Quantum Monte Carlo Calculations of Density Matrices of the Helium Atom: What Can We Learn from the Exact Solution?

Author: Carlos Rodriguez Perez

Supervisors: Vitaly Gorelov & Francesco Sottile

Département
de Physique
École normale
supérieure

Contents

1 Introduction

Electronic correlations pose one of the most significant challenges in the current state of the art of Condensed Matter Physics. They are the main driver behind many incredibly interesting phenomena such as, spin-charge separation, the behaviour of transition metal oxides (for example, high-temperature cuprate superconductors), Mott insulators or photovoltaic cells, to name a few. Their study can yield insight into these unusual properties, and hopefully turn them to technological use.

There is a class of materials recently termed metavalent in which electronic correlations are thought to be of importance. They exhibit large dielectric functions and possibly a new bonding scheme in which electrons appear to be delocalized inside a unit cell, yet with large correlations between electrons in different cells [1]. Their large dielectric functions make them very good light absorbers. We are interested in a possible application of these materials for photovoltaic cells, but the metavalent materials known thus far exhibit very small gaps, making them unsuiable for photovoltaics. One possible way to understand this new class of materials is to calculate their reduced density matrices and compare them with that of metals and semiconductors. The reduced density matrix is an object that can be obtained with Quantum Monte Carlo methods. These are a variety of stochastic methods that allow for extremely accurate calculations. To do this, a code has been developed by the author of this report in Python 3 to perform Variational Monte Carlo (VMC), Diffusion Monte Carlo (DMC), and the statistical analysis required to process the data for a single Helium atom (some of the code can be found at https://github.com/c-rodriguez-per). We needed to make our own code so that we could really understand the workings of the simulation and to tweak it to suit our needs. This has all been benchmarked on the helium atom, one of the simplest systems that exhibits electronic correlations. The helium atom has the advantage that the exact solution is known [2]. So we can benchmark the accuracy of QMC energies, wavefunctions, and reduced density matrices. This will be of high importance for a following study of metavalent materials.

The main questions that we want to address in this report are **how accurate are reduced density matrices calculated with QMC?** and if they are accurate **what is the cost of computing such an object with respect to the total energy calculations**? If possible we would like to know to which extend of accuracy this can be done with Variational Monte Carlo, as it is a much less computationally expensive technique.

This report is structured as follows. In Section 2 we introduce the theory behind the VMC and DMC we will be using, namely Variational Monte Carlo and Diffusion Monte Carlo and some general aspects of the technique. Then, in Section 3, we will discuss our results for the ground state energy of the Helium atom with these two methods in. We also introduce the results and the first excited state energy with VMC. Afterwards, in Section 4 we proceed to compare our results for the wavefunction, the reduced density matrix, and the electronic density. To the best of our knowledge, this is the first comparison of real space QMC reduced density matrices and wave functions with exact solutions, allowing for a benchmark of a method that is commonly used as a benchmark itself. We conclude the section with a summary of the calculations done. Then, in Section 5 we present our last remarks and some possible paths of further study.

2 Quantum Monte Carlo methods

2.1 Monte Carlo integration

Condensed matter and statistical physics are two of the main fields where many-body physics are explored, both domains require the calculation of high dimensional integrals. An example of such integrals is the calculation of the partition function (in a continuous system) of dimension $2dN$ or the expectation value of the energy dN (with N the number of particles and d the dimensionality of the problem).

Monte Carlo techniques are incredibly useful for the evaluation of integrals. This is especially true of highly dimensional integrals. While most methods perform numerical integration on a grid, Monte Carlo integration is done using random numbers drawn from some distribution that is integral-specific, this is the source of the computational advantage of Monte Carlo techniques. The way in which we perform such integration is by first splitting our function $g(x)$ into two functions $f(x)$ and $n(x)$, where the only constraint is that $n(x)$ must be a probability distribution (non-negative and normalized).

$$I = \int dx\ g(x) = \int dx\ f(x)n(x) = \lim_{N\to\infty} \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} f(x_i) \quad \text{such that } x_i \sim n(x) \tag{2.1.1}$$

Where $x_i \sim n(x)$ means "x_i distributed according to $n(x)$". With N_{MC} the number of points used in the sampling of the integral. This way of estimating the integral yields an error that scales like $1/\sqrt{N_{MC}}$. This means that in the infinite sample size limit we get the exact answer. However, this scaling is also one of the drawbacks of the technique, it means that if we want to reduce our error bars by a factor of 10, we must run for 100 times longer.

The main drawback of this technique is that we must ensure that our x_i are indeed distributed according to $n(x)$. It is not trivial to draw random numbers according to some arbitrary distribution $n(x)$. To solve this we use the Metropolis-Hastings algorithm explained in Appendix A.

2.2 Variational Monte Carlo

Variational Monte Carlo (VMC), introduced by W. L. McMillan [3], is a Quantum Monte Carlo method based on the well known Variational principle.

$$\boxed{E_v = \frac{\langle\psi|\, H\, |\psi\rangle}{\langle\psi|\psi\rangle} \geq E_0} \tag{2.2.1}$$

Which remains true for any well behaved wavefunction ψ. This means that we can choose a reasonable wavefunction that depends on some parameters. And then try to optimize the parameters to obtain as good a wavefunction as we can, this is,

$$E_v(\alpha) = \frac{\langle\psi(\alpha)|\, H\, |\psi(\alpha)\rangle}{\langle\psi(\alpha)|\psi(\alpha)\rangle} \geq E_0 \tag{2.2.2}$$

Where α is a set of parameters. If we rewrite this expectation value as an integral and perform some manipulations on it, we may use Equation 2.1.1. It is useful to remember that R here means the position of all the electrons of the system of interest.

$$E_v = \frac{\int dR dR'\, \langle\psi|R\rangle\, \langle R|\, H\, |R'\rangle\, \langle R'|\psi\rangle}{\int dR\, \langle\psi|R\rangle\, \langle R|\psi\rangle} = \frac{\int dR\ \psi^*(R) H \psi(R)}{\int dR\ \psi^*(R)\psi(R)} \tag{2.2.3}$$

In the second step we have assumed that the Hamiltonian is diagonal in position representation (which is the case for the helium atom). Let us now multiply and divide by the wavefunction before the Hamiltonian.

$$E_v = \frac{\int dR\ \psi^*(R)\psi(R)\psi^{-1}(R)H\psi(R)}{\int dR\ \psi^*(R)\psi(R)} \tag{2.2.4}$$

$$= \int dR\ \psi^{-1}(R)H\psi(R)\frac{\psi^*(R)\psi(R)}{\int dR\ \psi^*(R)\psi(R)} \tag{2.2.5}$$

$$= \int dR\ E_L(R)\ \rho(R) \tag{2.2.6}$$

Where we have grouped terms as our local energy and our probability distribution.

$$E_L(R) = \psi^{-1}(R)H\psi(R) \quad , \quad \rho(R) = \frac{\psi^*(R)\psi(R)}{\int dR\ \psi^*(R)\psi(R)} \tag{2.2.7}$$

And we can use the Monte Carlo method for integration together with Algorithm 2 to compute this integral.

$$\boxed{E_v \approx \frac{1}{N_{MC}} \sum_{R_i \sim \rho}^{N_{MC}} E_L(R_i)} \tag{2.2.8}$$

With a statistical error bar of, (let us point out that this error bar can indeed be taken to zero in the infinite N_{MC} limit)

$$\sigma = \sqrt{\frac{Var(E_L)}{N_{MC}}} \tag{2.2.9}$$

We could bypass the manipulations presented above and just calculate the two integrals presented in Equation 2.2.1, but by performing said manipulations we can make it such that we need only do one integral. If the wavefunction normalization were needed, both integrals would have to be performed.

2.3 Diffusion Monte Carlo

DMC is a projector approach to Monte Carlo simulations [4]. In this technique we use a projector operator that represents complex-time evolution.

$$\hat{G}_D = e^{-t\hat{H}} \tag{2.3.1}$$

It can be show that this projector applied to any state exponentially suppresses all states, with the lowest energy state being exponentially less suppressed.

$$|\psi_T\rangle = \sum_{i=0} c_i |\psi_i\rangle \to e^{-t\hat{H}} \sum_{i=0} c_i |\psi_i\rangle = \sum_{i=0} c_i e^{-tE_i} |\psi_i\rangle = c_0 e^{-tE_0} \left(|\psi_0\rangle + \sum_{i=1} c_i' e^{-t(E_i - E_0)} |\psi_i\rangle \right) \tag{2.3.2}$$

With $|\psi_i\rangle$ the i-th eigenstate of the Hamiltonian. Knowing this we can write that for sufficiently large t we have $|\psi_0\rangle = \hat{G}_D |\psi_T\rangle$.

Now we can write our exact ground state energy as, (the denominator would give $c_0 \exp\{-tE_0\}$, allowing us to do this as long as our trial wavefunction has non zero overlap with the ground state)

$$E_0 = \frac{\langle\psi_0| H |\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} = \frac{\langle\psi_T| H |\psi_0\rangle}{\langle\psi_T|\psi_0\rangle} = \frac{\langle\psi_T| H e^{-t\hat{H}} |\psi_T\rangle}{\langle\psi_T| e^{-t\hat{H}} |\psi_T\rangle} \tag{2.3.3}$$

This is the so called mixed estimator. when calculating energies this is exact. (If one wishes to compute an observable that does not commute with H one needs to resort to extrapolated estimators, see [5] or Equation 4.2.5.) Now, inserting the completeness relations we get,

$$E_0 = \frac{\int dR_0 dR_1 \langle\psi_T| H |R_0\rangle \langle R_0| e^{-t\hat{H}} |R_1\rangle \langle R_1|\psi_T\rangle}{\int dR_0 dR_1 \langle\psi_T|R_0\rangle \langle R_0| e^{-t\hat{H}} |R_1\rangle \langle R_1|\psi_T\rangle} \tag{2.3.4}$$

Now, we shall substitute using $\langle R_1|\psi_T\rangle = \psi_T(R_1)$ and also inserting $\psi_T^*(R_0)/\psi_T^*(R_0)$ in the above expression

$$E_0 = \int dR_0 dR_1 \frac{H\psi_T(R_0)}{\psi_T(R_0)} \frac{\psi_T^*(R_0) \langle R_0| e^{-t\hat{H}} |R_1\rangle \psi_T(R_1)}{\int dR_0 dR_1 \psi_T^*(R_0) \langle R_0| e^{-t\hat{H}} |R_1\rangle \psi_T(R_1)} \tag{2.3.5}$$

And we now group terms as follows,

$$E_L(R_0) = \frac{H\psi_T(R_0)}{\psi_T(R_0)} \quad , \quad \rho(R_0) = \int dR_1 \psi_T^*(R_0) \langle R_0| e^{-t\hat{H}} |R_1\rangle \psi_T(R_1) \tag{2.3.6}$$

We obtain exactly what we wanted,

$$E_0 = \int dR_0 E_L(R_0) \frac{\rho(R_0)}{\int dR_0 \rho(R_0)} \tag{2.3.7}$$

Using the normal expression for Monte Carlo integration this can be solved, provided we can create a series of independent configurations distributed as $\rho(R_0)$ which is not trivial at all.
Let us examine this distribution,

$$\rho(R_0) = \psi_T^*(R_0) \int dR_1 \langle R_0| e^{-t\hat{H}} |R_1\rangle \psi_T(R_1) \tag{2.3.8}$$

This projector cannot be solved directly, but we can use two tricks. Since the Hamiltonian commutes with itself we can break up this exponential into many small exponentials.

$$e^{-t\hat{H}} = \left(e^{-\tau\hat{H}}\right)^n, \qquad t = \tau n \tag{2.3.9}$$

Now, we use Trotter's formula, which states that in the $\tau \to 0$ limit this formula holds. (We split the potential in two so as to obtain a smaller time step error; this is called split two propagator.)

$$e^{-\tau(\hat{T}+\hat{V})} = e^{-\tau\hat{V}/2} e^{-\tau\hat{T}} e^{-\tau\hat{V}/2} + \mathcal{O}(\tau^2) \tag{2.3.10}$$

Now we can solve these small pieces. We introduce n completeness relations into ρ and write

$$\rho(R_0) = \psi_T^*(R_0) \int \prod_{i=1}^{n} dR_i \langle R_{i-1}| e^{-\tau\hat{H}} |R_i\rangle \psi_T(R_n) \tag{2.3.11}$$

To solve each of these pieces, we need only realize that the kinetic term is only a diffusion equation, and the potential one is just an evaluation. [4]

$$\langle R_{i-1}| e^{-\tau\hat{H}} |R_i\rangle = (2\pi\tau)^{-3N/2} \exp\left\{-\frac{(R_{i-1}-R_i)^2}{2\tau}\right\} \exp\{-\tau\left(V(R_{i-1})+V(R_i)\right)/2\} \tag{2.3.12}$$

The interpretation of this can be seen in Equation 2.3.12. We have a first part which just diffuses our walker, and a second which accumulates its weight.
For a fixed tau, the number of steps tell us how many of these diffusion moves and weights we get

for each walker. At infinite steps is when we fully project out the ground state (when our walkers are distributed according to the ground state). It must be noted that 'infinite steps' means a few thousand for our calculations. This can be seen in Figure 2 as it equilibrates after a few thousand moves.
For some interpretation, the weights tell us in some way, how promising one path seems. This will allow us to give more importance to this walker when doing our simulation, in the end we can perform a weighted sum.

$$E_0 \approx \frac{1}{\sum_i^{N_{MC}} w_i} \sum_{R_i \sim \rho}^{N_{MC}} E_L(R_i) w_i \tag{2.3.13}$$

Where $w_i = \exp\{-\tau \left(V(R_{i-1}) + V(R_i)\right)/2\}$ is the weight accumulated after many steps.

When it comes to actual code implementation, our starting wavefunction is of the sort $\psi_T(R_1) = \sum_i^{N_W} \delta(R_1 - R_i)$ (remember that we do not need normalized wavefunctions). This basically means that we are performing the projection method at the same time for all our walkers.
This is, each application of the projector with the small τ is one Monte Carlo step, and at each of these steps we do our computations simultaneously for all our independent walkers. This improves the efficiency of the calculation significantly.

2.4 Branching algorithm

This weighting presented in Equation 2.3.13 is computationally inefficient, as we are doing all the diffusion moves and computations for some walkers that will have negligible weights (some weights decrease exponentially with the number of steps). What we do is use these weights at each step to perform a 'birth-death' algorithm, this is called branching. After each step, we compute the weights and add a random number between 0 and 1. If the result is lower than 1 we kill the walker, if it is higher than one it has a probability to reproduce. This leads to the fact that our number of walkers is not constant (we call population our number of walkers), this results in problems such as population explosion or implosion. The way to address this problem will be discussed in the next section.

Algorithm 1: Branching

A basic pseudocode for branching is

i) The probability of surviving for each walker is $P = \exp\{-\tau \left(V(R) + V(R')\right)/2\}$. Let us remark that this is just the weight shown in Equation 2.3.12.

ii) If $P < 1$, then P is the probability of surviving to the next step. If $P \geq 1$ the walker survives and has a probability $P - 1$ to create a copy for the next step. Both possibilities can be computed at once by setting the number of copies for the next step as $M = P + \eta$ with η a random uniform number from 0 to 1.

To test the computational cost of branching and weighting, the same calculation has been performed for both methods and compared (we have tested on the ground state of atomic helium). The results can be found in Table 1, in which we can see that the branching method yields more accurate results for the same computer time.

	Weight	Branching
Acceptance	0.9938	0.9937
Computer time (s)	103.71	99.82
E (Ha)	-2.9052	-2.90391
Sigma (Ha)	0.0015	0.00044

Table 1: Results obtained for a 4000 walkers and time step $\tau = 0.01$ in atomic helium. We can see how for a same amount of computer time we get significantly better results with the branching method.

Due to its superior results the branching method will be the one used for all DMC calculations performed in this report.

2.5 Importance sampling

The branching method introduced in the previous section has some computational problems, this is, the fact that the potential energy has positive and negative divergences, which can lead to uncontrolled populations. One straightforward solution to this is to limit the divergence to a very large number, this solution will not be discussed in this report. Another method to tackle this problem is the so called importance sampling technique. [5]

In this technique a guiding wavefunction ψ_G is used. This has the effect of pushing the walkers towards regions where ψ_G is larger (In practice what we do is $\psi_G = |\psi_T|$). The effect of this is that our projector defined in Equation 2.3.12 is now changed to,

$$\langle R_{i-1}| \, e^{-\tau \hat{H}} \, |R_i\rangle = (2\pi\tau)^{-3N/2} \exp\left\{ -\frac{(R_{i-1} - R_i - \tau \boldsymbol{\nabla} \log \psi_G)^2}{2\tau} \right\} \exp\{-\tau \left(E_L(R_{i-1}) + E_L(R_i)\right)/2\} \tag{2.5.1}$$

This new projector presented in Equation 2.5.1, does solve our problem of population control. The choice $\psi_G = |\psi_T|$ means that we will converge even quicker to the desired distribution, thanks to the new term in the diffusion part of the projector.
It must also be said that in practice the second term is regularized with a trial energy E_T, which helps keep the population under control.

$$\exp\{-\tau \left(E_L(R_{i-1}) + E_L(R_i)\right)/2\} \to \exp\{-\tau(E_L(R_{i-1}) + E_L(R_i) - 2E_T)/2\} \tag{2.5.2}$$

This trial energy is updated so as to ensure that population is kept on average constant.

3 Application to He atom

3.1 Helium atom Hamiltonian

The helium atom is the simplest realistic system in which we can study electronic correlations. We are interested in this system because the ground state DMC is an exact method, and we also have access to an exact solution [2]. Within the Born-Oppenheimer approximation scheme, the helium atom Hamiltonian is, (during all this report atomic units will be used)

$$H = -\frac{1}{2} \sum_{i=1}^{N=2} \nabla_i^2 - \sum_{i=1}^{N=2} \frac{2}{r_i} + \frac{1}{r_{12}} \tag{3.1.1}$$

Where r_{12} is the electron electron distance and r_i is the position of the i-th electron.

3.2 Trial wavefunction for the Helium atom

When performing Quantum Monte Carlo calculations, one normally starts with a wavefunction obtained from a single particle Slater determinant using less expensive techniques such as DFT (Density Functional Theory) and adds another factor that accounts for the electronic correlations. In this report, we have decided to use a wavefunction of similar structure,

$$\psi = D_\uparrow(r_\uparrow) D_\downarrow(r_\downarrow) e^{J(R)} \tag{3.2.1}$$

In which $D_\uparrow$ is a Slater determinant of all the electrons with up spin and coordinates $r_\uparrow$, and $J(r)$ is a so called Jastrow factor in which we will incorporate electronic correlations. It is of form,

$$J(r) = \sum_{i=1}^{N} \chi(r_i) - \frac{1}{2} \sum_{i=1, j\neq i}^{N} u(r_i, r_j) \tag{3.2.2}$$

Where $\chi(r_i)$ is a electron-nucleus one body term, $u(r_i, r_j)$ is a electron electron two body term.
For the ground state of helium our Slater determinants reduce to 1s orbitals.

$$D_\uparrow(r_\uparrow) = exp(-\alpha r_\uparrow) \quad D_\downarrow(r_\downarrow) = exp(-\alpha r_\downarrow) \tag{3.2.3}$$

And we have decided to work with one- and two-body terms so that we can obtain as accurate a wavefunction as possible. These Jastrow factors have been selected because they are the ones used for solid molecular hydrogen due to their simplicity and good performance [6], so we expected them to have an acceptable performance in our case of atomic helium.

$$J(R) = -f_{ep} \sum_{i}^{N=2} e^{-r_i^2/\omega_{ep}^2} - f_{ee} e^{-r_{12}^2/\omega_{ee}^2} - f_{bf} e^{-r_{12}^2/\omega_{bf}^2} r_{12} \tag{3.2.4}$$

Using Variational Monte Carlo we can obtain the optimal parameters that minimize the energy. The optimization has been done by hand.

$$\alpha = 2, \quad f_{ep} = 0.777, \quad w_{ep} = 2.56, \quad f_{ee} = 0.41, \quad w_{ee} = 1.35, \quad f_{bf} = -0.5, \quad w_{bf} = 0.1 \tag{3.2.5}$$

It must be said that α and f_{bf} have not been optimized, as we need them to be these specific values so as to satisfy the cusp conditions.

As we can see in Equation 3.1.1, the potential energy exhibits a divergence in the Coulomb terms. The way this is treated is with the so called cusp conditions. Whenever there is a divergence in the potential energy, there is a cancelling divergence in the kinetic energy which comes from the wavefunction itself. If the cusp conditions are not satisfied, we will encounter divergences in our calculation. [7, 4]

$$\frac{1}{\psi}\frac{d\psi}{dr_{iI}}\bigg|_{r_{iI}=0} = -Z \quad , \quad \frac{dJ}{dr_{ij}}\bigg|_{r_{ij}=0, \sigma_i \neq \sigma_j} = -\frac{1}{2} \quad , \quad \frac{dJ}{dr_{ij}}\bigg|_{r_{ij}=0, \sigma_i = \sigma_j} = -\frac{1}{4} \tag{3.2.6}$$

Where r_{iI} is the electron ion distance, r_{ij} is the electron electron distance and σ_i is the spin of electron i. The first condition in Equation 3.2.6 will set our parameter α; the second is what sets f_{bf}. The last condition will be useful when dealing with the first excited state of helium, which is a triplet $2\,^3S$ (one electron in the first shell and another in the second, with parallel spins and no orbital angular momentum).

3.3 Variational Monte Carlo results

We must first explain how we obtained the parameters in Equation 3.2.5. We have optimized them in groups to reduce the load on the computer. To do this we selected them in pairs and run a VMC calculation for each possible pair value in a reasonable range, and studied the behaviour of the energy in this range. This was repeated several times, adjusting the range of the parameters so as to obtain as high an accuracy as possible. Then this was performed for the other parameters.

Using the parameters obtained, we can perform one Monte Carlo run. In VMC we generally want our acceptance probability to stay between 0.3 and 0.7. This is done to ensure that we properly sample all our space. The first part of our run is called the equilibration phase; this is the time it takes our initial configuration (in this case taken as Gaussian) to start sampling our target probability distribution. This part is discarded when calculating averages (see Figure 1). After this, the energies are accumulated and then averaged over to obtain the final energy with its error bar.

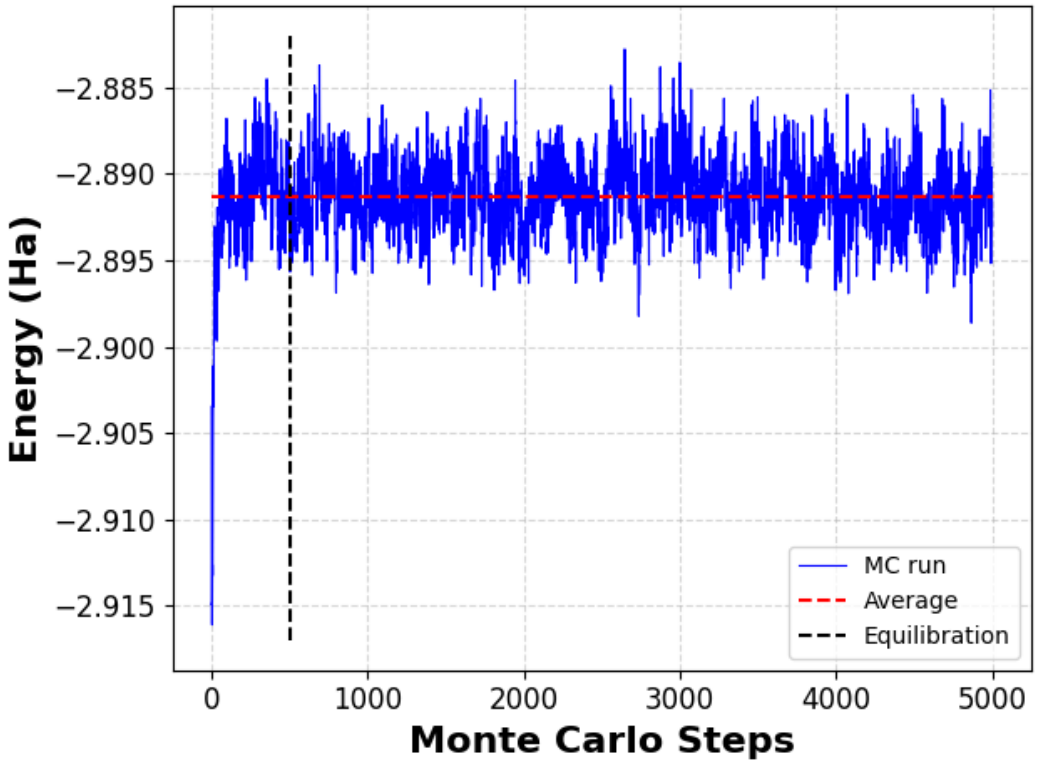

Figure 1: VMC run performed for $\tau = 0.35\ Ha^{-1}$, using 10^5 independent configurations and an acceptance ratio of 56%. The obtained energy is $E = -2.8913(1)\ Ha$.

From the data shown in Figure 1 we can conclude that our wavefunction yields an energy close to the one yielded by the true ground state of the system which is $E_0 = -2.9036(1)\ Ha$ (see Figure 5 or Section 3.5). Further improvements can be done adding more terms to the Jastrow factor or with the so-called backflow correlations introduced by Feynman and Cohen [8], which will not be discussed in this report. If a more accurate energy is sought, the Diffusion Monte Carlo technique is the way to go.

3.4 Diffusion Monte Carlo results

Diffusion Monte Carlo is an exact method for this problem, as the wavefunction does not have any nodes. We have performed several large computations, one of which can be seen in Figure 2. When

performing DMC runs, we want an acceptance ratio close to unity, due to the existence of time-step errors. These errors will be explored in Section 3.6, where we also shall perform an extrapolation to zero time-step.

In Figure 2 we can see one of the drawbacks of this method: the low time steps in our simulation mean that we take many steps to project the ground state. This means that about one quarter of our computer time had to be discarded. There are some solutions to this problem, like using an old distribution of walkers from VMC to start the simulation or allowing for large τ for a few steps at the beginning. None of these methods have been implemented.

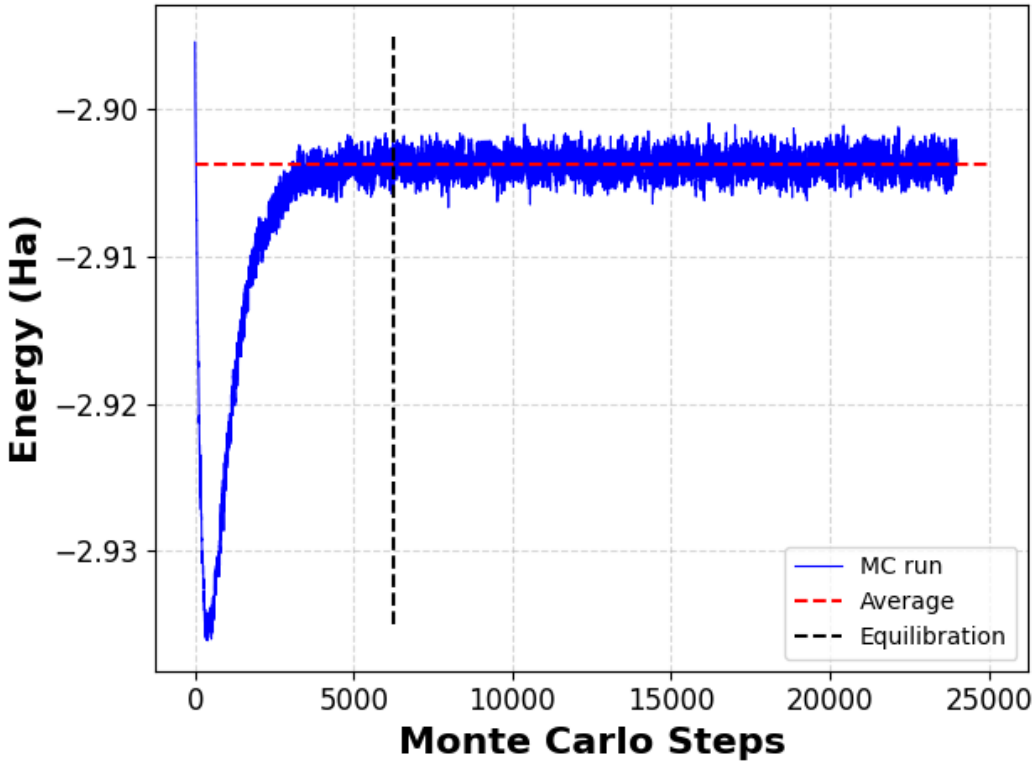

Figure 2: QMC trajectory for a time step of $\tau = 0.0015\ Ha^{-1}$, we can clearly see the equilibration phase and the cutoff for the warm-up. We obtain an energy of $E = -2.90374(3)\ Ha$. Calculation performed using $6.6 \cdot 10^5$ independent walkers and with an acceptance ratio over 99%.

One practical problem that arises in these computations is that of self correlation. Since our time step is very small, we cannot use all of the points in Figure 2 for the calculation of the average energy; we must use only uncorrelated points. To do this, we have calculated the self correlation and the correlation length, obtaining that the correlation length is $t_{corr} \sim 0.036 Ha^{-1}$, which in the case of Figure 2 is about 24 Monte Carlo steps. In practice, this means that in Figure 2 after the equilibration we have 790 independent points for the calculation of the energy of the run. This is the reason for the large amount of steps in comparison to Figure 1.

It is interesting now to take a small detour to explain the exact solution so as to compare the energies of both.

3.5 Hylleraas's solution

The ground state of helium is a problem that was solved by Hylleraas in 1929 [2]. To the best of our knowledge, there are no systematic comparisons of quantum Monte Carlo reduced density matrices with the exact ones given by the system.
The exact solution calculated by Hylleraas is not in a closed form. From the helium Hamiltonian, one can see that our problem can be reduced to three coordinates $s = r_1 + r_2$, $t = r_1 - r_2$ and $u = |\vec{r}_1 - \vec{r}_2||$. With these coordinates, our wavefunction can be exactly written as in Equation 3.5.1.

$$\psi(s,t,u) = e^{-ks} \sum_{j,l,m,n} c_{j,l,m,n} s^l (t/s)^m (u/s)^n ln^j(s) \tag{3.5.1}$$

Where k is the Kato cusp condition [7], and the logarithmic divergence was introduced by [9] to further improve the accuracy. It has been demonstrated that the solution shown in Equation 3.5.1 (including negative powers of s and t) is indeed a solution for the Hamiltonian in Equation 3.1.1 [10].
Using this solution, one can variationally obtain the parameters. This solution is exact in the sense that it can be systematically improved upon the addition of more terms, although we do not have a closed expression for it.

Both of the energies presented in [11] with the wavefunction in Equation 3.5.1 are significantly better than ours. One must take into account that they have used the already established code CASINO [12] with 42 parameters for the trial wavefunction. The usage of a superior code and a more tunable wavefunction is clearly expected to perform better.

3.6 Zero time step extrapolation

In Equation 2.3.10 we saw that when our time step goes to zero our results should be exact. We can perform several computations for different time steps and then perform an extrapolation.

$$U(\tau) = e^{-\tau(\hat{T}+\hat{V})} = e^{-\tau(\hat{V}-E_{ref})/2} e^{-\tau\hat{T}} e^{-\tau(\hat{V}-E_{ref})/2} \xrightarrow[sampling]{importance} e^{-\tau(\hat{E}_L-E_{ref})/2} e^{-\tau\hat{T}} e^{-\tau(\hat{E}_L-E_{ref})/2} \tag{3.6.1}$$

Which, when implemented in practice just means that we will propagate with the kinetic term. And our branching will be done with probability

$$P(R,R') = \exp\left(-\frac{\tau}{2}\left(E_L(R) + E_L(R') - 2E_{ref}\right)\right) \tag{3.6.2}$$

This is what is known as a Split-2 Propagator [13]. The error introduced by this propagator is quadratic in the time step and thus, the best way to perform the extrapolation is with a parabola. This what is done in Figure 3 where we can see that zero time step extrapolation yields very good results. Yet the result of the extrapolation is worse than that of the lowest time step calculation, this is to be expected, as these points are already highly accurate and the extrapolation is affected by points with less accuracy. It must be noted that our extrapolation result does not agree with the exact energy within the error bars, but the point with the lowest time step does (see Figure 3).

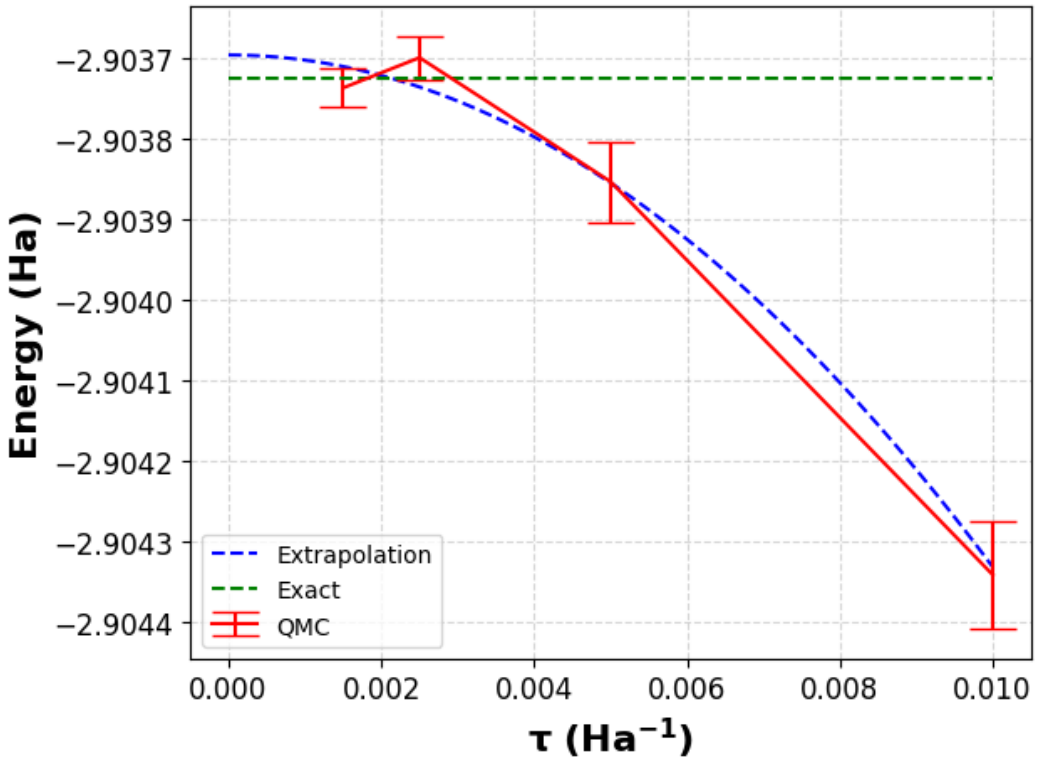

Figure 3: Zero time step extrapolation using Diffusion Monte Carlo. The fit is done with a second order polynomial and with the error bars as weights [13]. The obtained energy is $E_0 = -2.90370(1)\ Ha$. It can be seen that as we decrease our time step, the error introduced by the Trotterization is smaller than our error bars and thus can be considered negligible.

Zero time step extrapolation is an expensive yet powerful technique. As the time step is decreased, the number of steps must increase so as to keep the number of uncorrelated points similar, also the number of walkers must increase in the same way. This means that a reduction of τ by a factor of 10 increases the number of computations by 100.

3.7 Excited states with Quantum Monte Carlo

Quantum Monte Carlo also allows to calculate excited states. In our case, the first excited state of helium is a triplet, this offers some new challenges with respect to the ground state. Our excited state wavefunction $\psi^*(R)$ will be,

$$\psi^*(R) = \left[e^{-\alpha r_1}(2-\alpha r_2)e^{-\alpha r_2/2} - (2-\alpha r_1)e^{-\alpha r_1/2}e^{-\alpha r_2}\right] e^{J(R)} \tag{3.7.1}$$

And we keep the same Jastrow factor as before. Now, the wavefunction has a node, and thus it is negative at some points.

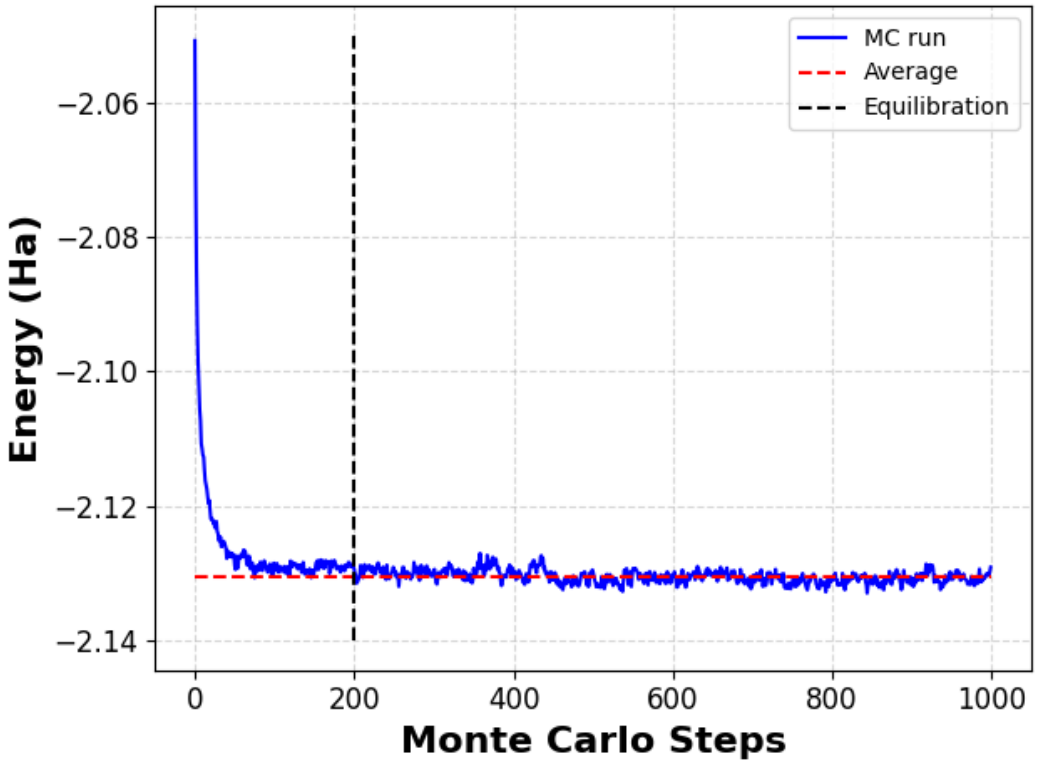

Figure 4: VMC calculation of the excited state energy. The energy obtained is $-2.13058(3)$ Ha. The run has been performed with 50k configurations and 1000 steps at $\tau = 0.35$

In Figure 4 we can see a run done for the excited state energy of Helium with Variational Monte Carlo. The energy obtained is not extremely accurate, this is mainly due to the fact that the Jastrow factor has not been re-optimized beyond enforcing the cusp conditions. Thus, we use the parameters presented in Equation 3.2.5 with the change $f_{bf} = -0.25$.

If one wants to run a DMC calculation, this is a problem as the wavefunction itself cannot be used as a probability distribution. It can be seen in Equation 2.3.11 that if one of these diffusion steps were to cross the node, we would acquire a minus sign that will be carried all the way to the computation of the energy.
The way to solve this is to use the so called fixed node approximation. In which we allow all walkers to freely move so long as they do not cross the nodes. There are two ways to implement this, (one can also allow walkers to cross the node and then make sure to take this minus sign into account, this is in theory exact, but dramatically increases the cost of the calculation).

i) Reject all moves that change the sign of the wavefunction.

ii) Kill a walker if its last move changed the sign of the wavefunction.

Out of both methods, we expect the rejection to be move efficient, and thus it is the method that we implemented for our calculations of excited states. It must also be said that importance sampling applies a force that is inversely proportional to the distance to the node. Thus the usage of importance sampling reduces the probability that a walker crosses a node dramatically. Nevertheless, this problem has to be treated by rejecting such moves. We have not performed a calculation of the excited state with DMC.

4 Benchmarking QMC

I would like to thank Valerio Olevano for providing us with the exact wavefunctions and solutions shown in this report.

4.1 Wavefunction comparison

It would be appropriate to first compare our trial wavefunction to the analytical one, this is done in Figure 5. This will allow us to see if correlations are being correctly treated, and will tell us how different both wavefunctions are. We can see that our trial wavefunction is mainly different at the electron-electron point, but quite similar at the nucleus. From this we can conclude that not only is our variational energy close to the ground state, but so is our wavefunction. This means that other observables could be calculated with this trial wavefunction. It also means that our trial wavefunction is not correlated enough.

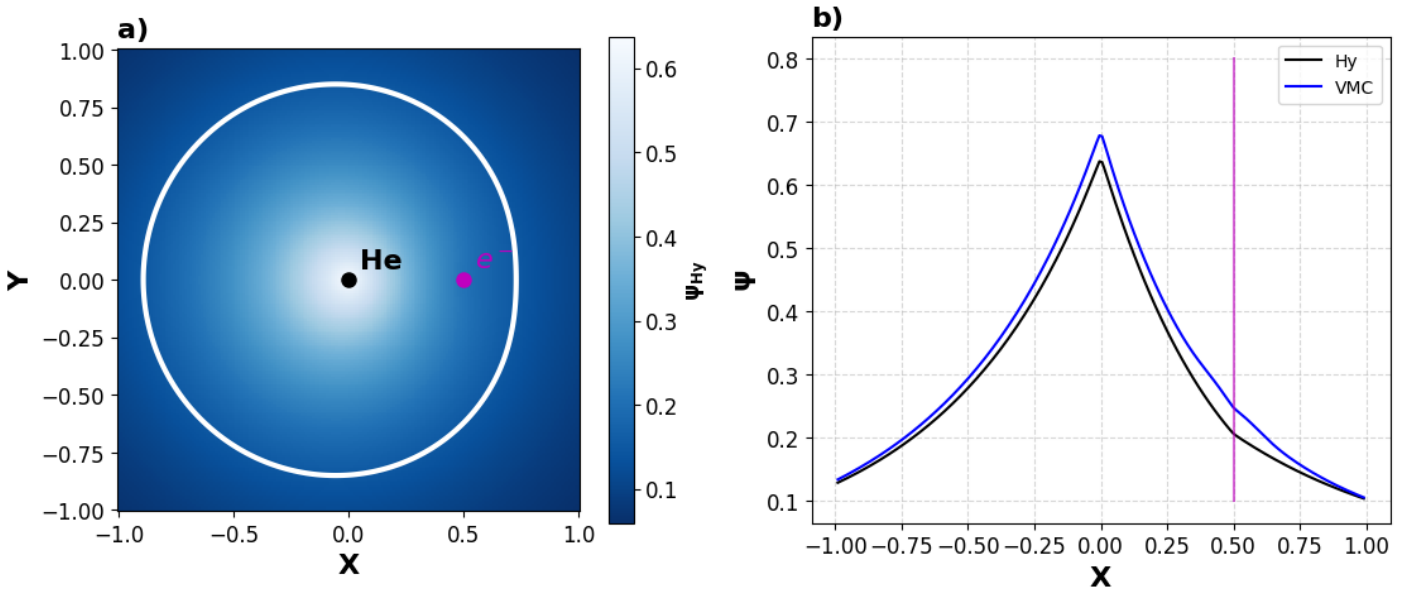

Figure 5: a) Hylleraas wavefunction with isocurve at 0.15 ($E = -2.9036(1)\ Ha$) b) Hylleraas and VMC wavefunction, both normalized ($E_{VMC} = -2.8913(1)\ Ha$). Both have represented the position of electron 1 (in pink) and of the nucleus. They are represented such that $r_1 = (0.5, 0, 0)$ and $z_2 = 0$. Both have been normalized by the computation of the six dimensional integral.

Both wavefunctions shown in Figure 5 are extremely similar, but some particular properties are indeed different. We can see that the electron ion cusp is well treated in both cases, as they are very similar and of almost equal slopes. The electron electron cusp is a different story, both wavefunctions satisfy Kato's cusp conditions, but the value at the e-e point is significantly different. This is a sign that the true ground state has higher correlations than those included in the trial wavefunction. The Hylleraas wavefunction exhibits a different slope to the right of the fixed electron, this resembles screening of the nuclear charge. Meanwhile the trial wavefunction does not show this change of slope. This is already telling us that our Jastrow factor is missing something, probably a three body term that changes behaviour when one electron is further from the nucleus than the other, this should be capable of reproducing a change of slope.

Another method of further improving this is to use the DMC scheme. This can be seen by comparing the energies of Figures 1 and 2, we can confidently say that DMC is very likely capturing the electronic screening behaviour present in Hylleraas' wavefunction, due to its highly accurate energy. The DMC wavefunction cannot be computed in general, but in two electron systems it can be calculated from the pair density (in Appendix B we present an approximation for the pair density, which we believe to be too crude to have a meaningful discussion).

4.2 One Body Reduced Density Matrix

The one body reduced density matrix (1RDM) is an incredibly useful object, as it allows to study the level of correlations and delocalization in a material. It is standard practice to use QMC to calculate this object and also its Fourier transform, the momentum distribution. We wanted to explore how accurate is QMC in the calculation of these objects since, to the best of our knowledge, there is no benchmark of 1RDM with real space QMC in the literature with respect to the exact result. Our case, atomic helium, is specially useful for this, as we have the exact solution.

The 1RDM is defined as,

$$\rho(r, r') = N \int dr_2 ... dr_N \ \psi(r, r_2, r_3 ... r_N) \psi^*(r', r_2, r_3 ... r_N) \tag{4.2.1}$$

The main problem that we faced is that this object is highly complex, requiring a large amount of computational resources. What we did was calculate elements fixing r'. This allows us to compute elements of this density matrix quite efficiently. One check that was done is that the diagonal of the 1RDM must equal the density, (with $n(r)$ the density at r)

$$\rho(r, r')\Big|_{r=r'} = n(r) \tag{4.2.2}$$

One way to compute this object with QMC is to use the ghost particle technique, in which we create a random position s to improve the computational efficiency of the calculation [3]. The price to pay for this is that our solution now depends on the trial wavefunction due to the mixed estimator error.

$$\rho(r, r') = \left\langle \frac{1}{f(s)\psi(r_1, r_2)} \left\{ \delta(r - r_1)\delta(r' - s)\psi(s, r_2) + \delta(r - r_2)\delta(r' - s)\psi(r_1, s) \right\} \right\rangle \tag{4.2.3}$$

Where $f(s)$ is the distribution from which we draw s, in this case taken uniform $f(s) = 1/\Omega$, and Ω is the volume of the simulation space. The calculation of this object is then straightforward: one has to perform a Monte Carlo calculation and accumulate this object instead of the energy. For the estimation of the error bars, we perform the same computation several times and then calculate the standard error on the mean for each point.

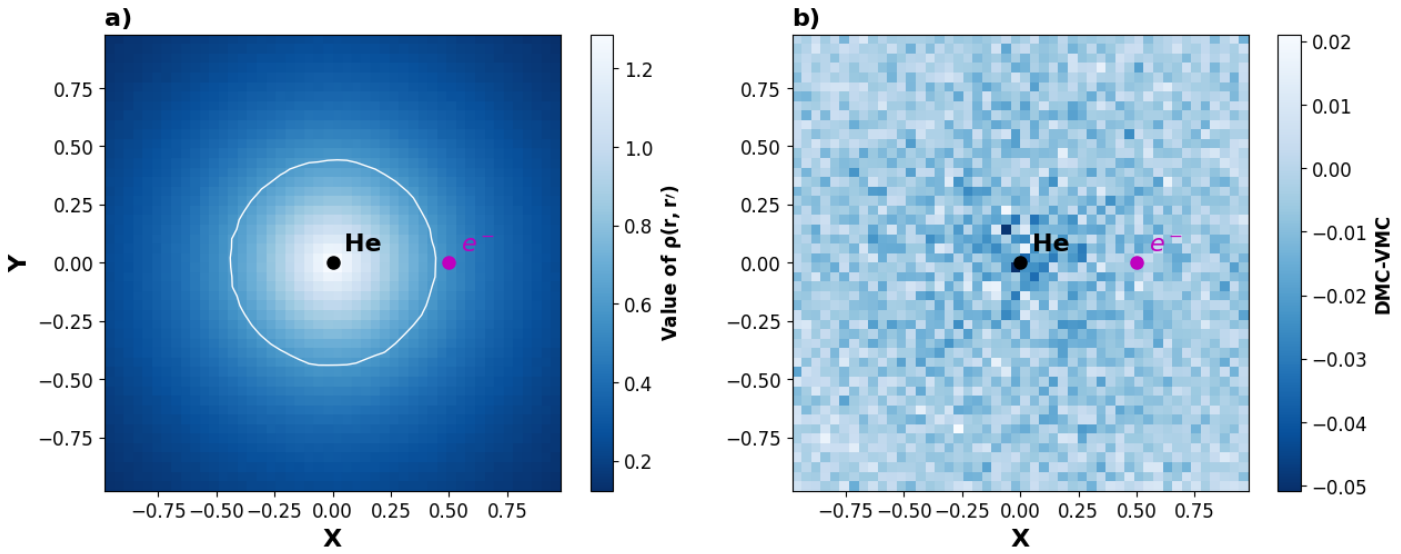

Figure 6: a) Reduced density matrix calculated with Diffusion Monte Carlo, b) Reduced density matrix with DMC minus the one calculated with VMC. It can quickly be seen that our trial wavefunction is of high quality, since the VMC and DMC results are very close to each other. Since our object $\rho(r, r')$ is a six dimensional tensor, we must slice it to obtain a 2d map, the slicing is done as follows, $r' = (0.5, 0, 0)$ and we set $z = 0$ so as to obtain a two-dimensional map. The error bars are of the order of 10^{-3} for the first figure and 10^{-2} for the second one.

The image shown in Figure 6a) is an off-diagonal slice of the reduced density matrix (only one point in the figure belongs to the diagonal), and thus it shows the correlation in the system. At first inspection, the slice presented in Figure 6a) seems to be spherically symmetric. We would expect the 1RDM not to be spherically symmetric, as the electronic correlation is distance dependant. We can see that the 1RDM does not present spherical symmetry in Figure 7. This non spherical terms can also be seen in the wavefunction in Figure 5. In Figure 7 we can see that both VMC and DMC allow for a good calculation of reduced density matrices, producing results that are in agreement with the exact one. We can conclude that at least in the helium atom, the 1RDM can be calculated with both QMC methods. As a further bechmark we have computed $\rho(r = r', r')$ and compared it against the electronic density $n(r')$, see Figure 8.

$$\rho_{VMC}(r', r') = 0.535(8),\ n_{VMC}(r') = 0.601(6),\ \rho_{DMC}(r', r') = 0.532(4),\ n_{DMC}(r') = 0.595(4) \quad (4.2.4)$$

We can see that while our results are not in agreement, they are close. One reason for this can be that we are not sampling on the exact same grid or the length of the calculation. For the DMC results, the problem could be due to the mixed estimator error.

The mixed estimator error in Diffusion Monte Carlo is an error that arises from the fact that what we are truly sampling is $\langle \psi_T | \psi_0 \rangle$. This problem appears for sure if one tries to calculate the expected value of an observable which does not commute with the Hamiltonian. A quick remedy is to use the so called extrapolated estimator [4].

$$\langle \psi_0 | \hat{S} | \psi_0 \rangle \approx 2 \langle \psi_0 | \hat{S} | \psi_T \rangle - \langle \psi_T | \hat{S} | \psi_T \rangle + \mathcal{O}[(\psi_0 - \psi_T)^2] \quad (4.2.5)$$

The main drawback of this technique is that we rely on the quality of our trial wavefunction. Since qualitatively we do not observe a difference in our DMC and VMC results we cannot at this stage perform a benchmark of the mixed estimator error. A more robust way of computing averages is to use path integral QMC methods, such as reptation QMC.

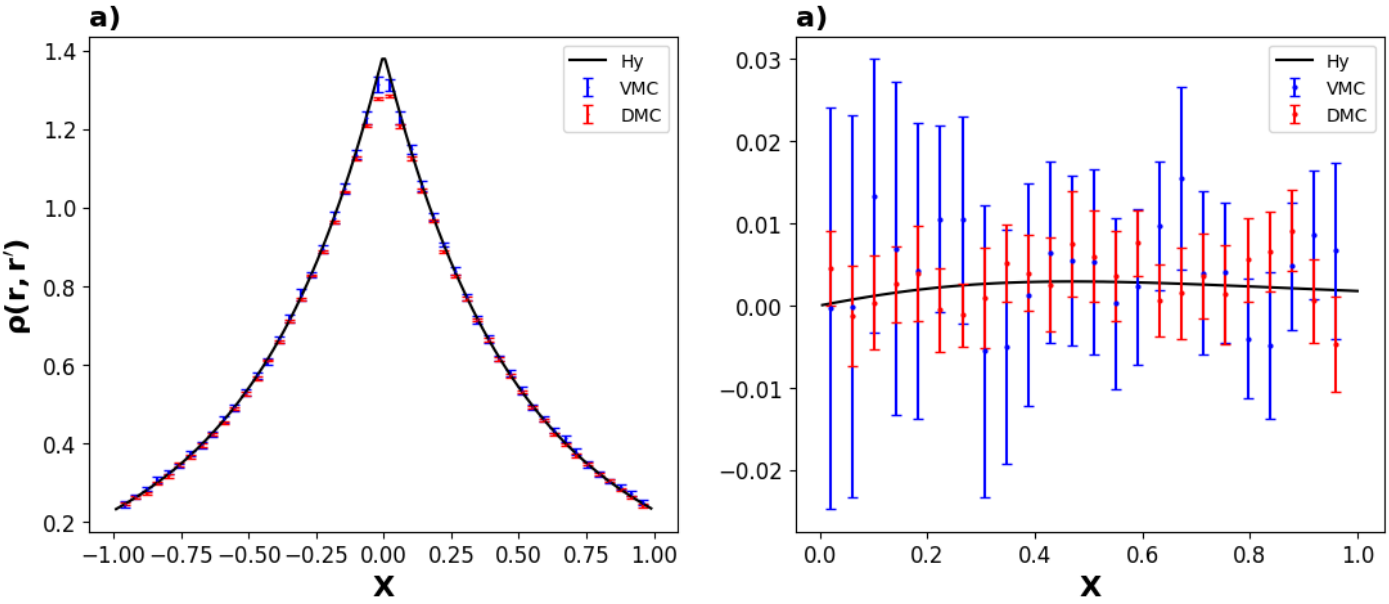

Figure 7: a) Slice of the 1RDM calculated with DMC, VMC and the exact one. b) $x > 0$ part minus $x < 0$ part of a). The slicing is done as follows, $r' = (0.5, 0, 0)$ and $z = y = 0$. We can clearly see that both methods work well, with small disparities at the nuclear position, this is to be expected as the $r = 0$ point is not included in the simulation, thus there is some smearing of the cusp.

In the case of the helium atom we can see that VMC is already qualitatively really good at computing the 1RDM. This is seen in Figure 7a. It is only by looking at features that are two orders of magnitude smaller (Figure 7b), that we can appreciate that DMC is more precise and slightly more accurate. If we were to work with a system in which correlations were more significant, VMC might prove a good option for the computation of the 1RDM. (In realistic systems we have much more electrons, this can lead to larger features that could be captured with VMC.)

The calculation of reduced density matrices is an important step towards understanding a material. It allows to study the degree of correlation, calculate orbital occupation numbers, calculate the natural orbitals and more. We can conclude that if a qualitative analysis is sought, VMC seems to be a viable method. However, if observables are to be computed with the 1RDM, we need much more precise calculations. For example, one calculation that can be done as a benchmark is the kinetic energy with the 1RDM.

$$K[\rho] = -\frac{1}{2}\int \nabla^2 \rho(r, r')\Big|_{r=r'} dr' \tag{4.2.6}$$

We have tried to compute the kinetic energy K, but the conclusion was not satisfactory since the result is way too sensitive to small changes in the 1RDM.

4.3 Electronic density

In condensed matter the density is quite an important object, as we know from the Hohenberg-Kohn theorem [14], we can in theory obtain any observable as a functional of the electronic density. This means that computing the density allows us, through the use of these functionals get observables. We can indeed compute the electronic density of helium with Quantum Monte Carlo methods.

The density is the integration of the wavefunction squared except on one coordinate. In our case, the density is, (it is also the diagonal of the one body reduced density matrix, which has been introduced

in Section 4.2, Equation 4.2.1)

$$n(r) = \int dr' |\psi(r, r')|^2 + \int dr' |\psi(r', r)|^2 = 2 \int dr' |\psi(r, r')|^2 \tag{4.3.1}$$

This is done because each electron has a different spin, and thus the spatial part of the wavefunction is symmetric under position exchange. This expression can quickly be turned into a Monte Carlo integral, as the integrand itself is a distribution; thus, our accumulator only has to be 1 to be able to use Equation 2.1.1.

$$n(r) = 2 \int dr' |\psi(r, r')|^2 = 2 \int dr' 1 \cdot |\psi(r, r')|^2 \approx 2 \frac{1}{N} \sum_{r_i}^{N} 1 \quad \text{with } r_i \sim |\psi(r, r')|^2 \tag{4.3.2}$$

What this equation means is that at each step, we just add one to where our electrons are. When we are done we divide by the number of steps, getting a density histogram of the positions, which is exactly what we want. This allows for a very quick calculation of the density on a grid, especially for the significantly faster VMC method. The error bars are obtained in the same way as those presented in Section 4.2.

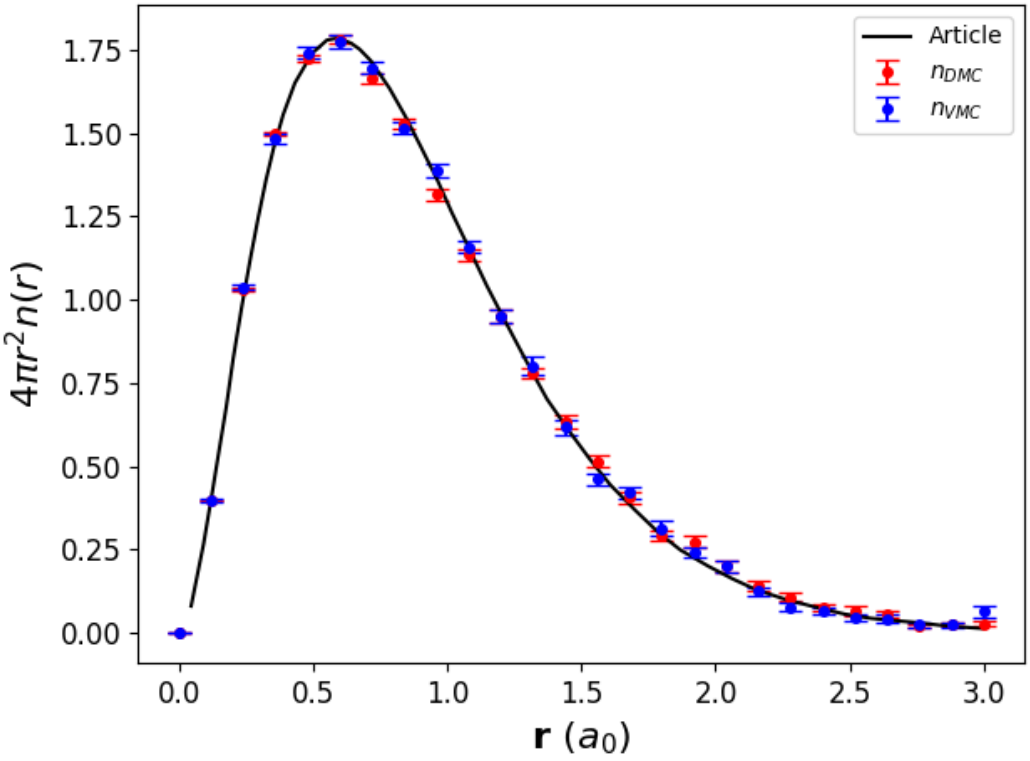

Figure 8: Radial density obtained with Diffusion Monte Carlo, Variational Monte Carlo and density obtained from [15]. Our calculations are done on a 51 points grid spanning from $[-3, 3]$ in the three dimensions, here we only see 26 points as we only show the radial part.

The results of our VMC and DMC calculations are shown in Figure 8 together with those obtained from [15]. We can quickly see that qualitatively at least for the density, Variational Monte Carlo is good enough as it is within error bars of the two other results. A small qualitative difference can be seen: DMC's value is slightly lower than VMC at the start and then a bit larger at the tail. This seems to point out that the ground state is a bit more delocalized than our trial wavefunction. It can also be understood as a screening of the nuclear charge.

The method that we have used is to place the electrons in whichever grid point was closer to them, this means that electrons beyond $3a_0$ are always placed at $3a_0$. This is the reason why both methods yield slightly larger values at $3a_0$ than expected.

4.4 QMC benchmark summary

In this section we will present a summary of the computational efforts and accuracy of different observables obtained with VMC and DMC during this work, see Table 2.

Quantity	Value	Relative error	N walkers (10^4)	N steps (10^3)	τ	Grid points
E_{VMC}	-2.89128	$4.5e^{-5}$	10	5	0.35	–
E_{DMC}	-2.90374	$7.9e^{-5}$	66	25	0.0015	–
$E_{0\ timestep}$	-2.903754	$1.9e^{-5}$	–	–	–	–
E^*_{VMC}	-2.130579	$1.6e^{-5}$	5	1	0.35	–
ρ_{VMC}	figure 6	0.019	2	20	0.35	50
ρ_{DMC}	figure 6	0.011	2	20	0.01	50
n_{VMC}	figure 8	0.091	1	5	0.35	51
n_{DMC}	figure 8	0.079	1	5	0.01	51

Table 2: Comparison table. The relative error is calculated as the error bar divided by the value. The relative error for the 2d objects is calculated as the average of the relative error at all points. It must be said that the large errors in the reduced density matrix indicate that we need more points to average, and more walkers to completely fill the whole space, as there are some points that are never visited and thus have an undefined relative error. E^*_{VMC} is our excited-state VMC energy. $E_{0\ timestep}$ is the energy obtained from 0 time step extrapolation, see Figure 3.

The results shown in Table 2 are extremely interesting. First, it tells us that calculating energies with DMC is indeed more expensive but much more accurate. We can also see that when it comes to density matrices and densities, DMC and VMC perform in a very similar way. This means that, for further studies, one can use the much less expensive VMC for obtaining these objects.

5 Conclusion

The basic theory of Quantum Monte Carlo methods has been explained. We have performed very accurate calculations of the energy of the ground state of helium, obtaining an energy of $E_{VMC} = -2.8913(1)\ Ha$ and $E_{DMC} = -2.90374(3)\ Ha$. The wavefunction used in VMC has been compared to the exact one, showing that the trial wavefunction used is of good quality. A way for improvement has been provided. We have also obtained the electronic density of the helium atom within error bars of [15]. Furthermore we have done a systematic comparison of QMC with an exact solution, allowing us to benchmark the accuracy of the 1RDM that QMC methods provide. We can see that even with a not exact trial wavefunction, a qualitatively good 1RDM can be obtained. We can also see in Table 2 that the computational cost for accurate and precise 1RDMs is not much larger than for the energies (This means that a run that accumulates the energy should also be able to obtain a precise 1RDM). We also want to highlight that if observables have to be computed with the 1RDM, much more precise and accurate calculations are needed. As a last remark we want to state that in the 1RDM and in the electronic density, the differences between VMC and DMC are only seen when one looks at the details

in the structure of the results. This means that VMC can be a useful tool for the calculation of these objects.

As a further study, it would be interesting to compare density matrices obtained from several methods, allowing for a full benchmark of the many-body theories that allow the calculation of this object. As of now, the one body reduced density matrix is an object that has not received a lot of attention, but it holds a lot of information. Most studies calculate it with QMC, but comparing other methods might tell us that we do not need the whole computational power of QMC for this object, and that maybe GW or DFT suffice. It would also be useful to benchmark the mixed estimator error and see to what degree it changes the real DMC solution.

Also, a similar study could be conducted for the alkaline earth metals. They should allow for a simple way of benchmarking pseudopotentials, as the valence electrons are the same as those of helium. Also the code developed can quickly be extended to the calculation of other atoms, with a natural next target being the lithium atom or the hydrogen molecule. Both would present different challenges and might lead to interesting results about the accuracy of QMC.

A natural continuation of this work is to compute the reduced density matrix of a metal, a semiconductor and a metavalent material and compare their behaviour. This could lead some insight into the interesting properties that these materials have and, hopefully, a deeper understanding of the large dielectric functions of metavalent materials. This could pave the way for new materials in photovoltaic cells.

References

[1] M. Wuttig, C.-F. Schön, J. Lötfering, P. Golub, C. Gatti, and J.-Y. Raty, "Revisiting the nature of chemical bonding in chalcogenides to explain and design their properties," *Advanced Materials*, vol. 35, no. 20, p. 2208485, 2023.

[2] E. Hylleraas, "Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium," *Z. Physik*, vol. 54, p. 347–366, 1929.

[3] W. L. McMillan, "Ground state of liquid He^4," *Phys. Rev.*, vol. 138, pp. A442–A451, Apr 1965.

[4] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," *Rev. Mod. Phys.*, vol. 73, pp. 33–83, Jan 2001.

[5] D. M. Ceperley, L. Reining, and R. M. Martin, *Interacting Electrons*. Cambridge: Cambridge University Press, 2016.

[6] C. Pierleoni, K. T. Delaney, M. A. Morales, D. M. Ceperley, and M. Holzmann, "Trial wave functions for high-pressure metallic hydrogen," *Computer Physics Communications*, vol. 179, no. 1, pp. 89–97, 2008.

[7] T. Kato, "On the eigenfunctions of many-particle systems in quantum mechanics," *Communications on Pure and Applied Mathematics*, vol. 10, no. 2, pp. 151–177, 1957.

[8] R. P. Feynman and M. Cohen, "Energy spectrum of the excitations in liquid helium," *Phys. Rev.*, vol. 102, pp. 1189–1204, Jun 1956.

[9] K. Frankowski and C. L. Pekeris, "Logarithmic terms in the wave functions of the ground state of two-electron atoms," *Phys. Rev.*, vol. 146, pp. 46–49, Jun 1966.

[10] T. Kinoshita, "Ground state of the helium atom," *Phys. Rev.*, vol. 105, pp. 1490–1502, Mar 1957.

[11] J. Li, N. D. Drummond, P. Schuck, and V. Olevano, "Comparing many-body approaches against the helium atom exact solution," *SciPost Phys.*, vol. 6, p. 040, 2019.

[12] R. J. Needs, M. D. Towler, N. D. Drummond, P. López Ríos, and J. R. Trail, "Variational and diffusion quantum Monte Carlo calculations with the CASINO code," *The Journal of Chemical Physics*, vol. 152, p. 154106, 04 2020.

[13] Z. Sukurma, M. Schlipf, M. Humer, A. Taheridehkordi, and G. Kresse, "Towards large-scale AFQMC calculations: Large time step auxiliary-field quantum Monte Carlo," 2024.

[14] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, pp. B864–B871, Nov 1964.

[15] S. Ragot and M. B. Ruiz, "Fourier–Legendre expansion of the one-electron density matrix of ground-state two-electron atoms," *The Journal of Chemical Physics*, vol. 129, no. 12, p. 124117, 2008.

[16] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *The Journal of Chemical Physics*, vol. 21, pp. 1087–1092, 06 1953.

[17] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

A Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm, introduced in 1953 by Metropolis et al. [16] is an algorithm that starting from an initial distribution, allows us to evolve it and obtain a target distribution proportional to the initial one. The evolution of our distribution is performed in subsequent steps, and the evolution of each of these steps depends only on the current distribution, making the Metropolis-Hastings algorithm a Markov chain [17]. Markov chains are stochastic processes that depend only on the current state of the system and not on previous steps.

A graphical example of this algorithm is shown in Figure 9 where we can see that this algorithm does evolve our positions to end up distributed according to some target distribution. Let us remark that in the Monte Carlo notation, $R = \{r_1, r_2, ..., r_N\}$ will denote the full position of all our particles.(We will call 'walker' to one set of electron positions)

Algorithm 2: Metropolis-Hastings

Let us examine a basic pseudo code for this algorithm.

i) Draw a set of random numbers for the initial position of our electrons R_0 in the way you think most resembles the target distribution $n(x)$. For example in Variational Monte Carlo (Section 2.2) will be the square of the modulus of the trial wavefunction.
ii) Propose a new configuration $R' = R_0 + \sqrt{\tau}\chi$ with χ a random Gaussian number.
iii) Compute the acceptance probability for each move $a = n(R')/n(R_0)$
iv) if $a > u$ we set $R_1 = R'$ else, $R_1 = R_0$ with u a uniform random number between 0 and 1.
v) Repeat steps ii-iv until you get the convergence you needed.

It can be seen how these moves allow the system to eventually reach all of configuration space, and thus the method is ergodic.

This algorithm is normally done with a large number of configurations at the same time, so as to make the code more efficient. The parameter τ is related to how large the moves can be, it will control the acceptance ratio of the algorithm. It can be seen that if $\tau = 0$ then $a = 1$ and thus all moves are accepted. The ideal acceptance ratio depends on the Monte Carlo method that is being implemented.

One important point to mention is that in the Metropolis algorithm we only need the ratio $n(R')/n(R_0)$, and thus we do not need to know the normalization constant of the probability distribution, this will be specially useful for wavefunctions as this means that we need not calculate the normalization.

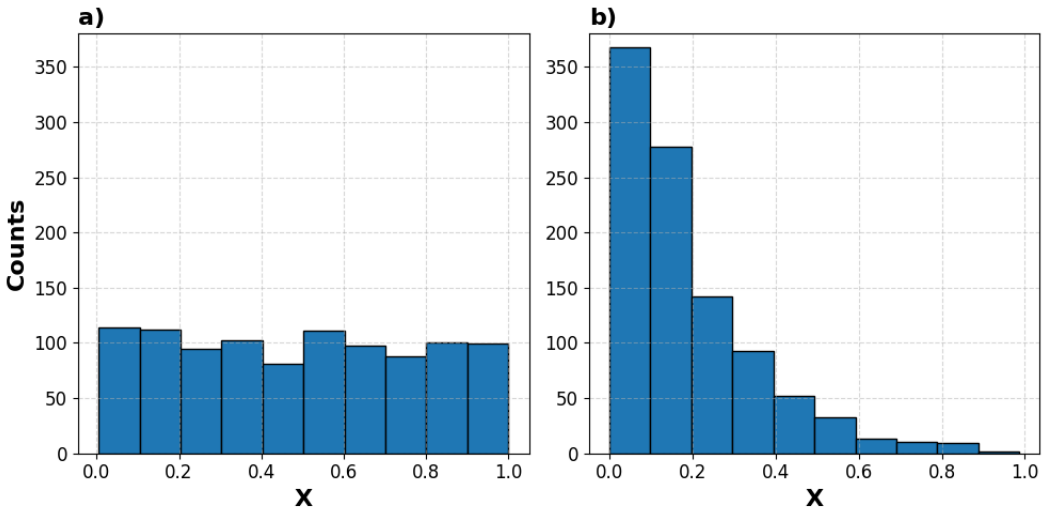

Figure 9: Example of the a metropolis algorithm run with 1000 configurations. a) Initial distribution taken to be a constant. b) Final distribution after 10 steps with $\tau = 0.5$. Target distribution is a decaying exponential $n(x) = e^{-5x}$ defined only in the interval $x \in [0, 1]$.

When performing a calculation, we will continue to perform this algorithm even after obtaining our distribution, this is done to get several data points that allow for averaging and error bar estimation. The way this algorithm is implemented introduces autocorrelation in our samples, this is, one step is not fully independent of the last one as they should in a Markov chain. There are two ways to mitigate this.

i) Use many independent configurations at once and get the average for each step over these configurations.

ii) Calculate the autocorrelation length t_c and perform a blocking of the data into blocks of this size. (One can also accumulate the energy every t_c steps)

In the calculations presented in this report both strategies have been implemented so as to increase the confidence of the error bars. Another strategy to reduce autocorrelation is to increase the time step, but this leads to large inefficiencies in Variational Monte Carlo, and to large errors in Diffusion Monte Carlo.

B Pair density approximation

Let us begin this section by stating that the approximation to the pair density presented here is not correct. We are currently working on finding a way to accurately sampling the pair density of helium.

The pair density $n^{(2)}(r, r')$ is an object which tells us how likely it is to find an electron in a position r if there is another in r'. In the case of two electrons, the pair density is conveniently defined as,

$$n^{(2)}(r, r') = 2|\psi(r, r')|^2 \tag{B.0.1}$$

In the general case it is the diagonal of the two body reduced density matrix, a much more complex object defined as (for two electrons we need not even integrate)

$$\rho^{(2)}(r, r' : r'', r''') = N(N-1) \int \psi(r, r', r_3..., r_N)\psi^*(r'', r''', r_3..., r_N)dr_3...dr_N \tag{B.0.2}$$

This object will allow us to directly compare our simulations with the Hylleraas wavefunction. To do this we will simply compute this object with one electron fixed and calculate the square root. This will yield a wavefunction comparable to that shown in Figure 5. The fixing of one electron is done to make this object smaller and more feasible in its calculation.

This has the advantage that it can be implemented for Diffusion Monte Carlo calculations, allowing us to compare two solutions which exhibit similar energies and accuracies. It is important to remember that the comparisons we are about to do work only for systems of two electrons, so this is not a general comparison that can be done. Nevertheless, for our purposes it will be extremely convenient.

First we must transform Equation B.0.1 into a Monte Carlo integral for one electron fixed at our target position $v = (0.5, 0, 0)$.

$$n^{(2)}(r, v) = 2|\psi(r, v)|^2 = 2\int dr_1 dr_2 |\psi(r_2, r_1)|^2 \delta(r_1 - v)\delta(r_2 - r) = 2\frac{1}{N}\sum_{r_{1i}, r_{2i}}^{N} \delta(r_{1i} - v)\delta(r_{2i} - r) \tag{B.0.3}$$

$$\text{such that } (r_{2i}, r_{1i}) \sim |\psi(r_2, r_1)|^2$$

As an approximation we decided to sample the conditional probability. For this our code already has an electron fixed at $r' = v$ and only moves the other electron. This means that what we actually obtain is the normalized conditional probability distribution of electron one $p(r|v)$ and then multiply by the electronic density at v. This is done so that our object $n^{(2)}(r, r')$ integrates to two over both coordinates.

$$n^{(2)}(r, v) = p(r|v)n(v) \tag{B.0.4}$$

By the use of this we obtain the results shown in Figure 10.

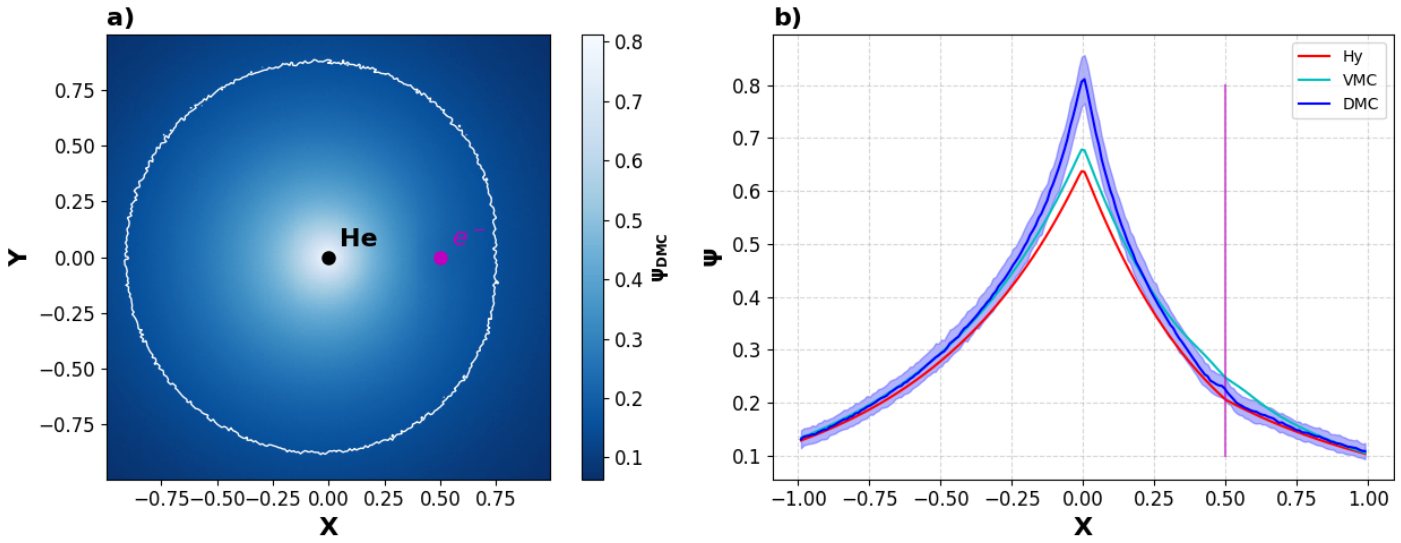

Figure 10: The DMC wavefunction presented here is not correct. a) Wavefunction obtained from pair density within DMC scheme, with isocurve at 0.15. b) All our wavefunctions in the $y = z = 0$ line (the DMC wavefunction has its errorbars as a shaded region), the position of the second electron is shown with the pink line. For subfigure a) the large similarity with Figure 5a) can be seen. The electronic repulsion can be seen in the difference of slopes of the positive and negative side. Nuclear screening can also be appreciated by the change of slope after $x = 0.5$.

The results shown in Figure 10 show an incredibly interesting result, our VMC wavefunction is superior

at the nucleus point, but it fails to capture the change of slope after the electron point (screening). This could be addressed with a different Jastrow factor. We must remark that the DMC problems at the cusp are due to the incorrectness of the approximation that we have used. DMC also exhibits a small bump at the electron-electron position, we believe this to also be an artifact of the sampling method used. Still our wrong approximation does capture the nuclear screening.