Variational Monte Carlo on Bosonic Systems

Winther-Larsen, Sebastian Gregorius[1, *] and Schøyen, Øyvind Sigmundson[1, *]

[1] *University of Oslo*

(Dated: April 5, 2018)

In this project we use Quantum Variational Monte Carlo on a toy model of a bosonic gas in an elliptic trap. The framework we have built is first tested against analytic solutions of simple non-interacting systems. From there the system is extended to include perturbed harmonic oscillator traps and particle interaction. We compute one-body densities for the interacting case and discuss why this tool is important and how our toy model has similar attributes to real-world systems.

CONTENTS

I. INTRODUCTION

We will in this project study the Variational Monte Carlo method, and use it to evaluate the ground state energy of a trapped, hard sphere Bose gas. [1] [2]

By building up an increasingly intricate system we eloquently illustrate the strength of the Variational Monte Carlo methods. Any Monte Carlo method draws on the fact that the computer will not complain when being tasked to perform a very similar task up to several million times in rapid succession. Monte Carlo methods is in essence a series of guesses where the current guess is only a minor alteration of the one prior to it. The trick lies in only accepting the subsequent guesses that improves upon our configuration up to a certain likelihood.

The method we have employed is called a *Variational* Monte Carlo method, because we make an educated guess at a trial wavefunction which we allow certain degrees of freedom called *variational parameters*. We shall see that the system converges to a proper [1] minimum for only a certain set of parameters.

First, we provide the theoretical foundation for the Variational Monte Carlo method. This includes explaining how Monte Carlo integration works, and introducing the important quantities *local energy*, *drift force* and *one-body densities*.

Second, we outline in more detailed algebra the nature of a systems we will be studying. For the non-interacting system we are able to derive an exact variational energy to which we can compare the results from our simulations. We furthermore give a description of an interacting system implemented as if each particle is a hard sphere.

Third, we provide a thorough description of algorithms and statistical methods employed in this study. Here we introduce the Metropolis-Hastings algorithm. This algorithm can by a deft trick be expanded with the method of

* Project code: https://github.com/Schoyen/FYS4411

[1] Global, as far as we can tell.

importance sampling yielding a more intelligent choice for the next step in the Monte Carlo cycling. The main point of note with regards to the statistical analysis is that the data generated by the Monte Carlo cycling is autocorrelated and computation of variance must be done with the *blocking method* for time series data. As an automated way of finding the optimal variational parameter, we sketch out the method of *gradient descent*.

Fourth, we put the machinery described in the previous sections to work. We test against the actual values in a non-interacting system before analysing the more interesting perturbed-trap system with interacting particles.

Lastly, we highlight interesting aspects of this study in a discussion of the results before ending on some summary remarks.

II. VARIATIONAL MONTE CARLO

The variational principle states that for a given *trial wavefunction*, $|\Psi_T\rangle$, the expectation value of the Hamiltonian, H, will be an upper bound to the ground state energy, E_0 of the system.

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

The complicated part of the variational principle is finding the minimum, or values close to the minimum. For a select few systems we can calculate this analytically, but for the interesting systems we quickly hit an intractable wall. To get around this we try to calculate the energy for a given Hamiltonian and a trial wavefunction for different variational parameters. The variance of the variational energy is given by

$$\sigma^2 = \langle\Psi_T| H^2 |\Psi_T\rangle - \langle\Psi_T| H |\Psi_T\rangle^2 . \tag{2}$$

If we have found the "true" trial wavefunction $|\Psi_T\rangle$, i.e., the eigenfunction to the Hamiltonian H, we get $\sigma^2 = 0$. We are thus interested in trying to find the variational parameters in the trial wavefunction that minimizes the variance σ^2.

A. Monte Carlo integration

Unfortunately the evaluation of the expectation value of the energy for a given trial wavefunction involves a large integral.

$$E = \frac{\langle\Psi_T| H |\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} = \frac{\int \mathrm{d}\mathbf{r}\Psi_T^*(\mathbf{r})H\Psi_T(\mathbf{r})}{\int \mathrm{d}\mathbf{r}|\Psi_T|^2}, \tag{3}$$

where $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$ are the positions[2] of the particles contained in the system.

To get around[3] the trouble of evaluating the multidimensional-integral we employ the method of Monte Carlo integration. This consists of approximating an integral of the expectation value of a certain quantitity by a sum.

$$\langle x\rangle = \int_{\mathbb{R}} xp(x)\mathrm{d}x \approx \frac{1}{N}\sum_{i=1}^{N} x_i p(x_i), \tag{4}$$

where x is a random variable with a probability density function $p(x)$ and $\{x_i\}$ for $i \in \{1, \ldots, N\}$ is the set of drawn random variables. Here N is the number of Monte Carlo samples chosen to approximate the integral. By the law of large numbers we expect that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} x_i p(x_i) = \langle x\rangle . \tag{5}$$

For us to be able to take Monte Carlo integration to the quantum realm we have to construct a probability density function from the wavefunction. Luckily the wavefunction squared is just such a quantity. Given a normalized wavefunction $|\psi\rangle$ we can compute the expectation value of a quantum observable $\mathcal{O}$ using Monte Carlo integration.

$$\langle\mathcal{O}\rangle = \langle\psi| \mathcal{O} |\psi\rangle = \int \mathrm{d}\mathbf{r}\psi^*(\mathbf{r})\mathcal{O}\psi(\mathbf{r}) \tag{6}$$

$$= \int \mathrm{d}\mathbf{r}|\psi(\mathbf{r})|^2 \frac{\mathcal{O}\psi(\mathbf{r})}{\psi(\mathbf{r})}, \tag{7}$$

where we have multiplied and divided by the wavefunction. We now define the probability density function, $p(\mathbf{r})$, as

$$p(\mathbf{r}) = |\psi(\mathbf{r})|^2, \tag{8}$$

and the local quantum observable

$$\mathcal{O}_L(\mathbf{r}) = \frac{\mathcal{O}\psi(\mathbf{r})}{\psi(\mathbf{r})}. \tag{9}$$

We see that if the chosen wavefunction is an eigenstate of the quantum operator, the local quantum observable becomes a constant. We can thus approximate the expecatation value of the quantum observable using Monte Carlo integration.

$$\langle O\rangle = \int \mathrm{d}\mathbf{r}p(\mathbf{r})\mathcal{O}_L(\mathbf{r}) \approx \frac{1}{N}\sum_{i=1}^{N} p(\mathbf{r}_i)\mathcal{O}_L(\mathbf{r}_i). \tag{10}$$

In the case of a non-normalized wavefunction we need to divide by a normalization factor in the probability density function.

[2] This can include both spatial and spin parts, but we will in the remainder of the project ignore spin.

[3] Not exactly *get around* per se, but make the problem solvable at all.

B. Local energy

As the many-body wavefunction creates a very large configuration space, where much of the wavefunction is small, we use the Metropolis algorithm in order to move towards regions in configuration space with "sensible" values. We define the *local energy*, $E_L(\mathbf{r})$, by

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

If $\Psi_T(\mathbf{r})$ is an exact eigenfunction of the Hamiltonian, E_L will be constant. The closer $\Psi_T(\mathbf{r})$ is to the exact wave function, the less variation in E_L as a function of $\mathbf{r}$ we get. This leads us to searching for values of $\mathbf{r}$ yielding the lowest variance in E_L.

When we are performing the Monte Carlo sampling we are interested in the expected value of the local energy as this will serve as our estimate of the true energy. Mathematically this is expressed as

$$\langle E_L \rangle = \frac{\int \mathrm{d}\mathbf{r}\Psi_T^*(\mathbf{r})H\Psi_T(\mathbf{r})}{\int \mathrm{d}\mathbf{r}|\Psi_T(\mathbf{r})|^2} = \frac{\int \mathrm{d}\mathbf{r}|\Psi_T(\mathbf{r})|^2 E_L(\mathbf{r})}{\int \mathrm{d}\mathbf{r}|\Psi_T(\mathbf{r})|^2}, \tag{12}$$

where we have multiplied and divided by the trial wavefunction to get the latter expression.

One of the most computationally intensive parts of the Variational Monte Carlo algorithm will be to compute E_L. We will therefore find an analytical expression for E_L in terms of the trial wavefunctions and the system we are exploring.

C. The drift force

A disadvantage in the use of the brute force Metropolis-Hastings algorithm is that we might be spending much computational resources in an uninteresting part of configuration space. To make smarter moves we will use the Metropolis-Hastings algorithm with the additions of *Importance Sampling* (which will be discussed in due time). This alteration of the algorithm is dependent on the *drift force* of the system[4].

$$\mathbf{F}(\mathbf{r}) = \sum_{k=1}^{N} \mathbf{F}_k(\mathbf{r}) = \sum_{k=1}^{N} \frac{2\nabla_k \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})}. \tag{13}$$

Using this expression we are able to move towards parts of configuration space where the gradient increases or decreases yielding a better choice of movements. We will mainly be interested in the drift force of a single particle k as the Monte Carlo sampling involves choosing a single particle at a time.

[4] Also known as the *quantum force*.

D. One-Body Densities

In order to make some conclusion towards where the particles in our systems are most prone to exist, we need a good way to visualise our wavefunctions. It is impossible for a human being to visualise a probability density in three dimensions, because this measure would have to be represented in four dimensions. The fact that we will be studying systems of up to $n = 500$ particles only adds to the headache. We therefore introduce the single-particle density function, or *one-body density* $\rho(x,t)$. The one-body density is defined by the integral over all coordinates except one,

$$\rho(x_1) \equiv \int \left|\Psi(x_1,\ldots,x_n)\right|^2 dx_2 \ldots dx_n. \tag{14}$$

The one-body density can be interpreted as one-dimensional probability density for finding a particle in x_1-direction. We also have the following relation,

$$n = \int \rho(x_1) dx_1, \tag{15}$$

in other words, the integral over the entire one-dimensional space in x_1-direction gives us the total number of particles in our system. This can be a bit odd at first sight as we are used to getting the number 1 when we integrate over the entire density.

It is best to think of the one-body density as the multivariate probability density for the entire system collapsed down to one dimension in a representation than looks very familiar to every physicist with a basic knowledge of quantum mechanics.

III. SYSTEMS

To model the trapped bosonic gas particles we use the potential

$$v(\mathbf{r}) = \begin{cases} \frac{1}{2}m\omega^2 r^2 & \text{(S)}, \\ \frac{1}{2}m\left[\omega^2(x^2+y^2)+\omega_z^2 z^2\right] & \text{(E)}, \end{cases} \tag{16}$$

where we can choose between a spherical (S) or an elliptical (E) harmonic oscillator trap. In the latter potential the oscillatory frequency is different in the z-direction, i.e., $\omega_z \neq \omega$. The full Hamiltonian of the system is given by

$$H = \sum_{i=1}^{N} h(\mathbf{r}_i) + \sum_{i<j}^{N} w(\mathbf{r}_i, \mathbf{r}_j), \tag{17}$$

where the single particle one-body operator, $h(\mathbf{r}_i)$, is given by

$$h(\mathbf{r}_i) = -\frac{\hbar^2}{2m}\nabla_i^2 + v(\mathbf{r}_i), \tag{18}$$

for equal masses, and the two-body interaction operator, $w(\mathbf{r}_i, \mathbf{r}_j)$, is

$$w(\mathbf{r}_i, \mathbf{r}_j) = \begin{cases} \infty & |\mathbf{r}_i - \mathbf{r}_j| \leq a, \\ 0 & |\mathbf{r}_i - \mathbf{r}_j| > a, \end{cases} \tag{19}$$

where a is the hard sphere of the particle. The trial wavefunction, $\Psi_T(\mathbf{r})$, we will be looking at is given by

$$\Psi_T(\mathbf{r}) = \Phi_T(\mathbf{r}) \prod_{j<k}^{N} f(a, \mathbf{r}_j, \mathbf{r}_k) \tag{20}$$

$$= \left(\prod_{i=1}^{N} g(\alpha, \beta, \mathbf{r}_i) \right) \prod_{j<k}^{N} f(a, \mathbf{r}_j, \mathbf{r}_k), \tag{21}$$

where α is a variational parameter and

$$\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, \alpha, \beta). \tag{22}$$

Here g are the single particle wavefunctions given by

$$g(\alpha, \beta, \mathbf{r}_i) = \exp\left[-\alpha(x_i^2 + y_i^2 + \beta z_i^2)\right] \equiv \phi(\mathbf{r}_i), \tag{23}$$

and $\Phi_T(\mathbf{r})$ is the *Slater permanent* consisting of the N first single particle wavefunctions, and the correlation wavefunction, f, given by

$$f(a, \mathbf{r}_j, \mathbf{r}_k) = \begin{cases} 0 & |\mathbf{r}_j - \mathbf{r}_k| \leq a, \\ \left(1 - \frac{a}{|\mathbf{r}_j - \mathbf{r}_k|}\right) & |\mathbf{r}_j - \mathbf{r}_k| > a. \end{cases} \tag{24}$$

We will for brevity use the notation $\phi(\mathbf{r}_i) = \phi_i$ and $r_{jk} = |\mathbf{r}_j - \mathbf{r}_k|$. In practice Equation 19 is unnecessary as the situtation where $r_{jk} \leq a$ automatically yields a probability density of zero from Equation 24 thus rejecting the state completely.

A. Non-interacting harmonic oscillators

We start by looking at a simple system of non-interacting harmonic oscillators, where $a = 0$ and $\beta = 1$. This means that there is no hard-shell interaction between the particles and that we use a spherical potential trap. We thus get the trial wavefunction

$$\Psi_T(\mathbf{r}) = \Phi_T(\mathbf{r}) = \prod_{i=1}^{N} \exp\left(-\alpha |\mathbf{r}_i|^2\right), \tag{25}$$

where $|\mathbf{r}_i| = r_i$. As $a = 0$ the interaction term, $w(\mathbf{r}_i, \mathbf{r}_j)$, vanishes and the Hamiltonian is given by (in the spherical case)

$$H = \sum_{i=1}^{N} h(\mathbf{r}_i) = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} m\omega^2 r_i^2 \right). \tag{26}$$

To find the drift force and the local energy we have to compute the gradient and the Laplacian of the trial wavefunction. The gradient is given by

$$\nabla_k \Psi_T(\mathbf{r}) = -2\alpha \mathbf{r}_k \Psi_T(\mathbf{r}), \tag{27}$$

whereas the Laplacian yields

$$\nabla_k^2 \Psi_T(\mathbf{r}) = \big(-2d\alpha + 4\alpha^2 r_k^2 \big) \Psi_T(\mathbf{r}), \tag{28}$$

where d is the dimensionality of the problem determined by $\mathbf{r}_k \in \mathbb{R}^d$. We can thus use the gradient to find an expression for the drift force for particle k.

$$\mathbf{F}_k(\mathbf{r}) = -4\alpha \mathbf{r}_k. \tag{29}$$

Using the Laplacian we can compute the kinetic term in the expression for the local energy. We get

$$E_L(\mathbf{r}) = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m} \big[-2d\alpha + 4\alpha^2 r_i^2 \big] + \frac{1}{2} m\omega^2 r_i^2 \right). \tag{30}$$

In natural units, $\hbar = c = 1$, and with the mass set to unity, $m = 1$, this reduces to

$$E_L(\mathbf{r}) = \alpha d N + \left(\frac{1}{2}\omega^2 - 2\alpha^2 \right) \sum_{i=1}^{N} r_i^2. \tag{31}$$

It is worth noting that for $\alpha = \omega/2$[5] we get a stable value for the local energy.

$$E_L(\mathbf{r}) \equiv E_L = \frac{\omega d N}{2}, \tag{32}$$

which is the stable minimum for $\alpha = \omega/2$.

1. Exact variational energy

As the system is non-interacting and consisting of Gaussians we can find an expression for the exact energy as a function of the variational parameter α, i.e.,

$$E(\alpha) = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \tag{33}$$

$$= \left(\frac{\hbar^2 \alpha}{2m} + \frac{m\omega^2}{8\alpha} \right) dN. \tag{34}$$

By minimizing this expression with respect to α yields

$$\frac{\mathrm{d}E(\alpha)}{\mathrm{d}\alpha} = 0 \implies \alpha_0 = \frac{m\omega}{2\hbar}, \tag{35}$$

which in natural units reduces to $\alpha_0 = \omega/2$. The energy at this value of α (in natural units) is then

$$E(\alpha_0) = \frac{\omega d N}{2}, \tag{36}$$

in perfect accordance with Equation 32 which yields high hopes for the latter equation and the expected minimum.

[5] Note that α is required to be positive.

B. Interacting hard sphere bosons

Moving to the full elliptical system for $\beta \neq 1$ and setting $a \neq 0$ we can write the trial wavefunction as

$$\Psi_T(\mathbf{r}) = \Phi_T(\mathbf{r})J(\mathbf{r}), \tag{37}$$

where $\Phi_T(\mathbf{r})$ is the same Slater permanent as in Equation 21 and $J(\mathbf{r})$ is the *Jastrow factor* given by

$$J(\mathbf{r}) = \exp\left(\sum_{j<l}^{N} u(r_{jl})\right), \tag{38}$$

where

$$u(r_{jk}) = \ln\big[f(a, \mathbf{r}_j, \mathbf{r}_k)\big] \equiv u_{jk} \tag{39}$$

The drift force for particle k in this system becomes

$$\mathbf{F}_k(\mathbf{r}) = 2\left(\frac{\nabla_k \phi_k}{\phi_k} + \sum_{m\neq k}^{N} \nabla_k u_{km}\right). \tag{40}$$

The Laplacian of particle k divided by the trial wavefunction is

$$\frac{\nabla_k^2 \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{\nabla_k^2 \phi_k}{\phi_k} + 2\frac{\nabla_k \phi_k}{\phi_k}\sum_{m\neq k}^{N} \frac{\Delta\mathbf{r}_{km}}{r_{km}}\frac{\partial u_{km}}{\partial r_{km}} + \sum_{m\neq k}^{N}\left(\frac{d-1}{r_{km}}\frac{\partial u_{km}}{\partial r_{km}} + \frac{\partial^2 u_{km}}{\partial r_{km}^2}\right) + \left(\sum_{m\neq k}^{N}\frac{\Delta\mathbf{r}_{km}}{r_{km}}\frac{\partial u_{km}}{\partial r_{km}}\right)^2, \tag{41}$$

where $\Delta\mathbf{r}_{km} = \mathbf{r}_k - \mathbf{r}_m$ and d is the dimensionality. Evaluating the gradient and Laplacian of the single particle functions yields

$$\frac{\nabla_k \phi_k}{\phi_k} = -2\alpha(x_k\mathbf{e}_i + y_k\mathbf{e}_j + \beta z_k \mathbf{e}_k), \tag{42}$$

$$\frac{\nabla_k^2 \phi_k}{\phi_k} = -2\alpha\big(d - 1 + \beta\big) + 4\alpha^2\big(x_k^2 + y_k^2 + \beta^2 z_k^2\big). \tag{43}$$

Repeating this for the correlation functions yields

$$\frac{\partial u_{km}}{\partial r_{km}} = \frac{a}{r_{km}(r_{km} - a)}, \tag{44}$$

$$\frac{\partial^2 u_{km}}{\partial r_{km}^2} = \frac{a^2 - 2ar_{km}}{r_{km}^2(r_{km} - a)^2}. \tag{45}$$

The local energy is thus

$$E_L(\mathbf{r}) = \sum_{k=1}^{N}\left\{-\frac{\hbar^2}{2m}\frac{\nabla_k^2\Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} + v(\mathbf{r}_k)\right\}, \tag{46}$$

where the Laplcian is given by Equation 41. We will avoid writing out the full expression for the local energy for this system due to the size of the equation.

1. Scaling the system

We now introduce a scaled distance $\mathbf{r}' = \mathbf{r}/a_{\text{ho}}$ [1], where

$$a_{\text{ho}} = \sqrt{\frac{\hbar}{m\omega}}. \tag{47}$$

We can then rewrite the Hamiltonian for the elliptic potential in terms of this new distance. By doing a variable substitution for each direction in the Laplace operator we get

$$\nabla_k^2 = \frac{1}{a_{\text{ho}}^2}{\nabla'_k}^2. \tag{48}$$

Looking at the one-body part of the Hamiltonian we get

$$H = \sum_{k=1}^{N}\left\{-\frac{\hbar^2}{2m}\nabla_k^2 + \frac{1}{2}m\left[\omega^2(x_k^2 + y_k^2) + \omega_z^2 z_k^2\right]\right\} \tag{49}$$

$$= \frac{\hbar\omega}{2}\sum_{k=1}^{N}\left\{-{\nabla'_k}^2 + \left[{x'_k}^2 + {y'_k}^2 + \lambda^2 {z'_k}^2\right]\right\}, \tag{50}$$

where we have introduced the dimensionless frequency $\lambda = \omega_z/\omega$. The single particle functions also get a scaling factor

$$\phi_k = \exp\left[-\frac{\alpha\hbar}{m\omega}({x'_k}^2 + {y'_k}^2 + \beta^2 {z'_k}^2)\right] \equiv \phi(\mathbf{r}'_k). \tag{51}$$

The correlation wavefunction and the interaction potential remain the same, but using the scaled hard core radius $a/a_{\text{ho}} = 0.0043$.

IV. ALGORITHMS

In this project we rely on a Monte Carlo approach of random sampling to obtain numerical results. We simulate random walks over a volume in order to find optimal parameters in our trial wavefunctions. The most common of such methods, which we make use of herein, is the Metropolis-Hastings algorithm.

A. The Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm can in our particular situation be condensed down to the following steps:

1. The system is initialised by a certain number N of randomly generated positions, or particles. This allows us to evaluate the wavefunction at these points and compute the local energy E_L.
2. The initial configuration is changed by setting a new position for one of these particles. The particle is picked at random.
3. A ratio between new wavefunction density and the previous (initial) density is computed and compared to a random number. This acceptance probability decides if the particle move is rejected or accepted. The particle is only allowed to move a predetermined step length.
4. If the particle movement is accepted and the local energy E_L is computed for the new system.
5. Repeat steps until convergence and an optimum is reached.

The algorithm described above can be applied in an "exhaustive" search of the parameter space in order to find the optimal parameters. Whether a proposed move is accepted or not is determined by a transition probability and the acceptance probability. The strength of the algorithm is that the transition algorithm need not be known. For example, the simplest case is to accept the new state, i.e., the new position for the random walker, if the ratio

$$q(\mathbf{r}_{i+1}, \mathbf{r}_i) = \frac{\left|\Psi_T(\mathbf{r}_{i+1})\right|^2}{\left|\Psi_T(\mathbf{r}_i)\right|^2}, \tag{52}$$

where $\mathbf{r}_{i+1}$ are all the positions at step $i+1$, is greater than a uniform probability $p \in [0, 1)$.

1. Importance Sampling

A problem with the naïve Metropolis-Hastings sampling approach is that the sampling of position space is done with no regard for where we are likely to find a particle. This problem can be remedied through *importance sampling*. It is reasonable to assume that the particles we erratically scatter in space are prone to movement towards the peaks of the probability density as dictated by the wave function. Consider therefore the Fokker-Planck equation,

$$\frac{\partial \Psi_T}{\partial t} = D\nabla \cdot (\nabla - \mathbf{F})\,\Psi_T, \tag{53}$$

which describes the evolution in time of a probability density function. In our case this is the trial wavefunction Ψ_T. Originaly an equation that models diffusion, we have a diffusion term D and the drift force Equation 13. In our case the diffusion term D is simply $1/2$ from the kinetic energy (in natural units).

We use the Langevin equation to find the new position of the particle.

$$\frac{\partial \mathbf{r}}{\partial t} = D\mathbf{F}(\mathbf{r}) + \boldsymbol{\eta}, \tag{54}$$

where $\boldsymbol{\eta}$ is a uniformly distributed stochastic variable for each dimenion. Solving Langevin's equation by Euler's method gives a recursive relation for the subsequent new positions of a particle.

$$\mathbf{r}_{i+1} = \mathbf{r}_i + D\mathbf{F}(\mathbf{r}_i)\delta t + \boldsymbol{\xi}\sqrt{\delta t}, \tag{55}$$

for a given time step δt [6] and a normally distributed stochastic variable $\boldsymbol{\xi}$.

Now we need to change the acceptance probability of the metropolis algorithm to something that takes the new sampling method into account.

$$q(\mathbf{r}_{i+1}, \mathbf{r}_i) = \frac{G(\mathbf{r}_{i+1}, \mathbf{r}_i, \delta t)\left|\Psi_T(\mathbf{r}_{i+1})\right|^2}{G(\mathbf{r}_i, \mathbf{r}_{i+1}, \delta t)\left|\Psi_T(\mathbf{r}_i)\right|^2}, \tag{56}$$

where G is the Green's function of the Fokker-Planck equation given by

$$G(\mathbf{r}_{i+1}, \mathbf{r}_i, \delta t) = \exp\left(-\frac{\left[\mathbf{r}_{i+1} - \mathbf{r}_i - D\mathbf{F}(\mathbf{r}_i)\delta t\right]^2}{4D\delta t}\right) \times \frac{1}{(4\pi D\delta t)^{dN/2}}, \tag{57}$$

where d is the dimensionality.

B. Statistical Analysis

If the results of the metropolis sampling were completely uncorrelated, it would be enough to compute the standard deviation in a familiar way,

$$\sigma = \sqrt{\frac{1}{N}(\langle E_L^2\rangle - \langle E_L\rangle^2}, \tag{58}$$

where N is the number of samples, or Monte-Carlo cycles, in the experiment. However, it is reasonable to assume that the data we are dealing with in this study is liable to suffer from *autocorrelation* and Equation 58 does not hold. The prevailing definition of autocorrelation in a data stream or signal is correlation between a delay of the signal and the original signal. One would be interested

[6] Bear in mind that Equation 54 is only valid as $\Delta t \to 0$, a property stemming from the use of Euler's method.

to find the delay, or lag, in the signal at which the "self-correlation" is highest. We refer to this spacing as d, and define the following correlation function,

$$f_d = \frac{1}{n-d}\sum_{k=1}^{n-d}(x_k - \bar{x}_n)(x_{k+d} - \bar{x}_n). \tag{59}$$

The keen reader would have noticed that the function f_d in Equation 59 would be equal to the sample variance for $d = 0$. We can now define the *autocorrelation function*

$$\kappa_d = \frac{f_d}{\mathrm{Var}(x)}, \tag{60}$$

which is equal to 1 if the data exhibits no autocorrealtions, equating to $d = 0$. From the autocorrelation function (60) we in turn define the *autocorrelation time*,

$$\tau = 1 + 2\sum_{d=1}^{n-1}\kappa_d, \tag{61}$$

notice that the autocorrelation time is 1 for a correlation free experiment.

We are now able to make a correction to the expression for the standard deviation improving on Equation 58 by taking correlation into account,

$$\sigma = \sqrt{\frac{1+2\tau/\Delta t}{N}(\langle E_L^2\rangle - \langle E_L\rangle^2)}, \tag{62}$$

where Δt is the time between each sample. The main problem at this point is to find τ, which we do not know for any given system and it is generally very expensive to compute. In order to find a good estimate of τ we use a procedure called *blocking*.

1. Blocking

In the method of blocking we group the samples into blocks of increasing size. If one where to compute the standard deviation for each block, one should see the variance increase with the block size. The standard deviation would only increase up to a certain point, from whence it would stay almost constant. What is happening is that we have reached a point where a particular sample from one block is no longer correlated with a corresponding sample from an adjacent block. The block size for this point of convergence now functions as an estimate for the autocorrelation time τ.

Instead of going by this "manual" method of looking at charts, we will construct a test statistic and let the computer do the necessary considerations in a more "automatic way.

The easiest way to do this is by use of the *automatic blocking scheme*, nicely illustrated with a flow chart in Figure 1. The input parameter for the method is an array of the local energies, but it can be any array of time series data, which is why it is herein referred to as $\mathbf{X}$. Any array of sequential data, where any particular data point depends on the previous data point inhibits all attributes of a time series. Our array of local energies is therefor most definitely a time series.

From $\mathbf{X}$ we compute the variance and sample covariance between the series itself and a one-point lag of the series. This covariance is usually referred to as the first order autocovariance.

$$\gamma_i = \frac{1}{n}\left(\sum_{k=1}^{n}\left((\mathbf{X}_i)_k - (\bar{\mathbf{X}}_i)_k\right)\right) \times \left(\sum_{k=0}^{n-1}\left((\mathbf{X}_i)_k - (\bar{\mathbf{X}}_i)_k\right)\right). \tag{63}$$

Now we "block"! The next step is to perform a transformation of the our time series $\mathbf{x}_i$, where we end up with an array that is half the length of the original time series,

$$(\mathbf{X}_i)_k = \frac{1}{2}\left((\mathbf{X}_i)_{2k-1} + (\mathbf{X}_i)_{2k}\right). \tag{64}$$

This process is were the term "blocking" stems from. There are a number of ways to perform the blocking - we could have picked the data points used in the new $(i+1)$ array randomly, or more orderly. Here we we have picked them sequentially. We have obtained $\frac{n}{2}$ new "primed" stochastic variables. Assuming that the reader is playing close attention, it is now needless to say that it is much easier to perform these sequential blocking steps if we have an array of length $n = 2^d$, where d is integer. Otherwise, problems would arise because one would have arrays of different lengths when the original array is "split in two". The ailment is easy to remedy by excluding an observation from the longest resulting array after splitting.

For each new time series $\mathbf{X}_{i+1}$ a new sample variance and sample covariance is computed. The blocking transformation is allowed to continue until we are we run out of data, i.e. the length of $\mathbf{X}_i < 2$. In other words, there is only a single number left. We now compute a test statistic,

$$M_j = \sum_{k=j}^{d-1} n_k \frac{\left[(n_k - 1)\frac{s_k^2}{n_k^2}y\gamma_k(1)\right]^2}{s_k^4}. \tag{65}$$

As luck would have it, this number M is χ^2-distributed. The following step would be to look up a statistical table and find the first k, such that $M_k \leq q_{d-k}(1-\alpha)$, where q is the statistical limit and α is the significance level that your hear desires. Finally, the correct sample standard deviation is $\frac{s_k^2}{n_k}$. Again, see Figure 1 for a summary of the algorithm.

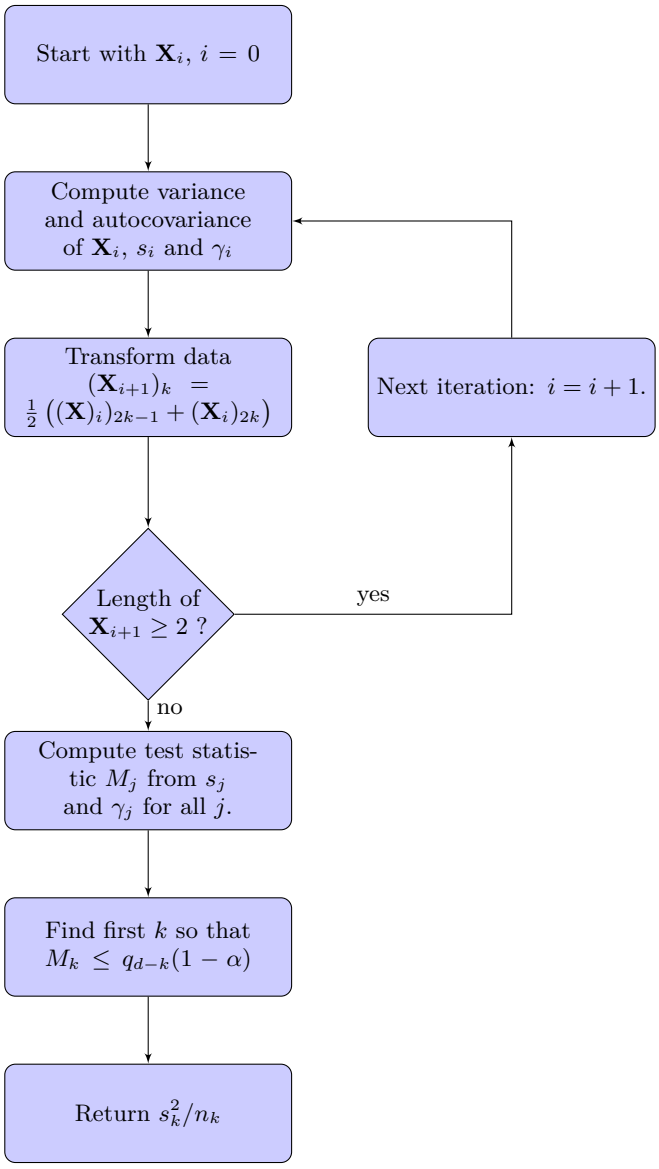

FIG. 1: Flow chart depicting the automatic blocking method for finding a better estimate for sample variance.

C. Gradient Descent

Scanning the entire space of variational parameters can be a somewhat despairing task. To aid us in the search we therefore implement the method of gradient descent[7]. In any optimum we generally have that the derivative should equal zero. In our case, we want to find the minimum of the expected local energy with respect to some variational parameter α, so that

$$\nabla_\alpha \left\langle E_L(\mathbf{R};\alpha)\right\rangle = 0. \tag{66}$$

[7] Also known as steepest descent, easily confused with the steepest descent method for integation.

The foundation for the method of gradient descent lies in the fact that some analytic [8] function $\mathbf{F}(\alpha)$ decreases fastest in the direction of $-\nabla_\alpha \cdot \mathbf{F}(\alpha)$. This means that we could eventually get to an ensuing minimum by trailing along the pathway of the following recursive relation,

$$\boldsymbol{\alpha}_{k+1} = \boldsymbol{\alpha}_k - \gamma \nabla_\alpha \cdot \mathbf{F}(\alpha_k). \tag{67}$$

If γ is small enough we should have $\mathbf{F}(\alpha_k) \geq \mathbf{F}(\alpha_{k+1})$ for all k and hopefully we will experience convergence to the desired minimum.

For our special case we only have one variational parameter, α. The expression for the derivative is,

$$\nabla_\alpha \langle E_L\rangle = 2\left(\left\langle E_L \frac{1}{\Psi_T}\nabla_\alpha \Psi_T\right\rangle - \left\langle \frac{1}{\Psi_T}\nabla_\alpha\Psi_T\right\rangle \langle E_L\rangle\right). \tag{68}$$

To evaluate this expression we sample the local energy, the parameter gradient of the wavefunction, and the combined expression of these at each Monte Carlo cycle.

1. *Stability of the algorithm*

When employing the method of gradient descent to locate a minimum value for the variational parameters we expect some issues due to stability of the Monte Carlo methods and the steepness of the expected energy curves. If our guess of the location of the minimum is far off we can expect the method to diverge as the gradient will become huge. This will yield a new value for $\boldsymbol{\alpha}$ in Equation 67 which will be further away and might result in a local minimum "trapping" the method. To prevent this from happening we have added a limit to the allowed change in $\boldsymbol{\alpha}$. If the change in $\boldsymbol{\alpha}$ is larger than a set order of magnitude we keep dividing $\boldsymbol{\alpha}$ with this number until this criteria is fulfilled.

V. RESULTS

A. Non-Interacting Systems

Initially we are interested in seeing how our implementation performs and to what extent it functions properly. the logical thing to do would be to look at a non-interacting system with a simple Gaussian wavefunction and harmonic oscillator potential. Such a system has a simple closed-form analytic solution[9], which is further simplified by the "god-given" analytic units (Equation 31). This expression is minimised for $\alpha = \frac{1}{2}\omega$. We

[8] We use the term analytic quite recklessly.

[9] A fact that every undergraduate physics student should be well aware of after finishing their first course in quantum physics

TABLE I: One particle in one dimension for the analytic expression with 2^{21} Monte Carlo cycles and step length of 0.5.

α	$\langle E_L \rangle$	σ	σ_b	A	t_{CPU}[s]
0.30	0.56244	0.00026	0.00129	0.892	0.43
0.34	0.53652	0.00019	0.00094	0.885	0.38
0.38	0.51863	0.00013	0.00063	0.878	0.37
0.42	0.50805	0.00009	0.00038	0.872	0.38
0.46	0.50201	0.00004	0.00017	0.866	0.37
0.50	0.50000	0.00000	0.00000	0.860	0.37
0.54	0.50183	0.00004	0.00015	0.855	0.38
0.58	0.50566	0.00007	0.00029	0.850	0.38
0.62	0.51175	0.00011	0.00041	0.845	0.37
0.66	0.51906	0.00014	0.00052	0.840	0.37
0.70	0.52883	0.00017	0.00062	0.836	0.37

TABLE II: One particle in one dimension for the numeric expression with 2^{21} Monte Carlo cycles and a step length of 0.5.

α	$\langle E_L \rangle$	σ	σ_b	A	t_{CPU}[s]
0.30	0.56675	0.00026	0.00135	0.891	0.53
0.34	0.53714	0.00019	0.00096	0.885	0.52
0.38	0.51836	0.00013	0.00062	0.878	0.52
0.42	0.50717	0.00009	0.00038	0.872	0.53
0.46	0.50184	0.00004	0.00017	0.866	0.52
0.50	0.50000	0.00001	0.00001	0.860	0.52
0.54	0.50133	0.00004	0.00015	0.855	0.52
0.58	0.50520	0.00007	0.00028	0.850	0.52
0.62	0.51120	0.00011	0.00039	0.845	0.52
0.66	0.51896	0.00014	0.00052	0.840	0.52
0.70	0.52794	0.00017	0.00061	0.835	0.51

also set $\omega = 1$ and get an expression for the expected minimum local energy,

$$E_L(\mathbf{r}) = \frac{dN}{2}, \tag{69}$$

which will be used as the benchmark we are ultimately aiming for.

We will start without importance sampling, that is by employing the simpler and more naïve "brute force" Monte Carlo sampling. Moreover, we will compute the value for the local energy analytically as well as with a numerical approach. In the numerical approach we employ the well-know approximation for the second derivative,

$$f''(x) \approx \frac{f(x-h) - 2f(x) + f(x+h)}{h^2}. \tag{70}$$

Table I shows the expected local energy $\langle E_L \rangle$, standard deviations σ, σ_b, the share of accepted Monte Carlo reorientations A and cpu time for the analytic computation at different values for the variational parameter α. Table II shows the same numbers but for the numeric scheme. The first column of standard deviations σ are the "naïve" standard deviations, while σ_b marks the standard deviations computed with the blocking method. One can clearly see, in both the analytic and numeric case, that the former standard deviations are lower than the latter, meaning that autocorrelated data can lead to a lower perceived uncertainty than what is reality. The blocking method takes this fact into account and for this reason, all other standard deviations will be computed with the blocking method from here on.

TABLE III: Comparison of energy, standard deviation and the CPU time for $\alpha = 0.5$ (the correct minimum) for the analytic and numeric schemes in three dimensions and N particles. The step length is 0.5.

	Analytic			Numeric		
N	$\langle E_L \rangle$	σ_b	t_{CPU}[s]	$\langle E_L \rangle$	σ_b	t_{CPU}[s]
1	1.5	0.0	0.50	1.49997	0.00008	0.73
10	15.0	0.0	0.73	14.98278	0.00867	6.40
100	150.0	0.0	4.13	149.81453	0.81501	387.60
500	750.0	0.0	19.16	773.85435	20.69789	9595.02

The first initial observation of note is that we have found an optimum for the system in terms of a minimum energy at $\alpha = \frac{1}{2}$. Moreover, notice how the standard deviation disappears completely. This should already be apparent from Equation 31, as the term involving particle positions disappears for the optimal α.

Systems of optimum variational parameter warrants further investigation, as summed up in Table III. The analytical results are quite predictable and monotonous; the expected local energy $\langle E_L \rangle$ is perfectly proportional by a factor $\frac{3}{2}$ to the number of particles and the standard deviation σ_b is zero for systems of any number of particles. The numerical results tells a more tortuous tale. As the number of particles increase, the uncertainty also increases rapidly. For the largest system, with $N = 500$ particles, the energy misses its target critically - the true energy is more than one standard deviation away. Lastly, the cpu time is notably higher for the numeric scheme.

In the appendices, Figure 7 shows how both the analytic and numeric approach matches the exact expression for the energy as a function of α for systems of increasing size. We see also here that the uncertainty increases in the numeric scheme as the system gets larger.

1. *Introducing Importance Sampling*

Next we will introduce importance sampling, the first prediction for importance sampling compared to brute, naïve method is that we should approach an equilibrium much faster because we are now making moves in a smarter way. This prediction is confirmed in Figure 2 and Figure 3.

Figure 2 shows the standard deviation of the local energies for increasing number of Monte Carlo cycles for both the regular sampling and the importance sampling methods. We see that the standard deviation decreases

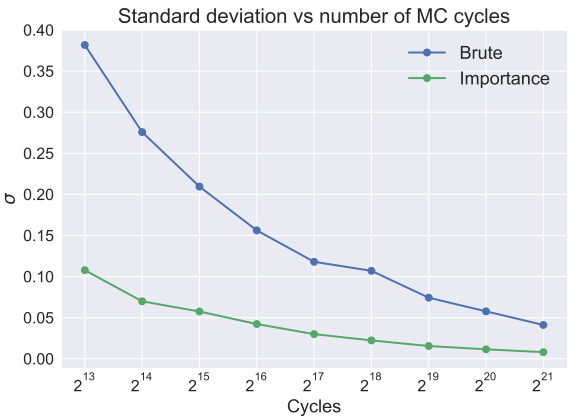

FIG. 2: As the number of Monte Carlo cycles increases the standard deviation of the local energies decrease. This effect is stronger when the importance sampling scheme is employed.

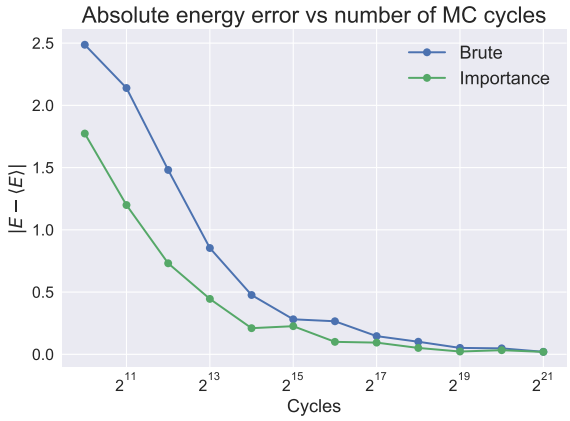

FIG. 3: The absolute difference between the expected energy and the exact energy for both brute force and importance sampling.

for both methods for a larger number of Monte Carlo cycles, but the standard deviation is always lower for the importance sampling method. This is in accordance with our expectations.

We have also compared the expected energy with the exact energy for different number of Monte Carlo cycles. The result of this analysis is shown in Figure 3. We see that both methods will evnetually result in expected energies that are equal to the exact energies as $\left|E - \langle E\rangle\right| \to 0$. However, for the method of importance sampling the error is consistently lower.

Table IV shows the standard deviations and acceptance ratios for different time steps. The simulations in this tabel are from the same span $\alpha \in [0.3, 07]$ as in Table I and Table II. The optimum energies for $\alpha = \frac{1}{2}$ gives the same result for importance sampling as for the naïve way of sampling. Notice that by lowering the time step, the uncertainty starts to decrease and then starts to increase again. Moreover, the acceptance ratios increases as the time step decreases.

TABLE IV: Importance sampled simulations of $n = 10$ particles in $d = 3$ dimensions. Average acceptance ratios and standard deviations for eleven alpha values between 0.3 and 0.7.

δt	$\bar{\sigma}_b$	$\bar{A}$	$\langle E\rangle\,(\alpha = 1/2)$
2^{+3}	0.09363	0.005	15.000
2^{+2}	0.01549	0.041	15.000
2^{+1}	0.00706	0.216	15.000
$2^{\ 0}$	0.00675	0.510	15.000
2^{-1}	0.00932	0.738	15.000
2^{-2}	0.01578	0.866	15.000
2^{-3}	0.02723	0.933	15.000
2^{-4}	0.05016	0.966	15.000
2^{-5}	0.08474	0.983	15.000
2^{-6}	0.11555	0.992	15.000
2^{-7}	0.17070	0.996	15.000

B. Interacting elliptical harmonic oscillator

We now make a change to the simple systems in the previous section in order to make everything a bit more interesting. The potential in the one-body Hamiltonian is perturbed in the z-direction according to Equation 50 by setting $\lambda = \sqrt{8} \approx 2.82843$. This corresponds to enclosing particles of the simulation in an elliptical trap instead of a spherical trap.

We also choose a new trial wavefunction with the elliptical gaussian single particle functions in Equation 51 and the Jastrow factor. The hard sphere radius is set to $a/a_{\text{ho}} = 0.0043$.

Due to similarities with the spherical harmonic oscillator system we make a guess that the true minimum of this new system should be situated close to the minimum of the previous system. Choosing seven values for $\alpha \in [0.2, 0.7]$ we run importance sampling on this system for $N = \{10, 50, 100\}$ particles in $d = 3$ dimensions using 2^{21} Monte Carlo cycles[10] and an additional 10% of the Monte Carlo cycles for thermalization of the system prior to doing any sampling. We have used $\delta t = 0.1$ yields a high acceptance ratio, but slower convergence.

In Table V the system is set to have $N = 10$ interacting particles. Notice first that the energy is higher in the interacting system than for a comparable non-interacting system (Table III). Moreover, the standard deviation gets lower as the system is close to an optimal variational parameter.

[10] This turned out to be a little bit overkill due to the fast convergence of importance sampling.

TABLE V: Simulation results for $N = 10$ bosonic interacting, elliptical, harmonic oscillators.

α	$\langle E_L \rangle$	σ_b	A	t_{CPU}[s]
0.2	35.17548	0.05399	0.990	24.5
0.3	27.62004	0.02311	0.981	13.6
0.4	24.97850	0.00827	0.972	13.5
0.5	24.39877	0.00030	0.961	13.4
0.6	24.83863	0.00604	0.950	13.3
0.7	25.82855	0.01059	0.938	13.2
0.8	27.22895	0.01436	0.924	13.1

TABLE VI: Simulation results for $N = 50$ bosonic interacting, elliptical, harmonic oscillators.

α	$\langle E_L \rangle$	σ_b	A	t_{CPU}[s]
0.2	181.69371	0.25533	0.987	199.9
0.3	142.61591	0.10782	0.978	195.8
0.4	129.88615	0.04051	0.969	190.4
0.5	127.29926	0.00595	0.957	195.1
0.6	129.97630	0.03300	0.946	193.2
0.7	135.67382	0.05785	0.933	192.0
0.8	143.23238	0.07286	0.921	192.7

The result of simulations with a higher number of particles, $N = 50$ and $N = 100$, are shown in Table VI and Table VII respectively. We observe that the energy for systems of interacting particles increases rapidly as the size of the system increases.

1. *Finding the minimum*

Using the method of gradient descent we choose a set of six starting $\alpha_0 \in [0.2, 0.8]$ and try to locate a value for α where the gradient goes to zero. We looked at $N = 10$ with $d = 3$, $\omega = 1$ and $\beta = \lambda = \sqrt{8}$. This yields the value for the minimum α to be

$$\alpha = 0.49744 \pm 0.00002. \tag{71}$$

In Figure 4 we can see how the conjugate gradient method "moves" to find the optimal value of α.

For this optimal, interacting α we have computed the expected energy and standard deviation for systems of

TABLE VII: Simulation results for $N = 100$ bosonic interacting, elliptical, harmonic oscillators.

α	$\langle E_L \rangle$	σ_b	A	t_{CPU}[s]
0.2	375.04401	0.51475	0.985	745.2
0.3	296.76879	0.20214	0.975	741.0
0.4	270.60830	0.08113	0.945	723.1
0.5	266.37263	0.02020	0.953	702.8
0.6	272.51171	0.07366	0.941	689.1
0.7	285.10285	0.11381	0.929	707.5
0.8	301.40609	0.15458	0.916	707.9

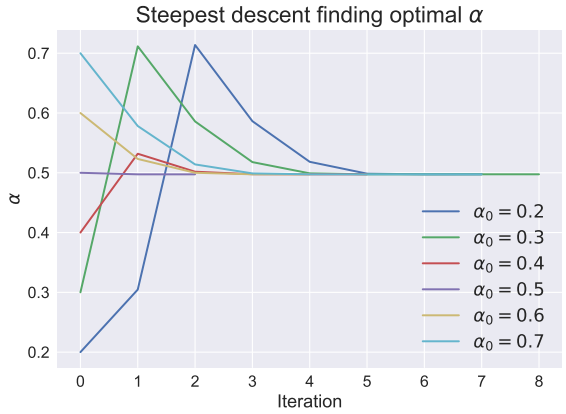

FIG. 4: In this figure we can see the convergence of different starting values for α towards the true value of α that minimizes the expected energy in the interacting case.

TABLE VIII: Comparison of the elliptical system with an elliptical gaussian trial wavefunction with and without the Jastrow factor.

	Interacting		Non-interacting	
N	$\langle E \rangle$	σ_b	$\langle E \rangle$	σ_b
10	24.3986	0.0003	24.1418	0.0002
50	127.2852	0.0077	120.7129	0.0012
100	266.2904	0.0297	241.4248	0.0025

$n = \{10, 50, 100\}$ particles. These results are shown in Table VIII, the non-interacting case, excluding Jastrow factor is included as a comparison.

2. *One-Body Densities*

We now look at one-body densities for the value of α found in Equation 71. The results for a system consisting of $n = 100$ particles are displayed in Figure 5. In this figure we see the densities for for a system confined to a perturbed elliptic trap with and without interaction. We see that for the interacting case, the one-body density is skewed towards a larger $|\mathbf{r}|$, as is expected. The "ideal" density, a system with spherical trap, is also included in the figure.

VI. DISCUSSION

A. Validity of Results

The non-interacting tests compared against the analytic "solution" showed to a very high degree the validity of the method. One can see this quite clearly from Figure 7. Because we get results from simulations that cor-

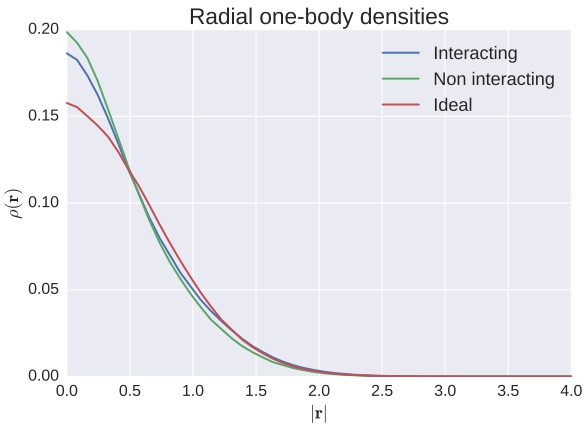

FIG. 5: In this figure the interacting and non-interacting densities correspond to the elliptical potential with and without the Jastrow factor respectively. The "ideal" system is the spherical potential with the non-interacting Gaussian trial wavefunction.

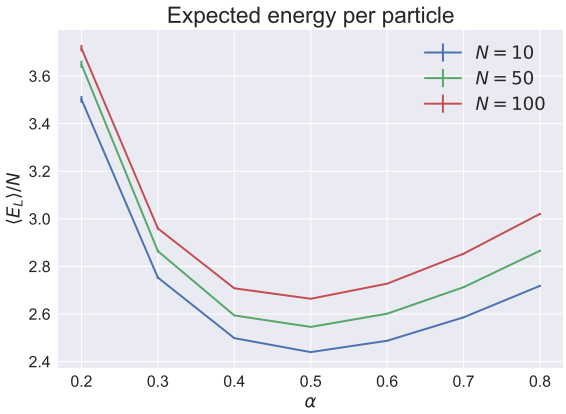

FIG. 6: Local energies as a function of variational parameter α for different number of particles.

respond to the analytic results we can be fairly certain that the method can be expanded upon to include attributes outside the realm problems that are analytically solvable. The systems with elliptic traps and interacting particles are such systems.

That said, the way to get the best confirmation of the results presented above is in comparison to experiment. This would be quite difficult as the systems presented are very theoretical in nature and most likely far from anything we would see in nature. Without being able to compare our result with anything happening in the real world this entire study would have no value except for the instrumental value, being no more than a theoretical exercise. However, the one-body density is a measure of the role of many-body correlations which in some cases can be related to experiment. For a system of charged particles it can be used to determine a charge distribution, which in turn can be linked to experiment. It would be very fruitful to simulate a system of such particles that is more natural, i.e. real, but we realise the purpose of this study is first and foremost to introduce variational Monte Carlo as a tool in computational physics. We assume, however, that this study is easy enough to expand into a study of alkali atoms like ^{87}Rb, ^{23}Na or ^{7}Li.

B. Energy per particle

For the non-interacting harmonic oscillator system there is no difference between a new "particle" and new dimensions [11]. This means that the inclusion of a new particle or dimension yields a linear scaling in the energy. Thus if we compute the energy per particle for systems of varying size we expect to get the same energy. This is most easily seen in Figure 7 and in the expression for the exact energy Equation 34. This, however, is *not* the case in the interacting system due to the Jastrow factor. Looking at Table V and Table VII we see that multiplying the expected energy of the former with 10 yields a lower number than the latter. We can also see this behavior in Figure 6, where the energy per particle for $N = \{10, 50, 100\}$ have been plotted in the same figure.

The reason for the increase in energy per particle can be ascribed to the repulsive interaction added by the Jastrow factor. The repulsive force can be explain thusly; The Jastrow factor increases the value of the wavefunctions for any given particle if that particle is some distance away from any other particle. An increased value in the wavefucntion is equivalent to an increase in probability amplitude, again correpsonding to an increase in probability density, giving rise the the fact that particles are more likely to appear further from one another.

C. Low density system

It is safe to say that the effect of adding a Jastrow factor is significant, i.e. the increase in energy for the system as the interaction is switched on is much higher than the uncertainty of the energy. A simple look at Figure 6 will be proof enough as the error bars of the figure are not discernible. On the other hand, the energy added to the system by including interaction is not very high compared to the energy of the system as a whole. For the larger of the interacting systems we have looked at,

[11] This is not entirely true in the case of our implementation of the Metropolis algorithms as we draw a single random particle and move all dimensions instead of iterating over all particles and all dimensions.

with $n = 100$ particles, The interaction add only a mere 9% extra energy (see Table VIII). Studying Figure 5 will reveal the same.

The reason for this stems from the assumption which justifies that adding only two-body interaction is sufficient to model all interaction to a reasonable extent. This assumption is justified by the fact that alkali and atomic hydrogen systems[12] are dilute. For gases of alkali atoms confined in magnetic or optical traps the average distance between atoms is much larger than the range of interatomic interaction. The probability is high that atoms very seldom interact and if so, just with one other atom.

VII. SUMMARY REMARKS

In this study we have built a functional variational Monte Carlo solver framework for a gas following Bose-Einstein statistics confined in an elliptic trap. Monte Carlo methods is a broad class of computational algorithms, whereas the variational Monte Carlo is a subset. Other applications of such methods within the physical sciences includes, but is not limited to, quantum chromodynamics, the modeling of aearodynamic forms, particle physics detector design and galaxy evolution. Needless to say, Monte Carlo method is a valuable tool for a computational physicist.

To prove the validity of the framework we tested it against an analytically solved case, an ideal non-interacting system. Thereafter we have shown how our implementation can handle large systems of interacting particles and perturbed traps. Even though this study is somewhat theoretic in nature, we have discussed the possibility of expanding our machinery to a case bearing a larger resemblance to a system found in nature. To this end we have computed one-body densities for the interacting system, which can be compared to real physical properties as measured in an experiment.

[1] J. DuBois and H. Glyde, Physical Review A **63**, 023602 (2001).
[2] J. Nilsen, J. Mur-Petit, M. Guilleumas, M. Hjorth-Jensen, and A. Polls, Physical Review A **71**, 053610 (2005).

[12] Systems similar to our toy model

Appendix A: Computing the drift force and local energy of the full system

Computing the gradient of the wavefunction we get

$$\nabla_k \Psi_T(\mathbf{r}) = \Big[\nabla_k \Phi_T(\mathbf{r})\Big] J(\mathbf{r}) + \Phi_T(\mathbf{r}) \nabla_k J(\mathbf{r}). \tag{A1}$$

The gradient of the Slater permament for particle k is given by

$$\nabla_k \Phi_T(\mathbf{r}) = \nabla_k \phi_k \prod_{i \neq k}^{N} \phi_i = \frac{\nabla_k \phi_k}{\phi_k} \Phi_T(\mathbf{r}). \tag{A2}$$

The gradient of the Jastrow factor is given by

$$\nabla_k J(\mathbf{r}) = J(\mathbf{r}) \nabla_k \sum_{m<n}^{N} u_{mn} \tag{A3}$$

$$= J(\mathbf{r}) \left(\sum_{m=1}^{k-1} \nabla_k u_{mk} \sum_{n=k+1}^{N} \nabla_k u_{kn} \right) \tag{A4}$$

$$= J(\mathbf{r}) \sum_{m \neq k}^{N} \nabla_k u_{km}, \tag{A5}$$

where the gradient of the interaction term splits the anti-symmetric sum into two parts. As $r_{ij} = r_{ji}$ we can combine these sums into a single sum. This in total yields the gradient

$$\nabla_k \Psi_T(\mathbf{r}) = \left(\frac{\nabla_k \phi_k}{\phi_k} + \sum_{m \neq k}^{N} \nabla_k u_{km} \right) \Psi_T(\mathbf{r}). \tag{A6}$$

The Laplacian of the trial wavefunction is found by finding the divergence of Equation A6.

$$\nabla_k^2 \Psi_T(\mathbf{r}) = \left(\nabla_k \left[\frac{\nabla_k \phi_k}{\phi_k} \right] + \sum_{m \neq k}^{N} \nabla_k^2 u_{km} \right) \Psi_T(\mathbf{r}) \tag{A7}$$

$$+ \left(\frac{\nabla_k \phi_k}{\phi_k} + \sum_{m \neq k}^{N} \nabla_k u_{km} \right)^2 \Psi_T(\mathbf{r}), \tag{A8}$$

where the squared term came from taking the gradient of the trial wavefunction. To further simplify we divide by the trial wavefunction. This yields

$$\frac{\nabla_k^2 \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{\nabla_k^2 \phi_k}{\phi_k} + 2 \frac{\nabla_k \phi_k}{\phi_k} \sum_{m \neq k}^{N} \nabla_k u_{km} + \sum_{m \neq k}^{N} \nabla_k^2 u_{km} + \left(\sum_{m \neq k}^{N} \nabla_k u_{km} \right)^2. \tag{A9}$$

From here, the next step is to find the gradient and the Laplacian of the single particle functions, ϕ_k, and the interaction functions u_{km}. For the single particle functions we use Cartesian coordinates when finding the derivatives while for the interaction functions we use spherical coordinates and do a variable substitution. Beginning with the gradient of the single particle functions we get

$$\nabla_k \phi_k = \nabla_k \exp\big[-\alpha(x_k^2 + y_k^2 + \beta z_k^2)\big] \tag{A10}$$

$$= -2\alpha(x_k \mathbf{e}_i + y_k \mathbf{e}_j + \beta z_k \mathbf{e}_k)\phi_k. \tag{A11}$$

Note that the subscripts on the unit vectors $\mathbf{e}_i$ are *not* the same as the subscripts used for its components. The Laplacian yields

$$\nabla_k^2 \phi_k = \Big[-2\alpha\big(d - 1 + \beta\big) + 4\alpha^2 \big(x_k^2 + y_k^2 + \beta^2 z_k^2\big) \Big] \phi_k, \tag{A12}$$

with d as the dimensionality of the problem. In order to derive the interaction functions we have to do a variable substitution using $r_{km} = |\mathbf{r}_k - \mathbf{r}_m|$. We can then rewrite the ∇_k-operator as

$$\nabla_k = \nabla_k \frac{\partial r_{km}}{\partial r_{km}} = \nabla_k r_{km} \frac{\partial}{\partial r_{km}} \tag{A13}$$

$$= \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\partial}{\partial r_{km}}. \tag{A14}$$

Applying this version of the ∇_k-operator to u_{km} yields

$$\nabla_k u_{km} = \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}}. \tag{A15}$$

For the Laplacian we switch a little back and forth between the two ways of representing the ∇_k-operator. We thus get

$$\nabla_k^2 u_{km} = \frac{\nabla_k \mathbf{r}_k}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \left[\nabla_k \frac{1}{r_{km}} \right] (\mathbf{r}_k - \mathbf{r}_m) \frac{\partial u_{km}}{\partial r_{km}} + \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \nabla_k \frac{\partial u_{km}}{\partial r_{km}} \tag{A16}$$

$$= \frac{d}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} - \frac{(\mathbf{r}_k - \mathbf{r}_m)^2}{r_{km}^3} \frac{\partial u_{km}}{\partial r_{km}} + \frac{(\mathbf{r}_k - \mathbf{r}_m)^2}{r_{km}^2} \frac{\partial^2 u_{km}}{\partial r_{km}^2} \tag{A17}$$

$$= \frac{d-1}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \frac{\partial^2 u_{km}}{\partial r_{km}^2}, \tag{A18}$$

where d is again the dimensionality of the problem. In total we can state an intermediate version of the Laplacian occuring in the local energy as

$$\frac{\nabla_k^2 \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{\nabla_k^2 \phi_k}{\phi_k} + 2\frac{\nabla_k \phi_k}{\phi_k} \sum_{m \neq k}^{N} \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \sum_{m \neq k}^{N} \left(\frac{d-1}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \frac{\partial^2 u_{km}}{\partial r_{km}^2} \right) + \sum_{m,n \neq k}^{N} \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\mathbf{r}_k - \mathbf{r}_n}{r_{kn}} \frac{\partial u_{km}}{\partial r_{km}} \frac{\partial u_{kn}}{\partial r_{kn}}. \tag{A19}$$

Moving on to the derivatives of the interaction terms, u_{km}, to get an explicit expression for the Laplacian.

$$\frac{\partial u_{km}}{\partial r_{km}} = \frac{a}{r_{km}(r_{km} - a)}, \tag{A20}$$

$$\frac{\partial^2 u_{km}}{\partial r_{km}^2} = \frac{a^2 - 2a r_{km}}{r_{km}^2 (r_{km} - a)^2}. \tag{A21}$$

The local energy and the drift force can now be found by combining these expressions. We will not write out the explicit expressions as these will be called by separated functions in our programs.

Appendix B: Brute Force Metropolis-Hastings

Appendix C: Variational parameter gradient of the expectation energy

Here we show how to arrive at the expression shown in Equation 68. This requires us to restrict our view to real trial wavefunctions, i.e., $\Psi_T^* = \Psi_T$, which is the case for the wavefunctions we are exploring.

$$\nabla_\alpha \langle E_L \rangle = \nabla_\alpha \left(\frac{\int \mathrm{d}\mathbf{r} \Psi_T^* H \Psi_T}{\int \mathrm{d}\mathbf{r} |\Psi_T|^2} \right) \tag{C1}$$

$$= -\frac{2\left(\int \mathrm{d}\mathbf{r} \Psi_T H \Psi_T\right)}{\left(\int \mathrm{d}\mathbf{r} \Psi_T\right)^2} \int \mathrm{d}\mathbf{r} \Psi_T \nabla_\alpha \Psi_T \tag{C2}$$

$$+ \frac{2}{\int \mathrm{d}\mathbf{r} \Psi_T^2} \int \mathrm{d}\mathbf{r} \Psi_T H \left[\nabla_\alpha \Psi_T\right] \tag{C3}$$

$$= -2 \langle E_L \rangle \frac{1}{\int \mathrm{d}\mathbf{r} \Psi_T^2} \int \mathrm{d}\mathbf{r} \Psi_T^2 \left(\frac{\nabla_\alpha \Psi_T}{\Psi_T} \right) \tag{C4}$$

$$+ \frac{2}{\int \mathrm{d}\mathbf{r} \Psi_T^2} \int \mathrm{d}\mathbf{r} \Psi_T^2 E_L \left(\frac{\nabla_\alpha \Psi_T}{\Psi_T} \right). \tag{C5}$$

We now use the definition of the expectation value to group terms together. We are thus left with

$$\nabla_\alpha \langle E_L \rangle = 2 \left(\left\langle E_L \frac{1}{\Psi_T} \nabla_\alpha \Psi_T \right\rangle - \langle E_L \rangle \left\langle \frac{1}{\Psi_T} \nabla_\alpha \Psi_T \right\rangle \right), \tag{C6}$$

which is what we wanted to show.

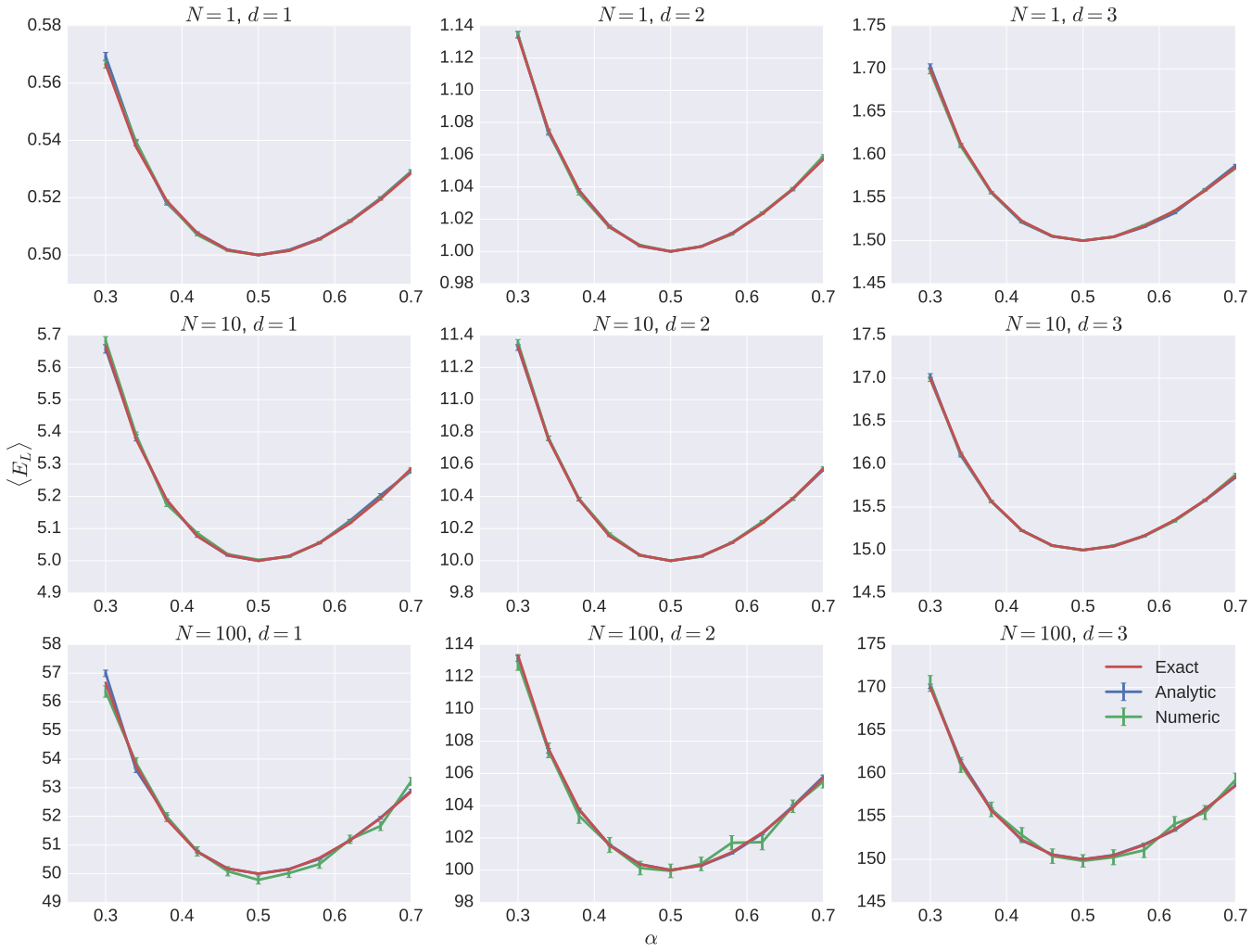

FIG. 7: Comparison of analytic and numeric results with the exact expression for the energy as a function of α. The plots show the standard deviation from the blocking method as error ticks. In the figure N is the number of particles and d the number of dimensions.