Project 2

Vilde Mari Reinertsen

Abstract

Many computational physiscs problems involve the many-body Schrödinger equation (SE) and modelling of quantum dots is one of them. The problem with the many-body SE is that it is very hard or impossible to solve it analytically. Therefore, numerical solution are necessary to gain insight into the behaviour of many-body systems of particles. In this particular project I am working with quantum Variational Monte Carlo (VMC) methods to solve the many-body SE for electrons in a harmonic oscillator trap, which here represents the quantum dot. I am looking at full-shell systems of two, six and twelve electrons in harmonic oscillator traps of different strengths, i.e. trap frequencies, and investigating the different systems by comparing expectation energies and one-body densities. Additionally, I look at how these properties change by introducing interaction between the particles.

Contents

1 Introduction

The main objective of this project is to investigate the many-body problem of simulating a quantum dot. A quantum dot is basically electrons that are trapped in an electrical potential. In this project, the potential is modelled as a harmonic oscillator potential. In this project Variational Monte Carlo (VMC) methods are used to solve the many-body SE for electrons. This project is about the so-called full-shell systems of two, six and twelve electrons in harmonic oscillator traps of different strengths, i.e. trap frequencies. The systems is investigated both for the case where interaction is neglected and included.

The report starts out by introducing the system with the representative equations and analysis tools. Most of the numerical tools used in the programming in this project has already been described in project 1 [1], but some additional things are explained or more thoroughly elaborated on in the theory and method part of this report. Furthermore, the results and discussion part first analyses how different important parameters have been chosen, i.e. choice of evaluation of the double derivative, step size of sampling techniques and method and parameters of the optimization. In addition the expectation energies and the one-body densities of the systems involves are compared and evaluated. At last, an evaluation of code's efficiently is made and some concluding remarks are stated.

2 Theory and method

In this project I have investigated a system of $N = 2$, 6 and 12 electrons, where N is the number of particles. It is a so-called closed shell-system. The Hamiltonian used to model this system is

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

where

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

is the single particle part and

$$\hat{H}_1 = \sum_{i<j} \frac{1}{r_{ij}}, \tag{2}$$

represent the interaction potential between particles. The Hamiltonian is written in atomic units, which implies that $\hbar = 1, m = 1$, the unit of length is $a_0 = {}^{4\pi\epsilon_0\hbar^2}/_{m_e e^2}$ and the unit of energy is $E_h = {}^{\hbar^2}/_{m_e a_0^2}$. In Eq. 2 $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ and in Eq. 1 ω is the oscillator frequency or the trap frequency.

2.1 Two particle system

The single particle wave function in two dimensions is

$$\phi_{n_x,n_y}(x, y) = AH_{n_x}(\sqrt{\omega}x)H_{n_y}(\sqrt{\omega}y) \exp\left(-\omega(x^2 + y^2)/2.\right. \tag{3}$$

where the functions $H_{n_x}(\sqrt{\omega}x)$ are Hermite polynomials, while A is a normalization constant. The relevant Hermite polynomials in this project are listed in Appendix C.

For the lowest-lying state, E_{00} (see Fig. 1), we have $n_x = n_y = 0$ and an energy $\epsilon_{n_x,n_y} = \omega(n_x + n_y + 1) = \omega$, the total energy of the lowest-lying state is hence 2ω because there is, due to Pauli's exclusion principle, room for two electrons with opposite spins. The total spin of the ground state for the system with two electron is therfore zero. The same applies for all closed shell systems in the ground state because the filled states all have two electrons with opposite spins that cancel each other. The energies of the other full-shell systems that include an increasing number of states are listed in Appendix B.

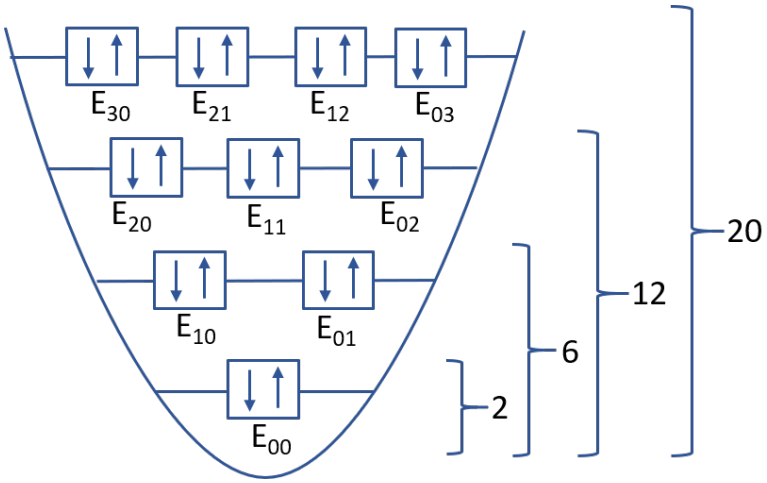

Figure 1: Illustration of the different electron states in a harmonic oscillator trap. The numbers E_{ij} represent the different single-particle states and the states at the same level have the same energy. The arrows show that two electrons with opposite spins can occupy the same state. On the right the full shell systems with the according number of particles are pointed out.

The expectation value can be found by solving the equation

$$\langle E_L \rangle = \frac{\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \psi_T^*(\boldsymbol{r}_1, \boldsymbol{r}_2) \hat{H}(\boldsymbol{r}_1, \boldsymbol{r}_2) \psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2)}{\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \psi_T^*(\boldsymbol{r}_1, \boldsymbol{r}_2) \psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2)}. \tag{4}$$

We use Variational Monte Carlo (VMC) methods to evaluate the Eq. 4 [1]. The exact wave function for two not interacting electrons in the ground state is given by

$$\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2) = C \exp\left(-\omega(r_1^2 + r_2^2)/2\right),$$

where $r_i = \sqrt{x_i^2 + y_i^2}$ and C is a normalization constant. The trial wave function we use for the not interacting case is

$$\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2) = C \exp\left(-\alpha\omega(r_1^2 + r_2^2)/2\right). \tag{5}$$

with the parameter α. From the exact wave function we know that $\alpha = 1$ for the situation without interaction. On the other hand, for the interacting case, the trial wave function for the two-electron system is

$$\psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2) = C \exp\left(-\alpha\omega(r_1^2 + r_2^2)/2\right) \exp\left(\frac{a r_{12}}{(1 + \beta r_{12})}\right), \tag{6}$$

where we introduce another parameter, β, and a spin factor, a. a is 1 when the two electrons have anti-parallel spins and $1/3$ when they have the parallel spins (this is not relevant before one includes more than two particles to the system, as can be seen from Fig. 1). This trial wave function is not the exact wave function, hence the simulation can only approximate the exact values for the ground state energy.

2.2 More particles

Since we are looking at closed shell systems, the next amount of particles are six. We can see this from Fig. 1, there are room for two electrons with opposite spin in the states here named E_{01} and E_{10}, in addition to the two in the lowest lying state. The trial wave function is now given by

$$\psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_6) = Det\left(\phi_1(\boldsymbol{r}_1), \phi_2(\boldsymbol{r}_2), \dots, \phi_6(\boldsymbol{r}_6)\right) \prod_{i<j}^{6} \exp\left(\frac{ar_{ij}}{(1+\beta r_{ij})}\right), \tag{7}$$

where

$$Det\left(\phi_1(\boldsymbol{r}_1), \phi_2(\boldsymbol{r}_2), \dots, \phi_6(\boldsymbol{r}_6)\right) = \begin{vmatrix} \phi_1(\boldsymbol{r}_1) & \phi_2(\boldsymbol{r}_1) & \cdots & \phi_6(\boldsymbol{r}_1) \\ \phi_1(\boldsymbol{r}_2) & \phi_2(\boldsymbol{r}_2) & \cdots & \phi_6(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{r}_6) & \phi_2(\boldsymbol{r}_6) & \cdots & \phi_6(\boldsymbol{r}_6) \end{vmatrix}$$

is the Slater determinant. This determinant occurs because electrons are indistinguishable particles and their wave function is hence antisymmetric. The functions, $\phi_i(\boldsymbol{r}_j)$, are given by Eq. 3 and the notation is explained in Tab. 1.

Table 1: The relation between the notation used in the determinant (left) compared to Eq. 3 (right).

ϕ_1	$\phi_{n_x=0,n_y=0}$	ϕ_7	$\phi_{n_x=2,n_y=0}$
ϕ_2	$\phi_{n_x=0,n_y=0}$	ϕ_8	$\phi_{n_x=2,n_y=0}$
ϕ_3	$\phi_{n_x=1,n_y=0}$	ϕ_9	$\phi_{n_x=1,n_y=1}$
ϕ_4	$\phi_{n_x=1,n_y=0}$	ϕ_{10}	$\phi_{n_x=1,n_y=1}$
ϕ_5	$\phi_{n_x=0,n_y=1}$	ϕ_{11}	$\phi_{n_x=0,n_y=2}$
ϕ_6	$\phi_{n_x=0,n_y=1}$	ϕ_{12}	$\phi_{n_x=0,n_y=2}$

Similarly if we include another shell in our system we get 12 particles and the trial wave function is

$$\psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{12}) = Det\left(\phi_1(\boldsymbol{r}_1), \phi_2(\boldsymbol{r}_2), \dots, \phi_{12}(\boldsymbol{r}_{12})\right) \prod_{i<j}^{12} \exp\left(\frac{ar_{ij}}{(1+\beta r_{ij})}\right). \tag{8}$$

The determinant have the same structure as for six particles and the relation to the single-particle wave functions are shown in Tab. 1.

2.3 Relevant topics for this project

For an explanation of the VMC method, sampling methods, optimization and gradient descent methods, statistical analysis and resampling techniques used in this project see the theory part of Project 1 [1]. However, below is some topics that are explained in further detail or new to this project.

2.3.1 Gradient descent - two parameters

In this project the gradient descent method also known as steepest descent method was used to evaluate the parameters that makes the trial wave function approximate the ground state wave function [2]. Compared to Project 1, there are two parameters to be evaluated. The derivative of the trial wave function with regards to the parameter α is

$$\frac{\frac{\partial \psi_T}{\partial \alpha}}{\psi_T} = -\frac{\omega}{2}\sum_i^N r_i^2 \tag{9}$$

and β is

$$\frac{\frac{\partial \psi_T}{\partial \beta}}{\psi_T} = -\sum_{i<j}^N \frac{a_{ij} r_{ij}^2}{(1+\beta r_{ij})^2}. \tag{10}$$

The derivative used in the gradient descent method is

$$\frac{\partial \langle E_L \rangle}{\partial \alpha} = 2\left(\left\langle \frac{\bar{\psi_T^\alpha}}{\psi_T} E_L \right\rangle - \left\langle \frac{\bar{\psi_T^\alpha}}{\psi_T} \right\rangle \langle E_L \rangle\right) \tag{11}$$

where $\left\langle \frac{\bar{\psi_T^\alpha}}{\psi_T} \right\rangle$ is the expectation value of Eq. 9 and $\left\langle \frac{\bar{\psi_T^\alpha}}{\psi_T} E_L \right\rangle$ is the expectation value of Eq. 9 multiplied with the local energy, and the same applies for the parameter β (except using Eq. 10).

2.3.2 One-body density

The radial one-body density is a measure of the spacial distribution of the electrons with respect to the distance from the origin of the harmonic oscillator trap. To calculate the radial one-body density, I sampled the position of the electrons at each accepted step in the MC cycles. To sample the positions, the distance from the origin to a set cut-off length was separated into bins with a length Δr. For every Monte Carlo step, the distance between the electron's position and the origin was calculated, and the bin that corresponded to the current distance got a count. In the end, one has an array corresponding to the different bins with counts corresponding to how many times an electron was found to have that particular distance to the origin. This array is normalized by dividing by the number of Monte Carlo steps. However, to get the density, one has to divide the number in the bins with the area or volume the bin represents. Because this project involve a two-dimensional problem, and I wanted to calculate the radial one-body density, I divided bin i with the area $A = \pi(r_i + \Delta r)^2 - \pi r_i^2$ where r_i is distance from the origin to bin i. This method to calculate the radial one-body density was found in Ref. [3].

2.3.3 Virial theorem

The virial theorem gives a relation between the time-average total kinetic energy, $\langle T \rangle$, and the time-average external potential energy, $\langle V_{ext} \rangle$, that is

$$2\langle T \rangle = n \langle V_{ext} \rangle \tag{12}$$

where $V(r) = Br^n$. In our case, when interaction is not included, $n = 2$ from the external potential term in Eq. 1, and therefore the average kinetic energy should be equal to the average potential energy. This can be used as a test to see if the simulations executed are correct.

2.3.4 Trap frequency

The trap frequency changes the external potential felt by the electrons. Figure 2 shows how a larger trap frequency, ω, results in a narrower external potential. In this narrow harmonic oscillator trap, the electrons are forced to be closer to each other. Later, I look at how ω affects the energy when the electrons are interacting with each other through the potential given in Eq. 2.

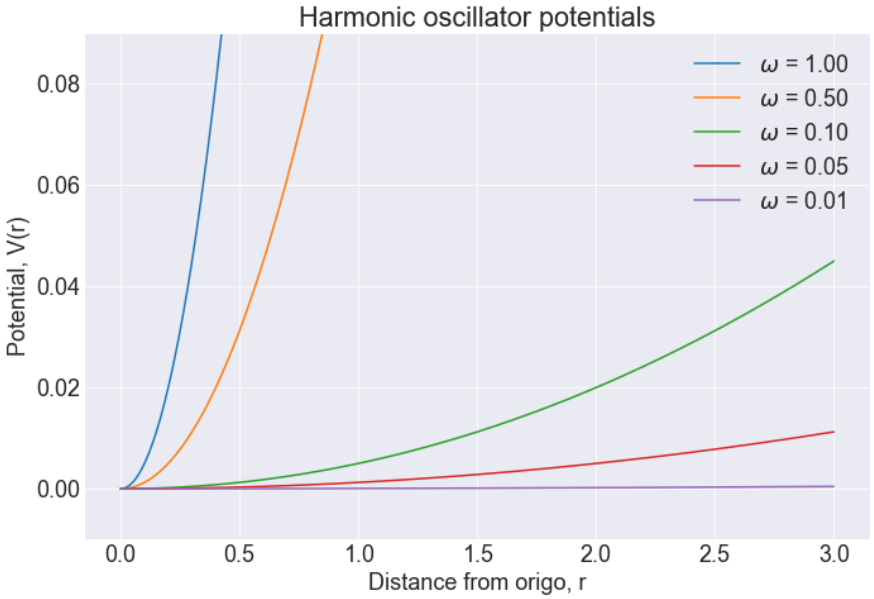

Figure 2: Illustration of how the potential changes when the trap frequency changes.

2.3.5 Evaluating the error

To evaluate the error of the calculated expectation energy one can use the standard error of the mean (SEM)

$$\text{SEM} = \frac{\sigma}{\sqrt{N}} \tag{13}$$

where N is the number of observations, in our case the number of Monte Carlo (MC) cycles. One can easily see that an increased number MC cycles will decrease the SEM.

2.4 The program

The program is found at this GitHub repository and includes comments to describe the code, but below are some general aspects of the program.

2.4.1 Initialization

First the different parts of the program is initialized based on what Hamiltonian that is used and what wave function is used. In this program, one can choose a Hamiltonian with and without interaction (`InteractionHarmonicOscillator` and `HarmonicOscillator`, respectively), and accordingly, a wave function with and without a Jastrow factor (`SlaterDeterminantInteraction` and `Slaterdeterminant`, respectively). Thereafter, the initial state is set up. The particles are distributed randomly after a set distribution. Two different distributions are used and chosen based on the sampling method. For importance sampling a Gaussian distribution is used (`GaussainDistribution`) and

for brute force sampling a random uniform distribution is used (`RandomUniform`). This initialization reflects in the way the sampling of new positions are made by the sampling method, which is explained in project 1 [1].

2.4.2 Metropolis steps

After the initialization the particles are moved one by one, chosen at random, and the new position is found through the sampling method. The metropolis ratio is calculated to determine whether the step is accepted or not. The metropolis ratio is also determined by the sampling method. If the step is accepted, the energy is sampled and if not, the former energy is sampled again. This is continued until all Monte Carlo (MC) cycles have finished.

2.4.3 Sampling

In this program one can choose to sample all local energies for all MC cycles. This makes it possible to performed resampling techniques on the dataset to improve estimate of the error by estimating the correlation between the different metropolis steps, i.e. the samples made of the local energy. For the other outputs, e.g. kinetic, potential and interaction energy and the mean distance, the average is calculated based on the sum of all the sampled values by dividing by the number of MC cycles.

2.5 Improving performance and efficiency

An important part of numerical simulations is computational performance and efficiency. Below are some topics related to this.

2.5.1 Vectorization

Compliers can optimize the code in different ways and one of them is called vectorization. With vectorization more of the memory is used simultaneously and operations are executed on vectors instead of scalars. This makes the code more efficient and can reduce the CPU time of the code. To include vectorization in my program I added

```
set(CMAKE_CXX_FLAGS_RELEASE "-O3")
```

in `CMakeLists`.

2.5.2 Parallelizing

Parallelizing the code is doing operations in parallel instead of sequentially. One can use different methods to parallelize code and combine different methods like MPI and openMP. The different methods can be easy or difficult to implement based on how much you control the the transfer of information between different processes.

In this case the parallelization needed is very simple. We want more MC cycles in less time. In this project I chose to parallelize using a Makefile and cmake. I ran the whole executable with different seeds for the random number generator and different file names to save all the local energies in four different threads on my own personal computer which is a quad-core laptop.

2.5.3 Reducing computational cost

An important part of making the program more efficient is to reduce the number of floating point operations (flops) by implementing a smart code and use expressions for ratios, e.g. the metropolis ratio.

In this particular project one can do many different things to make the code more efficient. The first thing is to split the Slater determinant into a spin-up determinant and a spin down determinant. By doing this one only have to update either of them when a spin up/spin down particle is moved. How this is done is explained further in Appendix A.1. Another method to reduce flops is to find an expression for the metropolis ratio, instead of calculating individual parts and afterwards calculate the ratio. This is explained in Appendix A.2. Other implementations are to updating the inverse instead of calculating the inverse for every step (see Appendix A.3). All of these changes results in another structure when performing the metropolis step (when importance sampling is used):

- Pick particle at random and find new position
- Update Slater determinant row (spin up or spin down one based on the particle that was picked)
- Calculate both old and new quantum force (gradient in a specific position) using the old inverse Slater determinant (see Appendix A.4).
- Calculate the metropolis ratio using the new Slater determinant and the old inverse determinant.
- Update the inverse Slater determinant (only spin up or spin down one based on the particle that was moved) only if the move is accepted.

Using this structure saves many flops since the inverse Slater determinant do not have to be calculated or updated at all if the move is not accepted and, as an example, updating the inverse Slater determinant takes $O(N^2)$ instead of $O(N^3)$ when the inverse is calculated.

3 Results and discussion

3.1 Two electrons in two dimensions

First, I looked at the simple case of two electrons in a harmonic oscillator trap. These electrons do not interact with each other and the trial wave function is given by Eq. 5.

3.1.1 Brute force sampling

Initially, brute force sampling was used to calculate the new position and evaluate the metropolis ratio. The double derivative of the wave function, used to calculate the kinetic energy part of the expectation energy, was evaluated both analytically and numerically to make a comparison. Table 2 shows the energy for different values for the parameter α. The numbers show that the standard error of the mean (SEM) is smaller compared to the estimate of the error from the blocking resampling method, σ_B. This implies that including the correlations, i.e. mainly correlations between one observation and the next where only one particle is moved, increases the estimate of the error. Considering the correlation between the observations gives a more correct estimate of the error, but one has to remember that the deviation from the blocking method only estimate the correlation aspects.

From Tab. 2 one can also observe that $\alpha = 1.0$ gives zero standard deviation and is therefore the optimal parameter. By comparing the results for the analytical and the numerical cases one can notice that the SEM and σ_B is similar for both cases, especially around the ground state ($\alpha = 1.0$). If the expectation energies from the analytical case and the numerical cases are compared, they differ with values at the scale of 10^{-3}, which is reasonable with a σ_B around 10^{-3} to 10^{-2}. At last one can observe, both from the individual CPU time measurements and the mean CPU time of these 10 measurements (though with different αs), that the analytical solution of the double derivative is much faster and more efficient than the numerical case.

Table 2: Comparing the results for analytical/numerical evaluation of the double derivative. Here $\langle E_L \rangle$ is the expectation value for the energy given in atomic units (a.u.) and CPU time is in units of seconds. σ_B is the standard deviation after resampling with the blocking method and SEM is the standard deviation of the mean. Number of MC cycles are 2^{21}.

Analytical:					Numerical:				
α:	$\langle E_L \rangle$:	SEM:	σ_B:	CPU time:	α:	$\langle E_L \rangle$:	SEM:	σ_B:	CPU time:
0.50	2.49402	0.00073	0.01022	5.57812	0.50	2.49991	0.00073	0.01093	18.20310
0.60	2.26441	0.00052	0.00690	5.76562	0.60	2.26412	0.00053	0.00727	18.21880
0.70	2.13118	0.00035	0.00448	5.92188	0.70	2.13039	0.00036	0.00436	18.62500
0.80	2.05016	0.00022	0.00263	5.67188	0.80	2.04993	0.00022	0.00269	18.46880
0.90	2.01015	0.00010	0.00116	5.96875	0.90	2.01160	0.00010	0.00118	18.46880
1.00	2.00000	0.00000	0.00000	5.62500	1.00	2.00000	0.00000	0.00000	18.31250
1.10	2.00871	0.00009	0.00102	6.20312	1.10	2.00825	0.00009	0.00097	18.35940
1.20	2.03402	0.00018	0.00175	6.34375	1.20	2.03308	0.00018	0.00170	18.31250
1.30	2.07259	0.00026	0.00244	5.95312	1.30	2.06460	0.00026	0.00243	20.23440
1.40	2.11041	0.00034	0.00311	6.15625	1.40	2.11803	0.00033	0.00308	19.00000
Mean CPU time: 5.91875					Mean CPU time: 18.62033				

3.1.2 Including importance sampling

Figure 3 compares the expectation value of the energy and the acceptance rate of brute force sampling and importance sampling. It can be observed from the right part of the figure that the acceptance rate of both methods increase with decreasing step size. These observations could indicate that a small step size would be ideal for both methods. Another observation is that even though the acceptance rate is increasing with smaller step sizes which is a good thing, σ_B is also increasing with the step size for both sampling methods which is not a good thing if σ_B is a good estimate of the true error. It could be that this is a balance between number of MC cycles and step size. That, with a smaller step size, one needs a larger number of MC cycles to move the particles enough around to get a true picture of the system, and hence decrease the error.

Furthermore, from the left part of the figure it can be observed that two of the expectation values for the energies are lower than the ground state energy ($dl = 0.005$ and $dl = 0.01$ with brute force sampling) even though these calculations where done with $\alpha = 0.9$. This should not be possible because the minimum in energy occur for the exact ground state wave function which has $\alpha = 1.0$. In addition, in Tab. 3 which compare the result of brute force sampling and importance sampling for

different step sizes, one can observe that σ_B is larger for the case of brute force sampling with a step size of 0.005 and 0.01. However, the SEM is smaller at least for $dl = 0.005$ compared to the larger step sizes with brute force sampling and the same step sizes with importance sampling.

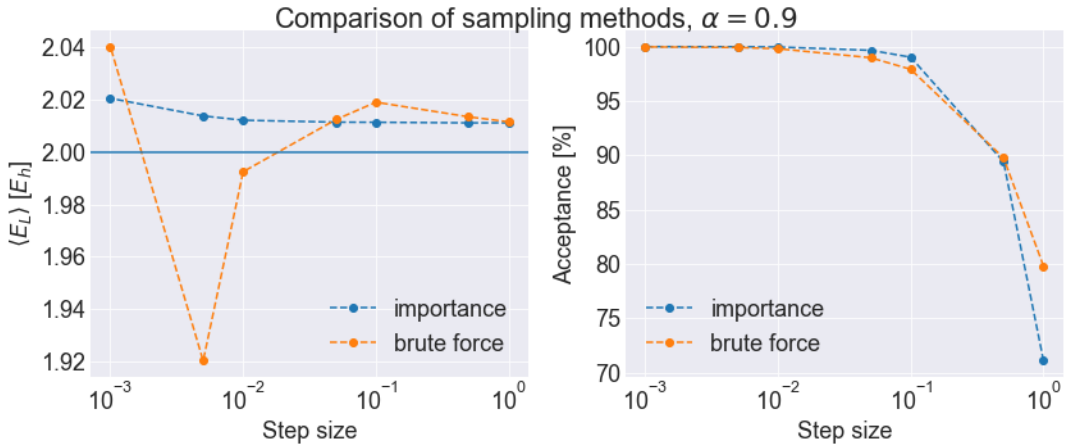

Figure 3: Left: Expectation energies after 2^{21} MC cycles for different step sizes. Right: Percentage of accepted steps for different step sizes. Here importance sampling and brute force sampling is compared.

Table 3: Comparing the results for importance/brute force sampling. Here the parameter α is set to 0.9 and number of MC cycles are 2^{21}. Acc. is short for acceptance and is here given in % and t_{CPU} is the CPU time used by the program in units of seconds. The rest of the values are as described in Tab. 2.

	Brute force:					Importance:				
dl:	$\langle E_L \rangle$:	SEM:	σ_B:	Acc.:	t_{CPU}:	$\langle E_L \rangle$:	SEM:	σ_B:	Acc.:	t_{CPU}:
1.000	2.011	0.00010	0.00067	79.794	7.953	2.011	0.00010	0.00022	71.078	8.531
0.500	2.013	0.00010	0.00113	89.762	7.719	2.011	0.00010	0.00023	89.343	8.672
0.100	2.019	0.00011	0.00540	97.911	7.438	2.011	0.00010	0.00047	99.049	8.859
0.050	2.012	0.00010	0.00801	98.980	7.922	2.011	0.00010	0.00068	99.662	8.188
0.010	1.992	0.00010	0.02282	99.809	8.141	2.007	0.00010	0.00136	99.968	8.281
0.005	1.920	0.00004	0.01082	99.922	7.672	2.010	0.00010	0.00199	99.989	8.547
0.001	2.040	0.00003	0.01043	99.976	7.516	2.007	0.00010	0.00410	99.999	8.016
	Mean CPU time: 7.766					Mean CPU time: 8.442				

I took a closer look at the actual local energies for the brute force sampling method to investigate the cases that gave $\langle E_L \rangle < 2.0$. Figure 4 shows that the energy is not stable for steps sizes smaller than 0.01, so even though the step size 0.001 seems to give reasonable expectation values for the energy (see Tab. 3 and Fig. 3), Fig. 4 shows that it is a lucky shot. I also saw that it was a lucky shot by running the calculation with brute force sampling and the step size, 0.005, with different seeds for the random number generator. The expectation energy for five different runs where $\langle E_L \rangle = 1.91487$, 2.03452, 1.90805, 1.88356 and 1.9284. From Fig. 4 one can observe that even $dl = 0.1$ seems to be too small since it also results in the local energy varying slowly and taking longer "trips" to higher energies and using many steps to get back down again, but for this step size the "trips" to higher energies are more frequent than for the smaller step sizes. I concluded that a step size of 0.5 is the

best choice for the brute force sampling because it gives reasonable changes of the local energy and an acceptance rate of $\sim$ 90 % (see Fig. 3).

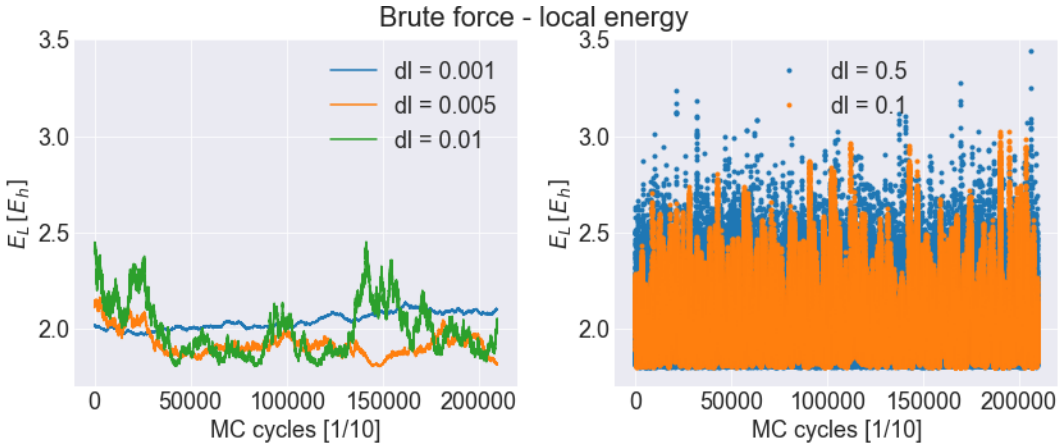

Figure 4: The local energy for every tenth MC cycle for brute force sampling at different step sizes, dl. a) shows the smaller step sizes and b) some that are a bit larger.

Proceeding to evaluate the energy, Table 4 shows how the energy changes with different trap frequencies, ω. The energy in the ground state for two electrons is given by 2ω and this is also the case in Tab. 4. The table also shows that the kinetic and potential energy follow the virial theorem, i.e. they are equal. They only differ by approximately $\pm 10^{-2}$ E_h.

From Tab. 4 one can observe that the mean distance is increasing with decreasing trap frequency. This is as expected from Fig. 2, where the potential is broadened with decreasing trap frequency and hence is not forcing the particles closer together to the same degree. The mean distance is a bit different for brute force sampling compared to importance sampling, but the similarity might improve if more MC cycles are used. The other value show, however, no large difference between the results from brute force sampling and importance sampling. But to be able to compare the sampling methods more thoroughly it is better to look at the case where the system is not in the ground state.

Table 4: Ground state energy of two electrons in a harmonic oscillator trap. Here $\overline{r}_{12}$ is the mean distance between the two electrons at positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ given in units of a_0 and $\langle E_L \rangle$, $\langle T \rangle$, $\langle V_{ext} \rangle$ and $\langle V_{int} \rangle$ are the expectation value of the local energy, kinetic energy, potential energy and interaction energy, respectively, given in units of E_h. The rest of the values are as explained in Tab. 2. Number of MC cycles are 2^{23}.

	Brute force:					Importance:				
ω	α	$\langle E_L \rangle$	$\overline{r}_{12}$	$\langle T \rangle$	$\langle V_{ext} \rangle$	α	$\langle E_L \rangle$	$\overline{r}_{12}$	$\langle T \rangle$	$\langle V_{ext} \rangle$
1.00	1	2.00	1.250	1.0008	0.9992	1	2.00	1.254	0.9982	1.0018
0.50	1	1.00	1.775	0.4971	0.5029	1	1.00	1.781	0.4973	0.5027
0.10	1	0.20	3.967	0.0996	0.1004	1	0.20	4.046	0.0987	0.1013
0.05	1	0.10	5.638	0.0497	0.0503	1	0.10	5.534	0.0512	0.0488
0.01	1	0.02	12.631	0.0099	0.0101	1	0.02	12.488	0.0101	0.0099

Table 5 shows the expectation value of the energy for various parameters, α. The calculated expectation values for the energy for the two different sampling methods are similar, especially close to the correct

parameter α, varying only by $\pm 10^{-2}$. However, what is different is that σ_B is larger for importance sampling than for brute force sampling. I do not know why there is a difference there. Maybe the estimate of the error is more correct, but the blocking code is the same, hence it would have to be that the sampling results in a different correlation. On the other hand, it could be that there is a larger error when using importance sampling. But it could also be that I should have chosen a larger step size for importance sampling since σ_B increased with the decrease of the step size.

At last, one can observe that the CPU time of the importance sampling method is larger than brute force sampling which is as expected because of the calculation of the gradient in the quantum force. But we know from Fig. 3 that the acceptance rate of brute force sampling is around 90 % compared to importance sampling which should be close to 100 % and this makes the importance sampling technique more effective in terms of MC cycles.

Table 5: Comparing the results for brute force sampling/importance sampling. Values are as explained in Tab. 2. Number of MC cycles are 2^{21}.

	Brute force:				Importance:			
α:	$\langle E_L \rangle$:	SEM:	σ_B:	CPU time:	$\langle E_L \rangle$:	SEM:	σ_B:	CPU time:
0.50	2.49074	0.00072	0.01030	6.26562	2.52950	0.00075	0.01952	8.53125
0.60	2.27645	0.00053	0.00721	6.53125	2.26531	0.00052	0.01281	8.62500
0.70	2.12481	0.00036	0.00451	6.59375	2.12449	0.00036	0.00813	8.76562
0.80	2.04873	0.00022	0.00265	6.68750	2.04797	0.00022	0.00454	9.37500
0.90	2.01135	0.00010	0.00115	6.82812	2.01171	0.00010	0.00203	8.81250
1.00	2.00000	0.00000	0.00000	6.75000	2.00000	0.00000	0.00000	8.76562
1.10	2.00863	0.00009	0.00097	6.59375	2.00954	0.00009	0.00171	8.40625
1.20	2.03343	0.00018	0.00174	7.29688	2.03276	0.00018	0.00305	8.50000
1.30	2.07103	0.00026	0.00246	6.75000	2.07363	0.00026	0.00440	8.67188
1.40	2.11725	0.00033	0.00317	6.70312	2.10697	0.00034	0.00536	8.67188
	Mean CPU time: 6.7000				Mean CPU time: 8.7125			

3.1.3 Including optimization

In this project there are two parameters to optimize and hence it is a more complicated problem. I chose, in this project compared to the previous one, to experiment with the minimization rate during the actual optimization. I anticipated that the ideal minimization rate would be different for the different size systems and therefore an analysis of one system would not transfer to the other systems. I started out using a minimization rate, $\gamma = 0.5$. It resulted in few steps until the parameter value stabilized both for guesses close to the optimal value and for guesses far away from the optimal value, but for the smallest trap frequencies I had to use $\gamma = 0.1$ or 0.2. For the case of two interacting fermions, the parameters were optimized by trying out different first guesses for α and β and tuning γ so that the parameters stabilized during the first 200 iterations. The optimal parameters were extracted from the mean of the last 50 iterations. An example run is shown in Fig. 5 for $\omega = 0.5$.

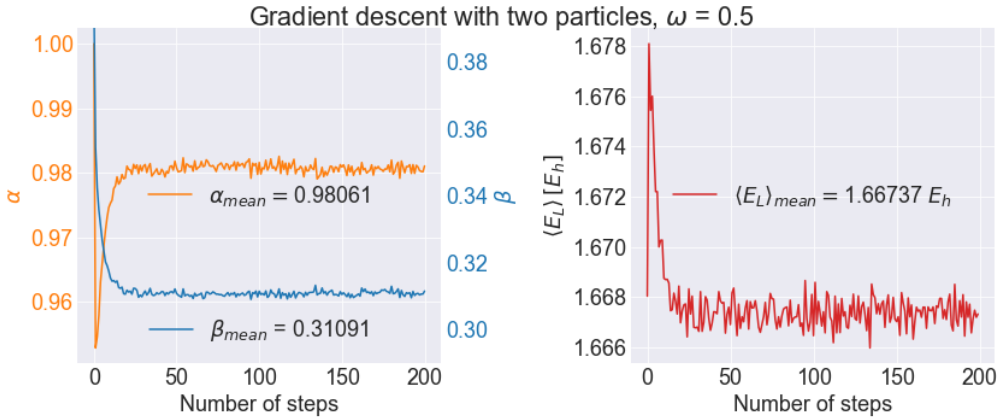

Figure 5: Left: The development of the parameters during the steps of gradient descent. The values α_{mean} and β_{mean} printed on the figure is the mean of the last 50 values. Right: The expectation value after 2^{19} number of MC cycles. Here also, the value printed on the figure is the mean of the last 50 iterations.

For the system with six interacting electrons, the method described above was used for the largest ωs (i.e. $\omega = 1.0, 0.5, 0.1$), but for the smaller ones I had to use a smaller minimization rate (i.e. 0.01-0.05) and I also had to move step by step from $\omega = 0.1$ to $\omega = 0.01$ with the step size of $\Delta\omega = 0.01$. I found the parameters for the current ω and used that as a guess for the next ω. I tried to do it in a more efficient way and let the simple gradient descent method find the minimum on its own, but with this unguided method, the optimization ended up in local minima at higher energies or going to infinite energies. To improve the code, I attempted to use the extended gradient descent method, described in project 1, which utilize the previous gradient to find the new parameter [1]. But the attempt did not improve the behaviour described earlier. Compared to project 1, this system's local energy dependence on the parameters ($\langle E_L \rangle (\alpha, \beta)$), is more complicated, and also involve two parameters instead of one, which makes it harder to optimize the parameters with this simple gradient descent method.

3.1.4 Including interaction

Here is the results of the expectation values approximating the ground state energy for two interacting electrons in a harmonic oscillator trap. The equations used to model the system is as described in the theory part of this report. That includes the trial wave function (see Eq. 6) with the Jastrow factor that is used to model the many-body wave function. Table 6 shows the results of the calculations with the optimal parameters found with the gradient descent method with importance sampling.

Table 6: Ground state energy of two interacting electrons in harmonic oscillator trap found with importance sampling. Here $\overline{r}_{12}$, $\langle T \rangle$, $\langle V_{ext} \rangle$ and $\langle V_{int} \rangle$as explain in Tab. 4. The rest of the values are as explained in Tab. 2. Number of MC cycles are 2^{23}

ω	α	β	$\langle E_L \rangle$	SEM	σ_B	$\overline{r}_{12}$	$\langle T \rangle$	$\langle V_{ext} \rangle$	$\langle V_{int} \rangle$
1.00	0.98846	0.39954	3.0069	0.00001	0.00018	1.643	0.8931	1.3052	0.8086
0.50	0.98082	0.31068	1.6674	0.00001	0.00019	2.481	0.4547	0.6997	0.5130
0.10	0.94734	0.17810	0.4486	0.00001	0.00023	6.724	0.0989	0.1787	0.1710
0.05	0.92262	0.14090	0.2610	0.00000	0.00021	10.333	0.0495	0.1024	0.1091
0.01	0.88305	0.07366	0.0776	0.00000	0.00010	29.293	0.0131	0.0283	0.0362

Table 6 shows that the expectation energy for the ground state with $\omega = 1.0$ is $3.0069 \pm 0.0002\ E_h$. This result is not that far away from the exact value of $3\ E_h$ from Taut's work [4]. The energy of the system is decreasing with decreasing trap frequency, ω. One can also observe that the potential energy dominates for large trap frequencies ($\omega = 1.0$ and $\omega = 0.5$), but for the smaller ωs the interaction energy and the potential energy is approximately equal. This seems to be reasonable since the potential energy is sort of stored in the force from the trap. The potential keeps the electrons together which means that a larger trap frequency gives a stronger trapping force and hence a larger potential energy. For all ω, the kinetic energy is the smallest one and it also decrease with the decreasing trap frequency. This also makes sense because the energy of the individual electron will oscillate between potential energy (either from the external potential or the potential from the interaction) and kinetic energy and when the potential energy decrease, the kinetic energy will decrease with it.

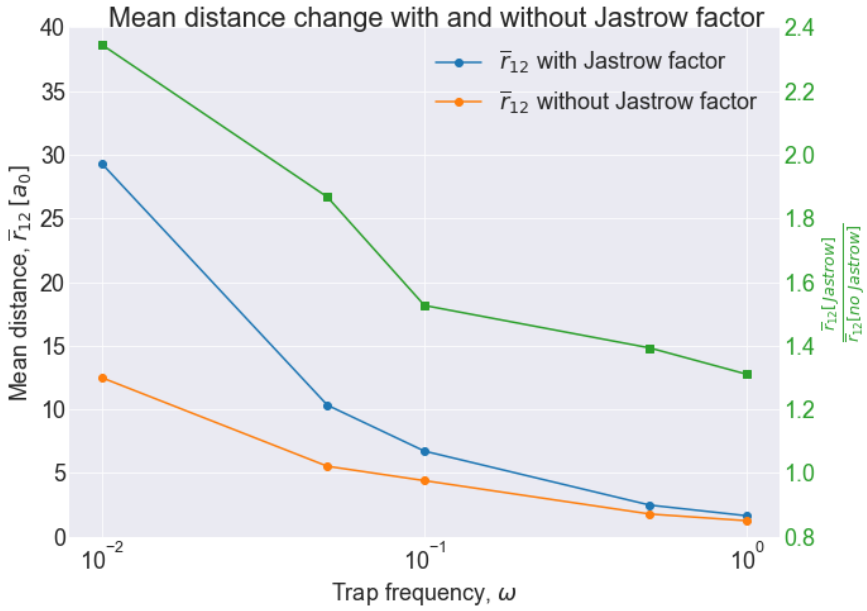

Figure 6: The mean distance between the two particles, calculated with importance sampling, compared for the situation with and without the Jastrow factor and at different trap frequencies.

Figure 6 shows the combined results from Tab. 4 and 6. One can observe that the mean distance is larger for the case with interaction. This is expected since the interaction potential (see Eq. 2) is a repulsive potential, and will hence force the electrons further apart. One can also observe that the ratio of the two different cases (right axis and green color) increase with decreasing trap frequency, i.e. the mean distance increases more for smaller trap frequencies. This could imply that the trap frequency is influencing the mean distance even more than the inclusion of interaction.

3.1.5 One-body density

The one-body density tells how the particles are distributed in space, and Fig. 7 shows the result of the one-body density of the system of two not interacting fermions for different trap frequencies. One can observe, as one saw in the development of the mean distance (see Fig. 6), that the particle distribution or density is spread out in space for smaller trap frequencies. From Fig. 7 one can also observe that the electrons spend most of their time around the origin of the harmonic oscillator trap. The shape of the curve seems not to change for the different trap frequencies, only the height and width.

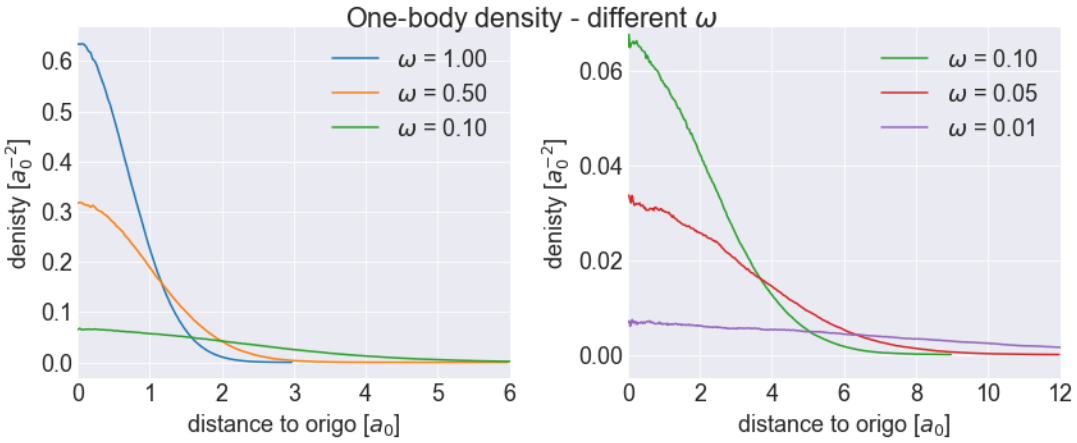

Figure 7: One-body density of the two electron system without the Jastrow factor, but with different trap frequencies. Left: The larger trap frequencies. Right: The smaller trap frequencies. The system with $\omega = 0.1$ is in both plots to show the difference in scale for the two plots. Used 2^{24} MC cycles.

To get a smooth curve, I found out that you either have to use larger bin sizes or more MC cycles. Since the one-body density is found by counting how many times a particle is situated within a certain bin, i.e. with a certain distance to origin, more samples of the positions will give a more true distribution. So if the bin size is small, you need more MC cycles to smooth out the random difference based on the number of samples between two neighbouring bins. On the other hand, if the bin size is too large real differences or fast changes in the distribution might not be detected.

Furthermore, Fig. 8 shows the one-body densities with and without the Jastrow factor. Here, one can also observe the same as was seen with the mean distance, that the Jastrow factor spreads out the distribution of the particles, but the difference is a lot clearer in the one-body density plot. Another interesting observation is that, for the interacting case, the maximum does not seem to lay at the origin. It is not that easy to see for the larger trap frequencies, but one can easily see it for $\omega = 0.1$ in the right part of Fig. 8. This seems to be reasonable because there is a repulsive force between the particles, they spend more time further away from the origin. It could be that for larger trap frequencies there is a larger energy related gain (i.e. the energy is lower) to be situated closer to origin, hence the maximum is not shifted as far from origin for those frequencies, but for the lower trap frequencies the gain obtained closer to origin does not compensate enough of the repulsive forces (or interaction energy), so the maximum is shifted away from origin. That should imply that the position of the maximum reflects a balance between the potential energy and the interaction energy.

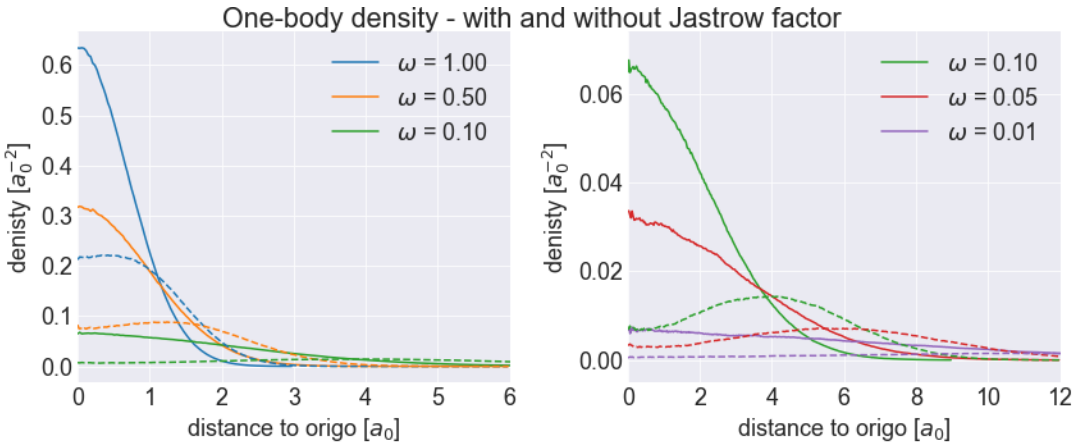

Figure 8: One-body density of the two electron system with and without the Jastrow factor and with different trap frequencies (indicated by the color of the lines). Dashed lines are with the Jastrow factor and filled lines are without Jastrow factor. Left: The larger trap frequencies. Right: The smaller trap frequencies. The system with $\omega = 0.1$ is in both plots to show the difference in scale for the two plots. Used 2^{24} MC cycles.

3.2 Six particles

To investigate the quantum dot system further, the next shell (illustrated in Fig. 1) is filled and this results in a system of six particles. The expectation value for the total energy (i.e. the local energy), the kinetic energy and the potential energy for the different trap frequencies without interaction are listed in Tab. 7. One can observe that for all the trap frequencies, the kinetic and potential energy seems to follow the virial theorem (see Eq. 12), if one allows for a small difference of maximum ± 0.03 a.u.. The energies are as expected from the calculated exact energies in Tab. 14 in Appendix B.

Table 7: Ground state energy of six electrons in a harmonic oscillator trap. Here $\langle E_L \rangle$, $\langle T \rangle$ and $\langle V_{ext} \rangle$ are as explain in Tab. 4. Number of MC cycles are 2^{23}.

ω	$\langle E_L \rangle$	$\langle T \rangle$	$\langle V_{ext} \rangle$
1.00	10.00	4.9865	5.0135
0.50	5.00	2.4973	2.5027
0.10	1.00	0.4861	0.5139
0.05	0.50	0.2483	0.2517
0.01	0.10	0.0454	0.0546

Furthermore, the result after including interaction is showed in Tab. 8. One can observe, if one compares the result with Tab. 6, that the both the SEM and σ_B are larger for the system with six electrons. This could be explained by the fact that the total energies are larger for this system, and hence the errors will be larger. If one compare the error as a fraction for the case with $\omega = 1.0$ $\sigma_B/\langle E_L \rangle = 5.99 \cdot 10^{-5}$ for two electrons and $3.98 \cdot 10^{-4}$ for six electrons and one can see that the error is larger for the six electron case as a fraction as well. This made me wonder if σ_B would be worse if the optimal parameters where worse, but σ_B indicate how good the expectation value is for the parameters given. It tells if the expectation value that comes out can be relied upon. It does not tell us how far the expectation value is from the real energy.

Table 8: Ground state energy of six interacting electrons in harmonic oscillator trap found with importance sampling. Number of MC cycles are 2^{23}.

ω	α	β	$\langle E_L \rangle$	SEM	σ_B	$\langle T \rangle$	$\langle V_{ext} \rangle$	$\langle V_{int} \rangle$
1.00	0.71567	0.49372	20.4492	0.00022	0.00813	2.3429	10.7076	7.3988
0.50	0.75823	0.34260	11.9868	0.00011	0.00522	1.3226	5.8094	4.8548
0.10	0.78852	0.15041	3.6542	0.00003	0.00416	0.2951	1.7035	1.6556
0.05	0.76518	0.10733	2.2223	0.00003	0.00436	0.1178	1.0882	1.0162
0.01	0.64907	0.05085	0.7191	0.00001	0.00586	0.0021	0.3803	0.3367

I wanted to take a closer look at the optimized parameters for the two different systems too, especially since I had so much trouble finding the parameters for the six electron system. Figure 9 compares the optimized parameters for the two systems with different amounts of particles at the different trap frequencies. One can observe that the β parameter which is involved in the Jastrow factor is similar for both systems and also show a similar development from one trap frequency to the next. This is expected, I guess, since the interaction between the particles do not change that much by increasing the number of particles. The other parameter α, however, which is involved in the Slater determinant, shows completely different values and development for the two systems. This is also reasonable because the wave function, i.e. the Slater determinant part, changes a lot when we increase the number of particles. Because when the number of particles is increased higher states with very different single-particle wave functions are filled and the Slater determinant is changed a lot. The development of α though is strange, I would have expected it to either increase or decrease with an increasing trap frequency, but in the right part of Fig. 9 one can observe that the largest α is found for $\omega = 0.1$ which is the middle ω investigated. But my expectations might be wrong, maybe the strange development is due to the single-particle states that are introduced with the higher energy states.

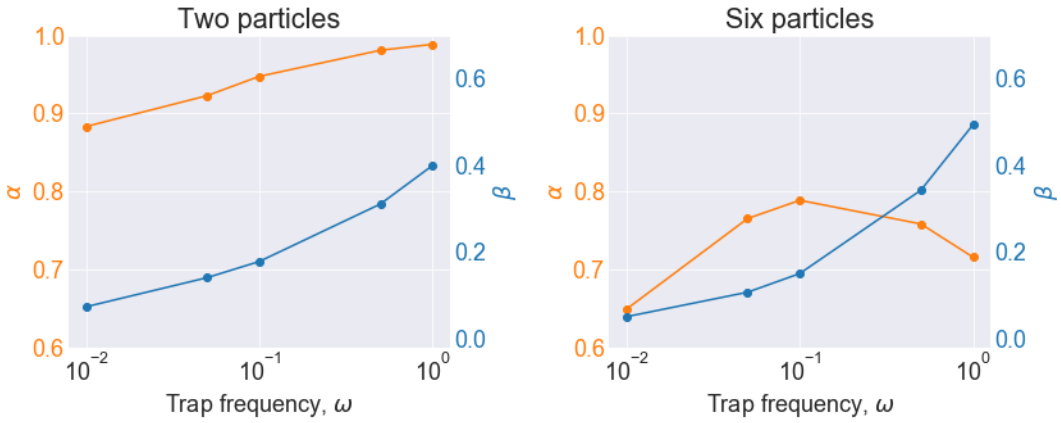

Figure 9: Left: The optimized parameters for the system with two particles and with the different trap frequencies. Right: The optimized parameters for for the system with six particles and with the different trap frequencies. All found with the simple gradient descent method.

Next, Fig. 10 compares the energy per particle of the two systems when interaction is included. The left part of the figure shows that the magnitude of kinetic energy per particle and the development of the kinetic energy with trap frequency is very similar for both systems (two and six particles). The

right part of the figure shows that both the potential and interaction energy per particle is larger for the six particle system at the same trap frequency, than the two particle system. This is as expected since the trap frequency determine the "size" of the trap. The six particle system's potential energy is hence larger for the same size system when there are more particles involved. The interaction energy is also larger when more particles, that are repulsed by each other, are forced together in the "same amount of space" (it is not exactly the same amount of space since the space is not strictly limited, but it is intuitive to image it like that). Apart from this, the six particle system show the same behaviour as the two particle system, with regards to energy.

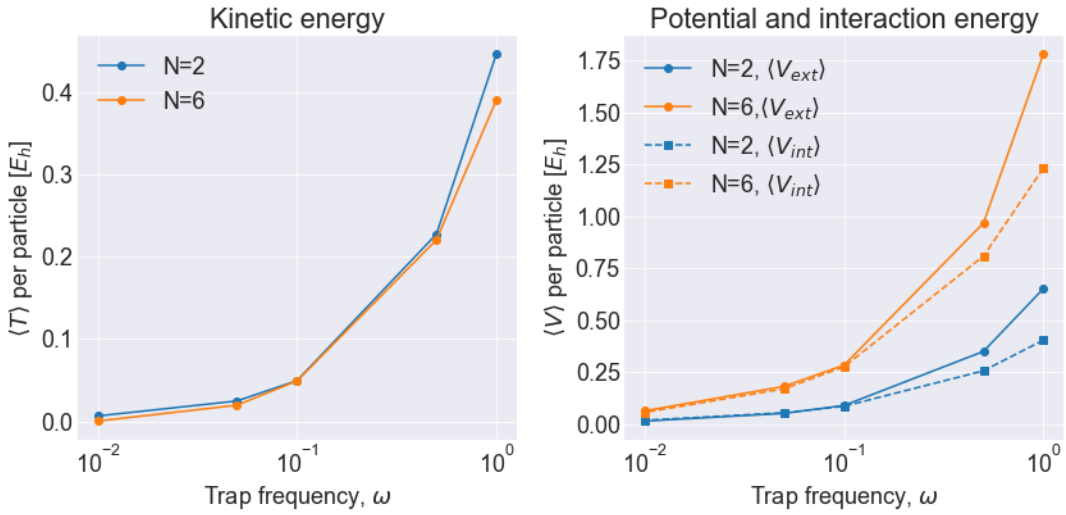

Figure 10: Left: The expectation value for the kinetic energy per particle for the two different systems examined at the different trap frequencies. Right: The expectation value of the potential and interaction energy per particle for the two different systems at the different trap frequencies.

Proceeding to the one-body density, Fig. 11 shows the one-body density for the six electron system. First, one can notice that the shape is different from the two electron system. Here the maximum density is not at the origin of the trap, but approximately a_0 away from the origin for $\omega = 1.0$ without the Jastrow factor. Furthermore, one can observe that the maximum moves further and further away from the origin when ω increases. Regarding the inclusion of interaction, the same behaviour is seen here, compared to the two electron one-body density, that including interaction broadens the whole curve and lowers the maximum density. The shape of the one-body density curve seems to be similar with and without interaction except for $\omega = 0.1$ and maybe also smaller ω where there is another maximum closer to the origin. Maybe the broadening of the trap with decreasing ω opens up for a redistribution of the particles when the repulsive interaction is included.

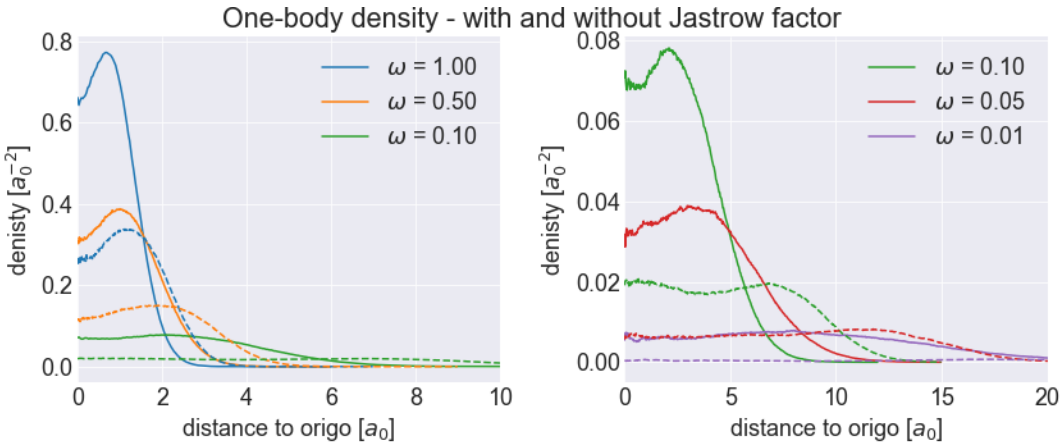

Figure 11: One-body density of the six electron system with and without the Jastrow factor and with different trap frequencies (indicated by the color of the lines). Dashed lines are with the Jastrow factor and filled lines are without Jastrow factor. Left: The larger trap frequencies. Right: The smaller trap frequencies. The system with $\omega = 0.1$ is in both plots to show the difference in scale for the two plots. Used 2^{24} MC cycles.

3.3 Twelve particles

By including another shell in the two dimensional harmonic oscillator trap, the system consists of twelve particles. The ground state energy at different ωs are listed in Tab. 9 and the energies are the same as the exact energies calculated from the energy of the different states (see Tab. 14 in Appendix B).

Table 9: Ground state energy of twelve electrons in a harmonic oscillator trap. Here $\langle E_L \rangle$, $\langle T \rangle$ and $\langle V_{ext} \rangle$ are as explain in Tab. 4. Number of MC cycles are 2^{23}.

ω	$\langle E_L \rangle$	$\langle T \rangle$	$\langle V_{ext} \rangle$
1.00	28.00	14.0117	13.9883
0.50	14.00	7.0463	6.9537
0.10	2.80	1.4084	1.3916
0.05	1.40	0.6901	0.7099
0.01	0.28	0.1419	0.1381

I tried to use the same method, the simple gradient descent method, to evaluate the parameters that gives me the closest thing to the ground state I can get with the trial wave function used in this project, but it seems like the system got too complicated with twelve particles. An example of the behaviour of the gradient descent method is showed in Fig. 12. One can observe, from the right part of the figure that the energy seems to be converging towards a minimum, but it is a very slow process. However, the left part of the figure shows how much the parameters are varying and that the parameter α is negative. I do not think that a negative α is a good sign especially after comparing my results with Ref. [3] where $\langle E_L \rangle = 65.7026(4)$ was found for the same system. It is strange that I get an energy that much smaller than the energy calculated in Ref. [3] considering that all wave functions except the ground state wave function should give larger energies than the ground state energy. However, the negative α could be the explanation if it makes the wave function not physical.

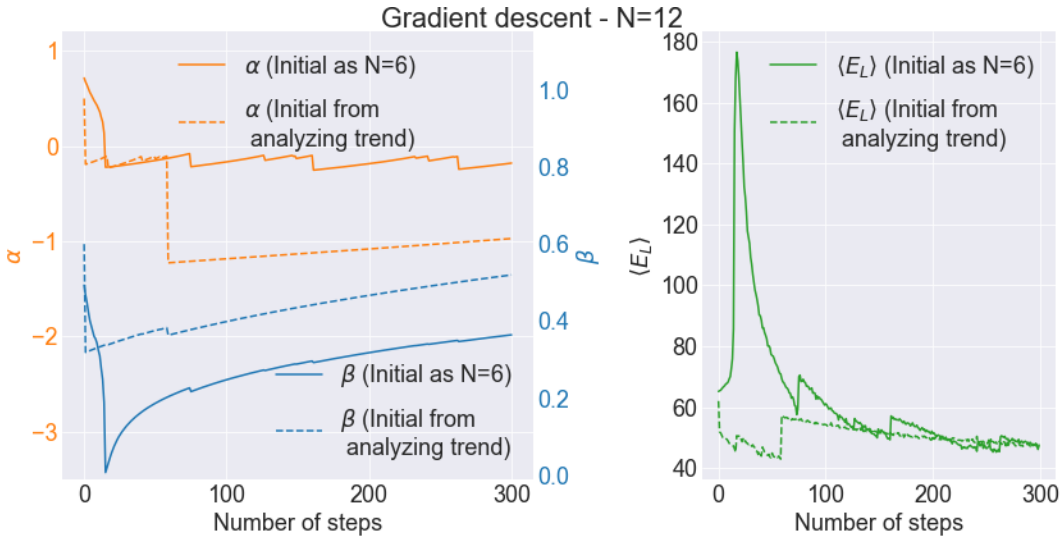

Figure 12: Two examples of the gradient descent method for twelve particles with different initial guesses for the two parameters. One of them is a guess based on the parameters found for the six electron system and the other is a guess based on the difference and trend found when the two electron and the six electron parameters was compared. Left: The development of the parameters. Right: The corresponding expectation value for the energy.

Continuing to the one-body density, Fig. 13 shows the one-body density for the twelve electron system, but only without interaction since a ground state for the interacting system was not found. Compared to the two smaller systems, this system shows another curve for the one-body density. Here, the maximum is once again at the origin of the trap, but there is also another local maximum approximately $1.5a_0$ away from the origin for $\omega = 1.0$. The scaling of the density seems to be similar to the former two systems where one see that the distribution broadens and the maximum is hence lowered, but the shape remains similar when ω is decreased.

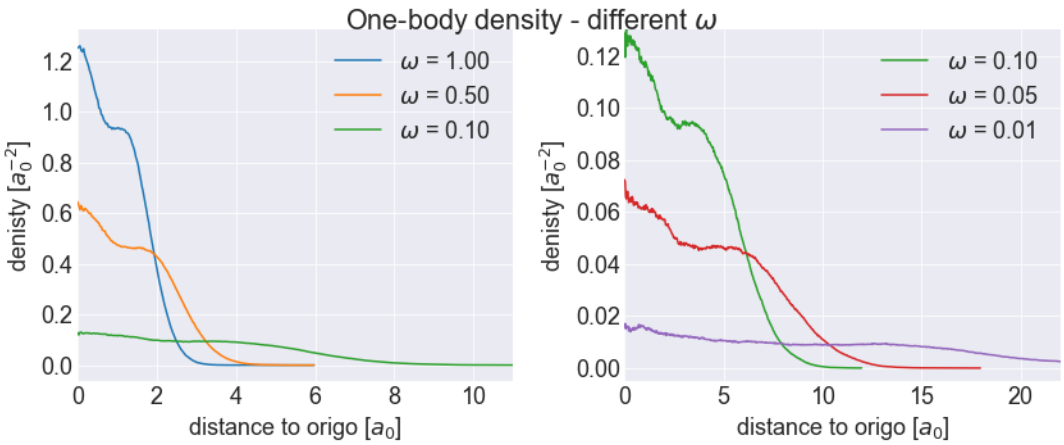

Figure 13: One-body density of the twelve electron system without the Jastrow factor, but with different trap frequencies. Left: The larger trap frequencies. Right: The smaller trap frequencies. The system with $\omega = 0.1$ is in both plots to show the difference in scale for the two plots. Used 2^{24} MC cycles.

3.4 Efficiency

In this project I wanted to write an efficient code, but I ran into some problems and after a while I figured out that it was best to use the not so efficient, but working code to get some results and use time to analyse them. Therefore, I have two branches in my repository at GitHub. One called

`no_optimization` and one master branch. The master branch includes my attempts to make the code more efficient as described in the theory part of this report, and the other branch is the code used to generate the results in this project. Both branches do include an efficient way of calculating the gradient and the double derivative though.

3.5 Performance analysis

At last, I include a small analysis of the performance of the code with and without vectorization and with and without parallelizing.

3.5.1 Vectorization

Table 10 lists the average CPU time of ten runs of the code with and without vectorization. The mean was found four times for both cases at different times during a working day. There is not a clear difference in CPU time, and one can also see from the four different attempts, that the CPU time changes based on the computer. I think one can conclude that the program is just as efficient with and without the vectorization. But this is just comparing the CPU time. As I have understood, vectorization can also help exploit the computer memory in a more efficient way.

Table 10: The average CPU time in seconds after 10 runs. The calculation is of the six electron system with interaction and involves importance sampling, saving and printing all local energies to file. Here I used 2^{22} MC cycles.

	CPU time	
Run	With vectorization	Without vectorization
1	145.930	148.531
2	149.653	142.881
3	155.731	153.519
4	146.769	146.834

3.5.2 Parallellization

The CPU times for the different threads for five different runs are listed in Tab. 11. The maximum CPU time of the four different threads is counted as the total CPU time. From the different runs, one can observe that the CPU time can vary based on the laptop and other processes on it, and it shows that it is important to time several runs to get a proper comparison of the CPU time of different methods.

Table 11: Four threads 2^{22} MC cycles on each.

Thread	1	2	3	4	Total:
CPU time	206.520	207.370	209.260	210.400	210.400
	228.270	229.930	229.930	230.110	230.110
	231.880	234.690	235.030	235.730	235.730
	231.330	234.060	234.170	234.300	234.300
	230.200	231.280	231.770	232.070	232.070
				Mean:	228.522

Table 12 compare the CPU time of parallelized runs and the not parallelized runs. One can see clearly that the parallelized runs are faster, but even though the parellelized run is done on four threads it is only between two and three times faster. One reason might be that the number of MC cycles used to reach an equilibrium is the same ($1 \cdot 10^5$) for both cases. For the single thread case it is done only once and afterwards all the local energies are sampled, but for the four thread case this is done independently on all four threads.

Table 12: Not parallelized: 2^{24} MC cycles and one thread. Parallelized: 2^{22} MC cycles and four threads. The CPU times are given in seconds and as an average over five runs.

	Not parallelized	Parallelized	Times faster
CPU time	656.210	228.522	2.87
	602.816	257.922	2.34

At last I compared the expectation value, the SEM and σ_B for one of the runs with and without parallelizing. The results are listed in Tab. 13 and shows a difference in the expectation value of the energy at approximately ± 0.01 which is larger than the calculated σ_B. One would expect a difference because of the different seeds in the random number generator, but maybe not that large. This might improve with a larger number of MC cycles though.

Table 13: A comparison of the expectation value, the SEM and σ_B for one of the runs with and without parallelizing. Units are as described in Tab. 4.

	$\langle E_L \rangle$	SEM	σ_B
Parallelized	20.4601	0.00016	0.00617
Not parallelized	20.4572	0.00016	0.00629

4 Conclusion

In this project I have used VMC methods to investigate quantum dots. I have looked at two dimensional full-shell systems of two and six electrons both with and without interaction and seen how the energy and one-body densities change when the trap frequency is changed. The energy of the system decrease with the decreasing trap frequency and the broader harmonic oscillator potential broadened the

one-body density as the trap frequency was decreased. I also looked at the full-shell system of twelve electrons without interaction and compared this system with the two smaller ones. The one-body density showed how the density of electrons changed when more shells where occupied by electrons.

Some aspects of the project that could have been done differently was to use a more robust optimization technique so that the twelve particle system could have been investigated with interaction as well. I also could have used a supercomputer to improve the errors through increasing the number of MC cycles. The code I used to generate the results could have been more efficient, I understood how to implement the changes to make it more efficient and I learned a lot with the attempt even though I did not have time to make the code work properly. That is why I chose to still include the different expressions in the Appendix of this report. Another thing that would make the program more elegant and probably more efficient is to change to the Eigen library for all vectors and matrices in the code instead of using many flops to convert the matrices to the Eigen matrix when I calculate the inverse and the determinant.

Appendices

A Improving calculations of important parts of code

A reduction of flops can be made to improve the effectiveness of the code. The expressions below are from Ref. [5].

A.1 Slater determinant

The Slater determinant contains the single-particle wave function of the number of particles included in the system evaluated at the posistion for all particles included because electrons are indistinguishable. The determinant is written as

$$D = Det\left(\phi_1(\boldsymbol{r}_1), \phi_2(\boldsymbol{r}_2), \ldots, \phi_N(\boldsymbol{r}_N)\right) = \begin{vmatrix} \phi_1(\boldsymbol{r}_1) & \phi_2(\boldsymbol{r}_1) & \cdots & \phi_N(\boldsymbol{r}_1) \\ \phi_1(\boldsymbol{r}_2) & \phi_2(\boldsymbol{r}_2) & \cdots & \phi_N(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{r}_N) & \phi_2(\boldsymbol{r}_N) & \cdots & \phi_N(\boldsymbol{r}_N). \end{vmatrix}$$

Hence the rows represent different positions, r_i, and the coloumns represent different states. To simplify the calculations one can put all states with spin up in one determinant and all states with spin down in another determinant. For six particles one then gets

$$D = D_\uparrow D_\downarrow = \begin{vmatrix} \phi_1(\boldsymbol{r}_1) & \phi_3(\boldsymbol{r}_1) & \phi_5(\boldsymbol{r}_1) \\ \phi_1(\boldsymbol{r}_2) & \phi_3(\boldsymbol{r}_2) & \phi_5(\boldsymbol{r}_2) \\ \phi_1(\boldsymbol{r}_3) & \phi_3(\boldsymbol{r}_3) & \phi_5(\boldsymbol{r}_3) \end{vmatrix} \begin{vmatrix} \phi_2(\boldsymbol{r}_4) & \phi_4(\boldsymbol{r}_4) & \phi_6(\boldsymbol{r}_4) \\ \phi_2(\boldsymbol{r}_5) & \phi_4(\boldsymbol{r}_5) & \phi_6(\boldsymbol{r}_5) \\ \phi_2(\boldsymbol{r}_6) & \phi_4(\boldsymbol{r}_6) & \phi_6(\boldsymbol{r}_6) \end{vmatrix}.$$

With this move, one looses the anti-symmetry with regards to spin, but the expectation value of the

energy remains the same.

The trial wavefunction can therefore be rewritten to

$$\psi_T = D_\uparrow D_\downarrow \psi_C$$

where ψ_C is the correlation part/interaction part of the trial wave function. Now only one of these matrises have to updated when a particle, depending on which spin the particle has.

A.2 The Metropolis ratio

In the metropolis test one calculates the ratio between the wave function before and after a proposed move, but now the wave function includes a determinant which is costly to calculate. I therefore want to utilize some relations from linear algebra to simplify the ratio and make the algorithm more efficent. The ratio between the Slater determinant part of the wavefunction, ψ_{SD}, is

$$R = \frac{\psi_{SD}(\boldsymbol{r}^{new})}{\psi_{SD}(\boldsymbol{r}^{old})} = \frac{\sum_i^N d_{ij}(\boldsymbol{r}^{new}) C_{ij}(\boldsymbol{r}^{new})}{\sum_i^N d_{ij}(\boldsymbol{r}^{old}) C_{ij}(\boldsymbol{r}^{old})}. \tag{14}$$

where $d_{ij} = \psi_i(j)$.

Here I have used the fact that when you calculate a determinant, you break it down into a sum of smaller determinants times a factor:

$$D = \begin{vmatrix} d_{11} & d_{12} & \cdots & d_{1N} \\ d_{21} & d_{22} & \cdots & d_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ d_{N1} & d_{N2} & \cdots & d_{NN} \end{vmatrix} = \sum_i^N d_{ij} C_{ij}.$$

So if $dij = d_{11}$ then

$$C_{11} = \begin{vmatrix} d_{22} & d_{23} & \cdots & d_{2N} \\ d_{32} & d_{33} & \cdots & d_{3N} \\ \vdots & \vdots & \ddots & \vdots \\ d_{N2} & d_{N3} & \cdots & d_{NN} \end{vmatrix}.$$

One can observe in Eq. 14 that if I move particle j from r_j^{old} to r_j^{new} the matrix C_{ij} is unchanged beacause the only change is with d_{ij} in the original determinant D that is not included in C_{ij}. Equation 14 is then

$$R = \frac{\sum_i^N d_{ij}(\boldsymbol{r}^{new})}{\sum_i^N d_{ij}(\boldsymbol{r}^{old})} \tag{15}$$

One can simplify this even further with the relation

$$\sum_{k=1}^{N} d_{ik} d_{kj}^{-1} = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \tag{16}$$

The ratio can be rewritten as

$$R = \frac{\sum_i^N d_{ij}(\boldsymbol{r}^{new}) d_{ij}(\boldsymbol{r}^{old})^{-1}}{\sum_i^N d_{ij}(\boldsymbol{r}^{old}) d_{ij}(\boldsymbol{r}^{old})^{-1}} = \sum_i^N d_{ij}(\boldsymbol{r}^{new}) d_{ij}(\boldsymbol{r}^{old})^{-1}. \tag{17}$$

The consequence of these calculations is that one only have to calculate the inverse values of the determinant once to know the values for $d_{ij}(\boldsymbol{r}^{old})^{-1}$ and then update only the row of the position that was changed in the Slater determinant and calculate the invers of the determinant again only if the move is accepted.

A.3 Updating the inverse of the Slater determinant

After a move is accepted in the Metropolis test, the row in the Slater determinant representing that particle is updated, but the inverse of the Slater determinant also needs to be updated because the Slater determinant has changed. This could be done by simply calculating the inverse of the determinant, but this is costly and there is a more efficient way. The elements of the determinant d_{kj}^{-1} can be found through

$$d_{kj}^{-1}(\boldsymbol{r}^{new}) = \begin{cases} d_{kj}^{-1}(\boldsymbol{r}^{old}) - \frac{d_{kj}^{-1}(\boldsymbol{r}^{old})}{R} \sum_{l=1}^{N} d_{il}^{-1}(\boldsymbol{r}^{new}) d_{lj}^{-1}(\boldsymbol{r}^{old}) & \text{if } i \neq j \\ \frac{d_{kj}^{-1}(\boldsymbol{r}^{old})}{R} \sum_{l=1}^{N} d_{il}^{-1}(\boldsymbol{r}^{old}) d_{lj}^{-1}(\boldsymbol{r}^{old}) & \text{if } i = j \end{cases},$$

where i is the number of the row representing the particle that was moved.

A.4 Gradients and double derivatives

The structure to calculate the gradient and the double derivative is the same as in project 1 [1]. However, the different parts have changed. The part that was the one-body part is different in this project and the gradient of the Slater determinant part replacing it is

$$\frac{\nabla_i |D(\boldsymbol{r})|}{|D(\boldsymbol{r})|} = \sum_{j=1}^{N} \nabla_i d_{ij}(\boldsymbol{r}) d_{ij}(\boldsymbol{r})^{-1} = \sum_{j=1}^{N} \nabla_i \phi_j(\boldsymbol{r}_i) d_{ij}(\boldsymbol{r})^{-1}. \tag{18}$$

When using importance sampling, the gradient is calculated to find the quantum force and the expression

$$\frac{\nabla_i |D(\boldsymbol{r^{new}})|}{|D(\boldsymbol{r^{new}})|} = \frac{1}{R} \sum_{j=1}^{N} \nabla_i \phi_j(\boldsymbol{r}_i^{\boldsymbol{new}}) d_{ij}(\boldsymbol{r^{old}})^{-1} \tag{19}$$

can then be used to calculate the gradient using the old inverse Slater determinant elements. This way one does not have to calculate the inverse Slater determinant if the move is not accepted.

If the move is accepted the double derivative, included in the calcualtion of the kinetic energy, is calculated using

$$\frac{\nabla_i^2|D(\boldsymbol{r})|}{|D(\boldsymbol{r})|} = \sum_{j=1}^{N} \nabla_i^2 d_{ij}(\boldsymbol{r}) d_{ij}(\boldsymbol{r})^{-1} = \sum_{j=1}^{N} \nabla_i \phi_j^2(\boldsymbol{r}_i) d_{ij}(\boldsymbol{r})^{-1}. \tag{20}$$

The Jastrow factor has also changed in this project compared to project 1 and if

$$e^{\frac{ar_{ij}}{1+\beta r_{ij}}} = e^{u(r_{ij})} \tag{21}$$

as in project 1. We can replace $u'(r_{ij})$ and $u''(r_{ij})$ in the expression of the double derivative of the trial wave function in project 1 with

$$u'(r_{ij}) = \frac{a}{(1+\beta r_{ij})^2} \tag{22}$$

and

$$u''(r_{ij}) = \frac{-2a\beta}{(1+\beta r_{ij})^3}. \tag{23}$$

B Energies

$$E_{n_x n_y} = \hbar\omega(n_x + n_y + \frac{d}{2}) \tag{24}$$

where d is the number of dimensions. In this project $d = 2$.

Table 14: The exact energies for the non-interacting case with different number of particles in a closed shell system.

Energies	
E_{00}	$\hbar\omega$
$E_{10} = E_{01}$	$2\hbar\omega$
$E_{20} = E_{02} = E_{11}$	$3\hbar\omega$
$E_{30} = E_{03} = E_{21} = E_{12}$	$4\hbar\omega$
$E_{N=2} = 2E_{00}$	$2\hbar\omega$
$E_{N=6} = E_{N=2} + 2E_{10} + 2E_{01}$	$10\hbar\omega$
$E_{N=12} = E_{N=6} + 2E_{20} + 2E_{02} + 2E_{11}$	$28\hbar\omega$
$E_{N=20} = E_{N=12} + 2E_{30} + 2E_{03} + 2E_{21} + 2E_{12}$	$60\hbar\omega$

C Hermite polynomials and the wavefunction derivatives

The relevant Hermite polynomials	
$H_0(\sqrt{\omega}x)$	1
$H_1(\sqrt{\omega}x)$	$2\sqrt{\omega}x$
$H_2(\sqrt{\omega}x)$	$4\omega x^2 - 2$
$H_3(\sqrt{\omega}x)$	$8\omega\sqrt{\omega}x^3 - 12\sqrt{\omega}x$

$$\phi_{n_x,n_y}(x,y) = AH_{n_x}(\sqrt{\omega}x)H_{n_y}(\sqrt{\omega}y)\exp\left(-\omega(x^2+y^2)/2.\right.$$

Table 15: The single-particle wavefunctions, $\psi_{n_x n_y}$, for this two dimensional harmonic ocsillator system.

Trial wavefunctions for the different states	
ψ_{00}	$A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{01}	$2\sqrt{\omega}xA\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{10}	$2\sqrt{\omega}yA\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{20}	$(4\omega x^2 - 2)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{02}	$(4\omega y^2 - 2)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{11}	$4\omega xyA\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{30}	$(8\omega\sqrt{\omega}x^3 - 12\sqrt{\omega}x)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{03}	$(8\omega\sqrt{\omega}y^3 - 12\sqrt{\omega}y)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{21}	$(8\omega\sqrt{\omega}x^2y - 4\sqrt{\omega}y)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
ψ_{12}	$(8\omega\sqrt{\omega}xy^2 - 4\sqrt{\omega}x)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$

Table 16: The derivatives of the single-particle wavefunctions, $\psi_{n_x n_y}$, for this two dimensional harmonic ocsillator system.

The derivative of the trial wavefunctions for the different states	
$\nabla\psi_{00}$	$(-\alpha\omega x, -\alpha\omega y)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{01}$	$-(\sqrt{\omega}(a\omega x^2-1), \alpha\omega^{3/2}xy)2A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{10}$	$-(\alpha\omega^{3/2}xy, \sqrt{\omega}(a\omega y^2-1))2A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{20}$	$-(2\alpha\omega^2x^3-\alpha\omega x-4\omega x, 2\alpha\omega^2x^2y-\alpha\omega y)2A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{02}$	$-(2\alpha\omega^2xy^2-\alpha\omega x, 2\alpha\omega^2y^3-\alpha\omega y-4\omega y)2A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{11}$	$(-4\omega y(\alpha\omega x^2-1), -4\omega x(\alpha\omega y^2-1))A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{30}$	$(-4\sqrt{\omega}(2\alpha\omega^2x^4-3(\alpha+2)\omega x^2+3), -4\alpha\omega^{3/2}xy(2\omega x^2-3))A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{03}$	$(-4\sqrt{\omega}(-4\alpha\omega^{3/2}xy(2\omega y^2-3), 2\alpha\omega^2y^4-3(\alpha+2)\omega y^2+3))A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{21}$	$(-4\sqrt{\omega}(\alpha\omega x^2(2\omega xy-1)-4\omega xy+1), -4\omega^{3/2}x(2x(\alpha\omega y^2-1)-\alpha y))A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla\psi_{12}$	$(-4\omega^{3/2}y(2y(\alpha\omega x^2-1)-\alpha x), -4\sqrt{\omega}(\alpha\omega y^2(2\omega xy-1)-4\omega xy+1))A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$

Table 17: The double derivatives of the single-particle wavefunctions, $\psi_{n_x n_y}$, for this two dimensional harmonic ocsillator system.

The double derivative of the trial wavefunctions for the different states	
$\nabla^2\psi_{00}$	$(\alpha^2\omega^2r^2-\alpha\omega)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{01}$	$2\alpha\omega^{3/2}x(\alpha\omega r^2-4)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{10}$	$2\alpha\omega^{3/2}y(\alpha\omega r^2-4)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{20}$	$2\omega(\alpha^2\omega(2\omega x^2-1)r^2+\alpha(2-12\omega x^2)+4))A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{02}$	$2\omega(\alpha^2\omega(2\omega y^2-1)r^2+\alpha(2-12\omega y^2)+4))A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{11}$	$4\alpha\omega^2xy(\alpha\omega r^2-6)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{30}$	$4\omega^{3/2}x(\alpha^2\omega(2\omega x^2-3)r^2-4\alpha(4\omega x^2-3)+12)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{03}$	$4\omega^{3/2}y(\alpha^2\omega(2\omega y^2-3)r^2-4\alpha(4\omega y^2-3)+12)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{21}$	$4\omega^{3/2}(\alpha^2\omega xr^2(2\omega xy-1)+4\alpha x(1-4\omega xy)+4y)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$
$\nabla^2\psi_{12}$	$4\omega^{3/2}(\alpha^2\omega yr^2(2\omega xy-1)+4\alpha y(1-4\omega xy)+4x)A\exp\left(\frac{-\alpha\omega r^2}{2}\right)$

References

[1] V. M. Reinertsen. Project 1, 2020.

[2] Wikipedia contributors. Gradient descent, Jun 2020.

[3] E. M. Nordhagen. Studies of Quantum Dots using Machine Learning. Master's thesis, University of Oslo, 2019.

[4] M. Taut. Two electrons in an external oscillator potential: Particular analytic solutions of a coulomb correlation problem. *Physical Review A*, 48(5):3561, 1993.

[5] M. Hjorth-Jensen. Computational Physics. Lecture Notes Fall 2015., 2015.