QUANTUM MONTE CARLO CALCULATIONS OF ELECTRONIC EXCITATIONS

By
Andrew James Williamson
Robinson College, Cambridge

Theory of Condensed Matter Group
Cavendish Laboratory
Madingley Road
Cambridge
CB3 0HE

A dissertation submitted for the degree of Doctor of Philosophy

September 1996

PREFACE

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Cambridge. It contains an account of research carried out between October 1993 and September 1996 in the Theory of Condensed Matter Group of the Department of Physics at the University of Cambridge, under the supervision of Dr. R.J. Needs.

Parts of the material have been published elsewhere as follows:

Chapter 4	Phys. Rev. B **53**, 9640 (1996), [1]
	Phys. Rev. Lett. **77**, 1099 (1996), [2]
Chapter 5	Phys. Rev. B **53**, 1814 (1996), [3]

Except where reference is made to the work of others, this dissertation is the result of my own work and contains nothing which is the outcome of work done in collaboration. This thesis has not been submitted in whole or in part for any degree or diploma at this or any other university.

Andrew James Williamson
September 1996

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Richard Needs, for all his encouragement and guidance throughout the course of this work and Dr. Guna Rajagopal for teaching me many things, most importantly, to understand just exactly "what is the deal", at all times. I would also like to thank my other colleague at Cambridge, Dr. Steven Kenny, for all the help he gave me, especially during the first few months of my Ph.D., and for convincing me that any error bar greater than 0.000001 Hartrees is unacceptable!

I would like to thank my colleagues Dr. Matthew Foulkes, Dr. Louisa Fraser and Dr. Alan James at Imperial College for many helpful discussions.

Finally, I wish to thank my wife and family for all their support and encouragement.

Contents

List of Figures

List of Tables

SUMMARY

This thesis is concerned with developing quantum Monte Carlo techniques to calculate electronic excitations.

Variational and diffusion quantum Monte Carlo calculations for the ground-state electronic properties of atoms and solids are reported. The importance of the choice of trial wavefunction in these calculations is highlighted. Methods for optimising this trial wavefunction, based on the minimisation of the variance of the local energy are developed for both atomic and solid wavefunctions. New forms of variational trial wavefunction are introduced that are more suitable for optimisation and are also much faster to evaluate.

The existence of finite size effects when simulating periodic systems is illustrated. The source of these finite size errors is tracked down to the Ewald method which introduces spurious correlations between electrons in different simulation cells. A new electron-electron interaction which consists of the Coulomb interaction between point particles at short range, and a long ranged averaged (Hartree) interaction is proposed. Hartree-Fock, variational and diffusion quantum Monte Carlo results for the energy of diamond-structure silicon are presented which demonstrate the effectiveness of this new method.

Variational and diffusion quantum Monte Carlo results for the quasiparticle band-structure of diamond-structure silicon using both highly optimized trial wavefunctions and the new electron-electron interaction are reported. These are based on two separate methods of calculating excited states, (i) the addition and subtraction of electrons, (ii) the promotion of electrons from the valence band into the conduction band. These two methods of calculating excited states are compared and contrasted with each other and other more traditional methods of calculating quasiparticle band-structures, such as the local density approximation to density functional theory.

The results show that if a careful choice of trial/guiding wavefunction is made, and the electron-electron interaction is chosen to reduce the Coulomb finite size effects, it

is possible to calculate electronic excitation energies within diffusion quantum Monte Carlo that represent a significant improvement over those calculated using the local density approximation.

Chapter 1

Introduction

Quantum Monte Carlo techniques provide a practical method for solving the many-body Schrödinger equation. They are closely related to the well established classical Monte Carlo methods that have been successfully applied to a wide range of problems involving stochastic behaviour ranging from scientific problems, engineering problems and modelling the financial markets. The common link between classical and quantum Monte Carlo techniques is the use of random numbers to evaluate multi-dimensional integrals.

In its simplest form, the variational Monte Carlo (VMC) technique is based on evaluating a high-dimensional integral by sampling the integrand using a set of randomly generated points. It can be shown that the integral converges faster using a Monte Carlo technique than more conventional techniques based on sampling the integrand on a regular grid for problems involving more than a few dimensions. Moreover, the statistical error in the estimate of the integral decreases as the square root of the number of points sampled, irrespective of the dimensionality of the problem.

In this chapter the QMC techniques are compared with more established methods of solving the many-body Schrödinger equation. The relative merits of the different techniques for calculating the electronic structure of atoms, molecules and solids are considered. The progress made over the past decade in developing QMC methods as a tool for tackling realistic continuum electronic structure problems is described. At the end of the chapter, a summary of the layout of the thesis is given.

1.1 The Many-Electron Problem

Within the Born-Oppenheimer approximation[4], the time independent Schrödinger equation for a fully interacting many-electron system is

$$\hat{H}\Psi = \sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m}\nabla_i^2 - \sum_{\alpha}\frac{Z_\alpha e^2}{4\pi\epsilon_0|\mathbf{r}_i - \mathbf{d}_\alpha|}\right)\Psi + \frac{1}{2}\sum_{i}\sum_{j\neq i}\frac{e^2}{4\pi\epsilon_0|\mathbf{r}_i - \mathbf{r}_j|}\Psi = E\Psi \quad , \tag{1.1}$$

where Ψ is the N-electron wavefunction, $\mathbf{r}_i$ are the electron positions, $\mathbf{d}_\alpha$ are the positions of the ions and Z_α are the ionic charges. This equation is impossible to solve exactly so approximate solutions must be sought. One of the main challenges of condensed matter physics is to try to find good, workable approximations that contain the essence of the physics involved in a particular problem and to obtain the most accurate solutions possible. For the rest of this thesis all equations will be written in atomic units, $e = m_e = \hbar = 4\pi\epsilon_0 = 1$.

1.1.1 One-Electron Methods

One of the most common ways of dealing with the many-fermion problem is to assume that each electron can be considered separately. Each electron is treated as moving in a mean field potential, $U(\mathbf{r})$. This potential models the effects of all the other particles in the system, as well as any external potential acting on the system.

The one-electron equations are of the form

$$-\frac{1}{2}\nabla^2\psi_i(\mathbf{r}) + U(\mathbf{r})\psi_i(\mathbf{r}) = \epsilon\psi_i(\mathbf{r}) \quad , \tag{1.2}$$

where $\psi_i(\mathbf{r})$ is a one-electron wavefunction and ϵ_i are Lagrange multipliers which arise from the fact that the one-electron wavefunctions are normalised. Choosing an appropriate $U(\mathbf{r})$ for the single electron is still a very complicated problem. $U(\mathbf{r})$ depends upon the interactions between the electrons and therefore on the one-electron wavefunctions. Since initially neither of these quantities, $U(\mathbf{r})$ or $\psi_i(\mathbf{r})$, is known, it is necessary to solve Eq.(1.2) in a self-consistent manner.

The Hartree Approximation

This approximation starts from the one-electron equations of Eq(1.2). $U(\mathbf{r})$ is chosen to try to model the interaction terms in this equation. The ions contribute a potential

$$U_{ion}(\mathbf{r}) = -\sum_{\alpha} \frac{Z_\alpha}{|\mathbf{r} - \mathbf{d}_\alpha|} \quad . \tag{1.3}$$

All the other electrons in the system also contribute to the potential. The potential due to the electrons is approximated by the electrostatic interaction with all the others, which can be written in terms of the electron density, $\rho(\mathbf{r})$, as

$$U_H(\mathbf{r}) = \int d\mathbf{r}'[\rho(\mathbf{r}') - \rho_i(\mathbf{r}')]\frac{1}{|\mathbf{r} - \mathbf{r}'|} \quad , \tag{1.4}$$

where the self-interaction potential due to electron i has been removed.

To actually calculate the Hartree potential it is necessary to know the electronic charge distribution of the system. If the electrons are assumed to be independent of each other, then it is straightforward to construct $\rho(\mathbf{r})$ from the single electron eigenstates

$$\rho(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2 \quad , \tag{1.5}$$

where the summation over i includes all occupied states. Using this charge density the total one-electron potential is

$$U_i(\mathbf{r}) = U_{ion}(\mathbf{r}) + \sum_{j \neq i} \int d\mathbf{r}' |\psi_j(\mathbf{r}')|^2 \frac{1}{|\mathbf{r} - \mathbf{r}'|} \quad . \tag{1.6}$$

The potential $U_i(\mathbf{r})$ is different for each orbital, and therefore the orbitals are not orthogonal. Note that $U_i(\mathbf{r})$ depends on all the other orbitals, ψ_j, and so the solution of Eq.(1.2) must be found self-consistently.

The choice of $U(\mathbf{r})$ in Eq.(1.6) all seems a bit like guesswork, but it can also be derived using the variational principle. We start with Eq.(1.1). The electrons are assumed to be non-interacting, and so the N-electron wavefunction is just the product of the one-electron wavefunctions,

$$\Psi = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\dots\psi_N(\mathbf{r}_N) \quad . \tag{1.7}$$

This Ψ can be used with Eq.(1.1) to find the expectation value of $\hat{H}$

$$\begin{aligned} \left\langle \Psi|\hat{H}|\Psi \right\rangle &= \sum_i \int d\mathbf{r}\psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 + U_{ion}(\mathbf{r})\right) \psi_i(\mathbf{r}) \\ &+ \frac{1}{2}\sum_i \sum_{j(\neq i)} \int d\mathbf{r}d\mathbf{r}' \frac{|\psi_i(\mathbf{r})|^2|\psi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} \quad . \end{aligned} \tag{1.8}$$

Introducing a Lagrange multiplier, ϵ_i , for the condition that the one-electron wavefunctions are normalised, and minimising the above equation with respect to the wavefunctions, so that

$$\frac{\delta}{\delta\Psi}\left[< \hat{H} > - \sum_j \epsilon_j \int |\psi_j|^2 d\mathbf{r}\right] = 0 \quad , \tag{1.9}$$

leads to a set of single particle equations,

$$\left(-\frac{1}{2}\nabla^2 + U_{ion}(\mathbf{r})\right)\psi_i(\mathbf{r}) + \sum_{j(\neq i)} \int d\mathbf{r}' \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \quad , \tag{1.10}$$

which are the same as substituting Eq.(1.6) in Eq.(1.2). These equations are known as the Hartree equations.

The Hartree-Fock Approximation

The Hartree-Fock approximation is an extension of the above Hartree approximation to include the permutation symmetry of the wavefunction which leads to the exchange interaction. Exchange is due to the Pauli exclusion principle, which states that the total wavefunction for the system must be antisymmetric under particle exchange.

This means that when two arguments are swapped the wavefunction changes sign as follows:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_N) = -\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N) \quad , \tag{1.11}$$

where $\mathbf{x}_i$ includes coordinates of position and spin. Therefore no two electrons can have the same set of quantum numbers, and electrons with the same spin cannot occupy the same state simultaneously.

Instead of using the simple product form of the wavefunction shown in Eq.(1.7), a Slater determinant wavefunction[5, 6] which satisfies antisymmetry is used

$$D = \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \cdots & \psi_1(\mathbf{x}_N) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_N) \\ \vdots & \vdots & & \\ \psi_N(\mathbf{x}_1) & \psi_N(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \quad , \tag{1.12}$$

where $\psi_i(\mathbf{x}_j)$ are the one-electron wavefunctions.

Following exactly the same method of minimising the expectation value of $\hat{H}$ with respect to the one-electron wavefunctions as was used in the derivation of the Hartree equations, results in the following set of one-electron equations, the Hartree-Fock equations;

$$\begin{aligned} \epsilon_i \psi_i(\mathbf{r}) &= \left(-\frac{1}{2}\nabla^2 + U_{ion}(\mathbf{r})\right)\psi_i(\mathbf{r}) + \sum_j \int d\mathbf{r}' \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}\psi_i(\mathbf{r}) \\ &- \sum_j \delta_{s_i s_j} \int d\mathbf{r}' \frac{\psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{r}) \quad , \end{aligned} \tag{1.13}$$

where s_i labels the spin of particle i. Note the self-interaction cancels out from the second and third terms. The extra term in these equations, when compared to Eq.(1.10), is known as the exchange term and is only non-zero when considering electrons of the same spin. The effect of exchange on the many-body system is that electrons of like spin tend to avoid each other. As a result of this, each electron has a “hole” associated with it which is known as the exchange hole (or the Fermi hole).

This is a small volume around the electron which like-spin electrons avoid. The charge contained in the exchange hole is positive and exactly equivalent to the absence of one-electron.

Unlike all the other terms acting on ψ_i , the exchange term is a non-local integral operator and this makes the Hartree-Fock equations hard to solve in all but a few special cases.

1.1.2 Configuration Interaction

The Configuration Interaction (CI) method is often employed by chemists when studying atoms or small molecules. In this method $< \Psi|\hat{H}|\Psi >$ is minimised by searching through all Ψ which are linear combinations of N-electron determinants made up from a total of M basis states ($M >> N$). This is obviously impractical to do thoroughly since there are

$$ {}^{M}C_{N} = \frac{M!}{N!(M-N)!} \tag{1.14} $$

N-electron determinants and therefore the matrix to be diagonalised has dimension ${}^{M}C_{N}$. Hence Configuration Interaction calculations usually only include determinants close to the Hartree-Fock ground state determinant. This means ignoring all high energy states, so the accuracy of these Configuration Interaction calculations is not as good as results found using Quantum Monte Carlo. Configuration Interaction calculations are not possible in solids due to the large number of determinants which would be required, however they can provide highly accurate results for small systems. In other words, CI calculations scale very badly with system size.

1.1.3 Density Functional Methods

Density Functional theory[7, 8] is a formally exact theory based on the charge density of a system. Kohn-Sham Density Functional theory[9] is a formally exact one-electron theory. Working within the Born-Oppenheimer approximation, the many-body Schrödinger equation,

$$\hat{H}\Psi = \sum_{i=1}^{N}\left(-\frac{1}{2}\nabla_i^2 - \sum_{\alpha}\frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{d}_\alpha|}\right)\Psi + \frac{1}{2}\sum_{i}\sum_{j\neq i}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\Psi = E\Psi \quad , \tag{1.15}$$

where Ψ is the many-body wavefunction, is replaced by a set of N one-electron equations of the form

$$\left(-\frac{1}{2}\nabla^2 + V(\mathbf{r})\right)\psi_i(\mathbf{r}) = \epsilon\psi_i(\mathbf{r}) \quad , \tag{1.16}$$

where $\psi_i(\mathbf{r})$ is a single-electron wavefunction. These one-electron equations contain a potential $V(\mathbf{r})$ produced by all the ions and the electrons. Density Functional theory properly includes all parts of the electron-electron interaction, i.e. the Hartree potential

$$V_H(\mathbf{r}) = \int d\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \quad , \tag{1.17}$$

where ρ is the charge density of all the electrons, a potential due to exchange and correlation effects, $V_{XC}(\mathbf{r})$, and the external potential due to the ions, $V_{\text{ext}}(\mathbf{r})$,

$$V(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r}) \quad . \tag{1.18}$$

Hohenberg and Kohn[10] originally developed Density Functional theory for application to the ground state of a system of spinless fermions. In such a system the particle density is given by

$$\rho(\mathbf{r}) = N\int |\Psi_o(\mathbf{r},\mathbf{r}_2,\ldots,\mathbf{r}_N)|^2 d\mathbf{r}_2\ldots d\mathbf{r}_N \quad , \tag{1.19}$$

with Ψ_0 being the many-body ground state wavefunction of the system. It can be shown that the total ground state energy of the system is a functional of the density, $E[\rho(\mathbf{r})]$, and that if the energy due to the electron-ion interactions is excluded the remainder of the energy is a universal functional of the density, $F[\rho(\mathbf{r})]$ (i.e. $F[\rho(\mathbf{r})]$ does not depend on the potential from the ions). The most elegant proof of Density Functional theory is due to Levy[11] and is as follows:

For a particular N-representable density[1] (i.e. any density given by an antisymmetric N-electron wavefunction), a functional of the density corresponding to any operator

[1] It was shown by Harriman[12] that all densities are N-representable

$\hat{O}$ can be defined via

$$O\left[\rho(\mathbf{r})\right] = \min_{|\Psi>\Rightarrow\rho(\mathbf{r})} \left\langle \Psi|\hat{O}|\Psi \right\rangle \quad . \tag{1.20}$$

The right hand side is evaluated by searching over wavefunctions, Ψ, which give rise to the density $\rho(\mathbf{r})$ and looking for the one which makes the expectation value of the operator $\hat{O}$ a minimum.

We can define $F[\rho(\mathbf{r})]$ in the same way, where

$$\hat{F} = \sum_i -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad , \tag{1.21}$$

and

$$F\left[\rho(\mathbf{r})\right] = \min_{|\Psi>\Rightarrow\rho(\mathbf{r})} \left\langle \Psi|\hat{F}|\Psi \right\rangle \quad . \tag{1.22}$$

Now let Ψ_0 be the ground state of an N-electron system and Ψ a state which yields a density $\rho(\mathbf{r})$ and minimises $< \Psi|\hat{F}|\Psi >$. Then, from the definition of $E[\rho(\mathbf{r})]$,

$$E[\rho(\mathbf{r})] = F[\rho(\mathbf{r})] + \int \rho(\mathbf{r})V_{\text{ext}}(\mathbf{r})d\mathbf{r} =< \Psi|\hat{F} + V_{\text{ext}}|\Psi > \quad . \tag{1.23}$$

Now $\hat{F} + V_{\text{ext}}$ is the electronic Hamiltonian, from Eq.(1.15), and so $E[\rho(\mathbf{r})]$ must obey the variational principle (see section 2.1),

$$E[\rho(\mathbf{r})] \geq E_0 \quad . \tag{1.24}$$

Also, from the definition of $F[\rho(\mathbf{r})]$, in Eq.(1.22), we have

$$F[\rho_0(\mathbf{r})] \leq< \Psi_0|\hat{F}|\Psi_0 >, \tag{1.25}$$

since Ψ_0 is just one of the trial wavefunctions that yield $\rho_0(\mathbf{r})$. Adding $\int \rho(\mathbf{r})V_{\text{ext}}(\mathbf{r})d\mathbf{r}$ to the above equation gives

$$E[\rho_0(\mathbf{r})] \leq E_0 \quad , \tag{1.26}$$

which in combination with Eq.(1.24) yields the desired result

$$E[\rho_0(\mathbf{r})] = E_0 \quad , \tag{1.27}$$

hence completing the proof.

Kohn-Sham Equations

Kohn and Sham [9] introduced a method based on the Hohenberg-Kohn theorem that enables one to minimise the functional $E[\rho(\mathbf{r})]$ by varying $\rho(\mathbf{r})$ over all densities containing N electrons. This constraint is introduced by the Lagrange multiplier, μ, chosen so that $\int \rho(\mathbf{r})d\mathbf{r} = N$,

$$\frac{\delta}{\delta\rho(\mathbf{r})}\left[E[\rho(\mathbf{r})] - \mu\int\rho(\mathbf{r})d\mathbf{r}\right] = 0$$

$$\Rightarrow \quad \frac{\delta E[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} = \mu \quad . \tag{1.28}$$

Kohn and Sham chose to separate $F[\rho(\mathbf{r})]$ into three parts, so that $E[\rho(\mathbf{r})]$ becomes

$$E\left[\rho(\mathbf{r})\right] = T_s\left[\rho(\mathbf{r})\right] + \frac{1}{2}\int\int\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{XC}[\rho(\mathbf{r})] + \int\rho(\mathbf{r})V_{\text{ext}}(\mathbf{r})d\mathbf{r} \quad , \tag{1.29}$$

where $T_s\left[\rho(\mathbf{r})\right]$ is defined as the kinetic energy of a *non-interacting* electron gas with density $\rho(\mathbf{r})$,

$$T_s\left[\rho(\mathbf{r})\right] = -\frac{1}{2}\sum_{i=1}^{N}\int\psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})d\mathbf{r} \quad . \tag{1.30}$$

Eq.(1.29) also acts as a definition for the *exchange-correlation energy functional*, $E_{XC}[\rho(\mathbf{r})]$. We can now rewrite Eq.(1.28) in terms of an effective potential, $V_{\text{eff}}(\mathbf{r})$, as follows

$$\frac{\delta T_s\left[\rho(\mathbf{r})\right]}{\delta\rho(\mathbf{r})} + V_{\text{eff}}(\mathbf{r}) = \mu \quad , \tag{1.31}$$

where

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + \int\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + V_{XC}(\mathbf{r}) \quad , \tag{1.32}$$

and

$$V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \quad . \tag{1.33}$$

Now, if one considers a system that really contained *non-interacting* electrons moving in an external potential equal to $V_{\text{eff}}(\mathbf{r})$, as defined in Eq.(1.32), then the same analysis would lead to exactly the same Eq.(1.31). Therefore, to find the groundstate energy and density, E_0 and $\rho_0(\mathbf{r})$ all one has to do is solve the one-electron equations

$$\left(-\frac{1}{2}\nabla_i^2 + V_{\text{eff}}(\mathbf{r}) - \epsilon_i\right)\psi_i(\mathbf{r}) = 0 \quad . \tag{1.34}$$

As the density is constructed according to

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 \quad , \tag{1.35}$$

these equations (1.32-1.34) must be solved self-consistently with Eq.(1.35).

The above derivation assumes that the exchange-correlation functional is known. At present numerical exchange-correlation potentials have only been determined for a few simple model systems, and so most current density functional calculations use the Local Density Approximation (LDA). The LDA approximates the XC functional to a simple function of the density at any position, $\mathbf{r}$. The value of this function is the XC energy per electron in a uniform homogeneous electron gas of density $n(\mathbf{r})$. The LDA expression for $E_{XC}[n(\mathbf{r})]$ is

$$E_{XC}[n(\mathbf{r})] \approx \int \epsilon_{XC}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r} \quad . \tag{1.36}$$

The LDA is remarkably accurate, but often fails when the electrons are strongly correlated, as in systems containing d and f orbital electrons.

1.2 Quantum Monte Carlo Calculations

The absence of any exchange or correlation between electrons in the Hartree method leave this technique too inaccurate for performing modern electronic structure calculations.

Hartree-Fock calculations, which include the exchange interaction between electrons, are most useful for performing calculations on relatively small systems as they are considerably more computationally expensive than Hartree and DFT-LDA calculations, due to the non-local exchange term. Even for atoms, however, Hartree-Fock theory is not ideal. For example, H^- is predicted to be unstable in contradiction to reality.

Various improvements to Hartree-Fock theory have been attempted. Unrestricted Hartree-Fock theory ignores some of the simplifying restrictions which are normally applied to Hartree-Fock wavefunctions. The exchange interaction is allowed to make the spatial parts of spin up and spin down electron wavefunctions different for the same state. However, although for some systems this results in an improvement [13], especially for open shell systems, it also sometimes produces worse results than conventional Hartree-Fock theory [14]. In general, Hartree-Fock theory is most useful as a tool for providing qualitative answers. It is also used as the starting point for methods, such as some Quantum Monte Carlo calculations (see chapter 2).

The success of the local density approximation is currently understood to be due to two points. (i) The sum rule on the exchange-correlation hole is conserved, i.e. within the LDA, the exchange-correlation hole contains exactly an equal and opposite amount of charge to the electron it surrounds. (ii) The exchange-correlation energy only depends on the spherical average of the exchange-correlation hole, i.e.

$$E_{XC}[\rho] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\overline{\rho}_{XC}(\mathbf{r},|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \quad , \tag{1.37}$$

where $\overline{\rho}_{XC}(\mathbf{r},|\mathbf{r}-\mathbf{r}'|)$ is the spherical average of $\rho_{XC}(\mathbf{r},\mathbf{r}')$ about $\mathbf{r}'=\mathbf{r}$ for each $\mathbf{r}$. Therefore, the fact that in the LDA, ρ_{XC} is constrained to be spherically symmetric about $\mathbf{r}'=\mathbf{r}$ is not a handicap. However, in strongly correlated systems, i.e. those containing d and f orbitals, the correlations may change the whole nature of the ground state and the Local Density approximation, derived from homogeneous electron gas results, is not successful. For example, the high T_c superconductor La_2CuO_4 is an anti-ferromagnetic insulator but the LDA finds it to be metallic. Also FeO, MnO and NiO all have Mott metal-insulator transitions but the LDA predicts that they are either semiconductors or metals. The LDA is only expected to be accurate for systems with slowly varying electronic charge densities, which is not the case in most real systems, but despite this it has been surprisingly successful. Other failings of the LDA are that it tends to overbind atoms in solids, that it finds stable negative ions to be unstable and that it predicts iron to be *fcc* paramagnetic, when it is actually *bcc* ferromagnetic.

The main problem with Hartree, Hartree-Fock and LDA methods is approximations they introduce in the process of reducing the many-body problem to a one-electron problem. Hartree and Hartree-Fock calculations do not, in general, provide satisfactory results and are best used as a qualitative guide to the expected ground state properties. The Configuration Interaction method, while in principle exact, is in practice, only useful for small systems; for condensed matter systems it is not of practical value.

Density Functional theory within the LDA provides the current staple method of performing electronic structure calculations and for many purposes gives good results. However, it fails for highly correlated systems and tends to underestimate band gaps and overestimates cohesive energies and hence is not ideal. Many-body approaches have been successful in some calculations, particularly of band gaps, but they are difficult to implement and it is hard to go beyond the low order GW [15] approximation.

It is therefore clear that there is room for a straightforward, accurate approach to many-body systems: the Quantum Monte Carlo method!

1.3 Layout of Thesis

In chapter 2, the two QMC methods used for the calculations in the rest of this thesis, the Variational quantum Monte Carlo (VMC) and Diffusion quantum Monte Carlo (DMC) method are introduced. Details of the algorithms used to implement these methods on serial and parallel computers are given.

In chapter 3, the specific details of how to perform a QMC calculation on a solid using the supercell technique are given. The choice of wavefunction and evaluation of Coulomb interactions are discussed.

Chapter 4 describes our application of the variance minimisation, optimisation technique to the problem of producing trial/guiding wavefunctions for use in QMC calculations of solids. Again details are given of the algorithm used and how to implement it on both serial and parallel computers. New functional forms of wavefunction are introduced, that yield comparable if not improved accuracy over traditional functional forms, are more suitable for optimisation and are considerably faster to evaluate within a QMC code.

Chapter 5 describes new forms of electron-electron interaction that are designed to dramatically reduce the troublesome Coulomb finite size effects present in QMC supercell calculations.

The technical advances made to the QMC technique described in chapters 4 and 5 are brought together in chapter 6 to enable us to attempt a new application of QMC - the calculation of excitation energies. Two separate methods of evaluating excitation energies within QMC are, (i) the addition and subtraction of electrons and (ii) the promotion of electrons. Both these techniques require energies to be evaluated to at least an order of magnitude higher accuracy than previous QMC calculations. The results obtained from the two techniques are compared both with each other and with the results of more established electronic structure techniques.

Finally, in chapter 7, some conclusions on the work are drawn.

Chapter 2

Quantum Monte Carlo Methods

2.1 The Variational Principle

The Variational Principle[16, 17] provides the starting point for almost all methods which aim to find an approximate solution to Schrödinger's equation. It may be stated as follows:

> *The expectation value of a Hamiltonian, $\hat{H}$, calculated using a trial wavefunction, Ψ_T, is never lower in value than the true ground state energy, ϵ_0, which is the expectation value of $\hat{H}$ calculated using the true ground state wavefunction, Ψ_0.*

Obviously this is extremely valuable because it means that it is always possible to find an upper bound for the ground state energy. It is also possible to use variational methods to study excited states (see chapter 6), but the real strength of this principle lies in finding ground state energies. If a trial function, Ψ_T, is guessed for the ground state then the expectation value of the Hamiltonian is

$$< \Psi_T|\hat{H}|\Psi_T >= \frac{\int \Psi_T^* \hat{H} \Psi_T d\tau}{\int \Psi_T^* \Psi_T d\tau} \quad . \tag{2.1}$$

The true normalized eigenfunctions, Ψ_n, of $\hat{H}$ form a complete basis, so the trial wavefunction, Ψ_T, may be expanded as a linear combination of these eigenfunctions,

$$\Psi_T = \sum_n c_n \Psi_n \quad , \tag{2.2}$$

with

$$\sum_n |c_n|^2 = 1 \quad . \tag{2.3}$$

As the Ψ_n are normalised it follows that Ψ_T is normalised. Using the expansion of Ψ_T to calculate $\left\langle \hat{H} \right\rangle$ from Eq.(2.1) gives

$$\begin{aligned}
\left\langle \Psi_T | \hat{H} | \Psi_T \right\rangle &= \int \left(\sum_n c_n^* \Psi_n^* \right) \hat{H} \left(\sum_m c_m \Psi_m \right) d\tau \\
&= \sum_n \sum_m c_n^* c_m \int \Psi_n^* \hat{H} \Psi_m d\tau \\
&= \sum_n \sum_m c_n^* c_m \epsilon_m \delta_{nm} \\
&= \sum_n |c_n|^2 \epsilon_n \quad , \qquad (2.4)
\end{aligned}$$

where ϵ_n is the eigenvalue corresponding to eigenstate Ψ_n . Since $\epsilon_n \geq \epsilon_0$ for all n, it is clear that

$$\left\langle \Psi_T \left| \hat{H} \right| \Psi_T \right\rangle \geq \epsilon_0 \quad . \tag{2.5}$$

Variational calculations rely on making a physically plausible guess at the form of the ground state wavefunction, Ψ_T, of the Hamiltonian, $\hat{H}$. This guess will be referred to as the trial/guiding wavefunction throughout this thesis. The "trial" part of the name refers to the use of the wavefunction as a guess of the true groundstate wavefunction to be used as the input wavefunction in a Variational quantum Monte Carlo (VMC) calculation. The "guiding" part refers to the use of the same wavefunction as an input wavefunction in the Diffusion quantum Monte Carlo (DMC) algorithm as part of the mechanism to introduce importance sampling. This will described in more detail in section 2.5.1. The trial/guiding wavefunction depends on a number of variable parameters which can be adjusted to minimise the energy expectation value. If the guessed values of these parameters are good and the chosen functional form builds in enough variational freedom to adequately describe the physics of the system being studied, then very accurate estimates of the ground state energy can be obtained. Variational quantum Monte Carlo (VMC) calculations are direct applications of the above variational principle.

2.2 Monte Carlo Methods

Monte Carlo methods are a way of using random numbers to perform numerical integrations. By way of example consider the integral

$$I = \int_{x_1}^{x_2} f(x)dx \quad . \tag{2.6}$$

There are many quadrature methods, with varying degrees of accuracy, which can be used to evaluate this integral. The trapezium rule and Simpson's method (see "Numerical Recipes", [18]) are both quadrature methods which involve evaluating $f(x)$ at evenly spaced points, x_i, on a grid. A weighted average of these values $f(x_i)$ gives an estimate of the integral

$$I_{\text{estimate}} = (x_2 - x_1)\frac{\sum_i w_i f(x_i)}{\sum_i w_i} \quad , \tag{2.7}$$

where the w_i are the weights. The weights and the sampling points are different for different methods of quadrature but all the methods sample the function $f(x)$ using pre-determined weights and sampling points.

Monte Carlo methods do not use specific sampling points but instead we choose points at random. The Monte Carlo estimate of the integral is then,

$$\begin{aligned} I_{\text{estimate}} &= (x_2 - x_1)\frac{1}{N}\sum_{i=1} f(x_i) \\ &= (x_2 - x_1)\overline{f} \quad , \end{aligned} \tag{2.8}$$

where the x_i are randomly sampled points and $\overline{f}$ is the arithmetic mean of the values of the function $f(x)$ at the sampling points. The standard deviation of the mean is given by

$$\sigma_m = \frac{\sigma}{\sqrt{N}} \quad , \tag{2.9}$$

where

$$\sigma^2 = \frac{\sum_i [f(x_i) - \overline{f}\,]^2}{N-1} \tag{2.10}$$

gives an estimate of the statistical error in the Monte Carlo estimate of the integral. Note that the error goes as $\frac{1}{\sqrt{N}}$, independent of the dimensionality of the integral.

2.3 Importance Sampling

Monte Carlo calculations can be carried out using sets of random points picked from any arbitrary probability distribution. The choice of distribution obviously makes a difference to the efficiency of the method. In most cases, Monte Carlo calculations carried out using uniform probability distributions give very poor estimates of high-dimensional integrals and are not a useful method of approximation. In 1953, however, Metropolis et. al. [19] introduced a new algorithm for sampling points from a given probability function. This algorithm enabled the incorporation of "importance sampling" into Monte Carlo integration. Instead of choosing points from a uniform distribution, they are now chosen from a distribution which concentrates the points where the function being integrated is large. Eq.(2.6) can then be rewritten as

$$I = \int_a^b \frac{f(x)}{g(x)} g(x) dx \quad , \tag{2.11}$$

where the function $g(x)$ is chosen to be a reasonable approximation to $f(x)$. The integral can be calculated by choosing the random points from the probability distribution $g(x)$ and evaluating $f(x_i)/g(x_i)$ at these points. To enable $g(x)$ to be act as a distribution function it must be of one sign everywhere, and the best possible choice[1]

[1] The choice of $g(x) = |f(x)|$ minimises the variance of the estimate of the integral

is $g(x) = |f(x)|$. The average of these evaluations gives an estimate of I. Another way of looking at this new integral is to define $dy = g(x)d\mathrm{x}$, in which case

$$I = \int_A^B \frac{f(x(y))}{g(x(y))} dy \quad , \tag{2.12}$$

where the limits of integration are changed to correspond to the change of variable. In this case, random points are chosen from a uniform distribution in the domain $A < y < B$. The new integrand, f/g, is close to unity and so the variance (i.e. the value of σ^2 as defined in Eq.(2.10)) for this function is much smaller than that obtained when evaluating the function by sampling from a uniform distribution. Sampling from a non-uniform distribution for this function should therefore be more efficient than doing a crude Monte Carlo calculation without importance sampling.

2.3.1 The Metropolis Algorithm

The method we have used to sample points from the chosen probability distribution, $g(x)$, is the Metropolis algorithm[19]. In this algorithm, a random walk is performed through the configuration space of interest. The walk is designed so that the points on the walk are distributed according to the required probability distribution. At each point on the walk a random trial move from the current position in configuration space is selected. This trial move is then either accepted or rejected according to a simple probabilistic rule. If the move is accepted then the "walker" moves to the new position in configuration space; otherwise the "walker" remains where it is. (By a "walker" we mean a point in the $3N$-dimensional configuration space of the problem.) Another trial step is then chosen, either from the new accepted position or from the old position if the first move was rejected, and the process is repeated. In this way it should be possible for the "walker" to explore the whole configuration space of the problem. The Metropolis algorithm provides a prescription for choosing which moves in configuration space to accept or reject. Suppose that the current position on the random walk is $\mathbf{R}$, where $\mathbf{R}$ defines the positions of all the electrons in the system

$$\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_N) \quad , \tag{2.13}$$

and that the move, chosen randomly, would make the new position the point $\mathbf{R}'$. Each of these points has a number density associated with it, $N(\mathbf{R})$ and $N(\mathbf{R}')$. The number density is simply proportional to the probability distribution of that point in configuration space. If the average over many steps of the random walk follows the specified probability distribution then the walk is said to have reached equilibrium. In this case, the average number densities of points on the walk at $\mathbf{R}$ and $\mathbf{R}'$, $N(\mathbf{R})$ and $N(\mathbf{R}')$, should be constant. That means that the probability of making a transition from $\mathbf{R}$ to $\mathbf{R}'$ has to be equal to the probability of making a transition in the opposite direction, from $\mathbf{R}'$ to $\mathbf{R}$. The transition probability of a trial move being made from $\mathbf{R}$ to $\mathbf{R}'$ is denoted by $P_{\text{trial}}(\mathbf{R} \to \mathbf{R}')$ and the transition probability of a trial move being made from $\mathbf{R}'$ to $\mathbf{R}$ by $P_{\text{trial}}(\mathbf{R}' \to \mathbf{R})$. The trial moves are chosen from a fixed probability distribution around the current position and since there is nothing special about the points $\mathbf{R}$ or $\mathbf{R}'$,

$$P_{\text{trial}}(\mathbf{R} \to \mathbf{R}') = P_{\text{trial}}(\mathbf{R}' \to \mathbf{R}) \quad . \tag{2.14}$$

The probability of a trial move from $\mathbf{R}$ to $\mathbf{R}'$ being accepted is $P_{\text{accept}}(\mathbf{R} \to \mathbf{R}')$ and the probability of accepting the reverse move is $P_{\text{accept}}(\mathbf{R}' \to \mathbf{R})$. Total probabilities for moves occurring in either direction are then

$$\begin{aligned} P(\mathbf{R} \to \mathbf{R}') &= P_{\text{trial}}(\mathbf{R} \to \mathbf{R}')P_{\text{accept}}(\mathbf{R} \to \mathbf{R}') \\ P(\mathbf{R}' \to \mathbf{R}) &= P_{\text{trial}}(\mathbf{R}' \to \mathbf{R})P_{\text{accept}}(\mathbf{R}' \to \mathbf{R}) \quad , \end{aligned} \tag{2.15}$$

and the equilibrium condition implies that

$$N(\mathbf{R})P(\mathbf{R} \to \mathbf{R}') = N(\mathbf{R}')P(\mathbf{R}' \to \mathbf{R}) \quad . \tag{2.16}$$

In the Metropolis algorithm the probability of accepting the random move, $P_{\text{accept}}(\mathbf{R} \to \mathbf{R}')$, is chosen to be

$$P_{\text{accept}}(\mathbf{R} \to \mathbf{R}') = \min\left\{1, \frac{N(\mathbf{R}')}{N(\mathbf{R})}\right\} \quad . \tag{2.17}$$

This satisfies the equilibrium conditions for the distribution. The mechanism for accepting a move with probability $\frac{N(\mathbf{R}')}{N(\mathbf{R})}$, within the QMC code, is to generate a random number in the range $[0,1]$. If this random number is less than $\frac{N(\mathbf{R}')}{N(\mathbf{R})}$ then the move is accepted. If this random number is greater than $\frac{N(\mathbf{R}')}{N(\mathbf{R})}$ then the move is rejected.

The above description was for a configuration space in which any one point could be reached from any other point in one move. If the configuration space is very large, then if small moves are made it is not possible to reach any other point in one move. As long as any one point can be reached from any other point the random walk is said to be *ergodic* and the Metropolis algorithm is still valid. A straightforward extension of the above argument can be made to justify the use of the Metropolis algorithm in this situation. To ensure that points are sampled correctly from the probability distribution, the random walk has to be allowed to proceed from some arbitrary initial starting point until the average over an ensemble of moves represents the distribution to be sampled. At this point equilibrium has been reached. It is only at this point that the importance sampling Monte Carlo calculation can be carried out. It should be noted that the Metropolis algorithm is just one of a large number of such algorithms, but we have not found any reason to choose a different algorithm.

2.4 Variational Quantum Monte Carlo

The Variational quantum Monte Carlo (VMC) method is the simpler of the two quantum Monte Carlo methods used in this thesis. It is based on a combination of the ideas described in the two previous sections, namely the variational principle and Monte Carlo evaluation of integrals using importance sampling based on the Metropolis algorithm.

Within the Born-Oppenheimer approximation [4], the Hamiltonian for a many body system can be written as

$$\hat{H} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 - \sum_i \sum_\alpha \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{d}_\alpha|} + \frac{1}{2}\sum_i \sum_{j\neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_\alpha \sum_{\beta\neq\alpha} \frac{Z_\alpha Z_\beta}{|\mathbf{d}_\alpha - \mathbf{d}_\beta|} \quad . \tag{2.18}$$

The expectation value of the exact groundstate wavefunction, Ψ_0, with this Hamiltonian, is the exact groundstate energy.

$$\begin{aligned} E_0 &= \frac{\left\langle \Psi_0 \left| \hat{H} \right| \Psi_0 \right\rangle}{\langle \Psi_0 | \Psi_0 \rangle} \\ &= \int \frac{\Psi_0^*(\mathbf{R}) \hat{H} \Psi_0(\mathbf{R}) d\mathbf{R}}{\Psi_0^*(\mathbf{R}) \Psi_0(\mathbf{R}) d\mathbf{R}} \quad , \end{aligned} \tag{2.19}$$

where $\mathbf{R}$ denotes the $3N$-dimensional vector of electronic positions. The VMC method relies on one being able to construct a trial wavefunction, Ψ_T, that is a reasonably good approximation to the true groundstate wavefunction, Ψ_0. The subject of how to produce a good trial wavefunctions is dealt with in depth in chapter 4. The energy associated with the trial wavefunction is given by,

$$\begin{aligned} E_T &= \frac{\left\langle \Psi_T | \hat{H} | \Psi_T \right\rangle}{\langle \Psi_T | \Psi_T \rangle} \\ &= \int \frac{\Psi_T^*(\mathbf{R}) \hat{H} \Psi_T(\mathbf{R}) d\mathbf{R}}{\Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R}) d\mathbf{R}} \quad . \end{aligned} \tag{2.20}$$

The variational principle, described in section 2.1, ensures that the energy, E_T, is a rigorous upper bound to the true groundstate energy, E_0.

The VMC method is a Monte Carlo method for evaluating the multi-dimensional integral in Eq.(2.20). This is achieved by rewriting Eq.(2.20) in the following form,

$$E_T = \int \frac{|\Psi_T(\mathbf{R})|^2 \frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} d\mathbf{R}}{\int |\Psi_T(\mathbf{R})|^2} \quad . \tag{2.21}$$

The Metropolis algorithm is used to sample a series of points, $\mathbf{R}$, in configuration space. At each of these points the "Local Energy", $\frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}$, is evaluated. After a sufficient number of evaluations of the local energy have been made, the average is taken in the same way as in Eq.(2.7).

$$E_{VMC} = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{H}\Psi_T(\mathbf{R_i})}{\Psi_T(\mathbf{R_i})} \quad , \tag{2.22}$$

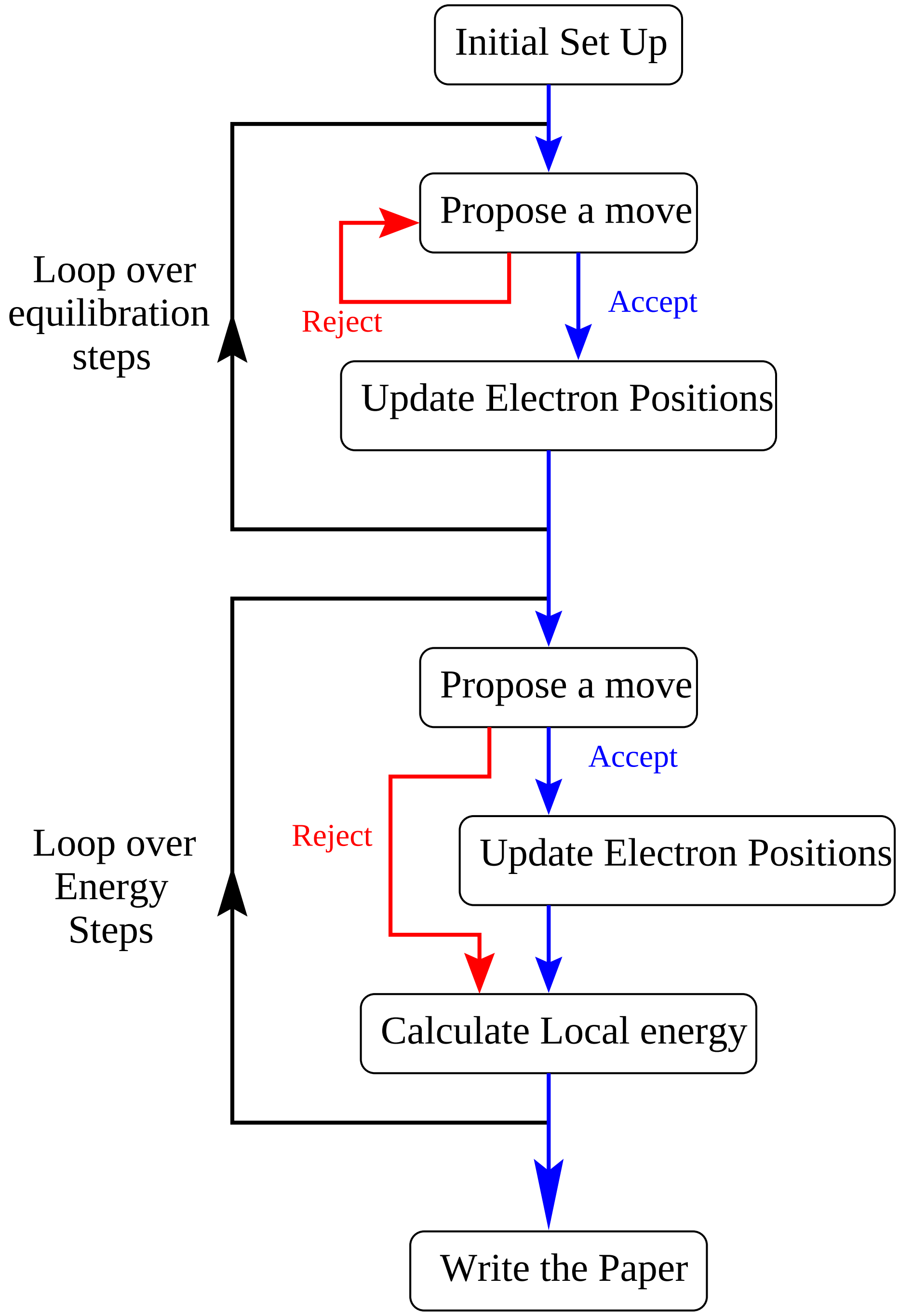


Figure 2.1: Flow chart illustrating the VMC algorithm.

where the Metropolis algorithm ensures that in the limit of large N, the $\mathbf{R}_i$ are sampled from $|\Psi_T(\mathbf{R})|^2$.

Figure 2.1 is a schematic flow chart illustrating how a typical VMC algorithm works. There are two distinct parts to the algorithm; an initial equilibration stage and an energy evaluation stage. During the initial equilibration stage, the walker is moved according to the Metropolis algorithm, but the local energy is not accumulated along the walk. This stage is required because the initial starting point of the walker is chosen randomly and therefore a set of Metropolis moves are required before the average along its walk is correctly sampling the distribution, $|\Psi_T|^2$. The required number of equilibration steps can be established by calculating the energy at each step from the beginning of the random walk and looking for the point at which there is no longer a drift in the average value of the local energy. During the energy evaluation stage, the energy of the walker is accumulated after each move. The method of evaluating the local energy within the QMC code is described in more detail in section 2.4.2. The method of updating the value of the wavefunction after each move is describing in appendix A.

2.4.1 Trial Wavefunctions

To perform a VMC calculation using the algorithm outlined in figure 2.1, one has to chose the form of the trial wavefunction, Ψ_T. This trial wavefunction should contain as much knowledge of the physics of the system being studied as possible. The choice of Ψ_T will completely determine the values of all the observables, such as the energy, obtained from the calculation.

For a bosonic system the many-body wavefunction is a symmetric function of the coordinates of the particles. McMillan, in his study of the ground state of liquid ^{4}He by the VMC method [20], used a many-body wavefunction given by a Jastrow function [21],

$$\psi_J(\mathbf{R}) = \psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) \tag{2.23}$$

$$= \exp\left[-\sum_{1\le i<j\le N} u(r_{ij})\right], \tag{2.24}$$

where $u(r)$ is a two-body function chosen to minimise the energy of the state. The function, $u(r)$, was chosen so as to enhance the probability of pairs of ^{4}He atoms being separated by a distance which minimises their interaction energy. The price to be paid for this is that the kinetic energy is increased due to the confinement, but the total energy is still reduced.

For fermionic systems the many-body wavefunction is antisymmetric under particle exchange. The simplest antisymmetric function one can choose is the Slater determinant, often referred to as the Hartree-Fock approximation.

$$\psi(\mathbf{R}) = D^{\uparrow}(\mathbf{R})D^{\downarrow}(\mathbf{R}), \tag{2.25}$$

where

$$D = \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) & \cdots & \phi_1(\mathbf{r}_{N/2}) \\ \phi_2(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) & \cdots & \phi_2(\mathbf{r}_{N/2}) \\ \vdots & \vdots & & \\ \phi_{N/2}(\mathbf{r}_1) & \phi_{N/2}(\mathbf{r}_2) & \cdots & \phi_{N/2}(\mathbf{r}_{N/2}) \end{vmatrix}. \tag{2.26}$$

For solids the single particle orbitals, ϕ_i are normally taken from either density-functional-theory, local-density-approximation calculations (DFT-LDA) or Hartree-Fock (HF) calculations. For all-electron atomic calculations the orbitals used are generally those obtained from some minimisation scheme [22]. The use of a separate determinant for up- and down-spin electrons means that the wavefunction is not antisymmetric on exchange of opposite spin electrons, however, this form gives the same expectation value as long as the operator is spin-independent[23]. The advantage of using two smaller determinants rather than one larger one is that it is computationally more efficient.

In this thesis, we adopt the definition of electron correlation as any further electron-electron interaction beyond that described by the exchange interaction in Hartree-Fock theory. According to this definition, the above form of fermionic wavefunction, Eq.(2.26), contains no correlation. In order to introduce correlation we multiply by a Jastrow factor which is symmetric under the exchange of particles, giving a wavefunction of the form

$$\psi(\mathbf{R}) = \psi_J(\mathbf{R})D^{\uparrow}(\mathbf{R})D^{\downarrow}(\mathbf{R}). \tag{2.27}$$

Two forms of the Jastrow factor are commonly used:

$$u(r) = \frac{A}{r}(1 - e^{-\frac{r}{F}}), \tag{2.28}$$

for solids, and

$$u(r) = \frac{ar}{1 + br}, \tag{2.29}$$

for atoms. The ratio of the two parameters (A/F) and the value of a are chosen such that the electron-electron "cusp" conditions [24] are obeyed, that is

$$\left.\frac{du}{dr}\right|_{r=0} = \begin{cases} -\frac{1}{2}, & \text{for opposite spins,} \\ -\frac{1}{4}, & \text{for parallel spins.} \end{cases} \tag{2.30}$$

The value of b can be chosen variationally. For solids the standard choice for fixing the remaining degree of freedom in the u function is made by considering the long-range behaviour of u [25, 26]. More optimised choices for this degree of freedom are discussed in chapter 4. For atoms this extra degree of freedom is used to either minimise the energy or the variance of the energy.

Recently more sophisticated Jastrow factors have been used. For atoms [22] this has been done by making the Jastrow factor a function of the electron-nucleus distance as well as the inter-electron distance. Similar schemes have also been implemented for solids [27, 1].

For solids it is found to be beneficial to introduce a one-body, χ function which attempts to reverse the effect of the Jastrow factor on the charge density. As the Jastrow function introduces an extra repulsion between electrons, this has the effect of smearing out the charge density. i.e. in regions where the original charge density is high, the effect of the Jastrow function is to reduce it and vice-versa. Methods for constructing the χ function and more optimised forms of Jastrow function are discussed in chapter 4.

Thus the final form of the fermionic wavefunction is

$$\psi(\mathbf{R}) = D^{\uparrow}(\mathbf{R})D^{\downarrow}(\mathbf{R}) \exp\left[\sum_{(s,i)=(\uparrow,1)}^{(\downarrow,N)} \chi_s(r_s^i) - \sum_{(\uparrow,1)\le(s,i)<(s',j)}^{(\downarrow,N)} u_{ss'}(r_{ij})\right] \quad . \tag{2.31}$$

This is referred to as the *Hartree-Fock-Jastrow-Chi* trial wavefunction.

2.4.2 Evaluating the Local Energy

A VMC calculation requires the evaluation of the local energy after moving each of the electrons. In practice, the energy is not evaluated after every move because the energies are strongly correlated from one move to the next. As the evaluation of the local energy for a given configuration is considerably more computationally expensive than the process of accepting or rejecting the set of moves to the next configuration, the local energy is only evaluated every few configurations. In our calculations the correlation length of the energy varies between 1 and 100 moves depending on the step size of the moves and the number of electrons in the system. Methods for calculating this correlation length are described in chapter 4.

The method used to calculate the local energy closely follows the work of Fahy, Wang and Louie [26]. The energy is calculated in two distinct parts: the kinetic energy and the electrostatic interaction energy. The kinetic energy is dependent upon the form of the wavefunction and the electronic positions whereas the electrostatic energy only depends upon the positions of the charges in the system.

Evaluating the Kinetic Energy

The single-particle kinetic energy operator for electron i is

$$\widehat{\mathrm{KE}}_i = -\frac{1}{2}\nabla_i^2 \quad , \tag{2.32}$$

The expected kinetic energy of electron i is therefore

$$\left\langle \widehat{\mathrm{KE}}_i \right\rangle = -\frac{1}{2}\frac{\langle\Psi|\nabla_i^2|\Psi\rangle}{\langle\Psi|\Psi\rangle} \quad . \tag{2.33}$$

This quantity is obtained using a Monte Carlo integration as described above. The Metropolis algorithm is used to sample the probability distribution $|\Psi|^2$, where Ψ is the wavefunction described in the previous section, and the estimator

$$\mathrm{KE}_i = -\frac{1}{2}\frac{\nabla_i^2\Psi}{\Psi} \tag{2.34}$$

is accumulated over the simulation to give the kinetic energy of electron i.

The calculation of $-\frac{1}{2}\frac{\nabla_i^2\Psi}{\Psi}$ is actually performed in two parts due of the form of the wavefunction being used. The trial wavefunction involves exponentials of the functions $u(r)$ and $\chi(\mathbf{r})$ which make it convenient to deal with logarithms of the wavefunction rather than differentiating the wavefunction directly. Defining

$$T_i = -\frac{1}{4}\nabla_i^2 \ln\Psi \tag{2.35}$$

and

$$F_i = \frac{1}{\sqrt{2}}\nabla_i \ln\Psi \quad , \tag{2.36}$$

then

$$2T_i - F_i^2 = -\frac{1}{2}\frac{\nabla_i^2\Psi}{\Psi} = \mathrm{KE} \quad . \tag{2.37}$$

If one considers the trial wavefunction in Eq.(2.31), introduced in the previous section, then

$$\begin{aligned}
\nabla_i \ln\Psi &= \nabla_i \ln D^{\uparrow} + \nabla_i \ln D^{\downarrow} + \nabla_i \ln[e^{-\sum_{i\neq j} u(r_{ij})}] + \nabla_i \ln[e^{\sum_i \chi(\mathbf{r}_i)}] \\
&= \frac{1}{D^{\uparrow}}\nabla_i D^{\uparrow} + \frac{1}{D^{\downarrow}}\nabla_i D^{\downarrow} - \sum_j \nabla_i u(r_{ij}) + \nabla_i \chi(\mathbf{r}_i) \quad ,
\end{aligned} \tag{2.38}$$

and

$$\nabla_i^2 \ln \Psi = -\left(\frac{1}{D^{\uparrow}}\nabla_i D^{\uparrow}\right)^2 + \frac{1}{D^{\uparrow}}\nabla_i^2 D^{\uparrow} - \left(\frac{1}{D^{\downarrow}}\nabla_i D^{\downarrow}\right)^2 + \frac{1}{D^{\downarrow}}\nabla_i^2 D^{\downarrow} - \sum_j \nabla_i^2 u(r_{ij}) + \nabla_i^2 \chi(\mathbf{r}_i) \quad . \tag{2.39}$$

T_i and F_i^2 are calculated from these equations at each step in the random walk. The kinetic energy as given by Eq.(2.37), is also calculated at each step and averages of all three quantities are found at the end of the simulation. The consistency of these three is checked using Green's relation, which shows that

$$< \mathrm{KE} >=< T_i >=< F_i^2 > \tag{2.40}$$

for all properly sampled wavefunctions. This consistency check is extremely useful when debugging a QMC code. If either the first or second derivative of the wavefunction has been calculated incorrectly, this will immediately show up in this consistency check and it is often clear which of the derivatives is being evaluated wrongly. The variances of $< T_i >$ and $< F_i^2 >$ are both much greater[26, 23] than the variance of the kinetic energy as given by Eq.(2.37), therefore it is this quantity which is used to estimate the kinetic energy in Monte Carlo calculations.

Electrostatic Energies

The three remaining terms in the many-body Hamiltonian are the electron-electron, electron-ion and ion-ion electrostatic interactions. When studying atomic and molecular systems, the evaluation of these terms simply involves summing up all the pairwise interactions present.

In a solid system, the situation is complicated by the use of periodic boundary conditions. Now the sum of pairwise interactions includes the infinite number of periodic images of each particle. The methods used to deal with these sums over periodic images are described in the following chapter, which deals with the details of performing QMC calculations on solids.

2.4.3 Accumulating Averages

In the flow chart of the VMC algorithm, (figure 2.1), it is simply stated that the energy of the walker is accumulated after moving each of the electrons. In fact we choose a slightly more complicated formula for updating each of the quantities being calculated[26]. After each proposed move, whether it is rejected or not, each quantity $< Q >$ is updated such that

$$< Q >= \frac{1}{\text{Total No. Moves}} \left[\, pQ_i(\mathbf{R}_{new}) + (1-p)Q_i(\mathbf{R}_{old}) \, \right] \quad , \tag{2.41}$$

where Q_i is any quantity of interest such as the kinetic energy, potential energy or total energy associated with particle i, and p is the probability of accepting the move from $\mathbf{R}_{old}$ to $\mathbf{R}_{new}$. It is possible simply to accumulate $Q_i(\mathbf{R})$ at just the new points on the walk, but the combination in Eq.(2.41), of values at the old and new points allows information about points which are rejected to be included and reduces the contribution from “unlikely” moves which are accepted. By this means, the variance of the expectation value of Q is reduced. It has been shown in Ref.[26] that the accumulation of $pQ_i(\mathbf{R}_{new}) + (1-p)Q_i(\mathbf{R}_{old})$ gives a correct sampling of the probability density, $|\Psi(\mathbf{R})|^2$. This can be demonstrated by considering each term, $pQ_i(\mathbf{R}_{new})$, and $(1-p)Q_i(\mathbf{R}_{old})$, separately and calculating the probability distribution which each term samples. By adding together these two probability distributions, it is shown that the combination of the two terms does indeed sample from the correct total probability distribution.

$Q_i(\mathbf{R})$ is the energy of particle i when the configuration is at $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N)$ The probability of being in configuration $\mathbf{R}$ and evaluating the energy of particle i, i.e. the probability of arriving at configuration $\mathbf{R}$ as a result of making a move of particle i to r_i, should obviously be $|\Psi(\mathbf{R})|^2/N$ if the sampling is being done correctly.

At the old position of electron i, $\mathbf{R}_{old} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{i,\text{old}}, \ldots, \mathbf{r}_N)$, electron i can move anywhere within the range of the maximum step size, i.e. within a volume V. The probability of arriving at $\mathbf{r}_i$ from $\mathbf{r}_{i,old}$ is $P_{trial}(\mathbf{r}_{i,old} \rightarrow \mathbf{r}_i)$. The probability of evaluating $Q_i(\mathbf{r})$ after accepting a move $\mathbf{r}_{i,old} \rightarrow \mathbf{r}_i$ is therefore

$$\begin{aligned} & \int_V d\mathbf{r}_{i,old} \, |\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{i,old}, \ldots, \mathbf{r}_N)|^2 P(\mathbf{r}_{i,old} \rightarrow \mathbf{r}_i) \\ = & \frac{1}{N} \int_V P_{trial}(\mathbf{r}_{i,old} \rightarrow \mathbf{r}_i) |\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{i,old}, \ldots, \mathbf{r}_N)|^2 \\ \times & \min\left\{1, \frac{|\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N)|^2}{|\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{i,old}, \ldots, \mathbf{r}_N)|^2}\right\} d\mathbf{r}_{i,old} \quad . \end{aligned} \tag{2.42}$$

The probability of being at $\mathbf{R}$ and rejecting any move $\mathbf{r}_i \rightarrow \mathbf{r}'$ is

$$\begin{aligned} & |\Psi(\mathbf{R})|^2 \frac{1}{N} \left\{1 - \int_V P(\mathbf{r}_i \rightarrow \mathbf{r}') d\mathbf{r}'\right\} \\ = & \frac{1}{N} |\Psi(\mathbf{R})|^2 - \frac{1}{N} \int_V P(\mathbf{r}_i \rightarrow \mathbf{r}') \\ \times & \min\left\{|\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N)|^2, |\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}', \ldots, \mathbf{r}_N)|^2\right\} d\mathbf{r}' \quad . \end{aligned} \tag{2.43}$$

The sum of these two probabilities, i.e. accumulating $[pQ_i(\mathbf{R}_{new}) + (1-p)Q_i(\mathbf{R}_{old})]$ as in Eq.(2.41), gives the probability density $|\Psi(\mathbf{R})|^2/N$, as required.

2.4.4 Performing VMC calculations of Parallel Computers

Monte Carlo calculations are inherently parallel in nature. At their most basic, they involve the calculation of a large set of independent random numbers and then the averaging over a set of results produced by each of these random numbers. Coupled with the fact that QMC calculations are relatively expensive to perform on todays workstation computers, they are an ideal candidate for porting to parallel architecture machines, which offer two to three orders of magnitude more computational power than a conventional workstation.

All the VMC calculations reported on in this thesis have been performed on the Cray T3D and Hitachi SR2001 Massively Parallel Processing (MPP) machines. Only a few changes to the general VMC algorithm outlined above are required to produce a VMC algorithm that will make efficient use of an MPP machine.

The modified parallel algorithm is illustrated in figure 2.2. Each processing element on the parallel machine runs its own version of the VMC algorithm. These are made independent of each other by starting the random number generator with a different seed on each node. Care must be taken to ensure that the period of the random number generator is sufficiently long that the sequences of random numbers on each of the nodes do not overlap. This ensures that by the time the equilibration process has finished, the positions of the electrons on each node are not correlated. Each node then accumulates its own set of observables such as the total energy and charge density. At the end of the calculation, these observables are all passed to one node that simply adds them up and takes the mean.

It is not strictly necessary to run through the equilibration stage on each of the nodes. A single node calculation could be used to produce a set of equilibrated configurations of electron positions that could then be used as a starting point for the energy evaluation stage of the process on each of the nodes. In practice, the equilibration stage of the calculation is only a small fraction of the total time and this was therefore not deemed necessary.

The time overhead required to perform the communication between processors for accumulating averages etc. is negligible. Therefore, the parallel efficiency of the algorithm is effectively 100%. This means that on parallel machines with a few hundred processing elements, the algorithm achieves almost linear scaling, i.e. if one doubles the number of processors on which the code is running, the overall wall-clock time required to perform a particular VMC calculation to within a certain specific statistical accuracy is halved.

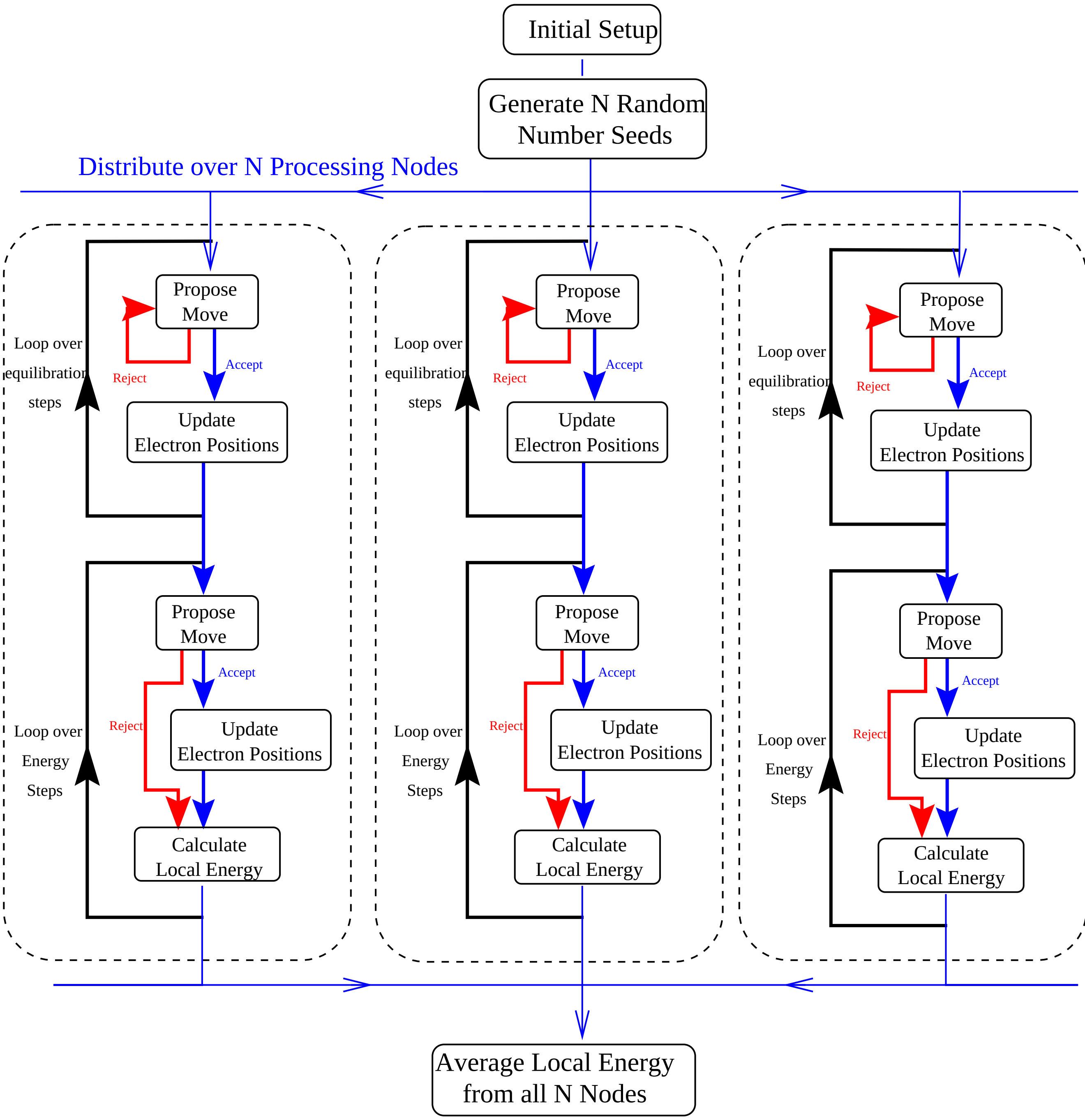


Figure 2.2: Flow chart illustrating the Parallel VMC algorithm. Each individual node is enclosed within the dashed boxes. Only three of the N nodes are shown in the diagram. The initial random seeds are generated on one node and then distributed out to all the other N-1 processing nodes. At the end of each block of energy evaluation moves, averages are taken across all the processors.

2.5 Diffusion Quantum Monte Carlo

Diffusion quantum Monte Carlo (DMC) is, in principle, an exact method for solving the Schrödinger equation for the ground-state of a many-body system; though in practice it is found that there are some approximations that have to be made.

2.5.1 The Method

The basis of DMC is to write the Schrödinger equation in imaginary time, taking

$$i\frac{\partial|\psi\rangle}{\partial t} = \hat{H}|\psi\rangle \tag{2.44}$$

$$\text{and letting} \quad \tau = it$$

$$\text{giving} \quad \frac{\partial|\psi\rangle}{\partial\tau} = -\hat{H}|\psi\rangle. \tag{2.45}$$

$|\psi\rangle$ can be expanded as:

$$|\psi\rangle = \sum_i c_i|\phi_i\rangle, \tag{2.46}$$

where

$$\hat{H}|\phi_i\rangle = \epsilon_i|\phi_i\rangle. \tag{2.47}$$

The $|\phi_i\rangle$'s and ϵ_i's are the energy eigenvectors and eigenvalues respectively of the time-independent Schrödinger equation.

We can write down the formal solution of the imaginary-time Schrödinger equation, Eq.(2.45),

$$|\psi(\tau_2)\rangle = e^{-\hat{H}(\tau_2-\tau_1)}|\psi(\tau_1)\rangle, \tag{2.48}$$

noting that $e^{-\hat{H}\tau}$ is the imaginary-time evolution operator. Furthermore, expressing $|\psi\rangle$ in the form of Eq.(2.46) implies that

$$|\psi(\tau_2)\rangle = \sum_i c_i e^{-E_i(\tau_2-\tau_1)}|\phi_i\rangle. \tag{2.49}$$

Letting $|\psi(\tau_1)\rangle$ be $|\psi(0)\rangle$ and hence $\tau_2 = \tau$, the evolution time for the system implies that, as long as $|\psi(0)\rangle$ is not orthogonal to the ground-state, $|\phi_0\rangle$, then

$$\lim_{\tau\to\infty} |\psi(\tau)\rangle = c_0 e^{-\epsilon_0\tau}|\phi_0\rangle, \tag{2.50}$$

where ϵ_0 is the ground-state energy. This can be seen by remembering that all other states have higher energies, ϵ_i, than the groundstate energy ϵ_0, and will therefore decay away faster. In the **R** representation, Eq.(2.50) becomes:

$$\lim_{\tau\to\infty} \psi(\mathbf{R},\tau) = e^{-E_0\tau}c_0\phi_0(\mathbf{R}), \tag{2.51}$$

where

$$\psi(\mathbf{R},\tau) = \langle\mathbf{R}|\psi(\tau)\rangle \quad \text{and} \quad \phi_i(\mathbf{R},\tau) = \langle\mathbf{R}|\phi_i(\tau)\rangle. \tag{2.52}$$

We now introduce an arbitrary energy offset term E_T, such that the imaginary-time Schrödinger equation, Eq.(2.45), is recast as:

$$\frac{\partial\psi(\mathbf{R},\tau)}{\partial\tau} = -\hat{H}\psi(\mathbf{R},\tau) + E_T\psi(\mathbf{R},\tau). \tag{2.53}$$

Then if E_T is adjusted to be the true ground-state energy, ϵ_0, the asymptotic solution is a steady-state solution.

We now use the fact that the Schrödinger equation in imaginary time looks like the diffusion equation. Explicitly writing out the Hamiltonian in Eq.(2.53), gives:

$$\frac{1}{2}\nabla^2_{\mathbf{R}}\psi(\mathbf{R},\tau) - V(\mathbf{R})\psi(\mathbf{R},\tau) + E_T\psi(\mathbf{R},\tau) = \frac{\partial\psi(\mathbf{R},\tau)}{\partial\tau}. \tag{2.54}$$

This is just a $3N$-dimensional diffusion equation, with $\psi(\mathbf{R},\tau)$ playing the role of the density of diffusing particles. The $[E_T - V(\mathbf{R})]\psi(\mathbf{R},\tau)$ term is a rate term, and describes the branching (or creation/annihilation) processes. The entire equation can be simulated by a combination of a diffusion and branching processes, in which the number of diffusing particles increases or decreases at a given point proportional to the density of diffusers and the potential energy at that point in configuration space.

It turns out that solving Eq.(2.54) this way is a very inefficient way to simulate the Schrödinger equation on a computer. This is because the branching rate, which is proportional to $V(\mathbf{R})$, can diverge to $\pm\infty$ for systems of particles interacting via the Coulomb interaction. This leads to large fluctuations in the number of diffusing particles which leads to a large variance in the estimate of the energy. These fluctuations can be dramatically reduced by the introduction of importance sampling [28] in a similar way to the implementation in the VMC algorithm (see section 2.3).

We will follow the scheme of Ref.[29] for the introduction of importance sampling. The first step is to introduce a guiding function, $\psi_G(\mathbf{R})$. We now define a new distribution $f(\mathbf{R},\tau) = \psi_G(\mathbf{R})\psi(\mathbf{R},\tau)$ which, if $\psi(\mathbf{R},\tau)$ satisfies the Schrödinger equation, is also a solution of the Schrödinger equation. Substituting $f(\mathbf{R},\tau)$ into Eq.(2.54) yields

$$-\frac{1}{2}\nabla^2 f(\mathbf{R},\tau) + \nabla.[\mathbf{F}(\mathbf{R})f(\mathbf{R},\tau)] - S(\mathbf{R})f(\mathbf{R},\tau) = -\frac{\partial f(\mathbf{R},\tau)}{\partial\tau}. \tag{2.55}$$

Where $\mathbf{F}(\mathbf{R})$ can now be interpreted as a "quantum force". We follow [30] in defining this term as

$$\mathbf{F}(\mathbf{R}) = \frac{\nabla\psi_G(\mathbf{R})}{\psi_G(\mathbf{R})}, \tag{2.56}$$

and $S(\mathbf{R})$, the *branching* term as

$$S(\mathbf{R}) = E_T - E_L(\mathbf{R}), \tag{2.57}$$

which is defined in terms of the local energy of the *guiding function*

$$E_L(\mathbf{R}) = \frac{\hat{H}\psi_G(\mathbf{R})}{\psi_G(\mathbf{R})} \quad . \tag{2.58}$$

We now have a drift-diffusion equation for $f(\mathbf{R}, \tau)$. The branching term is proportional to the "excess local energy", $|E_L(\mathbf{R}) - E_T|$, which with a good choice of ψ_G need not become singular when $V(\mathbf{R})$ does. To control branching we need to choose ψ_G such that E_L is everywhere as smooth as possible, i.e. we want as little variance as possible in $(E_L(\mathbf{R}) - E_T)$. Methods for optimising trial wavefunctions with exactly this property are described in detail in chapter 4. In general, the trial wavefunction, Eq.(2.31), used as the input to a VMC calculation, makes a suitable choice of guiding wavefunction.

As well as reducing the fluctuations in the number of diffusing particles, ψ_G also has another important role for fermionic systems. It determines the position of the nodes of the final wavefunction, due to the necessity of using the fixed-node approximation where the nodal structure of the exact groundstate wavefunction is assumed to be the same as the nodal structure of ψ_G to ensure that f is always of the same sign (see section 2.5.3). The accuracy of the position of the nodes in ψ_G therefore determines how good the estimate of the ground-state energy, E_0, is. This can be seen by considering the fact that at long (imaginary) times the distribution $f(\mathbf{R}, \tau)$ approaches $\psi_G(\mathbf{R})\phi_0(\mathbf{R})$, up to the constraint (imposed by the fixed-node approximation) that $\phi_0(\mathbf{R})$ must vanish at the nodes of $\psi_G(\mathbf{R})$. This implies that the long-time limit is the true fermionic ground-state if and only if the nodes of $\psi_G(\mathbf{R})$ correspond to the exact nodes of the ground-state wavefunction. The *fixed-node energy* is an upper bound to the exact fermionic energy[31]. We simulate the equation within only a small number of nodal regions. Each walker moves within one nodal region and rejects all moves that attempt to cross the nodal surface into another nodal region. This contradicts the requirement that we stipulated earlier that the walk must be ergodic. The tiling theorem [32, 33], however, states that the nodal regions of the true ground-state eigenfunction of a system of identical fermions are all related by permutation symmetry. Furthermore the nodal regions of the determinant of LDA eigenfunctions are also related by permutation symmetry because the wavefunction is the ground-state of a Hamiltonian, although it is not the many-body Hamiltonian of the system we are interested in. This means that we can simulate the equation within one nodal region and still be guaranteed to obtain the "best" variational result.

Eq.(2.55) can be written in integral form, in doing this we follow the procedure of Ref. [30]:

$$f(\mathbf{R}',\tau_0+\tau) = e^{\tau\times E_T(\tau_0+\tau)}\int \tilde{G}(\mathbf{R},\mathbf{R}',\tau)f(\mathbf{R},\tau_0)d\mathbf{R}, \tag{2.59}$$

where $\tilde{G}$ is the Green's function for the case $E_T = 0$. The energy shift $E_T(\tau_0+\tau)$ plays the role of an arbitrary time-dependent renormalisation, chosen in such a way that the probability distribution f remains finite and non-vanishing in the limit $\tau \to \infty$.

The three terms on the left-hand side of Eq.(2.55) describe respectively, diffusion, drift and growth/decay. An approximate Green's function can be formed, with an error of $\mathcal{O}(\tau^2)$ for small τ, by the product of Green's functions for diffusion, drift and growth/decay[34]:

$$\begin{aligned}\tilde{G}(\mathbf{R}',\mathbf{R},\tau) \;&=\; \frac{1}{(2\pi\tau)^{\frac{3N}{2}}}\int e^{-\frac{(\mathbf{R}'-\mathbf{R}'')^2}{2\tau}}\delta\left(\mathbf{R}''-\mathbf{R}-\mathbf{F}(\mathbf{R})\tau\right)\\ &\quad\times\, e^{-\frac{1}{2}[E_L(\mathbf{R}')+E_L(\mathbf{R})]\tau}d\mathbf{R}''+\mathcal{O}(\tau^2).\end{aligned} \tag{2.60}$$

In order to deal with the nodes in the fermion ground-state we must use the fixed-node approximation, as stated earlier (this and the various schemes to try to improve upon it are discussed in section 2.5.3). What this actually entails is that if a move is such that a walker would cross the node then it is immediately rejected. In other words, the node acts as an infinite potential barrier. It should be noted that the use of importance sampling does not introduce any extra approximations beyond the fixed node approximation already mentioned. With the exception of the fixed node approximation it is possible to calculate exact energies (and expectation values of any other operator that commutes with the Hamiltonian, $\hat{H}$) from the distribution f, using

$$E = \frac{\int f \frac{\hat{H}\Psi_T}{\Psi_T}}{\int f} = \frac{\int \Psi \hat{H} \Psi_T}{\int \Psi_T \Psi} = E_0 \quad . \tag{2.61}$$

2.5.2 A DMC Algorithm

The algorithm used in the DMC calculations reported in this thesis for solving Eq.(2.59) is illustrated in figure 2.3. It can be divided into the following steps:

1. Initialise a set of N_c configurations (typically 100-500), with configurations distributed according to a probability density $f(\mathbf{R}, 0)$, given by $|\psi_G(\mathbf{R})|^2$. These are generated by choosing configurations produced in a VMC run. The configurations are taken sufficiently far apart in the VMC run so as to ensure their independence, using the procedure described in section 4.2.6.

2. Each configuration in the list is taken in turn. The electrons in the j^{th} configuration are then moved, one at a time, by letting them diffuse independently for a time τ. If the current electron is the i^{th} electron in configuration j, it is moved according to

$$\mathbf{r}_i'^{(j)} = \mathbf{r}_i^{(j)} + \tau \mathbf{F}(\mathbf{r}_i^{(j)}) + \eta, \tag{2.62}$$

   where $\mathbf{r}_i^{(j)}$ is the three-dimensional coordinate of the electron being moved, and η is a three-dimensional Gaussian random variable with a mean of zero and a variance of τ. The time step τ is chosen to be small enough so that the time step error[30], introduced by the approximation to the Green's function in Eq.(2.60), is less than the statistical uncertainty. Usually τ is chosen so that the rate of acceptance of moves is $\geq 99\%$.

3. The fixed-node approximation is now applied. We check to see whether the move has caused the walker to cross the nodal surface. If this has occurred we reject the move and go to the next electron on the list.

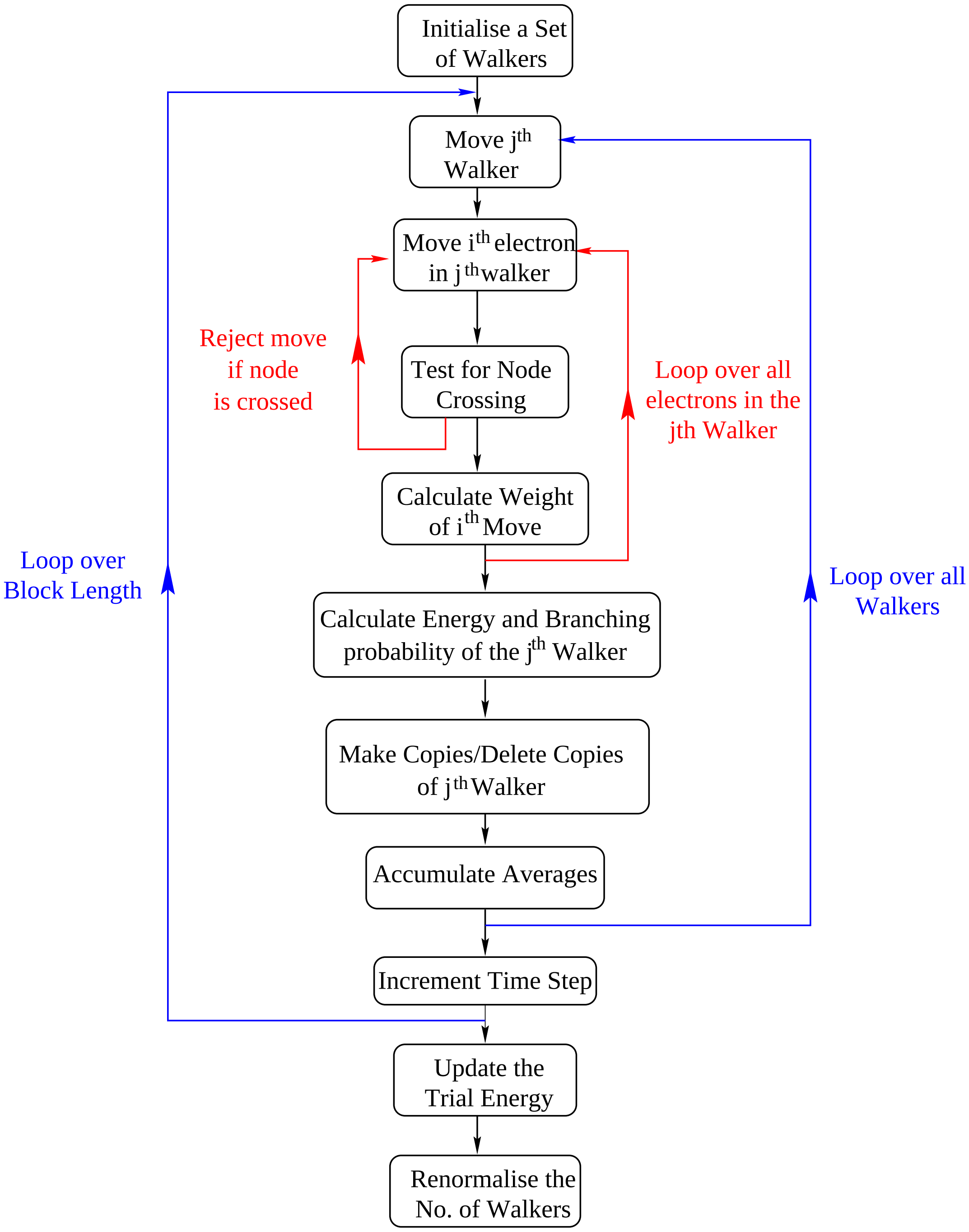


Figure 2.3: Flow chart illustrating the DMC algorithm.

4. Now calculate a weight for the move,

$$W(\mathbf{R}', \mathbf{R}) = \frac{|\psi_G(\mathbf{R}')|^2 \tilde{G}(\mathbf{R}', \mathbf{R}, \tau)}{|\psi_G(\mathbf{R})|^2 \tilde{G}(\mathbf{R}, \mathbf{R}', \tau)}. \tag{2.63}$$

   The move is then accepted with probability

$$P(\mathbf{R} \to \mathbf{R}', \tau) = min(1, W(\mathbf{R}', \mathbf{R})). \tag{2.64}$$

5. After all the electrons in the j^{th} configuration have been moved the time associated with this configuration is advanced by one time step, τ.

6. Calculate the branching probability for configuration j, using:

$$P_{B_j} = \exp(-\tau(\frac{1}{2}[E_L(\mathbf{R}') + E_L(\mathbf{R})] - E_T)). \tag{2.65}$$

   Then make

$$\overline{P}_{B_j} = int(P_{B_j} + \zeta) \tag{2.66}$$

   copies of the walker, where ζ is a uniform random number in the range [0,1].

7. If $\overline{P}_{B_j} \neq 0$ then weight $E_L(\mathbf{R}')$ and any other accumulated quantities by the branching probability, P_{B_j}. This branching probability is a measure of the importance weight of the walker.

8. E_{REF} is an adjustable energy used to stabilise the population of walkers. E_{REF} is recalculated according to

$$E_{\text{REF}} = E_T - \frac{C_{\text{EREF}}}{\tau} \log\left(\frac{N_w}{N_c}\right). \tag{2.67}$$

   Where C_{EREF} is a parameter adjusted so as to control population fluctuations. N_w is the number of configurations at the present time and N_c is the target number of configurations.

9. Repeat steps (2) through to (8) for all the configurations for a certain number of steps. Note that the total number of steps is divided up into blocks. The number of steps in each block is chosen such that the averages accumulated over the blocks of steps are statistically independent. This block length usually corresponds to diffusing each walker for a time of order a few atomic units.

10. Calculate the weighted mean of $E_L(\mathbf{R}')$ and all other desired averages. It is not necessary to use the more complicated scheme for updating the averages (section 2.4.3) used in VMC as the probability of accepting a given move is much higher (99%) due to the smaller time step used in DMC calculations.

11. Use the cumulative estimate of $E_L(\mathbf{R}')$ to update the trial energy E_T. We do this according to $E_T(\text{new}) = [E_T(\text{old}) + \langle E_L \rangle]/2$, where $< E_L >$ is the block average of the energy. Note, the initial value of E_T is usually taken as the VMC energy for the same system.

12. Renormalise the number of walkers to N_c; do this by randomly removing walkers ($N_w > N_c$), or creating copies of existing walkers ($N_w < N_c$), where N_c is the desired number of walkers.

13. Repeat steps (2) to (12) until there is no longer a detectable trend in the block average of $E_L(\mathbf{R}')$, i.e. the components of excited eigenstates in the wavefunction have effectively decayed away as shown in Eq.(2.48). At this point the steady-state has been reached.

14. Repeat steps (2) through to (12) until the variance of the cumulative averages has reached the desired level.

At the end of a DMC calculation the average values of E_{REF}, E_T, and E_L should be the same to within error bars.

The output from an example DMC calculation is shown in figure 2.4. It illustrates the initial exponential decay in the energy, taking place over the first 150 or so time steps, as the contribution to the distribution, f, from the excited states decays with time. When the graph levels out, the energy is oscillating due to statistical noise. Averages for observable quantities would be taken over the final 800 steps.

It is useful to note that if the weights in step (4) are set to

$$W(\mathbf{R}', \mathbf{R}) = \frac{|\psi_G(\mathbf{R}')|^2}{|\psi_G(\mathbf{R})|^2} \quad , \tag{2.68}$$

as in VMC, the drift term is removed, and the branching probabilities in step (7) are set to unity for all configurations, the algorithm reduces back to the VMC algorithm already described. This not only highlights the many similarities between the two algorithms, but also provides a useful method of checking whether a DMC algorithm is working correctly as it is can then be directly compared with known VMC results.

2.5.3 The Fixed-Node Approximation

The attraction of the DMC method is tempered by the nodal problem in many-fermion systems. The interpretation of the imaginary-time Schödinger equation as a diffusion equation rests upon the interpretation of the wavefunction as a probability density. The problem is that a probability density is positive by definition and corresponds to the concentration of the diffusing walkers.

The fact that the ground-state of the many-fermion system has nodes implies that some kind of external constraint must be assumed. We will now discuss such methods.

The first solution to the nodal problem was proposed by Anderson [35]. He took a wavefunction as a reference function then made the nodes of the reference function act as sinks for the walkers. This method is known as the fixed node approximation. The problem with the method is that if the nodal surfaces of the reference function do not coincide with the exact ones, a small bias is introduced in the diffusion process.

When we introduced the fixed-node approximation before we defined it slightly differently to above. This is because it has been shown [30] that the procedure of deleting walkers that cross a nodal surface introduces a bias proportional to the time step.

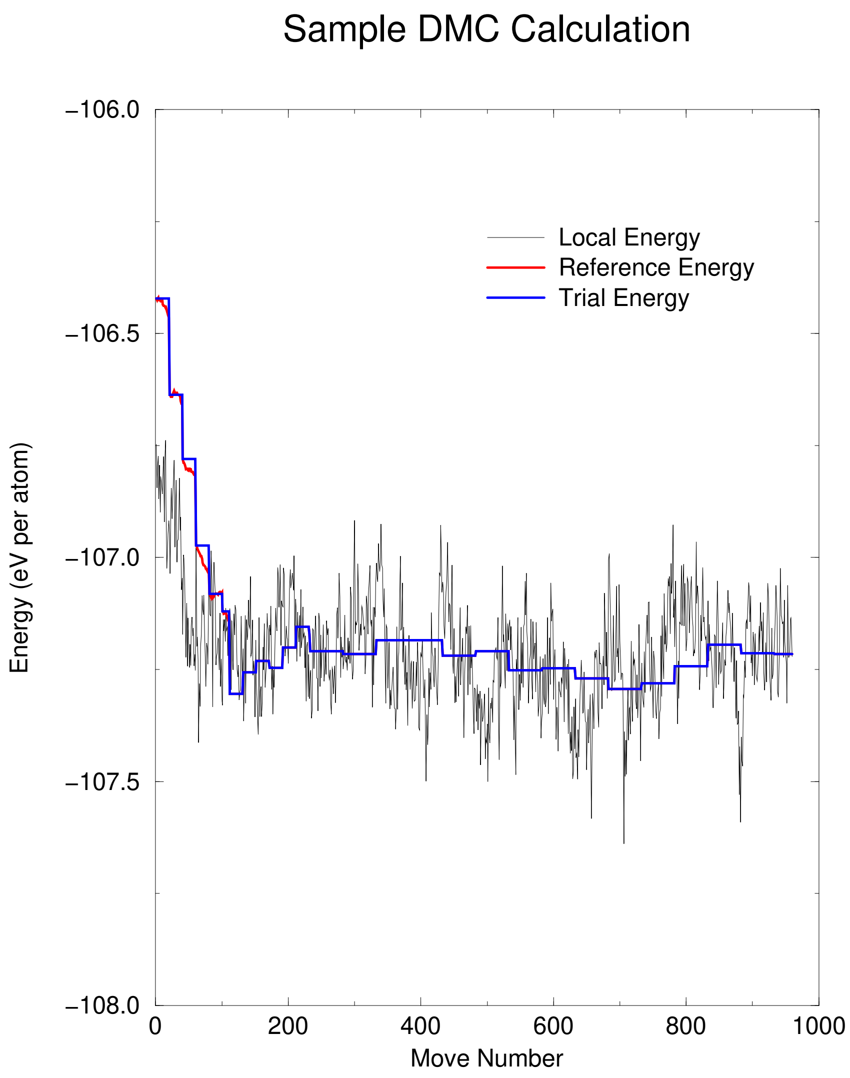


Figure 2.4: An example of results from a DMC calculation. The local energy of the guiding function is drawn in black. The reference energy is shown in red and the trial energy in blue. The reference energy is set equal to the trial energy after the initial energy propagation stage to avoid population control errors [30].

Instead the procedure that we use, and indeed the one that is now most generally used, is to reject moves that cross a node; that is a node acts as an infinite potential barrier. This method still introduces a bias but it is of the same order, second order in the time step, as the bias due to the finite time step Green function.

One method of refining this technique, proposed by Ceperley *et al.* [31] is to "release" the nodes. The antisymmetric wavefunction is then obtained via the difference of two populations of signed walkers, generated within the fixed node approximation, in the different regions corresponding to the positive and negative ones of a reference function. The problem with this is the population of the two sets of the walkers grows exponentially. This leads to large statistical noise and in general makes the technique very difficult to apply.

An alternative method for improving upon the fixed-node approximation was introduced for calculations on the two-dimensional electron gas by Kwon et al. [36]. They replace the positions of the orbitals in the Slater determinant by their *quasiparticle* coordinates, $\mathbf{x}_i$, given by

$$\mathbf{x}_i = \mathbf{r}_i + \sum_{j \neq i}^{N} \eta(r_{ij})(\mathbf{r}_i - \mathbf{r}_j) \quad , \tag{2.69}$$

where $\eta(r_{ij})$ is a *backflow* correlation function parametrised as

$$\eta(r) = a \frac{1 + br}{c + dr + r^{7/2}} \quad . \tag{2.70}$$

The idea of *backflow* was originally suggested by Feynman and Cohen[37]. By changing the coordinates of the particles for which the Slater determinant is being evaluated, one is effectively changing the nodal structure of the determinant. Kwon et al. found that the introduction of the backflow correction produced a significant reduction in the total energy of the two-dimensional electron gas at high densities. It is expected that the improvement will not be so large in three-dimensional systems, although this is yet to be fully tested.

2.5.4 DMC With Non-Local Pseudopotentials

The evaluation of non-local pseudopotentials within the QMC solid codes is dealt with in detail in chapter 3. Optimising wavefunctions with respect to these non-local pseudopotentials is discussed in chapter 4. In chapters 5 and 6, DMC calculations involving a non-local pseudopotential to describe silicon are reported. The use of such non-local pseudopotentials within DMC is a problem that is yet to be completely resolved. Here we discuss the approximations used later in the thesis to try and deal with this problem.

Consider Eq.(2.53), in which a pseudopotential is used to describe the valence electrons in a system;

$$\frac{\partial \psi(\mathbf{R},\tau)}{\partial \tau} = -\hat{H}_{\mathrm{val}}\psi(\mathbf{R},\tau) + E_T\psi(\mathbf{R},\tau). \tag{2.71}$$

We can divide the full valence Hamiltonian, $\hat{H}_{\mathrm{val}}$, into a local and non-local part

$$\hat{H}_{\mathrm{val}} = \hat{H}_{\mathrm{loc}} + W \quad , \tag{2.72}$$

where $\hat{H}_{\mathrm{loc}}$ includes the kinetic energy, the local part of the pseudopotential and the Coulomb interaction, and W includes the non-local part of the pseudopotential. The diffusion-drift equation for f, Eq.(2.55), can be split into two terms representing the local and non-local parts of the Hamiltonian

$$-\frac{1}{2}\nabla^2 f(\mathbf{R},\tau)+\nabla.[\mathbf{V}(\mathbf{R})f(\mathbf{R},\tau)]-\left(\frac{(E_T-H_{\mathrm{loc}})\Psi_G}{\Psi_G}\right)f(\mathbf{R},\tau)+\left(\frac{W\Psi}{\Psi}\right) = -\frac{\partial f(\mathbf{R},\tau)}{\partial \tau} \quad . \tag{2.73}$$

The first three terms on the left hand side of Eq.(2.73) can be interpreted as a local diffusion, drifting and branching process. However, the fourth term represents the operation of the non-local pseudopotential on the unknown wavefunction, Ψ, producing non-local branching. In the DMC calculations described in chapters 5 and 6, we have used the "locality approximation" which was introduced by Christiansen[38, 39, 40] and recently applied by Mitas[41, 42] to the problem of the unknown wavefunction in Eq.(2.73). In this approximation, the non-local pseudopotential acts not on the unknown wavefunction, Ψ, but on the guiding wavefunction, Ψ_G,

$$-\frac{1}{2}\nabla^2 f(\mathbf{R},\tau)+\nabla.[\mathbf{V}(\mathbf{R})f(\mathbf{R},\tau)]-\left(\frac{(E_T-H_{\rm loc})\Psi_G}{\Psi_G}\right)f(\mathbf{R},\tau)+\left(\frac{W\Psi_G}{\Psi_G}\right)=-\frac{\partial f(\mathbf{R},\tau)}{\partial\tau} \quad . \tag{2.74}$$

One is then free to impose the fixed-node approximation in the same way as before. In fact, when using the above local model potential, it is important that the fixed node approximation is applied as this steers the random walk away from the nodes of Ψ, where there will generally be divergences in the local model potential. Without the fixed node approximation, these divergences would cause large fluctuations in the population of diffusing particles.

It should be noted that within the "locality approximation", it is no longer generally true that the DMC estimate of the energy is an upper bound to the true groundstate energy. It must always be less than the VMC energy and it has been shown, [42] that the DMC energy converges quadratically to the exact groundstate energy as the guiding function Ψ_G approaches the true groundstate wavefunction.

2.5.5 Performing DMC Calculations on Parallel Computers

The argument for performing DMC calculations on a parallel computer is even more compelling than for VMC calculations, because DMC calculations require approximately an order of magnitude more CPU time than the equivalent VMC calculation.

As with VMC calculations, the DMC algorithm is intrinsically parallel. In the algorithm outlined in section 2.5.1, an ensemble of walkers is used to evaluate the local energy of the guiding function, Ψ_G, at each time step of the simulation. In our parallel version of the DMC algorithm, this ensemble of walkers is distributed across all $NNODES$ nodes of the parallel machine. Each node is responsible for performing stages (2)-(8) of the algorithm (the diffusion, drift and creation/annihilation of walkers) on its own subset of the total ensemble of walkers.

After all the walkers have been advanced for a block of time steps, the mean energy across all the walkers on all the nodes is used to update the trial energy as in stage (11) of the algorithm.

$$E_T^{\mathrm{NEW}} = \frac{1}{2}\left[E_T^{\mathrm{OLD}} + \frac{1}{NNODES}\sum_{i=1}^{NNODES} < E_L >_i\right] \quad , \tag{2.75}$$

where $< E_L >_i$ is the accumulated local energy of the subset of walkers on the ith node.

The renormalisation of the number of walkers at the end of a block is performed across all the nodes in the following way.

1. The number of 'live' walkers summed across all the nodes, N_w, at the end of the block is compared with the original number, N_c.
2. If there are too many walkers then $(N_w - N_c)$ walkers are selected at random and deleted. If there are too few walkers then $(N_c - N_w)$ walkers are chosen at random and copied.
3. The walkers are the redistributed across the nodes so that there are $N_c/NNODES$ walkers on each node at the start of the next block.

It is important to try and keep the number of walkers on each node equal, i.e. to 'load balance' the algorithm efficiently. The efficiency of the algorithm at any one time step is determined by

$$\text{efficiency} = \frac{\mathrm{NCONFIGS}_{\mathrm{total}}}{\mathrm{NCONFIGS}_{\mathrm{max}} * \mathrm{NNODES}} \tag{2.76}$$

where $\mathrm{NCONFIGS}_{\mathrm{total}}$ is the total number of walkers across all the nodes and $\mathrm{NCONFIGS}_{\mathrm{max}}$ is the number of walkers on the node which at that particular time step has the largest number of 'live' walkers. The efficiency of the algorithm can therefore be improved in two ways,

1. Keep the number of walkers as well 'balanced' across the nodes as possible.

2. Increase the average number of walkers per node, $N_c/NNODES$.

To understand (2), consider the case where at one time step all the nodes contain the same number of walkers. At the end of that step, one of the walkers on one of the nodes is duplicated by the branching process. That node then contains one more walker than all the other nodes. Therefore, at the next time step, all the other nodes will have to wait while that node moves its extra walker. The fraction of time wasted decreases as the average number of walkers per node increases. Option (1), involves balancing the time involved in redistributing walkers across the nodes with the time wasted due to reduced efficiency of the algorithm. On modern MPP machines such as the Cray T3D, the time to redistribute a walker is very small and so we choose to redistribute the walkers at the end of every time step. This places a lower bound on the efficiency of

$$\frac{\mathrm{N_c} * \mathrm{NNODES} + 1}{(\mathrm{N_c} + 1) * \mathrm{NNODES}} \;\leq\; \text{efficiency} \;\leq\; 1 \quad . \tag{2.77}$$

In other words any one node can never have more than one more walker than any other node. For the DMC calculations reported on in this thesis, there are typically 10 walkers per node, yielding an average efficiency of approximately 95%.

The parallel DMC algorithm requires a set of equilibrated configurations as an input, in the same way as the serial DMC algorithm. These configurations are produced by instructing each node in the parallel VMC algorithm to write out an equilibrated ensemble of configurations.

Chapter 3

Quantum Monte Carlo Calculations on Solids

The number of applications of QMC to solid systems is relatively small. The most noteworthy to date are the original Ceperley and Alder calculations of the total energy of the Homogeneous Electron Gas (HEG) [25, 43] and the HEG surface [44, 45], calculations for different phases of hydrogen [46, 47] and pseudopotential studies of heavier atoms such as carbon[48, 26], silicon[26, 49], germanium[50, 33] and nitrogen[51]. The reasons for the slow adoption of QMC as a tool for studying the electronic structure of solids is the intrinsic scaling of the algorithms with the fifth or sixth power of the atomic number and the large finite size effects present in traditional solid QMC calculations[3]. It was only with the introduction of pseudopotentials into QMC calculations[48, 26], that the study of solids with atoms heavier than lithium became feasible with current computing power. Even with this advance there are still several problems involved with performing solid calculations that are unique to QMC because of the real-space representation of the electrons as delta functions. The most severe of these is how to deal with the Coulomb interactions between particles. The traditional Ewald method for treating these interactions is described in section 3.4.

3.1 Supercell Calculations

As well as scaling with the fifth or sixth power of the atomic number of the atomic species being studied, QMC calculations also scale as the third power of the number of the electrons in the system. This is due to the process of updating the Slater determinant of one-electron orbitals after moving each electron (see appendix A). Therefore, even after the introduction of a pseudopotential to reduce the effective atomic number of the ionic cores, it is still important to keep the number of electrons in the system as small as possible.

One method of simulating a solid is to construct a cluster of atoms and then investigate the properties of the cluster as the number of atoms increases. As the size of the cluster increases, the collective behaviour of the atoms within the cluster should asymptotically approach those of the bulk solid. In practice, it turns out that the number of atoms that can be simulated in a QMC calculation is so small that any cluster constructed from such a small number of atoms would be completely dominated by surface effects and would not be able to reproduce the properties of atoms deep within the bulk of a true solid.

An alternative approach to simulating solids is the use of supercells[52]. Here one constructs a supercell containing relatively few atoms and electrons and then repeats the supercell throughout all space using periodic (or toroidal) boundary conditions. These boundary conditions mean that the supercell is wrapped around on itself and as an electron moves out of one side of the supercell it immediately moves back in through the opposite side. The advantage of using such a supercell is that there are no longer any "surface electrons" and hence the problems of the cluster method are removed. However, the supercell method itself still suffers from very significant finite size effects. These are due to the absence of long wavelength fluctuations in the charge density. For a simulation cell of linear dimension, L, the periodicity will remove any electron density waves with wavelength greater than L. One would expect this omission to be especially important in materials where long range effects are dominant such as superconductors containing Cooper pairs of electrons separated

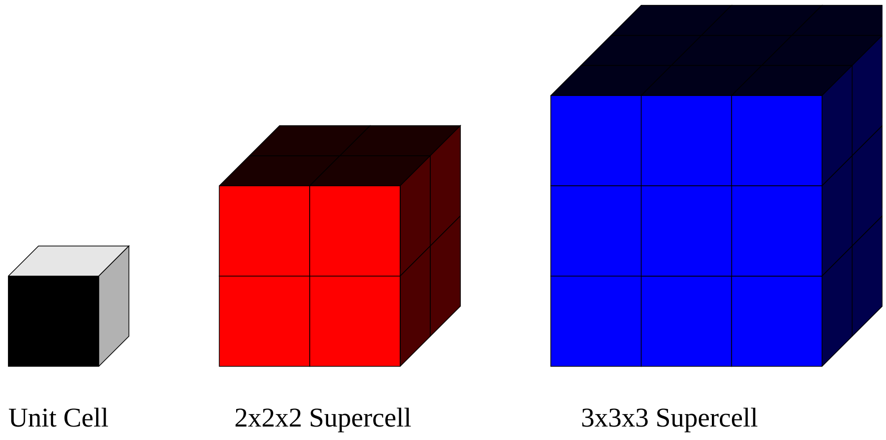


Figure 3.1: Illustration of different supercell sizes.

by many lattice constants. In these cases, the only cure for the finite size effects is to increase the size of the simulation cell being studied. Finite size effects are also present in the simplest systems such as the HEG[3]. Methods of dealing with these finite size effects are discussed in chapter 5.

In supercell calculations, the standard choice of supercell is an integer multiple of primitive unit cells. In the following work, we will refer to the size of supercell by an integer, n, where n=2 refers to a supercell consisting of a 2x2x2 array of primitive unit cells. This is illustrated in figure 3.1.

3.2 Wavefunctions for Solid Calculations

The general Hamiltonian of Eq.(2.18) can be adapted for a supercell calculation in the following way

$$\hat{H} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{\{\mathbf{R}_s\}} \sum_{i>j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}_s|} + \sum_{i=1}^{N} V(\mathbf{r}_i) \quad , \tag{3.1}$$

where $\{\mathbf{R}_s\}$ is the set of translation vectors of the supercell lattice, the potential $V(\mathbf{r})$

has the periodicity of $\{\mathbf{R}_s\}$, and N is the number of electrons in the supercell. In the case where the supercell is constructed from integer multiples of the primitive unit cells, as is the case for all the calculations described here, $V(\mathbf{r})$ also then has the periodicity of the set $\{\mathbf{R}_p\}$ of translation vectors of the underlying crystal lattice.

Trial wavefunctions for this supercell Hamiltonian are based on the general trial wavefunction introduced in Eq.(2.31),

$$\psi(\mathbf{R}) = D^{\uparrow}(\mathbf{R}) D^{\downarrow}(\mathbf{R}) \exp\left[\sum_{i}^{N} \chi(\mathbf{r}_i) - \sum_{i<j}^{N} u(r_{ij})\right] \quad . \tag{3.2}$$

The Slater determinants are constructed from one-electron orbitals obtained from an LDA calculation. The **k**-point sampling in the LDA calculation is chosen to produce the desired number of one-electron orbitals for constructing the Slater determinant in the QMC trial/guiding wavefunction. For example, if an n=1, 1x1x1 supercell is chosen for the QMC calculation then the LDA calculation is also performed on a single unit cell and the wavefunctions are sampled at one **k**-point. If an n=2, 2x2x2 supercell is chosen for the QMC calculation then the LDA calculation is again performed on a single unit cell but now the wavefunctions are sampled from a 2x2x2 mesh of **k**-points in the Brillouin Zone of the primitive lattice.

Recently, new insights have been made into the best choice for the **k**-points at which the one-electron wavefunctions should be calculated [33, 50] for use in QMC calculations. To understand these, first one should consider the translational symmetries of the above Hamiltonian.

1. $\hat{H}$ is invariant under the translation of any electron coordinate by a vector in $\{\mathbf{R}_s\}$.
2. $\hat{H}$ is invariant under the simultaneous translation of all electron coordinates by a vector in $\{\mathbf{R}_p\}$.

Symmetry (2) is a property of the truly infinite system, whereas (1) is a property of the supercell method. Both of these symmetries give rise to a Bloch type condition.

Symmetry (1) implies that the wavefunction can only change by a phase factor when any single electron is translated by a supercell lattice vector. The indistinguishability of the electrons ensures that this phase factor must be the same no matter which electron is moved. This can be demonstrated by applying Bloch's theorem separately to the first and second arguments of the wavefunction and, for the moment, assuming that the two $\mathbf{k}$-vectors are different;

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) = e^{-i\mathbf{k}_1 . \mathbf{R}_s} \Psi(\mathbf{r}_1 + \mathbf{R}_s, \mathbf{r}_2, \ldots) \tag{3.3}$$

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) = e^{-i\mathbf{k}_2 . \mathbf{R}_s} \Psi(\mathbf{r}_1, \mathbf{r}_2 + \mathbf{R}_s, \ldots) \quad . \tag{3.4}$$

We can then apply the permutation symmetry to Eq.(3.3):

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) = -e^{-i\mathbf{k}_1 . \mathbf{R}_s} \Psi(\mathbf{r}_2, \mathbf{r}_1 + \mathbf{R}_s, \ldots) \quad . \tag{3.5}$$

We now translate the second argument by $-\mathbf{R}_s$,

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) = -e^{-i(\mathbf{k}_1 - \mathbf{k}_2) . \mathbf{R}_s} \Psi(\mathbf{r}_2, \mathbf{r}_1, \ldots) \quad , \tag{3.6}$$

and then apply permutation symmetry once more, giving

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) = e^{-i(\mathbf{k}_1 - \mathbf{k}_2) . \mathbf{R}_s} \Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots) \quad . \tag{3.7}$$

It therefore follows that

$$\mathbf{k}_1 - \mathbf{k}_2 \in \{\mathbf{G}_s\} \quad , \tag{3.8}$$

where $\{\mathbf{G}_s\}$ is the set of vectors reciprocal to $\{\mathbf{R}_s\}$. The vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ can be reduced into the first Brillouin Zone (BZ) of the supercell reciprocal lattice, therefore we can choose $\mathbf{k}_1 = \mathbf{k}_2 = \mathbf{k}_s$ without loss on generality. The wavefunction can therefore be written in the form

$$\Psi_{\mathbf{k}_s}(\{\mathbf{r}_i\}) = U_{\mathbf{k}_s}(\{\mathbf{r}_i\}) \exp(i\mathbf{k}_s . \sum_{i=1}^{N} \mathbf{r}_i) \quad , \tag{3.9}$$

where $U_{\mathbf{k}_s}(\{\mathbf{r}_i\})$ is invariant under the translation of any electron coordinate by a vector in $\{\mathbf{R}_s\}$, and is antisymmetric under particle exchange.

Now consider the second symmetry of $\hat{H}$ which states that the wavefunction can only change by a phase factor when all the electrons are translated by a vector in $\{\mathbf{R}_p\}$. This allows us to write

$$\Psi(\{\mathbf{r}_i + \mathbf{R}_p\}) = e^{i\mathbf{k}_p.\mathbf{R}_p}\Psi(\{\mathbf{r}_i\}) \quad , \tag{3.10}$$

where $\mathbf{k}_p$ is the crystal momentum of the wavefunction and $\mathbf{k}_p$ can be reduced into the first BZ of the lattice reciprocal to $\{\mathbf{R}_p\}$. It therefore follows that Ψ can be written in the alternative form

$$\Psi(\{\mathbf{r}_i\}) = W_{\mathbf{k}_p}(\{\mathbf{r}_i\})\exp(i\mathbf{k}_p.\frac{1}{N}\sum_{i=1}^{N}\mathbf{r}_i) \quad , \tag{3.11}$$

where $W_{\mathbf{k}_p}$ in invariant under the simultaneous translation of all electron coordinates by a vector in $\{\mathbf{R}_p\}$ and is antisymmetric under particle exchange.

The operators which translate all the electrons by a vector in $\{\mathbf{R}_p\}$ and the operators which translate a single electrons by a vector in $\{\mathbf{R}_s\}$ commute with each other and with the Hamiltonian, i.e. they form a complete set of commuting operators. Therefore the eigenfunctions of the Hamiltonian in Eq.(3.1) can be chosen to satisfy both the above symmetries at the same time. We can obtain a relationship between the values of $\mathbf{k}_p$ and $\mathbf{k}_s$ by translating all the electrons by a vector in $\{\mathbf{R}_s\}$ (which is a subset of $\{\mathbf{R}_p\}$), and using Eq.(3.9) we find

$$\Psi(\{\mathbf{r}_i + \mathbf{R}_s\}) = e^{i\mathbf{k}_s.N\mathbf{R}_s}\Psi(\{\mathbf{r}_i\}) \quad . \tag{3.12}$$

This must agree with Eq.(3.10), which yields

$$N\mathbf{k}_s - \mathbf{k}_p \in \{\mathbf{G}_s\} \quad . \tag{3.13}$$

In a QMC trial wavefunction, the value of $\mathbf{k}_s$ is determined by the Slater determinant. If all the one-electron wavefunctions making up the determinant reduce to the same value of $\mathbf{k}_s$ in the supercell BZ, then the overall determinant and hence the wavefunction will have that value of $\mathbf{k}_s$. The value of $\mathbf{k}_p$ for a QMC trial wavefunction is determined by the sum of all the $\mathbf{k}_p$ values of the one-electron wavefunctions making up the determinant. Applications of QMC prior to Refs.[33, 50] used the conceptually simplest choice of $\mathbf{k}_s$ and $\mathbf{k}_p$, namely $\mathbf{k}_s = \mathbf{k}_p = 0$. This is achieved by choosing the $\mathbf{k}$-values for the one-electron orbitals on a uniform grid or mesh centred on the origin in reciprocal space, with a grid spacing $\mathbf{G}_s$.

In the limit of an infinite simulation cell, the value of $\mathbf{k}_s$ must tend to zero. However, for a finite simulation cell, the groundstate does not always take the values $\mathbf{k}_s = \mathbf{k}_p = 0$[33]. In the following section, we consider the specific systems of diamond structure germanium and silicon and explore what is the best choice of values of $\mathbf{k}_s$ and $\mathbf{k}_p$ for these systems.

3.3 Germanium and Silicon - The Diamond Structure

The solid QMC calculations reported on in chapters 4 to 6 all involve either germanium or silicon in the diamond structure. This is the stable structure for both materials at room temperature and pressure. The diamond structure consists of a face centred cubic (FCC) Bravais lattice with a two atom basis. The *conventional* unit cell is shown in figure 3.2 and contains four lattice sites and eight atoms. The tetrahedral bonding between atoms in the diamond structure can be clearly seen and is characteristic of the directional covalent bonding found in Group IV of the periodic table. The lattice constant of germanium is 5.65Å and that of silicon is 5.29Å. The diamond structure is relatively "empty", with a maximum of 34% of the available volume being filled by touching hard spheres. The *primitive* unit cell is shown in figure 3.3. Each primitive cell contains a basis of two atoms.

3.3.1 Choice of Slater Determinant

To determine the optimal choice of $\mathbf{k}_s$ for diamond-structure germanium, a series of LDA calculations were performed using $\mathbf{k}$-point meshes with different offsets from the origin[33]. The offset from the origin determines the value of $\mathbf{k}_s$ of the Slater deter-

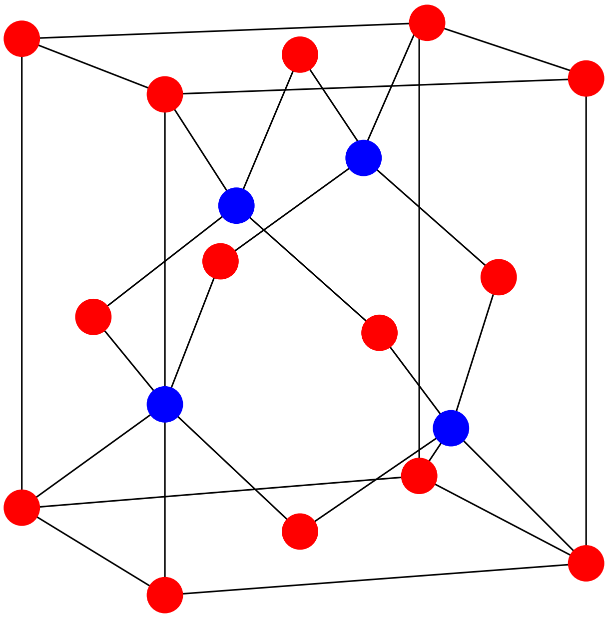


Figure 3.2: Conventional Unit Cell of the diamond structure. One of the atoms in the basis is shown in red and the other in blue.

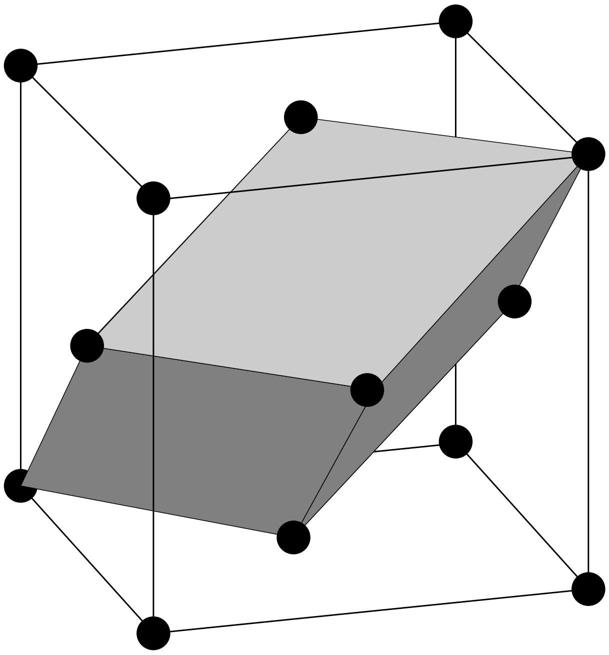


Figure 3.3: Primitive unit cell of the FCC Bravais Lattice

minant constructed from the orbitals. The results of these calculations are illustrated in figure 3.4. The results show the convergence of the total energy with supercell size for four different values of $\mathbf{k}_s$. For the Γ-point ($\mathbf{k}_s$=0) the energy converges slowly as the size of the supercell is increased. A far more rapid convergence is observed when L-point sampling is used. This means the $\mathbf{k}$-point grid is offset from the origin by $\mathbf{G}_{111}/2$, where $\mathbf{G}_{111} = \frac{1}{2}(\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3)$ and the $\mathbf{b}_i$ are the primitive reciprocal lattice translation vectors of the supercell. The best convergence is obtained when using an offset equal to the Baldereschi mean value point[53] of the supercell BZ.The Baldereschi mean point was not chosen for the QMC calculations, as wavefunctions at that $\mathbf{k}$-point are necessarily complex and the use of complex arithmetic would slow down the code considerably. It is possible to construct a real wavefunction using Baldereschi-point states by taking a linear combination of $\phi_{\mathbf{k}_B}$ and its complex conjugate, $\phi^*_{\mathbf{k}_B}$, as both $\phi_{\mathbf{k}_B}$ and $\phi^*_{\mathbf{k}_B}$ have the same energy eigenvalue. Unfortunately, this combination is no longer a Bloch function, and this results in a local energy function that does not have the periodicity of the supercell.

Although it would be too computationally expensive to repeat all the LDA calculations from figure 3.4 within the QMC formalism, VMC and DMC calculations for supercells with $n = 2$ and 3 [33] show that QMC calculations follow the LDA trend very closely. The calculations described in chapters 4 and 5 use Slater determinants where the one-electron orbitals have $\mathbf{k}$-points chosen on a mesh offset from the origin at the L-point. This corresponds to Monkhorst-Pack (MP)[54] sampling for n even supercells and better than MP sampling for n odd supercells.

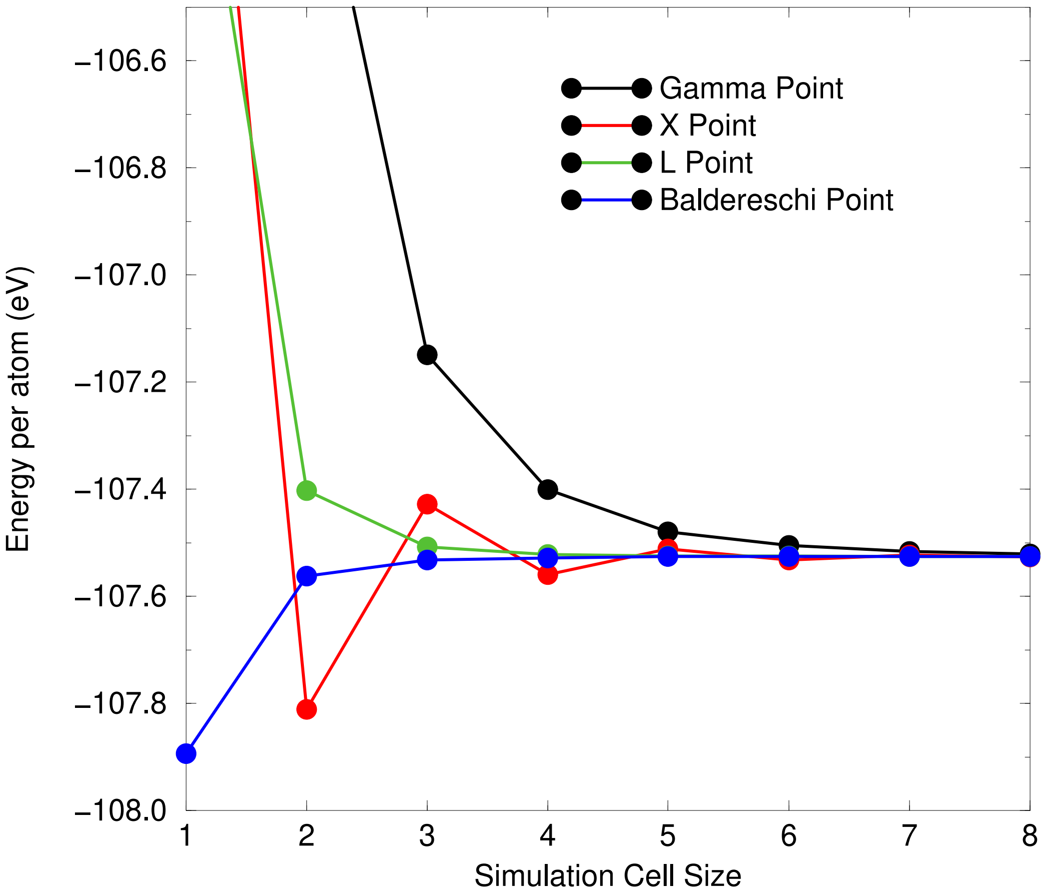


Figure 3.4: Convergence of the total energy with simulation cell size for different $\mathbf{k}$-point sampling schemes. The black line shows the results of $\mathbf{k} = 0$ sampling. The green line shows $\mathbf{k} = \mathbf{G}_{111}/2 = (\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3)/2$ sampling, where the $\mathbf{b}_i$ are the primitive reciprocal lattice translation vectors of the supercell. The red line shows $\mathbf{k} = \mathbf{G}_{100}$ sampling, and the blue line shows $\mathbf{k} = \mathbf{k}_B$ sampling, where $\mathbf{k}_B$ is the Baldereschi mean value point.

3.4 Coulomb Interactions in Supercell Calculations

As already mentioned, one of the most costly problems involved in performing a QMC calculation on a solid is the evaluation of the electrostatic interactions between the charged particles in the system. In this section we introduce the Ewald interaction, which has become the standard technique for evaluating Coulomb interactions in supercell QMC calculations. In chapter 5 a new interaction is proposed, which is designed to improve upon the results obtained with the Ewald interaction.

3.4.1 Isolated Simulation Cell

Before treating the full supercell system, let us consider the Coulomb energies of the particles in an isolated simulation cell[3]. This is exactly the situation one would be faced with when studying clusters of atoms within QMC.

The cell contains N electrons each with charge -1 at positions $\mathbf{r}_i$ and M nuclei with charges Z_α at positions $\mathbf{d}_\alpha$. When the Born-Oppenheimer approximation is used, the positions of the nuclei act only as parameters in the electronic Hamiltonian. This Hamiltonian can be written as

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 + U(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N, \mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_M) \quad . \tag{3.14}$$

For an isolated simulation cell, the term U is simply a superposition of the Coulomb energies for each particle,

$$U = \frac{1}{2}\left(-\sum_{i=1}^{N} \overline{\phi}(\mathbf{r}_i) + \sum_{\alpha=1}^{M} Z_\alpha \overline{\phi}(\mathbf{d}_\alpha) \right) \quad . \tag{3.15}$$

The Coulomb energy for each particle is the result of interactions with all the other charges. There is no self-interaction and so the electrostatic potentials, $\overline{\phi}(\mathbf{r})$, which appear in the equation for U, are the full Coulomb potentials, $\phi(\mathbf{r})$, minus the Coulomb potential of the particle situated at $\mathbf{r}_i$

$$\overline{\phi}(\mathbf{r}_i) = \lim_{\mathbf{r}\to\mathbf{r}_i} \left(\phi(\mathbf{r}) - \frac{1}{|\mathbf{r} - \mathbf{r}_i|} \right)$$

$$\overline{\phi}(\mathbf{d}_\alpha) = \lim_{\mathbf{r}\to\mathbf{d}_\alpha} \left(\phi(\mathbf{r}) - \frac{Z_\alpha}{|\mathbf{r}-\mathbf{d}_\alpha|} \right) \quad . \tag{3.16}$$

The full Coulomb potential, $\phi(\mathbf{r})$, may be calculated by solving Poisson's equation,

$$\nabla^2\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \quad , \tag{3.17}$$

where $\rho(\mathbf{r})$ is the charge density, and the boundary condition is that the potential tends to zero as $\mathbf{r} \to \infty$.

3.4.2 Periodic Boundary Conditions

We now need to extend the above analysis to calculate the potential in the simulation cell when it is just one of an infinite number of periodically repeated, identical simulation cells making up a bulk solid, as in a supercell calculation. The infinite summations of Coulomb potentials from all the charges in the infinite system are only conditionally convergent. It is therefore easier to solve Poisson's equation, again just for the simulation cell, but now imposing periodic boundary conditions to build in the effect of all the periodic images of the simulation cell.

Let us first consider such a crystal with a smoothly varying charge density. The charge density has the periodicity of the primitive lattice and can therefore be written as a Fourier expansion in the reciprocal primitive lattice vectors, $\mathbf{G}_p$, as follows

$$\rho(\mathbf{r}) = \sum_{\mathbf{G}_p} \rho(\mathbf{G}_p) e^{i\mathbf{G}_p.\mathbf{r}} \quad . \tag{3.18}$$

This charge density generates a potential, $\phi(\mathbf{r})$, satisfying Poisson's equation, Eq.(3.17). We can also expand the potential in a Fourier Series to obtain a reciprocal space version of Poisson's equation,

$$\mathbf{G}_p^2\phi(\mathbf{G}_p) = 4\pi\rho(\mathbf{G}_p) \quad . \tag{3.19}$$

Hence one can solve for $\phi(\mathbf{r})$

$$\phi(\mathbf{r}) = \sum_{\mathbf{G}_p\neq 0} \left(\frac{4\pi\rho(\mathbf{G}_p)}{\mathbf{G}_p^2} \right) e^{i\mathbf{G}_p.\mathbf{r}} \quad . \tag{3.20}$$

Eq.(3.19) is only soluble provided $\rho(\mathbf{G}_p = 0) = 0$, which implies that one can only solve Poisson's equation for an infinite periodic system if each unit cell is neutral. In this case, the value of $\phi(\mathbf{G} = 0)$ is completely arbitrary, corresponding to a constant shift in potential which does not affect the potential energy of the system as a whole because it applies in equal and opposite amounts to the positive and negative charges in the system. The solution of Poisson's equation subject to periodic boundary conditions is unique up to an arbitrary constant. This can be demonstrated as follows. Suppose that there were two different periodic solutions, $\phi_a(\mathbf{r})$ and $\phi_b(\mathbf{r})$, to Poisson's equation. The difference between these two solutions,

$$\phi(\mathbf{r}) = \phi_a(\mathbf{r}) - \phi_b(\mathbf{r}) \quad , \tag{3.21}$$

also satisfies the periodic boundary conditions. $\phi(\mathbf{r})$ must also satisfy Laplace's equation

$$\nabla^2 \phi(\mathbf{r}) = 0 \quad , \tag{3.22}$$

everywhere within the simulation cell and so $\phi(\mathbf{r})$ must be a constant throughout the simulation cell as this is the only periodic solution of Laplace's equation. Therefore the two solutions to Poisson's equation are indeed the same to within an arbitrary constant.

It has therefore been shown that the potential due to a *smooth* charge density can easily be evaluated in reciprocal space. However, in a QMC calculation, the instantaneous positions of the electrons are represented in real space by a set of delta functions. This array of delta functions cannot be represented by a convergent Fourier series in the same way that a smooth charge density can, and so the above method for evaluating the potential breaks down. The method of calculating the potential due to such a charge density of point charges that has been traditionally used in QMC calculations is Ewald summation. This method is described in the following section.

3.4.3 Ewald Summation

The Ewald summation method[55, 56, 57] is a technique for evaluating the potential, subject to periodic boundary conditions, due to a lattice of point charges, plus a screening background,

$$\rho_{\text{Ewald}}(\mathbf{r}) = \sum_{\mathbf{R}} \delta(\mathbf{r} - \mathbf{r}_n - \mathbf{R}) - \rho_{\text{background}} \quad , \tag{3.23}$$

where $\mathbf{r}_n$ is the position of the nth charge in the simulation cell and $\mathbf{R}$ is the set of supercell translation vectors. To solve Poisson's equation for the Ewald potential,

$$\nabla^2 \phi_{\text{Ewald}}(\mathbf{r}) = -4\pi \rho_{\text{Ewald}}(\mathbf{r}) \quad , \tag{3.24}$$

due to the above charge density, the density is split up into two components, the background charge and the array of delta functions. An array of Gaussian functions, centred at $\mathbf{r}_n + \mathbf{R}$ is added to each component of the Ewald charge density. The Gaussians are normalised to ensure that both of the individual components of the Ewald charge density are neutral. The two charge density components can be written as

$$\rho_{\text{Ewald}}(\mathbf{r}) = \rho_1(\mathbf{r}) + \rho_2(\mathbf{r}) \quad , \tag{3.25}$$

where

$$\rho_1(\mathbf{r}) = \left(\frac{1}{\mu \pi^{\frac{1}{2}}}\right) \sum_{\mathbf{R}} e^{-(\mathbf{r}-\mathbf{r}_n-\mathbf{R})^2/\mu^2} - \rho_{\text{background}} \tag{3.26}$$

$$\rho_2(\mathbf{r}) = \sum_{\mathbf{R}} \left(\delta(\mathbf{r} - \mathbf{r}_n - \mathbf{R}) - \left(\frac{1}{\mu \pi^{\frac{1}{2}}}\right) e^{-(\mathbf{r}-\mathbf{r}_n-\mathbf{R})^2/\mu^2} \right) \quad . \tag{3.27}$$

These charge densities are schematically represented in one dimension in figure 3.5.

The potential due to $\rho_1(\mathbf{r})$ is most conveniently calculated in reciprocal space. $\rho_1(\mathbf{r})$ has non-zero Fourier components on supercell reciprocal lattice vectors, $\mathbf{G}_s$, given by

$$\rho_1(\mathbf{G}_s) = \begin{cases} \left(\frac{1}{\Omega_s}\right) e^{-\mu^2 \mathbf{G}_s^2/4} e^{-i\mathbf{G}_s.\mathbf{r}_n}, & \mathbf{G}_s \neq 0 \\ 0, & \mathbf{G}_s = 0 \quad , \end{cases} \tag{3.28}$$

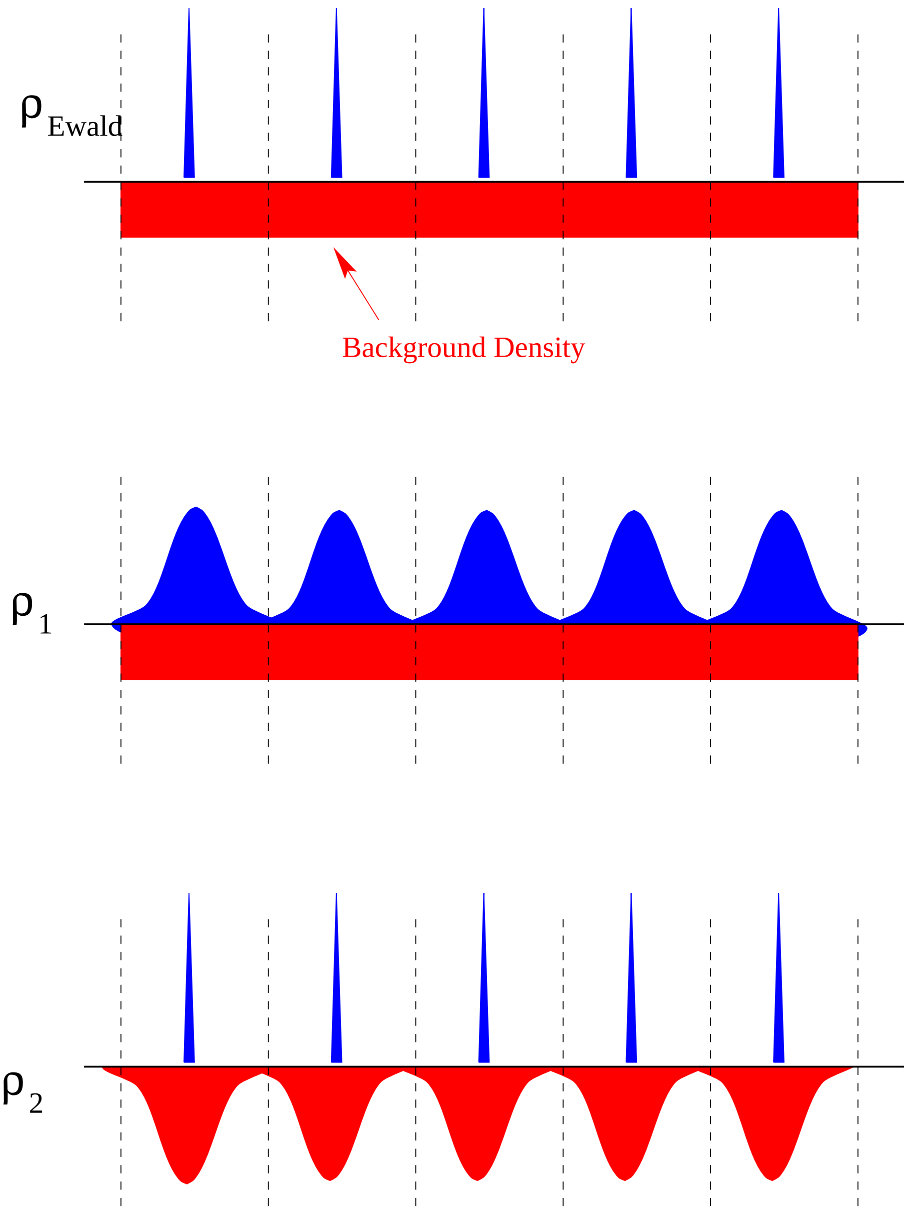


Figure 3.5: Schematic representation of the two components of the Ewald charge density. Blue indicates positive charge and red negative charge.

where Ω_s is the volume of the supercell. Solving the corresponding reciprocal space version of Poisson's equation, $\mathbf{G}_s^2\phi_1(\mathbf{G}_s) = 4\pi\rho_1(\mathbf{G}_s)$, gives

$$\phi_1(\mathbf{r}) = \sum_{\mathbf{G}_s \neq 0} \left(\frac{4\pi}{\mathbf{G}_s^2\Omega_s}\right) e^{-\mu^2\mathbf{G}_s^2/4} e^{i\mathbf{G}_s.(\mathbf{r}-\mathbf{r}_n)} \quad . \tag{3.29}$$

To calculate $\phi_2(\mathbf{r})$, the potential due to the array of point charges minus the screening Gaussians, evaluation in real space is more convenient since the coefficients in the Fourier expansion of a periodic array of delta functions do not decay for large $\mathbf{G}$ vectors. The resultant potential is the summed potential of the delta function point charges minus the sum of potentials due to the Gaussian charge distributions. One can show that the potential of the Gaussian charge distribution is given by

$$\phi(\mathbf{r}) = \frac{\text{erf}(r/\mu)}{r} \quad , \tag{3.30}$$

where the error function is defined as

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x'^2} dx' \quad . \tag{3.31}$$

Consequently, the real space sum generating $\phi_2(\mathbf{r})$ is

$$\phi_2(\mathbf{r}) = \sum_{\mathbf{R}} \frac{1 - \text{erf}(|\mathbf{r} - \mathbf{r}_n - \mathbf{R}|/\mu)}{|\mathbf{r} - \mathbf{r}_n - \mathbf{R}|} - \frac{\pi\mu^2}{\Omega_s^2} \quad , \tag{3.32}$$

where the last term, $\frac{\pi\mu^2}{\Omega_s^2}$, is added so that the average potential in the supercell is zero. Combining the reciprocal space sum for $\phi_1(\mathbf{r})$ and the real space sum for $\phi_2(\mathbf{r})$ gives the final result for $\phi_{\text{Ewald}}(\mathbf{r})$

$$\begin{aligned} \phi_{\text{Ewald}}(\mathbf{r}) &= \sum_{\mathbf{G}_s \neq 0} \left(\frac{4\pi}{\mathbf{G}_s^2\Omega_s}\right) e^{-\mu^2\mathbf{G}_s^2/4} e^{i\mathbf{G}_s.(\mathbf{r}-\mathbf{r}_n)} \\ &+ \sum_{\mathbf{R}} \frac{\text{erfc}(|\mathbf{r} - \mathbf{r}_n - \mathbf{R}|/\mu)}{|\mathbf{r} - \mathbf{r}_n - \mathbf{R}|} - \frac{\pi\mu^2}{\Omega_s^2} \quad . \end{aligned} \tag{3.33}$$

The value of $\phi_{\text{Ewald}}(\mathbf{r})$ is independent of the half width, μ, of the Gaussian charges. However, the value of μ affects the speed of convergence of the above real and reciprocal space sums.

The full potential of a simulation cell containing N electrons and M ions is found by superposing all the potentials for each charge component, since the full charge distribution is the superposition of all the point charges and their cancelling backgrounds,

$$\phi_{\text{total}}(\mathbf{r}) = -\sum_{n=1}^{N} \phi_{\text{Ewald}}(\mathbf{r}, \mathbf{r}_n) + \sum_{\alpha=1}^{M} Z_\alpha \phi_{\text{Ewald}}(\mathbf{r}, \mathbf{d}_\alpha) \quad . \tag{3.34}$$

Therefore $\overline{\phi}(\mathbf{r}_i)$, as defined in Eq.(3.16) is given by

$$\overline{\phi}(\mathbf{r}_i) = -\sum_{n=1}^{N} \phi_{\text{Ewald}}(\mathbf{r}, \mathbf{r}_n) + \sum_{\alpha=1}^{M} Z_\alpha \phi_{\text{Ewald}}(\mathbf{r}, \mathbf{d}_\alpha) - \xi \tag{3.35}$$

where

$$\begin{aligned} \xi &= \lim_{\mathbf{r} \to \mathbf{r}_n} \left(\phi_{\text{Ewald}}(\mathbf{r}, \mathbf{r}_n) - \frac{1}{|\mathbf{r} - \mathbf{r}_n|} \right) \\ &= \frac{1}{\Omega_s} \sum_{\mathbf{G}_s \neq 0} \frac{\exp\left(-\pi^2 \mu^2 \mathbf{G}_s^2 / 4\right)}{\pi \mathbf{G}_s^2} + \sum_{\mathbf{R}(\neq 0)} \frac{\operatorname{erfc}(|\mathbf{R}|/\mu)}{|\mathbf{R}|} - \frac{\pi \mu^2}{\Omega_s^2} - \frac{2}{\mu \sqrt{\pi}} \end{aligned} \tag{3.36}$$

is the self-image interaction, i.e. the potential at the unit point charge due to its own background and array of images. $\overline{\phi}(\mathbf{d}_\alpha)$ can be found in exactly the same way and the total electrostatic energy per simulation cell can then be written as

$$\begin{aligned} U &= \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \phi_{\text{Ewald}}(\mathbf{r}_i, \mathbf{r}_j) - \sum_{i=1}^{N} \sum_{\alpha=M}^{N} Z_\alpha \phi_{\text{Ewald}}(\mathbf{r}_i, \mathbf{d}_\alpha) \\ &+ \frac{1}{2} \sum_{\alpha=1}^{M} \sum_{\beta \neq \alpha}^{M} Z_\alpha Z_\beta \phi_{\text{Ewald}}(\mathbf{d}_\alpha, \mathbf{d}_\beta) + \frac{N\xi}{2} + \frac{\xi}{2} \sum_{\alpha=1}^{M} Z_\alpha^2 \quad . \end{aligned} \tag{3.37}$$

The charge neutrality of the simulation cell dictates that

$$N(-1) + \sum_{\alpha=1}^{M} Z_\alpha = 0 \quad . \tag{3.38}$$

Therefore the above expression for U is easily simplified to

$$\begin{aligned} U &= \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} (\phi_{\text{Ewald}}(\mathbf{r}_i, \mathbf{r}_j) - \xi) - \sum_{i=1}^{N} \sum_{\alpha=M}^{N} Z_\alpha (\phi_{\text{Ewald}}(\mathbf{r}_i, \mathbf{d}_\alpha) - \xi) \\ &+ \frac{1}{2} \sum_{\alpha=1}^{M} \sum_{\beta \neq \alpha}^{M} Z_\alpha Z_\beta (\phi_{\text{Ewald}}(\mathbf{d}_\alpha, \mathbf{d}_\beta) - \xi) \end{aligned} \tag{3.39}$$

In QMC solid calculations, the ionic coordinates are fixed throughout the calculation. Therefore the contribution to the total energy from the ion-ion Coulomb interaction need only be evaluated once at the beginning of the simulation.

As each of the electrons are moved in turn in a VMC calculation, the contribution to the total energy from the electron-electron and electron-ion Coulomb interactions needs to be recalculated for each electron after it is moved.

3.4.4 Electron-Ion Coulomb Interactions

Following the method of Ref.[26], we choose to represent the ionic cores in our germanium and silicon supercell calculations with pseudopotentials. This enables the number of valence electrons that are explicitly handled by the QMC algorithm to be reduced to four per atom in both cases.

The pseudopotential used to represent the Ge^{4+} ions in the germanium calculations described in chapter 4 was a local pseudopotential of the Starkloff-Joannopoulos form[58]. The pseudopotential used to represent the Si^{4+} ions in the silicon calculation described in chapters 5 and 6 is a norm-conserving, non-local pseudopotential generated using the method described by Kerker[59]. In this pseudopotential, the s and p potentials were generated from an s^2p^2 atomic groundstate and the d potential was generated from an $s^1p^{0.75}d^{0.25}$ atomic configuration as in Ref.[60]. In our calculations we chose the p potential to be the local potential as this results in a smaller contribution from the remaining non-local potential to the total energy than choosing either s or d to be local. A small non-local energy is desirable as the non-local energy is evaluated by a statistical integration within the QMC code. This integration is expensive to evaluate and can be evaluated more approximately (and cheaply) if the overall contribution from the non-local potential is small. Also, in DMC calculations, we would like the non-local energy to be as small as possible to reduce the effect of the "locality approximation".

Both these pseudopotentials feature a cutoff radius, beyond which the pseudopotential reduces to the full Z/r potential due to a $+Z$ point charge, where Z is the valence of the ion. To deal with the long ranged tails of the ionic potential the Ewald prescription, as described in the previous section, is used to evaluate the interaction energy between the lattice of charges representing the ionic core and all its periodic images and the lattice representing an electron and all its periodic images. This is illustrated in figure 3.6. For each electron-ion pair, if the electron is outside the cutoff radius of the pseudopotential (position 1 in figure 3.6), then the Ewald interaction is directly applied to calculate the Coulomb energy between the two corresponding lattices of charged particles and their screening background charges. If the electron falls within the cutoff radius of the pseudopotential, (position 2 in figure 3.6), then the Ewald interaction is still used to evaluate the interaction between the electron and all the periodic images of the ion, but a correction is applied to include the effect of the pseudopotential from the *ion in the simulation cell* on the *electron in the simulation cell* and identical "in cell" effects in all the periodic images of the simulation cell.

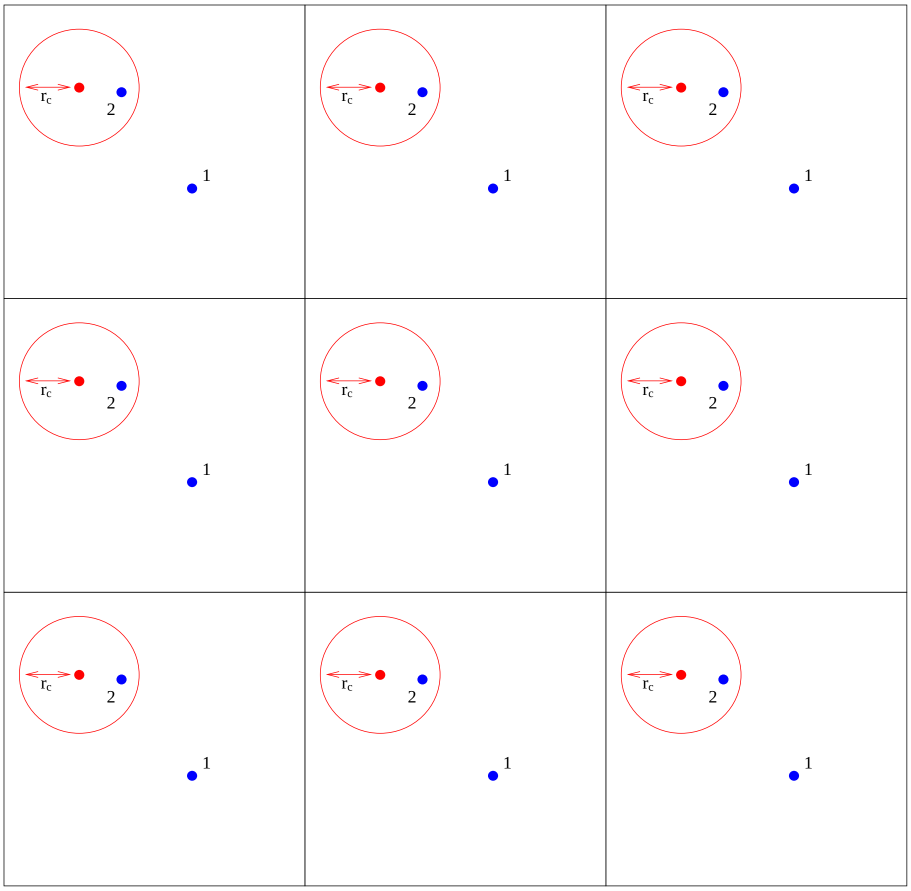


Figure 3.6: Schematic representation of the electron-ion interaction. The red point represents an ionic core with a cutoff radius, r_c, marked by the red circle. The blue circles marked 1 and 2 represent two different positions for an electron in the simulation cell. Only a central simulation cell and one periodic image in each direction is shown.

Chapter 4

Optimising Trial Wavefunctions

This chapter describes work carried out to optimise trial/guiding wavefunctions for use in VMC and DMC calculations. The motivation for optimising trial wavefunctions is introduced in the light of recent results for the ground state energy of solid germanium in the diamond structure using VMC and DMC[33, 50]. These results illustrate the difference in the quality of the standard trial/guiding wavefunction and the 'groundstate' wavefunction that the DMC algorithm propagates towards. The variance minimisation technique used to perform the optimisation and the various choices of functional form to optimise are then described in a similar order to that in which the work was originally performed. Preliminary results are given at the end of each stage of the optimisation as they had a major influence on the future direction of the project.

4.1 Motivation

The role of the trial/guiding wavefunction is crucial to both the VMC and DMC techniques.

4.1.1 VMC Calculations

The importance of the trial/guiding wavefunction in determining the accuracy of VMC calculations is illustrated by recent work [33] to calculate the total energy of solid germanium in the diamond structure using a 2x2x2 supercell containing 16 Ge atoms and 64 valence electrons within both the VMC and DMC formalisms. The results are illustrated in Figure 4.1.

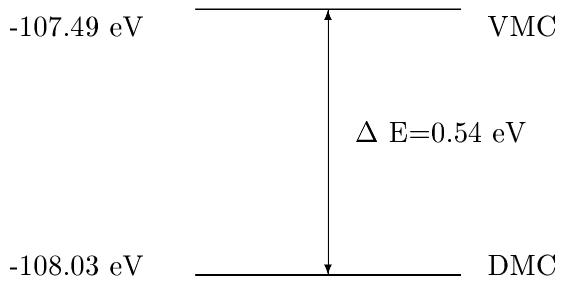


Figure 4.1: Difference in the energy of VMC and DMC results for 2x2x2 bulk germanium in the diamond structure.

The VMC energy and DMC energy are defined as

$$E_{VMC} = \frac{\int d\mathbf{R}\,(\Psi_T^*(\mathbf{R})\Psi_T(\mathbf{R}))\left(\Psi_T^{-1}(\mathbf{R})H\Psi_T(\mathbf{R})\right)}{\int d\mathbf{R}\,(\Psi_T^*(\mathbf{R})\Psi_T(\mathbf{R}))} \tag{4.1}$$

$$E_{DMC} = \frac{\int d\mathbf{R}\,(\Psi_T^*(\mathbf{R})\Phi_0(\mathbf{R}))\left(\Psi_T^{-1}(\mathbf{R})H\Psi_T(\mathbf{R})\right)}{\int d\mathbf{R}\,(\Psi_T^*(\mathbf{R})\Phi_0(\mathbf{R}))} \quad , \tag{4.2}$$

where Φ_0 is the wavefunction that Ψ_T propagates to during the diffusion process.

It has already been stressed in section 2.5.1 that the DMC energy is in principle exact (apart from the fixed-node approximation), i.e. for a given trial/guiding wavefunction Ψ_T , the DMC energy is equivalent to the lowest variational energy for all wavefunctions with the same nodal surface as Ψ_T . As the DMC and VMC [50] calculations share the same trial wavefunction, Ψ_T and hence have the same nodal surface, the difference in energy $|E_{DMC} - E_{VMC}|$ is due only to the difference between the VMC trial wavefunction Ψ_T and the wavefunction, Φ_0, to which the DMC calculation converges to in the long time approximation. This energy difference acts as a measure of the *quality* of Ψ_T , where the quality describes how closely Ψ_T matches the converged DMC ground state wavefunction, Φ_0, throughout all of configuration space. If one is able to optimise the trial/guiding wavefunction and therefore improve its quality, then the difference $|E_{DMC} - E_{VMC}|$ will be reduced and the accuracy of the VMC calculations is improved.

It is not only the value of the total energy which reflects the quality of a trial/guiding wavefunction. The intrinsic variance of the energy estimator, as defined by

$$\begin{aligned} \sigma^2_{VMC} &= \frac{\int d\mathbf{R}\,(\Psi^*_T(\mathbf{R})\Psi_T(\mathbf{R}))\left(\Psi_T^{-1}(\mathbf{R})H\Psi_T(\mathbf{R})\right)^2}{\int d\mathbf{R}\,(\Psi^*_T(\mathbf{R})\Psi_T(\mathbf{R}))} \\ &- \left(\frac{\int d\mathbf{R}\,(\Psi^*_T(\mathbf{R})\Psi_T(\mathbf{R}))\left(\Psi_T^{-1}(\mathbf{R})H\Psi_T(\mathbf{R})\right)}{\int d\mathbf{R}\,(\Psi^*_T(\mathbf{R})\Psi_T(\mathbf{R}))}\right)^2 \quad , \end{aligned} \tag{4.3}$$

is also important. If one is able to reduce the intrinsic variance, σ^2_{VMC}, of the VMC energy by improving the quality of Ψ_T , then the number of VMC moves required to achieve a specific variance of the mean, σ^2_M, decreases linearly with the variance. i.e. $\sigma^2_M = \sigma^2_{VMC}/N$, where N is the number of moves.

4.1.2 DMC Calculations

The quality of the trial/guiding wavefunction does not directly affect the final DMC estimate of the total energy of a given system (apart from the fixed node approximation). However, the intrinsic variance of E_{VMC} determines the variance of the estimate of the total energy at each step of the diffusion process. Therefore, as in the VMC technique, the number of DMC moves required to achieve a specific variance of the mean, σ^2_M, decreases linearly with the intrinsic variance of E_{VMC}. As the computational cost of a DMC calculation is roughly an order of magnitude greater than that of a comparable VMC calculation, reducing this intrinsic variance is very important.

Reducing the intrinsic variance of the trial/guiding wavefunction by optimisation also improves the numerical stability of a DMC calculation. As described in section 2.5.1, the multiplicity of a walker at the end of each DMC move is determined by its excess local energy, $[\frac{\hat{H}\Psi_T}{\Psi_T} - E_T]$. Therefore, any reduction in the intrinsic variance of Ψ_T will reduce the deviation of the individual multiplicities of the walkers from unity and help to produce a more stable population.

Finally, any improvement in the quality of Ψ_T should reduce the time taken for a DMC calculation to reach a converged energy. If Ψ_T is closer to Φ_0 the probability distribution $f(\mathbf{R}, t)$ (see section 2.5.1) starts closer to its converged form. See figure 2.4 and section 2.5.1 for more details on the convergence of the DMC algorithm.

4.2 Optimisation Method

4.2.1 Why Minimise the Variance of the Energy?

Perhaps the most intuitive method of wavefunction optimisation is to vary a set of parameters within the wavefunction so as to minimise the energy with respect to the values of those parameters. Proper application of this method for a parameterised wavefunction gives the best (lowest) value for the energy of the system, but it may give poor values for other expectation values. If the energy is minimised then the local energy may be too high in some regions of configuration space and too low in others, so that the overall quality of the wavefunction is poor. This type of behaviour contributes to the variance of the energy and this therefore suggests that minimisation of the variance may give a better fit for the wavefunction as a whole, so that satisfactory results are obtained for a range of quantities, including the energy[22]. Furthermore, the variance of the energy is zero for an eigenfunction and positive for an approximate wavefunction, and therefore the quantity to be minimised has a well defined minimum value.

4.2.2 Previous Applications of Variance Minimisation

The method of variance minimisation was first applied to quantum mechanical problems in the 1930's. It was first used in QMC calculations by Coldwell [61], and some of the most impressive QMC applications have been by Umrigar and coworkers [22, 30]. Umrigar developed the variance minimisation technique [22] to calculate wavefunctions for use in VMC and DMC calculations on the Be atom. He took the standard atomic trial wavefunction

$$\psi = D^{\uparrow} D^{\downarrow} e^{\sum_{i<j} \frac{a_{ij} r_{ij}}{(1+b r_{ij})}} \quad , \tag{4.4}$$

where a_{ij} is equal to 1/2 for antiparallel-spin electrons and 1/4 for parallel-spin electrons, to satisfy the cusp condition. He then optimised the value of the parameter b using variance minimisation.

He then attempted to generalise the Jastrow part of the wavefunction to take account of the individual positions of the electrons as well as the electron-electron separation.

$$\psi_J = \exp\left(\sum_{i<j} \frac{P([a], r_{ij}, s_{ij}, t_{ij})}{1 + P([b], r_{ij}, s_{ij}, t_{ij})}\right) \quad , \tag{4.5}$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|,\ s_{12} = |\mathbf{r}_1| + |\mathbf{r}_2|,\ t_{12} = |\mathbf{r}_1| - |\mathbf{r}_2|$ and P is a complete 4^{th} order polynomial in r, s and t with sets of coefficients $[a]$ and $[b]$. The sets of coefficients $[a]$ and $[b]$ were then optimised using the same variance minimisation technique. The error in the expectation value of the energy was reduced from 0.001 to 0.000003 Hartrees by the optimisation process.

Recently, Umrigar and Filippi[62] have extended their work to study first row diatomic molecules with QMC. They have used multiconfigurational wavefunctions and used the variance minimisation method to not only optimise the Jastrow and $\chi(\mathbf{r})$ functions but also to optimise the wavefunctions with respect to some variational parameters present in the one-electron wavefunctions that make up the Slater determinants.

Mitas and Martin have also optimised wavefunctions for use in atomic QMC calculations [41]. They also chose a two-body correlation function that depends on the electron-electron separation and an additional electron-ion term, $u(r_{iI}, r_{jI}, r_{ij})$

$$u(r_{iI}, r_{jI}, r_{ij}) = -\frac{c}{\gamma} e^{-\gamma r_{ij}} + \sum_{k<l,m} c_{klm}[a_k(r_{iI})a_l(r_{jI}) + a_k(r_{jI})a_l(r_{iI})]b_m(r_{ij}) \quad , \quad (4.6)$$

where

$$a_k(r) = \left(\frac{\alpha_k r}{1+\alpha_k r}\right)^2 \; , \; \alpha_k = \frac{\alpha_0}{2^{k-1}} \; , \; k > 0 \quad (4.7)$$

$$b_m(r) = \left(\frac{\beta_m r}{1+\beta_m r}\right)^2 \; , \; \beta_m = \frac{\beta_0}{2^{m-1}} \; , \; m > 0 \quad , \quad (4.8)$$

and r_{iI} is the separation between the i^{th} electron and the I^{th} ion, and $a_0(r) = b_0(r) = 1$. The largest range of k, l, m used was from 0 to 5 but only a subset of values were actually used.

When performing VMC calculations on atomic and molecular nitrogen [63], Mitas and Martin used 21 variational parameters in $u(r_{iI}, r_{jI}, r_{ij})$. These were optimised using the Umrigar minimisation of variance technique.

Mitas and Martin have also begun optimising wavefunctions for use in QMC calculations on solids. They used the same correlation function, Eq.(4.6) from the nitrogen atom calculations, but with only 6 optimised parameters, to perform calculations on solid nitrogen.

A few QMC trial/guiding wavefunctions have also been generated by methods other than the Umrigar minimisation of variance technique [22]. Tanaka [64] has performed VMC calculations on the cohesive energy of NiO using wavefunctions generated by minimising the total energy of the wavefunction. He used a trial wavefunction where the Jastrow factor took its standard form

$$u(r) = \frac{A}{r}\left(1 - e^{-\frac{r}{F}}\right) \quad , \quad (4.9)$$

and the Chi function was expressed as a sum of Gaussians

$$\chi(\mathbf{r}) = \sum_{\mathbf{R}} \sum_{i} c_i e^{-d_i(\mathbf{r}-\mathbf{R})^2} \quad , \quad (4.10)$$

where $\mathbf{r}$ is the electron position and $\mathbf{R}$ are the positions of the ionic cores.

4.2.3 The Variance Minimisation Method

We begin by writing the variance of the energy as

$$\sigma^2 = \sum_\alpha \left(\Psi_T^{-1}(\mathbf{R}_\alpha)\hat{H}\Psi_T(\mathbf{R}_\alpha) - < E_L >\right)^2 \left[\frac{w(\alpha)}{\sum_\beta w(\beta)}\right], \tag{4.11}$$

where $\hat{H}$ is the Hamiltonian as defined in chapter 3 and Ψ_T is the trial/guiding wavefunction which is to be optimised. The sum is over a set of $3N$-dimensional electron configurations, $\mathbf{R}_\alpha$, $< E_L >$ is an average energy,

$$< E_L > = \sum_\alpha \Psi_T^{-1}(\mathbf{R}_\alpha)\hat{H}\Psi_T(\mathbf{R}_\alpha) \left[\frac{w(\alpha)}{\sum_\beta w(\beta)}\right], \tag{4.12}$$

and the reweighting factors, $w(\alpha)$, are given by

$$w(\alpha) = \left|\frac{\Psi_T(\mathbf{R}_\alpha)}{\Psi_T^0(\mathbf{R}_\alpha)}\right|^2 . \tag{4.13}$$

The electron configurations are sampled from the starting distribution $|\Psi_T^0|^2$ and then kept fixed throughout the optimisation. This "correlated sampling" approach gives a good estimate of the difference in variance between wavefunctions corresponding to different sets of parameters. The process can be used iteratively by using the optimised set of parameters to regenerate a new set of configurations which are then used to perform a new optimisation. This is useful when the reweighting factors differ significantly from unity (see section 4.2.4). The non-linear optimisations over the multi-dimensional parameter spaces were performed using the NAG routine E04FCF. This works by finding the unconstrained minimum of a sum of squares, as in Eq.(4.11), using a modified Newton algorithm that requires the function values only.

4.2.4 Control of the Reweighting Factors

The variance minimisation procedure in Eq.(4.11) is stable for systems containing small numbers of electrons, but gradually becomes unstable as the numbers of electrons increases. The problem arises because the reweighting factors, $w(\alpha)$, in Eq.(4.11) deviate greatly from unity at the beginning of the optimisation process when the changes in the parameters are large due to the poor quality of the initial guess of the trial wavefunction. Note that for a small change, $\delta\chi$, in the $\chi(\mathbf{r})$ function, the local energy $(\Psi_T^{-1}\hat{H}\Psi_T)$ changes by an amount *proportional* to $N\delta\chi$, where N is the number of electrons, but the reweighting factor is multiplied by a factor which is *exponential* in $N\delta\chi$. For large N, this behaviour of the reweighting factors dominates the optimisation process. One way to counteract this is to increase the number of configurations used, but this rapidly becomes impractical and it was found that it is better to reduce the importance of the reweighting factor. For $N \sim 50$ and above, the reweighting factors were set to unity and several sets of configurations were generated with successively better $\chi(\mathbf{r})$ functions, until the procedure converged. This works because as the procedure converges the reweighting factors approach unity. At each iteration only limited variations of the parameters were allowed, which promotes the smooth convergence of the process. This optimisation is perfectly stable even up to the largest number of electrons studied ($N = 338$).

4.2.5 Choice of Average Local Energy, $< E_L >$

Several different definitions of the average local energy, $< E_L >$, in Eq.(4.11) were experimented with. The choice which most accurately describes the average local energy for a given set of variational parameters is the reweighted local energy, summed over all the configurations,

$$< E_L >= \sum_{\alpha} \left(\Psi_T^{-1}(\mathbf{R}_\alpha)\hat{H}\Psi_T(\mathbf{R}_\alpha)\right) \left[\frac{w(\alpha)}{\sum_\beta w(\beta)}\right] \quad . \tag{4.14}$$

If one adopts this definition, then the optimisation process can be pictured as minimising the variance of the energy about a mean energy which is also slowing decreasing as the quality of the trial wavefunction improves. This is illustrated in figure 4.2. It was found that this choice of E_L is the most unstable due to fluctuations in the reweighting factors, $w(\alpha)$. The overall stability of the optimisation process was improved by

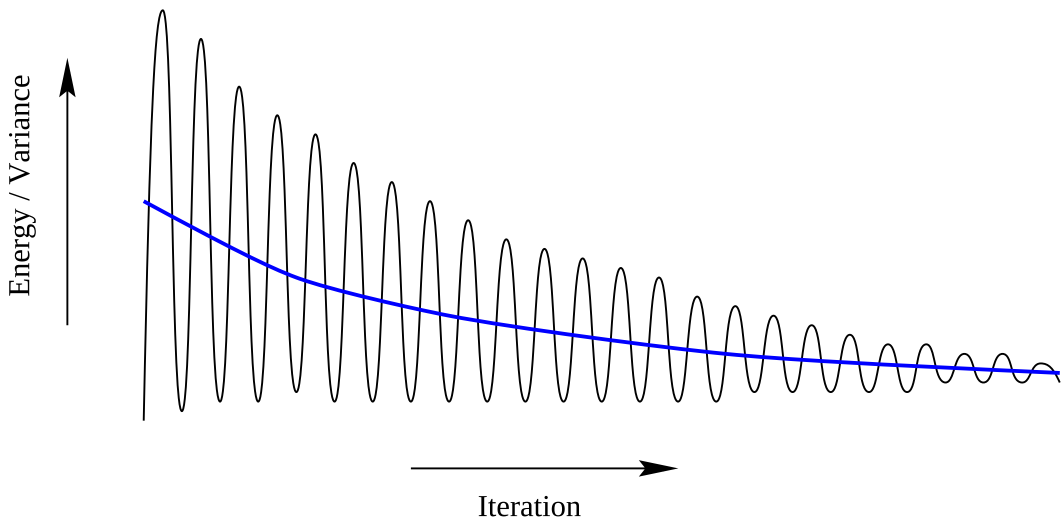


Figure 4.2: Schematic representation of the reduction in the energy and variance of the energy during the optimisation process. The blue line shows the mean local energy, summed over the ensemble of configurations. The black line schematically represents the variance of the local energy of the configurations.

choosing the simpler definition,

$$< E_L >= \frac{1}{N} \sum_{\alpha} \left(\Psi_T^{-1}(\mathbf{R}_\alpha) \hat{H} \Psi_T(\mathbf{R}_\alpha) \right) \quad . \tag{4.15}$$

Both definitions of E_L are equivalent at the end point of the optimisation where the reweighting factors are all unity. The simple definition of E_L in Eq.(4.15) has been used in all the optimisations described in this thesis. It is worth noting at this point that alternative choices for E_L have also been proposed. Umrigar[22] often chooses to keep E_L fixed, throughout the optimisation process, to the original VMC energy produced by the VMC calculation used to generate the configurations. Presumably, this is to improve the stability of the optimisation process still further. However, in our calculations, this choice of E_L produced a lower quality of trial wavefunction because as the optimisation proceeds and the average local energy decreases, one is no longer minimising the variance of the energy about the mean value. Finally, one could think of attempting to 'guide' the optimisation process in the right direction by choosing to minimise the variance about a fixed energy that was lower than the original VMC energy, such as the DMC energy. This also produced a lower quality trial wavefunction in our tests.

4.2.6 Generating Configurations

The configurations used in in the optimisation procedure Eq.(4.11) are generated by periodically storing the positions of all the electrons (a *configuration*) after a certain number of moves during a VMC calculation. To keep the computation time for generating the configurations to a minimum, the number of moves required to produce independent configurations (i.e. the correlation length) was investigated by a procedure based on that used by Jacucci and Rahman [65] which works as follows.

A series of N sequential configurations and their energies were generated by a VMC calculation. The N energies were divided up into b blocks each containing N_b energies. The average value for each block and the variance of the block averages are then given by

$$\overline{x}_b = \frac{1}{N_b}\sum_{i=1}^{N_b} x_i \quad , \quad \sigma^2(\overline{x}_b) = \frac{1}{b}\sum_{1}^{b}(\overline{x}_b - \overline{x})^2 \quad . \tag{4.16}$$

When the block size is large enough such that the individual block averages $\overline{x}_b$ can be considered as being independent, the value of $\sigma^2(\overline{x}_b)$ might be expected to be inversely proportional to N_b. This is because the individual error in the mean, $(\overline{x}_b - \overline{x})$ is proportional to $1/\sqrt{(N_b)}$. In an attempt to calculate this constant of proportionality, the *statistical inefficiency*, s, is defined as

$$s = \lim_{N_b \to \infty} \frac{N_b \sigma^2(\overline{x}_b)}{\sigma^2(x)} \quad . \tag{4.17}$$

It is possible to calculate a value for s from the N energies by plotting the value of s for a series of block sizes. Finally, from the definition of the variance in the mean for a series of N values and a correlation length τ, we have

$$\sigma^2(\overline{x}) = \tau \frac{\sigma^2(x)}{N} \quad . \tag{4.18}$$

It is clear that for the case of a single block that constitutes the whole sample (i.e. $N_b = N$, $b = 1$) s is equivalent to τ in (4.18). For the energies produced in a

VMC calculation the value of s proved to be much smaller than expected, about 2.8. This meant that every third move of all the electrons could be written out as an independent configuration. Where results are presented later in this report for fitting with 1000 configurations this required a 3000 move VMC calculation, and 30,000 moves were required to generate 10,000 configurations.

4.3 Optimising the χ function

4.3.1 Which Part of Ψ_T to Optimise?

As described in chapter 2, the many-body wavefunction can be written in the following form;

$$\Psi_T(R) = D^{\uparrow}(R)D^{\downarrow}(R)\exp\left(\sum_{(s,i)=(\uparrow,1)}^{(\downarrow,N)} \chi(\mathbf{r}_i)\right)\ \exp\left(-\sum_{(\uparrow,1)\leq(s,i)<(s',j)}^{(\downarrow,N)} u(r_{ij})\right) \quad . \tag{4.19}$$

The trial wavefunction in Eq.(4.19) is commonly referred to as the *Hartree-Fock-Jastrow-Chi* wavefunction. All three parts of the wavefunction, i.e. the HF Slater determinant of one-electron orbitals, the Jastrow factor, and the Chi function, would benefit from some form of optimisation of parameters, and so deciding which part to optimise is a matter of deciding which part should yield the greatest improvement in the quality of Ψ_T .

The Chi function has previously been constructed according to a scheme introduced by Fahy [26], using Eq.(4.20). Of the three parts of the Hartree-Fock-Jastrow-Chi trial wavefunction, this $\chi(\mathbf{r})$ function has the weakest theoretical justification for its form. It is constructed on an *ad hoc* basis by making the assumption that the LDA charge density is reasonably close to the real charge density. The $\chi(\mathbf{r})$ function is then constructed according to

$$\chi(\mathbf{r}) = \frac{1}{2}\ln\left[\frac{\rho_{LDA}(\mathbf{r})}{\rho_{\chi=0}(\mathbf{r})}\right] \quad , \tag{4.20}$$

where $\rho_{\chi=0}(\mathbf{r})$ is the density produced by a VMC calculation using a trial wavefunction, Ψ_T , with $\chi(\mathbf{r}) = 0$. The overall trial wavefunction should then reproduce a charge density close to the LDA form. Equation (4.20) certainly does produce a $\chi(\mathbf{r})$ that is able to reduce the energy of the system (compared to a Ψ_T with $\chi(\mathbf{r}) = 0$, see figure 4.4) and produce a charge density that is a reasonable reflection of the LDA charge density (see figure 4.3). However, there is nothing to suggest that a $\chi(\mathbf{r})$ function constructed using Eq.(4.20) is the best possible function either in terms of total energy or charge density. Indeed, there is little hard evidence [66] of exactly how accurate the LDA charge density is. It therefore seemed sensible to start optimising Ψ_T by optimising $\chi(\mathbf{r})$.

4.3.2 Choice of Parameters

In the solid QMC calculations described in chapter 3, $\chi(\mathbf{r})$ is expressed as a Fourier expansion in reciprocal lattice vectors, $\mathbf{G}$

$$\chi(\mathbf{r}) = \sum_{\mathbf{G}} \chi(\mathbf{G}) e^{i\mathbf{G}.\mathbf{r}} \quad . \tag{4.21}$$

These Fourier coefficients provide an obvious set of variational parameters to optimise the wavefunction with respect to. A total of 2554 $\mathbf{G}$-vectors was originally chosen [33] for the number of $\mathbf{G}$'s used in the expansion of $\chi(\mathbf{r})$ in the germanium solid. This is far too large a number of parameters for the optimisation procedure and so the first simplification is to reduce the number of free parameters by forcing all the coefficients of a star of $\mathbf{G}$'s to have the same magnitude. A star of $\mathbf{G}$'s is a group of $\mathbf{G}$'s related by the point group symmetry of the lattice. By grouping the $\mathbf{G}$'s in stars we can impose the full space group symmetry of the crystal on the $\chi(\mathbf{r})$ function. This grouping allows Eq.(4.21) to be rewritten as

$$\chi(\mathbf{r}) = \sum_{stars\ ,\ s} \chi_s \sum_{\mathbf{G}\ in\ s} e^{i\mathbf{G}.\mathbf{r}}.(\text{phase factor}) \quad . \tag{4.22}$$

For a crystal with the origin of coordinates at a centre of inversion symmetry, the phase factors in Eq.(4.22) are simply ± 1. Recently, Fahy et al.[67] have studied

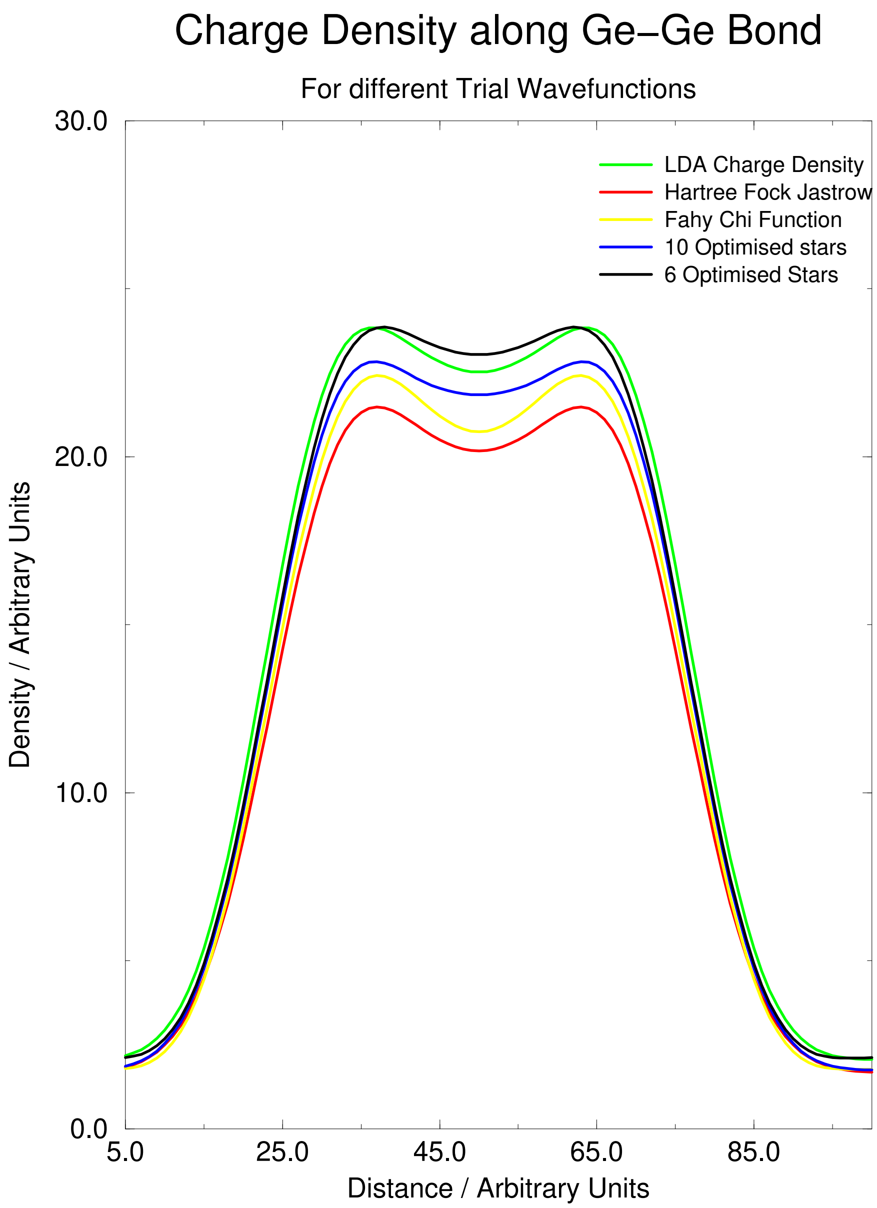


Figure 4.3: Charge density along the Ge-Ge bond for different $\chi(\mathbf{r})$ functions.

the Boron Nitride crystal within VMC. In this case the inversion symmetry no longer holds, but one can still choose a set of 'generalised stars' with which the $\chi(\mathbf{r})$ function can be represented by an independent, real, variational coefficients for each 'generalised star'. The grouping of $\mathbf{G}$ vectors into stars reduces the number of free parameters to around 130. This is still a very large space in which to perform an accurate optimisation, so it was decided to investigate the effect of reducing the number of stars in the expansion of χ, as it was suspected that stars for larger $\mathbf{G}$ vectors contained only noise. It was found that for solid germanium the $\chi(\mathbf{r})$ function could be described using 5-10 stars of $\mathbf{G}$ vectors.

4.3.3 Adding a new Function to Chi

In an attempt to further improve the quality of the trial wavefunction, it was proposed to add an extra function to $\chi(\mathbf{r})$ to increase the variational freedom available to the optimisation procedure. A spherically symmetric function centred on each of the ionic cores was chosen as this was simple to implement while still adding considerable extra variational freedom. The new function was designed to add some extra atomic quality to the trial wavefunction is these regions.

The requirements for the new function are only that it should be well behaved as an electron moves through the ionic core, i.e. there should be no cusp in $\chi(\mathbf{r})$ as $\mathbf{r}_i \to 0$.

Subject to these constraints, the most general polynomial function,

$$g(r) = \begin{cases} (L-r)^2 r^2 \sum_{i=0} \alpha_i T_i(r') + B(L-r)^2(\frac{L}{2}+r), & r < L \\ 0, & r > L \end{cases} , \tag{4.23}$$

was chosen. In Eq.(4.23), L is the range of the function, r' is a rescaled r,

$$r' = \frac{2r - L}{L} , \tag{4.24}$$

and the α_i are coefficients. Chebyshev polynomials, T_i, were chosen because they exhibit good numerical stability during the fitting procedure.

4.3.4 Results of Optimising the $\chi(\mathbf{r})$ Function in Germanium Solid

In previous work [33, 50] VMC and DMC calculations were performed on diamond-structure germanium using the techniques described in chapter 3, but variance minimisation was not used to optimise the $\chi(\mathbf{r})$ function in those calculations. The VMC calculations described in Refs.[33, 50] were repeated using the optimised $\chi(\mathbf{r})$ functions described above. A *fcc* simulation cell of diamond structure germanium containing 16 atoms was studied. The single particle orbitals were obtained from a LDA calculation using a plane-wave basis set containing all plane waves up to a kinetic energy cutoff of 40 Ry. The calculations were performed at a non-zero wavevector of the simulation cell Brillouin zone using the techniques described in chapter 3 and Refs.[33, 50].

The first set of calculations used a $\chi(\mathbf{r})$ function containing 6 stars of $\mathbf{G}$ vectors, whose coefficients had been optimised by the variance minimisation procedure, described in this chapter. The second set of optimisations of $\chi(\mathbf{r})$ were performed again with 6 stars of $\mathbf{G}$ vectors and with a further 6 parameters in the $g(r)$ function (5 for α_0 to α_4 and B). 10,000 configurations were used in both the optimisations. The configurations for the second optimisation were generated from a VMC calculation using the best $\chi(\mathbf{r})$ from the first optimisation. Values of L=3.0, 3.5, 4.0, 5.0 were used for the range of the $g(r)$ function. The results yielding the lowest variance of the local energy are illustrated in figure 4.4. This figure shows that the introduction of a $\chi(\mathbf{r})$ function according to the Fahy prescription, lowers the total energy by 0.15 eV per atom. Optimising the values of the coefficients in this $\chi(\mathbf{r})$ function using the variance minimisation scheme lowers the energy by a further 0.17 eV per atom. The introduction of $g(r)$ reduces the VMC energy by at most a further 0.05 eV per atom. The variance of the energy is only reduced by a few percent during the whole optimisation process.

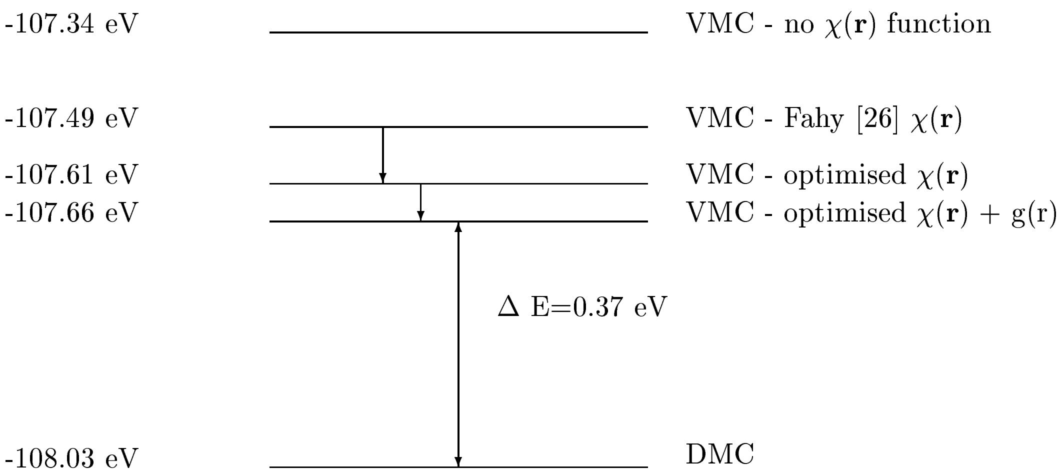


Figure 4.4: Reduction in energy from optimising the $\chi(\mathbf{r})$ function and adding g(r) to the $\chi(\mathbf{r})$ function.

4.4 Optimising the u Function

It is clear from the previous section that the optimisation procedure, Eq.(4.11), produces a trial wavefunction, Ψ_T , that has a VMC energy significantly closer to the DMC energy. However, there is still room for improvement (about 0.37 eV per atom). It was therefore decided to attempt to optimise another part of the trial/guiding wavefunction, Eq.(4.19) to see if the VMC energy could be reduced still further. The most obvious candidate for further optimisation was the Jastrow function as this also has a form in which optimisation of a relatively small set of parameters can be performed.

4.4.1 Choice of Functional Form for new term in the Jastrow Factor

In the case of optimising $\chi(\mathbf{r})$, the choice of parameters to optimise was obvious as the function is expressed as a Fourier expansion. In the case of the Jastrow function, this choice is not so clear. The current Jastrow factor has the form

$$\exp\left(-\sum_{(\uparrow,1)\leq(s,i)<(s',j)}^{(\downarrow,N)} u(r_{ij})\right) \quad , \tag{4.25}$$

where $u(r_{ij})$ is given by

$$u(r_{ij}) = \frac{A}{r_{ij}}\left(1 - e^{-\frac{r_{ij}}{F}}\right) \quad . \tag{4.26}$$

In an attempt to improve on the above function it was decided to add an extra term into the exponential in Eq.(4.25) to take account not only of the electron-electron separation, r_{ij}, but also of the individual positions of the electrons $\mathbf{r}_i$ and $\mathbf{r}_j$. It was suspected that the region where correlation effects are likely to deviate most strongly from the symmetric correlation described by the standard Jastrow function will be close to the ionic cores. Therefore, the any new function, $u_{sr}(r_i, r_j, r_{ij})$ should be short ranged and centred on each of the ions. For simplicity, the $u_{sr}(r_i, r_j, r_{ij})$ function was chosen to be a function of the distances of the 2 electrons from the ion, $|\mathbf{r}_i|$ and $|\mathbf{r}_j|$ and the electron-electron separation r_{ij}, with no angular dependence. It must also obey the following conditions:-

1. It should not cause the total Jastrow term to violate the cusp condition (see section 2.4.1), i.e. $\left.\frac{\partial u_{sr}}{\partial r_{ij}}\right|_{r\to 0} = 0$ for the extra term $u_{sr}(r_i, r_j, r_{ij})$.
2. $u_{sr}(r_i, r_j, r_{ij})$ should be well behaved as one of the electrons moves through an ion, i.e. there should be no cusp in $u_{sr}(r_i, r_j, r_{ij})$ or in the 1st derivative as $\mathbf{r}_i \to 0$.
3. $u_{sr}(r_i, r_j, r_{ij})$ should take the most general form possible, subject to the above 2 restrictions.

To check that condition 1 was satisfied, the new term, $u_{sr}(r_i, r_j, r_{ij})$ was expanded about $r_{ij} = 0$

$$u_{sr}(r_i, r_j, r_{ij}) = \alpha(r_i, r_j) + \beta(r_i, r_j)r_{ij} + \gamma(r_i, r_j)r_{ij}^2 + \dots \tag{4.27}$$

Now condition 1 specifies $\left.\frac{\partial u_{sr}}{\partial r_{ij}}\right|_{r_{ij}\to 0} = 0$, therefore

$$\frac{\partial \alpha(r_i, r_j)}{\partial r_{ij}} + \beta(r_i, r_j) = 0 \quad . \tag{4.28}$$

Finally, expanding $\frac{\partial \alpha(r_i, r_j)}{\partial r_{ij}}$ gives

$$\frac{\partial \alpha(r_i, r_j)}{\partial r_{ij}} = \frac{\partial \alpha}{\partial |r_i|}\frac{\partial |r_i|}{\partial r_{ij}} + \frac{\partial \alpha}{\partial |r_j|}\frac{\partial |r_j|}{\partial r_{ij}} \quad . \tag{4.29}$$

We chose to keep electron j fixed (i.e. $\frac{\partial r_j}{\partial r_{ij}} = 0$) and move electron i through it, to test the behaviour as $r_{ij} \to 0$. As the angle between $\mathbf{r}_i$ and $\mathbf{r}_j$ varies between 0 and π the value of $\frac{\partial r_i}{\partial r_{ij}}$ will vary smoothly between -1 and +1, (see figure 4.5).

Therefore, the only solution to Eq.(4.28) for all geometries of electrons is $\alpha = \beta = 0$.

The final form chosen for the new short range function $u_{sr}(r_i, r_j, r_{ij})$ was therefore

$$u_{sr}(r_i, r_j, r_{ij}) = r_{ij}^2 r_i^2 r_j^2 (r_i - L)^2 (r_j - L)^2 \sum_{i,j,k} \alpha_{ijk} T_i(r_i') T_j(r_j') T_k(r_{ij}') \tag{4.30}$$

The prefactor in Eq.(4.30) $r_{ij}^2 r_i^2 r_j^2 (r_i - L)^2 (r_j - L)^2$ performs the following functions; the r_{ij}^2 term removes any terms independent of r_{ij} or terms linear in r_{ij} as specified

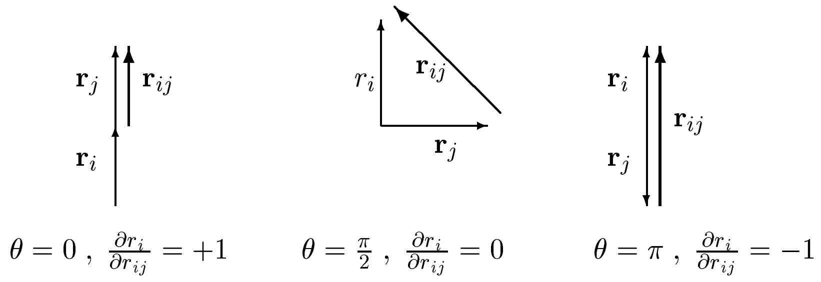


Figure 4.5: Dependence of $\frac{\partial r_i}{\partial r_{ij}}$ on the angle between $\mathbf{r}_i$ and $\mathbf{r}_j$

above. The $r_i^2 r_j^2$ term is required to satisfy condition 2, namely that the function be well behaved as one of the electrons moves through the ion. The need for this term in the prefactor was established by performing small simulations using different forms of $u_{sr}(r_i, r_j, r_{ij})$ as one electron moves through an ion. The $(r_i - L)^2(r_j - L)^2$ term enforces the short range nature of $u_{sr}(r_i, r_j, r_{ij})$ by forcing it to decay to zero with zero gradient when one of the electrons is a distance L from the ion. The remaining part of $u_{sr}(r_i, r_j, r_{ij})$ is a general Chebyshev expansion in all three variables; r_i , r_j, and r_{ij}.

It should be noted that there are in fact two separate $u_{sr}(r_i, r_j, r_{ij})$ functions required, one dealing with the case where the spins of electrons i and j are parallel and one where they are anti-parallel. This has no effect on the choice of functional form, but it does mean that there are twice as many parameters to be optimised and this reduces the maximum possible number of terms in the Chebyshev expansion.

It is also worth noting that the final form for $u_{sr}(r_i, r_j, r_{ij})$ is very similar to that proposed by Mitas [41], see Eq.(4.6). The difference between the functions is that Mitas only includes even powers of r_{ij} whereas the $u_{sr}(r_i, r_j, r_{ij})$ function used here contains odd and even powers and should therefore be more general.

4.4.2 Implementation of the new Jastrow Function

Another significant advantage of the optimisation procedure of Eq.(4.11) is that new functions such as $u_{sr}(r_i, r_j, r_{ij})$ can be tested without actually having to implement them into the mainstream solid VMC code. All one has to be able to do is calculate E_L^i and ψ_i for each configuration. The optimisation of Eq.(4.11) can then be performed to see if the proposed new function produces an improvement in the quality of Ψ_T , i.e. a reduction in the energy and variance of the energy. To calculate the new value for ψ_i with $u_{sr}(r_i, r_j, r_{ij})$ included, the Jastrow factor in Eq.(4.19) is written as

$$\exp\left(- \sum_{(\uparrow,1)\leq(s,i)<(s',j)}^{(\downarrow,N)} u(r_{ij}) - \sum_{ions,I} \sum_{r_i,r_j<L} u_{sr}(r_i, r_j, r_{ij}) \right) \quad , \tag{4.31}$$

where the sum over pairs of electrons is performed by storing lists of which electrons are within L of each ion. The intention was to then update these lists whenever an electron was moved in the VMC calculation.

To calculate the new value of the kinetic energy with $u_{sr}(r_i, r_j, r_{ij})$ included in the wavefunction, the gradient and Laplacian of the new $u_{sr}(r_i, r_j, r_{ij})$ also need to be calculated. This was done in a similar way to Eq.(4.31), by summing the contributions to the gradient and Laplacian of $u_{sr}(r_i, r_j, r_{ij})$ from each ion in the solid.

4.4.3 Results of Optimising the new Jastrow Factor

For a first test, the optimisation procedure was set up using 10,000 configurations generated by a VMC calculation, using the best parameters so far obtained for χ. The $\chi(\mathbf{r})$ function was hard wired to use those parameters. It was intended that if the inclusion of $u_{sr}(r_i, r_j, r_{ij})$ produced a large change in the wavefunction and therefore the charge density, $\chi(\mathbf{r})$ would also have to be included in the optimisation to allow it to adjust to these changes as there could be significant coupling between the $u_{sr}(r_i, r_j, r_{ij})$ and $\chi(\mathbf{r})$ functions.

Only the 0^{th} and 1^{st} orders of the Chebyshev expansion were used in Eq.(4.30). This gave a total of 12 parameters to be optimised (6 for the parallel and 6 the for anti-parallel $u_{sr}(r_i, r_j, r_{ij})$ functions). Four separate optimisations were performed for values of the range of $u_{sr}(r_i, r_j, r_{ij})$, L=3.0, 3.5, 4.0, 5.0. These correspond to an average number of electron pairs around each ion contributing to $u_{sr}(r_i, r_j, r_{ij})$ of 3, 5.5, 7, 12 respectively.

The overall size of $u_{sr}(r_i, r_j, r_{ij})$ naturally increased with L as more electron pairs were included, but the optimised parameters were still very small compared to the original part of the u function. The result with the lowest variance of the local energy was achieved by choosing L=4.0 (larger values lead to instabilities in the optimisation procedure.) This produced a lowering of the energy beyond the optimised $\chi(\mathbf{r})$ energy of about 0.01eV per atom. As with the optimisation of $\chi(\mathbf{r})$, the variance again showed little change.

4.4.4 Removing the $|\mathbf{r}_i|^2|\mathbf{r}_j|^2$ Prefactor in $u_{sr}(r_i, r_j, r_{ij})$

One possible explanation for the new $u_{sr}(r_i, r_j, r_{ij})$ function of Eq.(4.30) in the Jastrow factor being unable to reduce the energy and variance of energy significantly is that there are too many high powers in the prefactor of the Chebyshev expansion and these reduce its ability to produce good results. As the electron density close to the ion cores is very low compared to that in the bond (see Figure 4.3), it is possible that removing the $|\mathbf{r}_i|^2|\mathbf{r}_j|^2$ prefactor to the Chebyshev expansion will not cause any instability. The argument here being that even though without this prefactor there is a cusp in $u_{sr}(r_i, r_j, r_{ij})$ when one electron moves through the ion, as the density close to the ion is so low, this is a very unlikely event.

The optimisations performed above were repeated for the $u_{sr}(r_i, r_j, r_{ij})$ function with no $r_i^2 r_j^2$ term in the prefactor.

$$u_{sr}(r_i, r_j, r_{ij}) = r_{ij}^2 (r_i - L)^2 (r_j - L)^2 \sum_{i,j,k} \alpha_{ijk} T_i(r_i') T_j(r_j') T_k(r_{ij}') \tag{4.32}$$

The above assumption proved to be correct in that this did not appear to produce any serious instability into the optimisation. However, it also failed to produce any significant further reduction in either the energy or the variance of the energy.

4.5 A New u function

In a parallel development described in chapter 5 a new method for evaluating the expectation value of the Coulomb interaction was introduced which relies on the short range of the exchange-correlation hole [3]. This method allows accurate results to be obtained without summing the interactions over periodic images of the simulation cell, thereby reducing the computational cost significantly and leaving the evaluation of the Jastrow factor as the most time consuming part of the calculation. Therefore, it was decided to seek a new form for u which is as accurate as the previous form in Eq.(4.19), yet convenient for optimisation purposes and can be evaluated rapidly within a QMC calculation.

4.5.1 Form of the New u function

The new u function is similar to one used earlier for the HEG by Ortiz and Ballone [68, 69]. In common with Ortiz and Ballone a spherically symmetric u function is chosen, which is short ranged so that it need not be summed over simulation cells. This u function folds in the long range behaviour of the Jastrow factor in an approximate manner, and therefore it depends on the size of the simulation cell as well as on the electron density of the system. For each electron pair the separation vector $\mathbf{r}_{ij}$ is reduced to its minimum length (by subtraction of supercell lattice vectors) giving the vector between electron i and the nearest periodic image of electron j. This reduction procedure is illustrated in figure 4.6.

The precise form of the new u is different from that used by Ortiz and Ballone. It has certain advantages which will be described below. We demand that u obeys the following conditions:

i. $u(r)$ satisfies the cusp conditions as $r \to 0$;

ii. $u(r)$ is continuous and has a continuous first derivative for all $r > 0$;

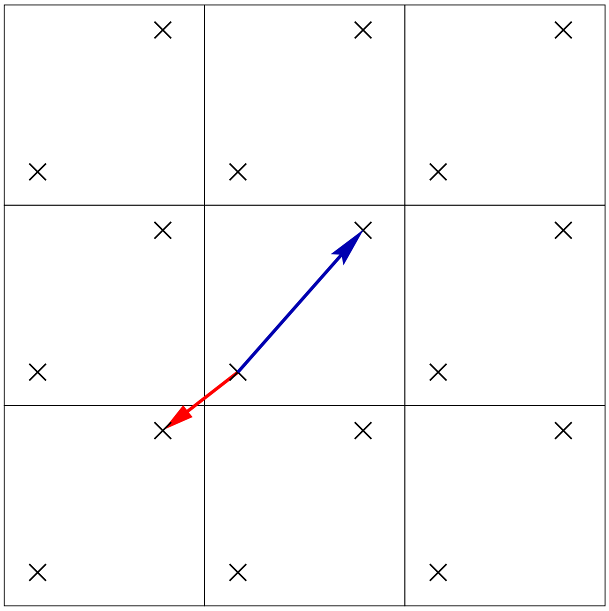


Figure 4.6: Reduction of the vector $\mathbf{r}_{ij}$ to its minimum length. The figure contains a square simulation cell and just one of the periodic images in each direction. The blue vector shows the original vector. The red vector has been been reduced to its minimum length by subtraction of a vertical and a horizontal lattice vector.

iii. $u(r)$ is linear in the variable parameters.

The only condition that $u(r)$ must satisfy for our QMC procedures to work is condition (ii) given above. If this condition is not obeyed then the kinetic energy estimator, $T_i = \frac{1}{4}\nabla_i^2 \ln \Psi$, will have δ-functions at the discontinuities, which will be missed by the sampling procedure. To ensure continuity of the first derivative of $u(r)$ for $r > 0$ it is required that u' goes (almost exactly) to zero at the surface of the sphere of radius L_{WS} inscribed within the Wigner-Seitz cell of the simulation cell. For $r > L_{WS}$, $u(r)$ and $u'(r)$ are set to zero. The cusp conditions are imposed on the first derivative of u at $r \to 0$ because this is a property of the exact wavefunction. In contrast to Ortiz and Ballone, continuity of the second derivative of u is not imposed. We write $u(r)$ as

$$u(r) = u_0(r) + f(r), \tag{4.33}$$

where u_0 is a fixed function and f contains the variable parameters. f is expanded as a linear sum of some basis functions, f_k:

$$f(r) = \sum_k b_k f_k(r). \tag{4.34}$$

For the fixed part of u, the following form was chosen,

$$u_0(r) = \frac{A}{r}\left(1 - \exp(-\frac{r}{F})\right) \exp\left(-\frac{r^2}{L_0^2}\right), \tag{4.35}$$

where F is chosen so that the cusp condition is obeyed and L_0 is chosen so that $u_0(L_{WS})$ is effectively zero ($< 10^{-6}$). Typically $L_0 = 0.25 L_{WS}$ and A is fixed by the plasma frequency[25]. The function u_0 is chosen to give a good description of the correlation so that the variable part of u is small. For the variable part we choose

$$\begin{array}{rcll} f(r) & = & B(\frac{L_{WS}}{2} + r)(L_{WS} - r)^2 & 0 \le r \le L_{WS} \\ & + & r^2 (L_{WS} - r)^2 \sum_{l=0}^{M} \alpha_l T_l(\overline{r}) & , \\ & = & 0 & r > L_{WS} \end{array} \tag{4.36}$$

where B and the α_l are variational coefficients, T_l is the lth Chebyshev polynomial, and

$$\overline{r} = \frac{2r - L_{WS}}{L_{WS}} \quad , \tag{4.37}$$

so that the range $(0, L_{WS})$ is mapped into the orthogonality interval of the Chebyshev polynomials, $[-1, 1]$. The use of Chebyshev polynomials rather than a simple polynomial expression improves the numerical stability of the fitting procedure. The function f is the most general polynomial expression containing powers up to r^{M+4} which satisfies the following conditions:

i. $f'(0) = 0$;

ii. $f(r \geq L_{WS}) = 0$;

iii. $f'(r \geq L_{WS}) = 0$.

Condition (i) ensures that $u(r)$ obeys the cusp conditions, which are incorporated in $u_0(r)$. Addition of a constant to $u(r)$ changes the normalisation of the wavefunction but not its functional form, and condition (ii) eliminates this unimportant degree of freedom. Condition (iii) ensures continuity of the first derivative of u at $r = L_{WS}$.

To start the optimisation process we perform a VMC run to produce the electron configuration data for the initial distribution $|\Psi_T^0|^2$ as described in section 4.2.6. For each electron configuration, $u(r)$ is summed over all distinct pairs of electron coordinates i and j in the simulation cell (with the separation vector reduced into the Wigner-Seitz simulation cell). For each configuration the following summation is performed

$$\begin{aligned} \sum_{i>j} u(r_{ij}) &= \sum_{i>j} u_0(r_{ij}) + \sum_k b_k \sum_{i>j} f_k(r_{ij}) \\ &= const. + \sum_k b_k g_k. \end{aligned} \tag{4.38}$$

Instead of storing the individual electron coordinates in each configuration we store the g_k, which is sufficient because the functional form for u is linear in the variable parameters. This reduces the storage and CPU time needed for the minimisation procedure, which requires no further summations over the electron coordinates when the values of the parameters, b_k, are altered. The first and second derivatives of u, which enter the expression for the energy, are dealt with in a similar manner. These savings are very significant when dealing with a large number of electrons in the simulation cell, and for the HEG we have performed full minimisations with up to 338 electrons.

4.5.2 Tests on Jellium

Wavefunctions were optimised for the HEG at a range of densities from $r_s = 0.1$ to $r_s = 10$. Excellent results were obtained at all densities, but for brevity only the results for r_s=1 are presented here. A wavefunction of Slater-Jastrow type (cf. Eq.(4.19)) was used, where the determinants, $D^{\uparrow(\downarrow)}$, were constructed from the lowest energy plane waves at zero wavevector within the simulation cell Brillouin zone. The one-body χ function was set to zero and the u function of Eqs.(4.33-4.37) was used. Separate u functions for parallel and antiparallel spins were used for *fcc* simulation cells containing N=30, 54, 178 and 338 electrons. In each case the numbers of up- and down-spin electrons were equal. Typically 10,000 electron configurations were sampled from a VMC run of sufficient length to ensure that the chosen configurations are statistically independent as described in section 4.2.6.

The variance minimisation procedure is stable for small N, but gradually becomes unstable as N increases. The technique described in section 4.2.4, was used to control this instability by fixing the reweighting factors to unity and regenerating configurations several times. This proved to be completely successful for all the system sizes studied. All calculations used 9 Chebyshev polynomials to represent $f(r)$, which tests show to give essentially complete convergence for the systems studied. The minimisation problem then has 20 parameters.

Table 4.5.2 shows the energy and standard deviation, σ, of the energy as a function of system size[2], comparing our u function Eqs.(4.33-4.37) with that of Eq.(4.19), which

N	E_{VMC} Eqs.(4.25-4.26)	σ Eqs.(4.25-4.26)	E_{VMC} Eqs.(4.33-4.37)	σ Eqs.(4.33-4.37)	E_{DMC}
30	0.4657	0.2	0.46979	0.22	
54	0.6085	0.24	0.6110	0.22	0.6069
178	0.6161	0.17	0.61679	0.17	0.6141
338	0.5772	0.14	0.57707	0.17	

Table 4.1: Energies, E, and standard deviations of the energy, σ, for the HEG at a density of $r_s = 1$ as a function of the number of electrons in the simulation cell, N. All entries are in Hartree atomic units per electron. The VMC energies are calculated with the Yukawa form of the Jastrow factor (Eqs.(4.25-4.26)) and our optimised spherically symmetric form (Eqs.(4.33-4.37)). The DMC energies do not depend on which of the two Jastrow factors is used.

includes a sum over simulation cells, and with DMC results. For the DMC calculations we used a time step of 0.01 au and an average population of 640 configurations. After equilibration the averages were collected over 5000 moves of all the electrons. The results obtained using our new u function are of similar quality to those obtained with the u function of Eqs.(4.25-4.26), but the new u function is much faster to evaluate. In figure 4.7 we show the optimised spin-parallel u function for $N = 338$, together with the u function of Eqs.(4.25-4.26) which is plotted in the [100] and [110] directions (for all other directions the u function lies between the values in these directions). In figure 4.7 the derivatives of the functions are shown. These figures show that the two functions are similar, but the optimised u function exhibits a slightly smaller derivative at intermediate distances. The optimised spin-antiparallel u function shows similar behaviour.

The reduction in computing cost from using the new u function is very significant. It is particularly effective when combined with our recently developed technique for evaluating the expectation value of Coulomb interactions in homogeneous systems [3], (see chapter 5). This combination of techniques entirely eliminates the need for time-consuming sums over simulation cells, and the resulting algorithm is extremely fast, with the most costly remaining operation being the calculation of determinants

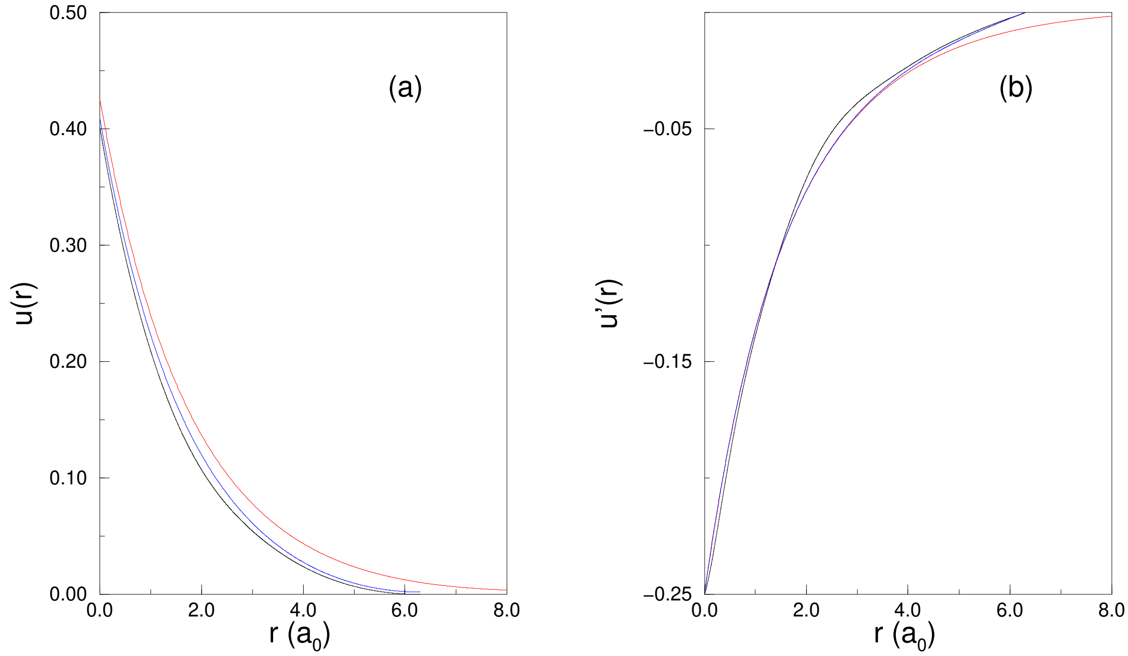


Figure 4.7: Comparison of spin-parallel u functions for the HEG at $r_s = 1$. The optimised function (black line) is shown along with the Ewald summed Yukawa form along the [100] direction (red line) and the [110] direction (blue line). Fig. 4.7a shows the u functions themselves while Fig. 4.7b shows the first derivatives.

which are evaluated for each electron move using the standard Sherman-Morrison[18] formula to update the matrix of cofactors.

4.6 Applying the New u Function to Solids

Having studied the HEG, the newly developed u function was applied to a crystalline solid. To enable direct comparison with the previous results, germanium in the diamond structure was used as a test material. The same *fcc* simulation cell of diamond structure germanium containing 16 atoms was studied. The same single-particle orbitals were used to construct the Slater determinant. The χ function was chosen to have the full symmetry of the diamond structure. Again the χ function was expanded in a Fourier series, grouping the **G** vectors into stars as in Eq.(4.22).

For the u function, the functional form of Eqs.(4.33-4.37) which was developed for the HEG was chosen. The u and χ functions were optimised simultaneously because they are strongly coupled. Typically 6 non-zero coefficients in Eq.(4.22) for the χ function and 8 parameters for both the parallel- and antiparallel-spin u functions in Eq.(4.36) were used, giving a total of 22 parameters in the minimisation problem. Variance minimisations were carried out using 10,000-100,000 independent N-electron configurations, which were regenerated several times. The final energy of -107.69 $\pm$ 0.01 eV per atom is 0.08 eV lower than the result obtained using the (Ewald summed) Yukawa potential of Eq.(4.25) and the variance minimisation procedure for χ, and 0.20 eV lower than the result obtained in our previous work using the Yukawa potential and Fahy's original prescription for χ [48, 26]. The energy of -107.69 eV per atom is only 0.34 eV per atom higher than the DMC result for this system of -108.03 $\pm$ 0.07 eV per atom quoted in Table I of Ref. [50]. (As discussed in Refs. [33, 50], we estimate that about 0.12 eV of this energy difference is due to the basis set incompleteness error in the single-particle orbitals, which affects the VMC much more than the DMC result, and which could be eliminated by the use of a larger basis set or a smoother pseudopotential.)

The optimised spin-parallel and spin-antiparallel u functions for germanium are similar to the Yukawa form in all directions. However, they have a smaller derivative at intermediate distances, exactly as observed in the HEG (see figure 4.7). The optimised χ function differs significantly from the original Fahy form, with some parameters changing by an order of magnitude. Altering the number of parameters in the optimisation scheme revealed that 6 non-zero coefficients was again sufficient to converge the χ function.

4.7 Optimising Wavefunctions for Atoms

The energy calculated within a VMC framework is an upper bound to the true energy of the system. When calculating energy differences it is important to use wavefunctions of similar quality for the different systems so that there is a strong cancellation of errors. This feature is common to many electronic structure methods, for instance in bandstructure methods the errors due to incomplete basis sets are often large, but if sufficient care is taken it is often possible to calculate energy differences with high precision. A severe test of such cancellation is provided by the cohesive energy of a solid, which is the difference in energy between the solid and its constituent atoms. This section describes calculations performed for the ground-state energy of the germanium pseudo-atom using the same pseudopotential as for our solid calculations, and using a wavefunction of, as near as possible, the same quality as in the solid calculations. This allows us to test the extent to which the remaining errors/omissions in the trial/guiding wavefunction cancel between atomic and solid germanium.

4.7.1 Choice of Atomic Wavefunction

The ground state of the germanium atom is a ^{3}P configuration which can be described by a single determinant. The orbitals for the determinant were obtained from a LSDA calculation for the atomic ground state. A spherically-symmetric χ function was used for both up- and down-spin electrons of the form of the function f of Eq.(4.36) with

L_{WS}=10 a_0. It is appropriate to choose this form for a pseudo-atom because the lack of electron-ion cusp condition implies that χ should be smooth at the origin. For the u functions we used the same form as for the HEG and germanium solid calculations, Eqs.(4.33-4.37), with L_{WS}=10 a_0.

4.7.2 Results

Convergent results were obtained using 8 parameters for both the χ and u functions. The energy obtained with the optimised wavefunction was -103.22 $\pm$ 0.01 eV and the standard deviation was 0.52 eV. The χ functions for up- and down-spin electrons were found to be very similar. This result is perhaps surprising because there is a single down-spin electron in an s-orbital, while there are three up-spin electrons, one in an s-orbital and two in p-orbitals. This point was investigated further by performing calculations using a single χ function, which gave an energy of -103.20 $\pm$ 0.01 eV, and a standard deviation of 0.54 eV, which are almost identical to the values obtained using separate functions for up- and down-spin electrons. The resulting χ and u functions were almost unchanged.

Figure 4.8 shows LDA and QMC charge densities for the pseudo-atom, calculated by Alan James from Imperial College. In figure 4.8a the VMC charge density from a wavefunction consisting of a determinant of LDA orbitals and a u function, but no χ function, is compared with the LDA density. This shows that the inclusion of the correlation factor, u, smears out the charge density considerably. In Fig. 4.8b we plot the charge density from an optimised wavefunction, containing both u and χ. This plot shows that the variance minimisation procedure results in a wavefunction whose charge density is very close to the LDA form. In figure 4.8c we plot the DMC charge density calculated with the optimised wavefunction as guiding function. The DMC charge density is very close to both the VMC and LDA charge densities, which shows that the true charge density is close to the LDA form. It seems that the physical idea behind the original Fahy prescription for χ, i.e. returning the charge density to the LDA form, is extremely good. However, the number of configurations required to fit the variational parameters in a well parameterised wavefunction via variance minimisation is much less than the number required to obtain an accurate charge density. Therefore it appears that for a given computational effort it is more efficient to generate χ via variance minimisation than to construct the charge density and generate χ from it. Moreover, the method of variance minimisation is more general in the sense that it does not rely on the separate determination of an accurate charge density from, for example, the LDA.

4.7.3 Cohesive Energies

In order to compare the calculated cohesive energy with experiment various correction terms must be added to the solid calculations to account for: (i) Coulomb finite-size effects; (ii) single-particle finite-size effects; (iii) the use of a local pseudopotential; and (iv) zero-point motion. These corrections are discussed in detail in reference [33]. All these corrections are essentially independent of the optimisation of the wavefunction and it therefore suffices to compare directly the VMC and DMC results for the same system. A DMC calculation for the germanium pseudo-atom gave an energy of -103.42 $\pm$ 0.03 eV, which is only 0.20 eV below our best VMC result, while for the solid the VMC calculation (without u_s) gave an energy 0.34 eV above the DMC result. Therefore the VMC cohesive energy, of 3.80 ev per atom, is only 0.14 eV less than the DMC result, of 3.85 eV per atom, which amounts to an error of only 4%. As mentioned earlier we believe that 0.12 eV of the difference in energy between the VMC and DMC results for the solid is due to the incomplete basis set used for the single particle orbitals. If this is corrected for, the VMC cohesive energy differs from the DMC value by only 1%. This indicates that although there is still a significant difference in the VMC and DMC energies, the difference in the atomic energies transfers almost completely into the solid calculations and therefore cancels out of the cohesive energies.

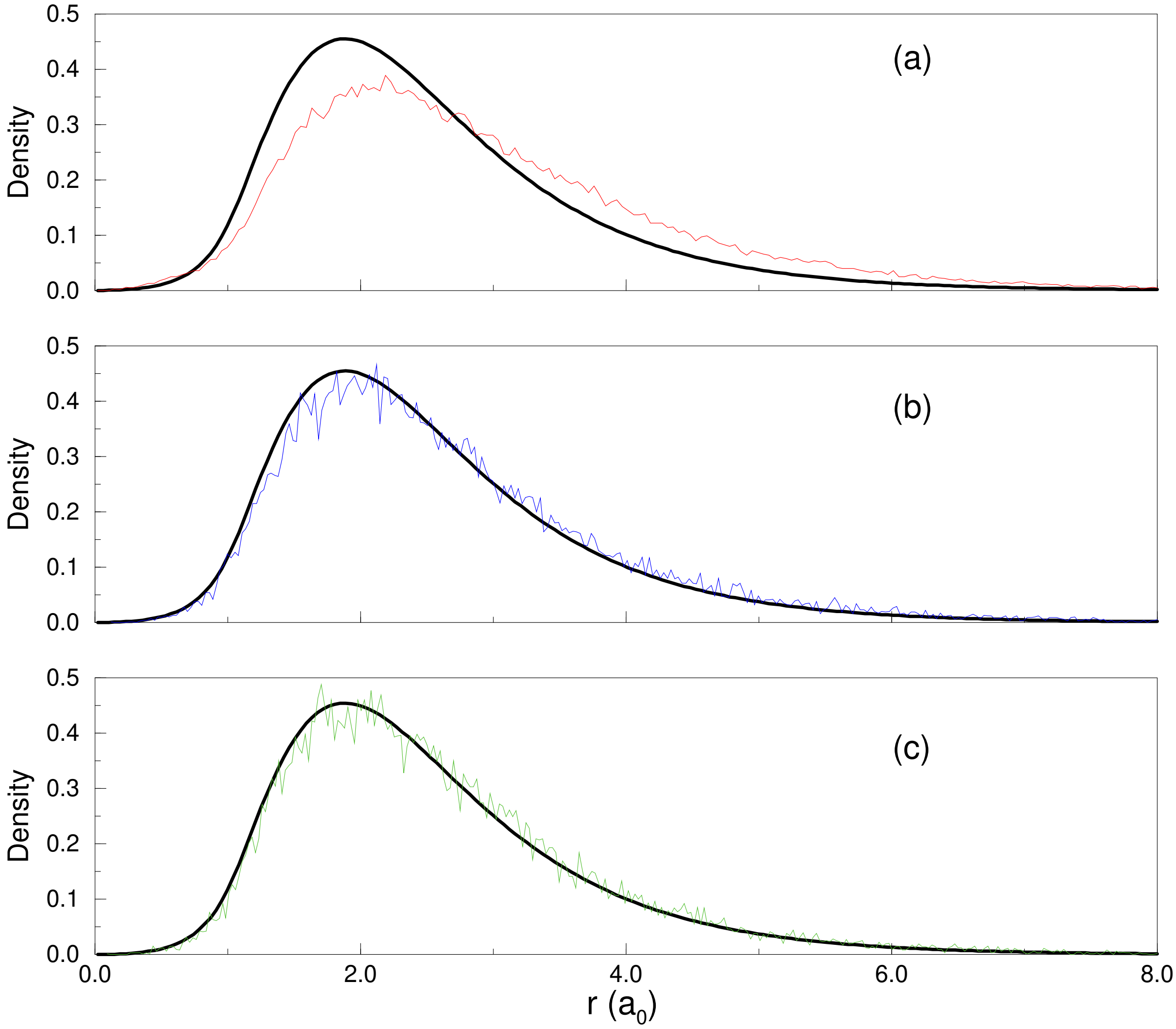


Figure 4.8: Comparison of QMC densities (coloured lines) and the LDA density (black lines) of the germanium pseudo-atom. Fig.a shows the VMC density (red line) calculated using a wavefunction consisting of a determinant of LDA orbitals and a u function, but with no χ function. Fig.b shows the VMC (blue line) density calculated using a wavefunction containing optimised u and χ functions. Fig.c shows the DMC density (green line) calculated using the same optimised wavefunction as the guiding wavefunction.

4.8 Variance Minimisation on Parallel Computers

An important factor in determining the accuracy of the wavefunction obtained by variance minimisation is the size of the ensemble of independent configurations used in the optimisation procedure, Eq.(4.11). During this work large ensembles of up to 100,000 configurations have been used. Each of these ensembles contained 64 electrons and had 20-40 parameters associated with it. The CPU and memory requirements for these problems are such that the advantages of using a parallel computer are considerable.

Therefore a parallel version of the variance minimisation procedure was developed using the "master-slave" programming model where one processor, the "master", delegates work to the other processors, the "slaves". The master processor sends work to the slaves who complete the required work and return the results back to the master. The numerical optimisation routine runs on the master processor and the ensemble of configurations is divided out among the slaves. The master processor broadcasts the values of the variational parameters to the slaves. Each of the slaves evaluates the reweighting factors, $w(\alpha)$, and the contributions to the variance (see Eq.(4.11) and Figure 4.9) for its subset of configurations. These are returned to the master which, via the NAG minimisation routine, determines new values for the variational parameters. The procedure is repeated until the imposed limit on the change in the values of the parameters is reached, or until a minimum in the variance is found.

4.9 Variance Minimisation with Non-Local Pseudopotentials

In the following chapter, results are presented for calculations performed on solid silicon in the diamond structure. The variance minimisation process for silicon is very similar to that for germanium except that a non-local pseudopotential is used in the silicon calculations. The inclusion of a non-local pseudopotential complicates the

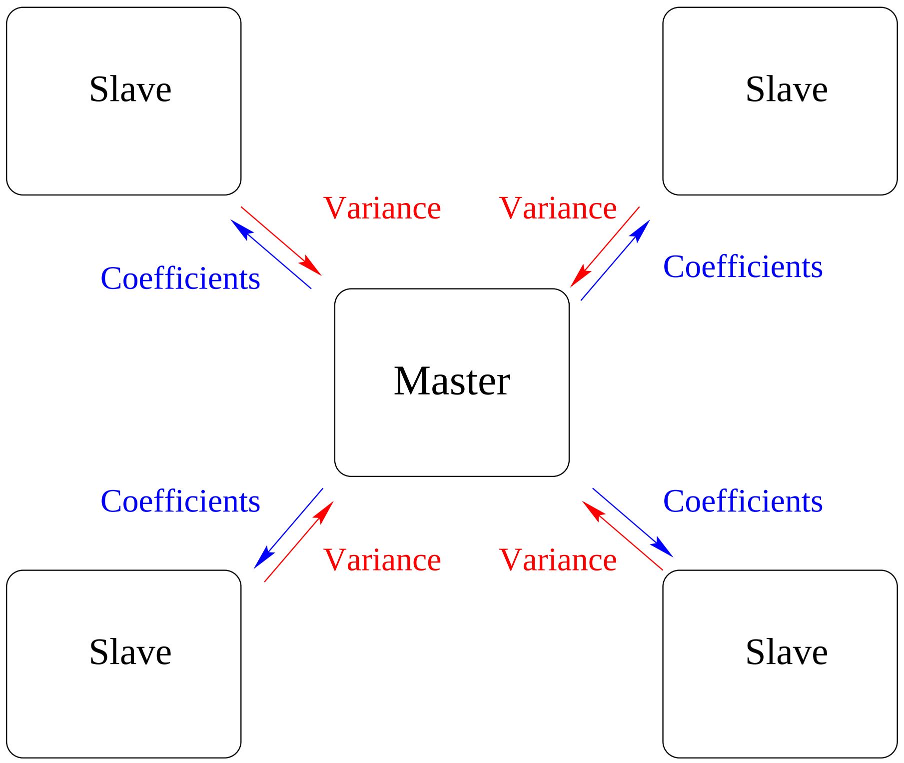


Figure 4.9: Variance minimisation on a parallel machine using the "master-slave" programming model. The master processor runs the numerical optimisation routine and farms out the evaluation of the variance of each configuration to the slave processors.

optimisation process somewhat because its expectation value depends of the value of the wavefunction throughout all space. This means that the value of the non-local potential energy cannot be stored as a constant during the optimisation procedure in the same way as the local potential energy. Any change in the values of the variational parameters in the wavefunction will effect the value of the non-local energy. This change then requires a full $3N$-dimensional integral to be evaluated for each change in the parameters. Several different techniques have been experimented with to try and deal with the non-local pseudopotential.

4.9.1 Keeping the Non-Local part Fixed during Optimisation

As the expectation value of the non-local pseudopotential taken with respect to the trial wavefunction is a small fraction of the total energy,[1] one option is to store the non-local energy associated with each configuration at the beginning of the optimisation procedure and keep it fixed. Care has to be taken when adopting this method to be sure to evaluate the non-local energy to sufficient accuracy at the beginning of the optimisation procedure. During a VMC calculation, the non-local pseudopotential is evaluated by the method proposed by Fahy et al. [26]. The non-local Hamiltonian for the ith electron is given by

$$\hat{H}^i_{nl} = \sum_l \int_0^\infty dr V_l(r) P^i_{l,r} \quad , \tag{4.39}$$

where the ion is at the origin. $P^i_{l,r}$ is the angular momentum l projection operator acting at a distance r from the origin;

$$P^i_{l,r} : \Psi(\mathbf{r}^\uparrow_1, ..., \mathbf{r}^s_i, ..., \mathbf{r}^\downarrow_N) \rightarrow \delta(r^s_i - r) \sum_{m=-l}^{l} Y_{lm}(\Omega_{\mathbf{r}^s_i}) \int_{r'=r} d\Omega_{\mathbf{r}'} Y^*_{lm}(\Omega_{\mathbf{r}'}) \Psi(\mathbf{r}^\uparrow_1, \ldots, \mathbf{r}^s_i = \mathbf{r}', \ldots, \mathbf{r}^\downarrow_N) \quad . \tag{4.40}$$

The non-local energy is then evaluated along the Monte Carlo random walk of points sampled from $|\Psi|^2$ according to

[1]For silicon the total energy per atom is of the order of -107eV and the contribution from the non-local part of the pseudopotential is of the order of 2eV per atom (with p local).

$$E_{nl} = \sum_i \frac{< \Psi|\hat{H}^i_{nl}|\Psi >}{< \Psi|\Psi >} = \int \frac{d\mathbf{R}}{\int d\mathbf{R}|\Psi|^2}|\Psi|^2 \sum_i \left(\frac{\hat{H}^i_{nl}\Psi}{\Psi} \right) \quad . \tag{4.41}$$

The z-axis is chosen along $\mathbf{r}^s_i$ to use the fact that $Y_{lm}(0,0) = 0$ for $m \neq 0$ to simplify $\frac{\hat{H}^i_{nl}\Psi}{\Psi}$ such that

$$\frac{\hat{H}^i_{nl}\Psi}{\Psi} = \sum_l V_l(\mathbf{r}) Y_{l0}(0,0) \int_{\mathbf{r}'=\mathbf{r}} Y^*_{l0}(\Omega_{\mathbf{r}'}) \frac{\Psi(\mathbf{r}^\uparrow_1, \ldots, \mathbf{r}^s_i = \mathbf{r}', \ldots, \mathbf{r}^\downarrow_N)}{\Psi(\mathbf{r}^\uparrow_1, \ldots, \mathbf{r}^s_i = \mathbf{r}, \ldots, \mathbf{r}^\downarrow_N)} d\Omega_{\mathbf{r}'} \quad . \tag{4.42}$$

The integral over $\Omega_{\mathbf{r}'}$ in Eq.(4.42) is performed statistically with the QMC code. A series of points are sampled on a grid surrounding each ion and the ratio of the value of the wavefunction with all other electrons fixed, and the i^{th} electron at $\mathbf{r}$ and at each of the grid points is used to evaluate Eq.(4.42).

During a VMC calculation, it is not necessary to sample enough points from the spherical grid to evaluate the non-local integral to high precision. Any variance in the value of Eq.(4.42) will be averaged out over the duration of the run. However, if the non-local energy is to be kept fixed throughout a variance minimisation run, it is important to ensure the sampling in the non-local integral is sufficient for each *individual* value of the non-local energy to have a small variance. Typically, up to 8 times as many sampling points are used to ensure accurate individual values of the non-local potential energy compared with a normal VMC calculation.

4.9.2 Evaluating the non-local Integral during Optimisation

If the non-local energy is strongly coupled to the expectation values of other observables via the trial/guiding wavefunction, then even though the non-local potential energy itself is a small fraction of the total energy, it is necessary to include it in the variance minimisation procedure. This can be achieved by storing all the information required to evaluate the change in the ratio of the wavefunction as the variational parameters change, at each of the grid points on which the original value of the non-local energy is stored. Schematically, one can write the total non-local energy for a single configuration as

$$E_{nl}^{\text{config}} = \sum_{\text{electrons, } i} \sum_{\text{ions, } I} \sum_{\text{grid points, } g} \sum_{l} \frac{\Psi(\mathbf{r}'_{iIgl})}{\Psi(\mathbf{r}_{iIgl})} \quad . \tag{4.43}$$

All the other electrons in the ratio $\frac{\Psi(\mathbf{r}')}{\Psi(\mathbf{r})}$ have been dropped from the notation and are assumed to be kept fixed. Eq.(4.43) can be simplified by using the fact that the non-local pseudopotential is relatively short ranged. Each electron only feels the effect of the non-local pseudopotential from one or two ions. The first two summations can therefore be replaced with a summation over *hits*. A *hit* is the event where an electron in the configuration is close enough to an ion to feel the effect of its non-local pseudopotential:

$$E_{nl}^{\text{config}} = \sum_{\text{hits, } h} \sum_{\text{grid points, } g} \sum_{l} \frac{\Psi(\mathbf{r}'_{hgl})}{\Psi(\mathbf{r}_{hgl})} \quad . \tag{4.44}$$

In a similar approach to that of Eq.(4.38) the contribution to the ratio $\frac{\Psi(\mathbf{r}')}{\Psi(\mathbf{r})}$ for each grid point of each hit in each configuration from the $\chi(\mathbf{r})$ function, u function and determinant are stored at the start of the optimisation procedure. For example, the contribution to each ratio of $\chi(\mathbf{r})$ functions can summed over all the $\mathbf{G}$ vectors in each star, s, such that

$$\chi_{hgs}^{\text{store}} = \sum_{\mathbf{G}\ in\ s} \left(\cos(\mathbf{G}.\mathbf{r}') - \cos(\mathbf{G}.\mathbf{r})\right) \quad . \tag{4.45}$$

The ratio of the $\chi(\mathbf{r})$ function at the points $\mathbf{r}$ and $\mathbf{r}'$ can then be quickly reconstructed for each set of trial parameters χ_s from

$$\chi_{h,g}^{\text{ratio}} = \exp\left(\sum_{s} \chi_s \chi_{hgs}^{\text{store}}\right) \quad . \tag{4.46}$$

A similar method of storage is adopted for the u function and the Slater determinant.

4.10 Limits of Variance minimisation

The preceding sections have outlined the development of the variance minimisation procedure. They describe how to perform optimisations on progressively more complicated systems such as those with new u functions and non-local pseudopotentials and how to increase the accuracy of the optimisation procedure by increasing the size of the ensemble of configurations and using parallel computers.

Variance minimisation itself is only a means to an end. Namely, providing wavefunctions for VMC and DMC calculations.

4.10.1 VMC Calculations

In the case of systems where the intention is only to perform VMC calculations, the above optimisation scheme has proved to be essential. When using the new u function from section 4.5, there is no approximate way to obtain the values of the variational parameters in the way that the original Fahy prescription allowed $\chi(\mathbf{r})$ functions to be approximately calculated. Therefore some form of variance minimisation is required to produce the wavefunction. To achieve state of the art accuracy in these VMC calculations it is clear that significant effort is required to ensure the quality of the trial wavefunction before performing the calculation. As the accuracy of the final VMC calculation is completely controlled by the accuracy of trial wavefunction, it is necessary to perform the optimisation procedure over an ensemble of configurations of equivalent size to the number of moves in the final VMC calculation. If several regenerations of configurations are required, this leaves the wavefunction optimisation as the most computationally expensive part of performing a VMC calculation.

4.10.2 DMC Calculations

If the intention is to use variance minimisation and then VMC as an initial guide to performing more accurate DMC calculations then the situation is not so clear cut. Only the variance of the expectation value of an operator is dependent on the quality of the guiding wavefunction, not the expectation value itself.[2] It is therefore not necessary to perform the optimisation to the same level of accuracy as for the VMC calculation. Indeed, the optimisation process can be more computationally expensive than the DMC run itself. This balance in effort spent on optimising the initial trial/guiding wavefunction versus time spent performing actual calculations will be returned to in chapter 6, where DMC calculations are performed on a series of systems describing excited states without the need for re-optimising the wavefunction for each system.

[2]The fixed-node approximation does depend on the guiding wavefunction. None of the optimisation here changes the nodal structure because the Slater determinant is not being optimised.

Chapter 5

Finite Size Effects

5.1 Introduction

As has been shown in earlier chapters, many-body simulation techniques such as the variational and diffusion Quantum Monte Carlo (QMC) methods are capable of yielding highly accurate results for correlated systems. In chapter 3 the technique of modelling large systems using a finite simulation cell subject to periodic boundary conditions was introduced. The use of a finite cell introduces “finite size effects” which are often very important, particularly for systems with long ranged interactions such as the Coulomb interaction. In this chapter a method is introduced for dealing with long ranged interactions in quantum many-body simulations which greatly reduces these finite size effects.

The finite size effects encountered in QMC calculations for electronic systems can be divided into two terms: (i) the independent particle finite size effect (IPFSE)[33, 50], and (ii) the Coulomb finite size effect (CFSE) [3, 46]. The IPFSE and CFSE are most easily defined with reference to results of local density approximation (LDA) calculations. The IPFSE is the difference between the LDA energies per atom in the finite and infinite systems and the CFSE is the remainder of the finite size error.

Recently a method was introduced [33, 50] for reducing the IPFSE in insulating systems by using the "special k-points" method borrowed from bandstructure theory [70, 54]. This has already been described in detail in chapter 3, section 3.2. This method reduces the IPFSE by an order of magnitude in insulators and leaves the CFSE as the dominant finite size effect. The CFSE, which is the subject of this chapter, arises from the long range of the Coulomb interaction and is therefore of wide significance in many-body simulations.

5.2 Motivation

The CFSE is illustrated in figure 5.1. This shows the total energy calculated using VMC for diamond-structure silicon using a finite simulation cell with periodic boundary conditions plotted for simulation cells containing multiples, n=2,3,4,5 of primitive lattice cells, which corresponds to 16,54,128 and 250 atoms, respectively. For each of the calculations, a full Hartree-Fock-Jastrow-Chi wavefunction as in Eq.(2.31) was used where the χ and u functions were optimised using the variance minimisation techniques described in chapter 4. Six stars of $\mathbf{G}$ vectors were used to describe the χ function and 22 variational parameters were used in the u function. The single particle orbitals used to construct the Slater determinant have $\mathbf{k}$-points chosen from a reciprocal space grid that is offset from the origin by $\mathbf{G}_{111}/2 = (\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3)/2$ sampling, where the $\mathbf{b}_i$ are the primitive reciprocal lattice vectors of the supercell, i.e. L-point sampling. This choice of sampling almost totally removes the IPFSE, leaving the CFSE as the dominant finite size effect. A norm-conserving, non-local pseudopotential was used to describe the silicon cores. The extensions to the variance minimisation scheme introduced in chapter 4 to deal with non-local pseudopotentials were experimented with in both the simple "fixed non-local" and "full optimisation" forms. The two resulting wavefunctions produced variational energies that were indistinguishable at the level of the statistical noise. This suggests that in silicon the "fixed non-local" approximation in the variance minimisation scheme is sufficient.

Figure 5.1 shows the total energy per atom asymptotically approaching a value that can be taken as the energy per atom in the bulk solid. The CFSE can then be defined as the difference in the total energy (once any IPFSE have been removed) at a specific system size and this bulk value.

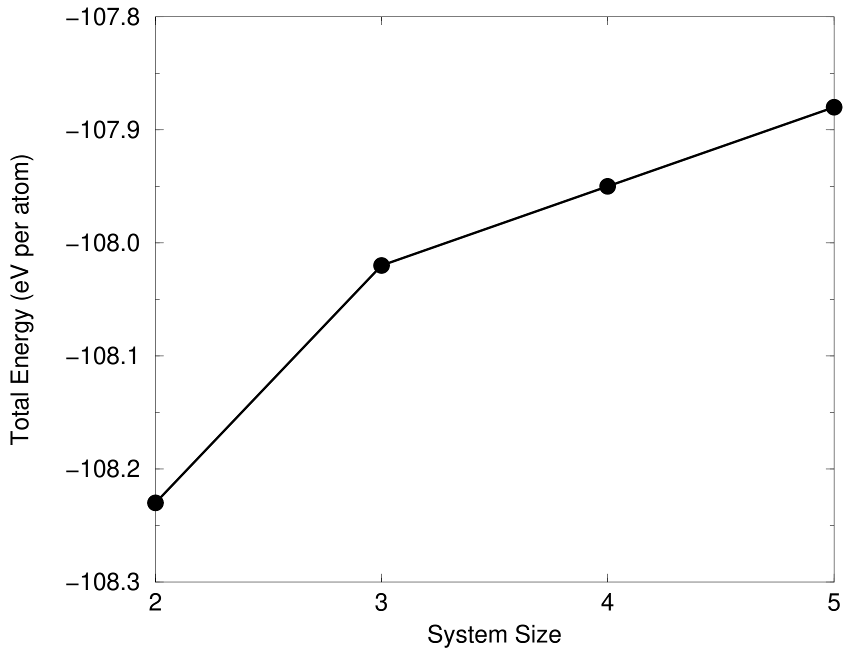


Figure 5.1: Total energy per atom calculated using VMC as a function of system size. The statistical error bars are smaller than the size of the symbols.

The motivation for reducing the CFSE is twofold. Firstly, if the desired level of accuracy in a specific calculation can be achieved by performing that calculation on a smaller system as a result of reducing the CFSE, then the computational benefits will be large as the computational time scales as approximately the third power of the number of electrons in the system. In Figure 5.1, the number of electrons scales as the cube of the system size, n and hence the total computational time scales as the ninth power of the x-axis.

Secondly, there are some problems in which even if one is able to perform the calculation on a large system size this still fails to reduce the CFSE. A standard problem in electronic structure theory is to calculate the energy required to create a point defect. This is done by subtracting the energy of the perfect crystal from that of a large simulation cell containing a single defect. Taking the example of the n=3 (54 atom) simulation cell of silicon, we find (see figure 5.1.) that the CFSE error in the VMC energy of the whole simulation cell is -5 eV. This is much larger than the energies of interest, which are often tenths of an eV per simulation cell. Moreover, it has been observed in previous work [46, 33] that these CFSE are approximately inversely proportional to the number of atoms in the simulation cell and hence the CFSE for the whole cell is *almost independent of* N[1]. Of course there will be a cancellation between the CFSEs in the perfect and defective solids, which will become more complete as the size of the simulation cell increases, so that eventually the energy difference will converge. However, we must expect that the incomplete cancellation of errors for finite simulation cells will lead to a significant uncertainty in the defect energy. Therefore, the only guaranteed way of improving the accuracy of such calculations is to significantly reduce the CFSE at all system sizes.

5.3 Previous Methods of Removing Finite Size Effects

In the past, corrections for the CFSE in QMC simulations have been applied using results for different simulation cell sizes and extrapolating to the infinite cell size limit [46, 33, 43, 47]. This empirical procedure is very costly as it requires several calculations at increasingly large system sizes. As mentioned in section 5.2, the CFSE is approximately inversely proportional to the number of atoms in the simulation cell. A reasonable approximation is

$$\mathrm{CFSE} = \frac{\alpha}{N r_s} \quad , \tag{5.1}$$

[1]In the limit of an infinitely large simulation cell the residual value of the CFSE reflects the fact that different choices of the boundary conditions at infinity give total energies which differ by a finite amount.

where N is the number of electrons in the system and r_s is defined as the radius of a sphere whose volume equals the volume per electron so that

$$\frac{1}{\rho} = \frac{4}{3}\pi r_s^3 \quad . \tag{5.2}$$

To obtain a value for α in Eq.(5.1), a series of calculations is performed for different system sizes. This fitting scheme is not very accurate when results for only 2 or 3 system sizes are available, especially when the estimate for the CFSE is subject to statistical noise at each system size.

5.4 Analysis of e-e energy

5.4.1 Comparison of Hartree-Fock and LDA Results

In an attempt to gain further understanding on the nature of these CFSE, the VMC calculations shown in figure 5.1, were repeated using the Local Density Approximation (LDA) to Density Functional Theory (DFT) and Hartree-Fock (HF) theory. The k-point sampling in the LDA and HF calculations was chosen to be consistent with choosing simulation cell sizes of n = 2,3,4 and 5 as in the VMC calculations. The results of these LDA and HF calculations are plotted in figures 5.2 and 5.3. Note the scale of the y-axis is the same in both graphs.

To facilitate comparison between the LDA and HF results the LDA orbitals were used to calculate the HF energies, so that the energy differences arise solely from the difference between the LDA exchange-correlation (XC) energy and the HF exchange energy. Figure 5.2 shows that the LDA energy converges very rapidly with simulation cell size, whereas in figure 5.3, the HF exchange energy converges very slowly with simulation cell size as did the VMC in figure 5.1. For n=3 the finite size error in the LDA energy (IPFSE) is 0.012 eV per atom, which is much smaller than the HF finite size error of -0.211 eV per atom. The slow convergence of the HF exchange energy with the density of BZ sampling (which is equivalent to the size of the simulation

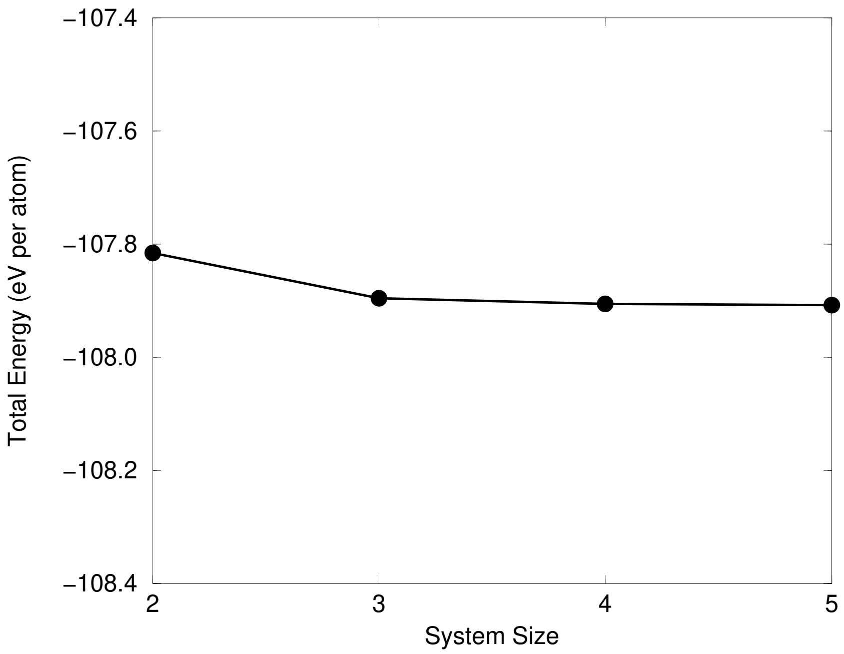


Figure 5.2: Total energy per atom calculated using LDA as a function of system size.

cell) is well known[71, 72, 73] and is usually solved by increasing the quality of the BZ integration.

The rapid convergence of the LDA energy with simulation cell size is easily understood. In the LDA, the total energy can be written as a functional of the charge density

$$E_{LDA}[n(\mathbf{r})] = K.E. + \int V_{ion}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{XC}[n(\mathbf{r})] \quad . \quad (5.3)$$

The LDA charge density $n(\mathbf{r})$ has been shown [3] to converge rapidly with simulation cell size, hence the total energy also converges rapidly. As the LDA orbitals were used to calculate the HF energies, the kinetic and external potential energy in the LDA and HF calculations in figures 5.2 and 5.3 are identical for each system size. If one considers the Hamiltonian for the system as in Eq.(5.4), the expectation value of the

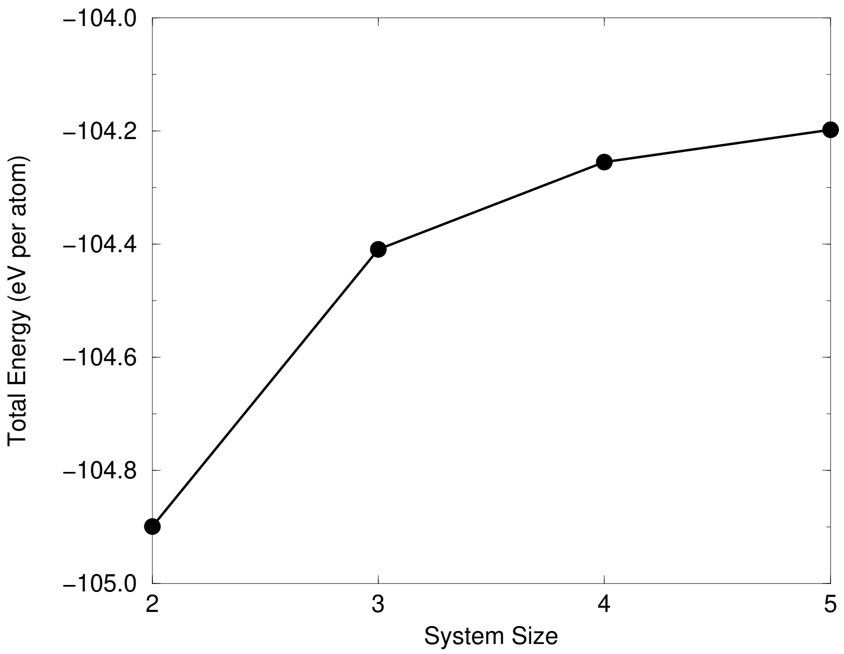


Figure 5.3: Total energy per atom calculated using HF, as a function of system size.

first and second terms of the Hamiltonian with respect to the trial wavefunction are the same in the LDA and HF and only differ in the final electron-electron interaction term;

$$\hat{H} = -\frac{1}{2}\sum_{i}^{N} \nabla_i^2 + \sum_{i}^{N} V(\mathbf{r}_i) + \sum_{\mathbf{R}} \sum_{i>j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{R}|} \quad . \tag{5.4}$$

Hence, the original interpretation of the the residual finite size effect after the subtraction of the IPFSE as a *Coulomb* FSE is consistent with these results.

5.4.2 Electron-Electron Interaction in More Detail

In an attempt to understand the origin of these CFSE, it is instructive to write the electron-electron interaction in terms of a Hartree interaction, which describes the electrostatic energy of the system and an Exchange-Correlation interaction, which describes the interaction of each electron with its exchange-correlation hole, via an interaction $\hat{\nu}(\mathbf{r}-\mathbf{r}')$;

$$E_{\rm e-e} = E_{\rm Hartree} + E_{\rm XC} \quad , \tag{5.5}$$

where

$$\begin{aligned} E_{\rm Hartree} &= \frac{1}{2}\int\int n(\mathbf{r})n(\mathbf{r}')\hat{\nu}(\mathbf{r}-\mathbf{r}')d\mathbf{r}d\mathbf{r}' \quad , \\ E_{\rm XC} &= \frac{1}{2}\int\int n_{XC}(\mathbf{r},\mathbf{r}')n(\mathbf{r}')\hat{\nu}(\mathbf{r}-\mathbf{r}')d\mathbf{r}d\mathbf{r}' \quad . \end{aligned} \tag{5.6}$$

Note the exchange-correlation energy in Eq.(5.6) is the full *many-body* exchange-correlation energy, not the exchange-correlation energy as defined in density functional theory (see chapter 1). It has been established [3, 74] that both the charge density, $n(\mathbf{r})$ and the shape of the exchange-correlation hole [2] $n_{XC}(\mathbf{r},\mathbf{r}')$ converge rapidly with system size and therefore exhibit a very small finite size effect. An extreme example is jellium (see section 5.6), where the charge density is exact for all simulation cell sizes but the CFSE is still present. It therefore appears that the likely source of the CFSE present in the VMC and HF calculations in figures 5.1 and 5.3 is due to the choice of the interaction, $\hat{\nu}(\mathbf{r}-\mathbf{r}')$, used in Eq.(5.6). In the calculations shown in figures 5.1 and 5.3 the Ewald interaction has been used to represent the Coulomb $\frac{1}{r}$ interaction between each electron and all the periodic images of all the other electrons produced by the periodic boundary conditions. The Ewald interaction has already been described in detail in chapter 3. Expanding the Ewald interaction for small r yields

$$\hat{\nu}_{Ewald}(\mathbf{r}) = \frac{1}{r} + \frac{2\pi}{3\Omega}\,\mathbf{r}^T\cdot\mathbf{D}\cdot\mathbf{r} + \mathcal{O}(\frac{r^4}{\Omega^{5/3}}) + \text{constant} \quad , \tag{5.7}$$

where Ω is the volume of the simulation cell and the tensor, $\mathbf{D}$, depends on the geometry of the simulation cell (for a cubic cell $\mathbf{D}$ is the identity matrix) and the constant is defined in chapter 3 so that the average value of the potential is zero. The deviations from $1/r$ model the effects of charges "outside" the simulation cell.

[2] Calculations of pair correlation functions [74] show they converge rapidly with simulation cell size.

In the XC integral, however, the interaction between each electron and its XC hole should be exactly $1/r$, independent of the size of the simulation cell. For very large simulation cells the $1/r$ term in the expansion of the Ewald interaction dominates, but for typical cell sizes such as those used in figures 5.1 and 5.3 the second term is significant and produces a finite size error proportional to $1/\Omega$ in the XC energy per electron. The XC energy is negative and the extra unphysical interaction makes the XC energy more negative. These observations explain why the HF energies in figure 5.3 converge with increasing simulation cell size (i) from below, and (ii) with an error which is roughly inversely proportional to the number of electrons in the simulation cell as originally proposed in section 5.3.

5.5 New Proposed Energy Expression

5.5.1 Short range of Exchange-Correlation Hole

Recent results [3, 74, 68, 69, 75] have provided strong numerical evidence that not only does the *shape* of the exchange-correlation hole converge rapidly with simulation cell size, it is also a relatively *short ranged* quantity. This is illustrated in figure 5.4, which shows the exchange-correlation hole calculated using VMC [74], for diamond-structure silicon, using a simulation cell containing 54 atoms. One electron is placed at the centre of a silicon-silicon covalent bond and the other electron position is within the (110) plane. Figure 5.5 shows a slice through the QMC charge density in the same (110) plane. The position of the central electron from figure 5.4 has been marked with a large white circle.

The short range nature of the exchange-correlation hole ensures that the exchange and correlation energy associated with each electron as written in Eq.(5.6) is well described purely within the simulation cell surrounding each electron. The use of the Ewald interaction to try and describe the exchange-correlation interaction between electrons in different simulation cells therefore appears unnecessary. The essential re-

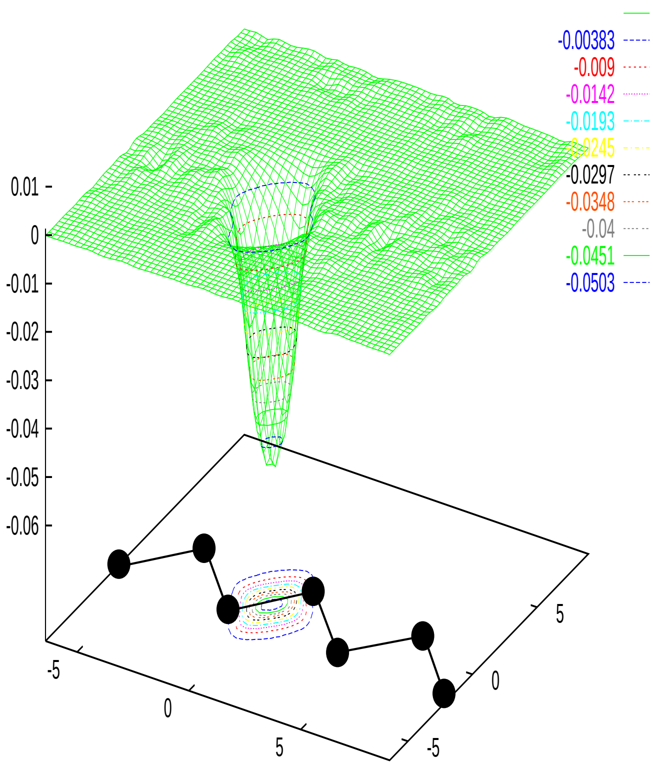


Figure 5.4: Exchange-Correlation hole in diamond-structure silicon from Ref.[74], with $\mathbf{r}$ at the bond centre and $\mathbf{r}'$ ranging over the (110) plane. The black circles represent the positions of the nuclei.

quirements of the electron-electron interaction are simply that (i) it correctly describes the Hartree energy and (ii) each electron interacts with its exchange-correlation hole via the full $1/r$ interaction. In fact it appears that it is the extra terms in the expansion of the Ewald interaction, Eq.(5.7) that are introducing finite size effects into the calculations.

5.5.2 Use $\frac{1}{r}$ Interaction

Having identified that almost all the exchange and correlation interaction between electrons occurs over a short range, it seems sensible to use the exact Coulomb inter-

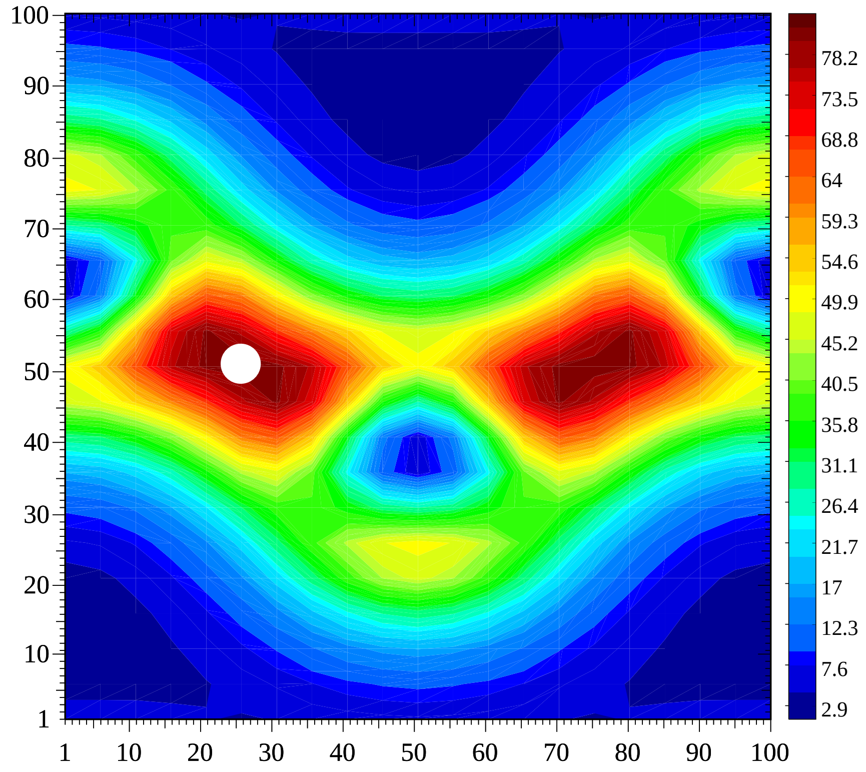


Figure 5.5: VMC Charge density calculated for 3x3x3 diamond structure silicon plotted in the (110) plane through the centre of a silicon-silicon covalent bond.

action, i.e. $\frac{1}{r}$ at short range, to capture this exchange and correlation as accurately as possible. The exchange and correlation energy can then be written as

$$E_{\mathrm{XC}} = \int_s |\Psi|^2 \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k - \frac{1}{2} \int_a \int_s n(\mathbf{r}) n(\mathbf{r}') f(\mathbf{r} - \mathbf{r}') d\mathbf{r} d\mathbf{r}' \tag{5.8}$$

where

$$f(\mathbf{r}_i - \mathbf{r}_j) = \begin{cases} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} & \text{within WS cell} \\ 0 & \text{otherwise} \end{cases}, \tag{5.9}$$

and the limit s on an integral describes an integral over the supercell and the limit a describes an integral over all space. The first term describes each electron interacting via the full $\frac{1}{r}$ Coulomb interaction with all other electrons within its Wigner-Seitz cell. This is similar to the evaluation of the new Jastrow function described in chapter 4 where each electron-electron separation vector is reduced into the Wigner-Seitz cell centred on the electron being considered by subtracting reciprocal lattice vectors from the electron-electron separation vector as illustrated in figure 4.6. The second term in Eq.(5.8) is required to cancel out the contribution from the first term to the Hartree energy. The Ewald interaction still correctly describes the Hartree energy,

$$E_{\text{Hartree}} = \frac{1}{2}\int\int n(\mathbf{r})n(\mathbf{r}')\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}')d\mathbf{r}d\mathbf{r}' \quad . \tag{5.10}$$

Combining Eqs.(5.8 and 5.10) produces a general expression for the electron-electron energy

$$\begin{aligned} E_{e-e} &= \int_s |\Psi|^2 \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \frac{1}{2}\int_s\int_a n(\mathbf{r})n(\mathbf{r}')\left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \quad . \end{aligned} \tag{5.11}$$

5.5.3 Choices for the f function

Before settling on the final choice of f as described in Eq. (5.9) three separate forms were experimented with.

i. A spherical f function that was cut off by a Fermi function at the edge of the largest sphere that could be inscribed within the Wigner-Seitz cell, i.e.

$$f(\mathbf{r}) = \frac{1}{|r|}\,\frac{1}{e^{-\frac{(r-L_{WS})}{T}} - 1} \quad , \tag{5.12}$$

where L_{WS} is the radius of the above sphere. This function was thought at the time to be more elegant as it removed any discontinuity in the gradient of f as an electron moves out of one side of the simulation cell and back in the opposite side.

ii. A spherical f function with a sharp cut off at the edge of the largest sphere that could be inscribed within the Wigner-Seitz cell, i.e.

$$f(\mathbf{r}) = \begin{cases} \frac{1}{|\mathbf{r}|} & |\mathbf{r}| < L_{WS} \quad , \\ 0 & \text{otherwise} \quad . \end{cases} \tag{5.13}$$

iii. The f function described by Eq. (5.9), which is not spherically symmetrical, and has a sharp cut off at the edge of the Wigner-Seitz cell.

Hartree-Fock calculations using LDA orbitals and VMC calculations showed that for the larger simulation cells, n=3,4,5 the energies obtained from all 3 interactions were virtually identical. For the smallest system size (n=2), the total energy was reduced by choosing the interaction, (iii), of Eq. (5.9). This is to be expected as the extra regions outside the cutoff sphere in this interaction allow more correlation between electrons. This choice of f also preserves the sum rule in the interaction of each electron with its exchange-correlation hole, i.e. each electron is interacting with the whole of the exchange-correlation hole via the $1/r$ interaction. If f is chopped off before the edge of the Wigner-Seitz cell as in (i) and (ii), then each electron does not interact via $1/r$ with the whole of the exchange-correlation hole.

The discontinuity in the derivative of the interaction introduced by this choice of f only introduces a discontinuity into the third derivative of the exact wavefunction. This does not contribute to the kinetic energy of the system and is therefore harmless.

5.6 Tests on the Homogeneous Electron Gas

In an earlier piece of work[3] Louisa Fraser performed a series of QMC calculations on the homogeneous electron gas (HEG). In these calculations she used a simplified version of the new interaction of Eq.(5.11).

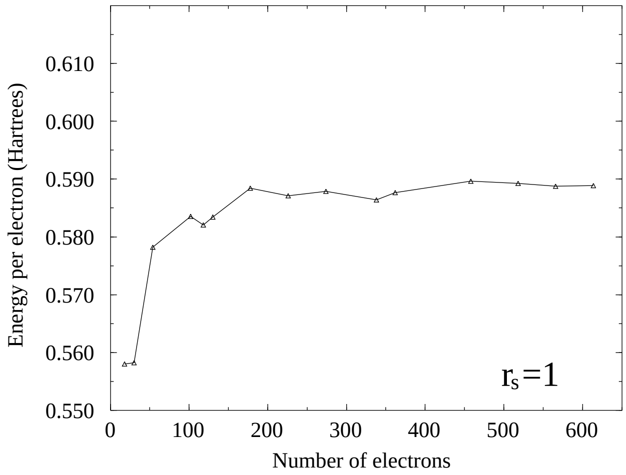
$$E_{e-e} = \int_s |\Psi|^2 \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k + \text{constant} \quad . \tag{5.14}$$

The charge density, $n(\mathbf{r})$, in the HEG is by definition uniform and so the integral in the second term of Eq.(5.11) reduces to a constant that is dependent only on the geometry of the Wigner-Seitz cell and the value of r_s. Therefore, the calculations used in Ref.[3] provide a first, simplified test of the full new interaction in Eq.(5.11).

5.6.1 VMC Results

Louisa Fraser performed a series of VMC calculations[3] for systems with r_s equal to 1 and 3. For comparison, the average value of r_s in silicon is approximately 2.6. A series of system sizes containing 18 to 614 electrons were used to examine the finite size scaling. In figure 5.6 the results of these calculations are shown using the old electron-electron interaction in the Hamiltonian which is based on the Ewald interaction. The energy at each system size has been "corrected" using the LDA finite size errors to remove the IPFSE. It can then be seen quite clearly that there is a residual CFSE that appears to decay approximately as $1/N$.

The calculations of figure 5.6 were then repeated using the new energy expression described by Eq.(5.11). The results are shown in figure 5.7. They show that the CFSE has been dramatically reduced and it is no longer possible to detect a trend in the remaining finite size effects.



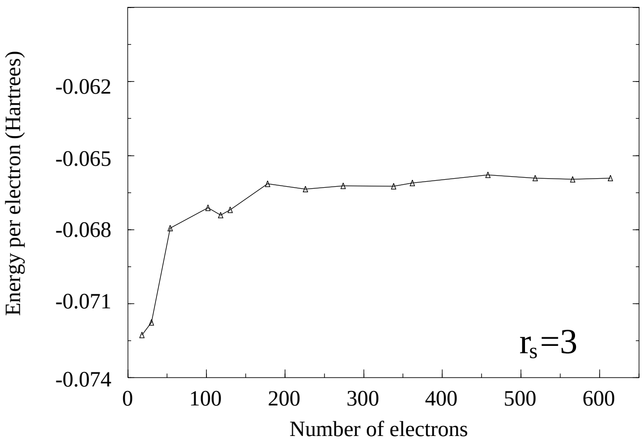


Figure 5.6: VMC results corrected using finite size errors from LDA calculations. Results are for densities corresponding to $r_s = 1$ and $r_s = 3$ and systems containing 18 to 614 electrons.

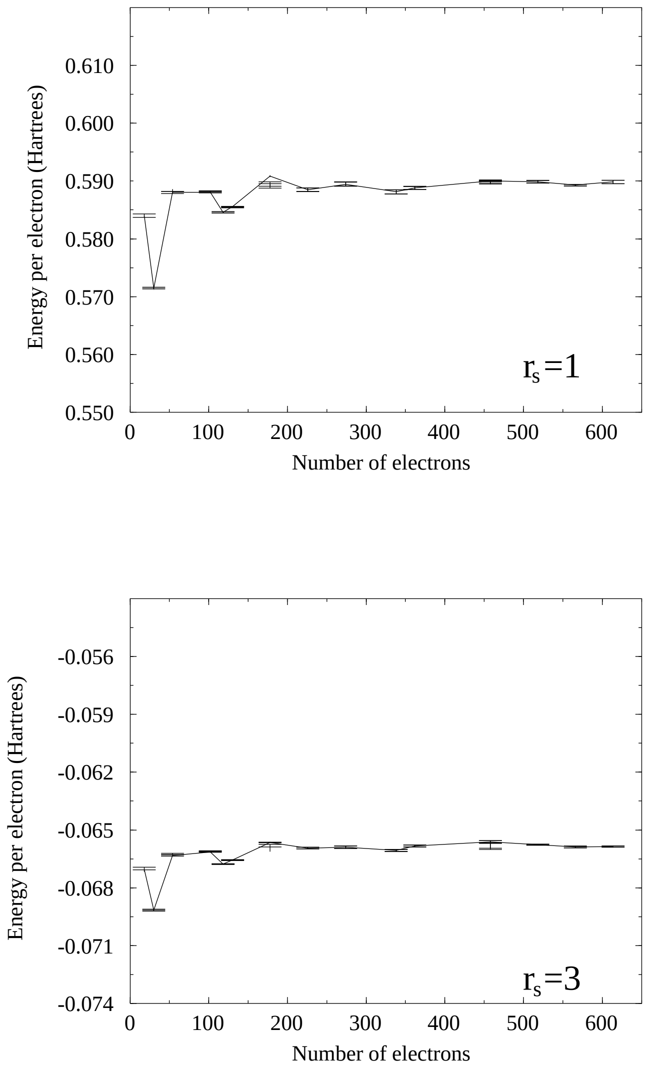


Figure 5.7: VMC results using the new electron-electron potential. Results have been corrected using finite size errors from LDA calculations. Results are for densities corresponding to $r_s = 1$ and $r_s = 3$ and systems containing 18 to 614 electrons.

5.7 Finite Size Effects in Inhomogeneous Solids

With the new expression in Eq. (5.11) for the electron-electron interaction proving so successful for calculations on the HEG, it was decided to attempt to produce a more general expression for use in inhomogeneous solids. Now the long range interaction term is non-trivial because the charge density is no longer uniform. In a VMC calculation the electronic charge density, $n(\mathbf{r})$, appearing in Eq. (5.11) could be accumulated during the simulation and the interaction energy evaluated afterwards. In a DMC calculation this is not possible because the total energy needs to be evaluated at the end of each step of the simulation. Therefore, one must know the charge density before starting the calculation. To overcome this problem calculations were performed with the LDA charge density, $n_{LDA}(\mathbf{r}')$, as the reference density in Eq. (5.11) i.e.

$$\begin{aligned} E_{e-e} &= \int_s |\Psi|^2 \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \frac{1}{2} \int_s \int_a n(\mathbf{r}) n_{LDA}(\mathbf{r}') \left[\hat{v}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \end{aligned} \tag{5.15}$$

It would be possible to update the input charge density afterwards and perform a self-consistent calculation. However, LDA charge densities are normally remarkably good and, moreover, the interaction energy is insensitive to the quality of the charge density used because $[\hat{v}_{Ewald}(\mathbf{r}) - f(\mathbf{r})]$ differs significantly from zero only when $|\mathbf{r}|$ is large. This approximation[3] was found to be so successful that for convenience it was also used for VMC calculations and in the variance minimisation procedures for optimising wavefunctions.

To implement the new interaction into the QMC code for solids described in chapter 3 the following changes were made

[3]Note that a successful QMC calculation requires a reasonable approximation to the wavefunction itself, so that the requirement of a reasonable approximation to the charge density is not a serious limitation.

- The $1/r$ interaction is evaluated as in the HEG by reducing each electron-electron separation vector into the Wigner-Seitz cell of the electron being considered.

- The Hartree-like term is treated as an extra one-body potential felt by all the electrons;

$$V_{Hartree}(\mathbf{r}) = \int n_{LDA}(\mathbf{r}')\left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}' \quad . \tag{5.16}$$

  This extra one-body potential is stored in reciprocal space and accumulated on the same set of primitive cell reciprocal lattice vectors that the χ function is evaluated on and the density is accumulated on.

The interaction leading to the energy expression in Eq. (5.15) is illustrated in Fig. 5.8. This depicts a rhombohedral simulation cell containing two electrons, on one of which is centred a hexagonal window corresponding to the Wigner-Seitz cell of the simulation cell. The (red) electron at the centre of the window experiences a $1/r$ interaction with all the other electrons within the window (one (blue) in this case) and an electrostatic interaction with the electronic charge density of the shaded region outside of the window.

5.7.1 Electron-Ion and Ion-Ion Interactions

For consistency, one should apply the same prescription to all the electrostatic interactions in the system, i.e. not only the electron-electron terms, but also the electron-nucleus and nucleus-nucleus interactions. These can be included by re-writing Eq. (5.11) to involve all charged particles as

$$\begin{aligned} E_{int} &= \frac{1}{2}\int_{\text{cell}} \rho(\mathbf{r})\rho(\mathbf{r}')[\hat{v}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')]d\mathbf{r}d\mathbf{r}' \\ &+ \int_{\text{cell}} |\Psi|^2 \left[\sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) - \sum_i \sum_\alpha Z_\alpha f(\mathbf{r}_i - \mathbf{R}_\alpha) \right. \end{aligned}$$

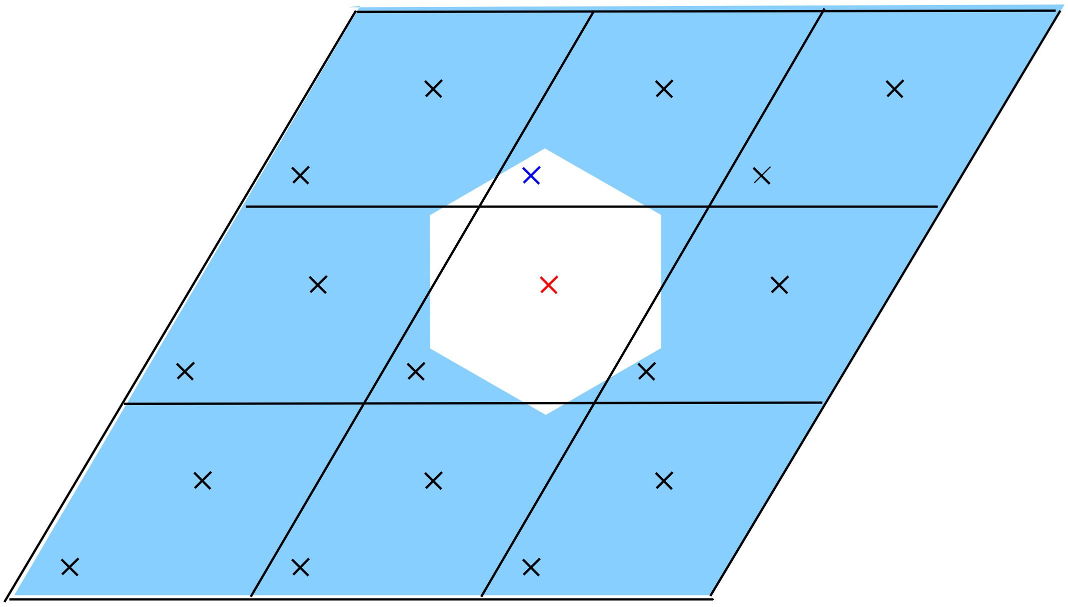


Figure 5.8: An illustration of the new interaction for a rhombohedral simulation cell containing two electrons (crosses). The hexagonal clear window centred on one of the electrons has the shape of the Wigner-Seitz cell of the simulation cell.

$$+ \sum_{\alpha>\beta} Z_\alpha Z_\beta f(\mathbf{R}_\alpha - \mathbf{R}_\beta) \Bigg] \Pi_k d\mathbf{r}_k \Pi_\gamma d\mathbf{R}_\gamma \quad , \tag{5.17}$$

where Ψ is the many-body wavefunction for the electrons and nuclei and $\rho(\mathbf{r})$ now also includes the charge density due to the nuclei. In all our calculations we have made the adiabatic approximation to separate the electronic and nuclear dynamical variables

$$\Psi(\mathbf{r}_i; \mathbf{R}_\alpha) = \psi(\mathbf{r}_i; \mathbf{R}_\alpha)\phi(\mathbf{R}_\alpha) \quad , \tag{5.18}$$

where the $\mathbf{R}_\alpha$ appear only as parameters in ψ. To make further progress we must assume a form for the nuclear part of the wavefunction, ϕ. The simplest assumption is that ϕ can be written as an appropriately symmetrised product of single-nucleus functions which are very strongly localised and therefore non-overlapping. The above equation then reduces to

$$E_{int} = \frac{1}{2}\int_{\text{cell}} n(\mathbf{r})n(\mathbf{r}')[\hat{v}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')]d\mathbf{r}d\mathbf{r}'$$

$$
\begin{aligned}
&+ \int_{\text{cell}} |\psi|^2 \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) \Pi_k d\mathbf{r}_k \\
&- \int_{\text{cell}} |\psi|^2 \sum_i \sum_\alpha Z_\alpha \hat{v}_{Ewald}(\mathbf{r}_i - \overline{\mathbf{R}}_\alpha) \Pi_k d\mathbf{r}_k \\
&+ \sum_{\alpha>\beta} Z_\alpha Z_\beta \hat{v}_{Ewald}(\overline{\mathbf{R}}_\alpha - \overline{\mathbf{R}}_\beta) \quad ,
\end{aligned}
\tag{5.19}
$$

where the $\overline{\mathbf{R}}_\alpha$ denote the centres of the single-nucleus functions, and n is the electron density. Note that the first two terms of the above equation correspond exactly to the electron-electron interaction in Eq. (5.11), and that the electron-nucleus and nucleus-nucleus terms now involve only the Ewald interaction. The above result can be understood in the following way. We are treating the ions as classical particles with well defined positions and therefore expect no exchange-correlation terms involving these particles. This leaves only the Hartree interaction which is correctly described by the Ewald interaction. One consequence of this is that as the Ewald interaction has a continuous derivative, the forces on the ions are continuous functions of the ionic positions, which means that this scheme is suitable for use in quantum molecular dynamics calculations.

5.8 Results for Silicon

5.8.1 VMC Results

To illustrate the effectiveness of the new electron-electron interaction the set of VMC calculations for the silicon simulation cells shown in figure 5.1 were repeated using the new interaction energy expression of Eq.(5.15). The wavefunctions were optimised using Hamiltonians based on both the Ewald interaction and with the new interaction, but even for the smallest simulation cells the wavefunctions obtained were virtually identical. This shows that properties other than the energy are not significantly affected by the change in the interaction term. The sensitivity to the input charge density used in Eq. (5.15) was tested using the charge densities obtained from an

n=2 LDA calculation and from a fully converged LDA calculation. This made no detectable difference to the calculated energies. Figure 5.9 shows the energy per atom obtained from VMC calculations using the new interaction and the Ewald interaction. It is clear that the use of the new interaction greatly reduces the CFSE.

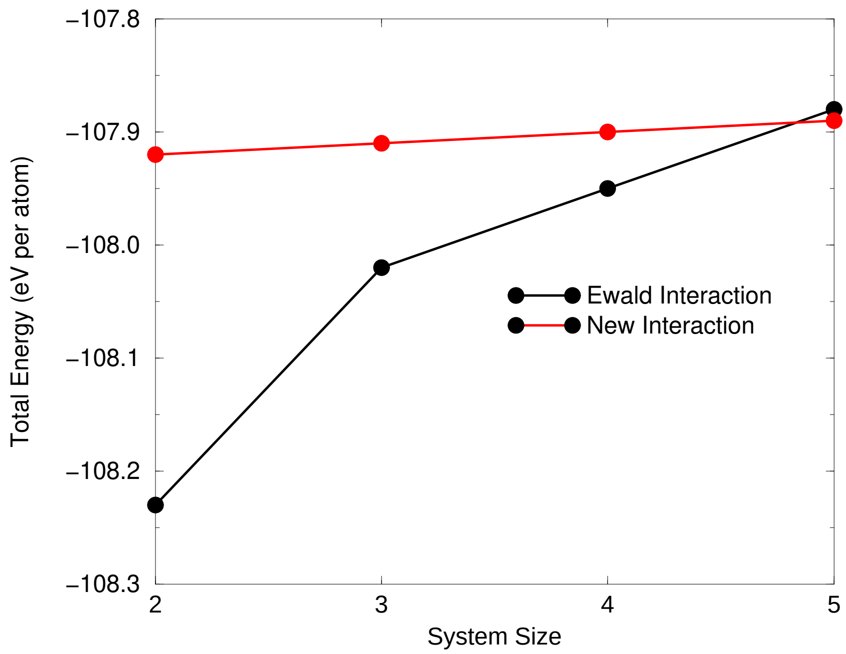


Figure 5.9: The energy per atom in diamond-structure silicon as a function of simulation cell size, from VMC calculations using the Ewald electron-electron interaction and the new interaction. The statistical error bars are smaller than the size of the symbols.

5.8.2 HF Results

The Hartree-Fock calculations from figure 5.3 were also repeated to ensure that the new interaction reduces the CFSE present in these calculations. The results are illustrated in Figure 5.9. The new interaction is clearly also successful in HF calculations, reducing the CFSE by a factor of about 3.

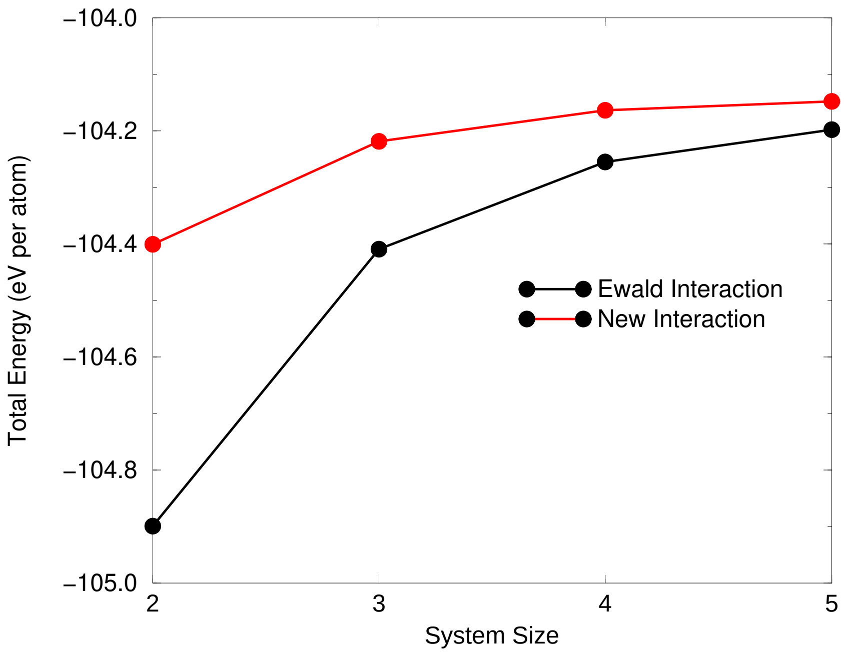


Figure 5.10: The energy per atom of diamond-structure silicon as a function of simulation cell size, from HF calculations using the Ewald electron-electron interaction and the new interaction.

5.8.3 DMC Results

Although it would be computationally prohibitive to repeat all the calculations of figure 5.9 within DMC to confirm that the new interaction behaves in a similar way, it is possible to compare DMC calculations at the smallest system size, n=2, to see the effect on the diffusion algorithm of switching to the new electron-electron interaction.

To perform a DMC calculation, one requires not only an energy expression but also an expression for the Hamiltonian of the system being studied. The Schrödinger equation may be "derived" by minimising an energy functional, $E[\Psi] = \langle\Psi|\hat{H}|\Psi\rangle$, where Ψ is a normalised wavefunction. If a similar procedure is carried out for a functional including the electron-electron interaction of Eq. (5.15), the electron-electron interaction operator in the resulting Schrödinger-like equation is

$$\hat{H}_{e-e} = \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) + \sum_i \int_{\text{cell}} \left[\hat{v}_{Ewald}(\mathbf{r}_i - \mathbf{r}) - f(\mathbf{r}_i - \mathbf{r})\right] n(\mathbf{r}) d\mathbf{r} \quad . \tag{5.20}$$

Again, as in Eq. (5.15), we chose to use the LDA charge density, $n_{LDA}(\mathbf{r})$, as the input density, $n(\mathbf{r})$, to the second term in the Hamiltonian. The total electron-electron energy, E_{e-e}, is then the expectation value of $\hat{H}_{e-e}$ minus a double counting term for the electrostatic interactions;

$$E_{e-e} = \langle \Psi | \hat{H} | \Psi \rangle - \frac{1}{2} \int \int n(\mathbf{r}) n_{LDA}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r} - \mathbf{r}') - f(\mathbf{r} - \mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \quad . \tag{5.21}$$

This double counting term can itself be accumulated during the DMC calculation and then subtracted off at the end of the simulation as it is a fixed constant that will not affect the diffusion algorithm.

Two DMC calculations were performed on the n=2 system of diamond-structure silicon to compare the effect of the new interaction in DMC and VMC. The DMC calculations were performed using a Slater determinant of single-particle orbitals with $\mathbf{k}$-points chosen on a reciprocal space grid centred at the origin, i.e. Γ-point sampling. This sampling was chosen as DMC calculations with Γ-point sampling are required in the next chapter as well.

The first DMC calculation was performed with the standard Hamiltonian as described in chapter 2. The second calculation used the new DMC Hamiltonian from Eq.(5.20). The same changes were made to the DMC algorithm (described in chapter 2) as were made to the VMC algorithm in section 5.7 to implement the new electron-electron interaction. The new DMC calculation exhibited the same stability in the population of walkers as the original Hamiltonian. It required a similar number of steps to diffuse to a state where energies could be accumulated and the intrinsic variance of the energy over the run was also very similar. Table 5.1 shows a comparison of the total energies obtained in VMC and DMC using the Ewald and new electron-electron interactions. The VMC results in table 5.1 were performed using Γ-point sampling to facilitate the comparison.

	Ewald Interaction (eV per atom)	New Interaction (eV per atom)
VMC	-106.88 ± 0.03	-106.70 ± 0.03
DMC	-107.41 ± 0.03	-107.30 ± 0.03
$\Delta_{\mathrm{VMC-DMC}}$	0.53 ± 0.04	0.60 ± 0.04

Table 5.1: Comparison of VMC and DMC results using the Ewald and New interactions. All calculations use Γ-point sampling of the one-electron wavefunctions. DMC results obtained using the energy expression of Eq.(5.21).

The results show that the reduction in energy obtained by performing a DMC calculation rather than a VMC calculation is similar for the two electron-electron interactions. This suggests that, as expected, the finite size effects in DMC broadly follow those in VMC and that using the new electron-electron interaction yields a similar improvement in DMC calculations to VMC calculations.

5.9 Variance Minimisation with the New Interaction

In section 5.8 results are presented for VMC calculations performed using both the Ewald interaction and the new interaction. To provide the fairest possible comparison of the two interactions, the trial wavefunctions for each calculation were optimised, using the variance minimisation technique described in chapter 4, with respect to the Hamiltonian to be used in the subsequent VMC calculation. This enables both calculations to be performed using the optimal values of the variational parameters in the wavefunction for the given Hamiltonian. The standard expression for the intrinsic variance

$$\sigma^2 = \sum_{\alpha} \left(\Psi_T^{-1}(\mathbf{R}_\alpha) \hat{H} \Psi_T(\mathbf{R}_\alpha) - < E_L > \right)^2 \left[\frac{w(\alpha)}{\sum_\beta w(\beta)} \right] \quad , \tag{5.22}$$

was used. The form of $\hat{H}$ was chosen to match the form to be used in the VMC calculation. For the calculations using the Ewald interaction, $\hat{H}_{e-e}$ was chosen to be

$$\hat{H}_{e-e} = \sum_{i>j} \hat{v}_{Ewald}(\mathbf{r}_i - \mathbf{r}_j) \quad . \tag{5.23}$$

For the calculations using the new electron-electron interaction, $\hat{H}_{e-e}$ was chosen to be

$$\begin{aligned} \hat{H}_{e-e} &= \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) \\ &+ \sum_i \int_{\text{cell}} \left[\hat{v}_{Ewald}(\mathbf{r}_i - \mathbf{r}) - f(\mathbf{r}_i - \mathbf{r})\right] n_{LDA}(\mathbf{r}) d\mathbf{r} \quad . \end{aligned} \tag{5.24}$$

In both cases, the electron configurations are still fixed throughout the optimisation process. Therefore the total electron-electron energy for each configuration is still just calculated once and stored throughout the optimisation procedure.

5.10 Conclusions

In summary, we have traced the source of the troublesome Coulomb finite size errors in quantum many-body calculations for periodic simulation cells to the use of the Ewald interaction, which gives a spurious cell-size-dependent contribution to the electron-electron interaction energy. A new model electron-electron interaction is introduced, which eliminates this problem, based on the idea that the exchange-correlation hole is short ranged. VMC calculations with up to 1000 electrons show that the new interaction gives much smaller finite size effects than the Ewald interaction. Preliminary DMC calculations suggest the new interaction is equally successful in the more sophisticated DMC formalism. This development will allow for significantly more accurate supercell calculations of correlated electron systems.

Chapter 6

QMC Calculations of Excitation Energies

It is well known that many-body correlations have a significant impact on important band structure features such as band gaps, band widths and other excitation energies. One of the advantages of the quantum Monte Carlo (QMC) method over other electronic structure techniques is that it can include these correlation effects. In this chapter two separate methods for using QMC to calculate excitation energies are described. These are the addition and subtraction of electrons from the bulk solid and the promotion of electrons from the valence band into the conduction band. Results are presented for the band gaps of silicon in the diamond structure obtained using both methods. Silicon was chosen to serve as a useful test material for experimenting with the different methods of calculating excitation energies using QMC. There is a wealth of both experimental and theoretical electronic structure results for excitation energies in silicon that can be used as a measure of the relative successes of the techniques described in this chapter.

The efficiency and accuracy of the two QMC methods are compared and contrasted with recent QMC results published by the Illinois group[51], and with other electronic structure techniques. Finally, some alternative techniques for calculating excitation energies within QMC are discussed.

6.1 Previous Work on Excitation Energies

6.1.1 QMC Calculations

The literature to date contains only a few references to the calculation of excited states within the QMC formalism. The most notable are those of Mitas and Martin[51, 76], who calculated the electronic energy gap for a compressed molecular nitrogen solid and carbon and compared them with LDA and Hartree-Fock estimations, Knorr and Godby[77] who performed DMC calculations on semiconductor wires, Fahy et al.[78] who calculated the quasiparticle energies within a single-mode approximation in diamond, and Engel et al.[79], who compared the quasiparticle bands in a two-dimensional crystal calculated using GW and QMC. QMC calculations of excited states via indirect methods[80, 81, 82], such as extracting excitation energies from the exponential decay curves in DMC, are discussed in section 6.6.2.

6.1.2 Hartree-Fock Calculations

Excitation energies can be calculated within Hartree-Fock (HF) theory via two methods. Koopmans' theorem[83] applies if the number of electrons in the system is large. Then adding or removing a single electron from the system will not affect the orbitals of the other electrons and they can be assumed fixed. The energy required to remove the kth electron from the system is then just $-\epsilon_k$. Hence the energy to transfer an electron from the ith to the jth state is $\epsilon_j - \epsilon_i$. Therefore, within Koopmans' theorem, the single-particle energies, ϵ_i from the Hartree-Fock equations, Eq.(1.13), can be interpreted as the excitation energies of the system.

Koopmans' theorem is only valid if the one-electron wavefunctions in the N-electron and the $(N \pm 1)$-electron Slater determinants are the same, i.e. the single-particle orbitals do not *relax* when an electron is added to or removed from the system. In a finite supercell calculation, such as those performed in chapters 4 to 6 of this thesis, it is not clear to what extent Koopmans' theorem still holds. In this case it is more appropriate to use the alternative method of performing total energy calculations for both the N and the $N \pm 1$ electron systems and then subtracting to find the ionisation energy and electron affinity. It has been shown [84] that for core levels in atoms there are significant differences between the results obtained using Koopmans' theorem and the method of performing two total energy calculations. There are many examples[71, 72, 73] of applications of these techniques to calculations of the HF bandstructure of semiconductors.

6.1.3 Density Functional Calculations

The local density approximation (LDA) is presently the most successful method for the determination of the ground state properties of solids. The eigenvalues of the LDA equation, though being *a priori* of no physical meaning are nevertheless commonly interpreted as single-particle energies. The energy gaps obtained from these single-particle energies are generally too small in comparison with experiment[85]. The deviation of the LDA gap from experiment can be anywhere between 100% in the case of germanium (in which the LDA predicts a negative gap) to fairly small percentage errors in wide gap insulators. Since modern bandstructure codes have reached a stage where the calculations are well converged, it is clear that this error indicates either a shortcoming of the exchange-correlation functionals currently used in the LDA or a more fundamental incapability of Kohn-Sham (KS) Density Functional theory (DFT) to calculate excitation energies. It has been shown[86, 87] that the exchange-correlation potentials for an N-particle and $N+1$-particle system differ by a finite quantity Δ, known as the discontinuity of the exchange-correlation potential. As this discontinuity in the exchange-correlation potential is a feature of exact KS-DFT, it is not immediately clear to what extent the errors in the energy gaps obtained from single-particle energies are due to limitations in the LDA or to the presence of this discontinuity. Godby, Schlüter and Sham[88, 89] have calculated an exchange-correlation potential for several semiconductors using the GW approximation for the self-energy, which can be expected to agree very closely with the exact exchange-correlation potential. This potential and the resulting KS-DFT bandstructure turned out to be in remarkably close agreement with the local density approximation. It therefore appears likely that the discontinuity in the exchange-correlation potential is responsible for over 80% of the errors observed in these gaps. It has also been demonstrated[90] that for a two-band semiconductor model, every state-independent exchange-correlation potential is bound to fail to describe excitation energies because some essential features are missing. However, the opposite situation (i.e. a very small discontinuity in the exchange-correlation functional) was found in simple one-dimensional Hubbard-like model for semiconductors[91, 92] where the exchange-correlation potential and its discontinuity can be calculated exactly.

A possible method for overcoming the problem of calculating excited states within DFT was proposed by Kohn[93]. He suggested determining excitation energies by calculating the ground state energies of the N and the $N \pm 1$ electron systems and then subtracting to find the ionisation energy. This is exactly the method outlined above for use in HF calculations and employed in QMC calculations in section 6.3.

6.1.4 Experimental Determination of Excited States

There is a wealth of experimental results for excitations in bulk diamond-structure silicon. A good summary is provided in the book by Chelikowsky and Cohen[94].

Traditionally, optical reflectivity measurements were the popular[95, 96, 97, 98, 99, 100] method of measuring band structures. They have two main advantages over other optical probes. First, they are not over sensitive to surface conditions. The photon sampling length in semiconductors is of the order of a few thousand angströms. Since intrinsic surface perturbations in semiconductors heal within tens of angströms, optical measurements are insensitive to surface states. Secondly, reflectivity measurements have better resolution than other techniques with respect to band structure features which lie in or above the optical region. In some cases, this resolution may approach a few meV.

The more modern tool for probing the electronic structure of solids is photo-electron spectroscopy[101]. In this technique the surface of a solid is bombarded with either X-ray or Ultraviolet radiation. The photons are absorbed in the bulk by a process which corresponds to a bulk excitation. The excited electrons then propagate to the crystalline surface and escape. By measuring the number of photoelectrons ejected with a specific energy for a given incident photon energy, one is able to probe the valence band density of states within the bulk solid.

One of the most useful implementations of photoelectron spectroscopy is the angle-resolved photoemission spectroscopy technique. This technique uses the conservation of momentum parallel to the surface to calculate the component of the bulk electron's momentum parallel to the surface. Hence one is able to measure not only the valence band density of states but also the valence bandstructure. It is also possible to measure the conduction bandstructure using the technique of inverse photoemission, which yields the energy and momentum of a photon emitted when an electron makes the transition from the conduction band to the valence band. All the experimental results for the bandstructure of silicon[102, 103, 104] quoted later in this chapter have been obtained using photoemission experiments.

6.2 QMC Methods for Calculating Excitation Energies

Two separate methods for calculating excited states within QMC, namely the addition and subtraction of electrons from the bulk solid and the promotion of individual electrons to the conduction band, have been applied to bulk silicon in the diamond structure. The following sections describe the technical details of implementing these techniques within both the variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) frameworks. Both methods are effectively '$\frac{1}{N}$' techniques, i.e. the change in the energy due to the excitation is inversely proportional to the number of electrons in the simulation cell. The calculations therefore have to be of sufficient accuracy to resolve this energy change amid the statistical noise produced by all N electrons. This is in contrast to the ground state calculations described in chapters 4 and 5 where one is interested in the total energy per atom averaged over all the atoms in the simulation cell. The increased accuracy required in these excitation energy calculations makes them especially demanding and we must be careful to use high quality trial/guiding wavefunctions and to correct for the finite size effects. Hence, the progress made in these areas that has already been described in chapters 4 and 5 is directly applicable to the techniques described here.

6.3 Addition and Subtraction of Electrons

Excitation energies can be found by taking the difference between the ground state energy of systems with different numbers of electrons. For example, a gap energy E_{gap} can be written as

$$E_{gap} = (E_{N+1} - E_N) - (E_N - E_{N-1}) \quad , \tag{6.1}$$

where $(E_{N+1} - E_N)$ is the energy to add an electron (the electron affinity[1]) and $(E_N - E_{N-1})$ is the energy to remove an electron (the ionization potential).

To obtain an estimate of the electron(hole) energies, $E_{N+1} - E_N$ $(E_N - E_{N-1})$, within a QMC calculation, all that is required is to include (remove) the corresponding conduction (valence) orbital in the Slater Determinant part of the trial wavefunction given in Eq.(2.31). The **k**-point at which the electron(hole) energy is calculated, is determined by the **k**-value of this conduction (valence) orbital.

6.3.1 Trial Wavefunction

In the following calculations it has been assumed that the addition (removal) of a single electron does not significantly affect the overall shape of the exchange correlation

[1]Strictly, the electron affinity of the N electron system is defined as $A(N) = E_{N-1} - E_N$

hole and so the same Jastrow correlation function that was introduced in chapter 4 has again been used. It was also assumed that although the addition(removal) of a single electron will reduce the symmetry of the charge density, this will not be a strong effect. The same form of the 1-body $\chi(\mathbf{r})$ function, grouped into stars according to the higher point group symmetry of the crystal in its ground state as described in section 4.3, should therefore still be applicable. The validity of these approximations in the trial wavefunction is tested in section 6.3.6, where the calculations are repeated in DMC. The value of the energy calculated in DMC is independent of the quality of the u and $\chi(\mathbf{r})$ parts of the trial wavefunction. Therefore, if the quality of the u and $\chi(\mathbf{r})$ parts of the trial wavefunction is much lower for the calculations where an electron has been added or removed, one would expect to see a bigger reduction in the energy when moving from VMC to DMC.

The trial wavefunctions for each of the E_N, E_{N+1} and E_{N-1} systems were individually optimised using the variance minimisation techniques described in chapter 4. Six stars of $\mathbf{G}$ vectors were used to describe the $\chi(\mathbf{r})$ functions and 22 parameters were used to describe the u functions, as in chapter 5. Large ensembles of up to 1 million electron configurations were used in each iteration of the optimisation procedure and 3-5 iterations were performed to obtain ground state wavefunctions with an accuracy of approximately $\pm$ 0.01eV per atom within the parameter space of the optimisation. This is equivalent to an accuracy of approximately $\pm$ 0.3eV in the energy of the gap as given by Eq.(6.1).

6.3.2 Electron-Electron Interaction

It has already been shown in chapter 5 that the choice of electron-electron interaction is fundamental, if one is to control the Coulomb finite size effects introduced when attempting to simulate a bulk system using a model system consisting of a few 10's of electrons in a simulation cell with periodic boundary conditions. Therefore, when performing QMC calculations for the gap energy as given in Eq.(6.1), the new electron-electron interaction described in section 5.5 was used both in the wavefunction optimisation procedure for each of the systems and then in the subsequent QMC calculations. The new energy expression for an N-electron system as stated in Eq.(5.15) naturally extends to systems containing $N \pm 1$ electrons in the following way,

$$\begin{aligned} E_{e-e}^{X} &= \int\limits_{s} |\Psi|^2 \sum_{i>j}^{X} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \frac{1}{2} \int\limits_{s} \int\limits_{a} n_{QMC}^{X}(\mathbf{r}) n_{LDA}^{X}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \quad , \end{aligned} \tag{6.2}$$

where n_{LDA}^{X} is the LDA charge density of a simulation cell containing $X = N$ or $N \pm 1$ electrons. The sum in the first term now extends over all electron-electron pairs including those involving the additional electron. The Hamiltonian associated with this energy expression is

$$\begin{aligned} \hat{H}_{e-e}^{X} &= \sum_{i>j}^{X} f(\mathbf{r}_i - \mathbf{r}_j) \\ &+ \sum_i \int_{\text{cell}} \left[\hat{\nu}_{Ewald}(\mathbf{r}_i - \mathbf{r}) - f(\mathbf{r}_i - \mathbf{r})\right] n_{LDA}^{X}(\mathbf{r}) d\mathbf{r} \quad , \end{aligned} \tag{6.3}$$

and the electron-electron energy in a VMC calculation is then given by

$$E_{e-e} = \langle \Psi | \hat{H}_{e-e}^{X} | \Psi \rangle - \frac{1}{2} \int \int n_{VMC}^{X}(\mathbf{r}) n_{LDA}^{X}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \quad . \tag{6.4}$$

6.3.3 New Electron-Electron Interaction

The expression for the gap energy in Eq.(6.1) is based on the change in the ground state energy of the bulk solid when a single electron is added or removed. In all our QMC calculations for bulk solids, we choose to model this system using a finite simulation cell to which we apply periodic boundary conditions. When an additional electron is added to this simulation cell, the periodic boundary conditions effectively introduce an equivalent extra electron into each of the periodic images of the simulation cell. This adds an extra electrostatic energy into the system, which when combined with a compensating background charge is just the Madelung energy of an electron (hole) crystal with the periodicity of the simulation cell. A similar effect has been observed in LDA calculations[105, 106, 107, 108]. Leslie and Gillan[109] proposed a correction term to the Hartree energy of the system to account for the additional electrostatic energy due to an array of charged defects,

$$E \approx E_0 - \frac{1}{2}\frac{\alpha q^2}{2\epsilon_r L} \quad , \tag{6.5}$$

where α is the Madelung constant for the supercell geometry, L is the length of the simulation cell, ϵ_r is a dielectric constant for the material, and q is the charge on the defect. The problem with this correction is how to choose the dielectric constant, ϵ_r. In general experimental values have been used and these have not been found to work particularly well.

It has also been speculated by Engel et al.[79] that similar effects may be present in their VMC calculations of the band structure of a two-dimensional model crystal. In their calculations, an extra electron was added into an orbital in the conduction band. However, this orbital is actually spread throughout the simulation cell and so can be regarded as contributing a much smaller term to the Hartree energy than a point defect plus background would. In the limit of infinite simulation cell size L, any additional energy due to interactions between the array of additional electrons would disappear. Therefore Engel et al. treat this as a finite size effect and deal with it by fitting results for a series of VMC calculations at different system sizes to the expression

$$E_{gap}(N) = E_{gap}(\infty) - \tilde{\alpha}/L \quad , \tag{6.6}$$

where L is the length of the simulation cell, and $\tilde{\alpha}$ is a parameter that represents a reduced α due to the screening of the other valence electrons.

In our QMC calculations[110, 3, 111] we no longer use the Ewald interaction to evaluate the electron-electron interaction between pairs of electrons and therefore we are not necessarily restricted to including all the periodic images of the additional electron(hole) in our system in the same way as Engel et al.

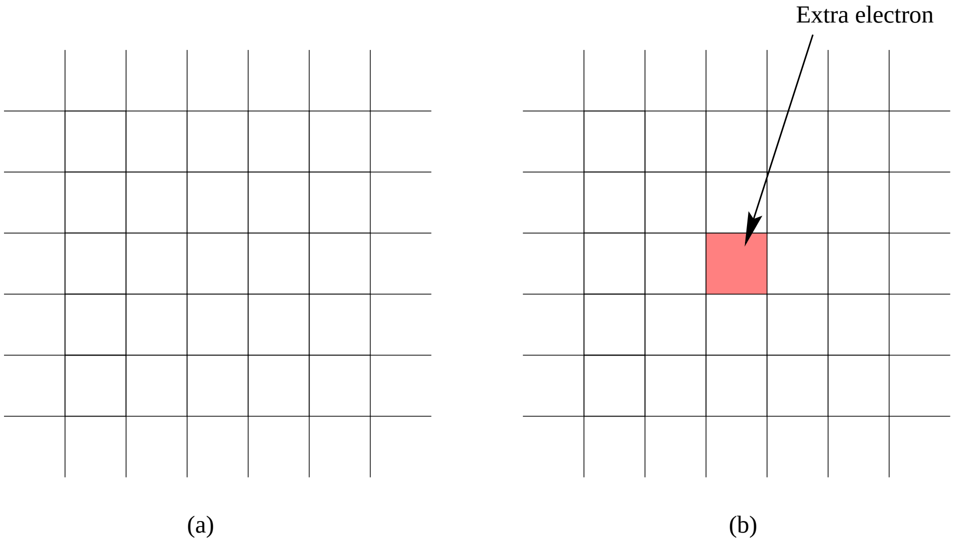


Figure 6.1: Addition of a single electron to the simulation cell. Figure (a) shows an N electron simulation cell periodically repeated. Figure (b) shows the same bulk system with an additional electron added only to the simulation cell (red).

Consider the two systems illustrated in figure 6.1. Figure (a) schematically represents the standard simulation cell for the N electron system and a few of the periodic images of the simulation cell. The electron-electron energy associated with this system can be defined as in section 5.5 by the new electron-electron energy expression for N electron systems,

$$\begin{aligned} E^N_{e-e} &= \int_s |\Psi|^2 \sum_{i>j}^{N} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \frac{1}{2} \int_s \int_a n^N_{QMC}(\mathbf{r}) n^N_{LDA}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \quad . \end{aligned} \tag{6.7}$$

In Figure 6.1(b) the same system is shown with an extra electron added *only to the* *actual simulation cell*, not to any of its periodic images. We can represent the change in the charge density of the *whole* system due to the additional electron by $\Delta(\mathbf{r})$, and we would like to confine $\Delta(\mathbf{r})$ to within the central simulation cell, i.e. there should be no additional electrons in the periodic images of the simulation cell. This effect can be achieved by altering the interaction so that each electron 'feels' the full $1/r$ interaction with all $N+1$ electrons within the simulation cell surrounding it[2], but only feels the Hartree interaction with the charge density due to N electrons in each periodic image outside the simulation cell. We re-write Eq.(6.7) to take account of the extra electron and the change in the charge density, $\Delta(\mathbf{r})$, which is confined to the central simulation cell, as

$$\begin{aligned} E^X_{e-e} &= \int_s |\Psi|^2 \sum_{i>j}^X f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \frac{1}{2}\int_s\int_a \Big(n^N(\mathbf{r}) + \Delta(\mathbf{r})\Big)\Big(n^N(\mathbf{r}') + \Delta(\mathbf{r}')\Big)\left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \end{aligned} \tag{6.8}$$

and expand out the product $\Big(n^N(\mathbf{r}) + \Delta(\mathbf{r})\Big)\Big(n^N(\mathbf{r}') + \Delta(\mathbf{r}')\Big)$, in the second term. We can discard the term in $\Delta(\mathbf{r})\Delta(\mathbf{r}')$ which is small as $\Delta(\mathbf{r})$ is a short ranged function and
$[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')]$ is a small for $|\mathbf{r}-\mathbf{r}'|$ small. This can be understood physically in the following way; $\Delta(\mathbf{r})$ represents the change in the charge density due to adding an electron. The $\Delta(\mathbf{r})\Delta(\mathbf{r}')$ term represents the interaction of this change with itself. In an infinite system $\Delta(\mathbf{r})$ is virtually zero and so this term should disappear. Removing this term yields

$$\begin{aligned} E^X_{e-e} &= \int_s |\Psi|^2 \sum_{i>j}^X f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \int_s\int_a n^X(\mathbf{r}) n^N(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \end{aligned}$$

[2]This is achieved by reducing the distances to all the other electrons into the Wigner-Seitz cell centred on the electron being considered.

$$- \frac{1}{2} \int_s \int_a n^N(\mathbf{r}) n^N(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r} - \mathbf{r}') - f(\mathbf{r} - \mathbf{r}') \right] d\mathbf{r} d\mathbf{r}' \quad , \tag{6.9}$$

where $n^X = n^N + \Delta(\mathbf{r})$.

As in Eq.(6.2), the first term describes the full Hartree and exchange/correlation interaction between all X electrons in the simulation cell. The second two terms can be interpreted as representing the Hartree interaction between the charge density due to X electrons inside the simulation cell and the charge density due to N electrons outside the simulation cell. Therefore, as far as the electron-electron interaction is concerned there is *only one extra electron* present in the system rather than the whole periodic array which is normally introduced. The use of this new energy expression removes the need for *ad hoc* corrections to the finite size effects such as those used by Engel et al. in Eq.(6.6). Note, when using either of the two energy expressions, Eq.(6.2) and Eq.(6.9), we include background charges so there is no contribution to the total energy or any gap energies from the $\mathbf{G} = 0$ component of the f interaction. This is equivalent to ensuring that each cell is neutral, as would be the case when a single electron is added to the infinite system.

6.3.4 Hartree-Fock Analysis of New Interactions

The difference between the new interaction introduced in chapter 5, Eq.(6.2), and the enhanced version of the interaction in Eq.(6.8) can be clearly seen within Hartree-Fock (HF) theory.

New Interaction, Eq.(6.2)

Within the HF approximation the interaction of Eq.(6.2) leads to the following expression for the total energy;

$$E_N^{HF} = \sum_i \int \phi_i^*(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 \right) \phi_i(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \int n(\mathbf{r}) n(\mathbf{r}') \hat{\nu}_{\text{Ewald}}(\mathbf{r} - \mathbf{r}') d\mathbf{r} d\mathbf{r}'$$

$$- \frac{1}{2}\sum_{i\neq j}^{N} \delta_{s_i s_j} \int\int \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}')f(\mathbf{r}-\mathbf{r}')\phi_j^*(\mathbf{r}')\phi_j(\mathbf{r})d\mathbf{r}d\mathbf{r}' + \int V_{\rm ext}(\mathbf{r})n(\mathbf{r})d\mathbf{r}$$

$$, \tag{6.10}$$

where s_i is the spin of the i^{th} electron. The two charge densities in the second (Hartree) term have been chosen to be the same for simplicity. In the case of HF calculations performed using fixed LDA orbitals, this corresponds to using the LDA charge density as the 'input' charge density to the interaction in the same way as is done for VMC calculations.

The resultant HF equations obtained from minimising E_N^{HF} in Eq.(6.10) with respect to the $\{\phi_i\}$ are

$$-\frac{1}{2}\nabla^2\phi_k(\mathbf{r}) + \sum_i \int \phi_i^*(\mathbf{r}')\phi_i(\mathbf{r}')\phi_k(\mathbf{r})\hat{\nu}_{\rm Ewald}(\mathbf{r}-\mathbf{r}')d\mathbf{r}'$$

$$- \sum_i^N \delta_{s_i s_k} \int \phi_i^*(\mathbf{r}')\phi_i(\mathbf{r})\phi_k(\mathbf{r}')f(\mathbf{r}-\mathbf{r}')d\mathbf{r}' + V_{\rm ext}(\mathbf{r})\phi_k(\mathbf{r}) = \epsilon_k\phi_k(\mathbf{r}) \quad . \tag{6.11}$$

Therefore the eigenvalue, ϵ_k, is given by

$$\begin{aligned}\epsilon_k \quad = \quad & \int \phi_k^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\phi_k(\mathbf{r})d\mathbf{r} + \int\int \sum_i \phi_i^*(\mathbf{r}')\phi_i(\mathbf{r}')\phi_k^*(\mathbf{r})\phi_k(\mathbf{r})\hat{\nu}_{\rm Ewald}(\mathbf{r}-\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ - \quad & \sum_i^N \delta_{s_i s_k} \int\int \phi_i^*(\mathbf{r}')\phi_i(\mathbf{r})f(\mathbf{r}-\mathbf{r}')\phi_k^*(\mathbf{r})\phi_k(\mathbf{r}')d\mathbf{r}d\mathbf{r}' + \int V_{\rm ext}(\mathbf{r})\phi_k^*(\mathbf{r})\phi_k(\mathbf{r})d\mathbf{r}\end{aligned}$$

$$. \tag{6.12}$$

Eqs.(6.10) and (6.12) yield an analogue of Koopmans' theorem for adding an electron

$$E_k - E_N = \epsilon_k + \frac{1}{2}\int \phi_k^*(\mathbf{r})\phi_k(\mathbf{r})\phi_k^*(\mathbf{r}')\phi_k(\mathbf{r}')\left[\hat{\nu}_{\rm Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right]d\mathbf{r}d\mathbf{r}' \quad , \tag{6.13}$$

where E_k is the total energy of the system with an extra electron added into the k^{th} orbital. Note that if $f(\mathbf{r}-\mathbf{r}')$ is replaced with the standard $\hat{\nu}_{\rm Ewald}(\mathbf{r}-\mathbf{r}')$ we retrieve the standard Koopmans' theorem.

The equivalent expression for removing an electron is

$$E_N - E_k = \epsilon_k - \frac{1}{2}\int \phi_k^*(\mathbf{r})\phi_k(\mathbf{r})\phi_k^*(\mathbf{r}')\phi_k(\mathbf{r}')\left[\hat{\nu}_{\text{Ewald}}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \quad , \tag{6.14}$$

and therefore the HF energy gap, obtained using the expression for the electron-electron interaction of Eq.(6.2) is given by

$$\begin{aligned}(E_k - E_N) - (E_N - E_j) &= \epsilon_k - \epsilon_j + \frac{1}{2}\int \phi_k^*(\mathbf{r})\phi_k(\mathbf{r})\phi_k^*(\mathbf{r}')\phi_k(\mathbf{r}')\left[\hat{\nu}_{\text{Ewald}}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \\ &+ \frac{1}{2}\int \phi_j^*(\mathbf{r})\phi_j(\mathbf{r})\phi_j^*(\mathbf{r}')\phi_j(\mathbf{r}')\left[\hat{\nu}_{\text{Ewald}}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \quad . \end{aligned} \tag{6.15}$$

Koopmans' theorem has therefore been modified. The interaction is, in a sense, including self-interaction like terms.

Enhanced Version from Eq.(6.9)

Let us now repeat the above analysis using the enhanced expression for the electron-electron interaction of Eq.(6.9), designed to remove the effect of the unwanted periodic array of additional electrons. The HF equivalent of this energy expression is

$$\begin{aligned} E_X^{HF} &= \sum_i \int \phi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\phi_i(\mathbf{r})d\mathbf{r} \\ &+ \frac{1}{2}\int\int n_X(\mathbf{r})n_X(\mathbf{r}')f(\mathbf{r}-\mathbf{r}')d\mathbf{r}d\mathbf{r}' - \frac{1}{2}\sum_{i\neq j}^{X}\delta_{s_is_j}\int\int \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}')f(\mathbf{r}-\mathbf{r}')\phi_j^*(\mathbf{r}')\phi_j(\mathbf{r})d\mathbf{r}d\mathbf{r}' \\ &+ \int\int n_X(\mathbf{r})n_N(\mathbf{r}')\left[\hat{\nu}_{\text{Ewald}}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \\ &- \frac{1}{2}\int\int n_N(\mathbf{r})n_N(\mathbf{r}')\left[\hat{\nu}_{\text{Ewald}}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \\ &+ \int V_{\text{ext}}(\mathbf{r})n_X(\mathbf{r})d\mathbf{r} \quad , \end{aligned} \tag{6.16}$$

The resultant HF equations obtained from minimising E_X^{HF} in Eq.(6.16) with respect to the $\{\phi_i\}$ are exactly the same as Eq.(6.11), obtained from the original energy expression. The eigenvalues, ϵ_k, are therefore also exactly the same as those given in Eq.(6.12). However, if one considers the change in the total energy on adding an electron to state k, using the energy expression in Eq.(6.16), one obtains a similar expression to Eq.(6.13) but without the term arising from the interaction between the k^{th} electron and its images, i.e.

$$\begin{aligned} E_k - E_N &= \int \phi_k^*(\mathbf{r}) - \frac{1}{2}\nabla^2 \phi_k(\mathbf{r}) d\mathbf{r} + \int\int \sum_i \phi_i^*(\mathbf{r}')\phi_i(\mathbf{r}')\phi_k^*(\mathbf{r})\phi_k(\mathbf{r})\hat{\nu}_{\text{ewald}}(\mathbf{r}-\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ &- \sum_i^N \delta_{s_i s_k} \int\int \phi_i^*(\mathbf{r}')\phi_i(\mathbf{r})f(\mathbf{r}-\mathbf{r}')\phi_k^*(\mathbf{r})\phi_k(\mathbf{r}')d\mathbf{r}d\mathbf{r}' + \int V_{\text{ext}}(\mathbf{r})\phi_k^*(\mathbf{r})\phi_k(\mathbf{r})d\mathbf{r} \\ &= \epsilon_k \quad . \end{aligned} \tag{6.17}$$

Therefore, when using the enhanced version of the electron-electron interaction from Eq.(6.9), one recovers the standard version of Koopmans' theorem where the eigenstates of the HF equations correspond to the excitation energies of the system,

$$E_k - E_N = \epsilon_k \quad , \tag{6.18}$$

and hence the HF energy gap, obtained using the expression for the electron-electron interaction of Eq.(6.8) is given by

$$(E_k - E_N) - (E_N - E_j) = \epsilon_k - \epsilon_j \quad . \tag{6.19}$$

The comparison of the two energy expressions, Eq.(6.2) and Eq.(6.9), within HF theory can therefore be summarised as follows. The enhanced version of the electron-electron interaction, Eq.(6.8), improves over the original expression in two ways, (i) it removes the self-term due to the interaction of an electron and its images, (ii) it recovers a proper version of Koopmans' theorem. Both these results provide additional support for the use of the enhanced interaction, Eq.(6.8), in the following QMC calculations.

6.3.5 Addition and Subtraction of Electrons in VMC

The same n=2, 16 atom, 64 electron system of diamond-structure silicon that was used in the Coulomb finite size calculations described in section 5.7 was used to calculate a series of energy gaps using the prescription given in Eq.(6.1) within the VMC framework by the method of addition and subtraction of electrons. Calculations were performed using the new electron-electron interaction designed to remove the effects of the periodic images of the additional (removed) electrons as described in Eq.(6.9).

VMC Results

The VMC results for the energy gaps at all possible **k**-points in an n=2 simulation cell are summarised in Table 6.1. For each of the **k**-points the$n$=2 LDA results and n=2 Hartree-Fock results have been included for comparison. The Hartree-Fock results were obtained using the fixed orbitals from the corresponding LDA calculation rather than performing a fully relaxed Hartree-Fock calculation. The LDA bandstructure exhibits virtually no finite size effect (< 0.2 eV at all points across the band), whereas the HF bandstructure exhibits a large finite size effect[112] (> 3.0 eV at all point across the band).

Band	VMC (eV)	HF (eV)	LDA (eV)	Experiment[94] (eV)
Γ_v	0	0	0	0
Γ_c	2.6 $\pm$0.4	5.0	2.43	3.4
X_v	-4.1 $\pm$0.4	-3.31	-2.93	-2.9
X_c	3.7 $\pm$0.4	2.73	0.46	1.17
L_v	-1.6 $\pm$0.4	-1.36	-1.23	-1.2 $\pm$ 0.2
L_c	2.0 $\pm$0.4	4.0	1.39	2.23

Table 6.1: VMC results for the addition and subtraction of electrons.

The results for VMC, LDA, HF and experiment have all been aligned by setting the value of $\Gamma_v = 0$ in each case. The values at each of the points in the bandstructure can then be obtained relative to this point, for example

$$\begin{aligned}
\Gamma_c &= \Gamma_v + (E_{N+1\Gamma} - E_N) - (E_N - E_{N-1\Gamma}) \\
X_c &= \Gamma_v + (E_{N+1X} - E_N) - (E_N - E_{N-1\Gamma}) \\
X_v &= X_c - \text{Gap at X} \\
&= (E_{N-1\Gamma} - E_{N-1X}) \\
&\mathit{etc.} \quad ,
\end{aligned} \tag{6.20}$$

where $E_{N+1\Gamma}$ refers to the ground state energy of a system with an extra electron added into an orbital that describes the bottom of the conduction band at the Γ-point and E_{N-1X} refers to the ground state energy of a system with an electron removed from an orbital that describes the top of the valence band at the X-point.

It can be clearly seen that the LDA gives the valence band energies (relative to Γ_v) at each of the **k**-points more accurately than the conduction band energies. The LDA significantly underestimates all the band gaps.

The quality of the VMC results is mixed. At the Γ- and L-points, the results are broadly in agreement with experiment. At the X-point the VMC overestimates the size of the gap. The errors in the VMC calculations are due to the quality of the trial wavefunction used and finite size effects. As mentioned in section 6.3, the $\chi(\mathbf{r})$ function for the excited states has a higher symmetry than the excited state charge density and the u function takes no account of any changes in the shape of the exchange-correlation hole due to the addition/removal of electrons. All the VMC trial wavefunctions contain a Slater determinant constructed using orbitals from a groundstate LDA calculation. The effect of allowing these orbitals to relax were suspected to be too small to resolve in VMC, but has been examined in DMC in the following section.

6.3.6 Addition and Subtraction of Electrons in DMC

In an attempt to improve on the VMC results from section 6.3.5 a selected set of the VMC calculations were repeated within DMC. As described in chapter 5, the DMC algorithm requires not only an energy expression but also the associated Hamiltonian. The Hamiltonian corresponding to the new electron-electron energy expression designed to remove the long range finite size effects introduced by the periodic boundary conditions acting on the additional electron is

$$\hat{H}_{e-e} = \sum_{i>j}^{X} f(\mathbf{r}_i - \mathbf{r}_j)$$
$$+ \sum_{i} \int_{s} \left[\hat{\nu}_{Ewald}(\mathbf{r}_i - \mathbf{r}) - f(\mathbf{r}_i - \mathbf{r})\right] n_{LDA}^{N}(\mathbf{r}) d\mathbf{r} \quad . \tag{6.21}$$

The Hamiltonian is physically very reasonable. It describes each electron 'feeling' the full $1/r$ interaction with all the other X electrons within a Wigner-Seitz cell centred on the electron and the Hartree interaction with a charge density due to N electrons outside the Wigner-Seitz cell.

The total electron-electron energy, E_{e-e}, in DMC is then

$$E_{e-e} = \langle\Psi|\hat{H}|\Phi\rangle - \frac{1}{2}\int\int n_{DMC}^{N}(\mathbf{r}) n_{LDA}^{N}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r} - \mathbf{r}') - f(\mathbf{r} - \mathbf{r}')\right] d\mathbf{r}d\mathbf{r}' \quad . \tag{6.22}$$

The second term in Eq.(6.22) can be accumulated during a ground state, N electron, DMC calculation and then just subtracted from the energies calculated with $X = N \pm 1$ electrons.

In the following DMC calculations each of the E_N, E_{N+1} and E_{N-1} calculations use the VMC trial wavefunction as the guiding wavefunction, Ψ_G, optimised for that system using the Hamiltonian of Eq.(6.21).

DMC Results

As DMC calculations are considerably more computationally expensive than VMC calculations, only a selected set of the VMC results were repeated in DMC. It was decided to use the addition and removal of electrons at the Γ-point as a test for comparing the DMC with VMC and experiment. Two separate modifications to the DMC algorithm were experimented with,

i. DMC calculations were performed using the original Ewald expression for the electron-electron interaction and the new version of the interaction described in section 6.3.3.

ii. DMC calculations were performed using single-particle orbitals in the Slater determinant from LDA calculations where the orbitals are kept fixed when an electron is added or removed and where the orbitals are allowed to relax when an electron is added or removed. This relaxation will change the nodal structure of the guiding wavefunction.

The first set of DMC calculations were performed using the new electron-electron interaction described in section 6.3.3 and a Slater determinant containing fixed LDA orbitals. The value of the gap at the Γ-point obtained from adding and removing an electron under these conditions was 3.95 $\pm$0.4 eV. This is almost within error bars of the experimental gap of 3.4 eV.

The same set of three calculations were repeated, but this time allowing the LDA orbitals used in the Slater determinant to relax. The value of the gap at the Γ-point was reduced to 3.59 $\pm$0.4 eV.

Finally, the calculations using the relaxed LDA orbitals were repeated using the original Ewald expression for the electron-electron interaction. The value of the gap at the Γ-point was reduced further to 3.34 $\pm$0.4 eV.

From these results it appears that

i. The DMC performs at least as well, if not better than the VMC. This is to be expected because the DMC calculations involving $N \pm 1$ electrons do not suffer from the limitations in the one- and two-body terms in the trial wavefunction.

ii. Relaxing the LDA orbitals used to construct the Slater determinant appears to have a small beneficial effect on the DMC results. This could be due to the relaxed orbitals reproducing the nodal surface of the true many-body wavefunction more accurately than fixed orbitals.

iii. There appears to be very little difference in DMC results for gaps obtained using the new interaction described in section 6.3.3 and the original Ewald interaction. As the new interaction is designed to remove the long range finite size errors this suggests that the finite size errors present in the *gap* are a short range effect not a long range effect as suggested in Ref.[79]. The considerable finite size errors present in the individual N, $N+1$ and $N-1$ calculations performed with the Ewald interaction almost entirely cancel from the gap.

   The relative insensitivity of the gap energy to the choice of electron-electron interaction is confirmed by the HF results. These show a small but consistent improvement when using the new interaction in preference to the Ewald interaction. For example in an n=2 system, the gap at the Γ point is improved by 15% when using the new interaction.

6.4 Promoting Electrons

The second method used to calculate excited states within QMC is that of promoting a single electron from a state in the valence band to a state in the conduction band (see figure 6.2). An excitation energy or gap can then be defined as the difference in energy between this new excited state and the ground state of the system,

$$E_{gap} = E_N^{Excited} - E_N \quad . \tag{6.23}$$

6.4.1 Trial Wavefunctions for Promoted States

To promote a single electron from the valence band to the conduction band, all that is required is to replace one of the one-electron orbitals in the Slater determinant part of the trial wavefunction that describes a state in the valence band with one that describes a state in the conduction band. i.e.

$$\begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) & \cdots \\ \phi_2(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) & \cdots \\ \vdots & \vdots & \\ \phi_i(\mathbf{r}_1) & \phi_i(\mathbf{r}_2) & \cdots \\ \vdots & \vdots & \\ \phi_N(\mathbf{r}_1) & \phi_N(\mathbf{r}_2) & \cdots \end{vmatrix} \xrightarrow{\text{promote}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) & \cdots \\ \phi_2(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) & \cdots \\ \vdots & \vdots & \\ \phi'_i(\mathbf{r}_1) & \phi'_i(\mathbf{r}_2) & \cdots \\ \vdots & \vdots & \\ \phi_N(\mathbf{r}_1) & \phi_N(\mathbf{r}_2) & \cdots \end{vmatrix} \tag{6.24}$$

where the ith ground state orbital has been replaced in the excited state Slater determinant by the $\phi'_i(\mathbf{r})$ excited state orbital.

Spin Contamination

The ground state of the silicon atoms in the bulk solid is a closed shell configuration. If a single up- or down-spin electron is excited from the valence band to the conduction band, the wave function then has a spin contamination, i.e. it is no longer an eigenstate of the $\hat{S}^2$ operator. It is possible to avoid this contamination either by exciting simultaneously the electron with opposite spin, to create a biexciton, or to use a superposition of determinantal products to obtain a spin singlet state. This superposed trial wavefunction is a generalisation of the trial wavefunction introduced in chapter 2,

$$\Psi_T(R) = \sum_n \alpha_n D_n^{\uparrow} D_n^{\downarrow} \exp\left(\sum_{(s,i)=(\uparrow,1)}^{(\downarrow,N)} \chi(\mathbf{r}_i) \right) \exp\left(- \sum_{(\uparrow,1)\leq(s,i)<(s',j)}^{(\downarrow,N)} u(r_{ij}) \right) \quad . \tag{6.25}$$

A spin-singlet wavefunction for diamond-structure silicon can be obtained from Eq.(6.25), by the superposition of two determinantal products,

$$\text{Determinant} = \frac{1}{\sqrt{2}} \left(\tilde{D}_1^{\uparrow} D_1^{\downarrow} + D_2^{\uparrow} \tilde{D}_2^{\downarrow} \right) \quad , \tag{6.26}$$

where $\tilde{D}$ indicates a Slater determinant in which one of the ground state orbitals has been replaced by an excited state orbital. In the first product, the down-spin determinant contains the ground state orbitals and the up-spin determinant has one of the ground state orbitals replaced by an excited state orbital. In the second product the up-spin determinant contains the ground state orbitals and the down-spin determinant has one of the ground state orbitals replaced with an excited state orbital.

One and Two-body Functions for Promotion Calculations

To perform a VMC calculation using an excited state Slater determinant, it is necessary to optimise the Jastrow and $\chi(\mathbf{r})$ functions to minimise the energy/variance of the energy of the wavefunction containing this new determinant. In an attempt to speed up the turn around for performing QMC calculations on excited states, it was decided to experiment with omitting the extra optimisation required to convert the ground state Jastrow and $\chi(\mathbf{r})$ functions to excited state Jastrow and $\chi(\mathbf{r})$ functions. Instead of performing this extra optimisation for each excitation we proceed straight to the DMC calculations. This is possible as long as the trial/guiding wavefunction is of sufficient quality. In the case of DMC calculations this implies two basic requirements;

(i). The trial/guiding wavefunction for the excited state must be orthogonal to the ground state wavefunction and all lower energy exact eigenstates or the DMC algorithm will simply propagate out the lowest energy solution.

(ii). The trial/guiding wavefunction must have an acceptably low variance of the local energy to ensure that the fluctuations in the population of walkers are manageable (see chapters 2 and 4 for details).

In all the excitation calculations described here, the excited orbitals in the excited state Slater determinants have different $\mathbf{k}$ values to the ground state orbital they replace, i.e. they are indirect excitations. This ensures that, on the grounds of translational symmetry, all the excited state wavefunctions are orthogonal to the ground state wavefunction, hence satisfying condition (i). Condition (ii) is not so clear cut. In section 6.3, where one is adding and removing electrons from the system, the relaxation in the Jastrow and $\chi(\mathbf{r})$ functions between systems with $N-1$, N and $N+1$ electrons was significant. This made it necessary to re-optimise the trial/guiding wavefunction for each system. When one promotes an orbital within the Slater determinant as in these calculations, the resulting change in the charge density and hence the change in the Jastrow and $\chi(\mathbf{r})$ functions is not as severe and so it was decided to attempt the DMC calculations for promoted states using the same one- and two-body functions as those in the trial/guiding wavefunction used for the ground state DMC calculation. This will not affect the DMC estimate of the total energy, but it will increase its variance by an amount dependent on the quality of the trial/guiding wavefunction.

6.4.2 Electron-Electron Interaction for Promoted States

The new electron-electron interaction introduced in chapter 5 and the enhanced version of this interaction described in section 6.3.3, which was designed to remove the interaction between the periodic array of additional electrons (holes) both have equivalent forms for dealing with systems containing promoted electrons. The equivalent of the new electron-electron interaction from chapter 5, designed to deal with N electron systems in their ground state, can also be written to deal with N electrons, where one is in an excited state, in the following way;

$$\begin{aligned} E_{\mathrm{e-e}}^{\mathrm{excited}} &= \int\limits_s |\Psi|^2 \sum_{i>j}^{N} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \frac{1}{2} \int\limits_s \int\limits_a n_{DMC}^{\mathrm{excited}}(\mathbf{r}) n_{LDA}^{\mathrm{excited}}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \end{aligned} \tag{6.27}$$

There also exists an equivalent of the enhanced interaction introduced in section 6.3.3 to remove the extra electrostatic energy introduced when an electron (hole) is added to the simulation cell. In the case of promoting electrons rather than addition and subtraction, this enhanced interaction (see Eq.(6.28)) corresponds to evaluating the Hartree interaction between the charge density due to the promoted system in the simulation cell and the charge density due to the ground state in all the periodic images of the simulation cell. In other words, the enhanced form removes any additional electrostatic energy due to the Hartree interaction between changes in charge density due to the promotion of an electron in the simulation cell and the same changes in the periodic images of the simulation cell

$$\begin{aligned} E_{\mathrm{enhanced\ e-e}}^{\mathrm{excited}} &= \int\limits_s |\Psi|^2 \sum_{i>j}^{N} f(\mathbf{r}_i - \mathbf{r}_j) \prod_k d\mathbf{r}_k \\ &+ \int\limits_s \int\limits_a n_{DMC}^{\mathrm{excited}}(\mathbf{r}) n_{LDA}^{N}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \\ &- \frac{1}{2} \int\limits_s \int\limits_a n_{DMC}^{N}(\mathbf{r}) n_{LDA}^{N}(\mathbf{r}') \left[\hat{\nu}_{Ewald}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')\right] d\mathbf{r} d\mathbf{r}' \quad . \end{aligned}$$

(6.28)

6.4.3 DMC Results

All of the possible excitations accessible in an n=2, 16 atom simulation cell are illustrated in figure 6.2. The promotion of an electron from the orbital representing the top of the valence band at the Γ-point to the orbital representing the bottom of the conduction band at the X-point ($E_N^{\Gamma\rightarrow X} - E_N^0$) was used as a test bed for the following variations in the DMC technique:

(i). As described in section 6.4.1, the use of a single product of Slater determinants for up and down spin electrons in which either the up-spin determinant contains an excited state orbital and the down-spin determinant contains ground state orbitals or vice versa, is spin contaminated. There are no calculations in the literature to indicate how severe the effect of this contamination is. Using the dual determinant wavefunction given in Eq.(6.26) slows down the code by almost a factor of two (as operations on the determinant(s) dominate the calculation), so it was decided to perform tests using single and dual determinants to determine whether the effect of the spin contamination is resolvable from the statistical noise and therefore whether the use of dual determinant wavefunctions is necessary.

(ii). Separate DMC calculations were performed using the two electron-electron interactions for the excited state described in Eqs.(6.27) and (6.28). These correspond to using either the ground state or excited state LDA charge density in the Hamiltonian of Eq.(6.21).

(iii). The one-electron orbitals used in the Slater determinant part of the trial/guiding wavefunction are generated from an LDA calculation (see chapter 2). Within the LDA there is only a very small change in the excitation energy if one relaxes the orbitals after promoting an electron to the conduction band compared with using fixed ground state orbitals to calculate all energy differences. Two separate

DMC calculations were performed on the $E_N^{\Gamma \to X}$ excited state to check whether using relaxed rather than fixed LDA orbitals in the Slater determinant had a significant effect on the DMC excitation energy. This could happen if the process of relaxing the LDA orbitals significantly altered the nodal structure produced by the Slater determinant.

The following DMC calculations were performed:

(i). A wavefunction containing a full dual determinant, as shown in Eq.(6.26), was used to represent the state in which one electron is promoted from the top of the valence band to the bottom of the conduction band. The LDA orbitals were not relaxed from their ground state forms. The enhanced interaction of Eq.(6.28) was used to model the electron-electron interaction. The total energy for the excited state system was -107.21 $\pm$ 0.05 eV per atom.

(ii). A wavefunction containing a single determinant of one-electron orbitals was used. Again the enhanced interaction of Eq.(6.28) was used to model the electron-electron interaction. Again the LDA orbitals were not relaxed from their ground state forms. The total energy for the excited state system was -107.22 $\pm$ 0.02 eV per atom.

(iii). A wavefunction containing a single determinant of one-electron orbitals was used again. This time the less sophisticated interaction of Eq.(6.27) was used to model the electron-electron interaction and the one-electron orbitals were relaxed in the LDA. The total energy for the excited state system was -107.22 $\pm$ 0.02 eV per atom.

The above results suggest that for the test case of promoting a single electron from the top of the valence band at the Γ-point to the bottom of the conduction band at the X-point, (i) The effect of spin contamination is not resolvable from the statistical noise, (ii) The two choices of electron-electron interaction yield the same energy, and (iii) the effect of relaxing the one-electron orbitals within the LDA has no significant effect on the total energy.

In the light of the above results, all the following promotion calculations are based on the enhanced interaction of Eq.(6.28). Although this is the more sophisticated interaction, it is actually slightly simpler to implement, because it only relies on the LDA ground state charge density as an input, whereas the less sophisticated interaction of Eq.(6.27) requires a separate LDA calculation of each excited state charge density for use as an input. Also, in all the following calculations, the same LDA orbitals obtained from a ground state calculation, have been used to construct the Slater determinant, again to simplify the setup procedure. A single determinantal product was used to represent the excited state to speed up the computation. The DMC calculations were performed using 384 configurations distributed over 128 nodes of the parallel computer. The diffusion algorithm used between 1500 and 2000 time steps. Approximately 250 of these time steps were required for the initial propagation stage of the algorithm (see chapter 2) and the remainder were used to accumulate statistics.

The results are shown in table 6.2. Again, the equivalent n=2 HF and LDA results have been included for comparison. It should be noted that is in the addition and subtraction of electron results, the LDA results contain only a small finite size effect, whereas the HF results contain a large finite size effect.

On the whole, the calculations appear extremely successful, with a significant fraction of the results in agreement with experiment to within error bars. Those calculations which significantly disagree with experiment all overestimate the size of the gap. This would be consistent with the quality of the trial wavefunction for the excited state is not being as good as that for the ground state. In particular the nodal structure of the excited states may not resemble the true nodal structure as closely as that of the ground states. These approximations to the excited states will tend to increase the estimate of the energy of the excited state and hence produce estimates of the gap that are too large.

The average deviation of the DMC energies from experiment is 0.3 eV. For the LDA it is -1.0 eV and for HF it is +6.8 eV.

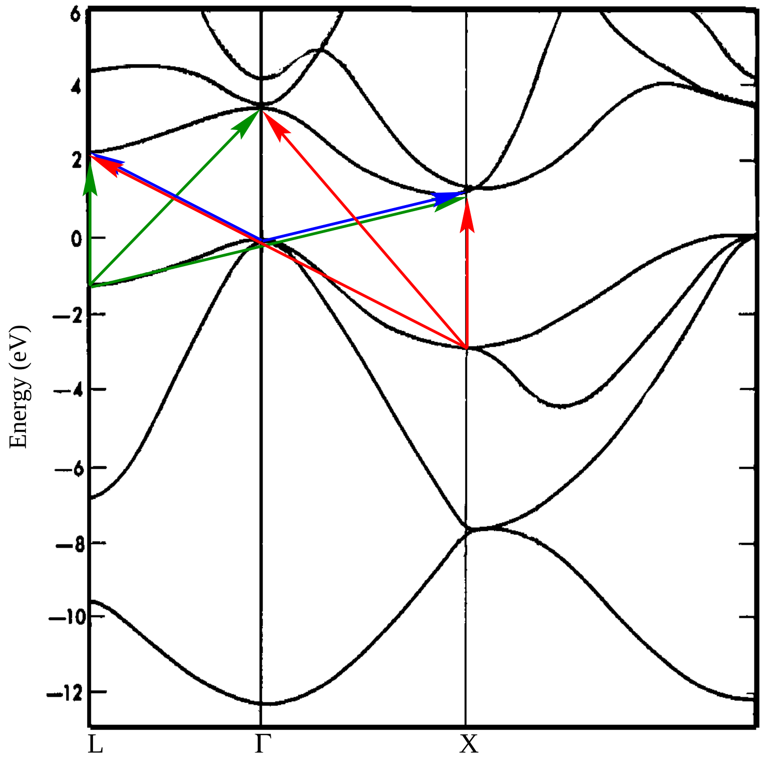


Figure 6.2: Pseudopotential band structure of silicon showing the Γ, X and L-points, taken from Ref.[94]. All possible excitations from the top of the valence band to the bottom of the conduction band are shown. Excitations from the Γ-point are shown in blue, from the X-point in red and from the L-point in green.

Promotion	DMC Gap (eV)	Expt. Gap (eV)[94]	LDA Gap (eV)	HF Gap (eV)
$\Gamma_v \to X_c$	1.2 ± 0.3	1.2	0.46	2.73
$X_v \to \Gamma_c$	7.0 ± 0.4	6.3	5.36	8.81
$\Gamma_v \to L_c$	2.24 ± 0.4	2.4	1.39	4.00
$L_v \to \Gamma_c$	5.6 ± 0.4	4.6	3.66	6.36
$X_v \to L_c$	5.6 ± 0.4	5.2	4.32	7.31
$L_v \to X_c$	2.6 ± 0.4	2.4	1.69	4.09
$X_v \to X_c$ (*)	4.9 ± 0.4	4.1	3.39	6.04
$L_v \to L_c$ (*)	3.4 ± 0.4	3.4	2.62	5.36

Table 6.2: DMC calculations for promoting electrons. Those excitations marked with a (*) are between distinct but equivalent **k**-points. The excited state is therefore still orthogonal to the ground state due to the different translational symmetry. All DMC calculations have been corrected to remove the exciton energy using Eq.(6.30)

Calculating Band Widths

As well as calculating energy gaps, it is also possible to calculate a band width via indirect excitations. In figure 6.3, two sets of excitations are shown for calculating the width of the valence band at the Γ-point.

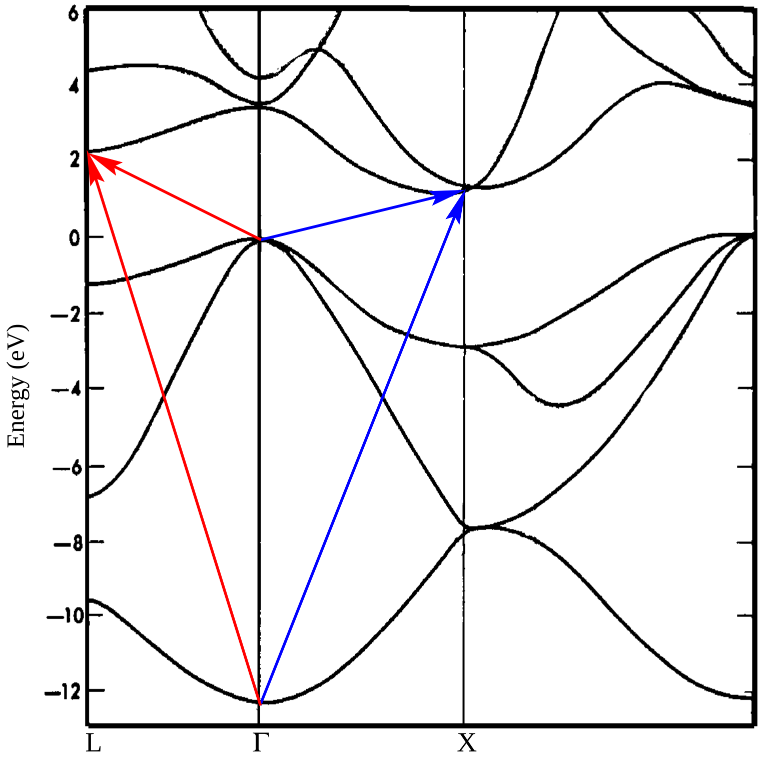


Figure 6.3: Indirect excitations to calculate the width of the valence band at the Γ-point. Excitations from $\Gamma \to X$ are shown in red and from $\Gamma \to L$ are shown in blue. The black lines represent a pseudopotential bandstructure[94].

The band width can be calculated as the difference in energy between exciting from the bottom of the valence band at Γ to the bottom of the conduction band at X or L and exciting from the top of the valence band at Γ to the bottom of the conduction band at X or L,

$$\text{Band Width} \;=\; (E_N^{\Gamma_1 \to X_5} - E_N^0) - (E_N^{\Gamma_4 \to X_5} - E_N^0)$$

$$
\begin{aligned}
&= E_N^{\Gamma_1 \to X_5} - E_N^{\Gamma_4 \to X_5} \\
\text{or Band Width} &= (E_N^{\Gamma_1 \to L_5} - E_N^0) - (E_N^{\Gamma_4 \to L_5} - E_N^0) \\
&= E_N^{\Gamma_1 \to L_5} - E_N^{\Gamma_4 \to L_5} \quad . \qquad (6.29)
\end{aligned}
$$

where $E_N^{\Gamma_4 \to L_5}$ represents the energy of the state where the orbital in the Slater determinant representing the top of the valence band at the Γ-point(Γ_4) has been replaced by one representing the bottom of the conduction band at the L-point(L_5). The results of these DMC calculations of the band width are shown in table 6.3. The n=2 LDA and HF band widths have been included for comparison. As with excitation energies the LDA band widths contain a very small finite size effect. The fully converged HF result for the band width is 18.5 eV[112].

The DMC results show a reasonable agreement with experiment, if not quite within error bars. As with the ordinary excitations, the DMC again overestimates the size of the band width. This overestimation could be due to the inferior quality of the nodal structure of the guiding wavefunction for the excitation from the bottom of the valence band to the bottom of the conduction band or to finite size effects.

1^{st}Promotion	2^{nd}Promotion	DMC (eV)	Expt. (eV)	LDA (eV)	HF (eV)
$\Gamma_4 \to X_5$	$\Gamma_1 \to X_5$	13.3 ± 0.6	12.5	12.03	16.2
$\Gamma_4 \to L_5$	$\Gamma_1 \to L_5$	14.1 ± 0.6	12.5	12.03	16.2

Table 6.3: Calculations of band widths. $\Gamma_4 \to X_5$ refers to the promotion of an electron from the 4^{th} band at the Γ-point to the 5^{th} band at the X-point

Excitonic Effects

Band gap energies calculated by the method of promoting an electron will be reduced in comparison to those calculated by addition and subtraction by an amount equal to the binding energy of the exciton. If, following Mitas[27], we suppose the exciton is of the Mott-Wannier [113, 114] type, then its binding energy is given by [76],

$$E_{\text{exciton}} = \frac{1}{2\epsilon r_0} \quad . \tag{6.30}$$

As the simulation cells used here are relatively small, the exciton is artificially localised. This can be taken account of in an approximate manner by choosing r_0 in Eq.(6.30) equal to the edge length of the simulation cell. For the $n = 2$ simulation cell used in these calculations, the exciton energy is estimated to be ≈ 0.1 eV for the *whole* simulation cell. This correction has already been added to the band gap results shown in Table 6.2. It cancels out of the band width results in Table 6.3.

6.5 Summary and Comparison of the Methods

The method of addition and subtraction of electrons to calculate a gap energy within QMC produced only mixed quality results within VMC. Even with the use of the enhanced electron-electron interaction to reduce the finite size effects in the gap energies, the VMC results are only of broadly comparable accuracy to the LDA and if anything a little worse. In contrast to the LDA, the VMC results all overestimate the band gaps and this is attributable to the inferior quality of the trial wavefunctions for the systems with $N \pm 1$ electrons.

Within DMC the results from addition and subtraction of electrons are much improved over VMC and represent a significant improvement over the LDA. It appears that contrary to our initial expectations, the results are not strongly sensitive to the choice of electron-electron interaction. In fact both the gap energy calculated using the Ewald interaction and the gap energy calculated using the new interaction agree with the experimental result to within error bars. This suggests that the finite size effects in the gap energies are a short range phenomena, not long range as suggested by Ref.[79].

The method of promoting electrons to calculate excited state energies proved successful with DMC. The results show that for the relatively simple excitations performed here, a single determinantal product is sufficient to represent the excited state. As with the addition and subtraction method, the results appear relatively insensitive to the choice of electron-electron interaction. Also, it appears that when promoting electrons, there is no need to relax the single-particle orbitals within the LDA before constructing the Slater determinant. The additional simplification of using the one- and two-body functions from the ground state trial wavefunction in the guiding wavefunction for the excited state proved successful as all the DMC calculations were numerically stable and did not exhibit any large fluctuations in the population of walkers indicative of a poor quality guiding wavefunction.

When comparing the results of all the possible excitations from the top of the valence band to the bottom of the conduction band, accessible in a n=2 supercell, the DMC shows a significant improvement over the LDA for all the excitations. On average, the DMC reduced the the difference in the excitation energy between the LDA and experiment by a factor of 3.

6.6 Alternative Methods for Calculating Excited States within QMC

6.6.1 Spectrum Folding

It is theoretically possible to use the fact that the variance of the local energy of an eigenstate of the Hamiltonian is zero to calculate excitation energies within QMC. The zero variance property of the ground state has already been utilised within the variance minimisation procedure described in chapter 4. In that case, the zero variance provides a useful lower bound to the quantity being minimised, namely the variance.

It is possible to extend this principle to search for the excited eigenstates of the same Hamiltonian. An ensemble of independent configurations were sampled from the ground state wavefunction using the procedure described in chapter 4. The variance of the local energy of this ensemble was then evaluated using the following expression,

$$\sigma^2 = \sum_{\alpha} \left(\Psi_T^{-1}(\mathbf{R}_\alpha) \hat{H} \Psi_T(\mathbf{R}_\alpha) - E_{\text{fixed}} \right)^2 \left[\frac{w(\alpha)}{\sum_\beta w(\beta)} \right] \quad , \tag{6.31}$$

where E_{fixed} is the energy about which the variance is to be evaluated. The variance about E_{fixed} was then minimised with respect to the variational parameters in the wavefunction for a series of values of E_{fixed} (i.e. scan over the energy range). The hope was that around each eigenvalue of the Hamiltonian, the variance should decrease considerably hence indicating the presence of such an eigenvalue. This is represented schematically in figure 6.4.

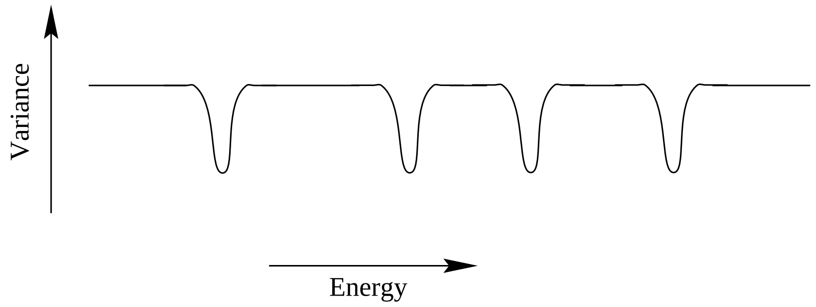


Figure 6.4: Schematic representation of the relationship between the local energy and its variance.

Unfortunately, we found that any small changes in the variance due to changes we made in the value of E_{fixed} were not distinguishable from the statistical noise present. To make this procedure effective, a considerably larger number of configurations would be required than are normally used in the variance minimisation procedure. This makes the technique prohibitively expensive.

6.6.2 DMC Decay Curves

The Monte Carlo solution to the diffusion equation can be written as a function of position, $\mathbf{R}$, and imaginary time, τ, as follows

$$\Psi(\mathbf{R}, \tau) = \sum_{i=0}^{\infty} c_i \phi_i(\mathbf{r}) e^{-(\epsilon_i - E_T)\tau} \quad , \tag{6.32}$$

where the coefficients, c_i, are the overlap integrals of $\Psi(\mathbf{R},0)$ with the eigenfunctions of the many-body Hamiltonian, ϕ_i. The DMC method relies on the fact that in the limit of large imaginary time $\Psi(\mathbf{R},\tau)$ is dominated by the lowest energy solution, ϕ_0.

However, in the initial short imaginary time regime, it is clear that the above equation contains information about the energy differences, $[\epsilon_i - E_T]$. For example the time-dependence of the energy estimate is given by

$$E(\tau) = \frac{\int f(\mathbf{R},\tau)E_L(\mathbf{R})d\mathbf{R}}{\int f(\mathbf{R},\tau)d\mathbf{R}} = \frac{\int \Psi(\mathbf{R},\tau)\hat{H}\Psi_G(\mathbf{R})d\mathbf{R}}{\int \Psi(\mathbf{R},\tau)\Psi_G(\mathbf{R})d\mathbf{R}}$$

$$= \frac{\sum_{i=0}^{\infty} c_i a_i e^{-(\epsilon_i - E_T)\tau}\epsilon_i}{\sum_{i=0}^{\infty} c_i a_i e^{-(\epsilon_i - E_T)\tau}} \quad , \tag{6.33}$$

where the a_i are the overlap integrals of the guiding wavefunction, $\Psi_G(\mathbf{R})$, with the eigenfunctions of the many-body Hamiltonian, ϕ_i. Therefore, if one was to compute the energy estimate as a function of time, then standard curve fitting methods could be used to extract the excited state energies, ϵ_i. In practice however, obtaining energies from Eq.(6.33) would be extremely difficult due to the statistical noise of the Monte Carlo simulation.

More sophisticated methods have been devised[80, 81, 82] to specifically measure the time dependence. Instead of the energy in Eq.(6.33), consider the expectation value of the Green's function,

$$I(\tau) = \left\langle \Psi_G | e^{-(\hat{H}-E_T)\tau} | \Psi_G \right\rangle \quad , \tag{6.34}$$

which can be sampled from the random walk by evaluating

$$I(\tau) = \langle W(\tau) \rangle_{\Psi_G^2} \quad , \tag{6.35}$$

where W is the cumulative branching weight,

$$W(\tau) = \prod_{k=0}^{N} e^{-[E_L(\mathbf{R},\tau)-E_T]\tau} \quad , \tag{6.36}$$

which is essentially the total population. If we insert a complete set of eigenstates into Eq.(6.36), the time behaviour of $I(\tau)$ is

$$I(\tau) = \sum_i a_i^2 e^{-(\epsilon_i - E_T)\tau} \quad . \tag{6.37}$$

This is a simpler expression to attempt to fit than Eq.(6.33). However, in practice it is still dominated by the statistical noise[23].

Chapter 7

Conclusions

In chapter 4 we have developed a technique for producing optimised trial wavefunctions based on the method of minimising the variance of the local energy. This technique has been successfully applied to both atoms and periodic solids. Additional enhancements to the technique enabled wavefunctions to be optimised with respect to different electron-electron interactions and with Hamiltonians containing non-local pseudopotentials.

The variance minimisation algorithm has been extended to work on parallel computer architectures and this allows optimisations to be performed over very large ensembles of up to one million configurations. These large ensembles of configurations enable wavefunctions to be optimised, within their parameter spaces, to an order of magnitude more accuracy than those currently described in the literature. New functional forms of one- and two-body functions have been introduced that are both more suited to optimisation and considerably faster to evaluate with the QMC code.

VMC calculations performed using these optimised trial wavefunctions have considerably lower variational energies. In most cases the energy difference between VMC and DMC has been approximately halved. When calculating cohesive energies, VMC calculations using the optimised trial wavefunctions yield almost identical results to the more sophisticated DMC technique.

In chapter 5 we have identified the source of the Coulomb finite size error present in QMC calculations as the Ewald interaction. We introduce a new electron-electron interaction which dramatically reduces the Coulomb finite size effects present in QMC and HF calculations. This interaction has been successfully tested in both a homogeneous system (HEG) and an inhomogeneous system (silicon). The interaction is equally successful when used in VMC and DMC calculations.

In chapter 6 we combined the technical advances made in the QMC methodology of chapters 4 and 5 to tackle the problem of calculating excitation energies within QMC.

Two separate methods of calculating excitation energies within QMC were introduced, (i) the addition and subtraction of electrons and (ii) the promotion of electrons. In both cases an enhanced version of the new electron-electron interaction, introduced in chapter 5, was used to reduce the effect of using a finite sized supercell to calculate excitation energies.

The results indicate that the VMC technique is not sufficiently accurate to resolve the $1/N$ changes in the energy when an electron is added(removed) from the system. However, within DMC the results are extremely encouraging. Using DMC we have calculated excitation energies in diamond-structure silicon using both of the above techniques. In both cases the results show a significant improvement over results obtained using the LDA.

As QMC is a non-perturbative many-body technique (in contrast to, for instance, the GW approximation) it is hoped that these preliminary calculations on the weakly correlated silicon system can be extended to more strongly correlated systems. A challenging group of systems to study would be the Mott insulating 3d monoxides such as MnO, FeO, CoO, and NiO.

Appendix A

Updating the Slater Determinant

Each time a single electron move is accepted, the value of the Slater determinant changes. Recalculating the determinant after each move is potentially an extremely costly procedure. Fortunately, there is a method for updating the wavefunction[115, 26] which is much less costly than recalculating the whole determinant every time an electron is moved. The procedure is described here for the specific case of moving a single electron, with spin $\uparrow$, from position $\mathbf{r}_{i,old}$ to position $\mathbf{r}_{i,new}$. The method is the same for all electrons and spins.

The quantity that is calculated is the ratio of the new and old determinants. From this ratios, the probability of acceptance of the move can be calculated. The cost of updating the quantities involved in the determinant, if the move is accepted, is greatly reduced by making use of this ratio.

The determinant $D^{\downarrow}(\mathbf{R})$ only involves spin $\downarrow$ electrons and therefore is not changed by the move. $D^{\uparrow}(\mathbf{R})$ is changed. The matrix $\underline{\underline{D}}^{\uparrow}$ has elements

$$D^{\uparrow}_{ji} = \phi_j(\mathbf{r}^{\uparrow}_i) \quad , \tag{A.1}$$

where ϕ_j are the set of N single-particle wavefunctions making up the Slater determinant. When the electron is moved it is only the ith column of the matrix which changes.

One way of calculating the determinant of any matrix, $\underline{\underline{A}}$, is to use the relation

$$det(\underline{\underline{A}})\underline{\underline{I}} = \underline{\underline{A}}^T \underline{\underline{A}}^C \tag{A.2}$$

where $\underline{\underline{I}}$ is the unit matrix, $\underline{\underline{A}}^T$ is the transpose of matrix $\underline{\underline{A}}$ and $\underline{\underline{A}}^C$ is the matrix of cofactors of matrix $\underline{\underline{A}}$. The determinant can be calculated by choosing any column, i, of the matrix and summing

$$det(\underline{\underline{A}}) = \sum_{j=1}^{N} \underline{\underline{A}}^T_{ij} \underline{\underline{A}}^C_{ji} \quad . \tag{A.3}$$

For the particular case of moving electron i, it is only the i^{th} column of $\underline{\underline{D}}^{\uparrow}$ which is altered and this does not change any elements of the i^{th} column of the matrix $\underline{\underline{D}}^C$. The determinant after the move has been made, $\underline{\underline{D}}^{\uparrow}_{new}$, can be written as

$$\underline{\underline{D}}^{\uparrow}_{new} = \sum_{j=1}^{N} \underline{\underline{D}}^{\uparrow\, T}_{ij,new} \underline{\underline{D}}^{\uparrow\, C}_{ji,new} \quad . \tag{A.4}$$

If the transpose of the inverse matrix is denoted by $(\underline{\underline{D}}^{\uparrow\, T})^{-1}$ then, from the definition of the inverse of a matrix,

$$\overline{\underline{\underline{D}}}^{\uparrow} = (\underline{\underline{D}}^{\uparrow\, T})^{-1} = \frac{1}{D^{\uparrow}} \underline{\underline{D}}^{\uparrow\, C} \quad , \tag{A.5}$$

so

$$D^{\uparrow\, C}_{ji,old} = D^{\uparrow}_{old} \overline{D}^{\uparrow}_{ji,old} \quad . \tag{A.6}$$

The new determinant, $D^{\uparrow}_{new}$, is written as

$$D^{\uparrow}_{new} = D^{\uparrow}_{old} \sum_{j=1}^{N} D^{\uparrow}_{ji,new} \overline{D}^{\uparrow}_{ji,old} \tag{A.7}$$

and the calculation of the ratio $q_i = \sum j = 1^N D^{\uparrow}_{new} / D^{\uparrow}_{old}$ is simply

$$q_i = \sum j = 1^N D^{\uparrow}_{ji,new} \overline{D}^{\uparrow}_{ji,old} \quad . \tag{A.8}$$

This ratio, q_i, is the quantity which is required when calculating the probability of the move being accepted.

If the move of electron i, with spin s, is accepted then the matrix $\overline{\underline{\underline{D}}}^s$ is updated in the following manner

$$\overline{D}^s_{jk,new} = \begin{cases} \frac{\overline{D}^s_{jk,old}}{q^s_i}, & \text{if } k = i \\ \overline{D}^s_{jk,old} - \frac{\overline{D}^s_{jk,old}}{q^s_i}\left[\sum_{l=1}^{N} \overline{D}^s_{lk,old}\phi_l(\mathbf{r}^s_{i,new})\right], & \text{if } k \neq i \end{cases} \quad (A.9)$$

The case where $k = i$ comes simply from the definition of $\overline{\underline{\underline{D}}}^\uparrow$ in Eq.(A.5). Finding the updated $\overline{\underline{\underline{D}}}^\uparrow$ matrix elements when $k \neq i$ is more complicated.

The new determinant matrix, $\underline{\underline{D}}^s_{new}$, is written as the old determinant matrix, $\underline{\underline{D}}^s_{old}$, plus a perturbation matrix, Δ. Therefore

$$D^s_{jk,new} = D^s_{jk,old} + \Delta_{jk} \quad (A.10)$$

where

$$\Delta_{jk} = \delta_{ik}\left(\phi_j(\mathbf{r}^s_{i,new}) - \phi_j(\mathbf{r}^s_{i,old})\right) = \delta_{ik}(\Delta\phi)_j \quad . \quad (A.11)$$

The following matrix identity is used to find the inverse of a sum of two matrices,

$$\left(\underline{\underline{A}} + \underline{\underline{B}}\right)^{-1} = \underline{\underline{A}}^{-1} - \left(\underline{\underline{A}} + \underline{\underline{B}}\right)^{-1} \underline{\underline{BA}}^{-1} \quad . \quad (A.12)$$

[This identity can be proved by multiplying on the left by (A +B).] Hence,

$$\begin{aligned} \overline{\underline{\underline{D}}}^s_{new} &= \left(\underline{\underline{D}}^{s\,T}_{old} + \underline{\underline{\Delta}}^T\right)^{-1} \\ &= \left(\underline{\underline{D}}^{s\,T}_{old}\right)^{-1} - \left(\underline{\underline{D}}^{s\,T}_{old} + \underline{\underline{\Delta}}^T\right)^{-1} \underline{\underline{\Delta}}^T \left(\underline{\underline{D}}^{s\,T}_{old}\right)^{-1} \\ &= \overline{\underline{\underline{D}}}^s_{old} - \overline{\underline{\underline{D}}}^s_{new}\underline{\underline{\Delta}}^T\overline{\underline{\underline{D}}}^s_{old} \quad , \end{aligned} \quad (A.13)$$

which can be written in terms of matrix elements as

$$\begin{aligned} \overline{D}^s_{jk,new} &= \overline{D}^s_{jk,old} - \sum_l \sum_m \overline{D}^s_{jm,new}\Delta^T_{ml}\overline{D}^s_{lk,old} \\ &= \overline{D}^s_{jk,old} - \sum_l \sum_m \overline{D}^s_{jm,new}\Delta_{lm}\overline{D}^s_{lk,old} \\ &= \overline{D}^s_{jk,old} - \sum_l \sum_m \overline{D}^s_{jm,new}\delta_{im}(\Delta\phi)_l\overline{D}^s_{lk,old} \\ &= \overline{D}^s_{jk,old} - \overline{D}^s_{ji,new}\sum_l(\Delta\phi)_l\overline{D}^s_{lk,old} \quad . \end{aligned} \quad (A.14)$$

Using the definition of the matrix elements of $\overline{D}^s_{new}$ for the case of $k = i$, as defined in Eq.(A.9) and expanding out $(\Delta\phi)_l$, the equation

$$\overline{D}^s_{jk,new} = \overline{D}^s_{jk,old} - \frac{\overline{D}^s_{ji,old}}{q^s_i} \sum_{l=1}^{N} \overline{D}^s_{lk,old}\phi_l(\mathbf{r}^s_{i,new}) + \frac{\overline{D}^s_{ji,old}}{q^s_i} \sum_{l=1}^{N} \overline{D}^s_{lk,old}\phi_l(\mathbf{r}^s_{i,old}) \qquad \text{(A.15)}$$

is obtained. The last term in the above equation $\sum_{l=1}^{N} \overline{D}^s_{lk,old}\phi_l(\mathbf{r}^s_{i,old})$ is simply the matrix multiplication $(D^s)^{-1}D^s$ which gives δ_{ik} . However in this case $k \neq i$ so the last term is zero. The update equation reduces to

$$\overline{D}^s_{jk,new} = \overline{D}^s_{jk,old} - \frac{\overline{D}^s_{jk,old}}{q^s_i} \sum_{l=1}^{N} \overline{D}^s_{lk,old}\phi_l(\mathbf{r}^s_{i,new}) \quad , \qquad \text{(A.16)}$$

as required.

Bibliography

[1] A. J. Williamson, S. D. Kenny, G. Rajagopal, A. J. James, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, and P. Maccallum. *Phys. Rev. B*, **53**, 9640, 1996

[2] S. D. Kenny, G. Rajagopal, R. J. Needs, W. K. Leung, M. J. Godfrey, A. J. Williamson, and W. M. C. Foulkes. *Phys. Rev. Lett.*, **77**, 1099, 1996

[3] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson. *Phys. Rev. B*, **53**, 1814, 1996

[4] N. Ashcroft and N. Mermin. *Solid State Physics*. Holt-Saunders, 1976

[5] J. C. Slater. *Phys. Rev.*, **81**, 385, 1951

[6] J. C. Slater. *Quantum Theory of Matter, 2nd edition*. McGraw-Hill, 1964

[7] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, 1989

[8] J. Callaway and N. March. *Density Functional Methods: Theory and Applications*, volume 38. Solid State Physics, 1984

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

[10] . P. Hohenberg and W. Kohn. *Inhomogeneous Electron Gas*. *Phys. Rev.*, **136**, B864, 1964

[11] M. Levy. *Proc. Nat. Acad. Sci.*, **76**, 6062, 1979

[12] J. E. Harriman. *Phys. Rev. A*, **24**, 680, 1981

[13] L. M. Sachs. *Phys. Rev.*, **117**, 1504, 1960

[14] D. A. Goodings. *Phys. Rev.*, **123**, 1706, 1961

[15] L. Hedin and S. Lundquist. *Solid State Physics*, volume 23. Academic Press, New York, 1969

[16] P. Davies. *Quantum Mechanics*. Chapman and Hall, 1994

[17] P. Atkins. *Molecular Quantum Mechanics*. Oxford University Press, 1983

[18] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes - The Art of Scientific Computing*. Cambridge University Press, 1986

[19] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. *J. Chem. Phys.*, **21**, 1087, 1953

[20] W. L. McMillan. *Phys. Rev.*, **138**, A442, 1965

[21] R. Jastrow. *Phys. Rev.*, **98**, 1479, 1955

[22] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins. *Phys. Rev. Lett.*, **60**, 1719, 1988

[23] B. L. Hammond, W. A. Lester, and P. J. Reynolds. World Scientific, Singapore, 1994

[24] T. Kato. *Comm. Pure Appl. Math.*, **10**, 151, 1957

[25] D. Ceperley. *Phys. Rev. B*, **18**, 3126, 1978

[26] S. Fahy, X. W. Wang, and S. G. Louie. *Phys. Rev. B*, **42**, 3503, 1990

[27] L. Mitas and R. M. Martin. *Phys. Rev. Lett.*, **72**, 2438, 1994

[28] M. H. Kalos, D. Levesque, and L. Verlet. *Phys. Rev. A*, **9**, 2178, 1974

[29] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester Jr. *Monte Carlo Methods in ab initio Quantum Chemistry*. *J. Chem. Phys.*, **77**, 5593, 1982

[30] C. J. Umrigar, M. P. Nightingale, and K. J. Runge. *J. Chem. Phys.*, **99**, 2865, 1993

[31] D. Ceperley and B. Alder. *J. Chem. Phys.*, **81**, 5833, 1984

[32] D. Ceperley. *J. Stat. Phys*, **63**, 1237, 1991

[33] G. Rajagopal, R. J. Needs, A. James, S. D. Kenny, and W. M. C. Foulkes. *Phys. Rev. B*, **51**, 10591, 1995

[34] J. Vrbik and S. M. Rothstein. *J. Chem. Phys.*, **63**, 130, 1986

[35] J. Anderson. *J. Chem. Phys.*, **65**, 4121, 1976

[36] . Y. Kwon, D. M. Ceperley, and R. M. Martin. *Phys. Rev. B*, **48**, 12037, 1993

[37] R. P. Feynman and M. Cohen. *Phys. Rev.*, **102**, 1189, 1956

[38] M. M. Hurley and P. A. Christiansen. *J. Chem. Phys.*, **86**, 1069, 1987

[39] P. Christiansen. *J. Chem. Phys.*, **88**, 4867, 1988

[40] P. Christiansen. *J. Chem. Phys.*, **94**, 7865, 1990

[41] L. Mitas. *Computer Simulation Studies in Condensed Matter Physics*. Springer Verlag, 1992

[42] L. Mitas, E. L. Shirley, and D. M. Ceperley. *J. Chem. Phys.*, **95**, 3467, 1991

[43] D. Ceperley and B. Alder. *Phys. Rev. Lett.*, **45**, 566, 1980

[44] X. P. Li, R. J. Needs, R. M. Martin, and D. M. Ceperley. *Phys. Rev. B*, **45**, 6124, 1992

[45] P. Acioli and D. Ceperley. *unpublished*

[46] D. Ceperley and B. Alder. *Phys. Rev. B*, **36**, 2092, 1987

[47] V. Natoli, R. M. Martin, and D. M. Ceperley. *Phys. Rev. Lett.*, **70**, 1952, 1993

[48] S. Fahy, X. W. Wang, and S. G. Louie. *Phys. Rev. Lett.*, **61**, 1631, 1988

[49] X. P. Li, D. M. Ceperley, and R. M. Martin. *Phys. Rev. B*, **44**, 10929, 1991

[50] G. Rajagopal, R. J. Needs, S. D. Kenny, W. M. C. Foulkes, and A. James. *Phys. Rev. Lett.*, **73**, 1959, 1994

[51] L. Mitas and R. M. Martin. *Phys. Rev. Lett.*, **72**, 2438, 1994

[52] G. K. Wannier. *Rev. Mod. Phys.*, **64**, 1045, 1992

[53] A. Baldereschi. *Phys. Rev. B*, **7**, 5212, 1973

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

[55] M. Allen and D. Tildesley. *Computer Simulation of Liquids*. Oxford, 1987

[56] M. P. Tosi. *Solid State Physics*, volume 16. Springer, 1964

[57] C. Kittel. *Introduction to Solid State Physics*. Wiley, 1986

[58] T. H. Starkloff and J. D. Joannopoulos. *Phys. Rev. B*, **16**, 5212, 1977

[59] G. P. Kerker. *J. Phys. C*, **13**, L189, 1980

[60] G. Bachelet, D. Hamann, and M. Schlüter. *Phys. Rev. B*, **26**, 4199, 1982

[61] R. Coldwell. *Int. J. Quant. Chem. Symp.*, **11**, 215, 1977

[62] C. Filippi and C. J. Umrigar. *J. Chem. Phys.*, **105**, 213, 1996

[63] L. Mitas and R. M. Martin. *to be published*

[64] S. Tanaka. *J. Phys. Soc. Japan*, **62**, 2112, 1993

[65] G. Jacucci and A. Rahman. *Nuovo Cimento*, **D4**, 341, 1984

[66] Z. W. Lu, A. Zunger, and M. Deutsch. *Phys. Rev. B*, **47**, 9385, 1993

[67] A. Malatesta, S. Fahy, and G. B. Bachelet. *submitted to Phys. Rev. B*

[68] G. Ortiz and P. Ballone. *Europhys. Lett.*, **23**, 7, 1993

[69] G. Ortiz and P. Ballone. *Phys. Rev. B*, **50**, 1391, 1994

[70] A. Baldereschi. *Phys. Rev. B*, **7**, 5212, 1973

[71] F. Gygi and A. Baldereschi. *Phys. Rev. B*, **34**, 4405, 1986

[72] W. v. d. Linden and P. Horsch. *Phys. Rev. B*, **37**, 8351, 1988

[73] W. v. d. Linden, P. Fulde, and K. P. Bohnen. *Phys. Rev. B*, **34**, 1063, 1986

[74] R. Hood, M.-Y. Chou, A. J. Williamson, G. Rajagopal, R. J. Needs, and W. M. C. Foulkes. *Submitted to Phys. Rev. Lett.*

[75] G. K. Wannier. *Phys. Rev. B*, **48**, 14859, 1993

[76] L. Mitas. *Electronic Properties of Solids Using Cluster Methods*. Plenum Press, 1994

[77] W. Knorr and R. W. Godby. *Phys. Rev. B*, **50**, 1779, 1994

[78] S. Fahy, X. W. Wang, and S. G. Louie. *Phys. Rev. Lett.*, **65**, 1478, 1990

[79] G. Engel, Y. Kwon, and R. Martin. *Phys. Rev. B*, **51**, 13538, 1995

[80] M. Caffarel and D. Ceperley. *J. Chem Phys.*, **97**, 8415, 1992

[81] D. Ceperley and B. Bernu. *J. Chem. Phys.*, **89**, 6316, 1988

[82] B. Bernu, D. Ceperley, and W. lester. *J. Chem. Phys.*, **93**, 552, 1990

[83] T. Koopmans. *Physica*, **1**, 104, 1934

[84] P. Bagus. *Phys. Rev.*, **139**, A619, 1965

[85] R. Dreizler and E. Gross. *Density Functional Theory*. Springer, 1990

[86] J. P. Perdew and M. Levy. *Phys. Rev. Lett.*, **51**, 1884, 1983

[87] L. J. Sham and M. Schlüter. *Phys. Rev. Lett.*, **51**, 1888, 1983

[88] R. W. Godby, M. Schlüter, and L. J. Sham. *Phys. Rev. Lett.*, **56**, 2415, 1986

[89] R. W. Godby, M. Schlüter, and L. J. Sham. *Phys. Rev. B*, **35**, 4170, 1987

[90] W. v. d. Linden, P. Horsch, and W. D. Lukas. *Solid State Commun.*, **59**, 485, 1986

[91] K. Schönhammer and O. Gunnarsson. *J. Phys. C*, **20**, 3675, 1987

[92] O. Gunnarsson and K. Schönhammer. *Phys. Rev. Lett.*, **56**, 1968, 1986

[93] W. Kohn. *Phys. Rev. B*, **33**, 4331, 1986

[94] J. Chelikowsky and M. Cohen. *Electronic Structure and Optical properties of Semiconductors*. Springer-Verlag, 1988

[95] D. Aspnes. *Advances in Solid State Physics*, volume 17. Pergamon, Oxford, 1976

[96] D. Aspnes. *Nuovo Cimento*, **39**, 337, 1977

[97] F. Bassani and G. Pastori. *Electronic States and Optical Transition in Solids*. Pergamon, Oxford, 1975

[98] D. L. Greenaway and K. Harbeke. *Optical Properties and Band Structures of Semiconductors*. Pergamon, Oxford, 1968

[99] M. Cardona. *Solid State Physics*. Suppl., 1969

[100] J. Tauc. *Optical Properties of Solids*. Academic, New York, 1966

[101] N. V. Smith. *Photoelectron Spectroscopy of Solids*. Springer, 1977

[102] . L. Ley, S. P. Kowalczyk, R. A. Pollak, and D. A. Shirley. *Phys. Rev. Lett.*, **29**, 1088, 1972

[103] W. D. Grobman and D. E. Eastman. *Phys. Rev. Lett.*, **29**, 1508, 1972

[104] W. E. Spicer and R. C. Eden. *Proc. of the 9th Int. Conf. on the Phys. of Semicon.*, page 61, 1968

[105] A. De Vita, M. Gillan, J. Lin, M. Payne, I. Stich, and L. Clarke. *Phys. Rev. B*, **46**, 12964, 1992

[106] A. De Vita, M. Gillan, J. Lin, M. Payne, I. Stich, and L. Clarke. *Phys. Rev. Lett.*, **68**, 3319, 1992

[107] R. Shah, A. De Vita, and M. C. Payne. *J. Phys. : Condensed Matter*, **7**, 6981–6992, 1995

[108] R. Shah, A. De Vita, V. Heine, and M. C. Payne. *Phys. Rev. B*, **53**, 8257–8261, 1996

[109] M. Leslie and M. J. Gillan. *J. Phys. C*, **18**, 973, 1985

[110] A. J. Williamson, G. Rajagopal, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, Y. Wang, and M. Y. Chou. *Submitted to Phys. Rev. Lett.*

[111] A. J. Williamson, R. J. Needs, and G. Rajagopal. *in preparation*

[112] P. Fulde. *Electron Correlations in Molecules and Solids*. Springer Verlag, 1995

[113] N. F. Mott. *Trans. Faraday Society*, **34**, 822, 1938

[114] G. K. Wannier. *Phys. Rev.*, **52**, 191, 1937

[115] D. Ceperley, G. Chester, and M. Kalos. *Phys. Rev. B*, **16**, 3081, 1971