

# **FYS4411 - Computational Physics II: Quantum Mechanical Systems**

## **Project 2 - Variational Monte Carlo Methods**

**Alfred Alocias Mariadason**

<https://www.github.com/Oo1Insane1oO/FYS4411>

April 30, 2017

## Abstract

## I INTRODUCTION

Using the Variational Monte Carlo, this project aims to find and analyze quantities such as the ground state energy and single-particle densities of quantum dots for so-called closed shell systems.

We use the usual approach by estimating expectation value of the ground state energy with the variational principle and minimizing. The algorithm used for the Monte Carlo method is the well known Metropolis algorithm.

The reason for using a Monte Carlo method for minimizing the trial ground state energy is because the expectation value would in general be a multi-dimensional integral depending on the number of particles and number of parameters involved in the total wave function. Such an integral is not adequately solved by traditional methods(i.e Gaussian-quadrature).

The desired result is that the Metropolis algorithm with importance sampling yields a better result both from a computational point of view. That is it finds a good estimate for the ground state energy efficiently without wasting too much time on the configuration space. The wave function only has small values in this large space meaning a homogeneous distribution of calculation points would yield a poor result, a non-homogeneous approach(such as with the Metropolis algorithm) would then, hopefully, give a better result.

## II THEORY

;;INSERT DESCRIPTION;;

### II.A HERMITE POLYNOMIALS

Hermite polynomials  $H(x)$  are solutions to the differential equation

$$\frac{d^2 H}{dx^2} - 2x \frac{dH}{dx} + (\lambda - 1) H = 0 \quad (1)$$

The polynomials fulfill the orthogonality relation

$$\int_{-\infty}^{\infty} e^{-x^2} H_n^2 dx = 2^n n! \sqrt{\pi} \quad (2)$$

with the recurrence relation

$$H_{n+1} = H_n - 2nH_{n-1} \quad (3)$$

### II.B HARMONIC OSCILLATOR

#### II.B.1 Cartesian Coordinates

The harmonic oscillator system in 2 dimensions and in natural units is given by the following Hamiltonian

$$\hat{H}_0 = \frac{1}{2} \sum_{i=1}^N (-\nabla_i^2 + \omega^2 r_i^2) \quad (4)$$

The wave functions in this case is then:

$$\phi_{n_x, n_y}(x, y) = A H_{n_x}(\sqrt{\omega}x) H_{n_y}(\sqrt{\omega}y) \exp\left(-\frac{\omega}{2}(x^2 + y^2)\right) \quad (5)$$

where  $H_n$  is a Hermite polynomial of order  $n$  and  $A$  is a normalization constant. The quantum numbers  $n_x$  and  $n_y$  go as  $n_x, n_y = 0, 1, 2, \dots$ . While  $\omega$  is the oscillator frequency.

The energies is

$$E = \hbar\omega (n_x + n_y + 1) \quad (6)$$

### II.C FOURIER TRANSFORMATION

Given an integrable function  $f : \mathbb{R} \rightarrow \mathbb{C}$  the Fourier transform to variable  $x$  is defined to be

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{ixs} dx \quad (7)$$

with  $x, s \in \mathbb{R}$ .

The inverse transformation is given as

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{-isx} ds \quad (8)$$

## II.D CONJUGATE GRADIENT METHOD

Suppose we have a linear system defined as

$$A\vec{x} = \vec{b} \quad (9)$$

where  $A$  is a  $n \times n$  real, symmetric and positive definite matrix while  $\vec{x}$  and  $\vec{b}$  are a non-zero real vectors.

Define now the conjugate directions of two vectors  $\vec{x}_i, \vec{x}_j : \mathbb{R} \rightarrow \mathbb{R}$  giving the constraint

$$\vec{x}_i^T A \vec{x}_j = 0 \quad (10)$$

This means that in the iterative process of finding  $\vec{x}$  we perform searches within the conjugate directions of  $\vec{x}$ .

Since the conjugate constraint (equation 10) is defined by an inner product within the space  $A$  is defined, the two vectors  $\vec{x}_i$  and  $\vec{x}_j$  are orthogonal. If we now define a span  $P$

$$P = \{\vec{p}_1, \dots, \vec{p}_n\} \quad (11)$$

consisting of  $n$  mutually orthogonal conjugate directions  $\vec{p}_i$ , we have a basis for  $\mathbb{R}^n$ . Expanding the solution  $x_{i+1}$  to equation 9 in the mentioned basis

$$\vec{x} = \sum_{i=1}^n C_i \vec{p}_i \quad (12)$$

the linear system can be rewritten as

$$A\vec{x} = \sum_{i=1}^n C_i A\vec{p}_i = \vec{b} \quad (13)$$

Giving the inner product

$$\vec{p}_k^T A\vec{x} = \sum_{i=1}^n C_i \vec{p}_k^T A\vec{p}_i \quad (14)$$

and we can define the coefficients  $C_k$  as

$$C_k = \frac{\vec{p}_k^T \vec{b}}{\vec{p}_k^T A\vec{p}_k} \quad (15)$$

The problem at hand is then to choose a sequence of  $n$  conjugate directions  $P$  and compute the coefficients  $C_k$ .

### II.D.1 Iterative Method

The conjugate gradient method can in a similar manner be used on an iterative basis. We start with an initial guess  $\vec{x}_0$  for the solutions and consider the linear system

$$A\vec{x} = \vec{b} - \vec{r} \quad (16)$$

where  $A, \vec{x}$  and  $\vec{b}$  are defined as before and  $\vec{r}$  is a so-called residual.

Let  $\vec{r}_k$  be the residual at the  $k$ -th step with a negative gradient. Using a similar approach as before we get that the conjugate direction  $\vec{p}_{i+1}$  is

$$\vec{p}_{i+1} = \vec{r}_k - \frac{\vec{p}_k^T A \vec{r}_k}{\vec{p}_k^T A \vec{p}_k} \vec{p}_k \quad (17)$$

The iterative process is then to compute the directions and then solve the arising linear system by computing the coefficients as given in equation 15.

The iterative process is then to compute

$$\vec{x}_{k+1} = \vec{x}_k + C_k \vec{p}_{k+1} \quad (18)$$

with

$$C_k = \frac{\vec{p}_k^T \vec{r}_k}{\vec{p}_k^T A \vec{p}_k} \quad (19)$$

The residual  $\vec{r}_k$  at the  $k$ -th step is defined from equation 16 as

$$\vec{r}_k = \vec{b} - A\vec{x}_k \quad (20)$$

## II.E METROPOLIS-HASTINGS ALGORITHM

The Metropolis algorithm bases itself on moves (also called transitions) as given in a Markov process(or Markov chain). Define a probability distribution function(PDF)  $w_j(t)$  with a transition probability  $w(i \rightarrow j)$  which for a given time-step yields in the Markov formula

$$w_i(t + \varepsilon) = \sum_j w(j \rightarrow i) w_j(t) \quad (21)$$

The transition probability is defined with an acceptance probability distribution  $A(j \rightarrow i)$  and a proposed probability distribution  $T(j \rightarrow i)$  as

$$w(j \rightarrow i) = A(j \rightarrow i) T(j \rightarrow i) \quad (22)$$

The acceptance  $A$  is the probability for the move to be accepted and the proposal  $T$  is different for each problem.

In order for this transition chain to reach a desired convergence and reversibility we have the well known condition for detailed balance  $w_{ij}A_{ji} = w_{ji}A_{ij}$ . This condition gives us that the probability distribution functions satisfy the following condition

$$w_i T_{i \rightarrow j} A_{j \rightarrow i} = w_j T_{j \rightarrow i} A_{i \rightarrow j} \Rightarrow \frac{w_i}{w_j} = \frac{T_{j \rightarrow i} A_{j \rightarrow i}}{T_{i \rightarrow j} A_{i \rightarrow j}} \quad (23)$$

We now need to choose an acceptance which fulfills equation 23 and a common choice is the Metropolis condition

$$A_{j \rightarrow i} = \min \left( 1, \frac{w_i T_{i \rightarrow j}}{w_j T_{j \rightarrow i}} \right) \quad (24)$$

The Metropolis-Hastings algorithm is thus

- (i) Pick initial state  $i$  at random.
- (ii) Pick proposed state at random in accordance to  $T_{j \rightarrow i}$ .
- (iii) Accept state according to  $A_{j \rightarrow i}$ .
- (iv) Jump to step (ii) until a specified number of states have been generated.
- (v) Save the state  $i$  and jump to step (ii).

## II.F VARIATIONAL PRINCIPLE

The variational principle states the following restriction on the ground state energy for a given symmetry

$$E_0 \leq \langle E[\Phi_T] \rangle = \int \phi_T^* \hat{H} \phi_T d\tau = \langle \phi_T | \hat{H} | \phi_T \rangle \quad (25)$$

that is the ground state energy  $E_0$  is bounded by the expectation value of the trial energy.

## II.G IMPORTANCE SAMPLING

In order to use the Metropolis algorithm as explained in section II.E, we need to find the proposal probability distribution labeled  $T_{j \rightarrow i}$ . This is what is known as importance sampling.

This section will derive the importance sampling by using the Fokker-Planck equation for one particle

$$\frac{\partial P}{\partial t} = D \frac{\partial}{\partial x} \left( \frac{\partial}{\partial x} - F \right) P(x, t) \quad (26)$$

where  $F$  is a drift term and  $D$  is a diffusion constant, and the Langevin equation

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta \quad (27)$$

where  $\eta$  is a Gaussian distributed random variable.

### II.G.1 Quantum Force

Since we are working with a isotropic diffusion characterized by a time-dependant probability density our system must obey the summed total Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \sum_i D \frac{\partial}{\partial x_i} \left( \frac{\partial}{\partial x_i} - F_i \right) P(x, t) \quad (28)$$

where  $F_i$  is now the  $i$ 'th component of the drift velocity term(given by an external potential). Since the probability is assumed to be convergent, that is it converges to a stationary probability density the time dependence at this point is zero for all  $i$ . We also know that the drift should be of form  $F = g(x) \frac{\partial P}{\partial x}$  giving

$$\frac{\partial^2 P}{\partial x_i^2} = P \frac{\partial g}{\partial P} \left( \frac{\partial P}{\partial x_i} \right)^2 + P g \frac{\partial^2 P}{\partial x_i^2} + g \left( \frac{\partial P}{\partial x_i} \right)^2 \quad (29)$$

Now we may use the condition for stationary density meaning the left hand side of equation 29 must equal zero giving that  $g = 1/P$ (only possibility the derivatives cancel). Inserting in  $P = \Psi_T$ (see [INSERT REF11](#)) we get that the expression for the quantum force is

$$F = \frac{2}{\Psi_T} \nabla \Psi_T \quad (30)$$

### II.G.2 Solution

Using Eulers method(Euler-Maruyama method [INSERT REF11](#)) on the Langevin equation(equation 27) one obtains the new positions

$$y = x + DF(x)\Delta t + \xi \sqrt{\Delta t} \quad (31)$$

with  $D = 1/2$  in natural units due to the kinetic energy term and  $\Delta t$  is a time-step parameter. The random variable  $\xi$  is within a Gaussian distribution of variance one and standard deviation zero.

For the differential equation (equation 26) we insert in for the quantum force given in equation 30 and arrive at the following diffusion equation

$$\frac{\partial P}{\partial t} = -D \frac{\partial^2 P}{\partial r^2} \quad (32)$$

This equation is solved by using a Fourier transform in the spatial coordinate  $r$  according to equation 7 giving the equation

$$\frac{\partial \tilde{P}(s, t)}{\partial t} = -Ds^2 \tilde{P}(s, t) \quad (33)$$

with solution

$$\tilde{P}(s, \Delta t) = \tilde{P}(s, 0) e^{-Ds^2 \Delta t} \quad (34)$$

Initially the probability density is centered at  $D\Delta t F(x)$ , that is the drift term. This can be expressed terms of a Dirac-delta function, meaning

$$P(y, x, 0) = \delta(y - D\Delta t F(x)) \quad (35)$$

where  $y$  is given in equation 31. Making an inverse transformation as described in equation 8 and solving the subsequent transcendental integral yields in

$$P(y, x, \Delta t) = \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left(-\frac{(y - x - D\Delta t F(x))^2}{4D\Delta t}\right) \quad (36)$$

which gives us the acceptance

$$A_{y \rightarrow x} = \min\left(1, \frac{|\psi_T(y)|^2 P(y, x, \Delta t)}{|\psi_T(x)|^2 P(x, y, \Delta t)}\right) \quad (37)$$

We may rearrange the acceptance ratio giving

$$\frac{P(y, x, \Delta t)}{P(x, y, \Delta t)} = \exp\left(\frac{1}{2} (F(x) + F(y)) \left(\frac{D\Delta t}{2} (F(x) - F(y)) - y + x\right)\right) \quad (38)$$

which is computationally less expensive.

## II.H VMC

This section will explain and derive the equations involved in the Variational Monte Carlo method. The whole section will assume that we have the following trial wave function,  $\psi_T$

$$\psi_T(\vec{r}_1, \dots, \vec{r}_N) \equiv \det(\phi_1(\vec{r}_1, \alpha), \dots, \phi_N(\vec{r}_N, \alpha)) \prod_{i < j}^N \exp\left(\frac{ar_{ij}}{1 + \beta r_{ij}}\right) \quad (39)$$

with the  $\vec{r}$ 's being the position of the electrons and the  $\phi$ 's being the wave function to some known system(i.e harmonic oscillator). The position  $r_{ij}$  is a relative distance  $|\vec{r}_i - \vec{r}_j|$  while  $\alpha$  and  $\beta$  are variational parameters and  $a$  is a specific constant dependant of the total spin symmetry of electron  $i$  and  $j$  as

$$a = \begin{cases} 1, & \text{anti-parallel spin} \\ \frac{1}{3}, & \text{parallel spin} \end{cases} \quad (40)$$

This is also known as a Pade-Jastrow factor.

We also define the total Hamiltonian of the system for the quantum dot case as

$$\hat{H} = \hat{H}_O + \hat{H}_I \quad (41)$$

with  $\hat{H}_O$  being the harmonic oscillator defined in equation 4 and  $\hat{H}_I$  being the Hamiltonian for the electron interactions(Coulomb interaction) defined as

$$\hat{H}_I = \sum_{i < j} \frac{1}{r_{ij}} \quad (42)$$

Lastly, we work in natural units setting  $\hbar = c = 1$ , and all the above equations(equations 39, 40, 41 and 42) also assume natural units.

### II.H.1 Expectation Value and Local Energy

Given the Hamiltonian equation 41 and a trial wave function  $\Psi_T(R, \Lambda)$  and using the variational principle, as given in equation 25 the upper bound for the ground state energy  $E_0$  if  $H(r)$  is

$$E[\hat{H}(R, \Lambda)] \leq \langle \hat{H} \rangle = \frac{\langle \Psi_T | \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \quad (43)$$

where  $R = (r_1, \dots, r_N)$  is the positions to  $N$  particles and  $\Lambda = (\lambda_1, \dots, \lambda_M)$  are the  $M$  variational parameters.

Now we can expand the trial wave function  $\Psi_T(R, \Lambda)$  in the orthonormal eigenstates of the Hamiltonian  $\hat{H}$ (which form a complete set)

$$\Psi_T(r) = \sum_i c_i \Psi_i(r) \quad (44)$$

and the upper bound given in equation 43 is

$$E_0 \leq \frac{\sum_{ij} c_i c_j^* \langle \Psi_j | \hat{H} | \Psi_i \rangle}{\sum_{ij} c_i c_j^* \langle \Psi_j | \Psi_i \rangle} = \frac{\sum_n a_n^2 E_n}{\sum_n a_n^2} \quad (45)$$

where the eigenequation for the Hamiltonian  $\hat{H}\Psi_n = E_n\Psi_n$  was used. The expression given in equation 43 is the expectation value we evaluate in each variational step that is we choose  $\alpha$  according to some minimization algorithm and re-evaluate the expectation value.

In order to introduce the transition probability as given in the Metropolis algorithm(see section II.E) the expectation value, equation 45, needs to be rewritten in terms of a PDF. We can define this as

$$P(R) \equiv \frac{|\Psi_T(R)|^2}{\int |\Psi_T(R)|^2 dR} \quad (46)$$

Now we observe that if we define a quantity

$$E_L(R, \Lambda) \equiv \frac{1}{\Psi_T(R, \Lambda)} \hat{H} \Psi_T(R, \Lambda) \quad (47)$$

which is the so-called local energy. The expectation value given in equation 45 can be rewritten as

$$E[H] = \int P(R) E_L(R, \Lambda) dR \approx \frac{1}{N} \sum_{i=1}^N P(r_i, \Lambda) E_L(r_i, \Lambda) \quad (48)$$

which is of the form given in equation 21 and  $N$  is the number of states(or Monte Carlo cycles).

### II.H.2 Analytical Expression for Local Energy

We use the Metropolis algorithm to find an estimate for the expectation value to the energy. In this expression we have a so-called local energy defined as

$$E_L = \frac{1}{\psi_T} \hat{H} \psi_T \quad (49)$$

This expression shows up in the integrand as the multiplied function to the PDF which is used in the Metropolis algorithm.

### II.H.3 Two Electron Case

We start by finding the local energy in the case with two electrons. The trial wave function is in this case (related to equation 5) using equation 39

$$\psi_T(\vec{r}_1, \vec{r}_2) = A \exp\left(-\frac{\alpha\omega}{2} (r_1^2 + r_2^2)\right) \exp\left(\frac{ar_{12}}{1 + \beta r_{12}}\right) \quad (50)$$

Using the definition of the trial wave function, equation 47 and the total Hamiltonian (equation 41) the local energy with equation 47 is

$$E_L = \frac{1}{\psi_T} (\hat{H}_O \psi_T + \hat{H}_I \psi_T) \quad (51)$$

we solve the first part  $\hat{H}_O \psi_T$

$$\hat{H}_O \psi_T = \frac{1}{2} (-\nabla_1^2 - \nabla_2^2 + \omega^2 (r_1^2 + r_2^2)) \psi_T \quad (52)$$

Starting with the Laplacian for electron 1 and solving the second derivative with respect to  $x_1$  we have

$$\frac{\partial^2 \psi_T}{\partial x_1^2} = A \exp\left(-\frac{\alpha\omega}{2} (r_2^2 + y_1^2)\right) \frac{\partial^2}{\partial x_1^2} \left[ \exp\left(-\frac{\alpha\omega}{2} x_1^2 + \frac{a}{\beta + \frac{1}{r_{12}}}\right) \right] \quad (53)$$

Starting with the first derivative on the exponential we get

$$\left. \begin{aligned} \frac{\partial}{\partial x_1} \left[ -\frac{\alpha\omega}{2} x_1^2 \right] &= -\alpha\omega x_1 \\ \frac{\partial}{\partial x_1} \left[ \frac{a}{\beta + \frac{1}{r_{12}}} \right] &= \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \end{aligned} \right\} \Rightarrow \frac{\partial \psi_T}{\partial x_1} = \left( -\alpha\omega x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \psi_T \quad (54)$$

meaning equation 53 is

$$\begin{aligned} \frac{\partial^2 \psi_T}{\partial x_1^2} &= A \exp\left(-\frac{\alpha\omega}{2} (r_2^2 + y_1^2)\right) \frac{\partial}{\partial x_1} \left[ -\alpha\omega x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right] \exp\left(-\frac{\alpha\omega}{2} x_1^2 + \frac{a}{\beta + \frac{1}{r_{12}}}\right) \\ &= \frac{\partial}{\partial x_1} \left[ \left( -\alpha\omega x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \psi_T \right] \end{aligned} \quad (55)$$

Using the product rule for differentiation and starting with the first expression we get that

$$\frac{\partial}{\partial x_1} \left[ -\alpha\omega x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right] = -\alpha\omega + \frac{a}{r_{12}(1 + \beta r_{12})^2} - \frac{a(x_1 - x_2)^2(1 + 3\beta r_{12})}{r_{12}^3(1 + \beta r_{12})^3} \quad (56)$$

giving

$$\begin{aligned} \frac{\partial^2 \psi_T}{\partial x_1^2} &= \psi_T \frac{\partial}{\partial x_1} \left[ -\alpha\omega x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right] + \left( -\alpha\omega x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right) \frac{\partial \psi_T}{\partial x_1} \\ &= \left[ -\alpha\omega + \frac{a}{(1 + \beta r_{12})^2} - \frac{a(x_1 - x_2)^2(1 + 3\beta r_{12})}{r_{12}^3(1 + \beta r_{12})^3} + \left( -\alpha\omega x_1 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right)^2 \right] \psi_T \end{aligned} \quad (57)$$

The second derivative with respect to  $y_1$  yields with a similar derivation

$$\frac{\partial^2 \psi_T}{\partial y_1^2} = \left[ -\alpha\omega + \frac{a}{(1 + \beta r_{12})^2} - \frac{a(y_1 - y_2)^2(1 + 3\beta r_{12})}{r_{12}^3(1 + \beta r_{12})^3} + \left( -\alpha\omega y_1 + \frac{a(y_1 - y_2)}{r_{12}(1 + \beta r_{12})^2} \right)^2 \right] \psi_T \quad (58)$$

The second derivatives with respect to  $x_2$  and  $y_2$ , are derived in a similar manner, only we get a change in signs when differentiating  $r_{12}$ . This gives

$$\frac{\partial^2 \psi_T}{\partial x_2^2} = \left[ -\alpha\omega + \frac{a}{r_{12}(1 + \beta r_{12})^2} - \frac{a(x_1 - x_2)^2(1 + 3\beta r_{12})}{r_{12}^3(1 + \beta r_{12})^3} + \left( \alpha\omega x_2 + \frac{a(x_1 - x_2)}{r_{12}(1 + \beta r_{12})^2} \right)^2 \right] \psi_T \quad (59)$$

and

$$\frac{\partial^2 \psi_T}{\partial y_2^2} = \left[ -\alpha\omega + \frac{a}{r_{12}(1 + \beta r_{12})^2} - \frac{a(y_1 - y_2)^2(1 + 3\beta r_{12})}{r_{12}^3(1 + \beta r_{12})^3} + \left( \alpha\omega y_2 + \frac{a(y_1 - y_2)}{r_{12}(1 + \beta r_{12})^2} \right)^2 \right] \psi_T \quad (60)$$

gathering equations 57, 58, 59 and 60 we get

$$\begin{aligned} (\nabla_1^2 + \nabla_2^2)\psi_T &= \frac{\partial^2 \psi_T}{\partial x_1^2} + \frac{\partial^2 \psi_T}{\partial y_1^2} + \frac{\partial^2 \psi_T}{\partial x_2^2} + \frac{\partial^2 \psi_T}{\partial y_2^2} \\ &= \left[ -4\alpha\omega + \frac{4a}{r_{12}(1 + \beta r_{12})^2} - \frac{2a(1 + 3\beta r_{12})}{r_{12}(1 + \beta r_{12})^3} + \alpha^2\omega^2(r_1^2 + r_2^2) - \frac{2a\alpha\omega r_{12}}{(1 + \beta r_{12})^2} + \frac{2a^2}{(1 + \beta r_{12})^4} \right] \psi_T \end{aligned} \quad (61)$$

$$= \left[ \alpha^2\omega^2(r_1^2 + r_2^2) - 4\alpha\omega - \frac{2a\alpha\omega r_{12}}{(1 + \beta r_{12})^2} + \frac{2a}{(1 + \beta r_{12})^2} \left( \frac{a}{(1 + \beta r_{12})^2} + \frac{1}{r_{12}} - \frac{2\beta}{1 + \beta r_{12}} \right) \right] \psi_T \quad (62)$$

and the local energy (equation 51) is finally

$$E_L = \frac{1}{2}\omega^2(1 - \alpha^2)(r_1^2 + r_2^2) + 2\alpha\omega - \frac{a}{(1 + \beta r_{12})^2} \left( \frac{a}{(1 + \beta r_{12})^2} - \alpha\omega r_{12} + \frac{1}{r_{12}} - \frac{2\beta}{1 + \beta r_{12}} \right) + \frac{1}{r_{12}} \quad (63)$$

## II.I Optimization

In the method described in section II.H the heavy load in terms of calculation lies within the calculation of the determinant ratio given in equations 37 and 30. This section will derive an expression for these ratios in terms of computation time. We will in the whole section assume we only move one particle at a time in the Monte Carlo cycle.

### II.I.1 Determinant Ratio

In the Metropolis algorithm we calculate a ratio of determinants in the Metropolis test. Starting by defining a Slater determinant Matrix  $D$  with entries defined as

$$D_{ij} \equiv \phi_j(r_i) \quad (64)$$

where the  $\phi$ 's are defined as in section II.H.

In terms of cofactors  $C_{ij}$  we have

$$\det(D) = \sum_{j=1}^N D_{ij} C_{ji} \quad (65)$$

If we now take into light the mentioned assumption about moving only one particle at a time the determinant given in equation 64 only gets a change of one row.

Defining the ratio as

$$R \equiv \frac{\det(D(x^{\text{new}}))}{\det(D(x^{\text{old}}))} \quad (66)$$

Using the fact that when moving the particle at position  $r_i$  the cofactors remain unchanged and inserting in equation 65 into equation 66 we have

$$R = \frac{\sum_{j=1}^N D_{ij}(r^{\text{new}}) C_{ji}(r^{\text{old}})}{\sum_{j=1}^N D_{ij}(r^{\text{old}}) C_{ji}(r^{\text{old}})} \quad (67)$$

Since the Slater is square (closed shell), we have the following ratio  $\frac{\det(D^{\text{new}})}{\det(D^{\text{old}})}$

$$\det(D) = \frac{D^{\dagger}}{D} \quad (68)$$

inserting this into equation 67 we get

$$R = \frac{\sum_{j=1}^N D_{ij}(r^{\text{new}}) D_{ji}^{-1}(r^{\text{old}}) |D|}{\sum_{j=1}^N D_{ij}(r^{\text{old}}) D_{ji}^{-1}(r^{\text{old}}) |D|} \quad (69)$$

Since  $D$  is invertible  $\frac{\det(D^{\text{new}})}{\det(D^{\text{old}})}$  we have

$$\sum_{k=1}^N D_{ik} D_{kj}^{-1} = \delta_{ij} \quad (70)$$



meaning the denominator in equation 69 is equal to 1 and the ratio is finally (with equation 64 inserted)

$$R = \sum_{j=1}^N \phi_j(r_i^{\text{new}}) \phi_{ji}^{-1}(r^{\text{old}}) \quad (71)$$

We can follow a similar approach to find the ratio given in equation 30 for the quantum force. The expression is simply

$$\frac{\nabla_i |D(r_i)|}{|D(r_i)|} = \sum_{j=1}^N \nabla_i \phi_j(r_i^{\text{new}}) \phi_{ji}^{-1}(r^{\text{old}}) \quad (72)$$

## II.J Inverse of Matrix

In section II.I.1 we derived a formula for calculating the ratio of determinants by the inverse of the old determinant when only one row is changed. This section gives a formula for updating the inverse of a matrix in that case. The formula is

$$D_{kj}^{-1}(r^{\text{new}}) = \begin{cases} D_{kj}^{-1}(r^{\text{old}}) - \frac{D_{ik}^{-1}(r^{\text{old}})}{R} \sum_{l=1}^N D_{il}(r^{\text{new}}) D_{lj}^{-1}(r^{\text{old}}), & j \neq i \\ \frac{D_{ik}^{-1}(r^{\text{old}})}{R} \sum_{l=1}^N D_{il}(r^{\text{old}}) D_{lj}^{-1}(r^{\text{old}}), & j = i \end{cases} \quad (73)$$

as described by Sherman and Morris [3, 8].

## III STATISTICAL ANALYSIS

Since Monte Carlo simulations can be considered to be computer experiments the resulting data can be analysed with the same statistical theory as one would with experimental data. This section will give a brief overview of some of the statistical concepts and explain the method of Blocking (for estimating the standard deviation).

### III.1 Standard Deviation, Variance and Covariance

Given a PDF  $P(x)$  we can write the probability for  $P(X)$  to be within a (non-zero) interval  $[a, b]$  as

$$\text{prob}(a \leq X \leq b) = \int_a^b P(x) dx \quad (74)$$

with  $X$  being a stochastic variable.

The probability for  $P(x)$  being less than some  $x$  is

$$C(X \leq x) = \int_{-\infty}^x p(x') dx' \quad (75)$$

The function  $C$  is known as a cumulative probability function, abbreviated as CDF. The relation between a PDF and CDF is

$$P(x) = \frac{d}{dx} C(x) \quad (76)$$

With this the PDF and CDF defined we can look at a special class of expectation values called moments. The  $n$ -th moment of a PDF  $P$  is defined as

$$\langle x^n \rangle \equiv \int x^n P(x) dx \quad (77)$$

This means the zeroth moment ( $n = 0$ ) is the normalization for  $P$  and the first moment is called the mean value for  $P$ . A special case of moments is the so-called central moments, the  $n$ -th central moment is defined to be

$$\langle (x - \langle x \rangle)^n \rangle \equiv \int (x - \langle x \rangle)^n P(x) dx \quad (78)$$

The second central moment is known as the variance of  $P$ . Given a stochastic variable  $X$  the variance labeled as  $\sigma_X^2$  is

$$\text{var}(X) = \sigma_X^2 = \langle (x - \langle x \rangle)^2 \rangle = \int (x - \langle x \rangle)^2 P(x) dx \quad (79)$$

$$= \int (x^2 - 2x \langle x \rangle + \langle x \rangle^2) P(x) dx \quad (80)$$

$$= \langle x^2 \rangle - 2 \langle x \rangle \langle x \rangle + \langle x \rangle^2 \quad (81)$$

$$= \langle x^2 \rangle - \langle x \rangle^2 \quad (82)$$

The square root of the variance is known as the standard deviation and is equal to the RMS(root-mean-square) of  $P$ , which is interpreted as the spread of  $P$ .

The last important statistical quantity is the covariance. For two stochastic variables  $X_i$  and  $X_j$  the covariance between them is define to be

$$\text{cov}(X_i, X_j) \equiv \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \quad (83)$$

where

$$\langle x_i \rangle = \int \cdots \int x_i P(x_1, \dots, x_n) dx_1 \dots dx_n \quad (84)$$

and  $P$  being a multivariate PDF.

For a finite size sample we replace the integrals with sums giving

$$\bar{x}_n \equiv \frac{1}{n} \sum_{k=1}^n x_k \quad (85)$$

$$\text{var}(x) \equiv \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x}_n)^2 \quad (86)$$

$$\text{cov}(x) \equiv \frac{1}{n} \sum_{kl} (x_k - \bar{x}_n) (x_l - \bar{x}_n) \quad (87)$$

The method of blocking estimates the standard deviation given in

## IV SETUP

## V RESULTS

## VI DISCUSSION

## VII CONCLUSION

## VIII References

- [1] David J. Griffiths Chapters 2, 4 and 7, *Introduction to Quantum Mechanics Second Edition*, Pearson, 2005, ISBN 0-13-111892-7.
- [2] Morten Hjort-Jensen, *Computational Physics: Hartree-Fock methods and introduction to Many-Body Theory*. <https://www.github.com/CompPhysics/ComputationalPhysics2/blob/gh-pages/doc/pub/basicMB/pdf>, 2017
- [3] Morten Hjort-Jensen, *Computational Physics: Variational Monte Carlo methods*. <https://www.github.com/CompPhysics/ComputationalPhysics2/blob/gh-pages/doc/pub/vmc/pdf>, 2017
- [4] M. Taut, *Two Electrons in an External Oscillator Potential; Particular analytic solutions of a Coulomb Correlation Problem*, Physical Review A 48, 3561, 1993.
- [5] Negele and Orland, *Quantum Many-Particle Systems*, Addison-Wesley.
- [6] Fetter and Walecka, *Quantum Theory of Many-Particle Systems*, Mcgraw-Hill, 1971.
- [7] Dickhoss and Van Neck, *Many-Body Theory Exposed*, World Scientific, 2006
- [8] Jack Sherman and Winifred J. Morrison, *Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix*. Annals of Mathematical Statistics 21, 1950. doi: <http://dx.doi.org/10.1214/aoms/1177729893>.