
Project 2; Variational Monte Carlo studies of electronic systems

Github repository:

<https://github.com/filiph1/FYS4411.git>

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

Filip Henrik Larsen
filiphenriklarsen@gmail.com

Date: May 25, 2016

Contents

1	Introduction	1
2	Theory and Methods	1
2.1	Preliminary derivations	1
2.2	Singlet electron state	1
2.3	Closed, 2-dimensional shell states	2
2.4	The Metropolis ratio test	2
2.5	Importance sampling	3
2.6	Local energy	4
2.7	Flowchart of metropolis algorithm	5
2.8	Benchmarking	6
3	Results	6
4	Comments	7

1 Introduction

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:

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

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 $n = n_x + n_y + \dots$. For two non-interacting electrons, the energy is $\epsilon_{n_1, n_2} = \omega(n_{x,1} + n_{y,1} + \dots + n_{x,2} + n_{y,2} + \dots + 2)$. 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:

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

The Laplacian of which (for particle i) is:

$$\nabla_i^2 \Psi_T = \nabla_i (\nabla_i \Psi_T) \quad (3)$$

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

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

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

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

which means the Laplacian is

$$\begin{aligned} \nabla_i^2 \Psi_T &= [\nabla_i [\dots]] \Psi_T + [\dots] \nabla_i \Psi_T \\ &= [\nabla_i [\dots]] \Psi_T + [\dots]^2 \Psi_T \end{aligned} \quad (6)$$

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

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

which gives

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

We therefore have:

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

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" problem¹, 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:

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

and therefore our trial wavefunction will be:

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

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 Schrödinger equation:

$$\phi_i(\mathbf{r}_i) = C H_{n_{x,i}}(\sqrt{\omega\alpha}x_i) H_{n_{y,i}}(\sqrt{\omega\alpha}y_i) e^{-\frac{\omega\alpha}{2}r_i^2} \quad (12)$$

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. Firstly, we move all spin-up particles to the "first" $\frac{N}{2}$ -positions, and the spin-down particles the remainder of positions.

2.4 The Metropolis ratio test

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:

$$R = \frac{|D_+^n| |D_-^n| \Psi_C^n}{|D_+^o| |D_-^o| \Psi_C^o} \quad (13)$$

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:

¹For every new tier in energy, we fill it up with electrons. So all considered tiers, or "shells", are full (or "closed")

$$R_D = \sum_{j=1}^{N/2} D_{ij}(\mathbf{r}^n) D_{ji}^{-1}(\mathbf{r}^o) \quad (14)$$

where D_{ij}^{-1} is element² 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:

$$D_{kj}^{-1}(\mathbf{r}^n) = \begin{cases} D_{kj}^{-1}(\mathbf{r}^o) - \frac{D_{ki}^{-1}(\mathbf{r}^o)}{R_D} \sum_{l=1}^{N/2} D_{il}(\mathbf{r}^n) D_{lj}^{-1}(\mathbf{r}^o) & \text{if } j \neq i \\ \frac{D_{ki}^{-1}(\mathbf{r}^o)}{R_D} & \text{if } j = i \end{cases} \quad (15)$$

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

$$\begin{aligned} \frac{\Psi_C^n}{\Psi_C^o} &= \prod_{i < j}^{N/2} e^{(f_{ij}^n - f_{ij}^o)} \\ &= \exp \left(\sum_{i < j}^{N/2} f_{ij}^n - f_{ij}^o \right) \\ &= \exp \left(\sum_{i=0}^{k-1} (f_{ik}^n - f_{ik}^o) + \sum_{j=k+1}^{N/2} (f_{kj}^n - f_{kj}^o) \right) \end{aligned} \quad (16)$$

where 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 ratios³.

2.5 Importance sampling

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

$$x_{new} = x_{old} + DF(x_{old})\Delta t + \xi\sqrt{\Delta t} \quad (17)$$

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:

$$G(y, x, \Delta t) = \frac{1}{(4\pi D \Delta t)^{3N/2}} \exp \left\{ -\frac{(y - x - D \Delta t F(x))^2}{4D \Delta t} \right\} \quad (18)$$

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

$$q(y, x) = \frac{G(x, y, \Delta t) |\Psi_T(y)|^2}{G(y, x, \Delta t) |\Psi_T(x)|^2} \quad (19)$$

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 written⁴:

$$F = 2 \left(\frac{\nabla |D_+|}{|D_+|} + \frac{\nabla |D_-|}{|D_-|} + \frac{\nabla \Psi_C}{\Psi_C} \right) \quad (20)$$

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

$$|D| = \sum_{j=1}^{N/2} D_{kj} C_{jk} \quad (21)$$

²Where i is the row and j is the column.

³Or multiply and then square. Really, it's up to you.

⁴The 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:

$$\begin{aligned}\frac{\nabla_k|D|}{|D|} &= \frac{\nabla_k \sum_{j=1}^{N/2} D_{kj} C_{jk}}{|D|} \\ &= \sum_{j=1}^{N/2} \frac{(\nabla_k D_{kj}) C_{jk}}{|D|} \\ &= \sum_{j=1}^{N/2} (\nabla_k D_{kj}) D_{jk}^{-1}\end{aligned}\tag{22}$$

which, from equation 15, means:

$$\frac{\nabla_k|D^n|}{|D^n|} = \frac{1}{R_D} \sum_{j=1}^{N/2} (\nabla_k D_{kj}^n) (D_{jk}^o)^{-1}\tag{23}$$

The correlation function gradient can be expressed:

$$\begin{aligned}\frac{\nabla_k \Psi_C}{\Psi_C} &= \frac{1}{\Psi_C} \nabla_k e^{\sum_{i<j}^N f_{ij}} \\ &= \sum_{i=1}^{k-1} \nabla_k f_{ik} + \sum_{j=k+1}^N \nabla_k f_{kj}\end{aligned}\tag{24}$$

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:

$$\begin{aligned}\nabla_k &= \frac{1}{r_{ik}} \mathbf{r}_{ik} \frac{\partial}{\partial r_{ik}} \\ \text{or} \\ \nabla_k &= -\frac{1}{r_{jk}} \mathbf{r}_{jk} \frac{\partial}{\partial r_{jk}}\end{aligned}\tag{25}$$

which gives:

$$\frac{\nabla_k \Psi_C}{\Psi_C} = \sum_{i=1}^{k-1} \frac{\mathbf{r}_{ik}}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} - \sum_{j=k+1}^N \frac{\mathbf{r}_{kj}}{r_{kj}} \frac{\partial f_{kj}}{\partial r_{kj}}\tag{26}$$

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:

$$\frac{\nabla^2 \Psi_T}{\Psi_T} = \frac{\nabla^2 |D_+|}{|D_+|} + \frac{\nabla^2 |D_-|}{|D_-|} + \frac{\nabla^2 \Psi_C}{\Psi_C} + 2 \left(\frac{\nabla |D_+|}{|D_+|} + \frac{\nabla |D_-|}{|D_-|} \right) \cdot \frac{\nabla \Psi_C}{\Psi_C}\tag{27}$$

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.

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

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:

$$\frac{\nabla_k^2 \Psi_C}{\Psi_C} = \left(\frac{\nabla_k \Psi_C}{\Psi_C} \right)^2 + \sum_{i=1}^{k-1} \left[\frac{d-1}{r_{ik}} \frac{\partial f_{ik}}{\partial r_{ik}} + \frac{\partial^2 f_{ik}}{\partial r_{ik}^2} \right] + \sum_{j=k+1}^N \left[\frac{d-1}{r_{kj}} \frac{\partial f_{kj}}{\partial r_{kj}} + \frac{\partial^2 f_{kj}}{\partial r_{kj}^2} \right] \quad (29)$$

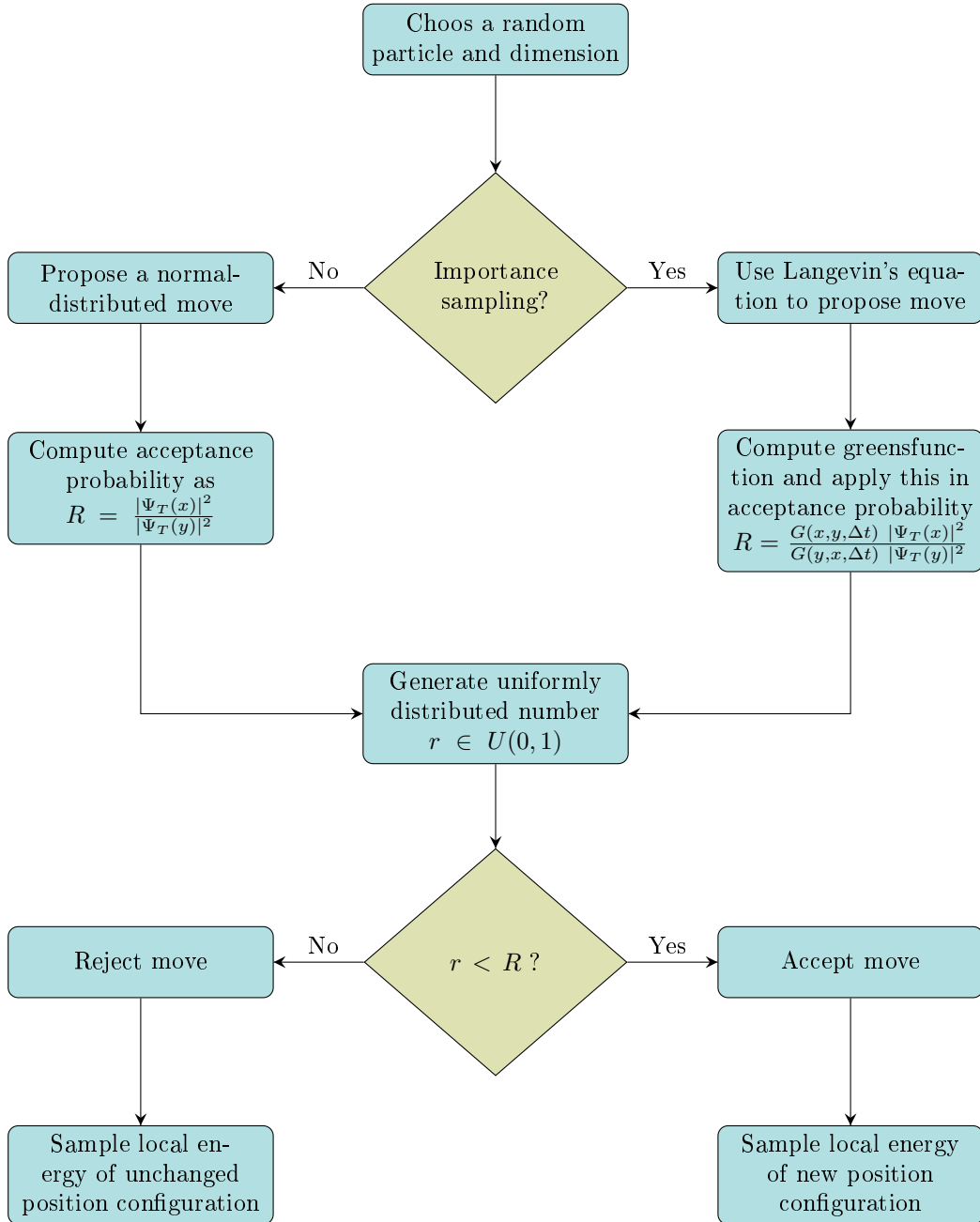
where d is the dimension we consider and:

$$\frac{\partial f_{ij}}{\partial r_{ij}} = \frac{a_{ij}}{(1 + \beta r_{ij})^2} \quad (30)$$

$$\frac{\partial^2 f_{ij}}{\partial r_{ij}^2} = -\frac{2a_{ij}\beta}{(1 + \beta r_{ij})^3} \quad (31)$$

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

2.7 Flowchart of metropolis algorithm



2.8 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 $a_{ij} = 0$, we should get the energies shown in table 1. If these values are reproduced, it means our Slater determinant expressions are correctly implemented.

Table 1: Harmonic oscillator energies

N	E [ω]
2	2
6	10
12	28
20	60

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. Obviously, this means that the full problem (Coulomb interactions and Jastrow correlations) works, at least for the two body system.

3 Results

Without any interaction in the system, no Coulomb potential or Jastrow factor, we are left with a simple harmonic oscillator with energies given as

$$E_{n_x, n_y} = \hbar\omega (n_x + n_y + 1). \quad (32)$$

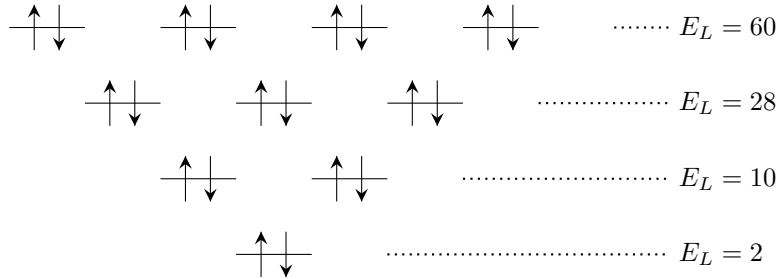


Figure 1: Illustrative figure of spin configurations in the closed shell system.

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

These result infer that our Slater determinants are implemented correctly.

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, though, 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.

4 Comments