Project 1 - Monte Carlo

Authors goes here

I. INTRODUCTION

The GitHub repository is available here: https://github.com/Caronthir/FYSSTK4155/projects/.

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. Gradient Descent

The optimal value for the variational parameter is found by gradient descent.

C. 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

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

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. Once 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

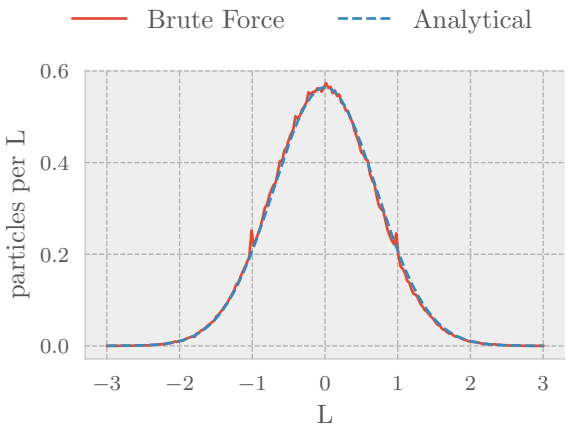
$$E_L = \frac{\alpha d}{m} N + \left(\frac{1}{2} m \omega^2 - \frac{2\alpha^2}{m} \right) \sum_{i}^{N} r_i^2$$

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

$$E_L = \frac{\omega d N}{2}$$

4. Interacting Case

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

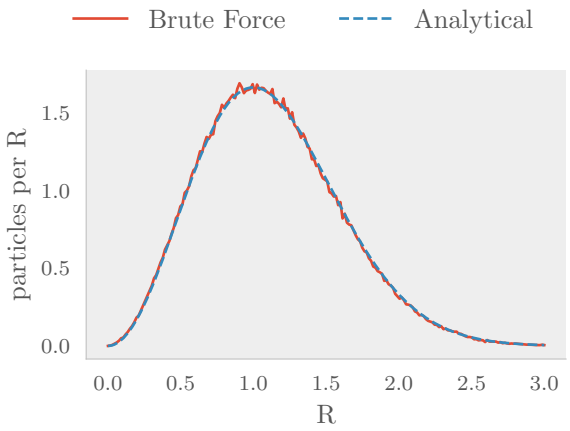


(a) Approximation of onebody density using Metropolis brute force sampling vs analytical(kilde)

(b) Approximation of onebody density using importance sampling vs analytical

Figure V.1: One-body density of 1 boson in 1 dimmension, using $N = 1e6$ cycles, $\omega = 1$, $\alpha = 0.5$.

In Figure V.1 we see that both brute force sampling and importance sampling manage to approximate the onebody density derived from the non-interacting trial wave function: Aside from the statistical noise introduced by the finite number of Monte-Carlo cycles, the approximated densities follow the analytical result closely. Figure Figure V.2 demonstrates that this also scales to more particles and dimensions, as seen from the radial onebody density of two particles in three dimensions.



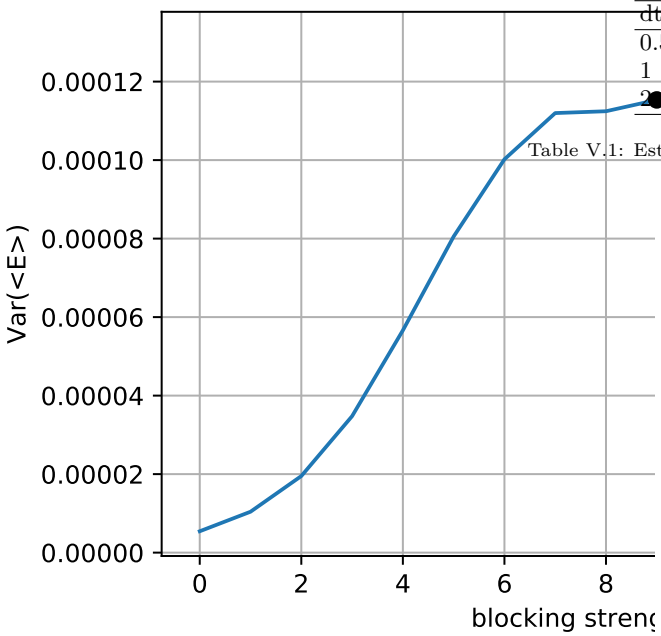
(a) Approximation of radial onebody density using Metropolis brute force sampling vs analytical

(b) Approximation

Figure V.2: Radial one-body density of 2 non-interacting bosons in 3 dimmensions, using $N = 1e6$ cycles, $\omega = 1$, $\alpha = 0.5$.

B. Blocking of Local Energy

In Figure V.3, we see the effect of applying blocking on the local energy data before calculating the variance of the estimator (kilde). Although the local energies are sampled from an approximately correct distribution, as indicated by Figure V.2, the data is produced by moving a single electron at a time. Moreover, the move may even be rejected by the metropolis algorithm. Because of this, the data is highly correlated, causing an underestimation of the variance as seen for the unblocked data(blocking strength 0). After repeated blocking, the variance stabilizes when the data is approximately uncorrelated. Note that this happens at different strengths of blocking for brute force sampling and importance sampling. This indicates that latter method creates data that are less correlated. This makes sense, as importance sampling uses additional information from the wave function and explores the space of states more efficiently than brute force. Figure V.3 shows that the variance stabilizes at strength 6 when using importance sampling, as opposed to 9 when using brute force. This grants us 8 times the effective amount of data and thus a smaller variance of the estimator.



(a) Brute Force Sampling

Figure V.3: Variance of the estimate of $< E >$ using blocked values of local energy at various strengths. The data has been produced for 40 non-interacting bosons i 3 dimension, using $N = 2^{20}$ cycles, $\alpha = 0.8$, $\omega = 1$, with and without importance sampling

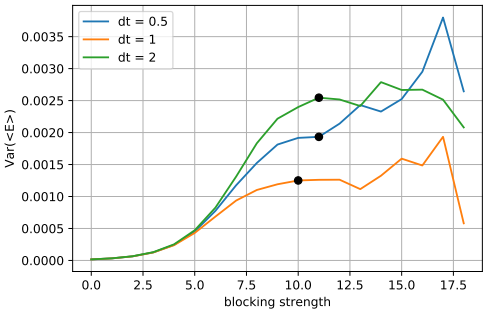


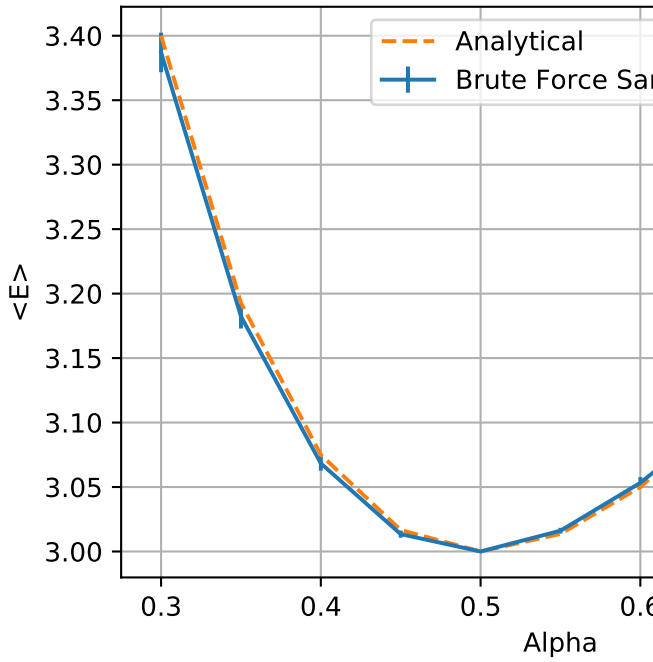
Figure V.4: Variance of the estimate of $< E >$ using blocked values of local energy at various strengths. The data has been produced for 40 non-interacting bosons i 3 dimension, using $N = 2^{20}$ cycles, $\alpha = 0.8$, $\omega = 1$, using importance sampling with step length 0.5, 1 and 2, respectively

Figure V.4 repeats the results of Figure V.3, but for a larger system of 40 non-interacting bosons using importance sampling only. Step lengths of 0.5, 1 and 2 was used. For a larger system, the local energies will tend to become more correlated as the moving of a single particle becomes a comparatively small change. This can be seen from the figure, as a higher order of blocking is required to stabilize the variance than in the two particle case.

Further, the least correlated data was produces for step length $\delta t = 1$, requiring blocking strength 10 as opposed to 11 for $\delta t = 0.5$ and 2. Using a step length

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



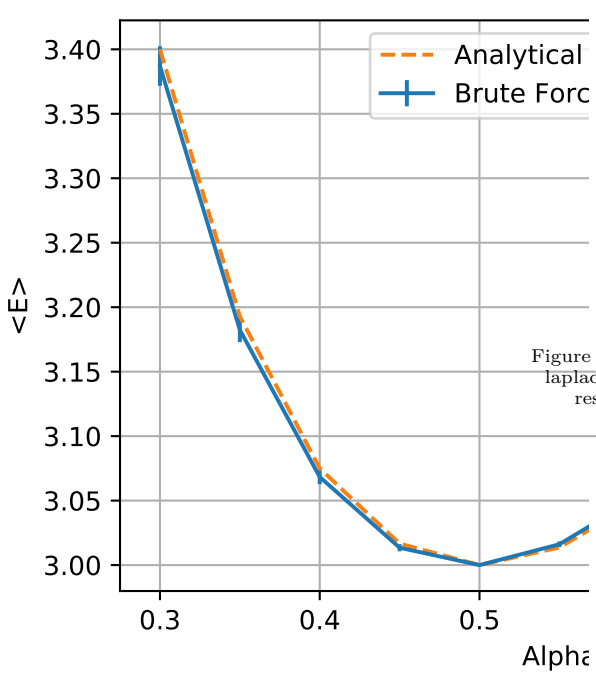
(a) Estimate of $\langle E \rangle$

Figure V.5: Estimated energy $\langle E \rangle$ and variance $V(E)$ for two non-interacting bosons, using 2^{17} cycles, $\omega = 1$ and brute force sampling. Errorbars was established using blocking. Teh analytical result is plotted for comparison.

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 drastic moves will tend to be rejected often, causing repeated values to be sampled. Table V.1 presents the resulting estimates of energy with the uncertainty established using the blocking. This shows that 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

As seen from Figure V.5, the estimated energy is consistent with the analytical within the uncertainty obtained using blocking. For $\alpha = 0.5$, the estimated value coincide with the analytical without any uncertainty, and $V(E)$ becomes zero. This is expected, as our trial



(a) Estimate of $\langle E \rangle$

Figure V.6: Estimated energy $\langle E \rangle$ and variance $V(E)$ for two non-interacting bosons, using 2^{17} cycles, $\omega = 1$ and importance sampling. Errorbars was established using blocking. The analytical result is plotted for comparison.

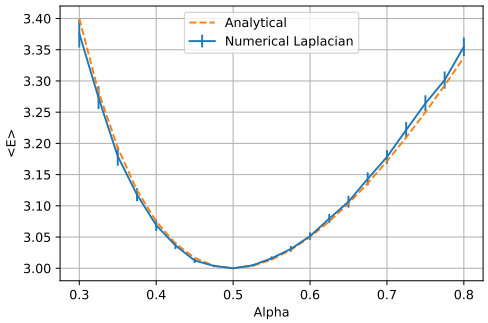


Figure V.7: Estimated energy $\langle E \rangle$ for two non-interacting bosons, using numerically evaluated laplacian, 2^{17} cycles, $\omega = 1$ and importance sampling. Errorbars was established using blocking.

wave function becomes the exact solution for $\alpha = 0.5$ for the non-interacting system.

Figure V.6 presents the same data as Figure V.5, but using importance sampling. The data is consistent with the previous discussion, showing that importance sampling reproduces the correct results with respect to the estimation of the energy.

D. Numerical and Analytical Laplacian

Yet again, Figure V.7 repeats the analysis of Figure V.5 and Figure V.6, but using finite difference for evaluating the Laplacian of the wave function rather than the analytical expression. The resulting estimate is within the statistical error of the Monte-Carlo simulation, showing that finite difference implementation is stable and of sufficient accuracy.

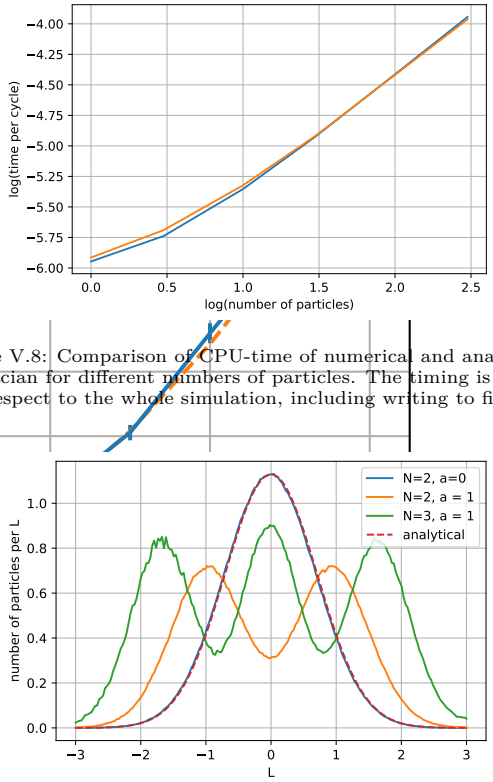


Figure V.8: Comparison of CPU-time of numerical and analytical laplacian for different numbers of particles. The timing is with respect to the whole simulation, including writing to file

Figure V.9: One body density in one dimension, for two non-interacting bosons, two interacting bosons, and three interacting bosons. $N = 400000$, $\alpha = 0.5$, $\omega = 1$, $a = 1$. The densities are compared to the analytical one body density for two non-interacting bosons.

Figure V.8 demonstrates that the CPU-time when using the analytical Laplacian is smaller than when using the finite difference method, for 1 to 30 non-interacting bosons. However, the speedup is small and hard to quantify, and even more negligible for systems of more particles. This is perhaps because, ultimately, in realistic simulations, a great portion of the CPU-time goes to other tasks such as writing the positions to file. The increase in CPU-time because of the numerical evaluation is not so big in comparison.

E. Interacting Potentials

?? presents the 1 dimmensional onebody density for two non-interacting bosons, two interacting bosons of radius $a = 1$, and three interacting bosons of the same radius. For the non-interacting case, the interacting model was used to produce the samples with the hard shell radius set to $a = 0$. Comparing it to the analyti-

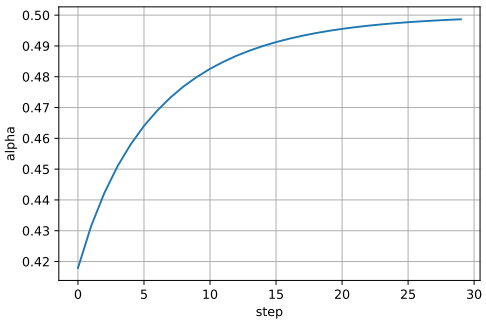
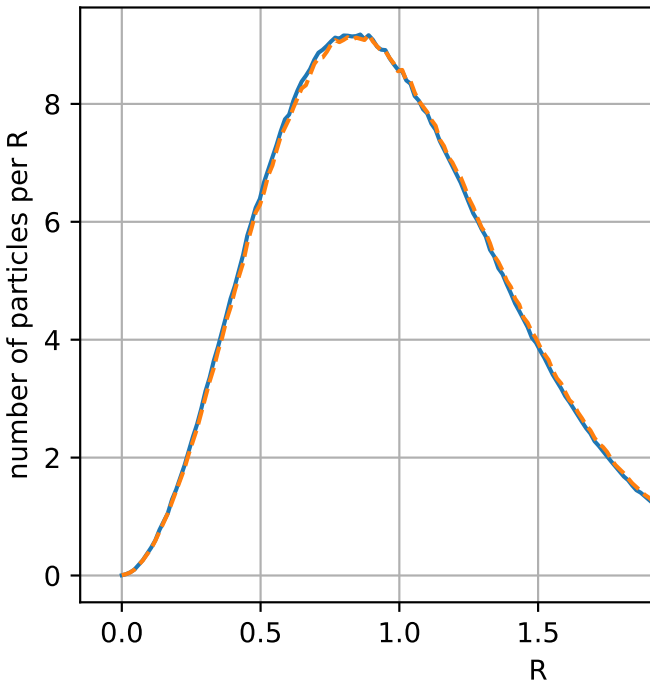


Figure V.10: Gradient decent on the parameter α for two non-interacting bosons, learning rate $\mu = 0.01$, $N = 10000$ steps and $\omega = 1$

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 V.2: 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



(a) Radial onebody density for 10 bosons in elliptical trap, interac

Figure V.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 $\beta = \gamma = 2.82843$.

cal result, we see that the implementation is consistent with the non-interacting case. For $a = 1$, we see that the correlating term produces a repulsive effect, creating two distinct maxima of relatively higher density. For the case of three interacting bosons, we instead get three maxima. Intuitively, repelling particles seek to maximize the distance from eachother, while still being confined by the harmonic potential.

F. Optimal Parameter α

Figure V.10 shows the result of gradient descent used to minimize the estimated energy $\langle E \rangle$ with respect to α. For the non-interacting case, the method converges to $\alpha = 0.5$, which we know to be the analytical value for the exact wave function. This establish confidence that the method of gradient decent also yields the optimal α for the interacting case, where a analytical solution is not known.

G. Interacting Bosons in Elliptical Potential

Table V.2 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 is it to spread out. This is consistent with the 1D result presented in Figure V.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 a accurat estimate for larger systems, a increase in the number of cycles is appropriate. However, because of limited computer power, this was neglected.

A comment on the interpretation of the error presented in Table V.2 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.

VI. CONCLUSION

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

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

Combining the results and diving through by Ψ_T then gives

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

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

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

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

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

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

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

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

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

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

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

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