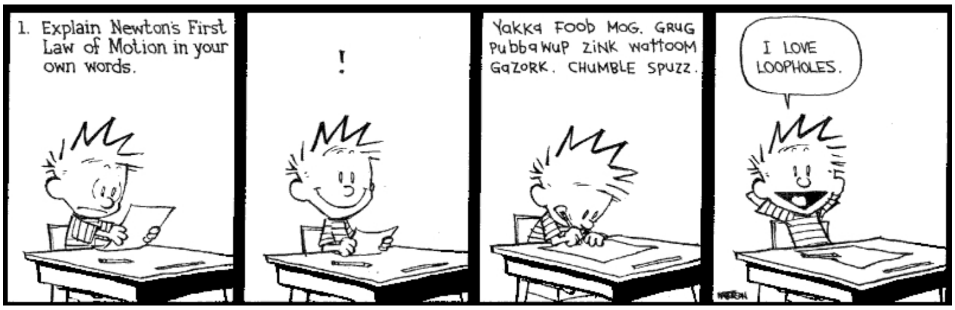
Variational Monte Carlo on bosonic systems

Winther-Larsen, Sebastian Gregorius[1] and Schøyen, Øyvind Sigmundson[1]
[1] *University of Oslo*
(Dated: March 1, 2018)

Something very abstract and clever should go here.



I. INTRODUCTION

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

II. THEORY

To model the trapped bosonic gas particles we use the potential

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

where we can choose between a spherical (S) or an elliptical (E) harmonic trap. The two-body Hamiltonian of the system is given by

$$H = \sum_{i=1}^{N} h(\mathbf{r}_i) + \sum_{i<j}^{N} w(\mathbf{r}_i, \mathbf{r}_j), \tag{2}$$

where the single particle one body operator, h, is given by

$$h(\mathbf{r}_i) = -\frac{\hbar^2}{2m}\nabla_i^2 + V_{\text{ext}}(\mathbf{r}_i), \tag{3}$$

(we assume equal mass) and the two-body interaction operator, w, is

$$w(\mathbf{r}_i, \mathbf{r}_j) = \begin{cases} \infty & |\mathbf{r}_i - \mathbf{r}_j| \leq a, \\ 0 & |\mathbf{r}_i - \mathbf{r}_j| > a, \end{cases} \tag{4}$$

where a is the hard sphere of the particle. The trial wavefunction, $|\Psi_T\rangle$, we will be looking at is given by

$$\Psi_T(\mathbf{r}) = \Phi_T(\mathbf{r}) \prod_{j<k}^{N} f(a, \mathbf{r}_j, \mathbf{r}_k) \tag{5}$$

$$= \left(\prod_{i=1}^{N} g(\alpha, \beta, \mathbf{r}_i)\right) \prod_{j<k}^{N} f(a, \mathbf{r}_j, \mathbf{r}_k), \tag{6}$$

where α and β are variational parameters and

$$\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \alpha, \beta). \tag{7}$$

Here g are the single particle wavefunctions given by

$$g(\alpha, \beta, \mathbf{r}_i) = \exp\left[-\alpha(x_i^2 + y_i^2 + \beta z_i^2)\right] \equiv \phi(\mathbf{r}_i), \tag{8}$$

and $|\Phi_T\rangle$ the *Slater permanent* consisting of the N first single particle wavefunctions, and the correlation wavefunction, f, given by

$$f(a, \mathbf{r}_j, \mathbf{r}_k) = \begin{cases} 0 & |\mathbf{r}_j - \mathbf{r}_k| \leq a, \\ \left(1 - \frac{a}{|\mathbf{r}_j - \mathbf{r}_k|}\right) & |\mathbf{r}_j - \mathbf{r}_k| > a. \end{cases} \tag{9}$$

We will for brevity use the notation $\phi(\mathbf{r}_i) = \phi_i$ and $r_{jk} = |\mathbf{r}_j - \mathbf{r}_k|$.

A. 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

$$E_L(\mathbf{r}) = \frac{H\Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})}. \tag{10}$$

If $|\Psi_T\rangle$ is an exact eigenfunction of the Hamiltonian, E_L will be constant. The closer $|\Psi_T\rangle$ is to the exact wave function, the less variation in E_L as a function of $\mathbf{r}$ we get. One of the most computationally intensive parts of the VMC algorithm will be to compute E_L. We therefore find an analytical expression for E_L in terms of the trial wavefunction.

B. The drift force

A disadvantage in the use of the brute-force Metropolis 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 (which will be discussed in due time). This algorithm is dependent on the drift force of the system.

$$\mathbf{F}(\mathbf{r}) = \sum_{k=1}^{N} \mathbf{F}_k(\mathbf{r}) = \sum_{k=1}^{N} \frac{2\nabla_k \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})}. \tag{11}$$

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.

III. NON-INTERACTING HARMONIC OSCILLATORS

We start by looking at a simple system of non-interacting harmonic oscillators. That is, where $a = 0$ and $\beta = 1$. We thus get the trial wavefunction

$$\Psi_T(\mathbf{r}) = \Phi_T(\mathbf{r}) = \prod_{i=1}^{N} \exp\left[-\alpha |\mathbf{r}_i|^2\right], \tag{12}$$

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)

$$H = \sum_{i=1}^{N} h(\mathbf{r}_i) = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m} \nabla_i^2 + \frac{1}{2} m\omega^2 r_i^2 \right). \tag{13}$$

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

$$\nabla_k \Psi_T(\mathbf{r}) = -2\alpha \mathbf{r}_k \Psi_T(\mathbf{r}), \tag{14}$$

whereas the Laplacian yields

$$\nabla_k^2 \Psi_T(\mathbf{r}) = \left(-2d\alpha + 4\alpha^2 r_k^2 \right) \Psi_T(\mathbf{r}), \tag{15}$$

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.

$$\mathbf{F}_k(\mathbf{r}) = -2\alpha \mathbf{r}_k. \tag{16}$$

Using the Laplcian we can compute the kinteic term in the expression for the local energy. We get

$$E_L(\mathbf{r}) = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m} \left[-2d\alpha + 4\alpha^2 r_i^2 \right] + \frac{1}{2} m\omega^2 r_i^2 \right). \tag{17}$$

In natural units, with $\hbar = c = m = 1$, this reduces to

$$E_L(\mathbf{r}) = \alpha d N + \left(\frac{1}{2}\omega^2 - 2\alpha^2 \right) \sum_{i=1}^{N} r_i^2. \tag{18}$$

It is worth noting that for $\alpha = \frac{1}{2}\omega$ (α is required to be positive) we will find a stable value which turns out to be the exact energy minimum. This happens as the entire sum over all the random walkers disappears.

A. 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.,

$$E(\alpha) = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}. \tag{19}$$

The final result for the energy is

$$E(\alpha) = \left(\frac{\hbar^2 \alpha}{2m} + \frac{m\omega^2}{8\alpha} \right) dN. \tag{20}$$

By minimizing this expression, i.e., finding the derivative of the energy with respect to α and equating this to zero, yields the expected minimum of variational energy to be

$$\frac{\mathrm{d}E(\alpha)}{\mathrm{d}\alpha} = 0 \implies \alpha_0 = \frac{m\omega}{2\hbar}, \tag{21}$$

which in natural units reduces to $\alpha_0 = \frac{1}{2}\omega$. The energy at this value of α (in natural units) is then

$$E(\alpha_0) = \frac{\omega d N}{2}. \tag{22}$$

IV. INTERACTING HARD SPHERE BOSONS

Moving to the full system allowing β to vary and setting $a \neq 0$ we can write the trial wavefunction as

$$\Psi_T(\mathbf{r}) = \Phi_T(\mathbf{r}) J(\mathbf{r}), \tag{23}$$

where $|\Phi_T\rangle$ is the same Slater permanent as in Equation 6 and $J(\mathbf{r})$ is the *Jastrow factor* given by

$$J(\mathbf{r}) = \exp\left(\sum_{j<l}^{N} u(r_{jl}) \right), \tag{24}$$

where $r_{jk} = |\mathbf{r}_j - \mathbf{r}_k|$ and

$$u(r_{jk}) = \ln\big[f(a, \mathbf{r}_j, \mathbf{r}_k)\big]. \tag{25}$$

To further shorten the notation we will use $u_{jk} = u(r_{jk})$. Computing the gradient of the wavefunction we get

$$\nabla_k \Psi_T(\mathbf{r}) = \Big[\nabla_k \Phi_T(\mathbf{r})\Big] J(\mathbf{r}) + \Phi_T(\mathbf{r}) \nabla_k J(\mathbf{r}). \tag{26}$$

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

$$\nabla_k \Phi_T(\mathbf{r}) = \nabla_k \phi_k \prod_{i\neq k}^{N} \phi_i = \frac{\nabla_k \phi_k}{\phi_k} \Phi_T(\mathbf{r}). \tag{27}$$

The gradient of the Jastrow factor is given by

$$\begin{aligned} \nabla_k J(\mathbf{r}) &= J(\mathbf{r}) \nabla_k \sum_{m<n}^{N} u_{mn} && (28) \\ &= J(\mathbf{r}) \left(\sum_{m=1}^{k-1} \nabla_k u_{mk} \sum_{n=k+1}^{N} \nabla_k u_{kn} \right) && (29) \\ &= J(\mathbf{r}) \sum_{m\neq k}^{N} \nabla_k u_{km}, && (30) \end{aligned}$$

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

$$\nabla_k \Psi_T(\mathbf{r}) = \left(\frac{\nabla_k \phi_k}{\phi_k} + \sum_{m\neq k}^{N} \nabla_k u_{km} \right) \Psi_T(\mathbf{r}). \tag{31}$$

The Laplcian of the trial wavefunction is found by finding the divergence of Equation 31.

$$\begin{aligned} \nabla_k^2 \Psi_T(\mathbf{r}) = &\left(\nabla_k \left[\frac{\nabla_k \phi_k}{\phi_k} \right] + \sum_{m\neq k}^{N} \nabla_k^2 u_{km} \right) \Psi_T(\mathbf{r}) && (32) \\ &+ \left(\frac{\nabla_k \phi_k}{\phi_k} + \sum_{m\neq k}^{N} \nabla_k u_{km} \right)^2 \Psi_T(\mathbf{r}), && (33) \end{aligned}$$

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

$$\begin{aligned} \frac{\nabla_k^2 \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = &\frac{\nabla_k^2 \phi_k}{\phi_k} + 2 \frac{\nabla_k \phi_k}{\phi_k} \sum_{m\neq k} \nabla_k u_{km} \\ &+ \sum_{m\neq k}^{N} \nabla_k^2 u_{km} + \left(\sum_{m\neq k}^{N} \nabla_k u_{km} \right)^2. && (34) \end{aligned}$$

To go from here we have 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 whereas we for the interaction functions will use spherical coordinates and do a variable substitution. Beginning with the gradient of the single particle functions we get

$$\begin{aligned} \nabla_k \phi_k &= \nabla_k \exp\big[-\alpha(x_k^2 + y_k^2 + \beta z_k^2)\big] && (35) \\ &= -2\alpha(x_k \mathbf{e}_i + y_k \mathbf{e}_j + \beta z_k \mathbf{e}_k)\phi_k. && (36) \end{aligned}$$

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

$$\begin{aligned} \nabla_k^2 \phi_k = \Big[&-2\alpha\big(d - 1 + \beta\big) \\ &+ 4\alpha^2\big(x_k^2 + y_k^2 + \beta^2 z_k^2\big)\Big]\phi_k, && (37) \end{aligned}$$

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

$$\begin{aligned} \nabla_k &= \nabla_k \frac{\partial r_{km}}{\partial r_{km}} = \nabla_k r_{km} \frac{\partial}{\partial r_{km}} && (38) \\ &= \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\partial}{\partial r_{km}}. && (39) \end{aligned}$$

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

$$\nabla_k u_{km} = \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}}. \tag{40}$$

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

$$\begin{aligned} \nabla_k^2 u_{km} &= \frac{\nabla_k \mathbf{r}_k}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \left[\nabla_k \frac{1}{r_{km}} \right] (\mathbf{r}_k - \mathbf{r}_m) \frac{\partial u_{km}}{\partial r_{km}} + \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \nabla_k \frac{\partial u_{km}}{\partial r_{km}} && (41) \\ &= \frac{d}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} - \frac{(\mathbf{r}_k - \mathbf{r}_m)^2}{r_{km}^3} \frac{\partial u_{km}}{\partial r_{km}} + \frac{(\mathbf{r}_k - \mathbf{r}_m)^2}{r_{km}^2} \frac{\partial^2 u_{km}}{\partial r_{km}^2} && (42) \\ &= \frac{d-1}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \frac{\partial^2 u_{km}}{\partial r_{km}^2}, && (43) \end{aligned}$$

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

$$\frac{\nabla_k^2 \Psi_T(\mathbf{r})}{\Psi_T(\mathbf{r})} = \frac{\nabla_k^2 \phi_k}{\phi_k} + 2\frac{\nabla_k \phi_k}{\phi_k} \sum_{m\neq k}^{N} \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \sum_{m\neq k}^{N} \left(\frac{d-1}{r_{km}} \frac{\partial u_{km}}{\partial r_{km}} + \frac{\partial^2 u_{km}}{\partial r_{km}^2} \right)$$
$$+ \sum_{m,n\neq k}^{N} \frac{\mathbf{r}_k - \mathbf{r}_m}{r_{km}} \frac{\mathbf{r}_k - \mathbf{r}_n}{r_{kn}} \frac{\partial u_{km}}{\partial r_{km}} \frac{\partial u_{kn}}{\partial r_{kn}} . \tag{44}$$

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

$$\frac{\partial u_{km}}{\partial r_{km}} = \frac{a}{r_{km}(r_{km} - a)}, \tag{45}$$

$$\frac{\partial^2 u_{km}}{\partial r_{km}^2} = \frac{a^2 - 2ar_{km}}{r_{km}^2 (r_{km} - a)^2}. \tag{46}$$

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

V. ALGORITHMS

In the 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

$$q(\mathbf{r}_{i+1}, \mathbf{r}_i) = \frac{\left|\Psi_T(\mathbf{r}_{i+1})\right|^2}{\left|\Psi_T(\mathbf{r}_i)\right|^2}, \tag{47}$$

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,

$$\frac{\partial \Psi_T}{\partial t} = D\nabla \cdot (\nabla - \mathbf{F})\, \Psi_T, \tag{48}$$

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 11. 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.

$$\frac{\partial \mathbf{r}}{\partial t} = D\mathbf{F}(\mathbf{r}) + \boldsymbol{\eta}, \tag{49}$$

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.

$$\mathbf{r}_{i+1} = \mathbf{r}_i + D\mathbf{F}(\mathbf{r}_i)\Delta t + \boldsymbol{\xi}\sqrt{\Delta t}, \tag{50}$$

for a given time step Δt [1] 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.

$$q(\mathbf{r}_{i+1}, \mathbf{r}_i) = \frac{G(\mathbf{r}_{i+1}, \mathbf{r}_i, \Delta t)\left|\Psi_T(\mathbf{r}_{i+1})\right|^2}{G(\mathbf{r}_i, \mathbf{r}_{i+1}, \Delta t)\left|\Psi_T(\mathbf{r}_i)\right|^2}, \tag{51}$$

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

$$G(\mathbf{r}_{i+1}, \mathbf{r}_i, \Delta t) = \exp\left(-\frac{\left[\mathbf{r}_{i+1} - \mathbf{r}_i - D\mathbf{F}(\mathbf{r}_i)\Delta t\right]^2}{4D\Delta t}\right) \times \frac{1}{(4\pi D\Delta t)^{dN/2}}, \tag{52}$$

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,

$$\sigma = \sqrt{\frac{1}{N}(\langle E_L^2\rangle - \langle E_L\rangle^2}, \tag{53}$$

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 53 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 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,

$$f_d = \frac{1}{n-d}\sum_{k=1}^{n-d}(x_k - \bar{x}_n)(x_{k+d} - \bar{x}_n). \tag{54}$$

The keen reader would have noticed that the function f_d in Equation 54 would be equal to the sample variance for $d = 0$. We can now define the *autocorrelation function*

$$\kappa_d = \frac{f_d}{\mathrm{Var}(x)}, \tag{55}$$

which is equal to 1 if the data exhibits no autocorrealtions, equating to $d = 0$. From the autocorrelation function (55) we in turn define the *autocorrelation time*,

$$\tau = 1 + 2\sum_{d=1}^{n-1}, \tag{56}$$

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 53 by taking correlation into account,

$$\sigma = \sqrt{\frac{1 + 2\tau/\Delta t}{N}(\langle E_L^2\rangle - \langle E_L\rangle^2)}, \tag{57}$$

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.

---

[1] Bear in mind that Equation 49 is only valid as $\Delta t \to 0$, a property stemming from the use of Euler's method.