Variational Monte-Carlo Simulations of Atomic Systems

Daniel Marelius Bjørnstad and Alexander Fleischer
https://github.com/lastis/FYS4411

Abstract

This is where the abstract goes. Write something smrt here.

I INTRODUCTION

To evaluate the results of the numerical methods, we investigated the possibility of finding a closed-form solution of the ground-state energy using the Variational Principle.

II METHODS

A. The Quantum Mechanical System

The Helium atom consists of two electrons orbiting a nucleus, where the distance between electron 1 and the nucleus, and electron 2 and the nucleus are labeled as $r_1 = \sqrt{x_1^2 + y_1^2 + z_1^2}$ and $r_2 = \sqrt{x_2^2 + y_2^2 + z_2^2}$ in cartesian coordinates.

The total potential energy of the system is modelled as

$$V(r_1, r_2) = -\frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}$$

where the interaction between each electron and the nucleus is given by the two first terms. The mutual electron-electron repulsion is given by the last. The distance between two electrons is $r_{ij} = |\boldsymbol{r}_j - \boldsymbol{r}_i|$.

The *Hamiltonian* of the system is thus

$$\widehat{\mathbf{H}} = -\frac{\nabla_1^2}{2} - \frac{\nabla_2^2}{2} - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}$$

scaled to *atomic units* (a.u.).

This is based on the simplified, analytically solvable, system for the Hydrogen atom, where the Hamiltonian is

$$\widehat{\mathbf{H}} = \widehat{\mathbf{K}} + \widehat{\mathbf{V}} = -\frac{\hbar^2}{2m_e}\hat{\nabla}^2 - \frac{e^2}{4\pi\epsilon_0 r} \tag{1}$$

with e as the electron charge, r is the radius from the proton to the electron and m_e as the mass of the electron.

This operator satisfies the *time-independent Schrödinger Equation*

$$\widehat{\mathbf{H}}|\psi\rangle = E|\psi\rangle \tag{2}$$

where E is the *energy eigenvalue* and ψ is the *eigenstate* of the system.

For a Hydrogen-like quantum mechanical system, we can model the general Hamiltonian of an atom as

$$\widehat{\mathbf{H}} = \sum_i (\hat{k}_i + \hat{v}_i) + \sum_{i<j} \hat{v}_{ij}$$

$$= -\sum_{i=1}^{N}\left[\frac{\hat{\nabla}_i^2}{2} + \frac{Z}{r_i}\right] + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{r_{ij}} \tag{3}$$

Here Z is the nuclear charge, and we have another sum for all the electron-electron repulsions.

We can add to this model by setting up a system of two atoms, a *diatomic molecule*, which gives us a Hamiltonian on the form

$$\widehat{\mathbf{H}} = -\sum_{i=1}^{N}\left[\frac{\hat{\nabla}_i^2}{2} + \frac{Z_1}{r_{ip1}} + \frac{Z_2}{r_{ip2}}\right] + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{r_{ij}} + \frac{Z_1 Z_2}{R}$$

We now want to look at the form of the wave function for the Hydrogen atom. The Hamiltonian (1) in atomic units is

$$\widehat{\mathbf{H}} = -\frac{\hat{\nabla}^2}{2} - \frac{1}{r}$$

and the energy eigenvalues in (2) are given by

$$E_n = -\frac{1}{2n^2}$$

Solving the Schrödinger equation is done by separation of variables in spherical coordinates

$$\psi_{nlm}(r, \theta, \varphi) = R_{nl}(r)Y_l^m(\theta, \varphi) \tag{4}$$

where n, l and m are quantum numbers (principal, azimuthal and magnetic respectively). The solution is separated into a radial part R_{nl} and an angular part Y_l^m called the spherical harmonics. A large part of this project is using the solutions of the *Hydrogen-like* wave functions as basis for the wave function for Helium, Beryllium and Neon.

The solutions of the first few *atomic orbitals* (and the ones we will use), are on the form

$$\psi_{100} = \phi_{1s} = e^{-\alpha r}$$
$$\psi_{200} = \phi_{2s} = \left(1 - \frac{\alpha r}{2}\right)e^{-\alpha r/2}$$
$$\psi_{211} = \phi_{2p_x} = xe^{-\alpha r/2}$$
$$\psi_{210} = \phi_{2p_y} = ye^{-\alpha r/2}$$
$$\psi_{21-1} = \phi_{2p_z} = ze^{-\alpha r/2}$$

where x, y and z are the cartesian coordinate positions of the electron, and α is some charge.

B. Slater Determinants

The Hamiltonian (3) introduced in the last section is invariant when you interchange two particles, and since electrons are *fermions*, it is required that the wave function is antisymmetric.

1

In practice, this means that for a system of two particles, 1 and 2, described by the wave functions ψ_a and ψ_b, we must have a total wave function on the form

$$\psi^{1,2} = \psi_a^1 \psi_b^2 - \psi_b^1 \psi_a^2 \tag{5}$$

This is a result that follows from the fact that particles are indistinguishable, and the minus sign comes from the antisymmetry condition. An exchange operator $\widehat{\mathbf{P}}$ acting on an antisymmetric wave function, will therefore change the sign, thus the wave function is an eigenfunction of $\widehat{\mathbf{P}}$ with eigenvalue $p = -1$.

For bosons however, the sign is positive. This means that they do not obey the *Pauli principle*, stating that two particles can not occupy the same state, which would imply $\psi^{1,2} = 0$.

If we rewrite equation (5), we see that this is actually a determinant

$$\psi^{1,2} = \begin{vmatrix} \psi_a^1 & \psi_b^1 \\ \psi_a^2 & \psi_b^2 \end{vmatrix}$$

that can be generallized to the N-particle case as

$$\psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1^1 & \cdots & \psi_N^1 \\ \vdots & \ddots & \vdots \\ \psi_1^N & \cdots & \psi_N^N \end{vmatrix} \tag{6}$$

These Slater determinants obey exchange operations, since switching two rows of the determinant changes the sign. In our project we will omit the factor in front of the determinant, as the important part is proportionality, and we will see later that constants get cancelled.

In this project, we are studying spin-$\frac{1}{2}$ particles, and we must address this problem in the Slater determinants we will use. The electrons have either spin $\uparrow$ or $\downarrow$, so our determinant in (6) have elements labeled $\uparrow$ or $\downarrow$. However, we must also label them by their atomic orbitals, so an element of the determinant will have the form

$$\phi_{j\sigma}(\boldsymbol{r}_i) = \phi_{j\sigma}^i$$

where j is the atomic orbital (1s, 2s, 2p in this project), σ is the spin, and i is the particle number. Determinants of matrices with elements like that for Helium, Beryllium and Neon, are unfortunately zero, since they are independent of spin (two and two columns would be equal). This can be bypassed by writing the determinant $|\hat{D}|$ as a product of two smaller ones for each spin[1], since

$$\psi_D = |\hat{D}| \propto |\hat{D}|_\uparrow |\hat{D}|_\downarrow \tag{7}$$

By doing this, the wave function will lose its antisymmetry when interchanging particles, but it doesn't affect the expectation values.

For Helium, we write the wave function as a product of the (reduced) 1×1 matrices to simplify the function. Therefore

$$\psi_D^{He} = \phi_{1s}(\boldsymbol{r}_1)\phi_{1s}(\boldsymbol{r}_2)$$

---
[1]Morten, 519-520

For Beryllium (four electrons), the corresponding Slater determinant is

$$\psi_D^{Be} = \begin{vmatrix} \phi_{1s\uparrow}^1 & \phi_{1s\uparrow}^2 & \phi_{1s\uparrow}^3 & \phi_{1s\uparrow}^4 \\ \phi_{1s\downarrow}^1 & \phi_{1s\downarrow}^2 & \phi_{1s\downarrow}^3 & \phi_{1s\downarrow}^4 \\ \phi_{2s\uparrow}^1 & \phi_{2s\uparrow}^2 & \phi_{2s\uparrow}^3 & \phi_{2s\uparrow}^4 \\ \phi_{2s\downarrow}^1 & \phi_{2s\downarrow}^2 & \phi_{2s\downarrow}^3 & \phi_{2s\downarrow}^4 \end{vmatrix} \tag{8}$$

$$= \begin{vmatrix} \phi_{1s\uparrow}^1 & \phi_{1s\uparrow}^2 \\ \phi_{2s\uparrow}^1 & \phi_{2s\uparrow}^2 \end{vmatrix} \begin{vmatrix} \phi_{1s\downarrow}^3 & \phi_{1s\downarrow}^4 \\ \phi_{2s\downarrow}^3 & \phi_{2s\downarrow}^4 \end{vmatrix} \tag{9}$$

and finally for Neon (ten electrons), the spin $\uparrow$ part

$$|\hat{D}|_\uparrow^{Ne} = \begin{vmatrix} \phi_{1s\uparrow}^1 & \phi_{1s\uparrow}^2 & \phi_{1s\uparrow}^3 & \phi_{1s\uparrow}^4 & \phi_{1s\uparrow}^5 \\ \phi_{2s\uparrow}^1 & \phi_{2s\uparrow}^2 & \phi_{2s\uparrow}^3 & \phi_{2s\uparrow}^4 & \phi_{2s\uparrow}^5 \\ \phi_{2p_x\uparrow}^1 & \phi_{2p_x\uparrow}^2 & \phi_{2p_x\uparrow}^3 & \phi_{2p_x\uparrow}^4 & \phi_{2p_x\uparrow}^5 \\ \phi_{2p_y\uparrow}^1 & \phi_{2p_y\uparrow}^2 & \phi_{2p_y\uparrow}^3 & \phi_{2p_y\uparrow}^4 & \phi_{2p_y\uparrow}^5 \\ \phi_{2p_z\uparrow}^1 & \phi_{2p_z\uparrow}^2 & \phi_{2p_z\uparrow}^3 & \phi_{2p_z\uparrow}^4 & \phi_{2p_z\uparrow}^5 \end{vmatrix} \tag{10}$$

and similarly for $\downarrow$

$$|\hat{D}|_\downarrow^{Ne} = \begin{vmatrix} \phi_{1s\uparrow}^6 & \phi_{1s\downarrow}^7 & \phi_{1s\downarrow}^8 & \phi_{1s\downarrow}^9 & \phi_{1s\downarrow}^{10} \\ \phi_{2s\downarrow}^6 & \phi_{2s\downarrow}^7 & \phi_{2s\downarrow}^8 & \phi_{2s\downarrow}^9 & \phi_{2s\downarrow}^{10} \\ \phi_{2p_x\downarrow}^6 & \phi_{2p_x\downarrow}^7 & \phi_{2p_x\downarrow}^8 & \phi_{2p_x\downarrow}^9 & \phi_{2p_x\downarrow}^{10} \\ \phi_{2p_y\downarrow}^6 & \phi_{2p_y\downarrow}^7 & \phi_{2p_y\downarrow}^8 & \phi_{2p_y\downarrow}^9 & \phi_{2p_y\downarrow}^{10} \\ \phi_{2p_z\downarrow}^6 & \phi_{2p_z\downarrow}^7 & \phi_{2p_z\downarrow}^8 & \phi_{2p_z\downarrow}^9 & \phi_{2p_z\downarrow}^{10} \end{vmatrix} \tag{11}$$

Here we have used equality instead of proportionality, since these are the actual wave functions we will use.

C. The Variational Principle

The Variation Principle states that if we have a Hamiltonian $\widehat{\mathbf{H}}$ and a trial wavefunction ψ_T, an upper bound for the ground state energy E_0 is given by equation (12).

$$E_0 \le \langle H \rangle = \frac{\int \mathrm{d}\boldsymbol{R}\, \psi_T^*(\boldsymbol{R})\widehat{\mathbf{H}}(\boldsymbol{R})\psi_T(\boldsymbol{R})}{\int \mathrm{d}\boldsymbol{R}\, \psi_T^*(\boldsymbol{R})\psi_T(\boldsymbol{R})} \tag{12}$$

To find such a wave function, we expand it in the eigenstates of the Hamiltonian (since they form a complete set) as follows

$$\psi_T(\boldsymbol{R}) = \sum_i c_i \psi_i(\boldsymbol{R})$$

and given that they are normalized, we get

$$E_0 \le \frac{\sum_{i,j} c_m^* c_n \int \mathrm{d}\boldsymbol{R}\, \psi_m^*(\boldsymbol{R})\widehat{\mathbf{H}}(\boldsymbol{R})\psi_n(\boldsymbol{R})}{\sum_{m,n} c_m^* c_n \int \mathrm{d}\boldsymbol{R}\, \psi_m^*(\boldsymbol{R})\psi_n(\boldsymbol{R})} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2}$$

since $E_0 = E_0 \sum_n |c_n|^2 \le \sum_n E_n |c_n|^2$

The key to the variational principle is to find suitable trial wave functions that live in the same Hilbert space as the Hamiltonian. Given a trial wave function, we can then vary som parameters $\boldsymbol{\alpha} = (\alpha, \beta, \ldots)$ to optimize the function, and the energy value E_T.

Another (analytical) quantity that we want to calculate, is the *local energy*

$$E_L(\boldsymbol{R}, \boldsymbol{\alpha}) = \frac{1}{\psi_T(\boldsymbol{R}, \boldsymbol{\alpha})} \widehat{\mathbf{H}} \psi_T(\boldsymbol{R}, \boldsymbol{\alpha}) \tag{13}$$

which together with (12) yields

$$\langle H \rangle = \frac{\int \mathrm{d}\boldsymbol{R}\, E_L |\psi_T|^2}{\int \mathrm{d}\boldsymbol{R}\, |\psi_T|^2} \tag{14}$$

D. Variational Monte Carlo (VMC)

The integrals to be solved in the variational method, does not scale well with traditional integral methods when increasing the number of particles and dimensions and using more complex wave functions . Therefore, we introduce the *brute-force Monte Carlo method* to solve the integrals.

We define

$$P(\boldsymbol{R}) \equiv \frac{|\psi_T|^2}{\int \mathrm{d}\boldsymbol{R}\,|\psi_T|^2}$$

and can then rewrite the integral in (14), and approximate it as

$$\int \mathrm{d}\boldsymbol{R}\, E_L P(\boldsymbol{R}) \approx \frac{1}{N} \sum_{i=1}^{N} E_L P(\boldsymbol{R})$$

Here N is the number of *Monte Carlo samples*. An outline of the algorithm can be seen in Algorithm 1.

Algorithm 1 VMC Algorithm

1: **procedure** INITIALIZATION
2:      Set a fixed number of MC steps.
3:      Choose initial position $\boldsymbol{R}$ and variational
           parameters $\boldsymbol{\alpha}$.
4:      Calculate $|\psi_T(\boldsymbol{R})|^2$.
5: **procedure** INITIALIZE ENERGY AND VARIANCE
      AND START THE MC CALCULATION
6:      Find the trial position $\boldsymbol{R}' = \boldsymbol{R} + \delta \times r$,
           where $r \in [0, 1]$ is randomly selected.
7:      Use the Metropolis algorithm to determine if the
           move $w = \frac{P(\boldsymbol{R}')}{P(\boldsymbol{R})}$ is accepted or rejected.
8:      Given that the move is accepted, set $\boldsymbol{R} = \boldsymbol{R}'$.
9:      Update averages.
10: **procedure** COMPUTE FINAL AVERAGES

E. The Metropolis Algorithm

In the VMC algorithm (Algorithm 1), one of the steps is to perform the Metropolis algorithm to determine wheter or not a step is rejected or accepted.

The Metropolis algorithm samples a normalized probability distribution by a stochastic process. In other words, it is a method for simultaing random walks, which we can use to do brute-force Monte Carlo computations. The algorithm uses *Markov chains*, where each state i, has some probability of going to another state j. Each state is only dependent on the last state it was in.

Define $P_i^{(n)}$ as the probability for finding the system in state i at some step n. This probability is given by the sum of the probability of previously being in a state j, transitioning to i, $T_{j \to i}$, and the probability of rejecting the transition to another state j, $1 - A_{i \to j}$. Where A is the acceptance ratio

This results in

$$P_i^{(n)} = \sum_j \left[P_j^{(n-1)} T_{j \to i} A_{j \to i} - P_i^{(n-1)} T_{i \to j} A_{i \to j} \right]$$

since the probability of making a transition must be 1, $\sum_j T_{i \to j}$.

See Algorithm 2 for an outline of the Metropolis algorithm.

Algorithm 2 Metropolis Algorithm

1: **procedure** METROPOLIS TEST
2:      Sample a possible new state j,
           with a probability $T_{i \to j}$.
3:      With probability $A_{i \to j}$, accept j as a new state.
4:      Set it as the new sample.
5:      The move is rejected with probability $1 - A_{i \to j}$.
6:      Set i as the sample again.

F. Importance Sampling

A random walker is the most efficient way to sample where the wave function is large. To further increase the efficiency, we can implement importance sampling, where the walk in space is governed by the trial wave function. The idea is to try and drive the walker in to an area with a higher probability density.

In the brute force method, one tries a new position

$$y = x + \xi\sqrt{\Delta t}$$

where ξ a *gaussian random variable* centered on x. With importance sampling, we add a term that takes into consideration where the probability density is higher.

The results are based on the *Fokker-Planck equation* and the *Langevin equation*. This makes the step look similar to a diffusion process, where the new position is given by

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

In this equation, ξ is a gaussian random variable and Δt is a chosen time step. In atomic units, D is $1/2$, and in our plots $\Delta t = 0.01$.

An important part of this result is that the walker is affected by $F(x)$ which contains the trial wavefunction. This force has been named the *quantum force*, given by

$$\boldsymbol{F} = 2\frac{1}{\psi_T}\nabla\psi_T.$$

This also has the result of changing the random number check in the Metropolis algorithm from: $q(y, x) = |\psi_T(y)|^2/|\psi_T(x)|^2$ to:

$$q(y, x) = \frac{G(x, y, \Delta t)|\psi_T(y)|^2}{G(y, x, \Delta t)|\psi_T(x)|^2}.$$

Where G is the Green function given by

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

This has its basis in our change of center for the gaussian distribution.

G. Gradient / Newton's Method

This section needs CONTENT!

H. Statistical Analysis

The MC calculations are a set of computational *experiments* with statistical errors. In these experiments, we are interested in the mean value of the ground energies and the density distribution. The individual samples are not as interesting to us, and we would rather like to look at the variance of the mean values. The variance of the mean value is closely connected with the correlation in the individual samples.

Both our random walker with and without importance sampling, gives correlated samples. In our case, we use *blocking*[2] as a technique to find out have many steps the walker has to do to be as uncorrelated as possible in comparison with its first step. This means that for a set of random samples A, one sample A_i will be correlated to some other sample A_{i+n}.

If we we group our data (N samples) in blocks of size n_b, then the number of blocks is $m = N/n_b$. The average of one block i (the average of all samples between $i n_b + 1$ and $(i+1)n_b$), is given by

$$\langle A \rangle_i = \frac{1}{n_b} \sum_{j=i n_b+1}^{(i+1)n_b} A_j$$

We can then calculate the average of all the blocks as

$$\langle A \rangle = \frac{1}{m} \sum_{i=1}^{m} \left(\frac{1}{n_b} \sum_{j=i n_b+1}^{(i+1)n_b} A_j \right) \equiv \frac{1}{m} \sum_{i=1}^{m} \langle A \rangle_i$$

and the variance as

$$\mathrm{Var}(A) = \frac{1}{m} \sum_{i=1}^{m} \langle A \rangle_i^2 - \left(\frac{1}{m} \sum_{i=1}^{m} \langle A \rangle_i \right)^2$$

By plotting the variance of the mean as a function of block size we will see that it reaches a plateau. This plateau means that increasing the sample length will no longer change the variance in the mean significantly. This happens at some block size $n_b > K_0$, where K_0 is called the *correlation length*. This makes us able to more easily calculate variance in the mean because we know how small block sizes we can use. This is used to make the standard deviation interval around our density plots.

---

[2] Ab Initio p. 60-61

Algorithm 3 Blocking Method
___
1: **procedure** COMPUTE VARIANCE OF MEAN
2:     Compute MC calculation, store samples in array.
3:     Loop over a set of block sizes n_b.
4:     For each n_b, calculate the mean of the block, and store these values in a new array.
5:     Take the mean and variance of this array.
6:     Store results.
___

I. The Hydrogen-Like Trial Wave Functions

The arguably most important part of the project is to find good trial wave functions. In the following, we will present the wave functions we have used in this project. In this section, we show the functions which are based on the solutions of (4).

Our trial wave function is a product of the Slater determinant ψ_D and a correlation part ψ_C that considers the electron-electron repulsion. The second part is also called the *Jastrow factor*[3]. The the general trial wave function is

$$\psi_T = \psi_D \psi_C$$

The trial wave functions are always dependent on the coordinates of every electron in some way. The Jastrow factor is generally written as

$$\psi_C = \prod_{i<j}^{n} \exp\left(\frac{a r_{ij}}{1 + \beta r_{ij}} \right) \tag{15}$$

where a is a factor that takes into account the spin of electron i and j, the *electron-electron cusp conditions*. When the particles have equal spin we have $a = 1/4$, and when they are different $a = 1/2$ (for example the ground state of Helium). For Beryllium and Neon, we set the spin of the first $n/2$ particles to $\uparrow$ and the last half to spin $\downarrow$. The Slater determinant part is given by (7).

We also want to determine the local energy E_L to approximate the ground state energy of the atom. The general expression for the local energy is

$$E_L = \frac{1}{\psi_T(\boldsymbol{R})} \widehat{\mathbf{H}} \psi_T(\boldsymbol{R})$$

and the Hamiltonian can generally be described as in (3).

From this expression, it's clear that $\widehat{\mathbf{K}}$ is the only operator that changes the trial wavefunction when we calculate the local energy. Therefore, we must calculate the following quantities

$$\frac{1}{\psi_T} \hat{k}_i \psi_T = -\frac{1}{2} \frac{\hat{\nabla}_i^2 \psi_T}{\psi_T}$$

The product rule of differentiation gives us

$$\frac{\hat{\nabla}^2 \psi_T}{\psi_T} = \frac{\hat{\nabla}^2 \psi_D}{\psi_D} + 2\frac{\hat{\nabla}\psi_D}{\psi_D} \cdot \frac{\hat{\nabla}\psi_C}{\psi_C} + \frac{\hat{\nabla}^2 \psi_C}{\psi_C}$$

So we need to calculate the four quantities

$$\frac{\hat{\nabla}^2 \psi_D}{\psi_D} \quad \frac{\hat{\nabla}\psi_D}{\psi_D} \quad \frac{\hat{\nabla}\psi_C}{\psi_C} \quad \frac{\hat{\nabla}^2 \psi_C}{\psi_C}$$

They are derived in the appendix, and they are not only useful when calculating the local energy.

---

[3] Ab Initio p. 280

i. The First Trial Wave Function of Helium

We first model the variational solution with a trial function of one variation parameter α (corresponding to the α in Hydrogen solution(4)). It has the form

$$\psi_{T1}^{He}(\mathbf{r_1}, \mathbf{r_2}) = \exp\left(-\alpha(r_1 + r_2)\right)$$

which is the Slater-determinant part of the wave function.

We now proceed to find the local energy, given by (13). The only part of the operator $\hat{H}$ from (3) that affects the wave function are the Laplace operators. Since ψ_{T1} is only spatially dependent on r_1 and r_2, we have

$$\nabla_i^2 \psi_{T1}^{He} = \left(\frac{\partial^2}{\partial r_i^2} + \frac{2}{r_i}\frac{\partial}{\partial r_i}\right)\psi_{T1}^{He} = \left(\alpha^2 - \alpha\frac{2}{r_i}\right)\psi_{T1}^{He}$$

for $i = 1,\ 2$, since

$$\frac{\partial}{\partial r_i}e^{-\alpha(r_1+r_2)} = -\alpha e^{-\alpha(r_1+r_2)}$$
$$\frac{\partial^2}{\partial r_i^2}e^{-\alpha(r_1+r_2)} = \alpha^2 e^{-\alpha(r_1+r_2)}$$

This gives us the following trial energy

$$E_{L1}^{He} = (\alpha - 2)\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{1}{r_{12}} - \alpha^2$$

The 2 in the $\alpha - 2$ term is the number of protons, Z.

ii. The second trial wavefunction

To approximate the closed-form solution even better, we assume another trial wavefunction based on the fact that the two electrons interact, namely

$$\psi_{T2}^{He}(\mathbf{r}_1, \mathbf{r}_2) = \exp\left(-\alpha(r_1 + r_2)\right)\exp\left(\frac{r_{12}}{2(1+\beta r_{12})}\right)$$

One can then in the same way as for ψ_{T1} calculate the local energy. The correlations part will give us some trouble when we try to calculate the Laplacian. This is due to the distance between $\mathbf{r}_1$ and $\mathbf{r}_2$, since this quantity is dependent on the angles φ and θ. The local energy now has the form

$$E_{L2}^{He} = E_{L1}^{He} + \frac{1}{2(1+\beta r_{12})^2}\left(\frac{\alpha(r_1+r_2)}{r_{12}}(1 - \frac{\mathbf{r}_1^T\mathbf{r}_2}{r_1 r_2})\right.$$
$$\left. - \frac{1}{2(1+\beta r_{12})^2} - \frac{2}{r_{12}} + \frac{2\beta}{1+\beta r_{12}}\right)$$

ADD CUSP CONDITIONS HERE!

iii. Beryllium

The trial wavefunction of Beryllium can be written as a product of a Slater determinant part and a correlation part on the form

$$\psi_T(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = \psi_D\psi_C \tag{16}$$

where the Slater determinant is

$$\psi_D = Det\left(\phi_1(\mathbf{r}_1), \phi_2(\mathbf{r}_2), \phi_3(\mathbf{r}_3), \phi_4(\mathbf{r}_4)\right) \tag{17}$$
$$= \left(\phi_{1s}^1\phi_{2s}^2 - \phi_{1s}^2\phi_{2s}^1\right)\left(\phi_{1s}^3\phi_{2s}^4 - \phi_{1s}^4\phi_{2s}^3\right)$$

and the correlation part is

$$\psi_C = \prod_{i<j}^4 g_{ij} = \prod_{i<j}^4 \exp\left(\frac{ar_{ij}}{1+\beta r_{ij}}\right) \tag{18}$$

Here $\phi_i(\mathbf{r}_i)$ are the hydrogen-like wavefunctions. They are given by the $1s$ and $2s$ orbital parts

$$\phi_{1s}^i = e^{-\alpha r_i}$$
$$\phi_{2s}^i = (1 - \alpha r_i/2)\,e^{-\alpha r_i/2}$$

which are dependent on the cartesian positions $\mathbf{r}_i = (x_i, y_i, z_i)$. The relative distance between two particles is

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

and obviously $r_{ij} = r_{ji}$.

iv. Neon

For Neon, in addition to

we need three more hydrogenic wavefunctions for the $2p$ orbital, given by $2p_x$, $2p_y$ and $2p_z$. These are given by

$$\phi_{2p_k}(\mathbf{r}_i) = \alpha k_i g_i \text{ for } k_i = x_i, y_i, z_i$$

so the wave function is

$$\psi_T^{Ne} = |\hat{D}|\uparrow\,|\hat{D}|\downarrow \prod_{i<j}^{10} \exp\left(\frac{ar_{ij}}{1+\beta r_{ij}}\right)$$

where the two Slater determinants are given in (10) and (11).

J. Introducing Gaussian-Type Orbitals

In this project, we have shown how one can use Hydrogen-like wave functions as the basis for the Slater determinant, and up until now, we have used so-called *Slater-Type Orbitals* (STOs). We will now replace the STOs with *Gaussian-Type Orbitals* (GTOs).

The spin-factorized Slater determinants (one for spin up and one for spin down), are given by

$$\psi_{D_{spin}} = \begin{vmatrix} \phi_1(\mathbf{r}_1) & \cdots & \phi_1(\mathbf{r}_n) \\ \vdots & \ddots & \vdots \\ \phi_n(\mathbf{r}_1) & \cdots & \phi_n(\mathbf{r}_n) \end{vmatrix}$$

where n is equal to half the number of particles, and $\phi_j(\mathbf{r}_i)$ are one-particle wave functions ϕ_{1s}, ϕ_{2s} and $\phi_{2p(3)}$ ((3) is short-hand notation for the x, y and z parts). These

spin independent, one-particle wave functions are linear combinations of basis functions, and we will now use *contracted* GTOs (CGTOs) for this purpose.

The functions can then be written as

$$\phi_j(\boldsymbol{r}_i) = \sum_p C_{pj}\varphi_p(\boldsymbol{r}_i)$$

The CGTOs are on the form

$$\varphi_p^{CGTO}(\boldsymbol{r}_i) = \sum_k N_k \varphi_k^{GTO}(\boldsymbol{r}_i)$$

with

$$\varphi_k^{GTO} = d_k x^m y^n z^o e^{-\alpha_k r^2}$$

which are called *cartesian* GTOs. The numerical coefficients α_k and d_k are chosen to optimize the basis functions, and we use values from the *Environmental Molecular Science Laboratory* (EMSL) website[4].

The normalization factors are given as[5]

$$N_k = \left(\frac{2\alpha_k}{\pi}\right)^{3/4}\sqrt{\frac{(4\alpha_k)^{m+n+o}}{(2m-1)!!\,(2n-1)!!\,(2o-1)!!}}$$

and since they are dependent on the total angular momentum, given by the quantum numbers $l = m + n + o$, we will write them as $N_k = N_{m,n,o}$.

K.   Efficient Computation of the Slater Determinant

For larger atoms, the evaluation of the gradient and the Laplacian of the Slater determinant becomes increasingly numerically demanding to compute. Computing these quantities with brute force, leads to $N \cdot d$ operations to find the determinant and thus multiplying this with our $O(N^3)$ operations. In the following, we derive a method that deals with this issue, and achieves a lower number of operations.

We can approximate the Slater determinant as

$$\Phi(\boldsymbol{r}_1,...,\boldsymbol{r}_N) \propto \det\uparrow \cdot \det\downarrow$$

where the spin determinants are the determinants which only depend on spin up and spin down respectively. The determinants are 2×2 for Beryllium and 5×5 for Neon. This is true only if $\widehat{\mathbf{H}}$ is spin independent.

Then, $\det \hat{D} = |\hat{D}| = |\hat{D}|_\uparrow \cdot |\hat{D}|_\downarrow$, where the Slater matrices are dependent on the positions of the electrons. Each time we update the positions and differentiate the Slater determinant, the Slater matrix is changed, but by calculating the determinant from scratch each time, we will certainly do unnecessary computations.

This is solved by the following algorithm, that instead of calculating the determinant, updates the *inverse* of the Slater matrix suitably.

[4]https://bse.pnl.gov/bse/portal
[5]Ab Initio, p. 280

We first express (i, j) elements of the inverse of $\hat{D}$ as

$$d_{ij}^{-1} = \frac{C_{ji}}{|\hat{D}|}$$

where C_{ji} is the transposed cofactor-matrix element of $\hat{D}$.

This motivates the ratio

$$R \equiv \frac{|\hat{D}(\boldsymbol{r}^{new})|_\uparrow}{|\hat{D}(\boldsymbol{r}^{old})|_\uparrow} = \frac{\sum_{j=1}^N d_{ij}^{new} C_{ij}^{new}}{\sum_{j=1}^N d_{ij}^{old} C_{ij}^{old}}$$

Every time we move particle i, the i-th row of $\hat{D}$ changes, and we have to update the inverse. However, the i-th row of $\hat{C}$ is independent[6] of the i-th row of $\hat{D}$, which means that we must have

$$\hat{C}_{ij}^{new} = \hat{C}_{ij}^{old} = (d_{ji}^{-1})^{old} \cdot |\hat{D}| \text{ for } j = 1,...,N$$

and using

$$\sum_{k=1}^N d_{ik}\,d_{kj}^{-1} = \delta_{ij}$$

The result is

$$R = \sum_{j=1}^N d_{ij}^{new}(d_{ji}^{-1})^{old} = \sum_{j=1}^N \phi_j(\boldsymbol{r}_i^{new})\,d_{ji}^{-1}(\boldsymbol{r}_i^{old})$$

The algorithm for updating the inverse of the matrix when a new position is accepted is then

---
Algorithm 4 Inverse of Slater Matrix

1: **procedure** UPDATE COLUMNS $j \neq i$
2:      **for** each column $i \neq j$ **do**
3:          $S_j = \sum_{l=1}^N d_{il}(\boldsymbol{r}^{new})d_{lj}^{-1}(\boldsymbol{r}^{old})$
4:          $(d_{kj}^{-1})^{new} = (d_{kj}^{-1})^{old} - \frac{S_j}{R}(d_{ki}^{-1})^{old}$
5: **procedure** UPDATE COLUMN i
6:      $(d_{ki}^{-1})^{new} = \frac{1}{R}(d_{ki}^{-1})^{old}$

We can then calculate the gradient and laplacian as

$$\frac{\hat{\nabla}_k|\hat{D}|}{|\hat{D}|} = \sum_{j=1}^N \nabla_k\phi_j(\boldsymbol{r}_i^{new})\,d_{ji}^{-1}(\boldsymbol{r}_i^{old})$$

$$\frac{\hat{\nabla}_k^2|\hat{D}|}{|\hat{D}|} = \sum_{j=1}^N \nabla_k^2\phi_j(\boldsymbol{r}_i^{new})\,d_{ji}^{-1}(\boldsymbol{r}_i^{old})$$

REMEMBER TO SAY SOMETHING ABOUT NUMBER OF OPERATIONS.

L.   Nonlinear Conjugate Gradient Method

Our aim is to calculate the lowest mean of the ground state energy E_0, but it depends on some variational parameters

[6]Since the cofactor-matrix elements c_{ij} is defined by removing i-th row and j-th column from a matrix $\hat{A}$, and then taking the determinant of the remaining matrix.

α. Therefore we need to find the global minimum with respect to these parameters, where

$$\frac{\partial E}{\partial \alpha_i} = 0$$

where α_i can take the values α and β.

For a function f that we wish to minimize, with

$$f(x) = \frac{1}{2}x^T A x + b^T x + c$$

we use the *nonlinear conjugate gradient method* (NGC)[7] to do it, see Algorithm 5. We proceed one step, decided by $x_{k+1} = x_k + ad_k$, where a is the step. This step is decided by *line search method*, for example the *bisection method* (outlined in Algorithm 6). The search is guided by the direction d_k.

Algorithm 5 NCG method

1: **procedure** INITIALIZE
2:     Let $k = 0$ and $x_k = x_0$ be our initial guess.
3:     Compute $d_k = d_0 = -\nabla f(x_0)$.
4: **procedure** FIND THE BEST STEP SIZE
5:     Compute a to minimize $f(x_k + ad_k)$
        by a line search in the direction d_k.
6: **procedure** UPDATE CURRENT GUESS
7:     Let $x_{k+1} = x_k + ad_k$.
8: **procedure** UPDATE THE SEACRH DIRECTION
9:     Let $d_{k+1} = -\nabla f(x_{k+1} + \beta_k d_k)$ where

$$\beta_k = \frac{\nabla f(x_{k+1}^T)(\nabla f(x_{k+1} - \nabla f(x_k))}{d_k^T(\nabla f(x_{k+1}) - \nabla f(x_k))}$$

10: **procedure** ITERATE
11:     Repeat procedures 2 to 4 until we have looked in all
        directions $(\dim x)$.

Given a function $g(x)$ and an interval $[a, b]$, where $g(a)$ and $g(b)$ has opposite signs (at least one zero crossing), we want to find the minimum. Calculate $f(c)$ for $c = (a + b)/2$, if $f(c)$ has opposite sign of $f(a)$ $(f(b))$ then the minimum is in $[a, c]$ $([c, b])$. Choose the new interval and continue until minimum is found (for some tolerance).

Algorithm 6 Bisection method

1: **procedure** ITERATION
2:     **while** $n < N_{max}$ **do**
3:         Compute c and $f(c)$.
4:         If minimum is found, return c. Success.
5:         If not: choose new interval $[a, c]$ or $[c, b]$.

---

[7]J. Nocedal, Numerical Optimization, 2006, Springer Science, ch. 5

M. Differentiation of the Energy

To optimize the conjugate gradient method, we need analytical expressions for the derivatives of the energy with respect to the variational paramaters α. We can calculate the derivative of the energy with respect to variational parameters in the following way[8]

$$\frac{\partial E}{\partial \alpha_i} = 2\left(\left\langle E_L \frac{1}{\psi_T}\frac{\partial \psi_T}{\partial \alpha}\right\rangle - E_L\left\langle \frac{1}{\psi_T}\frac{\partial \psi_T}{\partial \alpha}\right\rangle\right) \quad (19)$$

Since $(\ln u)' = u'/u$, we can simplify the calculations of the derivatives in (19) as

$$\frac{\partial \psi_T}{\partial \alpha_i} = \frac{\partial |\hat{D}|_\uparrow}{\partial \alpha_i} + \frac{\partial |\hat{D}|_\downarrow}{\partial \alpha_i} + \frac{\partial \psi_C}{\partial \alpha_i}$$

For $\alpha_i = \alpha$, the correlation part is vanishes, and for the correlation part, the determinant parts vanish.

For α, we can calculate the derivatives of the determinants using the efficient Slater method. We then get

$$\frac{\partial |\hat{D}|}{\partial \alpha} = \sum_{i=1}^{N}\sum_{j=1}^{N} d_{ji}^{-1}\frac{\partial d_{ij}}{\partial \alpha}$$

for each spin part.

For β, the exponential in (15) goes away due to the logarithm, and we have

$$\frac{\partial \ln \psi_C}{\partial \beta} = \frac{\partial}{\partial \beta}\sum_{i<j}^{n}\frac{ar_{ij}}{1 + \beta r + ij}$$

$$= \sum_{i=1}^{n}\sum_{j=i+1}^{n}\frac{-ar_{ij}^2}{(1 + \beta r_{ij})^2}$$

N. Onebody Density and Charge Density

The one-body density is computed from the form

$$\rho(\boldsymbol{R}) = |\psi(\boldsymbol{R})|^2$$

This is implemented as NOTE: ... in our program.

[8]Morten ch. 16

III IMPLEMENTATION

The methods described above are implemented in an object oriented C++ program which is simple to use, and does not have any dependencies.

Most of the code is contained in the class VMCSolver.cpp and VMCSolver.h. These were created to contain the full system and parameters for a single run.

To start a simulation one must instantiate the solver and initialize the system, either by file and/or manually. After this the integration can be run and all data will be contained in the solver object. The last step is to collect data from the solver.

All plot data are generated by individual programs using the mentioned solver. Plots are generated using python.

IV RESULTS

V CONCLUSION

VI RESULTS

Appendix A: Mathematical Derivations

The Laplace operator of a function f in three dimensions, $\nabla^2 f$, can be represented as

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) f + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) f + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial^2 \varphi} f$$

in spherical coordinates, and as

$$\left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial}{\partial z^2} \right) f$$

in cartesian coordinates.

i. The wavefunction of Beryllium

We start by finding the first and second derivative of the determinant part, which are only dependent on the radii of the particles. This means that the gradient can be written

$$\hat{\nabla} f = \hat{\mathbf{r}} \frac{\partial f}{\partial r} \quad \text{and} \quad \hat{\nabla}^2 f = \frac{\partial^2 f}{\partial r^2} + \frac{2}{r} \frac{\partial f}{\partial r}$$

where $\boldsymbol{r} = r\hat{\mathbf{r}}$.

We define the functions

$$f_i \equiv -\frac{\alpha r_i}{2}$$
$$g_i \equiv \exp(f_i)$$
$$F_{ij} \equiv \phi_{1s}(\boldsymbol{r}_i)\phi_{2s}(\boldsymbol{r}_j) - \phi_{1s}(\boldsymbol{r}_j)\phi_{2s}(\boldsymbol{r}_i)$$

and use them to simplify derivatives of the Hydrogen-like wavefunctions in the following way

$$\phi_{1s}^i = g_i^2$$
$$\frac{\partial \phi_{1s}^i}{\partial r_i} = -\alpha g_i^2$$
$$\frac{\partial^2 \phi_{1s}^i}{\partial r_i^2} = \alpha^2 g_i^2$$
$$\phi_{2s}^i = g_i(1 + f_i)$$
$$\frac{\partial \phi_{2s}^i}{\partial r_i} = -\alpha g_i(1 + f_i/2)$$
$$\frac{\partial^2 \phi_{2s}^i}{\partial r_i^2} = \frac{3\alpha^2}{4} g_i(1 + f_i/3)$$

and for the $2p$ part we need to differentiate in cartesian coordinates, since the Hydrogenic wave function is dependent on the angles, thus

$$\phi_{2p_k}^i = \alpha k_i g_i \text{ for } k_i = x_i, y_i, z_i$$
$$\frac{\partial \phi_{2p_x}^i}{\partial x_i} = \left(\alpha - \frac{\alpha^2 x_i^2}{2r_i}\right) g_i$$
$$\frac{\partial \phi_{2p_x}^i}{\partial y_i} = -\frac{\alpha^2 x_i y_i}{2r_i} g_i$$
$$\frac{\partial \phi_{2p_x}^i}{\partial z_i} = -\frac{\alpha^2 x_i z_i}{2r_i} g_i$$

which gives us a gradient on the form

$$\nabla \phi_{2p_x}^i = \left(\alpha g_i - \frac{\alpha^2 x_i^2}{2r_i} g_i, -\frac{\alpha^2 x_i y_i}{2r_i} g_i, -\frac{\alpha^2 x_i z_i}{2r_i} g_i\right)$$

The expressions for ϕ_{2p_y} and ϕ_{2p_z} are similarly

$$\nabla \phi_{2p_y}^i = \left(\frac{\alpha^2 x_i y_i}{2r_i} g_i, \alpha g_i - \frac{\alpha^2 y_i^2}{2r_i} g_i, -\frac{\alpha^2 y_i z_i}{2r_i} g_i\right)$$
$$\nabla \phi_{2p_z}^i = \left(-\frac{\alpha^2 x_i z_i}{2r_i} g_i, -\frac{\alpha^2 y_i z_i}{2r_i} g_i, \alpha g_i - \frac{\alpha^2 z_i^2}{2r_i} g_i\right)$$

Then the terms required to calculate the Slater determinant

$$F_{ij} = g_i^2(1+f_j)g_j - g_j^2(1+f_i)g_i$$

$$\frac{\partial F_{ij}}{\partial r_i} = -\alpha g_i g_j(g_i(1+f_j) - g_j(1+f_i/2))$$

$$\frac{\partial^2 F_{ij}}{\partial r_i^2} = \alpha^2 g_i g_j(g_i(1+f_j) - \frac{3}{4}g_j(1+f_i/3))$$

$$\frac{\partial F_{ij}}{\partial r_j} = \alpha g_i g_j(g_j(1+f_i) - g_i(1+f_j/2))$$

$$\frac{\partial^2 F_{ij}}{\partial r_j^2} = -\alpha^2 g_i g_j(g_j(1+f_i) - \frac{3}{4}g_i(1+f_j/3))$$

The determinant part is

$$\psi_D = F_{12}F_{34}$$

where the first is only affected by differentiation with respect to particle 1 or 2, and opposite for the second part

$$\frac{\hat{\nabla}_1 \psi_D}{\psi_D} = \frac{\hat{\nabla}_1 F_{12}}{F_{12}} \quad \frac{\hat{\nabla}_2 \psi_D}{\psi_D} = \frac{\hat{\nabla}_2 F_{12}}{F_{12}}$$

$$\frac{\hat{\nabla}_3 \psi_D}{\psi_D} = \frac{\hat{\nabla}_3 F_{34}}{F_{34}} \quad \frac{\hat{\nabla}_4 \psi_D}{\psi_D} = \frac{\hat{\nabla}_4 F_{34}}{F_{34}}$$

The result is two different quantities

$$\frac{\hat{\nabla}_k \psi_D}{\psi_D} = \begin{cases} \frac{\hat{\nabla}_i F_{ij}}{F_{ij}} = -\alpha \frac{\boldsymbol{r}_i}{r_i} \frac{(g_i(1+f_j)-g_j(1+f_i/2))}{g_i(1+f_j)-g_j(1+f_i)}, & \text{if } k=i=j-1 \\[3mm] \frac{\hat{\nabla}_j F_{ij}}{F_{ij}} = -\alpha \frac{\boldsymbol{r}_j}{r_j} \frac{(g_i(1+f_j/2)-g_j(1+f_i))}{g_i(1+f_j)-g_j(1+f_i)}, & \text{if } k=j=i+1 \end{cases}$$

one for particles 1 and 3, and one for 2 and 4, where $i \in 1,3$ and $j \in 2,4$.
For the second derivative part, we have

$$\frac{\hat{\nabla}_k^2 \psi_D}{\psi_D} = \begin{cases} \frac{\hat{\nabla}_i^2 F_{ij}}{F_{ij}} = \frac{\alpha^2(g_i(1+f_j)-\frac{3}{4}g_j(1+f_i/3))-\frac{2\alpha}{r}(g_i(1+f_j)-g_j(1+f_i/2))}{g_i(1+f_j)-g_j(1+f_i)} \\[3mm] \frac{\hat{\nabla}_j^2 F_{ij}}{F_{ij}} = \frac{\alpha^2(\frac{3}{4}g_i(1+f_j/3)-g_j(1+f_i))-\frac{2\alpha}{r}(g_i(1+f_j/2)-g_j(1+f_i))}{g_i(1+f_j)-g_j(1+f_i)} \end{cases}$$

with the same conditions as above. Now we move on to the calculation of the correlation parts, given by the function in (18).
First off, we define

$$f_{ij} \equiv \frac{ar_{ij}}{1+\beta r_{ij}}$$

with the corresponding derivatives (with respect to r_{ij})

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

$$f_{ij}'' = \frac{-2a\beta}{(1+\beta r_{ij})^3}$$

The gradient of the wavefunction, divided by the wavefunction, for particle k in the x-direction is then

$$\left[\frac{\hat{\nabla}_k \psi_C}{\psi_C}\right]_x = \frac{1}{\psi_C}\frac{\partial \psi_C}{\partial x_k}$$

If we look at the first derivative in the x-direction, we see that the parts of the wavefunction that is not dependent on k, will remain unaffected by the differentiation. When we split the expression for $i < k$ and $k > i$, we get that

$$
\begin{aligned}
\frac{\partial \psi_C}{\partial x_k} &= \prod_{i,j \neq k} g_{ij} \frac{\partial}{\partial x_k} \left[\prod_{i<k} g_{ik} \cdot \prod_{i>k} g_{ki} \right] \\
&= \prod_{i,j \neq k} g_{ij} \left[\prod_{i>k} g_{ki} \frac{\partial}{\partial x_k} \prod_{i<k} g_{ik} + \prod_{i<k} g_{ik} \frac{\partial}{\partial x_k} \prod_{i>k} g_{ki} \right] \\
&= \prod_{i,j \neq k} g_{ij} \left[\prod_{i>k} g_{ki} \sum_{i<k} \frac{\partial g_{ik}}{\partial x_k} \prod_{p \neq i} g_{pi} + \prod_{i<k} g_{ik} \sum_{i>k} \frac{\partial g_{ki}}{\partial x_k} \prod_{q \neq i} g_{iq} \right] \\
&= \prod_{i<j} g_{ij} \left[\frac{1}{\prod_{i<k} g_{ik}} \sum_{i<k} \frac{\partial g_{ik}}{\partial x_k} \prod_{p \neq i} g_{pi} + \frac{1}{\prod_{i>k} g_{ki}} \sum_{i>k} \frac{\partial g_{ki}}{\partial x_k} \prod_{p \neq i} g_{pi} \right] \\
&= \psi_C \left[\sum_{i<k} \frac{1}{g_{ik}} \frac{\partial g_{ik}}{\partial x_k} + \sum_{i>k} \frac{1}{g_{ki}} \frac{\partial g_{ki}}{\partial x_k} \right]
\end{aligned} \tag{20}
$$

Here we factorized the wavefunction outside the expression, and noticed that the only part that doesn't cancel is the ik-th and ki-th in the sums.
Dividing by the wavefunction, we get

$$
\begin{aligned}
\left[\frac{\hat{\nabla}_k \psi_C}{\psi_C} \right]_x &= \sum_{i=1}^{k-1} \frac{1}{g_{ik}} \frac{\partial g_{ik}}{\partial x_k} + \sum_{i=k+1}^{N} \frac{1}{g_{ki}} \frac{\partial g_{ki}}{\partial x_k} \\
&= \sum_{i=1}^{k-1} \frac{\partial f_{ik}}{\partial x_k} - \sum_{i=k+1}^{N} \frac{\partial f_{ki}}{\partial x_i} \\
&= \sum_{i=1}^{k-1} \frac{x_k - x_i}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} - \sum_{i=k+1}^{N} \frac{x_i - x_k}{r_{ki}} \frac{\partial f_{ki}}{\partial r_{ki}}
\end{aligned}
$$

since g_{ij} is an exponential function, so $\partial g_{ij}/\partial x_i = g_{ij} \partial f_{ij}/\partial x_j$. We also used the fact that $\partial g_{ij}/\partial x_i = -\partial g_{ij}/\partial x_j$ to differentiate with respect to the second index in both of the sums. Finally, we have used the chain rule to attain an expression that is dependent on the distance between the two particles

$$
\frac{\partial f_{ij}}{\partial x_j} = \frac{\partial f_{ij}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_j} = \frac{x_j - x_i}{r_{ij}} \frac{\partial f_{ij}}{\partial r_{ij}}
$$

Thus

$$
\begin{aligned}
\frac{\hat{\nabla}_k \psi_C}{\psi_C} &= \sum_{i=1}^{k-1} \frac{\boldsymbol{r}_{ik}}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} - \sum_{i=k+1}^{N} \frac{\boldsymbol{r}_{ki}}{r_{ki}} \frac{\partial f_{ki}}{\partial r_{ki}} \\
&= \sum_{i=1}^{k-1} \frac{\boldsymbol{r}_{ik}}{r_{ik}} \frac{a}{(1 + \beta r_{ik})^2} - \sum_{i=k+1}^{N} \frac{\boldsymbol{r}_{ki}}{r_{ki}} \frac{a}{(1 + \beta r_{ki})^2} \\
&= \sum_{i \neq k} \frac{\boldsymbol{r}_{ik}}{r_{ik}} \frac{a}{(1 + \beta r_{ik})^2}
\end{aligned}
$$

From (20), we can gather that the double derivative part is described by

$$
\left[\frac{\hat{\nabla}^2\psi_C}{\psi_C}\right]_x = \frac{1}{\psi_C}\frac{\partial}{\partial x_k}\left(\psi_C\left[\sum_{i<k}\frac{\partial f_{ik}}{\partial x_k} + \sum_{i>k}\frac{\partial f_{ki}}{\partial x_k}\right]\right)
$$

$$
= \left[\sum_{i<k}\frac{\partial^2 f_{ik}}{\partial x_k^2} + \sum_{i>k}\frac{\partial^2 f_{ki}}{\partial x_k^2}\right] + \frac{1}{\psi_C}\frac{\partial\psi_C}{\partial x_k}\left[\sum_{i<k}\frac{\partial f_{ik}}{\partial x_k} + \sum_{i>k}\frac{\partial f_{ki}}{\partial x_k}\right]
$$

$$
= \sum_{i\neq k}\frac{\partial^2 f_{ik}}{\partial x_k^2} + \left[\sum_{i=1}^{k-1}\frac{\partial f_{ik}}{\partial x_k} - \sum_{i=k+1}^{N}\frac{\partial f_{ki}}{\partial x_i}\right]^2
$$

$$
= \sum_{i\neq k}\frac{\partial}{\partial x_k}\left(\frac{\partial f_{ik}}{\partial r_{ik}}\frac{\partial r_{ik}}{\partial x_k}\right) + \left[\sum_{i\neq k}\frac{\partial r_{ik}}{\partial x_k}\frac{\partial f_{ik}}{\partial r_{ik}}\right]^2
$$

$$
= \sum_{i\neq k}\left[\frac{\partial r_{ik}}{\partial x_k}\frac{\partial}{\partial x_k}\frac{\partial f_{ik}}{\partial r_{ik}} + \frac{\partial f_{ik}}{\partial r_{ik}}\frac{\partial^2 r_{ik}}{\partial x_k^2}\right] + \left[\sum_{i\neq k}\frac{\partial r_{ik}}{\partial x_k}f'_{ik}\right]\left[\sum_{j\neq k}\frac{\partial r_{jk}}{\partial x_k}f'_{jk}\right]
$$

$$
\left[\frac{\hat{\nabla}^2\psi_C}{\psi_C}\right]_x = \sum_{i\neq k}\left[\left(\frac{\partial r_{ik}}{\partial x_k}\right)^2 f''_{ik} + f'_{ik}\frac{r_{ik}^2 - (x_k - x_i)^2}{r_{ik}^3}\right]
$$

$$
+ \sum_{j\neq k}\left[\frac{\partial r_{ik}}{\partial x_k}f'_{ik}\left(\frac{\partial r_{ik}}{\partial x_k}f'_{ik} + \sum_{j\neq k,i}\frac{\partial r_{jk}}{\partial x_k}f'_{jk}\right)\right]
$$

$$
\left[\frac{\hat{\nabla}^2\psi_C}{\psi_C}\right]_x = \sum_{i\neq k}\left[\left(\frac{\partial r_{ik}}{\partial x_k}\right)^2 f''_{ik} + f'_{ik}\frac{r_{ik}^2 - (x_k - x_i)^2}{r_{ik}^3}\right]
$$

$$
+ \sum_{j\neq k}\left[\left(\frac{\partial r_{ik}}{\partial x_k}f'_{ik}\right)^2 + \sum_{j\neq k,i}\frac{\partial r_{ik}}{\partial x_k}f'_{ik}\frac{\partial r_{jk}}{\partial x_k}f'_{jk}\right]
$$

$$
\left[\frac{\hat{\nabla}^2\psi_C}{\psi_C}\right]_x = \sum_{i\neq k}\left[\left(\frac{x_k - x_i}{r_{ik}}\right)^2 f''_{ik} + f'_{ik}\frac{r_{ik}^2 - (x_k - x_i)^2}{r_{ik}^3}\right]
$$

$$
+ \sum_{j\neq k}\left[\left(\frac{x_k - x_i}{r_{ik}}f'_{ik}\right)^2 + \sum_{j\neq k,i}\frac{(x_k - x_i)(x_k - x_j)}{r_{ik}r_{jk}}f'_{ik}f'_{jk}\right]
$$

$$
\left[\frac{\hat{\nabla}^2\psi_C}{\psi_C}\right]_x = \sum_{i\neq k}\left[\left(\frac{x_k - x_i}{r_{ik}}\right)^2 f''_{ik} + f'_{ik}\frac{r_{ik}^2 - (x_k - x_i)^2}{r_{ik}^3}\right]
$$

$$
+ \sum_{i,j\neq k}\frac{(x_k - x_i)(x_k - x_j)}{r_{ik}r_{jk}}f'_{ik}f'_{jk}
$$

If we now sum up for all dimensions, we get

$$
\frac{\hat{\nabla}^2\psi_C}{\psi_C} = \sum_{i\neq k}\left[\frac{r_{ik}^2}{r_{ik}^2}f''_{ik} + f'_{ik}\frac{3r_{ik}^2 - r_{ik}^2}{r_{ik}^3}\right] + \sum_{i,j\neq k}\frac{(\boldsymbol{r}_k - \boldsymbol{r}_i)(\boldsymbol{r}_k - \boldsymbol{r}_j)}{r_{ik}r_{jk}}f'_{ik}f'_{jk}
$$

$$
= \sum_{i\neq k}\left[f''_{ik} + \frac{2}{r_{ik}}f'_{ik}\right] + \sum_{i,j\neq k}\frac{(\boldsymbol{r}_k - \boldsymbol{r}_i)(\boldsymbol{r}_k - \boldsymbol{r}_j)}{r_{ik}r_{jk}}f'_{ik}f'_{jk}
$$