FYS-4411: Computational Physics II
Project 2

Gullik Vetvik Killie
Håkon Sebastian Bakke Mørk
Jose Emilio Ruiz Navarro

April 22, 2015

Abstract

By using Variational Monte Carlo methods with importance sampling, blocking and MPI we find the ground state energies of Beryllium and Neon. We also calculate the onebody density with and without the Jastrow factor. SOME RESULTS

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 create a lot of problems if the search range for the parameters is illy defined and is not close enough to the variational minimum, namely, the results will have a poor quality if this happens. 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.

2 Methods

2.1 Monte Carlo of the Helium Atom

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.

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.

2.2 Beryllium atom

It is fairly simple to extend the calculational 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 calculate a Slater determinant. However the computation of the Slater determinant can be simplified for beryllium. Sticking to hydrogen-like wave functions, we can write the trial wave function for beryllium as

$$\psi_T(\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}, \mathbf{r_4}) = Det\left(\phi_1(\mathbf{r_1}), \phi_2(\mathbf{r_2}), \phi_3(\mathbf{r_3}), \phi_4(\mathbf{r_4})\right) \prod_{i<j}^{4} \exp\left\{ \left(\frac{r_{ij}}{2(1+\beta r_{ij})} \right) \right\}, \tag{6}$$

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

$$\phi_{1s}(\mathbf{r_i}) = e^{-\alpha r_i}, \tag{7}$$

and

$$\phi_{2s}(\mathbf{r_i}) = (1 - \alpha r_i/2)\, e^{-\alpha r_i/2}. \tag{8}$$

The Slater determinant is calculated using these ansatzes.

Furthermore, for the Jastrow factor,

$$\Psi_C = \prod_{i<j} g(r_{ij}) = \exp\left\{ \sum_{i<j} \frac{a r_{ij}}{1 + \beta r_{ij}} \right\}, \tag{9}$$

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

$$a = \frac{1}{4}, \tag{10}$$

and if they have opposite spins we get a factor

$$a = \frac{1}{2}. \tag{11}$$

2.3 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. Therefore we also have to implement a better way to handle the Slater determinant than we did in the previous project. The trial wave function for Neon can be written as

$$\psi_T(\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_{10}}) = Det\left(\phi_1(\mathbf{r_1}), \phi_2(\mathbf{r_2}), \ldots, \phi_{10}(\mathbf{r_{10}})\right) \prod_{i<j}^{10} \exp\left\{ \left(\frac{r_{ij}}{2(1 + \beta r_{ij})} \right) \right\}, \tag{12}$$

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

$$\phi_{2p}(\mathbf{r_i}) = \alpha \mathbf{r_i} e^{-\alpha r_i/2}. \tag{13}$$

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

2.4 Derivation of local energies

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:

$$E_L = \sum_{i,i<j} \frac{1}{\Psi_T(\mathbf{r_i}, \mathbf{r_{ij}})} \hat{H} \Psi_T(\mathbf{r_i}, \mathbf{r_{ij}}) \tag{14}$$

$$= \sum_{i,i<j} \frac{1}{\Psi_T(\mathbf{r_i}, \mathbf{r_{ij}})} \left(-\frac{\nabla_i^2}{2} - \frac{Z}{r_i} - \frac{Z}{r_j} + \frac{1}{r_{ij}} \right) \Psi_T(\mathbf{r_i}, \mathbf{r_{ij}}) \tag{15}$$

$$= \sum_{i,i<j} -\frac{1}{2\Psi_T} \left(\nabla_i^2 \Psi_T \right) - \frac{Z}{r_i} - \frac{Z}{r_j} + \frac{1}{r_{ij}} \tag{16}$$

Let us change derivation variables:

$$-\frac{1}{2\Psi_T}\left(\nabla_i^2\Psi_T\right) = \sum_{m=1}^{3} -\frac{1}{2\Psi_T}\left(\frac{\partial^2\Psi_T}{\partial x_m^2}\right)_i \tag{17}$$

$$= \sum_{m=1}^{3} -\frac{1}{2\Psi_T}\left(\frac{\partial}{\partial x_m}\left(\frac{\partial\Psi_T}{\partial r_i}\frac{\partial r_i}{\partial x_m}\right)\right)_i \tag{18}$$

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

$$= \sum_{m=1}^{3} -\frac{1}{2\Psi_T}\left(\frac{\partial}{\partial x_m}\left(\frac{\partial\Psi_T}{\partial r_i}\frac{x_m}{r_i}\right)\right)_i \tag{19}$$

$$= \sum_{m=1}^{3} -\frac{1}{2\Psi_T}\left(\frac{\partial^2\Psi_T}{\partial x_m\partial r_i}\frac{x_m}{r_i} + \frac{\partial\Psi_T}{\partial r_i}\frac{\partial}{\partial x_m}\left(\frac{x_m}{r_i}\right)\right)_i \tag{20}$$

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 $\frac{2(x_1^2+x_2^2+x_3^2)}{r_i^3}$

$$= -\frac{1}{2\Psi_T}\left(\frac{\partial^2\Psi_T}{\partial r_i^2}\frac{x_1^2+x_2^2+x_3^2}{r_i^2} + \frac{\partial\Psi_T}{\partial r_i}\frac{2(x_1^2+x_2^2+x_3^2)}{r_i^3}\right)_i \tag{21}$$

$$= -\frac{1}{2\Psi_T}\left(\frac{\partial^2\Psi_T}{\partial r_i^2} + \frac{\partial\Psi_T}{\partial r_i}\frac{2}{r_i}\right) \tag{22}$$

Then the local energy becomes:

$$E_L = \sum_{i,i<j} -\frac{1}{2\Psi_T}\left(\frac{\partial^2\Psi_T}{\partial r_i^2} + \frac{\partial\Psi_T}{\partial r_i}\frac{2}{r_i}\right) - \frac{Z}{r_i} - \frac{Z}{r_j} + \frac{1}{r_{ij}} \tag{23}$$

2.5 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

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - F\right)P(x,t), \tag{24}$$

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

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta \tag{25}$$

whereη is a random variable. This gives us a new position

$$y = x + DF(x)\Delta t + \xi\sqrt{\Delta t}. \tag{26}$$

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

$$\frac{\partial P}{\partial t} = \sum_i D \frac{\partial}{\partial \mathbf{x_i}} \left(\frac{\partial}{\partial \mathbf{x_i}} - \mathbf{F_i} \right) P(\mathbf{x}, t), \tag{27}$$

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

$$\frac{\partial^2 P}{\partial \mathbf{x_i}^2} = P \frac{\partial}{\partial \mathbf{x_i}} \mathbf{F_i} + \mathbf{F_i} \frac{\partial}{\partial \mathbf{x_i}} P. \tag{28}$$

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

$$\frac{\partial^2 P}{\partial \mathbf{x_i}^2} = P \frac{\partial g}{\partial P} \left(\frac{\partial P}{\partial \mathbf{x}_i} \right)^2 + P g \frac{\partial^2 P}{\partial \mathbf{x}_i^2} + g \left(\frac{\partial P}{\partial \mathbf{x}_i} \right)^2 \tag{29}$$

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

$$\mathbf{F} = 2 \frac{1}{\Psi_T} \nabla \Psi_T, \tag{30}$$

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

$$G(y, x, \Delta t) = \frac{1}{(4\pi D \Delta t)^{3N/2}} \exp\left\{ \left(-(y - x - D\Delta t F(x))^2 / 4D\Delta t \right) \right\} \tag{31}$$

This means that the Metropolis algorithm

$$A(y, x) = \min(1, q(y, x))), \tag{32}$$

where

$$q(y, x) = \frac{|\Psi_T(y)|^2}{|\Psi_T(x)|^2}, \tag{33}$$

is replaced by the Metropolis-Hastings algorithm,

$$q(y, x) = \frac{G(x, y, \Delta t) |\Psi_T(y)|^2}{G(y, x, \Delta t) |\Psi_T(x)|^2} \tag{34}$$

2.6 Blocking

Blocking refers to a method to more accurately estimate the error of the values obtained by the VMC algorithm. It's actually a completely separate and independent process. 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* measurementes 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.

$$f_d = \frac{1}{n-d} \sum_{k=1}^{n-d} \left(x_k - \bar{x}_n\right)\left(x_{k+d} - \bar{x}_n\right) \tag{35}$$

This can be used to give an actual form to the correction factor:

$$\tau 1 + 2 \sum_{d=1}^{n-1} \frac{f_d}{var\left(x\right)} \tag{36}$$

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

$$err^2 = \frac{\tau}{n} var\left(x\right) \tag{37}$$

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

$$n_{eff} = \frac{n}{\tau} \tag{38}$$

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

$$\sigma = \sqrt{\left(\frac{1 + 2\tau/\Delta t}{n}\left(\bar{x^2} - \bar{x}^2\right)\right)} \tag{39}$$

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 actually quite simple. The total amount of measurements is divided into blocks (hence the name) 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.

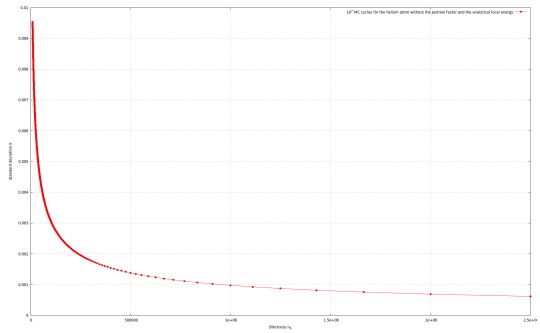

Figure 1: STD vs. blocksize. Here we see the plateuing trend of the STD as the blocksize increases, but the problem is that as the blocksize decreases it should decrease as well and not increase. This problem will be fixed in the future.

2.7 Calculating the Slater determinant

2.7.1 Setting up the Slater determinant

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

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \alpha, \beta, \gamma, \delta) = \frac{1}{\sqrt{4!}} \begin{vmatrix} \psi_{100\uparrow}(\mathbf{r}_1) & \psi_{100\uparrow}(\mathbf{r}_2) & \psi_{100\uparrow}(\mathbf{r}_3) & \psi_{100\uparrow}(\mathbf{r}_4) \\ \psi_{100\downarrow}(\mathbf{r}_1) & \psi_{100\downarrow}(\mathbf{r}_2) & \psi_{100\downarrow}(\mathbf{r}_3) & \psi_{100\downarrow}(\mathbf{r}_4) \\ \psi_{200\uparrow}(\mathbf{r}_1) & \psi_{200\uparrow}(\mathbf{r}_2) & \psi_{200\uparrow}(\mathbf{r}_3) & \psi_{200\uparrow}(\mathbf{r}_4) \\ \psi_{200\downarrow}(\mathbf{r}_1) & \psi_{200\downarrow}(\mathbf{r}_2) & \psi_{200\downarrow}(\mathbf{r}_3) & \psi_{200\downarrow}(\mathbf{r}_4) \end{vmatrix} \tag{40}$$

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

$$\det\uparrow(1,2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{100\uparrow}(\mathbf{r}_1) & \psi_{100\uparrow}(\mathbf{r}_2) \\ \psi_{200\uparrow}(\mathbf{r}_1) & \psi_{200\uparrow}(\mathbf{r}_2) \end{vmatrix} \tag{41}$$

and the Slater determinant for spin down as

$$\det\downarrow(3,4) = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{100\downarrow}(\mathbf{r}_3) & \psi_{100\downarrow}(\mathbf{r}_4) \\ \psi_{200\downarrow}(\mathbf{r}_3) & \psi_{200\downarrow}(\mathbf{r}_4) \end{vmatrix} \tag{42}$$

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

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N) \propto \det\uparrow \det\downarrow \tag{43}$$

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.

2.7.2 Calculation of the Slater determinant

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

$$|\hat{A}| = c_{11}c_{22}\dots c_{nn} \tag{44}$$

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

2.8 Verification of the model

2.9 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 we have found divided up the derivatives and found analytic expressions for all the parts.

The trialfunction can be factorized as

$$\Psi_T(\mathbf{x}) = \Psi_D\Psi_C = |D_\uparrow||D_\downarrow|\Psi_C \tag{45}$$

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

2.9.1 Gradient ratio

For quantum force we need to calculate the gradient ratio of the trialfunction which is given by

$$\frac{\nabla\Psi_T}{\Psi_T} = \frac{\nabla(\Psi_D\Psi_C)}{\Psi_D\Psi_C} = \frac{\nabla\Psi_D}{\Psi_D} + \frac{\nabla\Psi_C}{\Psi_C} \tag{46}$$

$$= \frac{\nabla|D_\uparrow|}{|D_\uparrow|} + \frac{\nabla|D_\downarrow|}{|D_\downarrow|} + \frac{\nabla\Psi_C}{\Psi_C} \tag{47}$$

2.9.2 Kinetic Energy

From the Hamiltonian the expectation value of kinetic energy for each electron is given by

$$K_i = -\frac{1}{2}\frac{\nabla_i^2\Psi}{\Psi} \tag{48}$$

Using the factorization of of the trialfunction from (45) we can calculate the ratio needed for the kinetic energy.

$$\frac{1}{\Psi_T}\frac{\partial^2\Psi_T}{\partial x_k^2} = \frac{1}{\Psi_D\Psi_C}\frac{\partial^2(\Psi_D\Psi_C)}{\partial x_k^2} = \frac{1}{\Psi_D\Psi_C}\frac{\partial}{\partial x_k}\left(\frac{\partial\Psi_D}{\partial x_k}\Psi_C + \Psi_D\frac{\partial\Psi_C}{\partial x_k}\right) \tag{49}$$

$$= \frac{1}{\Psi_D\Psi_C}\left(\frac{\partial^2\Psi_D}{\partial x_k^2}\Psi_C + 2\frac{\partial\Psi_D}{\partial x_k}\frac{\partial\Psi_C}{\partial x_k} + \Psi_D\frac{\partial^2\Psi_C}{\partial x_k^2}\right) \tag{50}$$

$$= \frac{1}{\Psi_D}\frac{\partial^2\Psi_D}{\partial x_k^2} + 2\frac{1}{\Psi_D}\frac{\partial\Psi_D}{\partial x_k}\cdot\frac{1}{\Psi_C}\frac{\partial\Psi_C}{\partial x_k} + \frac{1}{\Psi_C}\frac{\partial^2\Psi_C}{\partial x_k^2} \tag{51}$$

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

$$\frac{1}{\Psi_D}\frac{\partial^2\Psi_D}{\partial x_k^2} = \frac{1}{|D_\uparrow||D_\downarrow|}\frac{\partial^2|D_\uparrow||D_\downarrow|}{\partial x_k^2} = \frac{1}{|D_\uparrow|}\frac{\partial^2|D_\uparrow|}{\partial x_k^2} + \frac{1}{|D_\downarrow|}\frac{\partial^2|D_\downarrow|}{\partial x_k^2} \tag{52}$$

$$\frac{1}{\Psi_D}\frac{\partial\Psi_D}{\partial x_k} = \frac{1}{|D_\uparrow||D_\downarrow|}\frac{\partial|D_\uparrow||D_\downarrow|}{\partial x_k} = \frac{1}{|D_\uparrow|}\frac{\partial|D_\uparrow|}{\partial x_k} + \frac{1}{|D_\downarrow|}\frac{\partial|D_\downarrow|}{\partial x_k} \tag{53}$$

Inserting equations (53) and (52) into (51) we get

$$\frac{\nabla^2\Psi_T}{\Psi_T} = \frac{\nabla^2|D_\uparrow|}{|D_\uparrow|} + \frac{\nabla^2|D_\downarrow|}{|D_\downarrow|} + 2\left(\frac{\nabla|D_\uparrow|}{|D_\uparrow|} + \frac{\nabla|D_\downarrow|}{|D_\downarrow|}\right)\cdot\frac{\nabla\Psi_C}{\Psi_C} + \frac{\nabla^2\Psi_C}{\Psi_C} \tag{54}$$

Now we have 4 different types of ratios we need to find an expression for $\frac{\nabla^2|D|}{|D|}$, $\frac{\nabla|D|}{|D|}$, $\frac{\nabla^2\Psi_C}{\Psi_C}$ and $\frac{\nabla\Psi_C}{\Psi_C}$ to calculate both the gradient and Laplacian ratios of the wavefunction.

2.9.3 Determinant ratios

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

$$D_{ij} = \phi_j(\mathbf{r}_i) \tag{55}$$

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

$$|D| = \frac{\mathbf{D}^T}{\mathbf{D^{-1}}} = \sum_{j=1}^{N}\frac{C_{ji}}{D_{ij}^{-1}} = \sum_{j=1}^{N} D_{ij}C_{ji} \tag{56}$$

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

$$R_{SD} = \frac{|\mathbf{D}^{new}|}{|\mathbf{D}^{old}|} = \frac{\sum_{j=0}^{N} D_{ij}^{new}C_{ji}^{new}}{\sum_{j=0}^{N} D_{ij}^{old}C_{ji}^{old}} \tag{57}$$

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 (56) we get

$$R_{SD} = \frac{\sum_{j=0}^{N} D_{ij}^{new} (D_{ji}^{old})^{-1} |D^{old}|}{\sum_{j=0}^{N} D_{ij}^{old} (D_{ji}^{old})^{-1} |D^{old}|} \tag{58}$$

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

$$R_{SD} = \sum_{j=0}^{N} D_{ij}^{new} (D_{ji}^{old})^{-1} = \sum_{j=0}^{N} \phi_j(\mathbf{x}_i^{new}) D_{ji}^{-1}(\mathbf{x}^{old}) \tag{59}$$

2.9.4 Gradient determinant Ratio

2.9.5 Derivatives of single particle wavefunctions

Calculated in derivatives.py.

	ψ_i	$\frac{\partial \psi_i}{\partial r_i}$	$\frac{\partial^2 \psi_i}{\partial r_i^2}$
ψ_{1S}	$e^{-\alpha ri}$	$-\frac{\alpha}{r_i}\left(x_i + y_i + z_i\right) e^{-\alpha r_i}$	$\frac{\alpha}{r_i}\left(\alpha r_i - 2\right) e^{-\alpha r_i}$
ψ_{2S}	$\left(-\frac{\alpha r_i}{2} + 1\right) e^{-\frac{\alpha r_i}{2}}$	$\frac{\alpha e^{-\frac{\alpha r_i}{2}}}{4r_i}\left(\alpha r_i - 4\right)\left(x_i + y_i + z_i\right)$	$-\frac{\alpha e^{-\frac{\alpha r_i}{2}}}{8r_i}\left(\alpha^2 r_i^2 - 10\alpha r_i + 16\right)$
ψ_{2P}	$\alpha x_i e^{-\frac{\alpha r_i}{2}}$	$-\frac{\alpha e^{-\frac{\alpha r_i}{2}}}{2r_i}\left(\alpha x_i^2 + \alpha x_i y_i + \alpha x_i z_i - 2r_i\right)$	$\frac{\alpha^2 x_i}{4r_i}\left(\alpha r_i - 8\right) e^{-\frac{\alpha r_i}{2}}$

2.9.6 Gradient ratio of Padé-Jastrow factor

When derivating the Padé-Jastrow factor all the factors not involving the particle we are derivating with respect to will be canceled by the corresponding terms in the denominator.

$$\frac{1}{\Psi_C}\frac{\partial \Psi_C}{\partial x_k} = \sum_{i=1}^{k-1} \frac{1}{g_{ik}} \frac{\partial g_{ik}}{\partial x_k} + \sum_{i=k+1}^{N} \frac{1}{g_{ki}} \frac{\partial g_{ki}}{\partial x_k} \tag{60}$$

2.9.7 Correlation-to-correlation ratio

We have $N\left(N-1\right)/2$ 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 $g = g\left(r_{ij}\right)$.

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

$$R_C = \frac{\Psi_C^{new}}{\Psi_C^{cur}} = \prod_{i=1}^{k-1} \frac{g_{ik}^{new}}{g_{ik}^{cur}} \prod_{i=k+1}^{N} \frac{g_{ki}^{new}}{g_{ki}^{cur}} \tag{61}$$

or in the Padé-Jastrow form

$$R_C = \frac{\Psi_C^{\text{new}}}{\Psi_C^{\text{cur}}} = \frac{\exp U_{new}}{\exp U_{cur}} = \exp \Delta U \tag{62}$$

where

$$\Delta U = \sum_{i=1}^{k-1} \left(f_{ik}^{\text{new}} - f_{ik}^{\text{cur}} \right) + \sum_{i=k+1}^{N} \left(f_{ki}^{\text{new}} - f_{ki}^{\text{cur}} \right) \tag{63}$$

2.9.8 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

$$\frac{\nabla_i \Psi_C}{\Psi_C} = \frac{1}{\Psi_C} \frac{\partial \Psi_C}{\partial x_i} \tag{64}$$

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

$$\frac{1}{\Psi_C} \frac{\partial \Psi_C}{\partial x_k} = \sum_{i=1}^{k-1} \frac{1}{g_{ik}} \frac{\partial g_{ik}}{\partial x_k} + \sum_{i=k+1}^{N} \frac{1}{g_{ki}} \frac{\partial g_{ki}}{\partial x_k} \tag{65}$$

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

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

We now use the identity

$$\frac{\partial}{\partial x_i} g_{ij} = -\frac{\partial}{\partial x_j} g_{ij} \tag{67}$$

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

$$\frac{1}{\Psi_C} \frac{\partial \Psi_C}{\partial x_k} = \sum_{i=1}^{k-1} \frac{1}{g_{ik}} \frac{\partial g_{ik}}{\partial x_k} - \sum_{i=k+1}^{N} \frac{1}{g_{ki}} \frac{\partial g_{ki}}{\partial x_i} \tag{68}$$

and for the exponential case

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

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

$$\frac{\partial g_{ij}}{\partial x_j} = \frac{\partial g_{ij}}{\partial r_{ij}}\frac{\partial r_{ij}}{\partial x_j} = \frac{x_j - x_i}{r_{ij}}\frac{\partial g_{ij}}{\partial r_{ij}} \tag{70}$$

After substitution we get

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

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

$$\frac{\partial g_{ij}}{\partial r_{ij}} = g_{ij}\frac{\partial f_{ij}}{\partial r_{ij}} \tag{72}$$

and arrive at

$$\frac{1}{\Psi_C}\frac{\partial \Psi_C}{\partial x_k} = \sum_{i=1}^{k-1}\frac{\mathbf{r_{ik}}}{r_{ik}}\frac{\partial f_{ik}}{\partial r_{ik}} - \sum_{i=k+1}^{N}\frac{\mathbf{r_{ki}}}{r_{ki}}\frac{\partial f_{ki}}{\partial r_{ki}} \tag{73}$$

where we have the relative vectorial distance

$$\mathbf{r}_{ij} = |\mathbf{r}_j - \mathbf{r}_i| = (x_j - x_i)\mathbf{e}_1 + (y_j - y_i)\mathbf{e}_2 + (z_j - z_i)\mathbf{e}_3 \tag{74}$$

With a linear Padé-Jastrow we set

$$f_{ij} = \frac{ar_{ij}}{(1+\beta r_{ij})} \tag{75}$$

with the corresponding closed form expression

$$\frac{\partial f_{ij}}{\partial r_{ij}} = \frac{a}{(1+\beta r_{ij})^2} \tag{76}$$

2.9.9 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

$$\left[\frac{\nabla^2\Psi_C}{\Psi_C}\right]_x = 2\sum_{k=1}^{N}\sum_{i=1}^{k-1}\frac{\partial^2 g_{ik}}{\partial x_k^2} + \sum_{k=1}^{N}\left(\sum_{i=1}^{k-1}\frac{\partial g_{ik}}{\partial x_k} - \sum_{i=k+1}^{N}\frac{\partial g_{ki}}{\partial x_i}\right)^2 \tag{77}$$

But we have another, simpler form for the function

$$\Psi_C = \prod_{i<j}\exp f(r_{ij}) = \exp\left\{\sum_{i<j}\frac{ar_{ij}}{1+\beta r_{ij}}\right\} \tag{78}$$

and for particle k we have

Atom	VMC	References
Beryllium	-14.4127	-14.667
Neon	TODO	TODO

Table 1: Comparison of energies found with Variational Monte Carlo method and energies found in research papers.

$$\frac{\nabla_k^2 \Psi_C}{\Psi_C} = \sum_{ij\neq k} \frac{(\mathbf{r}_k - \mathbf{r}_i)(\mathbf{r}_k - \mathbf{r}_j)}{r_{ki} r_{kj}} f'(r_{ki}) f'(r_{kj}) + \sum_{j\neq k} \left(f''(r_{kj}) + \frac{2}{r_{kj}} f'(r_{kj}) \right) \tag{79}$$

We use

$$f(r_{ij}) = \frac{a r_{ij}}{1 + \beta r_{ij}} \tag{80}$$

and with

$$g'(r_{kj}) = dg(r_{kj})/dr_{kj} \qquad \text{and} \qquad g''(r_{kj}) = d^2 g(r_{kj})/dr_{kj}^2 \tag{81}$$

we find that for particle k we have

$$\frac{\nabla_k^2 \Psi_C}{\Psi_C} = \sum_{ij\neq k} \frac{(\mathbf{r}_k - \mathbf{r}_i)(\mathbf{r}_k - \mathbf{r}_j)}{r_{ki} r_{kj}} \frac{a}{(1+\beta r_{ki})^2} \frac{a}{(1+\beta r_{kj})^2} + \sum_{j\neq k} \left(\frac{2a}{r_{kj}(1+\beta r_{kj})^2} - \frac{2a\beta}{(1+\beta r_{kj})^3} \right) \tag{82}$$

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

$$\frac{\partial^2 f_{ij}}{\partial r_{ij}^2} = -\frac{2 a_{ij} \beta_{ij}}{(1 + \beta_{ij} r_{ij})^3} \tag{83}$$

2.10 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.

3 Results and discussion

Reference table 1.

Trialfunction	α	β
Beryllium ψ_{T2}	3.8	0.293
Neon ψ_{T2}	10.22	0.091

Table 2: The values for α and β where found by doing running Monte Carlo calculation with over a mesh of different α and β values. For Beryllium each Monte Carlo run went over 10000000, using importance sampling. For Neon each run went over 5000, also with importance sampling Then the run with the lowest energy gave the α and β values.

Num. of processes	1	2	3	4
Time to exec [s]	1.0	1.97	2.90	3.35

Table 3: MPI speedup

3.1 Alpha and Beta Values

We use Alpha and Beta values for Beryllium found in project 1, so the only Alpha and Beta values we have to find are for Neon. Table 2 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. FIGURES

3.2 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 3 and figure 2. We see that the speedup is good for 2 and 3 processes, but for 4 processes suffers somewhat because it also have to run the OS and other programs.

3.3 Onebody densities / Comparisons / Without Jastrow

4 Conclusions

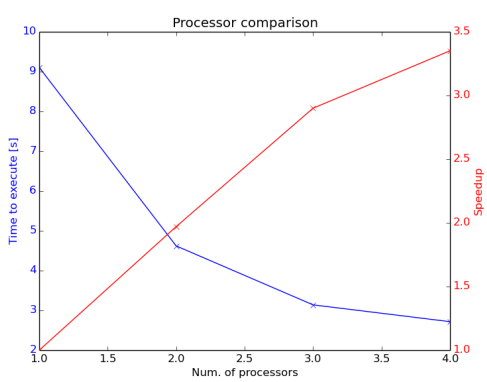

Figure 2: MPI speedup