Project 1 - Monte Carlo

Erlend Lima, Kristian Wold

A hard shell bosonic system in an elliptical harmonic oscillator investigated using variational Monte Carlo, with and without importance sampling. The variational parameter was optimized by gradient descent. Blocking is used to hinder underestimation of the statistical error due to correlated data, as popularized by [1].

For $\beta = 2.82843$, $a = 0.0043$, the ground state energy was found to be 24.3985 ± 0.0011, 127.37 ± 0.035 and 265.69 ± 0.27, for 10, 50 and 100 bosons, respectively. Importance sampling was shown to be more efficient than brute force sampling, in the sense that it created less correlated data and converged faster to the physical solution. The one-body density was shown to widen for a increasing number of particles, indicating that the hard shell potential behaves repulsively, as expected.

I. INTRODUCTION

Attention has in recent years been given to confined quantum mechanical systems, with the prospect of applications from qbits to energy storage. Among these are bosonic systems, due to the discovery of Bose-Einstein condensates in gases of alkali atoms [2].

Many-body systems are remarkably difficult to solve, rarely having any analytical solution. Variational Monte-Carlo offers a numerical approach, whereby parameters in a trial wave function are optimized to minimize the ground state energy. This approach is explored on a system of hard-shell bosons bound by an elliptical trap. The system is explored in several dimensions, differing number of particles and variations of the potentials. The implementation is compared to analytical results wherever possible, and its numerical results discussed.

The GitHub repository is available here: https://github.com/ErlendLima/FYS4411/tree/master/Project1.

II. THEORY

A. Variational Monte Carlo

In order to find a good candidate wavefunction for a given potential, one can employ the *variational principle*. One starts by guessing a trial wavefunction $|\Psi_T\rangle$ and estimating the trial energy, which is guaranteed to be equal to or higher than the true ground state energy E_0:

$$E_0 \leq E = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \tag{II.1}$$

If $|\Psi_T\rangle$ is an eigenfunction of the Hamiltonian, the variance σ^2 will be minimal

$$\sigma^2 = \frac{\langle \Psi_T | H^2 | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} - \left(\frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \right)^2 = 0$$

The variational principle expands on this idea by letting $|\Psi_T\rangle$ be a functional class of a *variational parameter* α. By varying α one can find the optimal trial wavefunction within the functional class by minimizing σ^2.

Only a small collection of potentials have analytical solution using the variational principle. For most potentials, one must numerically integrate (II.1) using Monte Carlo integration.

For a stochastic variable x with probability density function $p(x)$, the average $\langle x \rangle$ is defined as

$$\langle x \rangle = \int_{\mathbb{R}} x p(x) \mathrm{d}x$$

By sampling the stochastic variable M times, the average can be approximated by

$$\langle x \rangle = \int_{\mathbb{R}} x p(x) \mathrm{d}x \approx \frac{1}{M} \sum_{i=1}^{M} x_i p(x_i)$$

Applying this to an observable $\mathcal{O}$, we have

$$\begin{aligned}\langle \mathcal{O} \rangle &= \langle \Psi | \mathcal{O} | \Psi \rangle \\ &= \int \mathrm{d}\mathbf{r} \Psi^* \mathcal{O} \Psi \\ &= \int \mathrm{d}\mathbf{r} |\Psi|^2 \frac{1}{\Psi} \mathcal{O} \Psi \\ &= \frac{1}{M} \sum_{i=1}^{M} p(\mathbf{r}) \mathcal{O}_L\end{aligned}$$

where $|\Psi|^2$ is defined as the probability density function, and $\frac{1}{\Psi}\mathcal{O}\Psi$ the *local operator*.
The local trial energy can then be defined as

$$E_L = \frac{1}{\Psi_T} H \Psi_T$$

which can be computed using Monte Carlo integration as

$$\langle E_L \rangle \approx \frac{1}{M} \sum_{i=1}^{M} p(\mathbf{r}_i) E_L(\mathbf{r}_i)$$

The goal is therefore to minimize minimizing $\sigma^2 = \langle E_L^2 \rangle - \langle E_L \rangle^2$ over the variational parameter α.

B. The System

1. The Potentials

The Hamiltonian under investigation describes N bosons in a potential trap, and is on the form

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

where V_{ext} is the external potential of the trap while V_{int} is the internal potential between the particles. The external potential has an elliptical form, being anisotropic in the z-direction:

$$V_{\text{ext}}(\mathbf{r}) = \frac{1}{2} m \left(\omega \left[x^2 + y^2 \right] + \omega_z z^2 \right) \tag{II.2}$$

The internal potential is a hard shell potential, being infinite for distances where two bosons overlap:

$$V_{\text{int}} = \begin{cases} \infty, & \text{for } |r_i - r_j| \leq 0 \\ 0, & \text{otherwise} \end{cases}$$

2. The Trial Wavefunction

The elliptical spherical trap (II.2) represents a harmonic oscillator. As the trial wavefunction should be as close as possible to the expected true wavefunction, a reasonable guess at its shape is the eigenfunction of harmonic oscillators, namely Gaussian functions. For a N-bosonic system the trial wavefunction is therefore

$$\begin{aligned}h(\mathbf{r}_1, \ldots, \mathbf{r}_N, \alpha, \beta) &= \prod_{i=1}^{N} g(\mathbf{r}_i, \alpha, \beta) \\ &= \exp\left\{ -\alpha \sum_{i=1}^{N} \left(x_i^2 + y_i^2 + \beta z_i^2 \right) \right\}\end{aligned}$$

with g the onebody function. The internal potential should cause the wavefunction to decrease continuously down to zero as the distance of two particles goes to zero. One such possible function is

$$f(a, \mathbf{r}_i, \mathbf{r}_j) = \begin{cases} 0, & |\mathbf{r}_i - \mathbf{r}_j| \leq a \\ 1 - \frac{a}{|r_i - r_j|}, & \text{otherwise} \end{cases}$$

Combining both potential contributions, the complete trial wave function is therefore

$$\Psi_T(\mathbf{r}, \alpha, \beta, a) = \exp\left\{ -\alpha \sum_{i=1}^{N} \left(x_i^2 + y_i^2 + \beta z_i^2 \right) \right\} \prod_{i<j}^{N} f(a, \mathbf{r}_i, \mathbf{r}_j)$$

3. Non-interacting Case

For non-interacting bosons in a spherical with $\beta = 1$ and $a = 0$ the system reduces to spherical harmonic oscillators where analytical solutions are available. The trial wavefunction reduces to simply the product of one body elements

$$\Psi_T(\mathbf{r}) = \prod_{i}^{N} \exp\left[-\alpha \left(x_i^2 + y_i^2 + z_i^2 \right) \right] = \prod_{i}^{N} \exp\left(-\alpha |r_i|^2 \right)$$

while the Hamiltonian reduces to

$$H = \sum_{i}^{N} \frac{-\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} m \omega^2 r_i^2$$

which in natural units is

$$H = \frac{1}{2} \sum_{i}^{N} -\frac{1}{m} \nabla_i^2 + m \omega^2 r_i^2$$

Applying the Hamiltonian gives the local energy as

$$E_L = \frac{\alpha d}{m} N + \left(\frac{1}{2} m\omega^2 - \frac{2\alpha^2}{m}\right) \sum_i^N r_i^2$$

where d is the dimension. As the factor $\sum r_i^2$ is always positive, its term should be minimized, which is accomplished by setting $\alpha = \frac{m\omega}{2}$, giving a minimal local energy of

$$E_L = \frac{\omega d N}{2} \tag{II.3}$$

4. *Interacting Case*

The local energy for the full interacting case is much more complicated. The full computation is deferred to appendix A. The expression evaluates to

$$\begin{aligned} E_L = -\frac{1}{2m} \sum_i \Bigg[& 4\alpha^2 \left(x_k^2 \hat{\mathbf{i}} + y_k^2 \hat{\mathbf{j}} + \beta^2 z_k^2 \hat{\mathbf{k}} \right) \\ & - 2\alpha(d - 1 + \beta) \\ & - 4\alpha \left(x_k \hat{\mathbf{i}} + y_k \hat{\mathbf{j}} + \beta z_k \hat{\mathbf{k}} \right) \sum_{l \neq k} \frac{\mathbf{r}_k - \mathbf{r}_l}{r_{kl}} u'(r_{kl}) \\ & + \sum_{i \neq k} \sum_{j \neq k} \frac{(\mathbf{r}_k - \mathbf{r}_i)(\mathbf{r}_k - \mathbf{r}_j)}{r_{ki} r_{kj}} u'(r_{ki}) u'(r_{kj}) \\ & + \sum_{l \neq k} \left(u''(r_{kl}) + \frac{2}{r_{kl}} u'(r_{kl}) \right) \Bigg] \\ & + \sum_i V_{\text{ext}}(\mathbf{r}_i) + \sum_{i<j} V_{\text{int}}(\mathbf{r}_i, \mathbf{r}_i) \end{aligned}$$

C. Onebody Density

A useful way to understand a many body system is to integrate over all dimensions except for one, yielding the *one body density* $\rho(\mathbf{r})$, defined as

$$\rho(\mathbf{r}) = \int \mathrm{d}\mathbf{r}_2 \ldots \mathrm{d}\mathbf{r}_N \left|\Psi_T(\mathbf{r})\right|^2$$

It is a scaled probability density function giving the number of particles within the volume $\Delta\mathbf{r}$ as $\rho(\mathbf{r})\Delta\mathbf{r}$. By convention the integral over all $\mathrm{d}\mathbf{r}$ yields the total number of particles in the system, N.

D. Metropolis-Hastings

The estimate of the local energy relies on samples from the trial wave function. To get a physical value, the configuration of the particles must be as physical and probable. As the configuration achieving this is unknown, the configuration space must be explore. This is done through Monte Carlo simulation, more specifically by using the Metropolis-Hastings algorithm.

At each MC step a single particle is chosen at random, and a change to its position is proposed by moving it a fixed step length Δ and computing the ratio between new and old probability densities

$$\omega = \frac{P(\mathbf{r}_r, \ldots, \mathbf{r}_k^*, \ldots, \mathbf{r}_n)}{P(\mathbf{r}_r, \ldots, \mathbf{r}_k, \ldots, \mathbf{r}_n)} = \frac{\left|\Psi_T(\mathbf{r}_r, \ldots, \mathbf{r}_k^*, \ldots, \mathbf{r}_n)\right|^2}{\left|\Psi_T(\mathbf{r}_r, \ldots, \mathbf{r}_k, \ldots, \mathbf{r}_n)\right|^2}$$

where $\mathbf{r}^*$ denotes a modified position. If the ratio ω is greater than a uniformly distributed number $\theta \in [0, 1]$, the move is accepted. This ratio can often be reduced analytically to obviate the need for recomputing the entire probability density each step.

E. Drift Force

The disadvantage of the brute force approach is that the proposed move is random. The configuration space is explored more effectively by instead moving towards regions of higher probability density. This is achieved by adding a quantum force on the particle k:

$$\mathbf{F}_k(\mathbf{r}_1, \ldots, \mathbf{r}_N) = 2 \frac{1}{\Psi_T(\mathbf{r})} \nabla_k \Psi_T(\mathbf{r})$$

The force is incorporated into the probability ratio by the Green's function of the Fokker-Planck equation:

$$\begin{aligned} G(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{r}_1', \ldots, \mathbf{r}_N') = & \frac{1}{(4\pi D \Delta t)^{3/2}} \times \\ & \sum_{i=1}^N \exp\left(- \frac{[\mathbf{r}_i - \mathbf{r}_i' - D\Delta t \mathbf{F}_i]^2}{4D\Delta t} \right) \end{aligned}$$

giving a new expression of the transition probability as

$$\omega = \frac{G(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{r}_1', \ldots, \mathbf{r}_N') \left|\Psi_T(\mathbf{r}_1, \ldots, \mathbf{r}_N)\right|^2}{G(\mathbf{r}_1', \ldots, \mathbf{r}_N', \mathbf{r}_1', \ldots, \mathbf{r}_N') \left|\Psi_T(\mathbf{r}_1', \ldots, \mathbf{r}_N')\right|^2}$$

This is known as *importance sampling*. Far more computationally expensive than naive brute force sampling, by also leading to a larger number of accepted MC steps.

For the full system the drift force evaluates to

$$\mathbf{F}_k(\mathbf{r}_1,\ldots,\mathbf{r}_N) = -4\alpha\left(x_k\hat{\mathbf{i}} + y_k\hat{\mathbf{j}} + \beta z_k\hat{\mathbf{k}}\right) + 2\sum_{i\neq k}\frac{\mathbf{r}_k-\mathbf{r}_i}{r_{ki}}u'(r_{ki}) \tag{II.4}$$

See appendix A 3 for derivation.

III. METHOD

A. Outline of Program

The code is divided into two: a C++ backend and Python frontend. The Python code is a simple wrapper around the C++ and functions as a high level interface which is used in the Jupyter notebooks. The C++ is object oriented with System running the simulation, Sampler extracting interesting observables from the simulation, and Hamiltonian, WaveFunction and InitialState and their subclasses are composable objects for constructing different systems, such as spherical trap with non-interacting bosons.
The Jupyter notebooks are available in the GitHub repository, and run reproduce all presented herein.

B. Statistical Treatment

When dealing with estimation based on data, either produced experimentally or numerically, a proper statistical treatment is appropriate.
Perhaps the most common estimator in statistics is the mean:

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i, \tag{III.1}$$

where X_i are identically distributed random variables, but not necessarily independent. To establish the uncertainty of the estimator, it necessary to calculate $V(\bar{X})$. By using the linearity of the variance, we have

$$\begin{aligned} V(\bar{X}) &= V(\frac{1}{N}\sum_{i=1}^{N} X_i) \\ &= \frac{1}{N^2}V(\sum_{i=1}^{N} X_i) \\ &= \frac{1}{N^2}\sum_{i,j}\mathrm{Cov}(X_i,X_j) \end{aligned} \tag{III.2}$$

where $\mathrm{Cov}(X_i,X_j)$ is the covariance of X_i, and X_j. Under the assumption that the variables are independent, which is often approximately true, $\mathrm{Cov}(X_i,X_j)=0$, for $i\neq j$. The variance of the mean is thus simplified to

$$\begin{aligned} V(\bar{X}) &= \frac{1}{N^2}\sum_i Cov(X_i,X_i) \\ &= \frac{1}{N^2}\sum_i Var(X_i) \\ &= \frac{N}{N^2}Var(X) = \frac{Var(X)}{N} \end{aligned}$$

where it has been used that the variables are identically distributed. This is a convenient result, as one can calculate the variance of an estimator based on the variance of the random variables X that enter it. This is typically done using the sample variance

$$V(X) \approx \frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2,$$

where x_i is some sample and $\bar{x} = \frac{1}{N}\sum x_i$ is the sample mean. Problems arise when the data is correlated, as the neglect of cross terms present in Equation III.2 causes a underestimation of the variance of the estimator. This is very important to bear in mind when working with VMC, as data produced by simulating many particles tends to be highly correlated.
Made popular by H. Flyvbjerg [1], blocking is a renormalization method for treating correlated data when calculating the variance of $\bar{X}$. Blocking is defined by a recursive downsampling of data. Starting with a vector containing the original data $\vec{X}_0$, blocking of degree d is obtained by performing

$$\left(\vec{X}_{i+1}\right)_k = \frac{1}{2}\left(\left(\vec{X}_i\right)_{2k-1} + \left(\vec{X}_i\right)_{2k}\right),$$

for all $1\leq i\leq d-1$.
It can be shown that the quantity

$$V(\bar{X}_k) \approx \frac{\hat{\sigma}_k^2}{n_k} + e_k$$

is invariant with respect to the degree of blocking k, where $\hat{\sigma}_k^2$ is the sample variance of $\vec{X}_k$, and e_k is know as the truncation error. Under the assumption that the original data $\vec{X}_0$ is asymptotically uncorrelated, [1] showed that e_k is decreasing, and can be make small given big enough k. Therefore, for sufficiently large k

$$V(\bar{X}_0) = V(\bar{X}_k) \approx \frac{\hat{\sigma}_k^2}{n_k} + e_k \approx \frac{\hat{\sigma}_k^2}{n_k} \tag{III.3}$$

In this way, the sample variance of the sufficiently blocked data does not suffer from underestimation, as the correlation is first eliminated.

C. Gradient Descent

Variational Monte Carlo involves varying the variational parameter α to find its optimal value. This is accomplished using gradient descent. The iterative method updates its guess at α by steeping in the direction of the largest change in the local energy with respect to α

$$\alpha_{n+1} = \alpha_n - \eta \frac{\partial E_L}{\partial \alpha}(\alpha_n)$$

As optimizing α is comparatively simple, no more fancy a method is needed.

D. One-Body Density

To extract a one-body density from a Monte-Carlo simulation, we partition space into a number of bins in an appropriate range where the wave function is large. The bin size can be chosen small to get finer details of the density, but will require more data to mitigate statistical error.
For each particle configuration produced at every Metropolis step, the number of particles coinciding with each bin is checked. The one-body density is then produced by averaging over all Metropolis steps.

IV. RESULTS AND DISCUSSION

A. Validating the Model

N	Numerical	Analytical
1	1.5 ± 0.0	1.5
2	3.0 ± 0.0	3
10	15.0 ± 0.0	15
100	150.0 ± 0.0	150

Table IV.1: Expected local energy for non-interactive particles in 3D spherical potential compared to analytical solution for $\alpha = 0.5$ for different number of particles. By (II.3), there is no spatial dependency, making the comparison easy. The numerical results agree perfectly with the analytical value.

To validate that our implementation is working correctly, the model is run for a spherical potential without any interactions. The expected local energy is compared to the analytical solution, summarized in table IV.1. The computed energies are exactly $\frac{d}{2}$ of the number of particles, as expected from Equation (II.3) on page 3.
Further validation is performed by comparing the one body density found from brute force sampling, importance sampling and analytical solution Equation (II.4) shown in fig. IV.1. Aside from the statistical noise

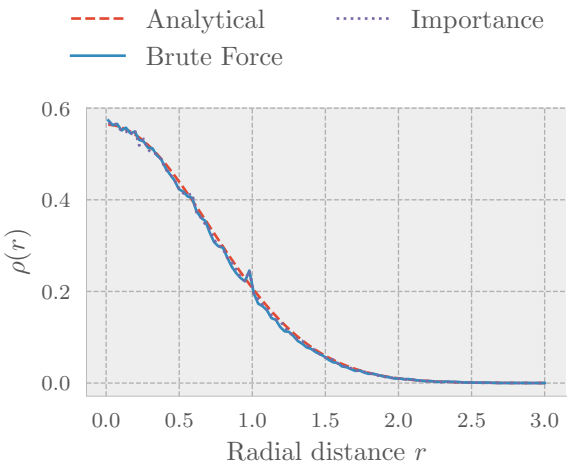

Figure IV.1: One-body density of 1 boson in 1 dimension, using $N = 1 \times 10^6$ cycles, $\omega = 1$, $\alpha = 0.5$.

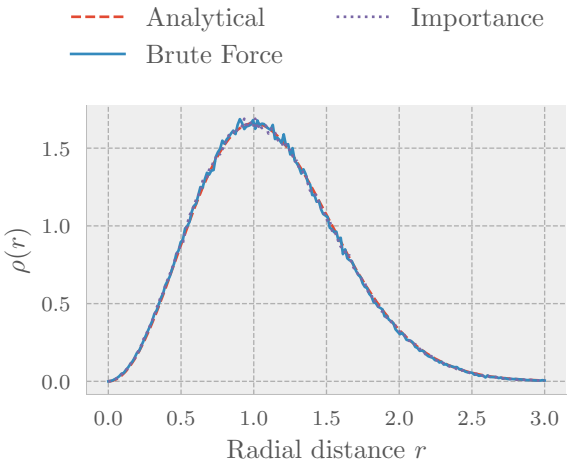

Figure IV.2: Radial one-body density of 2 non-interacting bosons in 3 dimensions, using $N = 1 \times 10^6$ cycles, $\omega = 1$, $\alpha = 0.5$.

introduced by the finite number of Monte-Carlo cycles, the approximated densities follow the analytical result closely. Figure IV.2 demonstrates how this scales with more particles and dimensions, as seen from the radial onebody density of two particles in three dimensions.

B. Blocking of Local Energy

The effect of applying blocking on the local energy before calculating the variance is shown in fig. IV.3. Although the local energies are sampled from an approximately correct distribution, the data is produced

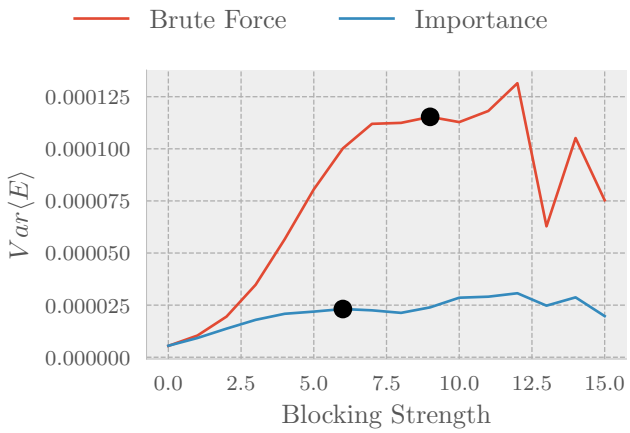

Figure IV.3: Variance of the estimate of $\langle E_L \rangle$ using blocked values of local energy at various strengths. As the blocking strength increases, the autocorrelation is decreased and variance increases before it plateaus, shown with the black dot. The data have been produced for 40 non-interacting bosons in three dimensions using $N = 2^{20}$ cycles, $\alpha = 0.8$, $\omega = 1$, with and without importance sampling

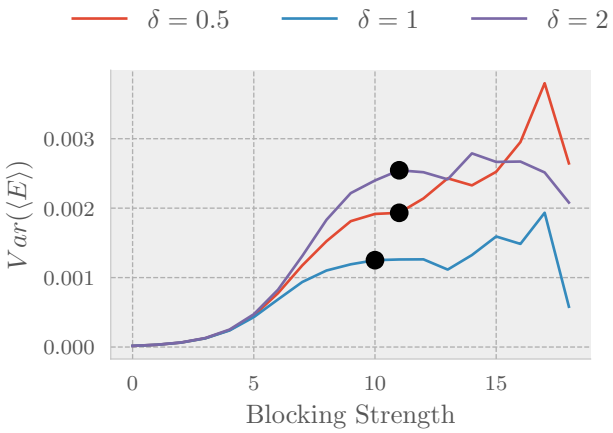

Figure IV.4: Variance of the estimate of $\langle E \rangle$ using blocked values of local energy at various strengths. The data has been produced for 40 non-interacting bosons in 3 dimensions, using $N = 2^{20}$ cycles, $\alpha = 0.8$, $\omega = 1$, using importance sampling with step length 0.5, 1 and 2, respectively

dt	$\langle E \rangle$	blocking strength
0.5	68.009 ± 0.044	11
1	68.039 ± 0.035	10
2	67.942 ± 0.050	11

Table IV.2: Estimate of energy $\langle E \rangle$ for 40 particles, $\alpha = 0.3$, $\omega = 1$

by moving a single boson at a time. The move may be rejected by the metropolis algorithm, causing no change in state. Because of this, the data is highly correlated, leading to an underestimation of the variance as seen for the unblocked data. The variance stabilizes after repeated blocking, when the data is approximately uncorrelated. Note that this happens at different strength for brute force sampling and importance sampling. From fig. IV.3 the latter requires a weaker blocking strength, indicating that the data created are less correlated. This is expected as importance sampling uses additional information from the wave function and explores the space of states more efficiently than brute force. The variance stabilizes at strength 6 when using importance sampling as opposed to 9 when using brute force, granting eight times the effective amount of data and thus a smaller variance of the estimator.

The effect of blocking for a larger system is shown in fig. IV.4 using only importance sampling and step lengths of 0.5, 1 and 2. As only a single particle is moved in each MC cycle, increasing the size of the system results in a comparatively smaller change. The local energies will invariably become more correlated, which is apparent when comparing the blocking strengths from fig. IV.4 to fig. IV.3.

Further, the least correlated data was produced with a step length of $\delta t = 1$, requiring blocking strength 10 as opposed to 11 for $\delta t = 0.5$ and 2. Using a step length too small results in the local energy changing minimally from cycle to cycle, increasing the correlation and minimizing the effective amount of data. The same happens if the step length is too big, as large moves will tend to be rejected often, causing repeated values to be sampled. Table IV.2 presents the resulting estimates of energy with the uncertainty established using the blocking. The estimated energy is only sensitive to the step length in the sense that more cycles may be needed to get an accurate result of the step length is taken too small or too big.

C. Energy of Non-Interacting System

Varying the variational parameter α gives different trial wave functions that have different expected local energy. For the non-interacting spherical system there are analytical solutions for the α-dependency of the local energy, and can be used to validate our model. This comparison is shown in fig. IV.5 and fig. IV.6. The curves overlap within the MC-uncertainty as expected, with the minimum occurring at $\alpha = 0.5$, giving confidence in the implementation.

D. Numerical and Analytical Laplacian

When evaluating the Hamiltonian, one can choose to use the analytical expression for the Laplacian or compute it numerically using finite differences. Both meth-

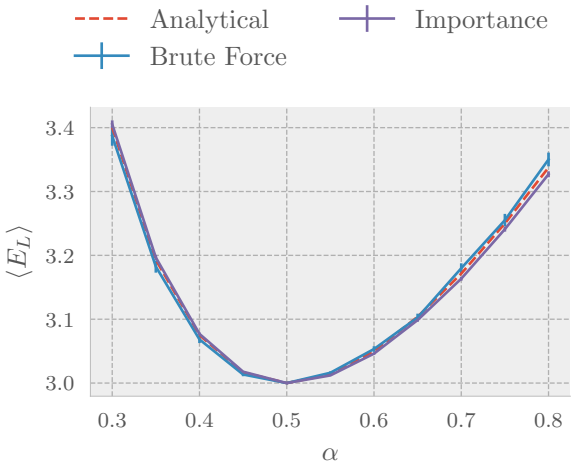

Figure IV.5: The expected local energy as function of the variational parameter α for both brute force and importance sampling, compared to the analytical solution. The system consists of two non-interacting bosons using 2^{17} cycles.

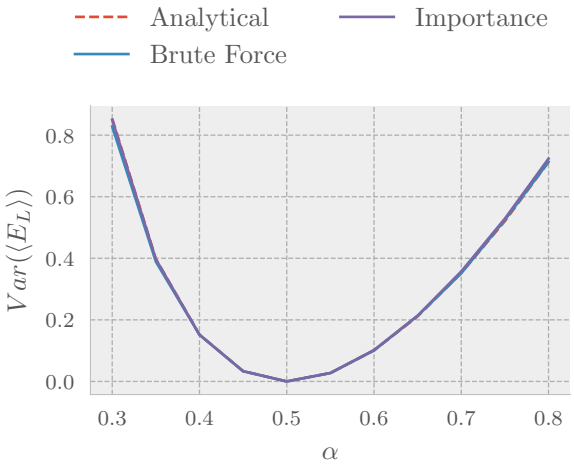

Figure IV.6: The variance of the expected local energy as function of the variational parameter α for both brute force and importance sampling, compared to the analytical solution. The system consists of two non-interacting bosons using 2^{17} cycles.

ods give the same solution, as illustrated by fig. IV.7. Suspecting that the computational cost differs, their CPU time are plotted in fig. IV.8. The difference is negligible with few particles, but the numerical Laplacian quickly outgrows the analytical, showing exponential dependency on the number of particles. Most potentials do now have analytical solution, having to rely on the much slower numerical computation.

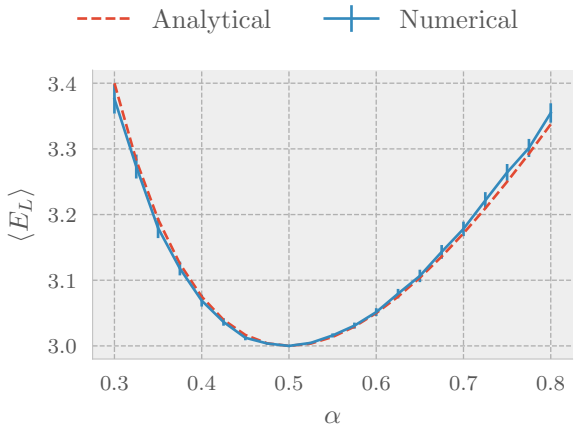

Figure IV.7: Estimated energy $\langle E_L \rangle$ for two non-interacting bosons, comparing the analytical solution to numerically evaluated Laplacian. 2^{17} cycles, $\omega = 1$ and importance sampling was used. Errorbars was established using blocking.

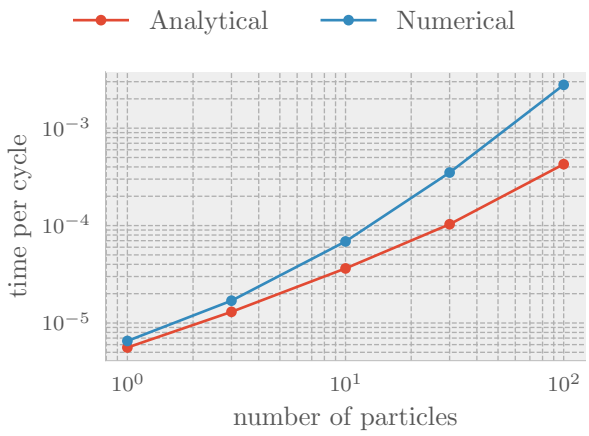

Figure IV.8: Comparison of CPU-time of numerical and analytical Laplacian for different numbers of particles. The timing is with respect to the whole simulation. However, as the uninteresting steps are of $\mathcal{O}(1)$, they irrelevant when comparing the graphs.

E. Interacting Potentials

Adding interaction between the bosons drastically changes the behavior of the system. In fig. IV.9 non-interacting bosons are compared to systems having two and three interacting bosons. The densities of the latter are spread out, with maxima directly proportional to the number of particles. The interpretation is that the interacting particles are repelled by each other and attempts to minimize their energy while staying within the potential, resulting in N bumps becoming more and more spread out as N grows.

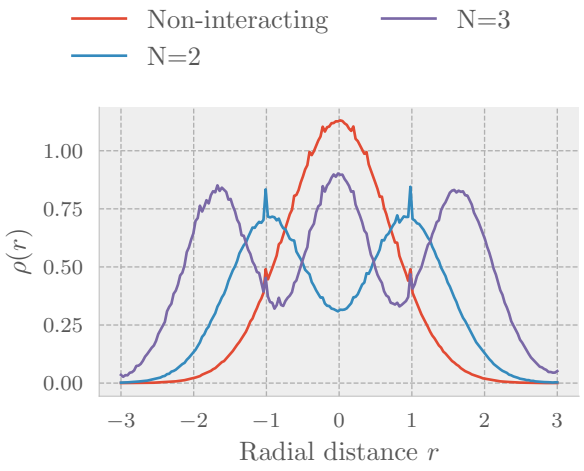

Figure IV.9: One body density ρ in one dimension for two non-interacting bosons, two interacting bosons, and three interacting bosons. Non-interacting bosons occupies the same space to minimize the energy, while the interacting bosons separate. The more bosons, the more they spread out.

F. Optimal Parameter α

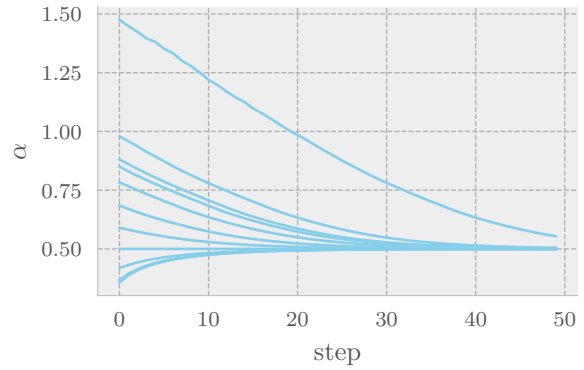

Figure IV.10: Gradient decent on the parameter α for two non-interacting bosons with different α_0, learning rate $\mu = 0.01$, $N = 10000$ steps and $\omega = 1$

Gradient descent was performed on the variational parameter α with non-interactive bosons. Figure IV.10 shows the results. The method is robust for different initial guesses α_0, all converging to 0.5 as expected. This establishes confidence that the method will converge to the optimal α in the interacting case where the analytical solution is not known.

Bosons	$\langle E \rangle$	optimal α
10	24.3985 ± 0.0011	0.49752
50	127.37 ± 0.035	0.48903
100	265.69 ± 0.27	0.48160

Table IV.3: Estimate of energy $\langle E \rangle$ for 10, 50 and 100 particles in elliptical potential. $\beta = \gamma = 2.8284$. $N = 2^20$ cycles per thread, of 12 threads. Importance sampling with a step length 0.5 was used. Gradient decent was used to optimize α

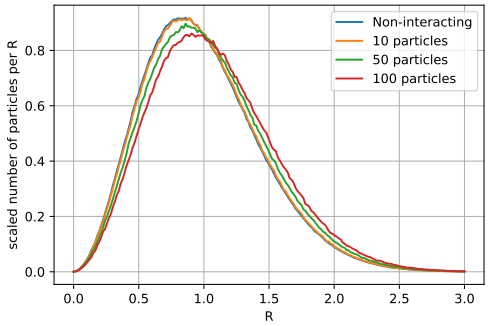

Figure IV.11: Comparison of the radial onebody density for interacting and non-interacting bosons. The hardshell radius of in the interacting case is $a = 0.0043$. The elliptical potential is defined by $\beta = \gamma = 2.82843$.

G. Interacting Bosons in Elliptical Potential

Table IV.3 presents the estimated ground state energies for 10, 50 and 100 interacting bosons confined by a elliptical potential. The optimal value for α was found using gradient decent. Notice that the optimal value is decreasing with the number of particles in the interacting case, whereas it is analytically $\alpha = 0.5$ in the non-interacting case for any number of particles. A smaller value of α indicates the onebody exponential functions widens, resulting in a wider distribution of particles. This makes sense, as hardshell bosons interact repulsively. Therefore, the more particles, the more energetically favorable it is to spread out. This is consistent with the 1D result presented in Figure IV.11. As the number of particles increase, the radial onebody density tends to migrate outward with respect to the non-interacting system.

The ground state energy was most accurately calculated for 10 bosons. This is because, as discussed earlier, data tend to be less correlated for smaller systems. To get an accurate estimate for larger systems, a increase in the number of cycles is appropriate. Due to the increasing CPU-demand, a trade-off must be made between time and accuracy.

A comment on the interpretation of the error presented in Table IV.3 is appropriate: The error is an estimate of the statistical error in the estimated ground state energy introduced by the Monte-Carlo simulation. It does not account for possible error introduced in the estimation of α that minimizes the energy. This is much harder to quantity, and is possibly large in the

case for 50 and 100 particles, as the gradient was hard to estimate consistently. Furthermore, the error does not in any way establish a confidence interval for the true ground state value of the Hamiltonian we are investigating, which is harder still to quantify. Our estimate a approximate minimum only in the subspace spanned by out trial wave function, which is hopefully, somewhat close the exact solution of the system.

V. CONCLUSION

Solving many-body systems numerically is a daunting task, but aided by Metropolis-Hastings it becomes feasible for small systems. It was shown that it is an accurate algorithm for sampling states and discovering the expected local energy for different trial wave functions.

As local energies produced in this fashion tends to be highly correlated, blocking proved invaluable for removing correlation and establishing faithful estimates of statistical error introduced by Monte-Carlo. The severe underestimation of variance when not using blocking was shown in fig. IV.3.

Importance sampling was shown in fig. IV.3 to produce data less correlated than that produced by brute force, in turn yielding more effective data, as a lower degree of blocking was required to make the data uncorrelated. The ground state energy in the space of our trial wave function was calculated to a precision of 0.005%, given that the optimal α was correct(table IV.3). The precision was lower for 50 and 100 bosons, 0.02% and 0.1% repectivly. As previously stated, a higher number of particles results in more strongly correlated local energies, and therefore less effective data.

The radial onebody density reveal an important characteristic of the system, namely that the hardshell potential acts repulsively, causing the distribution of particles to be wider in the interacting case than to the non-interacting. For a higher number of particles for a given confining potential, the widening was more pronounced, which goes to proving the same point.

VI. FUTURE WORK

Some manual labor is involved in picking the optimal degree of blocking and step length. These can be found automatically by seeing where the variance stabilizes and the acceptance rate, respectively. As the time spent manually performing these tasks was less than the time it would take to implement these solutions, this was not done, but remain low-hanging fruits for future work.

The project can be easily expanded to treat fermionic systems. As we have developed a general solver class for VMC problems, only the trial wave function must be replaced by a appropriate anti-symmetric function that fits the problem we want to investigate, such as electrons in a harmonic potential.

[1] H. Flyvbjerg and H. G. Petersen, The Journal of Chemical Physics **91**, 461 (1989), https://doi.org/10.1063/1.457480.
[2] J. K. Nilsen, J. Mur-Petit, M. Guilleumas, M. Hjorth-Jensen, and A. Polls, Physical Review A **71**, 053610 (2005).

Appendix A: Local Energy

1. The Gradient and Laplacian

The full trial wavefunction can be written as

$$\Psi_T(\mathbf{r}) = \prod_i^N \phi(\mathbf{r}_i) \cdot \exp\left(\sum_{j<k} u(r_{jk})\right)$$

where $u = \ln f(r_{ij})$ and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$.
By the chain rule we can focus on each factor separately when taking the gradient. The gradient of the first factor is simply the gradient of the one body element in question:

$$\nabla_k \prod_i \phi_i = \prod_{i\neq j} \phi_i \cdot \nabla_k \phi_k$$

The second factor is at face value slightly more arduous, but it is trivial. The summation is just a notational trick to ensure that each interaction between all particles are enumerated once and only once. As the gradient only affects particle k, only the interactions involving it will be non-zero. Hence the gradient is

$$\nabla_k \exp\left(\sum_{j<m} u(r_{jm})\right) = \exp\left(\sum_{j<m} u(r_{jm})\right) \sum_{l\neq k} \nabla_k u(r_{lk})$$

The total gradient of the trial wavefunction is therefore

$$\begin{aligned}\nabla_k \Psi_T(\mathbf{r}) &= \left[\underbrace{\prod_{i\neq k} \phi_i \cdot \nabla_k \phi_k}_{\bullet} + \underbrace{\prod_i \phi_i \sum_{l\neq k} \nabla_k u(r_{lk})}_{\bullet}\right] \\ &\quad \times \exp\left(\sum_{j<m} u(r_{jm})\right) \\ &= \left[\frac{\nabla_k \phi_k}{\phi_k} + \sum_{l\neq k} \nabla_k u(r_{lk})\right] \Psi_T(\mathbf{r}) \quad \text{(A.1)}\end{aligned}$$

The Laplacian is likewise computed termwise

$$\nabla_k^2 \bullet = \left[\nabla_k^2 \phi \prod_{i\neq k} \phi_i + \nabla_k \phi \sum_{i\neq k} \nabla_k u\left(r_{ki}\right)\right] \exp\left(\sum_{j<m} u(r_{jm})\right)$$

and

$$\begin{aligned}\nabla_k^2 \bullet = \Bigg[&\nabla_k \phi_k \prod_{i\neq k} \phi_i \sum_{l\neq l} \nabla_k u(r_{lk}) \\ &+ \prod_i \phi_i \left(\sum_{l\neq k} \nabla_k u\left(r_{lk}\right)\right)^2 \\ &+ \prod_i \phi_i \sum_{l\neq k} \nabla_k^2 u\left(r_{lk}\right)\Bigg] \exp\left(\sum_{j<m} u(r_{jm})\right)\end{aligned}$$

Combining the results and diving through by Ψ_T then gives

$$\begin{aligned}\frac{\nabla_k^2(\bullet + \bullet)}{\Psi_T} &= \frac{\nabla_k^2 \phi_k}{\phi_k} + \frac{\nabla_k \phi_k}{\phi_k} \sum_{l\neq k} \nabla_k u(r_{lk}) + \frac{\nabla_k \phi_k}{\phi_k} \sum_{l\neq k} \nabla_k u\left(r_{lk}\right) \\ &\quad + \left[\sum_{l\neq k} \nabla_k u\left(r_{lk}\right)\right]^2 + \sum_{l\neq k} \nabla_k^2 u\left(r_{lk}\right) \quad \text{(A.2)}\end{aligned}$$

To compute $\nabla_k u\left(r_{lk}\right)$, a change of variable is performed, with

$$\nabla_k = \nabla_k\left(r_{kl}\right) \frac{\partial}{\partial r_{kl}} = \frac{\mathbf{r}_k - \mathbf{r}_l}{r_{kl}} \frac{\partial}{\partial r_{kl}}$$

and

$$\begin{aligned}\nabla_k \left(\frac{\mathbf{r}_k - \mathbf{r}_l}{r_{kl}}\right) &= \frac{r_{kl} \nabla_k \left(\mathbf{r}_k - \mathbf{r}_l\right) - \left(\mathbf{r}_k - \mathbf{r}_l\right) \nabla_k r_{kl}}{r_{kl}^2} \\ &= \frac{d r_{kl}^2 - \left(\mathbf{r}_k - \mathbf{r}_l\right)\left(\mathbf{r}_l - \mathbf{r}_k\right)}{r_{kl}^3} \\ &= \frac{d r_{kl}^2 - \mathbf{r}_k^2 - \mathbf{r}_l^2 + 2\mathbf{r}_k \mathbf{r}_l}{r_{kl}^3} \\ &= \frac{d-1}{r_{kl}}\end{aligned}$$

Applying these relations on (A.2) and multiplying out the squared sum, we obtain the desired result

$$\frac{\nabla_k^2\Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{\nabla_k^2\phi_k}{\phi_k} + 2\frac{\nabla_k\phi_k}{\phi_k}\left(\sum_{l\neq k}\frac{\mathbf{r}_k-\mathbf{r}_l}{r_{kl}}u'(r_{kl})\right) + \sum_{i\neq k}\sum_{j\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_i)(\mathbf{r}_k-\mathbf{r}_j)}{r_{ki}r_{kj}}u'(r_{ki})u'(r_{kj}) + \sum_{l\neq k}\left(u''(r_{kl})+\frac{2}{r_{kl}}u'(r_{kl})\right) \tag{A.3}$$

The full expression involves computing each gradient and derivative. These are found to be

$$\nabla_k\phi_k = -2\alpha\left(x_k\hat{\mathbf{i}}+y_k\hat{\mathbf{j}}+\beta z_k\hat{\mathbf{k}}\right)\phi_k$$
$$\nabla_k^2\phi_k = \left[4\alpha^2\left(x_k^2\hat{\mathbf{i}}+y_k^2\hat{\mathbf{j}}+\beta^2 z_k^2\hat{\mathbf{k}}\right)-2\alpha(d-1+\beta)\right]\phi_k$$
$$u'(r_{ij}) = \frac{a}{r_{ij}(r_{ij}-a)}$$
$$u''(r_{ij}) = \frac{a^2-2ar_{ij}}{r_{ij}^2(r_{ij}-a)^2}$$

2. The Local Energy

The local energy for the full dimensional interactive case is found by using (A.3) and gradients in

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

giving

$$E_L = -\frac{1}{2m}\sum_i\left[4\alpha^2\left(x_k^2\hat{\mathbf{i}}+y_k^2\hat{\mathbf{j}}+\beta^2z_k^2\hat{\mathbf{k}}\right) - 2\alpha(d-1+\beta) - 4\alpha\left(x_k\hat{\mathbf{i}}+y_k\hat{\mathbf{j}}+\beta z_k\hat{\mathbf{k}}\right)\sum_{l\neq k}\frac{\mathbf{r}_k-\mathbf{r}_l}{r_{kl}}u'(r_{kl}) + \sum_{i\neq k}\sum_{j\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_i)(\mathbf{r}_k-\mathbf{r}_j)}{r_{ki}r_{kj}}u'(r_{ki})u'(r_{kj}) + \sum_{l\neq k}\left(u''(r_{kl})+\frac{2}{r_{kl}}u'(r_{kl})\right)\right] + \sum_i V_{\text{ext}}(\mathbf{r}_i) + \sum_{i<j}V_{\text{int}}(\mathbf{r}_i,\mathbf{r}_i)$$

3. The Drift Force

The drift force can likewise be found by inserting the expression for the gradient (A.1) into

$$\mathbf{F}_k = 2\frac{1}{\Psi_T(\mathbf{r})}\nabla_k\Psi_T(\mathbf{r})$$

giving

$$\mathbf{F}_k = 2\frac{\nabla_k\phi_k}{\phi_k} + 2\sum_{i\neq k}\frac{\mathbf{r}_k-\mathbf{r}_i}{r_{ki}}u'(r_{kj}) = -4\alpha\left(x_k\hat{\mathbf{i}}+y_k\hat{\mathbf{j}}+\beta z_k\hat{\mathbf{k}}\right) + 2\sum_{i\neq k}\frac{\mathbf{r}_k-\mathbf{r}_i}{r_{ki}}u'(r_{ki})$$

In the non-interactive case this reduces to

$$\mathbf{F}_k = -4\alpha\left(x_k\hat{\mathbf{i}}+y_k\hat{\mathbf{j}}+\beta z_k\hat{\mathbf{k}}\right)$$

Appendix B: Dimensionless Hamiltonian

To make the Hamiltonian dimensionless, the following change of variables are

$$\mathbf{r}\to\frac{\mathbf{r}}{a_{\text{ho}}} = \sqrt{\frac{m\omega}{\hbar}}\mathbf{r} \qquad E\to\frac{E}{\hbar\omega} \qquad \nabla\to\sqrt{\frac{\hbar}{m\omega}}\nabla$$

giving

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

where the relation $\gamma=\omega_z/\omega$ was made. Applying the E substitution and noting that V_{int} does not change if multiplied by any constant, the final result is obtained:

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