FYS3150 Computational Physics - Project 5
Quantum Dots & Variational Monte Carlo

Nils Johannes Mikkelsen
(Dated: December 14, 2018)

Three-dimensional Helium-like quantum dots are studied using Variational Monte Carlo. The project introduces two trial waves, one the ground state of the non-interacting quantum dot, and the other a perturbation of the former. The perturbation, a Jastrow factor, takes into account the separation between the electrons in the quantum dot. The energy minimum is found using a polynomial interpolation of the trial energy integrals from the Variational Monte Carlo algorithm. The second trial wave is found to be an improvement on the first trial wave with an optimal energy minimum of 3.730 Hartree.

All material written for project 5 may be found at:
https://github.com/njmikkelsen/comphys2018/tree/master/Project5
The project description was accessed from:
http://compphysics.github.io/ComputationalPhysics/doc/Projects/2018/Project5/QuantumMonteCarlo/pdf/QuantumMonteCarlo.pdf

I. INTRODUCTION

One of the greatest achievements of 20th century physics is the analytical treatment of the Hydrogen atom in quantum mechanics. Following the success of Hydrogen, quantum physicists turned their attention towards the heavier elements with equal enthusiasm, but quickly realised that this was a completely different ball game. It turns out that the interactions between electrons is very difficult to generalize, so much so in fact that no other element is exactly solvable.

A very early model of Helium approximated the Coulombic potential from the core via a harmonic oscillator. Today, this has largely abandoned, however, the model is frequently used to describe so-called *quantum dots* (also known as *artificial atoms*) in semi-conductors. These atom-like systems behave in many ways like an atom, but are not bound by a nucleus consisting of protons and neutrons such as normal atoms.

The main focus of this project is to study a three-dimensional case of a Helium-like quantum dot. The problem will be analysed using the varational principle in quantum mechanics, which is computationally realized using Markov Chain Monte Carlo methods.

II. THEORY

The theory on quantum mechanics presented here is largely adapted from Griffiths' *Introduction to Quantum Mechanics* [1], with some additional notes from Wikipedia on atomic units [2].

The theory on Markov Chain Monte Carlo methods and Variational Monte Carlo methods is based on the FYS 3150 lecture notes on Monte Carlo methods [3] and Variational Monte Carlo methods [4].

A. Quantum Mechanics - Miscellaneous

The reader is assumed to be familiar with most of the basic features of quantum mechanics such as Dirac notation, the notion of kets and Hilbert spaces, quantum spin, composite systems, etc. Nonetheless, the following sections introduce some of the basic aspects of quantum mechanics that are readily applied and studied in this project.

1. Notation

First and foremost, kets and bras are depicted as usual, e.g. $|\psi\rangle$ and $\langle\phi|$. The hermiation conjugate of some object c is denoted by a dagger: $c^{\dagger}$. In general, operators are denoted by upper case letters and constants by lower case letters, although exceptions such as the canonical operators and ladder operators do exist. Furthermore, operators are denoted by an additional "hat" in the absence of a particular basis. For example, the canonical operators $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ expressed in the position basis $|\mathbf{r}\rangle$, are given by:

$$\mathbf{r} \quad \text{and} \quad \mathbf{p} = -i\hbar\nabla \tag{1}$$

Moreover, the single-system Hamiltonian for a system with mass m in a potential V:

$$\hat{\mathcal{H}} = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{r}}, \hat{\mathbf{p}}) \tag{2a}$$

thus takes the following form in the position basis:

$$\mathcal{H} = \frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r}, -i\hbar\nabla) \tag{2b}$$

Unless otherwise specified, all operators are expressed in the continuous position basis. Provided a classical operator O, the corresponding quantum operator $\hat{O}$ is found by inserting the canonical operators: $\hat{O} = O(\hat{\mathbf{r}}, \hat{\mathbf{p}})$.

Composite systems are denoted either as a tensor product, say $|\psi\rangle\otimes|\phi\rangle$, or using some form of compact notation $|\psi, \phi\rangle$, depending on the system in question. This report will favour a compact notation in cases where such notation is natural. Each component of a composite system

is considered a degree of freedom, be it different spatial degrees of freedom of the same system or completely unrelated degrees of freedom (e.g. an electron's position and spin). The Hamiltonian of a composite system with N degrees of freedom is equal to the sum of the individual Hamiltonians for each component plus the interaction:

$$\hat{\mathcal{H}}_{\text{composite}} = \sum_{i=1}^{N} \hat{\mathcal{H}}_i + \hat{\mathcal{H}}_{\text{interaction}} \qquad (3)$$

where $\hat{\mathcal{H}}_{\text{interaction}}$ encompases all interactions between the N degrees of freedom.

For infinite-dimensional Hilbert spaces, the inner product between $|\psi\rangle$ and $|\phi\rangle$ is

$$\langle\psi|\phi\rangle = \int \mathrm{d}^3\mathbf{r}\, \psi(\mathbf{r})^* \phi(\mathbf{r}) \qquad (4)$$

Similarly, the expected value of some operator $\hat{A}$ for a system in a state $|\psi\rangle$ is

$$\langle\psi|\hat{A}|\psi\rangle = \int \mathrm{d}^3\mathbf{r}\, \psi(\mathbf{r})^* \hat{A}\psi(\mathbf{r}) \qquad (5)$$

In case $|\psi\rangle$ is an eigenket of $\hat{A}$ with corresponding eigenvalue a, (5) is simplified to

$$\langle\psi|\hat{A}|\psi\rangle = \langle\psi|a|\psi\rangle = a\langle\psi|\psi\rangle = a \qquad (6)$$

assuming $|\psi\rangle$ is normalised. Expanding the integral in (6) one finds that the wave functions simplify to $|\psi(\mathbf{r})|^2$: the probability density of finding the system at $\mathbf{r}$.

2. Natural (atomic) units

To simplify the mathematics and computations in this project, natural atomic units will be adapted. The particular system of units adapted in this project define

$$\hbar = e = m_e = k_e = k_B \quad \text{and} \quad c = \frac{1}{\alpha} \approx 137 \qquad (7)$$

where e is the elementary charge, k_e is Coulomb's constant, k_B is the Boltzmann constant and α is the fine structure constant. Written in terms of these units the Hamiltonian takes the form

$$\hat{\mathcal{H}} = -\frac{1}{2m}\nabla^2 + V(\mathbf{r}, -i\nabla) \qquad (8)$$

where m is given in terms of m_e. Energy is expressed in units of "Hartree energy" E_h:

$$\alpha = \frac{k_e e^2}{\hbar c} \implies E_h = \frac{m_e k_e^2 e^4}{\hbar^2} = m_e c^2 \alpha^2 = 1$$

In the framework of atomic units, it is common to measure the dimension of length in *atomic units*: a.u.

3. The variational principle & the variational method

The variational principle (not to be confused with the variational method, see below) is a simple statement that is true for all quantum systems. It states that the expected energy of a quantum system in a state $|\psi\rangle$ is greater than or equal to the ground state energy:

$$E_g \leq \frac{\langle\psi|\hat{\mathcal{H}}|\psi\rangle}{\langle\psi|\psi\rangle} \qquad (9)$$

Here $|\psi\rangle$ is not necessarily normalized (which is way the expected energy is divided by the norm). The equation becomes an equality only when $|\psi\rangle$ is the ground state.

The variational principle is the basis for the *variational method*, which exploits the simplicity of the varitional principle in order to approximate the ground state energy. The primary problem is to study some Hamiltonian $\hat{\mathcal{H}}$ whose energy eigenkets are unknown. The idea is to introduce a so-called *trial state* $|\psi_T(\gamma)\rangle$, which is dependent on some parameter γ (not necessarily scalar). If chosen correctly, the trial state may be able to reproduce the functional shape of the ground state, thus providing an approximate upper bound E_T of the ground state energy. The variational method is actually an umbrella term for several methods, however, their underlying idea is more or less the same. The process is most easily shown as an algorithm:

Algorithm 1 The Varitational Method

1: Suggest a trial state $|\psi_T(\gamma)\rangle$.
2: Compute the trial energy E_T as a function of γ:

$$E_T(\gamma) = \frac{\langle\psi_T(\gamma)|\hat{\mathcal{H}}|\psi_T(\gamma)\rangle}{\langle\psi_T(\gamma)|\psi_T(\gamma)\rangle}$$

3: Minimise E_T with respect to γ.
4: Choose the optimal $|\psi_T(\gamma)\rangle$ based on the γ that minimises E_T.
   The resulting optimal E_T is an upper bound on the ground state energy, while the corresponding trial state is the optimal approximation of the ground state.

Due to computational aspects, a common variant of the algorithm loops over the different γ values and minimizes E_T iteratively.

4. The Coulomb interaction

An interesting, although somewhat discouraging, reality is that one of the fundamental interactions of nature, the Coulomb interaction, is more or less analytically unsolvable for everything but the simplest of systems. The basic electron-electron interaction is:

$$V_{e^- e^-} = \frac{k_e e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \qquad (10)$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of electron 1 and 2.

B.   Quantum Mechanics - The Harmonic Oscillator

The harmonic oscillator is one of the finest achivements of quantum mechanics with applications in many areas of physics and in several other sciences. The following section will present the algebraic solution to the harmonic oscillator Hamiltonian, and generalise the results to an N-dimensional system of non-interacting oscillators.

1.   The single harmonic oscillator

The potential energy of a single-degree classical harmonic oscillator with oscillation frequency ω is $V = \frac{1}{2}\omega^2 x^2$, it follows that the Hamiltonian for a single-degree quantum harmonic oscillator with mass m is

$$\hat{\mathcal{H}}_{\text{HO}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}\omega^2 \hat{x} \tag{11}$$

where $\hat{p} = \hat{\mathbf{p}} \cdot \mathbf{e}_x$ is the system's momentum operator. The problem at hand is to find the energy eigenkets of the Hamiltonian. This will be done using the so-called *ladder operators*:[1]

$$\hat{a} = \frac{1}{\sqrt{2m\omega}}\big(+ i\hat{p} + m\omega\hat{x} \big) \tag{12a}$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2m\omega}}\big(- i\hat{p} + m\omega\hat{x} \big) \tag{12b}$$

One can easily show that the ladder operators satisfy the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$. Rewriting $\hat{x}$ and $\hat{p}$ in terms of $\hat{a}$ and $\hat{a}^\dagger$, one finds that the Hamiltonian can be written as

$$\hat{\mathcal{H}}_{\text{HO}} = \omega\left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) = \omega\left(\hat{a}\hat{a}^\dagger - \frac{1}{2} \right) \tag{13}$$

It turns out that the energy eigenkets can be uniquely labelled using a single non-negative integer index n: $|n\rangle$. Furthermore, the action of ladder operators on the energy eigenkets raises or lowers n as follows:

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad \text{and} \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{14}$$

Note that $n < 0$ is avoided as $\hat{a}|0\rangle = 0$.[2] It follows that $\hat{a}^\dagger \hat{a}|n\rangle = n|n\rangle$ such that

$$\hat{\mathcal{H}}_{\text{HO}}|n\rangle = \omega\left(n + \frac{1}{2} \right)|n\rangle = E_n|n\rangle \tag{15}$$

———

[1] These operators have several names, e.g. *raising/lowering* operators, *creation/annahilation* operators, etc.

[2] This is actually somewhat missleading, but it originates in the derivation of the harmonic oscillator energies. By requiring non-negative energies, one finds that the "ladder of energy eigenkets" must be truncated in order to avoid $\langle n|n\rangle < 0$. This truncation is also the origin of the index n.

The energy eigenkets can be uniquely described in terms of $|0\rangle$ and $\hat{a}^\dagger$ via

$$|n\rangle = \frac{\left(\hat{a}^\dagger\right)^n}{\sqrt{n!}}|0\rangle \tag{16}$$

where the factorial accounts for the $\sqrt{n+1}$ scaling in (14).

2.   The virial theorem

The virial theorem is an important result in both classical and quantum mechanics, a simple version of the virial theorem states that the expected total kinetic energy $\langle K \rangle$ is proportional to the expected potential energy $\langle V \rangle$. Peculiar to the harmonic oscillator, provided an energy eigenstate, the expected total kinetic energy is actually equal to the expected potential energy.

Suppose a quantum system subject to a harmonic potential is in the energy eigenket $|n\rangle$. The expected kinetic and potential energies are thus

$$\langle K \rangle_n = \langle n| \hat{K} |n\rangle = \langle n| \frac{\hat{p}^2}{2m} |n\rangle \tag{17a}$$

$$\langle V \rangle_n = \langle n| \hat{V} |n\rangle = \langle n| \frac{1}{2}m\omega^2 \hat{x}^2 |n\rangle \tag{17b}$$

Rewriting $\hat{x}$ and $\hat{p}$ in terms of $\hat{a}$ and $\hat{a}^\dagger$ results in:

$$\hat{x}^2 = \frac{1}{2m\omega}\left(\hat{a}^\dagger + \hat{a} \right)^2 \quad \text{and} \quad \hat{p}^2 = -\frac{m\omega}{2}\left(\hat{a}^\dagger - \hat{a} \right)^2$$

The ladder operator polynomials are:

$$\left(\hat{a}^\dagger + \hat{a} \right)^2 = \hat{a}^\dagger\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger + \hat{a}\hat{a}$$
$$= \hat{a}^\dagger\hat{a}^\dagger + \hat{a}\hat{a} + 2\hat{a}^\dagger\hat{a} + 1$$

$$\left(\hat{a}^\dagger - \hat{a} \right)^2 = \hat{a}^\dagger\hat{a}^\dagger - \hat{a}^\dagger\hat{a} - \hat{a}\hat{a}^\dagger + \hat{a}\hat{a}$$
$$= \hat{a}^\dagger\hat{a}^\dagger + \hat{a}\hat{a} - 2\hat{a}^\dagger\hat{a} - 1$$

Hence,

$$\hat{K}|n\rangle = -\frac{\omega}{4}\Big[\sqrt{(n+1)(n+2)}|n+2\rangle + \sqrt{n(n-1)}|n-2\rangle - 2n|n\rangle - |n\rangle \Big] \tag{18a}$$

$$\hat{V}|n\rangle = +\frac{\omega}{4}\Big[\sqrt{(n+1)(n+2)}|n+2\rangle + \sqrt{n(n-1)}|n-2\rangle + 2n|n\rangle + |n\rangle \Big] \tag{18b}$$

It follows then that

$$\langle K \rangle_n = \langle V \rangle_n = \frac{E_n}{2} = \frac{\omega}{2}\left(n + \frac{1}{2} \right) \tag{19}$$

That is, for a single harmonic oscillator with oscillation frequency ω in an energy eigenstate, the expected kinetic and potential energy is the same.

In case the harmonic oscillator is not in an energy eigenstate, but in some arbitrary state $|\psi\rangle$, it is not necessarily the case that $\langle K \rangle = \langle V \rangle$. Consider the fact that the energy eigenkets span the Hilbert space, thereby implying that

$$|\psi\rangle = \sum_n \alpha_n |n\rangle \qquad (20)$$

for an appropriate choice of α_n. Furthermore, the expected kinetic energy and potential energy is therefore

$$\langle\psi| \hat{K} |\psi\rangle = \sum_n \sum_m \alpha_n^* \alpha_m \langle n| \hat{K} |m\rangle \qquad (21a)$$

$$\langle\psi| \hat{V} |\psi\rangle = \sum_n \sum_m \alpha_n^* \alpha_m \langle n| \hat{V} |m\rangle \qquad (21b)$$

It is not immediately obvious from (21) whether $\langle K \rangle$ is equal to $\langle V \rangle$ or not. However if $\langle K \rangle = \langle V \rangle$, then

$$\langle\psi| \hat{K} |\psi\rangle - \langle\psi| \hat{V} |\psi\rangle = \langle\psi| (\hat{K} - \hat{V}) |\psi\rangle$$
$$= \sum_n \sum_m \alpha_n^* \alpha_m \langle n| (\hat{K} - \hat{V}) |m\rangle$$

must be equal to zero. The braket can be simplified by inserting (18):

$$-\frac{\omega}{2}\left[\sqrt{(m+1)(m+2)}\delta_{n,m+2} + \sqrt{m(m-1)}\delta_{n,m-2} \right]$$

which in turn yields:

$$\langle\psi| (\hat{K} - \hat{V}) |\psi\rangle =$$
$$-\frac{\omega}{2}\sum_n \alpha_n^*\left(\alpha_{n-2}\sqrt{(n-1)n} + \alpha_{n+2}\sqrt{(n+2)(n+1)} \right)$$

As this expression is not necessarily 0 for any state $|\psi\rangle$, it is not necessarily true that $\langle K \rangle$ is equal to $\langle V \rangle$.

3. The harmonic oscillator eigenfunctions

Having found the algebraic solution to the harmonic oscillator problem, the next step is to express the eigenkets in the position basis. The simplest eigenstate is the ground state, for which $E_0 = \omega/2$. Inserting $\hat{x} = x$ and $\hat{p} = -i\partial_x$ into (11) yields:

$$\mathcal{H}_{\text{HO}} = \frac{-1}{2m}\partial_x{}^2 + \frac{1}{2}m\omega^2 x^2$$

To avoid the plethora of constants and variables, consider the dimensionless change of variables:

$$\xi = \sqrt{m\omega}x \quad \text{and} \quad \partial_\xi = \sqrt{m\omega}\partial_x$$

Written in terms of ξ, the equation $\hat{\mathcal{H}} |0\rangle = \frac{1}{2}\omega |0\rangle$ may be written as

$$\frac{-\omega}{2}\partial_\xi{}^2\psi_0(\xi) + \frac{\omega}{2}\xi^2\psi_0(\xi) = \frac{\omega}{2}\psi_0(\xi)$$

where $\psi_0(\xi) = \langle\xi|0\rangle$ is the ground state wave function. The solution is

$$\psi_0(\xi) = \left(\frac{m\omega}{\pi}\right)^{1/4} e^{-\xi^2/2} = Ce^{-\xi^2/2} \qquad (22)$$

Because the ground state is known, equation (16) implies that the excited states may be found by successive application of $\hat{a}^\dagger$. In terms of ξ and ∂_ξ, $\hat{a}^\dagger$ is:

$$a^\dagger = \frac{1}{\sqrt{2m\omega}}\left(-\partial_x + m\omega x \right) = \frac{1}{\sqrt{2}}\left(\xi - \partial_\xi \right)$$

Computing $(a^\dagger)^n \psi_0/\sqrt{n!}$ becomes incredibly inefficient for large n. However, it is actually unecessary because there exists exact solutions in terms of the so-called *Hermite polynomials*. These polynomials, commonly denoted by H, are solutions to the differential equation:

$$\left[\partial_\xi{}^2 - 2\xi\partial_\xi + (\lambda - 1) \right] H(\xi) = 0 \qquad (23)$$

The differential equation has been generalised for any complex λ, but the solutions interesting to the harmonic oscillator have $\lambda = 2n + 1$, where $n = 0, 1, \ldots$ is the energy eigenket index. The Hermite polynomials can be explicitly expressed by Rodrigues' formula:

$$H_n(\xi) = (-1)^n e^{\xi^2}\left(\partial_\xi\right)^n e^{-\xi^2} \qquad (24)$$

from which one can show that the Hermite polynomials satisfy the following recurrence relations:

$$H_{n+1}(\xi) = 2\xi H_n(\xi) - \partial_\xi H_n \qquad (25a)$$

$$\partial_\xi H_n = 2n H_{n-1}(\xi) \qquad (25b)$$

Moreover, the Hermite polynomials satisfy the orthogonality condition:

$$\int_{-\infty}^{\infty} d\xi\, e^{-\xi^2} H_n(\xi)H_m(\xi) = \delta_{nm}2^n n!\sqrt{\pi} \qquad (26)$$

The harmonic oscillator energy eigenfunctions are expressed in terms of the Hermite polynomails as:

$$\psi_n(x) = \frac{C}{\sqrt{2^n n!}}H_n(\xi)e^{-\xi^2/2}, \quad \xi = \sqrt{m\omega}x \qquad (27)$$

(where C is the same as before.) While this equation may seem completely unmotivated, it can easily be shown from an argument of induction. The $n = 0$ case is easy:

$$\langle\xi|0\rangle = \frac{C}{\sqrt{2^0 0!}}H_0(\xi)e^{-\xi^2/2} = Ce^{-\xi^2/2} = \psi_0(\xi)$$

Now let $n = k$ and consider $\langle\xi|k+1\rangle$:

$$\langle\xi| \frac{\hat{a}^\dagger}{\sqrt{k+1}} |k\rangle = C\frac{\xi - \partial_\xi}{\sqrt{2(k+1)}}\frac{1}{\sqrt{2^k k!}}H_k(\xi)e^{-\xi^2/2}$$

It follows from (25) that $(\xi - \partial_\xi)H_k(\xi)e^{-\xi^2/2}$ is equal to $H_{k+1}(\xi)e^{-\xi^2/2}$, which completes the induction proof:

$$\langle\xi|k+1\rangle = C\frac{1}{\sqrt{2^{k+1}(k+1)!}}H_{k+1}(\xi)e^{-\xi^2/2} = \psi_{k+1}(\xi)$$

Note that the question of orthonormality is ensured by the orthonormality of the Hermite polynomials.

4. Composite systems of harmonic oscillators

Having studied the single harmonic oscillator, the final piece of the puzzle is to address a system of harmonic oscillators. Such a composite system could consist of different spatial degrees of freedom of the same particle, different particles or a mixture. Each individual harmonic oscillator behaves according to their Hamiltonian $\hat{\mathcal{H}}_{\text{HO}}$, and the system as a whole obeys equation (3).

Mathematically speaking, the Hamiltonian of a composite system with N harmonic oscillators is

$$\hat{\mathcal{H}}^{\text{sys}} = \hat{\mathcal{H}}_{\text{HO}} \otimes \cdots \otimes \hat{I} + \cdots + \hat{I} \otimes \cdots \otimes \hat{\mathcal{H}}_{\text{HO}} \quad (28)$$

But this is tedius to write, thus let the ith term (i.e., the term where the ith component of the tensor product is $\hat{\mathcal{H}}_{HO}$) be denoted by $\hat{\mathcal{H}}_{\text{HO}}^i$ such that the above expression simplifies to equation (3). The system's energy eigenkets are

$$|\psi\rangle = |n_1, \ldots, n_N\rangle = |n_1\rangle \otimes \cdots \otimes |n_N\rangle \quad (29)$$

where each $|n\rangle$ is the regular harmonic oscillator energy eigenkets. Because $\hat{\mathcal{H}}_{\text{HO}}^i$ only acts on the ith component of the tensor product, only the energy index n_i is affected:

$$\hat{\mathcal{H}}_{\text{HO}}^i |\psi\rangle = \omega_i \left(n_i + \frac{1}{2} \right) |\psi\rangle \quad (30)$$

where ω_i is the oscillation frequency of the ith harmonic oscillator. It follows that the composite system's energy eigenvalues are

$$\hat{\mathcal{H}}^{\text{sys}} |\psi\rangle = \left[\omega_1 \left(n_1 + \frac{1}{2} \right) + \cdots + \omega_N \left(n_N + \frac{1}{2} \right) \right] |\psi\rangle \quad (31)$$

This expression greatly simplifies when the oscillation frequency is uniform (i.e. $\omega_i = \omega$):

$$\hat{\mathcal{H}}^{\text{sys}} |\psi\rangle = \omega \left(n_1 + \cdots + n_N + \frac{N}{2} \right) |\psi\rangle \quad (32)$$

The ground state ($n_i = 0$) energy of the composite harmonic oscillator with uniform frequency is therefore

$$E_g = \omega \frac{N}{2} \quad (33)$$

If all oscillators are in an energy eigenstate, then $\langle K \rangle_n = \langle V \rangle_n$ implies that the total expected kinetic energy is equal to the total expected potential energy.

C. Markov Chain Monte Carlo Methods

1. Monte Carlo integration

One of the most versatile approaches to numerical integration is Monte Carlo integration. Its clever exploit of stochastic variables allows it to tackle high-dimensional integrals with little extra effort. There are several ways to develop a formalism for Monte Carlo integration, the one presented below is adapted to suit the needs of this project.

The goal of Monte Carlo integration is to evaluate integrals on the form:

$$I = \int_D \mathrm{d}x \, f(x), \quad x \in D \quad (34)$$

where x is not necessarily one-dimensional. Now let $p(x)$ denote a known distribution and consider the change of variables from x to $p(x)$:

$$\int_D \mathrm{d}x \, f(x) = \int_{x \in D} \mathrm{d}p(x) \, \frac{f(x)}{p(x)} = \left\langle \frac{f(x)}{p(x)} \right\rangle_{p(x)} \quad (35)$$

That is, the integral I is equal to the expected value of $f(x)/p(x)$ with respect to the probability distribution $p(x)$. The idea is to draw N_{MC} random numbers $X \sim p(x)$ and estimate the expected value with the sample expected value:

$$\left\langle \frac{f(x)}{p(x)} \right\rangle_X \approx \sum_{i=1}^{N_{\text{MC}}} \frac{f(x_i)}{p(x_i)} p(x_i) = \sum_{i=1}^{N_{\text{MC}}} f(x_i) \quad (36)$$

The basic Monte Carlo integration algorithm is summarised below:

Algorithm 2 Standard Monte Carlo Integration

1: Define N_{MC}.
2: Initialise SUM = 0.
3: **for** $i = 1, \ldots, N_{\text{MC}}$:
4:     Draw x_i at random according to $p(x)$.
5:     Evaluate $f(x_i)$.
6:     Add $f(x_i)$ to SUM.
7: **end for**
8: Normalise final value: $I = \text{SUM}/N_{\text{MC}}$.

---

2. Markov chains

A Markov chain is a stochastic model for the discrete evolution of a so-called "memoryless system",[3] i.e. a system whose current state only depends on the previous state of the system. Formally, such a memoryless system satisfies the *Markov property*, which requires that the probability of moving from a state X_n to state X_{n+1} is only dependent on X_n. The state space may either be continuous or discrete, each of which introduce a similar albeit distinct formalism. This project will only use a continuous state space, thus the discrete formalism will be ignored here.

Say the state space for a particular Markov chain is $\mathcal{S}$, then a *stochastic kernel* on $\mathcal{S}$ is a function $p : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ that satisfies

$$p(x, y) \le 0, \ \forall x, y \in \mathcal{S} \quad \text{and} \quad \int_{\mathcal{S}} \mathrm{d}y \, p(x, y) = 1, \ \forall x \in \mathcal{S}$$

Each Markov chain is parametrised by its corresponding stochastic kernel, which in this sense is usually referred to as the transition probability density. Now for every $p(x, y)$ there exists a so-called *Markov operator* P that is such that

$$[(\cdot)P](y) = \int_{\mathcal{S}} \mathrm{d}x \, p(x, y)(\cdot)$$

Note in particular that P is a left-acting operator, as per usual in the literature.

Furthermore, say the Markov chain has continued for n steps with previous measurements $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$, then the next measurement $X_{n+1} = y$ is drawn according to the distribution $p^{(n+1)}(y)$ whose evolution from the current distribution $p^{(n)}(x)$ is given by the Markov operator of the Markov chain:

$$p^{(n+1)}(y) = \left[p^{(n)}(x)P\right](y) = \int_{\mathcal{S}} \mathrm{d}x \, p(x, y)p^{(n)}(x) \quad (37)$$

A special case of (37) is when the distribution is invariant of P, that is, when $p^{(n+1)} = p^{(n)}$. This behaviour is called a stationary distribution and is usually denoted by $\pi(x)$:[4]

$$\pi(y) = \int_{\mathcal{S}} \mathrm{d}x \, p(x, y)\pi(x) \quad (38)$$

Such a state is guaranteed if the Markov chain obeys the so-called *detailed balance* condition. Provided a transition probability density $p(x, y)$ and a particular distribution $\pi(x)$, then the detailed balance condition states that

the Markov chain corresponding to $p(x, y)$ is *reversible* with respect to $\pi(x)$ if

$$\pi(x)p(x, y) = \pi(y)p(y, x), \ \forall x, y \in \mathcal{S} \quad (39)$$

It follows that

$$\int_{\mathcal{S}} \mathrm{d}x \, p(x, y)\pi(x) = \pi(y) \int \mathrm{d}x \, p(y, x) = \pi(y)$$

because $p(x, y)$ is normalised for all $x, y \in \mathcal{S}$. In conclusion, in case a Markov chain is "well-behaved", meaning it satisfies detailed balance, then the so-called *limiting distribution* of the Markov chain approaches π as $n \to \infty$:

$$\lim_{n \to \infty} \left[p^{(0)}(x)P^n\right](y) = \pi(y) \quad (40)$$

where P^n implies repeated operations on $p^{(0)}$, $p^{(1)}$, etc.

3. Markov Chain Monte Carlo: The Metropolis algorithm

As the name implies, Markov Chain Monte Carlo (MCMC) methods combine the concept of a Markov chain with Monte Carlo methods. There exists several MCMC methods, each with unique properties that may or may not be benefitial. This project will focus on the so-called *Metropolis* algorithm, which is a special case of the *Metropolis-Hastings* algorithm.

The basic idea behind the MCMC methods stems from equation (36). Ideally one would be able to draw a sufficiently large number of states x from $p(x)$ and follow algorithm 2 religiously. However, this is incredibly computation-ineffective and essentially infeasible for a standard computer. Enter Markov chains: A Markov chain with limiting distribution $\pi(x) = p(x)$ would necessarily generate states that are approximately distributed according to $p(x)$. It turns out that this may be exploited in order to avoid large computations.

The mathematics of MCMC methods are based on the detailed balance principle, consider the following rearranging of equation (39):

$$\frac{p(x, y)}{p(y, x)} = \frac{\pi(y)}{\pi(x)}$$

Now, suppose the current state of the Markov chain governed by $p(x, y)$ is X_n, i.e., the nth state in the chain. The idea is to separate the transition process into two steps: an initial *proposal*, and an *acceptance/rejection* step. These steps should be independent so that probability is independent, meaning the probability density of proposing a state y, $g(y|x)$, is independent of the probability of accepting the proposal, $\alpha(x, y)$. Because they are independent, it follows that

$$p(x, y) = g(y|x)\alpha(x, y) \quad (41)$$

meaning the above criteria may be rewritten as

$$\frac{\alpha(y, x)}{\alpha(x, y)} = \frac{p(y)}{p(x)} \frac{g(x|y)}{g(y|x)} \quad (42)$$

---

[3] A Markov chain may be generalised to continuous evolution, but this will not be used in this project.

[4] Becuase choosing π to represent something else other than π seemed like a great idea, obviously.

where $p(x) = \pi(x)$ is the assertion of MCMC methods that the limiting distribution of the Markov chain governed by $p(x, y)$ is the integrand of equation (34). This expression is particularly easy to implement in case the normalisation of p and g are independent on x and y, or if they happen to cancel each other (although this is very unlikely).

The last step it to choose an acceptance probability $\alpha(y, x)$ that satisfies 42, the *Metropolis choice* is

$$\alpha(y, x) = \min\left\{1, \frac{p(y)}{p(x)}\frac{g(x|y)}{g(y|x)}\right\} \qquad (43)$$

The Metropolis-Hastings algorithm may now be stated:

Algorithm 3 The Metropolis-Hastings Algorithm

1: Initialise the first state X_0.
2: **for** $n = 1, \ldots, N_{\text{MH}}$:
3:     Generate a proposal y according to $g(y|x)$.
4:     Evaluate acceptance probability $\alpha(y, x)$.
5:     **if** $\alpha(y, x) = 1$:
6:         Accept y.
7:     **else**
8:         Draw a uniformly distributed number $a \in [0, 1]$.
9:         **if** $a \leq \alpha(y, x)$:
10:             Accept y and set $X_{n+1} = y$.
11:         **else**
12:             Reject y and set $X_{n+1} = X_n$.
13:         **end if**
14:     **end if**
15: **end for**
---

where N_{MH} is the number of steps in the Markov chain. Note that the algorithm above does not factor in the Monte Carlo update, this is because the central purpose of the Metropolis-Hastings algorithm is namely just to sample x values according to $p(x)$. Nonethelesss, implementations of the algorithm usually include an intermediate step inside this loop in order to avoid multiple loops.

Furthermore, although the distribution is guaranteed to converge, it does not necessarily converge within the first few steps. Accordingly, common practice is to perform N_{burn} preparation cycles (before beginning the Monte Carlo integration) known as a "Metropolis burn-in". When a step in the Markov chain represents a forward step in time, the amount of necessary burn-in is commonly referred to as the equilibration time of the system.

As previously mentioned, this project will implement the Metropolis algorithm, which is a special case of the Metropolis-Hastings algorithm. The Metropolis assumes that g is symmetric in x and y, meaning $g(x|y) = g(y|x)$. This implies that the acceptance probability may be simplified to

$$\alpha_{\text{M}}(y, x) = \min\left\{1, \frac{p(y)}{p(x)}\right\} \qquad (44)$$

III. METHOD

The aim of this project is to study 2 electrons in a uniform harmonic oscillator potential who interact via the Coulomb interaction. Despite its simplicity, there does not exist an analytical solution to the problem, which is why the system has been extensively studied both analytically and computationally. Before continuing to the computational aspects of the project, some preparation is required.

A. Preparation

1. Some exact results

Say a single electron is located in a three-dimensional harmonic oscillator (with uniform oscillation frequency), it follows from (29) and (27) that the system's wave function $\psi(x, y, z)$ (ignore spin for the time being) is proportional to[5]

$$H_{n_x}(\sqrt{\omega}x)H_{n_y}(\sqrt{\omega}y)H_{n_z}(\sqrt{\omega}z)e^{-\frac{\omega}{2}(x^2 + y^2 + z^2)}$$

where n_x, n_y and n_z are the energy labels of the individual oscillators. This is clearly symmetric under the interchange of any two energy labels, hence the spatial state of the electron is symmetric. Introducing another electron to the picture results in a total wave function of

$$\Psi = \psi(x_1, y_1, z_1) \cdot \psi(x_2, y_2, z_2)$$

which clearly is symmetric under the interchange of particles. As electrons are fermions, it is necessary for their complete state to be anti-symmetric under the interchange of particles, and because the spatial state is symmetric, it is therefore necessary for the spin state to be anti-symmetric. Out of the four possible combinations of two spin-1/2 fermions, there is only one anti-symmetric state, the *singlet* state:

$$|\chi\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big)$$

whose total spin is 0. Lastly, it follows from (32) that the energy of this electron-electron system is

$$E_{\text{HO}}^{\text{tot}} = \omega\big(n_{x_1} + n_{x_2} + n_{y_1} + n_{y_2} + n_{z_1} + n_{z_2} + 3\big) \quad (45)$$

Hence, the ground state energy is 3ω.

Now consider the same pair of electrons, but with an additional Coulomb interaction. The Hamiltonian for this system is

$$\hat{\mathcal{H}}^{\text{sys}} = \hat{\mathcal{H}}_{\text{HO}}^1 + \hat{\mathcal{H}}_{\text{HO}}^2 + \hat{V}_{e^-e^-} \qquad (46)$$

---

[5] The normalization constant is ignored as the variational method accounts for unnormalized states.

where $\hat{\mathcal{H}}^i_{HO}$ denotes the three-dimensional harmonic oscillator Hamiltonian for particle i and $V_{e^- e^-}$ is the electron-electron Coulomb interaction potential. This is the system to be studied in the project, thus any future reference to the "*electron-electron system*" should be interpreted explicitly as this refering to (46). Written in the position basis the Hamiltonian becomes

$$\mathcal{H}^{\text{sys}} = -\frac{1}{2}\big(\nabla_1^2 + \nabla_2^2\big) + \frac{1}{2}\omega^2(r_1^2 + r_2^2) + \frac{1}{r_{12}} \quad (47)$$

2. Trial wave functions

The project will use the following two trial wave functions:

$$\psi_T^1(\mathbf{r}_1, \mathbf{r}_2) = e^{-\alpha\frac{\omega}{2}(r_1^2 + r_2^2)} \quad (48a)$$

$$\psi_T^2(\mathbf{r}_1, \mathbf{r}_2) = e^{-\alpha\frac{\omega}{2}(r_1^2 + r_2^2)} e^{\frac{r_{12}}{2(1+\beta r_{12})}} \quad (48b)$$

where $\mathbf{r}_i = (x_i, y_i, z_i)$, $r_i^2 = x_i^2 + y_i^2 + z_i^2$, $r_{12} = ||\mathbf{r}_1 - \mathbf{r}_2||$, and α and β are variational parameters. The former is the solution to the non-interacting system when $\alpha = 1$, while the latter is a perturbation of the former. The additional factor in ψ_T^2 is commonly known as a "*Jastrow factor*". Note that both waves technically include the non-interacting solution as $\psi_T^2 \to \psi_T^1$ when $\beta \to \infty$.

Clearly, ψ_T^1 and ψ_T^2 are not energy eigenstates of $\mathcal{H}^{\text{sys}}$, so they do not have sharp energies. The average energy is given by $\langle \psi_T | \hat{\mathcal{H}}^{\text{sys}} | \psi_T \rangle / \langle \psi_T | \psi_T \rangle$, which means it is interesting to know how $\mathcal{H}^{\text{sys}}\psi_T$ looks like. Applying (47) to the trial wave functions yields:

$$\mathcal{H}^{\text{sys}}\psi_T^1 = E^1 \psi_T^1 \quad \text{and} \quad \mathcal{H}^{\text{sys}}\psi_T^2 = E^2 \psi_T^2 \quad (49)$$

where

$$E^1 = 3\alpha\omega + \frac{1}{2}\omega^2(r_1^2 + r_2^2)(1 - \alpha^2) + \frac{1}{r_{12}} \quad (50a)$$

$$E^2 = E^1 + \frac{1}{2(1 + \beta r_{12})^2}\left[\alpha\omega r_{12} - \frac{2}{r_{12}}\right.$$
$$\left. + \frac{2\beta}{1 + \beta r_{12}} - \frac{1}{2(1 + \beta r_{12})^2}\right] \quad (50b)$$

Strictly speaking, E^1 and E^2 in (49) are not energies as
a) ψ_T^1 and ψ_T^2 are not normalized and
b) they depend on $\mathbf{r}_1$ and $\mathbf{r}_2$.

B. Variational Monte Carlo Algorithm

The "Variational Monte Carlo Algorithm" is a mixture of the variational method and Markov Chain Monte Carlo methods. Recall the definition of the trial energy $E_T(\gamma)$ in the variational method:

$$E_T(\gamma) = \frac{\langle \psi_T(\gamma) | \hat{\mathcal{H}} | \psi_T(\gamma) \rangle}{\langle \psi_T(\gamma) | \psi_T(\gamma) \rangle}$$

Inserting ψ_T^i and (50) into $E_T(\gamma)$ yields:

$$E_T^i(\alpha, \beta) = \frac{\int d\mathbf{r}_1 \, d\mathbf{r}_2 \, E^i(\mathbf{r}_1, \mathbf{r}_2)|\psi_T^i|^2}{\int d\mathbf{r}_1 \, d\mathbf{r}_2 \, |\psi_T^i|^2}$$

Within the contex of variational Monte Carlo, $E^i(\mathbf{r}_1, \mathbf{r}_2)$ is usually referred to as the *local* energy. Note that the bottom integral is just a number, specifically the norm of ψ_T^i. With this in mind, define the probability density:

$$p_{\alpha,\beta}^i = \frac{|\psi_T^i|^2}{\int d\mathbf{R} \, |\psi_T^i|^2} \quad (51)$$

where $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2)$ is the state of the electron-electron system. Inserting $p_{\alpha,\beta}^i$ back into $E_T^i(\alpha, \beta)$ yields:

$$E_T^i(\alpha, \beta) = \int d\mathbf{R} \, E^i(\mathbf{r}_1, \mathbf{r}_2) p_{\alpha,\beta}^i \quad (52)$$

This is on the form of equation (35) and thus the integral may be estimated by the sample exected value of $E_T^i(\mathbf{R})$ with random values $\mathbf{R}$ drawn from $p_{\alpha,\beta}^i$.

To sample random values directly from $p_{\alpha,\beta}^i$ is in general quite difficult (mostly because of the integral in the denominator), thus the (symmetric) Metropolis algorithm will be used instead. The stability of the Metropolis sampling will be verified using the variance of $E_T^i(\alpha, \beta)$. Apart from including a term E^2 in the algorithm, no additional adjustments are necessary to encompass the variance estimation.

Let $\mathbf{R}_n$ denote the state at step n in the Markov chain. The Metropolis choice is then

$$\alpha_M^i(\mathbf{R}_{n+1}, \mathbf{R}_n) = \min\left\{1, \frac{p_{\alpha,\beta}^i(\mathbf{R}_{n+1})}{p_{\alpha,\beta}^i(\mathbf{R}_n)}\right\} = \min\{1, A_i\}$$

where A_i is introduced for simplicity. By definition, A_i must be equal to $|\psi_T^i(\mathbf{R}_{n+1})|^2/|\psi_T^i(\mathbf{R}_n)|^2$, hence:

$$A_1 = \exp\left[-\alpha\omega\big(||\mathbf{R}_{n+1}||^2 - ||\mathbf{R}_n||^2\big)\right] \quad (53a)$$

$$A_2 = A_1 \exp\left[\frac{r_{12,n+1}}{1 + \beta r_{12,n+1}} - \frac{r_{12,n}}{1 + \beta r_{12,n}}\right] \quad (53b)$$

where $||\mathbf{R}_n||^2 = r_{1,n}^2 + r_{2,n}^2$.

The symmetric Metropolis algorithm requires that $p_{\alpha,\beta}^i(\mathbf{R}|\mathbf{R}') = p_{\alpha,\beta}^i(\mathbf{R}'|\mathbf{R})$. One way to accomodate this is to propose new states $\mathbf{R}'$ from $\mathbf{R}$ according to

$$\mathbf{R}' = \mathbf{R} + \hat{\delta} \quad (54)$$

where $\hat{\delta} \sim \mathcal{U}_6(-\delta, \delta)$ is a vector of random numbers. Now, the idea is to control the average Metropolis sampling acceptance ratio using δ. Ideally the acceptance ratio should be $\approx 50\%$, meaning $A \approx 1/2$. Furthermore, note that (53) treats $\alpha\omega$ as a single variable, hence, so should the choice of δ. Furthermore, the triangle inequality implies that:

$$\frac{1}{2} = e^{-\alpha\omega\big(||\mathbf{R}_{n+1}||^2 - ||\mathbf{R}_n||^2\big)} \leq e^{-\alpha\omega||\hat{\delta}^2||^2}$$

meaning $\ln(2) \geq \alpha\omega||\hat{\delta}||^2$. So far it is unclear how δ is related to $||\hat{\delta}||^2$, however, trial and error have shown that $\delta = \ln(2)$ yields an acceptance rate of about 52% when $\alpha\omega = 1$. For $\alpha\omega$ values below 1, $\ln(2)$ becomes a little to small as acceptance quickly soars above 80%. Knowing that (53) is exponential, it is likely that A is logarithmic in $\alpha\omega$. Hence, propose the relationship $\delta = \ln(\alpha\omega + 1)$, which becomes $\ln(2)$ when $\alpha\omega = 1$. It turns out that this is not enough for small $\alpha\omega$ values. What is needed is some functional that is zero for $\alpha\omega$ values close to 1, but peaks for small values. The easiest example of such a relationship is something like:

$$\frac{1}{x + a} - \frac{1}{1 + a}$$

where a is parameter. After some trial and error, the following $\delta(\alpha\omega)$ function was found to work very well for most $\alpha\omega$ values:

$$\delta(\alpha\omega) = \ln(\alpha\omega + 1) + \frac{1}{\alpha\omega + \frac{1}{4}} - \frac{1}{1 + \frac{1}{4}} \qquad (55)$$

In fact, the formula can be comfortably used for both A_1 and A_2 with $\alpha\omega$ values greater than 0.02, yielding an acceptance ratio of approximately $50 \pm 4\%$. For $\alpha\omega$ vaues below 0.02 the acceptance ratio increases rapidly, reaching somewhere between 70% and 80% already by $\alpha\omega = 0.01$.

The final Variational Monte Carlo Algorithm is listed in Algorithm 4. Note that it does not loop over α and β, this is intential (different experiments loop over different parameters). It is essentially a mixture of Algorithms 3 and 2, however with a clever manipulation of the double if-test inside the Metropolis loop. The initial if-test is actually uneccesary if α_M is directly defined as $p_{\alpha,\beta}^i(\mathbf{R}')/p_{\alpha,\beta}^i(\mathbf{R})$. This follows because of the fact that $b \leq 1$ (see algorithm). Consider the following: If A_i is greater than 1, then it is necessarily greater than b. Hence, defining α_M as the ratio of the probabilities and ignoring the first if-test produces the same result as what is described in Algorithm 3.

Algorithm 4 The Varitational Monte Carlo Algorithm

1: Initialize α, β and i.
2: Initialize N_{MC}, δ and $\mathbf{R}$.
3: Initialize $E = 0$ and $(E^2) = 0$.
4: **for** $n = 1, \ldots, N_{MC}$
5:     Propose $\mathbf{R}' = \mathbf{R} + \hat{\delta}$, where $\hat{\delta} \sim \mathcal{U}_6(-1, 1)$.
6:     Compute: $\alpha_M = A_i$
7:     Draw b according to $\mathcal{U}(0, 1)$.
8:     **if** $b \leq \alpha_M$:
9:         Update $\mathbf{R} = \mathbf{R}'$.
10:     **end if**
11:     Add $E^i(\mathbf{R})$ to E and $E^i(\mathbf{R})^2$ to (E^2).
12: **end for**
13: Normalize the Monte Carlo integrals:

$$E_T^i(\alpha, \beta) = \frac{E}{N_{MC}} \quad \text{and} \quad \text{Var}\big[E_T^i(\alpha, \beta)\big] = \frac{(E^2)}{N_{MC}} - \left(\frac{E}{N_{MC}}\right)^2$$

C. Experiments

The following section discusses the experiments run in this project. For simplicity, let "wave 1" and "wave 2" refer to trial wave functions (48a) and (48b) respectively.

1. Experiment 1 - Monte Carlo stability

Before analysing the variational parameter space it is important run a Monte Carlo cycle cost analysis such that computational power is not uneccessarily spent on immeasureable details. Because the Monte Carlo algorithm is only tied to the system through (55), which only regulates the acceptance ratio of the algorithm, it is thus reasonable to assume all systems (meaning all system-parameter combinations) will behave fairly similar with respect to the number of Monte Carlo cycles. For simplicity then, only the wave 1 system with $\alpha = \omega = 1$ is considered.

The goal of this experiment is to find a suitable balance between the accuracy of the Monte Carlo integrals and the efficiency of the algorithm. To this extent, 21 unique N_{MC} values logarithmically spaced between 10^2 and 10^7 are studied. Expectedly however, simulations with a small number of Monte Carlo cycles will probably not have their integrals converge close to the true values. Therefore, each simulation will be repeated 50 times in order to survey the distribution of integral values with respect to Monte Carlo cycles. Seeing that the point of the experiment is to quantify stability, no Metropolis burn-in is considered.

2. Experiment 2 - Wave 1 optimization

The aim of this experiment is to analyse the parameter space of wave 1, namely $\{\alpha\}$, in order to minimize the trial energy. Seeing that the energy minimum of the unperturbed system has $\alpha = 1$, it natural to investigate α values close to 1. The experiment consists of 5 repetitions of 100 α values between 0.5 and 1.5. The repetitions are introduced to ensure the integrals converge to a satisfactory margin.

It is unlikely that the optimal α is exactly one of the 100 values, thus the final minimum is found using a polynomial interpolation of the $E_T(\alpha)$ data points. The degree of the polynomial is chosen such that the functional form of the data is captured, while avoiding overfitting.

Having found the optimal α, the final task of this experiment is to study the ground state's electron separation, in particular as a function of the harmonic oscillator frequency. Using the optimal α, the simulation is run 5 times for 100 ω values between 0.05 and 1.00.

3. Experiment 3 - Wave 2 optimization

Mirroring experiment 2, the aim of this experiment is to analyse the parameter space of wave 2, namely $\{\alpha, \beta\}$, in order to minimize the trial energy. As opposed to wave 1's one-dimensional parameter space, wave 2's parameter space is two-dimensional, and thus requires many more simulation runs. The naive approach to this optimization problem is to select a very large grid of (α, β) values with very fine spacing and then wait for 100 years. Obviously this is unfeasible, hence the process must be divided into sections. The idea is still the same: run simulations for a large grid of (α, β) values, however with reasonable grid spacing. Next, find the region of the grid where the minimum is most likely located, then choose this region as the next grid with a more fine grid spacing. And so on. This iterative process will naturally end when the difference in trial energies between neighbouring grid points becomes less than the fluctuations in the energy integrals due to Monte Carlo instability.

Because wave 2 is a perturbation of wave 1, it is reasonable to assume that α plays a similar role for wave 2, hence the optimal α value found in the previous experiment is used as a center of the initial range of α values studied for wave 2. The radius of this range is set to approximately 25% of the optimal α. There is no particular range of β values that is more likely to contain the optimal β value. However, in complete faith of beautiful mathematics, the β value is probably closer to 1 than to say 10000, thus the range [0.25,1.50] is used in the initial grid.

Arriving on the final grid search, the final optimal α and β values are found using a polynomial interpolation of the trial energy data points. Again, the polynomial's degree is selected such that the functional form of the data set is preserved, while avoiding overfitting.

With the optimal variational parameters, the ground state's electron separation is again studied as a function of the harmonic oscillator frequency. Again, the simulation is run 5 times for 100 ω values between 0.05 and 1.00.

4. Experiment 4 - Energy balance

As a final test of the optimal variational wave functions, the virial theorem is put to the test. For this experiment, only the optimal trial wave functions are used. The goal is to estimate the ratio $\langle V \rangle / \langle K \rangle$ (potential to kinetic energy) as a function of the harmonic oscillator frequency. Moreover, of academic interest, the potential-kinetic energy ratio will be computed with and in the absence of the interaction term (i.e. $1/r_{12}$) in order to study its effects on the balance between kinetic and potential energy.

The simulations are run for 100 ω values between 0.05 and 1.00, and repeated 5 times in order to capture any spread in the kinetic and potential energies. Because there are 4 different scenarios (2 waves with or without the interaction), the total number of simulations is $4 \cdot 5 \cdot 100 = 2000$. Having produced the results, the energy ratios are displayed as functions of ω.

Note that the variational parameters are chosen using the previous experiments, which all included the interaction. It follows that the non-interacting scenarios will not achive $\langle V \rangle / \langle K \rangle = 1$, as is true for the harmonic oscillator in the ground state.

IV. RESULTS

A. Experiment 1

The distribution of the trial energy and energy variance integrals as functions of Monte Carlo cycles is shown in figure 1. As a general rule, the results seem to converge on $E_T \cong 3.8$ after about 10^5 cycles and on $\mathrm{Var}[E] \cong 0.35$ after just under 10^7 cycles. In addition to the energy integrals, figure 1 also features the simulations' corresponding distribution of average acceptance ratios (accepted cycles divided by total cycles). This acceptance ratios mirror the trial energy distribution and converges on about 50% after some 10^5 cycles.

A common trait shared by all distributions is that the convergence value is more or less the expected value of the distributions. This is important because it implies that there are two ways of improving the integral estimates: either by increasing the number of Monte Carlo cycles or by using the average of several simulations. For simulations with many Monte Carlo cycles, this is particularly important because running several simulations with the same number of cycles may be computationally benefitial to say increasing the number of cycles of a single run in order to achieve the same accuracy.

To strike a compromise between efficiency and precision, future experiments will be run using a Metropolis burn-in of 10^5 cycles and 10^5 cycles for the Monte Carlo integration. And if further accuracy is needed, the average over several runs is used instead of increasing the number of cycles.

B. Experiment 2

The results of the experiment are shown in figure 2. Most notably, there is good agreement between the trial energy minimum and the energy variance minimum, in addition the acceptance ratio is fairly well-behaved in the region surrounding the energy minimum (ca. $50 \pm 2\%$).

In experiment 1, the energy variance distribution was found to converge after some 10^7 cycles. Hence as expected, the tendency of energy variance to fluctuate is slightly greater than for the trial energy and acceptance ratio (seeing that only 10^5 cycles was used). Nevertheless, the additional runs successfully capture the shape of the energy variance curve (as predicted).

As seen in figure 3, the trial energy data was regressed against a polynomial of degree 5. In descending powers of α, the polynomial coefficients are approximately $[-2.83, 16.35, -38.41, 46.68, -29.03, 11.03]$. The polynomial has an approximate local energy minimum of about 3.773 Hartree using an optimal α of about 0.8794. There is very good agreement between the polynomial and the integrals, specifically there is little indication of overfitting.

The expected electron separation using the optimal α is shown in figure 4. As expected, the electrons are, on average, closer when the oscillator's strength is greater, and vice versa. As ω tends to zero, the separation increases rapidly. There are two factors likely responsible for this: the replusive Coulomb-interaction between the electrons, and the decreasing attraction of the harmonic oscillator.

The expected seperation flattens out around 2 a.u. for increasing ω. This is not unusual behaviour as the electrons repell eachother. The small gradient of the expected separation graph in figure 4 is also not unexpected as the electric repulsion goes as $1/r$.

Interestingly, the acceptance ratio also increases for small ω. This is mostly likely due to the behaviour of (55) for small $\alpha\omega$. The result of the increased acceptance can be seen in the spread of the expected separation graph in figure 4.
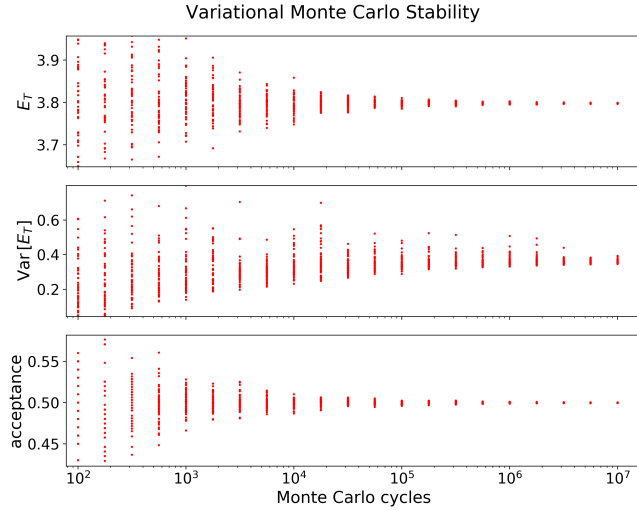

Figure 1: The stability of the Variational Monte Carlo algorithm, measured according to the convergence of the trial energy and energy variance integrals of wave 1. The lower plot shows the convergence of the acceptance ratio distribution, which is measure of the accuracy of the convergence.
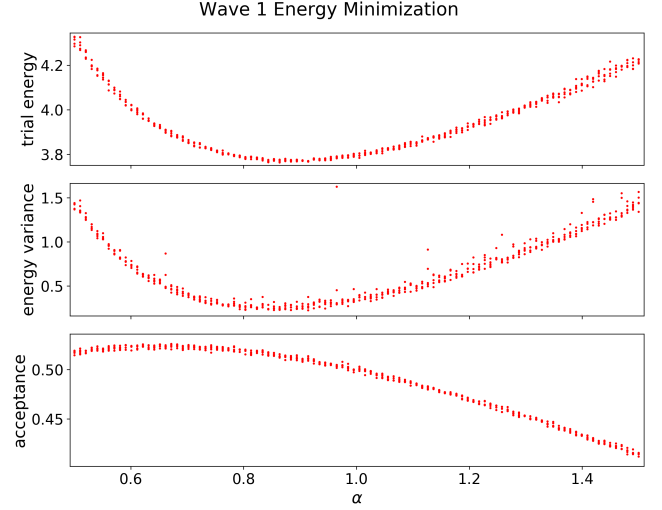

Figure 2: Variational Monte Carlo energy minimization of wave 1. The optimal α is approximately 0.8794, the corresponding energy minimum is about 3.773 Hartree.

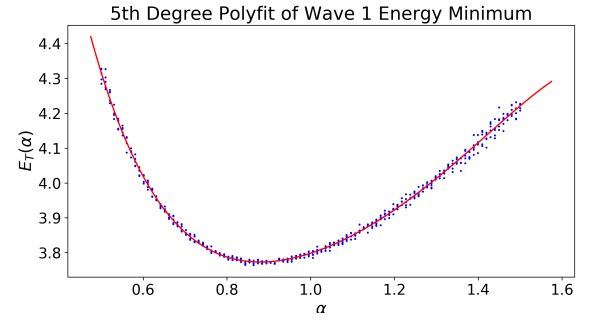Figure 3: 5th degree polynomial interpolation of the trial energy data in figure 2.
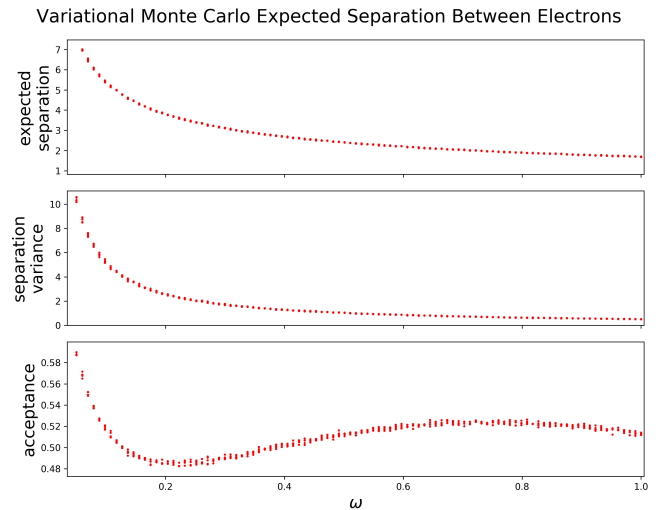

Figure 4: The expected separation between the electrons using the optimized wave 1.

C. Experiment 3

The results of the initial and final grid searches are shown in figure 6, the intermediate grid searches can be found on the GitHub repository. The grid search reveals a clear energy minimum located somewhere close to $(\alpha, \beta) = (0.95, 0.25)$.

The final grid search is shown in figure 7, it was found to behave quite well with respect to a 5th order polynomial regressor.[6] The polynomial interpolation is show in the upper image in figure 8, the lower image shows the interpolation residuals from the trial energy data. The size of the overal residual (i.e. the length of the residual vector) is about $6.24 \cdot 10^{-3}$. Note that the colour-coding in figure 7 does not correspond to the colour-coding in figure 8, despite having the same Matplotlib colormap.

The minimum energy of the polynomial interpolation was found to be about 3.730 Hartree using optimal variational parameters $\alpha = 0.9990$ and $\beta = 0.2715$. This energy minimum is about 0.043 Hartree less than the wave 1 energy minimum, suggesting wave 2 is a better approximation of the ground state wave function than wave 1.

Using the optimal parameters the expected distance between the electrons was computed for the various oscillator frequencies, the results are shown in figure 5. The results are very similar to the results of experiment 2 (figure 4). However, the wave 2 graph increases faster than the wave 1 graph as ω decreases. Otherwise, the same comments made on 4 also apply to 5; the graphs behave similarly with respect to ω and generally predict similar values.
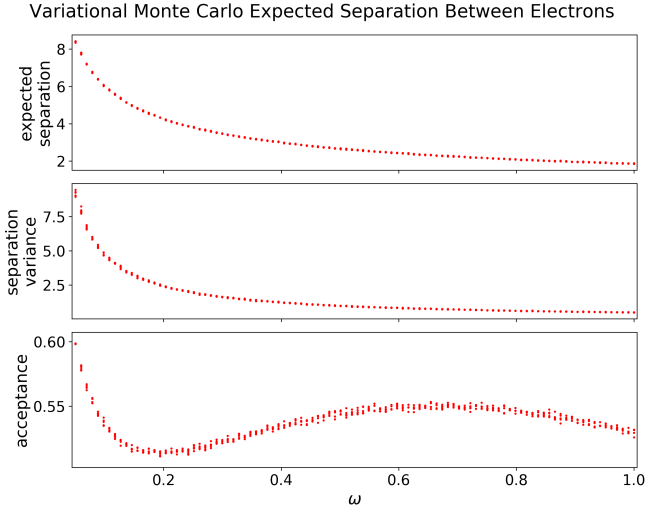

Figure 5: The expected separation between the electrons using the optimal wave 2 variational parameters $\alpha = 0.9990$ and $\beta = 0.2715$.

---

[6] If needed, the coefficients for this polynomial regressor can be accessed using code on the GitHub repository.
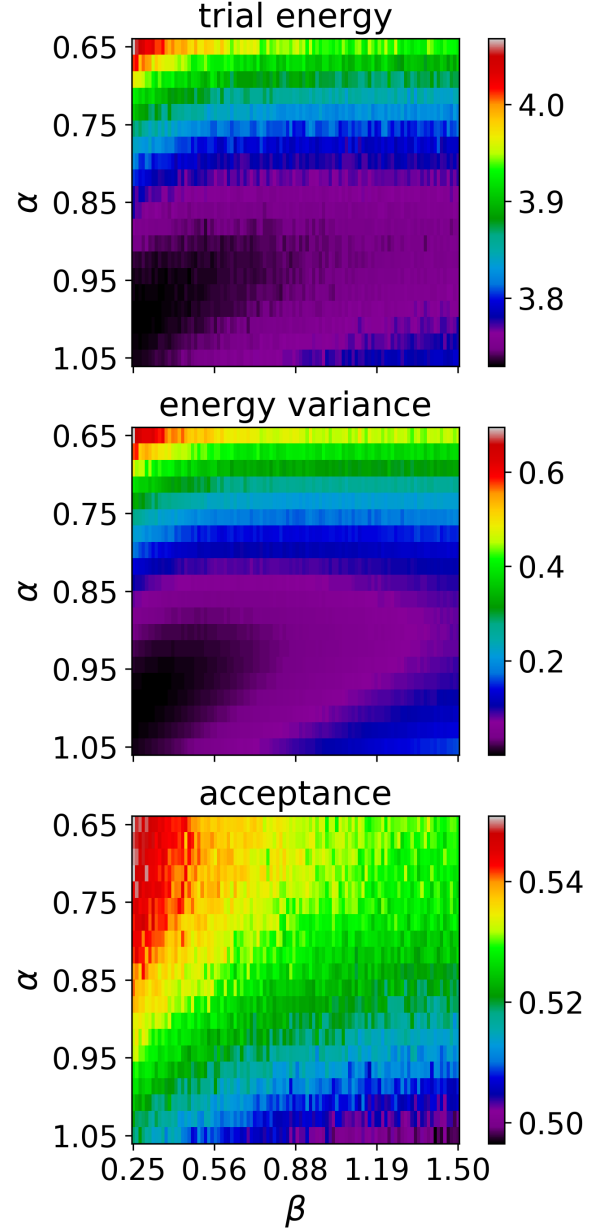

Figure 6: Variational Monte Carlo energy minimization of wave 2. The minimum seems to be centered close to $(\alpha, \beta) = (0.95, 0.25)$. The results after optimizing the parameter grid is shown in figure 7.

Wave 2 Energy Minimization

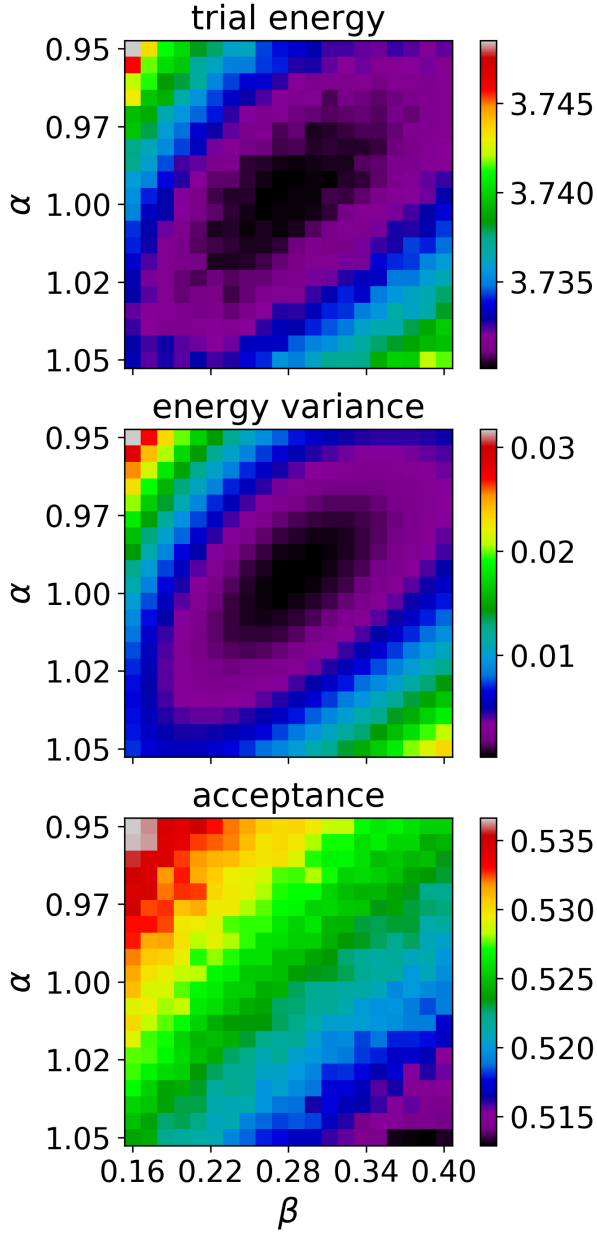### trial energy

energy variance

acceptance

Figure 7: Variational Monte Carlo energy minimization of wave 2. Here it is clear that the minimum is approximately $(\alpha, \beta) = (1, 0.28)$.
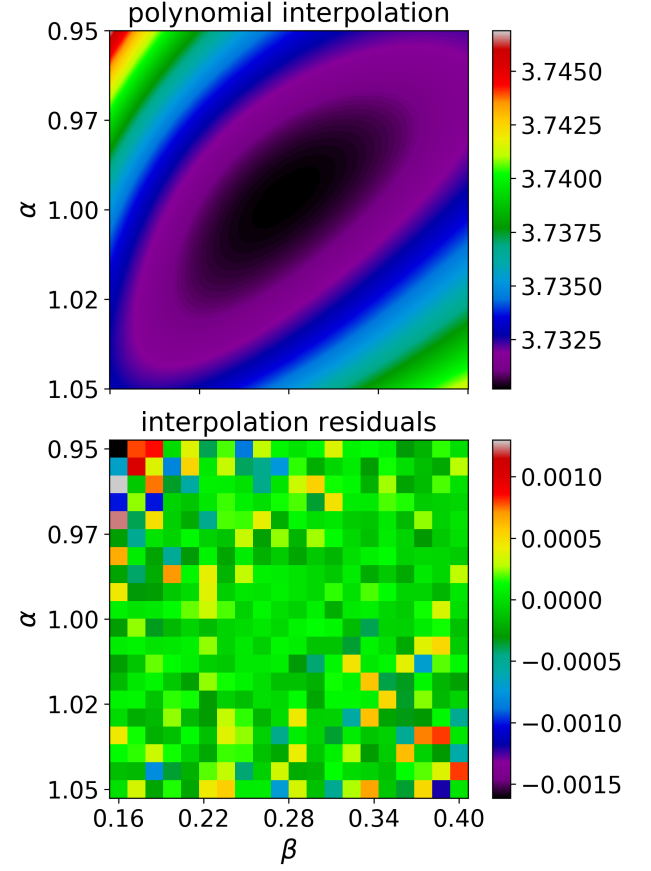
5th Degree Polyfit of
Wave 2 Energy Minimum

polynomial interpolation

interpolation residuals

Figure 8: 5th degree polynomial interpolation of the trial energy data in figure 7. The energy minimum is approximately located in $(\alpha, \beta) = (0.999, 0.271)$.

D.   Experiment 4

The electron-electron energy balance is shown in figure 9. The non-interacting data sets are very stable for all ω, albeit it slightly larger than 1, which is the true non-interacting ground state ratio. The interacting data sets both increase with decreasing ω, especially the interacting wave 1 dat set grows faster than the wave 2 data set. It should be noted that the true ground state is covered by wave 1 using $\alpha = 1$. However, because these data sets used the variational parameters optimized for the interacting Hamiltonian, neither will score particularly well with regards to the non-interacting energy balance.

The interacting energy balance is much greater than the non-interacting energy balance, this is clearly due to the Coulombic potential, which comes as an addition to the harmonic oscillator potential. Moreover, while the non-interacting energy balance decreases for the smaller ω, the interacting energy balance increases. This also most likely due to the addition of the Coulombic potential. Without having analysed the energy balance for higher values of the ω, the graphs seem to indicate that the interacting graphs converge on the non-interacting graphs for sufficiently large ω.
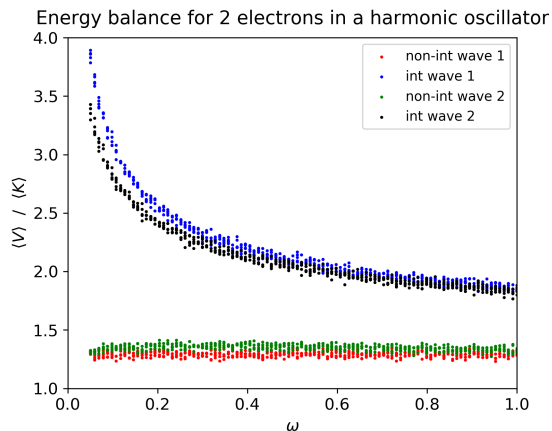

Figure 9: The balance between the potential and kinetic energy of two electrons in a harmonic oscillator. The four data sets refer to wave 1 and 2, with or without a Coulomb interaction.

V.   DISCUSSION

The energy minimum of wave 2 was about 3.730 Hartree, which is an approximate 1.1% improvement on Wave 1's energy minimum of about 3.773 Hartree. Hence, the Jastrow factor clearly introduced a correction to wave 1 that improved on the ground state approximation. The correction is very small, but does constitute a meaningful difference because it the Jastrow factor takes into account the electron-electron separation, which the non-interacting ground state (obviously) does not.

Hence, the ground state of the interacting system must therefore take also into account the electron-electron separation in some way.

Additionally, while the optimal wave 1 α was approximately 0.88, the optimal wave 2 α was approximately 1. This implies that the optimal Jastrow factor is merely a perturbation of the non-interacting ground state wave function, which is particularly interesting from the perspective of perturbation theory. That is, wave 2 can actually be thought of as a perturbed solution for some currently unknown electron-electron interaction. Hence, finding the correct variational ground state can be thought of as adjusting the unknown interaction to mimic the Coulomb interaction. Further study should consider improving on the Jastrow factor of wave 2 with either additional factors or adjustments to the Jastrow factor of wave 2 (e.g. replace the 1 in the Jastrow factor exponential with an additional variational parameter).

The Variational Monte Carlo algorithm has been very successful, that is, provided a reasonable number of Monte Carlo cycles. The Monte Carlo cost analysis of experiment 1 quantified such a "reasonable number" very effectively by looking at the evolution of the distribution of integral estimates. The reader is encouraged to perform a similar cost analysis, with special emphasis on the balance between increasing the number of cycles and running several simulations, in order to reduce the number of inefficient cycles spent on uneccessary accuracy.

With regards to the accuracy of using 10^5 Monte Carlo cycles in this project, the results are fairly reasonable, albeit some fluctuation in the energy variance integrals (most notably in figure 2). In fact, using 10^5 cycles did strike a balance between accuracy and efficiency, which was the intent of the cost analysis.

On a different note, how does the cost analysis relate to the accuracy of the ground state approximation? Well, the Monte Carlo accuracy deals only with the accuracy of the integral estimate, not the accuracy of the trial wave. Hence, to further improve the ground state approximation, it is probably more fruitful to explore other trail wave functions than to improve the optimization of the trial waves considered in this project.

Using a polynomial interpolation of the trial energy data proved itself very effective for both wave 1 and wave 2. However, the energy minimum was extracted using the local minimum of the polynomials, which technically is not an energy measurement per sé. A better approach would be to use the polynomials in order to extract the optimal variational parameters, but instead find the energy minimum by rerunning the simulations with the optimized parameters using a large number of Monte Carlo cycles. On the other hand, depending on the accuracy of the initial simulations, the difference between is not necessarily measurable.

The peaking behaviour seen in figures 4, 5 and 9 is largely due to the Coulombic electron repulsion: as the harmonic oscillator looses its grip on the electrons, the Coulomb interaction rips the electrons apart (see figures 4 and 5). This strong repulsion is widely dominant to the electron's kinetic energies, hence the large shift in the potential vs. kinetic energy balance. Moreover, the electrons necessarily carry more kinetic energy as the harmonic oscillator strength increases (as the electrons' oscillatory frequency increases).

VI.   CONCLUSION

Generally speaking, the project has been very successful with respect to its main goal: to approximate the ground state of a Helium-like quantum dot using the variational principle with trial wave functions based on the ground state wave function of the non-interacting system. The variational Monte Carlo algorithm was very successful in its estimate of trial energies, provided a reasonable choice of the number of Monte Carlo cycles per simulation. The two trial wave functions not only featured clear energy minimums in their respective parameter spaces, but wave 2 with an energy minimum of 3.730 Hartree was also found to be a 1.1% improvement on wave 1's energy minimum of 3.773 Hartree. Despite the energy difference being small, it clearly shows that the ground state must take the separation between the electrons into account.

[1] David J. Griffiths. *Introduction to Quantum Mechanics.* Prentice Hall, Inc., 2$^{\text{nd}}$ edition.
[2] Wikipedia: Atomic units. Online resource. https://en.wikipedia.org/wiki/Atomic_units.
[3] Morten Hjorth-Jensen. Fys 3150 lecture notes: Introduction to monte carlo methods, . http://compphysics.github.io/ComputationalPhysics/doc/pub/mcint/html/mcint.html.
[4] Morten Hjorth-Jensen. Fys 3150 lecture notes: Variational monte carlo methods, . http://compphysics.github.io/ComputationalPhysics/doc/pub/vmc/pdf/vmc-print.pdf.