Variational Monte Carlo studies of Atoms

by

HÅVARD SANDSDALEN

THESIS

for the degree of

MASTER OF SCIENCE

(Master in Computational Physics)

Faculty of Mathematics and Natural Sciences
Department of Physics
University of Oslo

June 2010

Det matematisk-naturvitenskapelige fakultet
Universitetet i Oslo

Contents

1 Introduction **7**

I Theory **9**

2 Quantum Physics **11**
2.1 Quantum Mechanics in one dimension 11
2.1.1 Probability and statistics 11
2.1.2 The time-independent Schrödinger Equation 12
2.2 Bra-Ket notation 15
2.3 Quantum Mechanics in three dimensions 15
2.3.1 Separation of variables - quantum numbers l and m 16
2.3.2 The Angular equation 16
2.3.3 The Radial equation and solution for the hydrogen atom 18

3 Many-particle Systems **21**
3.1 Atoms 21
3.1.1 Two-particle systems 21
3.1.2 Wave functions for N-particle atoms 23
3.1.3 Electron configuration 24
3.1.4 Hamiltonian and scaling 25
3.2 Hartree-Fock Theory 27
3.2.1 Hartree-Fock equations 29

4 Quantum Monte Carlo **35**
4.1 Markov chains 36
4.2 Random numbers 37
4.3 The Variational Principle 38
4.4 Variational Monte Carlo 39
4.4.1 VMC and the simple Metropolis algorithm 39
4.4.2 Metropolis-Hastings algorithm and importance sampling 41

5 Wave Functions **47**
5.1 Cusp conditions and the Jastrow factor 47
5.1.1 Single particle cusp conditions 47
5.1.2 Correlation cusp conditions 49
5.2 Rewriting the Slater determinant 49
5.3 Variational Monte Carlo wave function 50
5.4 Orbitals for VMC 51

5.4.1 S-orbitals . . . 51
5.4.2 P-orbitals . . . 52
5.5 Roothaan-Hartree-Fock with Slater-type orbitals . . . 55
5.5.1 Derivatives of Slater type orbitals . . . 57

II Implementation and results **59**

6 Implementation of the Variational Monte Carlo Method **61**
6.1 Optimizing the calculations . . . 62
6.1.1 Optimizing the ratio - $\Psi_T(r^{new})/\Psi_T(r^{old})$ . . . 63
6.1.2 Derivative ratios . . . 67
6.2 Implementation of Metropolis-Hastings algorithm . . . 73
6.3 Blocking . . . 74
6.4 Time step extrapolation . . . 76
6.5 Parallel computing . . . 76

7 Hartree-Fock implementation **77**
7.1 Interaction matrix elements . . . 77
7.2 Algorithm and HF results . . . 78

8 VMC Results **81**
8.1 Validation runs . . . 81
8.1.1 Hydrogen-like orbitals . . . 81
8.1.2 Slater-type orbitals - Hartree Fock results . . . 82
8.2 Variational plots . . . 83
8.2.1 Hydrogen-like orbitals . . . 83
8.2.2 Slater-type orbitals - Hartree Fock results . . . 87
8.3 Optimal parameters with DFP . . . 88
8.3.1 Hydrogen-like orbitals . . . 90
8.3.2 Slater-type orbitals - Hartree Fock results . . . 91
8.4 Time step analysis - extrapolated results . . . 91
8.4.1 Hydrogenic wave functions . . . 91
8.4.2 Slater-type orbitals . . . 91
8.5 Discussions . . . 96

9 Conclusion **99**

A Roothaan-Hartree-Fock results **101**

B Statistics **105**

C DFP and energy minimization **111**

Bibliography **114**

Acknowledgements

I would like to thank my advisor Morten Hjorth-Jensen for always being available and full of helpful adivce. Also, I would like to thank my trusted fellow students Magnus Pedersen Lohne, Sigurd Wenner, Lars Eivind Lervåg and Dwarkanath Pramanik for being positive and inspirational through the rough patches.

I would not have able to complete this work without the support of my family and friends. Thank you!

Håvard Sandsdalen

Chapter 1

Introduction

The aim of this thesis is to study the binding energy of atoms. Atoms of interest in this work are helium, beryllium, neon, magnesium and silicon. The smallest of all atoms, the hydrogen atom, consists of a nucleus of charge e and one electron with charge $-e$. The size of this atom is approximately the famous Bohr radius, $a_0 \approx 5 \cdot 10^{-11}$ (see [1]). To be able to describe such small systems, we must utilize the concepts of quantum physics.

Many-particle systems such as atoms, cannot be solved exactly using analytic methods, and must be solved numerically. The Variational Monte Carlo (VMC) method is a technique for simulating physical systems numerically, and is used to perform *ab initio* calculations on the system. The term "ab initio" means that the method is based on first principle calculations with strictly controlled approximations being made (e.g. the Born-Oppenheimer approximation, see section 3.1.4). For our case this means solving the *Schrödinger equation* (see e.g. [2]). Our main goal will be to use VMC to solve the time-independent Schrödinger equation in order to calculate the energy of an atomic system. We will study the *ground state* of the atom, which is the state corresponding to the lowest energy.

Variational Monte Carlo calculations have been performed on atoms on several occasions (see ref. [3]), but for the atoms magnesium and silicon this is not so common. In this work we will perform VMC calculations on both well-explored systems such as helium, beryllium and neon in addition to the less examined magnesium and silicon atoms. The helium, beryllium and neon calculations from [3] will serve as benchmark calculations for our VMC machinery, while for silicon and magnesium we have compared with results from [4].

Chapter 2 in this thesis will cover the quantum mechanics we need in order to implement a VMC calculation on atoms numerically. We start off by introducing the basic concepts of single-particle quantum mechanical systems, and move on to describe the quantum mechanics of many-particle systems in chapter 3. Furthermore, we will introduce the Variational Monte Carlo method in chapter 4, and a detailed discussion of how to approximate the ground state of the system by using Slater determinants is included in chapter 5.

A large part of this thesis will also be devoted to describing the implementation of the VMC machinery and the implementation of a simple Hartree-Fock method. This will be discussed in chapters 6 and 7. We will develop a code in the C++ programming language (see ref. [5]) that is flexible with respect to the size of the system and several other parameters introduced throughout the discussions.

Chapter 8 and 9 will contain the results produced by the VMC program we have

developed, as well as a discussion and analysis of the results we have obtained.

Part I

Theory

Chapter 2

Quantum Physics

In the 1920's European physicists developed quantum mechanics in order to describe the physical phenomena they had been discovering for some years. Such phenomena as the photoelectric effect, compton scattering, x-rays, blackbody radiation and the diffraction patterns (see ref. [1]) from the double-slit experiment indicated that physicists needed a new set of tools when handling systems on a very small scale, e.g. the behavior of single particles and isolated atoms.

This chapter will give an introduction to the relevant topics in quantum physics needed to describe the atomic systems in this project. Some parts will closely follow the discussions in the books [2] and [6], while other parts only contain elements from the sources listed in the bibliography.

2.1 Quantum Mechanics in one dimension

The general idea and goal of quantum mechanics is to solve the complex, time-dependent Schrödinger-equation (S.E.) for a specific physical system which cannot be described by classical mechanics. Once solved, the S.E. will give you the quantum mechanical wave function, Ψ, a mathematical function which contains all information needed about such a non-classical system. It introduces probabilities and statistical concepts which contradict with the deterministic properties of systems described by classical mechanics. The full one-dimensional S.E. for an arbitrary potential, V, reads:

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V\Psi. \tag{2.1}$$

The Schrödinger equation is the classical analogy to Newton's second law in classical physics, and describes the dynamics of virtually any physical system, but is only useable in small scale systems, quantum systems.

2.1.1 Probability and statistics

The wave function, Ψ, is now a function of both position, x, and time, t. Quantum mechanics uses the concept of probability and statistics via Ψ, and these solutions of the S.E. may be complex as the equation Eq. (2.1) is complex itself. To comply with the statistical interpretation we must have a function that is both real and non-negative. As described in [2], Born's statistical interpretation takes care of this problem by introducing the complex conjugate of the wave function. The product $\Psi^*\Psi = |\Psi|^2$ is interpreted as

the probability density for the system state. If the system consists of only one particle, the integral

$$\int_a^b |\Psi(x,t)|^2 \, \mathrm{d}x, \tag{2.2}$$

is then interpreted as the probability of finding the particle between positions a and b at an instance t. To further have a correct correspondence with probability, we need the total probability of finding the particle anywhere in the universe to be one. That is

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, \mathrm{d}x = 1. \tag{2.3}$$

In quantum mechanics, operators represent the observables we wish to find, given a wave function, Ψ. The operator representing the position variable, $\widehat{\mathbf{x}}$, is just x itself, while the momentum operator is $\widehat{\mathbf{p}} = -i\hbar(\partial/\partial x)$. All classical dynamical variables are expressed in terms of just momentum and position (see [2]). Another important operator is the Hamilton operator. The Hamilton operator gives the time evolution of the system and is the sum of the kinetic energy operator $\widehat{\mathbf{T}}$ and the potential energy operator $\widehat{\mathbf{V}}$, $\widehat{\mathbf{H}} = \widehat{\mathbf{T}} + \widehat{\mathbf{V}}$. The operator $\widehat{\mathbf{V}}$ is represented by the function V from Eq. (2.1), while the kinetic energy operator is

$$\widehat{\mathbf{T}} = \frac{\widehat{\mathbf{p}}^2}{2m} = \frac{(i\hbar)^2}{2m}\left(\frac{\partial}{\partial x}\right)^2 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}. \tag{2.4}$$

For an arbitrary operator, $\widehat{\mathbf{Q}}$, the expectation value, $\langle Q \rangle$, is found by the formula

$$\langle Q \rangle = \int \Psi^* \widehat{\mathbf{Q}} \Psi \mathrm{d}x. \tag{2.5}$$

The same goes for expectation values of higher moments, e.g. $\langle p^2 \rangle$:

$$\langle p^2 \rangle = \int \Psi^* \widehat{\mathbf{p}}^2 \Psi \mathrm{d}x. \tag{2.6}$$

The so-called variance of an operator or observable, σ_Q^2, can be calculated by the following formula:

$$\sigma_Q^2 = \langle Q^2 \rangle - \langle Q \rangle^2. \tag{2.7}$$

This quantity determines the standard deviation, $\sigma_Q = \sqrt{\sigma_Q^2}$. The standard deviation describes the spread around the expectation value. The smaller the standard deviation, the smaller the variation between the possible values of Q.

The wave functions, Ψ, exist in the so-called *Hilbert space*, a mathematical vector space of square-integrable functions.

2.1.2 The time-independent Schrödinger Equation

However, in many cases, we are only interested in a time-independent version of Eq. (2.1). A crucial point is then to demand the potential, V, to be a time-independent potential as well, viz. $V = V(x)$ (in one spatial dimension). This equation can be obtained by the well-known method of **separation of variables**. The trick is to write our wave function, $\Psi(x,t)$, as a product of a purely spatial function, $\psi(x)$, and another function, $\phi(t)$, depending only of time. That is, we assume: $\Psi(x,t) = \psi(x)\phi(t)$. By inserting

Ψ into the full S.E., Eq. (2.1), remembering that ψ and ϕ only depend on one variable each, then dividing by $\Psi = \psi\phi$, we get

$$i\hbar\frac{1}{\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V(x). \tag{2.8}$$

By inspecting Eq. (2.8), we see that the left side of the equation depends on t alone, while the right side depends only on x. This means that both sides must equal a constant. By varying t and thereby changing the left side, the right side would change without varying x. We call this separation constant E, giving us the two equations

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\frac{iE}{\hbar}\phi, \tag{2.9}$$

and

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V\psi = E\psi. \tag{2.10}$$

Equation (2.9) can be solved quite easily, and results in an exponential form for the time-dependent part,

$$\phi(t) = e^{-iEt/\hbar}. \tag{2.11}$$

The second equation, Eq. (2.10), is called the **time-independent Schrödinger equation**. By inspecting Eq. (2.11), we see that all expectation values will be constant in time because the time-dependent part from Ψ, $\phi(t)$, will only give a factor 1 when it is multiplied with its complex conjugate from Ψ^*. That is:

$$e^{-iEt/\hbar}e^{+iEt/\hbar} = 1. \tag{2.12}$$

The expectation values depend solely on the spatial parts, ψ. We call these separable solutions **stationary states**.

Stationary states and expectation values

Another point about the stationary solutions is the close relation with classical mechanics. The Hamilton function determines the total energy of the system, and is the sum of the kinetic and potential energy. The classical Hamiltonian function for any system with a time-independent potential is

$$H(x,p) = \frac{p^2}{2m} + V(x). \tag{2.13}$$

By using the canonical substitution $p \rightarrow (\hbar/i)(\mathrm{d}/\mathrm{d}x)$ for the quantum mechanical momentum operator, we get

$$\widehat{\mathbf{H}} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x). \tag{2.14}$$

This is identical to the time-independent Schrödinger equation, Eq. (2.10), and we can then obtain a much simplified Schrödinger equation:

$$\widehat{\mathbf{H}}\psi = E\psi. \tag{2.15}$$

The expectation value of the Hamilton operator, the total energy, is now given as

$$\langle H\rangle = \int \psi^*\widehat{\mathbf{H}}\psi\mathrm{d}x = E\int |\psi|^2\,\mathrm{d}x = E\int |\Psi|^2\,\mathrm{d}x = E. \tag{2.16}$$

Calculating $\langle H^2 \rangle$ gives

$$\langle H^2 \rangle = \int \psi^* \widehat{\mathbf{H}}^2 \psi \mathrm{d}x = E \int \psi^* \widehat{\mathbf{H}} \psi \mathrm{d}x = E^2 \int |\psi|^2 \, \mathrm{d}x = E^2 \int |\Psi|^2 \, \mathrm{d}x = E^2. \quad (2.17)$$

This means the variance and standard deviation of $\widehat{\mathbf{H}}$ are both zero

$$\sigma_H^2 = \langle H^2 \rangle - \langle H \rangle^2 = E^2 - E^2 = 0. \quad (2.18)$$

For these separable solutions, every measurement of the total energy will return the same value, E. Thus the spread around the expectation value is exactly zero.

General solutions

The general solution for these systems is a linear combination of different separable solutions. Different solutions with different separation constants, e.g.

$$\Psi_1(x,t) = \psi_1(x) e^{-iE_1 t/\hbar}, \quad \Psi_2(x,t) = \psi_2(x) e^{-iE_2 t/\hbar}, \quad (2.19)$$

which both are solutions of Eq. (2.10), can be used to construct the general solution

$$\Psi(x,t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_n t/\hbar} = \sum_{n=1}^{\infty} c_n \Psi_n(x,t), \quad (2.20)$$

where the factors c_n are probability weights for its corresponding stationary state.

Spin

An important, but difficult concept in quantum physics is spin. Spin is an intrinsic property of every elementary particle. Also composite systems will have a certain value of spin when imposing addition rules on the single particles that make up such a system. In this project we will deal with fermions, i.e. half integer spin particles (see section 3.1.1).

A particle will either have *spin up* or *spin down*. This is denoted by the spin states

$$\chi_+ = \uparrow, \quad (2.21)$$

and

$$\chi_- = \downarrow . \quad (2.22)$$

These spin states, $\chi_\pm$, are mutually orthogonal, but exist in another Hilbert space than the spatial wave functions, Ψ, and will not interfere with the integration $\int dx$.

This presentation of spin is very short due to the fact that we don't need much information about spin in this thesis. As shown later, the Hamiltonian will not depend on spin values, so the spin states, $\chi_\pm$, will only be used as a label to indicate which states are occupied or not. This will become more apparent in the discussion of many-particle systems.

2.2 Bra-Ket notation

As we will be dealing a lot with expectation values, which are integrals, it can be smart to introduce a compact way to describe the wave functions and the integrals. The so-called **ket**, $|\Psi\rangle$, represents our Ψ, while the **bra**, $\langle\Psi|$, represents the complex conjugate, Ψ^*. The expectation value

$$\langle\Psi|\,\widehat{\mathbf{H}}\,|\Psi\rangle\,, \tag{2.23}$$

is defined as

$$\langle\Psi|\,\widehat{\mathbf{H}}\,|\Psi\rangle = \int \Psi^*\widehat{\mathbf{Q}}\Psi \mathrm{d}x, \tag{2.24}$$

where x now represents all spatial dimensions and quantum numbers. Now we can write the expectation values of $\widehat{\mathbf{H}}$ as

$$\langle H\rangle = \langle\Psi|\,\widehat{\mathbf{H}}\,|\Psi\rangle\,. \tag{2.25}$$

2.3 Quantum Mechanics in three dimensions

While discussing quantum mechanics in one dimension is useful for getting in the basics, most real life systems occur in three dimensions. This section closely follows the discussion in [2].

In three dimensions the one-dimensional Hamiltonian, $\widehat{\mathbf{H}}(x, p_x) = p_x^2/2m_e + V(x)$ is replaced by

$$\widehat{\mathbf{H}}(x, y, z, p_x, p_y, p_z) = \frac{1}{2m_e}\left(p_x^2 + p_y^2 + p_z^2\right) + V(x, y, z). \tag{2.26}$$

with m_e being the electron mass. For quantum mechanical systems the momentum operators are substituted by

$$p_x \to \frac{\hbar}{i}\frac{\partial}{\partial x},\;\; p_y \to \frac{\hbar}{i}\frac{\partial}{\partial y},\;\; p_z \to \frac{\hbar}{i}\frac{\partial}{\partial z}. \tag{2.27}$$

We can write this in a more compact way on vector form as

$$\mathbf{p} = \frac{\hbar}{i}\nabla. \tag{2.28}$$

Introducing the **Laplacian** in Cartesian coordinates, $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$, we can write the full Hamiltonian as

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m_e}\nabla^2\Psi + V\Psi. \tag{2.29}$$

The normalization integral in three dimensions changes using the infinitesimal volume element $d^3\mathbf{r} = dx\,dy\,dz$. We now have

$$\int |\Psi(x, y, z, t)|^2\, dx\, dy\, dz = \int |\Psi(\mathbf{r}, t)|^2\, d^3\mathbf{r} = 1. \tag{2.30}$$

The general solutions in three dimensions can be expressed as

$$\Psi(\mathbf{r}, t) = \sum c_n\psi_n(\mathbf{r})e^{-iE_nt/\hbar}. \tag{2.31}$$

The spatial wave functions, ψ_n, satisfy the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m_e}\nabla^2\psi + V\psi = E\psi. \tag{2.32}$$

2.3.1 Separation of variables - quantum numbers l and m

For a central symmetrical potential where the function V only depends on the distance, $V = V(|\mathbf{r}|) = V(r)$, it is common to introduce **spherical coordinates**, (r, θ, φ) (see figure 2.1), and try the approach of separation of variables. The solutions are on the form (see [2])

$$\psi(r, \theta, \varphi) = R(r)Y(\theta, \varphi). \tag{2.33}$$

In spherical coordinates the time-independent Schrödinger equation, Eq. (2.32), is

$$-\frac{\hbar^2}{2m_e}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2\psi}{\partial\varphi^2}\right)\right] + V\psi = E\psi. \tag{2.34}$$

By performing the same exercise as in section 2.1.2, calling the separation constant $l(l+1)$, this will give rise to the radial and angular equations for a single particle in a three dimensional central symmetrical potential:

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - \frac{2m_e r^2}{\hbar^2}\left[V(r) - E\right] = l(l+1), \tag{2.35}$$

and

$$\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}\right] = -l(l+1). \tag{2.36}$$

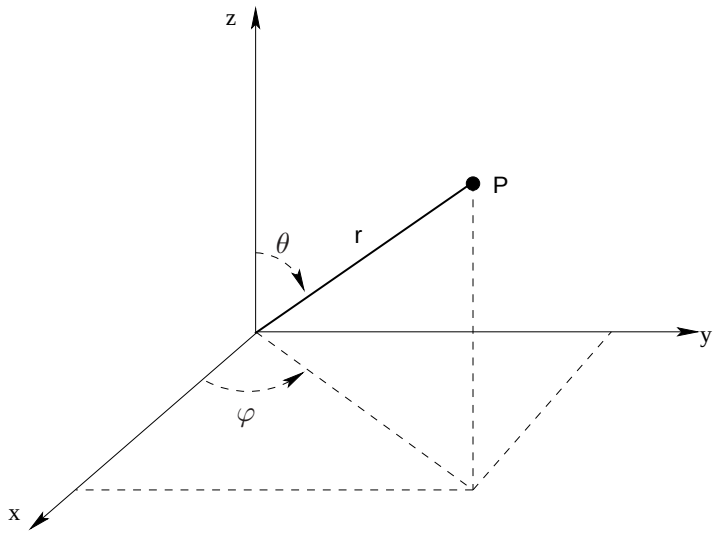

Figure 2.1: Visualizing the spherical coordinates of some point P; radius r, polar angle θ and azimuthal angle φ.

2.3.2 The Angular equation

If we multiply Eq. (2.36) by $Y\sin^2\theta$, again using separation of variables, with $Y(\theta, \varphi) = T(\theta)F(\varphi)$, and divide by $Y = TF$, we get

$$\frac{1}{T}\left[\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{dT}{d\theta}\right)\right] + l(l+1)\sin^2\theta = -\frac{1}{F}\frac{d^2F}{d\varphi^2}. \tag{2.37}$$

This time we call the separation constant m^2.

The φ equation

The equation for φ is calculated quite easily, with

$$\frac{d^2F}{d\varphi^2} = -m^2F \Rightarrow F(\varphi) = e^{im\varphi}, \tag{2.38}$$

letting the value m take both positive and negative values. Since the angle φ represents a direction in space, we must require that

$$F(\varphi + 2\pi) = F(\varphi). \tag{2.39}$$

It then follows that

$$e^{im(\varphi+2\pi)} = e^{im\varphi} \;\; \rightarrow \;\; e^{2\pi im} = 1. \tag{2.40}$$

For this to be fulfilled, m has to be an integer. Viz., $m = 0, \pm1, \pm2, \ldots$.

The θ equation

The differential equation for the polar angle, θ, reads

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{dT}{d\theta}\right) + [l(l+1)\sin^2\theta - m^2]T = 0. \tag{2.41}$$

This equation is more diffcult to solve by standard mathematics. In short, the solution is given by

$$T(\theta) = AP_l^m(\cos\theta), \tag{2.42}$$

where A is a constant, and P_l^m is the **associated Legendre function** (see e.g. [7])

$$P_l^m(x) = (1-x^2)^{|m|/2}\left(\frac{d}{dx}\right)^{|m|}P_l(x). \tag{2.43}$$

$P_l(x)$ is the lth degree **Legendre polynomial**, defined by the so-called **Rodrigues formula**:

$$P_l(x) = \frac{1}{2^l l!}\left(\frac{d}{dx}\right)^l\left(x^2-1\right)^l. \tag{2.44}$$

We see that l must be a positive integer for the differentiations to make sense, and furthermore, $|m|$ must be smaller or equal to l for $P_l^m(x)$ to be non-zero.

These solutions are the physically acceptable ones from the differential equation Eq. (2.36). Another set of non-physical solutions also exist, but these are not interesting for us.

Taking into account the normalization of the angular wave functions, a quite general expression for the functions Y_l^m is then

$$Y_l^m(\theta,\varphi) = \epsilon\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}e^{im\varphi}P_l^m(\cos\theta). \tag{2.45}$$

These functions are called the **spherical harmonics**. Here $\epsilon = (-1)^m$ for $m \geq 0$ and $\epsilon = 1$ for $m \leq 0$. Examples of these are:

$$Y_0^0 = \left(\frac{1}{4\pi}\right)^{1/2}, \tag{2.46}$$

$$Y_1^0 = \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta, \tag{2.47}$$

and

$$Y_1^{\pm 1} = \pm \left(\frac{3}{8\pi}\right)^{1/2} \sin\theta e^{\pm i\varphi}. \tag{2.48}$$

The quantum number l is commonly called the **azimuthal quantum number**, while m is called the **magnetic quantum number**. We observe that the spherical harmonics do not depend on the potential V, but we have required that the potential is spherically symmetric for the derivations.

2.3.3 The Radial equation and solution for the hydrogen atom

The radial equation reads

$$\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right) - \frac{2m_e r^2}{\hbar^2}[V(r) - E]R = l(l+1)R. \tag{2.49}$$

The first simple step is to change variables by using

$$u(r) = rR(r), \tag{2.50}$$

which will give the radial part of the Schrödinger equation on a much simpler form

$$-\frac{\hbar^2}{2m_e}\frac{d^2u}{dr^2} + \left[V + \frac{\hbar^2}{2m_e}\frac{l(l+1)}{r^2}\right]u = Eu. \tag{2.51}$$

The only difference between Eq. (2.51) and the one-dimensional Schrödinger equation, Eq. (2.10), is the potential. Here we have an effective potential

$$V_{eff} = V + \frac{\hbar^2}{2m_e}\frac{l(l+1)}{r^2}, \tag{2.52}$$

as opposed to the simple V in Eq. (2.10). The extra term in V_{eff}

$$\frac{\hbar^2}{2m_e}\frac{l(l+1)}{r^2}, \tag{2.53}$$

is called the centrifugal term.

The Hydrogen atom and the principal quantum number - n

An important quantum mechanical system is the hydrogen atom. It consists of an electron, charged $-e$, and a proton, charged e. When dealing with atomic systems, it is common to invoke the Born-Oppenheimer approximation (BOA). The BOA says that the kinetic energy of the nucleus is so small compared to the kinetic energy of the electron(s), that we can freeze out the nucleus' kinetic degrees of freedom. More on the Born-Oppenheimer approximation in section 3.1.4.

For the hydrogen atom, the potential function is given by Coulomb's law between charged particles:

$$V(r) = -\frac{e^2}{4\pi\epsilon_0}\frac{1}{r}, \tag{2.54}$$

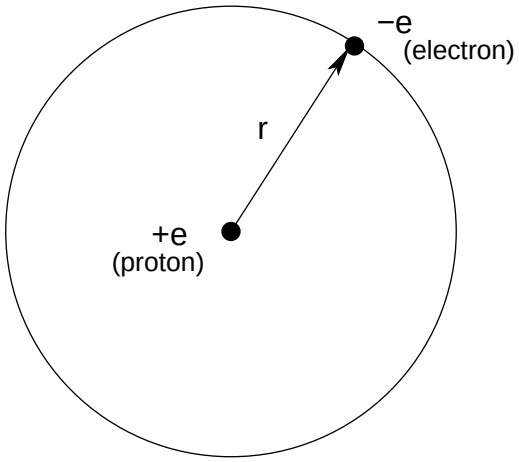

Figure 2.2: A shell model of the hydrogen atom.

where ϵ_0 is the vacuum permittivity. The radial equation for the hydrogen atom is then

$$-\frac{\hbar^2}{2m_e}\frac{d^2u}{dr^2} + \left[-\frac{e^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{\hbar^2}{2m_e}\frac{l(l+1)}{r^2}\right]u = Eu. \tag{2.55}$$

Figure 2.2 gives a schematic overview of the simple hydrogen atom.

In these derivations we will consentrate on the *bound* states, viz. $E < 0$.

The radial differential equation, Eq. (2.55), is quite tedious and difficult to solve, so in this section I will just state the important results. From solving the equation we will obtain a certain quantum number, n, also called the **principal quantum number**. This number determines the allowed energies for the system given by

$$E = -\left[\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right]\frac{1}{n^2} = \frac{E_1}{n^2} \tag{2.56}$$

for $n = 1, 2, 3, \ldots$. The ground state of hydrogen is then given as

$$E_1 = -\left[\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right] = -13.6 \text{ eV}., \tag{2.57}$$

The radial part of the wave function is now labeled by quantum numbers n and l, giving the full wave function the form

$$\psi_{nlm}(r, \theta, \varphi) = R_{nl}(r)Y_l^m(\theta, \varphi). \tag{2.58}$$

By solving the equation, we also find a constraint on l, namely

$$l = 0, 1, 2, \ldots, n-1. \tag{2.59}$$

The mathematical form of the radial wave function is given in [2] as

$$R_{nl} = \sqrt{\left(\frac{2}{na}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} e^{-r/na}\left(\frac{2r}{na}\right)^l \left[L_{n-l-1}^{2l+1}(2r/na)\right], \tag{2.60}$$

where a is the Bohr radius

$$a \equiv \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.529 \cdot 10^{-10}\text{m}, \tag{2.61}$$

and

$$L^p_{q-p}(x) \equiv (-1)^p \left(\frac{d}{dx}\right)^p L_q(x) \tag{2.62}$$

is the **associated** Laguerre polynomial. The function $L_q(x)$ is the qth Laguerre polynomial and is defined by

$$L_q(x) \equiv e^x \left(\frac{d}{dx}\right)^q (e^{-x}x^q) \tag{2.63}$$

with $p = 2l + 1$ and $q = n + l$. The radial wave functions needed in this thesis are as given in [2]

$$R_{10} = 2a^{-3/2}e^{-r/a}, \tag{2.64}$$

$$R_{20} = \frac{1}{\sqrt{2}}a^{-3/2}\left(1 - \frac{1}{2}\frac{r}{a}\right)\frac{r}{a}e^{-r/2a}, \tag{2.65}$$

$$R_{21} = \frac{1}{\sqrt{24}}a^{-3/2}\frac{r}{a}e^{-r/2a}, \tag{2.66}$$

$$R_{30} = \frac{2}{\sqrt{27}}a^{-3/2}\left(1 - \frac{2}{3}\frac{r}{a} + \frac{2}{27}\left(\frac{r}{a}\right)^2\right)e^{-r/3a}, \tag{2.67}$$

$$R_{31} = \frac{8}{27\sqrt{6}}a^{-3/2}\left(1 - \frac{1}{6}\frac{r}{a}\right)\left(\frac{r}{a}\right)e^{-r/3a}. \tag{2.68}$$

In sections 5.4 and 3.1.2 we will see how we can use these functions to construct the wave functions we need for the numerical calculations.

Chapter 3

Many-particle Systems

While single-particle systems are instructive and are fine examples to examine the properties of quantum mechanical systems, most real life systems consist of many particles that interact with each other and/or an external potential. The aim of this chapter is to describe the quantum mechanics of such systems, or more precisely, atomic systems.

The Hamiltonian of the atomic system will be discussed in section 3.1.4. This is the most important quantity when discussing the energy of a quantum mechanical system. In this part we will see how we can scale the Hamiltonian to have a much cleaner and more comfortable setup for our calculations.

We will also present the Hartree-Fock method (HF), a much used technique used in studies of many-particle systems. The HF method will mainly be used as a way to improve the single particle wave functions used in the Variational Monte Carlo calculations, which will be discussed in chapters 4 and 5.

3.1 Atoms

In quantum mechanics an atom can be viewed as a many-particle system. While the wave function for a single particle system is a function of only the coordinates of that particular particle and time, $\Psi(\mathbf{r}, t)$, a many-particle system will depend on the coordinates of all the particles.

3.1.1 Two-particle systems

The simplest example of a many-particle system is of course a two-particle system. The wave function will now be the function

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t), \tag{3.1}$$

and the Hamiltonian will take the form

$$\widehat{\mathbf{H}} = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1, \mathbf{r}_2, t) + V_{external}(\mathbf{r}_1, \mathbf{r_2}). \tag{3.2}$$

where $V_{external}$ depends on the system, e.g. the Coulomb interaction between electrons and the nucleus when dealing with atoms. The indices on the Laplacian operators indicate which electron/coordinates the derivative is taken with respect to. The

normalization integral is now

$$\int |\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \, d^3\mathbf{r}_1 d^3\mathbf{r}_2 = 1 \tag{3.3}$$

with the integrand

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \, d^3\mathbf{r}_1 d^3\mathbf{r}_2 \tag{3.4}$$

being the probability of finding particle 1 in the infinitesimal volume $d^3\mathbf{r}_1$, and particle 2 in $d^3\mathbf{r}_2$. For a time-independent potential, this will give the stationary solutions from separation of variables analogous to the single particle case in section 2.1.2. The full wave function will be

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, t) = \psi(\mathbf{r}_1, \mathbf{r}_2) e^{-iEt/\hbar}, \tag{3.5}$$

and the probability distribution function (PDF) will be time-independent

$$P(\mathbf{r}_1, \mathbf{r}_2) = |\psi(\mathbf{r}_1, \mathbf{r}_2)|^2. \tag{3.6}$$

Antisymmetry and wavefunctions

For general particles in a composite system, a wave function can be constructed by using the single particle wave functions the particles currently occupy. An example is if electron 1 is in state ψ_a and electron 2 is in state ψ_b. The wave function can be constructed as

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2). \tag{3.7}$$

The particles that make up an atom are electrons and the nuclei. The nuclei are the protons and neutrons that the atomic nucleus consists of. As we will see in section 3.1.4, the interesting parts are the electrons. Electrons and the nuclei are so-called fermions, that means particles with *half integer* spin, while particles with *integer* spin are called bosons (e.g. photons). An important point with respect to our quantum mechanical approach is how to construct the wave function for a composite system of particles. Quantum physics tells us that the particles are identical, and cannot be distinguished by some classic method like labelling the particles. This is taken care of by constructing a wave function that opens for the possibility of both electrons being in both states. That is,

$$\psi_\pm(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \pm \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)], \tag{3.8}$$

with A being a normalization constant. The plus sign is for bosons, while the minus sign applies for fermions. This also shows the **Pauli exclusion principle** which states: No two identical fermions can occupy the same state at the same time. E.g. if $\psi_a = \psi_b$ the wave function automatically yields zero for fermions, that is

$$\psi_-(\mathbf{r}_1, \mathbf{r}_2) = A[\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) - \psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] = 0, \tag{3.9}$$

and so the PDF will be zero everywhere. It also tells you that two bosons in **can** occupy the same state at the same time. In fact, any number of bosons may occupy the same state at the same time. Equation (3.8) also underlines an important property of fermionic wave function. That is the antisymmetric property when interchanging two particles. By switching particles 1 and 2 in ψ_-, you will get $-\psi_-$ in return, illustrated by

$$\begin{aligned}\psi_-(\mathbf{r}_2,\mathbf{r}_1) &= A[\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1) - \psi_b(\mathbf{r}_2)\psi_a(\mathbf{r}_1)] \\ &= -A[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) - \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] = -\psi_-(\mathbf{r}_1,\mathbf{r}_2). \end{aligned} \tag{3.10}$$

In atoms however, two partices are allowed to occupy the same spatial wave function state, but only if these have different spin states(ref. section 2.1.2). An example is helium, an atom consisting of two electrons and and a core with charge Z=2. Since we can't solve the helium problem, we use an ansatz for the wave function, with the two electrons occupying the ψ_{100}-functions found when solving the hydrogen problem. The two electrons in the helium case will have the wave functions

$$\Psi = 2a^{-3/2}e^{-r/a}\left(\frac{1}{4\pi}\right)^{1/2}\chi_\pm, \tag{3.11}$$

with $\chi_+ = \uparrow$ as the spin up state, and $\chi_- = \downarrow$ as the spin down state.

3.1.2 Wave functions for N-particle atoms

The many particle systems of most interest in this thesis are atomic systems. We can consider both the neutral atom, with the number of electrons equalling the number of protons in the nucleus, or as an isotope, where some of the outermost electrons have been excited. Examples which will be dealt with in this thesis are the neutral atoms helium, beryllium, neon, magnesium and silicon. As mentioned above, we have no closed form solution to any atomic system other than the hydrogen atom, and we have to make an ansatz for the wave function.

Slater Determinants

The two-electron wave function

$$\psi_-(\mathbf{r_1},\mathbf{r_2}) = A[\psi_a(\mathbf{r_1})\psi_b(\mathbf{r_2}) - \psi_b(\mathbf{r_1})\psi_a(\mathbf{r_2})], \tag{3.12}$$

can be expressed as a determinant

$$\psi_-(\mathbf{r_1},\mathbf{r_2}) = A\begin{vmatrix} \psi_a(r_1) & \psi_a(r_2) \\ \psi_b(r_1) & \psi_b(r_2) \end{vmatrix}. \tag{3.13}$$

This is called the **Slater determinant** for the two particle case where the available single particle states are ψ_a and ψ_b. The factor A is in this case $1/\sqrt{2}$. The general expression for a Slater determinant, Φ, with N electrons and N available single particle states is as given in [8]:

$$\Phi = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(r_1) & \psi_1(r_2) & \cdots & \psi_1(r_N) \\ \psi_2(r_1) & \psi_2(r_2) & \cdots & \psi_2(r_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(r_1) & \psi_N(r_2) & \cdots & \psi_N(r_N) \end{vmatrix}. \tag{3.14}$$

This will automatically comply with the antisymmetry principle and is suited as a wave function for a fermionic many-particle system. The general structure is to take the

determinant of an $N \times N$-matrix with available particle states in increasing order in the columns. The states are functions of the particle coordinates which are in increasing order in the rows.

More properties of the many-particle Slater determinant will be discussed in section 3.2.

3.1.3 Electron configuration

The electron configuration of an atom describes how the electrons are distributed in the given orbitals of the system. The rules for quantum numbers obtained in sections 2.3.1 and 2.3.3 give us the allowed combinations for the configuration. As shown, n can be $n = 1, 2, 3, \ldots$, while l has the allowed values $l = 0, 1, \ldots, n-1$. The quantum number m can take the values $m = -l, -l+1, \ldots, l-1, l$.

A common way to label the orbitals is to use a notation (see section 5.2.2 in [2]) which labels states with $l = 0$ as s-states, $l = 1$ as p-states and so on. An orbital with $n = 1$ and $l = 0$ is labelled $1s$, and an orbital with $n = 2$ and $l = 1$ is labelled $2p$. Table 3.1 shows the full list for the azimuthal quantum number, l. The letters s, p, d and f are historical

Quantum number l	Spectroscopic notation
0	s
1	p
2	d
3	f
4	g
5	h
6	i
7	k
$\vdots$	$\vdots$

Table 3.1: The table shows the spectroscopic notation for the azimuthal quantum number l.

names originating from the words *sharp*, *principal*, *diffuse* and *fundamental*. From g and out the letters are just given alphabetically, but actually skipping j for historical reasons.

The neon atom consists of $N = 10$ electrons and its orbital distribution is given as follows:

- When $n = 1$, l (and therefore m) can only assume the value 0. But an electron can either be in a spin up, or a spin down state. This means that two electrons can be in the $1s$-orbital. This is written as $(1s)^2$.
- For $n = 2$ and $l = 0$ there is only $m = 0$, but again two spin states, i.e. $(2s)^2$.
- For $n = 2$ and $l = 1$, we can have $m = -1, 0, 1$, and both spin states. This means 6 electrons in the $2p$-orbital, i.e. $(2p)^6$.
- The electron configuration in neon can then be written short and concise as: $(1s)^2(2s)^2(2p)^6$.

We can now use this notation to describe the distribution of electrons in the different orbitals for the atoms examined in this thesis. Table 3.2 shows a few important examples. All this tells us is which orbitals to put our electrons in, but it doesn't say anything about what the orbitals in fact are. This is taken care of in sections 5.4 and 5.5.

Atom	Full spectroscopic notation
Hydrogen	$(1s)$
Helium	$(1s)^2$
Beryllium	$(1s)^2(2s)^2$
Neon	$(1s)^2(2s)^2(2p)^6$
Magnesium	$(1s)^2(2s)^2(2p)^6(3s)^2$
Silicon	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^2$

Table 3.2: The table shows the electron configuration for hydrogen, helium, beryllium, neon, magnesium and silicon.

The spectroscopic notation is needed when dealing with shells that are not closed. Table 5.1 in [2] shows the configuration, ^{2S+1}L, for the first 36 atoms. Here S is the total spin of the system and L is the total angular momentum. Closed-orbital systems like helium, beryllium, neon and magnesium are labelled as ^1S (S here represents a state of zero angular momentum, and is not to be confused with the total spin). This means that the total spin is zero, corresponding to half of the electrons being in a spin-up state, and the rest in a spin down state. We also see that the angular momentum is zero, which simply means that all possible states are occupied. It is also common to denote the grand total angular momentum as a lower index J, but the values of interest for us are total spin and total angular momentum.

As an example we can consider neon. The first two electrons are in $1s$-states, which means that they have zero orbital momentum. The sum of the spin up and the spin down electrons is zero, $+\frac{1}{2} - \frac{1}{2} = 0$. This is also true for the two electrons in the $2s$-states. For the six remaining electrons in the $2p$-state, the sum of the angular momenta is zero, $(+1) + (+1) + 0 + 0 + (-1) + (-1) = 0$, corresponding to the six electrons being distributed in the states with $m_l = -1$, $m_l = 0$ and $m_l = +1$ (we count twice for spin up and spin down). From this we have that the total spin is zero, $2S + 1 = 1$, the total angular momentum is zero (it is represented by the capital S). This results in a configuration ^1S for neon.

For silicon however, we see that the notation is ^3P. This corresponds to the total spin being one and the total angular momentum being one as well. For the total spin to be one, the two outermost electrons must both have spin up. Since the total angular momentum is one, these two electrons can e.g. be in states $3p$ with $m_l = 0$ and $3p$ with $m_l = +1$.

Hund's rules (see [2]) are constructed to predict the configuration of these systems.

3.1.4 Hamiltonian and scaling

The Hamiltonian, $\widehat{\mathbf{H}}$, for an N-electron atomic system is subject to the Born-Oppenheimer approximation (BOA). The BOA effectively freezes out the nucleus' degrees of freedom given the electron mass is small compared to the mass of the nucleus. As a result, we approximate the kinetic energy of the nucleus to be zero during the

calculations. The Hamiltonian consists of only the kinetic energy of electrons in addition to the potential energy between electrons and the core, and between the electrons themselves. The Born-Oppenheimer Hamiltonian is given as

$$\widehat{\mathbf{H}} = \sum_i^N \frac{\widehat{\mathbf{p}}_i^2}{2m_e} - \sum_i^N \frac{Ze^2}{4\pi\epsilon_0 r_i} + \sum_{i<j} \frac{e^2}{4\pi\epsilon_0 r_{ij}}. \tag{3.15}$$

The operator $\widehat{\mathbf{p}}$ is the kinetic energy operator, m_e is the electron mass, Z is the number of protons in the nucleus (equal to the number of electrons in neutral atoms) and ϵ_0 the vacuum permittivity. The variables r_i and r_{ij} represent the distances between an electron and the nucleus and the distance between to different electrons respectively.

By substituting $\widehat{\mathbf{p}}$ with the quantum mechanical operator $\widehat{\mathbf{p}} = -i\hbar\nabla$, i being the imaginary unit and $\hbar$ Planck's original constant divided by 2π, we get

$$\widehat{\mathbf{H}} = -\frac{\hbar^2}{2m_e}\sum_i^N \nabla_i^2 - \sum_i^N \frac{Ze}{4\pi\epsilon_0 r_i} + \sum_{i<j} \frac{e^2}{4\pi\epsilon_0 r_{ij}}. \tag{3.16}$$

It is more desireable to work with a scaled version of the Hamiltonian. By defining

$$r_0 \equiv \frac{4\pi\epsilon_o\hbar^2}{m_e e^2}, \tag{3.17}$$

and

$$\Omega \equiv \frac{m_e e^4}{(4\pi\epsilon_0\hbar)^2}, \tag{3.18}$$

we can write dimensionless variables such as $H' = H/\Omega$, $\nabla' = \nabla/r_0$ and $r' = r/r_0$. The factors r_0 in ∇' comes from the factors $\mathrm{d}/\mathrm{d}x = \mathrm{d}/\mathrm{d}(x'r_0)$. By analyzing each term in Eq. (3.15) and writing out the original variables in terms of constants r_0 and Ω, and the dimensionless quantities r' and $\widehat{\mathbf{H}}'$ we get

$$\widehat{\mathbf{H}} = \widehat{\mathbf{H}}'\Omega = -\frac{\hbar^2}{2m_e r_0^2}\sum_i^N (\nabla_i')^2 + \frac{e^2}{4\pi\epsilon_0 r_o}\sum_i^N \frac{1}{r_i'} - \frac{Ze^2}{4\pi\epsilon_0 r_0}\sum_{i<j}\frac{1}{r_{ij}'}. \tag{3.19}$$

By inserting Eqs. (3.17) and (3.18) into Eq. (3.19) we obtain

$$\widehat{\mathbf{H}}'\Omega = -\frac{\Omega}{2}\sum_i^N (\nabla_i')^2 - \Omega\sum_i^N \frac{Z}{r_i'} + \Omega\sum_{i<j}\frac{1}{r_{ij}'}, \tag{3.20}$$

which divided by Ω will give the final dimensionless Hamiltonian

$$\widehat{\mathbf{H}} = -\frac{1}{2}\sum_i^N \nabla_i^2 - \sum_i^N \frac{Z}{r_i} + \sum_{i<j}\frac{1}{r_{ij}}. \tag{3.21}$$

Now the energy will be measured in the units of Hartree, E_h, which converts as, $1E_h = 2 \cdot 13.6$ eV. These are also called atomic units. We recognize the number $13,6$ as the ground state energy of the hydrogen atom (see section 2.3.3).

3.2 Hartree-Fock Theory

A popular and well used many-particle method is the Hartree-Fock method. Hartree-Fock theory assumes the Born-Oppenheimer approximation and can be used to approximate the ground state energy and ground state wave function of a quantum many body system. In Hartree-Fock theory we assume the wave function, Φ, to be a single N-particle Slater determinant. A Hartree-Fock calculation will result in obtaining the single-particle wave functions which minimizes the energy. These wave functions can then be used as optimized wave functions for a variational Monte Carlo machinery (see section 4.4). The single particle orbitals for an atom are given as

$$\psi_{nlm_l sm_s} = \phi_{nlm_l}(\mathbf{r})\chi_{m_s}(s), \tag{3.22}$$

where m_l is the projection of the quantum number l, given as merely m in section 3.1.3. The quantum number s represents the intrinsic spin of the electrons, which is $s = 1/2$. The quantum number m_s now represents the two projection values an electron can have, namely $m_s = \pm 1/2$. For a simple description, we label the single particle orbitals as

$$\psi_{nlm_l sm_s} = \psi_\alpha, \tag{3.23}$$

with α containing all the quantum numbers to specify the orbital. The N-particle Slater determinant is given as

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, \alpha, \beta, \ldots, \nu) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_\alpha(\mathbf{r}_1) & \psi_\alpha(\mathbf{r}_2) & \cdots & \psi_\alpha(\mathbf{r}_N) \\ \psi_\beta(\mathbf{r}_1) & \psi_\beta(\mathbf{r}_2) & \cdots & \psi_\beta(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_\nu(\mathbf{r}_1) & \psi_\nu(\mathbf{r}_2) & \cdots & \psi_\nu(\mathbf{r}_N) \end{vmatrix}. \tag{3.24}$$

The Hamiltonian (see section 3.1.4) is given as

$$\widehat{\mathbf{H}} = \widehat{\mathbf{H}}_1 + \widehat{\mathbf{H}}_2 = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} \right) + \sum_{i<j} \frac{1}{r_{ij}}. \tag{3.25}$$

or even as

$$\widehat{\mathbf{H}} = \sum_{i=1}^{N} \hat{h}_i + \sum_{i<j} \frac{1}{r_{ij}}, \tag{3.26}$$

with $\hat{h}_i$ being the one-body Hamiltonian

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i}. \tag{3.27}$$

The variational principle (section 4.3) tells us that the ground state energy for our Hamiltonian, E_0, is always less or equal to the expectation value of the Hamiltonian with a chosen trial wave function, Φ. That is

$$E_0 \leq E[\Phi] = \int \Phi^* \widehat{\mathbf{H}} \Phi d\tau, \tag{3.28}$$

where the brackets in $E[\Phi]$ tells us that the expectation value is a functional, i.e. a function of a function, here an integral-function. The label τ represents a shorthand notation for $d\tau = d\mathbf{r}_1 d\mathbf{r}_2 \ldots d\mathbf{r}_N$. We also assume the trial function Φ is normalized

$$\int \Phi^* \Phi d\tau = 1. \tag{3.29}$$

By introducing the so-called anti-symmetrization operator, $\mathcal{A}$, and the Hartree-function, $\Phi_H = \psi_\alpha(\mathbf{r}_1)\psi_\beta(\mathbf{r}_2)\dots\psi_\nu(\mathbf{r}_N)$, we can write the Slater determinant (Eq. 3.24) more compactly as

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \alpha, \beta, \dots, \nu) = \sqrt{N!}\mathcal{A}\Phi_H. \tag{3.30}$$

The Hartree-function is as shown a simple product of the available single particles states and the operator $\mathcal{A}$ is given as

$$\mathcal{A} = \frac{1}{N!}\sum_P (-)^P P, \tag{3.31}$$

with P being the permutation operator of two particles, and the sum spanning over all possible permutations of two particles. The operators $\widehat{\mathbf{H}}_1$ and $\widehat{\mathbf{H}}_2$ themselves do not depend on permutations and will commute with the anti-symmetrization operator $\mathcal{A}$, namely

$$[\widehat{\mathbf{H}}_1, \mathcal{A}] = [\widehat{\mathbf{H}}_2, \mathcal{A}] = 0. \tag{3.32}$$

The operator $\mathcal{A}$ also has the property

$$\mathcal{A} = \mathcal{A}^2, \tag{3.33}$$

which we can use in addition to Eq. (3.32) to show that

$$\int \Phi^* \widehat{\mathbf{H}}_1 \Phi d\tau = N! \int \Phi_H^* \mathcal{A}\widehat{\mathbf{H}}_1 \mathcal{A}\Phi_H d\tau \tag{3.34}$$

$$= N! \int \Phi_H^* \widehat{\mathbf{H}}_1 \mathcal{A}\Phi_H d\tau. \tag{3.35}$$

By using Eq. (3.30) and inserting the expression for $\widehat{\mathbf{H}}_1$ we arrive at

$$\int \Phi^* \widehat{\mathbf{H}}_1 \Phi d\tau = \sum_{i=1}^{N}\sum_P (-)^P \int \Phi_H^* \widehat{\mathbf{h}}_i P \Phi_H d\tau. \tag{3.36}$$

Orthogonality of the single-particle functions ensures us that we can remove the sum over P, as the integral will disappear when the two Hartree-Functions, Φ_H^* and Φ_H are permuted differently. As the operator $\widehat{\mathbf{h}}_i$ is a single-particle operator, all factors in Φ_H except one, $\phi_\mu(\mathbf{r}_i)$, will integrate out under orthogonality conditions, and the expectation value of the H_1-operator is written as

$$\int \Phi^* \widehat{\mathbf{H}}_1 \Phi d\tau = \sum_{\mu=1}^{N} \int \psi_\mu^*(\mathbf{r}_i)\widehat{\mathbf{h}}_i \psi_\mu(\mathbf{r}_i) d\mathbf{r}_i. \tag{3.37}$$

The two-body Hamiltonian, H_2, can be treated similarly:

$$\int \Phi^* \widehat{\mathbf{H}}_2 \Phi d\tau = N! \int \Phi_H^* \mathcal{A}\widehat{\mathbf{H}}_2 \mathcal{A}\Phi_H d\tau \tag{3.38}$$

$$= \sum_{i<j=1}^{N}\sum_P (-)^P \int \Phi_H^* \frac{1}{r_{ij}} P \Phi_H d\tau. \tag{3.39}$$

In order to reduce this to a simpler form, we cannot use the same argument as for H_1 for get rid of the P-sum, because of the form of the interaction, $1/r_{ij}$. The permutations of two electrons will not vanish and must be accounted for by writing

$$\int \Phi^* \widehat{\mathbf{H}}_2 \Phi d\tau = \sum_{i<j=1}^{N} \int \Phi_H^* \frac{1}{r_{ij}} (1 - P_{ij}) \Phi_H d\tau. \tag{3.40}$$

P_{ij} is the operator which permutes electrons i and j. However, by inspecting the parenthesis $(1 - P_{ij})$ we can again take advantage of the orthogonality condition in order to further simplify the expression:

$$\int \Phi^* \widehat{\mathbf{H}}_2 \Phi d\tau = \frac{1}{2} \sum_{\mu=1}^{N} \sum_{\nu=1}^{N} \left[\int \psi_\mu^*(\mathbf{r}_i) \psi_\nu^*(\mathbf{r}_j) \frac{1}{r_{ij}} \psi_\mu(\mathbf{r}_i) \psi_\nu(\mathbf{r}_j) d\mathbf{r}_i d\mathbf{r}_j \right.$$
$$\left. - \int \psi_\mu^*(\mathbf{r}_i) \psi_\nu^*(\mathbf{r}_j) \frac{1}{r_{ij}} \psi_\mu(\mathbf{r}_j) \psi_\nu(\mathbf{r}_i) d\mathbf{r}_i d\mathbf{r}_j \right]. \tag{3.41}$$

The first term is the *Hartree* term, while the second term is called the *Fock* term. Alternatively they are called the *direct* and *exchange* terms. We see the exchange term has permuted the particles i and j in the Φ_H-function. The factor of $1/2$ is due to the fact that we sum freely over μ and ν instead of using $\sum_{i<j}$.

Combining the results from $\widehat{\mathbf{H}}_1$ and $\widehat{\mathbf{H}}_2$ you get the full Hartree-Fock energy functional

$$E[\Phi] = \sum_{\mu=1}^{N} \int \psi_\mu^*(\mathbf{r}_i) \widehat{\mathbf{h_i}} \psi_\mu(\mathbf{r}_i) d\mathbf{r}_i + \frac{1}{2} \sum_{\mu}^{N} \sum_{\nu}^{N} \left[\int \psi_\mu^*(\mathbf{r}_i) \psi_\nu^*(\mathbf{r}_j) \frac{1}{r_{ij}} \psi_\mu(\mathbf{r}_i) \psi_\nu(\mathbf{r}_j) d\mathbf{r}_i d\mathbf{r}_j \right.$$
$$\left. - \int \psi_\mu^*(\mathbf{r}_i) \psi_\nu^*(\mathbf{r}_j) \frac{1}{r_{ij}} \psi_\nu(\mathbf{r}_i) \psi_\nu(\mathbf{r}_j) d\mathbf{r}_i d\mathbf{r}_j \right] \tag{3.42}$$

where now $\widehat{\mathbf{h_i}}$ are the one body Hamiltonians, that is the sum of the kinetic energy and the one-body interaction with the nucleus. The states ψ_μ and ψ_ν are the Hartree-Fock orbitals which minimizes the energy when we solve the Hartree-Fock equations. Using Bra-Ket notation (see section 2.2), the energy functional can be written more compactly as

$$E[\Phi] = \sum_{\mu=1}^{N} \langle \mu | \, h \, | \mu \rangle + \frac{1}{2} \sum_{\mu=1}^{N} \sum_{\nu=1}^{N} \left[\langle \mu\nu | \, \frac{1}{r_{ij}} \, | \mu\nu \rangle - \langle \mu\nu | \, \frac{1}{r_{ij}} \, | \nu\mu \rangle \right]. \tag{3.43}$$

3.2.1 Hartree-Fock equations

There are basically two approaches when it comes to minimizing the energy functional in Eq. (3.43). The most straight forward, but a bit more demanding method is to vary the single particle orbitals. The other approach is to expand the single particle orbitals in a well known basis as

$$\psi_a = \sum_{\lambda} C_{a\lambda} \psi_\lambda, \tag{3.44}$$

where now the ψ_λ are in a known orthogonal basis (e.g. hydrogen-like wave functions, harmonic oscillator functions etc.), and we vary the functions ψ_a with respect to the expansion coefficients $C_{a\lambda}$.

Varying the single particle wave functions

The background for these calculations are the principles of variational calculus as described in section 18.4.1 in [9]. By introducing so-called Lagrangian multipliers we can minimize a multivariable-variable functional with constraints. For this particular problem, we introduce N^2 such Lagrange multipliers, $\epsilon_{\mu\nu}$. The variational equation (section 18.4.1 in [9]) for the energy functional in Eq. (3.42) is written as

$$\delta E - \sum_{\mu=1}^{N}\sum_{\nu=1}^{N}\epsilon_{\mu\nu}\delta\int\psi_\mu^*\psi_\nu = 0. \tag{3.45}$$

We still assume the wave functions are orthogonal, so the variational equation can be written

$$\delta E - \sum_{\mu=1}^{N}\epsilon_{\mu}\int\psi_\mu^*\psi_\mu = 0. \tag{3.46}$$

The next step is to perform variation with respect to the single particle orbitals ψ_μ, which gives

$$\begin{aligned}
&\sum_{\mu}^{N}\int\delta\psi_\mu^*\widehat{\mathbf{h}}_i\psi_\mu d\mathbf{r}_i + \frac{1}{2}\sum_{\mu=1}^{N}\sum_{\nu=1}^{N}\left[\int\delta\psi_\mu^*\psi_\nu^*\frac{1}{r_{ij}}\psi_\mu\psi_\nu d\mathbf{r}_i d\mathbf{r}_j - \int\delta\psi_\mu^*\psi_\nu^*\frac{1}{r_{ij}}\psi_\nu\psi_\mu d\mathbf{r}_i d\mathbf{r}_j\right]\\
&+\sum_{\mu}^{N}\int\psi_\mu^*\widehat{\mathbf{h}}_i\delta\psi_\mu d\mathbf{r}_i + \frac{1}{2}\sum_{\mu=1}^{N}\sum_{\nu=1}^{N}\left[\int\psi_\mu^*\psi_\nu^*\frac{1}{r_{ij}}\delta\psi_\mu\psi_\nu d\mathbf{r}_i d\mathbf{r}_j - \int\psi_\mu^*\psi_\nu^*\frac{1}{r_{ij}}\delta\psi_\nu\psi_\mu d\mathbf{r}_i d\mathbf{r}_j\right]\\
&-\sum_{\mu=1}^{N}\epsilon_\mu\int\delta\psi_\mu^*\psi_\mu d\mathbf{r}_i - \sum_{\mu=1}^{N}\epsilon_\mu\int\psi_\mu^*\delta\psi_\mu d\mathbf{r}_i = 0.
\end{aligned} \tag{3.47}$$

The variations $\delta\psi$ and $\delta\psi^*$ are obviously not independent as they are only a complex transformation apart, but they can indeed be treated independently by replacing $\delta\psi$ by the imaginary $i\delta\psi$ and correspondingly $\delta\psi^*$ by $i\delta\psi^*$. The terms depending on $\delta\psi$ and $\delta\psi^*$ respectively can now be set equal to zero, yielding two independent sets of equations. By again combining them we will obtain the Hartree-Fock equations

$$\begin{aligned}
&\left[-\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} + \sum_{\nu=1}^{N}\int\psi_\nu^*(\mathbf{r}_j)\frac{1}{r_{ij}}\psi_\nu^*(\mathbf{r}_j)d\mathbf{r}_j\right]\psi_\mu(\mathbf{r}_i)\\
&\qquad - \left[\sum_{\nu=1}^{N}\int\psi_\nu^*(\mathbf{r}_j)\frac{1}{r_{ij}}\psi_\mu^*(\mathbf{r}_j)d\mathbf{r}_j\right]\psi_\nu(\mathbf{r}_i) = \epsilon_\mu\psi_\mu(\mathbf{r}_i),
\end{aligned} \tag{3.48}$$

where the integral $\int d\mathbf{r}_j$ also includes a sum over spin quantum numbers for electron j. The first two terms in the first square bracket is the one-body Hamiltonian while the third term is the direct term, representing a mean field of all the other electrons as seen by electron i. The term in the second square bracket is the exchange term, resulting from our antisymmetric wave function ansatz. This term also takes care of the so-called *self interaction* from the first term by cancellation when $i = j$. We can now define the direct and exchange operators

$$V_\mu^d(\mathbf{r}_i) = \int\psi_\mu^*(\mathbf{r}_i)\frac{1}{r_{ij}}\psi_\mu(\mathbf{r}_j)d\mathbf{r}_j, \tag{3.49}$$

and

$$V_\mu^{ex}(\mathbf{r}_i)\psi_\nu(\mathbf{r}_i) = \left(\int \psi_\mu^* \frac{1}{r_{ij}} \psi_\nu(\mathbf{r}_j) d\mathbf{r}_j\right) \psi_\mu(\mathbf{r}_i). \tag{3.50}$$

Then the Hartree-Fock equations may be written as

$$H_i^{HF}\psi_\nu(\mathbf{r}_i) = \epsilon_\nu \psi_\nu(\mathbf{r}_i), \tag{3.51}$$

with

$$H_i^{HF} = \widehat{\mathbf{h}}_i + \sum_{\mu=1}^{N} V_\mu^d(\mathbf{r}_i) - \sum_{\mu=1}^{N} V_\mu^{ex}(\mathbf{r}_i), \tag{3.52}$$

as the Hartree-Fock matrix.

In section 5.5, we will discuss the Roothaan-Hartree-Fock approach, which is based on the method discussed in this section, with wave functions

$$\psi = \sum_p \chi_p C_p, \tag{3.53}$$

where $C_p(\xi)$ are some coefficients and χ are chosen to be the *Slater type functions* on the form

$$\chi(r) \propto r^{n-1} \exp(-\xi r). \tag{3.54}$$

Here ξ is some factor to be determined during the minimization. We see that we must both vary the parameters and the exponents. The results are given in section 5.5 and the explicit expressions and results in [10].

Varying the coefficients

We see how the single particle functions in general, e.g. ϕ, are expressed as an expansion of known functions, ψ, on the form

$$\phi = \sum_i C_i \psi_i. \tag{3.55}$$

In this section we will describe a method where we only vary the coefficients, and let the actual basis functions remain unchanged. This will in time give us a more time-efficient way to perform a Hartree-Fock calculation as we will show at a later stage.

The single particle basis functions are here denoted as ψ_λ, where $\lambda = 1, 2, \ldots$ and represents the full set of quantum numbers in order to describe a single particle orbital. Our new single particle orbitals will then be an expansion of these functions ψ_λ and will have roman indices $a = 1, 2, \ldots$ to distinguish them from the chosen basis functions with Greek indices. That is

$$\psi_a = \sum_\lambda C_{a\lambda} \psi_\lambda, \tag{3.56}$$

where $C_{a\lambda}$ are the expansion coefficients that will be varied. By using the Bra-Ket notation we can introduce a more compact way to write the **direct** and **exchange** terms from the Hartree-Fock energy functional. They are

$$\langle \mu\nu | \, V \, | \mu\nu \rangle = \int \psi_\mu^*(\mathbf{r}_i)\psi_\nu^*(\mathbf{r}_j) V(r_{ij}) \psi_\mu(\mathbf{r}_i)\psi_\nu(\mathbf{r}_j) d\mathbf{r}_i d\mathbf{r}_j, \tag{3.57}$$

and

$$\langle \mu\nu | \, V \, | \nu\mu \rangle = \int \psi_\mu^*(\mathbf{r}_i)\psi_\nu^*(\mathbf{r}_j) V(r_{ij}) \psi_\nu(\mathbf{r}_i)\psi_\mu(\mathbf{r}_j) d\mathbf{r}_i d\mathbf{r}_j, \tag{3.58}$$

respectively. We have now only written V as a general potential instead of the explicit $V = 1/r_{ij}$.

The interaction does not change under a permutation of two particles, which means that

$$\langle \mu\nu | \, V \, | \mu\nu \rangle = \langle \nu\mu | \, V \, | \nu\mu \rangle \, . \tag{3.59}$$

This is also true for the general case

$$\langle \mu\nu | \, V \, | \sigma\rho \rangle = \langle \nu\mu | \, V \, | \rho\sigma \rangle \, . \tag{3.60}$$

We can have an even more compact notation by defining an antisymmetrized matrix element

$$\langle \mu\nu | \, V \, | \mu\nu \rangle_{AS} = \langle \mu\nu | \, V \, | \mu\nu \rangle - \langle \mu\nu | \, V \, | \nu\mu \rangle \, , \tag{3.61}$$

or for the general case

$$\langle \mu\nu | \, V \, | \sigma\rho \rangle_{AS} = \langle \mu\nu | \, V \, | \sigma\rho \rangle - \langle \mu\nu | \, V \, | \rho\sigma \rangle \, . \tag{3.62}$$

The antisymmetrized matrix element has the following symmetry properties:

$$\langle \mu\nu | \, V \, | \sigma\rho \rangle_{AS} = - \langle \mu\nu | \, V \, | \rho\sigma \rangle_{AS} = - \langle \nu\mu | \, V \, | \sigma\rho \rangle_{AS} \, . \tag{3.63}$$

It is also hermitian, which means

$$\langle \mu\nu | \, V \, | \sigma\gamma \rangle_{AS} = \langle \sigma\gamma | \, V \, | \mu\nu \rangle_{AS} \, . \tag{3.64}$$

Using these definitions we can have a short-hand notation for the expectation value of the H_2 operator (see Eq. (3.41)

$$\int \Phi^* \widehat{\mathbf{H}}_2 \Phi d\tau = \frac{1}{2} \sum_{\mu=1}^{N} \sum_{\nu=1}^{N} \langle \mu\nu | \, V \, | \mu\nu \rangle_{AS} \, . \tag{3.65}$$

which in turn will give the full energy functional

$$E[\Phi] = \sum_{\mu=1}^{N} \langle \mu | \, h \, | \mu \rangle + \frac{1}{2} \sum_{\mu=1}^{N} \sum_{\nu=1}^{N} \langle \mu\nu | \, V \, | \mu\nu \rangle_{AS} \, , \tag{3.66}$$

similar to the one in Eq. (3.43).

The above functional is valid for any choice of single particle functions, so by using the new single particle orbitals introduced above in Eq. (3.56), and by now calling our Hartree-Fock Slater determinant, Ψ, we then have

$$E[\Psi] = \sum_{a=1}^{N} \langle a | \, h \, | a \rangle + \frac{1}{2} \sum_{a=1}^{N} \sum_{b=1}^{N} \langle ab | \, V \, | ab \rangle_{AS} \, . \tag{3.67}$$

By inserting the expressions for ψ_a we will then have

$$E[\Psi] = \sum_{a=1}^{N} \sum_{\alpha\beta} C^*_{a\alpha} C_{a\beta} \langle \alpha | \, h \, | \beta \rangle + \frac{1}{2} \sum_{a=1}^{N} \sum_{b=1}^{N} \sum_{\alpha\beta\gamma\delta} C^*_{a\alpha} C^*_{b\beta} C_{a\gamma} C_{b\delta} \langle \alpha\beta | \, V \, | \gamma\delta \rangle_{AS} \, . \tag{3.68}$$

Similarly as for the previous case where we minimized with respect to the wave functions themselves, we now introduce N^2 Lagrange multipliers. As a consequence of choosing

orthogonal basis functions, i.e. $\langle\alpha|\beta\rangle = \delta_{\alpha,\beta}$, where $\delta_{x,y}$ is the Kronecker delta function, the new basis functions are orthogonal, and we have

$$\langle a|b\rangle = \delta_{a,b} = \sum_{\alpha\beta} C^*_{a\alpha}C_{a\beta}\langle\alpha|\,|\beta\rangle = \sum_{\alpha} C^*_{a\alpha}C_{a\alpha}. \tag{3.69}$$

The functional to be minimized will now be

$$E[\Psi] - \sum_a \epsilon_a \sum_\alpha C^*_{a\alpha}C_{a\alpha} \tag{3.70}$$

giving a variational equation

$$\frac{d}{dC^*_{k\lambda}}\left[E[\Psi] - \sum_a \epsilon_a \sum_\alpha C^*_{a\alpha}C_{a\alpha}\right] = 0, \tag{3.71}$$

where we have minimized with respect to some coefficient $C^*_{k\lambda}$ since $C^*_{k\lambda}$ and $C_{k\lambda}$ are independent. Inserting the energy functional from Eq. (3.66) we will get

$$\begin{aligned}\sum_\beta C_{k\beta}\langle\lambda|h|\beta\rangle - \epsilon_k C_{k\lambda} + \frac{1}{2}\sum_{b=1}^{N}\sum_{\beta,\gamma,\delta} C^*_{b\beta}C_{k\gamma}C_{b\delta}\langle\lambda\beta|V|\gamma\delta\rangle_{AS} \\ + \frac{1}{2}\sum_{a=1}^{N}\sum_{\alpha,\gamma,\delta} C^*_{a\alpha}C_{a\gamma}C_{k\delta}\langle\alpha\lambda|V|\gamma\delta\rangle_{AS} = 0\end{aligned} \tag{3.72}$$

as

$$\frac{\mathrm{d}}{\mathrm{d}C_{k\lambda}}\left[\sum_{a=1}^{N}\sum_{\alpha,\beta} C^*_{a\alpha}C_{a\beta}\langle\alpha|h|\beta\rangle\right] = \sum_{a=1}^{N}\sum_{\alpha,\beta} C_{a\beta}\langle\alpha|h|\beta\rangle\,\delta_{ka}\delta_{\lambda\alpha} \tag{3.73}$$

$$= \sum_\beta C_{k\beta}\langle\lambda|h|\beta\rangle\,. \tag{3.74}$$

In the last term in Eq. (3.72) we can use that $\langle\alpha\lambda|V|\gamma\delta\rangle_{AS} = -\langle\lambda\alpha|V|\gamma\delta\rangle_{AS} = \langle\lambda\alpha|V|\delta\gamma\rangle_{AS}$ and change summation variable in both sums from b to a which gives

$$\sum_\beta C_{k\beta}\langle\lambda|h|\beta\rangle - \epsilon_k C_{k\lambda} + \sum_{a=1}^{N}\sum_{\alpha,\gamma,\delta} C^*_{a\alpha}C_{k\gamma}Ca\delta\langle\lambda\alpha|V|\gamma\delta\rangle_{AS} = 0, \tag{3.75}$$

or by using

$$h^{HF}_{\lambda\gamma} = \langle\lambda|h|\gamma\rangle + \sum_{a=1}^{N}\sum_{\alpha,\delta} C^*_{a\alpha}C_{a\delta}\langle\lambda\alpha|V|\gamma\delta\rangle_{AS}\,,$$

we can write the Hartree-Fock equations for this particular approach as

$$\sum_\gamma h^{HF}_{\lambda\gamma}C_{k\gamma} = \epsilon_k C_{k\lambda}. \tag{3.76}$$

This is now a non-linear eigenvalue problem which can be solved by an iterative process, or a standard eigenvalue method. The advantage of this approach is the fact that we can calculate the matrix elements $\langle\alpha|\,h\,|\beta\rangle$ and $\langle\alpha\beta|\,V\,|\gamma\delta\rangle$ once and for all, since we don't change the basis functions, but only their coefficients. In the first approach, where we vary the actual wave functions, we have to calculate these matrix elements for each iteration. When we just vary the coefficients, we can calculate the matrix elements and store them in a table before we start the iterative process. This will lead to a quicker algorithm, but we will see that the method of varying both coefficients and basis functions actually produces the best results.

Chapter 4

Quantum Monte Carlo

Monte Carlo methods are designed to simulate a mathematical system, or in our case, a quantum mechanical system. By using random numbers, these methods are considered stochastic, i.e. non-deterministic unlike other simulation techniques, such as *Hartree Fock*, *Coupled Cluster Theory* and *Configuration Interaction*, ref. [11]. Monte Carlo methods can be used to simulate quantum mechanical systems, but are also well suited for calculating integrals, especially high-dimensional integrals. There are in fact several Quantum Monte Carlo techniques such as *Diffusion Monte Carlo*, *Green's function Monte Carlo* and *Variational Monte Carlo*, see [3]. In this thesis we will focus on the Variational Monte Carlo method.

The Quantum Monte Carlo calculations are so called *ab initio* methods, which are first principle calculations These methods have their basis on probability and statistics, and in order to get good expectation values and variances, the quantity in question must be sampled millions of times. Each such sample is called a Monte Carlo cycle, and can be mistaken for being linked with the dynamics of the system. However, the systems we will examine are stationary, meaning that they are time-independent systems.

To use the Quantum Monte Carlo technique, we will need a Probability Distribution Function (PDF), $P(x)$, to guide our sampling and which characterizes the system, with x being some set of variables, and of course a quantity to sample, in our case the energy of the system, E. The expectation value of the quantity of interest, e.g. an operator $\widehat{\mathbf{Q}}$, will now be

$$\langle Q \rangle = \int P(x)\widehat{\mathbf{Q}}dx. \tag{4.1}$$

We will see in section 4.4.1 that the operator we are looking for in our case will be the local energy operator

$$\widehat{\mathbf{E}}_L(\mathbf{R};\alpha) = \frac{1}{\psi_T}\widehat{\mathbf{H}}\psi_T(\mathbf{R};\alpha), \tag{4.2}$$

where $\mathbf{R}$ is the set of all spatial variables, ψ_T a trial wave function since we don't have the exact solution Ψ (if we did, we wouldn't have to go through all of this), and α a set of variational parameters which will be discussed in the following section.

Since the Monte Carlo method is a statistical method involving average values etc., a certain degree of uncertainty and error will always occur. A method to minimize the error estimate, the so-called *blocking* technique, will be discussed as a way to get the results as correct as possible. In addition to this, we will discuss the *Davidon-Fletcher-Powell*-algorithm (DFP), which is an improvement of the Conjugate Gradient Method (CGM), see [12]. DFP is a method which finds the minimum of a multivariable function.

This chapter will also discuss the variational principle and the Metropolis algorithm for the Variational Monte Carlo method. The Metropolis-Hastings algorithm will also be discussed including importance sampling which will improve the method further, with more relevant sample points.

4.1 Markov chains

A Markov chain is a way to model a system which evolves in time, and depends solely on it's previous state. We can consider a system that is described by its constituents having different probabilities to occupy some physical state. That is, the system is described by a PDF, $w_j(t)$, at an instance t, where the index j indicates that the system is discretized (for simplicity). We now have a transition probability, W_{ij}, which gives us the probability for the system going from a state given by $w_j(t)$ to a state $w_i(t+\epsilon)$, where ϵ is a chosen time step. The relation between the PDF's w_i and w_j will be

$$w_i(t+\epsilon) = \sum_j W_{ij} w_j(t), \tag{4.3}$$

where w_i and W_{ij} are probabilities and must fulfill

$$\sum_i w_i(t) = 1, \tag{4.4}$$

and

$$\sum_i W_{ij} = 1. \tag{4.5}$$

We can also write this as a matrix equation

$$\mathbf{w}(t+\epsilon) = \mathbf{W}\mathbf{w}(t). \tag{4.6}$$

The equilibrium state of the system can be described by

$$\mathbf{w}(\infty) = \mathbf{W}\mathbf{w}(\infty), \tag{4.7}$$

which implies a steady state, since the transition matrix now has no effect on the state vector, $\mathbf{w}$. Eq. (4.7) now reads as an eigenvalue equation. By solving this equation, the eigenvectors would then be $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \ldots$ and so on, while the corresponding eigenvalues would be $\lambda_1, \lambda_2, \lambda_3, \ldots$ etc. If the largest eigenvalue is λ_1, viz.

$$\lambda_1 > \lambda_2, \lambda_3, \lambda_4, \ldots, \tag{4.8}$$

the steady state would then be $\mathbf{w}_1$, as its eigenvalue would cause this state to dominate as time progresses.

However, we do not know the form of the transition probability, W_{ij}, and need a way to examine the conditions at equilibrium. We start with the so-called *Master equation*,

$$\frac{dw_i(t)}{dt} = \sum_j \left[W(j \to i) w_j - W(i \to j) w_i \right], \tag{4.9}$$

which expresses the time rate of the PDF as a balance between state transitional probabilities from states j to i and from state i to states j. At equilibrium this time

derivative is zero, $dw(t_{eq})/dt = 0$, implying that the sums of the transitions must be equal, namely

$$\sum_j W(j \to i) w_j = \sum_j W(i \to j) w_i. \tag{4.10}$$

By using Eq. (4.5) on the right-hand side of the equation, we get

$$w_i = \sum_j W(j \to i) w_j. \tag{4.11}$$

We do, however, need an even more stringent requirement to obtain the correct solutions, which is the so-called *detailed balance principle*:

$$W(j \to i) w_j = W(i \to j) w_i, \tag{4.12}$$

saying that the transition rate between any two states i and j must be equal as well. This gives the condition

$$\frac{w_i}{w_j} = \frac{W(j \to i}{W(i \to j)}. \tag{4.13}$$

The detailed balance principle is imposed on the system to avoid cyclic solutions. This principle ensures us that once the most probable state is reached, the one corresponding to the lowest eigenvalue, the system will not traverse back to any state, corresponding to higher eigenvalues, see ref. [13].

A Markov chain must also comply with the *ergodic principle*, which has as a criterion that a system should have a non-zero probability of visiting all possible states during the process (see [3]). A Markov chain does only depend on its previous step, and will fulfill this condition of ergodicity. The Metropolis algorithm (see section 4.4.1) is based on a Markovian process and fulfills this condition of the ergodic principle.

Our physical system, an atom consisting of N electrons, is a system that can be described by a PDF and which only depends on its previous system state. The PDF is here the absolute square of the wave function, namely $|\psi|^2$. Since this describes a Markov process, a wise choice would then be to use the Metropolis algorithm in order to simulate the system.

4.2 Random numbers

A main component of a Monte Carlo calculation is the use of random numbers. The most common random number distribution is the uniform distribution, which returns a random number r between zero and one, that is

$$r \in (0, 1)\,. \tag{4.14}$$

The name *uniform* (chapter 7.1 in [12]) tells us that every number between 0 and 1 could be returned with equal probability. Another much used distribution is the Gaussian, or normal distribution (see chapter 7.3.4 in [12]).

As we saw in section 4.1, a Markov chain is a mathematical model which relies only on it's previous step in the calculation. The true random numbers used to calculate i.e. new positions for particles etc., should be uncorrelated. In computers however, we cannot get true random numbers, and must therefore settle for so-called Pseudo-Random Number Generators (PRNG), where the numbers are correlated. Another problem with

these number generators is that they have a certain period, and this sequence of pseudo-random numbers will repeat itself after completing a cycle of the sequence. However, the pseudo-random number generators we use have quite large periods, and we will not suffer because of this.

4.3 The Variational Principle

The main target for us is to solve the time independent Schrödinger equation

$$\widehat{\mathbf{H}}\Psi = E\Psi \tag{4.15}$$

where $\widehat{\mathbf{H}}$ is the Hamiltonian, the operator representing the energy of the system, Ψ the quantum mechanical wave function, and E the total energy of the system. The variational principle tells us that for any given choice of trial wave function, ψ_T, the total energy

$$E = \frac{\langle\psi_T|\,\widehat{\mathbf{H}}\,|\psi_T\rangle}{\langle\psi_T|\psi_T\rangle} \tag{4.16}$$

will always be larger or equal to the true ground state energy, E_0, for the chosen Hamiltonian $\widehat{\mathbf{H}}$,

$$E_0 \leq \frac{\langle\psi_T|\,\widehat{\mathbf{H}}\,|\psi_T\rangle}{\langle\psi_T|\psi_T\rangle}. \tag{4.17}$$

To show this, we exploit the fact that $\widehat{\mathbf{H}}$ is Hermitian, and that the Hamiltonian will have a set of exact eigenstates forming a complete basis set, $|\Psi_i\rangle$. Our trial function can then be expanded as

$$\langle\psi_T| = \sum_i c_i\,|\Psi_i\rangle\,, \tag{4.18}$$

and the expectation value for $\widehat{\mathbf{H}}$ using the trial function ψ_T will be

$$\begin{aligned} E &= \frac{\langle\psi_T|\,\widehat{\mathbf{H}}\,|\psi_T\rangle}{\langle\psi_T|\psi_T\rangle} = \frac{\sum_{ij} c_i^* c_j\,\langle\Psi_i|\,\widehat{\mathbf{H}}\,|\Psi_j\rangle}{\sum_{ij} c_i^* c_j\,\langle\Psi_i|\Psi_j\rangle} \\ &= \frac{|c_i|^2 E_i\,\langle\Psi_i|\Psi_i\rangle}{|c_i|^2\,\langle\Psi_i|\Psi_i\rangle} = \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2}. \end{aligned} \tag{4.19}$$

The crucial point is to know that every excited state energy, E_i, is greater or equal to the ground state energy, E_0, that is

$$E = \frac{|c_i|^2 E_i\,\langle\Psi_i|\Psi_i\rangle}{|c_i|^2\,\langle\Psi_i|\Psi_i\rangle} = \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2} \geq \frac{|c_i|^2 E_0\,\langle\Psi_i|\Psi_i\rangle}{|c_i|^2\,\langle\Psi_i|\Psi_i\rangle} = \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2} = E_0. \tag{4.20}$$

This inequality gives an equality when you find the true ground state wave function, Ψ_0, that is

$$E_0 = \langle\Psi_0|\,\widehat{\mathbf{H}}\,|\Psi_0\rangle\,. \tag{4.21}$$

As our main goal is to approximate the ground state energy, this is the basis of our calculations. This gives us an idea to introduce variational parameters to our trial wave function in search for a good approximation. Instead of just having the wave functions depending on the spatial coordinates as $\psi_T = \psi_T(x, y, z)$, we can write ψ_T as

$$\psi_T = \psi_T(x, y, z, \alpha, \beta, \dots). \tag{4.22}$$

Our goal is to find the minimum of functional E in Eq. (4.16), viz. the set of parameters that gives the lowest energy when calculating E. A common approach in VMC is too also search for the minimum of the variance. *A priori*, we know that when the variance is exactly zero, we have found the true ground state. However, as we have no knowledge of the true form of the wave function, we must make an ansatz for it. This ansatz will in general not have a minimum for both the energy and the variance for the same set of parameters. As mentioned in [14], the most efficient wave function is found if we use a linear combination of the parameters that minimize the energy and the variance respectively.

In this work however, I have focused on finding the minimum of the energy only.

4.4 Variational Monte Carlo

The main exercise with the Variational Monte Carlo process for atoms is to move the particles in the system guided by a probability distribution and sample the energy at these states in order to calculate various expectation values (mainly $\langle E\rangle$ and $\langle E^2\rangle$). In our case we have a known Hamiltonian, $\widehat{\mathbf{H}}$, and a chosen many-particle trial wave function, ψ_T. The expectation value of the Hamiltonian is

$$\langle H\rangle = \frac{\int d\mathbf{R}\psi_T^*(\mathbf{R};\alpha)\widehat{\mathbf{H}}(\mathbf{R})\psi_T(\mathbf{R};\alpha)}{\int d\mathbf{R}\psi_T^*(\mathbf{R};\alpha)\psi_T(\mathbf{R};\alpha)}, \tag{4.23}$$

where $\mathbf{R} = (\mathbf{R_1}, \mathbf{R_2}, \ldots, \mathbf{R_N})$ are all the coordinates for the N particles, and α is the set of all the variational parameters in question. The general procedure will now be

1. Construct a trial wave function, ψ_T as a function of the N particles' coordinates and the chosen variational parameters
2. Calculate $\langle H\rangle$ using Eq. (4.23)
3. Vary the parameters according to some minimization technique

We will first focus on the second item in the list above, with a focus on the Metropolis algorithm.

4.4.1 VMC and the simple Metropolis algorithm

We assume first that our trial wave function, ψ_T, is not normalized so the quantum mechanical probability distribution is then given as

$$P(\mathbf{R};\alpha) = \frac{|\psi_T(\mathbf{R};\alpha)|^2}{\int |\psi_T(\mathbf{R};\alpha)|^2 d\mathbf{R}}. \tag{4.24}$$

Equation (4.24) gives the PDF for our system. We now define a new operator, called the local energy operator as

$$\widehat{\mathbf{E}}_L(\mathbf{R};\alpha) = \frac{1}{\psi_T}\widehat{\mathbf{H}}\psi_T(\mathbf{R};\alpha), \tag{4.25}$$

and the expectation value of the local energy is given as

$$\langle E_L(\alpha)\rangle = \int P(\mathbf{R};\alpha)\widehat{\mathbf{E}}_L(\mathbf{R};\alpha)d\mathbf{R}. \tag{4.26}$$

The dimensionless Hamiltonian in section 3.1.4 is given as

$$\widehat{\mathbf{H}} = -\frac{1}{2}\sum_{i}^{N}\nabla_i^2 - \sum_{i}^{N}\frac{1}{r_i} + \sum_{i<j}\frac{1}{r_{ij}}. \tag{4.27}$$

Inserting the Hamiltonian from Eq. (4.27) into Eq. (4.25), the local energy operator will be

$$\widehat{\mathbf{E}}_L = \frac{1}{\psi_T(\mathbf{R},\alpha)}\left(-\frac{1}{2}\sum_{i}^{N}\nabla_i^2 - \sum_{i}^{N}\frac{1}{r_i} + \sum_{i<j}\frac{1}{r_{ij}}\right)\psi_T(\mathbf{R},\alpha). \tag{4.28}$$

By observing that the last two terms don't affect the wave function in any way, we have

$$\widehat{\mathbf{E}}_L = \frac{1}{\psi_T(\mathbf{R},\alpha)}\left(-\frac{1}{2}\sum_{i}^{N}\nabla_i^2\right)\psi_T(\mathbf{R},\alpha) - \sum_{i}^{N}\frac{1}{r_i} + \sum_{i<j}\frac{1}{r_{ij}}. \tag{4.29}$$

This will be quantity we sample in each Monte Carlo cycle, following our PDF from Eq. (4.24). The more we improve the trial wave function the closer the expectation value of the local energy, $\langle E_L\rangle$, gets to the exact energy E.

As we sample the local energy E_L we also sample E_L^2 in order to get $\langle E_L^2\rangle$ for the variance

$$\sigma_{E_L}^2 = \langle E_L^2\rangle - \langle E_L\rangle^2. \tag{4.30}$$

We see from appendix B that we now can calculate the

$$\langle E_L\rangle = \frac{1}{n}\sum_{k=1}^{n}E_L(x_k), \tag{4.31}$$

where x_k are the points at which the local energy is sampled, and n are the number of sample points.

An important note is that in Metropolis algorithm only involves ratios between probabilities, so the denominator in Eq. (4.24) actually never needs to be calculated. We will see this in the next section.

The Metropolis algorithm

The Metropolis algorithm uses ratios between probabilities to determine whether or not a chosen particle is to be moved to a proposed position or not. When an electron is moved, the set of positions $\mathbf{R}$ change to positions $\mathbf{R}'$. The ratio $w = P(\mathbf{R}')/P(\mathbf{R})$ is now the transition probability from the state with particles being in positions $\mathbf{R}$ to a state where particles are in positions $\mathbf{R}'$. This ratio w will now be

$$w = \left|\frac{\psi_T(\mathbf{R}')}{\psi_T(\mathbf{R})}\right|^2. \tag{4.32}$$

The Metropolis algorithm tells us that if

$$w > 1, \tag{4.33}$$

we automatically accept the new positions. If $w < 1$, we compare w with a random number r, with $r \in (0,1)$. If

$$r \leq w, \tag{4.34}$$

we also accept the new positions. If neither of the two inequalities are true, the chosen electron remains at its position before the proposed move. The most effective way to use this method is to move one particle at a time before running the tests for w.

We now have enough information to describe the elements needed to calculate the expectation value for the local energy using the Metropolis algorithm. The specifics of the implementation will be discussed in the next chapter, but the following points still show how it's done.

1. Set all electrons in random positions initially, that is a set of positions $\mathbf{R}$
2. Start a Monte Carlo cycle and a loop over electrons
3. Calculate the new position,
$$\mathbf{R'}_i = \mathbf{R}_i + s * r,$$
for electron i using a uniform random number, r and a finite stepsize, s. This will change the entire set of positions from $\mathbf{R}$ to $\mathbf{R}$.
4. Calculate the ratio
$$w = |\psi_T(\mathbf{R'})/\psi_T(\mathbf{R})|^2,$$
5. Use the checks in Eqs. (4.33) and (4.34) to decide whether to accept new positions, $\mathbf{R} = \mathbf{R'}$, or to stay in initial positions $\mathbf{R}$.
6. If accepted, update positions
7. Repeat steps 3 through 6 for each electron
8. After looping over all particles, calculate the local energy and update $\langle E_L \rangle$ and $\langle E_L^2 \rangle$.
9. Repeat steps 3 through 8 for the chosen number of Monte Carlo cycles

The transition rules being used here is called a uniform symmetrical transition rule, and it is common to keep the acceptance ratio around 0.5 (this is achieved by choosing the step length $s \approx 1$). The acceptance ratio is the number of accepted steps divided by total number of particle moves. However, this simple Metropolis algorithm does not seem very efficient, due to the way the new positions are calculated. They depend solely on the previous position and not on any guiding mechanism. In order to make this method more efficient, we introduce the Metropolis-Hastings algorithm with so-called importance sampling. With importance sampling we will get a much higher acceptance ratio since the calculation of new positions and the Metropolis-test will depend on the gradient of the wave function. Figure 4.1 shows a float chart of the Metropolis method.

4.4.2 Metropolis-Hastings algorithm and importance sampling

As an improvement to our Metropolis algorithm, we introduce the Metropolis-Hastings algorithm, which involves importance sampling. A new term that has to be implemented is the **quantum force** of the system:

$$F = 2\frac{1}{\psi_T}\nabla\psi_T, \tag{4.35}$$

A physical interpretation of the quantum force, and therefore a good reason to implement it, is that the gradient of the wavefunction tells us which direction the electrons are

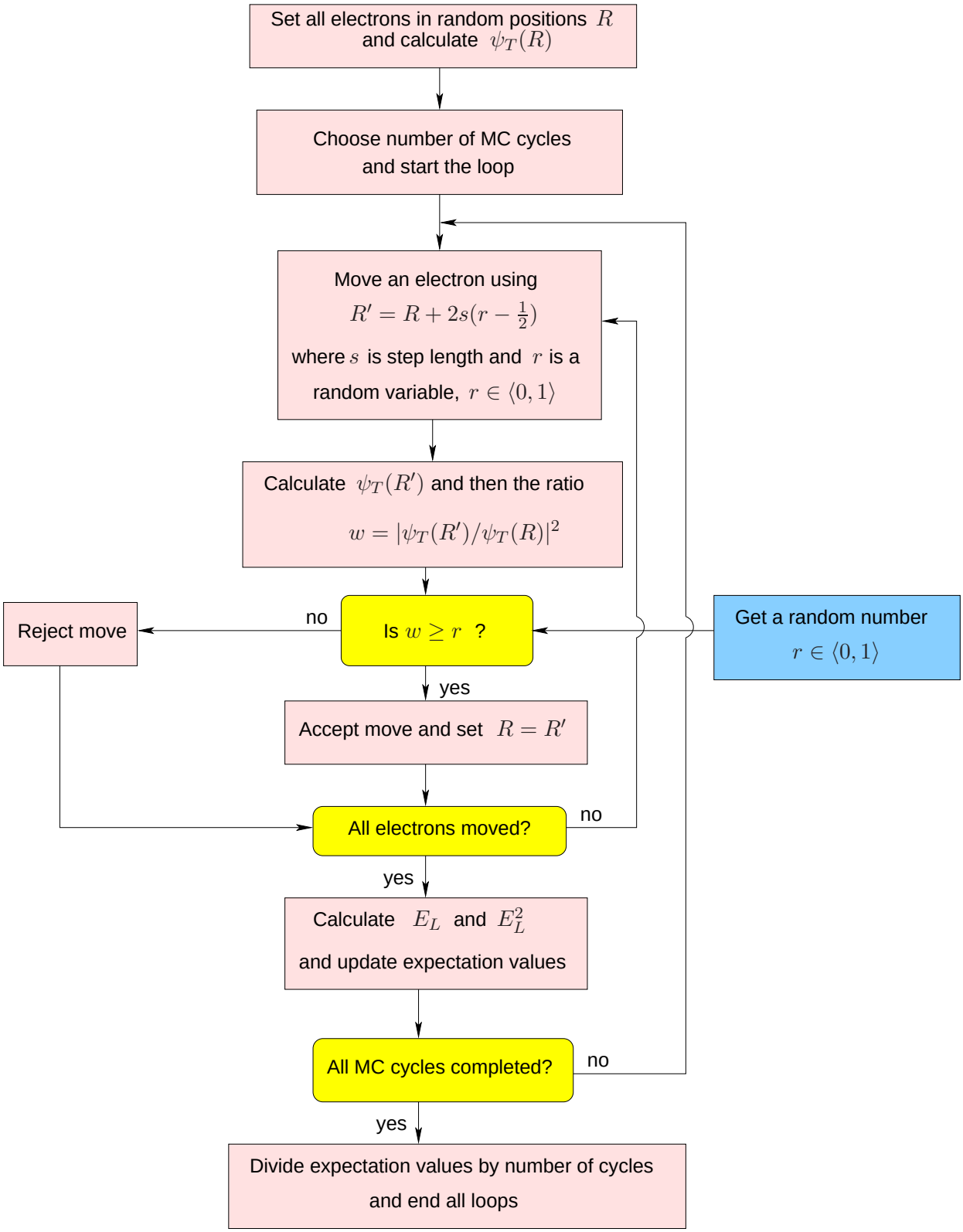

Figure 4.1: A float chart describing the Metropolis algorithm

moving. If we rather use the quantum force to steer us in the direction of the electron movement, rather than sampling random points around the nucleus, the points we sample will be more relevant and therefore our acceptance ratio increases to about 0.8-0.9. Despite the fact that computing the quantum force will increase the calculation time for each cycle, this will in time decrease our thermalization period, the period for the system to reach a most likely state. This will result in a more efficient algorithm.

Importance sampling

The one-dimensional diffusion equation is written as

$$\frac{\partial P(x,t)}{\partial t} = D\frac{\partial^2 P(x,t)}{\partial x^2}, \tag{4.36}$$

and describes how the probability distribution, P, of a random walker evolves in time. D is the diffusion constant (see chapter 14 in [9] or chapter 1 in [3]). The Green's function

$$G(y,x,t) = \frac{1}{(4\pi D\Delta t)^{3N/2}} \exp\left(-(y-x)^2/4Dt\right), \tag{4.37}$$

is a general solution to the diffusion equation for a single walker which starts off at position x at time $t = 0$. By using discrete time steps, it can be shown that a Markov process models a diffusion process, as the transition probability in a time step Δt from position x to position y, $G(y,x,\Delta t)$, only depends on the previous position, or state. The Fokker-Planck equation:

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - F\right)P(x,t), \tag{4.38}$$

describes the time evolution of a PDF. As opposed to the diffusion equation, the Fokker-Planck equation also includes a drift term, F, yielding a solution given as

$$G(y,x,\Delta t) = \frac{1}{(4\pi D\Delta t)^{3N/2}} \exp\left(-(y-x-D\Delta tF(x))^2/4D\Delta t\right). \tag{4.39}$$

This is very similar to the solution of the diffusion equation but includes the drift term, or external force, F. This can be interpreted as a probability distribution of a single random walker with a drift force term starting off at position x at time $t = 0$([15]). This Green's function serves as a transition probability for an electron when it is under influence of the drift term, F. This new propability must be multiplied with the probability from Eq. (4.24) and the ratio of probabilities from Eq. 4.32 is then replaced by the modified ratio

$$w = \frac{G(x,y,\Delta t)|\psi_T(y)|^2}{G(y,x,\Delta t)|\psi_T(x)|^2}, \tag{4.40}$$

or more explicitly

$$w = \frac{|\psi_T(y)|^2}{|\psi_T(x)|^2} \exp\left(-(y-x-D\Delta tF(x))^2 - (x-y-D\Delta tF(y))^2)\right). \tag{4.41}$$

The Langevin equation (see [3]) is a stochastic differential equation describing the dynamics of random walkers as

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta, \tag{4.42}$$

where η is a random variable following the uniform distribution between 0 and 1. Integrating this equation using the concepts of stochastic integration, will yield the equation

$$y = x + DF(x)\Delta t + \xi\sqrt{\Delta t}, \tag{4.43}$$

where ξ now is a Gaussian distributed random variable and Δt is a chosen timestep. This equation replaces Eq. (3) as the equation that calculates the new position for a particle. We see that this depends on the drift term, the quantum force, $F(x)$. This drift term "pushes" the walkers in the direction of more probable states because the gradient of our PDF will point in the direction of higher values of probability.

By using the Fokker-Planck equation (where P is a time-dependent probability density),

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - F\right)P(x,t), \tag{4.44}$$

and restraining a system to be in a stationary state, that is $\partial P/\partial t = 0$, we find the form of F as

$$F(x) = \frac{1}{P}\frac{dP}{dx}, \tag{4.45}$$

or

$$F(\mathbf{r}) = \frac{1}{P}\nabla P \tag{4.46}$$

in three dimension, giving us a constraint for the form of the quantum force. Our quantum mechanical system has a probability distribution given as

$$P(\mathbf{r}) = |\psi_T(\mathbf{r})|^2, \tag{4.47}$$

and the quantum force of our system will be

$$F(x) = 2\frac{1}{\psi_T}\nabla\psi_T. \tag{4.48}$$

This algorithm will speed up the process as more important or relevant states are sampled (or visited in state space), hence the name, importance sampling.

Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm can be summarized in the same way as for the simple Metropolis algorithm:

1. Set all electrons in random positions as an initial point, x
2. Start a Monte Carlo cycle and a loop over electrons
3. Calculate the new position,

   $$y_i = x_i + DF(x)\Delta t + \xi\sqrt{\Delta t}$$

   for the chosen electron i using a Gaussian distributed number, ξ and a chosen timestep, Δt
4. Calculate the quantum force for the new set of positions, y, and then the Green's function

5. Calculate the ratio w from equation Eq. (4.41)
6. Use the checks in Eqs. (4.33) and (4.34) to decide whether to accept new positions, y, or to stay in initial positions x
7. If accepted, update positions and quantum force
8. Repeat steps 3 through 7 for each electron
9. After looping over all particles, calculate the local energy and update $\langle E_L \rangle$ and $\langle E_L^2 \rangle$.
10. Repeat steps 3 through 9 for the chosen number of Monte Carlo cycles

Figure 4.2 shows a float chart of the Metropolis-Hastings algorithm. The detailed implementation as a computer program will be discussed in the next chapter.

Both the simple Metropolis and the Metropolis-Hastings algorithms demand $4N$ random numbers per Monte Carlo cycle, where N is the number of particles. Per particle move, we need 3 normal distributed random numbers to make a move in each of the spatial dimensions. We also need one uniformly distributed random number for the Metropolis-test itself.

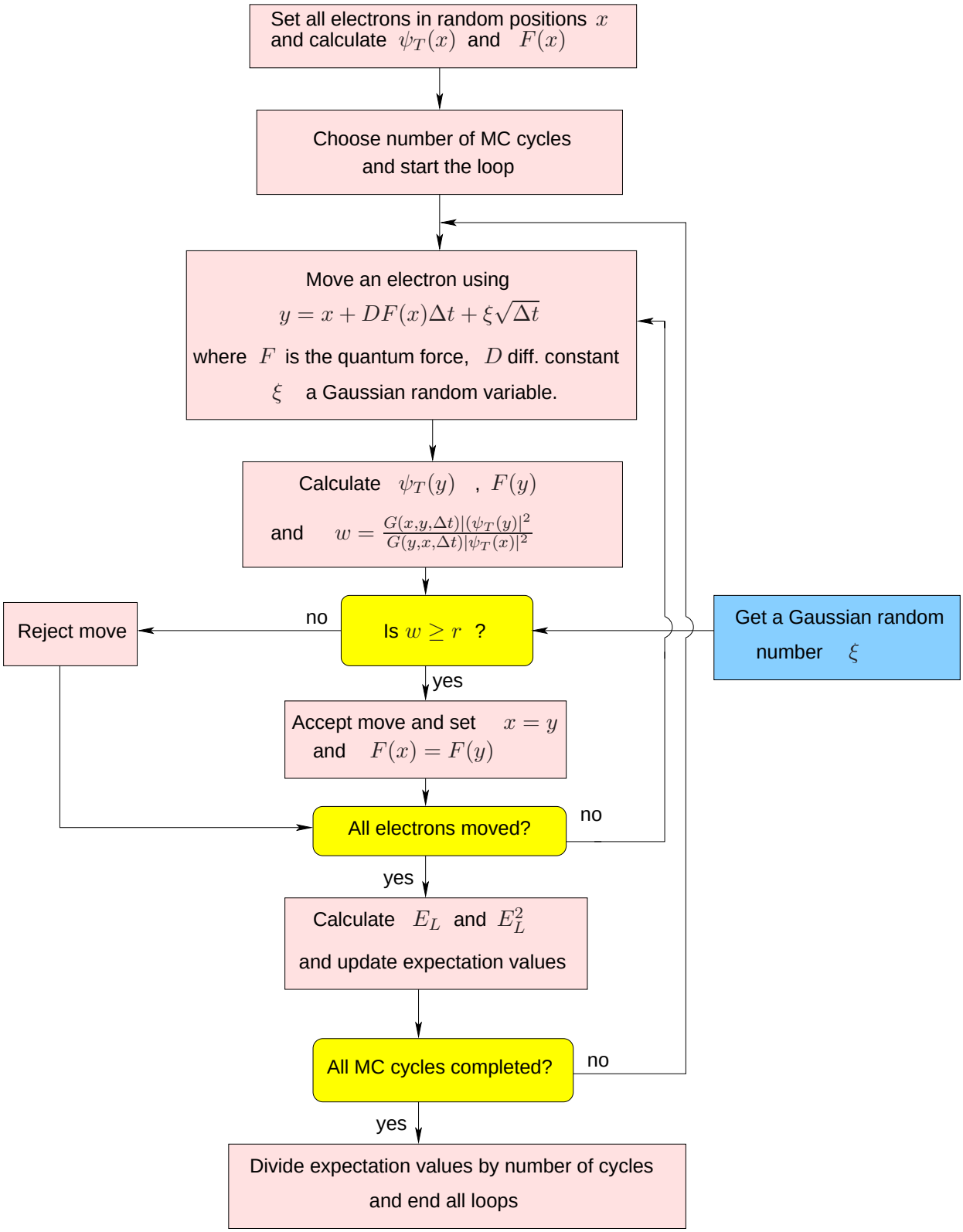

Figure 4.2: A float chart describing the Metropolis-Hastings algorithm

Chapter 5

Wave Functions

As the expectation values calculated with VMC depend on the chosen trial function, the results rely heavily on an educated guess. In this chapter we will examine the wave functions used for the Variational Monte Carlo method, and why they are good choices. This chapter will also introduce a correlation part, the Jastrow factor, to our trial wave function, since a pure Slater determinant is not enough to describe the inter-electronic correlations in an atomic system.

5.1 Cusp conditions and the Jastrow factor

The **cusp conditions** are constraints we get on the wave function when considering our system when distances between an electron and the core and between the electrons themselves become very small. The problems lie with the Coulomb force which depends on the inverse of the distance between charged particles. When these distances become small, we need some way to cancel the divergences that occur.

5.1.1 Single particle cusp conditions

In order to find the cusp conditions for the single particle wave functions, we start by examining the radial Schrödinger equation

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr}-\frac{l(l+1)}{r^2}\right)R(r)+V(r)R(r)=ER(r), \tag{5.1}$$

where $R(r)$ is the radial part of the full wave function. The dimensionless radial equation can be written as

$$\left(\frac{d^2}{d\rho^2}+\frac{2}{\rho}\frac{d}{d\rho}+\frac{2Z}{\rho}-\frac{l(l+1)}{\rho^2}+2E\right)R(\rho)=0, \tag{5.2}$$

with Z being the nuclear charge of the core. We now introduce

$$R(\rho)=\frac{1}{\rho}u(\rho) \tag{5.3}$$

as our radial wave function. This will give the radial equation as

$$\frac{d^2u(\rho)}{d\rho^2}+\frac{2Z}{\rho}u(\rho)-\frac{l(l+1)}{\rho^2}u(\rho)+2Eu(\rho)=0. \tag{5.4}$$

As ρ approaches zero, the terms that dominate will be

$$\frac{d^2u(\rho)}{d\rho^2} - \frac{l(l+1)}{\rho^2}u(\rho) = 0. \tag{5.5}$$

The general solution of this differential equation is

$$u(\rho) = C\rho^{l+1} + D\rho^{-l}. \tag{5.6}$$

As ρ^{-l} will blow up as ρ approaches zero, the only physical solution we can have is then

$$u(\rho) \sim \rho^{l+1}. \tag{5.7}$$

In general we must assume that

$$R(\rho) = \rho^{l+1}\gamma(\rho), \tag{5.8}$$

where $\gamma(\rho)$ is just some radial dependence function of ρ with the constraint that $\gamma(0) \neq 0$. For the kinetic energy not to diverge, we also need

$$\lim_{\rho\to 0} \frac{d^2\gamma(\rho)}{d\rho^2} = \Omega, \tag{5.9}$$

where Ω is finite. The dimensionless radial equation is now

$$\left(\frac{d^2}{d\rho^2} + \frac{2}{\rho}\frac{d}{d\rho} + \frac{2Z}{\rho} - \frac{l(l+1)}{\rho^2} + 2E\right)\rho^l\gamma(\rho) = 0, \tag{5.10}$$

By finding the derivatives

$$\frac{d^2}{d\rho^2}\left[\rho^l\gamma(\rho)\right] = (l-1)l\rho^{l-2}\gamma(\rho) + 2l\rho^{l-1}\frac{d\gamma(\rho)}{d\rho} + \rho^l\frac{d^2\gamma(\rho)}{d\rho^2}, \tag{5.11}$$

and

$$\frac{2}{\rho}\frac{d}{d\rho}\left[\rho^l\gamma(\rho)\right] = 2l\rho^{l-2}\gamma(\rho) + 2\rho^{l-1}\frac{d\gamma(\rho)}{d\rho}, \tag{5.12}$$

and inserting them into Eq. (5.10) while taking the limit when $\rho \to 0$, we will get

$$\gamma(\rho)\left[\frac{l(l+1)}{\rho^2} + \frac{2Z}{\rho} - \frac{l(l+1)}{\rho^2}\right] + \frac{d\gamma(\rho)}{d\rho}\left[\frac{2(l+1)}{\rho}\right] = 0 \tag{5.13}$$

as the terms that dominate. The second derivative from Eq. (5.9) is finite, as well as the term involving the energy E. By further simplification we will then have the cusp condition as

$$\frac{1}{\gamma(\rho)}\frac{d\gamma(\rho)}{d\rho} = -\frac{Z}{l+1}, \tag{5.14}$$

which indicates the exponential behavior a wave function must have. The solution of equation Eq. (5.14) is

$$\gamma(\rho) \sim e^{-Z\rho/(l+1)}. \tag{5.15}$$

The hydrogenic wave functions and the Roothan-Hartree-Fock orbitals both fulfill the cusp condition.

5.1.2 Correlation cusp conditions

To see how the wave function should behave with respect to correlation, we consider the helium system when the two electrons are very close, but far away from the core, that is, $r_1 \neq 0$, $r_2 \neq 0$ and $r_{12} \to 0$. Now the electron-nucleus terms in the Hamiltonian will be finite, while the repulsive electron-electron potential and the kinetic energy of the two electrons will diverge. When using center of mass coordinates for electron 1 and 2, but neglecting the center of mass motion since $r_{12} \to 0$, we will get a cusp condition also for the correlation part by using much of the same derivations as in section 5.1.1. The cusp condition for the two-particle part, $\mathcal{R}(r_{12})$, will then be

$$\frac{d\mathcal{R}(r_{12})}{dr_{12}} = \frac{1}{2(l+1)}\mathcal{R}(r_{12}), \tag{5.16}$$

as given in [3] and [9].

A correlation function that satisfies this cusp condition is the so-called Padé-Jastrow-function given by the two-particle functions

$$f(r_{ij}) = \exp\left(\frac{r_{ij}}{a(1+\beta r_{ij})}\right), \tag{5.17}$$

where β is a variational parameter. The factor a is a result of whether the intrinsic spins of particles i and j are parallel or not. That is

$$a = \begin{cases} 4 & \text{if spins are parallel,} \\ 2 & \text{if spins are anti-parallel.} \end{cases} \tag{5.18}$$

The full Padé-Jastrow function, J, for a system of N electrons will then be

$$J = \prod_{i<j}^{N} f(r_{ij}) = \prod_{i<j}^{N} \exp\left(\frac{r_{ij}}{a(1+\beta r_{ij})}\right), \tag{5.19}$$

which also satisfies the cusp condition.

5.2 Rewriting the Slater determinant

A general Slater determinant for an N-electron system will take the form:

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, \alpha, \beta, \ldots, \nu) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_\alpha(\mathbf{r}_1) & \psi_\alpha(\mathbf{r}_2) & \cdots & \psi_\alpha(\mathbf{r}_N) \\ \psi_\beta(\mathbf{r}_1) & \psi_\beta(\mathbf{r}_2) & \cdots & \psi_\beta(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_\nu(\mathbf{r}_1) & \psi_\nu(\mathbf{r}_2) & \cdots & \psi_\nu(\mathbf{r}_N) \end{vmatrix}, \tag{5.20}$$

where the ψ's are the available single particle orbitals. We could examine this more specifically by using the beryllium atom as an example. The Slater determinant for beryllium is written as

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \alpha, \beta, \gamma, \delta) = \frac{1}{\sqrt{4!}} \begin{vmatrix} \psi_\alpha(\mathbf{r}_1) & \psi_\alpha(\mathbf{r}_2) & \psi_\alpha(\mathbf{r}_3) & \psi_\alpha(\mathbf{r}_4) \\ \psi_\beta(\mathbf{r}_1) & \psi_\beta(\mathbf{r}_2) & \psi_\beta(\mathbf{r}_3) & \psi_\beta(\mathbf{r}_4) \\ \psi_\gamma(\mathbf{r}_1) & \psi_\gamma(\mathbf{r}_2) & \psi_\gamma(\mathbf{r}_3) & \psi_\gamma(\mathbf{r}_4) \\ \psi_\delta(\mathbf{r}_1) & \psi_\delta(\mathbf{r}_2) & \psi_\delta(\mathbf{r}_3) & \psi_\delta(\mathbf{r}_4) \end{vmatrix}, \tag{5.21}$$

with the quantum numbers representing different states as

$$|\mu\rangle = |n, l, m_l, \chi_\pm\rangle\,, \tag{5.22}$$

with $|\alpha\rangle = |1,0,0\uparrow\rangle$, $|\beta\rangle = |1,0,0\downarrow\rangle$, $|\gamma\rangle = |2,0,0\uparrow\rangle$ and $|\delta\rangle = |2,0,0\downarrow\rangle$. The values $\chi_\pm$ are given as $\chi_+ = \uparrow$ and $\chi_- = \downarrow$ (see section 2.1.2).

By writing out the determinant in Eq. (5.21) explicitly, we would get *zero*, as the spatial wave functions for both spin up and spin down states are identical, and all terms will cancel each other out. We could also try to rewrite the Slater determinant as a product of a spin up-Slater determinant and a spin down-Slater determinant as

$$\begin{aligned}\Phi(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4,\alpha,\beta,\gamma,\delta) &= Det\uparrow(1,2)Det\downarrow(3,4) - Det\uparrow(1,3)Det\downarrow(2,4)\\ -Det\uparrow(1,4)Det\downarrow(3,2) &+ Det\uparrow(2,3)Det\downarrow(1,4) - Det\uparrow(2,4)Det\downarrow(1,3)\\ &+ Det\uparrow(3,4)Det\downarrow(1,2),\end{aligned} \tag{5.23}$$

where we have defined

$$Det\uparrow(1,2) = \frac{1}{\sqrt{2}}\begin{vmatrix}\psi_{100\uparrow}(\mathbf{r}_1) & \psi_{100\uparrow}(\mathbf{r}_2)\\ \psi_{200\uparrow}(\mathbf{r}_1) & \psi_{200\uparrow}(\mathbf{r}_2)\end{vmatrix}, \tag{5.24}$$

as a spin up-determinant, and

$$Det\downarrow(3,4) = \frac{1}{\sqrt{2}}\begin{vmatrix}\psi_{100\downarrow}(\mathbf{r}_3) & \psi_{100\downarrow}(\mathbf{r}_4)\\ \psi_{200\downarrow}(\mathbf{r}_3) & \psi_{200\downarrow}(\mathbf{r}_4)\end{vmatrix}, \tag{5.25}$$

as a spin down-determinant. However, writing out all these smaller determinants, will still give zero, but the strategy of separating the Slater determinant in one spin up and one spin down determinant is still pursued. The Hamiltonian does not depend on spin, so we wish to avoid summation over spin variables. As described in [16], we can actually approximate the Slater determinant for a variational approach with a plain product of a spin up- and a spin down-Slater determinant

$$\Phi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N) \propto Det\uparrow Det\downarrow, \tag{5.26}$$

where the determinants $Det\uparrow$ and $Det\downarrow$ are $(N/2)\times(N/2)$-matrices for even numbered electron systems as examined in this project. We place the first $N/2$ electrons in the spin up determinant and the rest in the spin down determinant. This ansatz for the wave function is not antisymmetric under particle exchange as a wave function should be, but it will give the same expectation value as the correct Slater determinant for such a spin independent Hamiltonian.

5.3 Variational Monte Carlo wave function

By reviewing the last two sections we can write the full trial wave function for or Monte Carlo experiment. It can be written

$$\psi_T = \Phi \cdot J, \tag{5.27}$$

where the Slater determinant Φ can be approximated by

$$\Phi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N) \propto Det\uparrow Det\downarrow, \tag{5.28}$$

and J is

$$J = \prod_{i<j}^{N} \exp\left(\frac{r_{ij}}{a(1+\beta r_{ij})}\right). \tag{5.29}$$

For an N-particle system with $M = N/2$ possible spatial orbitals, the spin up determinant, $Det\uparrow$, can be written

$$Det\uparrow \propto \begin{vmatrix} \psi_{1\uparrow}(\mathbf{r}_1) & \psi_{1\uparrow}(\mathbf{r}_2) & \cdots & \psi_{1\uparrow}(\mathbf{r}_{N/2}) \\ \psi_{2\uparrow}(\mathbf{r}_1) & \psi_{2\uparrow}(\mathbf{r}_2) & \cdots & \psi_{2\uparrow}(\mathbf{r}_{N/2}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{M\uparrow}(\mathbf{r}_1) & \psi_{M\uparrow}(\mathbf{r}_2) & \cdots & \psi_{M\uparrow}(\mathbf{r}_{N/2}) \end{vmatrix}, \tag{5.30}$$

while the spin down determinant can be written

$$Det\downarrow \propto \begin{vmatrix} \psi_{1\downarrow}(\mathbf{r}_{N/2+1}) & \psi_{1\downarrow}(\mathbf{r}_{N/2+2}) & \cdots & \psi_{1\downarrow}(\mathbf{r}_N) \\ \psi_{2\downarrow}(\mathbf{r}_{N/2+1}) & \psi_{2\downarrow}(\mathbf{r}_{N/2+2}) & \cdots & \psi_{2\downarrow}(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{M\downarrow}(\mathbf{r}_{N/2+1}) & \psi_{M\downarrow}(\mathbf{r}_{N/2+2}) & \cdots & \psi_{M\downarrow}(\mathbf{r}_N) \end{vmatrix}. \tag{5.31}$$

These are the expressions needed for our trial wave function. The next two sections describe the single particle orbitals, the ψ's, we use in our Slater determinant.

5.4 Orbitals for VMC

In order to construct a trial Slater determinant for the Variational Monte Carlo method, our first choice is to use the hydrogenic single particle functions obtained when solving the Schrödinger equation for hydrogen. These functions comply with the cusp conditions described in section 5.1.1.

5.4.1 S-orbitals

An s-orbital is a spatial wave function with quantum number $l = 0$. Introducing the variational parameter α, the s-orbitals needed for calculating up to at least silicon, with $N = 14$ electrons, are $1s$, $2s$ and $3s$. Here we only need the radial part of the wave functions because the constant angular parts factor out in the Metropolis algorithm (see section 4.4.1). The spherical harmonics for $l = 0$ is

$$Y_0^0 = \left(\frac{1}{4\pi}\right)^{1/2}. \tag{5.32}$$

We now don't have to consider changing from spherical to Cartesian coordinates because the expression will remain the same using the simple relation $r = \sqrt{x^2+y^2+z^2}$. The hydrogenic radial functions are found in table 4.7 in [2] where non-vanishing factors(cf. Metropolis algorithm) $1/a$ have been replaced by α.

The $1s$ wave function

For the $1s$ function we only need the exponential function

$$\phi_{1s} = e^{-r\alpha}, \tag{5.33}$$

with the **first derivative** with respect to an arbitrary coordinate, x_i, given as:

$$\frac{\partial \phi_{1s}}{\partial x_i} = -\frac{\alpha x_i}{r} e^{-\alpha r}. \tag{5.34}$$

The Laplacian, $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$, needed for the kinetic energy is given as

$$\nabla^2 \phi_{1s} = \frac{\alpha}{r} (\alpha r - 2) e^{-\alpha r}. \tag{5.35}$$

The $2s$ wave function

The $2s$ wave function is given by

$$\phi_{2s} = \left(1 - \frac{\alpha r}{2}\right) e^{-\alpha r/2}, \tag{5.36}$$

with the **first derivative** with respect to. x_i is

$$\frac{\partial \phi_{2s}}{\partial x_i} = \frac{\alpha x_i}{2r} \left(2 - \frac{\alpha r}{2}\right) e^{-\alpha r/2}. \tag{5.37}$$

The **second derivative** of the $2s$ function gives the Laplacian of the function:

$$\nabla^2 \phi_{2s} = \frac{\alpha}{4r} \left(5\alpha r - \frac{\alpha^2 r^2}{2} - 8\right) e^{-\alpha r/2}. \tag{5.38}$$

The $3s$ wave function

The final s-function is the $3s$ wave function. It is

$$\phi_{3s} = \left(1 - \frac{2\alpha r}{3} + \frac{2\alpha^2 r^2}{27}\right) e^{-\alpha r/3}. \tag{5.39}$$

The **first derivative** is

$$\frac{\partial \phi_{3s}}{\partial x_i} = \left(-1 + \frac{10\alpha r}{27} - \frac{2\alpha^2 r^2}{81}\right) \alpha x_i e^{-\alpha r/3}, \tag{5.40}$$

while the **second derivative** summed up in all three dimensions is

$$\nabla^2 \phi_{3s} = \left(-2 + \frac{13\alpha r}{9} - \frac{2\alpha^2 r^2}{9} + \frac{2\alpha^3 r^3}{243}\right) \frac{\alpha}{r} e^{-\alpha r/3}. \tag{5.41}$$

5.4.2 P-orbitals

The p-orbitals are the single particle functions which have **azimuthal quantum number** $l = 1$. Now the allowed values for the **magnetic quantum number** are $m = -l, -l+1, \ldots, l-1, l$, i.e. $m = -1, 0, 1$. This is only possible for functions with quantum numbers $n = 1, 2, 3, \ldots$. In this thesis we only need s- and p-orbitals.

The angular parts for any p-orbital in spherical coordinates are

$$Y_1^0 = \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta, \tag{5.42}$$

$$Y_1^{\pm 1} = \pm \left(\frac{3}{8\pi}\right)^{1/2} \sin\theta e^{\pm i\varphi}. \tag{5.43}$$

In my program however, I am working with real wave functions and Cartesian coordinates. This is taken care of by introducing real solid harmonics, S_{lm}, as described in ref. [6]. The single particle wave functions, ψ_{nlm}, are written on a separable form with spherical coordinates as

$$\psi(r,\theta,\varphi)_{nlm} = R_{nl}(r)Y_{lm}(\theta,\varphi). \tag{5.44}$$

These are general solutions of the one-electron system in a central symmetric potential field. Our radial functions are on the form $R_{nl}(r) = r^l\mathcal{R}_{nl}(r)$. The solid harmonics, $\mathcal{Y}$, are related to the spherical harmonics by

$$\mathcal{Y}_{lm}(r,\theta,\varphi) = r^l Y_{lm}(\theta,\varphi). \tag{5.45}$$

We can then write the wave function as

$$\psi(r,\theta,\varphi)_{nlm} = R_{nl}(r)Y_{lm}(\theta,\varphi) = \mathcal{R}_{nl}(r)\mathcal{Y}_{lm}(r,\theta,\varphi). \tag{5.46}$$

Now there is an r-dependence in the angular part. We will only need the real-valued solid harmonics, S_{lm}, related to the solid harmonics by

$$S_{l,0} = \sqrt{\frac{4\pi}{2l+1}}\mathcal{Y}_{l,0}, \tag{5.47}$$

$$S_{l,m} = (-1)^m\sqrt{\frac{8\pi}{2l+1}}\text{Re }\mathcal{Y}_{l,m}, \tag{5.48}$$

and

$$S_{l,-m} = (-1)^m\sqrt{\frac{8\pi}{2l+1}}\text{Im }\mathcal{Y}_{l,m}. \tag{5.49}$$

The real solid harmonics needed for us are simply $S_{1,1} = x$, $S_{1,0} = z$ and $S_{1,-1} = y$. As mentioned earlier, the Metropolis algorithm will take care of any multiplicative factors in front of the wave functions, so we only need to consider one single spatial variable from the angular part of the wave function.

The $2p$ wave function

The expression for the $2p$ wave function in spherical coordinates is proportional to

$$R_{21} \propto re^{-\alpha r/2}. \tag{5.50}$$

By reviewing Eq. (5.45) and using the results from the real solid harmonics, we see the factor $r^l = r$ is replaced by a coordinate x_i. That is, $x_i = x, y, z$ depending on the quantum number m. The general case is then

$$\phi_{2p}(m) = x_i(m)e^{-\alpha r/2}. \tag{5.51}$$

with $x_i(1) = x$, $x_i(0) = z$ and $x_i(-1) = y$.

The first derivative depends on whether the coordinate being differentiated with respect to matches the coordinate x_i of the wave function or not. If these are the same, e.g. $m = 1 \rightarrow x_i = x$, the **first derivative** is

$$\frac{\partial \phi_{2p}}{\partial x} = \left(1 - \frac{\alpha x^2}{2r}\right) \frac{e^{-\alpha r/2}}{r}. \tag{5.52}$$

The expressions are analogous for y and z also.

If these coordinates do not match, e.g. $\phi_{2p}(1) = x \exp -\alpha r/2$ and the derivative is with respect to y, the expression is

$$\frac{\partial \phi_{2p}(1)}{\partial y} = -\frac{\alpha x y}{r} e^{-\alpha r/2}. \tag{5.53}$$

As for the first derivative, the second derivative also depends on which coordinate we differentiate with respect to. If the coordinates are the same, e.g. $m = 1 \rightarrow x_i = x$, the **second derivative** is

$$\frac{\partial^2 \phi_{2p}}{\partial x^2} = \frac{\alpha}{r} \left(\frac{x^3}{2r}\left(\frac{1}{r} + \frac{\alpha}{2}\right) - \frac{3x}{2}\right) e^{-\alpha r/2}, \tag{5.54}$$

while for different coordinates, e.g. $\phi_{2p}(1) = x \exp -\alpha r/2$ and the **second derivative** is with respect to y, the expression is

$$\frac{\partial^2 \phi_{2p}(1)}{\partial y^2} = \frac{\alpha x}{2r^3} \left(\frac{\alpha r y^2}{2} - r^2 + y^2\right) e^{-\alpha r/2}. \tag{5.55}$$

The final Laplacian is a sum of Eqs. (5.54) and (5.55) depending on which l-state is being used.

The $3p$ wave function

The $3p$ radial wave function is proportional to

$$R_{31} \propto \left(1 - \frac{r\alpha}{6}\right) \alpha r e^{-\alpha r/3}. \tag{5.56}$$

By using the real solid harmonics for the full single particle wave function, we get our expression for the $3p$ wave function;

$$\phi_{3p}(m) = x_i(m) \left(1 - \frac{\alpha r}{6}\right) \alpha e^{-\alpha r/3}, \tag{5.57}$$

with $x_i(1) = x$, $x_i(0) = z$ and $x_i(-1) = y$. The **first derivative** when differentiating with respect to. to the same coordinate x_i, e.g. x, is

$$\frac{\partial \phi_{3p}(1)}{\partial x} = \left(-\left(1 - \frac{\alpha r}{6}\right)\left(\frac{\alpha^2 x^2}{3r} + \alpha\right) - \frac{\alpha^2 x^2}{6r}\right) e^{-\alpha r/3}, \tag{5.58}$$

while for different coordinates, e.g. x and y, it is

$$\frac{\partial \phi_{3p}(1)}{\partial y} = \left(-\left(1 - \frac{\alpha r}{6}\right) \frac{\alpha^2 y x}{3r} - \frac{\alpha^2 x y}{6r}\right) e^{-\alpha r/3}, \tag{5.59}$$

and

$$\frac{\partial\phi_{3p}(1)}{\partial z} = \left(-\left(1-\frac{\alpha r}{6}\right)\frac{\alpha^2 zx}{3r} - \frac{\alpha^2 zx}{6r}\right)e^{-\alpha r/3}. \tag{5.60}$$

The **second derivative** with respect to the coordinate from the solid harmonic, e.g. x, is

$$\frac{\partial^2\phi_{3p}(1)}{\partial x^2} = \left(\left(1-\frac{\alpha r}{6}\right)\left(\frac{x^2}{3r^2}+\frac{\alpha x^2}{9r}-1\right)+\frac{\alpha x^2}{9r}+\frac{x^2}{6r^2}-\frac{1}{2}\right)\frac{\alpha^2 x}{r}e^{-\alpha r/3}, \tag{5.61}$$

while the **second derivative** with respect to a different coordinate is

$$\frac{\partial^2\phi_{3p}(1)}{\partial y^2} = \left(\left(1-\frac{\alpha r}{6}\right)\left(\frac{y^2}{3r^2}+\frac{\alpha y^2}{9r}-\frac{1}{3}\right)+\frac{\alpha y^2}{9r}+\frac{y^2}{6r^2}-\frac{1}{6}\right)\frac{\alpha^2 x}{r}e^{-\alpha r/3}, \tag{5.62}$$

and

$$\frac{\partial^2\phi_{3p}(1)}{\partial z^2} = \left(\left(1-\frac{\alpha r}{6}\right)\left(\frac{z^2}{3r^2}+\frac{\alpha z^2}{9r}-\frac{1}{3}\right)+\frac{\alpha z^2}{9r}+\frac{z^2}{6r^2}-\frac{1}{6}\right)\frac{\alpha^2 x}{r}e^{-\alpha r/3}. \tag{5.63}$$

By combining these we get the Laplacian of the wave function, $\nabla^2\phi_{3p}(1)$. The expression for this particular single partice wave function is

$$\nabla^2\phi_{3p}(1) = \frac{\partial^2\phi_{3p}(1)}{\partial x^2}+\frac{\partial^2\phi_{3p}(1)}{\partial y^2}+\frac{\partial^2\phi_{3p}(1)}{\partial z^2}. \tag{5.64}$$

These are all the explicit expressions needed for the Slater determinant with hydrogen-like wave functions.

5.5 Roothaan-Hartree-Fock with Slater-type orbitals

Another approach for the single particle basis is to use the single particle wave functions obtained using the Hartree-Fock method discussed in section 3.2.1. The Slater determinant is of the form

$$\Phi = \frac{1}{\sqrt{N!}}\begin{vmatrix} \phi_1(r_1) & \phi_1(r_2) & \cdots & \phi_1(r_N) \\ \phi_2(r_1) & \phi_2(r_2) & \cdots & \phi_2(r_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(r_1) & \phi_N(r_2) & \cdots & \phi_N(r_N) \end{vmatrix}, \tag{5.65}$$

where the functions ϕ are the single particle states. The Roothaan-Hartree-Fock approach (see [10]) is to choose *Slater type orbitals* (STO) as the basis functions in our wave function expansion. The STO are discussed in [6] and in [10]. In this work I have used the exact form from [10]. These are given as

$$\phi_{i\lambda\alpha} = \sum_p \psi_{p\lambda\alpha}C_{i\lambda p}, \tag{5.66}$$

where λ corresponds to quantum number l. The index i refers to the i^{th} orbital of symmetry λ, and p to the p^{th} orbital of symmetry λ. The index α corresponds to quantum number m. The **Slater functions**, ψ, are on the form

$$\psi_{p\lambda\alpha}(r,\theta,\varphi) = R_{\lambda p}(r)Y_{\lambda\alpha}(\theta,\varphi), \tag{5.67}$$

where

$$R_{\lambda p} = [(2n_{\lambda p})!]^{-1/2}(2\xi_{\lambda p})^{n_{\lambda p}+1/2}\, r^{n_{\lambda p}-1}e^{-\xi_{\lambda p}r}, \tag{5.68}$$

and Y are the normalized spherical harmonics. The solutions of this particular approach are given in appendix A, taken from [10].

By using the idea of real solid harmonics (see section 5.4) and the fact that all basis functions ψ in ϕ will have the same index α, hence the same spherical harmonics, we can ignore the pre-factor from the spherical harmonics (see section 4.4.1). Now the basis functions ψ will go from the form

$$\psi_i = N_i r^{n-1}\exp(-\xi r)Y_{\lambda\alpha}, \tag{5.69}$$

in spherical coordinates, to

$$\psi_i = N_i x_i r^{n-2}\exp(-\xi r), \tag{5.70}$$

in Cartesian coordinates, where the factor N_i is the normalization constant from Eq. (5.68). The results from helium, beryllium, neon, magnesium and silicon are listed in appendix A. As an example on how to construct the single particle orbitals from the tables in appendix A, I use the example given in [10]:

n,λ	Exponent, ξ	$1s$ exp.coeff.	$2s$ exp.coeff.	n,λ	Exponent, ξ	$2p$ exp.coeff.
1S	6.56657	0.19030	0.00754	2P	2.21734	0.21526
1S	4.24927	0.82091	−0.25055	2P	1.00551	0.84052
2S	1.41314	−0.00364	0.87099			
2S	0.87564	0.00251	0.18515			

Table 5.1: The table shows a possible solution for the boron atom.

Consider table 5.1 as a solution of the boron atom. Boron has five electrons distributed as $(1s)^2(2s)^2(2p)$. This means we need $1s$-, $2s$- and $2p$-orbitals. The first column tells us how many and which s-type Slater functions the $1s$- and $2s$-orbitals for the Slater determinant will consist of. In this case it will be two of each type. The second column gives us the factors, ξ, in the exponents of the exponential functions. These will be the same for both $1s$- and $2s$, and will also contribute to the prefactors N_i. The third column gives the coefficients for the $1s$ single particle orbital for the Slater determinant, while the fourth column gives the coefficients for the $2s$ single particle orbital. The single particle orbitals for $1s$ and $2s$ will then only differ by the constants in front of the Slater functions (the ψ's from Eq. (5.66)).

The fifth column indicates that there are only two basis functions needed to describe the p-orbital, namely the $2p$ single particle orbital, while the sixth and seventh give the exponents and $2p$-coefficients respectively.

The single particle orbitals are given in table 5.1 as

$$\phi(1s) = 0.19030\psi_1 + 0.82091\psi_2 - 0.00364\psi_3 + 0.00251\psi_4 \tag{5.71}$$

$$\phi(2s) = 0.00754\psi_1 - 0.25055\psi_2 + 0.87099\psi_3 + 0.18515\psi_4 \tag{5.72}$$

$$\phi(2p) = 0.21526\psi_5 + 0.84052\psi_6, \tag{5.73}$$

where now ψ_1 and ψ_2 are Slater functions of type $1s$, ψ_3 and ψ_4 of type $2s$, and ψ_5 and ψ_6 Slater functions of type $2p$. For larger atoms, the p-orbitals will be linear combinations of different p-type Slater functions. The s- and p-type Slater functions do not mix at any point, seen from Eq. (5.66).

The Slater functions themselves written in Cartesian coordinates are given as

$$\psi_1 = N_1 r^0 \exp(-6.56657r), \tag{5.74}$$
$$\psi_2 = N_2 r^0 \exp(-4.24927r), \tag{5.75}$$
$$\psi_3 = N_3 r^1 \exp(-1.41314r), \tag{5.76}$$
$$\psi_4 = N_4 r^1 \exp(-0.87564r), \tag{5.77}$$
$$\psi_5 = N_5 x_i \exp(-2.21734r), \tag{5.78}$$
$$\psi_6 = N_6 x_i \exp(-1.00551r), \tag{5.79}$$

where x_i corresponds to either x,y or z as described in section 5.4.2, and the factors N_i are the normalization factors from Eq. (5.68).

5.5.1 Derivatives of Slater type orbitals

The analytic expressions needed for the Slater type orbitals are of the form

$$\psi_{1s} \propto \exp(-ar), \tag{5.80}$$
$$\psi_{2s} \propto r \exp(-ar), \tag{5.81}$$
$$\psi_{3s} \propto r^2 \exp(-ar), \tag{5.82}$$
$$\psi_{2p} \propto x_i \exp(-ar), \tag{5.83}$$
$$\psi_{4p} \propto r^2 x_i \exp(-ar), \tag{5.84}$$

where a is some exponent determined by the Roothan-Hartree-Fock calculations and x_i the coordinate from the real solid harmonics. In order to calculate the quantum force (see section 4.4.2) and the kinetic energy we need both the first and second derivatives of the wave functions.

First derivatives

The s-orbitals have first derivatives proportional to

$$\frac{\partial \psi_{1s}}{\partial x_i} \propto -a x_i \exp(-ar), \tag{5.85}$$
$$\frac{\partial \psi_{2s}}{\partial x_i} \propto (1 - ar)\frac{x_i}{r} \exp(-ar), \tag{5.86}$$
$$\frac{\partial \psi_{3s}}{\partial x_i} \propto (2x_i - r a x_i) \exp(-ar), \tag{5.87}$$

while the $2p$-orbitals are

$$\frac{\partial \psi_{2p}(x)}{\partial x} \propto \left(1 - \frac{x^2 a}{r}\right) \exp(-ar), \tag{5.88}$$
$$\frac{\partial \psi_{2p}(x)}{\partial y} \propto -\frac{xya}{r} \exp(-ar), \tag{5.89}$$

where the parenthesis, (x), indicates which real solid harmonics the function depends on. Equations (5.88) and (5.89) are examples with the solid harmonics being the x-coordinate, and the derivatives with respect to x and then y (see the discussion in 5.4.2). The $4p$-orbitals are proportional to

$$\frac{\partial \psi_{4p}(x)}{\partial x} \propto \left(2x^2 - r^2 - rx^2a\right)\exp(-ar), \tag{5.90}$$

$$\frac{\partial \psi_{4p}(x)}{\partial y} \propto (2xy - rxya)\exp(-ar), \tag{5.91}$$

Second derivatives

The Laplacian of the s-orbitals are on the form

$$\nabla^2\psi_{1s} \propto \frac{a}{r}(ar-2)\exp(-ar) \tag{5.92}$$

$$\nabla^2\psi_{2s} \propto \left(\frac{2}{r} + a^2r - 4a\right)\exp(-ar) \tag{5.93}$$

$$\nabla^2\psi_{3s} \propto \left(6 - 6ar + a^2r^2\right)\exp(-ar), \tag{5.94}$$

while the Laplacians of the $2p$-orbitals consist of the second derivatives

$$\frac{\partial^2 \psi_{2p}(x)}{\partial x^2} \propto \frac{ax}{r}\left(-3 + \frac{x^2}{r^2} + \frac{x^2a}{r}\right)\exp(-ar), \tag{5.95}$$

$$\frac{\partial^2 \psi_{2p}(x)}{\partial y^2} \propto \frac{x}{r}\left(\frac{y^2}{r^2} - 1 + \frac{ay^2}{r}\right)\exp(-ar). \tag{5.96}$$

The $4p$-orbitals have the second derivatives

$$\frac{\partial^2 \psi_{4p}(x)}{\partial x^2} \propto x\left(6 - \frac{3x^2a}{r} + 3ar + x^2a^2\right)\exp(-ar), \tag{5.97}$$

$$\frac{\partial^2 \psi_{4p}(x)}{\partial y^2} \propto x\left(2 - \frac{3y^2a}{r} - ar + a^2y^2\right)\exp(-ar). \tag{5.98}$$

The expressions for $\partial^2\psi_{4p}(y)/\partial y^2$, $\partial^2\psi_{4p}(z)/\partial y^2$ etc. are identical to the ones above and are used to construct e.g.

$$\nabla^2\psi_{4p}(x) = \frac{\partial^2 \psi_{4p}(x)}{\partial x^2} + \frac{\partial^2 \psi_{4p}(x)}{\partial y^2} + \frac{\partial^2 \psi_{4p}(x)}{\partial z^2}. \tag{5.99}$$

All these expressions are needed for the implementation of the Slater determinant, the quantum force and the kinetic energy in our VMC calculation.

Part II

Implementation and results

Chapter 6

Implementation of the Variational Monte Carlo Method

The Variational Monte Carlo method was implemented by writing a C++-based code. The first thing was to implement a brute force code to handle the helium atom, and which could easily be reduced to solve the simple hydrogen system, with only one atom and hence no inter-electron potentials. Brute force means here that there was no importance sampling, so the simple Metropolis algorithm (see figure 4.1) was used. The next step was to implement importance sampling, in order to make the sampling more efficient, and hopefully getting better results for both energy and variance. This would mean we had to replace the simple Metropolis algorithm with the Metropolis-Hastings algorithm, which includes importance sampling.

However, the code was not efficient at this stage. By calculating the explicit wave function for each particle move, and using numerical derivatives, there was still room for improvement. For larger systems like beryllium, neon etc., we would need to implement a more efficient wave function with closed-form expressions for the Slater determinant and its derivatives, using the definitions given in section 5.2, in addition to an efficient computation of the Jastrow factor. By doing this we could improve the efficiency of each Monte Carlo cycle, resulting in less computation time. If this optimization of the wave function was made general enough, and we included closed form expressions for the single particle wave functions and their derivatives, we could easily use the program to calculate an approximation to the ground state energy for basically any even-numbered atom.

In this chapter I will describe the optimized version of the program, since this is of most interest, and the part that has taken the most time to develop. I will also focus on the Metropolis-Hastings algorithm and not the more simple Metropolis algorithm.

Initially, the code was developed using object-oriented programming (OOP), but complications and little programming experience prevented me from fully exploiting the strengths of OOP. However, the main Monte Carlo machinery is built within a single class, so at least one can argue that this approach has left me with a little more organized code. I will therefore not emphasize on the code structure, but rather on the key points of importance for the calculations.

At the end of this chapter I will describe the blocking method used to calculate the true standard error of the data sets produced by the VMC code. I will also discuss a way to estimate the true VMC results for $\Delta t = 0$ in the Metropolis-Hastings algorithm (see sections 4.4.2 and 4.4.2). A brief discussion of parallel computing is also included.

6.1 Optimizing the calculations

In order to reduce the time of thecalculations, we had to find a way to both rewrite the Slater determinant and the Jastrow factor. When increasing the system size, calculating the entire wave function,

$$\psi_T = \Phi \cdot J = Det \uparrow \cdot Det \downarrow \cdot J, \tag{6.1}$$

for each particle move, seemed inefficient. The determinants are defined as

$$Det \uparrow \ \propto \begin{vmatrix} \psi_{1\uparrow}(\mathbf{r}_1) & \psi_{1\uparrow}(\mathbf{r}_2) & \cdots & \psi_{1\uparrow}(\mathbf{r}_{N/2}) \\ \psi_{2\uparrow}(\mathbf{r}_1) & \psi_{2\uparrow}(\mathbf{r}_2) & \cdots & \psi_{2\uparrow}(\mathbf{r}_{N/2}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{M\uparrow}(\mathbf{r}_1) & \psi_{M\uparrow}(\mathbf{r}_2) & \cdots & \psi_{M\uparrow}(\mathbf{r}_{N/2}) \end{vmatrix}, \tag{6.2}$$

and

$$Det \downarrow \ \propto \begin{vmatrix} \psi_{1\downarrow}(\mathbf{r}_{N/2+1}) & \psi_{1\downarrow}(\mathbf{r}_{N/2+2}) & \cdots & \psi_{1\downarrow}(\mathbf{r}_N) \\ \psi_{2\downarrow}(\mathbf{r}_{N/2+1}) & \psi_{2\downarrow}(\mathbf{r}_{N/2+2}) & \cdots & \psi_{2\downarrow}(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{M\downarrow}(\mathbf{r}_{N/2+1}) & \psi_{M\downarrow}(\mathbf{r}_{N/2+2}) & \cdots & \psi_{M\downarrow}(\mathbf{r}_N) \end{vmatrix}. \tag{6.3}$$

for a system of N electrons with $M = N/2$ possible spatial wave functions. The Jastrow factor is given as

$$J = \prod_{i<j}^{N} f(r_{ij}) = \prod_{i<j}^{N} \exp\left(\frac{r_{ij}}{a(1+\beta r_{ij})}\right), \tag{6.4}$$

A crucial point is the fact that by moving one electron, as we do in the Metropolis algorithms (see sections 4.4.1 and 4.4.2), we actually only change one column in only **one** of the two Slater determinants. This should lead us to believe that an optimization should be obtainable.

We use the wave function, ψ_T, for:

- Calculating the transition probability in the Metropolis-Hastings algorithm as

$$w = \frac{G(x,y)|\psi_T(y)|^2}{G(y,x)|\psi_T(x)|^2}, \tag{6.5}$$

  where y are the new positions, and x are the old ones, and $G(y,x)$ is the Green's function in Eq. (4.39).

- Calculating the quantum force,

$$F(x) = 2\frac{\nabla\psi_T}{\psi_T}, \tag{6.6}$$

  needed for the Green's function involving the gradient of the wave function.

- Calculating the local energy,

$$E_L = \frac{\nabla^2\psi_T}{\psi T} - V, \tag{6.7}$$

  involving the Laplacian of the wave function.

This means that we need to optimize the ratios $\Phi(r^{new})/\Phi(r^{old})$, $\nabla\Phi(r)/\Phi(r)$, $\nabla^2\Phi(r)/\Phi(r)$, $J(r^{new})/J(r^{old})$, $\nabla J/J$ and $\nabla^2 J/J$.

Since our wave function is on the form

$$\psi_T = \Phi \cdot J, \tag{6.8}$$

the total expression for the gradient ratio will be

$$\frac{\nabla\psi_T}{\psi_T} = \frac{\nabla\Phi}{\Phi} + \frac{\nabla J}{J}. \tag{6.9}$$

For the kinetic energy term in the local energy we will have to calculate the Laplacian of ψ_T, divided by ψ_T itself as

$$\frac{\nabla^2\psi_T}{\psi_T} = \frac{\nabla^2\Phi}{\Phi} + \frac{\nabla^2 J}{J} + \left(\frac{\nabla\Phi}{\Phi}\cdot\frac{\nabla J}{J}\right). \tag{6.10}$$

6.1.1 Optimizing the ratio - $\Psi_T(r^{new})/\Psi_T(r^{old})$

When optimizing the wave function part of the transition ratio, it's a good idea to split it into a correlation part, and a Slater determinant part. This means

$$R = \frac{\psi_T(r^{new})}{\psi_T(r^{old})} = R_{SD}R_J = \frac{\Phi(r^{new})}{\Phi(r^{old})}\frac{J(r^{new})}{J(r^{old})}, \tag{6.11}$$

where Φ is the Slater determinant, and J is the correlation part, the Jastrow factor. In this section I will label the positions of all particles before a particle move as r^{old}, and the set of particle positions after the move as r^{new}.

Slater determinant ratio - $\Phi(\mathbf{r^{new}})\Phi(\mathbf{r^{old}})$

We start by defining the so-called Slater *matrix*, $\mathcal{D}$, the matrix used to calculate the Slater determinant, and its elements as $D_{ij} = \phi_j(r_i)$, where $\phi_j(r_i)$ are electrons i occupying a single particle state j. The inverse of the Slater matrix, $\mathcal{D}^{-1}$, is related to the Slater determinant, Φ or just $|D|$, as

$$\mathcal{D}^{-1} = \frac{\text{adj}\mathcal{D}}{|D|}, \tag{6.12}$$

or as matrix elements as

$$D^{-1}_{ij} = \frac{C_{ji}}{|D|}, \tag{6.13}$$

where adj$\mathcal{D}$ is the so-called *adjugate* of a matrix. The adjugate of a matrix $\mathcal{A}$ is the transposed of the co-factor matrix of $\mathcal{A}$ (indices are interchanged from Eq. (6.12) to Eq. (6.13)). For the matrix $\mathcal{D}$ to be invertible, we must have $\mathcal{D}^{-1}\mathcal{D} = \mathcal{I}$, where $\mathcal{I}$ is the identity matrix, or written out as

$$\sum_{k=1}^{N} D_{ik}D^{-1}_{kj} = \delta_{ij}. \tag{6.14}$$

Multiplying Eq. (6.13) with D_{ji} from the left and summing over i, will give

$$\sum_{i=1}^{N} D_{ji}D^{-1}_{ij} = \frac{\sum_{i=1}^{N} D_{ji}C_{ji}}{|D|}, \tag{6.15}$$

resulting in

$$|D| = \sum_{i=1}^{N} D_{ji} C_{ji}. \tag{6.16}$$

We can now express the ratio between the new and old Slater determinants, R_{SD}, as

$$R_{SD} = \frac{\sum_{j=1}^{N} D_{ij}(r^{new}) C_{ij}(r^{new})}{\sum_{j=1}^{N} D_{ij}(r^{old}) C_{ij}(r^{old})}. \tag{6.17}$$

When moving a particle i, only the i-th row of the Slater matrix changes, and therefore only the i-th row of the matrix of co-factors remains unchanged, which means

$$C_{ij}(r^{new}) = C_{ij}(r^{old}). \tag{6.18}$$

Using Eq. (6.13) gives

$$R_{SD} = \frac{\sum_{j=1}^{N} D_{ij}(r^{new}) C_{ij}(r^{old})}{\sum_{j=1}^{N} D_{ij}(r^{old}) C_{ij}(r^{old})} = \frac{\sum_{j=1}^{N} D_{ij}(r^{new}) D_{ji}^{-1}(r^{old})}{\sum_{j=1}^{N} D_{ij}(r^{old}) D_{ji}(r^{old})}. \tag{6.19}$$

This denominator will only give unity, so the ratio can be expressed as

$$R_{SD} = \sum_{j=1}^{N} D_{ij}(r^{new}) D_{ji}^{-1}(r^{old}) = \sum_{j=1}^{N} \phi_j(r_i^{new}) D_{ji}^{-1}(r^{old}). \tag{6.20}$$

When only moving the i-th particle, we only need to calculate a dot product, $\phi(r_i^{new}) \cdot D_i^{-1}$, where $\phi(r_i^{new})$ is the vector

$$\phi(r_i^{new}) = (\phi_1(r_i^{new}), \phi_2(r_i^{new}), \ldots, \phi_N(r_i^{new})), \tag{6.21}$$

and D_i^{-1} is the i-th column of the inverse matrix $\mathcal{D}^{-1}$ where the electron i is at the *old* position. We also have splitted the determinants in two separate determinants, so we actually only change **one** row in **one** of the Slater determinants when a particle i is moved.

The following code shows the function that calculates the ratio R_{SD}:

```
double atom::getRatio(double **D_up, double **D_down, int i, double alpha, double
    beta){
  double ratio=0;
  if(i<no_of_particles/2){
    for(int j=0; j<no_of_particles/2; j++){
      ratio += phi(r_new,alpha,beta,j,i)*D_up[j][i];
    }
    return ratio;
  }
  else{
    for(int j=0; j<no_of_particles/2; j++){
      ratio += phi(r_new,alpha,beta,j,i)*D_down[j][i-no_of_particles/2];
    }
    return ratio;
  }
}
```

Here the function `phi(r,alpha,beta,j,i)` calculates the single particle wave function for electron `i` in orbital `j`.

As described in [3] we can use an algorithm to update the inverse of the Slater matrix without calculating the entire Slater matrix again. After moving particle i, and if the move is accepted, we must for each column $j \neq i$ calculate a quantity

$$S_j = \sum_{l=1}^{N} D_{il}(r^{new}) D_{lj}(r^{old}). \tag{6.22}$$

The new elements of $\mathcal{D}^{-1}$ are as given on page 277 in [3]:

$$D_{kj}^{-1}(r^{new}) = \begin{cases} D_{kj}^{-1}(r^{old}) - \frac{S_j}{R_{SD}} D_{ki}^{-1}(r^{old}) & \text{if } j \neq i, \\ \frac{1}{R_{SD}} D_{kj}^{-1} & \text{if } j = i. \end{cases} \tag{6.23}$$

The following code shows how we first update the spin up determinant, and then the spin down determinant

```
void atom::update(double **D_up, double **D_down, int i, double ratio, double alpha,
    double beta){

  //SPIN UP ELECTRONS
  if(i<no_of_particles/2){
    for(int j=0; j<no_of_particles/2; j++){
      if(j!=i){
  for(int l=0; l<no_of_particles/2; l++){
    s_j[j]+=phi(r_new,alpha,beta,l,i)*D_up[l][j];
  }
      }
    }
    for(int j=0; j<no_of_particles/2; j++){
      if(j!=i){
  for(int k=0; k<no_of_particles/2; k++){
    D_up[k][j]=D_up[k][j]-s_j[j]*D_up[k][i]/ratio;
  }
      }
    }
    //i'th column, i=j
    for(int k=0; k<no_of_particles/2; k++){
      D_up[k][i]=D_up[k][i]/ratio;
    }
  }
//SPIN DOWN ELECTRONS
  else{
    i=i-no_of_particles/2;//This is because the spin-down matrix have indices 0,1,...,
        N/2
    for(int j=0; j<no_of_particles/2; j++){
      if(j!=i){
  for(int l=0; l<no_of_particles/2; l++){
    s_j[j]+=phi(r_new,alpha,beta,l,i+no_of_particles/2)*D_down[l][j];
  }
      }
    }
    for(int j=0; j<no_of_particles/2; j++){
      if(j!=i){
  for(int k=0; k<no_of_particles/2; k++){
    D_down[k][j]=D_down[k][j]-s_j[j]*D_down[k][i]/ratio;
  }
```

```
      }
    }
    //i'th column, i=j
    for(int k=0; k<no_of_particles/2; k++){
      D_down[k][i]=D_down[k][i]/ratio;
    }
    i=i+no_of_particles/2;
  }
}
```

Jastrow ratio - $\mathbf{J}(\mathbf{r}^{\mathbf{new}})/\mathbf{J}(\mathbf{r}^{\mathbf{old}})$

We now have to find a functional way to store the inter-electron distances, or for simplicity, the individual Jastrow functions,

$$g_{i,j} = g(r_{ij}) = \exp\left(\frac{r_{ij}}{a(1+\beta r_{ij})}\right), \tag{6.24}$$

between the different electrons i and j, where a is given by

$$a = \begin{cases} 4 & \text{if spins are parallel,} \\ 2 & \text{if spins are anti-parallel.} \end{cases} \tag{6.25}$$

From section 5.1.2, we have our Jastrow factor, J, on the form

$$J = \prod_{i<j} g_{i,j}. \tag{6.26}$$

A way to store these functions is to construct a matrix with elements above the diagonal as:

$$\mathcal{G} = \begin{bmatrix} \cdot & g_{1,2} & g_{1,3} & \cdots & g_{1,N} \\ & \cdot & g_{2,3} & \cdots & g_{2,N} \\ & & \cdot & \ddots & \vdots \\ & & & \cdot & g_{N-1,N} \end{bmatrix}. \tag{6.27}$$

The Jastrow ratio, R_J is

$$R_J = \frac{J(r^{new})}{J(r^{old})}. \tag{6.28}$$

From section 5.1.2, our Jastrow factor, J, is on the form

$$J = \prod_{i<j} g_{i,j} = \prod_{i<j} \exp f_{i,j}, \tag{6.29}$$

and when moving only particle k, only the elements with k as an index will change, that is $N-1$ elements. Because of our exponential form of the correlation part the ratio can be written

$$\frac{J(r^{new})}{J(r^{old})} = e^{\Delta J}, \tag{6.30}$$

where the intriguing part is ΔJ, given as

$$\Delta J = \sum_{i=1}^{k-1} (f_{i,k}^{new} - f_{i,k}^{old}) + \sum_{i=k+1}^{N} (f_{k,i}^{new} - f_{k,i}^{old}), \tag{6.31}$$

with the functions f_{ij} given as

$$f_{ij} = \frac{r_{ij}}{a(1+\beta r_{ij})}. \tag{6.32}$$

The following code shows the calculation of the Jastrow factor ratio:

```
double atom::getJastrowRatio(double** distance_old, double** distance_new,double beta){
 for(int k=0; k<no_of_particles; k++){
   for(int l=0; l<k; l++){
     jastrowRatio += distance_new[l][k]-distance_old[l][k];
   }
 }
 for(int k=0; k<no_of_particles; k++){
   for(int l=k+1; l<no_of_particles; l++){
     jastrowRatio += distance_new[l][k]-distance_old[l][k];
   }
 }
 return jastrowRatio;
}
```

The two-dimensional arrays `distance_new` and `distance_old` have been calculated using:

```
void atom::getDistance(double** distance, double** r_old, double beta){
 for(int k=0; k<no_of_particles; k++){
   for(int l=k+1; l<no_of_particles; l++){
     temp = diffR(r_old, k,l);
     //spin up
     if(((k < no_of_particles/2) && (l <no_of_particles/2)) || ((k>=no_of_particles/2
         && l>=no_of_particles/2))){
 a=0.25;
 distance[k][l] = a*temp/(1+beta*temp);
     }
     //spin down
     else{
 a=0.5;
 distance[k][l] = a*temp/(1+beta*temp);
     }
   }
 }
}
```

We see that the names of the arrays are a bit misleading, as we actually store the values $f(r_{ij})$ (see Eq. (6.32)) and not r_{ij} themselves.

6.1.2 Derivative ratios

In order to optimize the calculations, we need to find optimized expressions for the ratios $\nabla\Phi/\Phi$, $\nabla^2\Phi/\Phi$, $\nabla J/J$ and $\nabla^2 J/J$, as seen in Eqs. (6.9) and (6.10). This will lead to a more efficient evaluation of the quantum force and the kinetic energy.

Optimizing the ratios - $\nabla\Phi/\Phi$ and $\nabla^2\Phi/\Phi$

We use a similar procedure as in the previous section to get the explicit expressions for the gradient and Laplacian of the Slater determinant. When differentiating with respect

to the coordinates of a single particle i, we will get

$$\frac{\nabla_i \Phi}{\Phi} = \sum_{j=1}^{N} \nabla_i D_{ij}(r) D_{ji}^{-1}(r) = \sum_{j=1}^{N} \nabla_i \phi_j(r_i) D_{ji}^{-1}(r), \tag{6.33}$$

for the gradient, remembering that the operator ∇ actually is a vector

$$\nabla = \frac{\partial}{\partial x}\hat{\imath} + \frac{\partial}{\partial y}\hat{\jmath} + \frac{\partial}{\partial z}\hat{k}, \tag{6.34}$$

with $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ are the unit vectors for x-, y- and z- dimensions respectively.

For the Laplacian-expression we will obtain

$$\frac{\nabla_i \Phi}{\Phi} = \sum_{j=1}^{N} \nabla_i^2 D_{ij}(r) D_{ji}^{-1}(r) = \sum_{j=1}^{N} \nabla_i^2 \phi_j(r_i) D_{ji}^{-1}(r), \tag{6.35}$$

where ∇^2 is a scalar

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{6.36}$$

We see that the expressions only rely on the elements from the inverse Slater matrix, $\mathcal{D}^{-1}$, and the closed form expressions for the derivatives of the single particle orbitals (see sections 5.4 and 5.5).

The following code shows the calculation of the ratio $\nabla^2\Phi/\Phi$ in the local energy function:

```
//LAPLACE SLATER-DETERMINANT ANALYTIC
 for(int i=0; i<no_of_particles; i++){
   if(i<no_of_particles/2){
     for(int j=0; j<no_of_particles/2; j++){
 E_kinetic -= laplace_phi(r,alpha,beta,j,i)*D_up[j][i];
     }
   }
   else{
     for(int j=0; j<no_of_particles/2; j++){
 E_kinetic -= laplace_phi(r,alpha,beta,j,i)*D_down[j][i-no_of_particles/2];
     }
   }
 }
```

The function `laplace_phi(r,alpha,beta,j,i)` simply returns the Laplacian of the single particle wave function where electron i is in orbital j.

Optimizing the ratios - $\nabla J/J$ and $\nabla^2 J/J$

The **first derivative** results for the Padé-Jastrow function (as described in Eq. (6.26)) are taken from chapter 19 in [9] and is for a single coordinate, x, for particle k, given by

$$\frac{1}{J}\frac{\partial J}{\partial x_k} = \sum_{i=1}^{k-1} \frac{1}{g_{ik}}\frac{\partial g_{ik}}{\partial x_k} + \sum_{i=k+1}^{N} \frac{1}{g_{ki}}\frac{\partial g_{ki}}{\partial x_k}, \tag{6.37}$$

where g_{ij} is some correlation function between two particles. We have

$$g_{ij} = \exp f(r_{ij}) = \exp\left(\frac{r_{ij}}{a(1+\beta r_{ij})}\right), \tag{6.38}$$

a function of distances between the electrons. We now use the chain rule to get

$$\frac{\partial g_{ij}}{\partial x_j} = \frac{\partial g_{ij}}{\partial r_{ij}}\frac{\partial r_{ij}}{\partial x_j} = \frac{x_j - x_i}{r_{ij}}\frac{\partial g_{ij}}{\partial r_{ij}}. \tag{6.39}$$

The exponential form of g_{ij} also gives us

$$\frac{\partial g_{ij}}{\partial r_{ij}} = g_{ij}\frac{\partial f_{ij}}{\partial r_{ij}}, \tag{6.40}$$

so that

$$\frac{1}{J}\frac{\partial J}{\partial x_k} = \sum_{i=1}^{k-1}\frac{\partial f_{ik}}{\partial x_k} + \sum_{i=k+1}^{N}\frac{\partial f_{ki}}{\partial x_k}. \tag{6.41}$$

It is easy to show that

$$\frac{\partial f_{ij}}{\partial r_{ij}} = \frac{1}{a(1+\beta r_{ij})^2}, \tag{6.42}$$

and combining these results we will get

$$\frac{1}{J}\frac{\partial J}{\partial x_k} = \sum_{i=1}^{k-1}\frac{x_k - x_i}{a r_{ik}(1+\beta r_{ik})^2} + \sum_{i=k+1}^{N}\frac{x_k - x_i}{a r_{ik}(1+\beta r_{ik})^2}, \tag{6.43}$$

as the final result for the first derivative with respect to a chosen coordinate x, and particle k. The quantum force for a given particle k in coordinate x,

$$F(x_k) = \frac{2}{J}\frac{\partial J}{\partial x_k} = 2\left(\sum_{i=1}^{k-1}\frac{x_k - x_i}{a r_{ik}(1+\beta r_{ik})^2} + \sum_{i=k+1}^{N}\frac{x_k - x_i}{a r_{ik}(1+\beta r_{ik})^2}\right). \tag{6.44}$$

The following code shows the calculation of the quantum force with the first part focusing on the Slater determinant part discussed in the previous section, while the second part focuses on the correlation part of the quantum force (see Eq. (6.9)):

```
void atom::quantum_force(double**r,double**qm_force, double **D_up, double **D_down,
    double alpha, double beta){
  //SLATER PART
  for(int p=0; p<no_of_particles; p++){
    if(p<no_of_particles/2){
      for(int q=0; q<dimension; q++){
  for(int l=0; l<no_of_particles/2; l++){
    qm_force[p][q] += 2*gradient_phi(r,alpha,q,l,p)*D_up[l][p];
  }
      }
    }
    else{
      for(int q=0; q<dimension; q++){
  for(int l=0; l<no_of_particles/2; l++){
    qm_force[p][q] += 2*gradient_phi(r,alpha,q,l,p)*D_down[l][p-no_of_particles/2];
  }
      }
    }
  }
  //JASTROW PART
 for(int p=0; p<no_of_particles; p++){
    for(int q=0; q<dimension; q++){
      for(int l=p+1; l<no_of_particles; l++){
```

```
  qm_force[p][q] += 2*gradient_jastrow(r,p,l,q,beta);
      }
      for(int l=0; l<p; l++){
  qm_force[p][q] += 2*gradient_jastrow(r,p,l,q,beta);
      }
    }
  }
}//end quantum_force
```

The function `gradient_phi(r,alpha,q,l,p)` returns the first derivative of the single particle orbital ϕ_l with respect to coordinate `q` of electron `p`.
The function `gradient_jastrow(r,p,l,q,beta)`is given as

```
double atom::gradient_jastrow(double** r,int p, int l, int d, double beta){
  if(((p < no_of_particles/2) && (l <no_of_particles/2)) || ((p>=no_of_particles/2 &&
      l>=no_of_particles/2))){
    a=0.25;//More efficient to multiply by 0.25 than to divide by 4
    temp1 = diffR(r,p,l);
    temp2 = 1+beta*temp1;
    temp3 = r[p][d]-r[l][d];
    return a*temp3/temp1/temp2/temp2;
  }
  else{
    a=0.5;//More efficient to multiply by 0.5 than to divide by 2
    temp1 = diffR(r,p,l);
    temp2 = 1+beta*temp1;
    temp3 = r[p][d]-r[l][d];
    return a*temp3/temp1/temp2/temp2;
  }
}
```

The function `diffR(r,p,l)` simply returns the distance between electrons `p` and `l`.

For the **second derivative**, I refer to [9] where it is shown that for a correlation function of our form

$$g_{ij} = \exp f_{ij}, \tag{6.45}$$

the full expression for the Laplacian of the correlation part can be written as

$$\begin{aligned}\frac{\nabla^2 J}{J} = \left(\frac{\nabla J}{J}\right)^2 &+ \sum_{i=1}^{k-1}\left(\frac{(d-1)}{r_{ik}}\frac{\partial f_{ik}}{\partial r_{ik}} + \frac{\partial^2 f_{ik}}{\partial r_{ik}^2}\right) \\ &+ \sum_{i=k+1}^{N}\left(\frac{(d-1)}{r_{ki}}\frac{\partial f_{ki}}{\partial r_{ki}} + \frac{\partial^2 f_{ki}}{\partial r_{ki}^2}\right),\end{aligned} \tag{6.46}$$

where the first term is just the dot product of the gradient (see the previous section) with itself, and d is the number of spatial dimensions, i.e. $d = 3$. We have the explicit expression for the Padé-Jastrow correlation function

$$f_{ij} = \frac{r_{ij}}{a(1+\beta r_{ij})}, \tag{6.47}$$

where a is either 2 or 4 depending on the spins of particles i and j. The second derivative of f_{ij} with respect to r_{ij} is now

$$\frac{\partial^2 f_{ij}}{\partial r_{ij}^2} = -\frac{2\beta}{a(1+\beta r_{ij})^3}, \tag{6.48}$$

since the first derivative is

$$\frac{\partial f_{ij}}{\partial r_{ij}} = \frac{1}{a(1+\beta r_{ij})^2}, \tag{6.49}$$

as in the previous section. Inserting Eqs. (6.49) and (6.48) into Eq. (6.46) will then give

$$\frac{\nabla^2 J}{J} = \left(\frac{\nabla J}{J}\right)^2 + \sum_{i=1}^{k-1}\left(\frac{2}{r_{ik}}\frac{1}{a(1+\beta r_{ik})^2} - \frac{2\beta}{a(1+\beta r_{ik})^3}\right) + \sum_{i=k+1}^{N}\left(\frac{2}{r_{ki}}\frac{1}{a(1+\beta r_{ki})^2} - \frac{2\beta}{a(1+\beta r_{ki})^3}\right), \tag{6.50}$$

which can be reduced to

$$\frac{\nabla^2 J}{J} = \left(\frac{\nabla J}{J}\right)^2 + \sum_{i=1}^{k-1}\frac{2}{ar_{ik}}\frac{1}{(1+\beta r_{ik})^3} + \sum_{i=k+1}^{N}\frac{2}{ar_{ki}}\frac{1}{(1+\beta r_{ki})^3}. \tag{6.51}$$

In the local energy function we need to calculate the cross-term

$$\frac{\nabla\Phi}{\Phi}\cdot\frac{\nabla J}{J}, \tag{6.52}$$

and $\nabla^2 J/J$. The following code shows how I compute these quantities.

The Slater term, $\nabla\Phi/\Phi$ is computed as

```
//SLATER TERM:
for(int p=0; p<no_of_particles; p++){
  if(p<no_of_particles/2){
    for(int q=0; q<dimension; q++){
for(int l=0; l<no_of_particles/2; l++){
  temp1[p][q] += gradient_phi(r,alpha,q,l,p)*D_up[l][p];
}
    }
  }
  else{
    for(int q=0; q<dimension; q++){
for(int l=0; l<no_of_particles/2; l++){
  temp1[p][q] += gradient_phi(r,alpha,q,l,p)*D_down[l][p-no_of_particles/2];
}
    }
  }
}
```

The following code shows how I compute the Jastrow term, $\nabla J/J$, using the `gradient_jastrow`-functions explained previously.

```
//JASTROW TERM
for(int p=0; p<no_of_particles; p++){
  for(int q=0; q<dimension; q++){
    for(int l=p+1; l<no_of_particles; l++){
temp2[p][q] += gradient_jastrow(r,p,l,q,beta);
    }
    for(int l=0; l<p; l++){
temp2[p][q] += gradient_jastrow(r,p,l,q,beta);
    }
  }
}
```

We can now compute the cross-term by a simply dot product calculation using the following code:

```
for(int p=0; p<no_of_particles; p++){
  for(int q=0; q<dimension; q++){
    tempE += temp1[p][q]*temp2[p][q];
  }
}
```

The functions `gradient_jastrow(...)` and `gradient_phi(...)` have been explained earlier. The array `temp2[p][q]` now contains the ratio $\nabla J/J$ and is used to calculate the ratio $\nabla^2 J/J$ as given in Eq. (6.51). The code for this equation is:

```
//Gradient ratio squared:
for(int p=0; p<no_of_particles; p++){
   for(int q=0; q<dimension; q++){
     tempE2 += temp2[p][q]*temp2[p][q];
   }
 }
 //Second term, involving df/dx and d^2f/dx^2
 for(int p=0; p<no_of_particles; p++){
   for (int k = 0; k < no_of_particles; k++) {
     if ( k != p) {
 if(((p < no_of_particles/2) && (k <no_of_particles/2)) || ((p>=no_of_particles/2 &&
     k>=no_of_particles/2))){
    a=0.25;
    tempE3 += 2*a/diffR(r,p,k)/pow((1+beta*diffR(r,p,k)),3);
  }
  else{
    a=0.5;
    tempE3 += 2*a/diffR(r,p,k)/pow((1+beta*diffR(r,p,k)),3);
  }
     }
   }
 }
```

We now have all the pieces to calculate the local energy with `E_kinetic`, `tempE`,`tempE2` and `tempE3`. We only need to multiply the different term by scalar factors, and calculate the potential energy. The following code shows this.

```
//KINETIC ENERGY
 E_kinetic -= 2*tempE;
 E_kinetic -= tempE2;
 E_kinetic -= tempE3;

 E_kinetic *= 0.5;

 //POTENTIAL ENERGY Coulomb potential
   //electron-proton-interaction
 for(int i=0; i<no_of_particles; i++){
   r_single_particle=0;
   for(int j=0; j<dimension; j++){
     r_single_particle += r[i][j]*r[i][j];//r^2=x^2+y^2+z^2
   }
   E_potential -=charge/sqrt(r_single_particle);
 }
 //electron-electron-interaction
 for(int i=0; i<no_of_particles-1; i++){
   for(int j=i+1; j<no_of_particles; j++){
     r_12=0;
```

```
    for(int k=0; k<dimension; k++){
 r_12 +=(r[i][k]-r[j][k])*(r[i][k]-r[j][k]);
    }
    E_potential +=1/sqrt(r_12);
  }
 }//*/
 E_local = E_potential + E_kinetic;

 return E_local;
```

6.2 Implementation of Metropolis-Hastings algorithm

We can now use the results from the previous sections to show how the Metropolis-Hastings algorithm is implemented. This algorithm is discussed in section 4.4.2, and figure 4.2 sums up the main points quite nicely.

After calculating the initial positions for the electrons, the inverse spin-up and -down Slater matrices and the quantum force `F`, I start the first Monte Carlo cycle and move the first electron as

```
for(int j=0; j<dimension; j++){
  r_new[i][j]=r_old[i][j]+D*time_step*qm_force_old[i][j]
    +gaussian_deviate(&idum)*sqrt(time_step)
```

where the function `gaussian_deviate` returns a normal distributed random number. I then calculate the distances between the electrons to find both the Slater- and Jastrow-ratios for the full ratio of probabilities, $\psi_T(r^{new})/\psi_T(r^{old})$ as

```
//SETTING UP THE DISTANCE MATRIX
getDistance(distance_new, r_new, beta);
//SLATER DETERMINANT RATIO
ratio = getRatio(D_up, D_down, i, alpha, beta);

jastrowRatio = getJastrowRatio(distance_old, distance_new, beta);
jastrowRatio = exp(jastrowRatio);
```

The next point is to update the inverse Slater matrices, and then calculate the quantum force for the Green's function in the Metropolis-Hastings test:

```
 //temporary update for quantum force
    update(temp_up, temp_down, i, ratio, alpha, beta);
    quantum_force(r_new, qm_force_new, temp_up, temp_down, alpha, beta);

    //IMPORTANCE SAMPLING
    double greensfunction = 0.0;
    for(int k=0; k<no_of_particles; k++){
 for(int j=0; j < dimension; j++) {
   greensfunction += 0.5*(qm_force_old[k][j]+qm_force_new[k][j])*
     (D*time_step*0.5*(qm_force_old[k][j]-qm_force_new[k][j])
   -r_new[k][j]+r_old[k][j]);
 }
    }
    greensfunction = exp(greensfunction);
```

I use temporary arrays in case the Metropolis-Hastings test does not accept the particle move. The test itself is implemented as

```
 //METROPOLIS TEST
if(ran1(&idum)<= greensfunction*ratio*ratio*jastrowRatio*jastrowRatio){
  for(int k=0; k<no_of_particles; k++){
    for(int j=0; j<dimension; j++){
      r_old[k][j]=r_new[k][j];
      qm_force_old[k][j] = qm_force_new[k][j];
    }
  }
  //UPDATE INVERSE SLATER MATRICES
  for(int m=0; m<no_of_particles/2; m++){
    for(int n=0; n<no_of_particles/2; n++){
      D_up[m][n]=temp_up[m][n];
      D_down[m][n]=temp_down[m][n];
    }
  }
  //UPDATE THE DISTANCE MATRIX
  for(int m=0; m<no_of_particles; m++){
    for(int n=0; n<no_of_particles; n++){
      distance_old[m][n]=distance_new[m][n];
    }
  }
}//end metropolis test
```

The random number generator `ran1` is taken from [12]. If now all particles for the cycle have been moved, we can go on to sample the local energy for the expectation values $\langle E_L \rangle$ and $\langle E_L^2 \rangle$:

```
delta_E = local_energy(r_old, D_up, D_down, alpha, beta);
energy +=delta_E;
energy2 +=delta_E*delta_E;
```

At the end of the Monte Carlo loop, we then divide the expectation values by the number of Monte Carlo cycles we have performed.

Furthermore, an important note must be made regarding the calculation of the local energy. At the beginning of the simulation, the system will in general not be in its most probable state. By reviewing section 4.1, we see that a system must evolve before reaching its largest eigenvalue, corresponding to the most probable state. As a result of this, we must include some *thermalization* cycles before we start sampling the local energy. As seen in the algorithm, we initialize the system in random positions. If we perform a certain number of Monte Carlo cycles before we start sampling the local energy, the system will hopefully be in its most probable state when we actually **do** start sampling the energy. This will improve our results, since we will calculate the energy where the PDF has greater values, leading to more efficient calculations.

The number of thermalization cycles is not known *a priori*, and must be dealt with experimentally. In this thesis we found that by using $200,000$ thermalization cycles, we will safely start sampling the energy when the system has reached its most probable state, even for large systems.

6.3 Blocking

The simulation of physical systems using Monte Carlo methods, involve pseudo-random numbers and will produce correlated results. That is, as the pseudo-random numbers themselves will show correlated behavior, functions evaluated with respect to variables

dependent on these pseudo-random numbers will become correlated as well. As discussed in appendix B, the statistical error for set of correlated stochastic values is given by

$$\mathrm{err}_X^2 = \frac{1}{n}\mathrm{Cov}(x), \tag{6.53}$$

while for uncorrelated values this will reduce to

$$\mathrm{err}_X^2 = \frac{1}{n}\mathrm{Var}(x) = \frac{1}{n^2}\sum_{k=1}^{n}(x_k - \bar{x}_n)^2. \tag{6.54}$$

As given in appendix B, we can write the error for correlated cases as

$$\mathrm{err}_X^2 = \frac{\tau}{n}\mathrm{Var}(x), \tag{6.55}$$

where τ is the autocorrelation time

$$\tau = 1 + 2\sum_{d=1}^{n-1}\kappa_d, \tag{6.56}$$

describing the correlation between measurements. However, calculating κ_d, as given in Eqs. (B.34) and (B.35), will be rather time consuming.

By using the blocking technique, or 'data bunching', as described in [17] and [18], we actually have a way to make the measurements 'uncorrelated'. By grouping the data in equally sized blocks, and treating the mean, or sample mean, of each block as a measurement, we can find an error estimate by merely plotting the error as a function of block size as if the data were uncorrelated.

Suppose we have 1000 measurements in our sample. We can then divide these values into 10 blocks with 100 measurements in each block. Furthermore, we calculate the mean in each block, and treat them as our measurements. Then we calculate the error as if the data were uncorrelated using Eq. (6.54), where the block means are our x_k's and $n = 10$. The total sample mean,

$$\bar{x}_n = \frac{1}{n}\sum_{k=1}^{n} x_k, \tag{6.57}$$

will of course be identical whether we calculate it directly, that is for block sizes equal 1 and n equal total number of measurements, or if we group them into blocks.

The blocking method tells us that for block sizes smaller than the correlation time, the distance between correlated measurements, calculating the error using Eq. (6.54) will underestimate the error. When the block size is larger than the correlation time, the measurements are no longer correlated, and the error will no longer be affected by varying the block size. This becomes apparent as there will be a clear plateau when the data becomes uncorrelated. When the error reaches the plateau we have found an estimate for the true error,

A possible problem with relying on blocking is if the plateau is reached at a point where the number of blocks is too small to accurately estimate the variance. In this thesis we are lucky enough for that not to occur.

6.4 Time step extrapolation

By applying the Metropolis-Hastings algorithm, which introduces concepts and quantities from diffusion and random walk theory (see section 4.4.2), we are forced to deal with a time step, Δt, which for our practical calculations must be non-zero. We can understand this by inspecting the equation we use in the program for proposing a move for an electron,

$$y = x + DF(x)\Delta t + \xi\sqrt{\Delta t}, \tag{6.58}$$

where y is the proposed new position, x is the old position, D is the diffusion constant, $F(x)$ is the quantum force (see section 4.4.2) and ξ is normal distributed random variable. By having $\Delta t = 0$, the electrons would never move according to this model. However, our system is stationary, which means that we must have $\Delta t = 0$ in order to simulate the system correctly.

As done for the Diffusion Monte Carlo calculations (see e.g. [3]) by Sarsa in [19], it is possible to extrapolate towards $\Delta t = 0$ by plotting results for non-zero Δt and fitting a curve to the data plot.

In this work we have used a **linear** fitting which also accommodates for the standard error of the data. This fitting model is based on the method in chapter 15.4 in [12], and can both estimate the energy for $\Delta t = 0$ and the standard error for $\Delta t = 0$.

6.5 Parallel computing

In this program I have also implemented the possibility of running the code in parallel. This gives us the opportunity to run the code on several nodes simultaneously in order to obtain more samples than just one node would give us in the same space of time.

A Monte Carlo calculation is especially easy to program in parallel. We can run the program separately on different nodes with a thermalization period that ensures us we have reached the most probable state on all nodes. Furthermore, we can simply collect the individual mean values from all the nodes into one common mean value, as if the individual mean values were individual measurements (see section B in appendix B).

In this work I use the **Titan** cluster for large parallel jobs, such as variation of silicon, magnesium etc. The Titan computer cluster is described on its homepage http://hpc.uio.no.

Chapter 7

Hartree-Fock implementation

The Hartree-Fock method (HF) is discussed in section 3.2, and the results supply a great improvement to the single particle wave functions. Using HF, we find the optimal single particle wave functions (SPWF) for our chosen basis functions. By using the optimal SPWF, we no longer have to vary the Slater determinant in a VMC calculation, and can consentrate on finding the minimum by varying the parameter(s) in the Jastrow function. As an example, I have implemented Hartree-Fock for the Beryllium atom using the method of varying coefficients (see section 3.2.1). The basis functions are here chosen to be the hydrogen-like s-orbitals.

As seen in section 3.2.1, in order to get our single particle solutions

$$\psi_a = \sum_\lambda C_{a\lambda}\phi_\lambda, \tag{7.1}$$

we have to solve the Hartree-Fock equations

$$\sum_\gamma h^{HF}_{\lambda\gamma} C_{k\gamma} = \epsilon_k C_{k\lambda}, \tag{7.2}$$

where $h^{HF}_{\lambda\gamma}$ is

$$h^{HF}_{\lambda\gamma} = \langle\lambda|h|\gamma\rangle + \sum_{a=1}^{N}\sum_{\alpha,\delta} C^*_{a\alpha}C_{a\delta}\left\langle\lambda\alpha|V|\gamma\delta\right\rangle_{AS}. \tag{7.3}$$

We see how the coefficients $C_{k\gamma}$ are involved on both sides of the equation, and therefore leading to an non-linear eigenvalue problem.

7.1 Interaction matrix elements

As mentioned in section 3.2.1, this particular Hartree-Fock approach gives us the advantage of computing the matrix elements $\langle\lambda\alpha|V|\gamma\delta\rangle$ once and for all. We have chosen the hydrogenic s-waves as the basis functions ϕ_λ from Eq. (7.1).

The interaction is given by the Coulomb force between the electrons, and the matrix elements are given by

$$\langle\lambda\alpha|V|\gamma\delta\rangle = \int r_1^2 dr_1 \int r_2^2 dr_2 R^*_{n_\alpha 0}(r_1)R^*_{n_\beta 0}(r_2)\frac{1}{r_<}R_{n_\gamma 0}(r_1)R_{n_\delta 0}(r_2), \tag{7.4}$$

This is the direct term of the interaction piece. By calling the innermost integral over r_2 for I,

$$I = \int r_2^2 dr_2 R_{n_\beta 0}(r_2) \frac{1}{r_<} R_{n_\delta 0}(r_2),$$

we can split this up into two pieces. One where $r_2 < r_1$ and $r_< = r_1$ and one where $r_2 > r_1$ with $r_< = r_2$, giving

$$I = \int_0^{r_1} r_2^2 dr_2 R_{n_\beta 0}(r_2) \frac{1}{r_1} R_{n_\delta 0}(r_2) + \int_{r_1}^{\infty} r_2 dr_2 R_{n_\beta 0}(r_2) R_{n_\delta 0}(r_2),$$

The matrix element $\langle \lambda\alpha | V | \gamma\delta \rangle$ can then be written as

$$\langle \lambda\alpha | V | \gamma\delta \rangle = \int_0^{\infty} r_1^2 dr_1 R_{n_\alpha 0}(r_1) R_{n_\gamma 0}(r_1) \left[\int_0^{r_1} r_2^2 dr_2 R_{n_\beta 0}(r_2) \frac{1}{r_1} R_{n_\delta 0}(r_2) + \int_{r_1}^{\infty} r_2 dr_2 R_{n_\beta 0}(r_2) R_{n_\delta 0}(r_2) \right]. \tag{7.5}$$

The exchange term $\langle \lambda\alpha | V | \delta\gamma \rangle$ can be written the same way only by exchanging $\delta \leftrightarrow \gamma$ which gives

$$\langle \lambda\alpha | V | \delta\gamma \rangle = \int_0^{\infty} r_1^2 dr_1 R_{n_\alpha 0}(r_1) R_{n_\delta 0}(r_1) \left[\int_0^{r_1} r_2^2 dr_2 R_{n_\beta 0}(r_2) \frac{1}{r_1} R_{n_\gamma 0}(r_2) + \int_{r_1}^{\infty} r_2 dr_2 R_{n_\beta 0}(r_2) R_{n_\gamma 0}(r_2) \right] \tag{7.6}$$

In the equations above the functions R_{n_0} are defined by the radial part of the s-states, as

$$R_{n_0} = \left(\frac{2Z}{n} \right)^{3/2} \sqrt{\frac{1}{2n^2}} L_{n-1}^1 \left(\frac{2Zr}{n} \right) \exp\left(-\frac{Zr}{n} \right), \tag{7.7}$$

with $L_{n-1}^1 \left(\frac{2Zr}{n} \right)$ being the generalized Laguerre polynomials.

7.2 Algorithm and HF results

The Hartree-Fock program calculates the energy and optimized single particle wave functions to plug into the VMC machinery. We see from equation 7.2 that we are dealing with an eigenvalue problem. However, as the h^{HF}-matrix depends on the matrix elements of C, we have a non-linear problem, and must use an iterative self-consistency approach.

- The first thing to do is to choose the number basis functions we wish to expand the functions ψ in Eq. (7.1) with.
- We then calculate the matrix elements from Eq. (7.5). The elements were stored in a 4-dimensional array, and each integral was calculated using Gauss-Legendre-quadrature. One should really use Gauss-Laguerre to match the interval, but the Gauss-Legendre approach seemed to work well.

- The next step is to start the iteration process. The first thing we need is an ansatz for the C-matrix and we choose the identity matrix, corresponding to our new basis functions just being the old hydrogen s-waves. We then calculate the h^{HF}-matrix from Eq. (7.3) which we then diagonalize using library C++-functions. We now put the eigenvectors corresponding to the 4 lowest eigenvalues in the 4 first rows as we only deal with Beryllium. We can now calculate the energy, and use the modified diagonalized matrix as the C-matrix for next the iteration.

- When the energy difference between two iterations is low enough (determined by a preset tolerance $\sim 10^{-6}$) we stop the iterations and we have reached a desired set of coefficients for the new basis.

Table 7.1 shows the Hartree-Fock energy results for different number of s-waves included in the wave function expansion. It also shows the number of iterations needed for the energy difference between two succeeding iterations to be smaller than the chosen tolerance. We see from table 7.1 that the ground state energy for 15 s-waves produces approximately the same result as for 10 and 7 basis functions, but is improved in the 5th digit. The table also presents the exact result as given in [20]. [20]

We can compare the best Hartree-Fock result when only varying coefficients, $E = -14.5185\ E_h$, with the results from appendix A, $E = -14.573021\ E_h$, where Slater-type orbitals have been used. If we compare these with the exact result, $E_0 = -14.6674\ E_h$, it can lead us to believe that the Slater-type orbitals in the Roothaan-Hartree-Fock approach are a better choice for improving the VMC calculations.

In section 6.5.2 in [6], an argument is given for not using hydrogenic functions as basis functions for larger many-body atomic systems. They do not constitute a complete set by themselves, since the hydrogenic functions only describe bound states, and will not describe the unbound continuum states needed for a complete set. They also spread out quickly and become diffuse because of the $1/n$-factors in the polynomials and the exponentials. The Slater-type orbitals have a more compact exponential radial form, and we will avoid the problems involving diffuseness and the completeness of the basis set.

HF basis functions	Energy	Iterations
2	-13.7159	2
3	-14.5082	9
4	-14.5115	10
5	-14.5141	11
7	-14.5165	10
10	-14.5178	11
15	-14.5185	10
Exact	-14.6674	-

Table 7.1: The table shows results from different runs with different number of basis functions in addition to the exact result. Energies are given in the Hartree unit.

Chapter 8

VMC Results

In this chapter I will present the results produced by the Variational Monte Carlo code. The first part shows the variational plots for both sets of single particle orbitals, that means both hydrogen-like and Roothaan-Hartree-Fock orbitals. The next sections will be dedicated to finding the optimal parameters using the DFP-algorithm. The third part will present results of the time step extrapolation discussed in section 6.4 for the finely tuned parameter(s) produced by the DFP method. These results will also include the error estimates computed by the blocking method. The final part will compare the extrapolated results with calculations from [3] and [4].

All calculations are performed by the VMC code with the Metropolis-Hastings algorithm which includes importance sampling. There is a possibility that the minimum energy is found using different variational parameters for different time steps when analyzing the same system. However, by doing a simple check for helium, I could not find any significant difference for time steps $\Delta t = 0.05$ and $\Delta t = 0.01$. I then assume that the optimal parameters give the minimum in energy for all time steps within a reasonable range. Unless something else is stated, all results are calculated with time step $\Delta t = 0.05$.

8.1 Validation runs

In order to confirm that the Slater determinant is working properly, it is necessary to test the code. In most cases this means comparing the computed results with some closed form- or analytic expressions. For the hydrogen-like orbitals, I have compared with the closed form hydrogen energy solutions, while the Slater-type orbitals have been compared with the Roothaan-Hartree-Fock calculations (ref. [10]).

8.1.1 Hydrogen-like orbitals

To validate the Slater determinant with hydrogenic single particle orbitals, I run the program with no electron interaction. In this case this means using a pure Slater determinant, without the Jastrow factor, as the trial wave function. If we also neglect the electron-electron interaction energies, we can compare our results with the analytic results obtained using hydrogenic wave functions.

A hydrogen-like atom, is any atom with charge Z and only one electron. For this system the energy is given as

$$E_{NI} = -\frac{Z^2}{2}\frac{1}{n^2}, \tag{8.1}$$

where n is the principle quantum number of the orbital the single electron is occupying, ψ_{nlm} (see chapter 6.5.2 in ref. [6]).

For non-interacting atomic systems, we can use the hydrogen results to get a closed form expression for the non-interacting energy, E_{NI}. Since the electrons are non-interacting we can just add the energies of the orbitals as

$$E_{NI} = -\frac{Z^2}{2}\sum_{i=1}^{N}\frac{1}{n_i^2}, \tag{8.2}$$

where Z is the charge of the nucleus, N is the number of electrons, and n_i is the principal quantum number for the state occupied by electron i. The energy is in units of Hartree. Table 8.1 shows how the electrons for our systems of interest are distributed in the different energy orbitals.

Atom	Spectroscopic notation
Hydrogen	$(1s)$
Helium	$(1s)^2$
Beryllium	$(1s)^2(2s)^2$
Neon	$(1s)^2(2s)^2(2p)^6$
Magnesium	$(1s)^2(2s)^2(2p)^6(3s)^2$
Silicon	$(1s)^2(2s)^2(2p)^6(3s)^2(3p)^2$

Table 8.1: The table shows the electron configuration for hydrogen, helium, beryllium, neon, magnesium and silicon.

As an example we can calculate the analytic non-interacting helium energy as

$$E_{NI}(\text{He}) = -\frac{2^2}{2}\left(\frac{1}{1^2}+\frac{1}{1^2}\right) = -4. \tag{8.3}$$

By running the non-interacting version of the program for helium with, we get the exact energy

$$E = -4. \tag{8.4}$$

Table 8.2 shows that the exact energies calculated using Eq. (8.2) and the VMC results are identical for all atoms of interest. It is safe to say that the orbitals and the Slater determinant have been implemented correctly. By running a sufficient number of Monte Carlo cycles ($\approx 10^8$), the variance will reduce to

$$\sigma_E^2 = 10^{-5} - 10^{-6} \tag{8.5}$$

for all such non-interacting systems.

8.1.2 Slater-type orbitals - Hartree Fock results

To validate the Slater-type orbitals (STO) found by using Hartree-Fock in ref. [10], I ran the code with STO but no Jastrow factor. Then I could compare the computed energies with the energy results from ref. [10]. Table 8.3 shows this. We see how the energies correspond quite nicely for the smaller systems, while the deviations increase

Atom	E_{NI}	E
Helium	-4.00	-4.00
Beryllium	-20.00	-20.00
Neon	-200.00	-200.00
Magnesium	-304.00	-304.00
Silicon	-435.56	-435.56

Table 8.2: The table shows the analytic results, E_{NI}, the computed ground state energies, E, number of Monte Carlo cycles and the variance, σ_E^2. The energy results are in units of Hartree.

slightly for larger systems. A variational Monte Carlo process should in theory reproduce the Hartree-Fock energy when using the corresponding Hartree-Fock orbitals. We could have performed a time step extrapolation to hopefully get even closer to the Hartree-Fock energies, but the results in table 8.3 make us quite confident that the Slater-type orbitals form the Roothaan-Hartree-Fock calculations have been implemented correctly.

As an example I time extrapolated the results for neon, and got the value $E = -128.55$, which is in very good accordance with the HF energy from table 8.3.

Atom	E_{HF}	E	MC cycles	σ_E^2
Helium	-2.8616799	-2.8617823	$3 \cdot 10^7$	0.57097060
Beryllium	-14.573021	-14.573862	$3 \cdot 10^7$	2.9477341
Neon	-128.54705	-128.81981	$1 \cdot 10^8$	31.003216
Magnesium	-199.61461	-199.84100	$1 \cdot 10^8$	41.645320
Silicon	-288.85431	-289.11902	$5 \cdot 10^8$	58.958016

Table 8.3: The table shows the Roothaan-Hartree-Fock results, E^{HF} (see appendix A and ref. [10]), compared with the computed energies using the Roothaan STO, E, number of Monte Carlo cycles and the variance, σ_E^2. The energy results are in units of Hartree.

8.2 Variational plots

This section presents the plots of the variational runs for the different systems. For the hydrogen-like orbitals in section 8.2.1, the plots will be two-dimensional since the total wave function consists of two variational parameters. The parameters are α in the Slater determinant and β from the Jastrow function. For VMC calculations using the Slater-type orbitals presented in section 5.5, the plots will only be one-dimensional, since the single particle orbitals have already been optimized by the Roothaan-Hartree-Fock method. The variational parameter, β, comes from the Jastrow function.

8.2.1 Hydrogen-like orbitals

Figures 8.2, 8.3, 8.4, 8.5 and 8.6 show the variational results for helium, beryllium, neon, magnesium and silicon respectively. For the different variational calculations I have used between 10 and 100 million samples per parameter-set depending on the size of the system (smaller systems demand less computational time and we can more easily compute smoother plots). However, the main goal with these plots is to find the

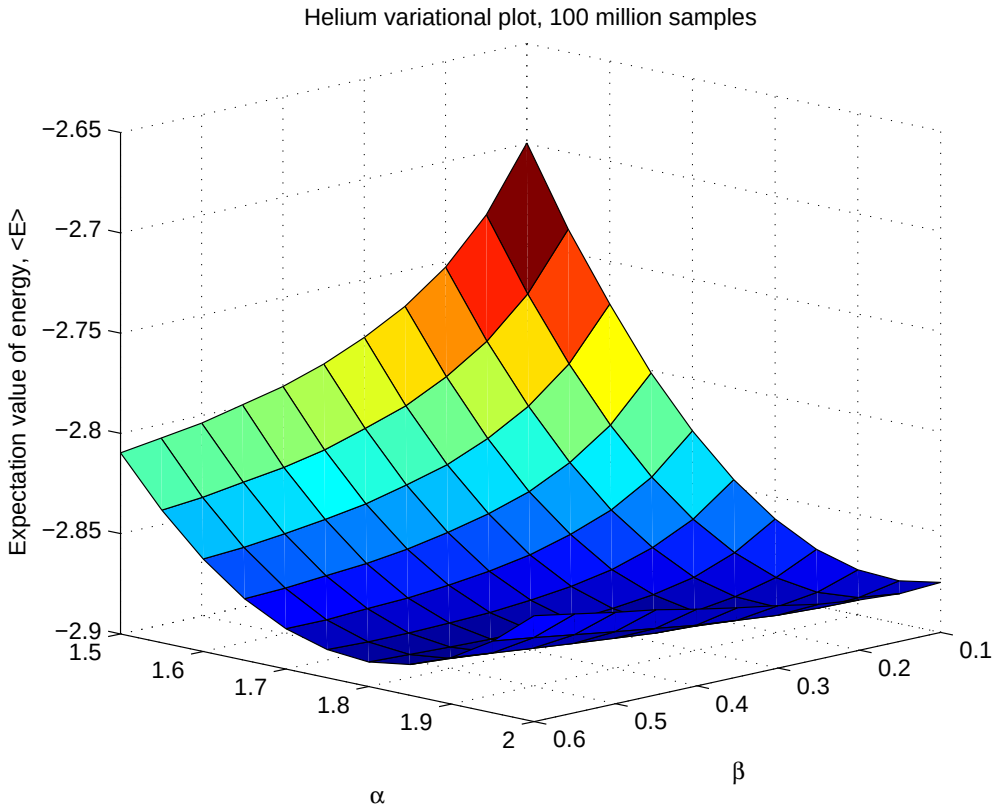

Figure 8.1: A typical variational plot. This plot shows the results for helium with 100 million samples. The energies are in units of Hartree.

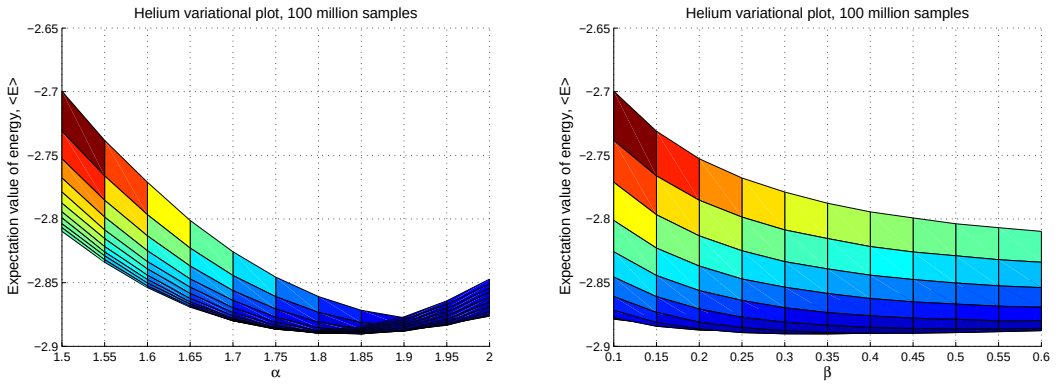

Figure 8.2: The figures show a variational plot of the helium ground state energy with 100 million samples using hydrogen-like orbitals. The figure on the left shows the variation of the α parameter, while the figure on the right shows the variation of the β parameter. The energies are in units of Hartree.

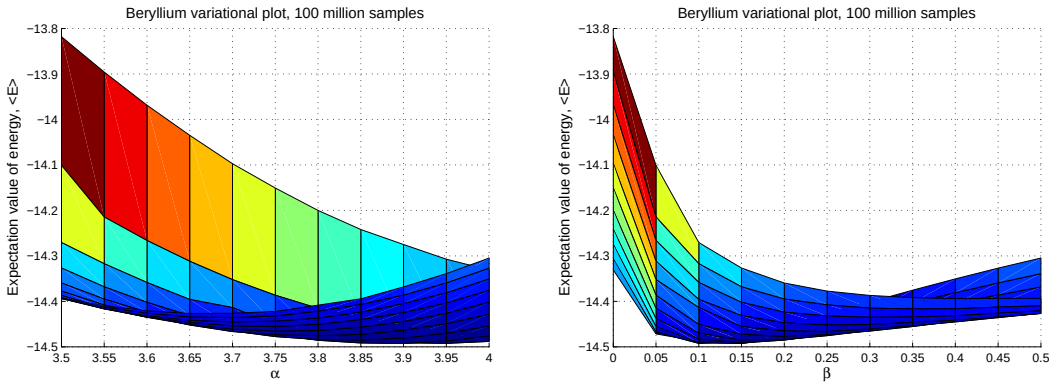

Figure 8.3: The figures show a variational plot of the beryllium ground state energy with 100 million samples using hydrogen-like orbitals. The figure on the left shows the variation of the α parameter, while the figure on the right shows the variation of the β parameter. The energies are in units of Hartree.

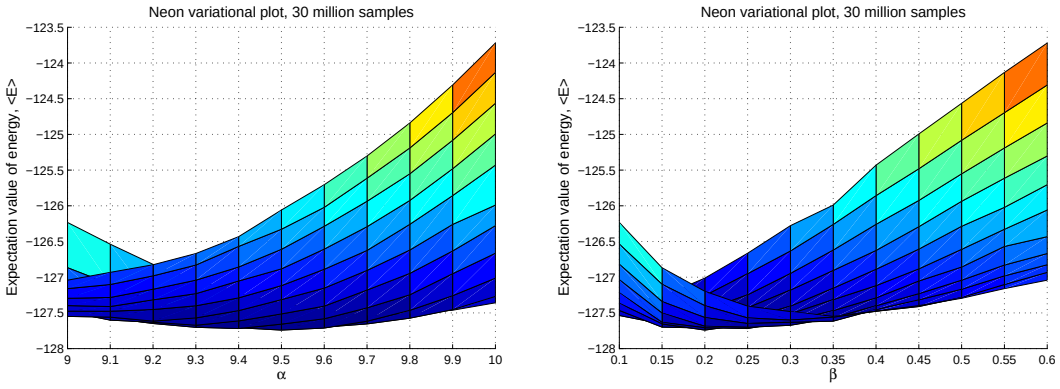

Figure 8.4: The figures show a variational plot of the neon ground state energy with 30 million samples using hydrogen-like orbitals. The figure on the left shows the variation of the α parameter, while the figure on the right shows the variation of the β parameter. The energies are in units of Hartree.

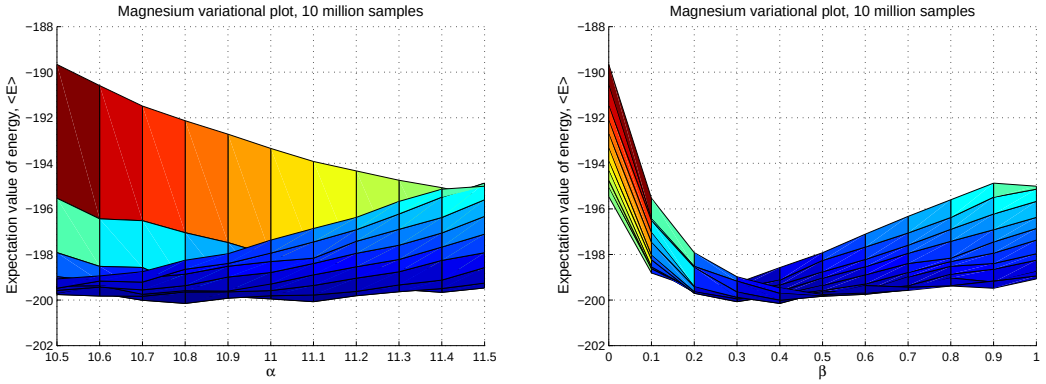

Figure 8.5: The figures show a variational plot of the magnesium ground state energy with 10 million samples using hydrogen-like orbitals. The figure on the left shows the variation of the α parameter, while the figure on the right shows the variation of the β parameter. The energies are in units of Hartree.

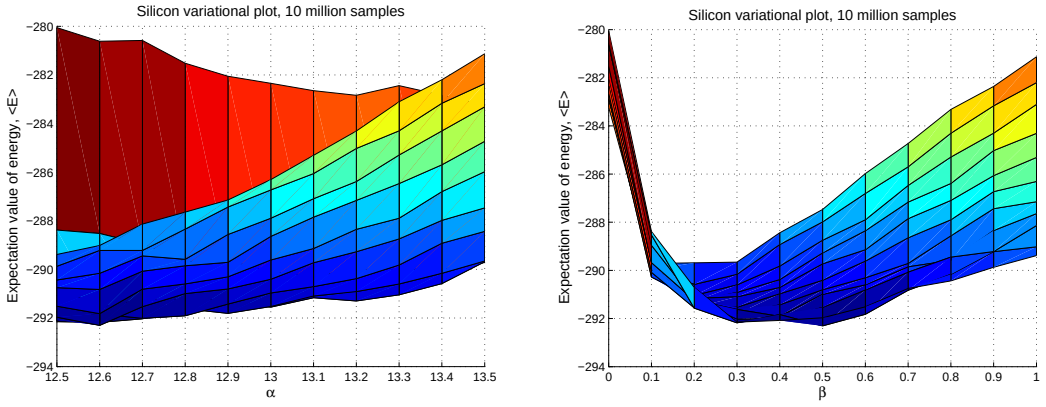

Figure 8.6: The figures show a variational plot of the silicon ground state energy with 10 million samples using hydrogen-like orbitals. The figure on the left shows the variation of the α parameter, while the figure on the right shows the variation of the β parameter. The energies are in units of Hartree.

whereabouts of the optimal parameters, and we are not concerned about precise energy measurements. The plots does not precise optimal parameters, but by merely looking at them, and where the energy has its minimum, we should be able to find very good initial values for the DFP-method (see appendix C). The DFP-method will hopefully supply us with even better **optimal** parameters using initial values from table 8.4.

Atom	α	β	MC cycles	E
Helium	1.85	0.35	10^8	≈ -2.89
Beryllium	3.90	0.12	10^8	≈ -14.49
Neon	9.20	0.50	$3 \cdot 10^7$	≈ -127.7
Magnesium	10.8	0.40	10^7	≈ -200
Silicon	12.6	0.50	10^7	≈ -292

Table 8.4: The table shows the apparent optimal parameters from the variational Monte Carlo plots. It also shows the number of Monte Carlo cycles and the approximate energy for the parameter set. The energies are in units of Hartree.

Table 8.4 shows the results for the variational parameters obtained by inspecting figures 8.2, 8.3, 8.4, 8.5 and 8.6. Some plots are very flat, e.g. the magnesium plot with respect to the α-parameter (see figure 8.5), and it is hard to determine the minimum by just inspecting the plots.

We also see that the silicon plot is difficult to read as a function of the α-parameter (see figure 8.6). It seems as though the minimum lies at about $\alpha = 12.6$, but this could be a local minimum since the plot does not show results for α-values less than 12.5. However, as we will see in section 8.3.1, the DFP-algorithm finds a minimum that lies within the plotted area.

To obtain smoother plots for both silicon and magnesium, I would have needed more samples, hence more computational time at the end of this work. These large parallel jobs tended to crash due to an error in the MPI (Message Passing Interface) system at Titan. When running the exact same code at a local computer, the program never crashed. However, the amount of computational time needed to get a *smooth* variational plot of such large systems would be very large on a local computer. Due to these time limitations and parallel jobs crashing at the Titan cluster (http://hpc.uio.no), I was not able to plot the silicon energy for parameter values lower than $\alpha = 12.5$. Again, as mentioned in the previous paragraph, the DFP method does return a minimum for α-values larger than 12.5, so the range of the plot will not be a problem.

8.2.2 Slater-type orbitals - Hartree Fock results

In this case, the trial wave function is reduced to a one-parameter function, since the Slater determinant has already been optimized by the Hartree-Fock method. When presenting the results, I will also include the variance plots to show that the parameters for minimum in the energy and minimum in the variance in general do not coincide. Since we have chosen the DFP method to search for the parameter that gives a minimum in the energy, this parameter value is the one we choose for further calculations.

Figures 8.7, 8.8, 8.9, 8.10 and 8.11 show the variation in the β parameter both for energy and variance for helium, beryllium, neon, magnesium and silicon respectively. We see that the minimum in energy and variance does not occur for the same β parameter. Especially the magnesium and silicon plots (figures 8.10 and 8.11) show this behavior,

with the minimum of variance not even being within the parameter interval we have plotted.

Again, the time limitation does not allow us to get perfectly smooth plots for larger systems, but the plots still supply us with good information concerning the parameter that gives a minimum in energy.

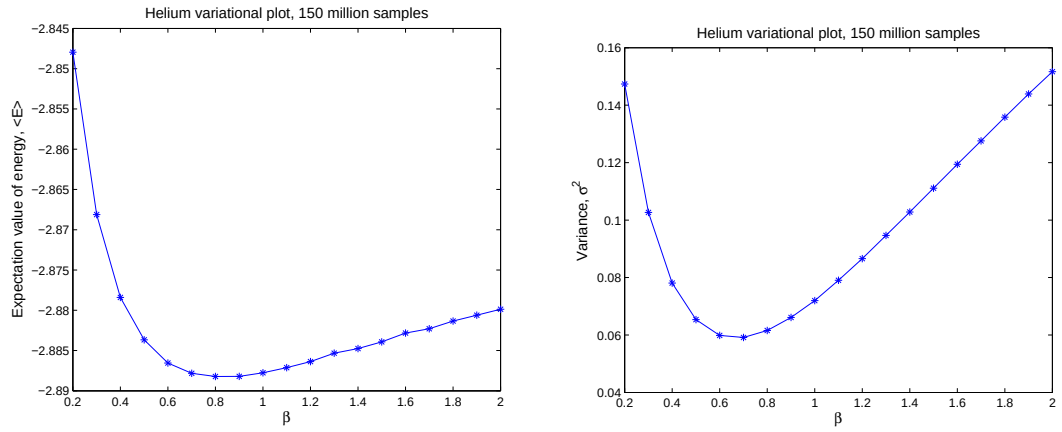

Figure 8.7: The figures show a variational plot of the helium ground state energy with 150 million samples using Slater type orbitals. The figure on the left shows the energy variation of the β parameter, while the figure on the right shows the variance variation of the β parameter. The energy results are in units of Hartree.

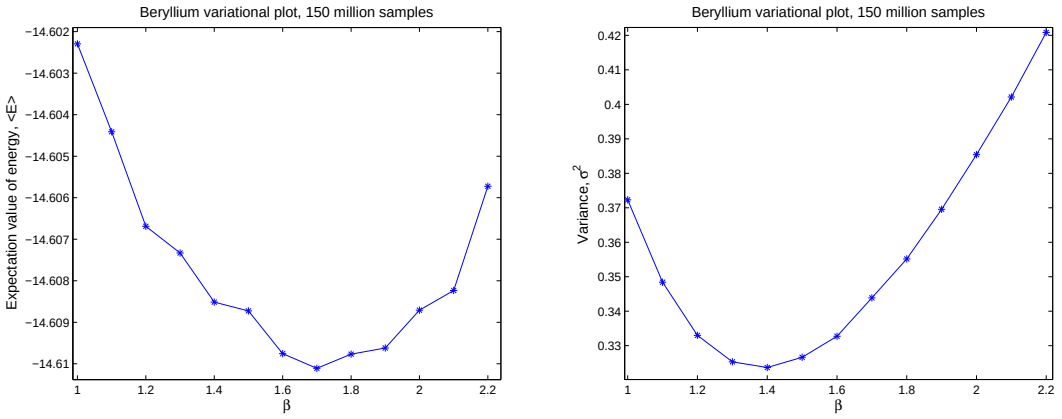

Figure 8.8: The figures show a variational plot of the beryllium ground state energy with 150 million samples using Slater type orbitals. The figure on the left shows the energy variation of the β parameter, while the figure on the right shows the variance variation of the β parameter. The energy results are in units of Hartree.

8.3 Optimal parameters with DFP

This section presents the results we get when using the minimization algorithm discussed in appendix C. The values in table 8.5 and 8.6 are average results of different DFP

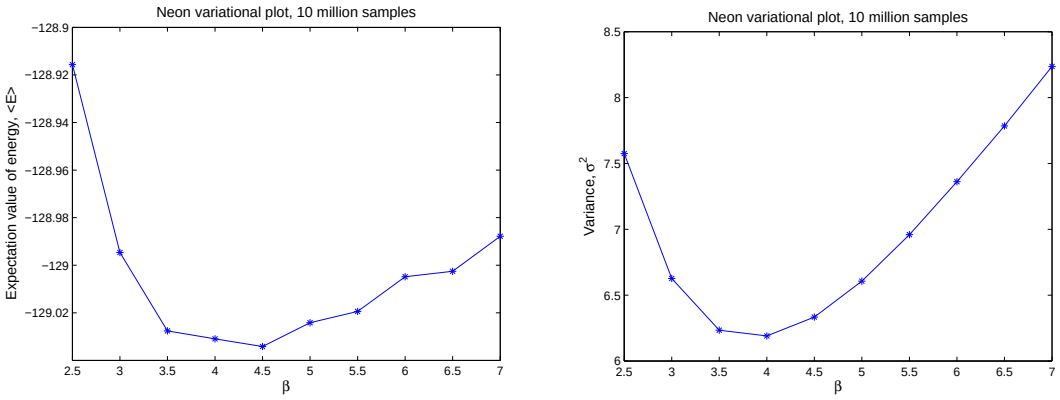

Figure 8.9: The figures show a variational plot of the neon ground state energy with 10 million samples using Slater type orbitals. The figure on the left shows the energy variation of the β parameter, while the figure on the right shows the variance variation of the β parameter. The energy results are in units of Hartree.

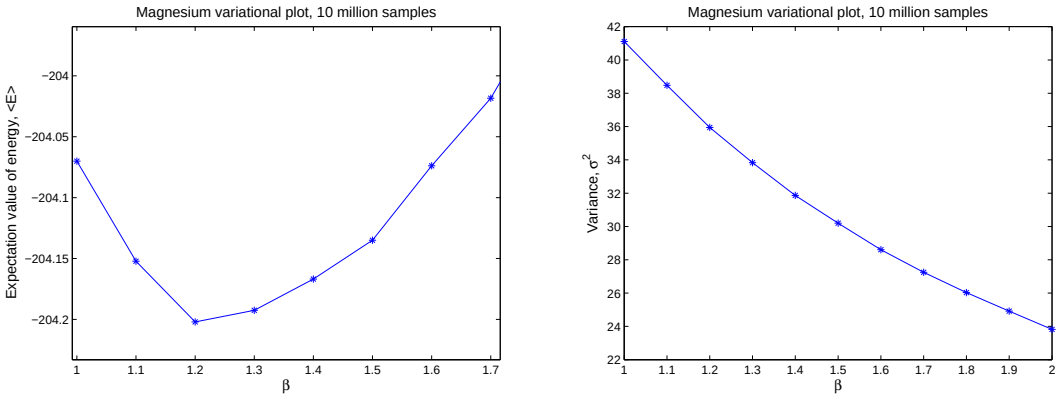

Figure 8.10: The figures show a variational plot of the magnesium ground state energy with 10 million samples using Slater type orbitals. The figure on the left shows the energy variation of the β parameter, while the figure on the right shows the variance variation of the β parameter. There is a big difference between energy and variance minimum. The energy results are in units of Hartree.

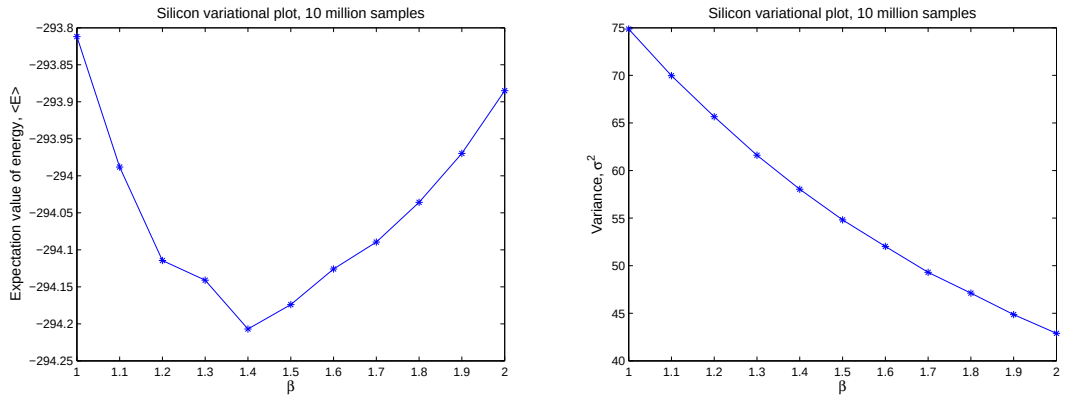

Figure 8.11: The figures show a variational plot of the silicon ground state energy with 10 million samples using Slater type orbitals. The figure on the left shows the energy variation of the β parameter, while the figure on the right shows the variance variation of the β parameter. There is a big difference between energy and variance minimum. The energy results are in units of Hartree.

minimization runs that all find their minimum for approximately the same parameters. The parameters usually differ in the second decimal so I decide to include three decimals in the final optimal parameters.

A problem with the minimization algorithm is that it sometimes makes the calculation crash. The reason is that the `dfpmin`-function includes inverting a matrix using LU-decomposition (see ref. [12]), and for some reason this matrix becomes singular. This happens more often for larger systems.

In this part I have used 10^7 Monte Carlo cycles for systems helium, beryllium and neon. For magnesium and silicon I unfortunately had to reduce this to 10^6 cycles since the calculations tended to crash more often due to the singular matrix problem.

8.3.1 Hydrogen-like orbitals

This part presents the results from the DFP algorithm when using hydrogen-like orbitals in the Slater determinant. All the results are given in table 8.5.

If we compare the apparent optimal parameters from figure 8.5 and the α_{min} and β_{min} in table 8.5

Atom	α_{start}	β_{start}	α_{min}	β_{min}	Average $\langle E \rangle$
Helium	1.85	0.35	1.839	0.348	-2.891
Beryllium	3.90	0.12	3.925	0.109	-14.493
Neon	9.50	0.20	9.546	0.177	-127.720
Magnesium	10.8	0.40	11.029	0.263	-200.217
Silicon	12.6	0.50	12.787	0.299	-292.319

Table 8.5: The table lists the apparent optimal parameters obtained from the plots in section 8.2.1and the average of the tuned DFP parameters resulting from several different DFP runs.

8.3.2 Slater-type orbitals - Hartree Fock results

Atom	β_{start}	β_{min}
Helium	0.8	0.811
Beryllium	1.7	1.697
Neon	4.50	4.527
Magnesium	1.20	1.379
Silicon	1.40	1.480

Table 8.6: The table lists the apparent optimal parameters obtained from the plots in section 8.2.2 and the average of the tuned DFP parameters resulting from several different DFP runs.

Table 8.6 shows the optimal parameters found by using the DFP algorithm. A noticeable thing is that the β parameter for helium, beryllium and neon is not changed significantly during the DFP minimizing process. For magnesium and silicon however, the DFP method does in fact find the energy minimum a bit further away from the minimum in the plots (figures 8.10 and 8.11). A major reason for this could obviously be that the plots for these systems are not smooth enough.

8.4 Time step analysis - extrapolated results

Through the calculations in the previous sections, we have now hopefully found the parameters that gives us the minimum in the energy. The final step in the calculations is then to calculate the energy for different time steps as discussed in section 6.4. These calculations will also implement the blocking method (see section 6.3) so that we can extrapolate to $\Delta t = 0$ for both the energy and the standard error of the data set.

Depending on the results of small test calculations on the different systems for different time steps, I have chosen three Δt values for each system that appear to produce sensible results. The values vary for the different systems and have been chosen from the values $\Delta t = 0.01$, $\Delta t = 0.025$, $\Delta t = 0.05$ and $\Delta t = 0.1$. Figure 8.13 shows all blocking results for the neon atom, for both hydrogenic wave functions and Slater-type orbitals.

8.4.1 Hydrogenic wave functions

In this part the results for the time step extrapolation is listed using the hydrogenic wave functions as single particle orbitals in the Slater determinant (see section 5.3). Tables 8.7, 8.8, 8.9, 8.10 and 8.11 shows the results for helium, beryllium, neon, magnesium and silicon respectively.

8.4.2 Slater-type orbitals

This part presents the time step extrapolated energy results when using Slater-type orbitals. Tables 8.12, 8.13, 8.14, 8.15 and 8.16 show the results for helium, beryllium, neon, magnesium and silicon respectively.

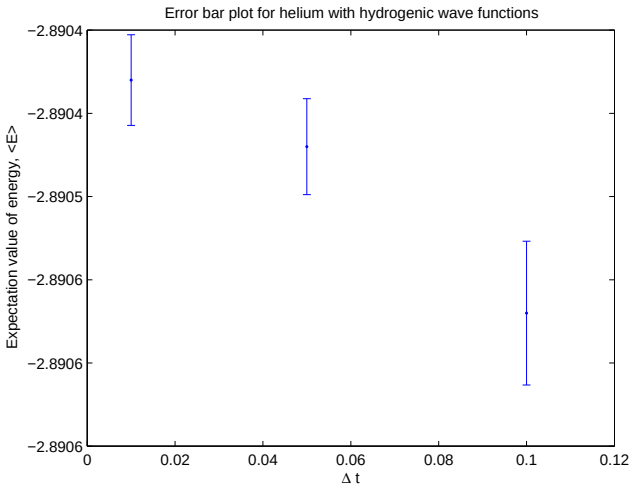

Figure 8.12: This figure shows the time step extrapolation for helium with hydrogenic wave function with error bars corresponding to the blocking errors.

Helium with hydrogenic wave functions

Δt	Energy	Standard error
0.10	-2.89057	$2.7 \cdot 10^{-5}$
0.05	-2.89047	$2.9 \cdot 10^{-5}$
0.01	-2.89043	$4.3 \cdot 10^{-5}$
Extrapolated	-2.89040	$3.9 \cdot 10^{-5}$

Table 8.7: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for helium with hydrogen-like wave functions.

Beryllium with hydrogenic wave functions

Δt	Energy	Standard error
0.05	-14.49349	$9.6 \cdot 10^{-5}$
0.25	-14.49852	$1.2 \cdot 10^{-4}$
0.01	-14.49960	$1.9 \cdot 10^{-4}$
Extrapolated	-14.50220	$1.8 \cdot 10^{-4}$

Table 8.8: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for beryllium with hydrogen-like wave functions.

Neon with hydrogenic wave functions

Δt	Energy	Standard error
0.05	-127.732	$5.4 \cdot 10^{-3}$
0.25	-127.437	$2.6 \cdot 10^{-3}$
0.01	-127.371	$1.7 \cdot 10^{-3}$
Extrapolated	-127.284	$2.5 \cdot 10^{-3}$

Table 8.9: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for neon with hydrogen-like wave functions.

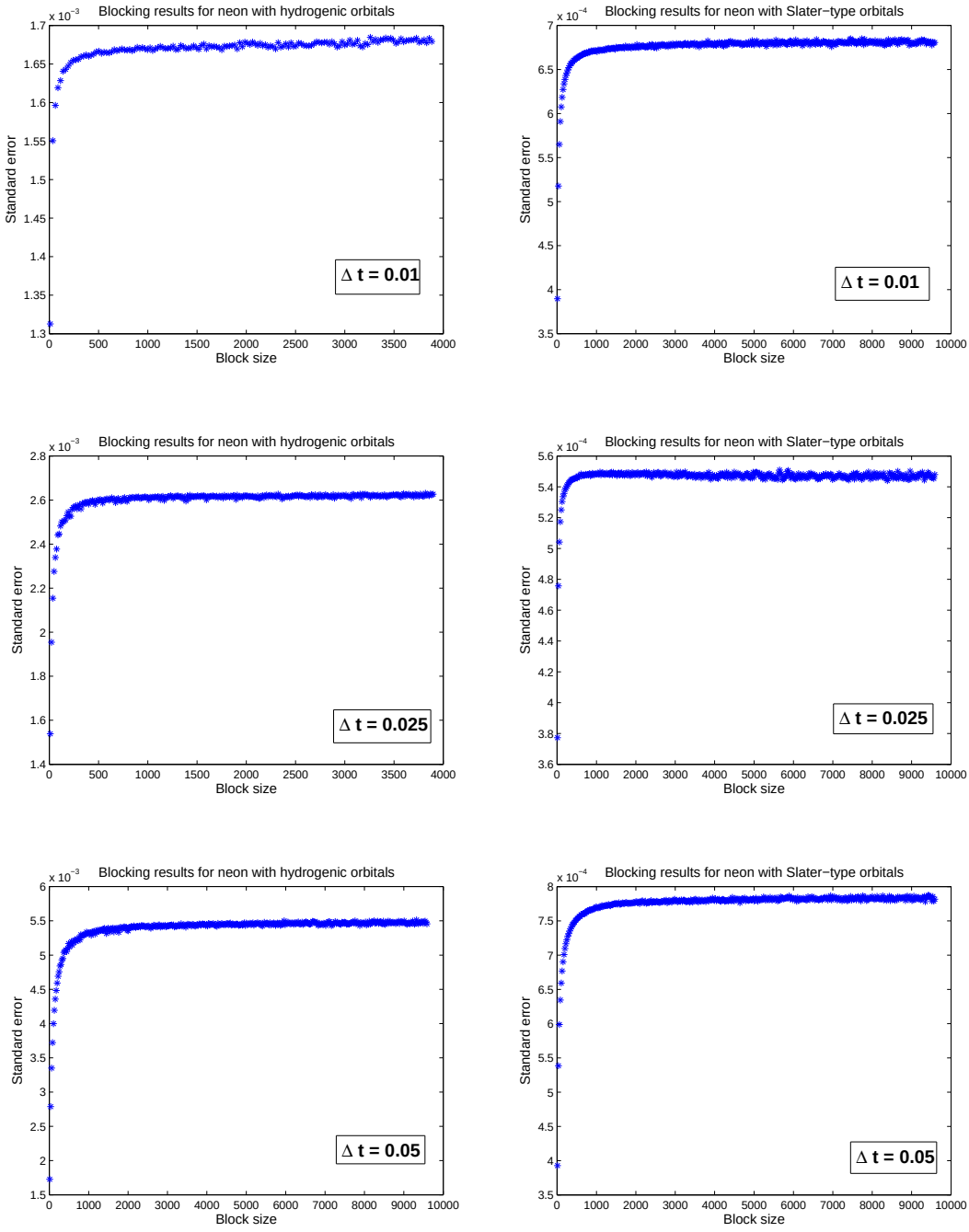

Figure 8.13: The figures show the blocking results for neon for both hydrogenic wave functions (left) and the Slater-type orbitals (right). The standard error as a function of block size is in units of Hartree.

Magnesium with hydrogenic wave functions

Δt	Energy	Standard error
0.05	-199.85	$2.1 \cdot 10^{-2}$
0.25	-197.530	$8.1 \cdot 10^{-3}$
0.01	-196.818	$3.8 \cdot 10^{-3}$
Extrapolated	-196.152	$6.6 \cdot 10^{-3}$

Table 8.10: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for magnesium with hydrogen-like wave functions.

Silicon with hydrogenic wave functions

Δt	Energy	Standard error
0.1	-298.00	$5.9 \cdot 10^{-2}$
0.05	-291.87	$4.2 \cdot 10^{-2}$
0.01	-284.352	$6.7 \cdot 10^{-3}$
Extrapolated	-282.750	$9.5 \cdot 10^{-3}$

Table 8.11: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for silicon with hydrogen-like wave functions.

Helium with Slater-type orbitals

Δt	Energy	Standard error
0.1	-2.88876	$2.3 \cdot 10^{-5}$
0.05	-2.88833	$3.1 \cdot 10^{-5}$
0.01	-2.88783	$6.8 \cdot 10^{-5}$
Extrapolated	-2.88782	$5.0 \cdot 10^{-5}$

Table 8.12: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for helium with Slater-type orbitals.

Beryllium with Slater-type orbitals

Δt	Energy	Standard error
0.1	-14.61260	$5.3 \cdot 10^{-5}$
0.05	-14.60941	$6.0 \cdot 10^{-5}$
0.01	-14.6088	$1.2 \cdot 10^{-4}$
Extrapolated	-14.60725	$9.5 \cdot 10^{-5}$

Table 8.13: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for beryllium with Slater-type orbitals.

Neon with Slater-type orbitals

Δt	Energy	Standard error
0.05	-128.9343	$7.8 \cdot 10^{-4}$
0.025	-128.7021	$5.5 \cdot 10^{-4}$
0.01	-128.6164	$6.4 \cdot 10^{-4}$
Extrapolated	-128.5187	$7.7 \cdot 10^{-4}$

Table 8.14: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for neon with Slater-type orbitals.

Magnesium with Slater-type orbitals

Δt	Energy	Standard error
0.05	-204.174	$1.9 \cdot 10^{-3}$
0.025	-204.482	$1.7 \cdot 10^{-3}$
0.01	-204.737	$2.1 \cdot 10^{-3}$
Extrapolated	-204.852	$2.3 \cdot 10^{-3}$

Table 8.15: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for neon with Slater-type orbitals.

Silicon with Slater-type orbitals

Δt	Energy	Standard error
0.05	-294.180	$2.6 \cdot 10^{-3}$
0.025	-294.665	$2.2 \cdot 10^{-3}$
0.01	-295.131	$2.4 \cdot 10^{-3}$
Extrapolated	-295.315	$2.8 \cdot 10^{-3}$

Table 8.16: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for silicon with Slater-type orbitals.

8.5 Discussions

In this section we compare our VMC results for helium, beryllium and neon with the VMC results from table 2.4 in [3]. For magnesium and silicon we compare with the results from table I in [4].

Table 8.17 lists the energy results compared to our reference energies. We see how for our case the helium energy is best calculated with the hydrogenic wave functions, rather than with the Slater-type orbitals. It seems as though the hydrogenic orbitals model the helium atom better than the Slater-type orbitals. But as we shall see for larger systems with more correlation, the Slater-type orbitals gives us an advantage.

For beryllium, the Slater-type orbitals seem to model the correlation between the electrons more correctly, and produce a more correct ground state energy estimation than by using hydrogenic wave functions. It seems as though the Jastrow factor is good for modelling the correlations that the Hartree-Fock optimization of wave functions has not already taken care of.

If we are to compare with the energies from [4], we clearly see that the Slater determinants with hydrogenic wave functions do not estimate the silicon energy in a good way. A good argument for this is the fact that we have chosen a Slater determinant as if the two outmost electrons were bound to a $3p$-state with $m_l = +1$, corresponding to the real solid harmonics being coordinate x. This is because our Slater determinant code is restricted to describe atoms with an equal number of spin up electrons as spin down electrons. Since the Hamiltonian does not interact with spin or orbital momentum, we simply choose $m_l = +1$, disregarding Hund's rules (see [2]). As seen in section 3.1.3 however, the two outermost electrons in silicon both have spin up and have a total angular momentum of one. The electrons could be configured in different ways, e.g. one in a $3p$-orbital with $m_l = +1$ and one in $3p$ with $m_l = 0$, both with spin up. The electrons could also be excited into $3d$-orbitals etc. To accommodate for this, we would have needed to implement linear combinations of Slater determinants and also include more single particle wave functions. Because of this, we see that our wave function probably is too primitive for this matter.

A troubling aspect however, is the behavior of the the extrapolated energies for magnesium and silicon using a Slater determinant with the Slater-type orbitals obtained by the Hartree-Fock calculations by Clementi and Roetti in [10]. Our VMC calculations actually give too attractive energies compared to the calculations in [4]. But if we

Energy comparison			
Atom	Reference energy	Extr. energy with HWF	Extr. energy with STO
Helium	−2.903726(4)	−2.89040(3.9)	−2.88722(5.0)
Beryllium	−14.6664(3)	−14.50220(1.8)	−14.60725(9.5)
Neon	−128.884(4)	−127.284(2.5)	−128.5187(7.7)
Magnesium	−200.0002(5)	−196.152(6.6)	−204.852(2.3)
Silicon	−289.1970(10)	−282.750(9.5)	−295.315(2.8)

Table 8.17: The table shows the energies from the time step extrapolation, and standard errors from the blocking method for neon with Slater-type orbitals. Helium, beryllium and neon energies are compared with VMC calculations from [3], while magnesium and silicon energies have been compared with VMC results from [4]. Energies are in units of Hartree.

consider the VMC calculations from [4] to be accurate, our results contradict with the variational principle (see section 4.3). If this is the case, there must be something wrong with the code that we have developed. However, we have seen in table 8.3 that our Slater determinant reproduces the Roothaan-Hartree-Fock energies from [10] quite nicely. We also saw the same when calculating a non-interacting system using a Slater determinant with hydrogenic orbitals, so we might be wrong to question the Slater determinant at this point.

By running the code with only a Slater determinant with STO and no Jastrow factor, we reproduce the Hartree-Fock energies. These energies are already very close to the results from [4], so we should not expect any significant improvement from an additional correlation factor. However, the energies are somehow underestimated when including the Jastrow factor.

We cannot overlook the possibility that the Jastrow function we have chosen in this work (see 5.1.2) has not been implemented correctly, but this still seems somewhat unlikely as it produces good improvements when working with hydrogenic wave functions. With that being said, we see a trend for the larger systems neon, magnesium and silicon when we include the Jastrow factor. The calculations then tend to produce less accurate results. For magnesium and silicon the energies are much too low for our liking, while for neon the Jastrow factor hardly improves the Hartree-Fock energy. In these three cases, the pure Hartree-Fock (Slater-type orbitals) Slater determinant without the Jastrow functions produces the best results.

Chapter 9

Conclusion

In this thesis we have performed Variational Monte Carlo calculations on the helium, beryllium, neon, magnesium and silicon atoms in order to estimate the ground state energy of the systems. We have developed a code in C++ that uses either hydrogen-like wave functions or Slater-type orbitals as single-particle wave functions. The Slater-type orbitals have been optimized by a Hartree-Fock calculation in [10]. The code uses the Metropolis-Hastings algorithm (see e.g. [3]) with importance sampling.

We have seen that the code that we have developed produces very good results for helium and beryllium in addition to sensible results for neon. For magnesium and silicon however, we have struggled to get good results. We clearly see that when introducing the Jastrow factor in addition to Slater-type orbitals, the energies are underestimated compared to the VMC calculations from [4].

Despite the problems we have encountered when dealing with larger systems, we still see that the Variational Monte Carlo machinery we have developed is working. By reviewing the results from the pure Slater determinant calculations (without any Jastrow factor), we see that the results come out perfectly.

In order to improve this model we can examine more closely the Jastrow function we have used in this work to see whether or not there could be an error in the implementation. Another approach would be to introduce other Jastrow factors in order to model the correlation more correctly, such as the ones used in [4].

To improve calculations for silicon and other open-shell systems, we can also implement the possibility of having linear combinations of Slater determinants. We see from the discussion in 8.5 that for these systems, the Slater determinant we have implemented might not be good enough to describe the physics.

When we get our full VMC machinery up and running, with correlations and all, the next step would be to perform a Green's function Monte Carlo calculation (see [3]). In order to perform Green's function Monte Carlo (GFMC) calculation, the method needs a starting point for the energy. An optimized Variational Monte Carlo energy is a good choice for this. A GFMC calculation is in theory an exact method, and will result in us being able to determine an accurate ground state wave function. With this we can define the quantum mechanical density, which can be used to construct a density functional for atoms using the adiabatic-connection method described in [21]. By starting with ab initio calculations, we can then hopefully improve the density functionals that are used to model materials.

Appendix A

Roothaan-Hartree-Fock results

In this part I will present Clementi and Roetti's results from their Roothaan-Hartree-Fock calculations for atoms helium, beryllium, neon, magnesium and silicon. The tables' applications are explained in section 5.5.

Helium

Helium consists of 2 electrons and a core of charge $2e$ and has the electron distribution $(1s)^2$. Table A.1 shows the Roothaan-Hartree-Fock(RHF) solution for Helium. The

n,λ	Exponent, ξ	$1s$ exp.coeff.
1S	1.41714	0.76838
1S	2.37682	0.22356
1S	4.39628	0.04082
1S	6.52699	−0.00994
1S	7.94525	0.00230

Table A.1: The table shows the Roothan-Hartree-Fock results for Helium.

tabulated Roothaan-Hartree-Fock ground state energy for Helium given in [10] is

$$E = -2.8616799\ E_h,$$

where E_h is the *Hartree* unit given in section 3.1.4.

Beryllium

Beryllium has 4 electrons with a nucleus charged $4e$ and has electrons distributed as $(1s)^2(2s)^2$. Table A.2 shows the RHF-solution for beryllium. In [10] the total ground state energy is calculated as

$$E = -14.573021\ E_h.$$

Neon

The ten Neon electrons are distributed as $(1s)^2(2s)^2(2p)^6$. Table A.3 shows the RHF-solution. The Roothaan-Hartree-Fock energy is

$$E = -128.54705\ E_h.$$

n,λ	Exponent, ξ	$1s$ exp.coeff.	$2s$ exp.coeff.
1S	3.47116	0.91796	−0.17092
1S	6.36861	0.08724	−0.01455
2S	0.77820	0.00108	0.21186
2S	0.94067	−0.00199	0.62499
2S	1.48725	0.00176	0.26662
2S	2.71830	0.00628	−0.09919

Table A.2: The table shows the Roothan-Hartree-Fock results for Beryllium.

n,λ	Exponent, ξ	$1s$ exp.coeff.	$2s$ exp.coeff.	n,λ	Exponent, ξ	$2p$ exp.coeff.
1S	9.48486	0.93717	−0.23093	2P	1.45208	0.21799
1S	15.56590	0.04899	−0.00635	2P	2.38168	0.53338
2S	1.96184	0.00058	0.18620	2P	4.48489	0.32933
2S	2.86423	−0.00064	0.66899	2P	9.13464	0.01872
2S	4.82530	0.00551	0.30910			
2S	7.79242	0.01999	−0.13871			

Table A.3: The table shows the Roothan-Hartree-Fock results for Neon.

Magnesium

The Magnesium atom has 12 electrons in orbitals $(1s)^2(2s)^2(2p)^6(3s)^2$. The RHF-solution is given in table A.4. The RHF-energy is given in [10] as

n,λ	Exponent, ξ	$1s$ e.c.	$2s$ e.c.	$3s$ e.c.	n,λ	Exponent, ξ	$2p$ e.c.
1S	12.01140	0.96430	−0.24357	0.04691	2P	5.92580	0.52391
3S	13.91620	0.03548	−0.00485	0.00144	4P	7.98979	0.07012
3S	9.48612	0.02033	0.08002	−0.01850	4P	5.32964	0.31965
3S	6.72188	−0.00252	0.39902	−0.07964	4P	3.71678	0.20860
3S	4.24466	0.00162	0.57358	−0.13478	4P	2.59986	0.03888
3S	2.53466	−0.00038	0.05156	−0.01906			
3S	1.46920	0.00015	−0.00703	0.48239			
3S	0.89084	−0.00004	0.00161	0.60221			

Table A.4: The table shows the Roothan-Hartree-Fock results for Magnesium.

$$E = -199.61461\ E_h.$$

Silicon

The silicon atom consists of a nucleus with charge $14e$ and 14 electrons distributed in orbitals as $(1s)^2(2s)^2(2p)^6(3s)^2(3p)^2$. Table A.5 shows the s-orbitals for the RHF-solution. Table A.6 shows the RHF results for the p-orbitals. The ground state energy calculated with the RHF-method gives

$$E = -288.85431\ E_h.$$

n,λ	Exponent, ξ	$1s$ exp.coeff.	$2s$ exp.coeff.	$3s$ exp.coeff.
1S	14.01420	0.96800	−0.25755	0.06595
3S	16.39320	0.03033	−0.00446	0.00185
3S	10.87950	0.02248	0.11153	−0.03461
3S	7.72709	−0.00617	0.40339	−0.10378
3S	5.16500	0.00326	0.55032	−0.19229
3S	2.97451	−0.00143	0.03381	−0.06561
3S	2.14316	0.00081	−0.00815	0.59732
3S	1.31306	−0.00016	0.00126	0.55390

Table A.5: The table shows the s-orbital Roothan-Hartree-Fock results for Silicon.

n,λ	Exponent, ξ	$2p$ exp.coeff.	$3p$ exp.coeff.
2P	7.14360	0.54290	−0.11535
4P	16.25720	0.00234	−0.00189
4P	10.79720	0.04228	−00473
4P	6.89724	0.32155	−0.07552
4P	4.66598	0.22474	0.01041
4P	2.32046	0.00732	0.46075
4P	1.33470	−0.00105	0.57665
4P	0.79318	0.00041	0.06274

Table A.6: The table shows the p-orbital Roothan-Hartree-Fock results for Silicon.

Appendix B

Statistics

This appendix will give a short introduction to many key aspects regarding the statistical terms used for calculating expectation values and its statistical error in the Variational Monte Carlo process. We will follow the work in [22] quite closely in order to get all the important points across.

A probability distribution function(PDF), $p(x)$, is a function that describes the frequency of a stochastic value X to occur, that is

$$p(x) = \text{Prob}(X = x). \tag{B.1}$$

This is for the discrete case.

For the continuous case, $p(x)$ represents a probability distribution, where we must deal with probabilities of values being within some interval. That is, the probability of the stochastic variable X taking a value on the finite interval $[a, b]$ is defined by an integral

$$\text{Prob}(a \leq X \leq b) = \int_a^b p(x)dx. \tag{B.2}$$

All PDF's must take values that are both positive and less or equal to unity, the absolute maximum for any probability function to make sense. This means

$$0 \leq p(x) \leq 1, \tag{B.3}$$

In addition to this, both the discrete and continuous PDF's must be normalized to one as

$$\sum_i p(x_i) = 1, \tag{B.4}$$

$$\int p(x)dx = 1. \tag{B.5}$$

Moments and expectation values

Since our wave functions are continuous, our PDF will be so as well, and the expectation value of a function $f(x)$ is given by

$$\langle f \rangle \equiv f(x)p(x)dx. \tag{B.6}$$

We also have the *moments*, as special case of expectation values, where the n-th moment is

$$\langle x^n \rangle \equiv \int x^n p(x)dx. \tag{B.7}$$

The *mean*, μ, is defined by the first moment, $\langle x \rangle$.

The *central moments* are defined by

$$\langle (x - \langle x \rangle)^n \rangle \equiv \int (x - \langle x \rangle)^n p(x) dx. \tag{B.8}$$

The second central moment is called the *variance*, σ_X^2 or $\mathrm{Var}(X)$, and can be written as

$$\begin{aligned} \sigma_X^2 = \mathrm{Var}(X) &= \langle (x - \langle x \rangle)^2 \rangle \\ &= \langle x^2 \rangle - \langle x \rangle^2. \end{aligned} \tag{B.9}$$

The standard deviation is then $\sigma = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$.

Correlated measurements

For now, we have been dealing with functions of one stochastic variable, so-called *univariate* PDF's. However, a PDF may as well consist of many variables. These PDF's are called *multivariate* PDF's. The variables themselves are independent, or *uncorrelated* is the multivariate PDF, $P(x_1, x_2, ..., x_n)$ can be factorized as

$$P(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} p_i(x_i), \tag{B.10}$$

where $p_i(x_i)$ are the univariate PDF's.

The so-called *covariance* of two stochastic variables, X_i and X_j, is defined as

$$\mathrm{Cov}(X_i, X_j) \equiv \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle. \tag{B.11}$$

We see that for $j = i$ this reduces to the variance of X_i, that is

$$\mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i) = \langle x_i^2 \rangle - \langle x_i \rangle^2. \tag{B.12}$$

If the variables X_i and X_j are independent, or uncorrelated, the PDF's will factorize as in Eq. (B.10), and we will have $\langle x_i x_j \rangle = \langle x_i \rangle \langle x_j \rangle$. This will give $\mathrm{Cov}(X_i, X_j) = 0$, when $i \neq j$.

Consider linear combinations of stochastic variables X_i and Y_j as

$$U = \sum_i a_i X_i, \tag{B.13}$$

and

$$V = \sum_j b_j Y_j, \tag{B.14}$$

where a_i and b_j are scalar coefficients. As given in [23], we will have

$$\mathrm{Cov}(U, V) = \sum_{i,j} a_i b_j \mathrm{Cov}(X_i, Y_j). \tag{B.15}$$

Since $\mathrm{Var}(X_i) = \mathrm{Cov}(X_i, X_i)$, the variance of U will be

$$\mathrm{Var}(U) = \sum_{i,j} a_i a_j \mathrm{Cov}(X_i, X_j), \tag{B.16}$$

which for uncorrelated variables will be reduced to

$$\mathrm{Var}(U) = \sum_i a_i^2 \mathrm{Cov}(X_i, X_i) = \sum_i a_i^2 \mathrm{Var}(X_i), \tag{B.17}$$

and

$$\mathrm{Var}\left(\sum_i a_i X_i)\right) = \sum_i a_i^2 \mathrm{Var}(X_i). \tag{B.18}$$

Approximations for finite data sets

Consider a computational experiment which produces a sequence of n stochastic values

$$\{x_1, x_2, \ldots, x_n\}, \tag{B.19}$$

which we will call the n *measurements* of our *sample*. We can hereby define the *sample mean* as

$$\bar{x}_n \equiv \frac{1}{n}\sum_{k=1}^{n} x_k, \tag{B.20}$$

with the *sample variance* and *sample covariance* being

$$\mathrm{Var}(x) \equiv \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x}_n)^2 \tag{B.21}$$

and

$$\mathrm{Cov}(x) \equiv \frac{1}{n}\sum_{kl}(x_k - \bar{x}_n)(x_l - \bar{x}_n), \tag{B.22}$$

respectively. This sample covariance is a measure of the correlation between succeeding values in the sample. These values are not the same as the mean μ_X, $\mathrm{Var}(X)$ and $\mathrm{Cov}(X)$ defined by the exact PDF, $p_X(x)$, but approximations to these quantities.

In the limit where $n \to \infty$ it can be shown that the sample mean, $\bar{x}_n$ approaches the true mean μ_X, that is

$$\lim_{n\to\infty} \bar{x}_n = \mu_X, \tag{B.23}$$

and $\bar{x}_n$ can be seen as an estimate of μ_X. But how good is this approximation? In order to find this we also need the error estimate of our measurement. We can view the sample means themselves as measurements in a collection of sample means (several experiments). To calculate the statistical error, we need the PDF of the sample means, $p_{\bar{X}_n}(x)$. The exact error will be given by the standard deviation of the sample means, err_X, also called sample error. However, we don't know the exact PDF, so we can only calculate an estimate of err_X.

The sample mean, $\bar{x}_n$, can be treated as a stochastic variable $\bar{X}_n$, because $\bar{x}_n$ is a linear combination of stochastic variables X_i with $1/n$ as common coefficients. This means we can write

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{B.24}$$

We will now have $p_{\bar{X}}(x)$ as the probability distributions of the sample means $\bar{X}_n$, but we cannot get a closed form expression for this. However, the *central limit theorem* (see [23]) opens for an approximation of $p_{\bar{X}}(x)$ as $n \to \infty$. It states that

$$\lim_{n\to\infty} p_{\bar{X}}(x) = \left(\frac{n}{2\pi Var(X)}\right)^{1/2} e^{-\frac{n(x-\bar{x}_n)^2}{2Var(X)}}, \tag{B.25}$$

which means that the PDF will approach a Gaussian distribution. This Gaussian distribution will have a mean and variance which equal to the true mean and variance, μ_X and σ_X^2.

The error is given by

$$\mathrm{err}_X^2 = \mathrm{Var}(\bar{X}_n) = \frac{1}{n^2}\sum_{ij}\mathrm{Cov}(X_i, X_j). \tag{B.26}$$

This exact error needs the true means μ_X in order to be calculated. We will not be able to do this unless we have the exact PDF's of the variables X_i. However, we only have the measurements in a sample, and not the actual PDF.

Instead, we estimate the true means μ_X by the sample means

$$\mu_{X_i} = \langle x_i \rangle \approx \frac{1}{n}\sum_{k=1}^{n} x_k = \langle x \rangle. \tag{B.27}$$

The estimate of the covariance will then be

$$\begin{aligned}\mathrm{Cov}(X_i, X_j) &= \langle (x_i - \langle x_i\rangle)(x_j - \langle x_j\rangle)\rangle \approx \langle (x_i - \bar{x})(x_j - \bar{x})\rangle \\ &\approx \frac{1}{n}\sum_{l}^{n}\left(\frac{1}{n}\sum_{k}^{n}(x_k - \bar{x}_n)(x_l - \bar{x}_n)\right) = \frac{1}{n}\frac{1}{n}\sum_{kl}(x_k - \bar{x}_n)(x_l - \bar{x}_n) \\ &= \frac{1}{n}\mathrm{Cov}(x).\end{aligned} \tag{B.28}$$

The error estimate can be written using Eq. (B.28) as

$$\begin{aligned}\mathrm{err}_X^2 &= \frac{1}{n^2}\sum_{ij}\mathrm{Cov}(X_i, X_j) \\ &\approx \frac{1}{n^2}\sum_{ij}\frac{1}{n}\mathrm{Cov}(x) = \frac{1}{n^2}n^2\frac{1}{n}\mathrm{Cov}(x) \\ &= \frac{1}{n}\mathrm{Cov}(x).\end{aligned} \tag{B.29}$$

The same approximation goes for the variance, as

$$\begin{aligned}\mathrm{Var}(X_i) = &= \langle (x_i - \langle x_i\rangle)^2\rangle \approx \langle (x_i - \bar{x})^2\rangle \\ &= \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x}_n)^2 \\ &= \mathrm{Var}(x),\end{aligned} \tag{B.30}$$

which for uncorrelated stochastic variables will give an error estimation as

$$\begin{aligned}\mathrm{err}_X^2 &= \frac{1}{n^2}\sum_{ij}\mathrm{Cov}(X_i, X_j) \\ &= \frac{1}{n^2}\sum_{i}\mathrm{Var}(X_i) \approx \frac{1}{n^2}\sum_{i}\mathrm{Cov}(x) \\ &= \frac{1}{n}\mathrm{Var}(x).\end{aligned} \tag{B.31}$$

Autocorrelation function and correlation time

As given in [22], we can rewrite the error as

$$\begin{aligned}\mathrm{err}_X^2 &= \frac{1}{n}\mathrm{Cov}(x) = \frac{1}{n}\mathrm{Var}(x) + \frac{1}{n}\left(\mathrm{Cov}(x) - \mathrm{Var}(x)\right) \\ &= \frac{1}{n^2}\sum_{k=1}^{n}(x_k - \bar{x}_n)^2 + \frac{2}{n^2}\sum_{k<l}(x_k - \bar{x}_n)(x_l - \bar{x}_n). \end{aligned} \tag{B.32}$$

We see the first term is just the variance, so the second term must be the error correction when the variables are correlated. By now introducing partial sums on the form

$$f_d = \frac{1}{n}\sum_{k=1}^{n-d}(x_k - \bar{x}_n)(x_{k+d} - \bar{x}_n), \tag{B.33}$$

we can write the correlation term as

$$\frac{2}{n^2}\sum_{k<l}(x_k - \bar{x}_n)(x_l - \bar{x}_n) = \frac{2}{n}\sum_{d=1}^{n-1} f_d, \tag{B.34}$$

where now f_d is a measure of how correlated values separated by a distance d are. By dividing f_d by $\mathrm{Var}(x)$, we can define the *autocorrelation function*:

$$\kappa_d = \frac{f_d}{\mathrm{Var}(x)}. \tag{B.35}$$

We see that f_d equals the sample variance, $\mathrm{Var}(x)$, when $d = 0$, so the autocorrelation function has a value of 1 for $d = 0$.

We can now rewrite the sample error as

$$\begin{aligned}\mathrm{err}_X^2 &= \frac{1}{n}\mathrm{Var}(x) + \frac{2}{n}\mathrm{Var}(x)\sum_{d=1}^{n-1}\frac{f_d}{\mathrm{Var}(x)} \\ &= \left(1 + 2\sum_{d=1}^{n-1}\kappa_d\right)\frac{1}{n}\mathrm{Var}(x) \\ &= \frac{\tau}{n}\mathrm{Var}(x), \end{aligned} \tag{B.36}$$

where τ,

$$\tau = 1 + 2\sum_{d=1}^{n-1}\kappa_d, \tag{B.37}$$

is called the *autocorrelation time* and can be used to define the effective number of measurements, $n_{eff} = n/\tau$, for us to simply approximate the error by the sample variance.

Appendix C

DFP and energy minimization

Since we wish to minimize the energy using our Variational Monte Carlo machinery, an important part is to find the optimal parameters for the wave function. That is, the parameters for the wave function which minimizes the energy.

A rather brute force first approach is to manually vary the parameters and plot the results as in section 8.2. This will give a good estimate of what the optimal parameters are, but will in general not be precise enough for more quantitatively precise measurements.

The method we use is a so-called quasi-Newton method, the *Davidon-Fletcher-Powell*-algorithm (DFP) as given in [12]. This method builds on the conjugate gradient method (CGM) and the steepest descent method (both are described in [12]). These techniques however, are respectively too costly numerically, or too brute force for our liking.

Both CGM and DPF are based on approximating a function of variational parameters, $f(\mathbf{x})$, by a quadratic form, where now $\mathbf{x}$ is the set of i variational parameters. This is done by writing the Taylor series of $f(\mathbf{x})$ at some point $\mathbf{P}$, as

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{P}) + \sum_i \frac{\partial f}{\partial x_i} + \frac{1}{2}\frac{\partial^2 f}{\partial x_i \partial x_j} \cdots \\ &\approx c - \mathbf{b}\cdot\mathbf{x} + \frac{1}{2}\mathbf{x}\cdot\mathbf{A}\cdot\mathbf{x} \end{aligned} \tag{C.1}$$

where now

$$c \equiv f(\mathbf{P}) \qquad \mathbf{b} \equiv -\nabla f|_{\mathbf{P}} \qquad [\mathbf{A}]_{ij} \equiv \frac{\partial^2 f}{\partial x_i \partial x_j}\Big|_{\mathbf{P}}. \tag{C.2}$$

The matrix $\mathbf{A}$ is called the *Hessian* matrix, the matrix containing all second derivatives. The gradient can easily be calculated as

$$\nabla f = \mathbf{A}\cdot\mathbf{x} - \mathbf{b}. \tag{C.3}$$

In the DFP method, we don't have to calculate the exact Hessian, $\mathbf{A}$, which takes a lot of computation time. The goal is then to iteratively construct a sequence of matrices, $\mathbf{H}_i$, such that

$$\lim_{i\to\infty} \mathbf{H}_i = \mathbf{A}^{-1}. \tag{C.4}$$

By its name, it should be clear that the Newton method for finding the zero of the gradient of function is similar to the quasi-Newton method. For the Newton method,

we have to second order, near a current iteration point $\mathbf{x}_i$, the following

$$f(\mathbf{x}) = f(\mathbf{x}_i) + (\mathbf{x} - \mathbf{x}_i) \cdot \nabla f(\mathbf{x}_i) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_i) \cdot \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}_i) \tag{C.5}$$

and

$$\nabla f(\mathbf{x}) = \nabla f(\mathbf{x}_i) + \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}_i). \tag{C.6}$$

In order to find the next iteration point using Newton's method, we use that $\nabla f(\mathbf{x}) = 0$, giving

$$\mathbf{x} - \mathbf{x}_i = -\mathbf{A}^{-1} \cdot \nabla f(\mathbf{x}_i). \tag{C.7}$$

The quasi-Newton method is introduced since we do not compute the exact Hessian and its inverse, but an approximation $\mathbf{H} \approx \mathbf{A}^{-1}$. As explained in [12], this approach can often be better than using the true Hessian, as the matrix $\mathbf{H}$ is constructed in such a way that it is always positive-definite and symmetric. The true Hessian is generally not always positive-definite when we are far from the minimum, which can lead to moving towards *increasing* values of f instead of decreasing values.

If we subtract Eq. (C.7) at $\mathbf{x}_{i+1}$ from the same equation at $\mathbf{x}_i$, we get

$$\mathbf{x}_{i+1} - \mathbf{x}_i = \mathbf{A}^{-1} \cdot (\nabla f_{i+1}) - \nabla f_i)), \tag{C.8}$$

where now $f_j = f(\mathbf{x}_j)$. At the point x_{i+1} we also have that $\mathbf{H}_{i+1}$ is a good approximation for $\mathbf{A}^{-1}$, yielding

$$\mathbf{x}_{i+1} - \mathbf{x}_i = \mathbf{H}_{i+1} \cdot (\nabla f_{i+1}) - \nabla f_i)). \tag{C.9}$$

The "outer" product of two vectors, $\mathbf{u}$ and $\mathbf{v}$, is given as, $\mathbf{u} \otimes \mathbf{v}$. As opposed to the "inner" product of two vectors

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} \tag{C.10}$$

which returns a scalar, the "outer" product,

$$\mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^T, \tag{C.11}$$

returns a matrix. Component $(\mathbf{u} \otimes \mathbf{v})_{ij}$ is simply given as $(\mathbf{u} \otimes \mathbf{v})_{ij} = u_i v_j$.

The DPF updating formula for $\mathbf{H}_{i+1}$ is given in [12] as

$$\begin{aligned} \mathbf{H}_i = \mathbf{H}_{i+1} &+ \frac{(\mathbf{x}_{i+1} - \mathbf{x}_i) \otimes (\mathbf{x}_{i+1} - \mathbf{x}_i)}{(\mathbf{x}_{i+1} - \mathbf{x}_i) \cdot (\nabla f_{i+1} - \nabla f_i)} \\ &- \frac{[\mathbf{H}_i \cdot (\nabla f_{i+1} - \nabla f_i)] \otimes [\mathbf{H}_i \cdot (\nabla f_{i+1} - \nabla f_i)]}{(\nabla f_{i+1} - \nabla f_i) \cdot \mathbf{H}_i \cdot (\nabla f_{i+1} - \nabla f_i)}. \end{aligned} \tag{C.12}$$

We have used the function `dfpmin.c` from [12] which uses this quasi-Newton method for finding a minimum for $f(\mathbf{x})$, while finding an approximation for the inverse Hessian. This method takes as input a function that returns the mean energy, and a function that calculates the gradient of the energy.

In this thesis we have a function $f = \bar{E}$, the mean of the local energy operator, $\widehat{\mathbf{E}}_L$, as given in 4.4.1. To use the DFP method we must have the gradient of f, that is the first derivatives of the energy expectation value with respect to the variational parameters, x_i. The first derivatives of the energy expectation value are given in [14] as

$$\begin{aligned} \frac{\partial \bar{E}}{\partial x_i} = \bar{E}_i &= \left\langle \frac{\psi_i}{\psi} E_L + \frac{H\psi_i}{\psi} - 2\bar{E}\frac{\psi_i}{\psi} \right\rangle \\ &= 2\left\langle \frac{\psi_i}{\psi}(E_L - \bar{E}) \right\rangle = 2\left\langle \frac{\psi_i}{\psi} E_L \right\rangle - 2\left\langle \frac{\psi_i}{\psi} \right\rangle \bar{E} \end{aligned} \tag{C.13}$$

where $\psi_i = \partial\psi/\partial x_i$. The expectation values in Eq. (C.13) cannot be calculated easily using closed form expressions, and must be evaluated numerically. In this work we have just used the simple *three-point estimation* of the first derivative

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}, \tag{C.14}$$

with $h = 0.001$.

Bibliography

[1] J. J. Brehm and W. J. Mullin, *Introduction to the Structure of Matter* (Wiley, 1989).

[2] D. Griffiths, *Introduction to Quantum Mechanics - International Edition* (Pearson Education International, 2005).

[3] B. Hammond, W. Lester, and P. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, 1994).

[4] M. Casula and S. Sorella, Journal of Chemical Physics **119**, 6500 (2003).

[5] B. Stroustrup, *The C++ Programming Language* (Addison-Wesley Longman Publishing Co., Inc., 2000).

[6] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, 2000).

[7] M. L. Boas, *Mathematical Methods in the Physical Sciences* (Wiley International, 2006).

[8] S. Raimes, *Many-Electron Theory* (North-Holland Pub. Co., 1972).

[9] M. Hjorth-Jensen, *Computational Physics* (2010).

[10] E. Clementi and C. Roetti, Atomic Data and Nuclear Data Tables **14**, 177 (1974).

[11] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, 2009).

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, 2007).

[13] M. E. Newman and G. E. Barkema, *Monte Carlo Methods in Statistical Phyiscs* (Oxford University Press, 1999).

[14] C. J. Umrigar and C. Filippi, Physical Review Letters **94**, 150201 (2005).

[15] J. M. Thijssen, *Computational Physics* (Cambridge, 2007).

[16] J. Moskowitz and M. Kalos, International Journal of Quantum Chemistry **20**, 1107 (1981).

[17] W. Krauth, *Statistical Mechanics: Algorithms and Computations* (Oxford University Press, 2006).

[18] H. Flyvbjerg and H. G. Petersen, Journal of Chemical Physics **91**, 461 (1989).

[19] A. Sarsa, J. Boronat, and J. Casulleras, Journal of Chemical Physics **116**, 5956 (2002).

[20] C. Barbieri, D. V. Neck, and M. Degroote, ArXiv e-print: arXiv:1004.2607 (2010).

[21] A. M. Teale, S. Coriani, and T. Helgaker, Journal of Chemical Physics **130**, 104111 (2009).

[22] M. Røstad, *Efficient Quantum Monte Carlo Calculations of Many-Body Systems* (University of Oslo, 2004).

[23] J. Rice, *Mathematical Statistics and Data Analysis* (Duxbury, 1995).