Variational Monte Carlo studies of Bose gas in harmonic oscillator trap

Kaspara Skovli Gåsvær & Peder Lon Hauge*
University of Oslo - Department of Physics

(Dated: April 6, 2021)

The main focus of this report is the study of how variational Monte Carlo methods can be used to describe a boson gas trapped in a harmonic oscillator well. In particular, we estimate the expectation value of the Hamiltonian when using a trial wave function with a single variational parameter α. We perform calculations for non-interacting bosons in spherical traps as well as interacting bosons in elliptical traps, optimizing α using gradient descent. In the case of the non-interacting bosons, estimates of $\langle E_L \rangle$ were produced with a relative standard error of order 10^{-7}, and in accordance with analytical expressions for all N. The implementation of the interacting case was not optimal but resulted in $\langle E_L \rangle$ estimates with standard errors of magnitude 10^{-3} less than the magnitude of $\langle E_L \rangle$ itself. We calculated one-body densities for both types of particles and found that the correlations introduced by the Jastrow factor narrowed the radial density curve, skewing it closer to the origin. The energy per particle for interacting particles did however grow as N increased, suggesting larger particle spread compared to non-interacting particles.

I. INTRODUCTION

Analytical solutions to quantum mechanical problems are generally hard to find. In order to study systems beyond the simplest ones (such as the hydrogen atom or a single particle in an energy potential), numerical methods are necessary. This is especially true when allowing for interaction between particles. In this articles we aim to study the ground state energy of an interacting particle Bose gas, which is done using the variational Monte Carlo (VMC) method.

The specific system to be studied is an arbitrary number of hard-sphere bosons trapped in a harmonic oscillator potential. By guessing the shape of the system's state function, the variational principle guarantees that the calculated ground state energy $\langle H \rangle$ will be an upper bound of the real one. This energy $\langle H \rangle$ is minimized by varying parameters in the trial wave function, hence the name "variational method".

The calculation of $\langle H \rangle$ itself is done by sampling an appropriate probability distribution function. The sampling is done with a Markov chain Monte Carlo method utilizing the Metropolis-Hastings algorithm. With this method, the task of doing a number of $2^{20} \approx 10^6$ such samples will be of no problem for a computer.

After establishing the main principles behind the variational principle and Monte Carlo calculations, we introduce several new refinements to the basic theory. This includes *importance sampling*, which provides a more efficient way of sampling configuration space than random walks, and *gradient descent*, which makes the search for optimal variational parameters much faster. In addition, the method known as *blocking* will be used for improving the statistical analysis of produced data.

Our results will include comparisons to both analytical solutions for simplified systems, and to relevant figures in articles by DuBois [1] and Nilsen [2]. Finally, we will compute a quantity known as the *one-body density*, which illustrates how the particles in the boson gas are distributed through the potential trap.

II. THEORY

A. The variational principle

The variational principle is a way to approximate to ground state energy E_{gs} of a quantum mechanical system. We assume that the system has a known Hamiltonian H, but that its wave function Ψ (with the corresponding ground state energy E_{gs}) is unknown.

It can be shown that the expectation value for H with *any* normalizable state function Ψ_T will be an upper bound for the true ground state energy, i.e.

$$\langle H \rangle \equiv \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = \frac{\int \mathrm{d}\mathbf{r} \Psi_T^* H \Psi_T}{\int \mathrm{d}\mathbf{r} |\Psi|^2} \geq E_{\text{gs}}. \quad (1)$$

A proof of the above statement is given in Griffiths [3]. The function Ψ_T is called a trial wave function, and it is usually dependent on a set of variational parameters $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, so that $\Psi_T = \Psi_T(\mathbf{r}; \boldsymbol{\alpha})$. After calculating $\langle H \rangle$ for some general $\boldsymbol{\alpha}$, values of these parameters are chosen so that $\langle H \rangle$ is minimized. This will be the best approximation of E_{gs}.

In theory, any wave function Ψ_T can be used with this method, but it's preferable to use physical intuition to chose a Ψ_T that resembles the true state in order to minimize the energy [3].

Since the true ground state Ψ is an eigenstate of H, the variance in energy σ^2_{std} is zero. Assuming Ψ to be normalized ($\langle \Psi | \Psi \rangle = 1$), we have

$$\sigma^2_{\text{std}} = \langle \Psi | H^2 | \Psi \rangle - \langle \Psi | H | \Psi \rangle^2 = E^2_{\text{gs}} (\langle \Psi | \Psi \rangle - \langle \Psi | \Psi \rangle^2) = 0. \quad (2)$$

* Code repository: https://github.com/pederlh/FYS4411

A benchmark of how close σ^2_{std} is to zero (for a trial wave function Ψ_T) is thus a benchmark of how close $\langle H \rangle$ is to the real ground state energy E_{gs}.

B. Hamiltonian of system

The system to be studied is a hard-sphere Bose gas (boson gas) trapped in a harmonic oscillator (HO) potential. In the most general case, we consider N identical particles in D spacial dimensions. The Hamiltonian of the system can then be stated as

$$H = \sum_{i=1}^{N} \left(\frac{-\hbar^2}{2m} \nabla_i^2 + V_{\text{ext}}(\mathbf{r}_i) \right) + \sum_{i<j}^{N} V_{\text{int}}(\mathbf{r}_i, \mathbf{r}_j), \quad (3)$$

with m as the particle mass, $\hbar$ the reduced Planck constant, V_{ext} the potential from the external HO potential, and V_{int} the interaction potential between the particles. The shorthand notation

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

signifies running over the all the pairwise interactions once.

Modeling the bosons with a hard-core diameter a, the inter-boson interaction is represented by a repulsive potential

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

The harmonic oscillator potential V_{ext} will typically be spherical, but in the three-dimensional cases ($D = 3$) we will also consider an elliptical harmonic trap. The potential can then be written as

$$V_{\text{ext}}(\mathbf{r}) = \begin{cases} \frac{1}{2} m \omega_{\text{ho}}^2 r^2, & \text{(S)} \\ \frac{1}{2} m \left[\omega_{\text{ho}}^2 (x^2 + y^2) + \omega_z^2 z^2 \right], & \text{(E)}, \end{cases} \quad (5)$$

where ω_{ho} and ω_z are trap frequencies, and (S) and (E) signifies "spherical" and "elliptical", respectively.

C. Trial wave function

The trial wave function used for the ground state of the N particle Bose gas is given by

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

with α and β as variational parameters.

We assume the ground state to be a product state of single-particle HO ground states (up to a proportionality factor)

$$g(\alpha, \beta, \mathbf{r}_i) = \begin{cases} \exp\left(-\alpha r_i^2\right), & \text{(S)} \\ \exp\left[-\alpha (x_i^2 + y_i^2 + \beta z_i^2)\right], & \text{(E)}, \end{cases} \quad (7)$$

and correlation wave functions

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

The product of correlation wave functions is also called the *Jastrow factor*.

If the trap potential is spherical, we can immediately set $\beta = 1$. The correlation wave function respects the the interaction potential V_{int} given in Eq. (4); if the distance between any two particles is less then a, the trial wave function collapses to $\Psi_T = 0$.

D. Scaling and fixed parameters

The mean square displacement for a particle in a harmonic oscillator trap (in the zero temperature ground state) is $\langle x^2 \rangle = 2\hbar / m\omega_{\text{ho}}$. The system can be scaled by defining a characteristic length $a_{\text{ho}} \equiv (\hbar / m\omega_{\text{ho}})^{\frac{1}{2}}$, so that the dimensionless positions are

$$\mathbf{r}' = \frac{\mathbf{r}}{a_{\text{ho}}}.$$

The Laplace operator for term k in the Hamiltonian is then

$$\nabla_k'^2 = \frac{1}{a_{\text{ho}}^2} \nabla_k^2,$$

as its components are on the form $\frac{\partial}{\partial x_k'} = \frac{\partial}{\partial (a_{\text{ho}} x_k)}$.

In addition, we denote the ratio of frequencies $\lambda \equiv \omega_z / \omega_{\text{ho}}$. The interaction part of the Hamiltonian is unaffected by the scaling, but the one-body part of it changes to

$$\begin{aligned} H_{\text{OB}} &= \sum_{i=1}^{N} \left[\frac{-\hbar^2}{2m} \frac{1}{a_{\text{ho}}^2} \nabla_i'^2 + \frac{1}{2} \omega_{\text{ho}}^2 a_{\text{ho}}^2 \left(x_i'^2 + y_i'^2 + \lambda^2 z_i'^2 \right) \right] \\ &= \hbar \omega_{\text{ho}} \sum_{i=1}^{N} \frac{1}{2} \left(-\nabla_i'^2 + \left(x_i'^2 + y_i'^2 + \lambda^2 z_i'^2 \right) \right). \end{aligned}$$

From the above equation it's evident that the Hamiltonian itself becomes dimensionless if scaled by a factor $\hbar \omega_{\text{ho}}$, such that $H' = H / \hbar \omega_{\text{ho}}$. The prime symbols can be removed, resulting in the full dimensionless Hamiltonian (in three dimensions) below:

$$H = \sum_{i=1}^{N} \frac{1}{2} \left(\nabla_i^2 + x^2 + y^2 + \lambda^2 z^2 \right) + \sum_{i<j}^{N} V_{\text{int}}(|\mathbf{r}_i - \mathbf{r}_j|). \quad (9)$$

The single-particle states of the wave function do not need to be scaled (Eq. (7)); the main change is in the hard-sphere radius a in Eq. (8). The scaling is done by letting $a \mapsto a/a_{\text{ho}}$, a value which will be fixed to $a/a_{\text{ho}} = 0.0043$ (as done in Refs. [1, 2]).

In addition, Refs. [1, 2] set $\lambda = \beta = \sqrt{8} \approx 2.82843$. We will also implement these values as this will allow for a later comparison of results.

E. Rewrite in terms of probability distribution

By introducing the probability distribution function (PDF)

$$P(\mathbf{r}) = \frac{|\Psi_T(\mathbf{r})|^2}{\int \mathrm{d}\mathbf{r}\, |\Psi_T(\mathbf{r})|^2} \tag{10}$$

and a quantity called the local energy

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

we can rewrite the expectation value $\langle H \rangle$ from Eq. (1) as

$$\langle H \rangle = \int \mathrm{d}\mathbf{r}\, P(\mathbf{r}) E_L(\mathbf{r}). \tag{12}$$

It can readily be seen that $P(\mathbf{r})$ fulfills the postitivity and normalization ($\int \mathrm{d}\mathbf{r}\, P(\mathbf{r}) = 1$) requirements for a PDF.

Instead of calculating $\langle H \rangle$ by numerical integration with Newton-Cotes (fixed interval) formulas, we can find $\langle H \rangle$ by sampling the local energy from the PDF. The law of large numbers tells us that

$$\langle H \rangle = \lim_{M \to \infty} \left[\frac{1}{M} \sum_{m=1}^{M} E_L(\mathbf{r}) \right] = \int \mathrm{d}\mathbf{r}\, P(\mathbf{r}) E_L(\mathbf{r}) = \langle E_L \rangle, \tag{13}$$

where M is the total number of samples. The above equation also shows that $\langle E_L \rangle = \langle H \rangle$.

F. Non-interacting bosons: analytical solution

If we study the case of non-interacting bosons in a spherical trap, we can set $a = 0$ (and $\beta = 1$) in the trial function (Eq. (6)). This simplification of $\Psi_T(\mathbf{r})$ makes it possible to find both an analytical expression local energy $E_L(\mathbf{r})$ and the value of variational parameter α that minimizes $\langle H \rangle$.

It can be shown that for N particles in D spacial dimensions, the local energy is given by

$$E_L(\mathbf{r}, \alpha) = DN\alpha + \frac{(1 - 4\alpha^2)}{2} \sum_{n=1}^{N} r_n^2. \tag{14}$$

One can also show that $\langle H \rangle$ is minimized when $\alpha = \frac{1}{2}$, and this energy becomes an upper bound of the system's ground state energy. This bound is found to be

$$\langle H \rangle_{\text{min}} = \frac{D}{2} N. \tag{15}$$

The above statements are proved in Appendix A. By choosing $\alpha = \frac{1}{2}$, the local energy from Eq. (14) simplifies to a constant $E_L = \frac{D}{2} N$, the same expression found for $\langle H \rangle_{\text{min}}$ above.

G. Interacting bosons: local energy

In the case of interacting bosons we will restrict ourselves to study systems in three-dimensional space, but allow for elliptical traps ($\beta > 0$). This case do not have analytical values of $\langle H \rangle$, but we can still find an expression for parts of the local energy $E_L(\mathbf{r})$.

Using the definition of the local energy given in Eq. (11) and the Hamiltonian from Eq. (3) (in natural units), a general expression for $E_L(\mathbf{r})$ is

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

It is possible to find an analytical expression for $\frac{\nabla_k^2 \Psi_T}{\Psi_T}$, which describes how the Laplacian for a given particle k operates on the trial wave function. In order to do this, we rewrite $\Psi_T(\mathbf{r})$ to the form

$$\Psi_T(\mathbf{r}) = \left[\prod_{i=1}^{N} \phi(\mathbf{r}_i) \right] \exp\left(\sum_{j<k} u(r_{jk}) \right), \tag{16}$$

where we have defined $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j|$, $\phi(\mathbf{r}_i) \equiv g(\alpha, \beta, \mathbf{r}_i)$ and $\exp(u(r_{ij})) \equiv f(a, |\mathbf{r}_i - \mathbf{r}_j|)$.

In Appendix B, we show that

$$\frac{\nabla_k^2 \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{\nabla_k^2 \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2 \frac{\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} \left(\sum_{l \neq k} \frac{(\mathbf{r}_k - \mathbf{r}_l)}{r_{kl}} u'(r_{kl}) \right) + \left(\sum_{l \neq k} \frac{(\mathbf{r}_k - \mathbf{r}_l)}{r_{kl}} u'(r_{kl}) \right)^2 + \sum_{l \neq k} \left(u''(r_{kl}) + \frac{2u'(r_{kl})}{r_{kl}} \right), \tag{17}$$

which includes the quantities

$$\nabla_k \phi(\mathbf{r}_k) = -2\alpha (x_k \mathbf{e}_1 + y_k \mathbf{e}_2 + \beta z_k \mathbf{e}_3) \phi(\mathbf{r}_k), \tag{18}$$

$$\nabla_k^2 \phi(\mathbf{r}_k) = \left[4\alpha^2 (x_k^2 + y_k^2 + \beta^2 z_k^2) - 2\alpha(2 + \beta) \right] \phi(\mathbf{r}_k), \tag{19}$$

$$u'(r_{kl}) = \frac{\mathrm{d}u(r_{kl})}{\mathrm{d}r_{kl}} = \frac{a}{r_{kl}(r_{kl}-a)}, \tag{20}$$

and

$$u''(r_{kl}) = \frac{\mathrm{d}^2u(r_{kl})}{\mathrm{d}r_{kl}^2} = \frac{a^2 - 2ar_{kl}}{r_{kl}^2(r_{kl}-a)^2}. \tag{21}$$

H. Analytical expression for quantum force

When introducing the concept of importance sampling later on, a central quantity is the quantum force defined as

$$\mathbf{F} = \frac{2\nabla\Psi_T}{\Psi_T}. \tag{22}$$

This quantity can be found analytically. For the non-interacting particles ($a = 0$), the quantum force on the particle indexed k is

$$\mathbf{F}_{k,\mathrm{non-int}} = -4\alpha\mathbf{r}_k, \tag{23}$$

while for interacting particles ($a \neq 0$) we can find

$$\mathbf{F}_{k,\mathrm{int}} = 2\frac{\nabla_k\phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2\sum_{l\neq k}\frac{(\mathbf{r}_k - \mathbf{r}_l)}{r_{kl}}u'(r_{kl}). \tag{24}$$

with expressions for $\nabla_k\phi(\mathbf{r}_k)$ and $u'(r_{kl})$ given in Eq. (18) and Eq. (20), respectively. The two equations (23) and (24) are derived in Appendix C.

I. One-body density

One way of describing the distribution of particles in the harmonic oscillator trap is through the calculation of one-body densities. In particular, we can find this probability density when centered at any arbitrary particle. The expression for the one-body density for a particle located at position $\mathbf{r}_1$ is

$$\rho(\mathbf{r}_1) = \int_{-\infty}^{\infty}\mathrm{d}\mathbf{r}_2\cdots\mathrm{d}\mathbf{r}_N\,|\Psi(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_N)|^2. \tag{25}$$

For a normalized wave function, integrating up the one-body density should result in the total number of particles N by convention, so that $\int_{-\infty}^{\infty}\mathrm{d}\mathbf{r}_1\,\rho(\mathbf{r}_1) = N$. Note that due to the bosons being identical and indistinguishable particles, the position $\mathbf{r}_1$ in the above definition can refer to any of the particles.

While the volume element $|\Psi|^2\,\mathrm{d}\mathbf{r}_k$ describes the probability of finding particle k at some given space, the one-body density volume element $\rho(\mathbf{r}_k)\,\mathrm{d}\mathbf{r}_k$ tells the probability of finding *any* of the N particles at the same location.

We will study the radial one-body density $\rho_r(r_1)$, which integrates out all coordinates besides a radial one (after transforming to spherical coordinates.)

Given the trial wave function corresponding to the spherical trap with non-interacting particles, the analytical expression for the radial one-body density can be found to be

$$\rho_r(r) = \frac{N}{\sqrt{\pi}}2^{\frac{7}{2}}\alpha^{\frac{3}{2}}e^{-2\alpha r^2},$$

and inserting the solution $\alpha = 1/2$ results in

$$\rho_r(r) = \frac{4N}{\sqrt{\pi}}e^{-r^2}. \tag{26}$$

We will in the Method section discuss how to calculate the one-body density in a practical manner.

III. METHOD

A. Variational Monte Carlo

1. *Markov Processes*

We need a way to sample random walkers for our system, which we do by a Markov process. If presented with a system in state i, a Markov process generates a new state j with the help of a transition probability $P(i \to j)$ which tells us the probability of the system moving from state i to state j. An important characteristic of a Markov process is that $\sum_j P(i \to j) = 1$, which is to say that if the system is in a state i then the Markov process must put it in a state j. This includes the case $j = i$, and so it can be put in a new state equal to the old state. Other important aspects of Markov processes are the principles of *ergodicity* and *detailed balance*. The condition of ergodicity states that with a long enough Markov chain all states j can be reached from state i, while the condition of detailed balance states that in equilibrium, the transition $i \to j$ should happen just as often as $j \to i$. This is equivalent to

$$p_iP(i \to j) = p_jP(j \to i), \tag{27}$$

where p_i is the probability of the system being in a state i if it were to be sampled from the desired distribution.

2. *Metropolis algorithm*

The first way of performing Monte Carlo calculations is by brute force. We now introduce the Metropolis algorithm. The idea is to generate a new state with some transition probability, and then determine if the state is

to be accepted or rejected using an acceptance rate. We therefore rewrite the transition probability defined by the Markov process as

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

where $T(i \to j)$ is the transition probability for moving fra state i to state j and $A(i \to j)$ is the acceptance rate. We then sample new states using the transition probability and accept/reject the transition using the acceptance rate. The goal of using the Metropolis algorithm is to ensure that after enough steps are taken, the probability of being in a state i at step n , P_i^n, becomes the desired probability distribution p_i. This requires us to find suitable probabilities T and A. Inserting our new expression for $P(i \to j)$ into the equation for detailed balance, (Eq. (27)), and rearranging a bit gives us

$$\frac{A(j \to i)}{A(i \to j)} = \frac{T(i \to j)}{T(j \to i)}\frac{p_i}{p_j}. \tag{29}$$

Now a common choice of acceptance rate which fulfills the above condition is the Metropolis choice, which is to maximize the value of $A(i \to j) = 1$. This let's us set

$$A(j \to i) = \min\left(1, \frac{T(i \to j)}{T(j \to i)}\frac{p_i}{p_j}\right). \tag{30}$$

By choosing this acceptance rate we have ensured that once our system equilibrates, the probability P_i^n both becomes and stays p_i.

3. Brute Force

Now we move on to what happens when using the brute force method. Given an initial state i with corresponding initial positions $\boldsymbol{r}_i$ which are used to calculate the value of the trial function Ψ_i. We calculate a proposed move of one of the random walkers in $\boldsymbol{r}_i$, and the coordinates of this arbitrarily chosen kth particle is $\boldsymbol{r}_i^k$. Defining a step-size h and a number g generated from a uniform distribution, the proposed move is

$$\boldsymbol{r}_j^k = \boldsymbol{r}_i^k + hg. \tag{31}$$

This new proposed state j has positions $\boldsymbol{r}_j$, and is used to calculate a new trial function value Ψ_j. We sample from a uniform distribution where all moves are equally likely, which leads to a symmetric transition probability, $T(i \to j) = T(j \to i)$. The acceptance rate is proportional to the value of the trial wave function's magnitude. Thus, we let

$$A(j \to i) = \min\left(1, \frac{|\Psi(\mathbf{r}_i)|^2}{|\Psi(\mathbf{r}_j)|^2}\right). \tag{32}$$

Whether or not the move is accepted or rejected we sample the local energy and from Eq. (13) its clear that these samples can be used to calculate the expectation value of the Hamiltonian.

4. Importance sampling

In order to achieve more efficient calculations, we move on from the basic brute-force Monte Carlo method by introducing importance sampling. The key element is that the random walkers used in every MC cycle are replaced with walkers guided in configuration space. The direction of the introduced bias is determined by the trial wave function itself.

This approach is based is based on the Fokker-Planck and Langevin equations, which are both intimately connected to diffusion processes and Markov Chains. The subsequent paragraphs will only state the main results needed for the implementation of importance sampling; a theoretical justification can be found in textbooks in stochastic processes (see Refs. [4] and [5]).

The Fokker-Planck equation describes the time-dependent probability density $P(\mathbf{r}, t)$ that characterizes an isotropic diffusion process. The equation is given as

$$\frac{\partial P}{\partial t} = \sum_i D\frac{\partial}{\partial r_i}\left(\frac{\partial}{\partial r_i} - F_i\right)P(\mathbf{r}, t), \tag{33}$$

where D is the diffusion coefficient and F_i is the drift term for the ith component (which is caused by an external potential). The time-derivative vanishes if the probability distribution is stationary or converges at $t \to \infty$, which means that all terms in the above sum adds up to zero. This yields the condition

$$\frac{\partial^2 P}{\partial r_i^2} = P\frac{\partial}{\partial r_i}F_i + F_i\frac{\partial}{\partial r_i}P.$$

Going on, we let the drift vector $\mathbf{F}$ have components of the form $F_i = g_i(P, \mathbf{r})\frac{\partial P}{\partial r_i}$. When inserting this into the above equation it can be shown that the only possible solution which fulfills the requirement of convergence is $g = \frac{1}{P}$. Using the probability density defined in Eq. (10), we find the drift term for a real and positive trial function Ψ_T to be

$$\mathbf{F} = \frac{1}{P}\nabla P = \frac{2\nabla\Psi_T}{\Psi_T},$$

a quantity earlier introduced in Eq. (22). The drift term is also referred to as the *quantum force*, as it drives the walkers to parts of the configuration space where $|\Psi_T|^2$ is large.

The sampling of new positions is done by looking at the solutions of the Langevin equation

$$\frac{\partial \mathbf{r}_i}{\partial t} = D\mathbf{F}(\mathbf{r}_i(t)) + \boldsymbol{\eta}, \tag{34}$$

where $\boldsymbol{\eta}$ is a random variable. Discretization using Euler's method leads to an expression for the updated position

$$\mathbf{r}_j \equiv \mathbf{r}_{i+1} = \mathbf{r}_i + D\mathbf{F}(\mathbf{r}_i)\Delta t + \boldsymbol{\xi}\sqrt{\Delta t}. \tag{35}$$

Here, $\boldsymbol{\xi}$ is a Gaussian random variable, and Δt is a chosen time step. $\boldsymbol{\xi}$ should be centered at zero, and is commonly chosen to have a standard deviation of 1. Note that the diffusion coefficient has a value $D = \frac{1}{2}$ when working in natural units (from the kinetic energy term).

The solution of the Fokker-Planck equation yields a transition probability, which takes the form of a Green's function

$$\begin{aligned} G(\mathbf{r}_j, \mathbf{r}_i, \Delta t) = &\frac{1}{(4\pi D\Delta t)^{3N/2}} \\ &\times \exp\left[\frac{-(\mathbf{r}_j - \mathbf{r}_i - D\Delta t\mathbf{F}(\mathbf{r}_i))^2}{4\pi D\Delta t}\right]. \end{aligned} \tag{36}$$

In comparison to Eq. (32), the new acceptance rate for the Monte Carlo simulation becomes

$$A(j \to i) = \min\left(1, \frac{G(\mathbf{r}_j, \mathbf{r}_i, \Delta t)}{G(\mathbf{r}_i, \mathbf{r}_j, \Delta t)} \frac{|\Psi(\mathbf{r}_i)|^2}{|\Psi(\mathbf{r}_j)|^2}\right). \tag{37}$$

B. Gradient descent

A method for minimizing $\langle H \rangle$ through the parameters $\boldsymbol{\alpha}$ is to implement a gradient descent algorithm. In gradient descent, we initialize this parameter as some $\boldsymbol{\alpha}_0$ and iteratively update it in the direction of the steepest descent through the scheme

$$\begin{aligned} \boldsymbol{v}_t &= \eta \nabla_\alpha \langle E_L(\boldsymbol{\alpha}) \rangle \\ \boldsymbol{\alpha}_{t+1} &= \boldsymbol{\alpha}_{\boldsymbol{t}} - \boldsymbol{v}_t. \end{aligned} \tag{38}$$

As seen in the above equations, we make use of the gradient of the expectation value of the local energy with respects to the parameters $\boldsymbol{\alpha}$ at the time step t. The learning rate η determines the size of the step taken in the direction of the gradient, which makes it a parameter to be optimized in its own to ensure fast and precise convergence to a minimum.

Choosing a small learning rate is more computationally expensive, while choosing one too big increases the risk of overshooting the minimum depending on the convexity of the function to be minimized. In regular gradient descent, the learning rate is set to be a constant; in other words there is no stochasticity and all directions in the parameter space are treated uniformly.

As the trial wave funtion only has a single variational parameter α, the gradient in Eq. (38) reduces to an ordinary derivative.

Defining

$$\bar{E}_\alpha \equiv \frac{\mathrm{d}\langle E_L(\alpha) \rangle}{\mathrm{d}\alpha} \quad \text{and} \quad \bar{\Psi}_\alpha \equiv \frac{\mathrm{d}\langle \Psi(\alpha) \rangle}{\mathrm{d}\alpha}, \tag{39}$$

it is can be shown that

$$\bar{E}_\alpha = 2\left(\left\langle \frac{\bar{\Psi}_\alpha}{\Psi} E_L \right\rangle - \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} \right\rangle \langle E_L \rangle\right). \tag{40}$$

A derivation of this equation is given in Appendix D.

C. Blocking (resampling technique)

When samples measurements are independent and identically distributed (iid) for a probability distribution X, the standard error σ of the mean $\langle X \rangle$ is

$$\sigma = \frac{\sigma_{\text{std}}}{\sqrt{n}}, \tag{41}$$

where n is the number of samples and σ_{std} is the standard deviation of the distribution itself. When the above requirement of iid samples doesn't hold, using the distribution variance leads to an overly optimistic take on the standard error. A better estimate of the uncertainty is the covariance, but this can be tricky and/or computationally expensive to calculate. We will tackle this problem using a resampling technique named "blocking", which provide easier and more effective calculations of the covariance.

Blocking is more suited than methods such as bootstrapping for larger datasets, as it halves the size of the dataset in each iteration rather than producing new datasets of the same size as the original. Let a dataset consist of 2^N measurements. With each iteration of the blocking method one takes the average of all subsequent neighboring samples in the dataset. This produces a new dataset of size 2^{N-1}, halving the size of the previous set. One repeats this until there is only one sample left, which can be interpreted as an estimation of the mean. The method also provides a better estimate of the standard error, which we will name σ_B. We will implement the method for blocking described in Ref. [6].

D. Equilibration time

Due to the random initialization of random walkers and the finite number of Monte Carlo cycles performed, starting the random walkers in an improbable area of configuration space might unnecessarily skew the calculated expectation values further away from the real expectation values.

It is thus customary to first perform what's called *equilibration cycles* or *burn-in cycles*. A freshly initialized system will run for a set amount of cycles of before we actually start recording data. We chose to set the amount of equilibration cycles to 10 % of the "proper" cycles. This fraction also applies to the individual iterations of the gradient descent process.

E. Radial one-body density histogram

As explained in Ref. [7], we can calculate the radial one-body density through calculating a histogram $H(r)$. Centering ourselves at the coordinates of particles 1's origin in phase space, we can count how many of the other particles fall into concentric shells of radius Δr.

Since the concentric shells have increasing volume, the histogram have to be scaled by the shell thickness in order to yield the one-body density. An algorithm that describes how this histogram can be created is outlined in Algorithm 1. Due to the trial function not being normalized, the last step is to introduce a normalization constant A, and normalize the radial one-body density so that $\int_0^\infty \mathrm{d}r\, r^2(A\rho_r(r)) = N$.

Algorithm 1 Radial one-body density

```
Calculate r_1                          ▷ Find radial coordinate of particle 1
Initialize num_bins                    ▷ Number of bins in histogram
Init. max_r                            ▷ Max radius in OBD calculation
Init. zeros-array Histogram            ▷ length num_bins
Init. linspace array r_vals            ▷ [0, max_r], length num_bins
Set Δr = max_r/num_bins
for i = 1, 2, ..., MC cycles do
    for j = 1, 2, ..., N do
        (Metropolis sampling)
        for k = 2, 3, ..., N do
            Calculate r                ▷ Radial component particle k
            bin_idx = floor(r/Δr)
            Histogram[bin_idx] += 1
Histogram /= (MC cycles * N)           ▷ Use average value
Init. zeros-array OBD                  ▷ length num_bins
for i = 0, 1, ..., num_bins do
    OBD[i] = Histogram[i] / ((4π/3)*((i+1)**3 - i**3)*Δr)
                                       ▷ Divide by volume of spherical shells
Normalize OBD
```

IV. RESULTS

A. Brute-force MC, non-interacting particles

1. Search of variational parameter

Using evenly spaced α-values in the interval $\alpha \in [0.3, 0.65]$, we tested the brute force Monte Carlo algorithm for $N = 1, 10, 100, 500$ particles. Studying the non-interacting boson gas in a spherical potential, we did 10^6 Monte Carlo samples for all values of α and N, and calculated $\langle E_L \rangle$ using both numerical differentiation (with step size 10^{-4}) and the analytical expression from Eq. (14).

Next, the calculations were compared to the analytical solution (Eq. (15)) in order to verify a correct implementation of the MC algorithm. At this point we assume that the samples are iid, and that the standard error σ can be calculated using Eq. (41). These results are presented in Figure 1. In addition, the average spent CPU time for different values of N is presented in Table I.

TABLE I. Average CPU time running 10^6 Monte Carlo cycles for calculating $\langle E_L \rangle$ (with brute-force sampling). Algorithms utilizing the analytical derivative expression and numerical differentiation are compared for various numbers N of particles.

N	Time (numerical) [s]	Time (analytical) [s]
1	7.249 ± 0.039	7.231 ± 0.038
10	7.601 ± 0.001	7.414 ± 0.167
100	11.698 ± 0.020	7.130 ± 0.002
500	30.059 ± 0.057	7.277 ± 0.072

B. Importance sampling

1. Determining time step parameter

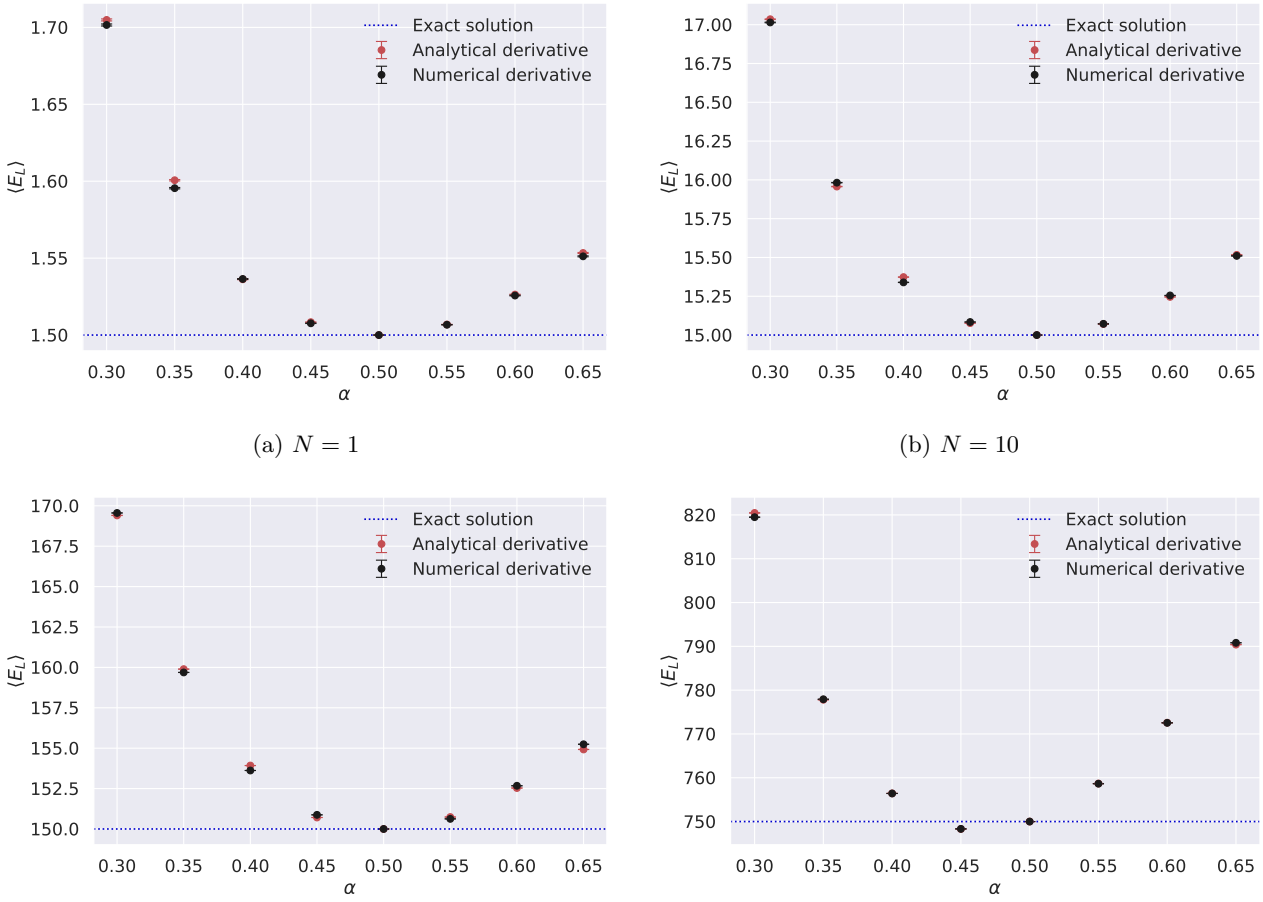
Implementing importance sampling requires assigning a value of the time step Δt in the Langevin equation solution (Eq. (35)). We performed Monte Carlo runs with values of $\Delta t \in [10^{-5}, 5]$, with the analytical expression for differentiation. The results found for $\langle E_L \rangle$ and the sample variance σ^2_{std} are presented in Table II.

TABLE II. Local energy calculated using analytical expression for the Laplacian with varying time steps Δt and variational parameters $\alpha \in \{0.45, 0.50, 0.55\}$. The $N = 100$ particles are placed in a 3D spherical HO-potential. Each experiment consists of 10^6 samples.

	$\alpha = 0.45$		$\alpha = 0.50$		$\alpha = 0.55$	
Δt	$\langle E_L \rangle$	σ^2_{std}	$\langle E_L \rangle$	σ^2_{std}	$\langle E_L \rangle$	σ^2_{std}
5.0	157.619	0.071	75.0	0.0[a]	−33.115	0.16
1.0	81.9	2.8	75.0	0.0	66.580	3.4
0.5	77.7	1.3	75.0	0.0	72.613	1.3
0.1	75.78	0.94	75.0	0.0	74.899	0.77
0.05	75.58	0.90	75.0	0.0	75.178	0.73
0.01	75.40	0.94	75.0	0.0	75.40	0.65
0.005	75.54	0.94	75.0	0.0	75.316	0.78
0.001	75.2	1.7	75.0	0.0	75.428	1.1
0.0005	74.7	1.5	75.0	0.0	75.47	1.2
0.0001	73.5	2.6	75.0	0.0	76.96	2.1
0.00005	72.3	3.1	75.0	0.0	78.1	2.1
0.00001	69.87	0.57	75.0	0.0	79.88	0.29

[a] Up to numerical precision. This applies to all entries in the column.

Looking at values of α close to the analytical solution



(a) $N = 1$

(b) $N = 10$

(c) $N = 100$

(d) $N = 500$

FIG. 1. The local energy $\langle E_L \rangle$ for various values of variational parameter α and number of particles N, calculated with brute force sampling. The standard errors assumes iid samples, and the analytical value for $\langle E_L \rangle$ is plotted as dotted blue lines.

$\alpha = 0.50$, one sees that choosing $\Delta t \in [0.005, 0.1]$ provides a sensible value of $\langle E_L \rangle$ and lower variance than for other choices. This justifies our use of $\Delta t = 0.005$ for all subsequent simulations with importance sampling.

2. *Search of variational parameter*

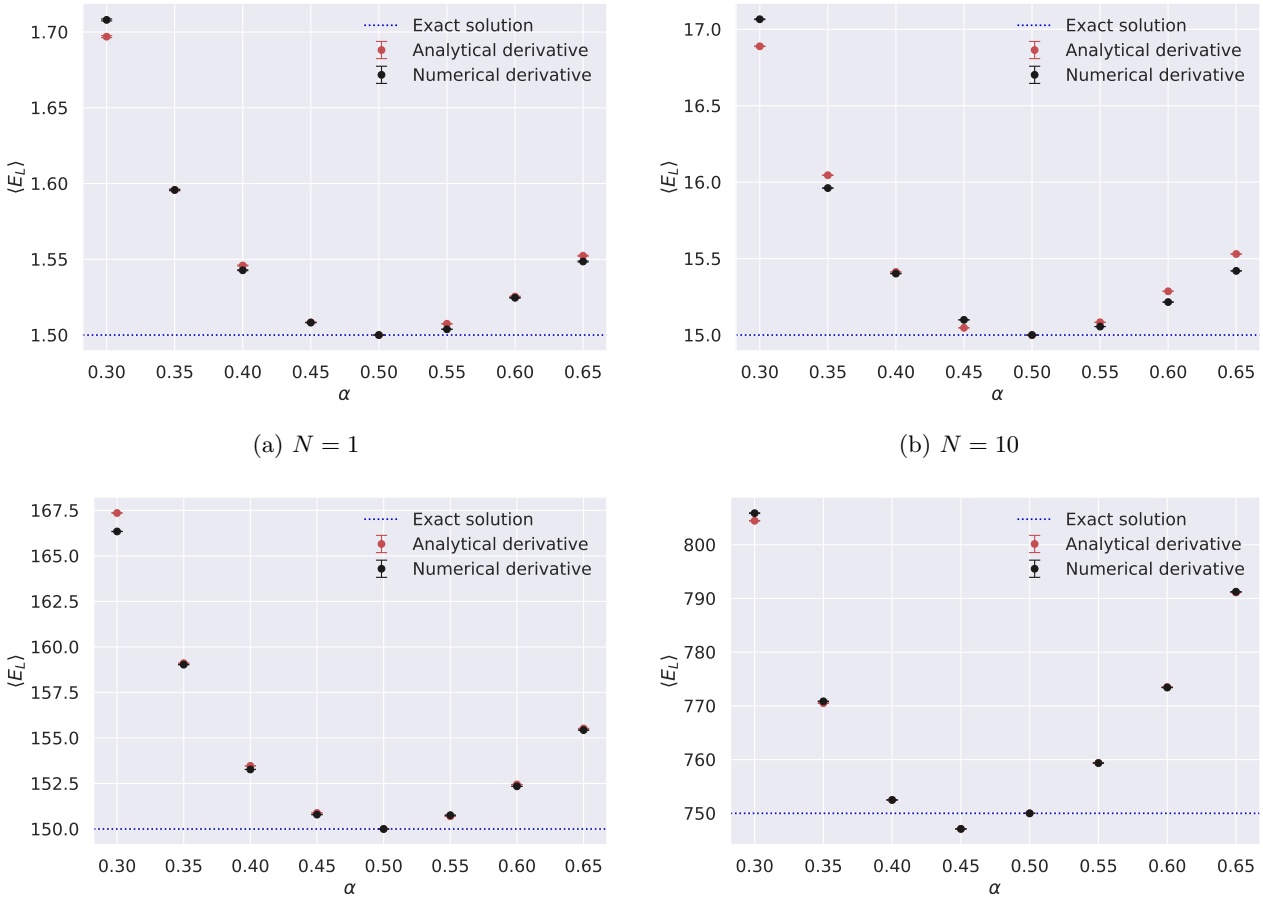
Setting $\Delta t = 0.005$, we repeated the experiment of finding the optimal value of α when presented with a list of evenly spaced $\alpha \in [0.3, 0.65]$, but this time using importance sampling in the Metropolis-Hastings algorithm. All other details are identical. As before, we made a figure showing calculations of $\langle E_L \rangle$ (Figure 2) and a table comparing CPU time (Table III).

TABLE III. Average CPU time running 10^6 Monte Carlo cycles for calculating $\langle E_L \rangle$ (with importance sampling). Algorithms utilizing the analytical derivative expression and numerical differentiation are compared for various numbers N of particles.

N	Time (numerical) [s]	Time (analytical) [s]
1	7.313 ± 0.004	7.237 ± 0.006
10	7.725 ± 0.001	7.244 ± 0.011
100	11.811 ± 0.004	7.251 ± 0.015
500	30.230 ± 0.084	7.231 ± 0.005

C. Gradient descent and blocking

While still studying the non-interacting particle gas in a spherical HO-potential, we implemented a gradient descent algorithm to find an optimal value for α. This was applied to the Monte Carlo method with importance Metropolis sampling and analytical derivative. When performing gradient descent we went with a rather



(a) $N = 1$

(b) $N = 10$

(c) $N = 100$

(d) $N = 500$

FIG. 2. The local energy $\langle E_L \rangle$ for various of variational parameters α and number of particles N, calculated with importance sampling. The standard errors assumes iid samples, and the analytical value for $\langle E_L \rangle$ is plotted as dotted blue lines.

low number of 1000 MC cycles per iteration.

The gradient descent algorithm was parallelized and run by 4 threads who each found their own optimal value for alpha. A variance of $\sigma^2_{\text{std}} < 10^{-9}$ was used as stopping criteria for the gradient descent. The mean $\bar{\alpha}$ was then used in doing a larger MC run of in total $2^{20} \approx 10^6$ samples in order to find $\langle E_L \rangle$.

We calculated the mean $\langle E_L \rangle$ with a standard error found using both the blocking algorithm (σ_B) using the standard derivative (σ from Eq. (41)). The results can be found in Table IV.

A plot illustrating the gradient descent evolution of a single thread for $N = 2, 16, 32, 64, 128$ non-interacting bosons is shown in Figure 3.

TABLE IV. Estimates of local energy $\langle E_L \rangle$ for various number of particles N for the non-interacting boson gas in a spherical HO-potential. The variational parameter $\bar{\alpha}$ is found from the gradient descent process. Two standard error estimates σ and σ_B for $\langle E_L \rangle$ are included.

$\bar{\alpha}$	N	$\langle E_L \rangle$	σ	σ_B
0.500007	2	3.0	0	$4 \cdot 10^{-7}$
0.5	16	24.0	0	$2 \cdot 10^{-7}$
0.500001	32	48.0	0	$1 \cdot 10^{-6}$
0.499997	64	96.0	$1 \cdot 10^{-7}$	$5 \cdot 10^{-6}$
0.500003	128	192.0	$2 \cdot 10^{-7}$	$2 \cdot 10^{-5}$

D. Interacting bosons

We repeated the gradient descent procedure using the full trial wave function and analytical local energy expression for interacting bosons. We used a less strict tolerance for breaking tailored to each N, as it became increasingly difficult to obtain a small variance as N got larger. With 4 threads performing 2^{18} samples each, we produced a data set of 2^{20}-$\langle E_L \rangle$ samples. As done above, we calculated $\langle E_L \rangle$ and the two standard errors σ_B and σ. The results are presented in Table V.

TABLE V. Estimates of local energy $\langle E_L \rangle$ for various number of particles N for the interacting boson gas in an elliptical HO-potential. The variational parameter $\bar{\alpha}$ is found from the gradient descent process. Two standard error estimates σ and σ_B for $\langle E_L \rangle$ and the energy per particle $\langle E_L \rangle / N$ is included.

$\bar{\alpha}$	N	$\langle E_L \rangle$	σ	σ_B	$\langle E_L \rangle / N$
0.4998	2	4.832421	$3 \cdot 10^{-6}$	$3 \cdot 10^{-5}$	2.42
0.4936	16	39.3117	$9 \cdot 10^{-4}$	$7 \cdot 10^{-2}$	2.46
0.4948	32	79.2114	$5 \cdot 10^{-4}$	$3 \cdot 10^{-2}$	2.48
0.5218	64	162.880	$3 \cdot 10^{-3}$	$1.5 \cdot 10^{-1}$	2.55

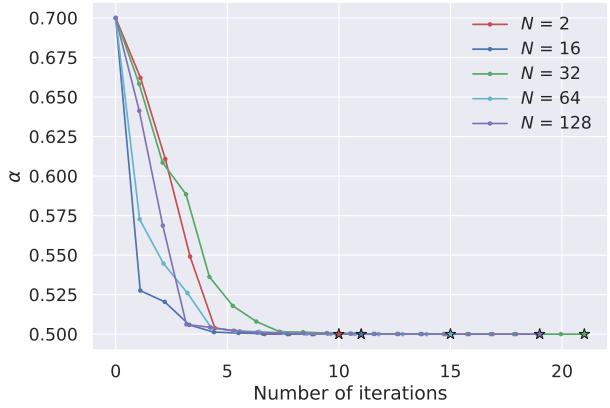


FIG. 3. The figure shows how the gradient descent method updates the value of α each iteration until reaching an optimal α value. We demanded a variance $\sigma^2_{\text{std}} < 10^{-9}$ for the gradient descent search to stop. The stars marks the iteration where the stopping criteria is met.

E. One body density

During the runs which produced the results in Tables IV and V, we also calculated the radial one-body density histogram of the system using Algorithm 1. Due to the narrow bin-width, we used the histogram to approximate a continuous function. The radial one-body density $\rho_r(r)$ is found for both the non-interacting gas in a spherical HO-potential and the interacting gas in an elliptical HO-potential.

In addition, the analytical solution for the non-interacting gas given in Eq. (26) is shown. The resulting plots are presented in Figure 4.

In order to illustrate proper normalization of the graphs in Figure 4, we have plotted $r^2 \cdot \rho_r(r)$ for $N = 16$ in Figure 5. Numerical integration of the three graphs using the trapezoidal rule yields 16 (up to numerical precision).

V. DISCUSSION

A. Non-interacting bosons

From Figure 1 it is evident that both the analytical and the numerical derivative yields satisfactory results for $\alpha = 0.5$ when compared to the exact solution for $\langle E_L \rangle$. The standard error in $\langle E_L \rangle$ is low throughout and we achieve a minimum of $\langle E_L \rangle$ for $\alpha = 0.5$ for all N except $N = 500$. For $N = 500$ we get a minimum for $\alpha = 0.45$ with $\langle E_L \rangle < \langle H \rangle_{\text{min}}$, which diverges from theory. We expect our calculation of the standard error to be overly optimistic, as we in this task did not take the covariance into account using only the "naive" way of calculating the standard deviation. This indicates that the error bars are too short, and could in fact be large enough to cover a value of $\langle E_L \rangle > \langle H \rangle_{\text{min}}$. In other words, the "true" value of $\langle E_L \rangle$ could still be greater than $\langle H \rangle_{\text{min}}$.

The differences between the calculated value of $\langle E_L \rangle$ far away from $\alpha = 0.50$ may be attributed to errors stemming from numerical differentiation. However, the big difference lies in CPU time when N grows large, as seen in Table I. The time spend by the analytical solution is more or less independent of N while the time spent by the numerical solution differs largely from $N = 10$ to $N = 500$.

B. Importance sampling

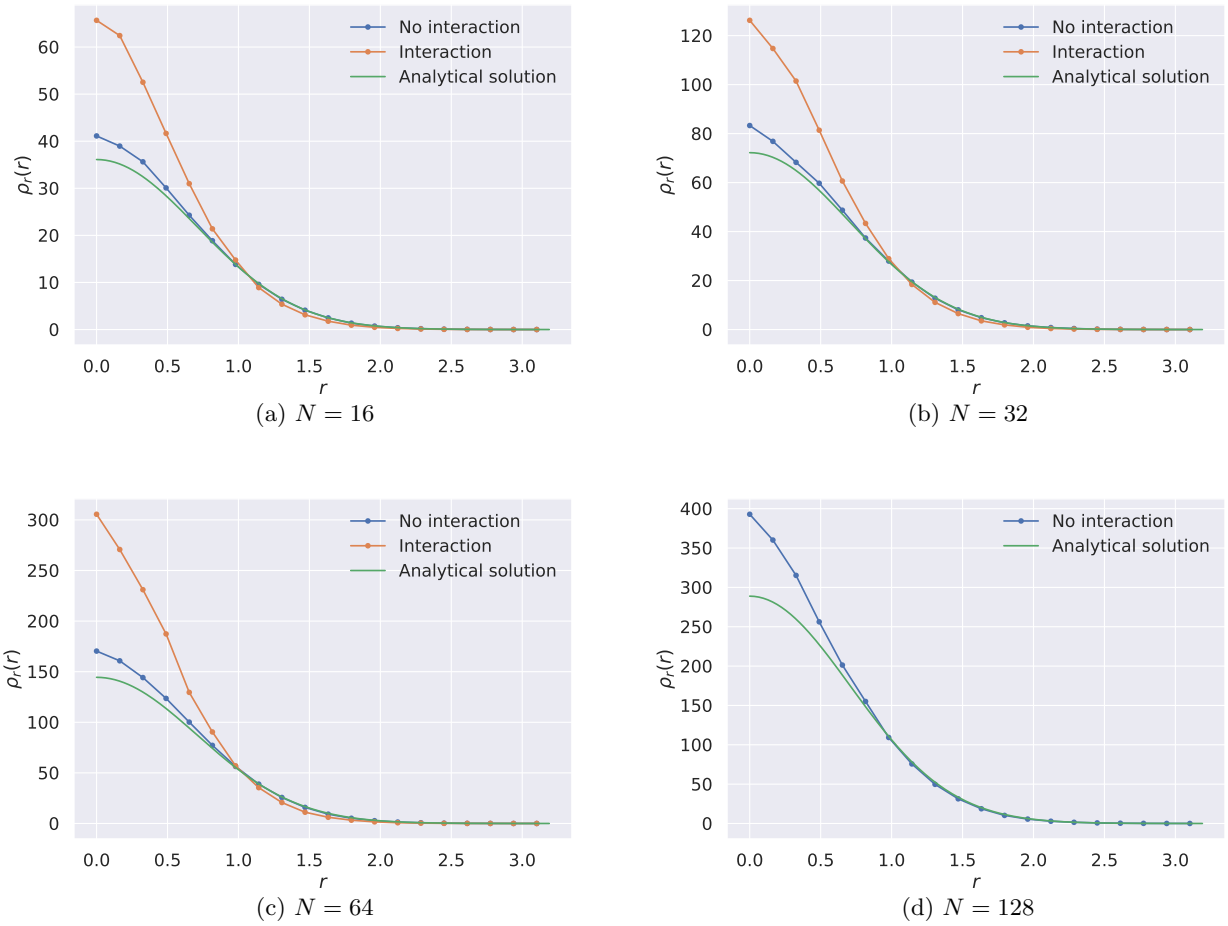
1. *Validity of results*

Studying Figure 2 one recognizes the pattern seen in the case without importance sampling in Figure 1. The standard error is large far away from the analytically correct value of $\alpha = 0.5$ and decreases towards it. For $\alpha = 0.5$ both the analytical and numerical expression for $\langle E_L \rangle$ is as good as spot on when comparing with the exact solution.Again, the largest difference between the two is CPU time used when N grows large, which is illustrated in Table III.

The zero variance for all entries with $\alpha = 0.50$ in Table II shows that this trial wave function indeed is the system's exact state. This conclusion is justified by Eq. (2).

Comparing Figures 1 and 2, it is not that clear what improvements introducing important sampling provided. This is possibly due to the minuscule size of the error bars making differentiating hard, but it is still possible to spot measurements with larger standard errors, and these do seem to appear randomly. We still think that that the improved sampling rules leads to better stability when studying the interacting particle gas, this is however not something that we have studied systematically.

If one takes a closer look at the subplot for $N = 500$, $\langle E_L \rangle$ for $\alpha = 0.45$ lies below the analytical solution, just as it did when we did not use importance sampling. As earlier, we believe that our method for calculating the



(a) $N = 16$ (b) $N = 32$

(c) $N = 64$ (d) $N = 128$

FIG. 4. Radial one-body density distributions for $N = 16, 32, 64, 128$ interacting and non-interacting particles. The analytical expression given in Eq. (26) is shown for comparison.

standard error at this point was to optimistic and that the true value of $\langle E_L \rangle$ could very well be above $\langle H \rangle_{\min}$.

2. *Choice of* Δt

From Eq. (35) we see that decreasing the value of the time step variable Δt translates to making smaller steps in configuration space. We expect this to lead to a higher probability of accepting new moves for the random walkers, however, we are ultimately only interested in choosing a Δt that minimizes the total variance.

This is the reasoning behind using Table II as guideline for choosing Δt; and from the Table it is evident that choosing Δt in the interval $[0.005, 0.1]$ gives a lower variance overall.

For the non-interacting case, using the analytical solution $\alpha = 0.50$ always leads to zero variance. This is due to to $\langle E_L \rangle$ being independent of the position of the random walkers (Eq. (14)). It is thus of more interest how the variance changes for values of α in the close vicinity of $\alpha = 0.50$.

Another way of picking a suitable Δt would be to study the variance for the interacting particles gas; however by exploring different values of Δt by trial and error, our choice of $\Delta t = 0.005$ was shown to give sufficiently stable results for the interacting case as well.

C. Gradient descent

As seen from Figure 3 the gradient descent algorithm converges nicely towards the optimal alpha for all choices of N. Even though we chose a rather small tolerance for "success", namely a variance of magnitude less than 10^{-9}, it took at most $20 - 22$ iterations for convergence. Choosing a less strict condition for breaking the gradient descent resulted in fewer iterations needed before success, but as we wanted a precise estimation of the optimal α value, we opted for a middle ground with some more iterations to ensure precision.

From Figure IV we see that the mean values for $\langle E_L \rangle$

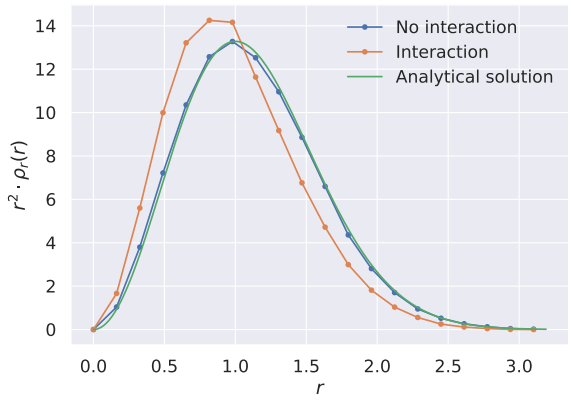


FIG. 5. Plot of radial one-body density $\rho_r(r)$ times r^2 for $N = 16$. Integrating below each individual curve return 16 up to numerical precision.

matches perfectly with the exact solution for all N. The standard error is overall very low, but we do see a small difference between the naive calculation where covariance does not factor in, and the blocking version where it does. The standard error from blocking is larger for all N, which fits well with our expectation of the naive way being overly optimistic.

D. Interacting bosons

In the case of interacting particles we ran into difficulties. As N grew larger, more and more unfavorable particle moves were accepted and the system went into turmoil. The local energy grew way too large in magnitude and the system could take long to recover, and even when it recovered the resulting set of samples was ruined. We therefore implemented a test to counteract this. If the sample of local energy turned out to be very far from the expected value, we rejected the move, re-initialized the random walkers and performed new equilibrium cycles. We see this as a less ideal solution to our problem, but after many many hour of debugging without identifying the source of the problem we deemed it as a lesser evil. We leave it as an exercise to the reader to download our repository and attempt to locate the reason for the faulty accepting of unfavorable moves.

The problem with an increasing instability seems connected to the fact that the particles on average have higher energy when N grows. This can be pictured as the hard-core potential pushing the particles up along the sides of the harmonic oscillator potential when the particles can't overlap. The increasing energy per particle for the interacting boson gas is shown in the last column of Table V. Other articles that have done similar experiments on boson gases also show the energy per particle increasing with N, an example by DuBois and Glyde can be seen in Figure 4 in Ref. [1].

Nonetheless, after implementing our test the algorithm worked just fine, yielding the results presented in Figure V. As in the case of the non-interacting bosons the blocking algorithm returns a less optimistic and likely more realistic standard error (in the order of $10 - 100$ times larger), just as we expected.

E. One body density

Even though the analytical solution for $\rho_r(r)$ seems to fit perfectly with the graph corresponding to the non-interacting case in Figure 5, this does not seem to be the case in Figure 4. The deviation is especially large around $r = 0$. This apparent discrepancy comes from the fact that when normalizing ρ_r, the contribution to the integral for low r is negligible. In order to arrive at better results for low r, it might be necessary to increase the number of samples and decrease the bin-width of the histogram.

Regarding the interacting particles, the results for the radial one-body densities (Figure 4) seems to suggest that the particles are packed more closely together when introducing the inter-particle repulsion. This is even more evident when studying Figure 5. The value of r where the area under the curves are equal on both sides (the radius of the sphere which contains half of the total particles) is clearly shifted towards lower rs when the Jastrow factor (correlation terms) is introduced.

We found the decreased spread of the particles to be curious, as it contradicts the reasoning behind the increased energy per particle when increasing N. We are unsure if this is due to a misinterpretation of the one-body density or an erroneous implementation of it. The general shape of $\rho_r(r)$ does however seem to resemble one-body density graphs found in the literature, such as Figure 1 in Ref. [2].

VI. CONCLUSION

In this report we looked at how the variational Monte Carlo method can be used to simulate both non-interacting and interacting bosons in a harmonic oscillator trap. We managed to produce a gradient descent method which for the non-interacting case managed to find an optimal value for the variational parameter α which diverged from the analytical value by only a factor of 10^{-5} for all N. This resulted in estimations of $\langle E_L \rangle$ fitting very well with the analytical expressions, with a relative error on the order of 10^{-7}.

For the case of interacting bosons we managed to find values of α which ensured a standard error below $10^{-3} \times \langle E_L \rangle$ for all N. The implementation of the interacting particles did not turn out perfectly and is thus a topic of interest for further work and development.

In addition, a thorough analysis was done studying the effects of introducing more efficient sampling rules and

analytical expressions for derivatives. Numerical differentiation is sufficient for low-particle systems, but the computational time blow up as N increases. Although importance sampling moves the random walkers in a more efficient manner in theory, we did not see this manifest into better results.

Finally, we calculated radial one-body densities for the system. Our results fit the analytical solution well for non-interacting gases, but adding repulsion between particles did seem to make the particles less spread out. Finding a good explanation for this is also a well-suited task for further research.

[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. Hjorth-Jensen, and A. Polls. Vortices in atomic Bose-Einstein condensates in the large-gas-parameter region. *Phys. Rev. A*, 71:053610, May 2005.

[3] D. J. Griffiths. *Introduction to Quantum Mechanics*, chapter 8. Cambridge University Press, 3 edition, 2018.

[4] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, 3 edition, 2007.

[5] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer, Berlin, 3 edition, 2004.

[6] Marius Jonsson. Standard error estimation by an automated blocking method. *Phys. Rev. E*, 98:043304, Oct 2018.

[7] M. N. Nordhagen. Studies of quantum dots using machine learning. http://urn.nb.no/URN:NBN:no-76863, 2019. Unpublished master thesis, University of Oslo.

Appendix A: Deriving analytical expressions for non-interacting bosons

1. Local energy

In this section we will derive the the analytical expressions for the local energy $E_L(\mathbf{r})$ and the minimum value of $\langle H \rangle$ as presented in Eqs. (14) and (15). We assume natural units, a spherical trap ($\beta = 1$), and no inter-particle repulsion ($a = 0$). We consider a case of N identical particles in D spacial dimensions. Then, the Hamiltonian (Eq. (3)) and the trial wave function (Eq. (6)) reduces to

$$H = \sum_{n=1}^{N} \frac{1}{2}\left(-\nabla_n^2 + r_n^2\right) \tag{A1}$$

and

$$\Psi_T(\mathbf{r}) = \prod_{i=1}^{N} e^{-\alpha r_i^2}. \tag{A2}$$

Both of the above expressions include the quantity $r_i^2 = \sum_{d=1}^{D} \left[x_i^d\right]^2$, where the subscript i denotes the index of the particle, and the superscript d the index of the spacial dimension ($d = 1$ corresponds to the x-axis, etc.).

We begin by finding out how the Laplacian for a specific particle k ($i \in \{1, 2, \ldots, k, \ldots, N\}$) acts on the trial wave function. In three dimensions we have

$$\nabla_k^2 \Psi_T(\mathbf{r}) = \left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2}\right) \Psi_T(\mathbf{r}),$$

which in D dimensions can be generalized to

$$\nabla_k^2 \Psi_T(\mathbf{r}) = \sum_{d=1}^{D} \frac{\partial^2}{\partial \left(x_k^d\right)^2} \Psi_T(\mathbf{r}).$$

We can begin by calculating the contribution from the term with $d = 1$:

$$\begin{aligned}
\frac{\partial^2}{\partial \left(x_k^1\right)^2} \Psi_T(\mathbf{r}) &= \frac{\partial}{\partial x_k^1}\left(\frac{\partial}{\partial x_k^1} \prod_{i=1}^{N} e^{-\alpha\left[(x_i^1)^2 + \cdots + (x_i^D)^2\right]}\right) = -2\alpha \frac{\partial}{\partial x_k^1}\left(x_k^1 \prod_{i=1}^{N} e^{-\alpha r_i^2}\right) \\
&= -2\alpha\left(\prod_{i=1}^{N} e^{-\alpha r_i^2} - 2\alpha\left(x_k^1\right)^2 \prod_{i=1}^{N} e^{-\alpha r_i^2}\right) = \left(4\alpha^2\left(x_k^1\right)^2 - 2\alpha\right) \Psi_T(\mathbf{r}).
\end{aligned}$$

Similar expressions are found for all other indices of d, so the Laplacian on $\Psi_T(\mathbf{r})$ is thus

$$\nabla_k^2 \Psi_T(\mathbf{r}) = \sum_{d=1}^{D} \left[\left(4\alpha^2\left(x_k^d\right)^2 - 2\alpha\right) \Psi_T(\mathbf{r})\right] = \left(4\alpha^2 r_k^2 - D \cdot 2\alpha\right) \Psi_T(\mathbf{r}). \tag{A3}$$

Using Eq. (A3), the local energy is found to be

$$\begin{aligned}
E_L(\mathbf{r}) &= \frac{1}{\Psi_T(\mathbf{r})} \sum_{n=1}^{N} \left[-\frac{1}{2}\nabla_n^2 + \frac{1}{2} r_n^2\right] \Psi_T(\mathbf{r}) \\
&= \frac{1}{\Psi_T(\mathbf{r})} \sum_{n=1}^{N} \left[-\frac{1}{2}\left(4\alpha^2 r_n^2 - D \cdot 2\alpha\right) + \frac{1}{2} r_n^2\right] \Psi_T(\mathbf{r}) \\
&= \frac{1}{2} \sum_{n=1}^{N} \left[2D\alpha + r_n^2(1 - 4\alpha^2)\right] = DN\alpha + \frac{(1 - 4\alpha^2)}{2} \sum_{n=1}^{N} r_n^2
\end{aligned}$$

The final expression is the local energy presented in Eq. (14).

2. Optimal choice of variational parameter and minimal state energy

We will now find the value of the variational parameter α that minimizes the expectation value of energy $\langle H \rangle$. The assumptions from the previous section still apply, so the Hamiltonian H and trial wave function $\Psi_T(\mathbf{r})$ will also be the ones given in Eqs. (A1) and (A2). Recalling Eq. (1), $\langle H \rangle$ is given by

$$\langle H \rangle \equiv \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}. \tag{A4}$$

We begin by finding the normalization factor $\langle \Psi_T | \Psi_T \rangle$, and we will rename this constant to C for brevity.

$$C \equiv \langle \Psi_T | \Psi_T \rangle = \int \mathrm{d}\mathbf{r}\, |\Psi(\mathbf{r})|^2 = \prod_{i=1}^{N} \prod_{d=1}^{D} \int \mathrm{d}x_i^d \, e^{-2\alpha \left(x_i^d\right)^2} = \prod_{i=1}^{N} \prod_{d=1}^{D} \sqrt{\frac{\pi}{2\alpha}} = \left(\sqrt{\frac{\pi}{2\alpha}}\right)^{ND}$$

Inserting H from Eq. (A1) into Eq. (A4), we find

$$\langle H \rangle = \frac{1}{C} \sum_{n=1}^{N} \langle \Psi_T | \left(-\frac{1}{2} \nabla_n^2 + \frac{1}{2} r_n^2 \right) | \Psi_T \rangle = \frac{1}{C} \sum_{n=1}^{N} \left[-\frac{1}{2} \langle \Psi_T | \nabla_n^2 | \Psi_T \rangle + \frac{1}{2} \langle \Psi_T | r_n^2 | \Psi_T \rangle \right].$$

Eq. (A3) tells how the Laplacian works on the function Ψ_T. We thus have

$$\begin{aligned}
\langle H \rangle &= \frac{1}{C} \sum_{n=1}^{N} \left[-\frac{1}{2} \langle \Psi_T | \left(4\alpha^2 r_n^2 - 2D\alpha \right) | \Psi_T \rangle + \frac{1}{2} \langle \Psi_T | r_n^2 | \Psi_T \rangle \right] \\
&= \frac{1}{C} \sum_{n=1}^{N} \left[\left(\frac{1}{2} - 2\alpha^2 \right) \langle \Psi_T | r_n^2 | \Psi_T \rangle + D\alpha \underbrace{\langle \Psi_T | \Psi_T \rangle}_{C} \right] \\
&= NDa + \frac{\left(1 - 4\alpha^2\right)}{2C} \sum_{n=1}^{N} \langle \Psi_T | r_n^2 | \Psi_T \rangle .
\end{aligned} \tag{A5}$$

The next step is to find the quantity $\langle r_n^2 \rangle \equiv \langle \Psi_T | r_n^2 | \Psi_T \rangle$ for a given index k. Solving this $N \cdot D$ dimensional integral, we use the notation

$$\langle r_k^2 \rangle = \int \mathrm{d}x_1^1 \cdots \mathrm{d}x_N^D \, e^{-2\alpha \left(x_1^1\right)^2} \cdots e^{-2\alpha \left(x_N^D\right)^2} r_k^2 = \prod_{d=1}^{D} \prod_{n=1}^{N} \int \mathrm{d}x_i^d \, e^{-2\alpha \left(x_i^d\right)^2} r_k^2.$$

We find

$$\begin{aligned}
\langle r_k^2 \rangle &= \left[\prod_{d=1}^{D} \int \mathrm{d}x_k^d \, e^{-2\alpha \left(x_k^d\right)^2} r_k^2 \right] \prod_{d=1}^{D} \prod_{i \neq k}^{N} \int \mathrm{d}x_i^d \, e^{-2\alpha \left(x_i^d\right)^2} \\
&= \left[\prod_{d=1}^{D} \int \mathrm{d}x_k^d \, e^{-2\alpha \left(x_k^d\right)^2} r_k^2 \right] \sqrt{\frac{\pi}{2\alpha}}^{D(N-1)} \\
&= \sqrt{\frac{\pi}{2\alpha}}^{D(N-1)} \prod_{d=1}^{D} \left[\int \mathrm{d}x_k^d \, e^{-2\alpha \left(x_k^d\right)^2} \sum_{d'=1}^{D} \left(r_k^{d'} \right)^2 \right] \\
&= \sqrt{\frac{\pi}{2\alpha}}^{D(N-1)} \sum_{d'=1}^{D} \left[\int \mathrm{d}x_k^{d'} \, e^{-2\alpha \left(x_k^{d'}\right)^2} \left(x_k^{d'} \right)^2 \prod_{d \neq d'}^{D} \int \mathrm{d}x_k^d \, e^{-2\alpha \left(x_k^d\right)^2} \right] \\
&= \sqrt{\frac{\pi}{2\alpha}}^{D(N-1)} \sum_{d'=1}^{D} \left[\int \mathrm{d}x_k^{d'} \, e^{-2\alpha \left(x_k^{d'}\right)^2} \left(x_k^{d'} \right)^2 \sqrt{\frac{\pi}{2\alpha}}^{D-1} \right] \\
&= \sqrt{\frac{\pi}{2\alpha}}^{D(N-1)} \sqrt{\frac{\pi}{2\alpha}}^{D-1} \sum_{d'=1}^{D} \left[\int \mathrm{d}x_k^{d'} \, e^{-2\alpha \left(x_k^{d'}\right)^2} \left(x_k^{d'} \right)^2 \right] \\
&= \sqrt{\frac{\pi}{2\alpha}}^{DN-1} \sum_{d'=1}^{D} \left[\sqrt{\frac{\pi}{2\alpha}} \frac{1}{4\alpha} \right] = \sqrt{\frac{\pi}{2\alpha}}^{DN} \frac{D}{4\alpha} = C \frac{D}{4\alpha}.
\end{aligned}$$

The last expression in the above line can be inserted into Eq. (A5), which yield

$$\langle H\rangle = ND\alpha + \frac{\left(1-4\alpha^2\right)}{2C}\sum_{n=1}^{N} C\frac{D}{4\alpha} = ND\alpha + \frac{\left(1-4\alpha^2\right)}{2}\frac{ND}{4\alpha} = \frac{ND\alpha}{2} + \frac{ND}{8\alpha}. \tag{A6}$$

This expectation energy is supposed to minimized with respect to α. We find that

$$\frac{\partial\langle H\rangle}{\partial\alpha} = \frac{ND}{2} - \frac{ND}{8\alpha^2} = 0 \implies \alpha = \frac{1}{2}$$

for $\alpha > 0$. It can readily be checked that $\frac{\partial^2\langle H\rangle}{\partial\alpha^2}\Big|_{\alpha=1/2} > 0$, so this does indeed give an energy minimum. Plugging $\alpha = \frac{1}{2}$ into Eq. (A6) results in $\langle H\rangle_{\min} = \frac{D}{2}N$, which is the result given in Eq. (15).

Appendix B: Deriving local energy for interacting bosons

In this section we will derive Eqs. (17 – 21), which are used to find the local energy for the interacting particle Bose-gas in three dimensions. Given the trial function in Eq. (16), we will begin by looking at the gradient for an arbitrary particle k. Note that the indices i, j, l, m and p all run from 1 to N.

$$\begin{aligned}
\nabla_k \Psi_T(\mathbf{r}) &= \nabla_k \left[\prod_i \phi(\mathbf{r}_i)\right] \exp\left(\sum_{j<m} u(r_{jm})\right) \\
&= \left(\nabla_k \left[\prod_i \phi(\mathbf{r}_i)\right]\right) \exp\left(\sum_{j<m} u(r_{jm})\right) + \left[\prod_i \phi(\mathbf{r}_i)\right] \nabla_k \exp\left(\sum_{j<m} u(r_{jm})\right) \\
&= \nabla_k \phi(\mathbf{r}_k) \left[\prod_{i\neq k} \phi(\mathbf{r}_i)\right] \exp\left(\sum_{j<m} u(r_{jm})\right) + \left[\prod_i \phi(\mathbf{r}_i)\right] \left[\sum_{l\neq k} \nabla_k u(r_{kl})\right] \exp\left(\sum_{j<m} u(r_{jm})\right). \qquad \text{(B1)}
\end{aligned}$$

The Laplacian can then be calculated as

$$\begin{aligned}
\nabla_k^2 \Psi_T(\mathbf{r}) &= \nabla_k \left[\nabla_k \Psi_T(\mathbf{r})\right] \\
&= \underbrace{\nabla_k \left[\nabla_k \phi(\mathbf{r}_k) \left[\prod_{i\neq k} \phi(\mathbf{r}_i)\right] \exp\left(\sum_{j<m} u(r_{jm})\right)\right]}_{\text{I}} + \underbrace{\nabla_k \left[\left[\prod_i \phi(\mathbf{r}_i)\right] \left[\sum_{l\neq k} \nabla_k u(r_{kl})\right] \exp\left(\sum_{j<m} u(r_{jm})\right)\right]}_{\text{II}}.
\end{aligned} \qquad \text{(B2)}$$

For simplicity we will begin by looking at the first term (I) in the above sum

$$\begin{aligned}
\text{I} &= \prod_{i\neq k} \phi(\mathbf{r}_i) \nabla_k \left[(\nabla_k \phi(\mathbf{r}_k)) \exp\left(\sum_{j<m} u(r_{jm})\right)\right] \\
&= \prod_{i\neq k} \phi(\mathbf{r}_i) \left[(\nabla_k^2 \phi(\mathbf{r}_k)) \exp\left(\sum_{j<m} u(r_{jm})\right) + (\nabla_k \phi(\mathbf{r}_k)) \nabla_k \exp\left(\sum_{j<m} u(r_{jm})\right)\right] \\
&= \prod_{i\neq k} \phi(\mathbf{r}_i) \exp\left(\sum_{j<m} u(r_{jm})\right) \left[\nabla_k^2 \phi(\mathbf{r}_k) + \nabla_k \phi(\mathbf{r}_k) \left(\sum_{l\neq k} \nabla_k u(r_{kl})\right)\right] \\
&= \frac{\Psi_T(\mathbf{r})}{\phi(\mathbf{r}_k)} \left[\nabla_k^2 \phi(\mathbf{r}_k) + \nabla_k \phi(\mathbf{r}_k) \left(\sum_{l\neq k} \nabla_k u(r_{kl})\right)\right].
\end{aligned}$$

Now, the second term (II) gives

$$\begin{aligned}
\text{II} &= \nabla_k \phi(\mathbf{r}_k) \left[\prod_{i\neq k} \phi(\mathbf{r}_i)\right] \left[\sum_{l\neq k} \nabla_k u(r_{kl})\right] \exp\left(\sum_{j<m} u(r_{jm})\right) \\
&\quad + \left[\prod_i \phi(\mathbf{r}_i)\right] \exp\left(\sum_{j<m} u(r_{jm})\right) \sum_{p\neq k} \nabla_k u(r_{kp}) \sum_{l\neq k} \nabla_k u(r_{kl}) \\
&\quad + \left[\prod_i \phi(\mathbf{r}_i)\right] \exp\left(\sum_{j<m} u(r_{jm})\right) \sum_{l\neq k} \nabla_k^2 u(r_{kl}) \\
&= \prod_i \phi(\mathbf{r}_i) \exp\left(\sum_{j<m} u(r_{jm})\right) \left[\frac{\nabla_k \phi(\mathbf{r}_k) \left[\sum_{l\neq k} \nabla_k u(r_{kl})\right]}{\phi(\mathbf{r}_k)} + \sum_{p\neq k} \nabla_k u(r_{kp}) \sum_{l\neq k} \nabla_k u(r_{kl}) + \sum_{l\neq k} \nabla_k^2 u(r_{kl})\right]
\end{aligned}$$

$$= \Psi_T(\mathbf{r}) \left[\frac{\nabla_k \phi(\mathbf{r}_k) \left[\sum_{l \neq k} \nabla_k u(r_{kl}) \right]}{\phi(\mathbf{r}_k)} + \sum_{p \neq k} \nabla_k u(r_{kp}) \sum_{l \neq k} \nabla_k u(r_{kl}) + \sum_{l \neq k} \nabla_k^2 u(r_{kl}) \right].$$

Inserting the expressions for the two terms into Eq. (B2) yields

$$\begin{aligned} \nabla_k^2 \Psi_T(\mathbf{r}) = \Psi_T(\mathbf{r}) \Bigg[& \frac{\nabla_k^2 \phi(\mathbf{r}_k) + \nabla_k \phi(\mathbf{r}_k) \left(\sum_{l \neq k} \nabla_k u(r_{kl}) \right)}{\phi(\mathbf{r}_k)} + \frac{\nabla_k \phi(\mathbf{r}_k) \left(\sum_{l \neq k} \nabla_k u(r_{kl}) \right)}{\phi(\mathbf{r}_k)} \\ & + \sum_{p \neq k} \nabla_k u(r_{kp}) \sum_{l \neq k} \nabla_k u(r_{kl}) + \sum_{l \neq k} \nabla_k^2 u(r_{kl}) \Bigg]. \end{aligned} \tag{B3}$$

We now go on to finding the gradient and Laplacian of $u(r_{kl})$. Performing the differentiation of u gives us

$$\nabla_k u_{kl} = \sum_{m=1}^{3} \mathbf{e}_m \frac{\partial}{\partial x_k^m} u(r_{kl}) = \sum_{m=1}^{3} \mathbf{e}_m \frac{\partial u(r_{kl})}{\partial r_{kl}} \frac{\partial r_{kl}}{\partial x_k^m} = u'(r_{kl}) \sum_{m=1}^{3} \mathbf{e}_m \frac{\partial r_{kl}}{\partial x_k^m}.$$

Since we define $r_{kl} \equiv |\mathbf{r}_{kl}|$, it can be written as

$$r_{kl} = |\mathbf{r}_k - \mathbf{r}_l| = \left| (x_k^1 - x_l^1)\mathbf{e}_1 + (x_k^2 - x_l^2)\mathbf{e}_2 + (x_k^3 - x_l^3)\mathbf{e}_3 \right|,$$

which can be used to calculate

$$\begin{aligned} \frac{\partial r_{kl}}{\partial x_k^m} &= \frac{\partial}{\partial x_k^m} |\mathbf{r}_{kl}| = \frac{\mathbf{r}_{kl}}{|\mathbf{r}_{kl}|} \frac{\partial}{\partial x_k^m} \mathbf{r}_{kl} = \frac{\mathbf{r}_{kl}}{|\mathbf{r}_{kl}|} \frac{\partial}{\partial x_k^m} [(x_k^1 - x_l^1)\mathbf{e}_1 + (x_k^2 - x_l^2)\mathbf{e}_2 + (x_k^3 - x_l^3)\mathbf{e}_3] = \frac{\mathbf{r}_{kl}}{r_{kl}} \cdot \mathbf{e}_m \\ &= \frac{(x_k^m - x_l^m)}{r_{kl}}. \end{aligned}$$

Inserting this into the above expression for $\nabla_k u_{kl}$, we arrive at

$$\nabla_k u_{kl} = u'(r_{kl}) \sum_{m=1}^{3} \mathbf{e}_m \frac{\partial r_{kl}}{\partial x_k^m} = \frac{u'(r_{kl})}{r_{kl}} \sum_{m=1}^{3} \mathbf{e}_m (x_k^m - x_l^m) = \frac{(\mathbf{r}_k - \mathbf{r}_l)}{r_{kl}} u'(r_{kl}). \tag{B4}$$

Similarly, we can calculate

$$\begin{aligned} \nabla_k^2 u(r_{kl}) &= \nabla_k \left(\frac{(\mathbf{r}_k - \mathbf{r}_l)}{r_{kl}} u'(r_{kl}) \right) = \sum_{m=1}^{3} \mathbf{e}_m \frac{\partial}{\partial x_k^m} \left[\frac{(\mathbf{r}_k - \mathbf{r}_l)}{r_{kl}} u'(r_{kl}) \right] \\ &= \sum_{m=1}^{3} \left[\frac{\partial}{\partial x_k^m} \mathbf{e}_m (\mathbf{r}_k - \mathbf{r}_l) \right] \frac{u'(r_{kl})}{r_{kl}} + \sum_{m=1}^{3} \mathbf{e}_m (\mathbf{r}_k - \mathbf{r}_l) \frac{\partial}{\partial x_k^m} \frac{u'(r_{kl})}{r_{kl}} \\ &= \sum_{m=1}^{3} \left[\frac{\partial}{\partial x_k^m} (x_k^m - x_l^m) \right] \frac{u'(r_{kl})}{r_{kl}} + \sum_{m=1}^{3} (x_k^m - x_l^m) \frac{\partial r_{kl}}{\partial x_k^m} \frac{\partial}{\partial r_{kl}} \frac{u'(r_{kl})}{r_{kl}} \\ &= \frac{3u'(r_{kl})}{r_{kl}} + \sum_{m=1}^{3} (x_k^m - x_l^m) \frac{(x_k^m - x_l^m)}{r_{kl}} \left[-\frac{u'(r_{kl})}{r_{kl}^2} + \frac{u''(r_{kl})}{r_{kl}} \right] \\ &= \frac{3u'(r_{kl})}{r_{kl}} + \sum_{m=1}^{3} \frac{(x_k^m - x_l^m)^2}{r_{kl}^2} \left[-\frac{u'(r_{kl})}{r_{kl}} + u''(r_{kl}) \right] \\ &= \frac{3u'(r_{kl})}{r_{kl}} + \left[-\frac{u'(r_{kl})}{r_{kl}} + u''(r_{kl}) \right] \frac{(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2}{r_{kl}^2} \\ &= \frac{3u'(r_{kl})}{r_{kl}} + u''(r_{kl}) - \frac{u'(r_{kl})}{r_{kl}} \\ &= u''(r_{kl}) + \frac{2u'(r_{kl})}{r_{kl}}, \end{aligned} \tag{B5}$$

and inserting the expressions found in Eqs. (B4) and (B5) into Eq. (B3), we find that

$$\nabla_k^2\Psi_T(\mathbf{r}) = \Psi_T(\mathbf{r})\Bigg[\frac{\nabla_k^2\phi(\mathbf{r}_k) + \nabla_k\phi(\mathbf{r}_k)\left(\sum_{l\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_l)}{r_{kl}}u'(r_{kl})\right)}{\phi(\mathbf{r}_k)} + \frac{\nabla_k\phi(\mathbf{r}_k)\left(\sum_{l\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_l)}{r_{kl}}u'(r_{kl})\right)}{\phi(\mathbf{r}_k)} + \sum_{p\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_p)}{r_{kp}}u'(r_{kp})\sum_{l\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_l)}{r_{kl}}u'(r_{kl}) + \sum_{l\neq k}\left(u''(r_{kl}) + \frac{2u'(r_{kl})}{r_{kl}}\right)\Bigg]. \tag{B6}$$

Finally, we divide the above expression with the trial wave function

$$\frac{\nabla_k^2\Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{\nabla_k^2\phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2\frac{\nabla_k\phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)}\left(\sum_{l\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_l)}{r_{kl}}u'(r_{kl})\right) + \left(\sum_{l\neq k}\frac{(\mathbf{r}_k-\mathbf{r}_l)}{r_{kl}}u'(r_{kl})\right)^2 + \sum_{l\neq k}\left(u''(r_{kl}) + \frac{2u'(r_{kl})}{r_{kl}}\right) \tag{B7}$$

in order to arrive at Eq. (17).

We finish by closer at the derivatives of $\phi(\mathbf{r}_k)$,

$$\nabla_k\phi(\mathbf{r}_k) = \nabla_k e^{-\alpha r_k^2} = -2\alpha e^{-\alpha r_k^2} = -2\alpha(x_k\mathbf{e}_1 + y_k\mathbf{e}_2 + \beta z_k\mathbf{e}_3)\phi(\mathbf{r}_k) \tag{B8}$$

$$\begin{aligned}\nabla_k^2\phi(\mathbf{r}_k) &= \nabla_k(-2\alpha r_k e^{-\alpha r_k^2}) = -2\alpha(1+1+\beta)\phi(\mathbf{r}_k) + (-2\alpha(x_k\mathbf{e}_1 + y_k\mathbf{e}_2 + \beta z_k\mathbf{e}_3))^2\phi(\mathbf{r}_k)\\ &= \phi(\mathbf{r}_k)\left[-2\alpha(2+\beta) + 4\alpha^2(x_k^2 + y_k^2 + \beta^2 z_k^2)\right],\end{aligned} \tag{B9}$$

and the derivatives of $u(r_{kl})$. Since $u(r_{kl}) = \ln f(r_{kl}) = \ln\left(1 - \frac{a}{r_{kl}}\right)$, we calculate

$$u'(r_{kl}) = \frac{1}{1-\frac{a}{r_{kl}}}\cdot\frac{a}{r_{kl}^2} = \frac{a}{r_{kl}(r_{kl}-a)} \tag{B10}$$

and

$$u''(r_{kl}) = a\cdot\frac{\mathrm{d}}{\mathrm{d}u}(r_{kl}(r_{kl}-a))^{-1} = -a(r_{kl}(r_{kl}-a))^{-2}(2r_{kl}-a) = \frac{a^2-2ar_{kl}}{r_{kl}^2(r_{kl}-a)^2}. \tag{B11}$$

The derivations of Eqs. (B8 – B11) shows the origin of Eqs. (18 – 21) in the report.

Appendix C: Analytical expression for the drift force

1. Non-interacting case

In this section we will be deriving the analytical expression for the drift force defined in Eq. (22). We will look at the case with only the harmonic oscillator potential for a spherical trap ($\beta = 1$). In addition, the two-body potential is discarded by setting $a = 0$, which reduces the trial function to the one presented in Eq. (A2). We first find the drift force for the three dimensional case, as further generalization to an arbitrary number D of spacial dimensions is simple. The drift force exerted on a particle k is

$$\mathbf{F}_k = \frac{2\nabla_k \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{2}{\Psi_T(\mathbf{r})}\left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial y_k}, \frac{\partial}{\partial z_k}\right)\Psi_T(\mathbf{r}) = \frac{2}{\prod_{i=1}^{N} e^{-\alpha\left(x_i^2+y_i^2+z_i^2\right)}}\left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial y_k}, \frac{\partial}{\partial z_k}\right)\prod_{i=1}^{N} e^{-\alpha\left(x_i^2+y_i^2+z_i^2\right)}.$$

The differentiation is equal in all three spacial components, so we choose to look at what happens when differentiating with respect to x:

$$\frac{\partial \Psi_T(\mathbf{r})}{\partial x_k} = \frac{\partial}{\partial x_k}\prod_{i=1}^{N} e^{-\alpha\left(x_i^2+y_i^2+z_i^2\right)} = -2\alpha x_k \prod_{i=1}^{N} e^{-\alpha\left(x_i^2+y_i^2+z_i^2\right)} = -2\alpha x_k \Psi_T(\mathbf{r}).$$

The two remaining partial derivatives will return a similar result as above, but with factors y_k and z_k instead of x_k. Now that we know how to find the gradient $\nabla_k \Psi_T$, we find the drift force to be

$$\mathbf{F}_k = -\frac{2}{\Psi_T(\mathbf{r})} 2\alpha \Psi_T(\mathbf{r})(x_k \mathbf{e}_x + y_k \mathbf{e}_y + z_k \mathbf{e}_z) = -4\alpha \mathbf{r}_k. \tag{C1}$$

Note that the expression $\mathbf{F}_k = -4\alpha \mathbf{r}_k$ is valid also in the D-dimensional case.

2. Interacting case

Finding the drift force corresponding to a given particle for an interacting particle Bose gas is straightforward, as the gradient $\nabla_k \Psi_T(\mathbf{r})$ has been calculated in Eq. (B1). Using the shape of the trial function (Eq. (16)), the gradient becomes

$$\nabla_k \Psi_T(\mathbf{r}) = \frac{\Psi_T(\mathbf{r})}{\phi(\mathbf{r}_k)} \nabla_k \phi(\mathbf{r}_k) + \Psi_T(\mathbf{r}) \left[\sum_{l \neq k} \nabla_k u(r_{kl})\right]$$

The drift force is thus

$$\mathbf{F}_k = \frac{2\nabla_k}{\Psi_T(\mathbf{r})} = \frac{2\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2\left[\sum_{l \neq k} \nabla_k u(r_{kl})\right],$$

and inserting $\nabla_k u(r_{kl})$ from Eq. (B4) results in

$$\mathbf{F}_k = \frac{2\nabla_k \phi(\mathbf{r}_k)}{\phi(\mathbf{r}_k)} + 2\sum_{l \neq k} \frac{(\mathbf{r}_k - \mathbf{r}_l)}{r_{kl}} u'(r_{kl}). \tag{C2}$$

Appendix D: Analytical expression for gradient descent

In this section we will derive Eq. (40), which describes how the expectation value of energy varies with changes in a variational parameter α. The quantities $\bar{E}_\alpha$ and $\bar{\Psi}_\alpha$ are defined in Eq. (39). We will remove the subscript T from the trial wave function and rewrite $\frac{\mathrm{d}}{\mathrm{d}\alpha} \to \nabla_\alpha$ for brevity (even though we are only considering a single variational parameter for the moment).

In the following derivation, we have to recall that the expectation value of any function X is given by

$$\langle X \rangle = \frac{\int \mathrm{d}\mathbf{r}\, \Psi^* X \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} \overset{\Psi \to \mathbb{R}}{=} \frac{\int \mathrm{d}\mathbf{r}\, \Psi X \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2}. \tag{D1}$$

Now, starting of by using the quotient rule of derivatives, we have

$$\begin{aligned}
\bar{E}_\alpha &= \frac{\partial}{\partial \alpha} \langle H \rangle = \frac{\partial}{\partial \alpha} \frac{\int \mathrm{d}\mathbf{r}\, \Psi^* H \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} \\
&= \frac{\nabla_\alpha \int \mathrm{d}\mathbf{r}\, \Psi^* H \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} - \frac{\left(\int \mathrm{d}\mathbf{r}\, \Psi^* H \Psi\right) \nabla_\alpha \int \mathrm{d}\mathbf{r}\, |\Psi|^2}{\left(\int \mathrm{d}\mathbf{r}\, |\Psi|^2\right)^2} \\
&= \frac{\int \mathrm{d}\mathbf{r} \left(\bar{\Psi}_\alpha^* H \Psi + \Psi^* H \bar{\Psi}_\alpha\right)}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} - \underbrace{\frac{\int \mathrm{d}\mathbf{r}\, \Psi^* H \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2}}_{\langle H \rangle = \langle E_L \rangle} \frac{\int \mathrm{d}\mathbf{r}\, \nabla_\alpha (\Psi^* \Psi)}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2}.
\end{aligned}$$

Assuming a real trial wave function, we can write

$$\nabla_\alpha (\Psi^* \Psi) = \nabla_\alpha \Psi^2 = 2\Psi \nabla_\alpha \Psi = 2\Psi \bar{\Psi}_\alpha = 2\Psi \frac{\bar{\Psi}_\alpha}{\Psi} \Psi,$$

which gives

$$\begin{aligned}
\bar{E}_\alpha &= \frac{\int \mathrm{d}\mathbf{r} \left(\bar{\Psi}_\alpha H \Psi + \Psi H \bar{\Psi}_\alpha\right)}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} - 2 \langle E_L \rangle \frac{\int \mathrm{d}\mathbf{r}\, \Psi \left(\frac{\bar{\Psi}_\alpha}{\Psi}\right) \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} \\
&= \frac{\int \mathrm{d}\mathbf{r} \left(\bar{\Psi}_\alpha H \Psi + \Psi H \bar{\Psi}_\alpha\right)}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} - 2 \langle E_L \rangle \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} \right\rangle.
\end{aligned}$$

Next, the Hermitian properties of H means that

$$\int \mathrm{d}\mathbf{r}\, \Psi H \bar{\Psi}_\alpha = \int \mathrm{d}\mathbf{r}\, \bar{\Psi}_\alpha (H\Psi)^* = \int \mathrm{d}\mathbf{r}\, \bar{\Psi}_\alpha H \Psi,$$

so

$$\begin{aligned}
\bar{E}_\alpha &= 2 \frac{\int \mathrm{d}\mathbf{r}\, \bar{\Psi}_\alpha H \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} - 2 \langle E_L \rangle \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} \right\rangle \\
&= 2 \frac{\int \mathrm{d}\mathbf{r}\, \Psi \left[\bar{\Psi}_\alpha \left(\frac{1}{\Psi} H \Psi\right) \frac{1}{\Psi}\right] \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} - 2 \langle E_L \rangle \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} \right\rangle \\
&= 2 \frac{\int \mathrm{d}\mathbf{r}\, \Psi \left(\frac{\bar{\Psi}_\alpha}{\Psi} E_L\right) \Psi}{\int \mathrm{d}\mathbf{r}\, |\Psi|^2} - 2 \langle E_L \rangle \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} \right\rangle \\
&= 2 \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} E_L \right\rangle - 2 \langle E_L \rangle \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} \right\rangle \\
&= 2 \left(\left\langle \frac{\bar{\Psi}_\alpha}{\Psi} E_L \right\rangle - \left\langle \frac{\bar{\Psi}_\alpha}{\Psi} \right\rangle \langle E_L \rangle \right),
\end{aligned}$$

which is the expression given in Eq. (40).