FYS-4411: Computational Physics II
Project 3

Gullik Vetvik Killie
Håkon Sebastian Bakke Mørk
Jose Emilio Ruiz Navarro

June 15, 2015

Abstract

In this work, a simple Variational Monte Carlo (VMC) method has been used to calculate the values of the energies of the ground states of three atoms: Helium, Beryllium and Neon, plus Hydrogen molecule and Beryllium Molecule. The program uses importance sampling to improve efficiency and make the results more precise. To further improve the efficiency, MPI and Gaussian Type Orbitals has been implemented as well. We provide a statistical analysis by the means of blocking so as to not understimate the error of our results. The one-body and charge densities were obtained to compare the effects of the Jastrow factor and provide insight into the electronic structure of the atoms.

1 Introduction

VMC methods pose a very attractive alternative to other more complex ways of finding the ground state energies of simple atoms and molecules, like configuration-interaction calculations. The price to be paid in exchange for this simplicity is the sensitivity to the trial wave functions that are used, a VMC algorithm is very sensitive to how these are constructed, so they are one of the most important aspects to be considered (in this work, given the simple nature of the atoms which we will be working with, it's not so important to worry about the quality of the trial wave functions because very simple and basic ones are more than enough to reproduce the actual results). It shouldn't be forgotten that it is a variational method, and this implies that finding the optimal set of variational parameters is going to be the most important part of the calculation itself because it would create a lot of problems if the search range for the parameters was illy defined and not close enough to the variational minimum, namely, the results would have a poor quality in this case. This means that the parameters need to be chosen very carefully, or a recursive search with decreasingly coarse spacing in the space of variational parameters is required if there is no deep knowledge about the system in question.

Instead of evaluating a very complex multidimensional integral to compute the expectation value of an operator, like the hamiltonian in this case, a VMC calculation exploits the fact that the majority of the configuration space where the wave function belongs can be regarded as much less important than other parts, the values of the wave function are too small there and can be mostly ignored during the integration of the algorithm. To capitalize this, the Metropolis algorithm is added to the VMC method, as well as importance sampling and Gaussian Type Orbitals.

2 Theory

2.1 Monte Carlo method with simple Metropolis sampling

In a quantum mechanical system the energy is given by the expectation value of the Hamiltonian, let Ψ_T be a proposal for a wavefunction that can describe the system.

$$E[\hat{H}] = \langle \Psi_T | \hat{H} | \Psi_T \rangle = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \hat{H} \Psi_T(\mathbf{R})}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})} \tag{1}$$

Let us introduce a local energy:

$$E_L(\hat{H}) = \frac{1}{\Psi_T(\mathbf{R})} \hat{H} \Psi_T(\mathbf{R})) \tag{2}$$

$$E[\hat{H}] = \frac{\int d\mathbf{R}\Psi_T^*(\mathbf{R})\Psi_T(\mathbf{R})E_L(\mathbf{R}))}{\int d\mathbf{R}\Psi_T^*(\mathbf{R})\Psi_T(\mathbf{R})} \tag{3}$$

Since the denumerator is a scalar constant after integrating it we can put it inside the integral in the numerator

$$E[\hat{H}] = \int d\mathbf{R}\frac{\Psi_T^*(\mathbf{R})\Psi_T(\mathbf{R})}{\int d\mathbf{R}'\Psi_T^*(\mathbf{R}')\Psi_T(\mathbf{R}')}E_L(\mathbf{R}) \tag{4}$$

$$E[\hat{H}] = \int d\mathbf{R}P(\mathbf{R})E_L(\mathbf{R}) \tag{5}$$

This probability function with $P(\mathbf{R})$ as the pdf, and we can use Monte Carlo integration to solve the integral. The algorithm for a Monte Carlo integration is given below.

1. Initialise system. Give particles a random position and decide how many Monte Carlo Cycles to run.
2. Start Monte Carlo Calculations
   (a) Propose a move of the particles according to an algorithm, for example $\mathbf{R_{new}} = \mathbf{R_{old}} + \delta * r$, where r is a random number in $[0, 1]$
   (b) Accept or reject move according to $P(\mathbf{R_{new}})/P(\mathbf{R_{old}}) \geq r$, where r is a new number. Update position values if accepted.
   (c) Calculate energy for this cycle.

See the git reposity in the reference list for a implementation of the Monte Carlo algorithm.

2.2 Atoms and Molecules

The dimensionless hamiltonian for the case of electrons around a nucleus is given by

$$\hat{H} = \sum_{i=1}^{N} -\frac{\nabla_i^2}{2} - \frac{Z}{r_i} + \sum_{i<j}\frac{1}{r_{ij}} \tag{6}$$

where r_i is the distance from electron i to the nucleus, Z is the nucleus charge, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. The kinetic energy for electron i is represented by $-\frac{\nabla_i^2}{2}$, $-\frac{Z}{r_i}$ the potential energy with respect to the nucleus and $\frac{1}{r_{ij}}$ the repulsive energy between the electrons i and j.

In the VMC calculation the local energy, E_L is a useful quantity and needs to be finite at all points to be normalizable. So by looking at the limits where the E_L diverges we can guess the form the wavefunctions should follow.

$$E_L(r_i, r_{ij}) = \frac{1}{\Psi_T}\hat{H}\Psi_T \tag{7}$$

In the cases where $r_i \to 0$ or $r_{ij} \to 0$ we need to make sure that the local energy does not diverge.

$$\lim_{r_i\to 0} E_L(r_i, r_{ij}) = \frac{1}{R_i(r_i)}\left(-\frac{1}{2}\frac{\partial^2}{\partial x_k^2} - \frac{Z}{r_i}\right)R_i(r_i) + G(r_i, r_{ij}) \tag{8}$$

$$\lim_{r_i\to 0} E_L(r_i, r_{ij}) = \frac{1}{R_i(r_i)}\left(-\frac{1}{2}\frac{\partial^2}{\partial r_k^2} - \frac{1}{r_i}\frac{\partial}{\partial r_i} - \frac{Z}{r_i}\right)R_i(r_i) + G(r_i, r_{ij}) \tag{9}$$

Given a well behaved wavefunction $\frac{1}{2}\frac{\partial^2}{\partial r_k^2}$ is finite.

$$\lim_{r_i\to 0} E_L(r_i, r_{ij}) = \frac{1}{R_i(r_i)}\left(-\frac{1}{r_i}\frac{\partial}{\partial r_i} - \frac{Z}{r_i}\right)R_i(r_i) \tag{10}$$

This is finite given when the following differential equation is fulfilled.

$$\frac{1}{R_i(r_i)}\frac{\partial}{\partial r_i}R_i(r_i) = -Z \qquad \text{With solution} \qquad R_i(r_i) = Ae^{-Z} \tag{11}$$

A similar calculation applies for $r_{12} \to 0$ and a trialfunction of the form

$$\Psi_T(r_i, r_j, r_{ij}) = e^{-\alpha\sum_N r_i}\prod_{i<j}^{N} e^{\beta r_{ij}} = e^{-\alpha\sum_N r_i}\prod_{i<j}^{N} e^{\frac{ar_{ij}}{1+\beta r_{ij}}}$$

should fulfill the condition that the local energy is finite.

	ψ_i	$\nabla\psi_i$	$\nabla^2\psi_i$
ψ_{1S}	$e^{-\alpha ri}$	$-\frac{\alpha}{r_i}\mathbf{r_i}e^{-\alpha r_i}$	$\frac{\alpha}{r_i}\left(\alpha r_i - 2\right)e^{-\alpha r_i}$
ψ_{2S}	$\left(-\frac{\alpha r_i}{2}+1\right)e^{-\frac{\alpha r_i}{2}}$	$\frac{\alpha}{4r_i}\left(\alpha r_i - 4\right)\mathbf{r_i}e^{-\frac{\alpha r_i}{2}}$	$-\frac{\alpha}{8r_i}\left(\alpha^2 r_i^2 - 10\alpha r_i + 16\right)e^{-\frac{\alpha r_i}{2}}$
ψ_{2Px}	$\alpha x_i e^{-\frac{\alpha r_i}{2}}$	$-\frac{\alpha x_i}{2r_i}\left(\alpha\mathbf{r_i} - 2r_i\hat{\mathbf{i}}\right)e^{-\frac{\alpha r_i}{2}}$	$\frac{\alpha^2 x_i}{4r_i}\left(\alpha r_i - 8\right)e^{-\frac{\alpha r_i}{2}}$
ψ_{2Py}	$\alpha x_i e^{-\frac{\alpha r_i}{2}}$	$-\frac{\alpha y_i}{2r_i}\left(\alpha\mathbf{r_i} - 2r_i\hat{\mathbf{i}}\right)e^{-\frac{\alpha r_i}{2}}$	$\frac{\alpha^2 x_i}{4r_i}\left(\alpha r_i - 8\right)e^{-\frac{\alpha r_i}{2}}$
ψ_{2Pz}	$\alpha x_i e^{-\frac{\alpha r_i}{2}}$	$-\frac{\alpha z_i}{2r_i}\left(\alpha\mathbf{r_i} - 2r_i\hat{\mathbf{i}}\right)e^{-\frac{\alpha r_i}{2}}$	$\frac{\alpha^2 x_i}{4r_i}\left(\alpha r_i - 8\right)e^{-\frac{\alpha r_i}{2}}$

Table 1: The different hydrogenic wavefunctions along with the gradients and laplacians. The derivatives is computed in the program derivatives.py G. V. Killie and Navarro (n.d.)

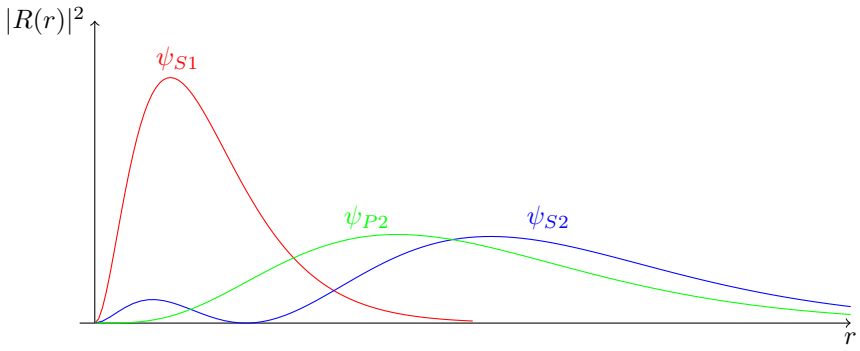

*Figure 1: The radial probability distribution of the first three orbitals of a hydrogen atom. The radial distribution functions is taken from (*Introduction to Quantum Mechanics by Pauling, Linus; Wilson, E.Bright *1935)*

2.2.1 Hydrogenic wavefunctions

A hydrogen atom is analytically solvable and we have exact wavefunctions corresponding to the electron being in the different shells. When we are building atoms containing more electrons turning it into a many body problem we base our guess at the trialfunction on the solutions to the hydrogen atom. So we need the wavefunctions for the five lowest states to calculate up to the 10 electron big neon atom. The hydrogenic wavefunctions along with their gradients and laplacians is contained in Table 1. The radial distribution of the first three hydrogen orbitals, which our trialfunctions is based on, is depicted in Fig. 1.

2.2.2 Helium atom

We wanted to do our Variational Monte Carlo, VMC, computations on helium first since it was a simpler system that the larger atoms and molecules. The hamiltonian for the helium atom is given by equation (6) and a trialfunction that fulfills the cusp conditions discussed earlier is

$$\Psi_{T2}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_{12}) = e^{-\alpha(r_1+r_2)}e^{\frac{r_{12}}{2(1+\beta r_{12})}} \tag{12}$$

We have also done calculations for a simplified version of the trialfunction by taking out electron-electron interaction from the trialfunction so we end up with the trialfunction

$$\Psi_{T1}(\mathbf{r}_1,\mathbf{r}_2) = e^{-\alpha(r_1+r_2)} \tag{13}$$

We calculated the local energies for these two trialfunctions and they are for Ψ_{T1} and Ψ_{T2}.

$$E_{L1} = (\alpha - Z)\left(\frac{1}{r_2}+\frac{1}{r_2}\right)+\frac{1}{r_{12}}-\alpha^2 \tag{14}$$

$$E_{L2} = E_{L1}+\frac{1}{2\left(1+\beta r_{12}\right)}\left[\frac{\alpha(r_1+r_2)}{r_{12}}\left(1-\frac{\mathbf{r_1}\mathbf{r_2}}{r_1 r_2}\right)-\frac{1}{2\left(1+\beta r_{12}\right)}-\frac{2}{r_{12}}+\frac{2\beta}{\left(1+\beta r_{12}\right)}\right] \tag{15}$$

2.2.3 Beryllium atom

It is fairly simple to extend the machinery of Variational Monte Carlo to other systems than the Helium atom. To show this we want to perform calculations on the beryllium atom. As beryllium has four electrons compared to the 2 of helium, we need to construct a Slater determinant from the hydrogenic wavefunctions. Sticking to hydrogen-like wave functions, we can write the trial wave function for beryllium as

$$\psi_T(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}, \mathbf{r_4}) = Det\left(\phi_1(\mathbf{r_1}), \phi_2(\mathbf{r_2}), \phi_3(\mathbf{r_3}), \phi_4(\mathbf{r_4})\right) \prod_{i<j}^{4} \exp\left\{\left(\frac{ar_{ij}}{(1+\beta r_{ij})}\right)\right\}, \tag{16}$$

where *Det* is a Slater determinant and the single-particle wave functions are the hydrogen wave functions for the 1s and 2s orbitals. With the variational ansatz these are

$$\phi_{1s}(\mathbf{r_i}) = e^{-\alpha r_i}, \tag{17}$$

and

$$\phi_{2s}(\mathbf{r_i}) = (1 - \alpha r_i/2)\, e^{-\alpha r_i/2}. \tag{18}$$

The Slater determinant is calculated using these ansatzes, and can for Beryllium be written out as

$$|D| \propto \left[\psi_{1s}(\mathbf{r_1})\psi_{2s}(\mathbf{r_2}) - \psi_{1s}(\mathbf{r_2})\psi_{2s}(\mathbf{r_1})\right]\left[\psi_{1s}(\mathbf{r_3})\psi_{2s}(\mathbf{r_4}) - \psi_{1s}(\mathbf{r_4})\psi_{2s}(\mathbf{r_3})\right] \tag{19}$$

Furthermore, for the Jastrow factor,

$$\Psi_C = \prod_{i<j} g(r_{ij}) = \exp\left\{\sum_{i<j} \frac{ar_{ij}}{1+\beta r_{ij}}\right\}, \tag{20}$$

we need to take into account the spins of the electrons. We fix electrons 1 and 2 to have spin up, and electron 3 and 4 to have spin down. This means that when the electrons have equal spins we get a factor

$$a = \frac{1}{4}, \tag{21}$$

and if they have opposite spins we get a factor

$$a = \frac{1}{2}. \tag{22}$$

2.2.4 Neon atom

Wishing to extend our variational Monte Carlo machinery further we implement Neon. Neon has ten electrons, so it is a big jump from Helium and Beryllium. Now we need better methods to handle the Slater determinant and is described in subsection 2.3. The trial wave function for Neon can be written as

$$\psi_T(\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_{10}}) = Det\left(\phi_1(\mathbf{r_1}), \phi_2(\mathbf{r_2}), \ldots, \phi_{10}(\mathbf{r_{10}})\right) \prod_{i<j}^{10} \exp\left\{\left(\frac{r_{ij}}{2(1+\beta r_{ij})}\right)\right\}, \tag{23}$$

Now we need to include the $2p$ wave function as well. It is given as

$$\phi_{2p}(\mathbf{r_i}) = \alpha \mathbf{r_i} e^{-\alpha r_i/2}. \tag{24}$$

where $r_i = \sqrt{r_{i_x}^2 + r_{i_y}^2 + r_{i_z}^2}$.

2.2.5 Hydrogen molecule

The VMC machinery can also handle simple molecules with some modifications considering that we now have two nuclei. So we need a slightly different Hamiltonian, where we we need include a few more terms in the potential energy of the system. If we let $\mathbf{R}$ be the vector between the nuclei we can write the positions, $\mathbf{r_{ip1}}$, of electron i in relation to nucleus 1.

$$\mathbf{r_{ip1}} = \mathbf{r_i} + \frac{\mathbf{R}}{2} \tag{25}$$

$$\mathbf{r_{ip2}} = \mathbf{r_i} - \frac{\mathbf{R}}{2} \tag{26}$$

Then we add all the terms for the potential energy with the kinetic energy and get the Hamiltonian

$$\hat{H} = \sum_{i=1}^{2} -\frac{\nabla_i^2}{2} - \frac{Z_1}{r_{ip1}} - \frac{Z_2}{r_{ip2}} + \sum_{i<j} \frac{1}{r_{ij}} + \frac{Z_1 Z_2}{|\mathbf{R}|} \tag{27}$$

Here Z_1 and Z_2 is given by the charge of the respective nuclei. Assuming that the trialfunction for the Hydrogen molecule is similar to two hydrogen atoms we can base our guess of the trialfunction on a linear combination of the two hydrogenic 1S wavefunctions around each nuclei. Disregarding factors, since they disappear in the VMC computation and assuming symmetry about the nuclei we end up with

$$\Psi_T(\mathbf{r_1}, \mathbf{r_2}, \mathbf{R}) = \psi(\mathbf{r_1}, \mathbf{R})\psi(\mathbf{r_2}, \mathbf{R}) \exp\left\{ \frac{r_{12}}{2(1+\beta r_{12})} \right\} \tag{28}$$

Where the hydrogenic wavefunctions is given by

$$\psi(\mathbf{r_i}, \mathbf{R}) = [\exp\{-\alpha r_{ip1}\} \pm \exp\{-\alpha r_{ip2}\}] \tag{29}$$

Here we should add together the $1S$ wavefunctions, as subtracting corresponds to the electrons having the same spin. In a Hydrogen molecule the electrons will have differents spin if possible.

2.2.6 Beryllium Molecule

The Beryllium molecule consists of four electrons shared between two nuclei with a charge of $Z = 4$, we will use the same method to calculate it as in Hydrogen molecule and it also shares the same Hamiltonian. As in the description of the Beryllium atom we will need to construct the trialfunction out of an Slater Determinant consisting of linear combinations of the hydrogenic wavefunctions ψ_{1s} and ψ_{2s}, and a correlation term.

$$\Psi_T(\mathbf{r_i}, \mathbf{R}) = |D| \prod_{i<j}^{4} \exp\left\{ \left(\frac{a r_{ij}}{(1+\beta r_{ij})} \right) \right\} \tag{30}$$

where the Slater determinant is construced by the following wavefunctions

$$\psi_{1S1}(\mathbf{r_i}) = [\exp\{-\alpha r_{ip1}\} + \exp\{-\alpha r_{ip2}\}] \tag{31}$$

$$\psi_{1S2}(\mathbf{r_i}) = [\exp\{-\alpha r_{ip1}\} - \exp\{-\alpha r_{ip2}\}] \tag{32}$$

$$\psi_{2S1}(\mathbf{r_i}) = \left[(1 - \alpha r_{i1p}/2)\, e^{-\alpha r_{i1p}/2} + (1 - \alpha r_{i2p}/2)\, e^{-\alpha r_{i2p}/2} \right] \tag{33}$$

$$\psi_{2S2}(\mathbf{r_i}) = \left[(1 - \alpha r_{i1p}/2)\, e^{-\alpha r_{i1p}/2} - (1 - \alpha r_{i2p}/2)\, e^{-\alpha r_{i2p}/2} \right] \tag{34}$$

Here we need to use both subtracting and addition in the combinations in the construction of the wavefunctions as each Beryllium will have two electrons in each shell so there will need to be two of the same spin in each of the shells.

2.3 Calculating the Slater determinant

To describe the wavefunction of multiple fermions we use a Slater determinant. The Slater determinant has the form

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \alpha, \beta, \gamma, \delta) = \frac{1}{\sqrt{4!}} \begin{vmatrix} \psi_{100\uparrow}(\mathbf{r}_1) & \psi_{100\uparrow}(\mathbf{r}_2) & \psi_{100\uparrow}(\mathbf{r}_3) & \psi_{100\uparrow}(\mathbf{r}_4) \\ \psi_{100\downarrow}(\mathbf{r}_1) & \psi_{100\downarrow}(\mathbf{r}_2) & \psi_{100\downarrow}(\mathbf{r}_3) & \psi_{100\downarrow}(\mathbf{r}_4) \\ \psi_{200\uparrow}(\mathbf{r}_1) & \psi_{200\uparrow}(\mathbf{r}_2) & \psi_{200\uparrow}(\mathbf{r}_3) & \psi_{200\uparrow}(\mathbf{r}_4) \\ \psi_{200\downarrow}(\mathbf{r}_1) & \psi_{200\downarrow}(\mathbf{r}_2) & \psi_{200\downarrow}(\mathbf{r}_3) & \psi_{200\downarrow}(\mathbf{r}_4) \end{vmatrix} \tag{35}$$

for a four-fermionic system. Because the spatial wave functions for spin up and spin down states are equal, this Slater determinant equals zero. We can rewrite the Slater determinant as a product of two Slater determinants, one for spin up and one for spin down. This gives us

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

Here we have defined the Slater determinant for spin up as

$$\det\uparrow(1,2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{100\uparrow}(\mathbf{r}_1) & \psi_{100\uparrow}(\mathbf{r}_2) \\ \psi_{200\uparrow}(\mathbf{r}_1) & \psi_{200\uparrow}(\mathbf{r}_2) \end{vmatrix} \tag{36}$$

and the Slater determinant for spin down as

$$\det\downarrow(3,4) = \frac{1}{\sqrt{2}}\begin{vmatrix} \psi_{100\downarrow}(\mathbf{r}_3) & \psi_{100\downarrow}(\mathbf{r}_4) \\ \psi_{200\downarrow}(\mathbf{r}_3) & \psi_{200\downarrow}(\mathbf{r}_4) \end{vmatrix} \tag{37}$$

And the total determinant is of course still zero.

Further, it can be shown that for the variational energy we can approximate the Slater determinant as

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N) \propto \det\uparrow \det\downarrow \tag{38}$$

We now have the Slater determinant as a product of two determinants, one containing the electrons with only spin up, and one containing the electrons of spin down. This approach has certain limits as the ansatz isn't antisymmetric under the exchange of electrons with opposite spins, but it gives the same expectation value for the energy as the full Slater determinant as long as the Hamiltonian is spin independent. We thus avoid summing over spin variables.

Now we have the Slater determinant written as a product of a determinant for spin up and a determinant for spin down. The next step is to invert the matrices using LU decomposition. We can thus rewrite a matrix $\hat{A}$ as a product of two matrices, $\hat{B}$ and $\hat{C}$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ b_{21} & 1 & 0 & 0 \\ b_{31} & b_{32} & 1 & 0 \\ b_{41} & b_{42} & b_{43} & 1 \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ 0 & c_{22} & c_{23} & c_{24} \\ 0 & 0 & c_{33} & c_{34} \\ 0 & 0 & 0 & c_{44} \end{pmatrix}$$

LU factorization exists for $\hat{A}$ if the determinant is nonzero. If $\hat{A}$ also is non-singular, then the LU factorization is unique and the determinant is given by

$$|\hat{A}| = c_{11}c_{22}\dots c_{nn} \tag{39}$$

Using this we can calculate the spin up determinant, the spin down determinant, and by putting them together, the Slater determinant.

2.4 Optimization

Since the raw computational cost of the VMC computation scales up very fast with the increase of particles considered we have implemented several different methods to achieve a faster speed allowing us to compute longer and with more precision. The optimization is done along three paths importance sampling, multithreaded computation and improvements to computing the ratios and local energy. Importance sampling improves the convergence speed of the algorithm, multithreaded computation allows several processors to compute at the same time and the algorithmic improvements to reduce computing saving time in computing the Slater determinants and removing the need for slow numerical derivations. Here we will go through the algorithmic optimizations first.

2.4.1 Metropolis-Hastings-Ratio

In the Metropolis-Hasting algorithm a ratio of the new probability and the old probability is calculated to evaluate if a suggested move is accepted or rejected.

$$\frac{P(\mathbf{r_{new}})}{P(\mathbf{r_{old}}} \geq r \tag{40}$$

For our probability function it can be written as

$$\frac{|D_\uparrow^{new}|}{|D_\uparrow^{old}|}\frac{|D_\downarrow^{new}|}{|D_\downarrow^{old}|}\frac{\Psi_C^{new}}{\Psi_C^{old}} \geq r \tag{41}$$

2.4.2 Slater-Determinant-Ratio

The Slater-Determinants ratios are slow to calculate by brute force, for example by LU-decomposition, but recalculation it after each move can be done in a computationally easier way.

To tackle the determinant ratios we need to introduce some notation. Let an element in the determinant matrix, $|D|$, be described by

$$D_{ij} = \phi_j(\mathbf{r}_i) \tag{42}$$

where ϕ_j is the j'th single particle wavefunction and $\mathbf{r}_i$ is the position of the i'th particle.

The inverse of a matrix is given by transposing it and dividing by the determinant, so the determinant can be written as

$$|D| = \frac{\mathbf{D}^T}{\mathbf{D^{-1}}} = \sum_{j=1}^{N} \frac{C_{ji}}{D_{ij}^{-1}} = \sum_{j=1}^{N} D_{ij} C_{ji} \tag{43}$$

This gives the ratio of the new and old Slater determinants the following

$$R_{SD} = \frac{|\mathbf{D}^{new}|}{|\mathbf{D}^{old}|} = \frac{\sum_{j=0}^{N} D_{ij}^{new} C_{ji}^{new}}{\sum_{j=0}^{N} D_{ij}^{old} C_{ji}^{old}} \tag{44}$$

Since we are only moving one particle at a time and the cofactor term relies on the other rows it doesn't change, $C_{ij}^{new} = C_{ij}^{old}$ in one movement. Combining this with equation (43) we get

$$R_{SD} = \frac{\sum_{j=0}^{N} D_{ij}^{new} (D_{ji}^{old})^{-1} |D^{old}|}{\sum_{j=0}^{N} D_{ij}^{old} (D_{ji}^{old})^{-1} |D^{old}|} \tag{45}$$

Since $\mathbf{D}$ is invertible, $\mathbf{D}\mathbf{D}^{-1} = \mathbf{1}$, the ratio becomes

$$R_{SD} = \sum_{j=0}^{N} D_{ij}^{new} (D_{ji}^{old})^{-1} = \sum_{j=0}^{N} \phi_j(\mathbf{x}_i^{new}) D_{ji}^{-1}(\mathbf{x}^{old}) \tag{46}$$

2.4.3 Correlation-to-correlation ratio

We have $N(N-1)/2$ relative distances r_{ij}. We can write these in a matrix storage format, where they form a strictly upper triangular matrix

$$\mathbf{r} \equiv \begin{pmatrix} 0 & r_{1,2} & r_{1,3} & \dots & r_{1,N} \\ \vdots & 0 & r_{2,3} & \dots & r_{2,N} \\ \vdots & \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & r_{N-1,N} \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

This upper triangular matrix form also applies to $g = g(r_{ij})$.

The correlation-to-correlation ratio, or ratio between Jastrow factors is given by

$$R_C = \frac{\Psi_C^{new}}{\Psi_C^{cur}} = \prod_{i=1}^{k-1} \frac{g_{ik}^{new}}{g_{ik}^{cur}} \prod_{i=k+1}^{N} \frac{g_{ki}^{new}}{g_{ki}^{cur}} \tag{47}$$

or in the Padé-Jastrow form

$$R_C = \frac{\Psi_C^{\text{new}}}{\Psi_C^{\text{cur}}} = \frac{\exp(U_{new})}{\exp(U_{cur})} = \exp(\Delta U) \tag{48}$$

where

$$\Delta U = \sum_{i=1}^{k-1} \left(f_{ik}^{\text{new}} - f_{ik}^{\text{cur}}\right) + \sum_{i=k+1}^{N} \left(f_{ki}^{\text{new}} - f_{ki}^{\text{cur}}\right) \tag{49}$$

2.4.4 Efficient calculation of derivatives

Calculating the derivatives involved in the VMC calculation numerically is slow in that they entail several calls to the wavefunctions in addition to introducing an extra numerical error. Here we will show how to divide up the derivatives and how to find analytical expressions for all the parts using the derivatives found in table 1.

The trialfunction can be factorized as

$$\Psi_T(\mathbf{x}) = \Psi_D \Psi_C = |D_\uparrow||D_\downarrow|\Psi_C \tag{50}$$

where $D_\uparrow$, $D_\downarrow$ and Ψ_C is the spin up and down part of the Slater determinant and the Jastrow factor respectively.

2.4.5 Gradient-Ratio

For the quantum force, and in the final expression for the local energy, we need to calculate the gradient ratio of the trialfunction which is given by

$$\frac{\nabla\Psi_T}{\Psi_T} = \frac{\nabla(\Psi_D\Psi_C)}{\Psi_D\Psi_C} = \frac{\nabla\Psi_D}{\Psi_D} + \frac{\nabla\Psi_C}{\Psi_C} \tag{51}$$

$$= \frac{\nabla|D_\uparrow|}{|D_\uparrow|} + \frac{\nabla|D_\downarrow|}{|D_\downarrow|} + \frac{\nabla\Psi_C}{\Psi_C} \tag{52}$$

2.4.6 Laplacian-Ratio

From the Hamiltonians and the expression for the local energy the local kinetic energy of electron i is given by the following

$$K_i = -\frac{1}{2}\frac{\nabla_i^2\Psi}{\Psi} \tag{53}$$

Using the factorization of the trialfunction from (50) we can calculate the ratio needed for the kinetic energy.

$$\frac{1}{\Psi_T}\frac{\partial^2\Psi_T}{\partial x_k^2} = \frac{1}{\Psi_D\Psi_C}\frac{\partial^2(\Psi_D\Psi_C)}{\partial x_k^2} = \frac{1}{\Psi_D\Psi_C}\frac{\partial}{\partial x_k}\left(\frac{\partial\Psi_D}{\partial x_k}\Psi_C + \Psi_D\frac{\partial\Psi_C}{\partial x_k}\right) \tag{54}$$

$$= \frac{1}{\Psi_D\Psi_C}\left(\frac{\partial^2\Psi_D}{\partial x_k^2}\Psi_C + 2\frac{\partial\Psi_D}{\partial x_k}\frac{\partial\Psi_C}{\partial x_k} + \Psi_D\frac{\partial^2\Psi_C}{\partial x_k^2}\right) \tag{55}$$

$$= \frac{1}{\Psi_D}\frac{\partial^2\Psi_D}{\partial x_k^2} + 2\frac{1}{\Psi_D}\frac{\partial\Psi_D}{\partial x_k}\cdot\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} + \frac{1}{\Psi_C}\frac{\partial^2\Psi_C}{\partial x_k^2} \tag{56}$$

Since the Slater determinant part of the trialfunction is separable into a spin up and down part we can simplify it further.

$$\frac{1}{\Psi_D}\frac{\partial^2\Psi_D}{\partial x_k^2} = \frac{1}{|D_\uparrow||D_\downarrow|}\frac{\partial^2|D_\uparrow||D_\downarrow|}{\partial x_k^2} = \frac{1}{|D_\uparrow|}\frac{\partial^2|D_\uparrow|}{\partial x_k^2} + \frac{1}{|D_\downarrow|}\frac{\partial^2|D_\downarrow|}{\partial x_k^2} \tag{57}$$

$$\frac{1}{\Psi_D}\frac{\partial\Psi_D}{\partial x_k} = \frac{1}{|D_\uparrow||D_\downarrow|}\frac{\partial|D_\uparrow||D_\downarrow|}{\partial x_k} = \frac{1}{|D_\uparrow|}\frac{\partial|D_\uparrow|}{\partial x_k} + \frac{1}{|D_\downarrow|}\frac{\partial|D_\downarrow|}{\partial x_k} \tag{58}$$

Inserting equations (58) and (57) into (56) we get

$$\frac{\nabla^2\Psi_T}{\Psi_T} = \frac{\nabla^2|D_\uparrow|}{|D_\uparrow|} + \frac{\nabla^2|D_\downarrow|}{|D_\downarrow|} + 2\left(\frac{\nabla|D_\uparrow|}{|D_\uparrow|} + \frac{\nabla|D_\downarrow|}{|D_\downarrow|}\right)\cdot\frac{\nabla\Psi_C}{\Psi_C} + \frac{\nabla^2\Psi_C}{\Psi_C} \tag{59}$$

So to calculate the laplacian-ratio and the gradient-ratio we need to find expressions for $\frac{\nabla|D|}{|D|}$, $\frac{\nabla^2|D|}{|D|}$, $\frac{\nabla\Psi_C}{\Psi_C}$ and $\frac{\nabla^2\Psi_C}{\Psi_C}$.

2.4.7 The $\nabla|D|/|D|$ ratio

By the same argument as in section 2.4.2 the the gradient-slater-ratio can be written as

$$\frac{\nabla|D|}{|D|} = \sum_{j=0}^{N}\nabla(D_{ij})D_{ji}^{-1} = \sum_{j=0}^{N}\nabla\phi_j(\mathbf{x}_i)D_{ji}^{-1}(\mathbf{x}) \tag{60}$$

2.4.8 The $\nabla^2|D|/|D|$ ratio

As in the previous section the laplacian can be calculated by a similar method to the one used in section 2.4.2, and we end up with

$$\frac{\nabla^2|D|}{|D|} = \sum_{j=0}^{N}\nabla^2(D_{ij})D_{ji}^{-1} = \sum_{j=0}^{N}\nabla^2\phi_j(\mathbf{x}_i)D_{ji}^{-1}(\mathbf{x}) \tag{61}$$

2.4.9 The $\nabla\Psi_C/\Psi_C$ ratio

We continue by finding a useful expression for the quantum force and kinetic energy, the ratio $\nabla\Psi_C/\Psi_C$. It has, for all dimensions, the form

$$\frac{\nabla_i\Psi_C}{\Psi_C} = \frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_i} \tag{62}$$

where i runs over all particles. Since the terms of the trialfunction that aren't aren't differentiated cancel with their corresponding terms in the denominator, so only $N-1$ terms survive the first derivative. We get

$$\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} = \sum_{i=1}^{k-1}\frac{1}{g_{ik}}\frac{\partial g_{ik}}{\partial x_k} + \sum_{i=k+1}^{N}\frac{1}{g_{ki}}\frac{\partial g_{ki}}{\partial x_k} \tag{63}$$

For the exponential form we get almost the same, by just replacing g_{ij} with $\exp(f_{ij})$ and we get

$$\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} = \sum_{i=1}^{k-1}\frac{\partial g_{ik}}{\partial x_k} + \sum_{i=k+1}^{N}\frac{\partial g_{ki}}{\partial x_k} \tag{64}$$

We now use the identity

$$\frac{\partial}{\partial x_i}g_{ij} = -\frac{\partial}{\partial x_j}g_{ij} \tag{65}$$

and get expressions where the derivatives that act on the particle are represented by the second index of g

$$\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} = \sum_{i=1}^{k-1}\frac{1}{g_{ik}}\frac{\partial g_{ik}}{\partial x_k} - \sum_{i=k+1}^{N}\frac{1}{g_{ki}}\frac{\partial g_{ki}}{\partial x_i} \tag{66}$$

and for the exponential case

$$\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} = \sum_{i=1}^{k-1}\frac{\partial g_{ik}}{\partial x_k} - \sum_{i=k+1}^{N}\frac{\partial g_{ki}}{\partial x_i} \tag{67}$$

Since we have that the correlation function is depending on the relative distance we use the chain rule

$$\frac{\partial g_{ij}}{\partial x_j} = \frac{\partial g_{ij}}{\partial r_{ij}}\frac{\partial r_{ij}}{\partial x_j} = \frac{x_j - x_i}{r_{ij}}\frac{\partial g_{ij}}{\partial r_{ij}} \tag{68}$$

After substitution we get

$$\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} = \sum_{i=1}^{k-1}\frac{1}{g_{ik}}\frac{\mathbf{r_{ik}}}{r_{ik}}\frac{\partial g_{ik}}{\partial r_{ik}} - \sum_{i=k+1}^{N}\frac{1}{g_{ki}}\frac{\mathbf{r_{ki}}}{r_{ki}}\frac{\partial g_{ki}}{\partial r_{ki}} \tag{69}$$

For the Padé-Jastrow form we set $g_{ij} \equiv g(r_{ij}) = e^{f(r_{ij})} = e^{f_{ij}}$ and

$$\frac{\partial g_{ij}}{\partial r_{ij}} = g_{ij}\frac{\partial f_{ij}}{\partial r_{ij}} \tag{70}$$

and arrive at

$$\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} = \sum_{i=1}^{k-1}\frac{\mathbf{r_{ik}}}{r_{ik}}\frac{\partial f_{ik}}{\partial r_{ik}} - \sum_{i=k+1}^{N}\frac{\mathbf{r_{ki}}}{r_{ki}}\frac{\partial f_{ki}}{\partial r_{ki}} \tag{71}$$

where we have the relative vectorial distance

$$\mathbf{r}_{ij} = |\mathbf{r}_j - \mathbf{r}_i| = (x_j - x_i)\mathbf{e}_1 + (y_j - y_i)\mathbf{e}_2 + (z_j - z_i)\mathbf{e}_3 \tag{72}$$

With a linear Padé-Jastrow we set

$$f_{ij} = \frac{ar_{ij}}{(1+\beta r_{ij})} \tag{73}$$

with the corresponding closed form expression

$$\frac{\partial f_{ij}}{\partial r_{ij}} = \frac{a}{(1+\beta r_{ij})^2} \tag{74}$$

2.4.10 The $\nabla^2\Psi_C/\Psi_C$ ratio

For the kinetic energy we also need the second derivative of the Jastrow factor divided by the Jastrow factor. We start with this

$$\left[\frac{\nabla^2\Psi_C}{\Psi_C}\right]_x = 2\sum_{k=1}^{N}\sum_{i=1}^{k-1}\frac{\partial^2 g_{ik}}{\partial x_k^2} + \sum_{k=1}^{N}\left(\sum_{i=1}^{k-1}\frac{\partial g_{ik}}{\partial x_k} - \sum_{i=k+1}^{N}\frac{\partial g_{ki}}{\partial x_i}\right)^2 \tag{75}$$

But we have another, simpler form for the function

$$\Psi_C = \prod_{i<j}\exp f(r_{ij}) = \exp\left\{\sum_{i<j}\frac{ar_{ij}}{1+\beta r_{ij}}\right\} \tag{76}$$

and for particle k we have

$$\frac{\nabla_k^2\Psi_C}{\Psi_C} = \sum_{ij\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_i)(\mathbf{r}_k-\mathbf{r}_j)}{r_{ki}r_{kj}}f'(r_{ki})f'(r_{kj}) + \sum_{j\neq k}\left(f''(r_{kj}) + \frac{2}{r_{kj}}f'(r_{kj})\right) \tag{77}$$

We use

$$f(r_{ij}) = \frac{ar_{ij}}{1+\beta r_{ij}} \tag{78}$$

and with

$$g'(r_{kj}) = dg(r_{kj})/dr_{kj} \qquad \text{and} \qquad g''(r_{kj}) = d^2g(r_{kj})/dr_{kj}^2 \tag{79}$$

we find that for particle k we have

$$\frac{\nabla_k^2\Psi_C}{\Psi_C} = \sum_{ij\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_i)(\mathbf{r}_k-\mathbf{r}_j)}{r_{ki}r_{kj}}\frac{a}{(1+\beta r_{ki})^2}\frac{a}{(1+\beta r_{kj})^2} + \sum_{j\neq k}\left(\frac{2a}{r_{kj}(1+\beta r_{kj})^2} - \frac{2a\beta}{(1+\beta r_{kj})^3}\right) \tag{80}$$

And for the linear Padé-Jastrow we get the closed form result

$$\frac{\partial^2 f_{ij}}{\partial r_{ij}^2} = -\frac{2a_{ij}\beta_{ij}}{\left(1+\beta_{ij}r_{ij}\right)^3} \tag{81}$$

2.4.11 Importance sampling

We now want to make the code more efficient, so we replace the brute force Metropolis algorithm with a walk in coordinate space biased by the trial wave function, an approach based on the Fokker-Planck equation and the Langevin equation for generating a trajectory in coordinate space.

For one particle or walker, a diffusion process characterized by a time-dependent probability density $P(x,t)$ in one dimension we have the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - F\right)P(x,t), \tag{82}$$

where F is a drift term and D is the diffusion coefficient.

The new positions in coordinate space are found using the Langevin equation with Euler's method. We go from the Langevin equation

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta \tag{83}$$

where η is a random variable. This gives us a new position

$$y = x + DF(x)\Delta t + \xi\sqrt{\Delta t}. \tag{84}$$

Here ξ is gaussian random variable and Δt is a chosen time step. D comes from the factor $1/2$ in the kinetic energy operator, and is therefore equal to $1/2$ in atomic units.

The process of isotropic diffusion characterized by a time-dependent probability density $P(\mathbf{x}, t)$ will, as an approximation, obey the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \sum_i D\frac{\partial}{\partial \mathbf{x_i}}\left(\frac{\partial}{\partial \mathbf{x_i}} - \mathbf{F_i}\right)P(\mathbf{x}, t), \tag{85}$$

where $\mathbf{F}_i$ is component number i of the drift term caused by an external potential, and D is the diffusion coefficient. We set the left hand side equal to zero and obtain the convergence to a stationary probability density

$$\frac{\partial^2 P}{\partial \mathbf{x_i}^2} = P\frac{\partial}{\partial \mathbf{x_i}}\mathbf{F_i} + \mathbf{F_i}\frac{\partial}{\partial \mathbf{x_i}}P. \tag{86}$$

Inserting the drift vector, $\mathbf{F} = g(\mathbf{x})\frac{\partial P}{\partial \mathbf{x}}$, we get

$$\frac{\partial^2 P}{\partial \mathbf{x_i}^2} = P\frac{\partial g}{\partial P}\left(\frac{\partial P}{\partial \mathbf{x}_i}\right)^2 + Pg\frac{\partial^2 P}{\partial \mathbf{x}_i^2} + g\left(\frac{\partial P}{\partial \mathbf{x}_i}\right)^2 \tag{87}$$

To meet the condition of stationary density the left hand side has to be zero. This means that the terms containing first and second order derivatives has to cancel each other, which is only possible if $g = \frac{1}{P}$. This yields

$$\mathbf{F} = 2\frac{1}{\Psi_T}\nabla\Psi_T, \tag{88}$$

known as the quantum force. This so-called force pushes the walker towards regions of configuration space where the trial wave function is large, thus increasing the efficiency of the simulation. This is a great improvement on the Metropolis algorithm where the walker has the same probability to move in every direction.

From the Fokker-Planck equation we get a transition probability given by Green's function

$$G(y, x, \Delta t) = \frac{1}{(4\pi D\Delta t)^{3N/2}}\exp\left(-\frac{(y - x - D\Delta t F(x))^2}{4D\Delta t}\right) \tag{89}$$

This means that the Metropolis algorithm

$$A(y, x) = \min(1, q(y, x))), \tag{90}$$

where

$$q(y, x) = \frac{|\Psi_T(y)|^2}{|\Psi_T(x)|^2}, \tag{91}$$

is replaced by the Metropolis-Hastings algorithm,

$$q(y, x) = \frac{G(x, y, \Delta t)|\Psi_T(y)|^2}{G(y, x, \Delta t)|\Psi_T(x)|^2} \tag{92}$$

2.5 Implementation of MPI

As the calculations now become increasingly complex and heavy we implement MPI to make use of multiple processors. Personal computers today usually have two, four or sometimes eight processors, which will give a fairly good speedup to our calculations. With bigger atoms or systems it is crucial to implement a way to distribute calculation to multiple processors.

Implementing MPI in the Monte Carlo method is very easy. As we deal with statistical values we can easily split up the problem. Each process will run its own set of samples. The number of samples used by each process is simply n/p, where n is the total number of samples we want to do, and p is the number of processes. In the end all processors send their results to the master process, which sums up the values and takes the average over all processes.

2.6 Blocking

Blocking refers to a method to more accurately estimate the error of the values obtained by the VMC algorithm. It is and independent method from the VMC computation that can be used afterwards to get a more robust estimate of the variance. The basic idea lies on the correlations between all the measurements. If these are important enough, they will produce an increase in the error that needs to be taken into account. The reason behind this is related to the effective amount of measurements, if there are correlations there will be measurements that will contain less information, so these won't be as valuable as the rest and it will be as if there are *less* measurements than we actually have. Obviously this is a problem, the usual identification of the error with $\sqrt{\frac{\sigma}{n}}$ will be overly optimistic and a correction is needed.

$$f_d = \frac{1}{n-d}\sum_{k=1}^{n-d}\left(x_k - \bar{x}_n\right)\left(x_{k+d} - \bar{x}_n\right) \tag{93}$$

Where f_d is the correlation between measurements separated by a distance of d. This can be used to give an actual form to the correction factor:

$$\tau = 1 + 2\sum_{d=1}^{n-1}\frac{f_d}{var\left(x\right)} \tag{94}$$

This is the autocorrelation time and it relates the error with the variance:

$$err^2 = \frac{\tau}{n} var\left(x\right) \tag{95}$$

And the inverse of the first factor is the number of effective measurements (that are useful since they contain information):

$$n_{eff} = \frac{n}{\tau} \tag{96}$$

The expression that relates the standard deviation with this correlation time is thus:

$$\sigma = \sqrt{\left(\frac{1 + 2\tau/\Delta t}{n}\left(\bar{x^2} - \bar{x}^2\right)\right)} \tag{97}$$

Where Δt is the time between samples, and it's commonly smaller than τ. The main problem is that to compute τ a lot of time is needed, and this is not feasible in most cases.

The solution is to use blocking, and the algorithm to do this is quite simple. The total amount of measurements is divided into blocks of a certain size, and for each block the standard deviation is obtained. When the standard deviation stops increasing as the block size does, the correlations are irrelevant and the value for it is ready.

2.7 Energy minimization

As it can be expected, the values of the energy can depend heavily with respect to the variational parameters, so the optimal value must be found. This usually means finding the minimum value, and the first option would obviously be just a brute force search, but this is not efficient and there are better alternatives. Among these, we can find the steepest descent method, the conjugate gradient method and the Newton-Raphson method (also known as just Newton's method). The first two offer the possibility of finding a minimum in a multivariate space, while Newton's method only allows searching in one dimension, but is simpler and faster.

Since these methods don't find minima but the zeros of a function, the derivative of said function is needed. But there is no analytical expression for the local energy, so a workaround must be used. It is possible to find an analytical expression of this derivative as a function of the values of the local energy and the logarithmic derivative of the wave function. Following (Hjorth-Jensen, 2013) section 16.11 the expression for the derivative with respect

to parameter c is:

$$\frac{\partial E}{\partial c} = 2\left[\langle E_L \frac{\partial \ln \Psi}{\partial c}\rangle - E\langle \frac{\partial \ln \Psi}{\partial c}\rangle\right] \tag{98}$$

Or more explicitly:

$$\frac{\partial E}{\partial c} = \frac{2}{N}\left[\sum_{i=1}^{N}\left(\left[E_L\left(c\right)\right]_i\left[\frac{\partial \ln \Psi}{\partial c}\right]_i\right) - \frac{1}{N}\sum_{i=1}^{N}\left(\left[E_L\left(c\right)\right]_i\sum_{j=1}^{N}\left[\frac{\partial \ln \Psi}{\partial c}\right]_j\right)\right] \tag{99}$$

Where the indices's i and j run through all the timesteps independently of each other. In our case, we are going to derive with respect to β only, because we are only interested in minimization in that direction. Since the derivative is composed of three parts, two for the Slater determinants and one for the Padé-Jastrow interaction part, we get:

$$\frac{\partial \ln \Psi}{\partial \beta} = \frac{\partial \ln \Psi_{SD\uparrow}}{\partial \beta} + \frac{\partial \ln \Psi_{SD\downarrow}}{\partial \beta} + \frac{\partial \ln \Psi_J}{\partial \beta} = \frac{\partial \ln \Psi_J}{\partial \beta} = \frac{\partial\left(\sum_{i<j}\frac{ar_{ij}}{1+\beta r_{ij}}\right)}{\partial \beta} = \sum_{i<j}\frac{-ar_{ij}^2}{\left(1+\beta r_{ij}\right)^2} \tag{100}$$

So the resulting expression is:

$$\frac{\partial E}{\partial \beta} = \frac{2}{N}\left[\sum_{i=1}^{N}\left(\left[E_L\left(\beta\right)\right]_i\left[\sum_{k<l}\frac{-ar_{kl}^2}{\left(1+\beta r_{kl}\right)^2}\right]_i\right) - \frac{1}{N}\sum_{i=1}^{N}\left(\left[E_L\left(\beta\right)\right]_i\sum_{j=1}^{N}\left[\sum_{k<l}\frac{-ar_{kl}^2}{\left(1+\beta r_{kl}\right)^2}\right]_j\right)\right] \tag{101}$$

With k and l running through the electrons. Newton's method will be used for its simplicity, and because we will only work with β. This method additionally requires the derivative of the function ($\frac{\partial E}{\partial \beta}$ in this case), but it's not possible to simply derive with respect to β again because there is no analytical expression for $E_L\left(\beta\right)$, so we must resort to numerical derivation. A simple, first oder finite difference method can be used for this:

$$\frac{\partial^2 E}{\partial \beta^2} = \lim_{h\to 0}\frac{\frac{\partial^2 E(\beta+h)}{\partial \beta^2} - \frac{\partial^2 E(\beta)}{\partial \beta^2}}{h} \tag{102}$$

And with this Newton's method can be implemented in a straight forward manner with one simple equation:

$$\beta_{i+1} = \beta_i - \frac{\frac{\partial E}{\partial \beta}}{\frac{\partial^2 E}{\partial \beta^2}} \tag{103}$$

The index i represents the iterations of Newton's method, not the iterations of the Monte Carlo loop, for each iteration of this index, a whole Monte Carlo loop is performed. In iteration 0 a seed must be provided. Depending on the guess for this seed, the method's performance will be better or worse.

There is one catch, the values of the first derivative are computed as a part of the Monte Carlo loop, and are thus susceptible to a certain degree of randomness. This variability introduces some uncertainty in each iteration. Basically, we are not applying the method to a function, but to a scattered point cloud from which we sample points one at a time. This, coupled with the fact that the second derivative is obtained numerically from two points of that cloud, made the method show poor results, namely slow convergence to values that apparently depended heavily on the seed choice. To circumvent this issue, a different, yet similar method was used.

The bisection method is similar to Newton-Raphson method, but it only needs the function that has the roots we want to find. And, instead of a point seed, it needs an interval seed. The method exploits Bolzano's theorem in said interval: if the values of the function (the derivative of the local energy in this case) in the extremes have different signs, the existence of at least one root is guaranteed. By evaluating the function in the midpoint of the interval, it's possible to know in which half of it the root lies, and thus the interval can be reduced to one of its halves. This process is repeated until a certain tolerance is reached and the root is obtained. In our case we don't have to worry about multiple roots because we know that there is only one minimum, the only problem is choosing the appropriate intervals so that the method can find the root.

The main advantage here compared to Newton's method is that the sensibility to the values of the function is much smaller because only the sign of the function values is important (and this only becomes a problem when we are already very close to the root). The other advantage is the robustness and simplicity, the method will converge no matter what if the appropriate interval is chosen, which is not something very complicated to guess and can be

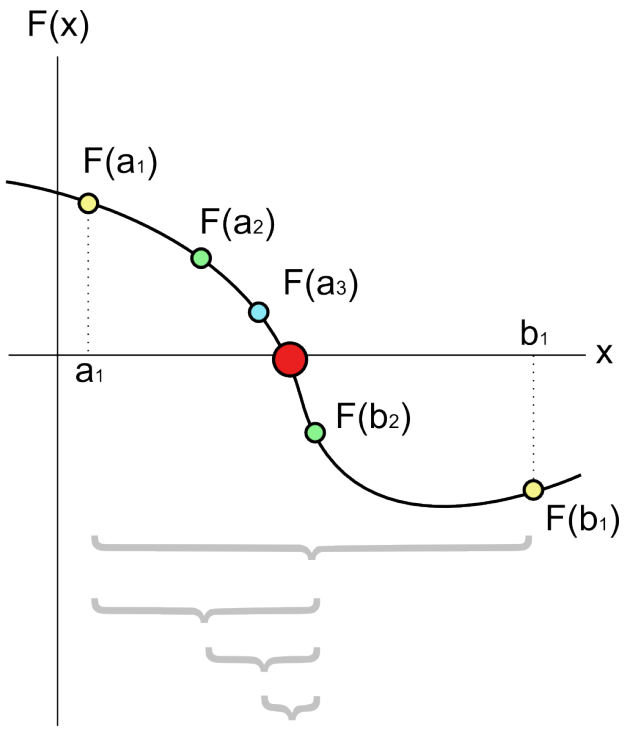

Figure 2: Schematic representation of the bisection method: in each iteration the interval is halved in two and it converges linearly to the solution

checked quite fast in case it's not so obvious. The downside is the linear, comparatively slow convergence; but in this case, due to the the imprecision in obtaining the values of the derivative of the local energy, Newton's method becomes relatively slow as well, so it's not actually a problem for our purposes.

2.8 Gaussian type orbitals

As proposed in a paper by Boys in 1950 (Boys, 1950), Gaussian Type Orbitals, or GTOs, can be used in electronic structure theory, and it is today common to use them, especially in computational chemistry. Most importantly for us they make us able to limit the variational calculations of complex atoms to one variational variable, thus reducing the number of variational possibilities significantly. This is obviously a great advantage, because as the number of particles in our calculations grow, the energy takes considerably longer time to compute for each combination of variables.

2.8.1 Using GTOs to replace the Slater type orbitals

To replace our current orbitals we need to calculate primitive GTOs and contract them to contracted GTOs. A contracted GTO, ϕ, is defined as

$$\phi(x,y,z) = \sum_i N_i \chi_i(x,y,z),$$

where χ_i is a primitive GTO and N_i is a normalization constant. The primitive GTO is defined as

$$\chi_i(x,y,z) = c_i x^m y^n z^o e^{-\alpha_i R^2}.$$

Here x, y, and z are Cartesian coordinates representing the distance to a nucleus, and $R^2{=}x^2 + y^2 + z^2$. The quantum numbers m, n and o depend on the angular momentum of the orbital. The numbers c_i and α_i are variational parameters, which can be fetched from an online library, namely EMSL (J. S. Binkley, 1980)(*EMSL Basis Set Exchange* n.d.). From this library we use the 3-21G basis to represent the orbitals.

When combining contracted primitives into orbitals that are to mimic and replace the Slater type orbitals additional constants are needed, one for each contracted GTO. These are calculated using Hartree-Fock calculations. For Helium, Beryllium and Neon these constants are given in table 7.

The contracted GTO primitives are not a perfect representation for the Slater Type Orbitals, however, as we can see in figure 3 and figure 4. The shape of the contracted GTO resembles the STO, but the top of the GTO is still not as sharp as the STO. A way to get a closer resemblance to the STO is to use another basis with a bigger set of primitives.

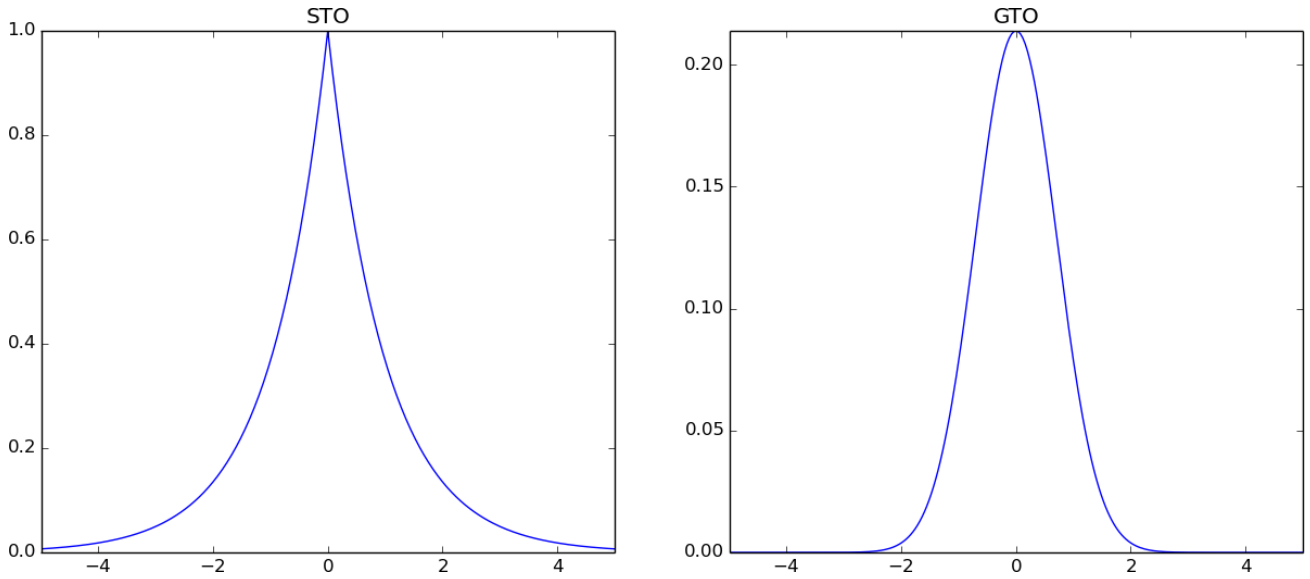

Figure 3: The shape of a GTO primitive compared to the shape of a Slater Type Orbital.

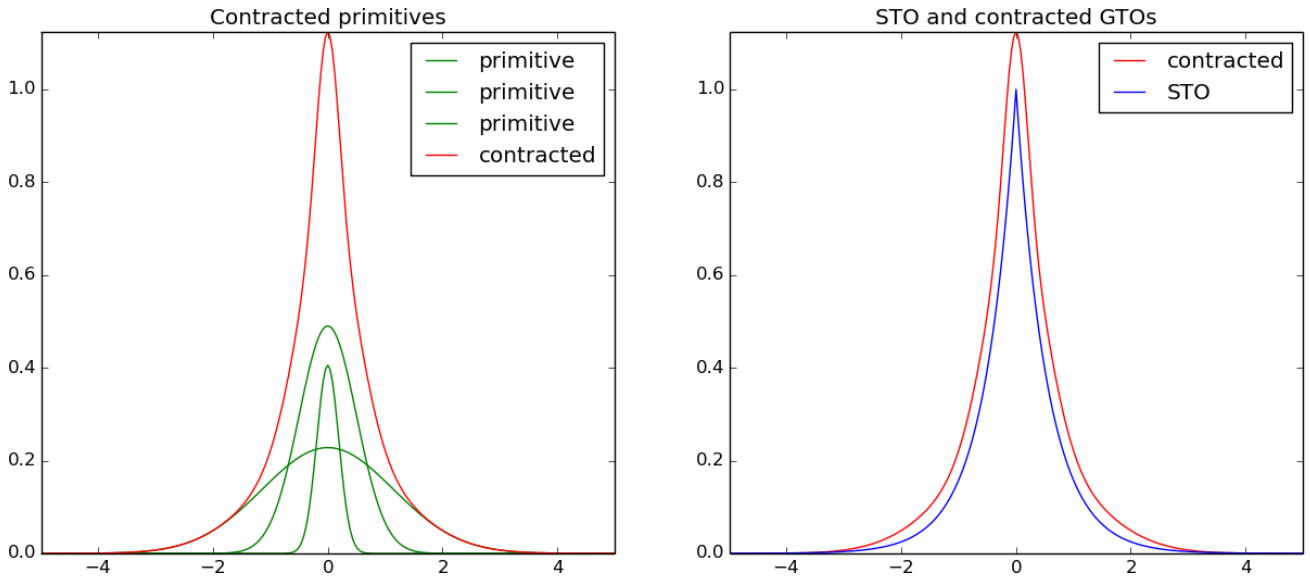

Figure 4: Construction of a contracted GTO from primitives and comparison of the contracted GTO to the Slater Type Orbital.

Atom	α	β	Cycles	VMC [au]	Variance	Reference energy [au]
Helium	1.843	0.34	10^8	-2.89024	3.77402×10^{-5}	-2.9037

Table 2: Comparison of the energy for Helium found with VMC without using Importance Sampling, and the reference energy found in research papers (Binkley and Pople, 1975).

Trialfunction	Numerical (s)	Analytical (s)	Ratio
Helium ψ_T	29.7288	20.1189	0.6767
Beryllium ψ_{T2}	58.9623	32.4622	0.5505

Table 3: The time to run a Monte Carlo run with 10^7 cycles for Helium, and 10^6 cycles for Beryllium. The closed expression for the local energy increased the computation time by a significant degree for each trialfunction.

As an example, for Helium, we find the parameters to be

	α_i	c_i
χ_1	13.62670	0.175230
χ_2	1.999350	0.893483
χ_3	0.382993	1.000000

Thus the Gaussian orbital becomes

$$\begin{aligned}\phi(x,y,z) &= 0.4579 \times \left(\left(\frac{2 \times 13.62670}{\pi} \right)^{3/4} 0.17523 e^{-13.62670 \times R^2} + \left(\frac{2 \times 1.99935}{\pi} \right)^{3/4} 0.893483 \times e^{-1.99935 \times R^2} \right) \\ &\quad + 0.6573 \times \left(\left(\frac{2 \times 0.382993}{\pi} \right)^{3/4} 1.0 e^{-0.382993 \times R^2} \right)\end{aligned}$$

For a complete list of the parameters used see Table 7 in Appendix C

3 Results

3.1 Variational Monte Carlo calculations of the helium atom

As a first attempt to solve the ground state energy for the helium atom we perform Variational Monte Carlo calculation with a brute force Metropolis sampling. We do this with two trial wave functions

$$\psi_T(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_{12}}) = \exp\{(-\alpha(r_1 + r_2))\} \exp\left\{ \left(\frac{r_{12}}{2(1+\beta r_{12})} \right) \right\},$$

using α and β as variational parameters. We run the Variational Monte Carlo calculation over different values for the two variables α and β, with 2×10^7 cycles in the Monte Carlo simulation, we get the results presented in figure 5.

We find the optimal values to be $\alpha = 1.843$ and $\beta = 0.34$, as we can see in the figures. Using these values for α and β we run the brute force variational Monte Carlo calculation. The program finds an optimal value for the steplength, δ, which results in roughly 50% accepted moves. Using 10^8 cycles the algorithm finds the steplength to be $\delta = 1.5$, giving 48.8% accepted moves. The energy found with this method is -2.89024 au, with a variance of 3.77402×10^{-5}, as presented in table 2. The parameter α can be interpreted as a parameter for the force pulling the electron to the nucleus.

3.1.1 Alpha and Beta Values

Table 4 shows the values, for α and β values, we got from running several Monte Carlo cycles with different values. As an algorithm to pick out the best values we minimized the energy found in the Monte Carlo runs. The values found are quite uncertain since the variance of the energy was quite high compared to the difference caused by varying the parameters. The variance was more smooth as a function of the parameters, see fig 5 and it is therefore easier to determine the best values for the variables from variance as a function of α and β.

3.1.2 Computational speed gain by using an analytical local energy

By using an analytical expression for the local energy instead of using a numerical derivation in the calculation the program becomes more efficient. For Helium we get a speedup of nearly 48%, while for Beryllium we get a speedup of approximately 80% was achieved, see table 3.

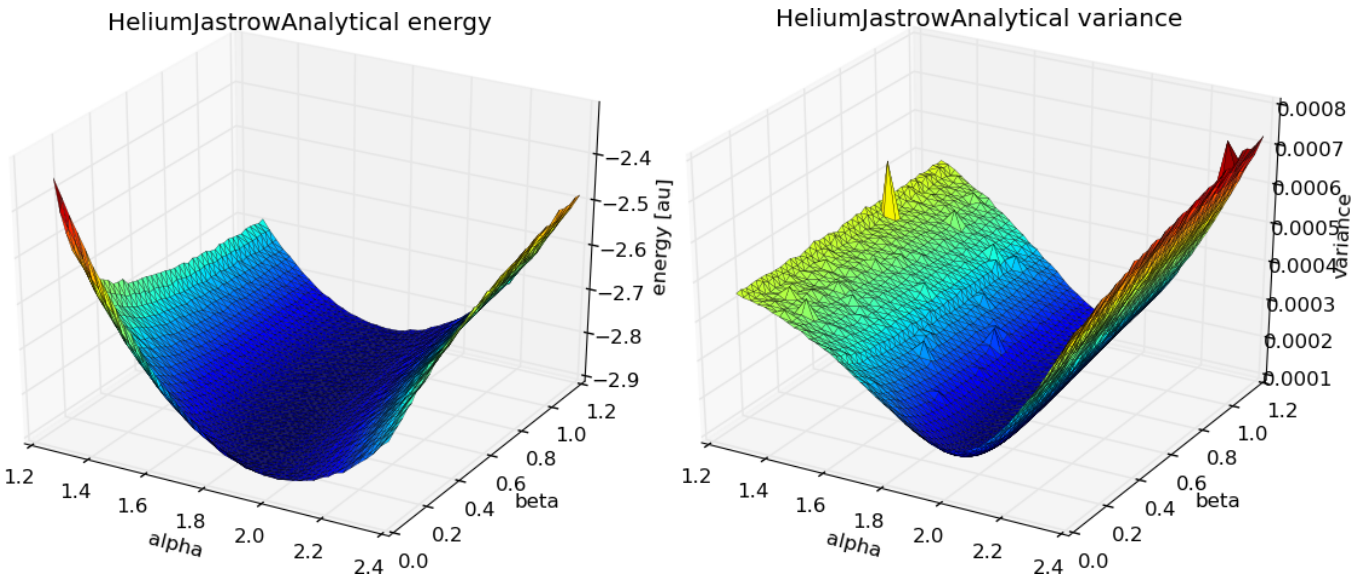

Figure 5: Using ψ_T, plot of the energy versus alpha and beta, and plot of the variance versus α and β.

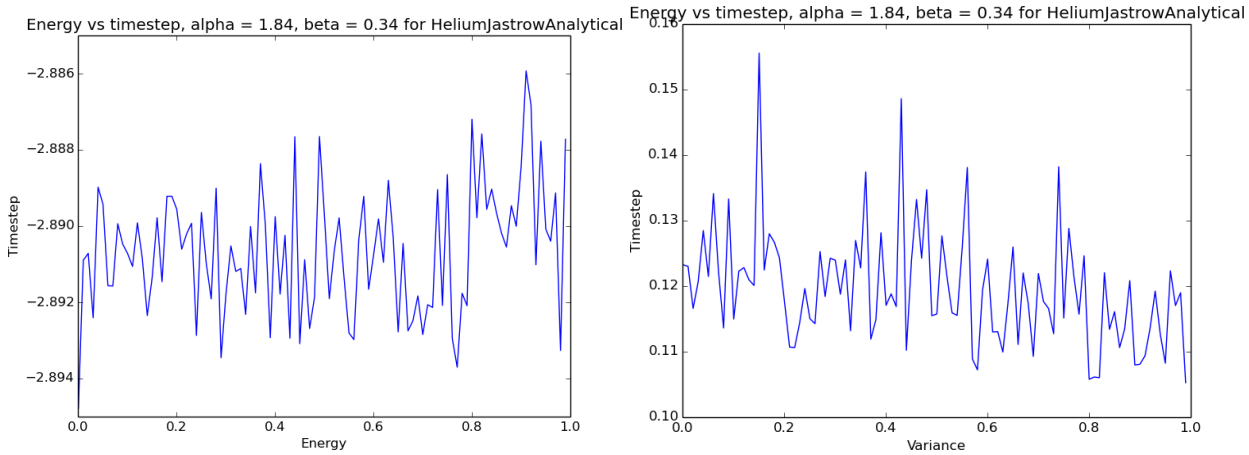

Figure 6: Plots for Helium ψ_T for the energy versus the timestep, and the variance versus the timestep.

3.1.3 Calculations using importance sampling

We now introduce importance sampling to our calculations. We search for the optimal variables and find them to be $\alpha = 1.843$ and $\beta = 0.34$. Using these values, with 10^8 cycles, we get an energy of -2.89012 au and a corresponding variance 7.76888×10^{-5}, as presented in table 4. The energy and variance as a function of the timestep, δt is shown in figure 6.

3.1.4 Bisection method and GTOs

Using GTOs we can nearly reproduce the energies we get from using Slater Type Orbitals, although we are unable to reach quite the same energy. This is due to the GTOs being an imperfect representation of the STOs. Using GTOs in stead of STOs is also considerably slower, as we see in 6. This can be due to inefficient code, and there is also room for improvements by calculating analytical derivatives of the GTOs. However using GTOs reduces the variables to only β, thus we don't have to look for energy minima by varying α. This means that it is possible to easily use a gradient method to more efficiently find the variable that gives minimum energy.

3.2 Variational Monte Carlo calculations of the Beryllium and Neon atoms

We attempt to solve the ground state energy for the Beryllium atom and Neon atom using a Variational Monte Carlo calculation with importance sampling. We have used the trial functions (16) for Beryllium and (23) for Neon which uses α and β as variational parameters.

For Beryllium we have the Alpha and Beta values $\alpha = 4.0$ and $\beta = 0.31$. We use 10^7 cycles and find the energy

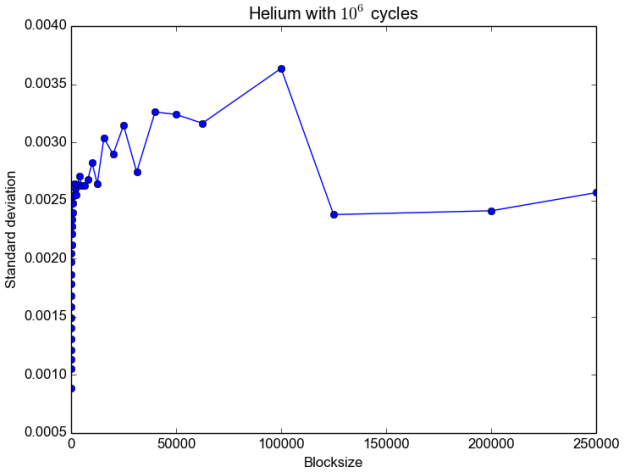

Figure 7: Variance vs. blocksize for Helium. The blocking behaviour can be seen very clearly as the variance plateaus as the blocksize grows.

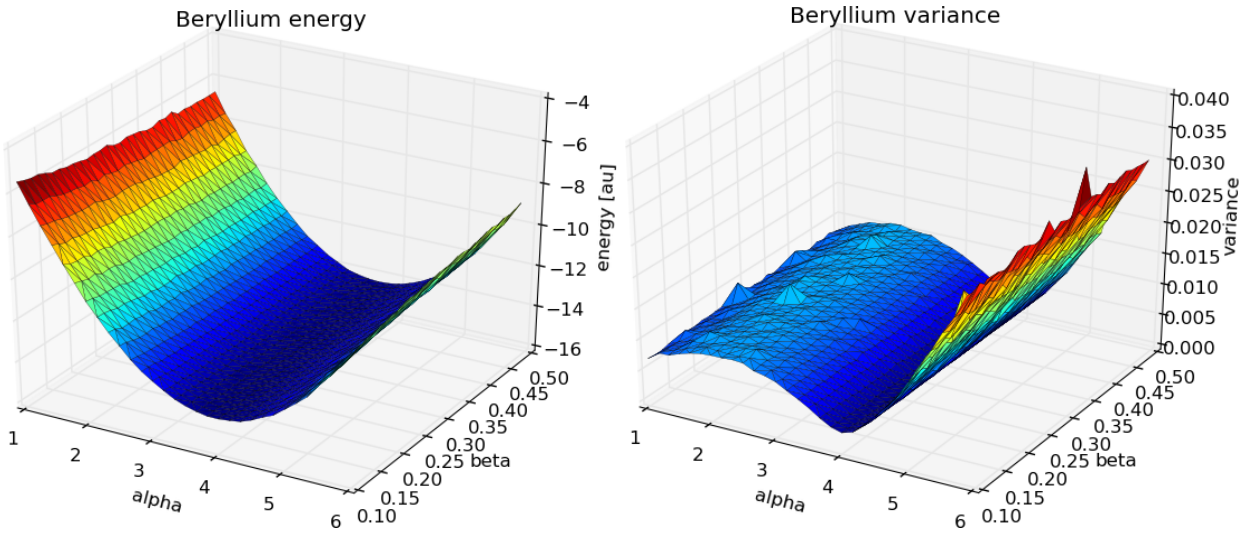

Figure 8: Energy (left) and variance (right) for different Alpha and Beta values for Beryllium, using 10^6 cycles.

to be -14.3902 au, with a variance of 9.08566×10^{-4}. For Neon, with 10^6 cycles, we get an energy of -127.875 au with a variance of 0.0131537.

However, as expected, by using Gaussian Type Orbitals we get less accurate energies compared to the reference values, as shown in table 6, and it is considerably slower. Nontheless, with using GTOs we only have one variable we need to vary to find the best ground state energy, β. Therefore we are also able to use bisection method to find the optimal β-value.

From reseach papers we find the value for energy in the ground state of Beryllium to be -14.667 au (Koput, 2011) and the value for energy in the ground state of Neon to be -128.928 au (Binkley and Pople, 1975). In Table 4 we compare results obtained with our Variational Monte Carlo method with results from various research papers.

3.2.1 Alpha and Beta Values

To find optimal Alpha and Beta values for the atoms we run VMC with ranges of different values for α and β. The resulting plots of variance and energy for different combinations are given in figure 8 for Beryllium and figure 9 for Neon. The optimal values are shown in table 4. As VMC runs slowly for Neon, because it has 10 electrons, we were only able to run over the range of Alpha and Beta values with 10^5 cycles. This is reflected in the higher variance, and the spikes in the variance plot.

3.2.2 Speedup with MPI

It is desirable to have a speedup as close as possible to the number of processors used. The speedup measured by our VMC program running 1, 2, 3 and 4 is shown in table 5 and figure 10. We see that the speedup is good for 2

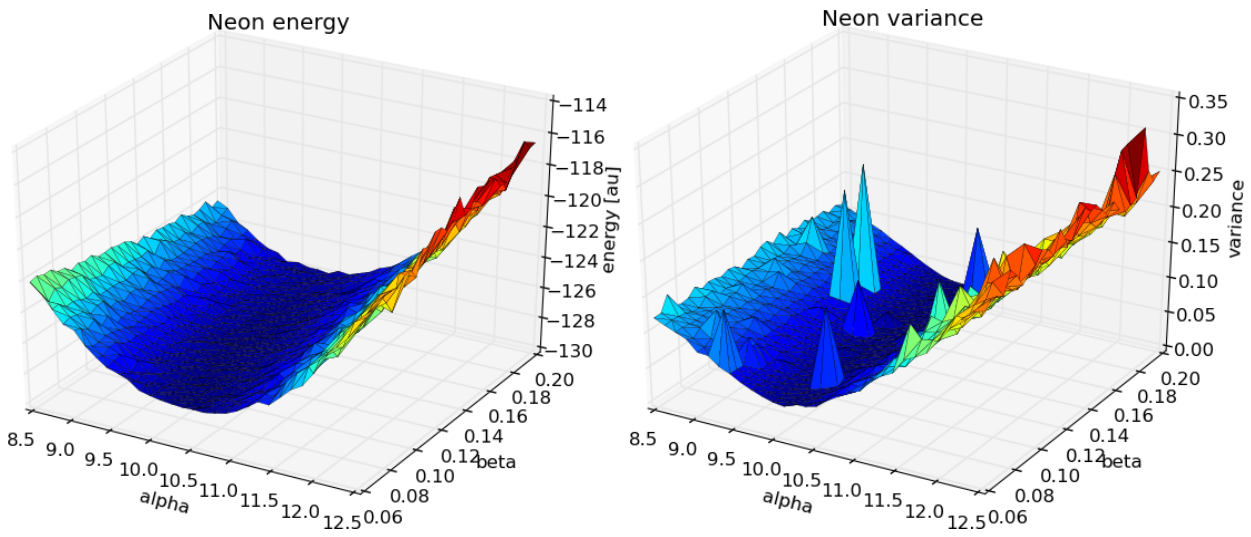

Figure 9: Energy (left) and variance (right) for different Alpha and Beta values for Neon, using 10^5 cycles.

Atom	α	β	Cycles	VMC [au]	Variance	Reference energy [au]
Helium	1.843	0.34	10^8	-2.89012	7.76888×10^{-5}	-2.9037
Beryllium	4.0	0.31	10^7	-14.3902	9.08566×10^{-4}	-14.667
Neon	10.22	0.091	10^6	-127.875	0.0131537	-128.928
H_2	1.289*	0.401	10^7	-1.15828	0.000225	-1.17540
H_2*	1.289	0.401*	10^7	0.320	0.001105	–
Be_2	3.725*	0.246	10^6	-31.349	0.00756	-29.339

Table 4: Comparison of energies resulting from our Variational Monte Carlo method with energies found in research papers (Koput, 2011) (Binkley and Pople, 1975). The values for α and β where found by doing running Monte Carlo calculation with over a mesh of different α and β values. The run with the lowest energy gave the α and β values. For H_2 and Be_2 we used α, β values, along withe the benchmark value, from Ihle and Ledum (2013) and used a nuclei distance of 1.40 a.u. and 4.63 a.u. respectively. The binding enery found for Be_2 is too low which is caused by a bug in the implementation of the molecule. H_2 is the Hydrogen molecule where the $1S$ wavefunctions was subtracted instead of added together.*

Num. of processes	1	2	3	4
Speedup	1.0	1.97	2.90	3.35

Table 5: MPI speedup

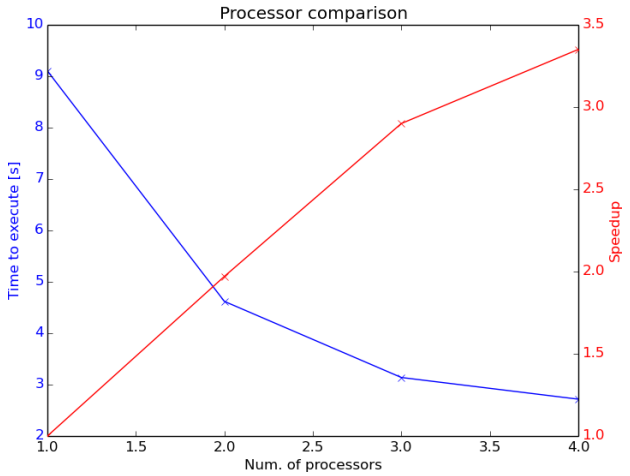

Figure 10: MPI speedup

Atom	β	Cycles	VMC [au]	Variance	Ref. energy [au]	GTO [s]	STO [s]
Helium	0.51625	2×10^6	-2.85482	0.00405383	-2.9037	15.007	6.48565
Beryllium	0.091797	7×10^7	-14.0182	0.00203359	-14.667	48321	4141
Neon	0.109375	1×10^6	-113.542	0.498411	-128.928	2821.16	203.012

Table 6: Comparison of energies found using bisection method with the refenrence energy (Koput, 2011) (Binkley and Pople, 1975) and comparison of the time used running the computation with the given number of cycles using GTOs and STOs.

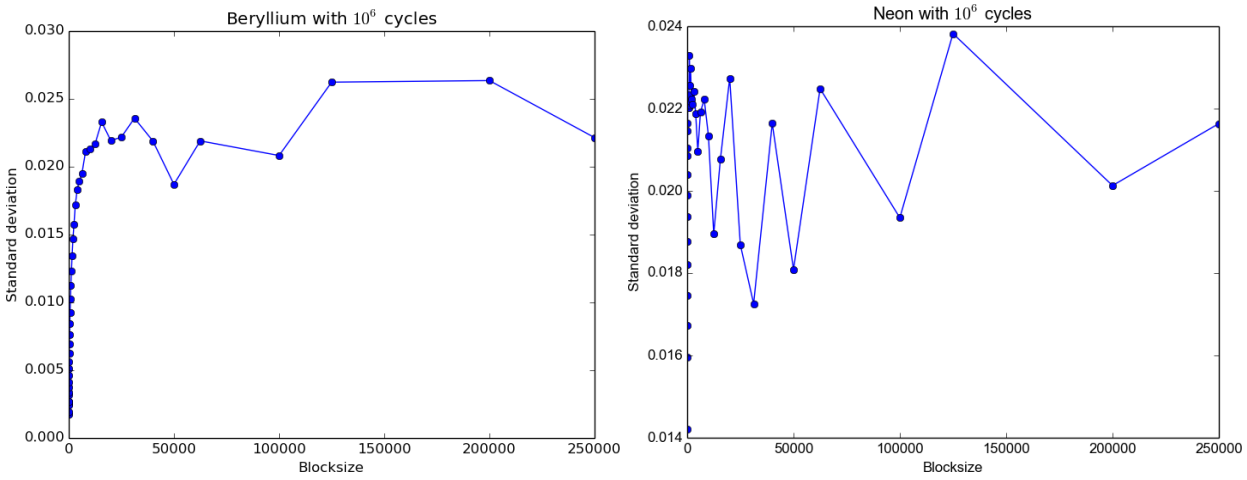

Figure 11: STD vs. blocksize with 10^6 MC cycles. The STD shows a clear plateauing behaviour trend as the blocksize increases after rising dramatically for small blocksizes. This is due to the blocksize catching up to the correlation lenght between the measurements, thus yielding the right STD.

and 3 processes, but for 4 processes suffers somewhat because it also have to run the OS and other programs.

3.3 Variational Monte Carlo calculation for Hydrogen and Beryllium molecules

For the Hydrogen molecule we did VMC computations of both two different trialfunctions as described in Section 2.2.5. With adding together the hydrogenic wavefunctions we found a binding energy of -1.15828 a.u., with $\sigma_*^2 = 0.000225$. For the trialfunction were the hydrogenic wavefunctions were subtracted the binding energy were positive and unrealistic, see Table 4 for values. For the second VMC calculation, which should be for two same spin electrons the nuclei distance, α and β values used were probably not correct and we should have spent more time making sure that they were.
In the computations of the Beryllium molecule there is some error in the code or procedure which causes faulty values, so the energy values are not to be trusted.

3.4 Charge density plots

3.4.1 Helium Atom

For the Helium atom we have ran simulations, with 10^6 Monte Carlo cycles, of several different trialfunctions and recorded the position of the electrons. We first did an experiment with the electrons being in a pure hydrogenlike wavefunction, then we ran an hydrogenlike wavefunction but optimized the α value to get a better ground state energy then at last we also included a Pàde-Jastrow correlation factor to include the effects of the electron-electron repulsion. For the pure hydrogenic wavefunction we got $E = -2.757$ and $\sigma_*^2 = 0.0030$, while with the optimized α value we got $\alpha = 1.65$, $E = -2.843$ and $\sigma_*^2 = 0.0027$, and lastly with the Pàde-Jastrow factor we got $\alpha = 1.843$, $\beta = 0.34$, $E = -2.887$, $\sigma_*^2 = 0.0010$. Here σ_*^2 is the variance for the local energy between each particle move in the VMC, not the true variance, which is obtained through the blocking routine described in Section 2.6. In figure 12 we plots of the positions the electrons occupied during the simulations and we can see that for the pure hydrogenic wavefunction the electrons was generally closer to the nucleus than in the other two simulations. The ground-state energy improved and got closer to the reference value Table 4 as the wavefunction got more sophisticated as well as the variance decreasing. In the wavefunction the α parameter represents the electrical attraction of the nucleus on the electron, and is equal to the charge in a Hydrogen atom, in a real Helium atom the attraction felt is smaller than the charge of the nucleus since there is also the effect of the other electrons. This is why we get a better simulation with a smaller α value. For the Helium atom the effect of the correlation factor is not very pronounced,

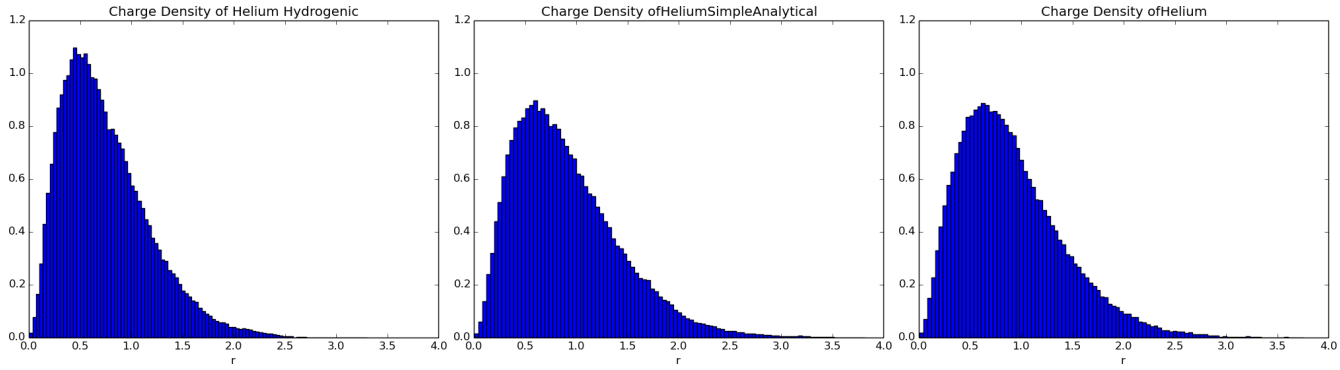

Figure 12: The electron-nucleus distance in an Helium atom simulated with different trialfunctions. From the left the plots represents pure hydrogenic wavefunctions ($\alpha = 2$, $E = -2.757$ and $\sigma_*^2 = 0.0030$), hydrogenic wavefunctions with an optimized alpha ($\alpha = 1.65$, $E = -2.843$ and $\sigma_*^2 = 0.0027$) and hydrogenic wavefunctions with a Jastrow factor and optimized parameters, ($\alpha = 1.843$, $\beta = 0.34$, $E = -2.887$, $\sigma_*^2 = 0.0010$). The simulations were run with 10^6 Monte Carlo cycles.

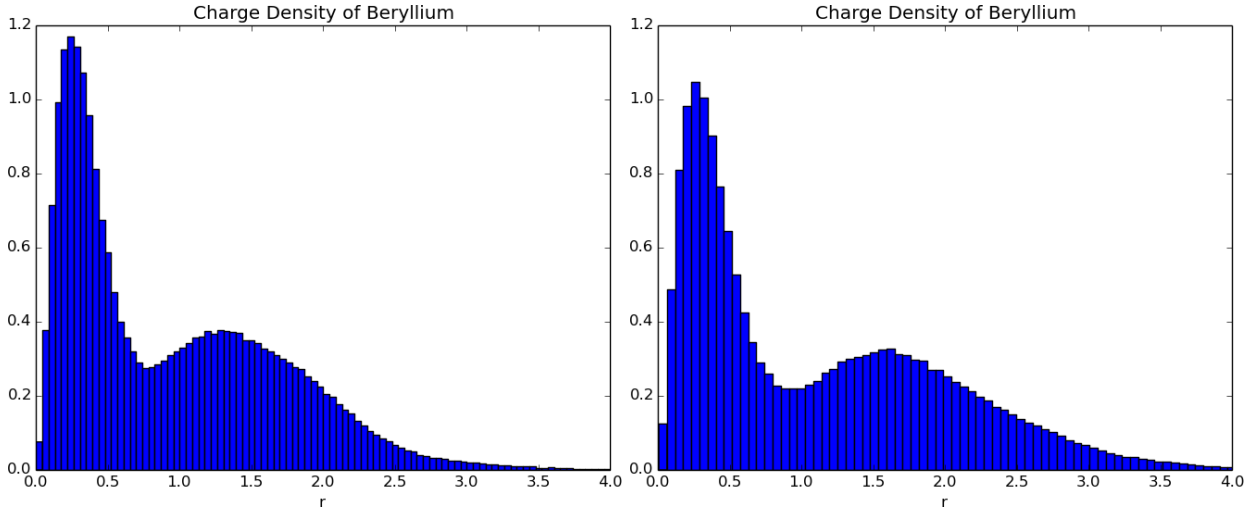

Figure 13: The electron-nucleus distance in an Beryllium atom simulated with different trialfunctions. The left plot is with hydrogenic wavefunctions ($\alpha = 4$, $E = -13.710$, $\sigma_*^2 = 0.0072$) and the right plot is withhydrogenic wavefunctions with a Jastrow factor and optimized parameters, ($\alpha = 4$, $\beta = 0.31$, $E = -14.385$, $\sigma_*^2 = 0.0029$). The simulations were run with 10^6 Monte Carlo cycles.

since the electrons can quite far away from each other, but it is still significant in the ground-state energy and the variance.

As expected since the trialfunctions used here is based only on the ψ_{1S} orbitals the radial distribution is similar to the ψ_{1S} hydrogen orbital in Fig. 1.

3.4.2 Beryllium

With the hydrogenic Beryllium trialfunction we got an ground-state energy of -13.710 with a σ_*^2 of 0.0072, when the correlation factor was included the energy dropped to -14.385 with a σ_*^2 of 0.0029. The hydrogenic wavefunction consists for Beryllium is made up of a Slater determinant constructed of the first two hydrogenic orbitals, plotted in Fig. 1. The traces of the $1S$ orbital is there in the first maximum and is closer to the nucleus than in an hydrogen atom, which is expected since the distance is a function of the nucleus charge, the contribution from the second orbital $2S$ is visible as the second maximum in the distribution Fig. 13. When the correlation factor is included the distribution is smeared out and the orbitals are not as sharp.

3.4.3 Neon

The ground state energy and the σ_*^2 got much better with the correlation factor included, from -110.301 to 127.888 for the energy and 0.0477 to 0.0190 for the estimation of the variance. The trialfunctions include all the three orbitals and the distribution is contracted closer to the nucleus than in the other atoms. Here we can see the largest effect due to the inclusion of the correlation factor, both in the ground-state energy and the radial distribution.

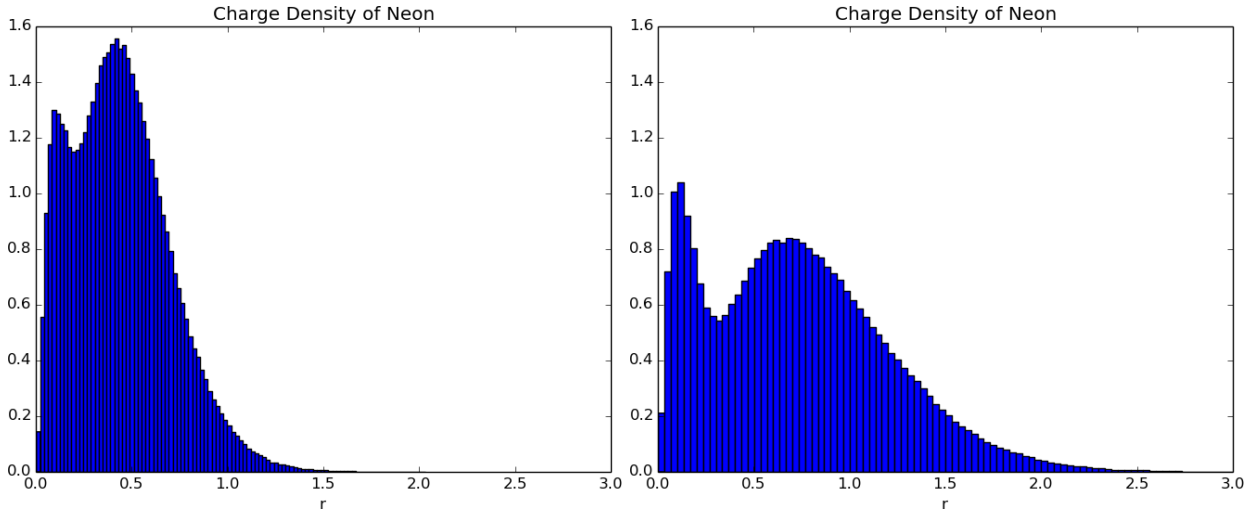

Figure 14: The electron-nucleus distance in an Neon atom simulated with different trialfunctions. The left plot is with hydrogenic wavefunctions ($\alpha = 10.22$, $E = -110.301$, $\sigma_*^2 = 0.0477$) and the right plot is withhydrogenic wavefunctions with a Jastrow factor and optimized parameters, ($\alpha = 10.22$, $\beta = 0.0.091$, $E = -127.888$, $\sigma_*^2 = 0.0190$). The simulations were run with 10^6 Monte Carlo cycles.

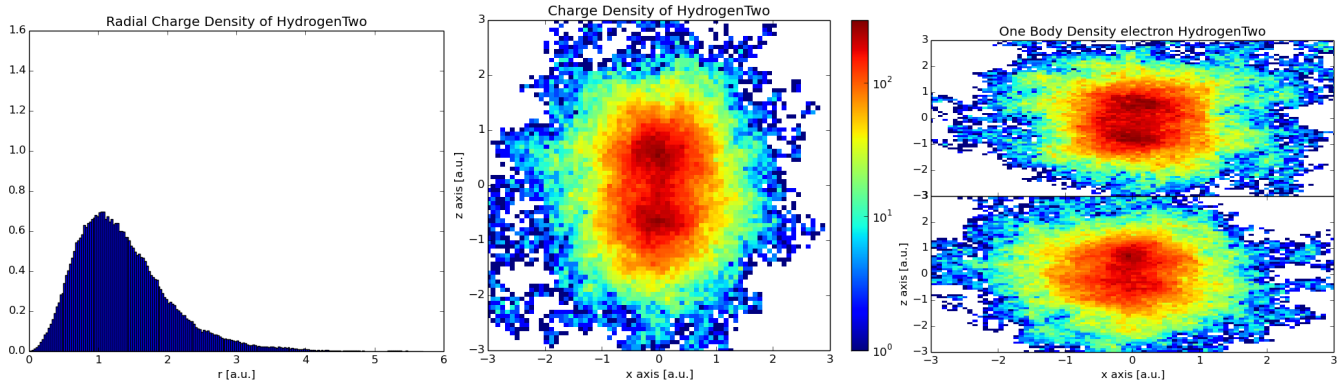

Figure 15: The left figure is the radial distribution of the Hydrogen Molecule from the origo, where both nuclei where placed $R = 1.40 a.u$ way from each other on the z axis. The figure in the center shows a slice of the xz plane with with of 1.0 a.u. On the right figure one body denseties off each of the electrons is shown. The run used had 10^6 VMC cycles. Both the nuclei are visible on the charge density plot as a denser concentration.

This is because the electrons are closer together than in the other atoms due to the higher charge of the nucleus.

3.4.4 Hydrogen Molecule

After extending the VMC calculation to also be able to work with simple atoms we did two runs and recorded the positions of the electrons and made charge density plots for the electrons. We used parameters from Ihle and Ledum, 2013. For the H_2 molecule, which is comparable to a Helium atom with the protons seperated, the radial distribution from the mass center was slightly more smeared out than in the Helium atom because the protons were slightly apart. In the charge density plot about the xz plane the electron density around the nuclei were higher. When the one-body density were plotted both the electron tended to move about both the nuclei and were not trapped along one nuclei.

3.4.5 Beryllium Molecule

For the Beryllium molecule the consentration of the charge density around the nuclei were sharper than in the Hydrogen molecule, which is likely caused by the higher charge of the nuclei. In the Beryllium atom we also see that the electron density is higher closer to the nucleus. The charge density plots here are not to be completely trusted because there is a bug in the implementation of the molecule which causes a too low binding energy, but much of the expected physics still seem to be captured by the VMC computation. In the one body plots following only one of the electrons we can see that, compared to the Hydrogen molecule, the nuclei in the molecule traps it's electrons more effieciently.

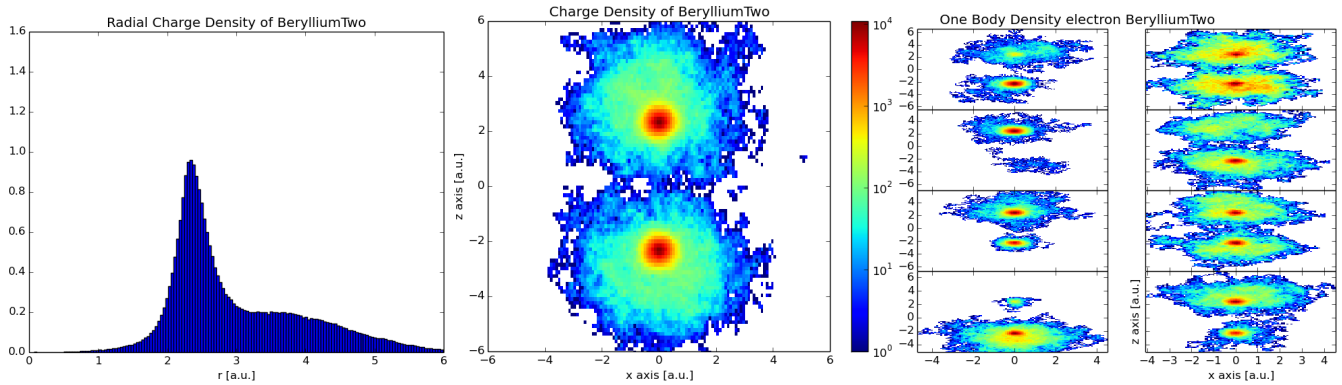

Figure 16: The left figure is the radial distribution of the Beryllium Molecule from the origo, where both nuclei where placed $R = 4.63 a.u$ *way from each other on the z axis. The figure in the center is shows a slice of the xz plane with with of* 1.0 *a.u. On the right figure one body denseties off each of the electrons is shown. The run used had* 10^6 *VMC cycles. The consentration of the charge density is very sharp around the two nuclei.*

4 Conclusions and discussion

We see that the Variational Monte Carlo method gives good values, both for small systems like Helium and Beryllium, and also for larger systems like Neon, as long as we deal with the interaction between the particles using the Jastrow factor. We have implemented MPI, making it possible to split up the Monte Carlo simulation in pieces and distribute it to multiple processors. This reduces the time to run simulations considerably. With a larger system we also needed a better way to handle the Slater Determinant. Using Slater Type Orbitals we get the best values, but using this makes the calculations dependent on two variables. This makes us able to use the bisection method to find the optimal beta value.

Taking a look at 4 it is apparent that the results are quite precise. In the case of Helium the experimental value is less than 0.5% higher than the computational one. For Beryllium the experimental value is just under 2% higher than the reference value, and for Neon, the difference is even less at just under 1%.

For all of the atoms the radial distribution of the wavefunctions showed traces of the hydrogenic wavefunctions they were constructed from. As the binding energy of the atoms during the VMC computations are close to the experimental values, it is likely that the atomic structure of the real atoms resemble the atomic structures of the VMC simulations.

These results are very good considering the amount of resources that went into calculating them: a simple standard PC and a handful of hours were enough. Considering that a relatively simple and very well tested method was used to obtain these results it is very easy to see why VMC is so popular for solving quite complicated problems like calculating the ground state energies of non-trivial quantum mechanical systems.

Looking ahead, to improve the Variational Monte Carlo program we could calculate more analytical functions for the Slater Determinant and for the derivatives of the GTO functions. These parts are used a lot during the program and switching from numerical to analytical solution would speed things up. Further, as the Gaussian Type Orbitals gave slightly dissapointing energies, we could implement another, larger basis set than the 3-21G basis, such as the 6-311G basis. This would give the GTO functions a shape that more closely resembles Slater Type Orbitals and therefore give better energies.

4.1 Critique on the exercise

The exercise is well made and has a good learning curve during the three projects. We feel that we have gained insight into calculation of physical properties using Monte Carlo methods. There are however a couple things that could be better. The charge density and GTOs could be better explained and feedback on project 1 and 2 could have come sooner.

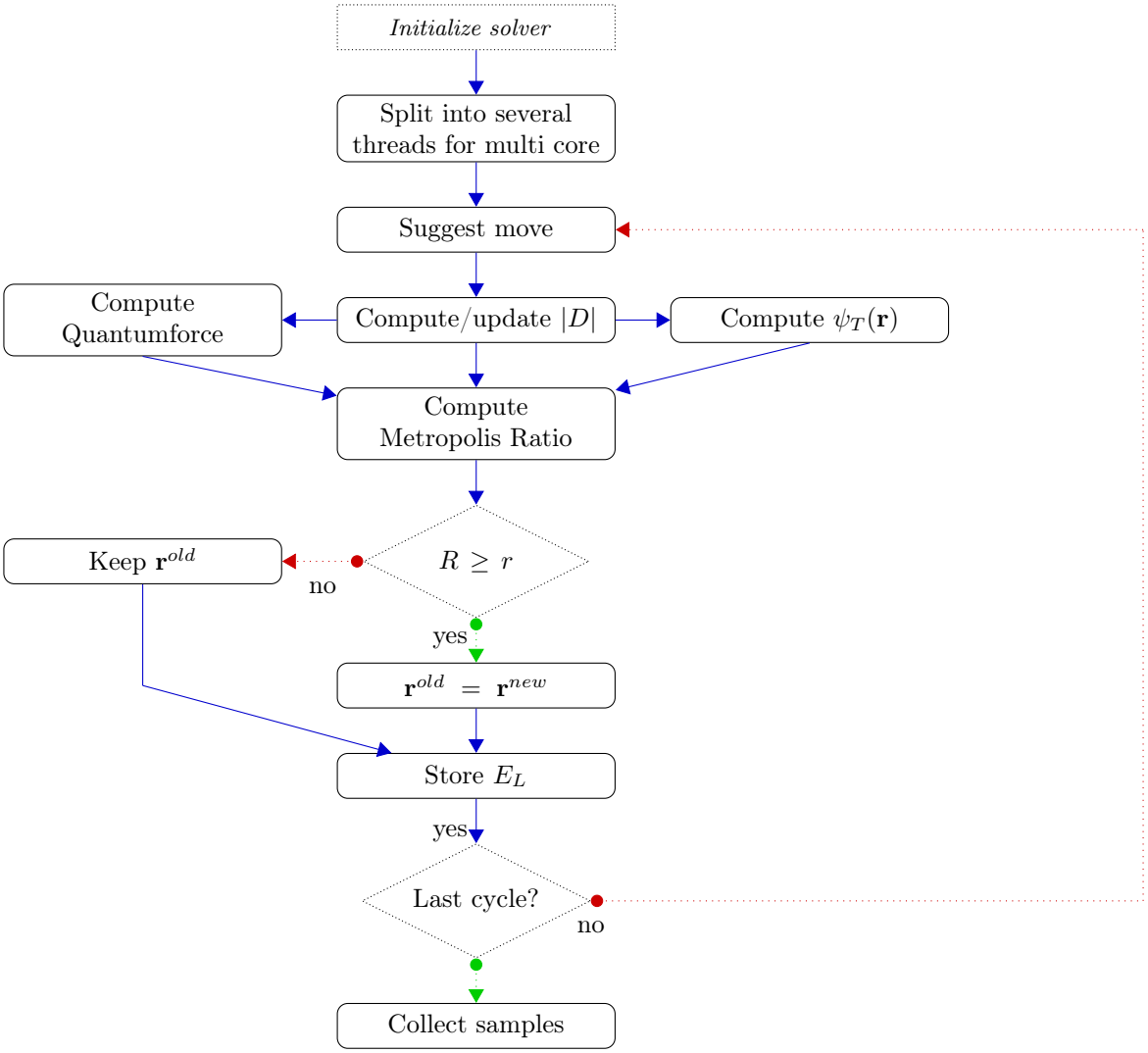

Figure 17: Schematic overview over the workflow of the VMC solver

A Program overview

A.1 The Solver

The variational monte carlo program can either be used directly as a cpp program, or with a python run.py script included in the folder ./python/ which calls the main cpp program with specific setting, the spesific settings is included in the README file. When running the VMC program it is ran with several parameters specified; number of processors to use, cycles to run, trialfunction to use and what it should do. There are several different options included as what it can do e.g. running several runs searching for te best α and β values, doing a long run writing down positions and several other options. The solver has several subclasses containing trialfunctions, derivatives and determinants. See Fig. 17 and Fig. 18 for an overview over the program.

A.2 Class structure

A.3 Python Programs

There is included several different python programs to help with processing the data, collected with the VMC solver, which is located in the python folder. There is several plotting scripts there and a script to calculate part of the derivatives of the wavefunctions.

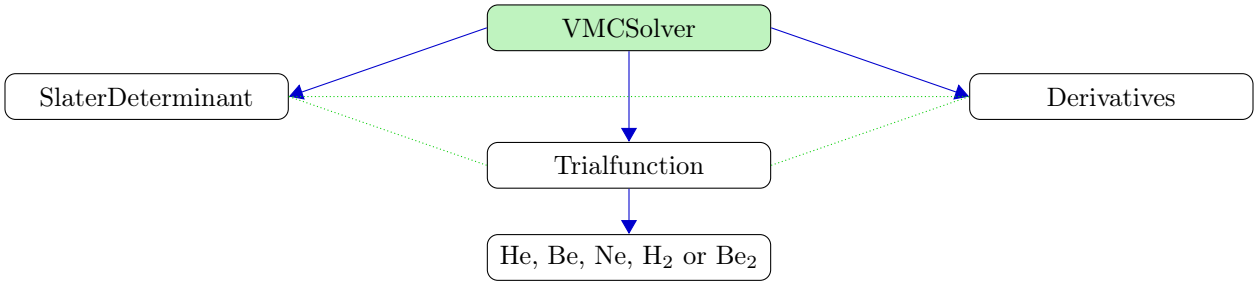

Figure 18: Class and subclass structure used by the program

B Closed expression for noncorrelation Helium trialfunction

B.1 Derivation of local energies, using radial coordinates

The local energy of is dependant on the Hamiltonian and the wavefunction describing the system, the Hamiltonian incorporates both a kinetic energy part given by $\frac{\nabla_i^2}{2}$ for each particle and a potential energy part given by $\frac{Z}{r_i}$ and $\frac{1}{r_{ij}}$, where Z is the charge of the center, r_i is the distance for electron i to the atom center and r_{ij} is the distance between electron l and m. Then the local energy is given by the following:

$$E_L = \sum_{i,i<j} \frac{1}{\Psi_T(\mathbf{r_i}, \mathbf{r_{ij}})} \hat{H} \Psi_T(\mathbf{r_i}, \mathbf{r_{ij}}) \tag{104}$$

$$= \sum_{i,i<j} \frac{1}{\Psi_T(\mathbf{r_i}, \mathbf{r_{ij}})} \left(-\frac{\nabla_i^2}{2} - \frac{Z}{r_i} - \frac{Z}{r_j} + \frac{1}{r_{ij}} \right) \Psi_T(\mathbf{r_i}, \mathbf{r_{ij}}) \tag{105}$$

$$= \sum_{i,i<j} -\frac{1}{2\Psi_T} \left(\nabla_i^2 \Psi_T \right) - \frac{Z}{r_i} - \frac{Z}{r_j} + \frac{1}{r_{ij}} \tag{106}$$

Let us change derivation variables:

$$-\frac{1}{2\Psi_T} \left(\nabla_i^2 \Psi_T \right) = \sum_{m=1}^{3} -\frac{1}{2\Psi_T} \left(\frac{\partial^2 \Psi_T}{\partial x_m^2} \right)_i \tag{107}$$

$$= \sum_{m=1}^{3} -\frac{1}{2\Psi_T} \left(\frac{\partial}{\partial x_m} \left(\frac{\partial \Psi_T}{\partial r_i} \frac{\partial r_i}{\partial x_m} \right) \right)_i \tag{108}$$

Since $r_i = \left(x_1^2 + x_2^2 + x_3^2 \right)^{1/2}$ then $\frac{\partial r_i}{\partial x_m} = \frac{\partial \left(x_1^2 + x_2^2 + x_3^2 \right)^{1/2}}{\partial x_m} = \frac{x_m}{r_i}$

$$= \sum_{m=1}^{3} -\frac{1}{2\Psi_T} \left(\frac{\partial}{\partial x_m} \left(\frac{\partial \Psi_T}{\partial r_i} \frac{x_m}{r_i} \right) \right)_i \tag{109}$$

$$= \sum_{m=1}^{3} -\frac{1}{2\Psi_T} \left(\frac{\partial^2 \Psi_T}{\partial x_m \partial r_i} \frac{x_m}{r_i} + \frac{\partial \Psi_T}{\partial r_i} \frac{\partial}{\partial x_m} \left(\frac{x_m}{r_i} \right) \right)_i \tag{110}$$

The term $\frac{\partial}{\partial x_m} \left(\frac{x_m}{r_i} \right)$ becomes for the different values for m, $\frac{\partial}{\partial x_1} \left(\frac{x_1}{\left(x_1^2 + x_2^2 + x_3^2 \right)^{1/2}} \right) = \frac{x_2^2 + x_3^2}{r_i^3}$ so all the values for m term it should sum up to $\frac{2(x_1^2 + x_2^2 + x_3^2)}{r_i^3}$

$$= -\frac{1}{2\Psi_T} \left(\frac{\partial^2 \Psi_T}{\partial r_i^2} \frac{x_1^2 + x_2^2 + x_3^2}{r_i^2} + \frac{\partial \Psi_T}{\partial r_i} \frac{2(x_1^2 + x_2^2 + x_3^2)}{r_i^3} \right)_i \tag{111}$$

$$= -\frac{1}{2\Psi_T} \left(\frac{\partial^2 \Psi_T}{\partial r_i^2} + \frac{\partial \Psi_T}{\partial r_i} \frac{2}{r_i} \right) \tag{112}$$

Then the local energy becomes:

$$E_L = \sum_{i,i<j} -\frac{1}{2\Psi_T} \left(\frac{\partial^2 \Psi_T}{\partial r_i^2} + \frac{\partial \Psi_T}{\partial r_i} \frac{2}{r_i} \right) - \frac{Z}{r_i} - \frac{Z}{r_j} + \frac{1}{r_{ij}} \tag{113}$$

We can apply this to the simple helium trialfunction with no electronic interaction to obtain the local energy.

1s
0.4579
0.6573

(a) Helium

1s	2s
-9.9281e-01	-2.1571e-01
-7.6425e-02	2.2934e-01
2.8727e-02	8.2235e-01
1.2898e-16	5.1721e-16
-2.3257e-19	4.5670e-18
5.6097e-19	-1.1040e-17
1.2016e-16	8.5306e-16
-4.6874e-19	7.0721e-18
1.1319e-18	-1.7060e-17

(b) Beryllium

1s	2s	$2p_x$	$2p_y$	$2p_z$
-9.8077e-01	-2.6062e-01	1.1596e-16	-8.3716e-18	-1.9554e-17
-9.3714e-02	2.5858e-01	-2.0106e-16	-9.7173e-17	-7.3738e-17
2.2863e-02	8.1619e-01	-3.2361e-16	1.3237e-16	1.5789e-16
-9.9519e-19	-5.6186e-18	2.7155e-02	-4.0320e-01	3.9171e-01
-1.2125e-18	-2.8615e-16	-5.6207e-01	-2.5833e-02	1.2375e-02
-4.1800e-19	4.6199e-17	9.1139e-03	-3.9180e-01	-4.0392e-01
-1.6696e-19	-4.2405e-18	2.8890e-02	-4.2895e-01	4.1673e-01
1.2125e-18	-2.9426e-16	-5.9797e-01	-2.7482e-02	1.3166e-02
3.8779e-19	5.0519e-17	9.6959e-03	-4.1683e-01	-4.2972e-01

(c) Neon

Table 7: Constants for combining contracted GTOs for Helium, Beryllium and Neon.

B.1.1 Helium: Simple trialfunction

The simple version of the trial function is only dependant on one parameter α and does not take into account interaction between the two electrons, it is of the form

$$\Psi_T(\mathbf{r_1}, \mathbf{r_2}) = \exp\{-\alpha(r_1 + r_2)\}$$

Let us set this trialfunction into the equation for the local energy (113).

$$E_L = \sum_{i,i<j} -\frac{1}{2\Psi_T}\left(\frac{\partial^2 e^{-\alpha(r_i+r_j)}}{\partial r_i^2} + \frac{\partial e^{-\alpha(r_i+r_j)}}{\partial r_i}\frac{2}{r_i}\right) - \frac{Z}{r_i} - \frac{Z}{r_j} + \frac{1}{r_{ij}} \tag{114}$$

$$E_L = -\frac{1}{2\Psi_T}\sum_{i=1}^{2}\left(\alpha^2 - \alpha\frac{2}{r_i}\right)\Psi_T - \frac{Z}{r_i} + \frac{1}{r_{ij}} \tag{115}$$

$$E_L = -\alpha^2 + (\alpha - Z)\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{1}{r_{12}} \tag{116}$$

C GTO tables

D Verification of the model

To ensure that the program produces valid results as it gets more complicated we have implemented several different tests of smaller parts of the program that all should be met if the VMC solver is functioning properly. This section consists of a list of the different tests.

D.1 Verification of the general Monte Carlo method

Since the wavefunction for a Hydrogen atom can be calculated analytically, see equation (117), a Monte Carlo calculation with that wavefunction should return an exact value for the energy.

Atom	Energy
Hydrogen	$E_{min} = -\frac{1}{2}$
Helium	$E_{min} = -4$
Beryllium	$E_{min} = -20$
Neon	$E_{min} = -200$

Table 8: Ground states for the different atoms without electron-electron interaction

$$\Psi(\rho) = \alpha\rho e^{-\alpha\rho} \qquad \text{With local energy:} \qquad E_L(\rho) = -\frac{1}{\rho} - \frac{\alpha}{2}\left(\alpha - \frac{2}{\rho}\right) \tag{117}$$

So a VMC calculation runrun with a Hydrogen atom with $\alpha = 1$ it should produce an energy of exactly -0.5 with 0 variance.

D.2 Verification of the Slater determinant and the laplacian Slaterdeterminant ratio

To verify the Slater determinant part of the trialfunctions we consider the atoms without any electron-electron interactions. Then it is a one-body system like hydrogen and can be calculated analytically, see Griffiths, 2005, and we get exact wavefunctions and energy with 0 variance.

In this case, where the correlation derivatives disappear, the kinetic part of the local energy gets simplified, from (59), to the following

$$\frac{\nabla^2\Psi_T}{\Psi_T} = \frac{\nabla^2|D_\uparrow|}{|D_\uparrow|} + \frac{\nabla^2|D_\downarrow|}{|D_\downarrow|}$$

To reproduce the correct results both the Slater determinant and it's laplacian needs to be correct.

D.3 The gradient

The gradient is used in the calculation of the quantum force, and all it's components is used in the calculation of the so by testing that this is correctly reproduced we get an inclination several terms are correct, $\frac{\nabla|D_\uparrow|}{|D_\uparrow|}$, $\frac{\nabla|D_\downarrow|}{|D_\downarrow|}$ and $\frac{\nabla\Psi_C}{\Psi_C}$.

The wavefunctions should be correct due to the earlier test, D.2, so a numerical derivation of the trialfunction should produce a correct gradient, which is then used to test the analytical version against.

D.4 The correlation laplacian

Due to the earlier two tests, D.2 and D.3, we can be fairly certain that all the terms except the laplacian-correlation-ratio, $\frac{\nabla^2\Psi_C}{\Psi_C}$ in the local energy equation (59) is correct. So by then comparing the analytical version of the local energy to the numerical version we can verify that the last term is also correct.

D.5 Local Energy in Helium

For Helium we have a complete closed expression for the local energy, E_l (116), we use this to check that the local energy calculation, see "Efficient calculation of stuff chapters"!!!!!!!!!, used on the more complicated atoms also replicates the local energy for the simpler atom which it does.

D.6 Verification of correlation gradient

The correlation gradient ratio is checked by calculating what calculating it directly for Helium, and then comparing this value against value produced by the program.

Let us consider the gradient ratio of the Padè-Jastrow factor in Helium, $\frac{\nabla\psi_C(\mathbf{r}_{12})}{\psi_C(\mathbf{r}_{12})}$ with $\psi_C(\mathbf{r_{12}}) = e^{\frac{r_{12}}{2(1+\beta r_{12})}}$. Using the results from equation (71) on Helium for the first electron we get

$$\left[\frac{\nabla\Psi_C}{\Psi_C}\right]_1 = \frac{1}{\Psi_c}\frac{\partial\Psi_C}{\partial x_k} = \frac{\mathbf{r_{12}}}{r_{12}}\frac{\partial}{\partial r_{12}}\left(\frac{r_{12}}{2(1+\beta r_{12})}\right) - \frac{\mathbf{r_{21}}}{r_{21}}\frac{\partial}{\partial r_{21}}\left(\frac{r_{21}}{2(1+\beta r_{21})}\right) \tag{118}$$

$$= 2\frac{\mathbf{r_{12}}}{r_{12}}\frac{\partial}{\partial r_{12}}\left(\frac{r_{12}}{2(1+\beta r_{12})}\right) \tag{119}$$

$$= \frac{\mathbf{r_{12}}}{r_{12}}\frac{1}{(1+\beta r_{12})^2} \tag{120}$$

Testing that the program reproduces this for the helium atom indicates the $\frac{\nabla\Psi_C}{\Psi_C}$ is being calculated correctly.

References

Binkley, J. S. and J. A. Pople (1975). "Møller–Plesset theory for atomic ground state energies". In: *International Journal of Quantum Chemistry* 9.2, pp. 229–236. ISSN: 1097-461X. DOI: 10.1002/qua.560090204. URL: http://dx.doi.org/10.1002/qua.560090204.

Boys, S. F. (1950). "Electronic Wave Functions. I. A General Method of Calculation for the Stationary States of Any Molecular System". In: *Proceedings of the Royal Society of London. Series A* 200.1063, pp. 542–554.

EMSL Basis Set Exchange. https://bse.pnl.gov/bse/portal. Accessed: 2015-05-13.

G. V. Killie, H. S. B. M. and J. E. R. Navarro. *"Git address for the code"*. URL: https://github.com/hakonsbm/CompPhys2_office306.

Griffiths, D. J. (2005). *Introduction to Quantum Mechanics*. Pearson.

Hjorth-Jensen, M. (2013). *Computational Physics: Lecture notes*.

Ihle, H. T. and L. Ledum (2013). "VMC-calculations of atoms and simple molecules." In:

Introduction to Quantum Mechanics by Pauling, Linus; Wilson, E.Bright (1935). *Introduction to Quantum Mechanics: McGraw-Hill Inc.,US 9780070489608*.

J. S. Binkley J. A. Pople, W. J. H. (1980). "Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements". In: *J. Am. Chem. Soc* 102, pp. 939–947. DOI: 10.1021/ja00523a008.

Koput, J. (2011). "The ground-state potential energy function of a beryllium dimer determined using the single-reference coupled-cluster approach". In: URL: http://www.researchgate.net/profile/Marco_Nascimento/publication/229696100_Ground_state_of_the_beryllium_atom_Reinvestigation_based_on_a_proper_independent_particle_model/links/0c96051d5d671b5eef000000.pdf.