


# COMPUTATIONAL PHYSICS II: QUANTUM MECHANICAL SYSTEMS

## Variational Monte Carlo for Bosons in a Harmonic Oscillator Trap

Mohamad Mahmoud

 [GitHub - click here -](#)

April 2021

### Abstract

A variational Monte-Carlo (VMC) approach for estimating the ground state energy of a Bose-Einstein condensate of alkali atoms such as  $^{87}\text{Rb}$ . The study was conducted for bosons in both a spherical and elliptical harmonic oscillator trap. The spherical trap creating the groundwork for the simple non-interacting case with a known analytical solution. Which we have used as a validity test for our numerical implementations within and around VMC. The analytical solution served as an essential component in determining the optimal parameters and the validity of our numerical differentiation results. The parameters were later used to study the behavior of a system in an elliptical trap with inter-boson interactions between particles. The one-body density along the z-axis showed the effect changing the shape of the trap had on the distribution of particles. While the one-body density in the xy-plane showed the repulsive effect of the *Jastrow factor*.

## Contents

<b>I</b>	<b>Introduction</b>	<b>2</b>
<b>II</b>	<b>Theory</b>	<b>2</b>
i	The System . . . . .	2
	The Simple Gaussian . . . . .	3
ii	The Local Energy . . . . .	3
	The Simple Gaussian Wave Function . . . . .	3
	The Full Wave Function . . . . .	4
	Scaling the Hamiltonian . . . . .	4
<b>III</b>	<b>Method</b>	<b>5</b>
i	Variational Monte Carlo with the Metropolis Algorithm . . . . .	5
ii	The Metropolis-Hastings Algorithm . . . . .	5
iii	Gradient Descent . . . . .	6
iv	Resampling Methods . . . . .	6
	Bootstrapping . . . . .	7
	Blocking . . . . .	7
v	Equilibration . . . . .	7
vi	One-Body Density . . . . .	7
<b>IV</b>	<b>Results &amp; Discussions</b>	<b>8</b>
i	The Simple Gaussian . . . . .	8
	Brute-Force Metropolis Sampling . . . . .	8
	Importance Sampling . . . . .	9
	Resampling Methods . . . . .	10
ii	The Full Wave Function . . . . .	11
	Gradient Descent & Ground State Energy . . . . .	11
	One-Body Density . . . . .	12
	The Jastrow Factor . . . . .	12
<b>V</b>	<b>Conclusion</b>	<b>13</b>
i	The Simple Gaussian . . . . .	13

ii	The Full Wave Function . . . . .	13
iii	Summary . . . . .	13
<b>References</b>		<b>15</b>
<b>VI Appendix</b>		<b>16</b>
i	Calculations for non-interacting bosons . . . . .	16
ii	Calculations for interacting bosons . . . . .	19
	Scaling the Hamiltonian . . . . .	19
	The Local Energy . . . . .	19
iii	Figures . . . . .	23
	The Simple Gaussian . . . . .	23
	Correlated . . . . .	25

## I. INTRODUCTION

A Bose-Einstein Condensate (BEC) is a gas composed of bosons, particles of integer spin that obey Bose-Einstein statistics, cooled down to a critical temperature  $T_C$  nearing absolute zero at which the gas goes through a phase transition. In this phase every particle falls into the lowest quantum state usually referred to as the ground state. Thus, every particle can be described by the same single particle wave function and therefor the expectation value of the energy is as follows

$$\langle E \rangle = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}, \alpha) H(\mathbf{R}) \Psi_T(\mathbf{R}, \alpha)}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}, \alpha) \Psi_T(\mathbf{R}, \alpha)} \quad (1)$$

where  $\mathbf{R}$  is the vector containing the position of all particles and  $\alpha$  is the variational parameter. The variational principle states that the expectation value  $\langle E \rangle$  is an upper bound for the ground state energy  $E_0$ .

$$E_0 \leq \langle E \rangle \quad (2)$$

This lays the groundwork for a field of high interest in physics. An example of a such system is Bose-Einstein Condensation in gases of alkali atoms such as  $^{87}\text{Rb}$ ,  $^{23}\text{Na}$ ,  $^7\text{Li}$  confined in magnetic traps. These confined Bose systems are dilute, [1]. This is can be shown by looking at the characteristic dimensions of a typical trap for the  $^{87}\text{Rb}$  gas  $a_{ho} = (\hbar/m\omega_{\perp})^{1/2} = 1 - 2 \times 10^4 \text{ \AA}$  and the interaction between atoms given by its s-wave scattering length  $a_{Rb}$  with the definite value usually set to  $100a_0$ , where  $a_0 = 0.5292 \text{ \AA}$  is the Bohr radius. This results in the definite ratio of atom size to trap size  $a_{Rb}/a_{ho} = 4.33 \times 10^{-3}$ . The atom density for  $^{87}\text{Rb}$  in the trap is typically  $n \simeq 10^{12} - 10^{14} \text{ atoms/cm}^3$  giving an inter-atom spacing  $l \simeq 10^4 \text{ \AA}$ . Thus, giving that the effective atom size is small compared to both the trap size and the inter-atomic spacing the system is indeed dilute. This alone means that the physics is dominated by the two-body collisions between particles, represented by the s-wave scattering length  $a_{Rb}$ , [2].

Many theoretical studies of Bose-Einstein condensates in gases of alkali atoms confined in magnetic or optical traps have been conducted in the frame work of the Gross-Pitaevskii (GP) equation [3]. We will consider a system of

$N$  bosons with mass  $m$  which are trapped in a harmonic oscillator potential.

Typically large multi-dimensional systems like these are very complex and cannot be solved analytically, therefore we turn to numerical approaches in order to simulate the system. We'll start off by studying the non-interacting case where an analytical solution is presented and venture off to the correlated case, which you will see is extremely convoluted and can only be solved numerically. The analytical non-interacting case works as a validity gauge for our numerical approaches. We will be using Variational Monte Carlo (VMC) in order to arrive to a good estimate of what the ground state for eq. 1 might be for any of the systems at hand.

In addition we will be contrasting brute-force VMC to an importance sampling technique, utilizing the Fokker-Planck equation that takes educated guesses in changing the particles position leading to better results overall. We will also be looking at two resampling methods *bootstrapping* and *blocking* to gain better understanding of our sample error as a result of our numerical approach and using gradient descent to determine the value of the variational parameter  $\alpha$  that yields the ground state energy.

## II. THEORY

### i. The System

We're evaluating bosons trapped in a harmonic oscillator potential either spherical ( $S$ ) or elliptical ( $E$ ). The system itself generally consist of  $N$  bosons with a fixed mass  $m$  where they traverse along a  $d$ -dimensional room. Our generalized equations will mainly focus on the 3 dimensional case. The external potential  $V_{ext}$  for both the spherical and the elliptical trap is given by

$$V_{ext}(\mathbf{r}) = \begin{cases} \frac{1}{2}m\omega_{ho}r^2 & (S) \\ \frac{1}{2}m \left[ \omega_{ho} (x^2 + y^2) + \omega_z^2 z^2 \right] & (E) \end{cases} \quad (3)$$

Where  $\omega_{ho}$  defines the trap's potential strength and as for the elliptical case  $\omega_{ho}$  is the frequency in the  $xy$ -plane or the perpendicular, whilst  $\omega_z$  is the frequency in the  $z$ -direction. At  $T = 0 \text{ K}$  the mean amplitude of a single boson is given by  $\langle x^2 \rangle = (\hbar/2m\omega_{ho})$  such

that  $a_{ho} = (\hbar/m\omega_{ho})^{1/2}$  defines the characteristic length of the trap. The ratio of the frequencies is denoted by  $\gamma = \omega_z/\omega_{ho}$ , leading to the ratio of the trap lengths  $(a_{ho}/a_z) = (\omega_z/\omega_{ho})^{1/2} = \sqrt{\gamma}$ , [1].

The inter-boson interactions between two particles is expressed as a repulsive potential  $V_{int}$  between the two.

$$V_{int}(\mathbf{r}) = \begin{cases} \infty & \text{for } |\mathbf{r}_i - \mathbf{r}_j| \leq a \\ 0 & \text{for } |\mathbf{r}_i - \mathbf{r}_j| > a \end{cases} \quad (4)$$

Where  $a$  is the hard-core diameter of the bosons so  $V_{int}$  naturally prevents the particles from ever occupying the same space in the case of a collision. The Hamiltonian for our harmonic oscillator trap, taking into account both of the potentials, is written as

$$H = \sum_i^N \left( \frac{-\hbar^2}{2m} \nabla_i^2 + V_{ext}(\mathbf{r}) \right) + \sum_{i < j} V_{int}(\mathbf{r}_i, \mathbf{r}_j) \quad (5)$$

Where we have used the following convention for the inner potential:

$$\sum_{i < j} V_{ij} = \sum_{i=1}^N \sum_{j=i+1}^N V_{ij}$$

Which is a double sum that runs over every pairwise interaction once.

For the  $\sum_i^N$ -part the first term is the kinetic energy of each particle and the second term is as previously mentioned the the external potential of the trap acting on each particle. The  $\sum_{i < j}$ -part is for the inter-boson interaction between a pair of particles.

Now that we laid the ground for the system, we are in need of one additional powerful tool that would describe the system in it's entirety throughout the whole process and that is a trial wave function. Our initial educated guess for the wave function consist of two parts one for the harmonic trap effecting every single particle each for its own and one for the inter-boson interaction or in other words the correlated aspect of the system.

$$\begin{aligned} \Psi_T(\mathbf{r}) &= \Psi_T(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, \alpha, \beta) \\ &= \left[ \prod_i g(\alpha, \beta, \mathbf{r}_i) \right] \left[ \prod_{j < k} f(a, |\mathbf{r}_j - \mathbf{r}_k|) \right] \end{aligned} \quad (6)$$

where  $\alpha$  and  $\beta$  are two variational parameters and  $a$  is the diameter in eq. 4 for the hard-core *shell* of the particle. The first term follows a Gaussian distribution and is a composition of the single-particle wave function  $g$ .

$$g(\alpha, \beta, \mathbf{r}_i) = e^{-\alpha(x_i^2 + y_i^2 + \beta z_i^2)} \quad (7)$$

Here  $\beta$  determines the elliptical shape of the distribution, so in the case where  $\beta = 1$  we end with a sphere and this indeed the case for the spherical trap. As for the correlation wave function  $f$ , also referred to as *the Jastrow factor*, it's represented by

$$f(a, |\mathbf{r}_j - \mathbf{r}_k|) = \begin{cases} 0 & |\mathbf{r}_j - \mathbf{r}_k| \leq a \\ 1 - \frac{a}{|\mathbf{r}_j - \mathbf{r}_k|} & |\mathbf{r}_j - \mathbf{r}_k| > a \end{cases} \quad (8)$$

### The Simple Gaussian

For the non-interacting case eq. 8 is a nonfactor, subsequently  $f = 1$ . This is done by setting the hard-core diameter to  $a = 0$ . Since in this case we are not interested in any boson-to-boson interaction and as such we get a product sequence over the Gaussian distribution for each particle. Furthermore to make so the trap follows a spherical form we set  $\beta = 1$ . From that we obtain the simplest form our wave function can take and it's given by

$$\Psi_T(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, a = 0, \beta = 1) = \prod_i e^{-\alpha r_i^2} \quad (9)$$

We will be referring to eq. 9 as *the simple Gaussian*.

A quantity used in our Metropolis-Hastings optimization is the quantum force and for the Simple-Gaussian wave function it takes the form

$$F = 2 \frac{\nabla \Psi_T}{\Psi_T} = -4\alpha \sum_i r_i \quad (10)$$

An entire derivation can be found in **VI Appendix i**.

### ii. The Local Energy

$$E_L(\mathbf{r}) = \frac{1}{\Psi_T(\mathbf{r})} H \Psi_T(\mathbf{r}) \quad (11)$$

### The Simple Gaussian Wave Function

An entire walk-through of the calculation can be found in **VI. Appendix i**.

For the Simple Gaussian case there's is no correlation and the Hamiltonian naturally becomes

$$H_{OB} = \sum_i^N \left( \frac{-\hbar^2}{2m} \nabla_i^2 + V_{ext}(\mathbf{r}) \right) \quad (12)$$

where OB stands for one-body. Since it's a spherical trap we use  $(S)$  in eq. 3. We first start by determining an analytical solution for the Laplacian of the wave function from eq. 11.

$$\begin{aligned}
\nabla_i^2 \Psi_T(\mathbf{r}) &= \nabla_i \cdot \nabla_i \Psi_T \\
&= -2\alpha \nabla_i \cdot [x_i, y_i, z_i] \Psi_T \\
&= [-2\alpha d + 4\alpha^2 r_i^2] \Psi_T(\mathbf{r})
\end{aligned} \tag{13}$$

where  $d$  is the number of dimensions, case being  $d = 3$ .

For the simple Gaussian trial function the local energy is given by

$$\begin{aligned}
E_L(\mathbf{r}) &= \sum_i^N \left( -\frac{\hbar^2}{2m} [-2\alpha d + 4\alpha^2 r_i^2] + V_{ext}(\mathbf{r}_i) \right) \\
&= \sum_i^N \left( -\frac{\hbar^2}{2m} [-2\alpha d + 4\alpha^2 r_i^2] + \frac{1}{2} m \omega_{ho} r_i^2 \right) \\
&= -\frac{\hbar^2}{2m} [-2\alpha N d + 4\alpha^2 \sum_i^N r_i^2] + \frac{1}{2} m \omega_{ho} \sum_i^N r_i^2
\end{aligned} \tag{14}$$

As for our upcoming numerical analysis we will be using natural units instead of the physical constants, meaning  $\hbar = m = \omega_{ho} = 1$ , which yields

$$E_L(\mathbf{r}) = \alpha N d - 2\alpha^2 \sum_{i=1}^N r_i^2 + \frac{1}{2} \omega_{ho}^2 \sum_{i=1}^N r_i^2 \tag{15}$$

For the case where  $\omega_{ho} = 1$  we find the ground state to occur for  $\alpha = 0.5$

$$E_L(\mathbf{r}) = \frac{Nd}{2} \tag{16}$$

which will work as a gauge for the validity of our numerical solutions later on.

### The Full Wave Function

*An entire walk-through of the calculation can be found in VI. Appendix ii.*

A more realistic case is the interacting case for bosons in an elliptical trap; ( $E$ ) in eq. 3. Meaning we're looking at the correlated wave function eq. 6, where now  $\beta$  could be different from 1 and the hard-core diameter  $a > 0$ . This ultimately means that our wave function contains the Jastrow factor to govern the boson-to-boson interaction and the Hamiltonian contains the repulsive potential eq. 4.

We rewrite the Jastrow factor as

$$f(r_{ij}) = \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \tag{17}$$

here  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  and  $u(r_{ij}) = \ln f(r_{ij})$ . In addition we express the one-body part as

$$\phi(\mathbf{r}_i) = g(\alpha, \beta, \mathbf{r}_i) \tag{18}$$

The final expression for the correlated wave function becomes

$$\Psi_T(\mathbf{r}) = \left[ \prod_i \phi(\mathbf{r}_i) \right] \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \tag{19}$$

It's the same as for the Simple Gaussian case we start by determining an analytical solution for the Laplacian then we can arrive to the Laplacian-term in the local energy

$$\begin{aligned}
\frac{1}{\Psi_T} \nabla_k^2 \Psi_T &= \frac{\nabla_k^2 \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2 \frac{\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} \sum_{l \neq k} u'(r_{kl}) \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \\
&\quad + \left( \sum_{l \neq k} u'(r_{kl}) \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \right)^2 \\
&\quad + \sum_{l \neq k} u''(r_{kl}) + u'(r_{kl}) \frac{2}{r_{kl}}
\end{aligned} \tag{20}$$

Where

$$\frac{\nabla_k^2 \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} = 4\alpha^2 \mathbf{r}_k^2 - 2\alpha(2 + \beta) \tag{21}$$

$$\frac{\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} = -2\alpha \mathbf{r}_k \tag{22}$$

$$u'(r_{kl}) = \frac{a}{r_{kl}(r_{kl} - a)} \tag{23}$$

$$u''(r_{kl}) = \frac{a^2 - 2ar_{kl}}{r_{kl}^2(r_{kl} - a)^2} \tag{24}$$

The drift force is given by

$$F_k = 2 \frac{\nabla_k \Psi_T}{\Psi_T} = 2 \frac{\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2 \sum_{l \neq k} u'(r_{kl}) \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \tag{25}$$

As already mentioned a complete walk-through for the gradient and the Laplacian can be found in the **Appendix**.

### Scaling the Hamiltonian

*An entire walk-through of the calculation can be found in VI. Appendix ii.*

Since we are in the process of rewriting equations. Often quantum systems are expressed in units of  $\hbar\omega$  for convenience. In the scaling for the Hamiltonian we aim to increase numerical stability by simplifying the expression for the tools we're working with and achieve a physical interpretation of the system by expressing the energy as units of  $\hbar\omega$ . This is done by factoring out  $\hbar\omega_0$  and dividing the Hamiltonian by that factor, which yields

$$H = \frac{1}{2} \sum_i^N \left( -\nabla_i^2 + x_i^2 + y_i^2 + \gamma^2 z_i^2 \right) + \sum_{i < j} V_{int}(\mathbf{r}_i, \mathbf{r}_j) \quad (26)$$

Following **VI. Appendix ii.** we have used:

$$a_{ho} = \sqrt{\hbar/m\omega_{ho}} \quad (27)$$

as the length scale and where we defined the parameter

$$\gamma = \omega_z/\omega_{ho} \quad (28)$$

Then scaled the particle coordinates and the hard-core diameter by  $a_{ho}$ , meaning

$$\mathbf{r}' = \mathbf{r}/a_{ho} \implies \mathbf{r} = a_{ho}\mathbf{r}' \quad (29)$$

$$a' = a/a_{ho} \implies a = a_{ho}a' \quad (30)$$

The hard-core diameter for the bosons was fixed to  $a/a_{ho} = 0.00433$  and  $\beta = \gamma = \sqrt{8} \approx 2.82843$  similar to what was done in [1] and [2].

### III. METHOD

#### i. Variational Monte Carlo with the Metropolis Algorithm

*Everything here is strictly derived from the lecture notes [4] for VMC.*

Given the expectation value of the energy as shown in the **Introduction**

$$\langle E \rangle = \frac{\int d\mathbf{R} \Psi_T^*(\mathbf{R}, \alpha) H(\mathbf{R}) \Psi_T(\mathbf{R}, \alpha)}{\int d\mathbf{R} \Psi_T^*(\mathbf{R}, \alpha) \Psi_T(\mathbf{R}, \alpha)} \quad (31)$$

Using the quantum properties of the wave function we devise a probability density function (PDF) defined as

$$P(\mathbf{R}) = \frac{|\Psi_T|^2}{\int |\Psi_T|^2 d\mathbf{R}} \quad (32)$$

and then another useful quantity referred to as the local energy

$$E_L(\mathbf{R}) = \frac{1}{\Psi_T(\mathbf{R})} H \Psi_T(\mathbf{R}) \quad (33)$$

Together they yield a new expression for the expectation value

$$\langle E \rangle = \int P(\mathbf{R}) E_L(\mathbf{R}; \alpha) d\mathbf{R} \approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i; \alpha) \quad (34)$$

$M$  being the number of Monte Carlo cycles. This approximation stems from the Bernoulli's law of large numbers, which states that the sample mean approaches the true mean as  $M \rightarrow \infty$ . Where the sample mean in this case is the expectation value of the energy,  $\langle E \rangle$ , and according to

the variational principle it's an upper bound for the true ground state energy  $E_0$ .

$$E_0 \leq \langle E \rangle \quad (35)$$

#### Performing a Variational Monte Carlo calculation

- For a value of  $\alpha$  evaluate the wave function  $\psi_{old}$  for the many-body system of  $N$  particles located in the position given by  $\mathbf{R}$  together with the local energy and it's squared counterpart.
- Draw a random particle and propose a change  $R' = R + r \cdot \text{step}$ , where  $r \in [-0.5, 0.5]$  is a random variable and step is a tuned MC step length.
- Evaluate the wave function  $\psi_{new}$  from that calculate the Metropolis test  $|\psi_{old}|^2/|\psi_{new}|^2$  in order to accept or decline the proposed change.

$$A_{i \rightarrow j} = \min \left( 1, \frac{|\psi_{old}|^2}{|\psi_{new}|^2} \right)$$

- If the proposed change is accepted calculate the local energy for the new wave function, store these values. Otherwise, store the local energy of the old wave function and revert to the old position  $R$ .
- Repeat  $M$  times and compute the averages at the end.
- A true energy minimum is reached when the variance  $\sigma_E(\alpha) = \langle E^2 \rangle - \langle E \rangle^2 = 0$

#### ii. The Metropolis-Hastings Algorithm

*Everything here is strictly derived from the lecture notes [5] for Importance Sampling.*

The Fokker-Planck equation

$$\frac{\partial P(x, t)}{\partial t} = D \frac{\partial}{\partial x} \left( \frac{\partial}{\partial x} - F \right) P(x, t) \quad (36)$$

$D$  is the diffusion coefficient,  $F$  is the drift term (the quantum force in our case) and  $P$  is the probability density. The goal here is to use eq. 36 to derive a better sampling rule with a higher acceptance ratio. Solving eq. 36 as shown in [5], yields a solution for the quantum force  $F$  given by

$$F = 2 \frac{\nabla \Psi_T}{\Psi_T} \quad (37)$$

The new position is determined by the Langevin equation

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta \quad (38)$$

where  $\eta$  is a random variable, yielding a new position

$$y = x + DF(x)\Delta t + \eta\sqrt{\Delta t} \quad (39)$$

The quantity  $D$  (in atomic units) is  $1/2$  which stems from the  $1/2$ -factor in the kinetic energy.  $\Delta t$  is a parameter in itself and values of  $\Delta t \in [0.001, 0.01]$  yield stable values of the ground state energy, according to [5].

A solution of the Fokker-Planck equation, yields the transition probability that will determine the acceptance ratio. This solution is known as the Green's function

$$G(y, x, \Delta t) = \frac{1}{(4\pi D\Delta t)^{3N/2}} e^{-(y-x-D\Delta t F(x))^2/4D\Delta t} \quad (40)$$

Thus Metropolis-Hastings substitute the transition probabilities from the prior ordinary Metropolis with the Green's function. Yielding an acceptance probability

$$A(y, x) = \min \left( 1, \frac{G(x, y, \Delta t) |\Psi_T(y)|^2}{G(y, x, \Delta t) |\Psi_T(x)|^2} \right) \quad (41)$$

In the ordinary Metropolis algorithm, we have a so-called brute-force sampling method, and to yield viable results far from all suggestions will actually be accepted (typically around 70%). By using the Langevin equation this ratio of accepted moves can be increased, fig. 11 in **VI. Appendix. iii**. A suggested move will still be done on a random particle, but the direction will be biased in the direction given by the Langevin equation.

### iii. Gradient Descent

*Everything here is strictly derived from the lecture notes [6] for Gradient Descent.*

For the non-interacting case we managed to determine analytically which value for  $\alpha$  coincides with the ground state. It's not always apparent which value of our variational parameter yields the ground state energy and for that we need a method to determine the minimum value. Here's where the Gradient Descent method comes in

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \gamma_k \nabla \mathbf{F}(\mathbf{x}_k)$$

The basic idea is that any function,  $\mathbf{F}(\mathbf{x})$ , for that matter decreases the fastest if one goes from  $\mathbf{x}$  in the direction of the negative gradient  $-\nabla \mathbf{F}(\mathbf{x})$ . Here  $\gamma_k > 0$  is the learning rate which gauges the severity in which we move in the direction of the negative gradient. Given a sufficiently small  $\gamma_k$  the condition  $\mathbf{F}(\mathbf{x}_{k+1}) \leq \mathbf{F}(\mathbf{x}_k)$  is fulfilled, which means with each iteration  $k$  we move towards a smaller value, i.e. a minimum.

Our trial wave function is dependent on  $\alpha$  as a variational parameter. The suggested approach is to treat the local energy,  $E_L$ , as a cost function. This is done by differentiating wrt.  $\alpha$  in an attempt at arriving to the

optimal value of that parameter, that yields the ground state energy.

$$\alpha_{i+1} = \alpha_i + \gamma \nabla \langle E_L(\alpha_i) \rangle \quad (42)$$

Here  $\nabla \langle E_L(\alpha_i) \rangle$  is the gradient of the expectation value for the local energy wrt.  $\alpha$  and  $\gamma$  is again the learning rate. So instead of having to manually find the best optimal variational parameter we will be using this iterative scheme to hopefully determine where the global minimum lies.

It's worth noting that gradient descent has a design flaw and that it is *deterministic*. Meaning, this approach converges to a minimum disregarding whether it's a local or a global one. This ultimately means that the initial guess is very important in regard to finding the global minimum and thus some trial and error must be expected.

As for the gradient of the cost function

$$\bar{E}_\alpha = \nabla \langle E_L(\alpha_i) \rangle = \frac{d \langle E_L(\alpha_i) \rangle}{d\alpha} \quad (43)$$

by the chain rule and the hermiticity of the Hamiltonian the derivative wrt.  $\alpha$  is given by

$$\bar{E}_\alpha = 2 \left( \left\langle \frac{\bar{\Psi}_\alpha}{\Psi_\alpha} E_L(\alpha) \right\rangle - \left\langle \frac{\bar{\Psi}_\alpha}{\Psi_\alpha} \right\rangle \langle E_L(\alpha) \rangle \right) \quad (44)$$

Where

$$\bar{\Psi}_\alpha = \frac{d\Psi_\alpha}{d\alpha} \quad (45)$$

is the derivative of the wave function wrt.  $\alpha$ , see **VI Appendix i**.

### iv. Resampling Methods

*Everything here is strictly derived from the lecture notes [7] for Resampling Methods.*

In the case when the sampled measurements shaping up the data are independent and identically distributed (**idd**), the variance of the mean is given by

$$\text{var}(\bar{\mathbf{X}}) = \frac{\sigma^2}{N} \quad (46)$$

where  $N$  is the number of samples and  $\sigma^2$  is the variance. When the requirement of **idd** samples isn't satisfied our estimate for the variance, eq. 46, becomes an overly optimistic estimate of the true sample error. Thus, the estimation of the variance is more complicated for correlated data.

In order to have a better assessment of the variance we are going to adopt two resampling techniques *bootstrapping* and *blocking*. The goal is to arrive at a better

estimate of the variance as a function of the variational parameter  $\alpha$  and thus being able to conclude which value of  $\alpha$  results in the ground state of the system. This is especially helpful when working with the correlated case and the numerical double derivative of the Hamiltonian operator.

### Bootstrapping

The bootstrap method randomly draws points from the original dataset, creating a smaller dataset to work with. The idea is to sample/draw random points, *with replacement*, from the original dataset  $k$  times and for each bootstrapping set we compute the variables we're interested in. That being the variance,  $\sigma_E^2$ , in our case. Thus, the bootstrapping variance would then be the mean variance of the  $k$ -times-sampled variance.

#### Bootstrapping Steps

- From a dataset  $\mathbf{x}_N$  containing  $N$  points. Draw, *with replacement*,  $n$  random points yielding a dataset  $x_n$ .
- From the drawn dataset compute the desired statistics.
- Repeat the process  $k$  times then compute the averages.

### Blocking

Here we want to thank Marius Jonsson, [8], for providing both the theoretical background and code needed for the automated blocking algorithm.

Jonsson's work [8] very basis is the blocking method popularized by Flyvbjerg and Petersen [9]. Blocking provides an alternative approach for estimating the variance. Its accuracy is similar to that of dependent bootstrapping with much better computational complexity of  $\mathcal{O}(n)$ ; as dependent bootstrapping scales poorly for large  $n$ .

The blocking algorithm requires the sample size to be  $N = 2^d$ , where  $d \geq 1$ . It's demanded since a *blocking transformation* is made by taking the mean of every pair of subsequent observations from  $\bar{\mathbf{X}}$  and constructs a new vector  $\bar{\mathbf{X}}_1$ . The process is repeated  $d$  times until  $\bar{\mathbf{X}}_{d-1}$ , resulting in  $d$  vectors in total,  $\bar{\mathbf{X}}_0, \bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_{d-1}$  containing the subsequent averages of observations. For each iteration the produced dataset is of the size  $N_i = 2^{d-i}$ , halving the number of observations with each iteration.

The variance as a result of a blocking transformation can be calculated by

$$\begin{aligned} V(\bar{\mathbf{X}}_k) &= \frac{\sigma_k^2}{n_k} + \frac{2}{n_k} \sum_{h=1}^{n_k-1} \left(1 - \frac{h}{n_k}\right) \gamma_k(h) \\ &= \frac{\sigma_k^2}{n_k} + e_k \end{aligned} \quad (47)$$

Here  $n_k = n/2^k$  is the size of the vector  $\bar{\mathbf{X}}_k$  after  $k$  blocking transformations,  $\gamma_k$  denotes the autocovariance and  $e_k$  is the truncation error. Flyvbjerg and Petersen show in their work that the sequence  $\{e_k\}$  becomes sufficiently small with a large number of iterations, hence large  $d$ . This means that the blocking transformation can be applied until the truncation error term is a relatively small quantity for our estimation of  $V(\bar{\mathbf{X}}_k)$  and thus end up with  $\sigma_k^2/n_k$  as our estimate.

### v. Equilibration

Due to the nature of how particles (walkers) are initiated at the start of the simulation, some equilibration in the direction governed by the acceptance probability of either the Metropolis or the Metropolis-Hastings step is needed. This is due to the fact that the random initialization of walkers could lead to an initial-positions configuration that is far from ideal and unlikely for the quantum case. Thus, leading to false estimates in what the energy expectation value could be.

Therefor, for each simulation we chose to run what amounts to 10 % of the number of MC-cycles as equilibration cycles before starting the sampled Monte-Carlo process in searching for the expectation values of the system.

### vi. One-Body Density

The one-body density, Nordhagen [10], represents the probability of finding a particle  $i$  at a certain position within the quantum system and entails the distribution of particles when calculated for all particles. Mathematically it is given by integrating over all particles except one

$$\rho(\mathbf{r}_i) = N \int_{-\infty}^{\infty} d\mathbf{r}_1 \cdots d\mathbf{r}_{i-1} d\mathbf{r}_{i+1} \cdots d\mathbf{r}_N |\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)|^2 \quad (48)$$

This is due to the indistinguishable nature of particles in quantum mechanics. The one-body density is the same regardless of which particle  $i$  we leave out. Any of the coordinates that we integrate over contain the information over all particles. Meaning  $\rho(\mathbf{r}_i)d\mathbf{r}_i$  represents the probability of finding any of the system's  $N$  particles within the volume element  $d\mathbf{r}_i$ . Thus, it should be normalized to the number of particles  $N$  that make up the system and conveniently

$$\int_{-\infty}^{\infty} d\mathbf{r}_i [\rho(\mathbf{r}_i)] = N \quad (49)$$

when working with normalized trial wave functions.

Finding an analytical expression for eq. 48 is generally difficult and require a rather simplified quantum system. But it can be solved numerically using Monte-Carlo integration by embedding it into what we have done so far with our VMC-analysis. This is done by devising bins for each of the dimensions and subsequently counting the

number of particles with a position that falls within that bin for each MC-cycle. At the end, scale down the integral to create a normalized histogram that shows the probability of a particle being within a certain bin of size  $\Delta l$  along one of the axes. Alternatively the radial density can be calculated as has been done in [10].

## IV. RESULTS & DISCUSSIONS

### i. The Simple Gaussian

#### Brute-Force Metropolis Sampling

We start by determining the best parameters to work with for our numerical analysis; such as the **number of threads** and the **steplength** utilized for each Metropolis step.

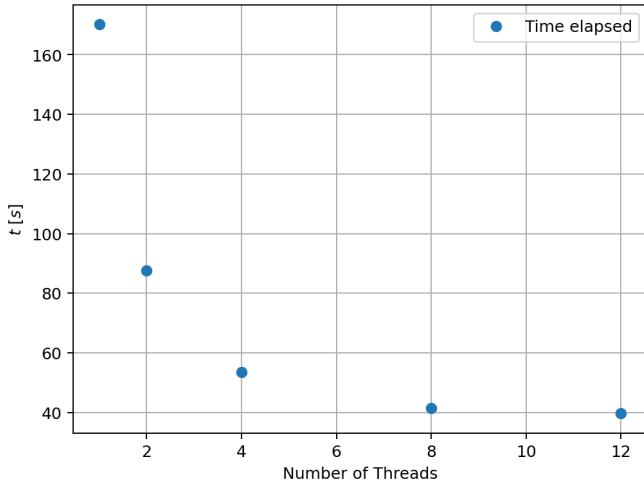


Figure 1. The time elapsed as a function of the number of threads for 100 particles in a 3-dimensional space and  $2^{21}$  MC-cycles.

Fig. 1 shows the time duration of an entire VMC brute force analysis,  $\alpha \in [0.2, 0.95]$ . Using the analytical solution with 100 particles in a 3-dimensional space and  $2^{21}$  Monte Carlo cycles. The analysis was done on a PC running Ubuntu equipped with a [Ryzen 5 PRO 4560U](#). This was done to demonstrate how effective parallelization is and why we chose to do the computation over 8 threads. It's worth noting that these results were obtained using the compiler optimization flag `-O3`. Through our testing we found that this yield a significant speedup in terms of runtime. Using the analytical solution for a 2 dimensional system with 10 particles, the time elapsed was 518 s and 64 s *with* and *without* the optimization flag respectively. In other words one can expect an *up to 8 times* speedup by just compiling using the optimization flag `-O3`.

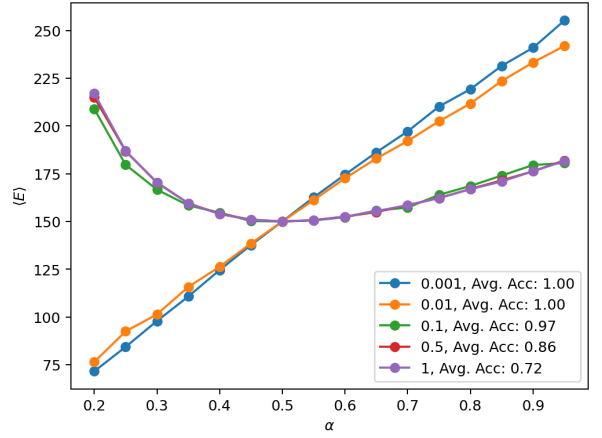


Figure 2. The energy as a function of the Metropolis step length: [0.001, 0.01, 0.1, 0.5, 1], with the average acceptance ratio **Avg. Acc.**, across all  $\alpha$  values.

Fig. 2 represents the energy of the system as a function of  $\alpha$ , along with the average acceptance ratio; both as a function of the **steplength** parameter of each Metropolis step. The acceptance ratio is defined as the number of steps carried over from each Metropolis step per. number of Monte Carlo-cycles. The energy and acceptance ratio were calculated for 100 particles in 3-dimensional space for  $2^{21}$  Monte Carlo cycles. The reason we chose to run with  $2^{21}$  MC-cycles is: 1) it is usual to set the number of MC-cycles to a relatively high number in the realm of one hundred thousand to two million, 2) the blocking technique requires data of sample size  $2^n$  to yield reliable results and 3) to study how the Metropolis step length effects the results.

As can be seen from fig. 2 the best results are obtained by setting **steplength**  $\geq 0.1$ . Setting **steplength** to less than that resulted in a stagnation of the optimization process of the Metropolis step, since the change asserted by the Metropolis step becomes minimal. Thus, further in our analysis we ran the computation over 8 threads and the Metropolis step length set to **steplength** = 1.

The analysis in fig. 3 was done for  $2^{21}$  Monte Carlo cycles. The figure represents the 3D-case for a system with 1, 10 and 100 particles respectively. The figures for the 1D and 2D case can be found in fig. 12 in **VI. Appendix iii**.



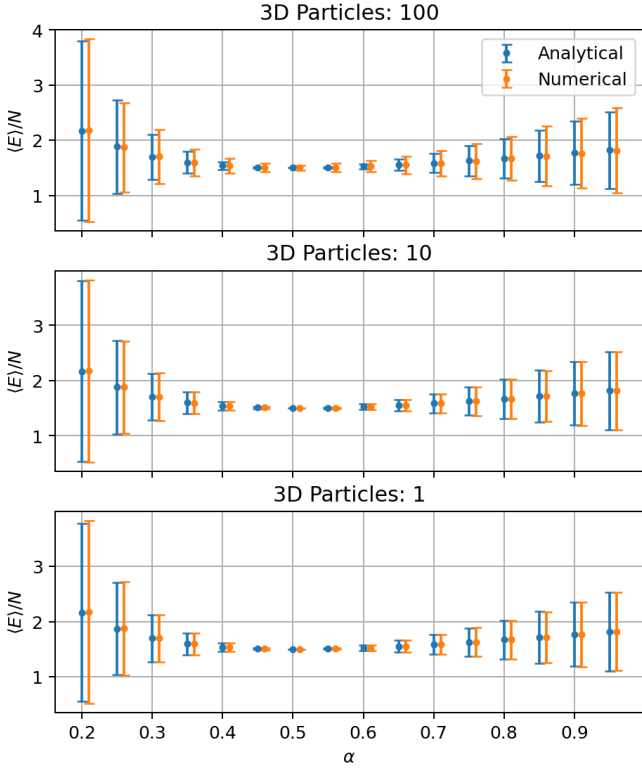


Figure 3. The energy expectation value per number of particles  $N$ , as a function of  $\alpha \in [0.2, 0.95]$  along with the standardized variance (per particle) as an errorbar, in three dimensions and for  $2^{21}$  MC-cycles. The numerical values are slightly shifted for better visibility. The numerical differentiation was done using  $h = 1e-06$ .

The calculation has been done for a value of  $\alpha \in [0.2, 0.95]$  for both the analytical expression, eq. 14, and the numerical differentiation of the Laplacian. The graphs show that  $\alpha = 0.5$  yields the lowest energy and that it corresponds to a true ground state energy, since the variance of the analytical solution is  $\sigma^2 = 0$ , see tab. I. And as for the numerical solution,  $\sigma^2 \approx 0$ , for 1 and 10 particle(s) and is relatively small for 100 particles. Fig. 3 also shows that the ground state for  $\alpha = 0.5$  follows eq. 16. Where for the 3D-case shown here  $\langle E_{1p} \rangle = 1.5$ ,  $\langle E_{10p} \rangle = 15$  and  $\langle E_{100p} \rangle = 150$ .

Table I. The energy expectation values and their corresponding variance for both the analytical and the numerical case. The values are for the 3-dimensional Simple Gaussian in it's ground state, i.e.  $\alpha = 0.5$ .

N	$\langle E \rangle_{Ana}$	$\sigma^2_{Ana}$	$\langle E \rangle_{Num}$	$\sigma^2_{Num}$
1	1.50	0.00	1.50	2.90e-07
10	15.00	0.00	15.00	5.64e-04
100	150.00	0.00	149.99	4.40

Table II. The CPU runtime for the *Brute Force* VMC of the Simple Gaussian, Analytical vs. Numerical. The time period shown is for the entire variational analysis, i.e.  $\alpha \in [0.2, 0.95]$ .

N	1D $t_{CPU}$		2D $t_{CPU}$		3D $t_{CPU}$	
	Analytical	Numerical	Analytical	Numerical	Analytical	Numerical
1	1.26	2.10	1.97	4.01	1.56	3.68
10	4.38	31.24	6.46	64.52	4.85	97.39
100	35.77	2195.92	36.36	4472.73	43.86	7669.45

The time it took for an entire VMC brute force analysis for  $N$  particles in an  $X$ -dimensional space over 8 threads is shown in tab. II. Clearly solving the problem numerically has a toll on our CPU-time. The reason for the overhead is the fact that the numerical double derivative increases the number of FLOPS significantly. From the derivation in **VI. Appendix i** we can see that in order to estimate the double derivative for all particles we have to reevaluate the wave function  $2*d*N$  times.

### Importance Sampling

Following the same form of analysis we seek to determine the best value for  $\Delta t$  (**timestep**) used to adjust the position in each Metropolis-Hastings step, as seen in eq. 39. This was done with the same preset: 100 particles in 3-dimensional space for  $2^{21}$  Monte Carlo cycles.

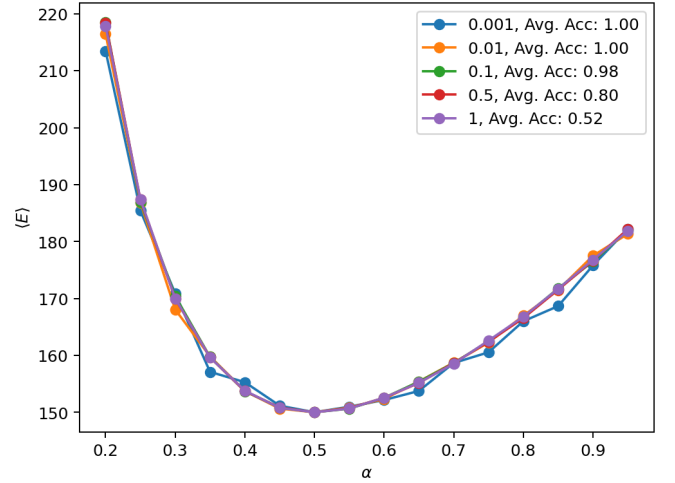


Figure 4. The energy as a function of the importance timestep: [0.001, 0.01, 0.1, 0.5, 1], with the average acceptance ratio Avg. Acc, across all  $\alpha$  values.

Fig. 4 represents the energy of the system as a function of  $\alpha$ , along with the average acceptance ratio; both as a function of the **timestep** parameter of each Metropolis-Hastings step. As can be observed the result for importance sampling differs a lot from the prior brute-force approach. Better convergence is achieved with smaller values, such as **timestep** = 0.01 *contra* the result for **steplength** = 0.01. The acceptance ratio as a function of **timestep** is on a par with the brute-force approach. It's also worth noting for **timestep** and **steplength** set to 1,

although the average acceptance ratio is lower,  $\text{Avg. Acc} = 0.52$ , than that of the brute force method,  $\text{Avg. Acc} = 0.72$ , the overall results are much more stable. This is do to the fact that the weighting of the acceptance probability through importance sampling serve as a much better *educated guess* than that of the brute-force.

As can be seen from fig. 4 the best results are obtained by setting `timestep` to 0.01 or 0.1. Thus further in our analysis we ran the computation over 8 threads and the Metropolis-Hastings step set to `timestep` = 0.1. Since, even if the Metropolis-Hastings is better in terms of results; logically it makes no sense to accept each *suggested* change for every step, since it's evident to encounter suggested changes that are suboptimal. And for the case of `timestep` = 0.01 every step is accepted, which we want to avoid.

The analysis in fig. 5 was done for  $2^{21}$  Monte Carlo cycles. The figure represents the 3D-case for a system with 1, 10 and 100 particles respectively. The figures for the 1D and 2D case can be found in fig. 13 in **VI. Appendix iii**.

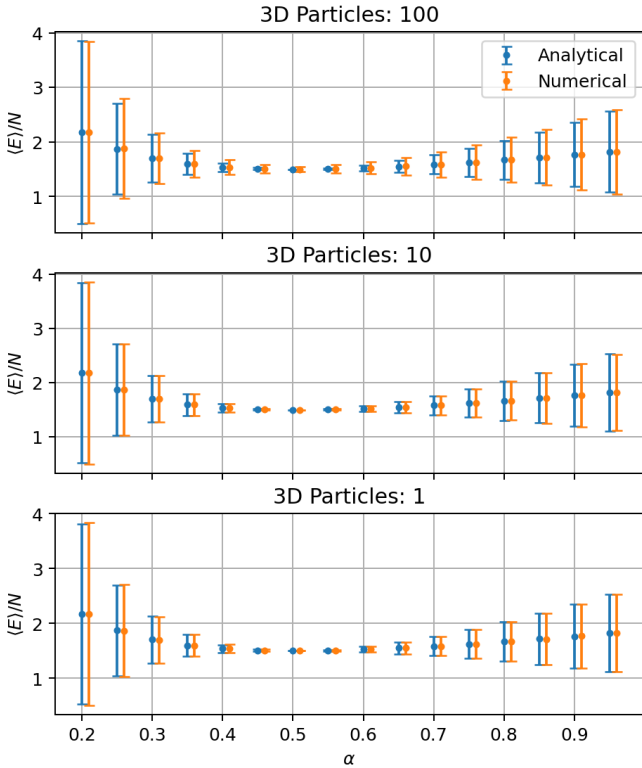


Figure 5. The energy expectation value per number of particles  $N$ , as a function of  $\alpha \in [0.2, 0.95]$  along with the standardized variance (per particle) as an errorbar, in three dimensions and for  $2^{21}$  MC-cycles. The numerical values are slightly shifted for better visibility. The numerical differentiation was done using  $h = 1e-06$ .

Fig. 4 is very similar to that of fig. 3 and follows the same trends. So naturally the conclusion is the same.

Both fig. 3 and fig. 4 show that both approaches yield viable and adequate results.

Table III. The energy expectation values and their corresponding variance for both the analytical and the numerical case. The values are for the 3-dimensional Simple Gaussian in it's ground state, i.e.  $\alpha = 0.5$ .

N	$\langle E \rangle_{Ana}$	$\sigma^2_{Ana}$	$\langle E \rangle_{Num}$	$\sigma^2_{Num}$
1	1.50	0.00	1.50	2.92e-07
10	15.00	0.00	15.00	5.64e-04
100	150.00	0.00	149.99	4.38

Table IV. The CPU runtime for the *Importance* VMC of the Simple Gaussian, Analytical vs. Numerical. The time period shown is for the entire variational analysis, i.e.  $\alpha \in [0.2, 0.95]$ .

N	1D $t_{CPU}$		2D $t_{CPU}$		3D $t_{CPU}$	
	Analytical	Numerical	Analytical	Numerical	Analytical	Numerical
1	3.22	4.22	3.10	4.64	3.76	5.84
10	7.38	34.21	6.30	67.14	7.26	102.23
100	38.90	2342.44	37.81	4839.65	44.37	7860.16

The values in both tab. IV and tab. III falls within the values of their brute-force counterparts. So what does importance sampling bring to the table exactly? We chose to run our analysis with  $2^{21}$  MC-cycles to avoid the shortcomings of the brute-force method; since it requires a high number of cycles to yield reliable and reproducible results. In fig. 2 and fig. 4 we demonstrated the standard Metropolis' reliance on the step-size and how the Metropolis-Hastings algorithm is overall better. But what about the number of MC-cycles? By looking at fig. 2 and fig. 4 we can see that both algorithm gave viable results for `step` = 0.1 with around the same average acceptance ratio. To demonstrate the behavior of both algorithms and show the shortcomings of the brute-force and the merits of the importance sampling we chose to conduct a similar analysis this time by varying the number of MC-cycles with a fixed step-size for both cases, set to `step` = 0.1. The results are shown in fig. 14 in **VI. Appendix iii**.

## Resampling Methods

Our analysis until now have shown that VMC with importance sampling is by far the more reliable approach. So from here on out our estimates are generated using importance sampling with `timestep` = 0.1. We start by analyzing the estimates of the sample variance generated by the bootstrapping and blocking method as a function of the number of MC-cycles. This is done to demonstrate the poor scalability of that of bootstrapping and why we chose to move further with blocking.

The bootstrapping was done such as for each bootstrap operation (`boot`) 0.1% of the data was drawn to calculate the variance,  $\sigma^2_{Boot}$ . The analysis was repeated

$2^n$  times (as many as we have data points: number of MC-cycles); in other words the number of boots equals the number of MC-cycles.

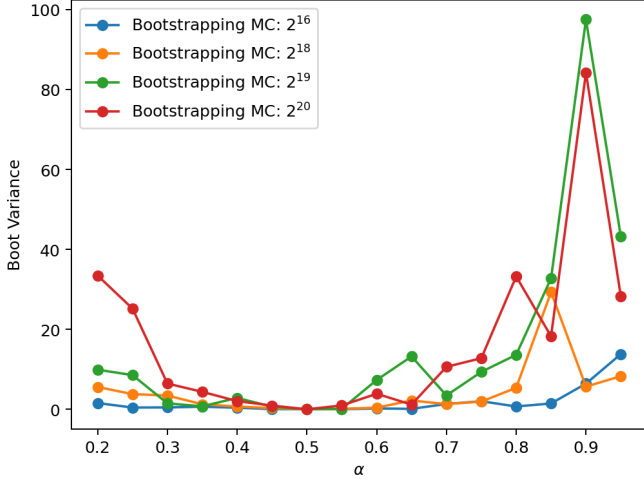


Figure 6. The boot variance as a function of the number of MC-cycles:  $[2^{16}, 2^{18}, 2^{19}, 2^{20}]$ , across all  $\alpha$  values.

As for blocking the C++-code provided by Jonsson [8] was used to calculate the blocking variance,  $\sigma_{Block}^2$ .

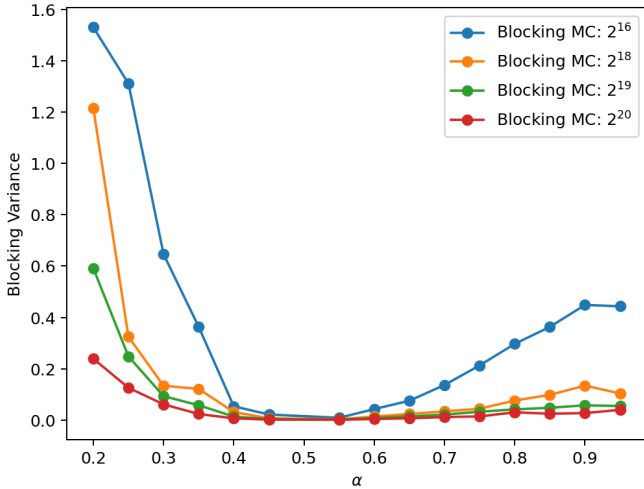


Figure 7. The blocking variance as a function of the number of MC-cycles:  $[2^{16}, 2^{18}, 2^{19}, 2^{20}]$ , across all  $\alpha$  values.

It's clear from fig. 6 that the bootstrapping method scales poorly with the number of data points. Thus, bootstrapping fails to be a good estimate of the sample variance as the sample size increases. Blocking on the other hand yields much more adequate results, fig. 7, and fares much better when there's more data to work with. Fig. 15 in **VI. Appendix iii** shows the variance of the quantum system  $\sigma^2$  as a function of  $\alpha$ . Which is what we have used so far as a metric to determine the ground state. It's clear from the analysis that  $\sigma^2$  is only of importance when

either a ground state is reached or nearing and does not represent the sample error.

## ii. The Full Wave Function

### Gradient Descent & Ground State Energy

Using the information we gathered so far we venture off to the correlated case. Where the Jastrow factor is in effect together with the scaled Hamiltonian, eq. 26, with the parameters  $\beta = \gamma = \sqrt{8}$  moving to an elliptical trap potential and hard sphere radius  $a = 0.0043$ . Using a gradient descent approach we search for the value of the variational parameter  $\alpha$  that hopefully correspond to the ground state energy, this is done by following eq. 42 and updating  $\alpha$  accordingly at the end of each VMC-simulation.

The gradient descent search was conducted using importance sampling with `timestep` = 0.1. The setup for the GD-search is as follows: 3 VMC-rounds with 20, 10 and 10 max number of iterations, respectively. A tolerance  $\epsilon = 1e-5$  acting as a stopper if  $\alpha$  converges. The VMC-rounds were calculated by first defining a starting point  $\alpha_0$  and conducting the computation for a number of MC-cycles:  $[2^{17}, 2^{18}, 2^{18}]$  with a learning rate:  $[0.001, 0.001, 0.0001]$ . This means the max allowed number of iterations is 40 if  $\alpha$  never converges.

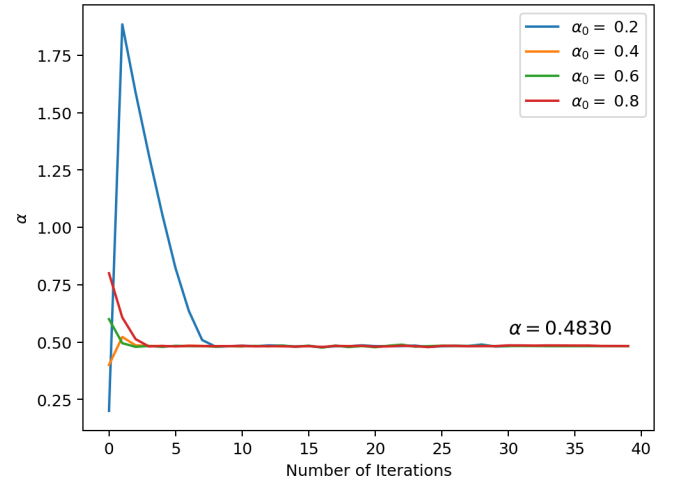


Figure 8. Gradient descent results for the 3 dimensional 100 particles case with  $\alpha_0 = 0.2, 0.4, 0.6, 0.8$  as initial values. The value of  $\alpha$  corresponding to the lowest energy is highlighted in the plot.

Fig. 8 shows which value of the variational parameter  $\alpha$  gradient descent eventually converges to. Our choice of the initial values  $\alpha_0$  was made with the intention of demonstrating that GD converges towards a certain value of  $\alpha$  no matter the initial value. The choice is mainly driven by the fact that  $\alpha = 0.5$  resulted in the ground state energy for the simple Gaussian and the starting points are either a distance 0.3 or 0.1 away on opposite ends. In addition, fig. 16 in **VI. Appendix iii** supports

our assumption that a value near  $\alpha = 0.5$  must correspond to the ground state energy. We did the same type of analysis for 10 and 50 particle. The gradient descent results can be found in fig. 17 in **VI. Appendix iii**.

Table V. The energy expectation values and their corresponding quantum system variance, blocking variance and acceptance ratio. The values are for the 3-dimensional interacting case with 100 particles.

$\alpha$	$\langle E \rangle$	$\sigma^2$	$\sigma_B^2$	Acc.
0.3	296.98	150.32	8.60e-02	9.75e-01
0.5	266.54	1.17	1.04e-03	9.53e-01
0.7	284.94	74.63	2.99e-02	9.29e-01

Tab. V reflects our findings from our resampling methods analysis fig. 7 and fig. 15 in **VI. Appendix iii**. The quantum variance  $\sigma^2$  hints that the ground state lies near  $\alpha = 0.5$  and the blocking variance  $\sigma_B^2$  shows our analysis driven choices for optimal parameters yield viable results. The values you see here are a product of fig. 16 in **VI. Appendix iii** and were run for  $2^{20}$  MC-cycles to insure adequate results.

### One-Body Density

Our gradient descent analysis fig. 8 shows that

$$\alpha_{min} = 0.483 \pm 0.0002 \quad (50)$$

is the value of  $\alpha$  that yielded the lowest energy.

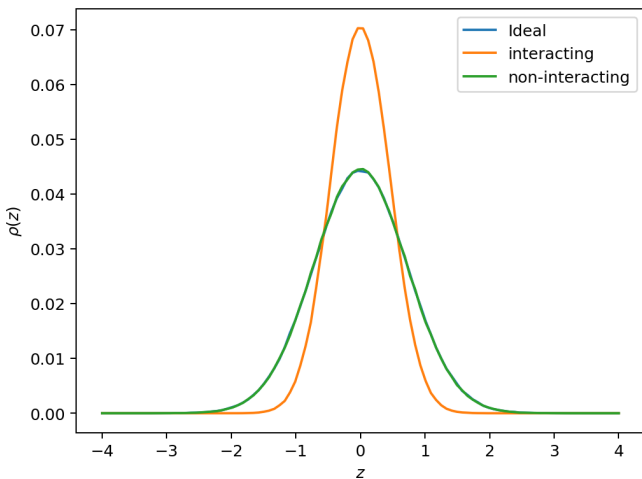


Figure 9. One-body density along the z-axis of the ideal, interacting and non-interacting case for 100 particles. Where **ideal**: Simple Gaussian, **interacting**: Correlated with  $\beta = \gamma = \sqrt{8}$  and  $a = 0.0043$ , **non-interacting**: correlated with  $\beta = \gamma = 1$  and  $a = 0$ .

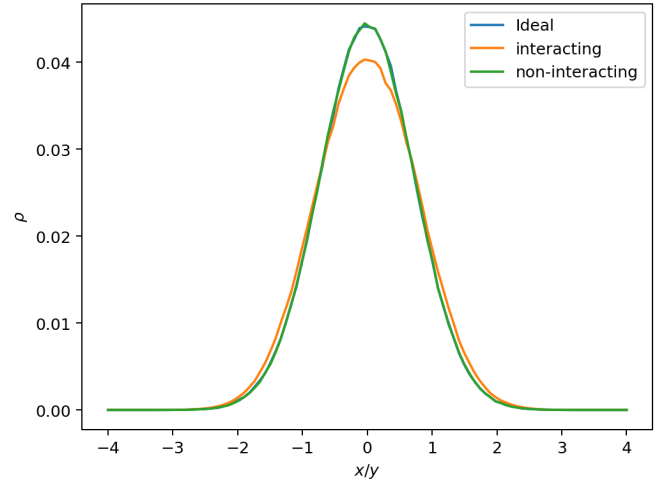


Figure 10. One-body density along the x-/y-axis of the ideal, interacting and non-interacting case for 100 particles.

Fig. 9 shows the distribution of particles along the z-axis it's clear that the elliptical parameter  $\gamma$  in eq. 26 skews the particles towards the center, which is why we observe a higher density for the interacting case. Fig. 10 on the other hand represents the distribution of particles along either the x- or y-axis. Here for the interacting case the particles are pushed away from the center, which is why the distribution favors larger values. This is due to both the circular nature of the trap (it's wider than the elliptical case) and the repulsive nature of the interaction between particles.

As for the ideal vs. the non-interacting case in fig. 9 and 10 the results are the same; which is to be expected. Reshaping the trap to a spherical trap by setting  $\gamma = 1$  and nullifying the boson-to-boson interaction by setting the hard-sphere diameter  $a = 0$  should yield back the same distribution as the ideal case.

The one-body density analysis for 10 and 50 particles can be found in fig. 18 and fig. 19 in **VI. Appendix iii**. It's worth noting that along the z-axis we will get a more concentrated distribution no matter the number of particles in the system. That is due to the elliptical nature of the trap along the z-axis. And as for the distribution of particles in the xy-plane we can see that by increasing the number of particles the repulsive interaction start to take effect.

### The Jastrow Factor

In order to study the effect of the Jastrow factor we compared the interacting case ( $a = 0.0043$ ) to the non-interacting case ( $a = 0$ ) in an elliptical trap ( $\beta = \gamma = \sqrt{8}$ ).

Table VI. The energy expectation values and their corresponding blocking variance for  $\alpha = 0.5$  (the ground-state of the ideal case). For both the interacting and non-interacting case.

N	Interacting		Non-interacting	
	$\langle E \rangle$	$\sigma_B^2$	$\langle E \rangle$	$\sigma_B^2$
10	24.40	1.79e-07	24.14	0
50	127.35	7.32e-05	120.71	0
100	266.54	1.04e-03	241.42	0

In an elliptical trap the energy no longer follows eq. 16 and thus the energy of the system is no longer directly linearly dependent on the number of particles and dimensions. For the non-interacting 3-dimensional case the energy follows the formula

$$\langle E \rangle = \frac{2 \cdot N + 1 \cdot N \cdot \sqrt{\gamma}}{2} \quad (51)$$

Tab. VI shows how the number of particles effects the total energy of the system for both the interacting and non-interacting case. The interaction between particles governed by the Jastrow factor had an increasing impact on the total energy of the system. This is due to the nature of the Jastrow factor within the correlated system. Given that two particles are within a certain distance of each other the repulsive effect, see fig. 10, between the two contributes to the increase in the system's local energy. This explains the gradual increase in the system's energy as a function of the number of particles. Since a more *densely populated* system equates to more frequent pair interactions.

## V. CONCLUSION

### i. The Simple Gaussian

The simple Gaussian, or the ideal case, although on its own it doesn't bring much value for theoretical quantum physics studies. It's a great tool in and of itself and acts as a validation gauge for the work we have implemented. With it we managed to determine that our numerical solution for the double derivative in the Hamiltonian is indeed viable, fig. 3 and 5. Since, having an analytical answer meant that we knew the behavior of the quantum variance when our trial function is an eigenvalue of the Hamiltonian ( $\sigma^2 = 0$ ) and as we observed the numerical results behaved in a similar way. Making it a viable approach for finding  $\nabla^2 \Psi_T$  when analytical expressions don't present themselves, which is often the case in quantum mechanics. But numeric differentiation usually comes at a cost of computation time, tab. II and tab. IV. Where one can expect several ours of computation to be added for very complex systems, in our case a 100 particles in a 3 dimensional trap took upwards of 2 hours with numerical differentiation.

In addition, the simple Gaussian made it possible for us to conduct a thorough analysis of both VMC- implementations, both brute-force sampling and importance sampling. We managed to determine which method is overall the better approach and that is importance sampling. Subsequently, it allowed us to study how tangible parameter like the step-size and number of MC-cycles can effect the validity of the results and to which degree. Effectively gifting us the optimal parameters to work with for the interacting case.

### ii. The Full Wave Function

In a spherical trap we observed that both the ideal case and the non-interacting case are clearly the same, see fig. 18 and fig. 19. The energy in this system scales proportionally with the number of particles and dimensions. This ultimately means that the energy per-particle remains the same no matter the configuration, as can be seen in fig. 13.

As for the elliptical trap through observing the results in tab. VI and some testing on our end we predicted that the local energy for the non-interacting case follows eq. 51, which seems very plausible. Since, the elliptical parameter  $\gamma$  only changes the shape of the trap along the z-axis. Thus, it makes sense that the energy's relationship to the number of dimensions and particles should change along one of these dimensions, namely the z-axis. In addition we saw the impact the Jastrow factor had on the local energy. Clearly, interaction plays a part in the system and especially for a relatively high number of particles as can be seen from tab. VI and fig. 16. When interaction between particles was taken under consideration the system's local energy increased by  $\approx 10\%$  for 100 particles.

Changing the shape of the trap changes the particles distributions as can be seen in fig. 9. The elliptical shape meant that particles got narrowly pushed together. But since the change was only applied along the z-axis we managed to observe the effect of the repulsive interaction between particles, fig. 10. The repulsive interaction governed by the Jastrow factor made it so particles were less likely to be within each other's vicinity and thus more spread out; increasing the system's local energy.

### iii. Summary

Clearly, VMC is a viable approach to estimating the ground state energy  $E_0$ , especially when the Metropolis-Hastings step is embedded into the work. Since, importance sampling proved itself to be very reliable in terms of the tangible parameters. In addition, one can always chose the numerical differentiation route for estimating the double derivative of the Hamiltonian. Only if that's the case, then we advise the reader to give it a little more thought than we have. Since, relying on such a solution opens the door for better optimizations and parallelization of the code. Which we did to a degree in our work but felt the need to mention. Since, our work with nu-

merical differentiation does not reflect what one should typically expect from better implementations of the finite difference method.

Resampling methods are vital for numerical approaches such as this. Automated blocking proved to be an essential tool in estimating the sample error of the numerical model. It was used to ensure that we are on the right path in terms of our numerical approach and that the results are indeed viable. As an example, with blocking we managed to confirm that the results obtained by the GD-search of the variational parameter  $\alpha$  can be trusted, eq. 50.

As for gradient descent our very basic implementation that relies on the essential ideas of how a gradient descent search works, needs to be changed. Better approaches such as stochastic gradient descent with mini-batches or preferably gradient descent with optimization algorithms such as Adagrad or Adam need to be considered. Especially, when working with many variational parameters and more complex quantum systems than we have. Our implementation is adequate for systems of Bose-Einstein Condensation in gases of alkali atoms such as  $^{87}\text{Rb}$ , with one variational parameter. But since both  $\gamma$  and  $\beta$  are by their very nature variational parameters, better and well-rounded gradient descent schemes need to be implemented in the latter case. This is under the assumption that the condition for diluteness is satisfied. Since, our model only considers the interaction between a pair of atoms. We cannot claim that our implementation is viable for denser systems that go beyond a dilute state.

## REFERENCES

- [1] J.L. Dubois and H. R. Glyde. Bose-Einstein condensation in trapped bosons: A variational Monte Carlo analysis. *Phys. Rev. A* **63**, 023602 (Jan 2001)
- [2] J.K. Nilsen, J. Mur-Petit, M. Guilleumas, M. Hjort-Jensen and A. Polls. Vortices in atomic Bose-Einstein condensates in the large-gas-parameter region. *Phys. Rev. A*, **71**, 053610 (May 2005).
- [3] A. Fabrocini and A. Polls. Beyond Gross-Pitaevskii: local density vs. correlated basis approach for trapped bosons *Phys. Rev. A* **60**, 2319 (Jan 1999)
- [4] M. Hjort-Jensen. Computational Physics II: Lecture notes on Building a Variational Monte Carlo program, Metropolis Algorithm and Markov Chains. Lecture Notes, (2022) <http://compphysics.github.io/ComputationalPhysics2/doc/pub/week2/html/week2-reveal.html> (accessed 05.04)
- [5] M. Hjort-Jensen. Computational Physics II: Lecture notes on Importance Sampling, Fokker-Planck and Langevin equation. Lecture Notes, (2022) <http://compphysics.github.io/ComputationalPhysics2/doc/pub/week4/html/week4-reveal.html> (accessed 05.04)
- [6] M. Hjort-Jensen. Computational Physics II: Lecture notes on Gradient Methods Stochastic Gradient Descent. Lecture Notes, (2022) <http://compphysics.github.io/ComputationalPhysics2/doc/pub/week6/html/week6-reveal.html> (accessed 05.04)
- [7] M. Hjort-Jensen. Computational Physics II: Lecture notes on Resampling Methods. Lecture Notes, (2022) <http://compphysics.github.io/ComputationalPhysics2/doc/pub/statanalysis/html/statanalysis.html>
- [8] M. Jonsson. Standard error estimation by an automated blocking method. *Phys. Rev. E*, **98**, 043304 (Oct 2018) <https://link.aps.org/doi/10.1103/PhysRevE.98.043304> (accessed 05.04)
- [9] Flyvbjerg, H. and Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **91**, 461-466 (1989) <https://doi.org/10.1063/1.457480> (accessed 05.04)
- [10] M. N. Nordhagen. Studies of quantum dots using machine learning. *Master Thesis*, University of Oslo (2019) <https://www.duo.uio.no/handle/10852/73753> (accessed 05.04)

## VI. APPENDIX

### i. Calculations for non-interacting bosons

$$\begin{aligned} E_L(\mathbf{r}) &= \frac{1}{\Psi_T(\mathbf{r})} H \Psi_T(\mathbf{r}) \\ &= \frac{1}{\Psi_T(\mathbf{r})} \left[ \sum_i^N \left( -\frac{\hbar^2}{2m} \nabla_i^2 + V_{ext}(\mathbf{r}_i) \right) \right] \Psi_T(\mathbf{r}) \end{aligned}$$

Starting with the  $H \Psi_T(\mathbf{r})$  part of the equation

$$\begin{aligned} H \Psi_T(\mathbf{r}) &= \left[ \sum_i^N \left( -\frac{\hbar^2}{2m} \nabla_i^2 + V_{ext}(\mathbf{r}_i) \right) \right] \Psi_T(\mathbf{r}) \\ &= \sum_i^N \left( -\frac{\hbar^2}{2m} \nabla_i^2 \Psi_T(\mathbf{r}) + V_{ext}(\mathbf{r}_i) \Psi_T(\mathbf{r}) \right) \end{aligned} \quad (52)$$

Solving  $\sum_i^N -\frac{\hbar^2}{2m} \nabla_i^2 \Psi_T$ , the sum over  $i$  for the Laplacian makes it so we compute the second derivative of the wave function with respect to each particle, for a particle  $k$  the second derivatives is expressed as

$$\nabla_k^2 \prod_i e^{-\alpha |\mathbf{r}_i|^2} = \nabla_k^2 \prod_i e^{-\alpha r_i^2} \quad (53)$$

where  $k$  is an element in  $\sum_i^N$ .

$$\prod_{i \neq k} e^{-\alpha |\mathbf{r}_i|^2} \nabla_k^2 e^{-\alpha |\mathbf{r}_k|^2} = \prod_{i \neq k} e^{-\alpha |\mathbf{r}_i|^2} \nabla_k^2 e^{-\alpha (x_k^2 + y_k^2 + z_k^2)} \quad (54)$$

We start by finding the general expression for the gradient ( $\nabla$ )

$$\begin{aligned} \nabla_k e^{-\alpha (x_k^2 + y_k^2 + z_k^2)} &= \left[ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right] e^{-\alpha (x_k^2 + y_k^2 + z_k^2)} \\ &= \left[ -2\alpha x_k e^{-\alpha |\mathbf{r}_k|^2}, -2\alpha y_k e^{-\alpha |\mathbf{r}_k|^2}, -2\alpha z_k e^{-\alpha |\mathbf{r}_k|^2} \right] \\ &= -2\alpha [x_k, y_k, z_k] e^{-\alpha |\mathbf{r}_k|^2} = -2\alpha \mathbf{r}_k e^{-\alpha |\mathbf{r}_k|^2} \end{aligned} \quad (55)$$

Bringing back the product sequence we get

$$\prod_{i \neq k} e^{-\alpha |\mathbf{r}_i|^2} \nabla_k e^{-\alpha |\mathbf{r}_k|^2} = -2\alpha \mathbf{r}_k \Psi_T(\mathbf{r}) \quad (56)$$

Now for the Laplacian ( $\nabla^2$ )

$$\begin{aligned} \nabla_k^2 e^{-\alpha |\mathbf{r}_k|^2} &= -2\alpha \nabla_k \mathbf{r}_k e^{-\alpha |\mathbf{r}_k|^2} = -2\alpha \left[ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right] [x_k, y_k, z_k] e^{-\alpha |\mathbf{r}_k|^2} \\ &= -2\alpha \frac{\partial}{\partial x} x_k e^{-\alpha |\mathbf{r}_k|^2} - 2\alpha \frac{\partial}{\partial y} y_k e^{-\alpha |\mathbf{r}_k|^2} - 2\alpha \frac{\partial}{\partial z} z_k e^{-\alpha |\mathbf{r}_k|^2} \\ &= -2\alpha \left( e^{-\alpha |\mathbf{r}_k|^2} - 2\alpha x_k^2 e^{-\alpha |\mathbf{r}_k|^2} \right) \\ &\quad - 2\alpha \left( e^{-\alpha |\mathbf{r}_k|^2} - 2\alpha y_k^2 e^{-\alpha |\mathbf{r}_k|^2} \right) \\ &\quad - 2\alpha \left( e^{-\alpha |\mathbf{r}_k|^2} - 2\alpha z_k^2 e^{-\alpha |\mathbf{r}_k|^2} \right) \\ &= -2\alpha \left[ (1 - 2\alpha x_k^2) + (1 - 2\alpha y_k^2) + (1 - 2\alpha z_k^2) \right] e^{-\alpha |\mathbf{r}_k|^2} \\ &= -2\alpha [3 - 2\alpha |\mathbf{r}_k|^2] e^{-\alpha |\mathbf{r}_k|^2} \end{aligned} \quad (57)$$



for a general  $d$ -dimensional case our expression then becomes

$$= -2\alpha \left[ d - 2\alpha |\mathbf{r}_k|^2 \right] e^{-\alpha |\mathbf{r}_k|^2} = \left[ -2\alpha d + 4\alpha^2 r_k^2 \right] e^{-\alpha |\mathbf{r}_k|^2} \quad (58)$$

where  $d$  is the number of dimension, 3 in our case. Bringing back the product sequence we get

$$\prod_{i \neq k} e^{-\alpha |\mathbf{r}_i|^2} \left[ -2\alpha d + 4\alpha^2 r_k^2 \right] e^{-\alpha |\mathbf{r}_k|^2} = \left[ -2\alpha d + 4\alpha^2 r_k^2 \right] \Psi_T(\mathbf{r}) \quad (59)$$

The final analytical expression for  $H\Psi_T(\mathbf{r})$  is

$$\begin{aligned} H\Psi_T(\mathbf{r}) &= \sum_i^N \left( -\frac{\hbar^2}{2m} \nabla_i^2 \Psi_T(\mathbf{r}) + V_{ext}(\mathbf{r}_i) \Psi_T(\mathbf{r}) \right) \\ &= \sum_i^N \left( -\frac{\hbar^2}{2m} \left[ -2\alpha d + 4\alpha^2 r_i^2 \right] \Psi_T(\mathbf{r}) + V_{ext}(\mathbf{r}_i) \Psi_T(\mathbf{r}) \right) \\ &= \sum_i^N \left( -\frac{\hbar^2}{2m} \left[ -2\alpha d + 4\alpha^2 r_i^2 \right] + V_{ext}(\mathbf{r}_i) \right) \Psi_T(\mathbf{r}) \end{aligned} \quad (60)$$

Using what we've obtained, the final expression for the local energy is then

$$\begin{aligned} E_L(\mathbf{r}) &= \frac{1}{\Psi_T(\mathbf{r})} H\Psi_T(\mathbf{r}) \\ &= \frac{1}{\Psi_T(\mathbf{r})} \left[ \sum_i^N \left( -\frac{\hbar^2}{2m} \left[ -2\alpha d + 4\alpha^2 r_i^2 \right] + V_{ext}(\mathbf{r}_i) \right) \right] \Psi_T(\mathbf{r}) \\ &= \sum_i^N \left( -\frac{\hbar^2}{2m} \left[ -2\alpha d + 4\alpha^2 r_i^2 \right] + V_{ext}(\mathbf{r}_i) \right) \end{aligned} \quad (61)$$

For the case of a spherical potential with  $\beta = 1$ ,  $\omega_{ho} = 1$

$$\begin{aligned} E_L(\mathbf{r}) &= \sum_{i=1}^N \left( -\frac{\hbar^2}{2m} \left[ -2\alpha d + 4\alpha^2 r_i^2 \right] + \frac{1}{2} m \omega_{ho}^2 r_i^2 \right) \\ &= \frac{-\hbar^2}{2m} \left[ -2\alpha N d + 4\alpha^2 \sum_{i=1}^N r_i^2 \right] + \frac{1}{2} m \omega_{ho}^2 \sum_{i=1}^N r_i^2 \end{aligned} \quad (62)$$

If we were to work with natural units, meaning  $\hbar = m = 1$

$$\begin{aligned} &= \frac{-1}{2} \left[ -2\alpha N d + 4\alpha^2 \sum_{i=1}^N r_i^2 \right] + \frac{1}{2} \omega_{ho}^2 \sum_{i=1}^N r_i^2 \\ &= \alpha N d - 2\alpha^2 \sum_{i=1}^N r_i^2 + \frac{1}{2} \omega_{ho}^2 \sum_{i=1}^N r_i^2 \end{aligned} \quad (63)$$

Choosing to set the oscillator frequency  $\omega_{ho} = 1$  the ground state occurs for  $\alpha = 0.5$ . Which in turn yields a simple analytical expression

$$= \frac{1}{2} N d - \frac{1}{2} \sum_{i=1}^N r_i^2 + \frac{1}{2} \sum_{i=1}^N r_i^2 = \frac{Nd}{2} \quad (64)$$

which we would use to test our numerical solution for the Hamiltonian.

Finding an analytical expression for the drift force

$$\begin{aligned}
F &= \frac{2\nabla\Psi_T}{\Psi_T} \\
\nabla\Psi_T &= \sum_i \nabla_i \Psi_T = -2\alpha \sum_i \mathbf{r}_i \Psi_T \\
F &= \frac{-4\alpha \sum_i \mathbf{r}_i \Psi_T}{\Psi_T} = -4\alpha \sum_i \mathbf{r}_i \\
F_k &= -4\alpha \mathbf{r}_k
\end{aligned} \tag{65}$$

Finding an expression for the numerical second derivative. We start with the second-order central derivative

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \tag{66}$$

Wave function second derivative

$$\nabla^2 \Psi_T(\mathbf{r}) = \sum_i^N \nabla_i^2 \Psi_T(\mathbf{r}) \tag{67}$$

for a particle  $k$  in a 3 dimensional case

$$\prod_{i \neq k} e^{-\alpha|\mathbf{r}_i|^2} \nabla_k^2 e^{-\alpha|\mathbf{r}_k|^2} = \prod_{i \neq k} e^{-\alpha|\mathbf{r}_i|^2} \left( \frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right) e^{-\alpha(x_k^2 + y_k^2 + z_k^2)} \tag{68}$$

$$\frac{\partial^2}{\partial x_k^2} e^{-\alpha(x_k^2 + y_k^2 + z_k^2)} = \frac{e^{-\alpha((x_k+h)^2 + y_k^2 + z_k^2)} - 2e^{-\alpha(x_k^2 + y_k^2 + z_k^2)} + e^{-\alpha((x_k-h)^2 + y_k^2 + z_k^2)}}{h^2} \tag{69}$$

$$\begin{aligned}
\left( \frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right) e^{-\alpha(x_k^2 + y_k^2 + z_k^2)} &= \left[ -2 \cdot 3e^{-\alpha(x_k^2 + y_k^2 + z_k^2)} + e^{-\alpha((x_k+h)^2 + y_k^2 + z_k^2)} + e^{-\alpha((x_k-h)^2 + y_k^2 + z_k^2)} \right. \\
&\quad + e^{-\alpha(x_k^2 + (y_k+h)^2 + z_k^2)} + e^{-\alpha(x_k^2 + (y_k-h)^2 + z_k^2)} \\
&\quad \left. + e^{-\alpha(x_k^2 + y_k^2 + (z_k+h)^2)} + e^{-\alpha(x_k^2 + y_k^2 + (z_k-h)^2)} \right] \cdot \frac{1}{h^2}
\end{aligned} \tag{70}$$

$$\begin{aligned}
\prod_{i \neq k} e^{-\alpha|\mathbf{r}_i|^2} \nabla_k^2 e^{-\alpha|\mathbf{r}_k|^2} &= \frac{-2 \cdot 3\Psi_T(\mathbf{r})}{h^2} + \prod_{i \neq k} e^{-\alpha|\mathbf{r}_i|^2} \left[ e^{-\alpha((x_k+h)^2 + y_k^2 + z_k^2)} + e^{-\alpha((x_k-h)^2 + y_k^2 + z_k^2)} \right. \\
&\quad + e^{-\alpha(x_k^2 + (y_k+h)^2 + z_k^2)} + e^{-\alpha(x_k^2 + (y_k-h)^2 + z_k^2)} \\
&\quad \left. + e^{-\alpha(x_k^2 + y_k^2 + (z_k+h)^2)} + e^{-\alpha(x_k^2 + y_k^2 + (z_k-h)^2)} \right] \cdot \frac{1}{h^2}
\end{aligned} \tag{71}$$

$$\begin{aligned}
\prod_{i \neq k} e^{-\alpha|\mathbf{r}_i|^2} \nabla_k^2 e^{-\alpha|\mathbf{r}_k|^2} &= \left[ -2 \cdot 3\Psi_T(\mathbf{r}) + \Psi(r_k: x_k + h) + \Psi(r_k: x_k - h) \right. \\
&\quad + \Psi(r_k: y_k + h) + \Psi(r_k: y_k - h) \\
&\quad \left. + \Psi(r_k: z_k + h) + \Psi(r_k: z_k - h) \right] \cdot \frac{1}{h^2}
\end{aligned} \tag{72}$$

**Not the best notation but I had to make lemonades out of L<sup>A</sup>T<sub>E</sub>X!**

Setting up for Gradient descent:

$$\begin{aligned}
\bar{\Psi}_T &= \frac{d\Psi_T}{d\alpha} = \frac{d}{d\alpha} \prod_i e^{-\alpha r_i^2} = \frac{d}{d\alpha} e^{-\alpha \sum_i r_i^2} = - \sum_i r_i^2 \cdot e^{-\alpha \sum_i r_i^2} \\
\frac{\bar{\Psi}_T}{\Psi_T} &= \frac{- \sum_i r_i^2 \cdot e^{-\alpha \sum_i r_i^2}}{e^{-\alpha \sum_i r_i^2}} = - \sum_i r_i^2
\end{aligned} \tag{73}$$

## ii. Calculations for interacting bosons

### Scaling the Hamiltonian

$$H = \sum_i^N \left( \frac{-\hbar^2}{2m} \nabla_i^2 + V_{ext}(\mathbf{r}) \right) + \sum_{i<j} V_{int}(\mathbf{r}_i, \mathbf{r}_j) \tag{74}$$

$$= \sum_i^N \left( \frac{-\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} m \left[ \omega_{ho}^2 (x_i^2 + y_i^2) + \omega_z^2 z_i^2 \right] \right) + \sum_{i<j} V_{int}(\mathbf{r}_i, \mathbf{r}_j) \tag{75}$$

For convenience we want to express the energy in units of  $\hbar\omega$ . Therefor we introduce a scaling factor  $\hbar\omega_{ho}$

$$= \frac{\hbar\omega_{ho}}{2} \sum_i^N \left( -\frac{\hbar}{m\omega_{ho}} \nabla_i^2 + \frac{m\omega_{ho}}{\hbar} \left[ x_i^2 + y_i^2 + \frac{\omega_z^2}{\omega_{ho}^2} z_i^2 \right] \right) + \sum_{i<j} V_{int}(\mathbf{r}_i, \mathbf{r}_j) \tag{76}$$

Inspecting the  $\hbar/m\omega_{ho}$ -factor we find that its physical SI unit is  $m$ . So naturally this our length scale and we define it as

$$a_{ho} = \sqrt{\hbar/m\omega_{ho}} \tag{77}$$

which helps shape up the scaled lengths

$$\mathbf{r}' = \mathbf{r}/a_{ho} \implies \mathbf{r} = a_{ho}\mathbf{r}' \tag{78}$$

$$a' = a/a_{ho} \implies a = a_{ho}a' \tag{79}$$

and naturally the Laplacian becomes

$$\nabla'^2_i = a_{ho}^2 \nabla_i^2 \implies \nabla_i^2 = \frac{\nabla'^2_i}{a_{ho}^2} \tag{80}$$

We also define the parameter

$$\gamma = \omega_z/\omega_{ho} \tag{81}$$

Dividing the Hamiltonian by the scaling factor  $\hbar\omega_{ho}$ , yields

$$= \frac{1}{2} \sum_i^N \left( -a_{ho}^2 \nabla_i^2 + a_{ho}^{-2} \left[ x_i^2 + y_i^2 + \gamma^2 z_i^2 \right] \right) + \sum_{i<j} V_{int}(\mathbf{r}_i, \mathbf{r}_j) \tag{82}$$

then scaling the Hamiltonian

$$= \frac{1}{2} \sum_i^N \left( -\nabla'^2_i + x'^2_i + y'^2_i + \gamma^2 z'^2_i \right) + \sum_{i<j} V_{int}(\mathbf{r}_i, \mathbf{r}_j) \tag{83}$$

### The Local Energy

$$\begin{aligned}
E_L(\mathbf{r}) &= \frac{1}{\Psi_T(\mathbf{r})} H \Psi_T(\mathbf{r}) \\
H &= \sum_i^N \left( \frac{-\hbar^2}{2m} \nabla_i^2 + V_{ext}(\mathbf{r}) \right) + \sum_{i<j} V_{int}(\mathbf{r}_i, \mathbf{r}_j)
\end{aligned}$$

We start by rewriting eq. 6 as

$$\Psi_T(\mathbf{r}) = \left[ \prod_i \phi(\mathbf{r}_i) \right] \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \quad (84)$$

where  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  and where we have rewritten the one-body part

$$g(\alpha, \beta, \mathbf{r}_i) = \phi(\mathbf{r}_i) \quad (85)$$

and the Jastrow factor

$$f(r_{ij}) = \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \quad (86)$$

for  $u$  is defined as  $u(r_{ij}) = \ln f(r_{ij})$ . To keep things tidy we define  $\Psi_{OB}$  to be the one body-part and  $\Psi_C$  to be the correlated part of eq. 84. The gradient  $\nabla_k$  of a particle  $k$  can be found by solving

$$\nabla_k \Psi_T = \Psi_C \nabla_k \Psi_{OB} + \Psi_{OB} \nabla_k \Psi_C \quad (87)$$

$$\nabla_k \Psi_{OB} = \nabla_k \prod_i \phi(\mathbf{r}_i) = \nabla_k \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \quad (88)$$

$$\nabla_k \Psi_C = \nabla_k \exp \left( \sum_{i < j}^N u(r_{ij}) \right) = \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \quad (89)$$

Where we've used the fact that all the terms  $r_{ij}$  that don't contain the position  $r_k$  of particle  $k$  are zero and we're left with  $r_k$  dependent terms. Thus, the first derivative can now be expressed as

$$\nabla_k \Psi_T = \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \cdot \nabla_k \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \quad (90)$$

The Laplacian

$$\nabla_k^2 \Psi_T = \nabla_k \left[ \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \cdot \prod_{i \neq k} \phi(\mathbf{r}_i) \nabla_k \phi(\mathbf{r}_k) + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \right] \quad (91)$$

Starting with the first term

$$\begin{aligned} \nabla_k \left[ \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \cdot \prod_{i \neq k} \phi(\mathbf{r}_i) \nabla_k \phi(\mathbf{r}_k) \right] &= \nabla_k^2 \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \\ &\quad + \nabla_k \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \nabla_k \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \end{aligned} \quad (92)$$

$$= \nabla_k^2 \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) + \nabla_k \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \quad (93)$$

Now for the second term

$$\begin{aligned}
\nabla_k \left[ \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \right] &= \nabla_k \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \nabla_k \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \nabla_k \sum_{l \neq k} \nabla_k u(r_{kl}) \\
&= \nabla_k \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{m \neq k} \nabla_k u(r_{km}) \cdot \sum_{l \neq k} \nabla_k u(r_{kl}) \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k^2 u(r_{kl})
\end{aligned} \tag{94}$$

In the second term both  $m$  and  $l$  are indices that iterate over the same set, which yields

$$\begin{aligned}
\nabla_k \left[ \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \right] &= \nabla_k \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \left( \sum_{l \neq k} \nabla_k u(r_{kl}) \right)^2 \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k^2 u(r_{kl})
\end{aligned} \tag{95}$$

The entire expression of the Laplacian is given by

$$\begin{aligned}
\nabla_k^2 \Psi_T &= \nabla_k^2 \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) + 2 \nabla_k \phi(\mathbf{r}_k) \prod_{i \neq k} \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k u(r_{kl}) \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \left( \sum_{l \neq k} \nabla_k u(r_{kl}) \right)^2 \\
&\quad + \prod_i \phi(\mathbf{r}_i) \cdot \exp \left( \sum_{i < j}^N u(r_{ij}) \right) \sum_{l \neq k} \nabla_k^2 u(r_{kl})
\end{aligned} \tag{96}$$

The Laplacian-term in the local energy

$$\frac{1}{\Psi_T} \nabla_k^2 \Psi_T = \frac{\nabla_k^2 \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2 \frac{\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} \sum_{l \neq k} \nabla_k u(r_{kl}) + \left( \sum_{l \neq k} \nabla_k u(r_{kl}) \right)^2 + \sum_{l \neq k} \nabla_k^2 u(r_{kl}) \tag{97}$$

Now for the  $\nabla_k \phi$ -term

$$\frac{\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} = \frac{\nabla_k e^{-\alpha \mathbf{r}_k^2}}{\phi(\mathbf{r}_k)} = \frac{-2\alpha \mathbf{r}_k e^{-\alpha \mathbf{r}_k^2}}{\phi(\mathbf{r}_k)} = -2\alpha \mathbf{r}_k \frac{\phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} = -2\alpha \mathbf{r}_k \tag{98}$$

as for the  $\nabla_k^2 \phi$ -term

$$\frac{\nabla_k^2 \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} = -2\alpha \frac{\nabla_k \mathbf{r}_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} = -2\alpha \frac{\nabla_k [x_k, y_k, \beta z_k] e^{-\alpha \mathbf{r}_k^2}}{\phi(\mathbf{r}_k)} \quad (99)$$

$$= -2\alpha \frac{(1+1+\beta)e^{-\alpha \mathbf{r}_k^2} - 2\alpha \mathbf{r}_k^2 e^{-\alpha \mathbf{r}_k^2}}{\phi(\mathbf{r}_k)} = \frac{4\alpha^2 \mathbf{r}_k^2 e^{-\alpha \mathbf{r}_k^2} - 2\alpha(2+\beta)e^{-\alpha \mathbf{r}_k^2}}{\phi(\mathbf{r}_k)} \quad (100)$$

$$= \frac{[4\alpha^2 \mathbf{r}_k^2 - 2\alpha(2+\beta)]\phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} = 4\alpha^2 \mathbf{r}_k^2 - 2\alpha(2+\beta) \quad (101)$$

as for  $\nabla_k u(r_{kl})$

$$\nabla_k u(r_{kl}) = u'(r_{kl}) \nabla_k r_{kl} = u'(r_{kl}) \nabla_k \sqrt{(\mathbf{r}_k - \mathbf{r}_l)^2} = u'(r_{kl}) \frac{\mathbf{r}_k - \mathbf{r}_l}{\sqrt{(\mathbf{r}_k - \mathbf{r}_l)^2}} = u'(r_{kl}) \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \quad (102)$$

$$u'(r_{kl}) = \frac{\partial u(r_{kl})}{\partial r_{kl}} = \frac{a}{r_{kl}(r_{kl} - a)} \quad (103)$$

and  $\nabla_k^2 u(r_{kl})$

$$\nabla_k^2 u(r_{kl}) = \nabla_k u'(r_{kl}) \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} = u''(r_{kl}) + u'(r_{kl}) \nabla_k \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} = u''(r_{kl}) + u'(r_{kl}) \frac{2}{r_{kl}} \quad (104)$$

$$u''(r_{kl}) = \frac{\partial}{\partial r_{kl}} u'(r_{kl}) = \frac{a^2 - 2ar_{kl}}{r_{kl}^2 (r_{kl} - a)^2} \quad (105)$$

The full expression for the Laplacian-term in the local energy becomes

$$\frac{1}{\Psi_T} \nabla_k^2 \Psi_T = 4\alpha^2 \mathbf{r}_k^2 - 2\alpha(2+\beta) \quad (106)$$

$$+ 2(-2\alpha \mathbf{r}_k) \sum_{l \neq k} u'(r_{kl}) \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \quad (107)$$

$$+ \left( \sum_{l \neq k} u'(r_{kl}) \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \right)^2 \quad (108)$$

$$+ \sum_{l \neq k} u''(r_{kl}) + u'(r_{kl}) \frac{2}{r_{kl}} \quad (109)$$

$$= 4\alpha^2 \mathbf{r}_k^2 - 2\alpha(2+\beta) \quad (110)$$

$$+ 2(-2\alpha \mathbf{r}_k) \sum_{l \neq k} \frac{a}{r_{kl}(r_{kl} - a)} \cdot \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \quad (111)$$

$$+ \left( \sum_{l \neq k} \frac{a}{r_{kl}(r_{kl} - a)} \cdot \frac{\Delta \mathbf{r}_{kl}}{r_{kl}} \right)^2 \quad (112)$$

$$+ \sum_{l \neq k} \left( \frac{a^2 - 2ar_{kl}}{r_{kl}^2 (r_{kl} - a)^2} + \frac{a}{r_{kl}(r_{kl} - a)} \cdot \frac{2}{r_{kl}} \right) \quad (113)$$

$$= 4\alpha^2 \mathbf{r}_k^2 - 2\alpha(2+\beta) \quad (114)$$

$$- 4\alpha \mathbf{r}_k \sum_{l \neq k} \frac{a}{(r_{kl} - a)} \frac{\Delta \mathbf{r}_{kl}}{r_{kl}^2} \quad (115)$$

$$+ \sum_{l \neq k} \frac{a^2}{r_{kl}^2 (r_{kl} - a)^2} \cdot \frac{\Delta \mathbf{r}_{kl}^2}{r_{kl}^2} \quad (116)$$

$$+ \sum_{l \neq k} \left( \frac{a^2 - 2ar_{kl}}{r_{kl}^2 (r_{kl} - a)^2} + \frac{2a}{r_{kl}^2 (r_{kl} - a)} \right) \quad (117)$$

### iii. Figures

#### The Simple Gaussian

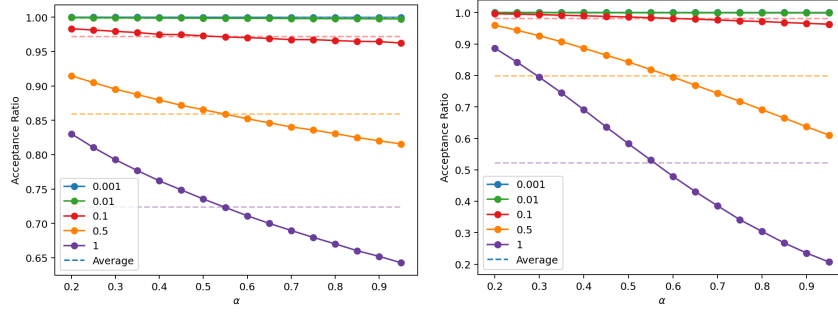


Figure 11. The acceptance ratio as a function of the Metropolis **steplength** (left) and Metropolis-Hastings **timestep** (right), step:  $[0.001, 0.01, 0.5, 1]$ , with the average acceptance ratio **Avg. Acc** for each step as the dashed line.

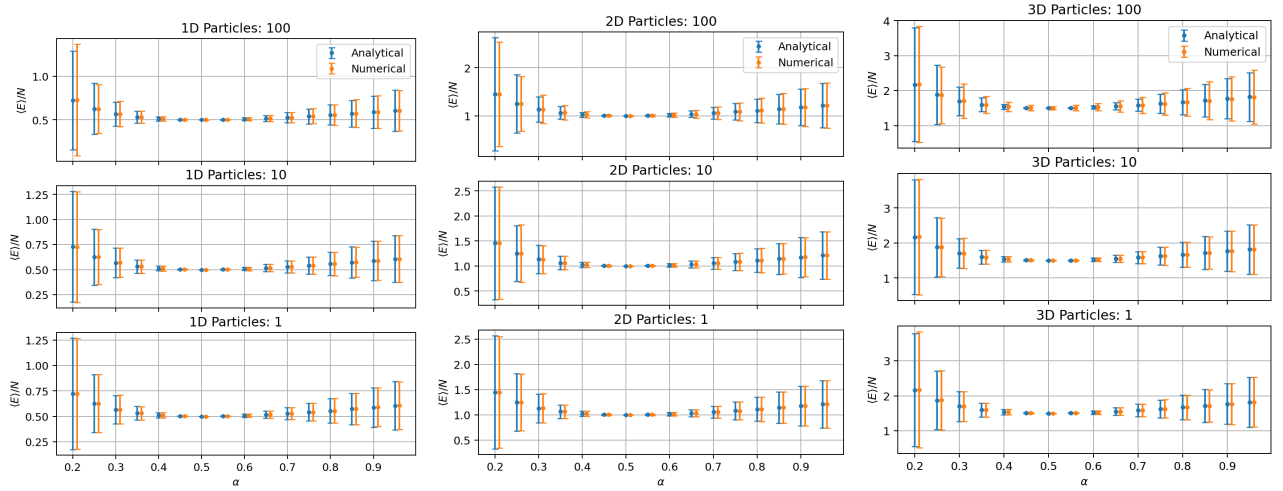


Figure 12. The energy expectation value per particle for the brute-force Metropolis as a function of  $\alpha \in [0.2, 0.95]$  for  $2^{21}$  MC-Cycles, in 1, 2 and 3 dimensions from left to right. The numerical values are slightly shifted for better visibility. The numerical differentiation was done using  $h = 1e-06$ .

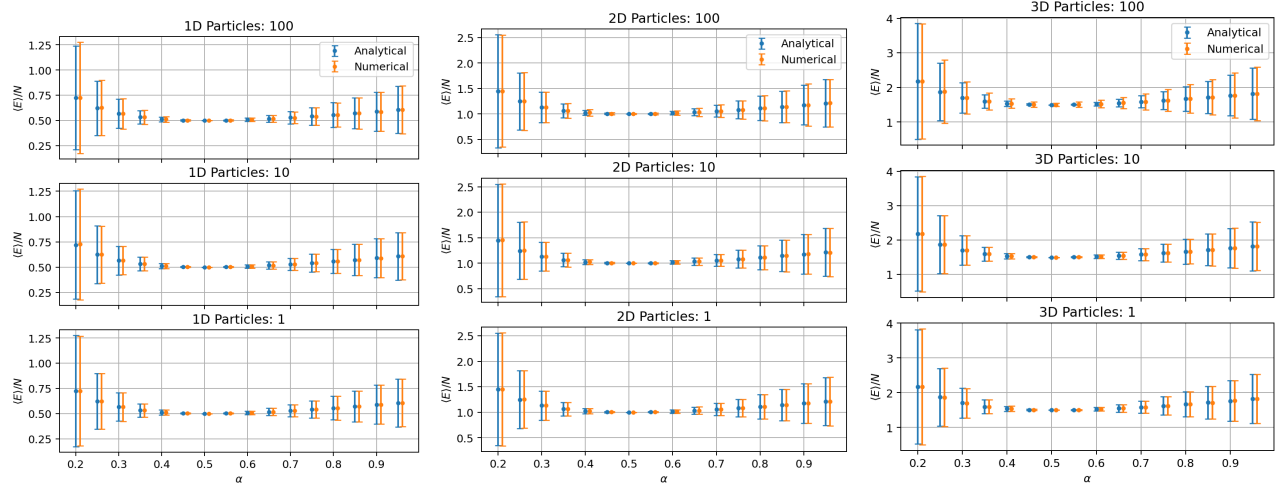


Figure 13. The energy expectation value per particle for the importance sampling Metropolis-Hastings as a function of  $\alpha \in [0.2, 0.95]$  for  $2^{21}$  MC-Cycles, in 1, 2 and 3 dimensions from left to right. The numerical values are slightly shifted for better visibility. The numerical differentiation was done using  $h = 1e-06$ .

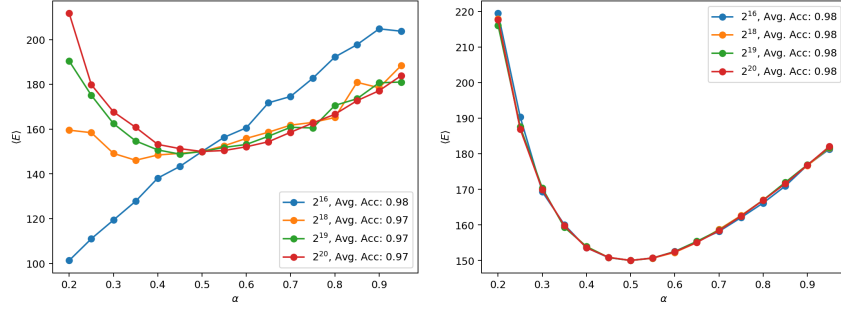


Figure 14. The energy as a function of the number of MC-cycles:  $[2^{16}, 2^{18}, 2^{19}, 2^{20}]$ , with the average acceptance ratio **Avg. Acc.**, across all  $\alpha$  values. Brute-force Metropolis (left) Metropolis-Hastings (right); both **steplength** and **timestep** were set to 0.1.

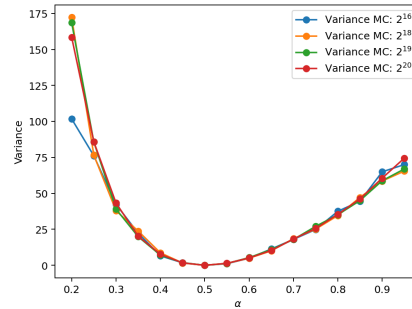


Figure 15. The variance of the quantum system  $\sigma^2$  as a function of the number of MC-cycles:  $[2^{16}, 2^{18}, 2^{19}, 2^{20}]$ , across all  $\alpha$  values.



## Correlated

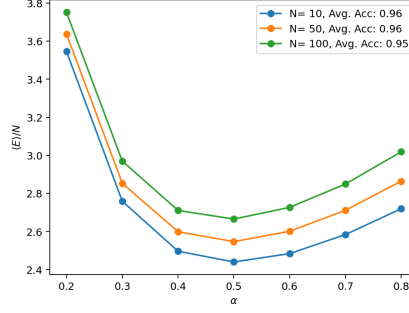


Figure 16. Energy per number of particles  $N$  for  $2^{20}$  MC-cycles as a function of  $\alpha$  with the average acceptance ratio Avg. Acc.

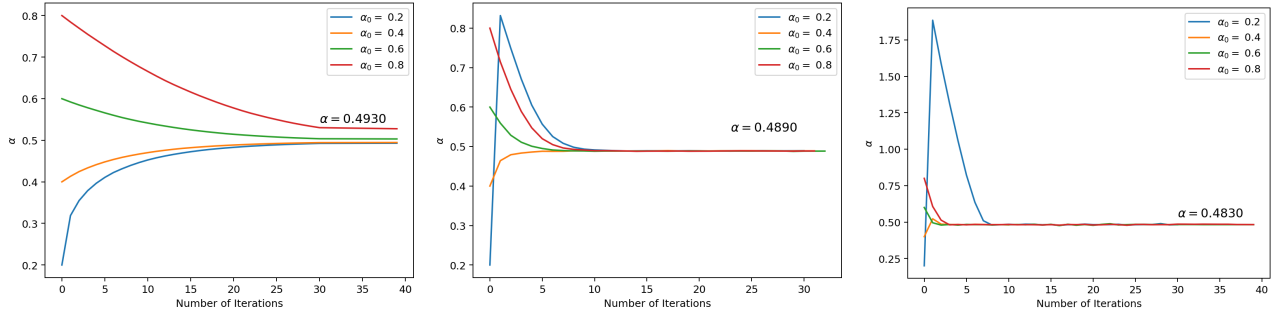


Figure 17. Gradient descent search of  $\alpha$  that minimizes the energy expectation value; with  $\alpha_0 = 0.2, 0.4, 0.6, 0.8$  as initial values. All for the 3 dimensional case with 10, 50 and 100 particles, respectively from left to right. The value of  $\alpha$  corresponding to the lowest energy is highlighted in each plot.

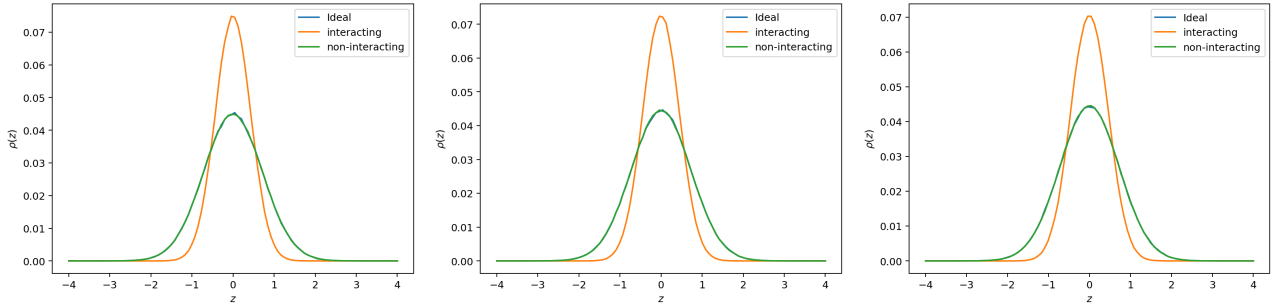


Figure 18. One-body density along the  $z$ -axis of the ideal, interacting and non-interacting case for 10, 50 and 100 particles respectively from the left. The calculation was done with values of  $\alpha$  determined by the GD-search. In other words for  $\alpha = 0.493, 0.489, 0.483$  respectively starting from the left.

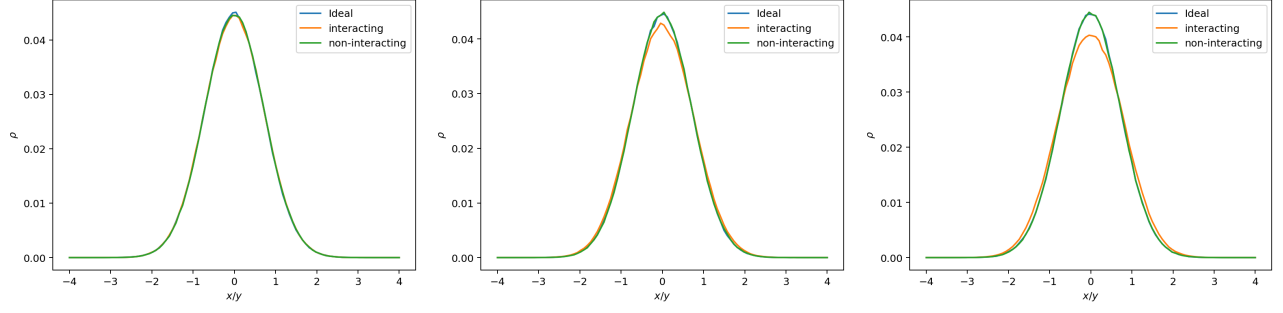


Figure 19. One-body density along the x-/y-axis of the ideal, interacting and non-interacting case for 10, 50 and 100 particles respectively from the left. The calculation was done with values of  $\alpha$  determined by the GD-search. In other words for  $\alpha = 0.493, 0.489, 0.483$  respectively starting from the left.