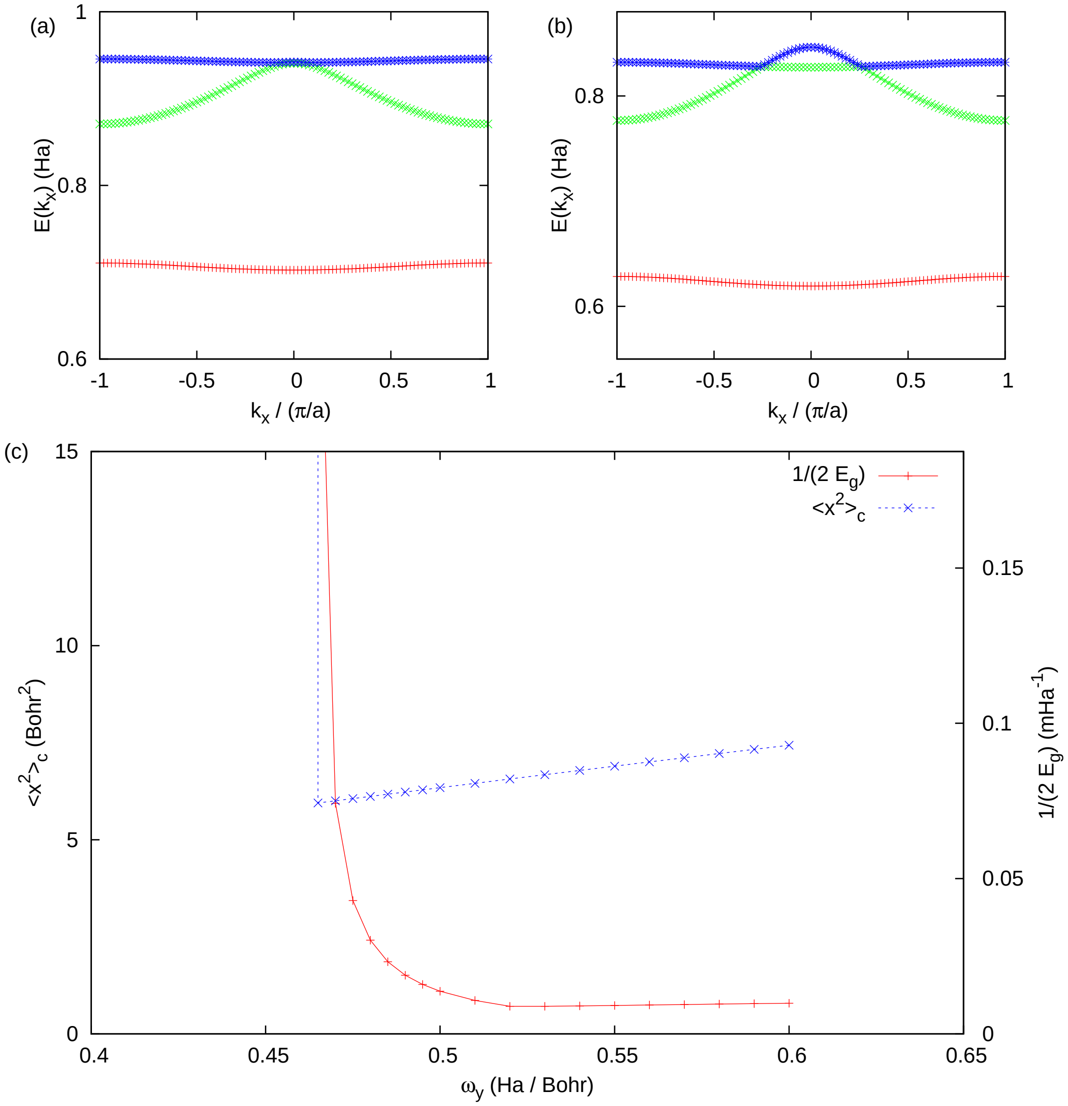

Figure 4.2: Quasi-one-dimensional chain of 4-electron dots: **(a)** Energy bands $E(k_x)$ just before band crossing. Topmost flat band is p_y-like, middle is p_x-like, lowest is s-like. **(b)** Energy bands $E(k_x)$ shortly after transition. **(c)** Localization length $\langle x^2\rangle_c$ (blue) and inverse direct energy $1/(2E_g)$ gap (red) to lowest unoccupied state as ω is reduced.

crossing bands, they cannot hybridize and the upper band has no effect on the lower until the moment at which they cross, at which point the lower band is no longer filled and the localization length becomes infinite. Beyond the transition, the bands pass straight through each other, remaining degenerate at the Fermi energy. Although at any k-point on the sampling grid the lowest three eigenvalues are occupied, the uppermost filled 'band' is in fact two separate bands overlapping; the matrix element $\langle u_{nk_x} | u_{n(k_x+\delta k_x)} \rangle$ of the periodic parts of the Bloch functions of the highest occupied states at the k points either side of the crossing point therefore vanishes, yielding an infinite $\langle x^2 \rangle_c$. In this case, therefore, $\langle x^2 \rangle_c$ gives no information about the onset of a metal-insulator transition from the insulating side. Figure 4.2(c) shows the localization length in the x direction as the y confinement is reduced. The value of $\langle x^2 \rangle_c$ falls slightly as the eigenfunctions spread out more along y, but because the uppermost unoccupied band does not influence the band below, there is no sign of the transition until it happens.

On the other hand, if instead we fill the three lowest bands and adjust ω_x and ω_y until the top of the p_x band touches the next band up (which has the same reflection symmetry as p_x), the states that become degenerate at the band crossing have the same symmetry under reflection in the x-axis. When k_x is exactly zero, the crossing bands have different y-reflection symmetries and cannot mix, but they hybridize and repel each other for other values of k_x. Figures 4.3(a) and 4.3(b) show the lowest four bands just as the bands touch and shortly after, as ω_x is reduced. The system is never truly metallic, although the gap becomes arbitrarily small at k_x=0. Nevertheless, the rapidly changing nature of the band with k_x means that the overlap integrals $\langle u_{nk_x} | u_{n(k_x+\delta k_x)} \rangle$ become very small around k_x=0, yielding a spike in the localization length near the 'transition', as shown in Fig.4.3(c). In this case, then, the localization length diverges even though no true metal-insulator transition occurs.

A simple one-dimensional tight-binding Hamiltonian that models this situation can be constructed by considering only the highest occupied and lowest unoccupied bands before and after the hybridization. The final eigenstates $\psi_{ik}(x) =$

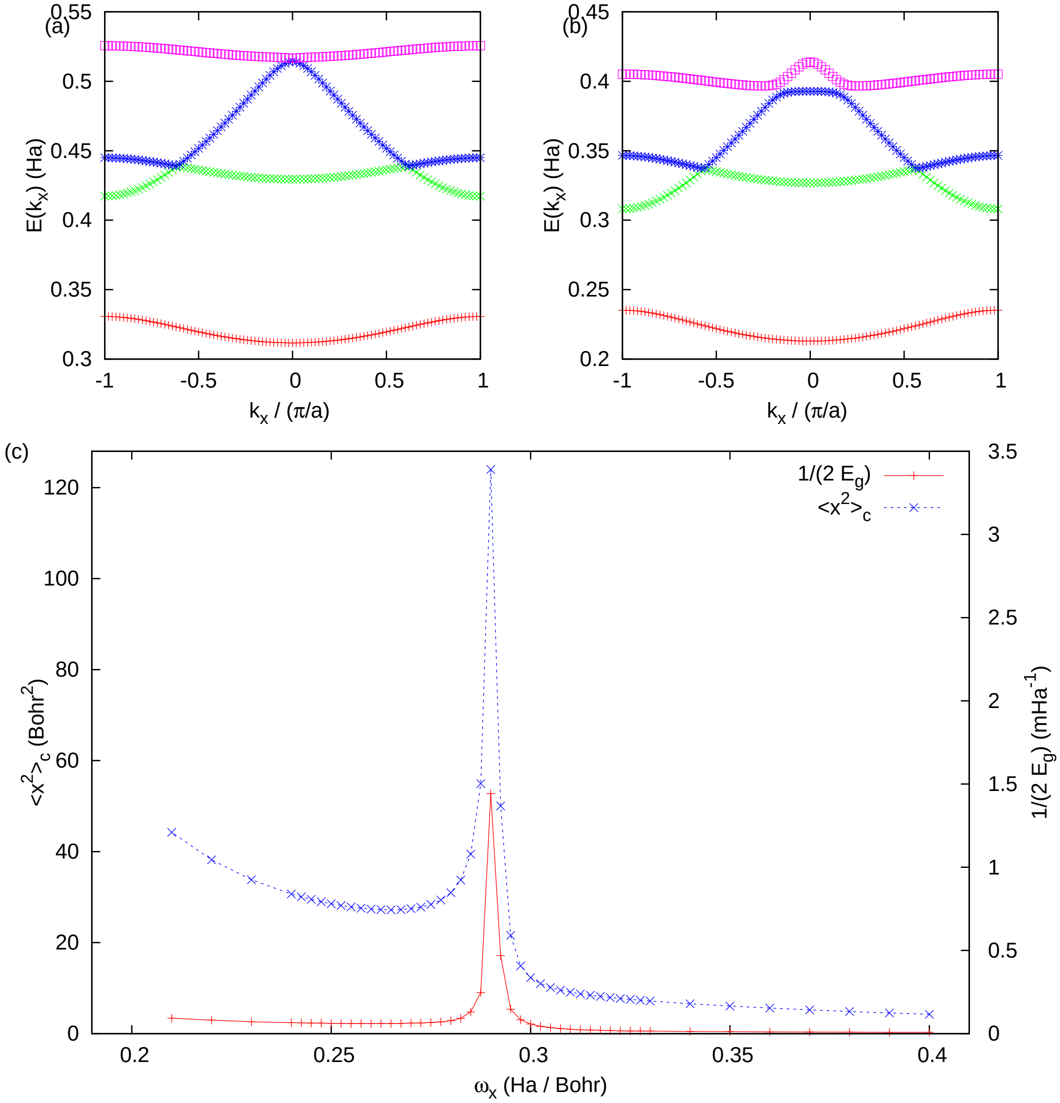

Figure 4.3: One-dimensional chain of 6-electron dots: **(a)** Energy bands $E(k_x)$ before band crossing. **(b)** Energy bands $E(k_x)$ shortly after transition **(c)** Localization length $\langle x^2\rangle_c$ (blue) and inverse direct energy gap (red) to lowest unoccupied state $1/(2E_g)$ as ω is reduced.

$e^{ikx}u_{ik}(x)$ for $i = 1, 2$ are constructed from a linear combination of basis functions $\chi_{ik}(x)$ which are not eigenstates of the Hamiltonian themselves but have a Bloch-like form with orthogonal periodic parts ϕ_{ik}:

$$\chi_{ik}(x) = \phi_{ik}(x)e^{ikx} \tag{4.63}$$

When expressed in the basis of the these functions the Hamiltonian matrix for small values of k takes the form (for an appropriate choice of external potential):

$$\hat{H} = \left(\begin{array}{cc} \Delta + \alpha k^2 & \Gamma k^2 \\ \Gamma k^2 & -\Delta - \beta k^2 \end{array} \right) , \tag{4.64}$$

where α and β describe the curvature of the uncoupled Bloch bands and Γ the matrix element between them.

Diagonalizing this Hamiltonian gives the eigenvalues

$$\epsilon^{\pm}(k) = \frac{(\alpha - \beta)}{2}k^2 \pm \sqrt{\left(\frac{(\alpha + \beta)}{2}k^2 + \Delta\right)^2 + \Gamma^2 k^4} \tag{4.65}$$

This shows that when $\Delta > 0$, the two bands are separated by a direct band gap of 2Δ at k=0. As the value of Δ reduces, the band gap also reduces. The two bands touch briefly when Δ=0, after which, as Δ becomes negative, they hybridize and repel in the manner shown The behaviour of the bands $\epsilon_{\pm}(k)$ and the two components of the eigenvector $\Psi_{-}(k)$ is shown in Fig.4.4.

As $\Delta \rightarrow 0_{-}$ around $k \sim 0$, the state at $k = 0$ remains strictly that of the upper band until $\Delta = 0$ at which point they are degenerate. Therefore, however small δk is, the overlap nearest to $k = 0$ is always between states originating from different bands.

Diagonalizing the Hamiltonian yields the eigenvectors

$$\mathbf{c}^{\pm}(k) = \frac{1}{\sqrt{1 + |p^{\pm}(k)|^2}} \Big[p^{\pm}(k),\ 1 \Big] , \tag{4.66}$$

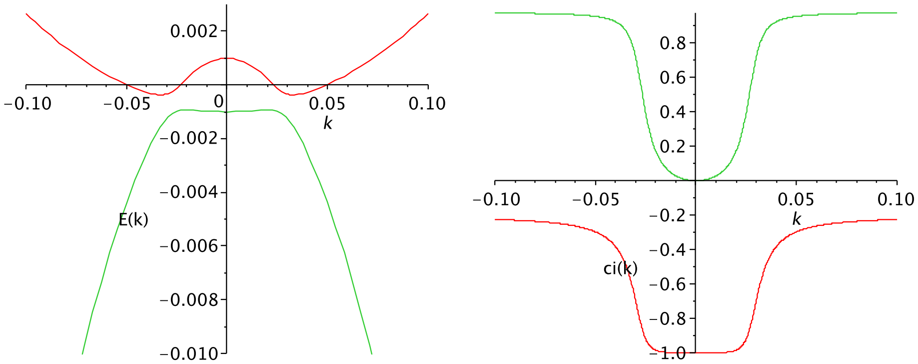

Figure 4.4: Eigenvalues $\epsilon_\pm(k)$ of the two bands of the linear chain model (left), and eigenvector components $c_1(k)$ (red) and $c_2(k)$ (green) of the occupied band $\Psi_-(k)$ in terms of the two basis functions $\phi_{1k}(x)$ and $\phi_{2k}(x)$, for small, negative Δ (right).

where

$$p^{\pm}(k) = \left(\frac{\alpha+\beta}{2\Gamma} + \frac{\Delta}{\Gamma k^2}\right) \pm \sqrt{\left(\frac{\alpha+\beta}{2\Gamma} + \frac{\Delta}{\Gamma k^2}\right)^2 + 1}\ . \tag{4.67}$$

The elements of the eigenvectors are the components of the periodic parts $u_k^{\pm}$ of the Bloch eigenstates $\psi_k^{\pm}$ along the basis functions $\phi_{1k}(x)$ and $\phi_{2k}(x)$:

$$u_k^{\pm}(x) = c_1^{\pm}(k)\phi_{1k}(x) + c_2^{\pm}(k)\phi_{2k}(x)\ . \tag{4.68}$$

Here we make the approximation that the dominant contribution to the localization length comes from the variation of the eigenvectors with k rather than the basis functions themselves. This is equivalent to neglecting the $\frac{\partial \phi_{ik}}{\partial k}$ terms in $\frac{\partial u_{ik}}{\partial k}$. This will inevitably become a valid approximation as $\Delta \to 0$ as the orbitals ϕ_{ik} are largely independent of Δ whereas the functions $p^{\pm}(k)$ vary very rapidly with it around $k = 0$.

Only the lower state is occupied, so the localization length can be calculated by substituting the expression for the periodic part $u_k^-(x)$ of the occupied eigenfunction $\psi_k^-(x)$ into Eq. 4.34. We can calculate the behaviour of the localization length in the limit that the dominant contribution comes from the integral around $k = 0$. We require

$$\langle x^2 \rangle_c = \frac{a}{2\pi} \int \mathrm{d}k \left\langle u^-(k) \middle| \frac{\partial^2 u}{\partial k^2} \right\rangle - \left| \left\langle u^-(k) \middle| \frac{\partial u}{\partial k} \right\rangle \right|^2 \tag{4.69}$$

around $k = 0$. We begin from

$$u(k) = \frac{1}{(1+p^2(k))^{1/2}}\,[p(k), 1] \tag{4.70}$$

dropping the superscripts on u and p, and we express the required derivatives with respect to k in terms of $p(k)$ and its derivatives:

$$\begin{aligned}\frac{\partial u}{\partial k} =& \frac{-p(k)}{(1+p^2(k))^{3/2}}\frac{\partial p}{\partial k}\,[p(k), 1] + \frac{1}{(1+p^2(k))^{1/2}}\left[\frac{\partial p}{\partial k}, 0\right]\\ =& \frac{1}{(1+p^2(k))^{3/2}}\frac{\partial p}{\partial k}\,[1, -p(k)]\end{aligned} \tag{4.71}$$

and

$$\begin{aligned}\frac{\partial^2 u}{\partial k^2} =& \frac{-3p(k)\left(\dfrac{\partial p}{\partial k}\right)^2}{(1+p^2(k))^{5/2}}\,[1, -p(k)] + \frac{\dfrac{\partial^2 p}{\partial k^2}}{(1+p^2(k))^{3/2}}\,[1, -p(k)]\\ &+ \frac{\dfrac{\partial p}{\partial k}}{(1+p^2(k))^{3/2}}\left[0, -\frac{\partial p}{\partial k}\right]\\ =& \frac{1}{(1+p^2(k))^{3/2}}\Bigg[\frac{\partial^2 p}{\partial k^2} - 3p(k)\left(\frac{\partial p}{\partial k}\right)^2/(1+p^2(k)),\\ &-p(k)\frac{\partial^2 p}{\partial k^2} - \left(1 - \frac{3p^2(k)}{(1+p^2(k))}\right)\left(\frac{\partial p}{\partial k}\right)^2\Bigg]\end{aligned} \tag{4.72}$$

Inserting these into the integrand of Eq. 4.69, most of the terms cancel:

$$\begin{aligned}\left\langle u(k)\middle|\frac{\partial u}{\partial k}\right\rangle =& \frac{1-1}{(1+p^2(k))}p(k)\frac{\partial p}{\partial k} = 0\\ \left\langle u(k)\middle|\frac{\partial^2 u}{\partial k^2}\right\rangle =& \frac{1}{(1+p^2(k))^2}\Bigg[p(k)\frac{\partial^2 p}{\partial k^2} - -\frac{3p^2(k)}{(1+p^2(k))}\left(\frac{\partial p}{\partial k}\right)^2\\ &- p(k)\frac{\partial^2 p}{\partial k^2} - \left(1 - \frac{3p^2(k)}{(1+p^2(k))}\right)\left(\frac{\partial p}{\partial k}\right)^2\Bigg]\\ =& \frac{-1}{(1+p^2(k))^2}\left(\frac{\partial p}{\partial k}\right)^2\end{aligned} \tag{4.73}$$

We can therefore express $\langle x^2 \rangle_c$ as

$$\langle x^2 \rangle_c = \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} \frac{1}{(1+(p(k))^2)^2} \left(\frac{\partial p(k)}{\partial k} \right)^2 dk \ . \tag{4.74}$$

For fixed values of α, β and Γ, the scaling behaviour with the gap parameter Δ may be extracted by changing variables to $u = \Delta/k^2$. We can write $p(k)$ as a function of u:

$$p(u) = \frac{(H+u) - \sqrt{(H+u)^2 + \Gamma^2}}{\Gamma} \ , \tag{4.75}$$

where $H = (\alpha + \beta)/2$. Noting that the integral is symmetric about $k = 0$, we obtain

$$\langle x^2 \rangle_c = \frac{2a}{\pi\sqrt{\Delta}} \int_{\Delta a^2/\pi^2}^{\infty} \frac{u^{3/2} \left(\frac{\partial p^-(u)}{\partial u} \right)^2}{(1+(p^-(u))^2)^2} \, \mathrm{d}u \ , \tag{4.76}$$

which, for small enough values of Δ, becomes independent of Δ except in the prefactor as long as the integrand remains finite. The integrand can be shown not to contain negative powers of u by considering the leading orders of u in each of its terms. Hence, for small Δ, the quadratic spread $\langle x^2 \rangle_c \propto \Delta^{-1/2}$. A numerical evaluation of Eq. 4.76 is shown in Figure 4.5, and is roughly consistent with the behaviour seen on the right hand side of the transition in Figure 4.3. However, because of finite size effects and convergence issues, it proves difficult to obtain data close enough to the point where the gap closes to determine if this relationship holds accurately. Additionally, the constant term present due to the variation of the ϕ_{ik} terms with k that will obscure this scaling behaviour except very close to the transition.

4.5.3 2D Arrays

This arrangement restores the two-dimensional symmetry and returns to potentials of the type described in Section 4.5.1, with a square lattice of unit cells each containing a single dot. In this situation, with the potential and the lattice symmetric in x and y, the p_x and p_y bands are degenerate along ΓM. By filling the first three bands and gradually reducing the strength of the confining potential, we can still drive the

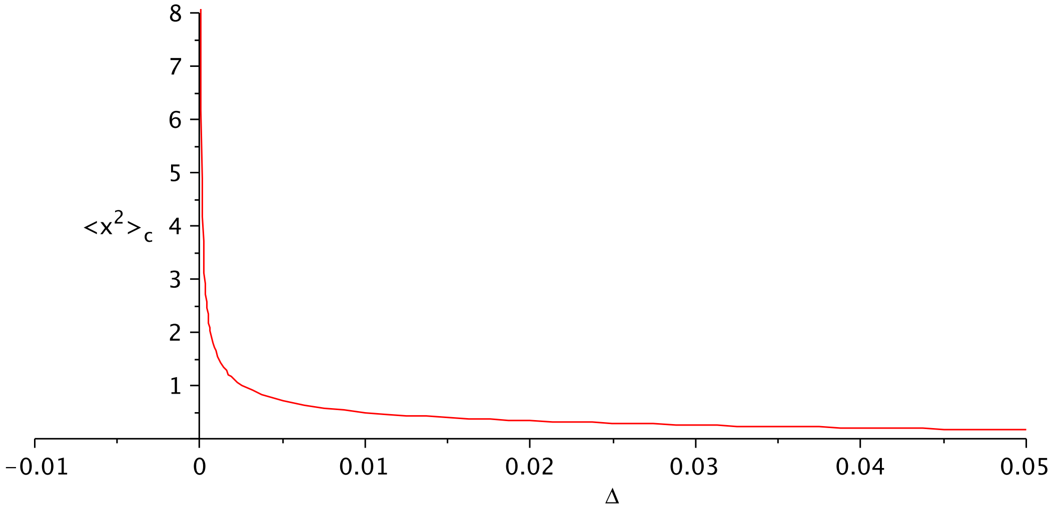

Figure 4.5: Two-band tight binding model corresponding to the system shown in Figure 4.3. As $\Delta \to 0$, the localization length diverges as $\Delta^{-1/2}$. The scale should not be taken as comparable to that of Figure 4.3 as it depends on unknown parameters H and Γ as seen in Eq. 4.75.

system through a transition to a metallic state, but this now occurs when the indirect gap above the p bands closes, as can be seen from Figure 4.6(a). Because the highest occupied and lowest unoccupied states are well separated in reciprocal space, the occupied states do not change nature or become strongly dependent on $\mathbf{k}$ as the gap closes, and the matrix elements between the periodic parts of the eigenfunctions at neighbouring $\mathbf{k}$ vectors behave smoothly. The localization length increases as the bands are brought together because the confining potential is being reduced, but it does not diverge on approach to the transition. This emphasizes again the importance of the approaching bands being able to 'see' each other if anything is to be observed in $\langle x^2 \rangle_c$.

The Kohn-Sham orbitals determined from this DFT simulation were combined with an optimized 1- and 2-electron Jastrow factor to evaluate energies and localization lengths within the many-electron variational quantum Monte Carlo (VMC) method, using the CASINO program [125]. The variance minimization procedure used to optimize the Jastrow factor resulted in close agreement between the DFT and VMC energies. As a consistency check, we confirmed that localization

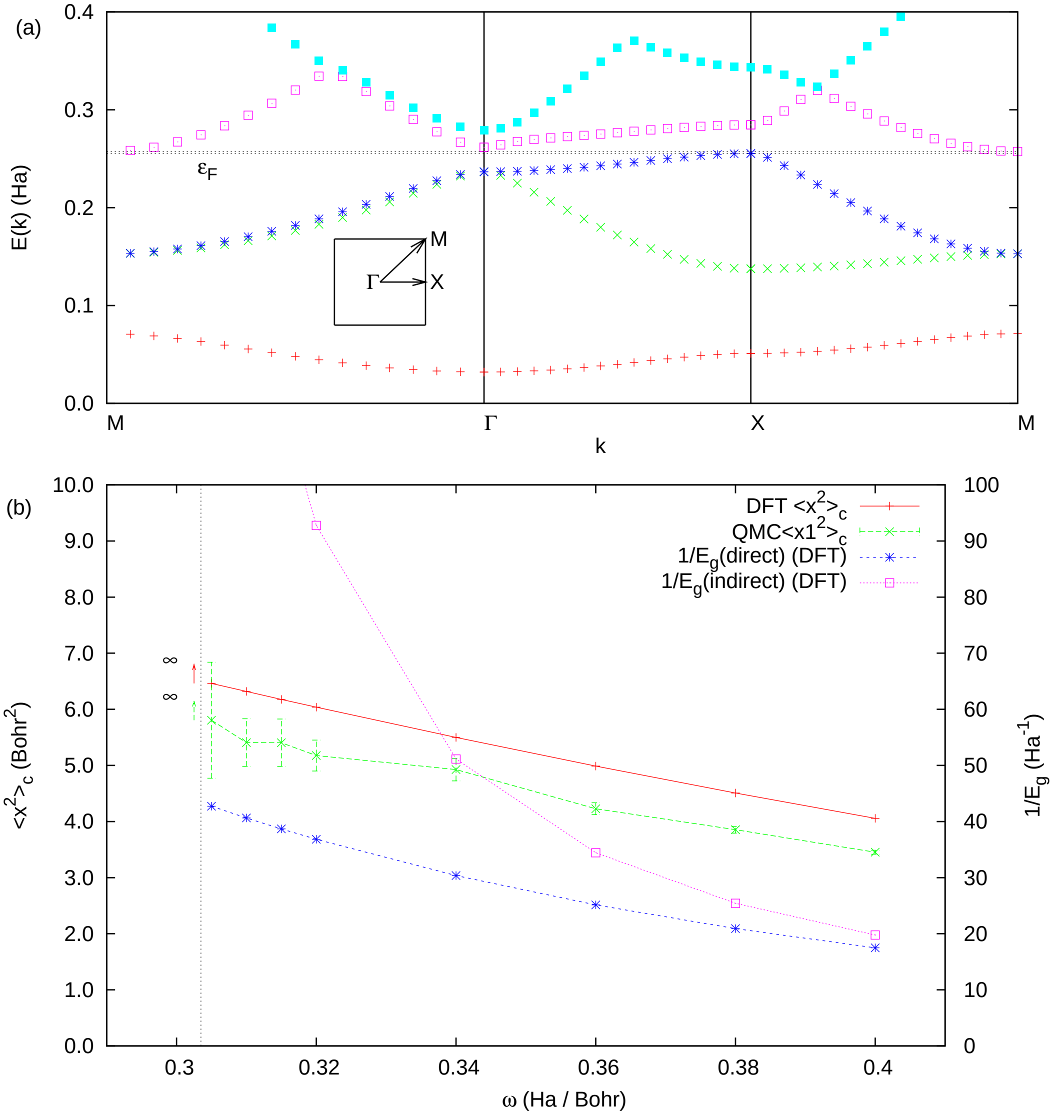

Figure 4.6: **(a)** Energy bands $E(\mathbf{k})$ along the high symmetry lines (inset: Brillouin zone of two-dimensional square lattice). Dotted lines denote the Fermi level. **(b)** Localization length $\langle x^2 \rangle_c$ in DFT and QMC as the dot confinement ω is varied (left scale). Inverse DFT direct and indirect energy gaps (right scale). Below ω =0.305 the system is metallic so $\langle x^2 \rangle_c$ becomes abruptly infinite. No divergence is seen while approaching the transition from the region of stronger confinement.

lengths calculated using VMC trial functions consisting of a Slater determinant of Kohn-Sham orbitals only, with no Jastrow factor, agreed with those obtained using the DFT-based method of 4.4.4 to within the statistical error of the Monte Carlo simulation.

Figure 4.6 shows the localization length of this indirect-gap system calculated using both DFT and VMC. As expected, decreasing the confinement by moving from right to left across the figure results in a gradual increase in the localization length, followed by a discontinuous jump to an infinite value when the indirect gap closes and some of the bands become partially filled. As discussed in 4.4.5, it is difficult to evaluate very small values of z_N (and hence very large localization lengths) using VMC because the statistical errors begin to overwhelm the result. It is clear, however, that the VMC localization length is shorter than the DFT localization length and appears to track it. To the extent that it is possible to judge, it appears that the DFT and VMC localization lengths jump discontinuously to infinity at the same point, presumably because the VMC Slater determinant is constructed using the occupied set of DFT orbitals, which changes discontinuously when the DFT gap closes. If the orbitals in the VMC determinant had been chosen to minimize the VMC energy, the VMC transition would presumably have occurred at a smaller value of ω.

4.5.4 Graphene Model

The hexagonal lattice model of a graphene sheet provides a realistic example in which a direct gap can be tuned down to zero. Real graphene consists of an equilateral triangular (hexagonal) two-dimensional lattice with a basis of two carbon atoms, labelled A and B in Figure 4.7(a), and four valence electrons per atom. Three electrons per atom form strong bonds with their nearest neighbours, leaving the electronic properties dominated by the one electron per atom in the π bonds. These form a bonding π band and and anti-bonding π^* band. A tight-binding analysis involving just these bands, as described in [174, 146], demonstrates the main features of the bandstructure: the two electrons per primitive cell fill the π band, which touches the π^* band at the corners of the hexagonal Brillouin zone, labelled $\mathbf{K}$ in Figure 4.7(b). Graphene is thus a semiconductor with a zero gap, and has interesting transport properties caused by the zero density of states at the Fermi level. Since the π band is fully occupied, its Wannier functions might be expected to be localized; on the other hand, since the band gap is zero, the upper bound on the localization length provided by the conductivity formula of Souza, Wilkens and Martin (Eq. 52 of [156]) is infinite.

If the potential is modified so that the energies of the two sites in the primitive cell are no longer equivalent, a gap opens up at the Fermi energy and grows in proportion to the difference of the on-site energies. We model this by placing Gaussian dots of the form

$$V_A(\mathbf{r}) = -V_d \exp(-|\mathbf{r} - \mathbf{r}_A|^2/\rho^2) \tag{4.77}$$

on the A sites and

$$V_B(\mathbf{r}) = -(1+\Delta)V_d \exp(-|\mathbf{r} - \mathbf{r}_B|^2/\rho^2) \tag{4.78}$$

on the B sites. Figure 4.7 shows the form of this potential when $\Delta = 0$.

At finite Δ there is an energy gap, caused by the attraction of the electrons to the lower-energy B sites. As $\Delta \to 0$, the energy gap reduces to zero and the localization length increases, as shown in Figure 4.8. The value of z_N at $\Delta = 0$ remains non-zero (and hence the localization length remains finite) for all practical system sizes. It is unclear whether the touching of the valence and conduction bands at a single point in k space is sufficient to force z_N to tend to zero as the system size tends to infinity.

It is again possible to derive a tight-binding description of this system — its derivation is very similar to that in [146] for the graphene sheet, except that the on-site energies of the A and B sites are allowed to differ by an amount Δ. However, in

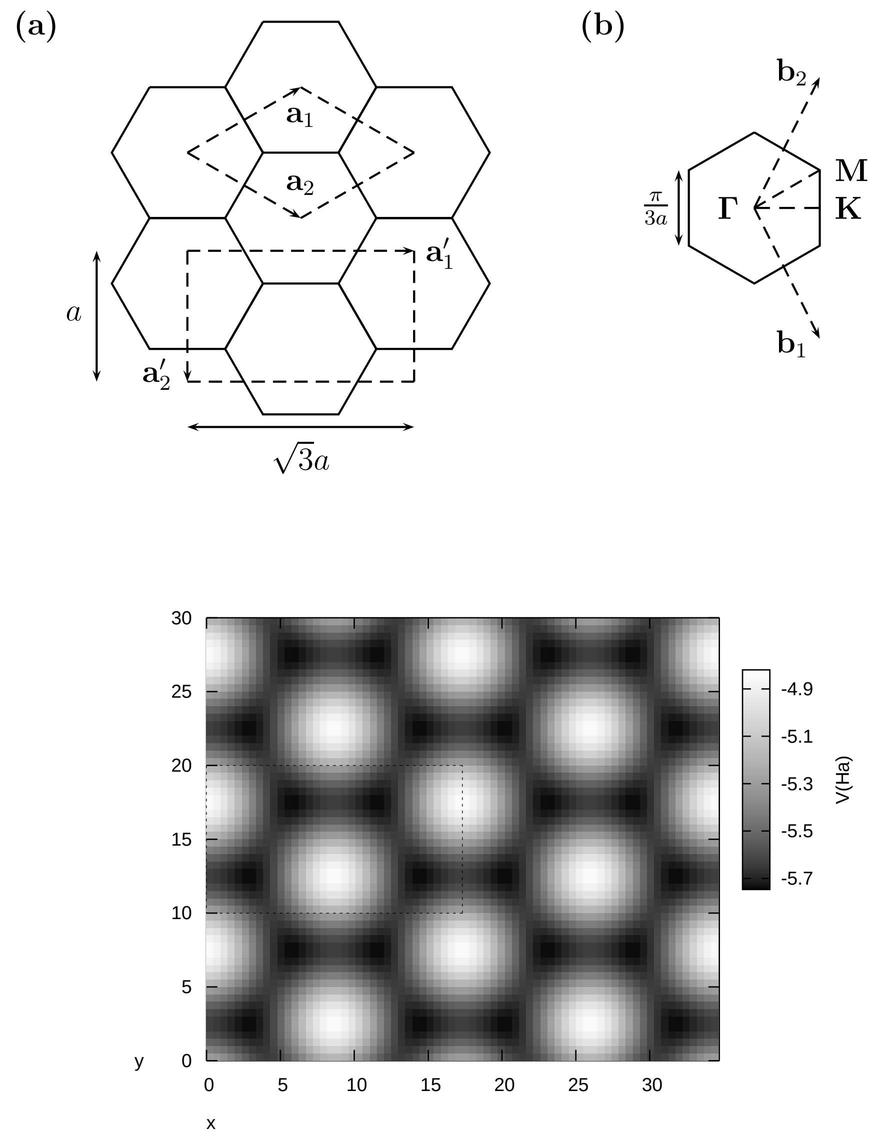

Figure 4.7: **(a)** The structure of the graphene-like model, showing the lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$ of the primitive cell (top), the lattice vectors $\mathbf{a}'_1$, $\mathbf{a}'_2$ of the larger rectangular cell used for simplicity (bottom), and the sites A and B on which the potential can be varied. **(b)** The Brillouin zone of the graphene-like model showing the reciprocal lattice vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and the high symmetry $\mathbf{k}$-points $\mathbf{K}$, $\mathbf{M}$ and $\mathbf{\Gamma}$. **(c)** The external potential of the graphene-like model for $a = 10$, $V_d = 3$, $\rho = 5$, and $\Delta = 0$. The dotted line encloses one rectangular unit cell.

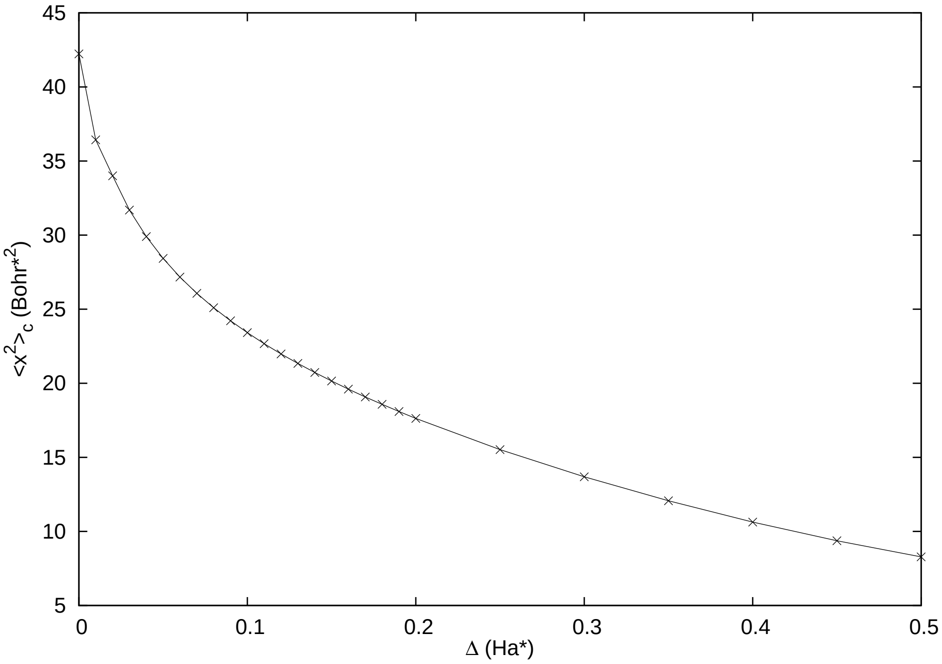

Figure 4.8: Localization length $\langle x^2\rangle_c$ as Δ is reduced to 0, at which point there is a transition to a semimetallic state. Available system sizes up to the $N_k = 36 \times 36$ system shown here are insufficient to resolve whether the localization length is truly diverging as the semimetallic state is approached. Even for $\Delta = 0$ a finite value of $\langle x^2\rangle_c$ is observed at all practical system sizes.

this situation the resulting 2D Brillouin zone integral is not analytically tractable. It could be numerically integrated, but this would not provide any insight not already given by the DFT simulation.

Chapter 5

The Surface Energy of the Electron Gas

5.1 The Jellium System and Surface

5.1.1 Introduction and Previous Investigations

We established in Chapter 2 the importance of the exchange-correlation functional in the accurate determination of energies with density functional theory. As mentioned, while the behaviour of the exchange part of the functional can be derived analytically in homogeneous systems, in inhomogeneous systems, and for the correlation part even in homogeneous systems, the functional must in general be either derived by approximations or fit to computational data from a higher level of theory — predominantly QMC.

It is therefore a topic of importance to test and benchmark the performance of such functionals. Historically, jellium, the homogeneous electron gas with a uniform positive background, has been the proving ground for all new electronic structure methods. For example, among the early uses of QMC [31] and GW [58] were such calculations — the former forming the basis for the Perdew-Zunger parameterization [133] of the Ceperley-Alder data for the correlation energy in the LDA functional. The jellium slab, or quasi-2D electron gas, is the simplest useful further test that methods can be subjected to beyond the homogeneous system. It is the simplest system with a surface, and thus provides a test of functionals over a wide range of densities, as the electron density must tail away from its bulk value deep inside the material, to zero far from the surface.

The first study of such a surface with DFT, that of Lang and Kohn [104], was also among the first real applications of LDA. This has been followed by studies with exchange-correlation functionals of gradually increasing sophistication [105, 131], and studies with the Fermi Hypernetted Chain method [96, 97], and the Random Phase Approximation (RPA) [98] and GW methods [44]. The paper by Yan [185] in 2000 summarizes the results of a variety of calculations as of that date, and the later paper by Staroverov [157] summarizes those as of 2004. The conclusion of tests of a wide range of DFT functionals is that the surface energy in DFT is only lightly dependent on the choice of exchange-correlation functional, but that designing a local functional to incorporate the correct long- and short-ranged behaviour simultaneously is a difficult prospect. The DFT results fall in a relatively narrow range, as can be seen from Table 5.1.

Applying QMC to this problem has long been an attractive prospect due to its high level of accuracy, and, in simple systems such as this, its relative absence of approximation. QMC should be able in principal to provide a benchmark against which the success of other approximations can be tested. However, for reasons that will be elaborated on in the following section, previous QMC work has been of mixed success. The first fixed-node DMC simulations were by Li et al. [111], and their work was expanded on by Acioli and Ceperley [1]. Both these groups studied slab systems to which periodic boundary conditions had been applied, and used the direct bulk vs slab comparison described in Section 5.2.2 to calculate surface energies. The results of these QMC simulations differed so substantially from the aforementioned DFT calculations that further investigation was required — this was provided by Sottile and Ballone [155], who performed DMC simulations on jellium spheres, for which the finite size effects are much less severe. Their results were much more in line with the DFT work — especially with DFT calculations of jellium spheres

Author and Reference	σ / erg cm^{-2}	Method
Yan et al. [185]	-610	DFT-LDA
Yan et al. [185]	-533	DFT-GGA + WVI*
Perdew et al. [131]	-690	DFT-GGA
Perdew et al. [131]	-567	DFT-metaGGA
Kurth and Perdew [98]	-533	DFT-LDA + RPA
Kurth and Perdew [98]	-587	DFT-GGA + RPA
Li et al. [111]	-465 ± 50	FN-DMC (slab vs bulk)
Acioli and Ceperley [1]	-429 ± 80	FN-DMC slab vs RN-DMC bulk
Correction to above [135]	-554 ± 80	FN-DMC slab vs FN-DMC bulk
Wood [178]	-600 ± 50	VMC slab only

Table 5.1: Comparison of surface energy σ at a specific density ($r_s = 2.07$, corresponding to the density of Aluminium) as calculated in a range of methods and variations thereof.

carried out by Almeida et al [5]. This presented something of a controversy, as while DMC was expected to be the most accurate method available, it did not appear to be giving reliable answers in slab systems. Subsequently, Pitarke [135] pointed out that the comparison in [1] was somewhat flawed as it compared release-node bulk calculations with fixed-node slab calculations. Correction of this error moved the DMC value somewhat towards the rest but the controversy remained.

Table 5.1 shows summarized results of surface energies of slab calculations at one specific density with a range of methods, highlighting the difference between the DMC results and all the other methods, and also the variation between various DFT functionals and other approximations. Surface energies are, for historical reasons, often quoted in units of erg cm^{-2}, from which the conversion factor to the more useful atomic units of surface energy, mHa bohr^{-2}, is 1/1556.8928.

The work on the surface energy of jellium presented here follows on from the work of Wood and Foulkes [178, 179, 180] which addressed the challenges faced by DMC simulations of Jellium slabs. Wood and Foulkes extended the Model Periodic Coulomb (MPC) interaction [177] to quasi-2D systems of this type [179] and designed a new form of wavefunction based on the plasmon normal modes present in the bulk and on the surface of a jellium slab [180]. However, while they were successful in identifying many of the sources of error in the original slab calculations, and in formulating a new method for the determination of the surface energy itself, their VMC-only calculations [178] were unable to obtain sufficient accuracy in total energies fully to resolve this controversy.

In this Chapter, we present an analysis of the sources of error in slab calculations in VMC and DMC, and present a comprehensive evaluation within DMC of the surface energy at a range of densities. In Section 5.1.2 we define the system under study. In Section 5.2.1 we discuss the generation of trial wavefunctions for DMC calculations of jellium slabs, and in Sections 5.2.2 and 5.2.3 we compare different methods of evaluating the surface energy itself. Finally, in Section 5.2.4 we present our results for the surface energy, and in 5.2.5 we discuss how these results can be used to estimate the magnitude of the fixed node errors, which have to date generally been thought of as an uncontrolled and unquantifiable approximation. Given the difference between the DMC results of Acioli and Ceperley, who mixed release-node DMC and fixed-node calculations, compared to the results when only fixed-node results were used, this is an important estimate to be able to make.

5.1.2 The Jellium Slab System

The homogeneous electron gas is characterized by a single parameter r_s, the radius of a sphere of size equal to the volume occupied by one electron. This volume is equal to the total volume over the number of electrons, which is of course the inverse of the density n, so

$$n = \frac{3}{4\pi r_s^3} \;\Rightarrow\; r_s = \left(\frac{4\pi n}{3}\right)^{-1/3} . \qquad (5.1)$$

The negative charge of the electrons is neutralized with a uniform positive background of the same density n, and the total energy per electron of the electrons in this system is defined as ϵ_{bulk}, which is a function of r_s — see for example [31].

To create a jellium slab, we carry over this definition of r_s in terms of the density, but cut off the positive background at a certain width s in one spatial direction (usually chosen as z). The background thus has the form (see also Figure

5.2)

$$n_b(z) = \begin{cases} 3/(4\pi r_s^3) & -s/2 < z < s/2 \; , \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

Because the density and the potential do not depend on x and y (the directions that lie in the plane of the slab), the system is homogeneous in these directions. In a DFT calculation, the in-plane symmetry allows us to consider, effectively, an infinite sized unit cell in those directions. This is achieved by performing the in-plane integration with respect to k_x and k_y analytically, and allows the number of electrons in the system to be infinite, negating the in-plane finite size effects examined in Section 5.2. In a QMC calculation, however, the unit cell has to be finite, with a fixed number of electrons N occupying a square in-plane unit cell of length L obeying periodic boundary conditions. The volume of background is thus $L^2 s$ and

$$N = nL^2 s = \frac{3L^2 s}{4\pi r_s^3} \tag{5.3}$$

Figure 5.1 shows a typical periodic unit cell for this system.

In previous QMC calculations, it has been standard to use periodic boundary conditions along all three axes, such that what is in fact simulated is an infinite stack of slabs separated by w. This leads to extra finite size errors as there is cross-talk between neighbouring slabs. Wood used the quasi-2D Ewald interaction originally proposed by Parry, enabling this error to be escaped by studying only a single slab. This eliminates cross-talk, but is slow and cumbersome to evaluate. Fortunately, Wood also showed that equivalent accuracy can be obtained by replacing the quasi-2D Ewald interaction with the quasi-2D form [179] of the Model Periodic Coulomb (MPC) interaction [177].

It is clear that deep enough inside a sufficiently wide slab system, the local value of the energy per electron must tend to its bulk value. In a finite slab system with two surfaces of area $A = L^2$ and N electrons, we can therefore define the energy

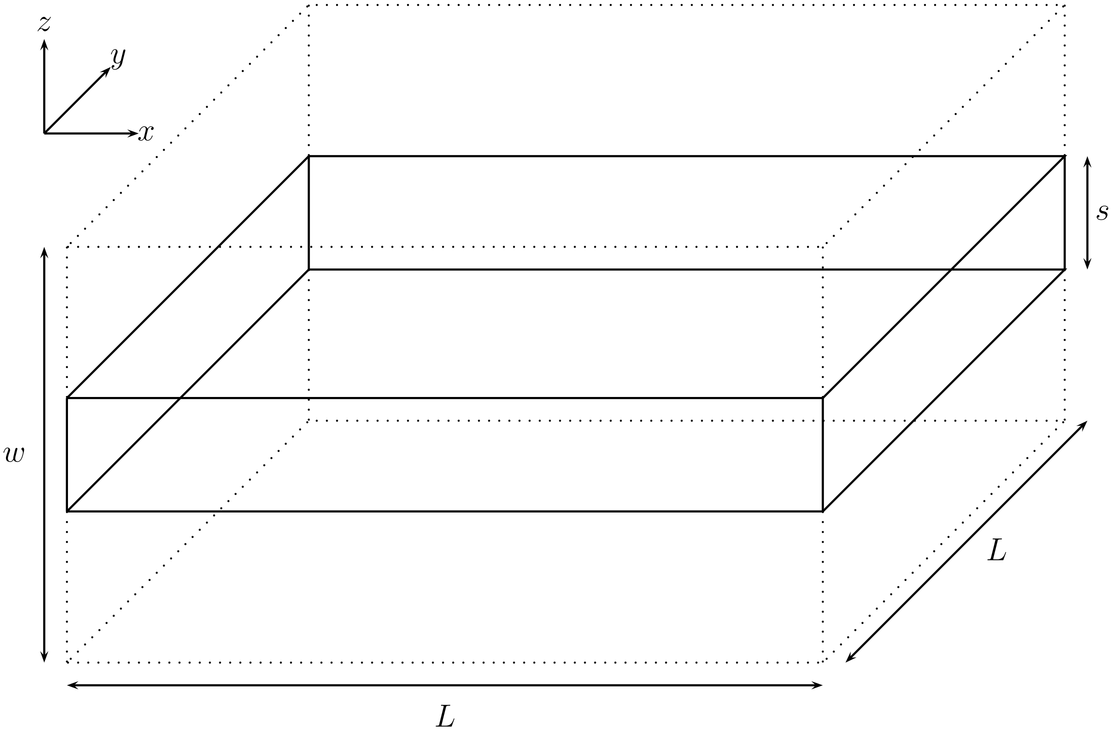

Figure 5.1: The simulation cell for a jellium slab calculation in QMC. The slab is of width s, enclosed in a box of width w, and is of size $L \times L$ in the plane. The density parameter r_s and the volume of slab enclosed, sL^2, determine the electron number N.

per electron in the slab in terms of that in the bulk and the surface energy σ, with

$$\epsilon_{slab} = \epsilon_{bulk} + \frac{2A\sigma}{N} \ . \tag{5.4}$$

Combining this with 5.3 gives us

$$\epsilon_{slab} = \epsilon_{bulk} + \frac{8\pi r_s^3 \sigma}{3s} \ . \tag{5.5}$$

Strictly, this formula only holds in the limit of an infinitely wide slab, as for finite-width slabs ϵ_{slab} displays oscillations as a function of slab width. Unfortunately, however, as $s \to \infty$, the bulk and slab values of ϵ become the same and the above equation cannot be used to find the surface energy. These out-of-plane finite size effects mean that surface energy calculations must be performed at a large but finite

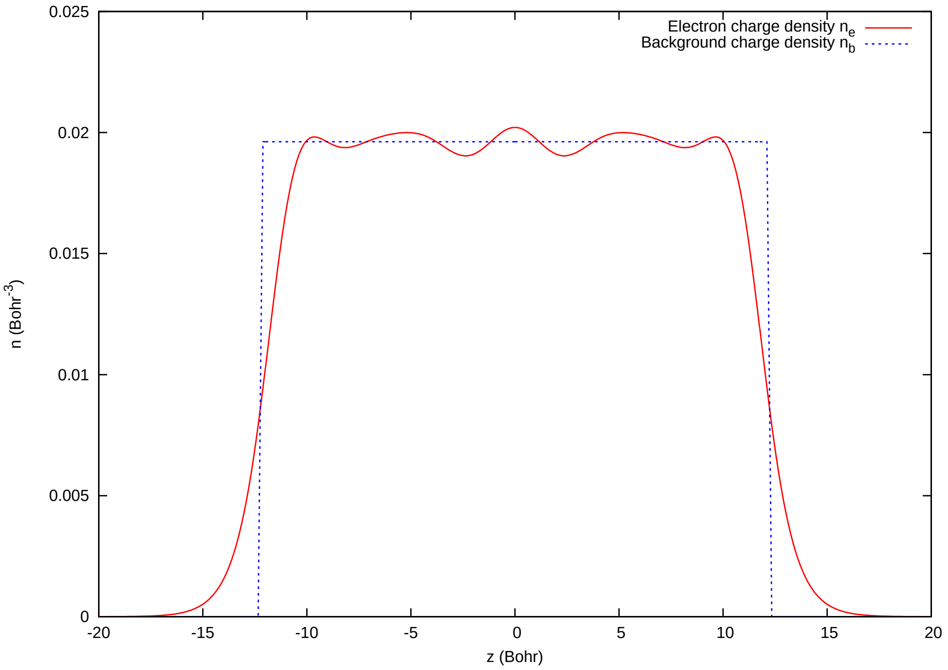

Figure 5.2: Electron density and background charge density along the z-direction, for a jellium slab at $r_s = 2.30$ Bohr and $s = 24.5$ Bohr. The electron density spills out of the slab by around 3 Bohr, and displays the oscillations characteristic of the first 5 occupied subbands.

value of s.

The electron density in a jellium slab is normally allowed to spill outside the region of the background charge to minimize the energy. Figure 5.2 shows a typical electron density compared to the corresponding background. It is also possible to introduce fixed hard boundaries at the edge of the slab — in effect an infinite potential well — so that the electrons are confined to the positive background region inside the slab. This is usually thought of as a different problem, however, as the behaviour of the electron density at the boundaries of the slab is very different from that of the jellium slab. It is referred to in the literature as the Infinite Barrier Model (see eg. [136]).

5.2 Surface Energy Calculations

The surface energy of jellium can be broken down, just like the total energy, into contributions from kinetic energy, external potential energy (the interaction between the electrons and the background charge), and electron-electron potential energy. The electron-electron energy can be further broken down, in a way that is explicit within density functional theory but must be inferred within quantum Monte Carlo, into a Hartree electrostatic term and an exchange-correlation term. We thus write the total surface energy σ_T, with σ_s the kinetic term, σ_{es} the electrostatic term from the Hartree electron-electron and background potentials, and σ_{XC} the exchange-correlation term, as

$$\sigma_T = \sigma_s + \sigma_{es} + \sigma_{XC} \tag{5.6}$$

One of the principal difficulties of making an accurate determination of σ in a jellium slab is that all three of these terms are individually large but nearly cancel to produce a relatively small surface energy. This should be no surprise, as the energy cost of cutting off a homogeneous background and a near-homogeneous electron gas, when the gas is allowed to spread out from the surface to minimize the energy, would be expected to be small. Example values for the breakdown of the surface energy into its components, taken from a DFT calculation with $r_s = 2.30\,\mathrm{bohr}$, are as follows:

- $\sigma_T = -104.50\,\mathrm{erg\,cm^{-2}}$ (total surface energy)
- $\sigma_s = -2801.60\,\mathrm{erg\,cm^{-2}}$ (kinetic energy contribution)
- $\sigma_{es} = 650.39\,\mathrm{erg\,cm^{-2}}$ (electrostatic energy contribution)
- $\sigma_{XC} = 2046.71\,\mathrm{erg\,cm^{-2}}$ (exchange-correlation energy contribution)

Often in the literature, given that it represents the only uncertain part of the calculation, only the exchange-correlation part σ_{XC} of σ_T is quoted. In QMC, there is no easy way to separate out the terms, as to disentangle the Hartree and XC contributions to the electron-electron interaction energy would require knowledge of the ground state density. Since the density operator does not commute with the Hamiltonian, only relatively inaccurate mixed estimator of its expectation value could be obtained. Moreover, since the ground state is anyway not an eigenstate of the electron-electron potential energy operator, the variance is high and the error bars large. We therefore stick, in this work, to quoting the full surface energy σ_T.

In the following subsections we discuss how to model the Jellium slab in DFT and QMC, and then how to extract the surface energy from these simulations in an accurate way.

5.2.1 Wavefunctions for Jellium Slabs in Quantum Monte Carlo

As ever, to perform accurate calculations in quantum Monte Carlo, we require accurate trial wavefunctions. In these calculations, we choose to build our wavefunctions on single particle orbitals from density functional theory. Because of the homogeneity in x and y, the DFT single-particle orbitals have the simple form

$$\phi_{n\mathbf{k}_{||}}(x,y,z) = u_n(z)\,\mathrm{e}^{i\mathbf{k}_{||}\cdot\mathbf{r}_{||}} \tag{5.7}$$

where n is the index of the subband in the out-of-plane direction, $\mathbf{k}_{||}$ is the in-plane k-vector, and $\mathbf{r}_{||} = (x, y, 0)$ is the in-plane position. The density is uniform in x, y and is easy to calculate from the orbitals, as

$$n_e(z) = \sum_{n}^{n_b} \omega_n |u_n(z)|^2 \tag{5.8}$$

where ω_n is a weight factor specifying the occupation of subband n. An example of the orbitals of a slab with five occupied subbands ($n_b = 5$) at $r_s = 2.30$ is shown in Figure 5.3. The filling of these subbands is responsible for the finite size effects discussed in the next section.

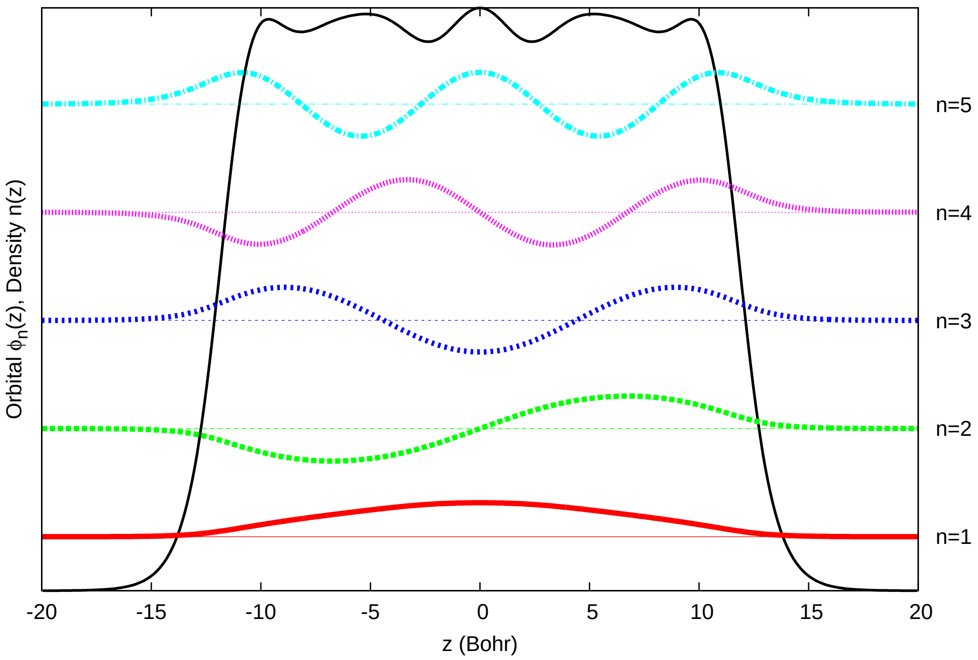

Figure 5.3: Orbitals of the 5 occupied subbands for the same slab as Figure 5.2 ($r_s = 2.30$, $s = 24.5$). Each subband is in a different colour and the axes of each have been shifted for clarity — the straight lines of the same colour mark the zero of each orbital. The density, $n_e(z) = \sum_i |\phi_i(z)|^2$ is shown in black.

The DFT simulations are carried out with a code supplied by Pablo Garcia-Gonzalez [57], which uses a grid in the z-direction extending outside the slab for some distance, giving a total cell size of w in the z-direction. In the xy-plane, the cell can be either of length L or infinite. Convergence with respect to parameters such as the grid spacing and cell width w is detailed in the work of Wood [178]. To perform the GGA calculations, the ability to use the PBE functional was added to this code by this author.

The QMC calculations use these DFT orbitals in a Slater determinant D and combine them with a Jastrow factor J to produce the many-electron wavefunction $\Psi(\mathbf{X})$, where $\mathbf{X} = \{(\mathbf{r}_i, \sigma_i)\}$:

$$\Psi(\mathbf{X}) = \mathrm{e}^{J(\mathbf{X})} D^{\uparrow}(\mathbf{R}^{\uparrow}) D^{\downarrow}(\mathbf{R}^{\downarrow}) \ . \tag{5.9}$$

An immediate difficulty is encountered when attempting to optimize the parameters in the Jastrow factors of these trial wavefunctions by applying variance minimization. If the standard power series expansion in electron-electron distances is used for the Jastrow, then the variance minimization, at least in the traditional scheme, fails spectacularly, for reasons documented previously [178]. The very recent scheme of Drummond and Needs [41] may meet with more success but was not available at the time of these calculations.

A possible solution to this optimization problem [180] involved deriving the form of the bulk and surface plasmons in the slab and basing the long-ranged part of the Jastrow factor on the behaviour of the plasmon normal modes [178]. While the approach successfully provided a theoretical description of the plasmons, attempts to use it as a basis for the long-ranged part of the Jastrow were unsuccessful. However, it did point the way to a usable form of short-range cusp term in the Jastrow and confirmed that the method of Malatesta et al. [114] could be used to ensure that the one-body terms $\chi(\mathbf{r})$ undo the unwanted spreading effect on the density of the two-body $u(r_{12})$ terms.

In this work, we use a Jastrow factor of the form

$$J(\mathbf{X}) = -\frac{1}{2}\sum_{i\neq j} u_{cusp}(\mathbf{x}_i,\mathbf{x}_j) + \sum_i \chi(\mathbf{r}_i)\,, \tag{5.10}$$

where

$$u_{cusp}(\mathbf{x}_i,\mathbf{x}_j) = \frac{\alpha}{2(1+\delta_{\sigma_i\sigma_j})}\mathrm{e}^{-r_{ij}/\alpha - r_{ij}^2/L_c^2}\,, \tag{5.11}$$

and χ takes the Malatesta [114] form,

$$\chi(\mathbf{r}_i) = \int_V u_{cusp}(\mathbf{r}_i,\mathbf{r}')\,n(z')\,\mathrm{d}^3\mathbf{r}'\,, \tag{5.12}$$

where r_{ij} is the separation of electrons i and j, $n(z)$ is the slab electron density along the z axis, and L_c and α are the only unknown parameters. The prefactor to the exponential ensures the cusp conditions as $r_{ij} \to 0$ are obeyed.

We fix L_c by using it to ensure that there is no unwanted cusp in u at the cell boundaries. Any such cusp would produce a contribution to the kinetic energy of

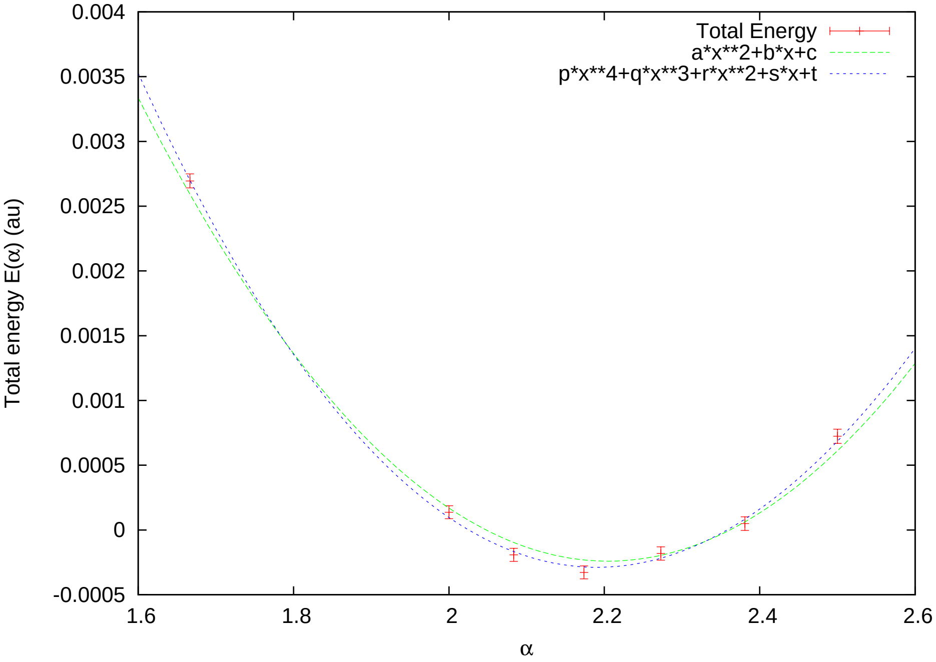

Figure 5.4: Optimization of Jastrow parameter α. Total energies for a range of values of α are calculated in VMC and a quartic fit to the curve is used to find the minimum where $\partial E/\partial\alpha = 0$.

the wavefunction that would be δ-function like, so would not be sampled in a QMC simulation. To ensure the cusp contributes negligibly, we require that as $r_{ij} \to L$, the u term has decayed to nearly zero, and thus fix L_c at a constant fraction of the cell size L. This leaves α as the only optimizable parameter, controlling the range of the short range correlation.

Because of the failure of variance minimization, the only practical way to optimize α is to do it “by-hand”. Figure 5.4 shows how this is performed: the u terms and corresponding χ terms for a series of values of α are generated according to Eqs. 5.11 and 5.12, and a minimal VMC run is performed for each. By fitting a quadratic or quartic function (the quartic usually provided a more accurate minimum) to the plot of $E_{VMC}(\alpha)$, a minimum could be calculated from the fit parameters. This value of α was then used in the main VMC and DMC runs.

The value of α was independently optimized for each simulation, as the optimal value depends on L, s and r_s. The effect of squashing the correlation hole into the cell with the r_{ij}^2/L_c^2 cutoff leads to a strong $1/L^2$ dependence of the VMC energy, which will be seen in the graphs in Section 5.2.4.

5.2.2 Bulk vs Slab Comparison

Given the form of 5.4, the most obvious method to extract a surface energy from slab and bulk energies would appear to be to rearrange for σ as follows:

$$\sigma = \frac{(\epsilon_{slab} - \epsilon_{bulk})N}{2A} = (\epsilon_{slab} - \epsilon_{bulk})\frac{3s}{8\pi r_s^3} \tag{5.13}$$

When using DFT it is easy to compare the bulk energies per electron with the slab energies per electron to obtain surface energy curves as a function of density parameter r_s. These are shown in Figure 5.5 for two very commonly used density functionals, the Perdew-Wang (PW) parameterization [132] of the Local Density Approximation (LDA) and the Perdew-Burke-Ernzerhof (PBE) parameterization [129] of the Generalized Gradient Approximation (GGA). However, the major disadvantage of attempting to use the same approach with QMC calculations can be seen as soon as we consider the errors involved if the energies per electron are inaccurate in any non-systematic way.

The simplest error to deal with is the statistical error that affects every DMC estimate of an energy per electron ϵ. Even if there were negligible error in the bulk energy (which is feasible as good wavefunctions are much more easily obtained in the bulk), using Eq. 5.13 to calculate σ would give an error $\Delta\sigma$ on the surface energy of

$$\Delta\sigma = \frac{3s}{8\pi r_s^3}\Delta\epsilon_{slab} \tag{5.14}$$

where $\Delta\epsilon_{slab}$ is the error on the estimate of the energy per electron in the slab. To demonstrate the accuracy required, an estimate can be made based on the corresponding DFT values. At $r_s = 2.07$ and with $s = 18.4$, the surface energy is of order $\sigma = -690\text{erg cm}^{-2}$ and the energy per electron is $\epsilon_{slab} = -9.31\text{mHa}$. If we decide we need an accuracy of 10% on σ, then we require an error of less than $\Delta\epsilon_{slab} = 0.2\text{mHa}$,

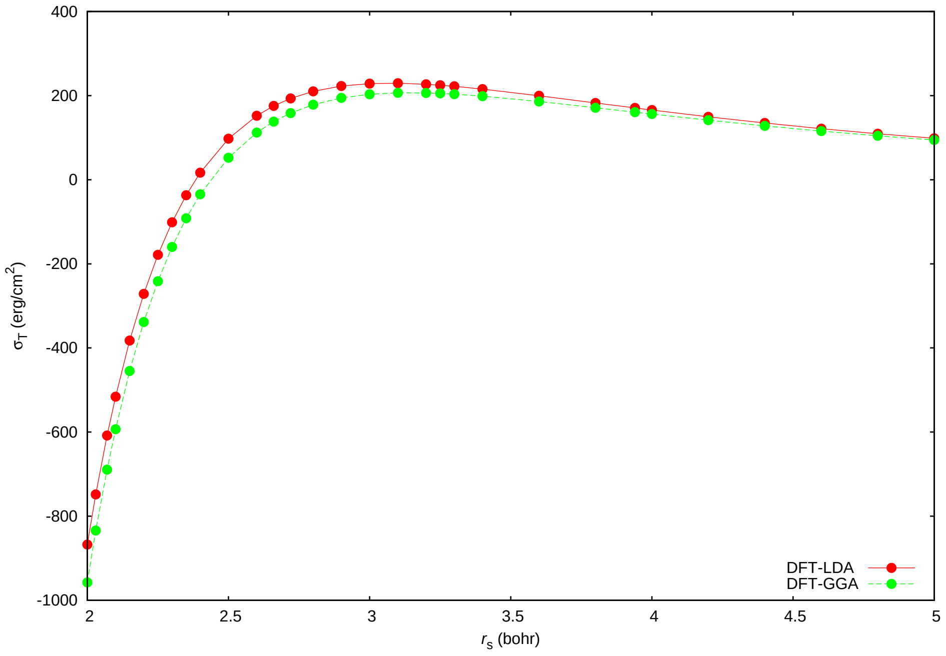

Figure 5.5: Surface energy $\sigma(r_s)$ against density parameter r_s in the range $r_s = 2$ to $r_s = 5$, calculated by comparing ϵ_{bulk} against ϵ_{slab} for slabs of width $s = 8.93r_s$ (chosen to minimize finite size fluctuations — see below). The surface energy with the LDA functional (red) is consistently marginally higher than the surface energy with the GGA functional (green) but otherwise agreement is good. Note that as $r_s \to 0$, $\sigma(r_s) \to -\infty$, and as $r_s \to \infty$, $\sigma(r_s) \to 0$.

which is already challenging. In principle, the stochastic error can be reduced to as low a level as required by increasing run times, but the slow decay of the errors ($1/\sqrt{M}$ with run length M) makes this unfeasible beyond a certain point if the trial wavefunction is poor.

Much more serious, however, are the unknown contributions of systematic errors. Pitarke and Eguiluz [136] demonstrated that the filling of subbands in the z direction causes an oscillation in σ with a wavelength of $\lambda_F/2$ where λ_F is the Fermi wavelength, related to the density parameter by

$$\frac{\lambda_F}{2} = \frac{1}{2} \times 2\pi r_s \left(\frac{4}{9\pi}\right)^{1/3} \simeq 1.63696 r_s \; . \tag{5.15}$$

Figure 5.6 shows these out-of-plane finite size effects: at $r_s = 2.07$ the wavelength of the oscillations is $\lambda = 1.62r_s$ and at $r_s = 3.94$ it is $\lambda = 1.63r_s$ so the above model is well-obeyed in this density range. Pitarke and Eguiluz's analysis was for the infinite barrier model so exact correspondence is unlikely.

The surface energy $\sigma(s)$ is thus an oscillatory function of s for finite slabs, but at large s it tends to a well-defined limit, $\sigma(s \to \infty)$. By analyzing the form of $\sigma(s)$ in DFT, it is possible to pick 'special' slab widths for which the calculated surface energy matches its infinite s limit. By examining the occupations of the subbands in each case, it can be seen that the special slab widths all have similar occupation patterns, with each successive special width adding electrons to a new, previously unoccupied subband. For example, s_1, s_2 and s_3 might correspond to the occupation of the $n = 1, 2, 3, 4$, $n = 1, 2, 3, 4, 5$ and $n = 1, 2, 3, 4, 5, 6$ subbands respectively. Figure 5.7 demonstrates the procedure of choosing these widths for $r_s = 2.07$, where this is the case. This DFT-based method for finding the special slab widths is expected to carry over accurately into QMC because it results from a subband-filling effect which is a property of the trial wavefunction, not a direct result of the way the energy calculations themselves are performed. This should eliminate any bias from the choice of values of s — a problem not addressed at all in previous QMC calculations of surface energy.

As already mentioned, in a DFT calculation we can perform the integration over the in-plane $\mathbf{k}_{||}$ vectors analytically, in effect taking the in-plane size L to infinity. In QMC, we are restricted to finite N we thus require finite L. We therefore introduce a second source of finite-size error, due to kinetic energy quantization from the finite set of allowed $\mathbf{k}_{||}$ vectors. The simplest method for ameliorating this in-plane error is to use arbitrary values of L and apply a standard correction based on the DFT finite size error, of the form:

$$\epsilon_{slab}^{QMC}(\infty) = \epsilon_{slab}^{QMC}(L) + \left(\epsilon_{slab}^{DFT}(\infty) - \epsilon_{slab}^{DFT}(L)\right) \ , \tag{5.16}$$

but it was found that using a similar procedure to the above, picking 'special' values of L where the in-plane error is near zero anyway, gives marginally better conver-

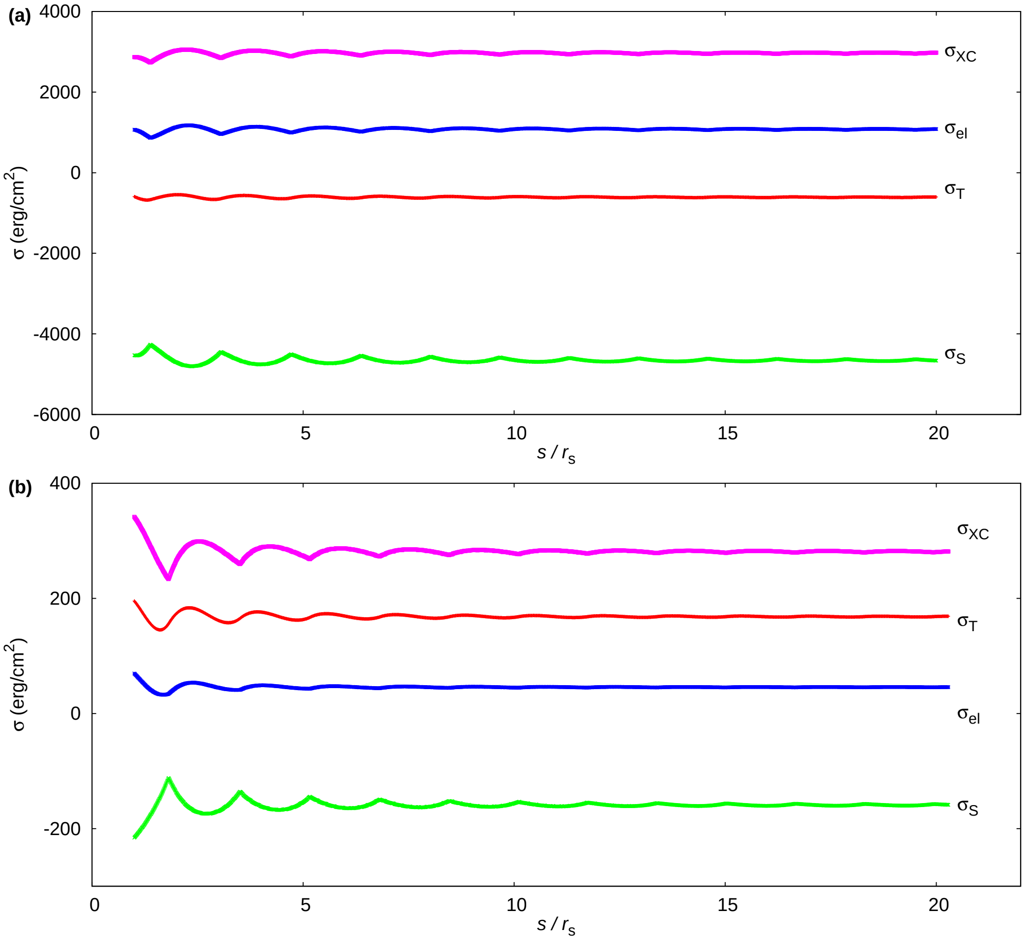

Figure 5.6: Breakdown of the DFT surface into its constituent parts for (a) $r_s = 2.07$ (b) $r_s = 3.94$. The total surface energy σ_T is the sum of the electrostatic, exchange-correlation and kinetic terms, σ_{es}, σ_{XC} and σ_s respectively. Out-of-plane finite size effects cause each term to oscillate with s as shells are filled, with cusps at points where new shells are occupied for the first time. The cusps cancel in σ_T to produce a smooth function of s. Note the differing scales — the oscillations become more severe at high density (low r_s). The value on the x-axis is s/r_s, so the fact that the wavelength looks the same for both r_s values (and all others tested) demonstrates that the period of the oscillations is in fact proportional to r_s.

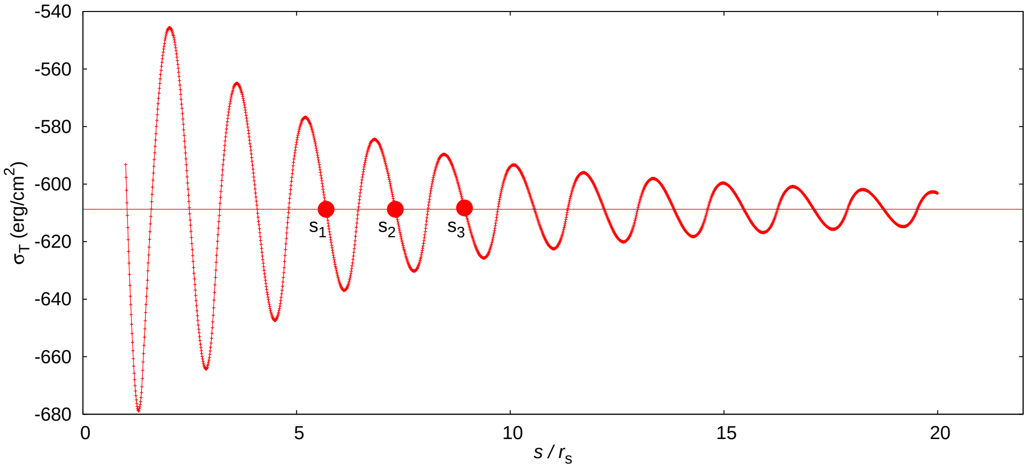

Figure 5.7: Choices of special slab width s_1, s_2, s_3 for $r_s = 2.07$. Values are picked such that the DFT surface energy at the special point matches that obtained in the infinite slab width limit.

gence. In practice we used a specific set of around 4-5 different values of N and thus L for each slab width.

The form of the in-plane finite size errors is shown in Figure 5.8. Their behaviour is unpredictable and non-trivial as they depend on the filling of stars of $\mathbf{k}_{||}$ vectors of equal magnitude. In general they are much more severe at high density, as can be seen from the magnitude of the energy fluctuations relative to the total energies in Figure 5.8(a) compared to 5.8(b).

The final major source of error in DMC calculations is the fixed node error discussed in Section 3.3.4. The trial wavefunction specifies the nodal surface, which determines the fixed node ground state energy E_0^{FN} within the fixed node approximation. How close this is to the true ground state energy E_0 depends on the accuracy of the nodal surface, which is extremely hard to quantify. Techniques such as backflow (see Section 3.2.4) are capable of improving the nodal surface but may yield very different levels of accuracy in different systems. In this case, homogeneous backflow in the bulk is much easier to describe, and thus more likely to lower the DMC energy, than the inhomogeneous backflow in the slab system. Even when backflow is not used, as in this work, the quality of the nodal surface in the much simpler bulk system is likely to be higher than that in the slab. When comparing

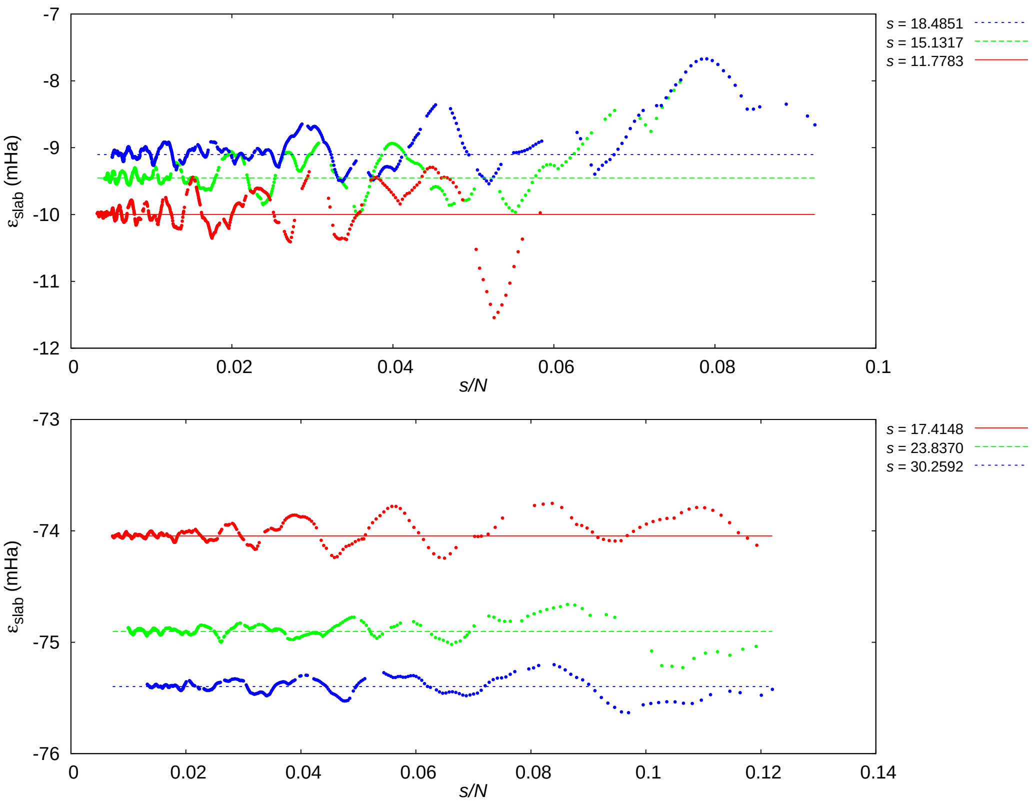

Figure 5.8: The total energy per electron in the slab ϵ_{slab} in DFT as a function of s/N for three choices of s at each density: (a) $r_s = 2.07$ (b) $r_s = 3.94$. The abscissa $s/N \propto 1/L^2$ so these oscillations demonstrate the in-plane finite size effects, which are much more severe at high density (low r_s). The slab widths s_1, s_2, s_3 (red, green and blue) are the special slab widths chosen as in Figure 5.7.

"like with like", such as in the defect formation energy calculations in the next chapter, or binding energy calculations in molecules, it is reasonable to assume, because of the similarity of the systems being compared, that the fixed-node error cancels to a high degree. Here, however, there is less reason to expect good cancellation between the slab and the bulk as they are described in very different ways. Indeed, Pitarke's improvement on Acioli's analysis, by removing the unreliable comparison of release-node bulk results with fixed-node slab results (see Table 5.1), shows how large a difference the fixed node error can make.

5.2.3 Slab-only Comparison

In light of all of the concerns presented above, we would like to formulate a method that can do away with the bulk calculation entirely and compare only like with like, using only slab results. We do this by returning to Eq. 5.5:

$$\epsilon_{slab} = \epsilon_{bulk} + \frac{8\pi r_s^3 \sigma}{3s} \tag{5.17}$$

and noting that it is of the form $y = b + cx$, where $x = 1/s$, and that we need not know the slope b to find the intercept c if we have calculations of $\epsilon_{slab}(1/s)$ at a range of values of $1/s$.

Taking into account the conclusions of the error analysis of the previous section, we adopt a scheme which minimizes all the identifiable sources of error in QMC calculations of surface energies [181]. This method can be summarized as follows:

1. Infinite-cell DFT calculations are performed, using the LDA and the PBE parameterization of the GGA, for a range of slab widths, as shown in Figure 5.6.

2. Three special slab widths are then chosen — consecutive values of s for which, within the infinite-cell LDA calculations, $\sigma(s) = \lim_{s' \to \infty} \sigma(s')$.

3. For each special slab width, DFT and QMC simulations are performed for a set of values of in-plane size L to obtain values of slab energy per electron ϵ_{slab}. Values of L are chosen such that the energy per electron for the finite cell, as calculated in the LDA, closely matches its infinite-cell value; this reduces the in-plane finite-size errors.

4. A graph of ϵ_{slab} against s/N, which is proportional to $1/L^2$, is plotted for each special slab width. If the L-dependent finite-size errors in QMC are proportional to $1/L^2$, then three straight lines will be obtained; $\epsilon_{slab}(1/L^2)$ can then be extrapolated to the infinite-L limit.

5. The extrapolated values of ϵ_{slab} are plotted against $1/s$. If Eq. 5.5 holds, this will generate a straight line, the gradient of which gives the surface energy.

The advantages of using this procedure are convincing: we sidestep the problems caused by the different fixed-node errors and wavefunction quality in bulk and slab calculations by only using the slab results; we obtain slab energies converged with slab width s and electron number N; and we can combine the results of a large number of independent simulations to produce a more accurate surface energy with a well-defined error, rather than produce a different surface energy for each simulation, with unknown error bars.

This latter point relies on an analysis of the errors accompanying the linear fits in steps (4) and (5) above. In both cases we are fitting something of the form $y = a + bx$, from quantities x_i, y_i where each y_i is accompanied by known statistical error of variance σ_i^2, and we require the statistical variance of the fitted parameters a and b. The procedure for analyzing these errors is well known (see eg [138]), and gives:

$$a = \frac{S_{xx}S_y - S_xS_{xy}}{SS_{xx} - (S_x)^2}; \qquad b = \frac{SS_{xy} - S_xS_y}{SS_{xx} - (S_x)^2} \tag{5.18}$$

$$\sigma_a^2 = \frac{S_{xx}}{SS_{xx} - (S_x)^2}; \qquad \sigma_b^2 = \frac{S}{SS_{xx} - (S_x)^2} \tag{5.19}$$

where $S = \sum_i \frac{1}{\sigma_i^2}$, $S_x = \sum_i \frac{x_i}{\sigma_i^2}$, $S_y = \sum_i \frac{y_i}{\sigma_i^2}$, $S_{xx} = \sum_i \frac{x_i^2}{\sigma_i^2}$, $S_{xy} = \sum_i \frac{x_iy_i}{\sigma_i^2}$. The assumption underlying this treatment is that all deviation from the optimal fitted line is stochastic in nature. This assumption is not strictly obeyed in this case, but as there is no predictable form to the remaining finite size effects, this should not present a particular problem.

Eqs. 5.18 and 5.19 are used to generate the values and error bar respectively of $\epsilon_{slab}(\frac{1}{L^2} \to 0)$ for each s, and thus errors on each of the points of $\Delta\epsilon_{slab}(\frac{1}{s})$ which go into calculating the gradient. The error on the fit to these points, again using these formulae, then gives the error in the total energy.

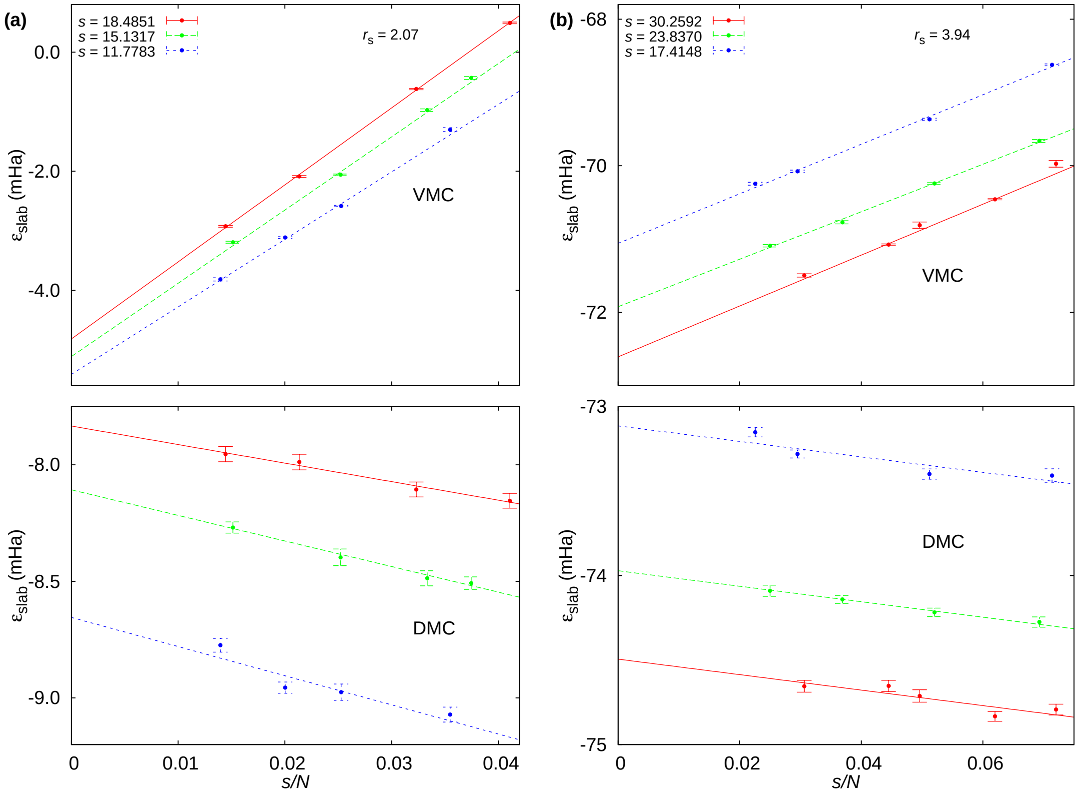

Figure 5.9: The energy per electron ϵ_{slab} calculated using VMC (top) and DMC (bottom) as a function of $s/N \propto 1/L^2$. The points have been chosen for their minimal in-plane finite size effects. Remaining finite-size effects are then corrected with DFT-derived finite size corrections. The remaining finite size effects are specific to QMC and are caused by the inadequacy of the long-range correlation included in our form of Jastrow in VMC, and residual kinetic energy finite size effects in DMC.

5.2.4 Results and Analysis

The results of the applying the method detailed above to calculate surface energies in VMC and DMC are shown in Figures 5.9, 5.10 and 5.11. Random Phase Approximation (RPA) and RPA+ results are also shown, courtesy of García-González [181]. The RPA+ combines the RPA with a correction to account for the short-range electron-electron correlation [98]; the correction used here is derived from the LDA.

To enable accurate comparisons to be made between methods, table 5.2 shows the data used to obtain Figure 5.11. Our new DMC results are consistent with results obtained using DFT in the LDA and using the RPA. They do not agree

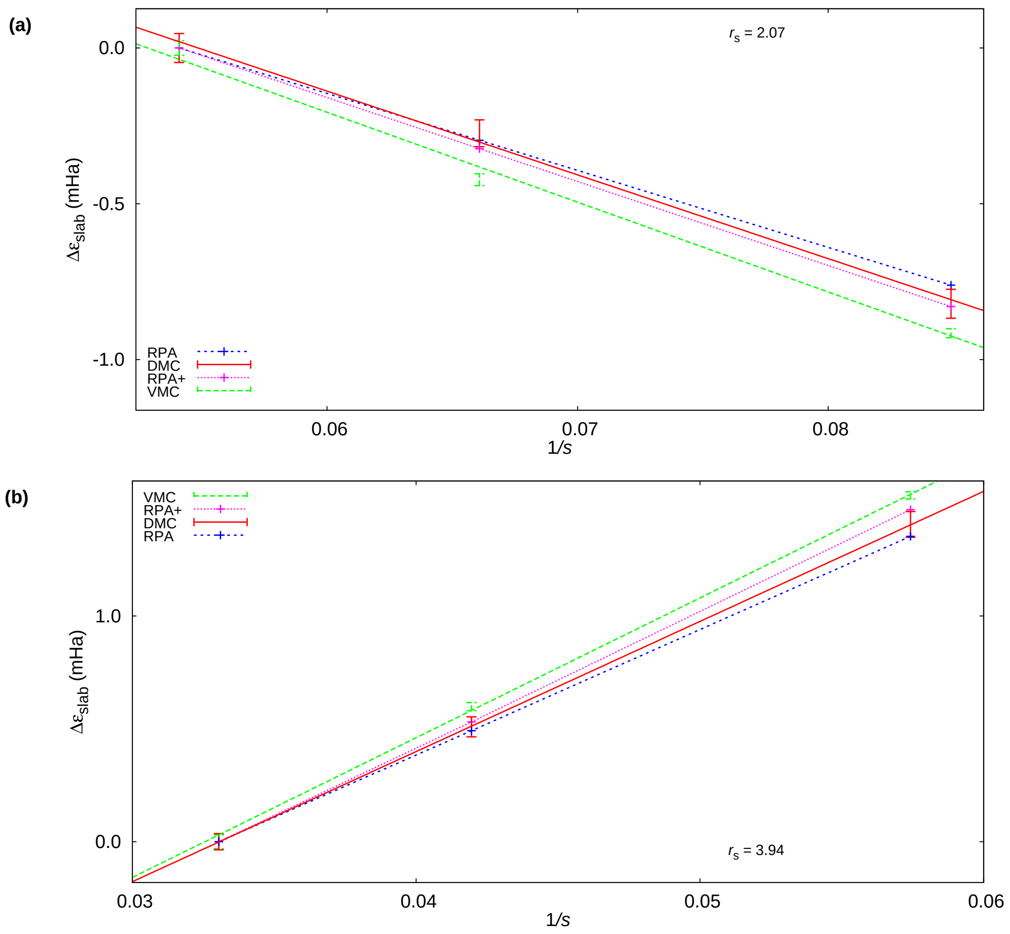

Figure 5.10: Change in the energy per electron ϵ_{slab} (extrapolated to infinite in-plane cell-size) as a function of inverse slab width $1/s$ for two densities: (a) $r_s = 2.07$; (b) $r_s = 3.94$. Because $\epsilon_{slab} = \epsilon_{bulk} + 8\pi r_s^3 \sigma/3s$, the gradient of $\epsilon_{slab}(1/s)$ gives us σ. The gradients are shown for four methods: RPA and RPA+ (blue and red, respectively), and VMC and DMC (green and purple respectively). In the latter two cases, the straight line is a linear fit to the points, which come with stochastic error bars.

with our GGA, RPA, or VMC values or previous calculations of the same. This contributes to the body of evidence suggesting that the GGA is inadequate for surfaces. The column labelled RPAc demonstrates the effectiveness of our method of selecting three slab widths for each density by applying it to RPA: we compare the RPA results obtained this way with the corresponding fully-converged RPA values including simulations of wider slabs, which are impractical in QMC. The agreement between the RPA and RPAc values indicates that the method is accurate to within

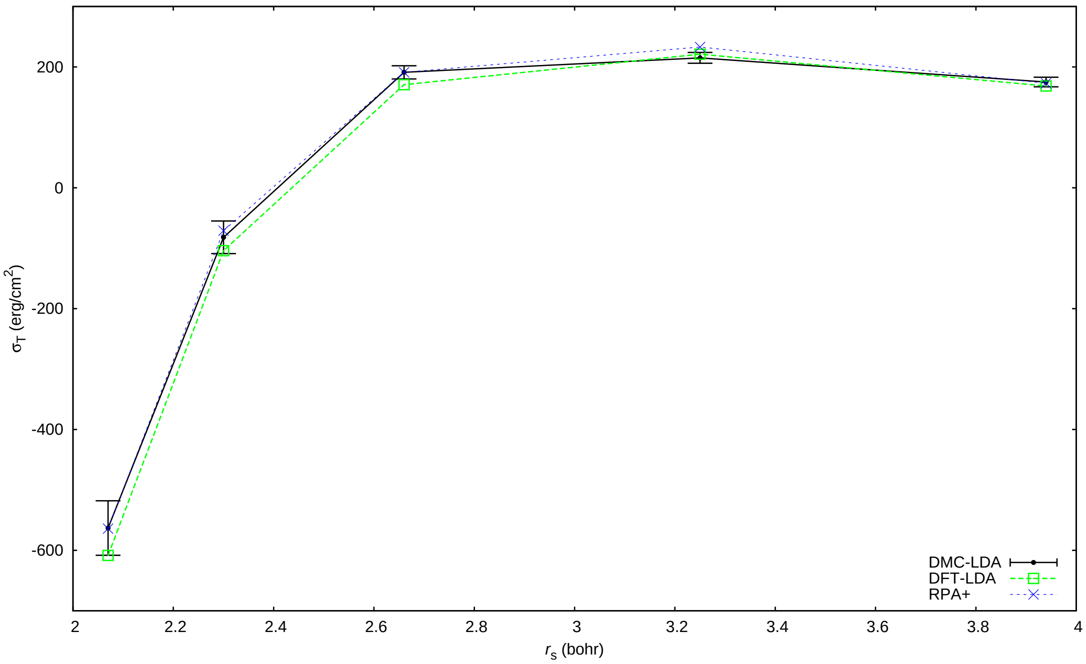

Figure 5.11: The surface energy σ of the Jellium slab as a function of the density parameter r_s for three different calculation methods. The points chosen correspond approximately to the electron densities of a range of metals: $r_s = 2.07$ to Aluminium, $r_s = 2.30$ to Zinc, $r_s = 2.66$ to Magnesium, $r_s = 3.25$ to Lithium, and $r_s = 3.94$ to Sodium. The lines are a guide to the eye.

the error bars introduced by QMC.

Also included in the table are DMC results based on GGA orbitals rather than LDA orbitals; these help to demonstrate that the result obtained was not solely a function of the nodal surface used. If DMC results based on GGA orbitals had followed the trend of the GGA, the above argument that the GGA is poor for surfaces would be rendered unconvincing. This does not appear to be the case, although the error bars make it hard to draw a firm conclusion.

Several comparisons of the jellium surface energy obtained by different methods exist in the literature [185, 157, 99], and they agree on the general trend followed by the results. This trend is that, over a wide range of densities, the following holds true:

$$\sigma_{GGA} > \sigma_{LDA} > \sigma_{mGGA} > \sigma_{RPA+} > \sigma_{RPA} \tag{5.20}$$

r_s (bohr)	LDA	GGA	VMC$_{LDA}$	DMC$_{LDA}$	DMC$_{GGA}$	RPA+	RPA	RPAc
2.07	-608.2	-690.6	-637 ± 50	-563 ± 45		-564	-517	-506
2.30	-104.0	-164.1	-131 ± 40	-82 ± 27		-71	-34	-25
2.66	170.6	133.0	200 ± 11	179 ± 13	158 ± 26	191	216	225
3.25	221.0	201.2	252 ± 14	216 ± 8		233	248	253
3.94	168.4	158.1	188 ± 3	175 ± 8	169 ± 9	173	182	187

Table 5.2: Table of surface energy σ at a range of density parameters r_s corresponding to various metals as in Figure 5.11, with a range of different methods. All surface energies in erg cm^{-2}.

where mGGA stands for meta-GGA. Our DMC results lie towards the middle of this range, consistent with the LDA, mGGA, and RPA values, closer to the RPA values for high densities and to the LDA values for low densities.

The controversy that led to the need for this research is thus resolved, and DMC just about retains its place as in principle the most accurate available method for such calculations. Although DMC calculations are plagued by systematic errors, they can all be eliminated if sufficient care is taken with the method.

5.2.5 Estimation of Fixed Node Errors

Considering the difference between $\epsilon_{slab}(r_s)$ and $\epsilon_{bulk}(r_s)$ gives us a rare opportunity to estimate the size of the fixed node error. Additionally, we can evaluate how large an error is made if the original bulk vs slab comparison used in previous calculations is employed.

To provide bulk energies for comparison, fixed-node DMC simulations were carried out on bulk jellium. Fully optimizable Jastrow factors were included to reduce the variance, but the nodal surface was simply that of a Slater determinant of plane waves. Both FCC cubic and simple cubic (SC) unit cells containing a range of numbers of electrons between $N = 120$ and $N = 650$, all corresponding to filled stars of k-points, were simulated. After the single-particle finite size corrections have been applied, the results lie close to a straight line in $1/N$. Although the fcc and sc cells followed slightly different trends, the plot of $\epsilon_{bulk}(1/N)$ could be extrapolated fairly accurately to the infinite N limit. Figure 5.12 shows $\epsilon_{bulk}(1/N)$ for $r_s = 2.07$.

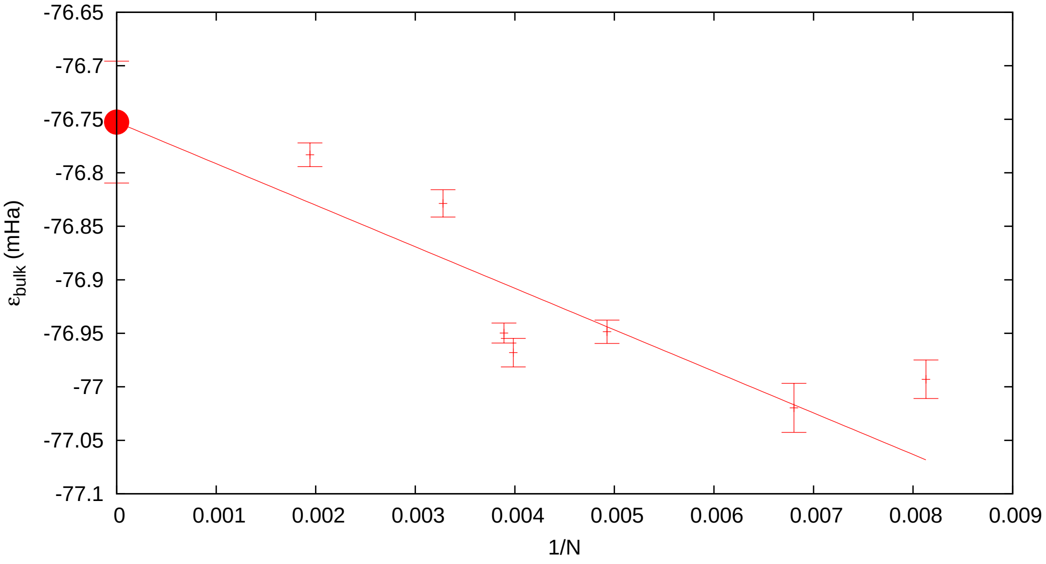

Figure 5.12: Bulk energy per electron ϵ_{bulk} as a function of inverse number of electrons $1/N$ for $r_s = 2.07$. While fcc and sc unit cells follow slightly different trends, they both nevertheless converge to a sufficiently accurate value in the limit of infinite sized cells. The large filled circle shows the extrapolated $N \to \infty$ value.

Results for all the densities are shown in Table 5.3.

Also in Table 5.3 are the surface energies by the bulk vs slab comparison method. While the errors are large, they clearly deviate both from each other (indicating a strong slab width bias) and from the slab-only, extrapolated value (indicating a strong fixed node error). Our slab-only method corrects both these errors.

These bulk energies per electron can then be compared with the values of ϵ_{bulk} obtained by extrapolating $\epsilon_{slab}(1/s)$ to infinite width. Table 5.4 shows the bulk and extrapolated slab energies, and the difference between the two methods, inferred from the difference between the two values of ϵ_{bulk}, which presumably results from the difference in absolute fixed node error. Figure 5.13 shows ΔE_{FN} as a function of r_s — there is no clear trend but the error appears to be getting larger at lower densities, although it decreases as a fraction of the total energy per electron. Nowhere in the range is it large enough to suggest that the fixed node error on the slab-only results is large enough to invalidate the surface energies calculated in this work.

r_s (bohr)	ε_{bulk} (mHa)	$\sigma(s_1)$	$\sigma(s_2)$	$\sigma(s_3)$	Extrap. σ
2.07	-6.48 ± 0.14	-536 ± 42	-516 ± 46	-524 ± 56	-563 ± 45
2.30	-31.39 ± 0.15	-52 ± 31	-50 ± 39	-33 ± 49	-82 ± 27
2.66	-54.63 ± 0.13	224 ± 15	243 ± 21	256 ± 26	179 ± 13
3.25	-71.34 ± 0.07	254 ± 6	267 ± 8	284 ± 11	216 ± 8
3.94	-76.75 ± 0.06	194 ± 5	200 ± 4	208 ± 9	175 ± 8

Table 5.3: Bulk energies per electron at a range of densities, and surface energies calculated at each of the special slab widths by comparing energies per electron in the bulk and the slab, compared to the surface energies from the extrapolation method of Section 5.2.3. The dependence on s of the former surface energies shows a finite size bias, which should be absent in our method if large enough values of s are considered. The difference between bulk vs slab and slab only results demonstrates that there must be a considerable fixed node error, which should at least be minimized with our method as the quality of the nodal surface should be roughly constant at different slab widths. All surface energies in erg cm^{-2}.

r_s (bohr)	ϵ_{bulk}(bulk calculation)	ϵ_{bulk}(slab extrapolation)	ΔE_{FN}
2.07	-6.48(14)	-6.36(16)	0.12(21)
2.30	-31.39(15)	-31.25(11)	0.14(19)
2.66	-54.63(13)	-54.23(10)	0.40(16)
3.25	-71.34(07)	-70.84(09)	0.50(11)
3.94	-76.75(06)	-76.4(14)	0.35(15)

Table 5.4: Comparison of bulk energies per electron obtained from bulk calculations, with those obtained by extrapolating $\epsilon_{slab}(1/s)$ to infinite s. The statistical errors on the bulk calculations and extrapolation errors on the slab calculations are shown in brackets. The difference gives an estimate of the difference in fixed node error between the two methods. All energies are in mHa.

5.2.6 Remaining DMC Finite Size Errors

The DMC energy per electron results in Figure 5.9, while correcting most of the strong $1/L^2$ error in the VMC results, nevertheless show a remaining finite size effect, which has not so far been explained. Examining the DMC plots in Figure 5.9, it appears that while as a function of $1/L^2$ each individual line is straight and can be well approximated by a linear fit, the gradients are not the same for each value of s as would be expected for a purely surface-based error. Since each line corresponds to a fixed s value, a change in slope of that line represents an additional error proportional to $1/L^2$, and at fixed s, $1/L^2 \propto 1/N$. It thus appears that the

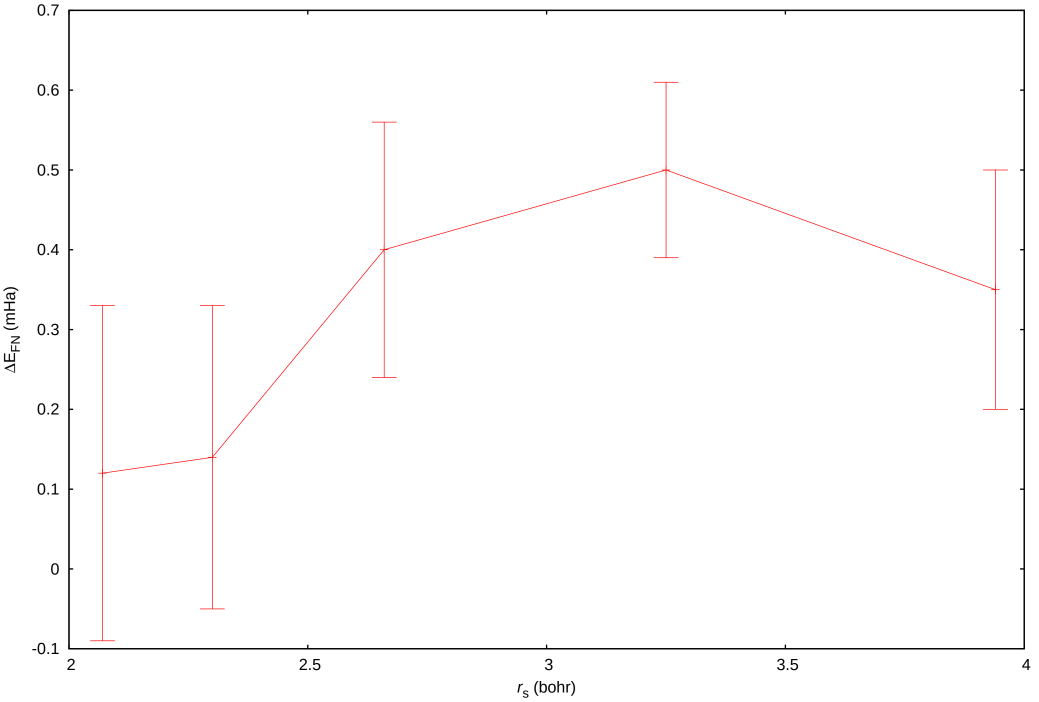

Figure 5.13: Fixed Node error ΔE_{FS} as a function of density parameter r_s. No clear trend is observed but as a fraction of the total energy the fixed node error falls rapidly with decreasing density.

remaining finite size errors in the DMC simulation are a function of $1/N$.

Recent work by Chiesa [33] has shed light on the form of the remaining finite size errors in both the kinetic energy (ΔT_N) and the potential energy (ΔV_N) in DMC calculations, once the independent particles finite size error has been accounted for by twist-averaging. The important result of this paper is that a finite sized system inevitably neglects a sometimes significant contribution to the kinetic energy from the long-ranged part of the Jastrow factor, which they derive in the homogeneous electron gas as:

$$\Delta T_N = \frac{\omega_p}{4N} \tag{5.21}$$

Appendix A gives a derivation of the Chiesa argument for the form of ΔT_N and expands upon it with a more involved corresponding treatment of inhomogeneous systems with long range Jastrows that, as in the case of the plasmon normal modes of a slab, are not simply a function of electron-electron separation. We then show how the plasmon normal mode formulae derived by Wood [180] can be used to treat the inhomogeneous jellium slab in a similar way to the HEG. The end result is a term which, while not easily directly calculable due to its dependence on terms such as $\rho(-G_z, G_z)$, has the same $1/N$ dependence as the above, and thus goes some way to explaining the remaining finite size effects. Further work could perhaps be undertaken to correct this error analytically, but given that the procedure of fitting lines to $\epsilon_{slab}(1/N)$ appears reliable and can be justified theoretically, it would be of limited practical value.

Chapter 6

Point Defects in Alumina

This chapter addresses the calculation of the formation energies of point defects in metal oxides, specifically alumina. Point defect energetics and properties are a topic with a very wide range of applications to fields as diverse as materials science, engineering, chemistry, electronics and more, and to which a great deal of effort in electronic structure theory has been applied in recent decades. There remain, however, many open questions and questionable methods are commonly used without adequate testing. As has already been discussed, quantum Monte Carlo is becoming acknowledged as one of the most accurate methods available for the determination of energies of solid systems, yet there have been few applications to date of this topic to point defects. Among the exceptions is work by Leung et al. [110], Hood et al. [76], Alfè et al. [4] and Batista et al. [15] which showed intriguing differences between defect formation energies obtained in quantum Monte Carlo and those from density functional theory.

Point defects are any irregularity in the structure of a crystalline solid that is not extended in space in any dimension. Possible point defects include include: vacancies — sites that are normally occupied but which become unoccupied; interstitials — sites that are normally vacant but which become occupied by an atom, either of one of the components of the host crystal (a self-interstitial) or of an impurity species; and substitutionals — where an impurity species replaces an atom of the host. A raft of more subtle combinations of the above are also possible.

Point defects are interesting because they are capable of very strongly influencing the material properties. Optical properties are strongly affected by the presence of free charges, for example at anion vacancies, where the charge previously on the anion often remains on the site to leave the crystal locally neutral. Catalysis often relies on surface defects as the active site where reactions occur. Electrical properties are often wholly or in part dependent on the presence of defects, which are responsible, for example, for the presence of free electrons or holes in semiconductors. Finally, mechanical properties often strongly depend on the changes in elasticity, ductility, brittleness and so on caused by the presence or absence of defects.

Alumina is an interesting material for a variety of reasons, and makes an excellent test case for a new method because it has a number of challenging features. Its bonding is part way between ionic and covalent, with the formal charge state of Al^{3+} and O^{2-} being misleading as to the level of ionicity. It has a complex, distorted structure due to its coordination, which will be described in more detail in Section 6.2.1. It has anisotropy in both its dielectric tensor and its elasticity, rendering analytic corrections to long range electrostatic and elastic forces potentially challenging. It has a very wide spectrum of possible properties, which can be strongly affected by doping and thus by the point defects it contains. Finally, while alumina has a wide range of industrial uses [38], and has been much studied both theoretically and experimentally, there remain a number of open questions about the origins of its properties, relating to point defects. For example, many aspects of the diffusion of oxygen through the lattice remain unexplained [70], even down to the identity of the dominant diffusing defect species in many situations. Defect complexes such as AlO have been proposed recently [37] but calculations have proved inconclusive.

A number of previous attempts have been made to apply electronic structure methods to defects in Al_2O_3 [119, 80, 23]. The results of empirical pair-potential models [66, 102] vary wildly according to the potential used [28], principally because of the mixed ionic-covalent bonding, the complexity and low symmetry of the crystal structure, and the major changes in nearest-neighbour coordination on creation of defects. Interestingly, ab initio methods have scarcely fared better. For example, the DFT formation energy of a neutral Oxygen vacancy has been variously quoted as 12.92 eV [80], 10.14 eV [23], 5.83 eV [184], 7.08 eV [158] and 13.3 eV [119]. Similar variation was seen in the geometries accompanying these energies, with the relaxation of the nearest neighbour Al ions away from the defect ranging from 0.4% to 12%.

Some of this uncertainty can be attributed to the different possible definitions of what is meant by a formation energy of a defect. We will address this question in Section 6.1. Even when the definition is established, however, there are a number of approximations that can be made for the chemical potentials of the species involved — we discuss our approach to this problem in Section 6.1.4. Section 6.2 describes the results of our DFT calculations of formation energies of point defects in alumina, which we believe make several substantial improvements over previous calculations. One critical uncertainty we hope to address with QMC is the issue of band gap corrections discussed in Section 6.2.4. Section 6.3 discusses DMC calculations of formation energies in charged vacancies — a first for QMC.

6.1 Thermodynamics of Defect Formation Energies

The formation energy of an isolated point defect is not a trivial quantity to define precisely, due to a number of theoretical and practical issues relating to the methods used. In recent years, one particular method, that of Zhang and Northrup [186] has become widely accepted as applicable. A summary of the use of the method, focussing on semiconductors but of wider relevance, can be found in [171]. In essence, the formalism involves treating a particular sample of the system (usually a supercell of many primitive cells, one of which contains the defect) as if it formed in equilibrium with reservoirs of each of its constituent atomic species at chemical potentials equal to those the perfect crystal, and a reservoir of electrons at the chemical potential of the top of the valence band. Defect formation energies are then the difference between the formation energy of the supercell containing the defect and the same supercell containing perfect crystal. This method has found extensive use and has generally been found to agree well with experimental defect formation energies in cases where these are available.

In this section we first derive the Law of Mass Action to demonstrate the usefulness of the formation energy, and then discuss various method-independent aspects of defect formation energy calculations.

6.1.1 The Law of Mass Action and Defect Concentrations

Equilibrium under constant temperature and pressure is defined as the state in which the system has the minimum possible value of the Gibbs Free Energy G. Because these are the conditions under which comparatively slow reactions occur, G is the most appropriate thermodynamic potential for most chemistry and crystal growth. It can be written

$$G = H - TS \ , \tag{6.1}$$

where in a crystalline solid, H is the enthalpy, primarily from chemical bonding, T is the temperature of the system and the entropy S is composed of two contributions $S = S_v + S_c$, vibrational and configurational entropy respectively. The latter term is simply $S_c = k \ln W$ where k is Boltzmann's constant and W is the number of possible distinguishable configurations of the system. For a perfect crystal $W = 1$ and so $S_c = 0$.

The number of ways of arranging n vacancies in a crystal of N_0 atoms is the number of ways of arranging n items at $N_0 + n$ sites, so

$$P = \frac{(N_0 + n)!}{N_0! n!} \ . \tag{6.2}$$

The change in G from creating n defects of formation energy h and vibrational entropy s_v per defect is therefore

$$\Delta G = nh - nTs_v - kT \ln \frac{(N_0+n)!}{N_0!n!} \ . \tag{6.3}$$

This is minimized when $\frac{\mathrm{d}\Delta G}{\mathrm{d}n} = 0$, so the equilibrium number of defects will be given by

$$h - Ts_v - kT\frac{\mathrm{d}}{\mathrm{d}n} \ln \frac{(N_0+n)!}{N_0!n!} = 0 \ . \tag{6.4}$$

Using Stirling's approximation, $\ln n! \simeq n \ln n - n$, for large n, this becomes

$$h - Ts_v - kT \ln \left(\frac{N_0+n}{n} \right) = 0 \ , \tag{6.5}$$

which gives us the familiar expression for the Law of Mass Action:

$$\frac{N_0+n}{n} \simeq \mathrm{e}^{s_v/k} \mathrm{e}^{-h/kT} \ . \tag{6.6}$$

Because the entropic term s_v is relatively unimportant and does not vary much between defect types, it is usually safely neglected — although this is not always valid.

The dominant term in the mass-action formula, $\exp(-\Delta H/kT)$ varies by many orders of magnitude between different species of defect (see eg [154]), implying that is is in general sufficient to consider only one dominant type of disorder. In highly ionic compounds such as Alumina, types of disorder which exchange anions and cations are energetically unfavourable because of the large electrostatic repulsion, so the principal forms of disorder are those which remove an ion from an occupied site and move it to an otherwise unoccupied position within the lattice, known as anion and cation Frenkel disorder, or remove charge-neutral combinations of ions leaving vacancies, known as Schottky disorder.

6.1.2 Defect Species and Notation

The standardized system of notation for denoting defect types is that of Kröger and Vink [95]. This system uses a main symbol to define the defect species itself, a subscript to define the site the defect occupies, and a superscript to define the charge state of the defect. If the defect is a vacancy, the main symbol is a V. If the defect occupies an interstitial site, the subscript is I. Otherwise, the subscript is the symbol of the element the defect replaces. In the superscript, dashes ($'$) represent negative charge relative to the same site in a perfect crystal, and dots ($^{\cdot}$) denote positive charge relative to the same site in a perfect crystal.

Some examples in the case of Alumina include: a doubly-charged oxygen vacancy, $\mathrm{V}_{\mathrm{O}}^{\cdot\cdot}$; an Al^{3+} ion interstitial at an otherwise unoccupied interstitial site, $\mathrm{Al}_{\mathrm{I}}^{\cdot\cdot\cdot}$; a missing Al^{3+} ion, $\mathrm{V}_{\mathrm{Al}}^{'''}$, and an O^{2-} ion interstitial, $O_I^{''}$. Dots and dashes are often replaced with numbers: $\mathrm{V}_{\mathrm{O}}^{\cdot\cdot}$ could equally be written $\mathrm{V}_{\mathrm{O}}^{+2}$.

In the case of divacancies or other more complex multiple defects, the charge state of a set of defects can be denoted by bracketting them and stating the charge state of the complex. The Al-O divacancy with overall charge -1 relative to the perfect crystal is then $(\mathrm{V}_{\mathrm{O}}^{\cdot\cdot}, \mathrm{V}_{\mathrm{Al}}^{'''})'$.

Finally, defect reactions can be written in terms of these defect notation symbols. For example, the formation of a cation Frenkel defect in Alumina could be written:

$$\mathrm{Al}_{\mathrm{Al}} + \mathrm{V}_{\mathrm{I}} \leftrightharpoons \mathrm{Al}_{\mathrm{I}}^{\cdot\cdot\cdot} + V_{Al}^{'''} \; , \tag{6.7}$$

or, to simplify this by omitting the normal components of the perfect crystal:

$$\mathrm{nil} \leftrightharpoons \mathrm{Al}_{\mathrm{I}}^{\cdot\cdot\cdot} + V_{Al}^{'''} \; . \tag{6.8}$$

6.1.3 Zhang-Northrup Formula for Formation Energies

There are a number of different ways of defining a formation energy: fundamentally, the formation energy is the energy difference between a volume of the crystal with the defect and without, but if one wishes to use the formation energy to predict equilibrium concentrations under particular formation conditions, a certain level of care is required to make sure the definition agrees with that required as input to the Law of Mass Action. The Zhang-Northrup method compares the free energy of a supercell containing the defect and sufficient surrounding bulk crystal to the sum of the chemical potentials (effectively the Gibbs free energies per particle) of all the component atoms that would have made up a corresponding sample of perfect crystal.

We imagine a supercell which, if it was full of perfect bulk crystal would contain n_i atoms of each species i, connected to a large reservoir of atoms at chemical potential μ_i for each species, and a reservoir of electrons at chemical potential μ_e. To form the defect, we remove or add Δn_i atoms of each species and charge q, and produce something with a total energy of $E_T^{def,q}$. The change in H_f to form this defect from its constituents can thus be written

$$\Delta H_f = E_T^{def,q} - \sum_i (n_i + \Delta n_i)\mu_i + q\mu_e \ ,$$

where we have neglected the vibrational contribution to the entropy in the solid phase. Since q is the actual charge (including the sign), its contribution is $+q\mu_e$ because the electron has charge -1.

Because the energy of the same supercell of perfect crystal would have been

$$E_T^{perf} = \sum_i n_i \mu_i \ ,$$

the formation energy can be written as

$$\Delta H_f = E_T^{def,q} - E_T^{perf} - \sum_i \Delta n_i \mu_i + q\mu_e \ .$$

For the specific case of alumina, we can write

$$\Delta H_f = E_T^{def,q} - E_T^{perf} - \Delta n_{Al}\mu_{Al} - \Delta n_O\mu_O + q\mu_e \ , \tag{6.9}$$

where μ_{Al} and μ_O are the chemical potentials of Al and O atoms in the perfect crystal, and μ_e is the chemical potential of an electron in bulk alumina. Following [103, 56], μ_e is chosen as $\mu_e = E_{VBM}^{def} + \epsilon_F$ where E_{VBM}^{def} is the position of the valence band in the defect cell and ϵ_F is the position of the Fermi level relative to this.

For formation energies of charged states, there may be a different constant background potential E_{VBM}^{def} in the defect cell from E_{VBM}^{perf} in the perfect cell, because of the influence of the defects of neighbouring cells. One way to correct for this would be to write E_{VBM}^{def} as $E_{VBM}^{def} = E_{VBM}^{perf} + \Delta V$ where, following [137, 118], ΔV can be determined by comparing the average potentials between the perfect and defect crystals in a region of the supercell far from the defect. We can thus write the formation energy as

$$\begin{aligned} \Delta H_f &= E_T^{def,q} - E_T^{perf} - \Delta n_{Al}\mu_{Al} - \Delta n_O\mu_O \\ &\quad + q(E_{VBM}^{perf} + \Delta V + \epsilon_F) \ . \end{aligned}$$

However, the the ΔV correction is not necessarily the best way to correct for defect-defect interactions. In Section 6.1.5 we will examine methods of removing the spurious contribution of defect-defect interactions from the energy $E_T^{def,q}$, and devise a method in which ΔV corrections are not required.

6.1.4 Chemical Potentials

The formula for the formation energy of a defect includes the chemical potentials of the species that have been added or removed to form the defect. These are difficult quantities to define unambiguously in a substance composed of more than one species as it is not possible to decide from the end product alone how much of the total energy should be attributed to one element and how much to the others. However, certain facts make it possible to put upper and lower bounds on the chemical potentials involved regardless of the formation conditions, and some knowledge of the conditions under which the substance was formed is often then sufficient to decide or calculate where in this range the chemical potential is likely to lie.

Taking Alumina as a specific example, we need to find ranges for the chemical potentials of the aluminium and oxygen ions, μ_{Al} and μ_O. All such thermodynamic quantities are temperature dependent, and we want to approximate the conditions under which the oxide was formed as closely as possible, so we will work at a fixed temperature T throughout this derivation, then attempt to calculate $\mu(T)$ in terms of $\mu(0)$, which we are more likely to be able to obtain with electronic structure methods.

From the fact that the solid forms from gaseous oxygen and bulk aluminium at standard pressure and temperature T, and is then stable, we can be sure that the chemical potential per atom of both species in the bulk is lower than it would be in their elemental forms. Denoting by $\mu^0_{\mathrm{O}}(T)$ the chemical potential of an oxygen atom in gaseous O_2 at standard pressure and temperature T, this gives us

$$\mu_O(T) \leq \mu^0_{\mathrm{O}}(T) \; . \tag{6.10}$$

The value of $\mu^0_{\mathrm{O}}(T)$ is a well-defined quantity at a specific temperature, the evaluation of which will be discussed below. Similarly

$$\mu_{\mathrm{Al}}(T) \leq \mu^0_{\mathrm{Al}}(T) \; , \tag{6.11}$$

where $\mu^0_{\mathrm{Al}}(T)$ is the Gibbs free energy per atom of metallic Aluminium.

The lower limits can be determined from the energy of formation of $\mathrm{Al_2O_3}$ from these same constituents. The Gibbs free energy of formation of the oxide, $\Delta G_f^{\mathrm{Al_2O_3}}(T)$, is given by the difference between the chemical potentials of the compound and the elemental forms that compose it:

$$\Delta G_f^{\mathrm{Al_2O_3}} = \mu_{\mathrm{Al_2O_3}} - 2\mu^0_{\mathrm{Al}} - 3\mu^0_{\mathrm{O}} \tag{6.12}$$

corresponding to the reaction

$$2\mathrm{Al(s)} + \frac{3}{2}\mathrm{O_2(g)} \rightleftharpoons \mathrm{Al_2O_3}(s)\ . \tag{6.13}$$

Because the chemical potential per formula unit is given by the chemical potentials of the components in the bulk compound, we know that

$$\mu_{\mathrm{Al_2O_3}} = 2\mu_{\mathrm{Al}} + 3\mu_{\mathrm{O}}\ , \tag{6.14}$$

so combining this with Eq. 6.12 gives

$$\Delta G_f^{\mathrm{Al_2O_3}} = 2\mu_{\mathrm{Al}} + 3\mu_{\mathrm{O}} - 2\mu^0_{\mathrm{Al}} - 3\mu^0_{\mathrm{O}}\ , \tag{6.15}$$

Multiplying Eq. 6.10 by three and adding the above equation to it gives

$$3\mu_{\mathrm{O}} + \Delta G_f^{\mathrm{Al_2O_3}} \leq 3\mu^0_{\mathrm{O}} + 2\mu_{\mathrm{Al}} + 3\mu_{\mathrm{O}} - 2\mu^0_{\mathrm{Al}} - 3\mu^0_{\mathrm{O}}\ . \tag{6.16}$$

Hence,

$$\frac{1}{2}\Delta G_f^{\mathrm{Al_2O_3}} + \mu^0_{\mathrm{Al}} \leq \mu_{\mathrm{Al}}\ , \tag{6.17}$$

and similarly

$$\frac{1}{3}\Delta G_f^{\mathrm{Al_2O_3}} + \mu^0_{\mathrm{O}} \leq \mu_{\mathrm{O}}\ . \tag{6.18}$$

Because it is always true that $\mu_{\mathrm{Al_2O_3}} = 2\mu_{\mathrm{Al}} + 3\mu_{\mathrm{O}}$, if μ_{O} is at its maximum value μ^0_{O}, under the same formation conditions μ_{Al} must be at its minimum, $\frac{1}{2}\Delta G_f^{\mathrm{Al_2O_3}} + \mu^0_{\mathrm{Al}}$, and similarly if μ_{Al} takes its maximum value μ^0_{Al}, we find $\mu_{\mathrm{O}} = \frac{1}{2}\Delta G_f^{\mathrm{Al_2O_3}} + \mu^0_{\mathrm{O}}$.

The chemical potential of an Aluminium atom in the metallic elemental phase is not strongly temperature or pressure dependent, and so can be well-approximated by the zero-temperature total energy per atom of metallic aluminium, as obtained using the same theoretical framework as was adopted for the alumina calculations. This requires, in a pseudopotential calculation, that we assume the same implicit reference state — that of separated electrons and pseudo-ions (with all their valance electrons) at rest.

For the oxygen, on the other hand, the elemental form of which is molecular O_2 gas, temperature and pressure will have a very strong effect on the chemical potential, so we need a model for the chemical potential $\mu_O(p_{\mathrm{O}_2}, T)$ which includes the behaviour of an ideal gas of O_2 molecules, treating translational, vibrational and rotational degrees of freedom explicitly. A second and more serious problem is that calculation of the ground state of the oxygen molecule is a very tough problem for electronic structure methods, particularly density functional theory, and the results are very much dependent on the method used. Here we modify and apply the method presented in [52] which invokes a thermodynamic cycle corresponding to the formation of an oxide, and thus avoids the ill-defined calculation of the total energy of the O_2 molecule.

The obvious thermodynamic cycle to consider is the reaction for the formation of Al_2O_3 itself, as the requisite calculations all have to be done anyway (although in principle any oxide should produce the same answers). The reaction is

$$2\mathrm{Al(s)} + \frac{3}{2}\mathrm{O_2(g)} \rightleftharpoons \mathrm{Al_2O_3(s)} \ , \tag{6.19}$$

where (s) refers to the bulk crystals of Aluminium metal and Alumina and (g) to gaseous oxygen. At constant temperature and pressure, the Gibbs free energy balance for this reaction is described by

$$g^0_{\mathrm{Al_2O_3}}(T) = 2\mu^0_{\mathrm{Al}}(T) + 3\mu^0_{\mathrm{O}}(T) + \Delta G^0_{\mathrm{Al_2O_3}}(T) \ , \tag{6.20}$$

where $\Delta G^0_{\mathrm{Al_2O_3}}(T)$ is the standard Gibbs free energy of formation of the oxide at temperature T and $g^0_{\mathrm{Al_2O_3}}(T)$ is the chemical potential per formula unit of Alumina. We now set T to the standard temperature T^0 at which we can obtain the relevant experimental data and rearrange for $\mu_{\mathrm{O}}(T^0)$

$$\mu_{\mathrm{O}}(T^0) = \frac{1}{3}\left(g^0_{\mathrm{Al_2O_3}}(T^0) - 2\mu^0_{\mathrm{Al}}(T^0) - \Delta G^0_{\mathrm{Al_2O_3}}(T^0)\right) \ . \tag{6.21}$$

The pressure dependence of μ_{O} can be modelled using the ideal gas expression:

$$\mu_{\mathrm{O}}(p_{\mathrm{O_2}}, T) = \mu^0_{\mathrm{O}}(T) + \frac{1}{2}k_B T \log(p_{\mathrm{O_2}}/p^0) \ , \tag{6.22}$$

where the superscript "0" on the symbols p and μ_{O} denotes the values of these quantities at standard pressure.

Denoting by $\Delta\mu^0_{\mathrm{O}}(T) = \mu^0_{\mathrm{O}}(T) - \mu^0_{\mathrm{O}}(T^0)$ the difference between the chemical potential at standard temperature and that at temperature T, we can therefore write:

$$\begin{aligned}\mu_{\mathrm{O}}(p_{\mathrm{O_2}}, T) \ &= \ \frac{1}{3}\left(g^0_{\mathrm{Al_2O_3}}(T^0) - 2\mu^0_{\mathrm{Al}}(T^0) - \Delta G^0_{\mathrm{Al_2O_3}}(T^0)\right) \\ &\quad + \Delta\mu^0_{\mathrm{O}}(T) + \frac{1}{2}k_B T \log(p_{\mathrm{O_2}}/p^0) \ . \end{aligned} \tag{6.23}$$

Two possible routes now exist, the most obvious of which is to look up $\Delta\mu^0_{\mathrm{O}}(T)$ in tables of thermodynamic data. As can be seen from the comparison in [52], however, it is adequate to use the formula for an ideal gas of rigid dumbells [82],

$$\Delta\mu^0_{\mathrm{O}}(T) = -\frac{1}{2}\left((S^0_{\mathrm{O_2}} - C^0_P)(T - T^0) + C^0_P T \log(T/T^0)\right) \ , \tag{6.24}$$

where $S^0_{\mathrm{O_2}}$ is the standard entropy of gaseous oxygen and C^0_P the corresponding heat capacity at constant pressure. Even when T is as high as 2000K, this approximation only incurs an error of 0.036eV.

Our scheme for a unique definition of μ_{Al} and μ_{O} is thus complete: we choose our formation temperature and pressure, we obtain $\mu_{\mathrm{O}}(p_{\mathrm{O_2}}, T)$ from Eq. 6.23 and Eq. 6.24 and then from Eq. 6.14 we know $\mu_{\mathrm{Al}}(T)$. The only quantities taken from experiment are $\Delta G^0_{\mathrm{Al_2O_3}}(T^0)$, $S^0_{\mathrm{O_2}}$ and C^0_P, whereas $\mu^0_{\mathrm{Al_2O_3}}$ and μ^0_{Al} are calculated within the theoretical framework we are using.

6.1.5 Defect-Defect Interactions

We are attempting to simulate an isolated defect surrounded by an infinite medium of bulk crystal. This can be looked at as a fictitious situation, as in any real solid many such defects will be present, and defects form clusters if it is energetically favourable for them to do so. Moreover, for reasons of stoichiometry, defects often appear as Frenkel pairs or matched sets of Schottky vacancies. However, as a theoretical construct, the isolated defect is entirely satisfactory, and we wish to be able to model isolated defects accurately.

When, in practice, we define a supercell of many copies of the primitive cell and apply periodic boundary conditions, we are imposing a repeat length on the system. If we create a defect in one of these supercells, then the potential created by this defect will be felt by all its periodic copies, and it will feel the potential of an infinite lattice of copies of itself. If the defect is charged, this presents a particular problem, as the energy of an infinite array of like-charged monopoles is of course infinite. In the real system, positively charged defects would be associated either with slight increases of electron density in the surrounding region, or with a corresponding number of nearby negatively charged defects.

To make the isolated defect calculation feasible it is therefore necessary to accompany the monopole charge on the defect with a homogeneous uniform background, of equal and opposite total charge. This makes the supercell neutral overall and allows the Ewald method to continue to be used. In DFT calculations, this charge neutralization step is in fact automatic and unheeded, because the solution of Poisson's equation for a periodic charge distribution is only defined to within an unknown constant which is always set to zero. This is equivalent to assuming that the cell is in fact neutral even if it is not, and that the $\mathbf{G} = 0$ component of the Fourier transform of the potential is zero.

For a system which is genuinely neutral, the average value of the potential, v_B, is irrelevant, as it appears as many times with a positive sign as with a negative one and so does not contribute to the total energy. This should also be the case in calculations involving charged systems, because as the background potential is genuinely undefined, it cannot appear in any quantity which corresponds to a real observable such as a defect formation energy. In practice, the cancellation follows because the electronic chemical potential μ_e is measured relative to the *same* background potential as the defect calculation: if the defect has charge q, and thus contributes qv_B to the total energy of the defect cell, the electron chemical potential term in the Zhang-Northrup formation energy formula contributes $-qv_B$ and ΔH_f is unchanged.

Two different approaches are commonly used to ensure this alignment of chemical potentials happens correctly in practice. The first is to try and retrospectively adjust the value of the total energy of the defect cell so that the spurious contribution from interacting periodic replicas of the defect is removed, whereas the second is to try to align the value of E_{VBM} we use in determining μ_e with the value it has in the defect supercell — which of course includes the contribution from the monopole charge. This latter approach constitutes the average potential alignment discussed in Section 6.1.3, and the former approach involves the application of adjustments along the lines of the Makov-Payne correction [115], the practical application of which will be discussed in Section 6.2.5. In past work, these two methods have frequently been considered separate and complementary corrections, but the above analysis makes clear that they are both attempts at the same correction and should not be applied together.

The third alternative is to try and correct the boundary conditions on the potential during the simulation — an *a priori* correction rather than an *a posteriori* post-processing step. To a limited extent, this has been tried before: the Local Moment Counter Charge (LMCC) method of Schulz [147] uses a model density constructed out of spherical Gaussians to match the badly-behaved moments of the density of the defect cell, subtracting them to leave a periodic density with no long-range multipole interactions, and calculating the interaction of the model part analytically. However, this method is highly unsatisfactory as it neglects the polarization response of the surrounding medium and assumes that the remaining periodic charge after the moments have been removed closely mirrors the true periodic charge of the crystal — which there is no guarantee will be the case. Wright and Modine [182] demonstrated that LMCC corrected energies often converge considerably more slowly to the correct answer than the uncorrected values as the supercell size increases.

A much more satisfying approach would be to apply the correct boundary conditions to each component of the system at the point of calculating the Hartree potential, during the simulation. If the charge density in the defect supercell is $n_{def}(\mathbf{r})$ whereas in the same supercell of perfect crystal it was $n_{perf}(\mathbf{r})$, then we can define the difference

$$\Delta n(\mathbf{r}) = n_{def}(\mathbf{r}) - n_{perf}(\mathbf{r}) \ , \tag{6.25}$$

which is the aperiodic part of the charge, and sums to the defect charge q. It is this part of the charge that should not be treated with periodic boundary conditions; instead we should insist that $\lim_{r\to\infty} V(\mathbf{r}) = 0$. Appendix B discusses how to apply these corrected boundary conditions to all the components of the DFT energy functional in such a way that they could be incorporated into a simulation.

In practice, calculations using this approach turned out to be less reliable than a relatively simple but novel interpolation scheme that will be described in Section 6.2.5. This interpolation scheme was thus used to produce the final results presented in this thesis.

6.2 Density Functional Theory Calculations

6.2.1 Geometry

The Corundum, or α-alumina, structure in which Al_2O_3 is found is a complex structure dictated by the demands of the coordination of the Aluminium and Oxygen ions and the preference for close packing. It is commonly described by either of two unit cells: although the primitive cell is rhombohedral, a hexagonal unit cell is frequently a more convenient choice for the purposes of calculations and graphical representations. To a first approximation, the oxygen ions occupy a hexagonal close packed structure and the aluminium ions occupy the octahedral interstitial sites, which themselves form an alternative hcp lattice. However, because of the 2:3 coordination required to maintain charge neutrality, there are only Al ions at two-thirds of these interstitial sites, which distorts the positions of the oxygen ions as they are drawn towards the sites which do contain metal ions. The close packed layers can be described by the sequence $AaBbAcBaAbBcA$, where A, B refer to oxygen layers in their two interlocking positions, and a, b, c refer to the three possible positions of the vacancies in the aluminium layers. The first and last A's are equivalent layers, so the structure repeats every six hcp layers of oxygen atoms — necessitating a large unit cell in the hexagonal representation. Figure 6.1 shows a supercell of $2 \times 2 \times 1$ copies of the hexagonal cell in side on and top down view.

The rhombohedral primitive cell contains 10 atoms (two formula units of $\mathrm{Al_2O_3}$) and can be represented with the lattice vectors $\mathbf{t}_i$:

$$\begin{pmatrix} t_{1x} & t_{1y} & t_{1z} \\ t_{2x} & t_{2y} & t_{2z} \\ t_{3x} & t_{3y} & t_{3z} \end{pmatrix} = \begin{pmatrix} s & 0 & r \\ -s/2 & \sqrt{3}s/2 & r \\ -s/2 & -\sqrt{3}s/2 & r \end{pmatrix} . \tag{6.26}$$

The lengths s and r are defined in terms of the rhombohedral lattice constant a_{rho} and the angle β between any two lattice vectors by $s = 2a_{rho}/\sqrt{3}\sin(\beta/2)$ and $r = \sqrt{a_{rho}^2 - s^2}$. Crystallographic measurements [183, 153, 6] give $a_{rho} = 9.694$ Bohr and $\beta = 55°3'$. These parameters can be optimized within DFT for consistency. Within this structure, the oxygen ions (left) and aluminium ions (right) are at positions $\mathbf{r} = u\mathbf{t}_1 + v\mathbf{t}_2 + w\mathbf{t}_3$ given in terms of fractions of the lattice vectors $\mathbf{t}_i$ by:

u	v	w
$2\delta_O$	$1/2 - 2\delta_O$	$3/4$
$1/4$	$1 - 2\delta_O$	$1/2 + 2\delta_O$
$1/2 - 2\delta_O$	$3/4$	$2\delta_O$
$1/2 + 2\delta_O$	$1/4$	$1 - 2\delta_O$
$3/4$	$2\delta_O$	$1/2 - 2\delta_O$
$1 - 2\delta_O$	$1/2 + 2\delta_O$	$1/4$

u	v	w
$1/6 - \delta_{Al}$	$1/6 - \delta_{Al}$	$1/6 - \delta_{Al}$
$2/6 + \delta_{Al}$	$2/6 + \delta_{Al}$	$2/6 + \delta_{Al}$
$4/6 - \delta_{Al}$	$4/6 - \delta_{Al}$	$4/6 - \delta_{Al}$
$5/6 + \delta_{Al}$	$5/6 + \delta_{Al}$	$5/6 + \delta_{Al}$

(6.27)

Crystallographic measurements [6] give values of δ_O and δ_{Al} of 0.027 and

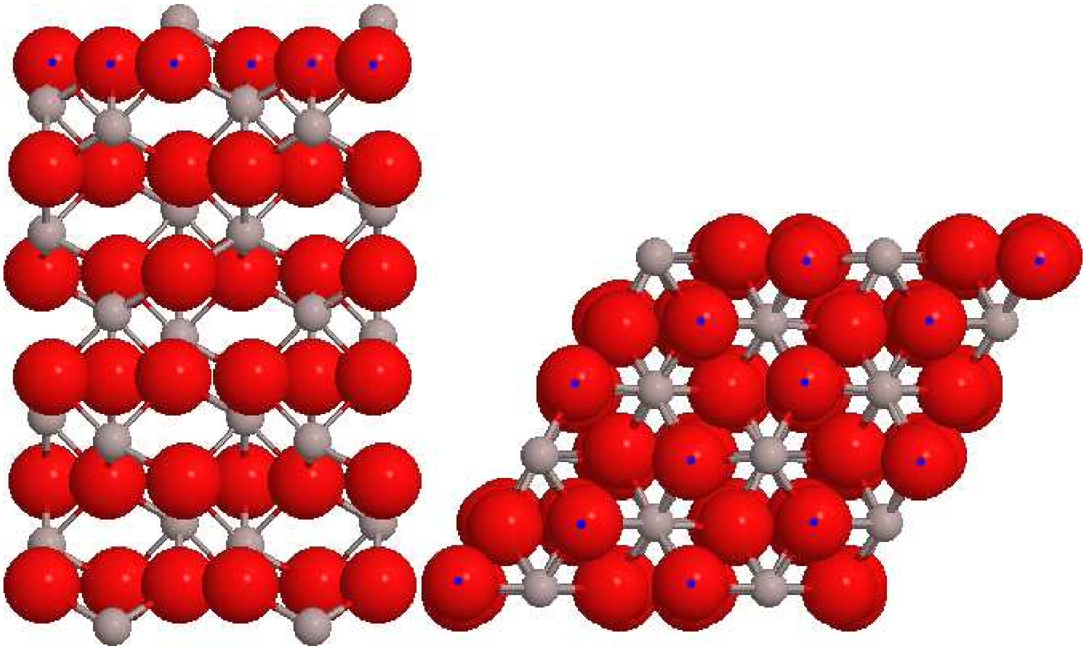

Figure 6.1: Supercell of Al_2O_3 in the hexagonal setting, with 2 × 2 × 1 copies of the hexagonal unit cell. Left: side view along the **a** axis, showing the planes of oxygen ions (red) and the aluminium ions (grey) only occupying 2/3rds of the sites. Right: Top down view along the **c** axis, showing the layers of triangles of oxygen ions slightly distorted out of their hcp positions by the absence of some of the aluminium ions. The blue dots denote those atoms forming a particular layer of O ions, to demonstrate the triangles of oxygen above and below each Al ion. This is the same supercell used for later defect formation energy calculations in QMC.

0.018 respectively. These are the fractional displacements along the axes from the hexagonal close packed structure due to the 'missing' Al ions. These would also need to be optimized within DFT.

From examining the positions above, the relation to a bipartite hcp structure is clear: if all the $\delta's$ were zero and there were extra Aluminium ions at (0,0,0) and (3/6,3/6,3/6), the bipartite hcp form would be restored.

The small angle between the lattice vectors means that the resulting rhombohedral cell is very elongated, so it is more common to work with the hexagonal cell defined by:

$$\begin{pmatrix} t_{1x} & t_{1y} & t_{1z} \\ t_{2x} & t_{2y} & t_{2z} \\ t_{3x} & t_{3y} & t_{3z} \end{pmatrix} = \begin{pmatrix} \sqrt{3}a_0/2 & a_0/2 & 0 \\ 0 & a_0 & 0 \\ 0 & 0 & c_0 \end{pmatrix}, \tag{6.28}$$

where the lengths a_0 and c_0 are related to the parameters of the rhombohedral primitive cell by $a_{rho} = \sqrt{(\frac{c_0}{3})^2 + (\frac{a_0}{\sqrt{3}})^2}$ and $\sin(\beta/2) = a_0/2a_{rho}$. Working with this cell, c_0 and a_0 are optimized within DFT and can be compared to a_{rho} and β as a consistency check.

This larger cell contains six formula units of Al_2O_3, with ions at:

u	v	w
0	$1/3-\delta_O$	$0/6$
$1/3-\delta_O$	0	$0/6$
$2/3+\delta_O$	$2/3+\delta_O$	$0/6$
δ_O	$2/3$	$1/6$
$1/3$	$1/3+\delta_O$	$1/6$
$2/3-\delta_O$	$1-\delta_O$	$1/6$
$1/3+\delta_O$	δ_O	$2/6$
$2/3$	$2/3-\delta_O$	$2/6$
$1-\delta_O$	$1/3$	$2/6$
0	$2/3+\delta_O$	$3/6$
$1/3-\delta_O$	$1/3-\delta_O$	$3/6$
$2/3+\delta_O$	0	$3/6$
δ_O	$1/3+\delta_O$	$4/6$
$1/3$	$1-\delta_O$	$4/6$
$2/3-\delta_O$	$2/3$	$4/6$
$1/3+\delta_O$	$1/3$	$5/6$
$2/3$	δ_O	$5/6$
$1-\delta_O$	$2/3-\delta_O$	$5/6$

u	v	w
0	0	$3/12-\delta_{Al}$
0	0	$5/12+\delta_{Al}$
0	0	$9/12-\delta_{Al}$
0	0	$11/12+\delta_{Al}$
$1/3$	$2/3$	$1/12+\delta_{Al}$
$1/3$	$2/3$	$5/12-\delta_{Al}$
$1/3$	$2/3$	$7/12+\delta_{Al}$
$1/3$	$2/3$	$11/12-\delta_{Al}$
$2/3$	$1/3$	$1/12-\delta_{Al}$
$2/3$	$1/3$	$3/12+\delta_{Al}$
$2/3$	$1/3$	$7/12-\delta_{Al}$
$2/3$	$1/3$	$9/12+\delta_{Al}$

(6.29)

Another common way of quoting the positions is to state the position of one atom of each species and the full set of symmetries, which can be used to generate the positions of all the rest. In this case, two parameters u_{Al} and v_O are usually specified, which are related to δ_{Al} and δ_O by $v_O = 1/3 - \delta_O$ and $u_{Al} = 1/3 + \delta_{Al}$. Then $u_{Al} \simeq 0.352$ and $v_O \simeq 0.306$, and there is an Al ion at $(0, 0, u_{Al})$ and an O ion at $(v_O, 0, 1/4)$, and all other equivalent positions.

6.2.2 Convergence Tests

In order to have confidence in quantities provided by DFT, the energies it provides must be converged with respect to a variety of criteria, and ideally should be independent of the various choices of approximation for exchange correlation (XC) functional and pseudopotential (PSP). In this section, we will summarize the results of the necessary tests and comparisons. The DFT calculations described here were performed with CASTEP [34], one of the longer-established plane-wave DFT codes.

In CASTEP, in common with all plane-wave DFT codes, the single particle orbitals are represented in a plane wave basis with coefficients $c_{i,\mathbf{k}+\mathbf{G}}$, such that

$$\psi_{i\mathbf{k}} = \sum_{|\mathbf{G}|<G_{cut}} c_{i,\mathbf{k}+\mathbf{G}} \mathrm{e}^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \ . \tag{6.30}$$

As before, it is important to ensure that the maximum plane-wave energy $E_{cut} = \frac{1}{2}G_{cut}^2$ is large enough that the energy is converged with respect to the size of the basis.

An additional constraint for orbitals to be used in quantum Monte Carlo calculations is that the variance of the local energy is more strongly dependent on the quality of the wavefunction, than the total energy. Because the variance is the primary factor in determining the number of steps for which the calculation must be run before an accurate energy and corresponding error bar is obtained, it is important to lower it as much as possible by improving the trial wavefunction. In the case of the orbitals, increasing the plane wave cut off continues to lower the variance considerably beyond the point at which total energies within DFT are fully converged. Although using a larger basis set than is necessary for total energy convergence within DFT slows down the DFT calculation, the CPU time required for the DFT calculations is negligible compared to that for the subsequent DMC calculations, so the sacrifice is easily justified.

Figure 6.2 shows the behaviour of the total energy within DFT of a rhombohedral cell of Al_2O_3 as the plane wave cutoff is varied, as calculated using the LDA and with Vanderbilt [169] Ultrasoft Pseudopotentials (USPs). The total energy was converged to 0.015 eV/atom at 550 eV, so this cutoff was used throughout the USP calculations. USPs cannot be used in QMC simulations, so we also need to test the convergence of the Trail-Needs Dirac-Fock (TN-DF) norm-conserving PSPs we will be using there [162, 163]. There are two available versions of the TN-DF

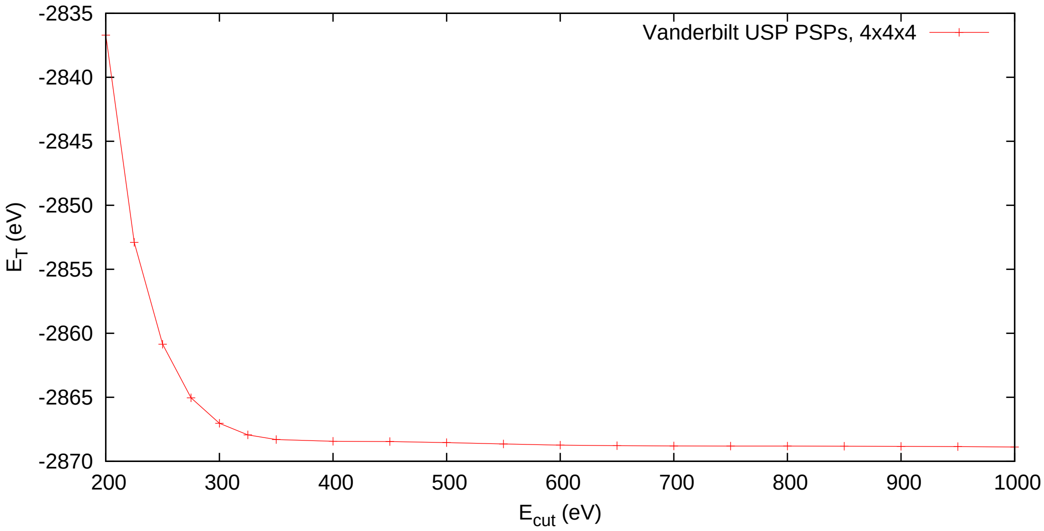

Figure 6.2: Convergence of total energy of a $4 \times 4 \times 4$ rhombohedral cell of Al_2O_3 with plane wave cut off E_{cut} with Vanderbilt ultrasoft pseudopotentials. The total energy is converged to 0.015 eV/atom by $E_{cut} = 550\,\text{eV}$. This type of pseudopotential requires considerably lower cutoffs than norm-conserving psps, so will be used throughout the geometry optimization. However, they cannot be used in QMC calculations.

pseudopotentials, which differ in their choices of core radii. The harder (shorter core radius) versions are intended for high-accuracy atomic calculations, whereas the softer (larger core radius) versions are adequate for all other purposes. Tests showed that using the harder pseudopotentials made no appreciable difference to geometry parameters or binding energies, although they required a notably higher cutoff energy. Figure 6.3 shows that the softer TN-DF PSPs require a cutoff of around 1700eV, whereas the harder versions require over 2500 eV for the same level of convergence.

The next important optimization is of the bulk cell geometry. This optimization was performed in the aforementioned rhombohedral unit cell with four parameters: a_{rho}, the rhombohedral lattice constant; α_{rho}, the angle between the lattice vectors; and u_{Al} and v_{O}, where $u_{Al} - \frac{1}{3}$ and $\frac{1}{3} - v_O$ are the fractional shifts along the lattice vectors of the Aluminium and Oxygen ions, respectively, due to the vacant sites. Starting from the experimental values for these parameters from

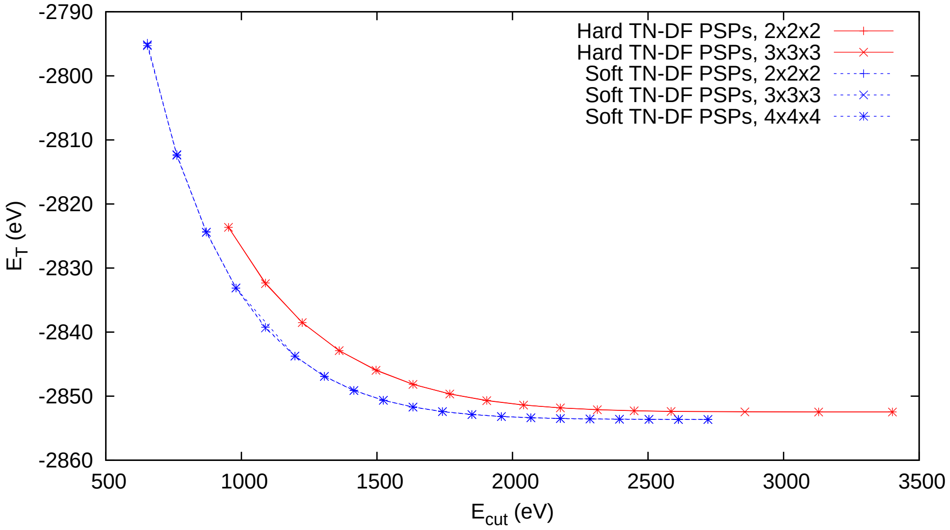

Figure 6.3: Convergence of total energy of 2×2×2 and 3×3×3 rhombohedral cells of Al_2O_3 with plane wave cut off E_{cut} with hard and soft Trail-Needs Dirac-Fock pseudopotentials (see text). Although the total energy is converged to 0.015 eV/atom by $E_{cut} = 1700$ eV, we generate the orbitals for QMC at $E_{cut} = 3000$ eV as the variance continues to reduce long after the energy is converged. 3000 eV happened to be the point at which the memory to represent the required 288 orbitals of the defect cell exceeded the available 2GB limit.

Ref. [6], the structure was optimized in DFT with a range of functionals and both types of pseudopotential described above, in order to estimate the effect of varying either choice. The results are summarized in Table 6.1, along with the mean lengths of the nearest neighbour (NN) Al-O bonds.

Because alumina is such a strong insulator, convergence of the total energy for fixed geometry with increasing k-point sampling is rapid, the difference between Monkhorst-Pack [122] grids of $2 \times 2 \times 2$ and $3 \times 3 \times 3$ being less than 0.02 eV/atom (see Figure 6.3) and between $3 \times 3 \times 3$ and $4 \times 4 \times 4$ less than 0.002 eV/atom. The optimized geometry parameters turn out to have some minor sensitivity to k-point sampling, so $4 \times 4 \times 4$ grids were used throughout the bulk investigations with DFT. This is sufficient to converge the parameters to the accuracy shown in Table 6.1.

It is important, when performing calculations within the framework of DFT, to establish that the answers are not strongly dependent the specific choices of approximation used. If they are, it sometimes suggests that the properties being mod-

Method	a_{rho} (A)	α_{rho}(°)	u_{Al}	v_O	a_{Al-O}(A)
Experiment [6]	5.136	55.28	0.352	0.306	1.915
LDA-USP	5.052	55.30	0.352	0.307	1.885
GGA-USP	5.150	55.29	0.352	0.306	1.921
LDA-TNDF	5.147	55.39	0.352	0.305	1.921
GGA-TNDF	5.169	55.29	0.352	0.305	1.928

Table 6.1: Optimized geometry parameters for bulk Alumina compared to experiment. The results for different choices of exchange-correlation functional (LDA and GGA) and different types of pseudopotential (Ultrasoft and Trail-Needs Dirac-Fock) are all in reasonably good agreement with the experimental parameters. As usual, the LDA overbinds the solid, while the GGA mildly underbinds it. Conveniently, since this is the combination used in the QMC, (but probably fortuitously) the TN-DF pseudopotentials in the LDA give the closest match to the experimental parameters.

elled are not accurately described within density functional theory. A good example is the bandgap, which comes out strongly functional-dependent but is always underestimated. Fortunately, the above geometry tests and the comparison of cohesive energies in Table 6.3 in Section 6.3.5, which also includes QMC results, demonstrate that while the formation and atomization energies do depend, as expected, on the choice of exchange-correlation functional used and mildly on the pseudopotential, the differences are not large. As is commonly observed in metal oxides [62, 83], LDA overbinds the dimers and the solid, whereas GGA underbinds them (mildly).

6.2.3 Improved Chemical Potentials

We now begin to examine the implementation in DFT of the various schemes presented in Section 6.1 for improving the accuracy of formation energy calculations. The results presented in these sections use all of the techniques (described in Sections 6.2.3, 6.2.4 and 6.2.5) even though by necessity they were introduced one by one. As previously mentioned, defect formation energies are functions of the chemical potentials of the atomic species involved and of the electrons. As alumina is an insulator, the Fermi energy must lie somewhere within the energy gap, but precisely where will depend on the level of doping with impurities of differing valence. If Al^{3+}

ions are replaced with, for example, Mg^{2+} ions, then the number of free electrons is reduced and the chemical potential falls. If Al^{3+} ions are replaced with Ti^{4+}ions, the chemical potential will rise. In both cases, some sort of charged compensating defect will also form, the energy of which will depend on the chemical potential, until equilibrium is obtained.

To illustrate the effect of each improvement, results for the formation energy of the oxygen vacancy V_O^q will be shown as a function of the Fermi energy ϵ_F, as they illustrate well many of the subtleties of the calculations. The Fermi energy is specified relative to the valence band maximum, and is defined by the equation $\mu_e = E_{VBM} + \epsilon_F$. ϵ_F lies between 0 (corresponding to the valence band maximum), and the bandgap E_g (corresponding to the conduction band minimum). We set the maximum allowed value of ϵ_F to the DFT bandgap rather than the real gap to prevent spurious high-charge states appearing stable within DFT.

The formation energy of an oxygen vacancy of charge q in the Zhang-Northrup formalism is

$$\Delta H_f = E^q_{def} - E_{perf} + \mu_O + q\mu_e \ . \tag{6.31}$$

As the formation energy contains a $q\mu_e$ term, it varies in energy with ϵ_F unless $q = 0$. Therefore, at any particular value of ϵ_F, one of the charge states $q = 0, 1, 2$ will be the most stable (lowest ΔH_f). In the following graphs we plot only the most stable states for each type of calculation, with crosses marking the changeovers.

Figure 6.4 shows a comparison of the two schemes for calculating the chemical potential. In the first, we rely, for the calculation of μ_O, on calculation of the total energy of the oxygen molecule using the same method as for the bulk. We then assume the formation conditions represented the oxidation limit (the lower bound of the chemical potential in Eq. 6.18), and simply take $\mu_O = \frac{1}{2}E_T(O_2)$. This gives a highly method-dependent result, as the total energy of O_2 represents a challenging problem for DFT for well-documented reasons. Therefore, when we compare formation energies, we find they vary by as much as 2 eV between different combinations of pseudopotential and exchange-correlation energy, as seen in Figure

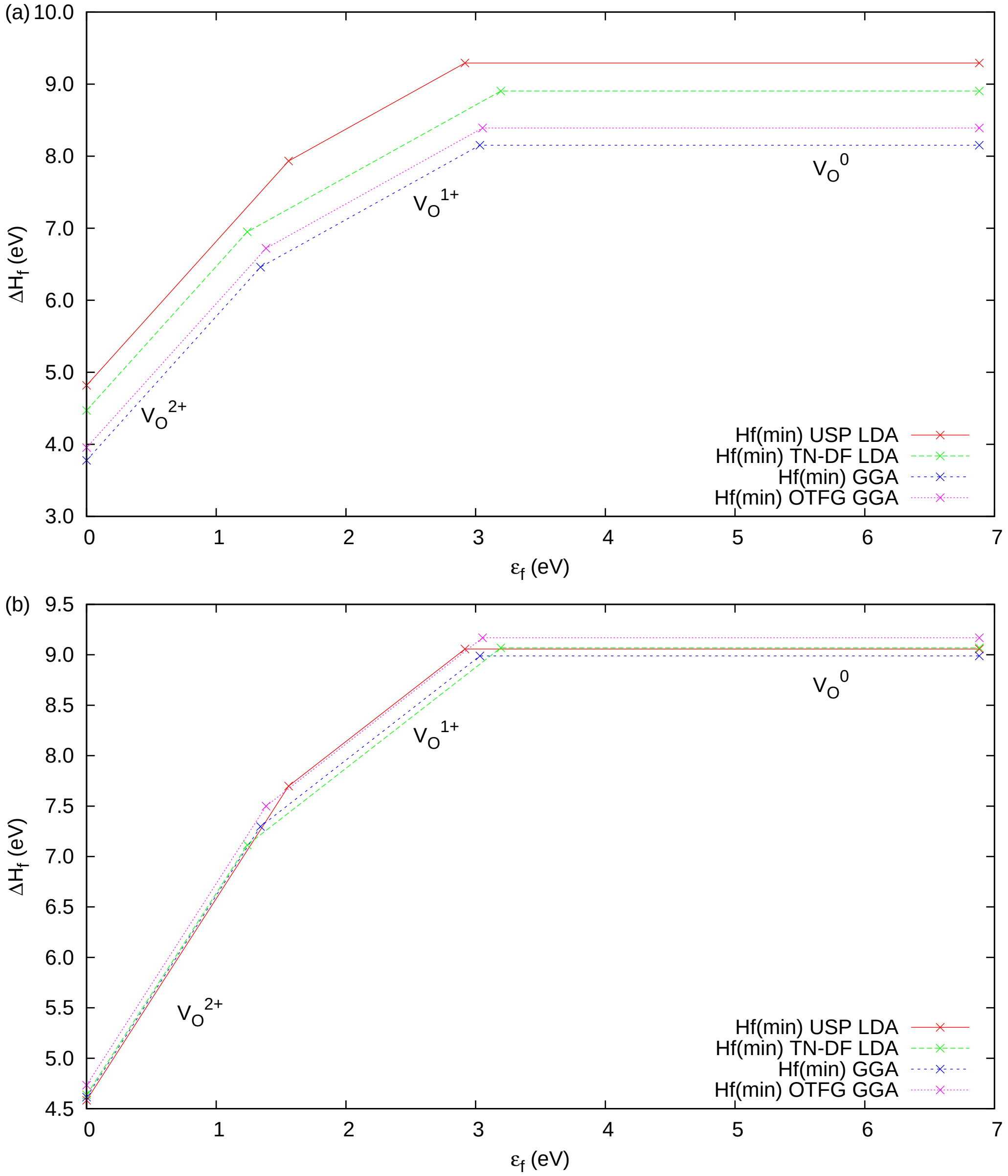

Figure 6.4: Comparison of formation energies of the Oxygen vacancy (a) without and (b) with the improved chemical potential scheme described in Section 6.1.4. Without it, there is a spread of values of H_f of nearly 2eV between different XC functionals and PSPs. With it, they all fall in a region of around 0.2eV.

6.4(a). This is clearly unsatisfactory. Figure 6.4(b) shows the same plots with the method of Finnis, Lozovoi and Alavi (FLA) used for the calculation of μ_O, bypassing the calculation of $E_T(O_2)$ by using the experimental oxide formation energy. Using this method, the lines fall almost on top of each other and vary by no more than around 0.2 eV. This suggests that the majority of the disparity in formation energies between different functionals and pseudopotentials originates from the calculations of the atomic or molecular systems required for the μ_i calculations rather than those for the solid systems. This is in line with our expectations of where DFT will produce accurate answers and where it will fail. For the rest of the work in this thesis, including the DMC results, we thus use the FLA method.

6.2.4 Band Gap Corrections

In the region surrounding a defect, the local geometry and electronic structure are significantly changed from the bulk and there will be electrons occupying levels which can be meaningfully described as 'defect levels', in that they are localized around the defect site. In these cases, the highest occupied states, which are in the band gap, can either be derived from conduction band states which have been lowered in energy by the presence of the defect, or from valence band states which are raised in energy by the presence of the defect. In the latter case, if the defect-induced states are not very strongly localized they are similar in character to the valence band states, which are occupied, and the resulting energy eigenvalue will be as correct as can be expected in DFT. In the former case, however, the defect states are similar to conduction band states, which are not occupied, and thus suffer from the well known DFT band-gap problem, whereby energy gaps to unoccupied levels are underestimated.

The resulting underestimation of conduction band-derived states is a common problem in DFT calculations on defects, and is intrinsic to the method so could be considered to be 'correct' at least within DFT. However, it is common to try and make the DFT values of ΔH_f match experimental results by applying an empirically

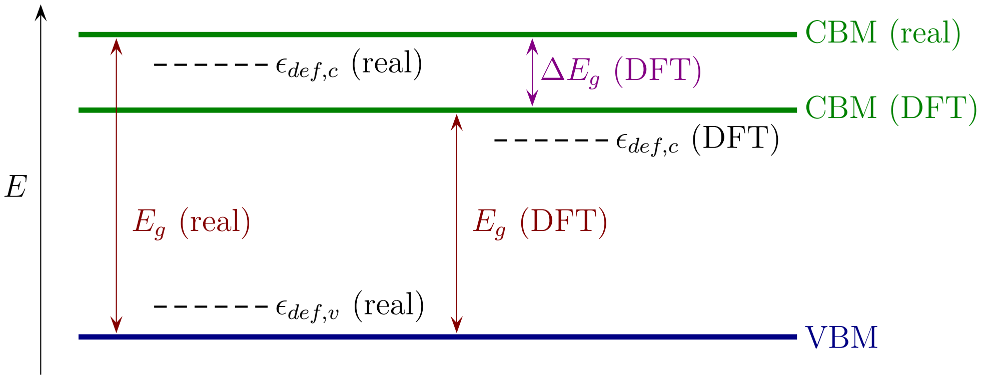

Figure 6.5: Band gap correction scheme. The Valence Band Maximum (VBM) and Conduction Band Minimum (CBM) are shown in blue and green respectively. The fact that the DFT energy gap differs from the real one (shown in red) by an amount ΔE_g means that while the eigenvalues of states tied to the valence band, such as that labelled $\epsilon_{def,v}$ can be believed, those tied to the conduction band, such as $\epsilon_{def,c}$, are assumed to be wrong by approximately ΔE_g.

justified correction based on a rigid shift the energy eigenvalues of the occupied defect levels. Figure 6.5 demonstrates how this works schematically. If there are m electrons occupying defect states derived from the conduction band and the energy gap is underestimated by ΔE_g, the usual remedy [187] is to add $m \times \Delta E_g$ to the energy of the defect cell. In semiconductors this adjustment appears to be a vital remedy to make the calculated defect formation energies agree with experiment. However, it represents a severely uncontrolled approximation, especially in the case of strongly ionic insulators where, rather than being delocalized like conduction-band-derived states in a semiconductor, the defect states are often strongly localized. Furthermore, it is often unclear whether a particular state is more conduction-band-like or valence-band-like as these are somewhat ill-defined terms. In these cases, it is not obvious that the band gap correction should be applied at all. Even worse, because such solids generally have large band gaps, ΔE_g is large if it is applied.

Figure 6.6 illustrates the effect of the band gap correction in calculations of the Oxygen vacancy. If the defect is neutral, there are two localized electrons left on the vacancy site, which was previously an anion site containing an O^{2-} ion with a filled $2p$ shell of 6 electrons. If the defect has +1 charge, there is one localized

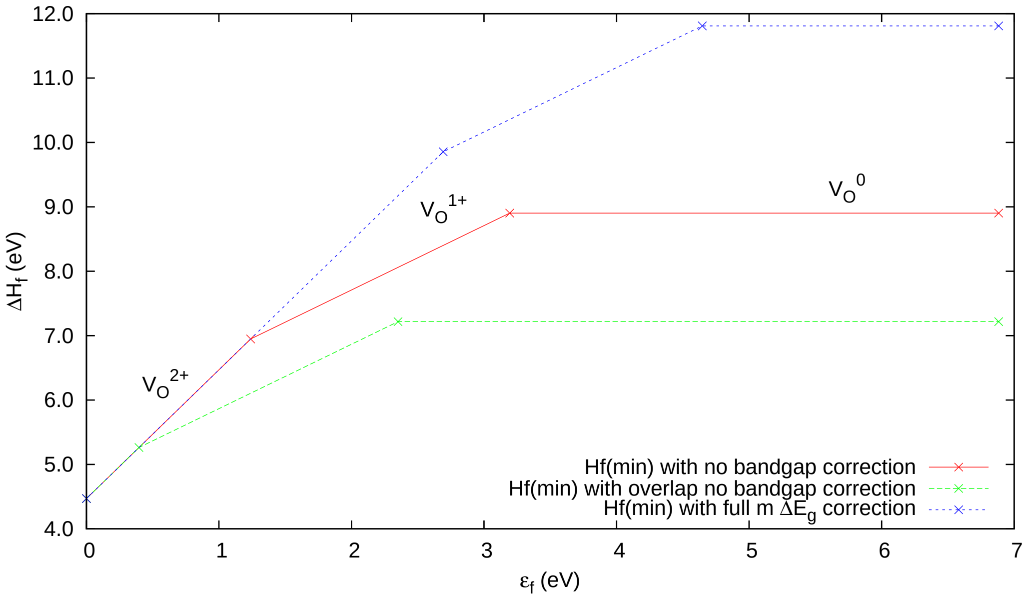

Figure 6.6: The effect of the band gap correction on the formation energies of an Oxygen vacancy. Shown in green are the uncorrected formation energies. The blue line applies the full $m \times \Delta E_g$, and the red line applies the correction of Eq. 6.32. Evidently, the variation depending on which method is chosen is huge. The latter one, however, should give the most reliable results.

electron. If it has +2 charge, there are none. In this case, therefore, $m = 2$ for V_O^0, $m = 1$ for V_O^{1+}, and $m = 0$ for V_O^{2+}. Note, however, that even though there is no longer an Oxygen nucleus at the defect site, it is not necessarily the case that the states in the region must derive from the unoccupied aluminium $3s$ orbitals at the bottom of the conduction band — they may still have the symmetry of the now-missing oxygen s-orbitals. The green curve in the figure shows the formation energy of the defect as a function of ϵ_F with no bandgap corrections applied. The blue curve shows it with a full $m \times \Delta E_g$ correction applied (E_g in LDA is 6.88 eV, compared to the real gap of 9.12 eV, so ΔE_g is 2.24 eV). The difference between the green and blue lines for the neutral vacancy is thus a full 4.48 eV, and our uncertainty about the true value of this correction would make a mockery of the high accuracy obtainable elsewhere in the calculation.

A significant improvement in the treatment of this problem may be obtained by attempting to estimate what a sensible value for m would be, rather than use the full number of occupied states [11]. In effect, one tries to calculate the percentage of conduction band nature in the defect states by calculating their overlap with all the conduction band states of the perfect crystal. The correction then becomes

$$\sum_{i\,def}\sum_{j\,cond}|\langle\psi_i^{def}|\psi_j^{perf}\rangle|^2\times\Delta E_g=\sum_{i\,def}(1-\sum_{j\,valence}|\langle\psi_i^{def}|\psi_j^{perf}\rangle|^2)\times\Delta E_g\ . \quad (6.32)$$

where the superscript labels $perf$ and def refer to the wavefunctions of the perfect crystal supercell and the defect-containing supercell respectively. This may be implemented as a post-processing step, loading in the self-consistent wavefunctions of both defect and perfect systems and calculating all the required overlaps. In the case of the oxygen vacancy, one obtains $m(V_O^0) = 2 \times 0.37$ and $m(V_O^{1+}) = 1 \times 0.37$. The formation energy with this band gap correction plotted is in red in Figure 6.6 and appears more believable than either extreme, and will be used where required in the rest of the DFT results in this work.

Despite this improvement, however, the band gap correction is still an uncontrolled approximation, and it is desirable to test its accuracy. Since quantum Monte Carlo does not suffer from the problem that makes the band gap correction necessary in DFT, comparison with QMC results may yield some information about how much of the band gap correction to apply to the DFT results.

6.2.5 Finite Size Scaling

The importance of finite size errors in periodic supercell calculations of charged defects has gained greater and greater recognition in recent years, and a very large number of attempts have been made to correct or negate them. Makov and Payne [115] proposed some of the most commonly used correction methods, expanding upon work by Leslie and Gillan [107]. In their approach, they calculate analytically the expansion of the energy of a periodic array of interacting multipoles. Their expression for the energy of a supercell of size L, relative to the infinite limit as

$L \to \infty$, E_0, is

$$E(L) = E_0 - \frac{q^2\alpha}{2\epsilon L} - \frac{2\pi q Q}{3\epsilon L^3} + \mathcal{O}[L^{-5}] \ , \tag{6.33}$$

where q is the monopole aperiodic charge, Q is the quadrupole moment of the aperiodic charge, ϵ is the relative dielectric constant of the medium in which the charges are embedded and α is the Madelung constant, which is a property of the shape of the primitive cell. For a simple cubic system, $\alpha = 2.8373$; for this hexagonal system, with different lengths in the **a** and **c** directions, α is a function of both lengths.

The Makov-Payne corrections are fairly widely used (see eg [45, 64] and many more), but have been shown to be unreliable in a variety of circumstances. This unreliability is relatively easy to verify, by studying the same defect in larger and larger cells and extrapolating the trend with L to $L \to \infty$ to find a converged value. This infinite-L result can then be compared with the corrected value obtained at fixed L. Castleton et al. [27] gave an exhaustive comparison of use of the first, and first and second terms of the Makov-Payne correction, and correction of the average potential. None of these methods were very reliable but the indications were that the form, though not the magnitude, of the Makov Payne correction, with a term $\propto 1/L$ and a term $\propto 1/L^3$, appeared accurate. Castleton et al. also found that if they used fitted parameters rather than the calculated or experimental values of α, ϵ and Q, they obtained much better results. This agrees with the conclusions of Shim, Lee and Nieminen [150] and Lento, Mozos and Nieminen [106], who found that Makov-Payne corrections had much greater reliability in situations where, as the charge q is increased, the charge density was becoming closer to the bulk distribution rather than further away. For example, if an ion that is normally +3 charged is removed from a site, the neutral vacancy corresponds to taking away the 3 electrons it previously donated to nearby anions, changing the nearby electron density significantly. As the charge on the defect is increased, electrons are returned, until at $q = -3$ the density around the neighbouring anions is back to where it was before. In this case the Makov-Payne corrections are reported to be accurate, as the charge could be regarded as pointlike and centered on the defect (in this case, it represents the missing cation). In other situations, where the nearby charge density is changed very significantly from its perfect crystal form and polarization is important, convergence of the Makov-Payne corrected results may be slower than using the convergence of the uncorrected results.

In this work, we introduce what at first appears a minor variant on the above fitting procedure but one that turns out to have a number of useful features. Examining Eq. 6.33 we see that the first order correction term, which dominates in single defects where the monopole-quadrupole and dipole-dipole terms are small, contains the term α/L, which is often referred to as the Madelung potential v_M. Previous methods have taken α as fixed for a particular lattice and used the behaviour of the total energy as a function of L for fixed lattice geometry to estimate ϵ. However, this is unnecessarily limiting: if we change the shape of our supercell by varying the repeat length along one axis at a time, we change α (and also make L rather ill-defined although v_M is still well defined). If we then plot the total energy against v_M for a wide range of cell shapes and sizes, we can still extrapolate to an infinite cell by taking the limit as $v_M \rightarrow 0$. In practice this is accomplished by fitting a straight line $E(v_M) = E_0 + bv_M$ to the calculated energies. Comparing with Eq. 6.33 gives $b = q^2/2\epsilon$, so we can also find an effective dielectric constant from this method. What makes this method especially useful is that it is possible to design unit cells where the Madelung constant becomes negative, so that we have points on either side of $v_M = 0$, and rather than an unreliable extrapolation to a point outside the range we are merely interpolating, which is much more reliable.

Figure 6.7 shows the results of this method applied to the four principal intrinsic point defects of alumina in their highest charge states. As expected, it works most reliably for V_{Al}^{3+} for the reasons given above, but remains perfectly adequate for V_0^{2+}, with uncertainty in the interpolation not exceeding around 0.2 eV. In every case, a reasonably accurate value can be obtained in the $v_M \rightarrow 0$ limit. The value of ϵ varies somewhat according to which defect is studied but it is always in the range $\epsilon = 3.4$ to 3.9. This variation may be because in reality the dielectric tensor

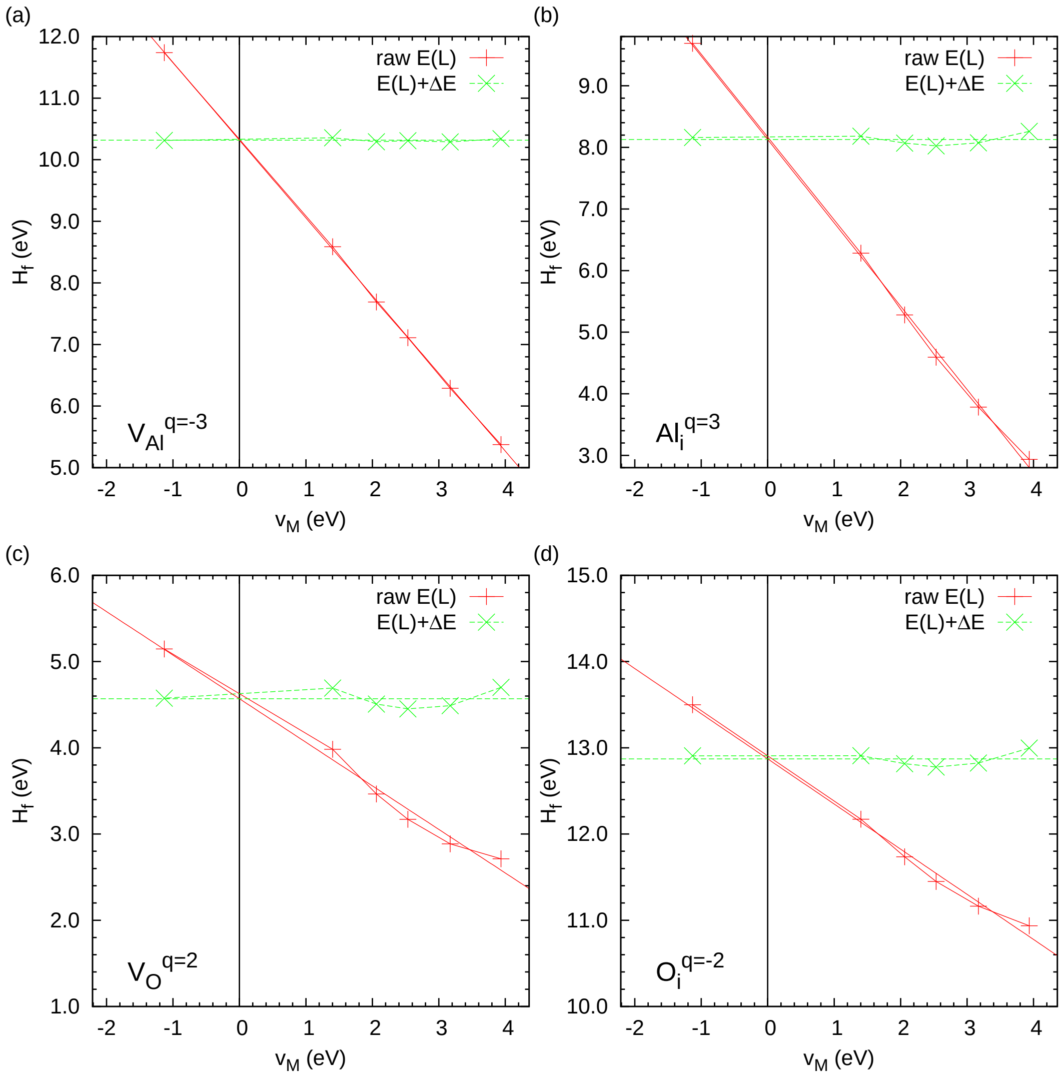

Figure 6.7: Finite size scaling graphs for the principal intrinsic defects of alumina, each in their highest charge states. (a) Aluminium vacancy, charge +3 (b) Aluminium Interstitial, charge +3 (c) Oxygen Vacancy, charge +2 (d) Oxygen Interstitial, charge −2. As expected, the method works best for V_{Al}^{3+} but causes an uncertainty of no more than around 0.2 eV on any of the defects. From left to right along the x-axis, the points represent the cells $2\times2\times3$, $2\times2\times2$, $4\times4\times1$, $3\times3\times1$ and $2\times2\times1$. To test that extending along different axes was not giving different trends, the challenging $3\times3\times2$ calculation (2592 electrons) was tried for V_{Al}^{3+} — it too falls on the same line.

in Alumina is anisotropic, with eigenvalues differing by around 20%. Although the formalism could easily be extended to cope with anisotropy it would probably not give any real benefit. We take the average value between the various defects, $\epsilon_{fit} = 3.7$, and thus apply the correction

$$\Delta E = q^2 v_M / 2\epsilon_{fit} \tag{6.34}$$

to all of our defect cell total energies.

6.2.6 Defect Formation Energies

A vital part of defect formation energy calculations is relaxation of the geometry. When an atom has been added or removed to form a defect, the surrounding atoms feel large forces, especially in cases where the defect site is charged, and relax towards or away from it to find their new equilibrium positions. In the defect supercells we need to leave a region of bulk-like perfect crystal surrounding the defect, so we only allow atoms within a radius of $3.4\,\AA$ to move. The moving atoms include 1st, 2nd and 3rd nearest neighbours of the defect site (referred to as 1NN, 2NN and 3NN respectively). Increasing the radius of relaxation to include 4NN as well changed the results by less than 0.05eV in most cases, so this range was considered adequate. Table 6.2 shows the bond lengths and changes in them for the optimized geometry for each defect species in all accessible charge states. We note good agreement with the geometries of Matsunaga et al [119].

From these relaxed geometries, the final defect supercell energies can be compared with the perfect crystal supercell energy and the chemical potentials calculated by the methods of Section 6.2.3 to find formation energies for each species and each charge state. Figure 6.8 shows the collated results of all of these calculations, with bandgap corrections in the form detailed in Section 6.2.4, and fitted Makov-Payne style defect-defect interaction corrections.

Our results differ in a number of significant ways from previous calculations of the same quantities. For example, Matsunaga et al applied full $m \times \Delta E_g$ bandgap

Species	1 NN	Change (to)	2 NN	Change (to)	3 NN	Change (to)
Bulk Al	0.189	(6 O)	0.272	(4 Al)	0.317	(3 Al, 3 O)
V_{Al}^{0}	0.201	6.82%	0.262	-3.50%	0.318	+0.11%
V_{Al}^{-1}	0.201	6.81%	0.250	-8.04%	0.317	+0.63%
V_{Al}^{-2}	0.201	6.81%	0.257	-5.43%	0.316	-1.58%
V_{Al}^{-3}	0.202	7.16%	0.253	-6.81%	0.314	-2.52%
I site	0.189	(2 Al)	0.195	(6 O)	0.272	(6 Al)
Al_{i}^{0}	0.214	12.96%	0.213	+9.23%	0.279	+2.57%
Al_{i}^{+1}	0.184	-2.65%	0.220	+12.82%	0.281	+3.31%
Al_{i}^{+2}	0.183	-3.17%	0.221	+13.33%	0.281	+3.31%
Al_{i}^{+3}	0.182	-3.70%	0.221	+13.33%	0.281	+3.31%
Bulk O	0.189	(4 Al)	0.268	(12 O)		
V_{O}^{0}	0.186	-1.59%	0.268	-0.06%		
V_{O}^{+1}	0.197	+4.51%	0.265	-1.18%		
V_{O}^{+2}	0.207	+9.55%	0.261	-2.55%		
I site	0.189	(2 Al)	0.195	(6 O)	0.272	(6 Al)
O_{I}^{0}	0.184	-2.65%	0.211	+8.21%	0.276	+1.47%
O_{I}^{-1}	0.176	-6.88%	0.216	+10.77%	0.273	+0.37%
O_{I}^{-2}	0.167	-11.64%	0.222	+13.85%	0.270	-0.74%

Table 6.2: Geometry relaxation of the region around each defect species. For each defect species, distance to and the identities of the the 1NN, 2NN and 3NN atoms in the perfect crystal are shown, followed by the corresponding distances after the geometry has been relaxed. The percentage changes in the 1NN, 2NN and 3NN distances are also shown. As expected, positively charged defects attract the anions and repel the cations, whilst negatively charged defects do the opposite, although this rule is not obeyed strictly far from the defect.

corrections and no defect-defect interaction corrections and all the species of defect existed in their highest charge states over almost the entire ϵ_F range. This is contrary to expectations from experiment and physical intuition, and in the context of the current work can be seen to result from over-stabilization of charged defects within periodic boundary conditions, when no finite-size corrections for defect-defect interactions are applied. Their large bandgap correction also caused there to be negligible regions of stability for certain defects, such as V_O^{+1}, which given the known existence of F^{1+} centers in alumina [61] is unlikely. Figure 6.8 shows a large region of stability for V_O^{+1} in line with expectations.

Several key points about the nature of defects in alumina can be inferred from

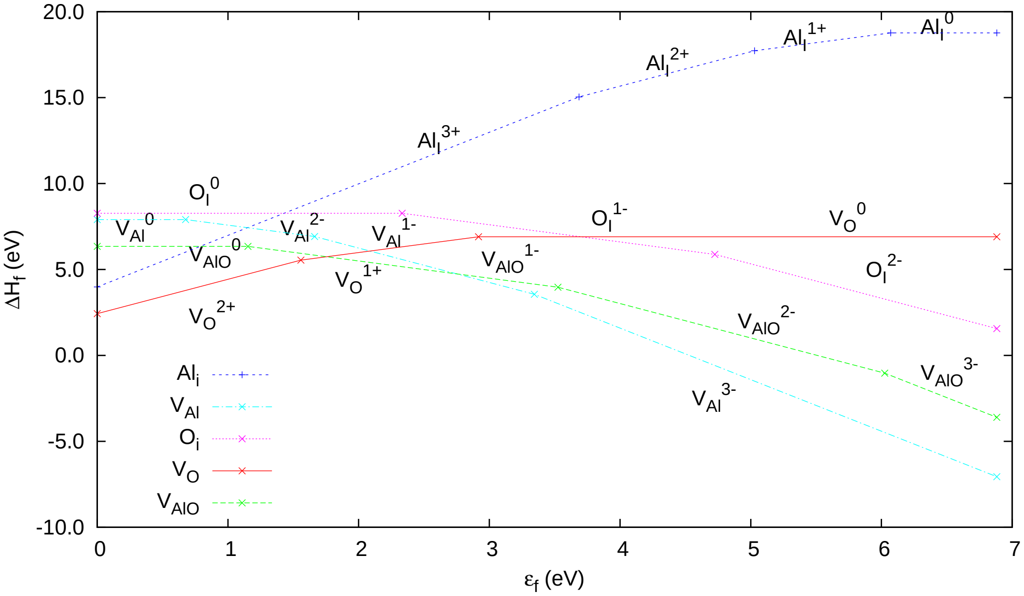

Figure 6.8: Defect Formation Energies of the four main types of intrinsic point defect in Alumina, calculated in DFT with ultrasoft pseudopotentials and the LDA. The determination of μ_O was carried out at $T = 2000\,\mathrm{K}$ and $p_{O_2} = 10\,\mathrm{atm}$, illustrative of a feasible point at which the defect concentrations could become frozen.

Figure 6.8. Firstly, as the Fermi energy varies over its accessible range, the dominant species of defect switches from being Oxygen vacancies at low ϵ_F to Aluminium vacancies at high ϵ_F. This suggests strongly that Schottky disorder (removal of quintets of vacancies) will be the prominent form of intrinsic disorder, with reactions of the form

$$\mathrm{nil} \leftrightharpoons 3V_O^{2+} + 2V_{Al}^{3-} \ . \tag{6.35}$$

Secondly, interstitials are greatly disfavoured relative to vacancies. The natural (and lowest energy) positions for an interstitial are the vacant octahedrally coordinated empty sites of the Al sublattice, where Al ions would sit were it not for the 2:3 coordination. However, this site has nearest neighbours of both types of ion at comparable distances (0.185 nm to 2 Al^{3+} ions, 0.195 nm to 8 O^{2-} ions), and inserting either ion into this site is strongly energetically unfavourable. Presumably, in the case of the oxygen ion, the large ionic radius plays a role in making this the case.

Finally, we address briefly a slightly more complex defect that has been suggested as candidate for oxygen diffusion in alumina [37]. Several other defect complexes have a comparable formation energy to those shown above, but the most prominent one is the AlO defect, V_{AlO}, involving removing adjacent Aluminium and Oxygen ions. As ϵ_F is varied, this is stable in states ranging from $q = 0$ to $q = -3$, and would be the most stable defect over some of the range. As can be seen from its formation energy in Figure 6.8, AlO is promising candidate structure for oxygen diffusion under certain circumstances, particularly if its migration barrier turned out to be lower than that of V_{O} and concentrations under reasonable formation conditions were high. Such issues will be the subject of further investigation beyond the scope of this thesis.

6.3 Quantum Monte Carlo Calculations

6.3.1 Generation of Quantum Monte Carlo orbitals

When generating orbitals for quantum Monte Carlo, one is faced with several constraints that do not apply to DFT. Firstly, the Vanderbilt ultrasoft pseudopotential formulation cannot be applied in QMC as they make use of an augmentation charge which means the orbitals do not represent valid full single-electron states. We are therefore limited to norm-conserving pseudopotentials, and we employ the Trail-Needs pseudopotentials already referred to, as they are designed for accurate QMC calculations. Experience suggests orbitals derived from DFT with the LDA work marginally better in DMC than those derived with the GGA, although there is no firm evidence either way. We use LDA orbitals in the QMC calculations in this work. The plane-wave cutoff was 3000eV as stated, to reduce the variance as much as possible within the available computer memory. Calculations of the DFT defect formation energies with the TN-DF PSPs at 3000eV give nearly identical answers to those obtained using the ultrasoft pseudopotentials at 550eV.

Plane waves are an immensely inefficient representation to use in QMC calculations, as the number of function evaluations to determine the value of the orbital

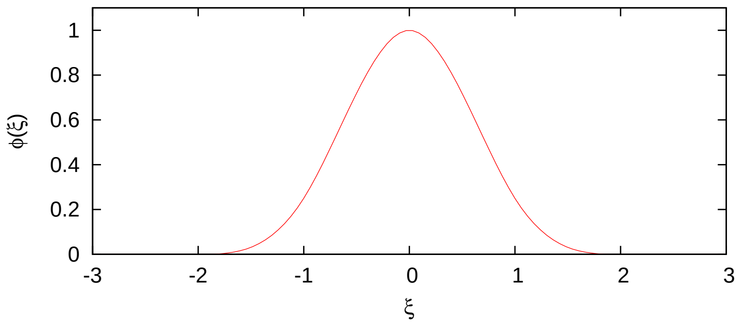

Figure 6.9: Blip function component along one axis, $\phi(\xi)$, centered on zero. A three-dimensional blip function is centered on each grid point and overlaps with the 8 nearest grid points (one in each direction).

at a particular point scales with the system size, and this introduces an unnecessary extra factor of N into the scaling with system size. The plane-wave representation is therefore converted to a blip representation using the methods described in [3]. Blip representations use a set of B-splines, which are localized cubic splines, centered on the points of a regular grid with spacings (a, b, c) along each axis. The blip function centered on $\mathbf{r}_s = (x_s, y_s, z_s)$ has the form

$$\Theta_s(\mathbf{r}) = \phi((x - x_s)/a)\phi((y - y_s)/b)\phi((z - z_s)/c) \ , \tag{6.36}$$

where $\phi(\xi)$ is a localized cubic function with continuous first and second derivatives (see Figure 6.9)

$$\phi(\xi) = \begin{cases} 1 - \frac{3}{2}\xi^2 + \frac{3}{4}|\xi|^3 & 0 \le |\xi| \le 1 \ , \\ \frac{1}{4}(2 - |\xi|)^3 & 1 \le |\xi| \le 2 \ , \\ 0 & |\xi| \ge 2 \ . \end{cases} \tag{6.37}$$

The orbital n is then built from the blip functions at each grid point multiplied by coefficients a_{ns}

$$\psi_n(\mathbf{r}) = \sum_s a_{ns}\Theta_s(\mathbf{r}) \ . \tag{6.38}$$

The blip coefficients are closely related to the plane wave coefficients and can be obtained from them by fast Fourier transforms and some linear algebra (see [3]).

Because the blip functions overlap, at any particular position in the system there are 64 nonzero blip functions. However, the cost to evaluate an orbital at any specific position in the system stays fixed no matter how large the system is made (unlike plane waves, where the number of operations required to evaluate an orbital at a given point scales as L^3). Attempts were also made to localize the orbitals [2] with a unitary transformation among the occupied orbitals. This is designed to obtain so-called “linear scaling”, by transforming the orbitals to a maximally localized form and then truncating them. However, it proved difficult to get reliable and consistent results without extending the truncation radius to a range at which the speedup became minimal as so many of the orbitals contributed in each part of the cell. Localized blip representations of the orbitals were therefore not used in this work.

The large computational cost of QMC calculations means that it is not feasible to work with numbers of electrons as large as those DFT can cope with. Additionally, because electron-electron interactions are calculated explicitly rather than via a mean-field approximation, we must work with a large supercell to converge the electron-electron energy. The dependence of total energy in DFT on k-point sampling in Section 6.2.2 shows that by $2 \times 2 \times 2$ supercells of rhombohedral bulk we have already obtained convergence of the single-particle finite size effects. This is already 80 atoms and 384 electrons, so going to $3 \times 3 \times 3$ is unfeasible in QMC. We thus use the $2 \times 2 \times 2$ supercell in QMC for the bulk investigations. The $2 \times 2 \times 1$ hexagonal supercell of the previous section remains feasible, with 576 electrons. It is expected that Coulomb finite size errors will be minimal, and in any case will approximately cancel between a defect supercell and a perfect supercell.

6.3.2 Jastrow Optimization

With the single particle orbitals constructed, the next step of the calculation is to create a Jastrow factor for each simulation, to express electron correlation. The

Jastrow factor has no effect on the nodes of the system, so in principle, in perfect diffusion Monte Carlo it would have no effect on the total energy. However, as mentioned in Chapter 3, the closer the trial wavefunction is to the true ground state, the lower the variance in DMC and hence the quicker the total energy converges to within a given statistical error. Given the extremely high computational demands of these calculations, this is of the utmost importance.

Given that the evaluation of Jastrow factors themselves is not expected to comprise a large fraction of the total computational effort, a fairly complex form can be chosen, providing sufficient variational freedom to minimize the energy effectively, while not providing so much freedom that finding a reliably optimized minimum proves impossible. Sufficient variational freedom was found to result from a Jastrow factor composed of 1-, 2- and 3-body terms (electron-nucleus χ terms, electron-electron u terms and electron-electron-nucleus f terms in the forms described in [43]). Each term is represented as a power series expansion with optimizable coefficients, multiplied by a step function with a set cutoff and a polynomial that ensures that the right cusp conditions are obeyed automatically. In the u terms, which are functions of electron-electron separation r_{ij}, we chose an expansion up to r_{ij}^8, hence 7 free linear parameters and one cutoff length, with separate parameters for $u_{\uparrow\uparrow}$ and $u_{\uparrow\downarrow}$. For the χ terms, which are functions of electron-ion distance r_{iI} we employed a different χ term for each type of atom in the simulation and again allowed 7 linear parameters and a cutoff length for the optimization. The f terms are functions of the distances r_{ij}, r_{iI} and r_{jI} between two atoms and an ion, and for these we used a power series with up to second order in each distance, although many of them are required to be equal because the electrons are indistinguishable. This leaves 8 free parameters and a cutoff length for the f terms.

All of these parameters were optimized via the accelerated variance minimization scheme of [41]. The optimization successfully recovers a large fraction of the correlation energy: by running separate simulations with and without the Jastrow, we find E_{VMC} and E_{HF} — the latter being the “Hartree-Fock” energy found by using simply the Slater determinant of Kohn-Sham orbitals as our trial wavefunction. Technically this is the DFT exact-exchange energy because the orbitals are not calculated within Hartree-Fock theory, so produce a higher energy than true HF orbitals would. If we assume that the DMC total energy E_{DMC} is very close indeed to the correct answer and recovers somewhere quite close to 100% of the correlation energy, then

$$\frac{E_{VMC} - E_{HF}}{E_{DMC} - E_{HF}} \tag{6.39}$$

estimates the fraction of the correlation energy recovered. In the case of the $2 \times 2 \times 2$ rhombohedral unit cell, the fraction of the correlation energy recovered by our Slater-Jastrow variational trial function is around 91%, which suggests we have a respectably accurate trial wavefunction. However, since the 'correlation energy' accounts for around 60eV of the total energy of −2864 eV, this already suggests our VMC description may not be of sufficient accuracy to resolve defect formation energies of the order of 5 eV.

The situation becomes worse when we try to describe the defect cells in VMC. Because of the low symmetry of alumina almost every atom in a defect supercell is in a slightly different position relative to the defect. The region around the defect is extremely inhomogeneous, which would necessitate a very large number of different χ terms for different sites, and preferably a spatially inhomogeneous u term. The latter is not very feasible with current technology, and the former is limited by the demands of variance minimization, in that it becomes impossible to find a good variational minimum as more and more free parameters are introduced. A good balance was found to be to keep the bulk χ terms fixed on all but the defect site itself and its nearest neighbours of each type. This allowed the correlation introduced by the Jastrow to vary in the local environment of the defect. In the case of vacancies, a χ term was centered on the point from which the atom had been removed, to allow the free charge remaining on the site to be optimized. Figure 6.10 shows the optimized Jastrow χ terms of a supercell containing an oxygen interstitial.

Several interesting conclusions can be drawn from the optimized Jastrows. For example, the vacancy χ terms are negative, showing that in the QMC simulation

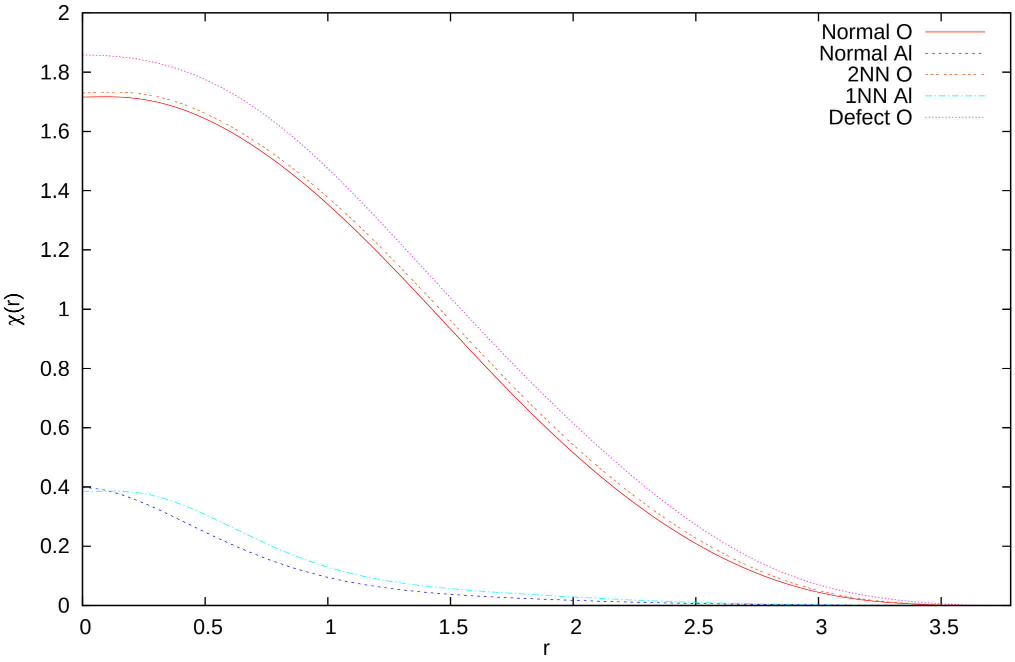

Figure 6.10: Optimized Jastrow χ terms for the supercell containing an oxygen interstitial. The red and dark blue lines are the oxygen and aluminium χ terms respectively, which do not differ significantly from their forms in the perfect crystal. The orange line is the nearest neighbour oxygen χ and the cyan line is the second nearest neighbour aluminium χ. The purple line shows the χ term on the defect oxygen itself. The variance minimization finds a slightly enhanced density on the oxygen site reduces the energy.

the electron density on the vacancy site is reduced relative to that in the DFT simulation. In the case of the interstitials, the χ term is larger on the oxygen interstitial than it is on bulk Oxygen, suggesting the QMC increases the electron density over that of the DFT simulation. Both could be taken as suggestions that the DFT simulation may not be giving an accurate description of the defect electronic structure — a possibility we will return to later.

6.3.3 Timestep Errors

As explained in Section 3.3, the DMC algorithm is only genuinely exact in the limit of infinitesimal timestep, otherwise there is a timestep error. However, this is a relatively well-behaved systematic error that can easily be corrected for. It is of course preferable to work with a timestep short enough that the error is negligible to begin with and can be neglected, but the shorter the timestep, the more steps it takes for a walker to move a large enough distance through configuration space that its energy decorrelates. As the timestep is decreased, therefore, the correlation time increases. The RMS distance diffused by a particle in time T is $\sqrt{2dDAT}$, where d is the dimensionality of the system, D is the diffusion constant $D = \frac{1}{2m}$ (hence $\frac{1}{2}$ for electrons in atomic units), and A is the acceptance ratio (close to 0.5 in VMC, close to 1 in DMC).

If λ is the distance a particle must travel to become decorrelated, which occurs in n moves or time $T = n\Delta\tau$ for timestep $\Delta\tau$, then n can be written

$$n = \frac{\lambda^2}{2dDA\Delta\tau} . \qquad (6.40)$$

Therefore, as long $\Delta\tau$ is small enough that $A \simeq 1$, as is always the case in accurate DMC simulations, $n_{corr} \propto (\Delta\tau)^{-1}$. The computational time to achieve a given error bar follows the same trend, so a tradeoff must be found between accuracy and computational feasibility.

Figure 6.11 shows the total energy of an Oxygen atom calculated using the same Dirac-Fock pseudopotential used in the Alumina calculations, for a range of timesteps. The behaviour, as expected, is roughly linear in $\Delta\tau$ and can be accurately extrapolated to a zero timestep value of -431.425 eV. Figure 6.12 shows the same graph for the $2\times2\times2$ rhombohedral supercell of Al_2O_3 (80 nuclei and 384 electrons). This time, the timestep error per atom converges much more rapidly with decreasing timestep — representing the greater homogeneity of the solid system compared to the atomic one. The timestep error at 0.005 au is already only 0.004eV/atom so no timestep correction need be applied to these results as it would be dwarfed by the statistical error.

There is no reason to believe timestep error would have a systematic effect on defect formation energies: given that the cells being compared are very similar except in the vicinity of one atom, it is likely that the errors almost exactly cancel

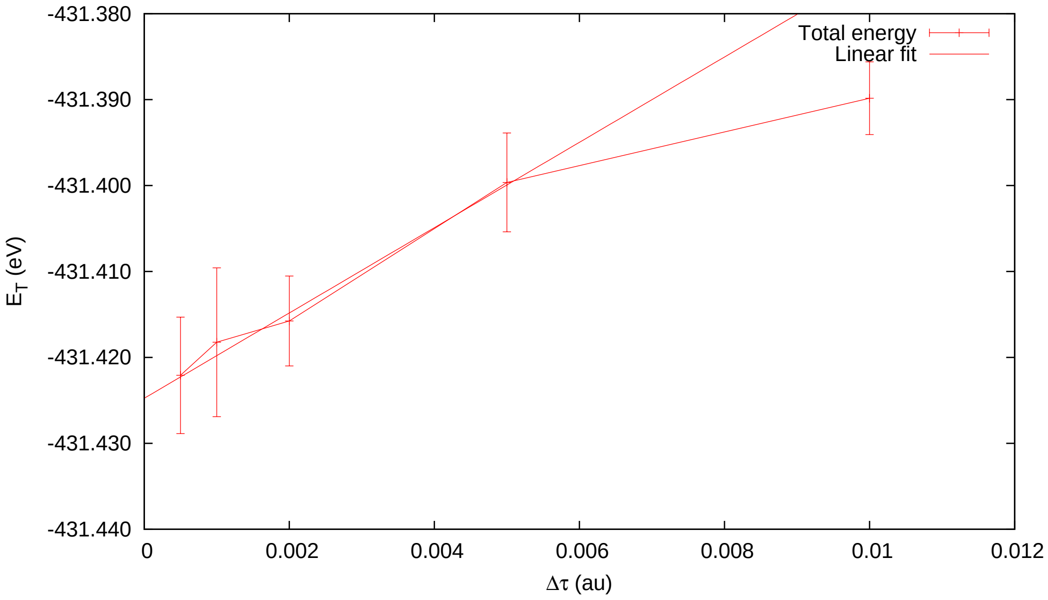

Figure 6.11: Timestep error on the total energy E_T of the oxygen atom as timestep $\Delta\tau$ is varied. With a timestep of $\Delta\tau = 0.005$ the timestep error is comparable to the statistical error on the energy, and is around than 0.025 eV/atom.

apart from a contribution comparable to that of one atom — which again is dwarfed by the statistical error.

On the evidence of these tests, a timestep of $\Delta\tau = 0.005\,\mathrm{au}$ was chosen for all the simulations presented here, as a balance between accuracy and computational demands. In the case of the single-atom results, where the timestep error is more significant, the linear fit shown in Figure 6.11 was used to extrapolate to zero timestep.

6.3.4 Pseudopotential Localization Error

As discussed in Section 3.3.6, the use of non-local pseudopotentials in DMC introduces an approximation into the otherwise exact DMC scheme, because the nonlocal operator is applied to the trial wavefunction rather than the true ground state. This error is known to be relatively small as long as the trial wavefunction is good, but since it is an uncontrolled approximation it is important to estimate its size. This can be done by applying the scheme of Casula [24] which forces the estimate of the

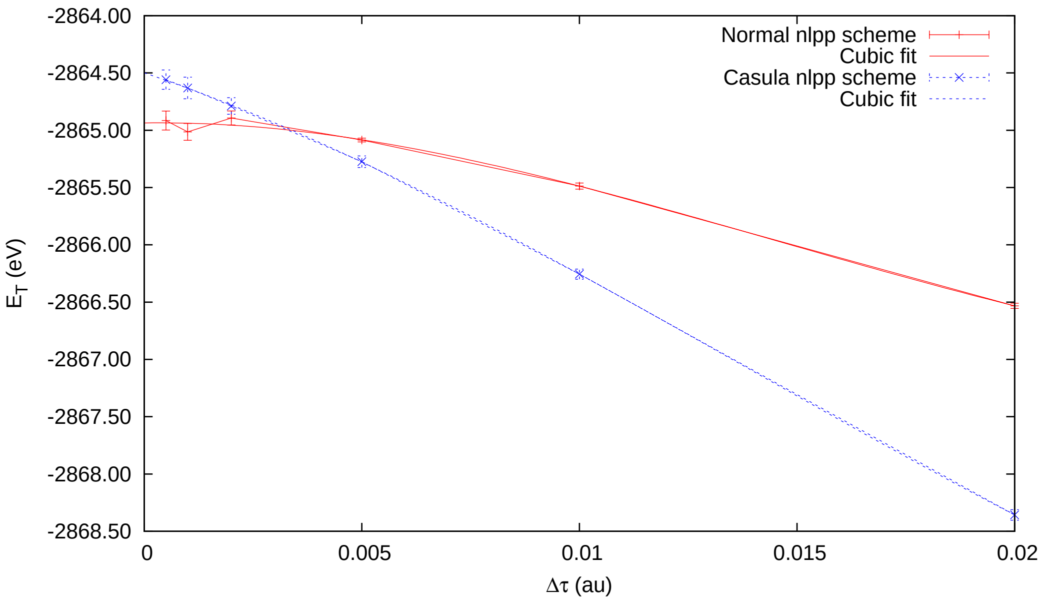

Figure 6.12: Timestep error on the total energy E_T of a rhombohedral cell of alumina as timestep $\Delta\tau$ is varied. The value chosen for the simulations described here was 0.005 au, which corresponded to a timestep error of less than 0.004 eV/atom with the normal nonlocal pseudopotential scheme (red). Also shown are results for the Casula scheme for nonlocal pseudopotentials (blue) - see Section 6.3.4, which demonstrate a much more significant timestep error. For this reason, the Casula scheme was not used in this research.

nonlocal pseudopotential energy to be variational with respect to the real energy (or at least its fixed-node value). In general, use of the Casula scheme raises the total energy, and the amount by which it does so gives an estimate of the pseudopotential localization error. From the y-intercepts of the fitted curves in Figure 6.12, we see that the Casula scheme, when extrapolated to zero timestep, gives an answer different from the standard nonlocal pseudopotential scheme by 0.49 eV, or 0.24 eV per formula unit.

However, the Casula scheme comes at the cost of a greatly increased timestep dependence of the total energy. If we continued to work at the timestep of 0.005 au which was accurate in the normal method, we incur an unacceptable timestep error of nearly 2 eV. Because of the ensuing increase in runtime, it was not feasible to reduce the timestep to a level sufficient to for adequate accuracy in the defect formation energy calculations when using the Casula scheme. Fortunately, there is every reason to expect that when comparing two almost identical cells and using the standard scheme, almost all of the pseudopotential localization error will cancel in the comparison, leaving approximately the contribution to the error from one atom. This, by the above logic, is no more than around 0.05 eV, and will be swamped by the statistical error. The pseudopotential localization error can therefore safely be neglected in the defect formation energies.

For thermodynamic formation energy calculations of the bulk materials and of the elemental forms, which will be discussed shortly, however, which rely on total energies of the whole cell, pseudopotential localization error should not be neglected as it skews the results notably. In Section 6.3.5 we include this correction to the total energy of alumina per formula unit to the formation and atomization energies.

6.3.5 Binding energies and Bulk Properties

To check the accuracy of our description of bulk alumina and to demonstrate the improvement in accuracy possible in DMC, we have calculated various properties of bulk alumina and its constituents. In Table 6.3 we show the formation energy of the atomic form (Al and O atoms) from their standard states (Al metal and O_2 gas respectively), plus the formation energy of an AlO dimer and the atomization energies and formation energies of bulk alumina. While the comparison to experiment is not an exact one as it neglects all temperature dependence (the experimental results are for standard pressure and temperature whereas the theoretical ones are effectively zero temperature) it can be seen that DMC does a very significantly better job of calculating all these quantities than any of the DFT formulations tested. This allows us considerable confidence in our description of the solid.

It is also important to verify that the optimized geometry does not differ to any significant degree within QMC and DFT. If it does, it invalidates our transferral of DFT defect geometries to QMC. While advances have been made recently in the QMC calculation of forces [9], no methods are currently available that allow the accurate determination of forces in large periodic systems. Additionally, due to

Method	$\Delta_f H_0$(Al)	$\Delta_f H_0$(O)	$\Delta_f H_0$(AlO)	$\Delta_a H_0(\mathrm{Al_2O_3})$	$\Delta_f H_0(\mathrm{Al_2O_3})$
LDA-USP	4.05	3.62	0.91	-37.09	-18.15
LDA-DF	4.10	3.67	1.13	-36.48	-16.95
GGA-USP	3.41	2.82	0.74	-30.22	-14.94
DMC	3.47(1)	2.54(1)	0.68(1)	-32.62(3)	-18.04(3)
Expt.	3.42	2.58	0.69	-31.95	-17.37

Table 6.3: Formation and Atomization Energies (all in eV). $\Delta_f H_0$(Al) and $\Delta_f H_0$(O) are the formation energies per atom of Al and O atoms in the gas phase. $\Delta_a H_0(\mathrm{Al_2O_3})$ and $\Delta_f H_0(\mathrm{Al_2O_3})$ are the atomization and formation (cohesive) energies per formula unit of alumina. $\Delta_f H_0$(AlO) is the formation energy of an Aluminium-Oxygen dimer. Experimental data is obtained from the NIST Webbook [32].

the computational demands of DMC, it is unfeasible to exhaustively search the large parameter space provided by the range of possible values of the four geometry parameters discussed in Section 6.2.1. Consequently only the effect of varying a_{rho} will be investigated here with QMC, as it was the parameter which had the largest effect on the total energy of the cell, and because the other three do not significantly change for small variations of a_{rho} in DFT.

Figure 6.13 shows the behaviour of the total energy in various methods to investigate the accuracy of using DFT geometries in QMC. Because QMC cannot at present perform automatic geometry optimization, one must take a slightly roundabout route to verifying this. Shown in red in Figure 6.13 is the DFT curve of total energy vs volume, with its minimum at V_0 corresponding to the optimized geometry of Section 6.2.2. The DFT orbitals for each lattice parameter were then used in QMC simulations, with Jastrow factors which were separately reoptimized for each system. In green is the curve of E_T vs V/V_0 obtained within VMC. The independent Jastrow optimization renders the determination of a minimum difficult as it introduces a certain amount of noise that is determined by how good a minimum the optimization found in each case. This noise is washed out at the DMC level, however, as the total energy is independent of Jastrow quality in the fixed node approximation. The DMC curve is shown in blue, and it can be verified both visually and by fitting a quadratic that the QMC and DFT minima are not at a significantly

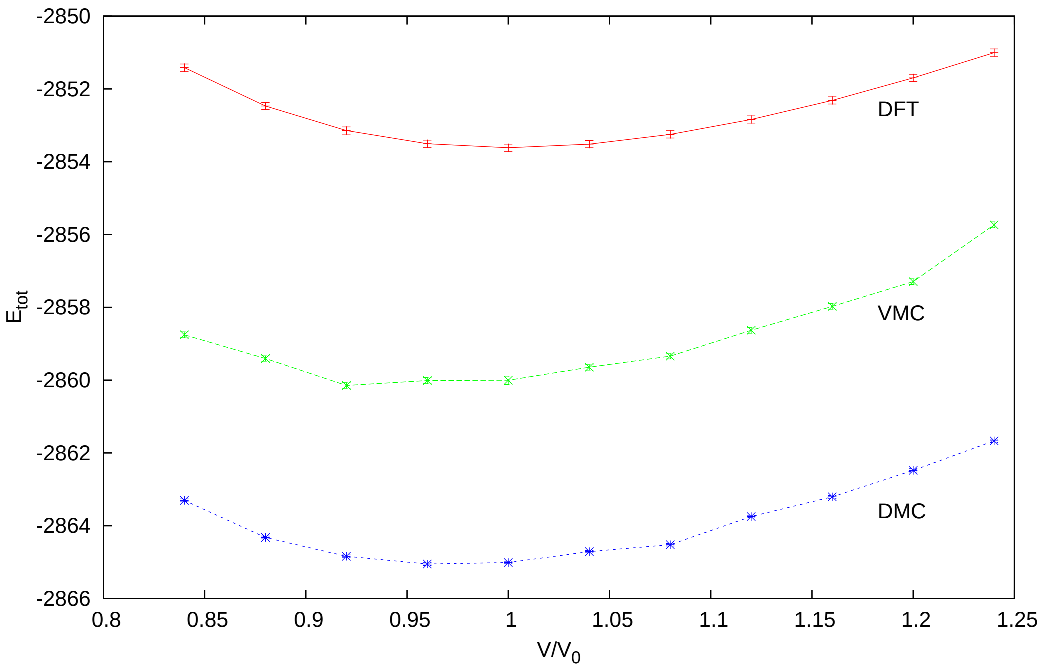

Figure 6.13: Total energies of a rhombohedral $2\times2\times2$ supercell of Al_2O_3 calculated using different methods, as a function of primitive cell volume V relative to its DFT equilibrium value V_0. The minima of the curves fall are very close, supporting the idea that DFT geometry can be accurately used in QMC calculations.

different position, suggesting that it is indeed sensible to use DFT geometries in our QMC simulations.

6.3.6 Band Gaps in QMC

As a final test of the bulk system, we calculate the band gap within QMC, to demonstrate its high accuracy for excited state energies. The experimental energy gap of alumina is difficult to determine for various reasons, but is generally accepted to be around 8.7 eV at 300 K [55]. It is also known to be quite strongly temperature dependent. This temperature dependence comes about from the strong dependence of the band gap on the lattice parameter. Using the parameterized temperature dependence given in [55] to extrapolate to 0 K gives a zero-temperature band-gap of 9.1 eV, which we take as our experimental value for comparison. There are two main methods that can be used for the calculation of bandgaps in QMC simulations, each with associated problems. The first is to relate the gap to the difference between the energies of adding and subtracting a particle from the system in its neutral state, as:

$$E_g = E(N+1) + E(N-1) - 2E(N)\ , \tag{6.41}$$

where $E(N)$ is the total energy of some supercell of the bulk and $E(N \pm 1)$ are the total energies of this system with one electron added or subtracted. The latter are subject to a large finite size effects because this extra or missing electron is squeezed into the volume of the supercell and repeated periodically. The results converge, in the limit of an infinite cell, but even in DFT the infinite simulation cell limit is not the same as the limit of perfect k-point sampling in this case. These finite size effects are easily demonstrated in a simpler system with a small unit cell, such as NaCl, where a range of sizes can easily be compared.

In QMC, there is a significant caveat to using Eq. 6.41. In charged periodic systems evaluated with the Ewald method there is a strong finite size effect in the Coulomb interaction which does not necessarily cancel between the $E(N+1)$ and $E(N-1)$ systems. Appendix C discusses the embedding of one Ewald calculation inside another, by calculating the energy of one system embedded inside a larger but still periodic one and taking the size of the larger system to infinity. This is of relevance both to the calculation of the energy of the perfect crystal with electrons added or removed and to the energies of the defect cells when the defect is charged. The conclusion arrived at after significant algebra is that while in DFT, the energies of charged systems can be evaluated approximately correctly by including a neutralizing background, this is not the case in QMC. In QMC simulations, the electron positions are explicit and fixed at specific values at any particular timestep, and this produces an error equivalent in magnitude to the Madelung term of the Makov-Payne correction but the sign of which depends on whether the charge was added or taken away relative to the bulk. We thus correct this error by adding $\pm\frac{1}{2}qv_M$, the energy of the point charge sitting in the cancelling background. Although in Eq. 6.41 the two corrections cancel each other exactly, this correction is vital in order

Method	Excitation	Charging
LDA-USP	6.78	6.84
GGA-USP	6.31	6.95
DMC	9.4(4)	10.6(5)
Experiment [55]	9.1	-

Table 6.4: Band gaps in DFT and DMC with both the methods described in this section (in eV). Only the DMC value by the excitation method is in agreement with experiment.

to calculate a sensible value of E_{VBM} in QMC, which is necessary for the formation energy of charged defects.

However, there is an alternative route that may be more reliable for calculating band gaps. We can compare the energy of the normal, neutral system with that of a version where a promotion has been made in the Slater determinant of single-particle orbitals, taking an electron from the highest occupied state and moving it to the lowest unoccupied one. The gap is then simply

$$E_g = E_T(\mathrm{PR}) - E_T(\mathrm{GS}) \ , \tag{6.42}$$

where $E_T(\mathrm{GS})$ and $E_T(\mathrm{PR})$ are the total energies of the system in the ground state and in the excited state respectively. The disadvantage of this method is that what one has created, effectively, with the promotion from the valence band to the conduction band is an exciton, which has some unknown exciton binding energy that contributes to the energy. However, since the exciton binding energy is likely to be a small contribution on the scale of the associated statistical error, this is less likely to be a problem than the Coulomb errors of the previous method.

Table 6.4 shows the band gap calculated using both methods, in DFT (with LDA and GGA) and in DMC, compared with the experimental value discussed above. Where excitation gaps are listed in DFT, this is simply the difference in the eigenvalues between the top of the valence band and the bottom of the conduction band.

The DFT values are all strongly affected by band gap errors in DFT, and the DMC band gap calculated with the charged systems also differs by 1.5 eV from the experimental value — possibly an uncancelled finite size effect. However, the value calculated by the promotion method agrees within the DMC error bar with the experimental value. This gives us confidence in the ability of QMC to give correct formation energies for defects involving electrons in conduction band-like states.

6.3.7 Formation Energies

Figures 6.14 and 6.15 are the central result of this chapter, showing the results of our DMC calculations of the formation energies of oxygen vacancies and interstitials. We chose to concentrate on the oxygen defects because it is the diffusion of oxygen that underlies the so-called ‘conundrum’ relating to the diffusion properties of alumina [70] — namely that theoretical results for formation energies are difficult to reconcile with the available experimental data. Most of the high-temperature properties of alumina (sintering, creep, formation and stability of oxide scales on metals containing aluminium) depend on the diffusion of oxygen around the lattice, so in order to try to resolve this ambiguity, we study the oxygen defects with DMC to determine if the DFT formation energies are accurate.

In the course of generation of trial wavefunctions, we performed a VMC study of the defects shown here. Because of the difficulty in describing the correlation in the defect cell accurately, the formation energies in VMC come out many eV too large. For example, the neutral oxygen defect has $\Delta H_f^{VMC} = 15.8(6)\,\text{eV}$, compared to the DMC value of 7.36 eV. The VMC simulations were therefore not used for anything other than the generation of an initial set of walker configurations for the ensuing DMC simulations.

Figure 6.14 shows the DMC formation energy of the oxygen vacancy compared to the LDA results for the TN-DF pseudopotentials over the range of Fermi energy for which both are valid ($\epsilon_F = 0 \rightarrow E_g^{DFT}$), with the oxygen chemical potential chosen to represent that at $T = 273\,\text{K}$ and $p = 1$ atm. The effects of changing μ_O are described by the $\Delta\mu_O$ and ϵ_F dependence of the values in Table 6.5. As

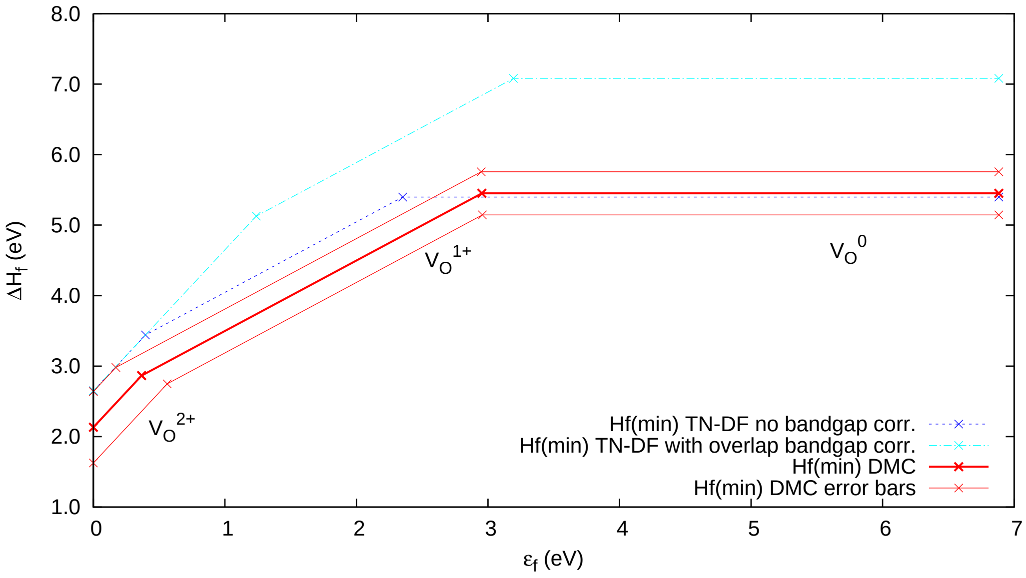

Figure 6.14: DMC formation energies, with associated error bars, of the Oxygen vacancy in its various charge states, compared to LDA calculations with TN-DF pseudopotentials, with and without bandgap corrections.

shown in Section 6.2.3, the DFT result is essentially independent of the functional and pseudopotential, unless bandgap corrections are taken into account. The DMC results, however, differ significantly from these results, coming out notably below the DFT lines in each case, even if no bandgap correction is included. If it is included, the DFT result is pushed higher and the DMC line is beneath the DFT line by $1-2\,\text{eV}$. This does not provide a clear answer to the question of whether the bandgap correction is a sensible approximation, as it could be interpreted either as evidence that it should not be applied (in which case $\Delta H_f^{DFT}(V_O^0)$ is close to $\Delta H_f^{DMC}(V_O^0)$), or that the band gap correction should be applied, but that the DFT result is uniformly too low across its range. Figure 6.16 shows a combined plot of both the DFT and DMC results for all the main defect species.

Figure 6.15 shows the formation energies of the oxygen interstitial, and is less ambiguous as there are no bandgap corrections in this case. Here, the DMC result is considerably higher than the DFT energies, by over $2\,\text{eV}$. The most important conclusion one can draw from this is that the DFT-LDA appears to significantly

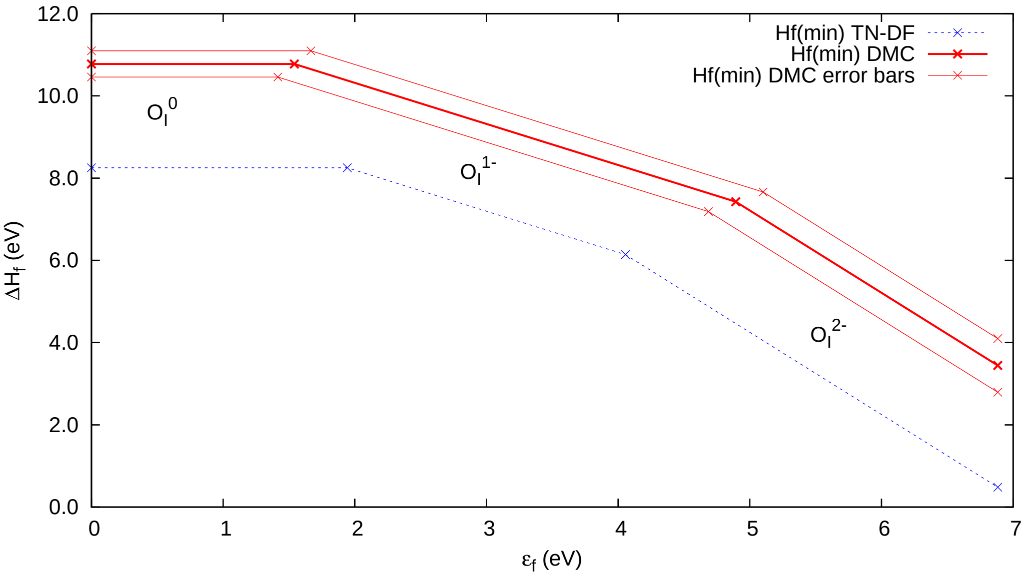

Figure 6.15: DMC formation energies, with associated error bars, of the Oxygen interstitial in its various charge states, compared to LDA calculations with TN-DF pseudopotentials, with and without bandgap corrections.

overbind the oxygen interstitial. This is in line with both the results for cohesive energies in Table 6.3, with general experience with DFT, and with previous studies of defects in DMC. Leung, Needs, et al. [110] found this when studying self-interstitials in bulk silicon: while LDA and GGA formation energies for the stablest interstitial were 3.3 and 3.8 eV respectively, the DMC value was 4.9 eV. The same conclusion was found by Batista et al., who studied a wider range of DFT functionals [15] and found uniformly that DFT overbound the interstitial. This conclusion is given further weight by the results presented here. It suggests that in the situation of the oxygen interstitial, where the local environment, bonding and coordination number of the interstitial oxygen is extremely different from that of the bulk oxygen atoms, DFT fails to describe its electronic structure correctly and reports it to be more stable than it really is. Interestingly, the only functional tested in [15] that was able to reproduce the DMC results to any degree was a hybrid function including screened exchange, the Heyd-Scuseria-Ernzerhof (HSE) functional [71, 72, 73].

The other main point to note from these results is that the DFT calculations

Defect	ΔH_f^{DFT} (eV)	ΔH_f^{DMC}(eV)
V_O^0	$7.38 + \Delta\mu_O$	$7.3(3) + \Delta\mu_O$
V_O^{+1}	$5.03 + \Delta\mu_O + \epsilon_F$	$4.4(3) + \Delta\mu_O + \epsilon_F$
V_O^{+2}	$4.63 + \Delta\mu_O + 2\epsilon_F$	$4.0(5) + \Delta\mu_O + 2\epsilon_F$
O_I^0	$6.26 - \Delta\mu_O$	$8.9(3) - \Delta\mu_O$
O_I^{-1}	$8.21 - \Delta\mu_O - \epsilon_F$	$10.4(4) - \Delta\mu_O - \epsilon_F$
O_I^{-2}	$12.26 - \Delta\mu_O - 2\epsilon_F$	$15.3(6) - \Delta\mu_O - 2\epsilon_F$

Table 6.5: Formation energies of the Oxygen defects in Alumina, in DFT and DMC, with the dependence on varying chemical potentials μ_O and μ_e shown explicitly.

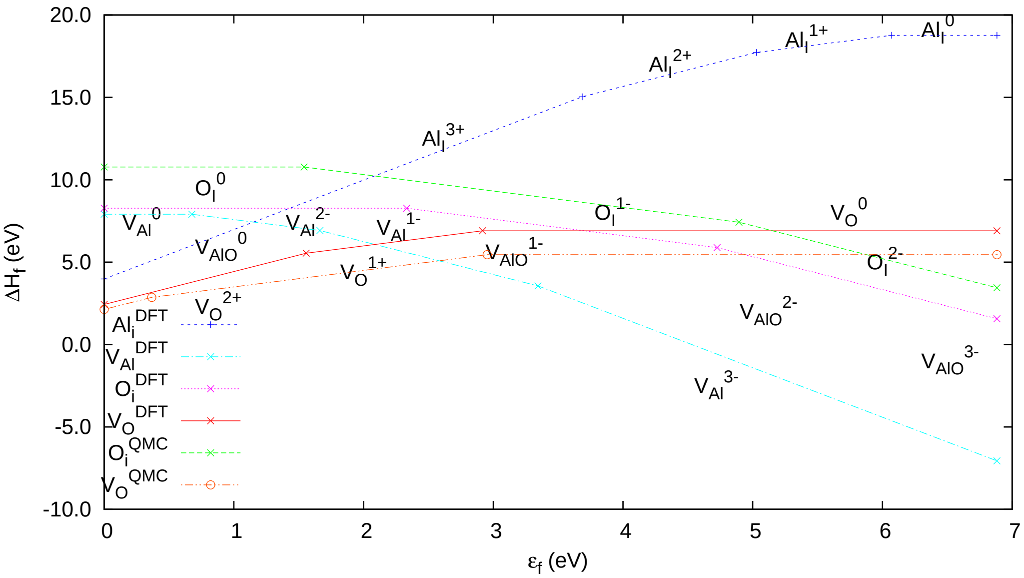

Figure 6.16: DFT and DMC formation energies of the four main types of intrinsic defect, showing the effects of DFT overbinding.

appear to be quite strongly affected by self-interaction errors. The localized charge density associated with defect states is an obvious situation where this would occur. If a single electron is localized on a defect, it should not have any contribution to its Hartree energy from its own charge distribution. In DFT this is not the case, as every electron feels the same effective potential, to which every electron contributes. This biases the DFT results upwards for cases where there is a lone localized electron (such as V_O^{1+}), resulting in a smaller region where this is the stablest defect. In DMC, it is observed that the region of stability for V_O^{1+} is considerably larger than in DFT, which we attribute to the correction of the self-interaction error.

Given the success of DMC in achieving highly accurate results for quantities such binding energies, it is reasonable to expect that its power to improve on DFT in those situations where DFT is of insufficient accuracy will prove to carry over to defect formation energies. If the indications provided by both these calculations and the previous defect studies are correct, then significant amounts of DFT work on defects may have to be re-thought.

Chapter 7

Conclusions

This thesis has presented work in three separate areas in which there was a demonstrable need for QMC simulations, whether to act as a benchmark, as a demonstration and proof of concept, or to investigate the differences between single-electron and many-electron phenomena. Additionally, the DFT studies, which in many cases were intended as a preliminary to the QMC simulations and to generate single-particle orbitals for use in QMC, have in some cases improved on previous similar calculations or investigated previously unexplored phenomena.

The areas investigated here, while ostensibly separate, are loosely connected by a number of similar themes. Firstly, they are all strongly affected by finite size effects. Solutions to the problems of finite size effects fall broadly into two categories: extrapolation (or interpolation, via certain tricks) to infinite size, or analytic corrections to the results of finite systems, and both are examined here in the course of the various topics covered. In the case of polarization and localization lengths, the results do not converge until the approximations used in the derivation of usable expressions become accurate, and this occurs only slowly with increasing k-point sampling. Fortunately, the exponential localization of the Wannier functions ensures that once the system size reaches a certain point, finite size effects on polarization and localization properties decay rapidly. In the Jellium slab calculations, it is finite-size errors that have rendered previous calculations unreliable. Errors arising from the slab-width dependence they easily controlled, and by working at special slab widths one can negate them. The in-plane effects are more of a challenge but we have presented a method that we believe has been able to overcome the worst of the resulting bias. Finally, in the case of point defects, it is the interaction between periodic replicas of the defect which we must attempt to negate, and we have presented both an empirical approach to this problem an attempt to correct the boundary conditions so as to remove the effect in the first place.

The second main linking theme is the benchmarking and testing of Density Functional Theory. Diffusion Monte Carlo is not, in the near future, likely to be able to take over from DFT as a powerhouse of large scale electronic structure calculations — the computing time costs remain simply too high for it to be viable in systems of the size required to study many phenomena. However, its roles in benchmarking and checking the accuracy of DFT, investigating the circumstances under which the approximations and corrections to DFT are valid, and evaluating the success of the DFT description of many-electron phenomena are well established and represented here in the motivation of all three studies.

In some ways, the quantum Monte Carlo side of the localization lengths study represents a something of a dead end, as the phenomenology seen in the correlated many-body systems was qualitatively similar to that seen in the DFT calculations. If one were to investigate a metal-insulator transition with quantum Monte Carlo, the localization length would be infinite on the metallic side and finite on the insulating side, but we have shown that it is only under certain specific circumstances that a divergence is seen as the transition is approached from the insulating side, and most transitions in real materials do not match these circumstances. Additionally, the determining factor in how the localization length behaves is the sensitivity to boundary conditions of the single particle orbitals. While correlation via a Jastrow factor can change the length itself, it cannot change boundary condition sensitivity, so while it is evaluated with an operator that makes clear that it is inherently a many-body phenomenon, it nevertheless remains essentially a one-electron property. In the course of the DFT investigation, we have also examined the properties of the localization tensor over a transition. It is normally assumed to diverge as an insulator to metal transition is approached from the insulating side. However, our results show that this is not the case: the band insulator to metal transition is normally first order and not associated with a diverging length scale. We have also identified examples where the localization length diverges but the system is insulating on both sides of the divergence. The usefulness of the localization length as an indicator of the approach of an insulator to metal transition is therefore limited to certain particular circumstances.

The surface energy of the electron gas investigation represents the most clear-cut success of this thesis, although it less novel in some ways than the other sections. It is undoubtedly the most accurate calculation to date of the surface energy of jellium slabs within DMC, and corrects most of the errors that have reduced the accuracy of previous calculations below the level usually expected of DMC. The only uncontrolled error remaining that could affect the surface energy values is a systematic variation of fixed-node error with slab width — which there is no reason to believe exists. Recent advances in DMC techniques might allow these calculation to be made even more accurate, and reduce the error bars that result from uncertainty on the fitting parameters of the extrapolation. These would include twist-averaging to eliminate the independent-particles finite-size error, inclusion of backflow to further reduce the fixed-node error, and a full treatment of the long-range kinetic energy finite size error in the manner of Ref. [33]. These additions do not necessarily require any advance in computing power to make them feasible, but would require some theoretical work in the case of backflow, to find a compact way of parameterizing a form of inhomogeneous backflow suitable for slab systems.

Finally, the investigation of point defects in alumina has reached a number of interesting conclusions: previous work within DFT has been found to have neglected the crucial effects of the interactions of the periodic replicas of charged defects, and also to have treated band gap corrections with a scheme only really applicable to semiconductors and not valid for the tightly-bound localized orbitals of an ionic metal oxide. Our calculations of the full spectrum of defect formation energies with respect to the chemical potentials of the components reveals that there is a richer spectrum of possible charge states than previously thought, as past calculations reported the higher charge states as over-stabilized for the reasons stated. Within QMC, we have overcome the hurdles to successful calculation of defect properties, and presented the first DMC calculation of the formation energy of charged defects. Our results show intriguing differences relative to the DFT energies, and suggest that in agreement with the work of Leung, Hood, Needs et al., overbinding of solids within the LDA leads to significant overestimates of the formation energies of vacancies and underestimates of the formation energies of interstitials.

Further work would be of value in a number of areas of this study of defect in metal oxides, with both computational methods used. Following up the work on boundary condition corrections by implementing it in a DFT code inside the SCF cycle could allow defect formation energies and geometries to be calculated without needing to apply extrapolative correction formulae. Further work, perhaps using linear scaling DFT [152] to simulate larger and larger systems, could be used to determine whether the method of varying the cell shape as well as size produces reliable results when extrapolating to infinite size. Comparison of other oxides would also be profitable — and this is indeed planned. TiO_2 in its rutile structure is one obvious candidate, if problems with the representation of its pseudopotential can be overcome. Comparison of a number of oxides may make it possible to determine more definite answers to two remaining questions, namely whether the overbinding suggested by the differences between the DMC and the DFT is a systematic effect, and whether or not band gap corrections should be applied and to what extent. If the former is true, it may suggest that DFT in the formulation used here may in many cases lack the accuracy required to be a useful predictive tool in the field of point defects in ionic crystals.

Appendix A

The Kinetic Energy Term in DMC Finite Size Errors

In this appendix, we derive an expression for the long-range kinetic energy terms of the finite size error in a jellium slab calculation of the type described in Chapter 5. The derivation follows a similar route to a recent paper by Chiesa [33], which examined the homogeneous electron gas. Here we use the same approach, employing the form of the plasmon normal modes presented by Wood for a jellium slab system [180].

Within the inhomogeneous Random Phase Approximation [60] we can describe the ground state as the product of a short range part and the long-ranged collective motion of the plasmon normal modes:

$$\Psi = \Psi_s \exp\left[-\frac{1}{2}\sum_{i,j} u(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \chi(\mathbf{r}_i)\right] \ . \tag{A.1}$$

We define the transform of the u term:

$$u_{\mathbf{G},\mathbf{G}'} = \int \mathrm{d}^3\mathbf{r} \int \mathrm{d}^3\mathbf{r}'\, u(\mathbf{r}, \mathbf{r}')\mathrm{e}^{-i\mathbf{G}.\mathbf{r}}\mathrm{e}^{-i\mathbf{G}'.\mathbf{r}'} \ , \tag{A.2}$$

$$u(\mathbf{r},\mathbf{r}') = \frac{1}{\Omega^2}\sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\mathrm{e}^{i\mathbf{G}.\mathbf{r}}\mathrm{e}^{i\mathbf{G}'.\mathbf{r}'} \ , \tag{A.3}$$

and of the chi term:

$$\chi_{\mathbf{G}} = \int \mathrm{d}^3\mathbf{r}\,\chi(\mathbf{r})\mathrm{e}^{-i\mathbf{G}.\mathbf{r}} \ , \tag{A.4}$$

We can now express the Jastrow Factor in terms of components of the momentum density $\rho_{\mathbf{G}} = \sum_i \mathrm{e}^{i\mathbf{G}.\mathbf{r}_i}$:

$$\Psi = \Psi_s \exp\left[-\frac{1}{2\Omega^2}\sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\sum_i \mathrm{e}^{i\mathbf{G}.\mathbf{r}_i}\sum_j \mathrm{e}^{i\mathbf{G}'.\mathbf{r}_j} + \frac{1}{\Omega}\sum_{\mathbf{G}}\chi_{\mathbf{G}}\sum_i \mathrm{e}^{i\mathbf{G}.\mathbf{r}_i}\right] \ , \tag{A.5}$$

$$\Psi = \Psi_s \exp\left[-\frac{1}{2\Omega^2}\sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\rho_{\mathbf{G}}\rho_{\mathbf{G}'} + \frac{1}{\Omega}\sum_{\mathbf{G}}\chi_{\mathbf{G}}\rho_{\mathbf{G}}\right] \ , \tag{A.6}$$

where the sums are over all reciprocal lattice vectors $\mathbf{G}$ commensurate with the periodically repeated simulation cell.

The kinetic energy can be written as $T_N = -\frac{1}{4}\sum_{j=1}^{N}\left\langle \nabla_j^2 \ln \Psi \right\rangle$ by Green's identity, so applying this to the above wavefunction gives us

$$T = T_s - \frac{1}{4}\left\langle \sum_k \nabla^2_{\mathbf{r}_k}\left(\frac{-1}{2\Omega^2}\sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\big(\sum_i \mathrm{e}^{i\mathbf{G}.\mathbf{r}_i}\big)\big(\sum_j \mathrm{e}^{i\mathbf{G}'.\mathbf{r}_j}\big)\right.\right. \tag{A.7}$$

$$\left.\left. + \frac{1}{\Omega}\sum_{\mathbf{G}}\chi_{\mathbf{G}}\sum_i \mathrm{e}^{i\mathbf{G}.\mathbf{r}_i}\right)\right\rangle . \tag{A.8}$$

where T_s is the kinetic energy associated with the short-ranged part of the wavefunction, Ψ_s, and the rest is the kinetic energy associated with the long-ranged Jastrow factor. The Laplacian acting on the $\rho_{\mathbf{G}}\rho_{\mathbf{G}'}$ term gives:

$$\begin{aligned}
\sum_k \nabla^2_{\mathbf{r}_k} \sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'} \big(\sum_i \mathrm{e}^{i\mathbf{G}.\mathbf{r}_i}\big)\big(\sum_j \mathrm{e}^{i\mathbf{G}'.\mathbf{r}_j}\big) \\
= \sum_k \sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'} \Big[-G^2 \mathrm{e}^{i\mathbf{G}.\mathbf{r}_k} \big(\sum_j \mathrm{e}^{i\mathbf{G}'.\mathbf{r}_j}\big) - G'^2 \mathrm{e}^{i\mathbf{G}'.\mathbf{r}_k} \big(\sum_i \mathrm{e}^{i\mathbf{G}.\mathbf{r}_i}\big) \\
+ 2i\mathbf{G}\mathrm{e}^{i\mathbf{G}.\mathbf{r}_k}.i\mathbf{G}'\mathrm{e}^{i\mathbf{G}'.\mathbf{r}_k}\Big] \\
= -\sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'} \Big[(G^2 + G'^2)\rho_{\mathbf{G}}\rho_{\mathbf{G}'} + 2\mathbf{G}.\mathbf{G}'\rho_{(\mathbf{G}+\mathbf{G}')}\Big]
\end{aligned} \tag{A.9}$$

According to RPA theory, the χ term is expected to cancel out the effect of the u term on the density, and can be written as

$$\chi(\mathbf{r}) = \int \mathrm{d}^3\mathbf{r}' u(\mathbf{r},\mathbf{r}')\langle\rho(\mathbf{r}')\rangle \ , \tag{A.10}$$

so its Fourier transform can be expressed in terms of that of the u term.

$$\begin{aligned}
\chi_{\mathbf{G}} &= \int\int \mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}'\, u(\mathbf{r},\mathbf{r}')\,\langle\rho(\mathbf{r}')\rangle\,\mathrm{e}^{i\mathbf{G}.\mathbf{r}} \\
&= \int\int \mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}'\,\frac{1}{\Omega^2}\sum_{\mathbf{G}'',\mathbf{G}'} u_{\mathbf{G}'',\mathbf{G}'}\mathrm{e}^{i(\mathbf{G}''-\mathbf{G}).\mathbf{r}}\,\mathrm{e}^{i\mathbf{G}'.\mathbf{r}'}\,\langle\rho(\mathbf{r}')\rangle \\
&= \frac{1}{\Omega}\sum_{\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\langle\rho_{\mathbf{G}'}\rangle \ .
\end{aligned} \tag{A.11}$$

The contribution of the long range Jastrow factor to the kinetic energy in the inhomogeneous system can therefore be expressed in terms of the density and the u

terms:

$$\begin{aligned}T &= T_s - \frac{1}{4}\left\langle \frac{1}{2\Omega^2}\sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\Big[(G^2+G'^2)\rho_{\mathbf{G}}\rho_{\mathbf{G}'} + 2\mathbf{G}.\mathbf{G}'\rho_{(\mathbf{G}+\mathbf{G}')} - 2G^2\rho_{\mathbf{G}}\langle\rho_{\mathbf{G}'}\rangle\Big]\right\rangle \\ &= T_s - \frac{1}{8\Omega^2}\left\langle \sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\Big[(G^2+G'^2)\big(\delta\rho_{\mathbf{G}}\delta\rho_{\mathbf{G}'} - \langle\rho_{\mathbf{G}}\rangle\langle\rho_{\mathbf{G}'}\rangle\big) + 2\mathbf{G}.\mathbf{G}'\rho_{(\mathbf{G}+\mathbf{G}')}\Big]\right\rangle \\ &= T_s - \frac{1}{8\Omega^2}\sum_{\mathbf{G},\mathbf{G}'} u_{\mathbf{G},\mathbf{G}'}\Big[(G^2+G'^2)\Big(\langle S(\mathbf{G},\mathbf{G}')\rangle - \langle\rho_{\mathbf{G}}\rangle\langle\rho_{\mathbf{G}'}\rangle\Big) + 2\mathbf{G}.\mathbf{G}'\langle\rho_{(\mathbf{G}+\mathbf{G}')}\rangle\Big]\end{aligned} \tag{A.12}$$

As an aside, we note at this point that in a homogeneous system, where $\rho(\mathbf{r}) = \text{const}$, $u(\mathbf{r},\mathbf{r}') = u(\mathbf{r}-\mathbf{r}')$ and $S(\mathbf{r},\mathbf{r}') = S(\mathbf{r}-\mathbf{r}')$, we have the following relations:

$$\begin{aligned}\langle\rho(\mathbf{G})\rangle &= \frac{N}{\Omega}\delta_{\mathbf{G},\mathbf{0}} \\ u(\mathbf{G},\mathbf{G}') &= \Omega\delta_{\mathbf{G},-\mathbf{G}'}\int \mathrm{d}\mathbf{r}\, u(\mathbf{r})\mathrm{e}^{-i\mathbf{G}.\mathbf{r}} = \Omega\delta_{\mathbf{G},-\mathbf{G}'}u(\mathbf{G}) \\ S(\mathbf{G},\mathbf{G}') &= \Omega\delta_{\mathbf{G},-\mathbf{G}'}\int \mathrm{d}\mathbf{r}\, S(\mathbf{r})\mathrm{e}^{-i\mathbf{G}.\mathbf{r}} = \Omega\delta_{\mathbf{G},-\mathbf{G}'}S(\mathbf{G})\end{aligned}$$

so we could recover the radially symmetric form of T as described in [33]:

$$T = T_s - \frac{1}{4\Omega}\sum_{\mathbf{G}\neq 0} NG^2 u_{\mathbf{G}}\big[S(\mathbf{G}) - 1\big] \tag{A.13}$$

As it is not clear in the original paper, we discuss briefly here the derivation of the homogeneous finite size error term itself here, to compare it with the derivation of the inhomogeous version.

The known behaviour of $S(\mathbf{k})$ suggests that the dominant contribution to the finite size error will come, in the limit of $\mathbf{k} \to 0$, from the “-1” term as $S(\mathbf{k}) \to 0$, so to find the kinetic energy finite size error we define

$$f(\mathbf{k}) = \frac{1}{4}k^2 u(\mathbf{k})\ .$$

The finite size error on T is defined as the difference between what the value in the limit of infinite system size, where the sum becomes an integral, and the value in the finite system where it remains a finite sum with the $\mathbf{G} = 0$ term excluded:

$$\Delta T_N = \frac{1}{(2\pi)^3} \int f(\mathbf{k}) \mathrm{d}\mathbf{k} - \frac{1}{\Omega} \sum_{\mathbf{G} \neq 0} f(\mathbf{G}) \; . \tag{A.14}$$

We can express the second term by using the Poisson sum formula $\sum_{\mathbf{L}} \widetilde{f}(\mathbf{L}) = \frac{1}{\Omega} \sum_{\mathbf{G}} f(\mathbf{G})$ and splitting off the $\mathbf{L} = 0$ and $\mathbf{G} = 0$ terms from the sums, so as to write the sum over $\mathbf{G} \neq 0$ as

$$\frac{1}{\Omega} \sum_{\mathbf{G} \neq 0} f(\mathbf{G}) = \sum_{\mathbf{L} \neq 0} \widetilde{f}(\mathbf{L}) + \widetilde{f}(0) - \frac{f(0)}{\Omega} \tag{A.15}$$

The integral term is by definition equal to $\widetilde{f}(0)$, so comparing the last two equations leaves

$$\Delta T_N = \frac{f(0)}{\Omega} - \sum_{\mathbf{L} \neq 0} \widetilde{f}(\mathbf{L}) \tag{A.16}$$

In the uniform electron gas, we know from the RPA that, using atomic units and thus omitting the physical constants, the u term of the Jastrow factor becomes

$$u(k) = \frac{4\pi}{\omega_p k^2} \tag{A.17}$$

in the small-k limit. The plasma frequency ω_p is given by $\omega_p^2 = 4\pi n$ in atomic units, so we have

$$f(k) = \frac{1}{4} k^2 u(k) = \frac{\omega_p}{4n} \tag{A.18}$$

For the first order correction we neglect the $\sum_{\mathbf{L} \neq 0}$ terms and the correction is simply

$$\frac{f(0)}{\Omega} = \frac{\omega_p}{4N} \tag{A.19}$$

and the total kinetic energy finite size correction is, as in Chiesa's paper:

$$\Delta T_N = \frac{\omega_p}{4N} \tag{A.20}$$

Application to Jellium Slab Bulk Plasmons

Now we switch back to the inhomogeneous case and return to attempting to evaluate Eq. A.12 for a specific form of u. Previous work by Wood [178] gives us the long ranged part of the u term in Jellium slabs, obtained from the plasmon normal modes. It takes the form

$$\begin{aligned} u_{bulk}(\mathbf{r},\mathbf{r}') =& \frac{4\pi}{\omega_p}\frac{1}{L^2 s}\sum_{\mathbf{G}}\frac{4}{G^2}\cos\mathbf{G}_{||}.(\mathbf{r}_{||}-\mathbf{r}'_{||})\sin G_z z\,\sin G_z z' \\ & \times\Theta(z)\Theta(s-z)\Theta(z')\Theta(s-z')\ , \end{aligned} \tag{A.21}$$

where $\Theta(z)$ is the step function and $\mathbf{G}=(\mathbf{G}_{||},G_z)$, with $\mathbf{G}_{||}=\frac{2\pi}{L}(n,m)$ for $n,m>=0$ and $G_z=\frac{p\pi}{s}$ for $p>0$.

This form depends only on the in-plane separation $\mathbf{r}_{||}-\mathbf{r}'_{||}$ not on the absolute in-plane positions $\mathbf{r}_{||}$ and $\mathbf{r}'_{||}$, but it does depend explicitly on z and z'. It can therefore be transformed as

$$u_{bulk}(\mathbf{G}_{||};k_z,k'_z)=\int_{L^2}\mathrm{d}^2\mathbf{r}_{||}\int\mathrm{d}z\int\mathrm{d}z'\,u(\mathbf{r}_{||};z,z')\mathrm{e}^{i\mathbf{G}_{||}.\mathbf{r}_{||}}\mathrm{e}^{ik_z z}\mathrm{e}^{ik'_z z'}\ . \tag{A.22}$$

The integrals are conveniently separable, leading to

$$u(\mathbf{G}_{||};k_z,k'_z)=\frac{4\pi}{\omega_p s}\sum_{G_z}\frac{4}{G_{||}^2+G_z^2}\widetilde{P}(G_z,k_z)\widetilde{P}(G_z,k'_z)\ , \tag{A.23}$$

where $P(G_z,z)=\sin G_z z\,\Theta(z)\Theta(s-z)$ and

$$\widetilde{P}(G_z,k_z)=\int_{-\infty}^{\infty}P(G_z,z)\,\mathrm{e}^{ik_z z}\mathrm{d}z\ , \tag{A.24}$$

which can be evaluated with the convolution theorem:

$$\begin{aligned} \mathscr{F}\big[\sin G_z z\big] &= \frac{1}{2i}\left[\delta(k_z-G_z)-\delta(k_z+Gz)\right] \\ \mathscr{F}\left[\Theta(z)\Theta(s-z)\right] &= \frac{1}{ik_z}\left[\mathrm{e}^{ik_z s}-1\right] \end{aligned} \tag{A.25}$$

giving

$$\mathscr{F}\left[\sin G_z z \Theta(z)\Theta(s-z)\right] = \int_{-\infty}^{\infty} \frac{1}{2i}\big[\delta(p_z - G_z) - \delta(p_z + Gz)\big]\frac{\mathrm{e}^{i(k_z-p_z)s}-1}{i(k_z-p_z)}\mathrm{d}p_z\ , \tag{A.26}$$

and thus

$$\widetilde{P}(k_z, G_z) = \frac{1}{i}\left[\mathrm{e}^{-i(k_z-G_z)s/2}\mathrm{sinc}((k_z-G_z)s) - \mathrm{e}^{-i(k_z+G_z)s/2}\mathrm{sinc}((k_z+G_z)s)\right] \tag{A.27}$$

We can therefore write the full Fourier transform of the Jastrow as

$$u(\mathbf{G}_{||}; k_z, k_z') = \frac{e^2}{\hbar\omega_p\epsilon_0}\frac{1}{s}\sum_{G_z}\frac{4}{G_{||}^2 + G_z^2}\widetilde{P}(k_z, G_z)\widetilde{P}(k_z', G_z) \tag{A.28}$$

To use Chiesa's method to find the kinetic energy finite size error ΔT_N we need to evaluate, according to Eq. A.12 and Eq. A.14,

$$\begin{aligned}
\Delta T_N = & -\frac{1}{8}\int\frac{\mathrm{d}^2\mathbf{k}_{||}}{(2\pi)^2}\int\frac{\mathrm{d}k_z}{2\pi}\int\frac{\mathrm{d}k_z'}{2\pi}\,u(\mathbf{k}_{||}, k_z, k_z') \\
& \times\Big[(2k_{||}^2 + k_z^2 + k_z'^2)\langle S(\mathbf{k}_{||}, k_z; -\mathbf{k}_{||}, k_z')\rangle + 2(-k_{||}^2 + k_z k_z')\langle\rho(\mathbf{0}, k_z + k_z')\rangle\Big] \\
& - -\frac{1}{8}\frac{1}{L^2}\sum_{\mathbf{G}_{||}\neq\mathbf{0}}\int\frac{\mathrm{d}k_z}{2\pi}\int\frac{\mathrm{d}k_z'}{2\pi}\,u(\mathbf{G}_{||}, k_z, k_z') \\
& \times\Big[(2G_{||}^2 + k_z^2 + k_z'^2)\langle S(\mathbf{G}_{||}, k_z; -\mathbf{G}_{||}, k_z')\rangle + 2(-G_{||}^2 + k_z k_z')\langle\rho(\mathbf{0}, k_z + k_z')\rangle\Big] \\
\simeq & -\frac{1}{8}\frac{1}{L^2}\int\frac{\mathrm{d}k_z}{2\pi}\int\frac{\mathrm{d}k_z'}{2\pi}\,u(\mathbf{0}, k_z, k_z')\Big[(k_z^2 + k_z'^2)\langle S(0, k_z; 0, k_z')\rangle \\
& + 2k_z k_z'\langle\rho(\mathbf{0}, k_z + k_z')\rangle\Big]\ .
\end{aligned} \tag{A.29}$$

By inserting the Fourier transform of u, we get

$$\Delta T_N = -\frac{1}{8}\left\langle \frac{1}{L^2}\int\frac{\mathrm{d}k_z}{2\pi}\int\frac{\mathrm{d}k'_z}{2\pi}\frac{4\pi}{\omega_p s}\sum_{G_z}\frac{4}{G_z^2}P(k_z,G_z)P(k'_z,G_z)\Big[\right. \tag{A.30}$$

$$\left.(k_z^2+k_z'^2)\rho(0,k_z)\rho(0,k'_z)+2k_zk'_z\rho(\mathbf{0},k_z+k'_z)\Big]\right\rangle \tag{A.31}$$

which, on reinserting the densities as $\rho(0,k_z)=\sum_i \mathrm{e}^{ik_zz_i}$ becomes:

$$\begin{aligned}\Delta T_N = \frac{-1}{2\omega_p}\Bigg\langle \frac{1}{sL^2}\sum_{G_z}\frac{1}{G_z^2}\Bigg(&\sum_i\int\frac{\mathrm{d}k_z}{2\pi}k_z^2\widetilde{P}(k_z,G_z)\mathrm{e}^{ik_zz_i}\sum_j\int\frac{\mathrm{d}k'_z}{2\pi}\widetilde{P}(k'_z,G_z)\mathrm{e}^{ik'_zz_j}\\ &+\sum_j\int\frac{\mathrm{d}k'_z}{2\pi}k_z'^2\widetilde{P}(k'_z,G_z)\mathrm{e}^{ik'_zz_j}\sum_i\int\frac{\mathrm{d}k_z}{2\pi}\widetilde{P}(k_z,G_z)\mathrm{e}^{ik_zz_i}\\ &+2\sum_i\int\frac{\mathrm{d}k'_z}{2\pi}k'_z\widetilde{P}(k'_z,G_z)\mathrm{e}^{ik'_zz_i}\int\frac{\mathrm{d}k_z}{2\pi}k_z\widetilde{P}(k_z,G_z)\mathrm{e}^{ik_zz_i}\Bigg)\Bigg\rangle\end{aligned} \tag{A.32}$$

Transforming the $\widetilde{P}$'s back again and using $\mathscr{F}[k\widetilde{f}(k)]=i\frac{\partial f}{\partial x}$ gives

$$\begin{aligned}\Delta T_N = \frac{\omega_p}{2nL^2s}\Bigg\langle\sum_{G_z}\frac{1}{G_z^2}\Big[2&\left(\sum_i(G_z^2)\sin G_zz_i\Theta(z_i)\Theta(s-z_i)\right)\\ &\times\left(\sum_j\sin G_zz_j\Theta(z_j)\Theta(s-z_j)\right)\\ &+2\sum_i\left(iG_z\cos G_zz_i\right)\left(iG_z\cos G_zz_i\right)\Theta(z_i)\Theta(s-z_i)\Big]\quad\Bigg\rangle\end{aligned} \tag{A.33}$$

which can be written as

$$\begin{aligned}\Delta T_N =&\frac{\omega_p}{2N}\left\langle\sum_{G_z}\left(\sum_i 2(\sin^2 G_zz_i-\cos^2 G_zz_i)+2\sum_{i\neq j}\sin G_zz_i\sin G_zz_j\right)\right\rangle\\ =&\frac{\omega_p}{2N}\left\langle\sum_{G_z}\sum_i\left(4\cos 2G_zz_i+2\sum_{j\neq i}\sin G_zz_i\sin G_zz_j\right)\right\rangle\end{aligned} \tag{A.34}$$

The pair density $\rho(\mathbf{r},\mathbf{r}')$ is defined by

$$\rho(\mathbf{r},\mathbf{r}') = \Big\langle \sum_{i,j}{}' \delta(\mathbf{r}-\mathbf{r}_i)\delta(\mathbf{r}'-\mathbf{r}_j) \Big\rangle \tag{A.35}$$

and its transform in the same system as used previously is

$$\rho(\mathbf{k},\mathbf{k}') = \int\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \Big\langle \sum_{i,j}{}' \delta(\mathbf{r}-\mathbf{r}_i)\delta(\mathbf{r}'-\mathbf{r}_j) \mathrm{e}^{-i\mathbf{k}.\mathbf{r}}\mathrm{e}^{-i\mathbf{k}'.\mathbf{r}'} \Big\rangle = \Big\langle \sum_{i,j}{}' \mathrm{e}^{-i\mathbf{k}.\mathbf{r}_i}\mathrm{e}^{-i\mathbf{k}'.\mathbf{r}_j} \Big\rangle \tag{A.36}$$

so we can write Δ_N as

$$\begin{aligned} \Delta T_N &= \frac{\hbar\omega_p}{2N} \Big\langle \sum_{G_z} 4\left[\rho(2G_z)+\rho(-2G_z)\right] && \text{(A.37)} \\ &\quad +2\left[\rho(G_z,G_z)-\rho(G_z,-G_z)-\rho(-G_z,G_z)+\rho(-G_z,-G_z)\right] \Big\rangle && \text{(A.38)} \end{aligned}$$

This gives the scaling with N observed in Chapter 5: $\Delta T_N \propto 1/N$. While the remaining terms involving transforms of the pair density could, in principle, be evaluated in QMC, this would be a challenging calculation and is unnecessary in the context examined here.

Appendix B

Correcting Periodic Boundary Conditions in Defect Cells

In an ideal world, we would be able, when estimating formation energies of charged defects, to calculate of the correct energy of a single supercell of the defect system embedded in an infinite lattice of copies of the corresponding supercell of perfect crystal, with all the interactions between electrons and ions taking the correct form, whether they result from localized charges or periodic ones. This aim is made difficult by the fact that this system contains ions and contributions to the charge density that are repeated periodically, and ions and charge density that are localized at the defect site. A localized 'ion' may be a real one (ie an interstitial), or an oppositely-charged version of one present in the perfect crystal calculation, there to represent a vacancy, and the overall change in charge density may be negative or positive depending on the charge on the defect. In this Appendix, we attempt to formulate the energy functional for the defect cell in such a way that the correct boundary conditions are applied to each component of each contribution to the energy, so as to arrive at the correct energy. We assume that it is possible to carry out accurate calculations of the density $n_{\text{perf}}(\mathbf{r})$ and Hartree potential $V_{\text{H}}^{\text{perf}}(\mathbf{r})$ associated with the perfect crystal, and store this for use in the defect calculation.

We write the total energy of the defect cell in terms of the density $n(\mathbf{r})$ as

$$E_{T(def)}[n] = T_{\mathrm{s}}[n] + E_{\mathrm{xc}}[n] + E_{\mathrm{H}}[n] + E_{\mathrm{psp}}[n] + E_{\mathrm{II}} \tag{B.1}$$

where all the terms have their standard meaning. We assume that in a calculation of a periodically repeated defect, we can obtained the correct density, but that because we were using the wrong interaction in parts of the functional, we obtain the wrong energy. To get the correct energy from this density, we can take the output density from the periodic defect calculation, and combine it with the perfect crystal density to get the correct form of the functional. Changing the boundary conditions of the interaction used for parts of the density and some of the atoms will change the Kohn-Sham effective potential in the cell and thus the single particle eigenvectors, so all the contributions except the exchange-correlation energy will change to some degree. We address each contribution term by term.

We write the electron density in the defect cell as $n_{\mathrm{def}}(\mathbf{r}) = n_{\mathrm{per}}(\mathbf{r}) + n_{\mathrm{loc}}(\mathbf{r})$, where $n_{\mathrm{loc}}(\mathbf{r})$ is the difference in charge density between the perfect supercell and the defect supercell. The defect cell is imagined as containing periodic versions of all the same ions in the perfect cell, and then added ones or subtracted ones, where a subtracted ion which was originally of pseudo-charge q_α and pseudopotential $V_{\mathrm{loc},\alpha}(r)$ is represented by an ion at the same point with charge$-q_\alpha$ and pseudopotential $-V_{\mathrm{loc},\alpha}(r)$ to cancel it out.

The periodic charge density, which can be expressed in Fourier components $n(\mathbf{G})$, produces a potential $V^{\mathrm{H}}_{\mathrm{per}}(\mathbf{r})$ given by

$$V^{\mathrm{H}}_{\mathrm{per}}(\mathbf{r}) = \sum_{\mathbf{G}\neq 0} \frac{4\pi n(\mathbf{G})}{\Omega G^2} \mathrm{e}^{i\mathbf{G}.\mathbf{r}} \ , \tag{B.2}$$

while the localized charge density produces a Hartree potential $V^{\mathrm{H}}_{\mathrm{loc}}(\mathbf{r})$ given by

$$V^{\mathrm{H}}_{\mathrm{loc}}(\mathbf{r}) = \int_{a.s.} \frac{n(\mathbf{r}')\mathrm{d}^3\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} \ . \tag{B.3}$$

This potential extends over all space, but since in practice the density change $n_{\mathrm{loc}}(\mathbf{r})$ is confined to one supercell, the integral need only be over that cell.

The full Hartree Energy of one defect cell is given by the sum of the interactions between all the contributions to $V^{\mathrm{H}}(\mathbf{r})$ and $n(\mathbf{r})$:

$$\begin{aligned} E_{\mathrm{H}}[n] =& \frac{1}{2}\int_{\mathrm{cell}} V^{\mathrm{H}}_{\mathrm{per}}(\mathbf{r}) n_{\mathrm{per}}(\mathbf{r}) \mathrm{d}^3\mathbf{r} + \frac{1}{2}\int_{\mathrm{cell}} V^{\mathrm{H}}_{\mathrm{loc}}(\mathbf{r}) n_{\mathrm{loc}}(\mathbf{r}) \mathrm{d}^3\mathbf{r} + \int_{\mathrm{cell}} V^{\mathrm{H}}_{\mathrm{per}}(\mathbf{r}) n_{\mathrm{loc}}(\mathbf{r}) \mathrm{d}^3\mathbf{r} \\ =& \frac{1}{2}\sum_{\mathbf{G}\neq 0} \frac{4\pi n_{\mathrm{per}}(\mathbf{G}) n_{\mathrm{per}}(-\mathbf{G})}{\Omega G^2} + \frac{1}{2}\int_{\mathrm{cell}}\int_{\mathrm{cell}} \frac{n_{\mathrm{loc}}(\mathbf{r}') n_{\mathrm{loc}}(\mathbf{r}) \mathrm{d}^3\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r} \\ &+ \int_{\mathrm{cell}} \sum_{\mathbf{G}\neq 0} \frac{4\pi n_{\mathrm{per}}(\mathbf{G}) \mathrm{e}^{i\mathbf{G}.\mathbf{r}}}{\Omega G^2} n_{\mathrm{loc}}(\mathbf{r}) \; , \end{aligned} \tag{B.4}$$

The first term has already been calculated in the perfect crystal calculation and is standard. For the second term, the localization of the charge density means a cutoff Coulomb interaction can be used, with a radius R greater than the distance between periodic images, and by padding a larger version of the density grid with zeros the integral can still be evaluated with Fourier transform methods. The final term has been written as the interaction of the localized charge with the periodic potential rather than vice versa in order that it can be evaluated as an integral within one unit cell rather than all space. Then, because the potential from the perfect crystal only exists on the **G**-vectors of the original cell, it does not matter that the Fourier transform of $n_{\mathrm{loc}}(\mathbf{r})$ exists for all values of **k**, so this contribution can easily be evaluated with the standard Fourier methods of calculating Hartree energies in periodic systems.

Comparing this correct energy with what we would actually have evaluated in a calculation of a periodically repeated defect, gives us a correction to the Hartree Energy of the defect cell of

$$\Delta E_{\mathrm{H}}[n] = \frac{1}{2}\int_{\mathrm{cell}}\int_{\mathrm{cell}} \frac{n_{\mathrm{loc}}(\mathbf{r}') n_{\mathrm{loc}}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3\mathbf{r}' \mathrm{d}^3\mathbf{r} - \frac{1}{2}\sum_{\mathbf{G}\neq 0} \frac{4\pi n_{\mathrm{loc}}(\mathbf{G}) n_{\mathrm{loc}}(-\mathbf{G})}{\Omega G^2} \tag{B.5}$$

Alternatively, we could use the corresponding formulae for the Hartree potential to correct the effective potential inside the SCF loop of a calculation.

We now address the contribution to the energy from the intereaction of the pseudopotential and the density. A periodically repeated ion of species α at position $\mathbf{R}_\alpha$ is represented by a pseudopotential with valence charge Z_α, the local part of which is divided into Coulombic (Z_α/r) and non-Coulombic parts. The non-Coulombic part is stored internally in a program such as CASTEP as a spherical Fourier series $V_{\mathrm{nc},\alpha}(q)$. The full pseudopotential is reconstructed using the formula:

$$V_{\mathrm{psp},\alpha}(q) = 4\pi \left(V_{\mathrm{nc},\alpha}(q) - \frac{Z_\alpha}{q^2} \right) \ , \tag{B.6}$$

where the 4π has been taken outside the brackets merely as a convention. The local part of the pseudopotential of this ion is then interpolated onto the grid of G-vectors of the unit cell, omitting the $\mathbf{G} = 0$ term of the Coulombic part and treating it separately, giving:

$$V^{\mathrm{psp}}_{\mathrm{per},\alpha}(\mathbf{r}) = 4\pi \left(\sum_{\mathbf{G}} V_{\mathrm{nc},\alpha}(G) \mathrm{e}^{i\mathbf{G}.(\mathbf{r}-\mathbf{r}_\alpha)} - \sum_{\mathbf{G}\neq 0} \frac{Z_\alpha}{G^2} \mathrm{e}^{i\mathbf{G}.(\mathbf{r}-\mathbf{r}_\alpha)} \right) . \tag{B.7}$$

So including the structure factor for all the ions I of each species, the total local pseudopotential is

$$V^{\mathrm{psp}}_{\mathrm{per}}(\mathbf{r}) = \sum_{\alpha} 4\pi \left(\sum_{\mathbf{G}} V_{\mathrm{nc},\alpha}(G) \sum_{I} \mathrm{e}^{-i\mathbf{G}.\mathbf{R}_{I,\alpha}} \mathrm{e}^{i\mathbf{G}.\mathbf{r}} - \sum_{\mathbf{G}\neq 0} \frac{Z_\alpha}{G^2} \sum_{I} \mathrm{e}^{-i\mathbf{G}.\mathbf{R}_{I,\alpha}} \mathrm{e}^{i\mathbf{G}.\mathbf{r}} \right) , \tag{B.8}$$

and the electron-ion interaction energy is

$$\begin{aligned} E^{\mathrm{psp}}_{\mathrm{per}}(\mathbf{r}) &= \int_{\mathrm{cell}} \sum_{\alpha} \sum_{\mathbf{G}\neq 0} 4\pi \left(V_{\mathrm{nc},\alpha}(G) - \frac{Z_\alpha}{G^2} \right) \left(\sum_{I} \mathrm{e}^{-i\mathbf{G}.\mathbf{R}_{I,\alpha}} \right) \mathrm{e}^{i\mathbf{G}.\mathbf{r}} \, n_{\mathrm{per}}(\mathbf{r}) \, \mathrm{d}^3\mathbf{r} \\ &\quad + \sum_{\alpha} V_{\mathrm{nc},\alpha}(0) \, \frac{N_e}{\Omega} \ . \end{aligned} \tag{B.9}$$

If we want to calculate the contribution of the localized ions, we need equations in real space, we would need $V_{\mathrm{psp},\alpha}(r)$, and we must reconstruct this from the known $V_{\mathrm{psp},\alpha}(q)$:

$$\begin{aligned} V_{\mathrm{psp},\alpha}(r) =& \frac{1}{(2\pi)^3}\int 4\pi\left(V_{\mathrm{nc},\alpha}(q)-\frac{Z_\alpha}{q^2}\right)\mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}}\,\mathrm{d}^3\mathbf{q} \\ =& \frac{4\pi}{(2\pi)^3}\int_0^\infty\left(V_{\mathrm{nc},\alpha}(q)-\frac{Z_\alpha}{q^2}\right)4\pi q^2\frac{\sin(qr)}{qr}\mathrm{d}q \end{aligned} \; . \tag{B.10}$$

We then split this into a short range part, from $q = q_{max}$to $q = \infty$, and a long range part from $q = 0$ to $q = q_{max}$:

$$\begin{aligned} V_{\mathrm{psp},\alpha}(r) \;=\; & \frac{(4\pi)^2}{(2\pi)^3}\Bigg[\int_0^{q_{max}} V_{\mathrm{nc},\alpha}(q)\frac{q\sin(qr)}{r}\mathrm{d}q - Z_\alpha\int_0^{q_{max}}\frac{\sin(qr)}{qr}\mathrm{d}q \\ & + \int_{q_{max}}^\infty\left(V_{\mathrm{nc},\alpha}(q)-\frac{Z_\alpha}{q^2}\right)\frac{q\sin(qr)}{r}\mathrm{d}q\Bigg] . \end{aligned}$$

The final term is zero as the pseudopotential has been constructed so that it is smooth around $r = 0$, so the non-Coulombic part must exactly cancel the divergence of Z_α/r at small r and thus large q, so the two parts cancel each other exactly. The second term is a Sine Integral $Si(q_{max})$ which is easily calculated with standard algorithms, and the first term can be evaluated numerically, so $V_{\mathrm{psp},\alpha}$ can be accurately reconstructed. We can then interpolate this on to the real space grid to get the localized version of the total ionic pseudopotential:

$$V_{\mathrm{loc}}^{\mathrm{psp}}(\mathbf{r}) = \sum_\alpha\sum_I V_{\mathrm{psp},\alpha}(|\mathbf{r}-\mathbf{r}_I|) \; , \tag{B.11}$$

and the energy is

$$E_{\mathrm{loc}}^{\mathrm{psp}} = \int_{\mathrm{cell}}\sum_\alpha\sum_I V_{\mathrm{psp},\alpha}(|\mathbf{r}-\mathbf{r}_I|)n_{\mathrm{loc}}(\mathbf{r})\mathrm{d}^3\mathbf{r} \; . \tag{B.12}$$

The boundary-condition correction to E_{psp} is thus obtained by considering the difference between Eq. B.9 and the above:

$$\Delta E_{\mathrm{psp}} = \int_{\mathrm{cell}} V^{\mathrm{psp}}_{\mathrm{loc}}(\mathbf{r}) n_{\mathrm{loc}}(\mathbf{r}) \, \mathrm{d}^3\mathbf{r} - \int_{\mathrm{cell}} V^{\mathrm{psp}}_{\mathrm{per}}(\mathbf{r}) n_{\mathrm{loc}}(\mathbf{r}) \, \mathrm{d}^3\mathbf{r} \; , \tag{B.13}$$

where we only consider $n_{\mathrm{loc}}(\mathbf{r})$ since the contribution from the periodic part of the density was already calculated correctly in the original calculation by the same argument as for the Hartree energy.

Finally we come to the ion-ion interaction. Within periodic boundary conditions, the interaction between ions (or pseudo-ions) of species α, β with charges Z_α, Z_β is given by the Ewald interaction

$$E^{\mathrm{II}}_{\mathrm{per}} = \frac{1}{2} \sum_{\alpha,\beta} \sum_{I, J \neq I} Z_{\alpha I} Z_{\beta J} (v_{EW}(\mathbf{R}_\alpha - \mathbf{R}_\beta) - v_M) \; , \tag{B.14}$$

while for ‘localized’ ions it is simply the Coulomb form:

$$E^{\mathrm{II}}_{\mathrm{loc}} = \sum_{\alpha,\beta} \sum_{I, J \neq I} Z_{\alpha I} Z_{\beta J} \frac{1}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \; . \tag{B.15}$$

Hence to correct the interactions we simply find the difference:

$$\Delta E_{\mathrm{II}} = E^{\mathrm{II}}_{\mathrm{loc}} - E^{\mathrm{II}}_{\mathrm{per}} \tag{B.16}$$

Appendix C

Embedding of Ewald Calculations

When we simulate a ‘charged’ periodic system in quantum Monte Carlo, we are necessarily including a neutralizing uniform background in our calculation, or the total Ewald energy would be infinite. The requirement of strict neutrality over a unit cell of volume Ω demands that each and every charge q be accompanied by a background charge density $-q/\Omega$ over the whole cell. If, on the other hand, what we really mean to simulate is a single charged cell embedded in an infinite periodic lattice of neutral copies of an uncharged version, then the cancelling background of the extra charge can legitimately be imagined as extending over a much larger, or potentially infinite volume, which will produce a significantly different energy. The error of confining the background to a single simulation cell can be thought of as a finite size error

Here we shall imagine a very large cell, which we shall call the supercell, composed of $M \times M \times M$ copies of the simulation cell. Each cell contains N charges q_i at positions $\mathbf{r}_i$. In one of these cells, we also either place an extra charge q_P at position $\mathbf{r}_P$, which may represent an extra electron, or we take away one of the original N charges q_G from position $\mathbf{r}_G$. We will treat these two situations separately, as contrary to what one might at first imagine, the situation of removing a charge is not equivalent to adding an extra charge of $-q$ on top of one of the original charge. Apart from anything else, that would lead to an infinite energy.

We want to know the error we are making by using the Ewald method on the ‘charged’ cell, so we compare two ways of evaluating the energy of an embedded version of the defect cell: a) by treating the whole supercell as periodic and evaluating the Ewald energy for the whole system; and b) by treating the ‘charged’ simulation cell as periodic, and then treating the neutral cell as periodic, and adding the energy of one of the charged simulation cells to $(M^3 - 1)$ of the neutral simulation cells.

We label the real and reciprocal lattice vectors and volume of the (smaller) simulation cell $\mathbf{R}_1$, $\mathbf{G}_1$, Ω_1 respectively, while those of the (larger) supercell are $\mathbf{R}_2$, $\mathbf{G}_2$, Ω_2. These are related in that $\Omega_2 = M^3\Omega_1$, that the set of vectors $\{\mathbf{R}_1\}$ contains $\{\mathbf{R}_2\}$ as a subset, and that the set of vectors $\{\mathbf{G}_2\}$ contains $\{\mathbf{G}_1\}$ as a subset. The copies of the simulation cell contained within the supercell are at positions labelled by $\mathbf{R}_s$.

Each charge q on the ‘1’ lattice has a cancelling background of $-q/\Omega_1$, so the charge distribution is

$$\rho_1(\mathbf{r}) = \frac{-q}{\Omega_1} + \sum_{\mathbf{R}_1} q\delta(\mathbf{r} - \mathbf{R}_1) \tag{C.1}$$

over all space. This can be inserted into Poisson’s Equation to give an expression involving the Ewald potential

$$\nabla^2 v_{\mathrm{Ew}_1}(\mathbf{r}) = -4\pi\rho_1(\mathbf{r}) \; , \tag{C.2}$$

which can be Fourier transformed to give

$$\tilde{v}_{\mathrm{Ew}_1}(\mathbf{G}_1) = \begin{cases} 0 & \mathbf{G}_1 = 0 \\ \dfrac{4\pi}{\Omega_1 G_1^2} & \mathbf{G}_1 \neq 0 \end{cases} \; . \tag{C.3}$$

In practice we use the Ewald method to split the badly behaved real space form of $v_{\mathrm{Ew}_1}(\mathbf{r})$ into a part that converges rapidly in real space and a part that converges rapidly in reciprocal space, giving the commonly-used formula:

$$v_{\mathrm{Ew}}(\mathbf{r}) = \frac{4\pi}{\Omega}\sum_{\mathbf{G}\neq 0}\frac{\exp\left(\frac{-G^2}{4\kappa^2}+i\mathbf{G}.\mathbf{r}\right)}{G^2} - \frac{\pi}{\kappa^2\Omega} + \sum_{\mathbf{R}}\frac{\mathrm{erfc}(\kappa|\mathbf{r}+\mathbf{R}|)}{|\mathbf{r}+\mathbf{R}|} \tag{C.4}$$

for the potential of a lattice of point charges and their cancelling backgrounds, assuming one charge is at the origin.

We will be distinguishing between the potential produced by particles repeated in each copy of the simulation cell, $v_{\mathrm{Ew}_1}(\mathbf{r})$, and the potential produced by particles repeated only in each copy of the whole simulation cell, $v_{\mathrm{Ew}_2}(\mathbf{r})$. It is intuitive but important to note that the potential felt on the simulation cell lattice is the same as the potential felt on the supercell lattice but copied to each of the individual copies of the simulation cell lattice. To show this, we write

$$\begin{aligned}\sum_{\mathbf{R}_s} v_{\mathrm{Ew}_2}(\mathbf{r}+\mathbf{R}_s) =& \frac{4\pi}{M^3\Omega_1}\sum_{\mathbf{G}_2\neq 0}\frac{\exp\left(\frac{-G_2^2}{4\kappa^2}+i\mathbf{G}_2.\mathbf{r}\right)}{G_2^2}\sum_{\mathbf{R}_s}\exp(i\mathbf{G}_2.\mathbf{R}_s) - \sum_{\mathbf{R}_s}\frac{\pi}{\kappa^2 M^3\Omega_1} \\ &+ \sum_{\mathbf{R}_s}\sum_{\mathbf{R}_2}\frac{\mathrm{erfc}(\kappa|\mathbf{r}+\mathbf{R}_2+\mathbf{R}_s|)}{|\mathbf{r}+\mathbf{R}_2+\mathbf{R}_s|}\ .\end{aligned} \tag{C.5}$$

The vectors in $\{\mathbf{R}_2\}+\{\mathbf{R}_s\}$ span all the vectors in $\{\mathbf{R}_1\}$, so the final term can be replaced with a sum over $\{\mathbf{R}_1\}$, reducing it to the equivalent term of $v_{\mathrm{Ew}_1}(\mathbf{r})$. There are M^3 vectors in $\{\mathbf{R}_s\}$, so the middle term is just $\frac{\pi}{\kappa^2\Omega_1}$. Finally, for each $\mathbf{G}_2$ which is not a member of $\{\mathbf{G}_1\}$, the sum over $\mathbf{R}_s$ gives zero, whereas if it is a member of $\{\mathbf{G}_1\}$, each $\mathrm{e}^{i\mathbf{G}_2.\mathbf{R}_s}$ term is 1 and the sum is M^3, so the first term can be rewritten in terms of a sum over $\mathbf{G}_1$, showing that

$$\sum_{\mathbf{R}_s} v_{\mathrm{Ew}_2}(\mathbf{r}+\mathbf{R}_s) = v_{\mathrm{Ew}_1}(\mathbf{r}) \tag{C.6}$$

The N particles in the uncharged simulation cell create a total potential of

$$\phi_{\mathrm{Ew}_1}(\mathbf{r}) = \sum_{i}^{N} q_i v_{\mathrm{Ew}_1}(\mathbf{r} - \mathbf{r}_i) \ , \tag{C.7}$$

which is periodic and already includes the potential from the full set of M^3N particles in the simulation cell as the periodic replicas of the M particles in the simulation cell.

If we now consider adding an extra charge, at position $\mathbf{r}_P$, to one (and only one) of the simulation cells in the large supercell, the extra electron, we add an additional potential of

$$\phi_{\mathrm{Ew}_2}(\mathbf{r}) = q_P v_{\mathrm{Ew}_2}(\mathbf{r} - \mathbf{r}_P) \ . \tag{C.8}$$

Each particle in the simulation cell feels the potential due to everything except itself, so the potential $\bar{\phi}_i(\mathbf{r}_i)$ felt by particle i can be written

$$\begin{aligned}\bar{\phi}_i(\mathbf{r}_i) &= \lim_{\mathbf{r}\to\mathbf{r}_i} \left(\phi_{\mathrm{Ew}_1}(\mathbf{r}) + \phi_{\mathrm{Ew}_2}(\mathbf{r}) - \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|} \right) \\ &= \lim_{\mathbf{r}\to\mathbf{r}_i} \left(\sum_{j}^{N} v_{\mathrm{Ew}_1}(\mathbf{r} - \mathbf{r}_j) + q_P v_{\mathrm{Ew}_2}(\mathbf{r} - \mathbf{r}_P) - \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|} \right) \ .\end{aligned} \tag{C.9}$$

For all particles i except the extra one P, we have

$$\bar{\phi}_i(\mathbf{r}_i) = q_i v_{M_1} + \sum_{j \neq i}^{N} q_j v_{\mathrm{Ew}_1}(\mathbf{r}_i - \mathbf{r}_j) + q_P v_{\mathrm{Ew}_2}(\mathbf{r}_i - \mathbf{r}_P) \ , \tag{C.10}$$

while for particle P itself we have

$$\bar{\phi}_P(\mathbf{r}_P) = q_P v_{M_2} + \sum_{j}^{N} q_j v_{\mathrm{Ew}_1}(\mathbf{r}_P - \mathbf{r}_j) \ . \tag{C.11}$$

In both cases, the Madelung potential

$$\begin{aligned} v_M =& \lim_{\mathbf{r}\to 0}\left(v_{\mathrm{Ew}}(\mathbf{r}) - \frac{1}{r}\right) \\ =& \frac{4\pi}{\Omega}\sum_{\mathbf{G}\neq 0}\frac{\exp\left(\frac{-G^2}{4\kappa^2}\right)}{G^2} - \frac{\pi}{\kappa^2\Omega} + \sum_{\mathbf{R}\neq 0}\frac{\mathrm{erfc}(\kappa|\mathbf{R}|)}{|\mathbf{R}|} - \frac{2\kappa}{\sqrt{\pi}}\ , \end{aligned} \tag{C.12}$$

is the potential felt by the particle due to its cancelling background and all the periodic replicas of itself. The value of v_M is dependent on cell size and geometry but is usually negative as the particle is always nearer to a comparable amount of its cancelling background, which is of the opposite sign, than it is to its nearest neighbour replicas. In highly elongated cells this is no longer the case.

The total Ewald interaction energy in the supercell is the sum of the energies of all the M^3 copies of the original arrangement of charges, plus the energy of the extra one. We assume here that the N electrons are in the same positions each cell, i.e. at $\mathbf{r}_i + \mathbf{R}_s$. In reality they would be in different places in each cell, both due to the fact that they are separate particles in the first place and that they will move to screen the charge of the charged cell. However, in a QMC simulation, this is the approximation we are forced to make, and for a large cell it will not be a large error. In this approximation we get:

$$\begin{aligned} U_{\mathrm{Ew}} =& \frac{1}{2}\sum_{\mathbf{R}_s}\sum_{i}^{N} q_i\bar{\phi}_i(\mathbf{r}_i + \mathbf{R}_s) + \frac{1}{2}q_P\bar{\phi}_P(\mathbf{r}_P) \\ =& \frac{1}{2}M^3\sum_{i}^{N} q_i^2 v_{M_1} + \frac{1}{2}\sum_{\mathbf{R}_s}\sum_{i}^{N}\left(\sum_{j\neq i}^{N} q_i q_j v_{\mathrm{Ew}_1}(\mathbf{r}_i + \mathbf{R}_s - \mathbf{r}_j)\right. \\ &\left. + q_i q_P v_{\mathrm{Ew}_2}(\mathbf{r}_i + \mathbf{R}_s - \mathbf{r}_j)\right) + \frac{1}{2}q_P^2 v_{M_2} + \frac{1}{2}\sum_{j}^{N} q_P q_j v_{\mathrm{Ew}_1}(\mathbf{r}_P - \mathbf{r}_j) \end{aligned} \tag{C.13}$$

Using Eq. C.6 and remembering that $v_{\mathrm{Ew}_1}(\mathbf{r}+\mathbf{R}_s) = v_{\mathrm{Ew}_1}(\mathbf{r})$ we can write this as

$$U_{\mathrm{Ew}}^{\mathrm{sup}} = \frac{M^3}{2}\sum_i^N q_i^2 v_{M_1} + \frac{1}{2}q_P^2 v_{M_2} + \frac{M^3}{2}\sum_i^N \sum_{j\neq i}^N q_i q_j v_{\mathrm{Ew}_1}(\mathbf{r}_i - \mathbf{r}_j) + \sum_i^N q_i q_P v_{\mathrm{Ew}_1}(\mathbf{r}_i - \mathbf{r}_P) \tag{C.14}$$

If we had used the Ewald interaction on the simulation cell to calculate the energy, effectively assuming that q_P is repeated in every cell, we would have had, for one cell:

$$U_{\mathrm{Ew}}^{\mathrm{sim}} = \frac{1}{2}\sum_i^N q_i^2 v_{M_1} + \frac{1}{2}\sum_i^N \sum_{j\neq i}^N q_i q_j v_{\mathrm{Ew}_1}(\mathbf{r}_i - \mathbf{r}_j) + \frac{1}{2}q_P^2 v_{M_1} + \sum_i^N q_i q_P v_{\mathrm{Ew}_1}(\mathbf{r}_i - \mathbf{r}_P) \tag{C.15}$$

whereas in the neutral cell we would simply have had

$$U_{\mathrm{Ew}}^{\mathrm{neut}} = \frac{1}{2}\sum_i^N q_i^2 v_{M_1} + \frac{1}{2}\sum_i^N \sum_{j\neq i}^N q_i q_j v_{\mathrm{Ew}_1}(\mathbf{r}_i - \mathbf{r}_j) \tag{C.16}$$

Hence, the error we would have made making by neutralizing q_P over Ω_1 rather than Ω_2 is

$$\Delta U_{\mathrm{Ew}} = U_{\mathrm{Ew}}^{\mathrm{sup}} - \left(U_{\mathrm{Ew}}^{\mathrm{sim}} + (M^3-1)U_{\mathrm{Ew}}^{\mathrm{neut}}\right) \tag{C.17}$$

which, after combining Eqs. C.14, C.15, C.16, almost entirely cancels to give

$$\Delta U_{\mathrm{Ew}} = \frac{1}{2}q_P^2\left(v_{M_2} - v_{M_1}\right) \tag{C.18}$$

If we let the supercell size tend to infinity, v_{M_2} vanishes and the error is simply one Madelung energy of the small cell:

$$\Delta U_{\mathrm{Ew}} = -\frac{1}{2}q_P^2 v_{M_1} \tag{C.19}$$

The derivation in the case where we have removed a charge proceeds very similarly except for the sign change when the q_P terms are replaced with q_G ones (because they now represent interactions which are no longer there relative to the neutral state). If $\bar{\phi}_G$ represents the potential missing from the supercell because of the removed charge, then the full Ewald energy of the charged cell is

$$\begin{aligned}
U_{\text{Ew}}^{\text{sup}} =& \frac{1}{2}\sum_{\mathbf{R}_s}\sum_{i}^{N} q_i\bar{\phi}_i(\mathbf{r}_i+\mathbf{R}_s) - \frac{1}{2}q_G\bar{\phi}_G(\mathbf{r}_G) \\
=& \frac{1}{2}M^3\sum_{i}^{N} q_i^2 v_{M_1} + \frac{1}{2}\sum_{\mathbf{R}_s}\sum_{i}^{N}\left(\sum_{j\neq i}^{N} q_iq_jv_{\text{Ew}_1}(\mathbf{r}_i+\mathbf{R}_s-\mathbf{r}_j)\right. \\
& \left. - q_iq_Gv_{\text{Ew}_2}(\mathbf{r}_i+\mathbf{R}_s-\mathbf{r}_j)\right) - \frac{1}{2}q_G^2v_{M_2} - \frac{1}{2}\sum_{j}^{N} q_Gq_jv_{\text{Ew}_1}(\mathbf{r}_G-\mathbf{r}_j) \qquad \text{(C.20)} \\
=& \frac{1}{2}M^3\sum_{i}^{N} q_i^2 v_{M_1} - \frac{1}{2}q_G^2v_{M_2} + \frac{1}{2}M^3\sum_{i}^{N}\sum_{j\neq i}^{N} q_iq_jv_{\text{Ew}_1}(\mathbf{r}_i-\mathbf{r}_j) \\
& - \sum_{i}^{N} q_iq_Gv_{\text{Ew}_1}(\mathbf{r}_i-\mathbf{r}_G)
\end{aligned}$$

Similarly for the smaller simulation cell:

$$U_{\text{Ew}}^{sim} = \frac{1}{2}\sum_{i}^{N} q_i^2 v_{M_1} + \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} q_iq_jv_{\text{Ew}_1}(\mathbf{r}_i-\mathbf{r}_j) - \frac{1}{2}q_G^2v_{M_1} - \sum_{i}^{N} q_iq_Gv_{\text{Ew}_1}(\mathbf{r}_i-\mathbf{r}_G) \qquad \text{(C.21)}$$

so that the difference $\Delta U_{\text{Ew}} = U_{\text{Ew}}^{sup} - (U_{\text{Ew}}^{sim} + (M^3-1)U_{\text{Ew}}^{neut})$ becomes

$$\Delta U_{\text{Ew}} = \frac{1}{2}q_G^2\left(-v_{M_2} + v_{M_1}\right) \qquad \text{(C.22)}$$

and in the limit as $M \to \infty$ we get

$$\Delta U_{\text{Ew}} = \frac{1}{2}q_G^2 v_{M_1} \qquad \text{(C.23)}$$

The conclusion is thus that removing an electron causes a Madelung error of opposite sign to adding one, whereas adding a positive charge would have caused an error of the same sign. This emphasises that unlike in the case of considering purely charge densities, we cannot cancel the effects of a point charge of one sign by placing a point charge of the opposite sign on top of it (this situation would make no physical sense anyway)

The assumption underlying both of these derivations is that the copies of the neutral cell nearby to the charged cell have their N charges in positions comparable to those in the charged cell. This is equivalent to saying that the charges in the charged cell do not move to screen the effect of the extra charge q_P or removed charge q_G. In something like a charged version of a perfect crystal calculation, where the charge is spread out over the whole simulation cell, there is not very much to screen and this will probably be quite close to true, but in a defect with a strongly localized charge, it will certainly not be. In this case, even the atoms move considerably from their positions in the perfect crystal, and the electron density moves still further, so the potential of the defect charge will then be quite strongly screened.

Bibliography

[1] P. H. Acioli and D. M. Ceperley, Phys. Rev. B **54**, 17199 (1996).

[2] D. Alfè and M J Gillan, J. Phys.: Condens. Matter **16** L305 (2004).

[3] D. Alfè and M. J. Gillan, Phys. Rev. B **70**, 161101(R) (2004).

[4] D. Alfè and M. J. Gillan, Phys. Rev. B **71**, 220101(R) (2005).

[5] L. M. Almeida, J. P. Perdew and C. Fiolhais, Phys. Rev. B **66**, 075115 (2002).

[6] H. d'Amour, D. Schiferl, W. Denner, H. Schulz, and W. B. Holzapfel, J. Appl. Phys. **49**, 4411 (1978).

[7] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, *LAPACK Users Guide*, Third Edition, Society for Industrial and Applied Mathematics, Philadelphia, PA (1999).

[8] C. Attaccalite, S. Moroni, P. Gori-Giorgi, and G. Bachelet, Phys. Rev. Lett. **88**, 256601 (2002).

[9] A. Badinski and R. J. Needs, Phys. Rev. E **76**, 036707 (2007).

[10] M. Bajdich, L. Mitas, G. Drobný, L. K. Wagner, and K. E. Schmidt, Phys. Rev. Lett. **96**, 130201 (2006).

[11] G. A. Baraff and M. Schlüter, Phys. Rev. B **30**, 1853 (1984).

[12] Ph. Baranek, C. M. Zicovich-Wilson, C. Roetti, R. Orlando, and R. Dovesi, Phys. Rev. B **64**, 125102 (2001).

[13] B. Barbiellini, J. Phys. Chem. Solids **61**, 341 (2000).

[14] R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., J. Comput. Phys. **96**, 258 (1991).

[15] E. R. Batista, J. Heyd, R. G. Hennig, B. P. Uberuaga, R. L. Martin, G. E. Scuseria, C. J. Umrigar, and J. W. Wilkins, Phys. Rev. B **74**, 121102 (2006).

[16] F. Bloch, Z. Phys. **52**, 555 (1928).

[17] E.I. Blount, Solid State Phys. **13**, 306 (1962).

[18] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. **98**, 046402 (2007).

[19] M. D. Brown, J. R. Trail, P. López Ríos, and R. J. Needs, J. Chem. Phys. **126**, 224110 (2007).

[20] C. G. Broyden, Math. Comput., **19**, 577 (1965).

[21] M. Caffarel and P. Claverie, J. Chem. Phys. **88**, 1088 (1988).

[22] M. Caffarel and P. Claverie, J. Chem. Phys. **88**, 1100 (1988).

[23] J. Carrasco, N. Lopez, and F. Illas, Phys. Rev. Lett. **93**, 225502 (2004).

[24] M. Casula, Phys. Rev. B **74**, 161102 (2006).

[25] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).

[26] M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).

[27] C. W. M. Castleton, A. Höglund and S. Mirbt, Phys. Rev. B, **73**, 035215 (2006).

[28] C. R. Catlow, R. James, W. C. Mackrodt, and R. F. Stewart, Phys. Rev. B **25**, 1006 (1982).

[29] D. M. Ceperley, Phys. Rev. Lett. **69**, 331 (1992).

[30] D. M. Ceperley. J. Stat. Phys, **63**, 1237, (1991).

[31] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[32] M. W. Chase, Jr., *NIST-JANAF Themochemical Tables*, Fourth Edition, J. Phys. Chem. Ref. Data, Monograph **9**, 1 (1998).

[33] S. Chiesa, D. M. Ceperley, R. M. Martin and M. Holzmann, Phys. Rev. Lett. **97**, 076404 (2006).

[34] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, M. C. Payne, Zeitschrift fr Kristallographie, **220**, 567 (2005).

[35] M. L. Cohen and V. Heine. *Pseudopotential theory of cohesion and structure.* In Solid State Physics, **24**, 250. Academic Press, New York, (1970).

[36] M. Dewing, J. Chem. Phys. **113**, 5123 (2000).

[37] R. H. Doremus, J. Appl. Phys. **100**, 101301 (2006).

[38] E. Dörre and H. Hübner, *Alumina: Processing, Properties and Applications*, Springer-Verlag (1984).

[39] R. M. Dreizler, and E. K. U. Gross, *Density Functional Theory* (Springer, Berlin) (1990).

[40] N. D. Drummond, P. López Ríos, A. Ma, J. R. Trail, G. G. Spink, M. D. Towler, and R. J. Needs, J. Chem. Phys. **124**, 224104 (2006).

[41] N. D. Drummond and R. J. Needs, Phys. Rev. B **72**, 085124 (2005).

[42] N. D. Drummond and R. J. Needs, Phys. Rev. Lett. **99**, 166401 (2007).

[43] N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. B **70**, 235119 (2004).

[44] A. G. Eguiluz, M. Heinrichsmeier, A. Fleszar and W. Hanke, Phys. Rev. Lett. **68**, 1359 (1992).

[45] P. Erhart, K. Albe, and A. Klein, Phys. Rev. B **76**, 075204 (2007).

[46] P. P. Ewald, Ann. Phys. **64**, 25 (1921).

[47] S. Fahy, X. W. Wang and S. G. Louie, Phys. Rev. B **42**, 3503 (1990).

[48] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. Wiley, New York, (1968).

[49] R. P. Feynman, Phys. Rev. **94**, 262 (1954).

[50] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956).

[51] A. Filippetti and N. Spaldin, Phys. Rev. B **68**, 045111 (2003).

[52] M. W. Finnis, A. Y. Lozovoi and A. Alavi, Annu. Rev. Mater. Res **35**, 167 (2005).

[53] H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).

[54] W. M. C. Foulkes, L. Mitas, R. J. Needs, G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[55] R. H. French, J. Am. Ceram. Soc. **73**, 477 (1990).

[56] A. García and J. E. Northrup, Phys. Rev. Lett. **74**, 1131 (1995).

[57] P. García-González, Program for density-functional theory calculations of jellium slabs (dftsurf4), (2001).

[58] P. García-González and R. W. Godby, Phys. Rev. B **63**, 075112 (2001) .

[59] R. Gaudoin and J. M. Pitarke, Phys. Rev. Lett. **99**, 126406 (2007).

[60] R. Gaudoin, M. Nekovee, W. M. C. Foulkes, R. J. Needs and G. Rajagopal, Phys. Rev. B **63**, 115115 (2001).

[61] M. Ghamnia, C. Jardin and M. Bouslama, J. Electron Spectrosc. Relat. Phenom. **133**, 55 (2003).

[62] J. Goniakowski, J. M. Holender, L. N. Kantorovich, and M. J. Gillan, J. A. White, Phys. Rev. B **53**, 957 (1996).

[63] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, D.C. Allan, Computational Materials Science **25**, 478 (2002).

[64] J. P. Goss, P. R. Briddon, and M. J. Shaw, Phys. Rev. B **76**, 075204 (2007).

[65] C. W. Greeff and W. A. Lester, Jr., J. Chem. Phys. **109**, 1607 (1998).

[66] R. W. Grimes, J. Am. Ceram. Soc. **77**, 378 (1994).

[67] E. K. U. Gross, E. Runge, and O. Heinonen, *Many Particle Theory* (Hilger, Bristol) (1991).

[68] D. R. Hamman, M. Schlüter, and C. Chiang, Phys. Rev. Lett., **43**, 1494, (1979).

[69] B. L. Hammond, P. J. Reynolds, and W. A. Lester, Jr., J. Chem. Phys. **87**, 1130 (1987).

[70] A. H. Heuer and K. P. D. Lagerlöf, Philosophical Magazine Letters **79**, 619 (1999).

[71] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).

[72] J. Heyd and G. E. Scuseria, J. Chem. Phys. **120**, 7274 (2004).

[73] J. Heyd and G. E. Scuseria, J. Chem. Phys. **121**, 1187 (2004).

[74] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[75] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, Phys. Rev. E **68**, 046707 (2003).

[76] R. Q. Hood, P. R. Kent, R. J. Needs, and P. R. Briddon, Phys. Rev. Lett. **91**, 076403 (2003).

[77] C. J. Huang, C. Filippi, and C. J. Umrigar, J. Chem. Phys. **108**, 8838 (1998).

[78] M. M. Hurley and P.A. Christiansen, J. Chem. Phys. **86**, 1069 (1987).

[79] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).

[80] F. Janetzko, R. A. Evarestov, T. Bredow, K. Jug, physica status solidi(b) **241**, 1032, (2004).

[81] W. Janke, *Statistical Analysis of Simulations: Data Correlations and Error Estimation*, published in J. Grotendorst, D. Marx, A. Muramatsu (Eds.), *Quantum Simulations of Many-Body Systems: From Theory to Algorithms*, John von Neumann Institute for Computing, Jülich, NIC Series **10**, 423 (2002).

[82] K. Johnston, M. R. Castell, A. T. Paxton, M. W. Finnis, Phys. Rev. B **70**, 085415 (2004).

[83] G. Jomard, T. Petit, A. Pasturel, L. Magaud, G. Kresse and J. Hafner, Phys. Rev. B **59**, 4044 (1999).

[84] R.O. Jones, and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[85] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, R1651 (1993).

[86] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[87] P. R. Kent, Randolph Q. Hood, A. J. Williamson, R. J. Needs, W. M. C. Foulkes, and G. Rajagopal, Phys. Rev. B **59**, 1917 (1999).

[88] P. R. C. Kent, R. J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12344 (1999).

[89] C. Kittel, *Introduction to Solid State Physics* 7th ed. (Wiley, New York, 1996).

[90] L. Kleinman and D. M. Bylander, Phys. Rev. Lett., **48**, 1425, (1982).

[91] W. Kohn, Phys. Rev. **115**, 809 (1959).

[92] W. Kohn, Phys. Rev. **133**, A171 (1964).

[93] W. Kohn and J. R. Onffroy, Phys. Rev. B **8**, 2485 (1973).

[94] W. Kohn, L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[95] F. A. Kröger and H. J. Vink, *Solid State Physics*, Vol. 3, F. Seitz and D. Turnbull Eds, New York: Academic Press, (1956).

[96] E. Krotschek, W. Kohn and G.-X. Qian, Phys. Rev. B, **32**, 5693 (1985).

[97] E. Krotschek and W. Kohn, Phys. Rev. Lett. **57**, 862 (1986).

[98] S. Kurth and J. P. Perdew, Phys. Rev. B **59**, 10461 (1985).

[99] S. Kurth, J. P. Perdew, and P. Blaha, Int. J. Quantum Chem. **75**, 889 (1999).

[100] Y. Kwon, D. M. Ceperley, and R. M. Martin, Phys. Rev. B **48**, 12037 (1993).

[101] Y. Kwon, D. M. Ceperley, and R. M. Martin, Phys. Rev. B **58**, 6800 (1998).

[102] K. P. D. Lagerlöf and R. W. Grimes, Acta Mater. **46**, 5689 (1998).

[103] D. B. Laks, C. G. Van de Walle, G. F. Neumark, P. E. Blöchl, and S. T. Pantelides, Phys. Rev. B **45**, 10965 (1992).

[104] N. D. Lang and W. Kohn, Phys. Rev. B **1**, 4555 (1970).

[105] D. C. Langreth and M. J. Mehl, Phys. Rev. Lett **46**, 446 (1981).

[106] J. Lento, J.-L. Mozos and R. M. Nieminen, J. Phys. Condens. Matter, **14**, 2637 (2002).

[107] M. Leslie and M. J. Gillan, J. Phys. C, **18**, 973 (1985).

[108] M. Levy, Proc. Natl. Acad. Sci. U. S. A., **76**, 6062 (1979).

[109] M. Levy, Phys. Rev. A, **26**, 1200 (1982).

[110] W.-K. Leung, R. J. Needs, G. Rajagopal, S. Itoh, and S. Ihara, Phys. Rev. Lett. **83**, 2351 (1999).

[111] X. P. Li, R. J. Needs, R. M. Martin and D. M. Ceperley, Phys. Rev. B **45**, 6124 (1992).

[112] E. H. Lieb, Int. J. Quant. Chem., **24**, 243 (1983).

[113] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **74**, 066701 (2006).

[114] A. Malatesta, S. Fahy, G. B. Bachelet, Phys. Rev. B **56**, 12201 (1997).

[115] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).

[116] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge (2004).

[117] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[118] T. Mattila and A. Zunger, Phys. Rev. B **58**, 1367 (1998).

[119] K. Matsunaga, T. Tanaka, T. Yamamoto and Y. Ikuhara, Phys. Rev. B **68**, 085110 (2003).

[120] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys., **21**, 1087 (1953).

[121] L. Mitáš, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).

[122] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976)

[123] J. W. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos, J. Chem. Phys. **77**, 5188, (1976).

[124] R. Needs and N. Drummond, Phys. Rev. B **72**, 085124 (2005).

[125] R.J. Needs, M.D. Towler, N.D. Drummond and P.R.C. Kent, *CASINO Version 2.0 User Manual* (Cambridge: University of Cambridge, 2004) .

[126] M. P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. **87**, 043401 (2001).

[127] G. Ortiz and R.M. Martin, Phys. Rev. B **49**, 14202 (1994).

[128] D. E. Parry, Surf. Sci. **49**, 433 (1975).

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

[130] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[131] J. P. Perdew, S. Kurth and A. Zapan, Phys. Rev. Lett **82**, 2544 (1999).

[132] J. P. Perdew and Y. Wang, Phys. Rev. B **46**, 12947 (1992).

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

[134] J. C. Phillips, Phys. Rev. **112**, 685 (1958).

[135] J. M. Pitarke, Phys. Rev. B **70**, 087401, (2004).

[136] J. M. Pitarke and A. G. Eguiluz, Phys. Rev. B **57**, 6329 (1998).

[137] S. Pöykkö, M. J. Puska, and R. M. Nieminen, Phys. Rev. B **53**, 3813 (1996).

[138] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge Univ Press (1986).

[139] G. Rajagopal, R. J. Needs, A. James, S. D. Kenny, and W. M. C. Foulkes, Phys. Rev. B **51**, 10591 (1995).

[140] R. Resta, Ferroelectrics **136**, 51 (1992).

[141] R. Resta, J. Phys. Condens. Matter **14**, R625 (2002) .

[142] R. Resta, Phys. Rev. Lett. **80**, 1800 (1998).

[143] R. Resta, Phys. Rev. Lett. **96**, 137601 (2006).

[144] R. Resta, S. Sorella, Phys. Rev. Lett. **82**, 370 (1999).

[145] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr., J. Chem. Phys. **77**, 5593, (1982).

[146] R. Saito, G. Dresselhaus, M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial, London, 1998).

[147] P. A. Schultz, Phys. Rev. B **60**, 1551 (1999).

[148] C. Sgiarovello, M. Peressi, R. Resta, Phys. Rev. B **64**, 115202 (2001).

[149] L. J. Sham and M. Schlter, Phys. Rev. Lett. **51**, 1888 (1983).

[150] J. Shim, E.-K. Lee, Y. J. Lee, R. M. Nieminen, Phys. Rev. B **71**, 035206 (2005).

[151] A. Shukla, Phys. Rev. B **61**, 13277 (2000).

[152] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, J. Chem. Phys. **122**, 084119 (2005).

[153] J. C. Slater, *Quantum Theory of Molecules and Solids*, New York: McGraw-Hill (1965).

[154] D.M. Smythe, *The Defect Chemistry of Metal Oxides*, Oxford University Press (2000).

[155] F. Sottile and P. Ballone, Phys. Rev. B **64**, 045105 (2001).

[156] I. Souza, T. Wilkens, R. Martin, Phys. Rev. B **62**, 1666 (2000).

[157] V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew, Phys. Rev. B **69**, 075102 (2004).

[158] I. Tanaka, K. Tatsumi, M. Nakano, H. Adachi, and F. Oba, J. Am. Ceram. Soc., **85**, 68 (2002).

[159] B. Tanatar and D. M. Ceperley, Phys. Rev. B **39**, 5005 (1989).

[160] A. Thom and A. Alavi, J. Chem. Phys. **123**, 204106, (2005).

[161] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007).

[162] J. R. Trail and R. J. Needs, J. Chem. Phys. **122**, 014112 (2005).

[163] J. R. Trail and R. J. Needs, J. Chem. Phys. **122**, 174109 (2005).

[164] P. Umari, X. Gonze, A. Pasquarello, Phys. Rev. B **69**, 235102 (2004).

[165] P. Umari, A. J. Willamson, G. Galli, and N. Marzari, Phys. Rev. Lett **95**, 207602 (2005).

[166] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

[167] C. J. Umrigar, M. P. Nightingale and K. J. Runge, J. Chem. Phys. 99, 2865 (1993).

[168] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).

[169] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[170] D. Vanderbilt and R. D. King-Smith, Phys. Rev. B **48**, 4442 (1993).

[171] C. G. van de Walle and J. Neugebauer, J. App. Phys **95**, 3851 (2004).

[172] M. Veithen, X. Gonze, and Ph. Ghosez, Phys. Rev. B **66**, 235113 (2002).

[173] U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).

[174] P. R. Wallace, Phys. Rev. **71**, 622 (1947).

[175] G. H. Wannier, Phys. Rev. **52**, 191 (1937).

[176] T. Wilkens and R. Martin, Phys. Rev. B **63**, 235108 (2001).

[177] A. J. Williamson, G. Rajagopal, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, Y. Wang and M-Y. Chou, Phys. Rev. B **55**, R4851 (1997).

[178] B. Wood, Ph.D, Thesis, Imperial College London (2005).

[179] B. Wood, W. M. C. Foulkes, M. D. Towler, N. D. Drummond, J. Phys: Condens. Matter **16**, 891 (2004).

[180] B. Wood and W. M. C. Foulkes, J. Phys: Condens. Matter **18**, 2305 (2006).

[181] B. Wood, N. D. M. Hine, W. M. C. Foulkes and P. García-González, Phys. Rev. B **76**, 035403 (2007).

[182] A. F. Wright and N. A. Modine, Phys. Rev. B **74**, 235209 (2006).

[183] R. W. G. Wyckoff, *Crystal Structures*, New York: Wiley (1964).

[184] Y. N. Xu, Z. Q. Gu, X. F. Zhong, and W. Y. Ching, Phys. Rev. B **56**, 7277 (1997).

[185] Z. Yan, J. P. Perdew, S. Kurth, C. Fiolhais and L. Almeida, Phys. Rev. B **61**, 2595 (2000).

[186] S. B. Zhang and J. E. Northrup, Phys. Rev. Lett. **67**, 2339 (1991).

[187] S. B. Zhang, S. H. Wei, A. Zunger, and H. Katayama-Yoshida, Phys. Rev. B **57**, 9642 (1998).