FYS4411 - Computational quantum mechanics

Spring 2016

Project 2;
Variational Monte Carlo studies of electronic systems

Github repository:

https://github.com/filiphl/FYS4411.git

Sean Bruce Sangolt Miller
s.b.s.miller@fys.uio.no

Filip Henrik Larsen
filiphenriklarsen@gmail.com

Roar Emaus
roarem@fys.uio.no

Date: May 30, 2016

Abstract

Contents

1 Introduction

Quantum dots are, possibly large, atomic systems packed so closely together that the Pauli principle forces the electrons to be in different states. These systems can be made approximately 2-dimensional[1] by strings of quantum dots. Quantum dots are much larger than atoms while possessing many of their characteristics, which means they are both easy and interesting to measure. Measurements provide means with which to test quantum theory and models. Up to 20 electrons, it is reasonable to approximate the system Hamiltonian by placing all electrons in a harmonic oscillators and having coulomb interactions. If only closed-shell quantum dots are considered, a model that splits the wavefunction can be used, which greatly reduces the number of calculations. The purpose of this project is see if this simple model can reproduce known results. This will be done through variational Monte-Carlo (VMC) by using a Slater determinant with a Pade-Jastrow factor

First, a detailed calculation of the two-body quantum dot will be performed, followed by deduction of algorithms for a general many-body numerical approach to a system of $N = \{2,\ 6,\ 12,\ 20\}$ electrons. Then, an explanation of the optimization of the variational parameters will given, followed by a small talk on the "blocking" method. Since the program script will be large, a few benchmarks will be mentioned. Lastly, results will be presented followed by some comments.

2 Theory and Methods

2.1 Preliminary derivations

While performing VMC it is of course favourable to use analytical expressions, should they not demand an unacceptable increase in CPU time. We will therefore need to calculate the local energy $E_L = \frac{1}{\Psi_T} H\Psi_T$ and the quantum force $F = \frac{2}{\Psi_T}\nabla\Psi_T$. The Hamiltonian H used will be:

$$H = H_0 + H_I = \sum_{i=1}^{N}\left(-\frac{1}{2}\nabla_i^2 + \frac{1}{2}\omega^2 r_i^2\right) + \sum_{i<j}\frac{1}{r_{ij}} \tag{1}$$

I.e. a harmonic oscillator potential with Coulomb interactions. The Laplacian will be the most demanding quantity to calculate.

2.2 Singlet electron state

For an electron in a harmonic oscillator potential, the energy is given by $\epsilon_n = \omega(n+1)$, where we have used natural units and $n = n_x + n_y + \ldots$. For two non-interacting electrons, the energy is $\epsilon_{n_1,n_2} = \omega(n_{x,1} + n_{y,1} + \ldots + n_{x,2} + n_{y,2} + \ldots + 2)$. Obviously the energy is lowest for $n_1 = n_2 = 0$, giving $\epsilon_{0,0} = 2\omega$.

Since $n_1 = n_2 = 0$ means the two electron are in the same spatial wavefunction, they must have different spins. Since electrons are spin-$\frac{1}{2}$ particles, they combine to give total spin zero, i.e. they form the singlet state.

For the singlet electron state we will use the trial wavefunction:

$$\Psi_T(\mathbf{r}_1, \mathbf{r}_2) = Ce^{-\frac{\alpha\omega}{2}(r_1^2 + r_2^2)} e^{\frac{ar_{12}}{1+\beta r_{12}}} \quad , \quad a = 1 \tag{2}$$

The Laplacian of which (for particle i) is:

$$\nabla_i^2 \Psi_T = \nabla_i(\nabla_i \Psi_T) \tag{3}$$

We will use the following change of coordinates when it simplifies calculations:

$$\begin{aligned} \frac{\partial}{\partial r_{i,j}} &= \frac{\partial r_{12}}{\partial r_{i,j}} \frac{\partial}{\partial r_{12}} \\ \rightarrow \nabla_i &= \frac{(-1)^i}{r_{12}}(x_1 - x_2, y_1 - y_2)\frac{\partial}{\partial r_{12}} \\ &= \frac{(-1)^i}{r_{12}}\mathbf{r}_{12}\frac{\partial}{\partial r_{12}} \end{aligned} \tag{4}$$

where $r_{i,j}$ is element j of $\mathbf{r}_i$. The gradient, which is needed for the quantum force as well, is then

[1] I.e. all electron movement is confined to a plane. This does not mean the space is considered two-dimensional, so the 3-dimensional Coulomb potential will still be used.

$$
\begin{aligned}
\nabla_i \Psi_T &= -\alpha\omega \mathbf{r}_i \Psi_T + \frac{(-1)^i}{r_{12}} \mathbf{r}_{12} \left[\frac{\partial}{\partial r_{12}} \left(\frac{a r_{12}}{1+\beta r_{12}} \right) \right] \Psi_T \\
&= \left[-\alpha\omega \mathbf{r}_i + \frac{(-1)^i}{r_{12}} \mathbf{r}_{12} \frac{a}{(1+\beta r_{12})^2} \right] \Psi_T
\end{aligned} \tag{5}
$$

which means the Laplacian is

$$
\begin{aligned}
\nabla_i^2 \Psi_T &= \left[\nabla_i [\ldots] \right] \Psi_T + [\ldots] \nabla_i \Psi_T \\
&= \left[\nabla_i [\ldots] \right] \Psi_T + [\ldots]^2 \Psi_T
\end{aligned} \tag{6}
$$

where $[\ldots]$ is the last parenthesis in equation 5. The parenthesis in the first term above is

$$
\begin{aligned}
\nabla_i \left[-\alpha\omega \mathbf{r}_i + \frac{(-1)^i}{r_{12}} \mathbf{r}_{12} \frac{a}{(1+\beta r_{12})^2} \right] &= -2\alpha\omega + \frac{(-1)^i}{r_{12}} \left(\frac{(-1)^i 2 a r_{12}}{(1+\beta r_{12})^2} - \frac{(-1)^i 2 a \beta r_{12}}{(1+\beta r_{12})^3} - \frac{(-1)^i a}{r_{12}(1+\beta r_{12})^2} \right) \\
&= -2\alpha\omega - \frac{a}{(1+\beta r_{12})^2} \left(\frac{1}{r_{12}} - \frac{2}{r_{12}} + \frac{2\beta}{1+\beta r_{12}} \right) \\
&= -2\alpha\omega + \frac{a}{r_{12}(1+\beta r_{12})^2} - \frac{2a\beta}{(1+\beta r_{12})^3}
\end{aligned} \tag{7}
$$

which gives

$$
\nabla_i^2 \Psi_T = \left[-2\alpha\omega + \frac{a}{r_{12}(1+\beta r_{12})^2} - \frac{2a\beta}{(1+\beta r_{12})^3} + \alpha^2\omega^2 r_i^2 + \frac{a^2}{(1+\beta r_{12})^4} - \frac{2\alpha\omega a (-1)^i}{r_{12}(1+\beta r_{12})^2} \mathbf{r}_i \cdot \mathbf{r}_{12} \right] \Psi_T \tag{8}
$$

We therefore have:

$$
\sum_{i=1}^{2} \frac{1}{\Psi_T} \nabla_i^2 \Psi_T = -4\alpha\omega + \frac{2a}{r_{12}(1+\beta r_{12})^2} - \frac{4a\beta}{(1+\beta r_{12})^3} + \alpha^2\omega^2 (r_1^2 + r_2^2) + \frac{2a^2}{(1+\beta r_{12})^4} - \frac{2\alpha\omega a}{(1+\beta r_{12})^2} r_{12} \tag{9}
$$

2.3 Closed, 2-dimensional shell states

If we again consider the non-interacting, harmonic oscillator confined electron system, we can increase the number of electrons beyond 2. If, in some wondrous universe, we had fermions with three different spins states (like a spin-1 boson), then for three electrons we could again set $n_1 = n_2 = n_3 = 0$ and have an anti-symmetric spin state. However, in our world, we must go up in energy for more than two electrons.
The $n = 0$ state was the ground state. The next state has degeneracy 2; $(n_1, n_2) = (1, 0), (0, 1)$. After that we have degeneracy 3;$(n_1, n_2) = (2, 0), (1, 1), (0, 2)$. Each of these states are doubly degenerate due to spin.

The reason in explaining this is because, assuming we have a so-called "closed shell" problem[2], we can do some manipulations that greatly reduce the number of calculations necessary to perform VMC. Firstly, we need to rewrite the trial wavefunction.
As is already known, the true wavefunction is approximated by an analytical solution to some simpler problem, and a Jastrow factor. The analytical part can be written as a Slater determinant:

$$
\Psi_D = \frac{1}{\sqrt{N}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \ldots & \phi_{N-1}(\mathbf{r}_1) & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \ldots & \phi_{N-1}(\mathbf{r}_2) & \phi_N(\mathbf{r}_2) \\ & & \vdots & & \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \ldots & \phi_{N-1}(\mathbf{r}_N) & \phi_N(\mathbf{r}_N) \end{vmatrix} \tag{10}
$$

and therefore our trial wavefunction will be:

$$
\Psi_T = \Psi_D \Psi_C \ , \ \ \Psi_C = \prod_{i<j}^{N} e^{f_{ij}} \quad , \quad f_{ij} \equiv \frac{a_{ij} r_{ij}}{1+\beta r_{ij}} \tag{11}
$$

[2]For every new tier in energy, we fill it up with electrons. So all considered tiers, or "shells", are full (or "closed")

where α, β are the variational parameters, a_{ij} is connected to particle spins, and N is the total number of particles. The single particle functions are solutions to the two dimensional, harmonic oscillator Scrödinger equation:

$$\phi_i(\mathbf{r_i}) = CH_{n_{x,i}}(\sqrt{\omega\alpha}x_i)H_{n_{y,i}}(\sqrt{\omega\alpha}y_i)e^{\frac{\omega\alpha}{2}r_i^2} \tag{12}$$

Obviously, Ψ_D is a time consuming object to calculate at every Metropolis step. We will therefore do some neat tricks that reduce the number of calculations.
The first is to rewrite Ψ_D by using that the Hamiltonian is spin-independent. If we now let ϕ_n be the spatial-component of the wavefunction, then it causes no effect on the energy if we write:

$$\Psi_T = \Psi_{D+}\Psi_{D-}\Psi_C \tag{13}$$

where $\Psi_{D+} \equiv |D_+|$ and $\Psi_{D-} \equiv |D_-|$ are Slater determinants consisting only of spin-up and spin-down particles, respectively. This means $|D_+|$ and $|D_-|$ are $\frac{N}{2}$-dimensional, and we see this method only works for closed shell systems. We will let the first $\frac{N}{2}$ particles be spin-up and the rest spin-down. That is, $\mathbf{r_1} - \mathbf{r}_{N/2}$ are the positions of the spin-up particles, while the rest are positions for spin-down particles. Obviously this is wrong, since we can't separate which particles are spin-up and which are spin-down in reality, due to the properties of identical particles. However, as stated, due to the spin-independent Hamiltonian, this causes no effect on the energy, which is all we're after.

2.4 The Metropolis ratio test

The basic principle behind the Metropolis algorithm is to make an assumption on the transition probability for a system to move from setting to another, as an exact, or even approximate, expression is lacking.

If the probability distribution for a state i is given by w_i, then from Markov chain theory the time derivative is:

$$\frac{\partial w_i(t)}{\partial t} = \sum_j W(j \to i)w_j(t) - W(i \to j)w_i(t) \tag{14}$$

where $W_{i\to j}$ is the probability of moving from a state i to another state j, i.e the rate of change in w_i is given by the probability for a state j to go to i minus the probability of state i going to j, summed over all j. The most likely state will fulfil $\frac{\partial w_i(t)}{\partial t} = 0$, giving:

$$\begin{aligned} W(j \to i)w_j(t) &= W(i \to j)w_i(t) \\ \Rightarrow \quad \frac{W(j \to i)}{W(i \to j)} &= \frac{w_i}{w_j} \end{aligned} \tag{15}$$

Since the transition probability W is unknown, we approximate it by guessing its form:

$$W(j \to i) = T(j \to i)A(j \to i) \tag{16}$$

where T is the transition moving probability, while A is the probability of accepting such a move. Furthermore, in brute force Metropolis, one guesses $T_{i\to j} = T_{j\to i}$. Therefore:

$$\frac{A_{j\to i}}{A_{i\to j}} = \frac{w_i}{w_j} \tag{17}$$

Since the probability densities are known, it is known whether or not this ratio is larger than one. If it's larger than one, then the acceptance probability from j to i is the biggest, i.e we are more likely to accept the move than not. Therefore, we simply say the move is accepted. However, if the probability ratio is smaller than one, then we are more likely to move from the new state to the one the system is currently in. To check whether or not we should accept the move, we may compare it to, say, a "coin toss"; if it's bigger, the move is accepted.

So, at each Metropolis step, we need the ratio of probabilities. We first define $R \equiv \frac{\Psi_T^n}{\Psi_T^o}$, where "$n$" means the new wavefunction and "o" means the old (or the current, but "c" could be confused with "correlation"). Written out, this is:

$$R = \frac{|D_+^n|}{|D_+^o|}\frac{|D_-^n|}{|D_-^o|}\frac{\Psi_C^n}{\Psi_C^o} \tag{18}$$

If we only move one position at a time, then only one row in either D_+ or D_- will change. This means if we move a spin-up position, then $|D_-^n| = |D_-^o|$, so we need only consider one of the determinant fractions for each R.
Through some simple steps, one can show the determinant fraction (R_D) reduces to:

$$R_D = \sum_{j=1}^{N/2} D_{ij}(\mathbf{r}^n) D_{ji}^{-1}(\mathbf{r}^o) \tag{19}$$

where D_{ij}^{-1} is element[3] ij of the inverse of D, D is either the spin-up or spin-down determinant, and $\mathbf{r}_i$ is the moved position. Since $D_{ij}(\mathbf{r}) = \phi_j(\mathbf{r}_i)$, the only difficulty remaining is to find the elements of the inverse matrix.

The elements of an inverse matrix are given by the Sherman-Morrison formula, which, when applied to the current case, gives:

$$D_{kj}^{-1}(\mathbf{r}^n) = \begin{cases} D_{kj}^{-1}(\mathbf{r}^o) - \frac{D_{ki}^{-1}(\mathbf{r^o})}{R_D} \sum_{l=1}^{N/2} D_{il}(\mathbf{r}^n) D_{lj}^{-1}(\mathbf{r}^o) & \text{if } j \neq i \\ \\ \frac{D_{ki}^{-1}(\mathbf{r^o})}{R_D} & \text{if } j = i \end{cases} \tag{20}$$

The ratio for the correlation function has a rather nice expression:

$$\begin{aligned} \frac{\Psi_C^n}{\Psi_C^o} &= \prod_{i<j}^{N/2} e^{(f_{ij}^n - f_{ij}^o)} \\ &= \exp\left(\sum_{i<j}^{N/2} f_{ij}^n - f_{ij}^o \right) \\ &= \exp\left(\sum_{i=0}^{k-1} (f_{ik}^n - f_{ik}^o) + \sum_{j=k+1}^{N/2} (f_{kj}^n - f_{kj}^o) \right) \end{aligned} \tag{21}$$

were we used $f_{ij}^n - f_{ij}^o = 0 \,\forall\, i,j \neq k$, and k is the moved position. The first sum is then for $j = k$ and the second sum is for $i = k$, with the restriction $i < j$.
The only step that remains is to square and multiply the two ratios[4].

2.5 Importance sampling

In importance sampling, each suggested move requires the calculation of the quantum force:

$$x_{new} = x_{old} + DF(x_{old})\Delta t + \xi\sqrt{\Delta t} \tag{22}$$

Solutions to the Fokker-Planck equation gives the transition probability, which must be multiplied with the probability density. The transition probability is therefore given by:

$$G(y, x, \Delta t) = \frac{1}{(4\pi D\Delta t)^{3N/2}} \exp\left\{ -\frac{(y - x - D\Delta t F(x))^2}{4D\Delta t} \right\} \tag{23}$$

where y is the new position and x the old, and the acceptance test becomes:

$$q(y, x) = \frac{G(x, y, \Delta t)|\Psi_T(y)|^2}{G(y, x, \Delta t)|\Psi_T(x)|^2} \tag{24}$$

The quantum force, given by $F = 2\frac{\nabla\Psi_T}{\Psi_T}$, requires the gradient of Ψ_T. Obviously, the quantum force can be written[5]:

$$F = 2\left(\frac{\nabla|D_+|}{|D_+|} + \frac{\nabla|D_-|}{|D_-|} + \frac{\nabla\Psi_C}{\Psi_C} \right) \tag{25}$$

Of course, when only position k is altered, only the k'th gradient in ∇ changes (recall the definition $\nabla \equiv (\nabla_1, \nabla_2, \ldots, \nabla_N)$, and needs to be re-evaluated. We know the determinant can written:

$$|D| = \sum_{j=1}^{N/2} D_{kj} C_{jk} \tag{26}$$

[3] Where i is the row and j is the column.
[4] Or multiply and then square. Really, it's up to you.
[5] The superscript "o" has been dropped since there will be no "mix" of new and old coordinates for the rest of this subsection.

where C_{jk} are the cofactors of D, and is independent of the i'th row in D, i.e. changing k does not change C_{jk}. It is therefore independent of the position change. Changing row i means all the other gradients in $\nabla|D|$ are the same as before, and we only need to re-evaluate $\nabla_k|D|$. This means:

$$\begin{aligned} \frac{\nabla_k|D|}{|D|} &= \frac{\nabla_k \sum_{j=1}^{N/2} D_{kj}C_{jk}}{|D|} \\ &= \sum_{j=1}^{N/2} \frac{(\nabla_k D_{kj})C_{jk}}{|D|} \\ &= \sum_{j=1}^{N/2} (\nabla_k D_{kj}) D^{-1}_{jk} \end{aligned} \tag{27}$$

which, from equation 20, means:

$$\frac{\nabla_k|D^n}{|D^n|} = \frac{1}{R_D} \sum_{j=1}^{N/2} (\nabla_k D^n_{kj})(D^o_{jk})^{-1} \tag{28}$$

where the factor $\nabla_k D_{kj} = \nabla_k \phi_j(\mathbf{r}_k)$ is:

$$\begin{aligned} \nabla_k D_{kj} = A\Big(& H'_{n_{j,x}}(\sqrt{\omega\alpha}x_k)H_{n_{j,y}}(\sqrt{\omega\alpha}y_k) - \alpha\omega x_k H_{n_{j,x}}(\sqrt{\omega\alpha}x_k)H_{n_{j,y}}(\sqrt{\omega\alpha}y_k) \\ & ,H'_{n_{j,y}}(\sqrt{\omega\alpha}y_k)H_{n_{j,x}}(\sqrt{\omega\alpha}x_k) - \alpha\omega y_k H_{n_{j,x}}(\sqrt{\omega\alpha}x_k)H_{n_{j,y}}(\sqrt{\omega\alpha}y_k)\Big) e^{-\frac{\alpha\omega}{2} r_k^2} \end{aligned} \tag{29}$$

The correlation function gradient can be expressed:

$$\begin{aligned} \frac{\nabla_k \Psi_C}{\Psi_C} &= \frac{1}{\Psi_C} \nabla_k e^{\sum_{i<j}^{N} f_{ij}} \\ &= \sum_{i=1}^{k-1} \nabla_k f_{ik} + \sum_{j=k+1}^{N} \nabla_k f_{kj} \end{aligned} \tag{30}$$

but since f_{ij} only depends on r_{ij}, it would preferable to express ∇_k in terms of r_{ij}. In equation 4, we showed this change for a simpler system. Applied to this problem, we can derive:

$$\begin{aligned} \nabla_k &= \frac{1}{r_{ik}} \mathbf{r}_{ik} \frac{\partial}{\partial r_{ik}} \\ &\text{or} \\ \nabla_k &= -\frac{1}{r_{jk}} \mathbf{r}_{jk} \frac{\partial}{\partial r_{jk}} \end{aligned} \tag{31}$$

which gives:

$$\frac{\nabla_k \Psi_C}{\Psi_C} = \sum_{i=1}^{k-1} \frac{\mathbf{r}_{ik}}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} - \sum_{j=k+1}^{N} \frac{\mathbf{r}_{kj}}{r_{kj}} \frac{\partial f_{kj}}{\partial r_{kj}} \tag{32}$$

We now have all the necessary tools to perform importance sampling.

2.6 Local energy

Lastly, the local energy needs to be calculated. As usual, the Laplacian fraction $\frac{\nabla^2 \Psi_T}{\Psi_T}$ is the most demanding object to calculate. the starting point is:

$$\frac{\nabla^2 \Psi_T}{\Psi_T} = \frac{\nabla^2|D_+|}{|D_+|} + \frac{\nabla^2|D_-|}{|D_-|} + \frac{\nabla^2 \Psi_C}{\Psi_C} + 2\left(\frac{\nabla|D_+|}{|D_+|} + \frac{\nabla|D_-|}{|D_-|}\right) \cdot \frac{\nabla \Psi_C}{\Psi_C} \tag{33}$$

which looks easy enough. The last term contain vectors already known, while the first two are derived in the exact same manner as for the gradients, i.e.

$$\frac{\nabla_k^2|D|}{|D|} = \sum_{j=1}^{N/2} (\nabla_k^2 D_{kj}) D_{jk}^{-1} \tag{34}$$

where (here we abbreviate $H_{n_{j,z}}(\sqrt{\omega\alpha} z_k)$ by $H_{n_j}(z)$ for compactness)

$$\begin{aligned} \nabla_k^2 D_{kj} = A\Big[& H''_{n_j}(x) H_{n_j}(y) + H''_{n_j}(y) H_{n_j}(x) \\ & - 2\alpha\omega \left(x H'_{n_j}(x) H_{n_j}(y) + y H'_{n_j}(y) H_{n_j}(x) \right) \\ & - \alpha\omega H_{n_j}(x) H_{n_j}(y) \left(d - \alpha\omega r_k^2 \right) \Big] e^{-\frac{\alpha\omega}{2} r_k^2} \end{aligned} \tag{35}$$

Unfortunately, the middle term is not so nice to find. The steps needed are many and tedious, but not difficult. They are therefore omitted and we show only the final result:

$$\frac{\nabla_k^2 \Psi_C}{\Psi_C} = \left(\frac{\nabla_k \Psi_C}{\Psi_C} \right)^2 + \sum_{i=1}^{k-1} \left[\frac{d-1}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} + \frac{\partial^2 f_{ik}}{\partial r_{ik}^2} \right] + \sum_{j=k+1}^{N} \left[\frac{d-1}{r_{kj}} \frac{\partial f_{kj}}{\partial r_{kj}} + \frac{\partial^2 f_{kj}}{\partial r_{kj}^2} \right] \tag{36}$$

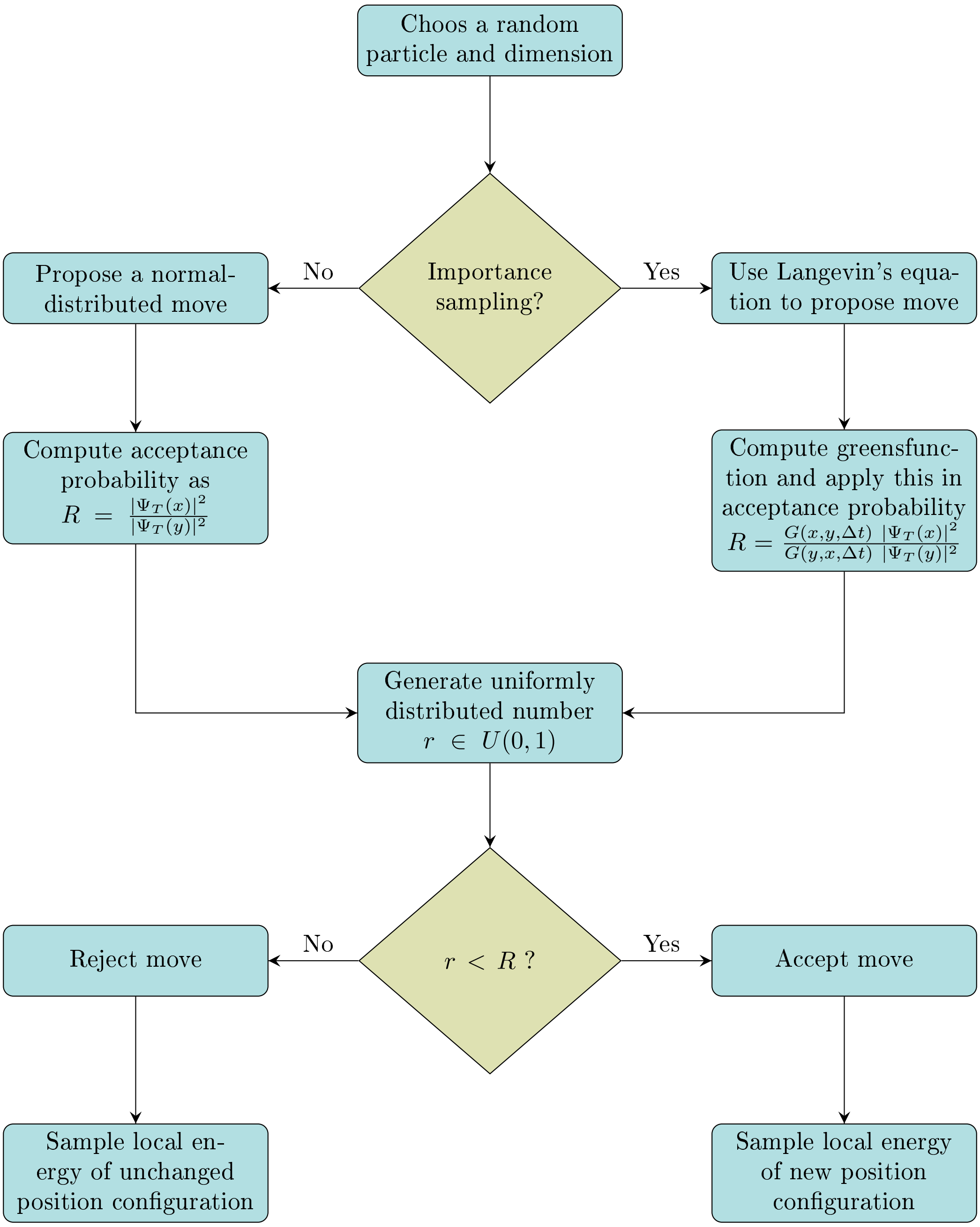
where d is the dimension we consider and:

$$\frac{\partial f_{ij}}{\partial r_{ij}} = \frac{a_{ij}}{(1+\beta r_{ij})^2} \tag{37}$$

$$\frac{\partial^2 f_{ij}}{\partial r_{ij}^2} = -\frac{2 a_{ij} \beta}{(1+\beta r_{ij})^3} \tag{38}$$

where a_{ij} equals 1 for anti-parallel spins and $\frac{1}{3}$ for parallel.

2.7 Flowchart of metropolis algorithm



2.8 Optimizing parameters

Our trial wavefunctions, both for the two-body system and many-body system (equations 2 and 11), contain two variational parameters, α and β. In the VMC approach, the minimal energy is sought. Therefore, the goal is to minimize $\langle E_L \rangle$ with respect to these variational parameters. There are several ways to minimize a value with respect to some parameters, and here we will use the method of steepest descent (SD). The SD method, in algorithm form, is:

$$\vec{x}_{n+1} = \vec{x}_n - \gamma_n \nabla f \tag{39}$$

where $\vec{x}$ is a vector containing the variables for which one wishes to find the minimum of f, γ is a steplength and n is an index expressing the number of iterations. In application to the current problem, $f = \langle E_L \rangle$, and $\vec{x} = (\alpha, \beta)$. The above equation can thus be rewritten as

$$(\alpha, \beta)_{n+1} = (\alpha, \beta)_n - \gamma_n (\frac{\partial}{\partial \alpha}, \frac{\partial}{\partial \beta}) \langle E_L \rangle. \tag{40}$$

However, since $\langle E_L \rangle$ is quite a time-consuming quantity we find numerically, and its derivatives ($\bar{E}_\alpha \equiv \frac{d\langle E_L\rangle}{d\alpha}$ and $\bar{E}_\beta \equiv \frac{d\langle E_L\rangle}{d\beta}$) even more so, an "analytical" expression is desirable. This can be found as follows:

$$
\begin{aligned}
\bar{E}_\alpha &= \frac{d}{d\alpha}\int dx P(x) E_L \\
&= \frac{d}{d\alpha}\int dx \frac{|\psi|^2}{\int dx'|\psi|^2}\frac{1}{\psi}H\psi \\
&= \frac{d}{d\alpha}\int dx \frac{\psi^* H\psi}{\int dx'|\psi|^2}
\end{aligned}
\tag{41}
$$

Since the Hamiltonian is hermitian, one has $\int dx\psi^* H\psi = \int dx H\psi^*\psi$, giving:

$$
\begin{aligned}
\bar{E}_\alpha &= \frac{d}{d\alpha}\int dx \frac{H\psi^*\psi}{\int dx'|\psi|^2} \\
&= \left[\int dx \frac{H\left(\psi^*\left(\frac{d\psi}{d\alpha}\right)+\left(\frac{d\psi^*}{d\alpha}\right)\psi\right)}{\int dx'|\psi|^2}\right] - \left[\int dx \frac{H\psi^*\psi}{\left(\int dx'|\psi|^2\right)^2}\int dx'\left(\psi^*\left(\frac{d\psi}{d\alpha}\right)+\left(\frac{d\psi^*}{d\alpha}\psi\right)\right)\right]
\end{aligned}
\tag{42}
$$

Again one may use the hermiticity of the Hamiltonian to get $\int dx H\psi^*\left(\frac{d\psi}{d\alpha}\right) = \int dx H\left(\frac{d\psi^*}{d\alpha}\right)\psi$. So:

$$
\begin{aligned}
\bar{E}_\alpha &= 2\left[\int dx \frac{H\psi^*\frac{d\psi}{d\alpha}}{\int dx'|\psi|^2}\right] - 2\left[\int dx \frac{H\psi^*\psi}{\left(\int dx'|\psi|^2\right)^2}\int dx'\psi^*\frac{d\psi}{d\alpha}\right] \\
&= 2\left[\int dx \frac{H\psi^*\frac{d\psi}{d\alpha}}{\int dx'|\psi|^2} - \int dx \frac{H\psi^*\psi}{\int dx'|\psi|^2}\int dx'\frac{1}{\int dx'|\psi|^2}\psi^*\frac{d\psi}{d\alpha}\right] \\
&= 2\left[\int dx \frac{\psi^*\left(\frac{E_L}{\psi}\frac{d\psi}{d\alpha}\right)\psi}{\int dx'|\psi|^2} - \int dx \frac{\psi^* E_L\psi}{\int dx'|\psi|^2}\int dx'\frac{\psi^*\left(\frac{1}{\psi}\frac{d\psi}{d\alpha}\right)\psi}{\int dx'|\psi|^2}\right] \\
\therefore \bar{E}_\alpha &= 2\left(\langle\frac{\bar{\psi}_\alpha}{\psi}E_L\rangle - \langle\frac{\bar{\psi}_\alpha}{\psi}\rangle\langle E_L\rangle\right)
\end{aligned}
\tag{43}
$$

where $\bar{\psi}_\alpha \equiv \frac{d\psi}{d\alpha}$. Obviously, the derivative with respect to β is the same.

In order to find the optimal parameters, the ones that give minimal energy, we set a criteria that if the new parameters, $\vec{x}_{n+1}$, suggested by equation equation 40 has a shorter gradient than the current, $\vec{x}_n$, we accept this change, whereas if it doesn't we reset the parameters to the current and shorten the steplength by a factor 0.7. We are content when the length of the gradient is sufficiently short and store these values as the optimal parameters.

The optimal parameters produced are not trivial however. As illustrated in figure 1 the minimum value of the local energy depends much more strongly on the value of α than the value of β in the two-electron system. Our approach will most likely produce a value of α that is in agreement with the true minimum, but our value of β may differ somewhat. However, this difference does not seem to have much of a consequence for this system. For the case of the two-electron system the optimal parameters found were $\alpha = 1.003$ and $\beta = 0.3$, which resulted in a local energy of $\langle E_L\rangle = 3.003$. This is sufficiently close to the real minimum of 3, and seems to be in agreement with the figure as well.

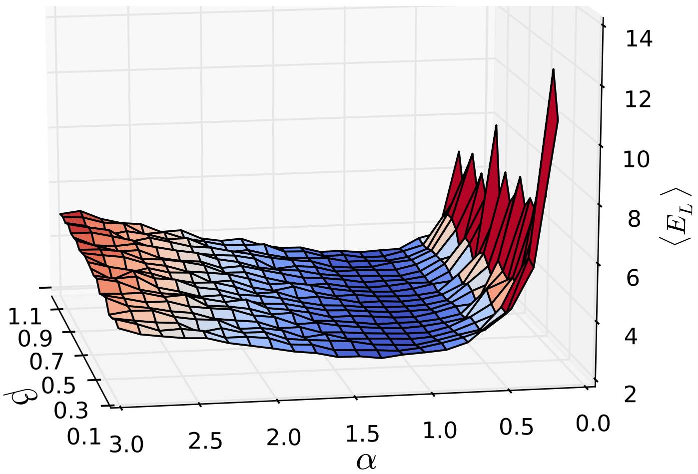


Figure 1: Local energy as a function of the parameters α and β in the two-electron system.

2.9 Blocking method

After having run a Monte Carlo simulation, the variation in estimated local energies are calculated as

$$\sigma = \sqrt{\frac{1}{n}\left(\langle E_L^2\rangle - \langle E_L\rangle^2\right)}. \tag{44}$$

However, these values will be much too low. This is because one assumes all data to be completely uncorrelated. Each energy is calculated by a small perturbation to the system setting[6], which means each new setting is very dependant on the previous setting. After sufficiently many perturbations, though, the system at step i will be so different from that of step j, that $\langle E_L\rangle_i$ is basically uncorrelated to $\langle E_L\rangle_j$, but not for all $\langle E_L\rangle_k$ between i and j. Ideally, one would like to find a *correlation time* τ such that i and j will be uncorrelated if a time greater than τ has passed. If Δt is the time between two Metropolis steps, then one would like to find $|i-j|$ in $\tau = |i-j|\Delta t$.
A method of dealing with this is the blocking technique. The set of $\langle E_L\rangle$ measurements is grouped into blocks, each of which will give an average of average local energies. Afterwards, one can calculate the variance of these averages of averages. If then the standard deviation (equation above) is plotted as a function of the number of blocks[7], one can find the lowest number of blocks where the curve approaches a plateau. Then one can calculate τ and find the correlation time. The true standard deviation is then:

$$\sigma = \sqrt{\frac{1+2\tau/\Delta t}{n}\left(\langle \mathbf{M}^2\rangle - \langle \mathbf{M}\rangle^2\right)} \tag{45}$$

2.10 Benchmarking

While the flowchart seems nice and compact, there is quite a bit code that needs implementing and some benchmarks would be nice. The first is so check if the program reproduces the energy expected for non-interacting electrons in a harmonic oscillator potential. This means that if we remove the Coulomb potential and set $a_{ij} = 0$, then the problem is simply several non-interacting harmonic oscillators and we are left with a simple harmonic oscillator with energies given as

$$E_{n_x,n_y} = \hbar\omega\left(n_x + n_y + 1\right). \tag{46}$$

In figure 2 these energies levels are shown. If sums of these values are reproduced, it means our Slater determinant expressions are correctly implemented.

[6] In the program discussed here, only a single dimension of a single particle will be moved each step.
[7] Inversely proportional to the block size by block size = $\frac{\text{number of samples}}{\text{number of blocks}}$

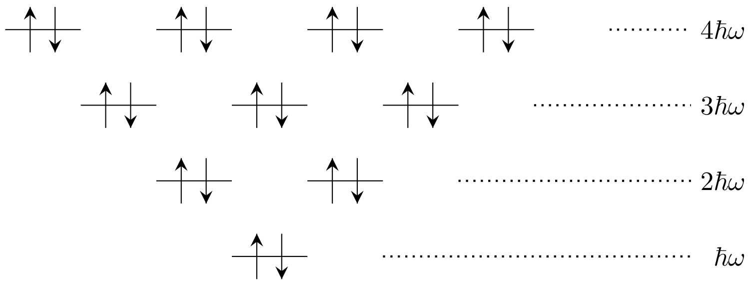


Figure 2: Illustrative figure of spin configurations in the closed shell system.

Another benchmark is to reproduce the results of our program for the two body quantum dot. This system was solved analytically, and thus would show that the many body program can reproduce analytical results for 2 particles. This means that the full problem (Coulomb interactions and Jastrow correlations) works, at least for the two body system.

In order to know if our two-body script works, we can compare with the work of Taut[8], where he calculated the analytical solution of two electrons in an external oscillator potential with Coulomb interactions. He found:

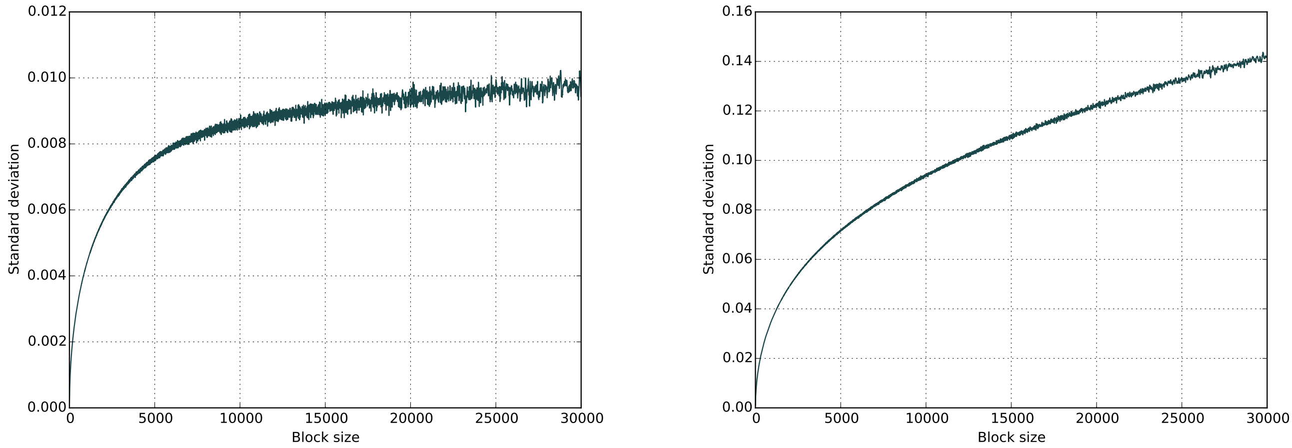
3 Results

Table 1: Expectation values of local energy, kinetic energy and potential energy for $N = \{2, 6, 12, 20\}$ and oscillator frequencies $\omega = \{1.0, 0.5, 0.1, 0.05, 0.01\}$. σ is the standard deviation of $\langle E_L \rangle$, found by blocking. α, β are the optimal parameters.

N	ω	$\langle E_L \rangle$	$\langle E_K \rangle$	$\langle E_P \rangle$	σ	α	β
2	1.00	3.0031	0.9017	2.1013	1.53360e-6	1.09975	0.29472
	0.50	1.6611	0.4559	1.2052	2.10014e-7	1.00015	0.29972
	0.10	0.4415	0.0964	0.3451	5.09917e-5	0.99534	0.16965
	0.05	0.2722	0.0480	0.2242	1.04670e-6	0.99591	0.22695
	0.01	0.0873	0.0107	0.0766	4.48831e-4	0.99777	0.19004
6	1.00	20.204	3.8059	16.398	1.43517e-3	1.00127	0.46939
	0.50	11.821	1.7797	10.042	1.60940e-3	0.97611	0.32474
	0.10	3.7466	0.2878	3.3077	2.22848e-3	0.99293	0.32392
	0.05	2.4792	0.2165	2.2627	3.35813e-3	0.96891	0.37989
	0.01	0.8972	0.0472	0.8500	1.84693e-3	1.00751	0.21825
12	1.00	65.776	8.7763	57.000	1.46457e-2	0.80173	0.80030
	0.50	39.223	4.0024	35.221	1.03759e-2	0.80003	0.50039
	0.10	14.683	1.1936	13.489	3.82364e-2	1.00071	0.69019
	0.05	7.6569	0.4908	7.1661	6.69810e-3	0.94721	0.15057
	0.01	3.2558	0.0937	3.1621	2.22688e-2	0.70992	0.34371
20	1.00	157.48	20.430	137.05	2.35715e-2	0.92930	0.80390
	0.50	93.981	7.9372	86.044	1.42877e-2	0.82140	0.47966
	0.10	30.157	1.1977	28.959	1.63797e-2	0.53094	0.39662
	0.05	18.667	0.5123	18.155	1.45351e-2	0.43081	0.31274
	0.01	6.2090	0.0821	6.1269	6.18577e-3	0.37267	0.12807

For the standard deviations (σ), the blocking plots were easily readable for $N = 2$ and $N = 6$, but the "plateau" for higher particle numbers were a bit hard to read. In figure 3, there are two blocking plots presented. As can be seen, for $N = 6$ the plateau starts at about 10000-15000, while for $N = 20$ it is not completely clear. However, both were deemed to lie in the region 10000-15000 for these specific cases.

[8]M. Taut, Phys. Rev. A 48, 3561 - 3566 (1993)



(a) $N = 6$, $\omega = 0.5$

(b) $N = 20$, $\omega = 0.5$

Figure 3: Blocking for $N = 6$ and $N = 20$.

Table 2: Local energy computed for systems without the Jastrow factor and Coulomb potential. This resembles a pure harmonic oscillator and the results are dead on.

N	E_L	σ
2	2	0
6	10	0
12	28	0
20	60	0

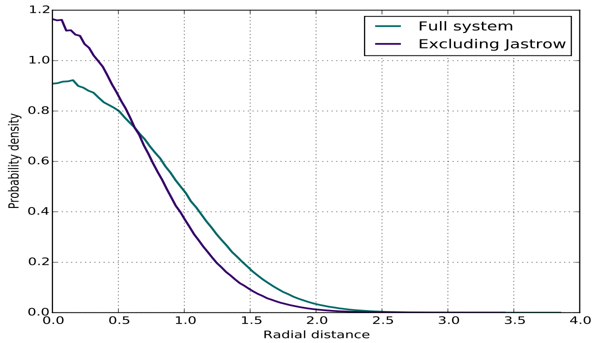
This matches the sums of the energy levels shown in figure 2, and infer that our Slater determinant expressions are implemented correctly.

In table 3, the energies for the two body system are presented, and was calculated using the analytical expressions from section 2.2. As is known, the energy for $\omega = 1.0$ and $N = 2$ is exactly 3, and the table confirms this to quite a high accuracy.
The average distance between the electrons for $\omega = 1.0$ is found to be $\langle r_{12} \rangle = 1.8$.

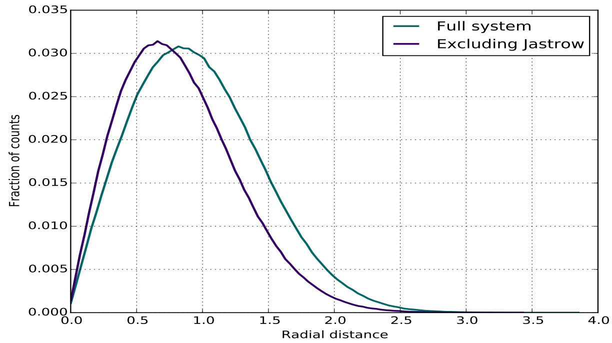
Table 3: Expectation value of the kinetic energy and potential energy for several values of ω.

ω	$\langle E_K \rangle$	$\langle E_P \rangle$	$\langle E_L \rangle$
0.01	0.0122	0.0969	0.0863
0.05	0.0511	0.2274	0.2769
0.1	0.0939	0.3544	0.4593
1.0	0.8983	2.1475	3.0002

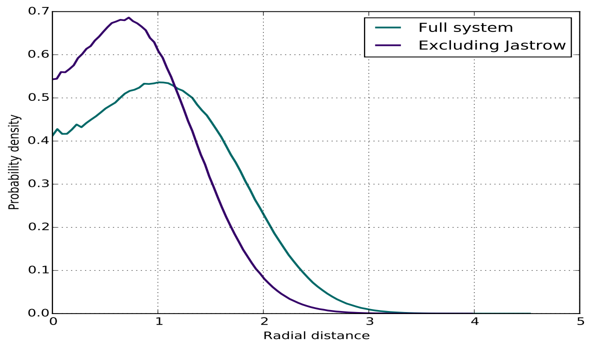
Note, however, that each of the values listed in table 3 are from individual runs. The expectation values of the kinetic, potential and total (local) energy are thus not from the same simulation, and their value might not add up perfectly as $E_L = E_K + E_P$. The total energy is only shown to give a sense of its magnitude.



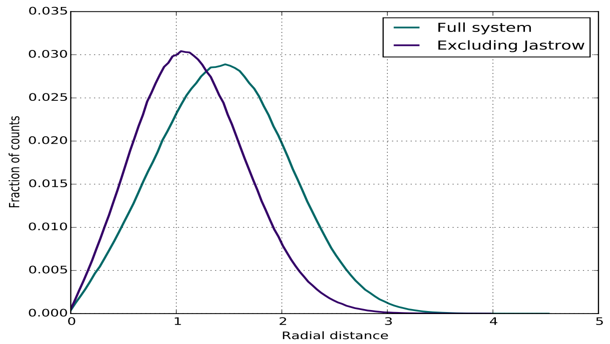
(a) $N = 2$, $\omega = 1.0$



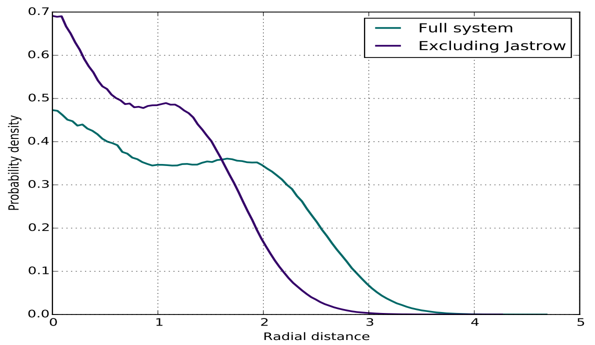
(b) $N = 2$, $\omega = 1.0$



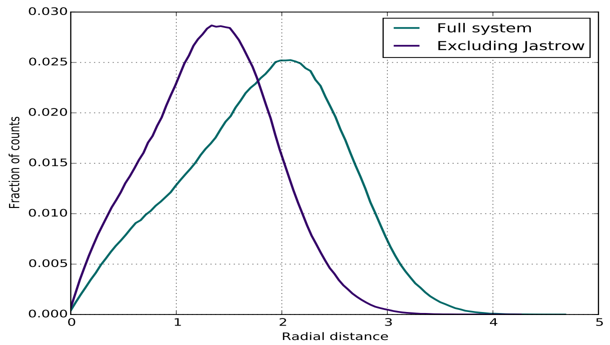
(c) $N = 6$, $\omega = 1.0$



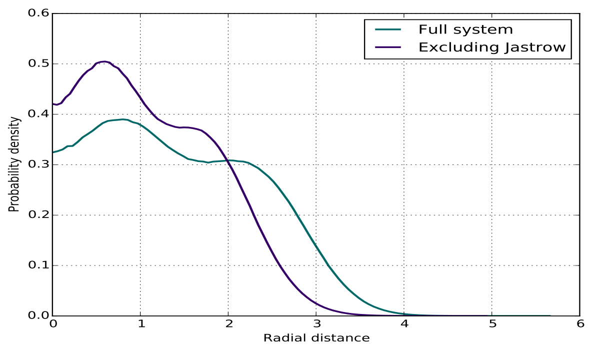
(d) $N = 6$, $\omega = 1.0$



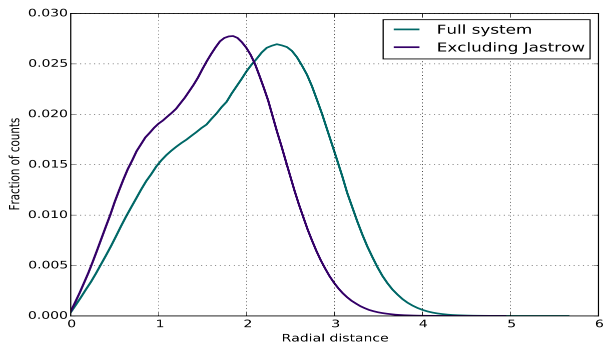
(e) $N = 12$, $\omega = 1.0$



(f) $N = 12$, $\omega = 1.0$



(g) $N = 20$, $\omega = 1.0$



(h) $N = 20$, $\omega = 1.0$

Figure 4: One-body densities (left) and radial distributions (right) for $N = \{2, 6, 12, 20\}$ and $\omega = 1.0$.

4 Comments

How to start?