

# Project 2 in FYS4411: Computational Physics 2

Mathias M. Vege

June 14, 2017

## Abstract

Results for a quantum mechanical system of  $N = 2, 6, 12$  and  $N = 20$  electrons simulated with the variational Monte Carlo method is presented in this paper. The electrons are confined in their ground state of a harmonic oscillator potential, interacting only through the Coulomb interaction. To account for this interaction, we have introduced a correlating Jastrow factor. The results have been obtained with both uniform sampling and importance sampling, although the latter proved problematic for  $N > 2$ . On the final results, a blocking analysis were performed. For  $N = 2$  we got an estimate for the ground state energy  $\langle E \rangle = 3.0005 \pm 10^{-5}$  a.u. For  $N = 6$  electrons we found  $\langle E \rangle = 20.207 \pm 3.0 \cdot 10^{-5}$  a.u., for  $N = 12$  electrons  $\langle E \rangle = 65.932 \pm 2.0 \cdot 10^{-4}$  a.u and for  $N = 20$  electrons we got  $\langle E \rangle = 156.31 \pm 10^{-3}$  a.u. Comparing this to corresponding results from the article by Pedersen Lohne et al. [7],  $\langle E \rangle_1 = 3.0$  a.u.,  $\langle E \rangle_1 = 20.174$  a.u.,  $\langle E \rangle_1 = 65.741$  a.u. and  $\langle E \rangle_1 = 155.96$  a.u., we see we have a remarkable good agreement in results.

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Theory</b>	<b>3</b>
2.1	Variational Monte Carlo . . . . .	4
2.2	The Metropolis Algorithm . . . . .	6
2.2.1	Uniform sampling . . . . .	7
2.2.2	Importance sampling . . . . .	7
2.2.3	Steepest descent . . . . .	10
2.3	Electron wave function . . . . .	11
2.3.1	Slater determinants . . . . .	11
2.3.2	The Jastrow factor . . . . .	12

2.4	Two electron case . . . . .	12
2.4.1	Unperturbed expectation value . . . . .	12
2.4.2	Local energy . . . . .	15
2.4.3	Quantum force . . . . .	17
2.4.4	$\alpha$ and $\beta$ derivatives in the steepest descent algorithm .	17
2.5	$N$ electron case . . . . .	18
2.5.1	Jastrow factor derivatives . . . . .	18
2.5.2	Slater determinant derivatives . . . . .	19
2.5.3	Local energy . . . . .	21
2.5.4	Quantum force . . . . .	22
2.5.5	$\alpha$ and $\beta$ derivatives in the steepest descent algorithm .	22
2.6	Blocking . . . . .	23
2.7	Parallelization . . . . .	24
<b>3</b>	<b>Implementation</b>	<b>24</b>
3.1	Blocking . . . . .	25
3.2	Program verification and tests . . . . .	25
3.2.1	Unperturbed energies . . . . .	25
3.2.2	2 electron case with and without Jastrow . . . . .	25
3.2.3	Importance sampling . . . . .	25
3.2.4	Parallelization versus non-parallelization . . . . .	25
<b>4</b>	<b>Results</b>	<b>26</b>
4.1	Uniform sampling . . . . .	26
4.2	Optimal parameters . . . . .	26
4.3	Importance sampling . . . . .	27
4.4	Blocking . . . . .	28
4.5	Program verification and tests . . . . .	29
4.5.1	Unperturbed energies . . . . .	29
4.5.2	2 electron case with and without Jastrow . . . . .	30
4.5.3	Parallelization versus non-parallelization . . . . .	31
<b>5</b>	<b>Discussion and conclusion</b>	<b>31</b>
5.1	Issues related to implementation . . . . .	32
5.2	Uniform sampling . . . . .	32
5.3	Optimal parameters from uniform sampling . . . . .	32
5.4	Importance sampling . . . . .	32
5.5	Blocking . . . . .	33
5.6	Verification and tests . . . . .	33
5.7	Closing thoughts and remarks. . . . .	34

<b>Appendices</b>	<b>34</b>
<b>A Hermite polynomials</b>	<b>34</b>
<b>B Integral formulas</b>	<b>35</b>
<b>C Jacobi's formula</b>	<b>36</b>

## 1 Introduction

The goal of this project is to study closed shell systems of electrons confined in a harmonic oscillator potential - a quantum dot. Expectation values for the ground state energies, kinetic and potential energies will be retrieved. We are interested in is a two dimensional system of  $N$  electrons, and since we have closed shell systems we will look at  $N = 2, 6, 12$  and 20 electrons. The program used can be viewed at GitHub [9].

In order to begin this project, let us first go through the theory behind variational Monte Carlo, and all that entails. The following theory part is lengthy, so if the reader feels so inclined, he or she is free to skip to the implementation section.

## 2 Theory

As tradition demands we begin by looking at the Hamiltonian of the system we are to solve,

$$\hat{H} = \sum_{i=1}^N \left( -\frac{1}{2} \nabla_i^2 + \frac{1}{2} \omega^2 r_i^2 \right) + \sum_{i<j}^N \frac{1}{r_{ij}} \quad (1)$$

In order to accommodate a modern notational benefits and simplifications, we use natural units( $\hbar = c = e = m_e = 1$ ). We can also observe that  $N$  is the number of particles we are using, and the  $\omega$  is the oscillator frequency for the harmonic oscillator part. The first part, we recognize as the unperturbed part of the Hamiltonian,

$$\hat{H}_0 = \sum_{i=1}^N \left( -\frac{1}{2} \nabla_i^2 + \frac{1}{2} \omega^2 r_i^2 \right), \quad (2)$$

and the last part is the perturbation to our system,

$$\hat{H}_1 = \sum_{i < j}^N \frac{1}{r_{ij}} \quad (3)$$

such that  $\hat{H} = \hat{H}_0 + \hat{H}_1$ . The distance between two electrons is defined as following,

$$r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4)$$

## 2.1 Variational Monte Carlo

In order to go through the basic steps of variational Monte Carlo, we begin by deriving the basic formulas needed. Let  $\psi_T$  be our trial wave function, which we can expand into an energy basis which we assume is normalized,

$$\psi_T = \sum_{i=0} c_i |\psi_i\rangle$$

The expectation value of the energy given by the Hamiltonian  $\hat{H}$  (from now on omitting the hat) is,

$$\begin{aligned} \langle E \rangle &= \langle \Psi_T | H | \psi_T \rangle \\ &= \sum_{i=0} \sum_{j=0} c_i^* c_j \langle \psi_i | H | \psi_j \rangle \\ &= \sum_{i=0} \sum_{j=0} c_i^* c_j E_j \langle \psi_i | \psi_j \rangle \\ &= \sum_{i=0} |c_i|^2 E_i \end{aligned}$$

From the variational principle in quantum mechanics, we have that the energy of the ground state can be approximated by a trial wave function,

$$E_0 \geq E[\psi_T] = \frac{\langle \psi_T | H | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} \quad (5)$$

Writing this on integral form, we get

$$\begin{aligned}
E[H] = \langle H \rangle &= \frac{\int d\mathbf{r} \psi_T^*(\mathbf{r}) H(\mathbf{r}) \psi_T(\mathbf{r})}{\int d\mathbf{r} \psi_T^*(\mathbf{r}) \psi_T(\mathbf{r})} \\
&= \frac{\sum_{ij} c_i^* c_j \int d\mathbf{r} \psi_i^*(\mathbf{r}) H(\mathbf{r}) \psi_j(\mathbf{r})}{\sum_{ij} c_i^* c_j \int d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r})} \\
&= \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2} \\
&\geq E_0
\end{aligned}$$

In order to make any progress with variational Monte Carlo, we need to get ourselves a wave function. This wave function will have to take  $N$  particles  $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ , and have can be varied by a set of variational parameters  $\alpha = \{\alpha_i\}$ . The integral we wish to solve is now,

$$E[H] = \frac{\int d\mathbf{r} \psi_T^*(\mathbf{r}, \alpha) H(\mathbf{r}) \psi_T(\mathbf{r}, \alpha)}{\int d\mathbf{r} \psi_T^*(\mathbf{r}, \alpha) \psi_T(\mathbf{r}, \alpha)} \quad (6)$$

As with Monte Carlo methods, we need to define ourselves a probability distribution function(PDF) which will be based off our trial wave function  $\psi_T(\mathbf{r})$ ,

$$P(\mathbf{r}) = \frac{|\psi_T(\mathbf{r})|^2}{\int |\psi_T(\mathbf{r})|^2 d\mathbf{r}} \quad (7)$$

We can now define  $E_L$  as our local energy,

$$E_L = \frac{1}{\psi_T} H \psi_T \quad (8)$$

We can use this and the PDF, we can rewrite our  $E[H]$  which now also depends on  $\alpha$ ,

$$E[H(\alpha)] = \int P(\mathbf{r}) E_L(\mathbf{r}) d\mathbf{r} \approx \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} P(\mathbf{r}_i, \alpha) E_L(\mathbf{r}_i, \alpha)$$

where  $N_{MC}$  is the number of Monte Carlo cycles which we run for. We now have an estimate for the ground state energy from the average of our Monte Carlo samples. The algorithm for the VMC(Variational Monte Carlo) method will be based of the Metropolis algorithm([5][3]), thus a quick review of this is in order.

## 2.2 The Metropolis Algorithm

The Metropolis algorithm ([5][3]) is a method for obtaining a sequence of samples when direct sampling is problematic. Given a PDF  $P_i^{(n)}$ , where  $n$  is the time step, and  $i$  is the current state, the algorithm is such that a transitioning probability to a new state  $j$  is given by  $T_{i \rightarrow j}$ . Then, the probability of accepting this state is given by  $A_{i \rightarrow j}$ . If rejected, no move is performed and we remain at state  $i$ . We will require that  $A$  and  $T$  is such that  $P_i^{(n \rightarrow \infty)} \rightarrow p_i$ . This gives rise to the relation

$$\begin{aligned}
P_i^{(n)} &= \sum_j \left[ P_j^{(n-1)} T_{j \rightarrow i} A_{j \rightarrow i} + P_i^{(n-1)} T_{i \rightarrow j} (1 - A_{i \rightarrow j}) \right] \\
&= \sum_j \left[ P_j^{(n-1)} T_{j \rightarrow i} A_{j \rightarrow i} - P_i^{(n-1)} T_{i \rightarrow j} A_{i \rightarrow j} + P_i^{(n-1)} T_{i \rightarrow j} \right] \\
&= \sum_j \left[ P_j^{(n-1)} T_{j \rightarrow i} A_{j \rightarrow i} - P_i^{(n-1)} T_{i \rightarrow j} A_{i \rightarrow j} \right] + P_i^{(n-1)} \sum_j T_{i \rightarrow j} \\
&= \sum_j \left[ P_j^{(n-1)} T_{j \rightarrow i} A_{j \rightarrow i} - P_i^{(n-1)} T_{i \rightarrow j} A_{i \rightarrow j} \right] + P_i^{(n-1)} \cdot 1
\end{aligned}$$

Since we require our system to make some transition into a new state  $j$ ,  $\sum_j T_{i \rightarrow j} = 1$ . Now using the requirement  $P_i^{(n \rightarrow \infty)} \rightarrow p_i$ , we get

$$\begin{aligned}
p_i &= \sum_j [p_j T_{j \rightarrow i} A_{j \rightarrow i} - p_i T_{i \rightarrow j} A_{i \rightarrow j}] + p_i \\
&\downarrow \\
0 &= \sum_j [p_j T_{j \rightarrow i} A_{j \rightarrow i} - p_i T_{i \rightarrow j} A_{i \rightarrow j}]
\end{aligned}$$

The requirement applies for the system as a whole when time tends to infinity. This is a rather weak requirement, so we seek to enforce a similar condition to each state. Namely,

$$p_j T_{j \rightarrow i} A_{j \rightarrow i} = p_i T_{i \rightarrow j} A_{i \rightarrow j}$$

which we observe implies that transitioning and accepting a state  $i \rightarrow j$  is equal to that of  $j \rightarrow i$ . By rearranging we get,

$$\frac{A_{j \rightarrow i}}{A_{i \rightarrow j}} = \frac{p_j T_{j \rightarrow i}}{p_i T_{i \rightarrow j}} \quad (9)$$

This is a much stronger requirement, and is called *detailed balance*. The Metropolis algorithm is now all about maximizing  $A$ , such that

$$A_{i \rightarrow j} = \min \left( 1, \frac{p_i T_{i \rightarrow j}}{p_j T_{j \rightarrow i}} \right)$$

The ratio which we use to accept or reject a step is defined as

$$q(\mathbf{r}^{\text{new}}, \mathbf{r}^{\text{old}}) = \frac{p_i T_{i \rightarrow j}}{p_j T_{j \rightarrow i}} \quad (10)$$

Let us then recapitulate what the Metropolis algorithm becomes as applied to our VMC problem.

---

**Algorithm 1** Metropolis algorithm for Variational Monte Carlo.

---

- 1: Set up initial conditions for our system.
  - 2: Sample a new state  $j$  with transition probability  $T_{i \rightarrow j}$  by the desired sampling method.
  - 3: Accept new state  $j$  with acceptance probability  $A_{i \rightarrow j}$ .
  - 4: If rejected, we go back to state  $i$ .
  - 5: Sample the system.
  - 6: Repeat step 2-5  $N_{MC}$  times.
- 

### 2.2.1 Uniform sampling

For sampling a new state (by which we mean a new position), we use

$$\mathbf{r}^{\text{new}} = \mathbf{r}^{\text{old}} + \xi \Delta \mathbf{r}^{\text{old}} \quad (11)$$

where  $\xi$  is a random number chosen from a uniform distribution, and  $\Delta \mathbf{r}$  is the step length which we update for. For this kind of uniform sampling, we seek to tune  $\Delta \mathbf{r}$  to such that we accept roughly 50% all suggested moves.

The acceptance ratio (10) as applied to our VMC calculation, is (dubbing it as  $R$  instead of  $q(\mathbf{r}^{\text{new}}, \mathbf{r}^{\text{old}})$ ),

$$R \equiv \frac{\psi_T^{\text{new}}}{\psi_T^{\text{old}}} \quad (12)$$

### 2.2.2 Importance sampling

If we wish to improve our VMC Metropolis algorithm, we can use something called importance sampling. It is based upon the Fokker-Planck equation,

which characterizes a move through coordinate space. For one dimension, it reads

$$\frac{\partial P}{\partial t} = D \frac{\partial}{\partial x} \left( \frac{\partial}{\partial x} - F \right) P(x, t) \quad (13)$$

The  $D$  is referred to as the diffusion constant since it is derived from the Fokker-Planck equation, and the  $F$  is called the drift term. A new position in coordinate space is chosen from solving the Langevin equation,

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta \quad (14)$$

$\eta$  is a random variable. This is solved using Euler's method. For our problem, the solution and sampling of a new position is defined by

$$\mathbf{r}^{\text{new}} = \mathbf{r}^{\text{old}} + D\mathbf{F}(\mathbf{r}^{\text{old}})\Delta t + \xi\sqrt{\Delta t} \quad (15)$$

Where  $D = 0.5$  in atomic units, and  $\mathbf{F}$  is the quantum force as defined by,

$$\mathbf{F} = \frac{2}{\psi_T} \nabla \psi_T \quad (16)$$

We typically choose  $\Delta t \in [0.001, 0.01]$ . The random variable  $\xi$  is chosen from a Gaussian distribution around 0 with a standard deviation of 1. Further, the acceptance ratio  $q(\mathbf{r}^{\text{new}}, \mathbf{r}^{\text{old}})$  now becomes,

$$q(\mathbf{r}^{\text{new}}, \mathbf{r}^{\text{old}}) = \frac{G(\mathbf{r}^{\text{old}}, \mathbf{r}^{\text{new}})|\psi_T(\mathbf{r}^{\text{new}})|^2}{G(\mathbf{r}^{\text{new}}, \mathbf{r}^{\text{old}})|\psi_T(\mathbf{r}^{\text{old}})|^2} \quad (17)$$

where the  $G(\mathbf{r}^{\text{new}}, \mathbf{r}^{\text{old}})$  is a Green's function which is a solution to the Fokker-Planck equation. This is a transition probability, and is defined by

$$G(\mathbf{r}^{\text{new}}, \mathbf{r}^{\text{old}}) = \frac{1}{(4\pi D\Delta t)^{3N/2}} \exp \left( -(\mathbf{r}^{\text{new}} - \mathbf{r}^{\text{old}} - D\delta t\mathbf{F}(\mathbf{r}^{\text{old}}))^2 / 4D\Delta t \right) \quad (18)$$

A more detailed derivation of importance sampling can be found in the lecture notes given of the Computational Physics [4]. As we can already observe, it is possible to simplify the Greens ratio (17) due to the exponential. We will begin by using that the squaring of the exponential implies a dot product for vectors. For simplicity, we dub  $\mathbf{x} = \mathbf{r}^{\text{old}}$  and  $\mathbf{y} = \mathbf{r}^{\text{new}}$ . Through a



lot of patience, we get

$$\begin{aligned}
\frac{G(\mathbf{x}, \mathbf{y})}{G(\mathbf{y}, \mathbf{x})} &= \exp \left[ \frac{1}{4D\Delta t} \left\{ - (x_x^2 - x_x y_x - x_x D\Delta t F_x(\mathbf{y}) + y_x^2 - y_x x_x \right. \right. \\
&\quad + y_x D\Delta t F_x(\mathbf{y}) - x_x D\Delta t F_x(\mathbf{y}) + y_x D\Delta t F_x(\mathbf{y}) + D^2 \Delta t^2 F_x^2(\mathbf{y})) \\
&\quad + (y_x^2 - y_x x_x - y_x D\Delta t F_x(\mathbf{x}) + x_x^2 - x_x y_x \\
&\quad + x_x D\Delta t F_x(\mathbf{x}) + D^2 \Delta t^2 F_x^2(\mathbf{x}) - y_x D\Delta t F_x(\mathbf{x}) + x_x D\Delta t F_x(\mathbf{x})) \\
&\quad - (x_y^2 - x_y y_y - x_y D\Delta t F_y(\mathbf{y}) - x_y y_y + y_y^2 \\
&\quad + y_y D\Delta t F_y(\mathbf{y}) - x_y D\Delta t F_y(\mathbf{y}) + y_y D\Delta t F_y(\mathbf{y}) + D^2 \Delta t^2 F_y^2(\mathbf{y})) \\
&\quad + (y_y^2 - y_y x_y - y_y D\Delta t F_y(\mathbf{x}) + x_y^2 - x_y y_y + x_y D\Delta t F_y(\mathbf{x}) \\
&\quad \left. \left. + D^2 \Delta t^2 F_y^2(\mathbf{x}) - y_y D\Delta t F_y(\mathbf{x}) + x_y D\Delta t F_y(\mathbf{x})) \right\} \right] \\
&= \exp \left[ \frac{1}{4D\Delta t} \left\{ - y_x D\Delta t F_x(\mathbf{x}) + x_x D\Delta t F_x(\mathbf{x}) + D^2 \Delta t^2 F_x^2(\mathbf{x}) \right. \right. \\
&\quad - y_x D\Delta t F_x(\mathbf{x}) + x_x D\Delta t F_x(\mathbf{x}) + x_x D\Delta t F_x(\mathbf{y}) - y_x D\Delta t F_x(\mathbf{y}) \\
&\quad + x_x D\Delta t F_x(\mathbf{y}) - y_x D\Delta t F_x(\mathbf{y}) - D^2 \Delta t^2 F_x^2(\mathbf{y}) - y_y D\Delta t F_y(\mathbf{x}) \\
&\quad + x_y D\Delta t F_y(\mathbf{x}) + D^2 \Delta t^2 F_y^2(\mathbf{x}) - y_y D\Delta t F_y(\mathbf{x}) + x_y D\Delta t F_y(\mathbf{x}) \\
&\quad + x_y D\Delta t F_y(\mathbf{y}) - y_y D\Delta t F_y(\mathbf{y}) + x_y D\Delta t F_y(\mathbf{y}) - y_y D\Delta t F_y(\mathbf{y}) \\
&\quad \left. \left. - D^2 \Delta t^2 F_y^2(\mathbf{y}) \right\} \right] \\
&= \exp \left[ \frac{1}{4D\Delta t} \left\{ - y_x D\Delta t [F_x(\mathbf{x}) + F_x(\mathbf{x}) + F_x(\mathbf{y}) + F_x(\mathbf{y})] \right. \right. \\
&\quad - D^2 \Delta t^2 [F_x^2(\mathbf{y}) - F_x^2(\mathbf{x})] + 2x_x D\Delta t [F_x(\mathbf{x}) + F_x(\mathbf{y})] \\
&\quad - 2y_y D\Delta t [F_y(\mathbf{x}) + F_y(\mathbf{y})] + 2x_y D\Delta t [F_y(\mathbf{x}) + F_y(\mathbf{y})] \\
&\quad \left. \left. - D^2 \Delta t^2 [F_y^2(\mathbf{y}) - F_y^2(\mathbf{x})] \right\} \right] \\
&= \exp \left[ - \frac{y_x}{2} (F_x(\mathbf{x}) + F_x(\mathbf{y})) - \frac{D\Delta t}{4} (F_x^2(\mathbf{y}) - F_x^2(\mathbf{x})) \right. \\
&\quad + \frac{x_x}{2} (F_x(\mathbf{x}) + F_x(\mathbf{y})) - \frac{y_y}{2} (F_y(\mathbf{x}) + F_y(\mathbf{y})) \\
&\quad \left. - \frac{D\Delta t}{4} (F_y^2(\mathbf{y}) - F_y^2(\mathbf{x})) + \frac{x_y}{4} (F_y(\mathbf{x}) + F_y(\mathbf{y})) \right]
\end{aligned}$$

Through the magic of simplification (but mostly patience), we arrive at

$$\begin{aligned} \frac{G(\mathbf{x}, \mathbf{y})}{G(\mathbf{y}, \mathbf{x})} = \exp & \left[ \frac{1}{2} (x_x - y_x) (F_x(\mathbf{x}) + F_x(\mathbf{y})) \right. \\ & + \frac{1}{2} (x_y - y_y) (F_y(\mathbf{x}) + F_y(\mathbf{y})) \\ & \left. - \frac{D\Delta t}{4} (F_x^2(\mathbf{y}) - F_x^2(\mathbf{x}) + F_y^2(\mathbf{y}) - F_y^2(\mathbf{x})) \right] \end{aligned}$$

Summing over dimensions, and we get

$$\frac{G(\mathbf{x}, \mathbf{y})}{G(\mathbf{y}, \mathbf{x})} = \exp \left[ \sum_i^{N_{\text{dim}}} \frac{1}{2} (F_i(\mathbf{x}) + F_i(\mathbf{y})) \left( \frac{D\Delta t}{2} (F_i(\mathbf{x}) - F_i(\mathbf{y})) - y_i + x_i \right) \right] \quad (19)$$

### 2.2.3 Steepest descent

The steepest descent method is a simple iterative method for finding a local minimum in parameters. It is used when we are tasked to solve systems of linear equations on the form of

$$A\mathbf{x} = \mathbf{b}.$$

Through an iterative process we seek to minimize the following,

$$\mathbf{r} = \mathbf{b} - A\mathbf{x},$$

where the exact solution implies that  $\mathbf{r} = 0$ . The process for solving this through steepest descent can be set up as,

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \nabla f(\mathbf{x}_i) \delta t \quad (20)$$

Applied to our VMC case with one or two variational parameters,  $\alpha$  and  $\beta$ , the equations will take on the form of (20),

$$\begin{pmatrix} \alpha_{i+1} \\ \beta_{i+1} \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} - \delta t \begin{pmatrix} \frac{d\langle E_L[\alpha] \rangle}{d\alpha} \\ \frac{d\langle E_L[\beta] \rangle}{d\beta} \end{pmatrix} \quad (21)$$

Where the  $\langle E_L \rangle$  is the expectation value of the local energy and  $\delta t$  is the step size of the method. The expectation value of the local energy is found by

$$\frac{d\langle E_L[\alpha_i] \rangle}{d\alpha_i} = 2 \left( \left\langle \frac{\bar{\psi}_{\alpha_i}}{\psi_T[\alpha_i]} E_L[\alpha_i] \right\rangle - \left\langle \frac{\bar{\psi}_{\alpha_i}}{\psi[\alpha_i]} \right\rangle \langle E_L[\alpha_i] \rangle \right) \quad (22)$$

with the derivative of  $\psi$  with respect to the variational parameter is defined as

$$\bar{\psi}_{\alpha_i} = \frac{d\psi[\alpha_i]}{d\alpha_i}$$

As evident, this method may oscillate around the solution, since the step size is fixed. Other more optimal methods such as the conjugate gradient method exist (see lecture notes [4]), but are not implemented in the project.

### 2.3 Electron wave function

Our wave function will consist of two parts: one that comes from the harmonic oscillator potential and is built up based on the fermionic nature of the system, and one that provides a correlation between two particles - the so-called Jastrow factor. The wave function we construct, will be called our *trial wave function*.

$$\psi_T(\mathbf{r}) = \psi_J(\mathbf{r})\psi_{SD}(\mathbf{r}) \quad (23)$$

#### 2.3.1 Slater determinants

The wave function for an electron in a two dimensional harmonic oscillator potential can be written as a Hermite polynomial,

$$\phi_{n_x, n_y}(x, y) = A H_{n_x}(\sqrt{\omega\alpha}x) H_{n_y}(\sqrt{\omega\alpha}y) \exp\left(-\frac{\omega\alpha}{2}(x^2 + y^2)\right) \quad (24)$$

The details surrounding the mysterious variational parameter  $\alpha$  will be explained further on. For now, we will place these  $\phi$ 's into a Slater determinant. The Slater determinant is a creature that describes the wave function of a fermionic system, while also enforcing anti-symmetrization and thus the Pauli principle. Our Slater determinant will take the following form, when we describe the specific state of a system by  $n_x, n_y$  and a specific particle  $\mathbf{r}_i = x_i\mathbf{i} + y_i\mathbf{j}$ ,

$$D = \det(\Phi(\mathbf{r})) \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \dots & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \dots & \phi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \dots & \phi_N(\mathbf{r}_N) \end{vmatrix} \quad (25)$$

From how we have described the position, we introduce following, useful notation,

$$r_i = \sqrt{x_i^2 + y_i^2} \quad (26)$$

Note that we have defined  $\mathbf{r} \equiv (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  in the Slater determinant (25). To pull this definition back to our trial wave function (23), we get

$$\psi_{SD}(\mathbf{r}) = \det(\Phi(\mathbf{r})) = D \quad (27)$$

Where the *OB* stands for one body, as in one body wave function. We define an element of the Slater matrix as  $d_{ij}$ . We now need to look into the part that accounts for many body effects, the Jastrow factor.

### 2.3.2 The Jastrow factor

The correlation term is called a *Jastrow factor* is as mentioned here to take into account many-body effects of our system. The general shape of it is

$$\psi_J(\mathbf{r}) = \prod_{i < j}^N \exp\left(\frac{ar_{ij}}{1 + \beta r_{ij}}\right) = \prod_{i=1}^N \prod_{j=i+1}^N \exp\left(\frac{ar_{ij}}{1 + \beta r_{ij}}\right) \quad (28)$$

where the *C* stands for correlation. The *a* is a parameter that is 1 for parallel spin, and 1/3 for anti-parallel spin. The  $\beta$  is another variational parameter and the  $r_{ij}$  is defined by the equation (4) as the distance between two electrons. For now, we shall begin by looking closer at the two-electron case.

## 2.4 Two electron case

For the two electron case the Hamiltonian takes on familiar form,

$$H = -\frac{1}{2} \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} \right) + \frac{1}{2} \omega^2 (x_1^2 + y_1^2 + x_2^2 + y_2^2) + \frac{1}{r_{12}} \quad (29)$$

We will assume the spin is anti-parallel and the total spin is 0 in the ground state for the 2 electron case due to Pauli's exclusion principle.

### 2.4.1 Unperturbed expectation value

We begin by looking at the unperturbed expectation value, in which we will try to convince ourselves that the energy comes out to be  $2\omega$  in atomic units. Given a wave function,

$$\Phi_0(\mathbf{r}_1, \mathbf{r}_2) = C \exp\left(-\frac{\omega}{2}(r_1^2 + r_2^2)\right) \quad (30)$$

with  $r_1$  and  $r_2$  being defined by the notation introduced in the Slater determinant(26).

$$\begin{aligned}\langle E_0 \rangle &= \langle \Psi_0 | H_0 | \Psi_0 \rangle \\ &= \int C \exp \left( -\frac{\omega}{2} (r_1^2 + r_2^2) \right) \left( -\frac{1}{2} (\nabla_1^2 + \nabla_2^2) + \frac{1}{2} \omega^2 (r_1^2 + r_2^2) \right) \\ &\quad \times C \exp \left( -\frac{\omega}{2} (r_1^2 + r_2^2) \right) dx_1 dy_1 dx_2 dy_2\end{aligned}$$

Since the derivatives are similar, we can look a single derivative, and then solve for that one.

$$\begin{aligned}\frac{\partial^2}{\partial x^2} \exp \left( -\frac{\omega}{2} x^2 \right) &= \frac{\partial}{\partial x} (-\omega x) \exp \left( -\frac{\omega}{2} x^2 \right) \\ &= (\omega^2 x^2 - \omega) \exp \left( -\frac{\omega}{2} x^2 \right)\end{aligned}$$

Inserting this, into our integral, we get,

$$\begin{aligned}I_{x_1} &= -\frac{1}{2} \int \exp \left( -\frac{\omega}{2} (x_1^2 + y_1^2 + x_2^2 + y_2^2) \right) \frac{\partial^2}{\partial x_1^2} \\ &\quad \times \exp \left( -\frac{\omega}{2} (x_1^2 + y_1^2 + x_2^2 + y_2^2) \right) dx_1 dy_1 dx_2 dy_2 \\ &= -\frac{1}{2} \int (\omega^2 x_1^2 - \omega) \exp \left( -\omega (x_1^2 + y_1^2 + x_2^2 + y_2^2) \right) dx_1 dy_1 dx_2 dy_2 \\ &= -\frac{\omega}{2} \left( \omega \int x^2 \exp \left( -\omega (x_1^2 + y_1^2 + x_2^2 + y_2^2) \right) dx_1 dy_1 dx_2 dy_2 \right. \\ &\quad \left. - \int \exp \left( -\omega (x_1^2 + y_1^2 + x_2^2 + y_2^2) \right) dx_1 dy_1 dx_2 dy_2 \right)\end{aligned}$$

We can now use the integral formulas seen in the appendix, integral (70) and (71). Noting that there are actually four integrations happening, we get

$$\begin{aligned}
& -\frac{\omega}{2} \left( \omega \int x^2 \exp(-\omega(x_1^2 + y_1^2 + x_2^2 + y_2^2)) dx_1 dy_1 dx_2 dy_2 \right. \\
& \quad \left. - \int \exp(-\omega(x_1^2 + y_1^2 + x_2^2 + y_2^2)) dx_1 dy_1 dx_2 dy_2 \right) \\
& = -\frac{\omega}{2} \left( \left( \frac{\pi^{\frac{1}{2}}}{2} \omega^{-\frac{3}{2}} \right) \left( \pi^{\frac{1}{2}} \omega^{-\frac{1}{2}} \right)^3 - \left( \pi^{\frac{1}{2}} \omega^{-\frac{1}{2}} \right)^4 \right) \\
& = -\frac{\omega}{2} \left( \frac{\pi^2}{2} \omega^{-3} + \pi^2 \omega^{-2} \right) \\
& = -\frac{\pi^2}{4\omega} + \frac{\pi^2}{2\omega} \\
& = \frac{\pi^2}{4\omega}
\end{aligned}$$

Putting this together with the other Gaussian symmetric integral coming from the harmonic oscillator, we get

$$\begin{aligned}
\langle E_0 \rangle & = C^2 \left( 4 \frac{\pi^2}{4\omega} + 4 \frac{\pi^2}{4\omega} \right) \\
& = C^2 \frac{2\pi^2}{\omega}
\end{aligned}$$

We now need to find the constant C.

$$\begin{aligned}
\langle \Phi_0 | \Phi_0 \rangle & = C^2 \int \exp(-\omega(x_1^2 + y_1^2 + x_2^2 + y_2^2)) dx_1 dy_1 dx_2 dy_2 \\
& = C^2 \left( \sqrt{\frac{\pi}{\omega}} \right)^4
\end{aligned}$$

Since our wave function is required to be normalized C becomes,

$$C = \frac{\omega}{\pi}$$

Inserting this, and we get our expectation value

$$\langle E_0 \rangle = \frac{\omega^2}{\pi^2} \frac{2\pi^2}{\omega} = 2\omega$$

Which, is the answer we expected. This will provide a useful test for us later on. The local energy for the unperturbed two electron case is thus given as,

$$E_L = 2\omega. \quad (31)$$

Similar results can be retrieved for  $N = 6, 12$  and  $20$  electrons, corresponding to  $10\omega$ ,  $28\omega$  and  $60\omega$ .

#### 2.4.2 Local energy

We now seek to find an expression for the local energy as given by equation (8), for the two electron case.

$$E_L = \frac{1}{\psi_T} H \psi_T \quad (32)$$

The trial wave function is given as

$$\psi_T = \psi_{SD} \psi_J = C \exp\left(-\frac{\omega\alpha}{2} (r_1^2 + r_2^2)\right) \exp\left(\frac{ar_{12}}{1 + \beta r_{12}}\right) \quad (33)$$

When the Hamiltonian is given as in equation (29), the main problem is to find the laplacian,

$$\nabla^2 = \nabla_1^2 + \nabla_2^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial y_2^2}$$

and it is thus where we shall begin. For simplicity, we write

$$\frac{\nabla^2 \psi_T}{\psi_T} = \frac{\nabla^2 \left( C \exp\left(-\frac{\omega\alpha}{2} (r_1^2 + r_2^2) + \frac{ar_{12}}{1 + \beta r_{12}}\right) \right)}{\psi_T}$$

We start by doing a single derivative.

$$\frac{\partial}{\partial x_1} \psi_T = \psi_T \left( -\omega\alpha x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \quad (34)$$

For the derivative with respect to  $x_2$  we get an additional minus sign in the last term. For the second derivative, we get

$$\begin{aligned} \frac{\partial^2}{\partial x_1^2} \psi_T &= \frac{\partial}{\partial x_1} \left( \psi_T \left( -\omega\alpha x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \right) \\ &= \psi_T \left( \frac{\partial}{\partial x_1} (-\omega\alpha x_1) + \frac{\partial}{\partial x_1} \left( \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \right) \\ &\quad + \left( -\omega\alpha x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \frac{\partial}{\partial x_1} (\psi_T) \end{aligned}$$

We now do the second term by itself,

$$\frac{\partial}{\partial x_1} \left( \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) = \frac{a}{r_{12}(1 + \beta r_{12})^2} - \frac{a(x_1 - x_2)^2}{r_{12}^3(1 + \beta r_{12}^2)} - \frac{2a\beta(x_1 - x_2)^2}{r_{12}^2(1 + \beta r_{12})^3}$$

An overall minus sign appears when taking the derivative with respect to  $x_2$ . We now look at the last term in our second derivative.

$$\begin{aligned} \left( -\omega\alpha x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \frac{\partial}{\partial x_1} (\psi_T) &= \left( \omega^2\alpha^2 x_1^2 - \frac{\omega\alpha a x_1(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right. \\ &\quad \left. - \frac{\omega\alpha a x_1(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} + \frac{a^2(x_1 - x_2)^2}{r_{12}^2(1 + \beta r_{12})^4} \right) \psi_T \\ &= \left( \omega^2\alpha^2 x_1^2 - \frac{2\omega\alpha a x_1(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} + \frac{a^2(x_1 - x_2)^2}{r_{12}^2(1 + \beta r_{12})^4} \right) \psi_T \end{aligned}$$

Putting all this together, the second derivative becomes

$$\begin{aligned} \frac{\partial^2}{\partial x_1^2} \psi_T &= \left( -\omega\alpha + \omega^2\alpha^2 x_1^2 - \frac{2\omega\alpha a x_1(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} + \frac{a^2(x_1 - x_2)^2}{r_{12}^2(1 + \beta r_{12})^4} \right. \\ &\quad \left. + \frac{a}{r_{12}(1 + \beta r_{12})^2} - \frac{a(x_1 - x_2)^2}{r_{12}^3(1 + \beta r_{12}^2)} - \frac{2a\beta(x_1 - x_2)^2}{r_{12}^2(1 + \beta r_{12})^3} \right) \psi_T \\ &= \left( -\omega\alpha + \omega^2\alpha^2 x_1^2 - \frac{2\omega\alpha a x_1(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} + \frac{2a}{(1 + \beta r_{12})^2} \right. \\ &\quad \left. \times \left( \frac{1}{2r_{12}} - \frac{(x_1 - x_2)^2}{2r_{12}^3} - \frac{\beta(x_1 x_2)^2}{r_{12}^2(1 + \beta r_{12})} + \frac{a(x_1 - x_2)^2}{2r_{12}^2(1 + \beta r_{12})^2} \right) \right) \psi_T \end{aligned}$$

We can now add the derivative with respect to  $x_1$  and  $x_2$ .

$$\begin{aligned} \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) \psi_T &= \left( -2\omega\alpha + \omega^2\alpha^2 (x_1^2 + x_2^2) - \frac{2\omega\alpha a(x_1 - x_2)^2}{r_{12}(1 + \beta r_{12})^2} \right. \\ &\quad \left. + \frac{2a}{(1 + \beta r_{12})^2} \left( \frac{1}{r_{12}} - \frac{(x_1 - x_2)^2}{r_{12}^3} - \frac{2\beta(x_1 - x_2)^2}{r_{12}^2(1 + \beta r_{12})} + \frac{a(x_1 - x_2)^2}{r_{12}^2(1 + \beta r_{12})^2} \right) \right) \psi_T \end{aligned}$$

Putting this together with the derivatives with respect to  $y_1$  and  $y_2$  and using the relation given in the beginning (4), we get

$$\begin{aligned} \frac{\nabla^2 \psi_T}{\psi_T} &= -4\omega\alpha + \omega^2\alpha^2(r_1^2 + r_2^2) - \frac{2\omega\alpha a r_{12}}{(1 + \beta r_{12})^2} \\ &\quad + \frac{2a}{(1 + \beta r_{12})^2} \left( \frac{2}{r_{12}} - \frac{r_{12}^2}{r_{12}^3} - \frac{2\beta r_{12}^2}{r_{12}^2(1 + \beta r_{12})} + \frac{a r_{12}^2}{r_{12}^2(1 + \beta r_{12})^2} \right) \end{aligned}$$



which becomes our final expression for the Laplacian,

$$\frac{\nabla^2 \psi_T}{\psi_T} = \omega^2 \alpha^2 (r_1^2 + r_2^2) - 4\omega\alpha - \frac{2\omega\alpha ar_{12}}{(1 + \beta r_{12})^2} \quad (35)$$

$$+ \frac{2a}{(1 + \beta r_{12})^2} \left( \frac{a}{(1 + \beta r_{12})^2} + \frac{1}{r_{12}} - \frac{2\beta}{(1 + \beta r_{12})} \right) \quad (36)$$

Which gives the total local energy by adding the harmonic oscillator term and coulomb interaction,

$$E_L = -\frac{1}{2} \left( (\alpha^2 - 1)\omega^2 (r_1^2 + r_2^2) - 4\omega\alpha - \frac{2\omega\alpha ar_{12}}{(1 + \beta r_{12})^2} \right) \quad (37)$$

$$+ \frac{2a}{(1 + \beta r_{12})^2} \left( \frac{a}{(1 + \beta r_{12})^2} + \frac{1}{r_{12}} - \frac{2\beta}{(1 + \beta r_{12})} \right) + \frac{1}{r_{12}} \quad (38)$$

### 2.4.3 Quantum force

We also have to find the quantum force given by equation (16), which involves taking the gradient of the wave function. As we already know the first derivatives from taking the gradient(34), the process is trivial.

$$\mathbf{F} = 2 \left( -\omega\alpha x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \mathbf{i} + 2 \left( -\omega\alpha y_1 + \frac{a(y_1 - y_2)}{r_{12}(1 + \beta r_{12})^2} \right) \mathbf{j} \quad (39)$$

For  $r_2$  do we get a minus in front of the last term.

### 2.4.4 $\alpha$ and $\beta$ derivatives in the steepest descent algorithm

For the steepest descent method (20), we have to find derivatives of the trial wave function (33) with respect to  $\alpha$  and  $\beta$ .

From the equation for the energy expectation value (22), we see that we have to find the derivative of the wave function(30) with respect to the variational parameters  $\alpha$  and  $\beta$ .

$$\frac{d\psi_T[\alpha]}{d\alpha} = -\frac{\omega}{2} (r_1^2 + r_2^2) \psi_T \quad (40)$$

For  $\beta$  we get,

$$\frac{d\psi_T[\beta]}{d\beta} = -\frac{ar_{12}^2}{(1 + \beta r_{12})^2} \psi_T \quad (41)$$

Which is what is need in order to apply steepest descent to the two-body Jastrow factor case.

## 2.5 $N$ electron case

For the  $N$ -electron case, we apply the Hamiltonian as states in equation (1) at the beginning of the theory section on the trial wave function as stated in (23). The  $N$  electron energy is once more given as the local energy (8). From looking at the general Hamiltonian (1), we see that a major task will be to once more find the laplacian of our trial wave function (23).

$$\nabla_i^2 \psi_T = \frac{\nabla_i^2 \psi_{SD}}{\psi_{SD}} + \frac{\nabla_i^2 \psi_J}{\psi_J} + 2 \frac{\nabla_i \psi_{SD}}{\psi_{SD}} \cdot \frac{\nabla_i \psi_J}{\psi_J} \quad (42)$$

Since we are calculating the ground state energy, all the lowest lying levels will be filled up according to the Pauli exclusion principle, and the total spin should be zero.

### 2.5.1 Jastrow factor derivatives

We begin by finding the Jastrow terms. The full derivation can be found in the lecture notes for Computational Physics [4].  $k$  is here used instead of  $i$  as in the equation to the Hamiltonian (1). Since the Pade-Jastrow factor is an exponential as given in (28), we make following definitions,

$$g_{ij} \equiv g(r_{ij}) = e^{f(r_{ij})} = e^{f_{ij}} \quad (43)$$

This yields a general equation for the gradient of the Jastrow factor,

$$\frac{1}{\psi_J} \nabla_k \psi_J = \sum_{i \neq k}^N \frac{1}{g_{ik}} \frac{\mathbf{r}_{ik}}{r_{ik}} \frac{\partial g_{ik}}{\partial r_{ik}} = \sum_{i \neq k}^N \frac{\mathbf{r}_{ik}}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} \quad (44)$$

the  $\mathbf{r}_{ik}$  is defined analogues to the electron distance (4), but as

$$\mathbf{r}_{ik} = \mathbf{r}_k - \mathbf{r}_i \quad (45)$$

By taking the first derivative of exponential content in (28), we find the derivative of  $f_{ij}$  as

$$\frac{\partial f_{ik}}{\partial r_{ik}} = \frac{a_{ik}}{(1 + \beta r_{ik})^2} \quad (46)$$

By using our previous equation (44), we get

$$\frac{1}{\psi_J} \nabla_k \psi_J = \sum_{i \neq k}^N \frac{\mathbf{r}_{ik}}{r_{ik}} \frac{a_{ik}}{(1 + \beta r_{ik})^2} \quad (47)$$

We can now look for the Laplacian. The general formula for this is one is given by taking the gradient once more,

$$\frac{\nabla_k^2 \psi_J}{\psi_J} = \left( \frac{\nabla_k \psi_J}{\psi_J} \right)^2 + \sum_{i \neq k}^N \left[ \left( \frac{d-1}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} + \frac{\partial^2 f_{ik}}{\partial r_{ik}^2} \right) \right] \quad (48)$$

The  $d$  stands for the dimension. The second derivative of  $f_{ik}$  is given by

$$\frac{\partial^2 f_{ik}}{\partial r_{ik}^2} = -\frac{2a_{ik}\beta}{(1 + \beta r_{ik})^3} \quad (49)$$

By setting  $d = 2$  (as we only are working in two dimensions), and inserting the partial derivative into the general formula for the Laplacian, we get,

$$\frac{1}{\psi_J} \nabla_k^2 \psi_k = \left( \frac{\nabla_k \psi_J}{\psi_J} \right)^2 + \sum_{i \neq k}^N \frac{a_{ik}(1 - \beta r_{ik})}{r_{ik}(1 + \beta r_{ik})^3} \quad (50)$$

### 2.5.2 Slater determinant derivatives

We can now look at the Slater determinant. As shown by Moskowitz and Kalos [6], we can split the spin parts into two Slater determinants, one for spin up and one for spin down since our Hamiltonian is spin independent. If we do not do this the determinant will be zero. By using the definition we presented in (27), we get

$$\psi_{SD} = D = D_\uparrow D_\downarrow \quad (51)$$

This gives following expressions for the gradient,

$$\frac{\nabla \psi_{SD}}{\psi_{SD}} = \frac{\nabla D_\downarrow}{D_\downarrow} + \frac{\nabla D_\uparrow}{D_\uparrow} \quad (52)$$

and the laplacian

$$\frac{\nabla^2 \psi_{SD}}{\psi_{SD}} = \frac{\nabla^2 D_\downarrow}{D_\downarrow} + \frac{\nabla^2 D_\uparrow}{D_\uparrow} \quad (53)$$

The gradient of the Slater determinant is given in the lecture notes 2015 [4] as

$$\frac{\nabla_k \psi_{SD}}{\psi_{SD}} = \sum_{i \neq k}^N \nabla_k d_{ik}(\mathbf{r}) d_{ki}^{-1}(\mathbf{r}) \quad (54)$$

Since  $d_{ik} = \phi_k(\mathbf{r}_i)$ , we must find the derivative of the single particle wave function(24). The gradient of this is given as,

$$\nabla_i \phi_j(\mathbf{r}_i) = \left( \frac{\partial}{\partial x_i} \mathbf{i} + \frac{\partial}{\partial y_i} \mathbf{j} \right) A H_{n_x}(\sqrt{\omega\alpha}x) H_{n_y}(\sqrt{\omega\alpha}y) \exp\left(-\frac{\omega\alpha}{2}(x^2 + y^2)\right)$$

Only looking at the  $x$ -derivative and ignoring the constant  $A$  gives us,

$$\begin{aligned} \frac{\partial}{\partial x_i} \phi_j(\mathbf{r}_i) &= \frac{\partial}{\partial x_i} \left( H_{j_x}(\sqrt{\omega\alpha}x_i) H_{j_y}(\sqrt{\omega\alpha}y_i) \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right) \right) \\ &= \left[ H_{j_y}(\sqrt{\omega\alpha}y_i) \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right) \frac{\partial}{\partial x_i} (H_{j_x}(\sqrt{\omega\alpha}x_i)) \right. \\ &\quad \left. + H_{j_x}(\sqrt{\omega\alpha}x_i) H_{j_y}(\sqrt{\omega\alpha}y_i) (-\omega\alpha x_i) \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right) \right] \\ &= \left[ \frac{\partial}{\partial x_i} (H_{j_x}(\sqrt{\omega\alpha}x_i)) - \omega\alpha x_i H_{j_x}(\sqrt{\omega\alpha}x_i) \right] \\ &\quad \times H_{j_y}(\sqrt{\omega\alpha}y_i) \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right) \end{aligned}$$

Adding this together with the  $y$ -derivative, and we get the gradient of the single particle wave function,

$$\begin{aligned} \nabla_i \phi_j(\mathbf{r}_i) &= \left[ \left( \frac{\partial}{\partial x_i} (H_{j_x}(\sqrt{\omega\alpha}x_i)) - \omega\alpha x_i H_{j_x}(\sqrt{\omega\alpha}x_i) \right) H_{j_y}(\sqrt{\omega\alpha}y_i) \mathbf{i} \right. \\ &\quad \left. + \left( \frac{\partial}{\partial y_i} (H_{j_y}(\sqrt{\omega\alpha}y_i)) - \omega\alpha y_i H_{j_y}(\sqrt{\omega\alpha}y_i) \right) H_{j_x}(\sqrt{\omega\alpha}x_i) \mathbf{j} \right] \\ &\quad \times \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right) \end{aligned} \tag{55}$$

We can now find the Laplacian of the Slater determinant.

$$\frac{\nabla_k^2 \psi_{SD}}{\psi_{SD}} = \sum_{i \neq k}^N \nabla_k^2 d_{ik}(\mathbf{r}) d_{ki}^{-1}(\mathbf{r}) \tag{56}$$

We now need to find the Laplacian of the single particle function.

$$\begin{aligned}
\frac{\partial^2}{\partial x_i^2} \phi_j(\mathbf{r}_i) &= \frac{\partial}{\partial x_i} \left( \left[ \frac{\partial}{\partial x_i} (H_{j_x}(\sqrt{\omega\alpha}x_i)) - \omega\alpha x_i H_{j_x}(\sqrt{\omega\alpha}x_i) \right] \right. \\
&\quad \times H_{j_y}(\sqrt{\omega\alpha}y_i) \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right) \Bigg) \\
&= \left( \frac{\partial^2}{\partial x_i^2} (H_{j_x}(\sqrt{\omega\alpha}x_i)) - 2\omega\alpha x_i \frac{\partial}{\partial x_i} (H_{j_x}(\sqrt{\omega\alpha}x_i)) \right. \\
&\quad \left. - \omega\alpha H_{j_x}(\sqrt{\omega\alpha}x_i) + \omega^2\alpha^2 x_i^2 H_{j_x}(\sqrt{\omega\alpha}x_i) \right) \\
&\quad \times H_{j_y}(\sqrt{\omega\alpha}y_i) \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right)
\end{aligned}$$

Combining the second derivatives of respectively  $x_i$  and  $y_i$ , we get

$$\begin{aligned}
\nabla_i^2 \phi_j(\mathbf{r}_i) &= \left( H_{j_y}(\sqrt{\omega\alpha}y_i) \frac{\partial^2}{\partial x_i^2} (H_{j_x}(\sqrt{\omega\alpha}x_i)) \right. \\
&\quad + H_{j_x}(\sqrt{\omega\alpha}x_i) \frac{\partial^2}{\partial y_i^2} (H_{j_y}(\sqrt{\omega\alpha}y_i)) \\
&\quad - 2\omega\alpha \left[ x_i H_{j_y}(\sqrt{\omega\alpha}y_i) \frac{\partial}{\partial x_i} (H_{j_x}(\sqrt{\omega\alpha}x_i)) \right. \\
&\quad \left. + y_i H_{j_x}(\sqrt{\omega\alpha}x_i) \frac{\partial}{\partial y_i} (H_{j_y}(\sqrt{\omega\alpha}y_i)) \right] \\
&\quad - 2\omega\alpha H_{j_x}(\sqrt{\omega\alpha}x_i) H_{j_y}(\sqrt{\omega\alpha}y_i) \\
&\quad \left. + \omega^2\alpha^2 (x_i^2 + y_i^2) H_{j_x}(\sqrt{\omega\alpha}x_i) H_{j_y}(\sqrt{\omega\alpha}y_i) \right) \exp\left(-\frac{\omega\alpha}{2}(x_i^2 + y_i^2)\right)
\end{aligned} \tag{57}$$

We now have all the ingredients to find the local energy as given by equation (8).

### 2.5.3 Local energy

The local energy is given by

$$E_L = \frac{1}{\psi_T} \left( -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \frac{1}{2} \omega^2 \sum_{i=1}^N r_i^2 + \sum_{i \leq j}^N r_{ij} \right) \psi_T \tag{58}$$

Where the Laplacian is defined as (42). By using the gradients and Laplacians we found for the Jastrow((44),(48)), Slater determinant((54),(56)) and

the single particle wave functions((55),(57)), we will have the full local energy for  $N$  electrons. This equation is a beast in the true meaning of the word, and we will therefore not write it in its full length.

#### 2.5.4 Quantum force

Finding the quantum force as stated in equation (16) is now a simple task, as we already know all the gradients needed,  $\nabla_i \psi_T = \nabla_i \psi_{SD} + \nabla_i \psi_J$ .

$$\mathbf{F}_k = 2 \left( \sum_{i \neq k}^N \frac{\mathbf{r}_{ik}}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} + \sum_{i \neq k}^N \nabla_k d_{ik}(\mathbf{r}) d_{ki}^{-1}(\mathbf{r}) \right) \quad (59)$$

where depending on particle  $k$ , we will either have the gradient of the Slater determinant to be for spin up or down.

#### 2.5.5 $\alpha$ and $\beta$ derivatives in the steepest descent algorithm

From the lecture notes [4], we have that the derivative of a variational parameter is given by

$$\frac{\partial E_L[c_m]}{\partial c_m} = 2 \left[ \left\langle E_L[c_m] \frac{\partial \ln \psi_L}{\partial c_m} \right\rangle - \langle E_L[c_m] \rangle \left\langle \frac{\partial \ln \psi_L}{\partial c_m} \right\rangle \right]$$

where  $c_m$  is the variational parameter we are studying. As for the two electron case, the brackets are indicating expectation values. The  $\psi_L$  is given as,

$$\frac{\partial \ln \psi_L}{\partial c_m} = \frac{\partial \ln(\psi_{SD\uparrow})}{\partial c_m} + \frac{\partial \ln(\psi_{SD\downarrow})}{\partial c_m} + \frac{\partial \ln(\psi_J)}{\partial c_m} \quad (60)$$

Since  $\frac{\partial \ln \psi}{\partial c_m} = \frac{1}{\psi} \frac{\partial \psi}{\partial c_m}$ , we have to find the partial derivatives of the wave functions with respect to the variational parameters. We begin by looking at the Jastrow factor.

$$\begin{aligned} \frac{1}{\psi_J} \frac{\partial \psi_J}{\partial \beta} &= \frac{1}{\psi_J} \exp \left( \sum_{i < j}^N \frac{a_{ij} r_{ij}}{1 + \beta r_{ij}} \right) \\ &= \frac{1}{\psi_J} \left( - \sum_{i \neq j}^N \frac{a_{ij} r_{ij}^2}{(1 + \beta r_{ij})^2} \right) \psi_J \end{aligned}$$

Which leads to,

$$\frac{1}{\psi_J} \frac{\partial \psi_J}{\partial \beta} = - \sum_{i \neq j}^N \frac{a_{ij} r_{ij}^2}{(1 + \beta r_{ij})^2} \quad (61)$$

Now, for the Slater determinant. Since the derivative is the same for either spin up or spin down, we will drop specifying the spin value. In finding this formula, the Jacobi's formula has been used.

$$\frac{\partial \psi_{SD}}{\partial \alpha} = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial \phi_j(\mathbf{r}_i)}{\partial \alpha} D_{ji}^{-1} \quad (62)$$

$$(63)$$

Thus, we find the derivative of the single particle function as

$$\begin{aligned} \frac{\partial \phi_j(\mathbf{r}_i)}{\partial \alpha} = & \left[ \frac{1}{2} \sqrt{\frac{\omega}{\alpha}} \left( x_i H_{j_y}(\sqrt{\omega \alpha} y_i) \frac{\partial}{\partial \alpha} (H_{j_x}(\sqrt{\omega \alpha} x_i)) \right. \right. \\ & \left. \left. - y_i H_{j_x}(\sqrt{\omega \alpha} x_i) \frac{\partial}{\partial \alpha} (H_{j_y}(\sqrt{\omega \alpha} y_i)) \right) \right. \\ & \left. - \frac{\omega}{2} (x_i^2 + y_i^2) H_{j_x}(\sqrt{\omega \alpha} x_i) H_{j_y}(\sqrt{\omega \alpha} y_i) \right] \exp \left( -\frac{\omega \alpha}{2} (x_i^2 + y_i^2) \right) \end{aligned} \quad (64)$$

The derivatives of the Hermite polynomials with respect to  $\alpha$  are given by the Hermite derivative recursion relation seen in the appendix(69), as

$$\frac{\partial}{\partial \alpha} (H_n(\sqrt{\omega \alpha} x)) = n x \sqrt{\frac{\omega}{\alpha}} H_{n-1}(\sqrt{\omega \alpha} x) \quad (65)$$

## 2.6 Blocking

When dealing with large Monte Carlo generated datasets, we will discover that they often are correlated. That is, the random number generator is not truly random. To remedy this, we introduce a statistical technique called *blocking*. The idea behind this, can be summarized as following:

- Define a set of different block sizes  $\{N_b\}$ , in which we will divide our data set into  $n_b$  such blocks.
- Calculate the mean of each block with size  $N_b$ , such that the size of the data set  $\{x_i\}$  is reduced the number of  $n_b$  blocks of size  $N_b$ , thus creating a new dataset  $\{x'_i\}$

- From the new dataset,  $\{x'_i\}$ , we calculate the variance  $\sigma_{N_b}$ .
- Repeat this procedure for all the different block sizes, and plot the resulting variance  $\sigma_{N_b}$  as a function of the block size or number of blocks.

The plot one ends up with will show that the variance stabilizes after some time for a certain range of block sizes. This range is where we expect to have an uncorrelated dataset. The full proof for this can be viewed in the article by Flyvbjerg and Petersen [2].

## 2.7 Parallelization

Using the package Open MPI (Message Passing Interface) for C++, we can parallelize the our program. It is thus beneficial to look into the type of speedup one may get from parallelizing. Let us define the time taken to run the program by one processor to be  $T_1$ , and for  $p$  processor  $T_p$ . We define the possible speed-up as

$$S(p) = \frac{T_1}{T_p} \quad (66)$$

In order to find the possible speedup  $S(p)$  for  $p$  processors, we have the following relation called Amdahl's law[1],

$$S(p) = \frac{1}{(1-f) + f/p} \quad (67)$$

We quickly observe, that if one can obtain a value for the speedup(66), one can find the fraction  $f$  of the program that is parallelizable by,

$$f = \frac{1 - S(p)}{S(p)(1/p - 1)} \quad (68)$$

## 3 Implementation

The Variational Monte Carlo method for solving this many-body problem was implemented in the programming language C++. The results gathered will be for different electron closed shell configurations, specifically  $N = 2, 6, 12, 20$  and  $\omega = 1.0, 0.5, 0.1, 0.05, 0.01$ . Further, the program was run for both uniform sampling and importance sampling as long as it was feasible. The program used can be viewed at GitHub [9].

The program was run on primarily two computers.



- MacBook Pro(Retina, 15-inch, Mid 2015), 2.2 GHz Intel Core i7(4 cores), 16 GB ram.
- Ubuntu 16.04 LTS, Intel Core i7-2600 3.4 GHz(8 cores), 16 GB ram.

### 3.1 Blocking

For all variances found, statistical blocking will have been performed as described in blocking section. An optimal block size to run statistics can be chosen from looking at the subsequent plots for the variance  $\sigma^2$  versus block size.

### 3.2 Program verification and tests

#### 3.2.1 Unperturbed energies

In the case we have unperturbed energies, we should be able to get exact solutions down to the very last decimal. This will provide a useful test to check if our program is correctly set up.

#### 3.2.2 2 electron case with and without Jastrow

For the hard coded case with the Jastrow factor, we should get an energy that equal 3 a.u. as found by Taut [8] when running with Jastrow factor and Coulomb interaction. We have hard coded the two electron case without a Jastrow factor, using wave function seen in the section for the unperturbed two electron case (30) and the local energy (31).

#### 3.2.3 Importance sampling

Assuming that the uniform sampling method has been implemented (11), we should get the exact same result as when running with importance sampling (15). Also, the acceptance rate for importance sampling should be around 99% (see lectures notes [4]).

#### 3.2.4 Parallelization versus non-parallelization

A test were made, running for  $4 \cdot 10^7$  Monte Carlo steps on a MacBook Pro computer. Statistics with all 4 cores and then only a single core were obtained.

## 4 Results

The results for running without importance sampling can be viewed in the following table. We will be comparing results with the paper by Pedersen Lohne et al. [7].

### 4.1 Uniform sampling

The results for using uniform sampling can be viewed in table 1.

Table 1: Results for uniform sampling. Note that for 12 and 20 electrons, we get significant instabilities for small  $\omega$ . The local energies  $\langle E \rangle_1$  is retrieved from the article by Pedersen Lohne et al. [7]. Energies are in atomic units.

$N$	$\omega$	$\langle E \rangle$	$\langle E \rangle_{\text{no-J}}$	$\langle E \rangle_1$	$\sigma^2$	$\langle K \rangle$	$\langle V \rangle$	$N_{MC}$	Acceptance
2	1	3.0004	0.74177	3	$5 \cdot 10^{-12}$	0.89511	2.1053	$10^8$	55.8
2	0.5	1.6607	0.67789	1.6597	$5 \cdot 10^{-12}$	0.43592	1.2248	$10^8$	55.3
2	0.1	0.44174	0.63311	-	$1 \cdot 10^{-12}$	0.092115	0.34962	$10^8$	55.5
2	0.05	0.25506	0.4889	-	$6 \cdot 10^{-13}$	0.045761	0.2093	$10^8$	56.8
2	0.01	0.080072	0.48552	-	$9 \cdot 10^{-13}$	0.0098149	0.070258	$10^8$	57.8
6	1	20.207	0.62917	20.174	$7 \cdot 10^{-10}$	3.7597	16.447	$10^8$	52.6
6	0.5	11.811	0.58145	11.806	$2 \cdot 10^{-10}$	1.7463	10.065	$10^8$	52.8
6	0.1	3.5695	0.55456	-	$3 \cdot 10^{-11}$	0.32922	3.2403	$10^8$	54.3
6	0.05	2.1681	0.56093	-	$1 \cdot 10^{-11}$	0.17099	1.9971	$10^8$	54.3
6	0.01	0.70304	0.5785	-	$4 \cdot 10^{-12}$	0.036174	0.66687	$10^8$	54.7
12	1	65.932	0.57152	65.741	$4 \cdot 10^{-08}$	9.5685	56.364	$10^7$	49.6
12	0.5	39.252	0.49623	39.219	$6 \cdot 10^{-9}$	4.2899	34.962	$10^7$	50.4
12	0.1	12.33	0.36874	-	$1 \cdot 10^{-9}$	0.80906	11.521	$10^7$	51.6
12	0.05	7.6106	0.28426	-	$1 \cdot 10^{-9}$	0.43071	7.1799	$10^7$	51.1
12	0.01	-0.75377	0.20099	-	0.007	-5.3587	4.6049	$10^7$	$2.8 \cdot 10^{-5}$
20	1	156.31	0.52079	155.96	$9 \cdot 10^{-7}$	18.556	137.75	$10^6$	47.9
20	0.5	94.109	0.4415	93.989	$1 \cdot 10^{-7}$	8.4593	85.65	$10^6$	48.3
20	0.1	30.395	0.28958	-	0.001	2.4679	27.927	$10^6$	0.0149
20	0.05	23.516	0.22152	-	0.02	-3.0374	26.553	$10^6$	0.00023
20	0.01	-208.3	0.13049	-	20	-326.16	117.87	$10^6$	$4.5 \cdot 10^{-5}$

### 4.2 Optimal parameters

The optimal parameters as found through 250 steepest descent iterations and with uniform sampling can be viewed in the table below. All uniform sampling results 1 used the optimal parameters as listed in following table.

Table 2: Optimal parameters obtained by running steepest descent for 250 iterations with a step size of  $\Delta t = 0.001$ .

$N$	$\omega$	$\alpha$	$\beta$	$\alpha_{\text{no-J}}$
2	1	0.9949	0.39597	0.74177
2	0.5	0.95089	0.34845	0.67789
2	0.1	0.94509	0.19661	0.63311
2	0.05	0.90622	0.1616	0.4889
2	0.01	0.90493	0.14358	0.48552
6	1	1.0242	0.4374	0.62917
6	0.5	0.93161	0.38809	0.58145
6	0.1	0.83303	0.21247	0.55456
6	0.05	0.84182	0.15325	0.56093
6	0.01	0.84112	0.077169	0.5785
12	1	1.0976	0.42928	0.57152
12	0.5	0.93292	0.41348	0.49623
12	0.1	0.83223	0.22546	0.36874
12	0.05	0.84117	0.1693	0.28426
12	0.01	0.86067	0.072959	0.20099
20	1	1.0597	0.50139	0.52079
20	0.5	0.94244	0.4374	0.4415
20	0.1	0.8492	0.26526	0.28958
20	0.05	0.84246	0.26443	0.22152
20	0.01	0.84	0.1	0.13049

### 4.3 Importance sampling

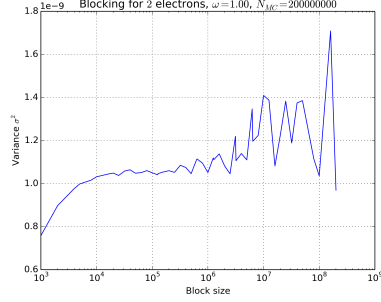
Assuming that the uniform sampling method has been implemented (11), we should get the exact same result as when running with importance sampling (15). Also, the acceptance rate for importance sampling should be around 99%[4]. The  $\Delta t = 0.001$  for the importance sampling runs(same as for the steepest descent). We observe that for  $\omega = 1.0$  with importance sampling, we have a local energy of  $\langle E \rangle = 3.0005 \pm 10^{-4}$  a.u., while without importance sampling we have  $\langle E \rangle = 3.0004 \pm 2 \cdot 10^{-6}$  a.u.

Table 3: Importance sampling results together with energies and variances obtained without importance sampling. Energies are in atomic units.

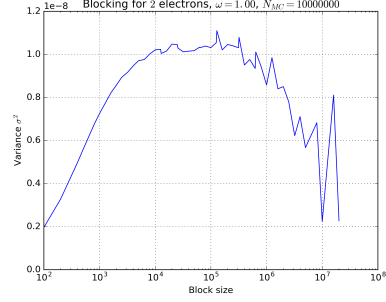
$\omega$	$\langle E \rangle$	$\langle E \rangle_{\text{no-J}}$	$\langle E \rangle_{\text{no imp.}}$	$\langle E \rangle_{\text{no J/imp.}}$	$\langle E \rangle_1$	$\sigma^2$	$\sigma^2_{\text{no imp.}}$	Imp. Acceptance
1	3.0005	3.1696	3.0004	3.1692	3	$6 \cdot 10^{-6}$	$10^{-8}$	99.983545
0.5	1.6607	1.813	1.6607	1.8062	1.6597	$6 \cdot 10^{-6}$	$5 \cdot 10^{-9}$	99.99067
0.1	0.44431	0.52961	0.44174	0.53671	0	$6 \cdot 10^{-6}$	$1 \cdot 10^{-9}$	99.99766
0.05	0.2599	0.32027	0.25506	0.32266	0	$5 \cdot 10^{-6}$	$4 \cdot 10^{-10}$	99.9989
0.01	0.087904	0.11449	0.080072	0.1128	0	$7 \cdot 10^{-6}$	$8 \cdot 10^{-11}$	99.99968

#### 4.4 Blocking

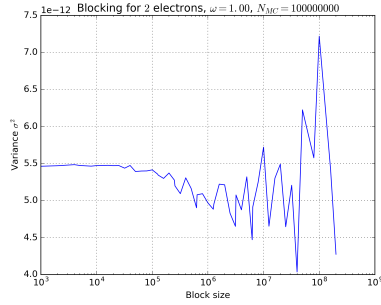
The results for blocking can be viewed in following figures. A subset of all available plots were chosen, as well as the choice to only block for a certain range of blocks. The smallest block-sizes were around  $10^2 - 10^3$  samples per block. This gave a *significant* speed-up when performing blocking. From the resulting plots, acceptable block-sizes for uncorrelated variances were determined to be around  $10^4 - 10^5$ .



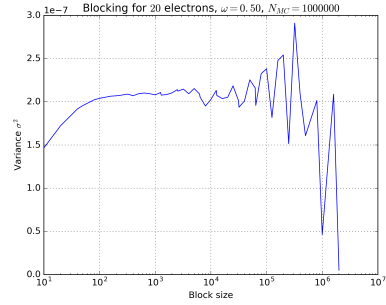
(a) With hard coded importance sampling.  $2 \cdot 10^8$  MC samples,  $\omega = 1.0$ .



(b) With importance sampling.  $10^7$  MC samples,  $\omega = 1.0$ .



(c) 2 electron case without importance sampling.  $10^8$  MC samples,  $\omega = 1.0$ .



(d) 20 electron case without importance sampling.  $10^6$  MC samples,  $\omega = 0.5$ .

Figure 1: Different plots of variance  $\sigma^2$  versus different block-sizes.

## 4.5 Program verification and tests

### 4.5.1 Unperturbed energies

The unperturbed energies can be viewed in the following table.

Table 4: Unperturbed cases where the variational parameter  $\alpha = 1.0$ . All results obtained using  $10^6$  Monte Carlo steps and with uniform sampling.

$N$	$\omega$	$\langle E \rangle$	$\sigma^2$	$\langle K \rangle$	$\langle V \rangle$	Accept.
2	1	2	0	0.99887	1.0011	52.8
2	0.5	1	0	0.4999	0.5001	51.6
2	0.1	0.2	0	0.099835	0.10016	52.7
2	0.05	0.1	0	0.049947	0.050053	54
2	0.01	0.02	$-10^{-20}$	0.0099913	0.010009	56.1
6	1	10	0	5.0015	4.9985	47.6
6	0.5	5	0	2.4999	2.5001	46.5
6	0.1	1	0	0.49971	0.50029	47.6
6	0.05	0.5	0	0.24996	0.25004	48.8
6	0.01	0.1	0	0.050012	0.049988	50.9
12	1	28	0	14	14	43.8
12	0.5	14	0	6.9998	7.0002	42.7
12	0.1	2.8	$2 \cdot 10^{-16}$	1.3993	1.4007	43.8
12	0.05	1.4	$6 \cdot 10^{-17}$	0.70004	0.69996	44.9
12	0.01	0.28	$3 \cdot 10^{-18}$	0.13999	0.14001	46.9
20	1	60	0	30.007	29.993	41
20	0.5	30	0	14.998	15.002	40
20	0.1	6	0	2.9986	3.0014	40.9
20	0.05	3	0	1.4999	1.5001	42
20	0.01	0.6	$-7 \cdot 10^{-18}$	0.29996	0.30004	43.9

#### 4.5.2 2 electron case with and without Jastrow

For the hard coded case with the Jastrow factor, we should get an energy that equal 3 a.u. as found by Taut [8] when running with Jastrow factor and Coulomb interaction. We have hard coded the two electron case without a Jastrow factor, using wave function seen in the section for the unperturbed two electron case (30) and the local energy (31).

We observe that for the hard coded case with Jastrow factor and importance sampling, we have a local energy of  $\langle E \rangle = 3.0004 \pm 10^{-5}$  a.u.

Table 5: Case (a) is for hard-coded 2 electron case with Jastrow and Coulomb included. Case (b) is for hard-coded 2 electron case with Jastrow but no Coulomb interaction. Case (c) is with neither Jastrow or Coulomb interaction included. The acceptance rate for (a) and (b) were both around 98.5%, while for case (c) the acceptance rate were at 56.6%. Energies are in atomic units.

Case	$\langle E \rangle$	$\sigma^2$	$\langle K \rangle$	$\langle V \rangle$	$\alpha$	$\beta$	$N_{MC}$
(a)	3.0004	$10^{-9}$	0.89959	2.1008	0.99651	0.39915	$2 \cdot 10^8$
(b)	2.1243	$2 \cdot 10^{-7}$	0.99493	1.1294	1.0905	0.49336	$2 \cdot 10^8$
(c)	2	0	0.54379	1.4562	0.68672	-	$2 \cdot 10^8$

#### 4.5.3 Parallelization versus non-parallelization

We made five runs with and without parallelization for  $4 \cdot 10^7$  Monte Carlo cycles on four cores.

Table 6: Times for parallelization. Right column is with program run with 1 core, right one is for 4 cores. Run for 2 electrons with  $\omega = 1.0$  for  $N_{MC} = 4 \cdot 10^7$ .

$T_1$ [s]	$T_p$ [s]
22.313027	75.566675
22.484323	74.749889
22.220841	75.247932
22.757712	75.280813

From the values listed in table 6, we get that the speed from equation (66) is  $S(p=4) = 3.351 \pm 0.036$ , and that the fraction of the program that is parallelizable is  $f = 0.702 \pm 0.003$ .

## 5 Discussion and conclusion

The road to achieving a satisfactory result have been long and a difficult one. The target I had in mind when starting was to create a program both flexible and efficient. While the first may be true, the latter still has a lot of potential(the author suspects another month might have been needed).

## 5.1 Issues related to implementation

The most pertinent issue remaining is the one regarding importance sampling for  $N > 2$  electrons, in which singular matrices appear. The main suspect for this was the quantum force. After careful analysis of the code, no faults could be found and root problem remain hidden.

## 5.2 Uniform sampling

However, usable results were achievable despite implementation issues. Results using uniform sampling(see equation (1)) were presented in table 1. One worrying aspect of these results is the low variance. For cases comparable with results from importance sampling, we have almost six orders of magnitude in difference. Since the results appear to be in good agreement with cited articles, it is difficult to pinpoint exact reason for this. It should be pointed out that results without importance sampling were run with  $8 \cdot 10^8$  Monte Carlo cycles, while with importance sampling for only  $8 \cdot 10^7$  cycles.

Instabilities for  $N = 12$  and  $N = 20$  electrons were found for small  $\omega$  values. One possible remedy to this might be to decrease step size  $\Delta t$ , which is a possible improvement for the future.

Let us now look at three cases. For  $N = 6$  electrons we have  $\langle E \rangle = 20.207 \pm 3.0 \cdot 10^{-5}$  a.u., compared to  $\langle E \rangle_1 = 20.174$  a.u. For  $N = 12$  electrons, we have  $\langle E \rangle = 65.932 \pm 2.0 \cdot 10^{-4}$  a.u., compared to  $\langle E \rangle_1 = 65.741$  a.u. At last, for  $N = 20$  electrons, we have  $\langle E \rangle = 156.31 \pm 10^{-3}$  a.u. and a reference of  $\langle E \rangle_1 = 155.96$  a.u. What this tells us, is that our method using *uniform* sampling is correctly implemented.

## 5.3 Optimal parameters from uniform sampling

The optimal parameters found through steepest descent can be viewed in table 2. As seen, there are no obvious outliers. No convergence criteria were implemented, nor any other methods for improving step size as one gradually creeps towards the local minimum. Had time been not been a constraint, methods for achieving an better step sizes  $\Delta t$  when in the vicinity of a minimum, would have been implemented.

## 5.4 Importance sampling

Due to the error with singular matrices and thus not being able to find their inverse, importance sampling was only achievable for  $N = 2$  electrons. Still,



we were able to produce results corresponding good with reference values as well as results obtained without importance sampling. For  $\omega = 1.0$ , we have results that are in close proximity for those from uniform sampling.

## 5.5 Blocking

The results from blocking can be viewed in figure 1. We see that we get significant oscillations in the variance as the block-sizes increases beyond  $10^6$  samples per block. As mentioned earlier, I chose block-sizes to be around  $\sim 10^4 - 10^5$  when retrieving the optimal variances. If we look close at the plot for 2 electrons and  $\omega = 1.0$  with importance sampling 1b, we see that from running importance sampling we get over a order of magnitude in improvement(assuming that the variance keeps increasing for continuously smaller block-sizes).

## 5.6 Verification and tests

A number of different tests were implemented. As we demonstrated in equation (31), the unperturbed energy should come out be exactly 2 a.u.(atomic unites) with zero variance for 2 electrons. For 6, 12 and 20 electrons we should get respectively  $10\omega$ ,  $28\omega$  and  $60\omega$  and zero variance. From the table 4 we observe that this is fulfilled to almost the exact decimal. By almost, I imply that we have certain instabilities for small  $\omega$  values. We can also see that the acceptance ratio is not necessarily at 50% either. This can be attributed to poor calibration, as one need to calibrate this to be 50% for *all* cases. If time were not of the essence, this could be automated. All in all, this is a good indication that our program is correct(at least for uniform sampling).

From table 5 we observe three different hard-coded cases. From comparing these values with the ones obtained from the general case we have a what appears to be a good one-to-one mapping. If we look at the 2 electron case with  $\omega = 1.0$ , for the hard coded case with importance sampling we have  $\langle E \rangle = 3.0004 \pm 10^{-5}$  a.u., and for the general method we have  $\langle E \rangle = 3.0005 \pm 10^{-5}$  a.u. This compared to the results obtained by Pedersen Lohne et al. [7]  $\langle E \rangle = 3.0$ , we observe we have a rather good result for both hard-coded and general method. This, and the high acceptance rate( $\sim 99\%$ ), leads us to believe that our method is working as intended.

We also achieved a speedup by parallelizing our programs(the blocking method was also parallelized). For our main program, we achieved a parallelizable fraction of  $f = 0.702 \pm 0.003$ . The low  $f$  is somewhat spurious,

as one should expect from a naively and simple parallelization with little cross-communication to have a fraction closer to 1. Further investigation is needed to determine the cause of this.

### 5.7 Closing thoughts and remarks.

As we now have observed, results have been obtained even though certain elements of the simulation remains out of scope. We can infer from comparing results with tests and other articles[7], that using uniform sampling have provided good results. Suspicion whether the variance is correct or not still remains, but the local energy  $\langle E \rangle$  appears to be within acceptable ranges. Further investigations into the root bug of using importance sampling would be interesting, but only when available time tends to infinity.

## Appendices

### A Hermite polynomials

The first few Hermite polynomials in the physicists definition that will be used are given as,

$$H_0(x) = 1$$

$$H_1(x) = 2x$$

$$H_2(x) = 4x^2 - 2$$

$$H_3(x) = 8x^3 - 12x$$

$$H_4(x) = 16x^4 - 48x^2 + 12$$

Their derivatives is given as

$$\frac{\partial H_0(x)}{\partial x} = 0$$

$$\frac{\partial H_1(x)}{\partial x} = 2$$

$$\frac{\partial H_2(x)}{\partial x} = 8x$$

$$\frac{\partial H_3(x)}{\partial x} = 24x^2 - 12$$

$$\frac{\partial H_4(x)}{\partial x} = 64x^3 - 96x$$

Their second derivatives goes as

$$\frac{\partial^2 H_0(x)}{\partial x^2} = 0$$

$$\frac{\partial^2 H_1(x)}{\partial x^2} = 0$$

$$\frac{\partial^2 H_2(x)}{\partial x^2} = 8$$

$$\frac{\partial^2 H_3(x)}{\partial x^2} = 48x$$

$$\frac{\partial^2 H_4(x)}{\partial x^2} = 192x^2 - 96$$

For finding the  $\alpha$  derivative, we can use the recursive derivative relation,

$$H'_n(x) = 2nH_{n-1}(x) \quad (69)$$

## B Integral formulas

Solution to a Gaussian integral is given as,

$$\int \exp(-\omega x^2) dx = \sqrt{\frac{\pi}{\omega}} = \sqrt{\pi} \omega^{-\frac{1}{2}} \quad (70)$$

Another symmetric Gaussian integral is given as,

$$\int x^2 \exp(-\omega x^2) dx = \frac{\sqrt{\pi}}{2} \omega^{-\frac{3}{2}} \quad (71)$$

## C Jacobi’s formula

When finding the derivative of the Slater determinant, we will need the Jacobi formula,

$$\frac{d}{dt} \det A(t) = \text{tr} \left( \text{adj}(A(t)) \frac{dA(t)}{dt} \right) \quad (72)$$

## References

- [1] Gene M. Amdahl. Validity of the single processor approach to achieving large scale computing capabilities. In *Proceedings of the April 18-20, 1967, Spring Joint Computer Conference, AFIPS '67 (Spring)*, pages 483–485, New York, NY, USA, 1967. ACM. doi: 10.1145/1465482.1465560. URL <http://doi.acm.org/10.1145/1465482.1465560>.
- [2] Henrik Flyvbjerg and Henrik Gordon Petersen. Error estimates on averages of correlated data. *The Journal of Chemical Physics*, 91(1):461–466, 1989.
- [3] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [4] M. Hjort-Jensen. Computational physics: Lecture notes fall 2015, 2015.
- [5] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 06 1953.
- [6] Jules W. Moskowitz and M. H. Kalos. A new look at correlations in atomic and molecular systems. i. application of fermion monte carlo variational method. *International Journal of Quantum Chemistry*, 20(5):1107–1119, 1981. ISSN 1097-461X. doi: 10.1002/qua.560200508. URL <http://dx.doi.org/10.1002/qua.560200508>.
- [7] M. Pedersen Lohne, G. Hagen, M. Hjorth-Jensen, S. Kvaal, and F. Pederiva. Ab initio. *Phys. Rev. B*, 84:115302, Sep 2011. doi: 10.1103/PhysRevB.84.115302. URL <https://link.aps.org/doi/10.1103/PhysRevB.84.115302>.

- [8] M. Taut. Two electrons in an external oscillator potential: Particular analytic solutions of a coulomb correlation problem. *Phys. Rev. A*, 48: 3561–3566, Nov 1993. doi: 10.1103/PhysRevA.48.3561. URL <https://link.aps.org/doi/10.1103/PhysRevA.48.3561>.
- [9] M. Vege. Project 2 github repository. <https://github.com/hmvege/FYS4411-Project-2>, June 2017.