Variational Monte Carlo methods for quantum dots

Matteo Seclì <mattes@mail.uio.no>

December 8, 2014

Abstract

The aim of this project is to use the Variational Monte Carlo (VMC) method to evaluate the ground state energy, one-body densities, expectation values of the kinetic and potential energies and single-particle energies of quantum dots with $N = 2$ and $N = 6$ electrons, so-called *closed shell systems*. We will begin with two electrons confined in a pure two-dimensional isotropic harmonic oscillator potential, with and without the repulsive interaction, developing a simple trial wave-function and a basic program that runs our calculations. Then we are going to improve the trial wave-function adding the so-called *Jastrow factor*, and finally we are going to extend the entire class of wave-functions and the program itself to a given number of electrons (obviously, for a closed-shell system). After that we are going to improve our sampling, switching from a brute-force approach to the clever *importance sampling*. In addition, we will present a method that implements analytical derivatives for calculating both the energy and the diffusive force of the importance sampling.

Contents

1 Introduction and theoretical basis

1.1 The Variational Principle

The Variational Principle is a method of general validity that can be used to gather information about a system with a Hamiltonian that we are unable to diagonalize (i.e., we can't solve the Schrödinger equation for that system). Specifically, this principle gives us an *upper bound* for the energy of the ground state, that we will call here E_{gs}. The formulation is astonishingly simple: if you pick *any state* $|\psi\rangle$ *whatsoever*, then

$$E_{\text{gs}} \leq \frac{\left\langle \psi \left| \hat{H} \right| \psi \right\rangle}{\langle \psi \,|\, \psi \rangle}. \tag{1}$$

In this project we are going to use just this formula, because you see that the normalization of the state is not necessary. But if you have a normalized state, then equation (1) is further simplified in

$$E_{\text{gs}} \leq \left\langle \psi \left| \hat{H} \right| \psi \right\rangle \equiv \left\langle \hat{H} \right\rangle. \tag{2}$$

This equation seems a kind of magic but actually the proof of this fact is really simple, and we are going to sketch here a proof to show the power of this method (roughly, as it appears in [Griffiths]).

PROOF: Since the (unknown) eigenstates $|\psi_n\rangle$ of $\hat{H}$ form a complete set, we can expand our random state $|\psi\rangle$ on this basis:

$$|\psi\rangle = \sum_n C_n |\psi_n\rangle, \qquad C_n \in \mathbb{C},$$

where the $|\psi_n\rangle$'s are such that

$$\hat{H} |\psi_n\rangle = E_n |\psi_n\rangle.$$

Then, we have

$$\left\langle \psi \left| \hat{H} \right| \psi \right\rangle = \sum_{nm} C_m^* C_n E_n \underbrace{\langle \psi_m \,|\, \psi_n \rangle}_{\delta_{mn}}$$

$$= \sum_n |C_n|^2 E_n,$$

$$\langle \psi \,|\, \psi \rangle = \sum_n |C_n|^2.$$

Since $E_n \geq E_{\text{gs}} \, \forall n$, it follows immediately that

$$E_T \doteq \frac{\left\langle \psi \left| \hat{H} \right| \psi \right\rangle}{\langle \psi \,|\, \psi \rangle} = \frac{\sum_n |C_n|^2 E_n}{\sum_n |C_n|^2} \geq \frac{\sum_n |C_n|^2 E_{\text{gs}}}{\sum_n |C_n|^2} = E_{\text{gs}}. \tag{3}$$

∎

In practice, one chooses a class of states for $|\psi\rangle$ parametrized by one or more parameters – the so-called *variational parameters*, and then calculates the quantity E_T for multiple sets of values of the variational parameters $\alpha_1 \ldots, \alpha_n$. The lower value of E_T obtained in this way is the required upper bound. If one manages to find an even lower value for E_T with a more clever wave-function, then he has found a better upper bound for the ground state energy.

The only trouble with this method is that we never know for sure how close we are to the *actual* ground state energy; all we get for sure is just an upper bound. However, if one manages to guess a *realistic* $|\psi\rangle$, he often gets values for the ground state energy that miraculously match the actual ones.

In the following, we represent $|\psi\rangle$ in its function-representation as

$$|\psi\rangle \simeq \psi_T,$$

where "$\simeq$" means "represented by" and the subscript T is just to remember that ψ_T is our *trial* wave-function.

In our case the Hamiltonian lies in an infinite-dimension Hilbert space, so we have to replace the sums in (3) with integrals over all the space. Indicating with $d\vec{\tau}$ a space element, we can write

$$E_T = \frac{\int d\vec{\tau} \, \psi_T^* \hat{H} \psi_T}{\int d\vec{\tau} \, |\psi_T|^2}. \tag{4}$$

Calculating E_T as it appears in (4) is not a piece of cake. However, we can recast that equation in a simpler form if we multiply and divide by ψ_T at the left of $\hat{H}$. In fact,

$$E_T = \frac{\int d\vec{\tau}\, \psi_T^* \frac{\psi_T}{\psi_T} \hat{H} \psi_T}{\int d\vec{\tau}\, |\psi_T|^2} \tag{5}$$

$$= \frac{\int d\vec{\tau}\, |\psi_T|^2 \frac{1}{\psi_T} \hat{H} \psi_T}{\int d\vec{\tau}\, |\psi_T|^2} \tag{6}$$

$$= \int d\vec{\tau}\, \frac{|\psi_T|^2}{\int d\vec{\tau}\, |\psi_T|^2} \frac{1}{\psi_T} \hat{H} \psi_T \tag{7}$$

$$= \int d\vec{\tau}\, \mathcal{P}(\vec{\tau}) E_L(\vec{\tau}) \tag{8}$$

$$\simeq \frac{1}{n} \sum_{i=1}^{n} E_L(\vec{\tau}) \tag{9}$$

where

$$\mathcal{P}(\vec{\tau}) = \frac{|\psi_T|^2}{\int d\vec{\tau}\, |\psi_T|^2} \tag{10}$$

and

$$E_L(\vec{\tau}) = \frac{1}{\psi_T} \hat{H} \psi_T. \tag{11}$$

The dependence of ψ_T on the position in space and on the variational parameters has been omitted just for better readability. The quantity E_L is called the *local energy*, and it's what we are going to calculate in this project. Obviously, for bigger n we obtain better results.

1.2   Considerations about the physical system

Our system is represented by a number N of electrons that have total energy [1]

$$\hat{H} = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 + \frac{1}{2}\omega^2 r_i^2 \right) + \sum_{i<j} \frac{1}{r_{ij}}, \tag{12}$$

where $r_{ij} = |\vec{r}_i - \vec{r}_j|$ and natural units ($\hbar = c = e = m_e = 1$) are used in order to have the energy in atomic units. You see that the Hamiltonian includes a standard part (an harmonic oscillator) plus a repulsion potential, that is the one that gives troubles in analytical calculations.

An important feature of our system is that it is made up of *identical particles*. Let's try to exploit this feature to simplify our trial wave-function.

Suppose that we have a system made up of two particles, let's call them 1 and 2. Then, the state $|\psi\rangle$ of the system can be expressed (for example) as

$$|\psi\rangle = |1\rangle \otimes |2\rangle. \tag{13}$$

Now we can introduce the *permutation operator* $\hat{P}$, defined as

$$\hat{P}|1\rangle \otimes |2\rangle \doteq |2\rangle \otimes |1\rangle. \tag{14}$$

One immediately sees that $\hat{P} = \hat{P}^\dagger$ and $\hat{P}^2 = \hat{I}$ ($\hat{I}$ is the identity operator), which means that $\hat{P}$ is hermitian and unitary. These two properties allow us to say that the eigenvalues of $\hat{P}$ are ± 1. In general, this operator $\hat{P}$ does not commute with $\hat{H}$; but – and here is the trick – it *does* commute with $\hat{H}$ for a system of *identical particles*. Since in that case $[\hat{H}, \hat{P}] = 0$, the eigenstates of $\hat{H}$ must also be eigenstates of $\hat{P}$. Recalling that the eigenvalue of $\hat{P}$ are ± 1, one can rewrite the state ψ as

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|1\rangle \otimes |2\rangle \pm |2\rangle \otimes |1\rangle \Big) \tag{15}$$

to make it also eigenstate of $\hat{P}$. You see that, in this form, the state is either symmetric or antisymmetric; particles with symmetric wave-function are called *bosons*, and particles with antisymmetric wave-function are

---

[1] Calling the Hamiltonian the total energy is improper. Rather it should be called the *generalized energy*, since in general it is *not* the total energy because it also contains terms of the 1st and 0th order in the generalized velocities. However in this case the system does not contain non-conservative forces (from which the extra terms stem), so we can safely identify the Hamiltonian with the total energy.

called *fermions*. It turns out that electrons are fermions, so we are interested only in the antisymmetric case. Rewriting the fermions wave-function in coordinate representation, we obtain:

$$\psi(\vec{r}_1, \vec{r}_2, \sigma_1, \sigma_2) = \Big(\langle\vec{r}_1| \otimes \langle\vec{r}_2| \Big)\, |\psi\rangle = \frac{1}{\sqrt{2}}\Big(\langle\vec{r}_1\,|\,1\rangle \langle\vec{r}_2\,|\,2\rangle - \langle\vec{r}_1\,|\,2\rangle \langle\vec{r}_2\,|\,1\rangle \Big) \tag{16}$$

$$= \frac{1}{\sqrt{2}}\Big(\phi_1(\vec{r}_1, \sigma_1)\phi_2(\vec{r}_2, \sigma_2) - \phi_2(\vec{r}_1, \sigma_1)\phi_1(\vec{r}_2, \sigma_1) \Big) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1(\vec{r}_1, \sigma_1) & \phi_1(\vec{r}_2, \sigma_1) \\ \phi_2(\vec{r}_1, \sigma_2) & \phi_2(\vec{r}_2, \sigma_2) \end{vmatrix} \tag{17}$$

where the ϕ_i's are the single-particle wave-functions, the σ_i's just indicate the explicit dependence on the spin, and the determinant in (17) is called the *Slater determinant*. For N particles the determinant has exactly the same shape (is just a $N \times N$ matrix) and the factor $\frac{1}{\sqrt{2}}$ is replaced with $\frac{1}{\sqrt{N!}}$. However, we don't care about the normalization factor because it's not required to apply the variational principle.

We can recast this equation in a simpler form. In fact, it's possible to rewrite $|\phi\rangle$ as

$$|\phi\rangle = |\eta\rangle \otimes |\chi\rangle, \tag{18}$$

where $|\eta\rangle$ is the *spatial part* and $|\chi\rangle$ is the *spin part*. Note that – since the Hamiltonian is spin-independent – the spin part just takes into account the spin configuration of the system, and it has nothing to do with the positions of the particles themselves! In other words, while the spatial part depends on the positions of the particles, the spin part does not at all and we can "forget" about it if we are smart.

In fact, the best way to minimize the energy is to put the two electrons in the same orbital. This means that the spatial part is *symmetric* under the interchange of two particles; but we have also said that this wave-function has to be antisymmetric. So, the spin part has to be *antisymmetric*. If you look at the Clebsch-Gordan table for combining two spin-1/2 (Figure 1), you see that the only antisymmetric combination is the total spin state $|0,0\rangle$, that can be written as

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle \right) \tag{19}$$

This fact permits a simplification of the Slater determinant; in fact, if we suppose – for example – that particle 1 has spin-up and particle 2 has spin-down, we can write

$$\Psi_T(\vec{r}_1, \vec{r}_2, \sigma_1, \sigma_2) = \begin{vmatrix} \phi_1(\vec{r}_1, \sigma_1) & \phi_1(\vec{r}_2, \sigma_1) \\ \phi_2(\vec{r}_1, \sigma_2) & \phi_2(\vec{r}_2, \sigma_2) \end{vmatrix} = \phi_1(\vec{r}_1)\phi_2(\vec{r}_2). \tag{20}$$

where the ϕ_i's are now independent on the spin. In other words, they are just the spatial parts η_i's of the single-particle wave-functions. You also see that, since now the single-particle wave-functions are independent on the spin, the are *exactly* the same function because they occupy the same energy level!

The generalization to N particles is straightforward, and will be given later.

Figure 1: The Clebsch-Gordan table for combining two spin-1/2. Source: http://pdg.lbl.gov/2002/clebrpp.pdf.

1.3   The Metropolis algorithm

Let's first briefly introduce what a *Markov chain* is. We call a Markov chain a stochastic process where the transition probability that characterizes the transition between states only depends on the previous step and not from the path to get to such step. We can then write (see [Hjorth-Jensen])

$$w_i(t + \delta t) = \sum_i W(j \to i) w_j(t), \tag{21}$$

where $w_i(t + \delta t)$ is our PDF – in our case $|\psi_T(R)|^2$ and $W(j \to i)$ is the probability of jumping from j to i, also called the *transition probability*. We also have

$$\sum_i w_i(t) = 1. \tag{22}$$

due to the normalization condition. We can see from equation (21) the probability density function at $t + \delta t$ only depends on its previous state t.

In the case of infinite dimensionality, representing W as a matrix, we have

$$\hat{w}(t + \delta t) = \hat{W}\hat{w}(t). \tag{23}$$

from which

$$\hat{w}(\infty) = \hat{W}\hat{w}(\infty). \tag{24}$$

Then, the problem is just an eigenvalue problem with eigenvalue 1.

Now, let's consider an *ergodic* system, where the average over time of a process is the same as the ensemble average. Or in other words, as one increases the steps, there exists a positive probability measurement at step n that is independent on the probability distribution at the initial step 0. It holds

$$W(i \to j)w_i = W(j \to i)w_j, \tag{25}$$

from which

$$R \doteq \frac{W(i \to j)}{W(j \to i)} = \frac{w_j}{w_i} = \frac{|\psi_T(r^{\text{new}})|}{|\psi_T(r^{\text{old}})|}. \tag{26}$$

Since $W(i \to j)$ is unknown, we model it as the product of two probabilities:

$$W(i \to j) = A(i \to j)\,T(i \to j), \tag{27}$$

where $A(i \to j)$ is the probability of accepting the move and $T(i \to j)$ is the probability to transition from the state i to the state j. In a similar way we can write:

$$W(j \to i) = A(j \to i)\,T(j \to i). \tag{28}$$

1.3.1   Brute force sampling

The *brute force* Metropolis algorithm simply consists in the assumption that the transition probability is symmetric, that is

$$T(i \to j) = T(j \to i), \tag{29}$$

which leads to

$$\frac{W(i \to j)}{W(j \to i)} = \frac{A(i \to j)}{A(j \to i)}. \tag{30}$$

Then, we can distinguish between two cases.

- If $\dfrac{w_j}{w_i} \geq 1$, then

$$A(i \to j) \geq A(j \to i). \tag{31}$$

 So, if we manage to move to such a state, we accept this move because we are moving to a higher probability region.

- If $\dfrac{w_j}{w_i} < 1$, then

$$A(j \to i) > A(i \to j). \tag{32}$$

 We can't blindly reject this move just because we are moving to a lower probability region, but we can't also accept all of this kind of moves. So, what one does is to compute a random number $r \in (0, 1)$ and to accept the move if $r \leq w_j/w_i$.

To summarize, one sets

$$A(i \to j) = \min\left(1, \frac{w_j}{w_i}\right) = \min\left(1, \frac{|\psi_T(r^{\text{new}})|^2}{|\psi_T(r^{\text{old}})|^2}\right) \tag{33}$$

and then compares with a random number $r \in (0, 1)$. If $A(i \to j) \geq r$ the step is accepted, otherwise it's rejected.

1.3.2   Importance sampling and Schödinger equation as a diffusion problem

Importance sampling consists in doing more clever assumptions on the shape of $T(i \rightarrow j)$. The discussion is the same as the brute force case, but this time we have to set

$$A(i \rightarrow j) = \min\left(1, \frac{w_j T(j \rightarrow i)}{w_i T(i \rightarrow j)}\right) = \min\left(1, \frac{|\psi_T(r^{\text{new}})|^2 T(j \rightarrow i)}{|\psi_T(r^{\text{old}})|^2 T(i \rightarrow j)}\right). \tag{34}$$

If we recast the Schrödinger equation as a diffusion problem, $T(i \rightarrow j)$ is going to be a Gaussian distribution (or a modified Gaussian); let's do that.

     The main idea is to model $T(i \rightarrow j)$ in order to reflect the shape of the trial wave-function itself and have in this way a more "intelligent" walk, that wastes less points. The walk is generated using the Fokker-Planck and the Langevin equations; the former – for a one-dimension diffusion problem – is

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - F\right)P(x,t), \tag{35}$$

where $P(x,t)$ is the time-dependent probability density, D is the *diffusion coefficient* (it's 0.5 for the Schrödinger equation, see [Høgberget]) and F is the *drift term*.

     The Langevin equation, on the other hand, reads as

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta, \tag{36}$$

being η a *random variable*. By solving equation (36) with Euler's method, it follows that the new position is given by

$$y = x + DF(x)\Delta t + \xi, \tag{37}$$

being ξ a *Gaussian* random variable and Δt a chosen time step.

     For two or more dimensions, the Fokker-Planck equation is just

$$\frac{\partial P}{\partial t} = \sum_i D\frac{\partial}{\partial \vec{x}_i}\left(\frac{\partial}{\partial \vec{x}_i} - \vec{F}_i\right)P(\vec{x}_i,t) \tag{38}$$

We can obtain a convergence to a stationary probability density setting the left hand side to zero, that means all the terms of the sum must be equal to zero. Expanding the product, we obtain

$$\frac{\partial^2 P}{\partial \vec{x}_i^2} = P\frac{\partial}{\partial \vec{x}_i}\vec{F}_i + \vec{F}_i\frac{\partial}{\partial \vec{x}_i}P. \tag{39}$$

Since the drift vector has the form $\vec{F} = g(\vec{x})\dfrac{\partial P}{\partial \vec{x}}$, we get

$$\frac{\partial^2 P}{\partial \vec{x}_i^2} = P\frac{\partial g}{\partial P}\left(\frac{\partial P}{\partial \vec{x}_i}\right)^2 + Pg\frac{\partial^2 P}{\partial \vec{x}_i^2} + g\left(\frac{\partial P}{\partial \vec{x}_i}\right)^2. \tag{40}$$

     If we want a stationary density the left hand side of equation (40) must be zero, and this is possible only if $g = \dfrac{1}{P}$. With this substitution, the drift vector reads as

$$\vec{F} = \frac{1}{P}\nabla P. \tag{41}$$

Since – as we said many times before – our probability density is the modulus squared of the trial wave-function, that is

$$P = |\psi_T|^2, \tag{42}$$

the drift vector in equation (41) finally reads as

$$\vec{F} = 2\frac{1}{\Psi_T}\nabla\Psi_T. \tag{43}$$

Written in this way, $\vec{F}$ is also called the *quantum force*. This force pushes the walker in those regions where the trial wave function is larger, increasing the efficiency of the simulation.

     Solving the Fokker-Plank equation finally gives an expression for the transition probability $T(x \rightarrow y)$, that turns out can be expressed as the Green function

$$G(y,x,\Delta t) = \frac{1}{(4\pi D\Delta t)^{3N/2}}\exp(-(y - x - D\Delta t F(x))^2/4D\Delta t). \tag{44}$$

     With this expression, the acceptance ratio in equation (34) becomes

$$R = \frac{|\psi_T(y)|^2 T(y \rightarrow x)}{|\psi_T(x)|^2 T(x \rightarrow y)} = \frac{|\psi_T(y)|^2 G(x,y,\Delta t)}{|\psi_T(x)|^2 G(y,x,\Delta t)}. \tag{45}$$
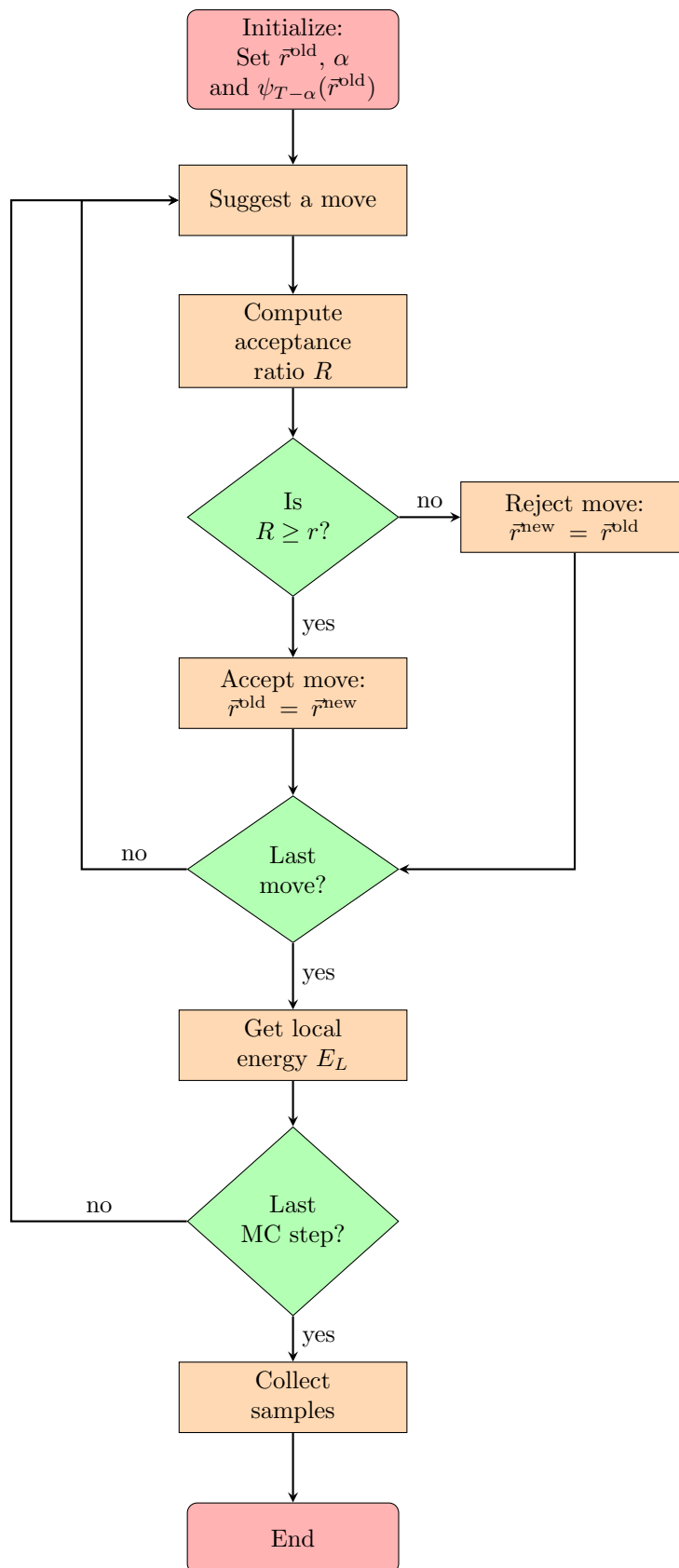
1.3.3   Implementation

Figure 2: Chart flow for the Quantum Variational Monte Carlo algorithm.

We can sum up the algorithm as follows. Given a random number generator that returns aleatory values in the interval $(0, 1)$ with a uniform (or Gaussian for the importance sampling) distribution, then the Metropolis

algorithm consists in these steps.

- Chosen a starting point x_0 and a step length l (or a time-step Δt for the importance sampling), a random number ε is generated in the interval $(0, 1)$.

- The proposed step is $x^{\text{new}} = x^{\text{old}} + l\varepsilon$ (or $x^{\text{new}} = x^{\text{old}} + D\vec{F}(x^{\text{old}})\Delta t + \varepsilon$ for the importance sampling).

- The probability ratio $R = \frac{|\psi_T(r^{\text{new}})|^2 T(j \to i)}{|\psi_T(r^{\text{old}})|^2 T(i \to j)}$ is computed.

- A new random number r is generated in the interval $(0, 1)$.

- If $R \geq r$, the step is accepted and the new position is stored by letting $x^{\text{old}} = x^{\text{new}}$.

- If $R < r$, the step is rejected and the new position is discarded by letting $x^{\text{new}} = x^{\text{old}}$.

A flow chart of this process is sketched in Figure 2.

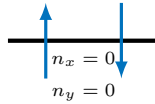# 2 The 2-electrons system

2.1 The unperturbed system

Figure 3: The 2-electrons system configuration.

We will first consider the unperturbed system (without the electron-electron repulsion) with $\omega = 1$, that has Hamiltonian

$$\hat{H}_0 = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 + \frac{1}{2}\omega^2 r_i^2 \right), \tag{46}$$

with $N = 2$ and in two dimensions. For a single electron in a two-dimensional (isotropic) harmonic oscillator, the energy is $E_s = \hbar\omega(n_x + n_y + 1)$, where n_x and n_y are the quantum numbers for the x and y dimensions. The ground state (See Figure 3) is obtained for $(n_x, n_y) = (0, 0)$ which gives $E_s = 1$ a.u.. Since we have two particles, the ground state energy of our system is in total

$$E_{\text{gs}} = 2 \text{ a.u.}.$$

In general, for $\omega \neq 1$, the energy will just be $E_{\text{gs}} = 2\omega$.

Now, we have to choose a trial wave-function. For a single electron, the exact wave-function of the ground state is

$$\phi_{n_x, n_y}(x, y) = AH_{n_x}(\sqrt{\omega}x)H_{n_y}(\sqrt{\omega}y)\exp\left(-\omega\left(x^2 + y^2\right)/2\right), \tag{47}$$

being A a normalization constant and H_n is the (physical) Hermite polynomial of order n. Since the 0-th order Hermite polynomials are just 1 and the two-particles wave-function can be seen as a product of the single-particle wave-functions, our total wave-function will just be

$$\phi(\vec{r}_1, \vec{r}_2) = C\exp\left(-\omega\left(r_1^2 + r_2^2\right)/2\right). \tag{48}$$

Adding a variational parameter α and forgetting about the normalization constant C, we obtain our trial wave-function for this case:

$$\psi_T(\vec{r}_1, \vec{r}_2) = \exp\left(-\alpha\omega\left(r_1^2 + r_2^2\right)/2\right). \tag{49}$$

Note that, since the correct solution of this problem is equation (48), we expect to obtain a variational energy *exactly* equal to 2 for $\alpha = 1$.

Figure 4: The program running for a two-electron system without repulsion.

The electron-electron repulsion can be simply turned off by commenting out the line

```
MySystem.add_potential(new eRepulsion(gP));
```

in the code. Running the program in Qt Creator shows the window in Figure 4, where some parameters can be specified before the actual calculation begins.

Just using a finer step for α, we obtain the energy graph shown in Figure 5.
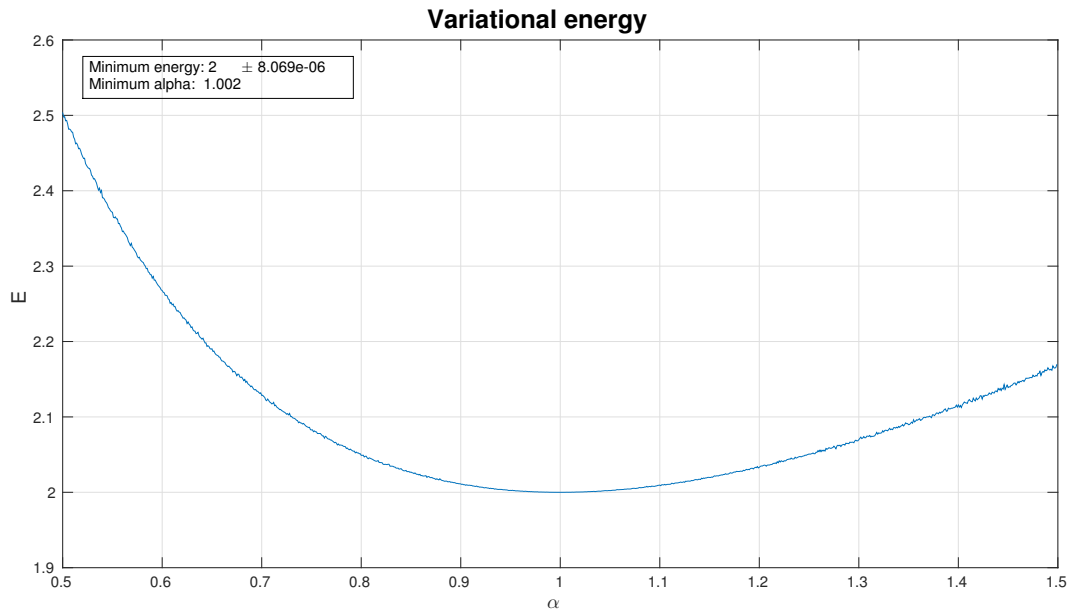

Figure 5: The variational energy versus the variational parameter α. The settings used are: brute force sampling with step length 2, no Jastrow factor, no parallelization, 1000 variations of α around 1 with step 0.001, 1×10^7 Monte Carlo steps. Acceptance ratio varies from 40 to 60 %.

As you can see in Figure 5, we get $E_{\mathrm{gs}} = 2$ with a high degree of precision for $\alpha = 1$[2].

We can also state that the total spin (better: its z-component) of the system is 0, as shown in equation (19). If you want a more "physical" reason than just reading off a table of coefficients, this is a direct consequence of the shape of the wave-function for a fermionic system, as shown in equation (17). In fact, in our case the spatial part of the two single-particle wave-functions is equal (they occupy the same energy level), yielding to

$$|1\rangle = |\phi\rangle \otimes |\chi_1\rangle, \qquad |2\rangle = |\phi\rangle \otimes |\chi_2\rangle. \tag{50}$$

---

[2]actually the minimum is for $\alpha = 1.002$, but in this case we are not using so many steps to trust the third decimal digit. The smaller α-step was just used to have a nicer plot.

Comparing with equation (15) one sees that $|\psi\rangle = 0$ unless $|1\rangle \neq |2\rangle$, and since the spatial parts are equal one has to conclude that $|\chi_1\rangle \neq |\chi_2\rangle$[3]. Now,

$$|\chi_1\rangle = |s, m_s\rangle \tag{51}$$

with $s = 1/2$, so we have to say that – for example – particle 1 has spin-up ($m_s = 1/2$) and particle 2 has spin-down ($m_s = -1/2$). This fact is also known as the *Pauli exclusion principle*, and applies for a generic system of fermions.

2.2 The complete system

We want now to analyse the system with the full Hamiltonian, namely the Hamiltonian in equation (12):

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{2}\omega r_1^2 + \frac{1}{2}\omega r_2^2 + \frac{1}{|\vec{r_1} - \vec{r_2}|}, \tag{52}$$

where $\vec{r_1}$ and $\vec{r_2}$ are the position operators of the two particles.

You see that this time we have a problem when two particles are very near to each other, since the term

$$\sum_{i<j} \frac{1}{|\vec{r_1} - \vec{r_2}|}$$

tends to be a division by zero. Introducing the substitution

$$\begin{cases} \vec{r} = \vec{r_1} - \vec{r_2} \\ \vec{R} = \frac{1}{2}(\vec{r_1} + \vec{r_2}) \end{cases} \tag{53}$$

the Hamiltonian can be rewritten in a decoupled way as[4]

$$\hat{H} = -\nabla_{\vec{r}}^2 - \frac{1}{4}\nabla_{\vec{R}}^2 - \frac{1}{4}\omega^2\vec{r}^2 + \omega^2\vec{R}^2 + \frac{1}{r}. \tag{54}$$

Since the Hamiltonian is independent on spin, we can factorize the wave-function as

$$\psi(\vec{r}, \vec{R}, m_{s_1}, m_{s_2}) = \varphi(\vec{r})\xi(\vec{R})\chi(m_{s_1}, m_{s_2}) \tag{55}$$

which separates the Schrödinger equation into

$$\left(-\frac{1}{2}\nabla_{\vec{r}}^2 + \frac{1}{2}\omega_r^2\vec{r}^2 + \frac{1}{2}\frac{1}{r}\right)\varphi(\vec{r}) = E_r\varphi(\vec{r}) \tag{56}$$

and

$$\left(-\frac{1}{2}\nabla_{\vec{R}}^2 + \frac{1}{2}\omega_R^2\vec{R}^2\right)\xi(\vec{R}) = E_R\xi(\vec{R}), \tag{57}$$

where $\omega_r = \frac{1}{2}\omega$ and $\omega_R = 2\omega$.

We are not interested in the R-part of the solution, since it gives no troubles; we are more interested in equation (56), because we can cancel the divergence for $r \to 0$ working on the solution. We can also discard about the oscillator potential because it gives no divergence. Since we are interested only in the radial part of the solution, the relevant contributions from the laplacian in spherical coordinates are

$$\nabla_r^2 = \frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}. \tag{58}$$

We also notice that $\frac{d^2}{dr^2}$ gives a finite contribution (is just the momentum part) and so does $\frac{d}{dr}$. The only piece that is not finite is the $\frac{1}{r}$ in front of $\frac{d}{dr}$, that we can use to cancel the divergence in the interaction potential. Using these facts, equation (56) can be rewritten in a way that is called the *cusp condition* for this system, namely

$$\lim_{r \to 0} \frac{1}{\varphi(r)}\left(-\frac{1}{2}\frac{\cancel{1}}{r}\frac{d}{dr} + \frac{1}{2}\frac{\cancel{1}}{r}\right)\varphi(r) = 0. \tag{59}$$

Solving for φ by partial integration gives

$$\varphi(r) = \exp(Cr), \qquad r \to 0 \tag{60}$$

---

[3]We don't want $|\psi\rangle = 0$ because it's not a physical state and it is not normalizable.
[4]See [Taut].

where C is a constant. For these types of Monte Carlo calculations, a so-called *Padé-Jastrow* factor that has the property (60) is often used. This factor has the form (for two electrons)

$$\exp\left(\frac{ar}{(1+\beta r)}\right),\tag{61}$$

where (in 2 dimensions)

$$\beta = \begin{cases} 1 & \text{anti-parallel spin} \\ 1/3 & \text{parallel spin} \end{cases}\tag{62}$$

and β is a variational parameter. With this factor, our trial wave-function reads as

$$\psi_T(\vec{r_1},\vec{r_2}) = \exp\left(-\alpha\omega\left(r_1^2+r_2^2\right)/2\right)\exp\left(\frac{ar}{(1+\beta r)}\right).\tag{63}$$

Using the ansatz in equation (63) in our program, we obtain a preliminary result shown in Figure 6.

Figure 6: The variational energy versus the variational parameters α and β. The settings used are: brute force sampling with step length 2, Jastrow factor, no parallelization, 200 variations of α and β with step 0.01, 1×10^5 Monte Carlo steps. Acceptance ratio varies from 45 to 55 %.

Refining around the minimum value to have better results, we finally obtain the plot in Figure 7.
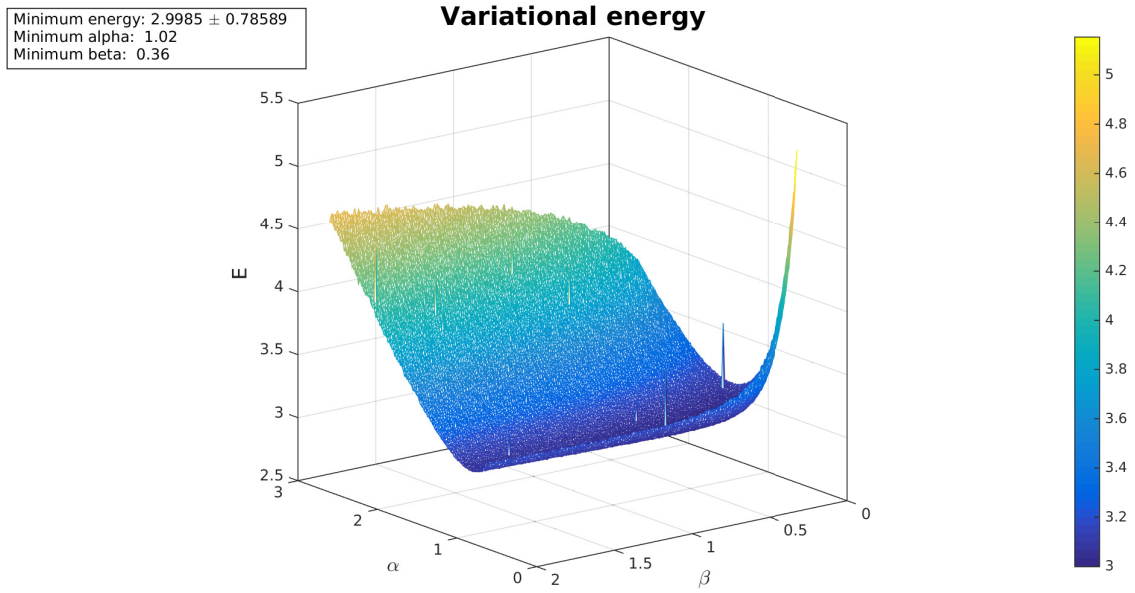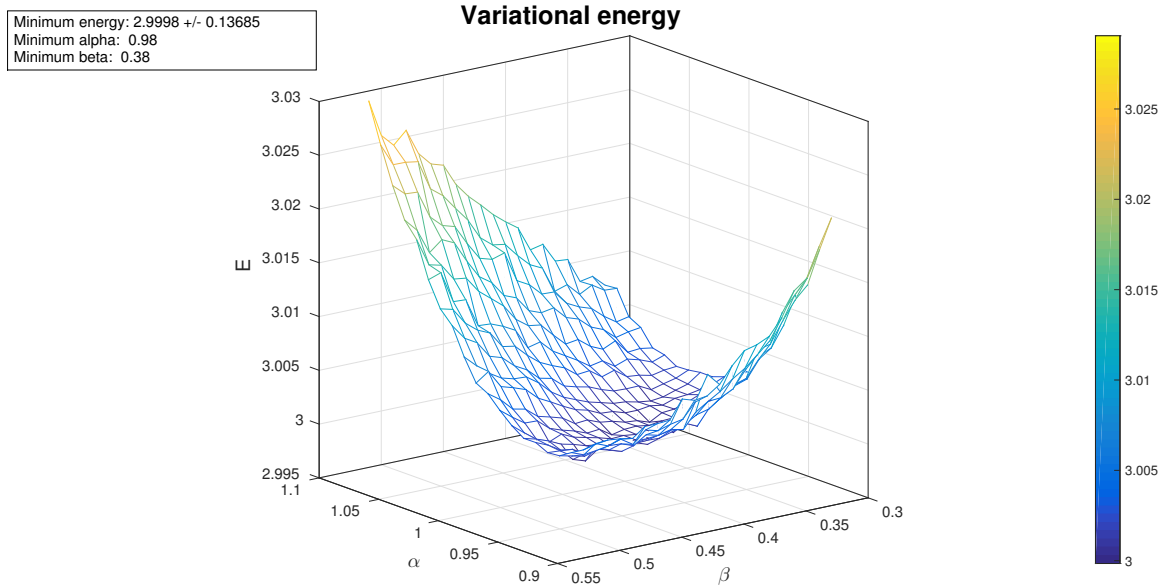
Figure 7: The variational energy versus the variational parameters α and β. The settings used are: brute force sampling with step length 2, Jastrow factor, no parallelization, 20 variations of α and β with step 0.01, 1×10^6 Monte Carlo steps. Acceptance ratio varies from 45 to 55 %. The relative distance is about 1.64 in natural units.

You see that we are very near to the exact result $E_{\mathrm{gs}} = 3$ a.u. calculated from [Taut] and reported in [Pedersen Lohne, et al.]. However, the energy error is still quite big; a way to improve the precision of our program is to *parallelize* it.

2.2.1 Parallelization

In this project, parallelization is done by using OpenMP[5]. The main idea is to run multiple computations of E_T for the same pair of variational parameters (α, β) and then take the mean. Since the "measurements" of E_T are not correlated, the error on its mean value falls like $1/\sqrt{n}$, where n is the number of measurements. To be sure that the measurements are not correlated we have to make *private* local copies of the variables used in the computation, one for each thread used. Then, we have to sync al the threads to be sure that all the measurements are complete, and finally we can take the mean. An example of how this is implemented in our program is given below:

```
// Begin parallelization
omp_set_num_threads(gP.num_threads);
#pragma omp parallel shared(cumulative_e, cumulative_e2)
{
    // Make a private copy of all the stuff
    mat cumulative_e_local = zeros<mat>(vmcParams.max_variations+1,vmcParams.
        max_variations+1);
    mat cumulative_e2_local = zeros<mat>(vmcParams.max_variations+1,vmcParams.
        max_variations+1);
    struct GeneralParams gP_local = gP;
    struct VariationalParams vP_local = vP;
    struct VMCparams vmcParams_local = vmcParams;

    // Build the orbitals
    Orbitals* MyWaveFunction_local = new AlphaHarmonicOscillator(gP_local, vP_local);

    // Set up the Jastrow factor
    Jastrow* MyJastrowFactor_local;
    if (vmcParams_local.jF_active) MyJastrowFactor_local = new Pade_Jastrow(gP,vP);
    else MyJastrowFactor_local = new No_Jastrow();

    // Make a local copy of the system for convenience
    System MySystem_local = MySystem;
```

---

[5]http://openmp.org/.

```
     // Do the mc sampling                                                   23
     mc_sampling(gP_local, vP_local, vmcParams_local,                        24
               cumulative_e_local, cumulative_e2_local,                      25
               MySystem_local, MyWaveFunction_local, MyJastrowFactor_local); 26
                                                                             27
     // Be sure to have all the contributions                                28
     #pragma omp barrier                                                      29
     #pragma omp critical                                                     30
     {                                                                        31
         // Add the contributions                                            32
         cumulative_e += cumulative_e_local;                                 33
         cumulative_e2 += cumulative_e2_local;                               34
     }                                                                        35
 }                                                                            36
                                                                             37
 // Normalize to the number of threads                                        38
 cumulative_e = cumulative_e/((double) gP.num_threads);                      39
 cumulative_e2 = cumulative_e2/((double) gP.num_threads);                    40
```

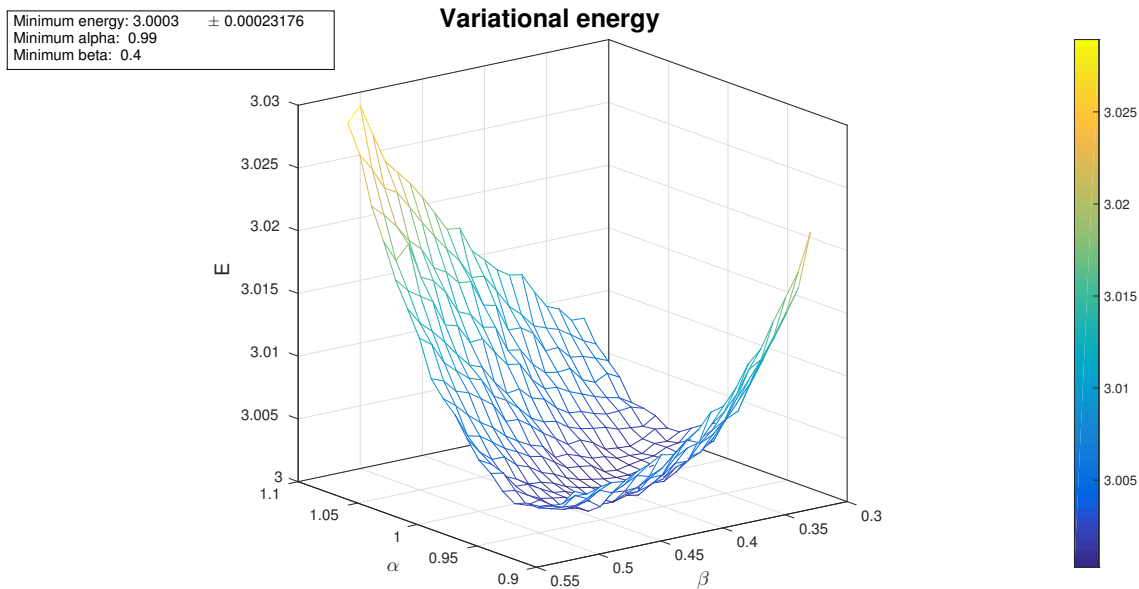Re-doing the calculations using parallelization gives the result in Figure 8.

Figure 8: The variational energy versus the variational parameters α and β. The settings used are: brute force sampling with step length 2, Jastrow factor, parallelization (8 threads), 20 variations of α and β with step 0.01, 2×10^5 Monte Carlo steps. Acceptance ratio varies from 45 to 55 %.

Note that this time, even using a lower number of Monte Carlo steps, we have an amazingly precise estimate for our energy, comparable with the one obtained in [Pedersen Lohne, et al.] using Diffusion Monte Carlo.

2.2.2   Importance sampling

Let's now try to implement importance sampling, as explained in Section 1.3.2. In order to show the differences with the brute force algorithm, we will repeat the calculations with exactly the same settings used in Figure 8, this time with importance sampling active. We will also try to do the calculations for different time-steps $\Delta t = 0.001, 0.01, 0.1$.

The results are shown in Figures 9, 10 and 11.

As you see, there is a huge difference between the result obtained with $\Delta t = 0.001$ and the one obtained with $\Delta t = 0.1$. This is because, if the time step is too small and we don't have enough Monte Carlo steps, our walker cannot span completely the integration space. It will remain near its starting point and all the points reached will be inside the integration space, close to each other. In fact, the acceptance ratio for the case $\Delta t = 0.001$ and $\Delta t = 0.01$ is *practically* 100 %. The third case, for $\Delta t = 0.1$, has a little lower acceptance ratio, that is about 99.79 %. This fact guarantees that our walker is reaching all the borders of our integration space, and
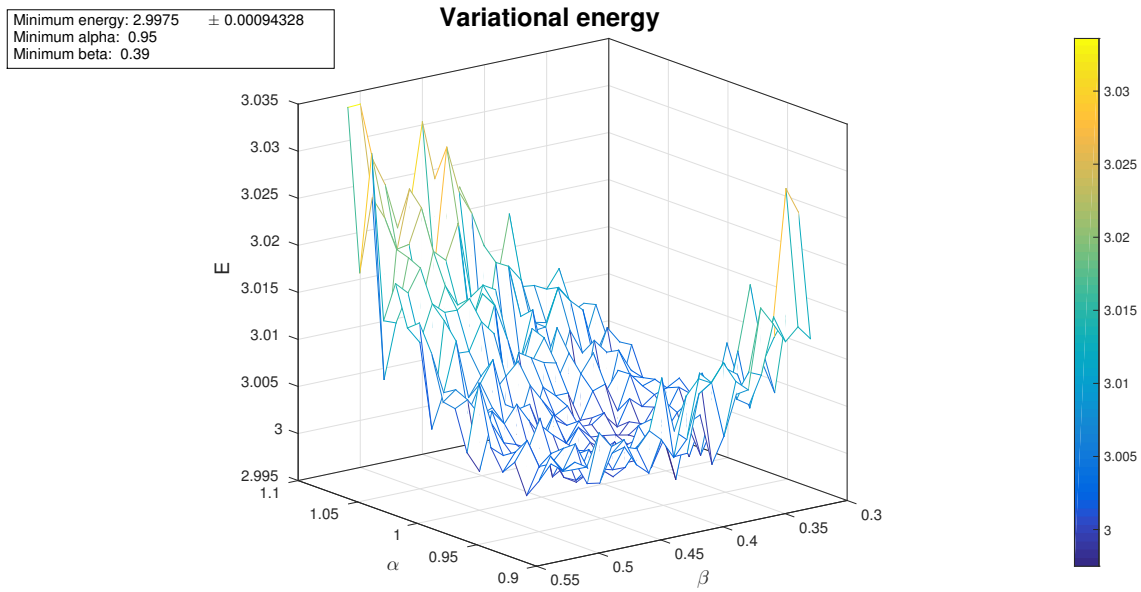
Figure 9: The variational energy versus the variational parameters α and β. The settings used are: importance sampling with $\Delta t = 0.001$, Jastrow factor, parallelization (8 threads), 20 variations of α and β with step 0.01, 2×10^5 Monte Carlo steps. Acceptance ratio is about 99.999 %.
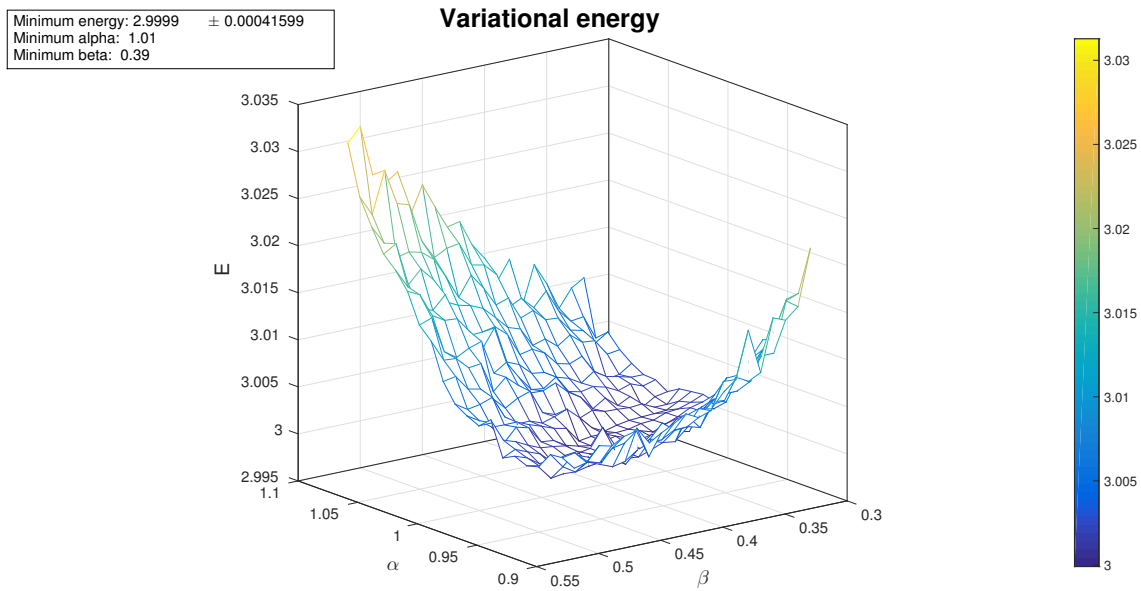

Figure 10: The variational energy versus the variational parameters α and β. The settings used are: importance sampling with $\Delta t = 0.01$, Jastrow factor, parallelization (8 threads), 20 variations of α and β with step 0.01, 2×10^5 Monte Carlo steps. Acceptance ratio is about 99.999 %.
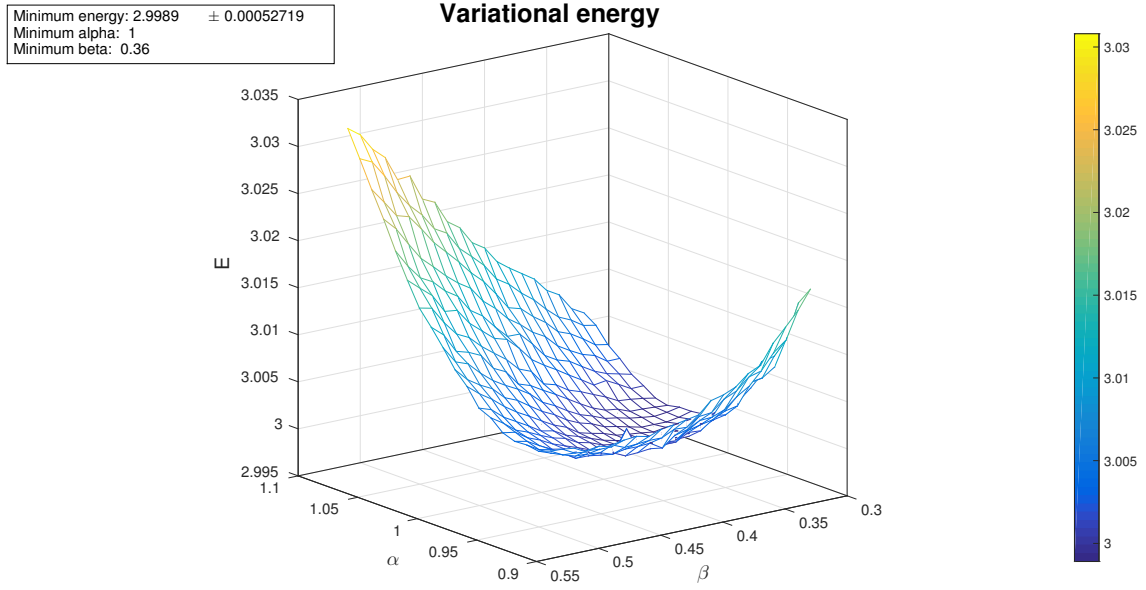
Figure 11: The variational energy versus the variational parameters α and β. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 20 variations of α and β with step 0.01, 2×10^5 Monte Carlo steps. Acceptance ratio is about 99.793 %.

in fact it also reaches some points outside it. In other words, for the given number of Monte Carlo iterations, we have a large enough time-step to cover all the space. If one wants to use a lower time-step to have better precision, he has to increase the number of Monte Carlo cycles; as an example, for $\Delta t = 0.001$, one has to use at least 1×10^9 Monte Carlo loops.
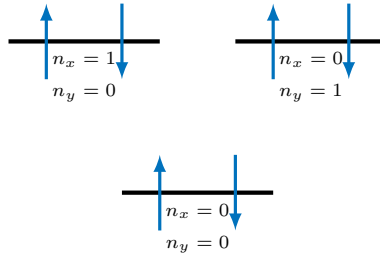
3   The 6-electrons system

Figure 12: The 6-electrons system configuration.

The formulae discussed in the previous sections can be generalized for a N electrons case. The trial wave function can be written as

$$\psi_T(\vec{r}_1 \dots \vec{r}_N) = A \operatorname{Det}(S) \prod_{i<j}^{N} e^{\frac{r_{ij}}{1+\beta r_{ij}}}, \tag{64}$$

where $\operatorname{Det}(S)$ is the Slater determinant

$$\operatorname{Det}(S) = \begin{vmatrix} \phi_1(\vec{r}_1) & \phi_2(\vec{r}_1) & \dots & \phi_N(\vec{r}_1) \\ \phi_1(\vec{r}_2) & \phi_2(\vec{r}_2) & \dots & \phi_N(\vec{r}_1) \\ \vdots & & & \vdots \\ \phi_1(\vec{r}_N) & \phi_2(\vec{r}_N) & \dots & \phi_N(\vec{r}_N) \end{vmatrix} \tag{65}$$

As we did in Section 1.2, this determinant can be further simplified if we label our electrons in a smart way, as shown in [Hjorth-Jensen].

Let's consider a spin-independent quantum mechanical operator $\hat{O}(\vec{r})$ acting on our trial wave function. The latter depends on $\vec{x} = (\vec{r}, \sigma)$ where $\vec{r}$ includes the space coordinates and σ the spin coordinates. We have that:

$$\left\langle \hat{O} \right\rangle = \frac{\left\langle \psi_T(\vec{x}) \left| \hat{O}(\vec{r}) \right| \psi_T(\vec{x}) \right\rangle}{\left\langle \psi_T(\vec{x}) \, | \, \psi_T(\vec{x}) \right\rangle} \tag{66}$$

Now we can replace the total antisymmetric wave function with one with permuted arguments and we can arbitrarily choose that the first $N/2$ arguments are spin up and the other half are spin down. We get:

$$\psi_T(\vec{x_1}, \ldots, \vec{x_N}) \longrightarrow \psi_t(\vec{x_{i1}}, \ldots, \vec{x_{iN}}) \tag{67}$$
$$= \psi_T(\{\vec{r_{i1}}, \uparrow\}, \ldots, \{\vec{r_{iN/2}}, \uparrow\}, \{\vec{r_{i1}}, \downarrow\}, \ldots, \{\vec{r_{iN/2}}, \downarrow\}) \tag{68}$$
$$= \psi_T(\{\vec{r_1}, \uparrow\}, \ldots, \{\vec{r_{N/2}}, \uparrow\}, \{\vec{r_{N/2+1}}, \downarrow\}, \ldots, \{\vec{r_N}, \downarrow\}) \tag{69}$$

The operator $\hat{O}$ is symmetric with the respect to the exchange of labels in a pair of particles, so:

$$\left\langle \hat{O} \right\rangle = \frac{\left\langle \psi_T(\vec{r}) \left| \hat{O}(\vec{r}) \right| \psi_T(\vec{r}) \right\rangle}{\left\langle \psi_T(\vec{r}) \, | \, \psi_T(\vec{r}) \right\rangle} \tag{70}$$

The wave function is now antisymmetric with respect to exchange of spatial coordinates of pairs of spin-up or spin-down electrons. Therefore, for spin-independent Hamiltonians, the Slater determinant can be splitted in a product of two Slater determinants, one for the single particle orbitals with spin up and the other with single particle orbitals with spin down. We have then:

$$\psi_S = S^\uparrow S^\downarrow \tag{71}$$

where

$$S^\uparrow = \begin{vmatrix} \phi_1(\vec{r}_1) & \phi_2(\vec{r}_1) & \ldots & \phi_{N/2}(\vec{r}_1) \\ \phi_1(\vec{r}_2) & \phi_2(\vec{r}_2) & \ldots & \phi_{N/2}(\vec{r}_1) \\ \vdots & & & \vdots \\ \phi_1(\vec{r}_{N/2}) & \phi_2(\vec{r}_{N/2}) & \ldots & \phi_{N/2}(\vec{r}_{N/2}) \end{vmatrix} \tag{72}$$

and $S^\downarrow$ is obviously defined in a similar way. The ϕ_i's functions are just as the one defined in equation (47).

So, our final trial wave-function is just

$$\psi_T = |S^\uparrow||S^\downarrow|J \tag{73}$$

being J the Padé-Jastrow factor.

The C++ implementation of ψ_S is therefore straightforward and it is shown below.

```cpp
double Orbitals::SlaterD(mat& r) {
    mat D_up(n_half,n_half), D_down(n_half,n_half);

    for (int i = 0; i < n_half; i++) {
        this->set_qnum_indie_terms(r, i);
        for (int j = 0; j < n_half; j++) {
            D_up(i,j) = this->phi(r, i, j);
        }
    }

    for (int i = 0; i < n_half; i++) {
        this->set_qnum_indie_terms(r, i+n_half);
        for (int j = 0; j < n_half; j++) {
            D_down(i,j) = this->phi(r, i+n_half, j);
        }
    }

    double slater = det(D_up)*det(D_down);

    D_up.reset();
    D_down.reset();

    return slater;
}
```
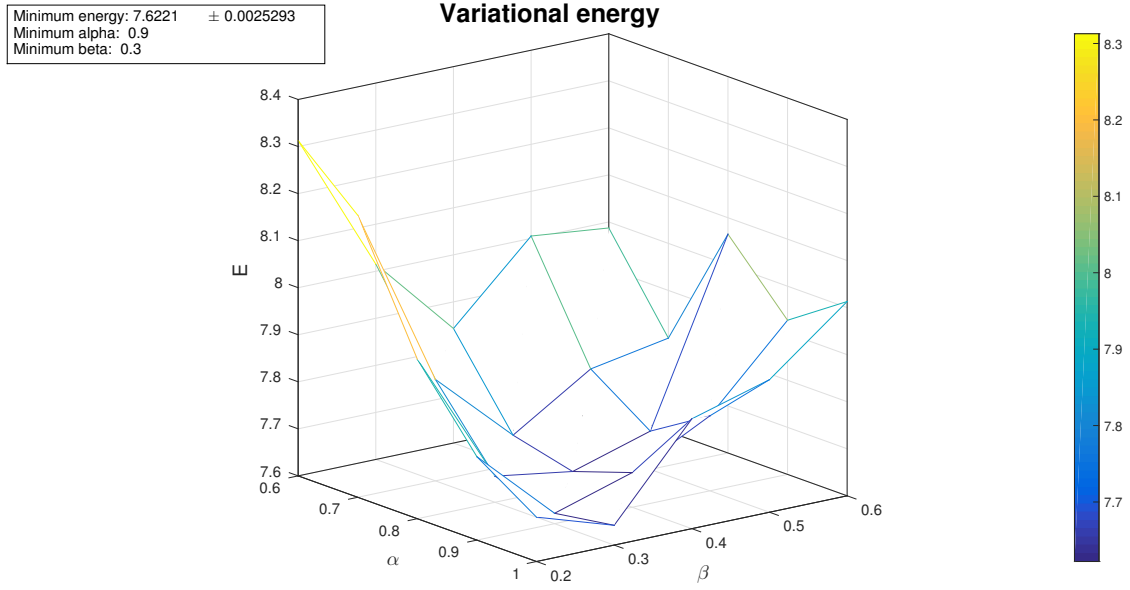
Figure 14: The variational energy versus the variational parameters α and β. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 5 variations of α and β with step 0.1, 2×10^5 Monte Carlo steps. $\omega = 0.28$.

3.1 $\omega = 0.28$

We start by studying our system for $\omega = 0.28$. A preliminary and very rough result is shown in Figure 13.
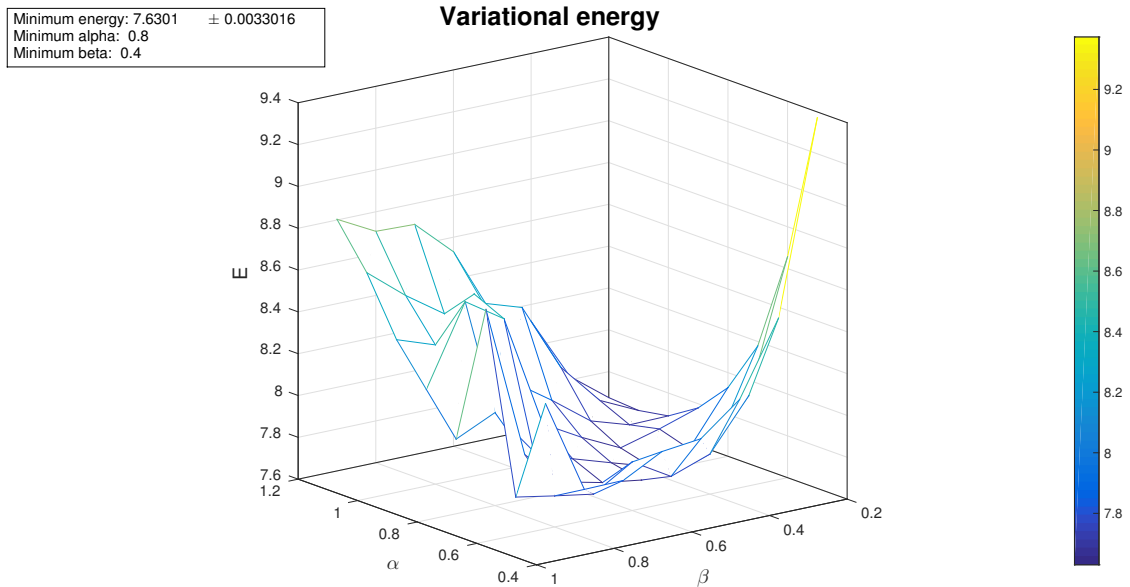

Figure 13: The variational energy versus the variational parameters α and β. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 8 variations of α and β with step 0.1, 1×10^5 Monte Carlo steps. $\omega = 0.28$.

Restricting to the significant region, we get the result in Figure 14.

You see that the plot is no more nice like the 2-electrons case, mainly because we didn't used enough Monte Carlo steps. The problem is that now the program has to calculate 3×3 determinants every time, so it's much slower than the previous case. To refine the result, we just doubled the number of Monte Carlo loops and

restricted a little bit the variational parameters grid; taking a finer grid with more Monte Carlo cycles would have required ages, literally.

Anyway, still with this configuration, we get a result that is really near the one calculated by DMC in [Pedersen Lohne, et al.], that turns out to be 7.6001(1) a.u..

A nice way to test the Slater determinant code is to turn off the electron-electron repulsion. In that case, ψ_S is the *exact* solution (for $\alpha = 1$). We already showed in Section 2.1 that the ground-state energy for the single electron is $E_s = \hbar\omega(n_x + n_y + 1)$. Since – in the 6 electrons case – we have 2 electrons with $n_x + n_y = 0$ and 4 electrons with $n_x + n_y = 1$ (See Figure 12), the total energy for the ground state is

$$E_{\text{gs}} = \hbar\omega\left[2(0+1) + 4(1+1)\right] = 10\omega. \tag{74}$$

Turning off the repulsion results in a minimum $E_{\text{gs}} = 2.799\,999\,999\,999\,9(2)$ a.u. for $\alpha = 1$, that confirms the prediction.

3.2 $\omega = 0.5$

We repeat the same calculations with $\omega = 0.5$, obtaining the results in Figure 15.
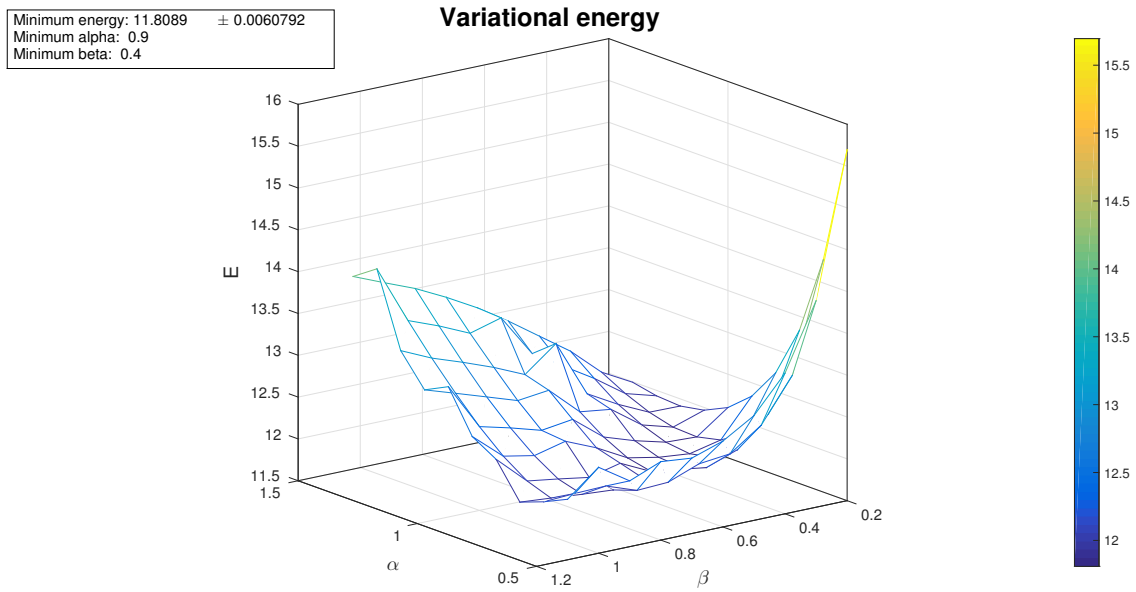

Figure 15: The variational energy versus the variational parameters α and β. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 10 variations of α and β with step 0.1, 2×10^5 Monte Carlo steps. $\omega = 0.50$.

Again, our result is pretty consistent with the one calculated by DMC in [Pedersen Lohne, et al.], that is 11.7888(2) a.u..

3.3 $\omega = 1$

Just with the same parameters used in Figure 15 but changing the frequency to $\omega = 1.00$, we have the results in Figure 16.
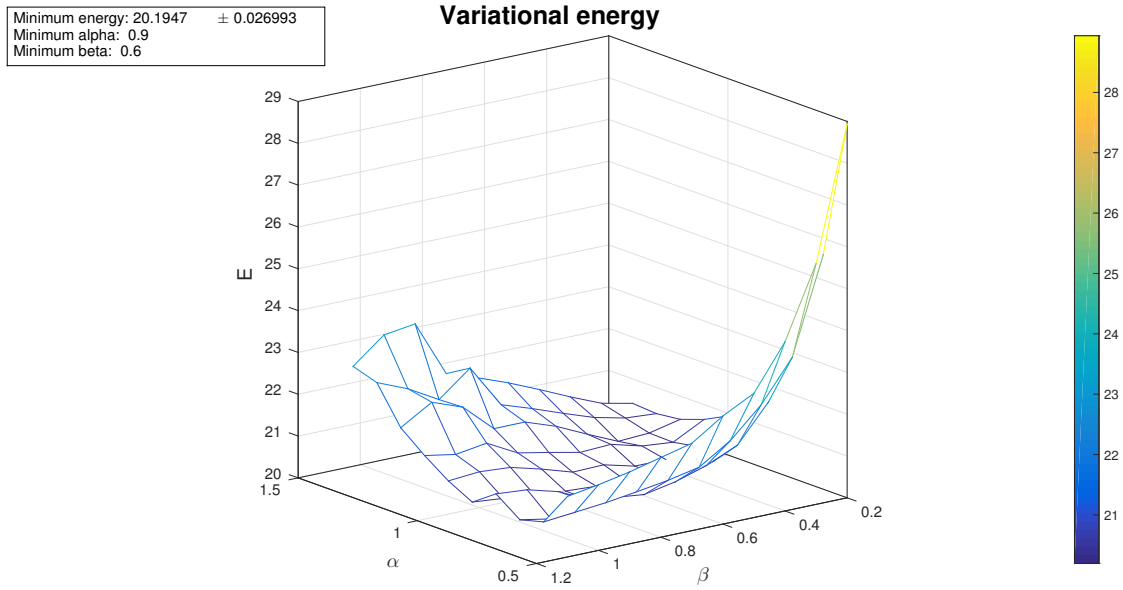
Figure 16: The variational energy versus the variational parameters α and β. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 10 variations of α and β with step 0.1, 2×10^5 Monte Carlo steps. $\omega = 1.00$.

Once again, our result is in very good accordance with the one calculated by DMC in [Pedersen Lohne, et al.], that is $20.1597(2)$ a.u..

4   The virial theorem

The virial theorem states that the expectation value of the total kinetic energy $\langle T \rangle$ is proportional to the expectation value of the total potential energy $\langle V \rangle$. For a pure harmonic oscillator, this theorem is simply the theorem of the *equipartition theorem*:

$$\langle T \rangle = \langle V \rangle. \tag{75}$$

This is quite easy to prove. Let's introduce the ladder operators

$$\hat{a} = \frac{1}{\sqrt{2\hbar\omega m}} \left(m\omega\hat{x} + i\hat{p} \right) \qquad \text{and} \qquad \hat{a}^\dagger = \frac{1}{\sqrt{2\hbar\omega m}} \left(m\omega\hat{x} - i\hat{p} \right) \tag{76}$$

It can be shown that these operators act on the harmonic oscillator eigenstates such that

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle \quad \text{and} \quad \hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \tag{77}$$

and we also have

$$N |n\rangle \doteq \hat{a}^\dagger \hat{a} |n\rangle = n |n\rangle. \tag{78}$$

Using these properties, the calculation of the expected value of the potential energy is straightforward:

$$\left\langle n \left| \frac{1}{2} m \omega^2 \hat{x}^2 \right| n \right\rangle = \frac{1}{2} m \omega^2 \left\langle n \left| \hat{x}^2 \right| n \right\rangle \tag{79}$$

$$= \frac{1}{2} \cancel{m} \cancel{\omega^2} \frac{\hbar}{2 \cancel{m\omega}} \left\langle n \left| \left(\hat{a}^\dagger + \hat{a} \right) \left(\hat{a} + \hat{a}^\dagger \right) \right| n \right\rangle \tag{80}$$

$$= \frac{\hbar \omega}{4} \left(\left\langle n \left| \hat{a}^\dagger \hat{a}^\dagger \right| n \right\rangle + \left\langle n \left| \hat{a} \hat{a}^\dagger \right| n \right\rangle + \left\langle n \left| \hat{a}^\dagger \hat{a} \right| n \right\rangle + \left\langle n \left| \hat{a} \hat{a} \right| n \right\rangle \right) \tag{81}$$

$$= \frac{\hbar \omega}{4} \left(\sqrt{n+2} \sqrt{n+1} \cancel{\langle n | n + 2 \rangle}^{\,0} \right. \tag{82}$$

$$\left. + \left\langle n \left| \hat{N} + 1 \right| n \right\rangle + \left\langle n \left| \hat{N} \right| n \right\rangle + \sqrt{n} \sqrt{n-1} \cancel{\langle n | n - 2 \rangle}^{\,0} \right) \tag{83}$$

$$= \frac{\hbar \omega}{4} \left((n+1) + n \right) = \frac{\hbar \omega}{4} (2n+1) \tag{84}$$

$$= \frac{\hbar \omega}{2} \left(n + \frac{1}{2} \right) \tag{85}$$

If we perform the same calculation for $\left\langle n \left| \frac{\hat{p}^2}{2m} \right| n \right\rangle$ (the expected value of the kinetic energy) we will obtain the result, yielding

$$\langle T \rangle = \langle V \rangle \tag{86}$$

and

$$\langle T \rangle + \langle V \rangle = \hbar \omega \left(n + \frac{1}{2} \right), \tag{87}$$

that is the total energy of the harmonic oscillator, as expected.

Let's now graph the ratio $\langle T \rangle / \langle V \rangle$ for different values of ω.

4.1 The 2-electrons system

We firstly computed $\langle T \rangle$ and $\langle V \rangle$ *without* the electron-electron repulsion, to check our code. We chose $\omega = 0.01, 0.28, 0.50, 0.70, 1.00$. The result – completely expected – is shown in Figure 17.
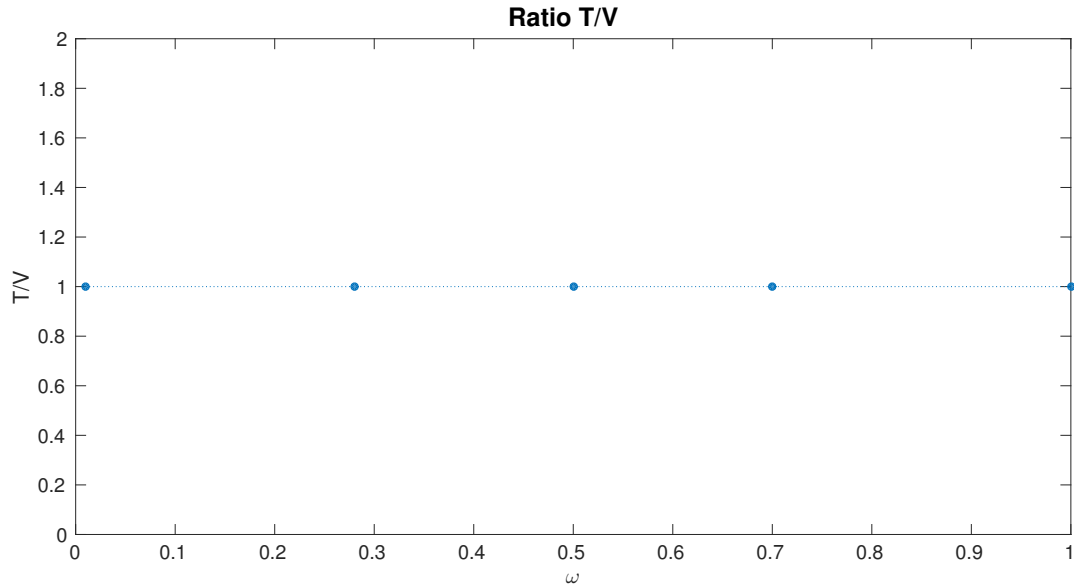

Figure 17: The ratio $\langle T \rangle / \langle V \rangle$ for different values of ω and no electron-electron repulsion. The settings used are: importance sampling with $\Delta t = 0.1$, parallelization (8 threads), 1×10^7 Monte Carlo steps.

After that, we added the electron-electron repulsion and we did again the calculations for the same ω values. The result, shown in Figure 18, is quite different.
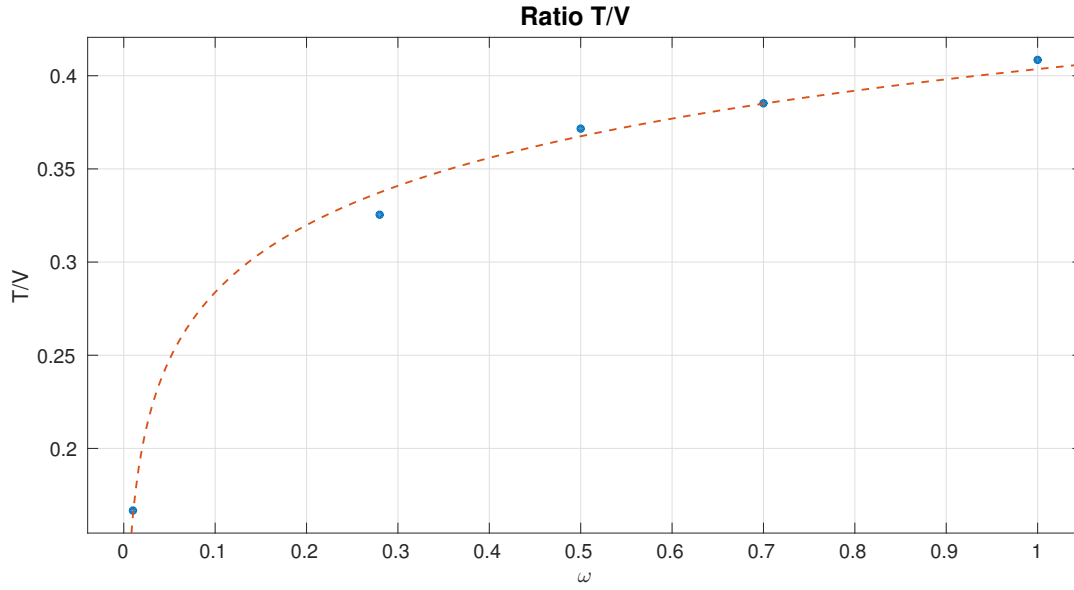
Figure 18: The ratio $\langle T \rangle / \langle V \rangle$ for different values of ω with electron-electron repulsion. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 1×10^7 Monte Carlo steps. The fit is logarithmic.

The dashed line is a fit of the type

$$\frac{\langle T \rangle}{\langle V \rangle} = b \log(c\omega), \tag{88}$$

which gives $b = 0.052(7)$ and $c = 2340(2021)$. We can't really say if the behaviour of our data is logarithmic without a theoretical basis and with just 5 points; the purpose of the dashed line in Figure 18 is just to show that our data is *no way* linear. This is due to the presence of the repulsive force, that can be dominant or not.

In a very intuitive way, we can at least say that what we obtain is coherent with the physics of the system. For $\omega \gg 1$ the oscillator potential is dominant and the repulsion can be considered as a small perturbation, which means that we have more kinetic energy than potential energy ($T > V$). For $\omega \ll 1$, we almost have only a repulsion potential plus a small harmonic perturbation; in other words, $T < V$.

If one really wants to find an analytical relation for the ratio $\langle T \rangle / \langle V \rangle$ of this system, he should compute the two terms separately in a way similar to the one shown in Section 4 and then take the ratio of them. Note that this is not always possible, since the necessary integrals could not be solved analytically.
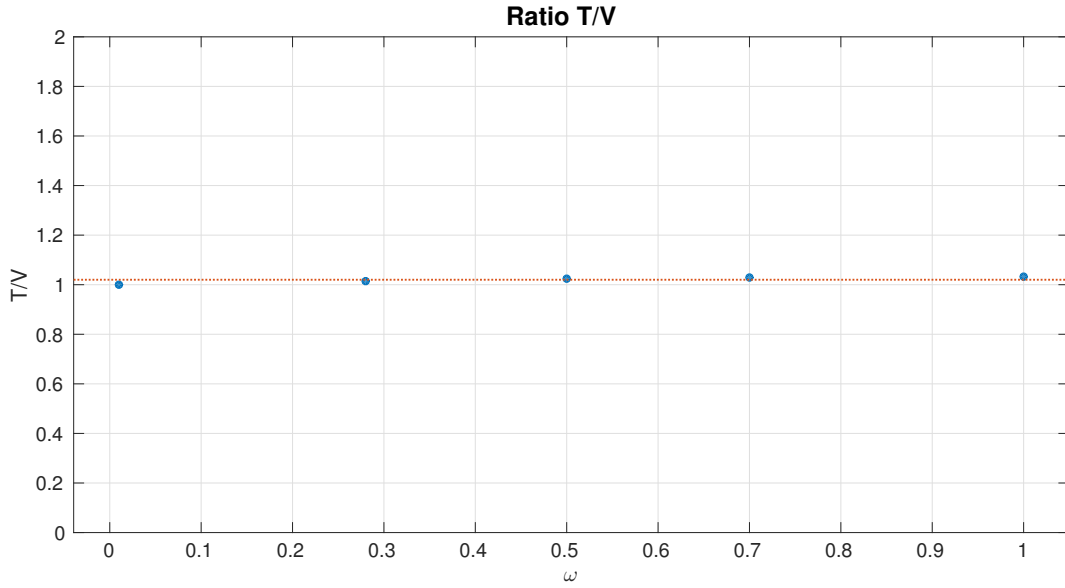
4.2 The 6-electrons system

Figure 19: The ratio $\langle T \rangle / \langle V \rangle$ for different values of ω and no electron-electron repulsion. The settings used are: importance sampling with $\Delta t = 0.1$, parallelization (8 threads), 1×10^6 Monte Carlo steps. A linear fit gives a ratio of 1.02(2), that is still compatible with the expected value 1.
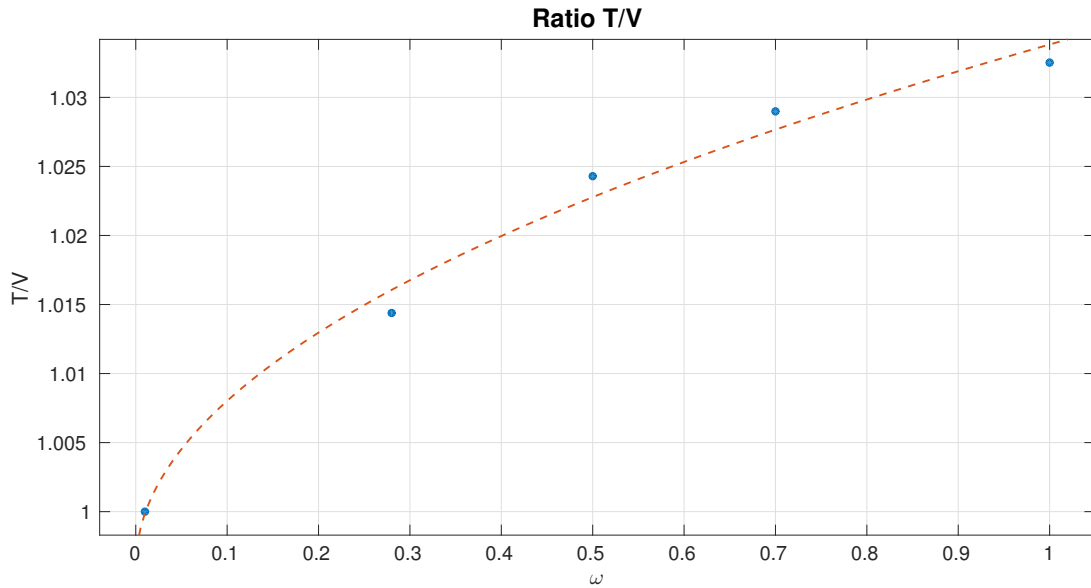

Figure 20: The ratio $\langle T \rangle / \langle V \rangle$ for different values of ω with electron-electron repulsion. The settings used are: importance sampling with $\Delta t = 0.1$, Jastrow factor, parallelization (8 threads), 1×10^6 Monte Carlo steps. The fit is a square root.

We did again the same calculations (with the same ω values) for 6 electrons. This time, we don't obtain a perfect 1 ratio for the non repulsive case (Figure 19); a linear fit gives 1.02(2), which is anyway still compatible with the expected value 1. A possible explanation for this behaviour is that we are not using enough Monte Carlo steps; they are ten times less than the previous case, when we are dealing with a number of electrons that is three times more the 2-electrons case.

The discussion of what we obtain adding the electron-electron repulsion, shown in Figure 20, is the same as

the 2-electrons case in the previous section. The fit (that *is not* particularly indicative) is of the type

$$\frac{\langle T \rangle}{\langle V \rangle} = b\sqrt{c\omega} + d, \tag{89}$$

which gives $b = 0.06(1)$, $c = 0.416(1)$ and $d = 1.00(1)$.

5 Analytical derivatives

The most time-consuming part of the Metropolis-Hastings algorithm is the computation of the acceptance ratio, the kinetic energy and the quantum force, that involve the calculation of several first and second derivatives. A way to – possibly – improve our program is to use analytical derivatives. Note that this is possible in very few cases – like this one, so we will take advantage of it.

Let's remember that the wave-function is

$$\psi_T = |\vec{S}^\uparrow||\vec{S}^\downarrow|J, \tag{90}$$

where $|\vec{S}^\uparrow|$ is the Slater determinant relative to the spin-up particles, $|\vec{S}^\downarrow|$ is the one relative to the spin-down particles and J is the Jastrow factor.

For the quantum force $\vec{F}_i$ relative to particle i we have that

$$\vec{F}_i = 2\frac{1}{\psi_T}\nabla_i \Psi_T \tag{91}$$

$$= 2\frac{\nabla_i(|\vec{S}^\uparrow||\vec{S}^\downarrow|J)}{|\vec{S}^\uparrow||\vec{S}^\downarrow|J} \tag{92}$$

$$= 2\left(\frac{\nabla_i|\vec{S}^\uparrow|}{|\vec{S}^\uparrow|} + \frac{\nabla_i|\vec{S}^\downarrow|}{|\vec{S}^\downarrow|} + \frac{\nabla_i J}{J}\right). \tag{93}$$

Since particle i has either spin up or spin down, one of the two terms with the Slater determinant will vanish, because it's not dependant on that particle. So, labelling with α the spin of particle i, we have that

$$\vec{F}_i = 2\left(\frac{\nabla_i|\vec{S}^\alpha|}{|\vec{S}^\alpha|} + \frac{\nabla_i J}{J}\right) \tag{94}$$

Doing the same with the local energy, we end up with

$$E_L = \frac{1}{2}\sum_i \frac{1}{\psi_T}\nabla_i^2 \psi_T + \sum_i V_i \tag{95}$$

Then,

$$\frac{1}{\psi_T}\nabla_i^2 \psi_T = \frac{\nabla_i^2|\vec{S}^\alpha|}{|\vec{S}^\alpha|} + \frac{\nabla_i^2 J}{J} + 2\frac{\nabla_i|\vec{S}^\alpha|}{|\vec{S}^\alpha|}\frac{\nabla_i J}{J}, \tag{96}$$

where the components relative to the spin opposite to the one of the particle taken into account vanish for the reason explained before.

According to [Høgberget], the four terms that appear in (96) are

$$\frac{\nabla_i J}{J} = \sum_{k\neq i=1}^N \frac{a_{ik}}{r_{ik}}\frac{\vec{r}_i - \vec{r}_k}{(1 + \beta\,r_{ik})^2} \tag{97}$$

$$\frac{\nabla_i^2 J}{J} = \left|\frac{\nabla_i J}{J}\right|^2 + \sum_{k\neq i=1}^N a_{ik}\frac{(d-3)(\beta\,r_{ik}+1)+2}{r_{ik}(\beta\,r_{ik}+1)^3} \tag{98}$$

$$\frac{\nabla_i|\vec{S}|}{|\vec{S}|} = \sum_k \nabla_i\phi_k(\vec{r}_i^{\text{new}})(\vec{S}_{ki}^{\text{new}})^{-1} \tag{99}$$

$$\frac{\nabla_i^2|\vec{S}|}{|\vec{S}|} = \sum_k \nabla_i^2\phi_k(\vec{r}_i^{\text{new}})(\vec{S}_{ki}^{\text{new}})^{-1} \tag{100}$$

with k that spans the Slater matrix relative to the ith particle. The quantities $\nabla_i\phi_k$ and $\nabla_i^2\phi_k$ are also tabulated in [Høgberget]. It's also possible to optimize the determinant of the inverse of the Slater matrix, but we opted here for a numerical computation of the inverse matrix using Armadillo. If one does that, it's also possible to optimize the acceptance ratio to have a even faster code.

5.1 Execution times

Let's see if we have a real gain in speed for the 2-electrons and the 6-electrons cases. Note that we have to run the program with a single thread to avoid extra waiting time on the CPU due to the line

```
        #pragma omp barrier                                                                    1
```
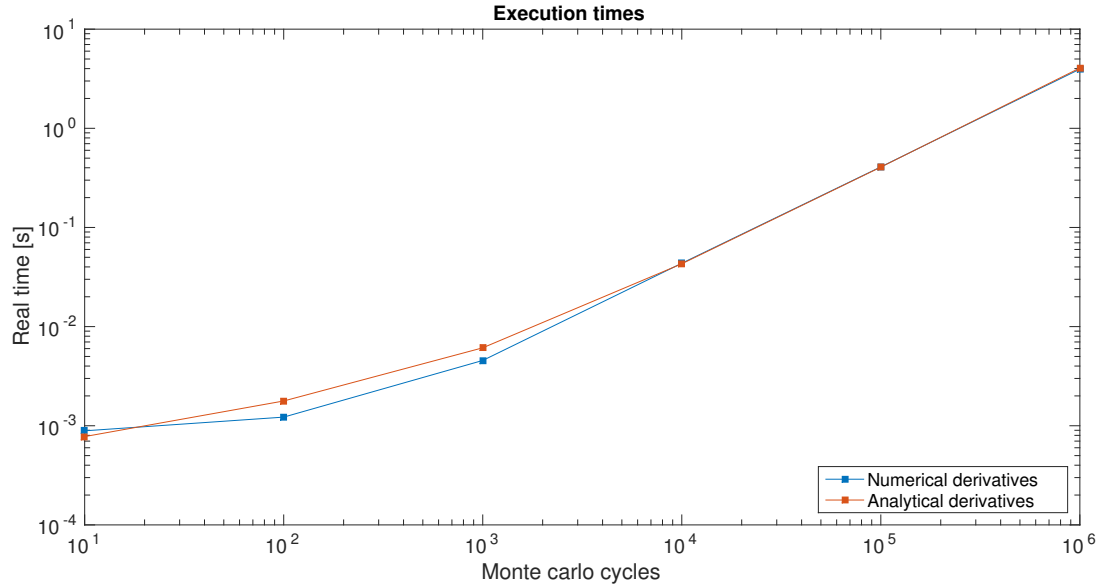
in the parallel cycles. Our benchmarks are shown in Figures 21 and 22.

Figure 21: Execution times for 2-electrons and a single pair of variational parameters (α, β). The GNU/Linux system tool `time` was used to take the measurements.
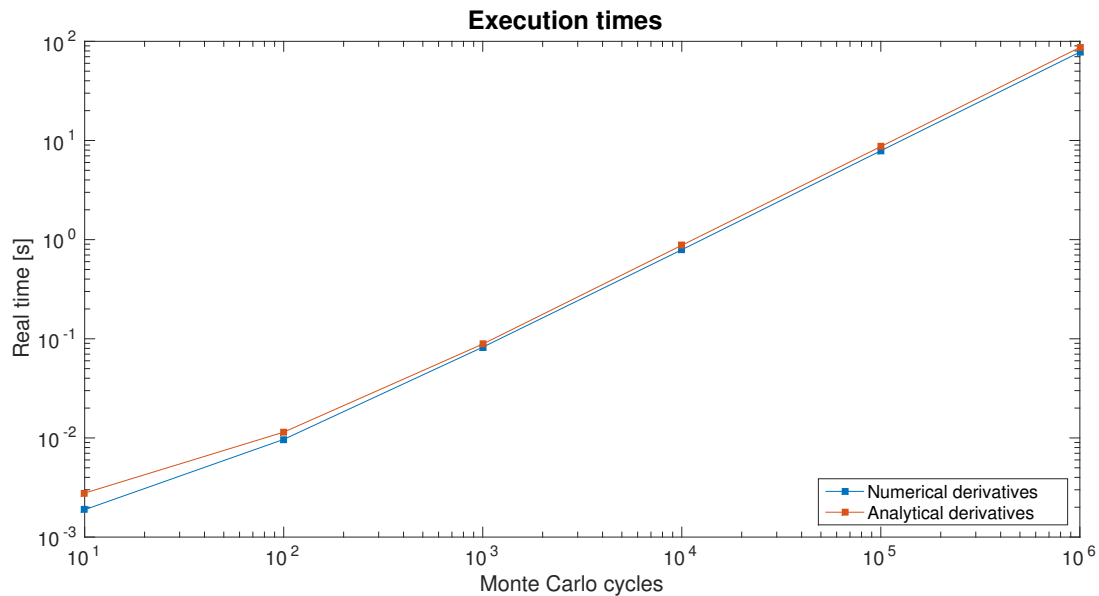

Figure 22: Execution times for 6-electrons and a single pair of variational parameters (α, β). The GNU/Linux system tool `time` was used to take the measurements.

As you see, we don't have great improvements; on the contrary, we obtain a code that is even a little bit slower than the numerical one for the 6-electrons case! However, there are some things that have to be pointed

out. First of all we decided to perform a numerical inversion of a matrix, that is really time-consuming (it's a $\mathcal{O}(n^3)$ operation). Secondly there is no optimization of the acceptance ratio, that can be further simplified – as shown again in [Høgberget]. Lastly, the benefits of numerical derivatives could not appear until one reaches 12 or even 20 particles; this is because the computational complexity of a determinant of a matrix is $\mathcal{O}(n^3)$, that for 3×3 matrices (as the 6-electrons case) is still acceptable. The struggle in calculating big determinants all the time is more evident for higher numbers of particles; for the 12-particles case the Slater determinant is 8 times slower than the 6-particles case, while for the 20-particles case is approximately 37 times slower.

Another thing that has to be noted is that – with the analytical derivatives – we obtain slightly lower values for the energy. Deactivating the electron-electron repulsion the discrepancy still persists, meaning that the reason of this fact has to be found in the Slater determinant derivatives implementation. A possible explanation is that we are (numerically) performing the inverse of a matrix that has very small elements, and this could cause a loss of precision.

6　Conclusions and perspectives

In this project we really showed the power of the VMC method, that even if it doesn't give automatically the ground state energy of a system can be used to approximate it – to excess – with a considerable degree of precision. If we have no clues about the properties of the system we can use a brute force sampling, trying to manually set a length-step that gives about 50%. But if we are smarter, we can use the information we have about the expected result to implement a more efficient sampling.

This program has been implemented in a highly object-oriented way, following the suggestions contained in [Høgberget], and can also fully handle the 12-electrons case (look at the implementation of the orbitals in the code). However, it is still *extremely slow* – practically unusable. So, I'm planning to improve this program using MPI and GPU parallelization, that would give me some extra speed.

I'm also planning to extend this program to handle a custom number of variational parameters, and to improve the trial wave-function with some extra terms as suggested by professor Francesco Pederiva[6] and/or using the Hartree-Fock method. This would require to abandon the analytical derivatives route and concentrate the efforts in the algorithm improvements in terms of efficiency.

Finally, for me this project has been a springboard into the quantum computational physics world; I enjoyed so much in doing this simulation that I've decided to goo deeper into this subject and start to explore of its full potential in advanced courses at UiO and NMBU.

---

[6]University of Trento, Department of Physics.

References

[Hjorth-Jensen] Hjorth-Jensen, Morten. (2014, Aug). *Computational Physics - Lecture Notes Fall 2014*. Oslo: University of Oslo. Retrieved from: `http://www.uio.no/studier/emner/matnat/fys/FYS3150/h14/undervisningsmateriale/Lecture%20Notes/lecture2014.pdf`

[Taut] Taut, M. (1993, Nov). Two electrons in an external oscillator potential: Particular analytic solutions of a Coulomb correlation problem. *Phys. Rev. A, 48*(5), 3561–3566. `doi:10.1103/PhysRevA.48.3561`, `http://link.aps.org/doi/10.1103/PhysRevA.48.3561`

[Pedersen Lohne, et al.] Pedersen Lohne, M. and Hagen, G. and Hjorth-Jensen, M. and Kvaal, S. and Pederiva, F. (2011, Sep). Ab initio computation of the energies of circular quantum dots. *Phys. Rev. B, 84*(11), 115302. `10.1103/PhysRevB.84.115302`, `http://link.aps.org/doi/10.1103/PhysRevB.84.115302`

[Høgberget] Høgberget, Jørgen. (2013). *Quantum Monte-Carlo Studies of Generalized Many-body Systems*. Oslo: University of Oslo. `http://hdl.handle.net/10852/37167`. Retrieved from: `https://duo.uio.no/bitstream/10852/37167/1/jorgenh_thesis.pdf`

[Griffiths] Griffiths, David Jeffery. (Second Ed.). (2005). *Introduction to Quantum Mechanics*. Pearson Education, Inc. `ISBN 978-81-7758-230-7`.