M6 Supervision Exercises

Oli Bridge, *St Catharine's College*

January 2023

1 Integrators

1.1 Momentum conservation without Noether

The total linear momentum of an N-particle system is given by

$$\mathbf{P} = \sum_{i=1}^{N} \mathbf{p}_i \tag{1}$$

where $\mathbf{p}_i$ is the momentum of particle i. Taking the time derivative of Equation (1) gives

$$\begin{aligned} \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{N} \mathbf{p}_i\right] \\ &= \sum_{i=1}^{N} \frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} = \sum_{i=1}^{N} \dot{\mathbf{p}}_i \\ &= \sum_{i=1}^{N} m_i \ddot{\mathbf{r}}_i \\ &= -\sum_{i=1}^{N} \frac{\partial U(\mathbf{r}^N)}{\partial \mathbf{r}_i} \end{aligned} \tag{2}$$

where the fourth line uses Newton's second law, $m_i\ddot{\mathbf{r}}_i = \mathbf{f}_i(\mathbf{r}^N)$. The potential $U(\mathbf{r}^N)$ for pairwise interactions is written as a sum over all pairs of particles:

$$U(\mathbf{r}^N) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i} u_{ij}(r_{ij}) \tag{3}$$

The derivative of $U(\mathbf{r}^N)$ with respect to the i^{th} particle position $\mathbf{r}_i$ is given by

$$\begin{aligned} \frac{\partial U(\mathbf{r}^N)}{\partial \mathbf{r}_i} &= \frac{1}{2}\sum_{j\neq i}\left(\frac{\partial u_{ij}(r_{ij})}{\partial \mathbf{r}_i} + \frac{\partial u_{ji}(r_{ji})}{\partial \mathbf{r}_i}\right) \\ &= \sum_{j\neq i} \frac{\partial u_{ij}(r_{ij})}{\partial \mathbf{r}_i} \\ &= \sum_{j\neq i} \mathbf{f}_{ij} \end{aligned} \tag{4}$$

where the second line uses the fact that $r_{ij} = r_{ji}$. The pairwise force $\mathbf{f}_{ij}$ satisfies Newton's third law, $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$. Thus, the time derivative for the total momentum is

$$\begin{aligned} \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} &= \sum_{i=1}^{N} \sum_{j \neq i} \mathbf{f}_{ij} \\ &= \boxed{0} \end{aligned} \tag{5}$$

which sums to zero due to cancellations from the pairwise additive forces. This shows that the total momentum is conserved, given that there are no external forces.

1.2 Leap-Frog algorithm

In the Leap-Frog algorithm, the velocities are given by

$$\mathbf{v}_i(t + \delta t/2) = \frac{\mathbf{r}_i(t + \delta t) - \mathbf{r}_i(t)}{\delta t} \tag{6}$$

$$\mathbf{v}_i(t - \delta t/2) = \frac{\mathbf{r}_i(t) - \mathbf{r}_i(t - \delta t)}{\delta t} \tag{7}$$

and the integration step is given by

$$\mathbf{v}_i(t + \delta t/2) = \mathbf{v}_i(t + \delta t/2) + \frac{\delta t}{m_i}\mathbf{f}_i(t) \tag{8}$$

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \delta t \mathbf{v}_i(t + \delta t/2) \tag{9}$$

Substituting (8) into (9) gives

$$\begin{aligned} \mathbf{r}_i(t + \delta t) &= \mathbf{r}_i + \delta t \left[\mathbf{v}_i(t - \delta t/2) + \frac{\delta t}{m_i}\mathbf{f}_i(t) \right] \\ &= \mathbf{r}_i + \delta t \left[\frac{\mathbf{r}_i(t) - \mathbf{r}_i(t - \delta t)}{\delta t} + \frac{\delta t}{m_i}\mathbf{f}_i(t) \right] \\ &= 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \delta t) + \frac{\delta t^2}{m_i}\mathbf{f}_i(t) \end{aligned} \tag{10}$$

which is exactly equivalent to the verlet algorithm:

$$\boxed{\mathbf{r}_i(t + \delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \delta t) + \frac{\delta t^2}{m_i}\mathbf{f}_i(t)} \tag{11}$$

1.3 Harmonic motion

The autocorrelation function for $A(t)$ is given by

$$C_{AA}(\tau) = \frac{1}{\Delta t - \tau} \int_0^{\Delta t - \tau} \mathrm{d}t \; A(t+\tau)A(t) \tag{12}$$

For $A(t) = \cos \omega t$ and $\Delta t \gg 2\pi/\omega$, the autocorrelation function is evaluated as

$$\begin{aligned}
C_{AA}(\tau) &= \frac{1}{\Delta t - \tau} \int_0^{\Delta t - \tau} \mathrm{d}t \; \cos(\omega(t+\tau)) \cos \omega t \\
&= \frac{1}{\Delta t - \tau} \int_0^{\Delta t - \tau} \mathrm{d}t \; \cos(\omega t + \omega\tau) \cos \omega t \\
&= \frac{1}{\Delta t - \tau} \int_0^{\Delta t - \tau} \mathrm{d}t \; \frac{1}{2}(\cos(2\omega t + \omega\tau) + \cos \omega\tau) \\
&= \frac{1}{\Delta t - \tau} \int_0^{\Delta t - \tau} \mathrm{d}t \; \frac{1}{2}(\underbrace{\cos(2\omega t + \omega\tau)}_{\text{this term integrates to zero on average for large } \Delta t} + \cos \omega\tau) \\
&= \frac{\cos \omega\tau}{2(\Delta t - \tau)} \int_0^{\Delta t - \tau} \mathrm{d}t \\
&= \boxed{\frac{1}{2} \cos \omega\tau}
\end{aligned} \tag{13}$$

For $A(t) = \underbrace{a_1 \cos \omega_1 t}_{B(t)} + \underbrace{a_2 \cos \omega_2 t}_{C(t)}$, the autocorrelation function is

$$\begin{aligned}
C_{AA}(\tau) &= \frac{1}{\Delta t - \tau} \int_0^{\Delta t - \tau} \mathrm{d}t \; (B(t+\tau) + C(t+\tau))(B(t) + C(t)) \\
&= C_{BB}(\tau) + C_{CC}(\tau) + C_{BC}(\tau) + C_{CB}(\tau) \\
&= \frac{a_1^2}{2} \cos \omega_1 t + \frac{a_2^2}{2} \cos \omega_2 t + C_{BC}(\tau) + C_{CB}(\tau)
\end{aligned} \tag{14}$$

The cross-correlation functions $C_{AB}(\tau)$ and $C_{BA}(\tau)$ are of the following form:

$$\begin{aligned} C_{BC}(t) &= \frac{a_1 a_2}{\Delta t - \tau} \int_0^{\Delta t - \tau} \mathrm{d}t \ \cos(\omega_1(t+\tau)) \cos \omega_2 t \\ &= 0 \qquad \text{for } \Delta t \gg 2\pi/\omega \text{ and } \omega_1 \neq \omega_2 \end{aligned} \tag{15}$$

Thus, $C_{AB}(\tau)$ and $C_{BA}(\tau)$ are equal to zero, since sinusoidal waves of differing frequencies are uncorrelated. $C_{AA}(\tau)$ is therefore

$$\boxed{C_{AA}(\tau) = \frac{a_1^2}{2} \cos \omega_1 t + \frac{a_2^2}{2} \cos \omega_2 t} \tag{16}$$

2 Structural descriptors: Radial distribution function

2.1 RDF of an ideal gas

The radial distribution function (RDF) for a system of N particles in a volume V can be calculated as follows:

$$g(r) = \frac{V}{N^2 v_{\text{shell}}(r, \Delta r)} \sum_{i=1}^{N} n_i(r, \Delta r) \tag{17}$$

where $n_i(r, \Delta r)$ is the average number of particles within a shell of radius r, thickness Δr and volume $v_{\text{shell}}(r, \Delta r)$. For the ideal gas system ('ideal.xyz') the following RDF was obtained:

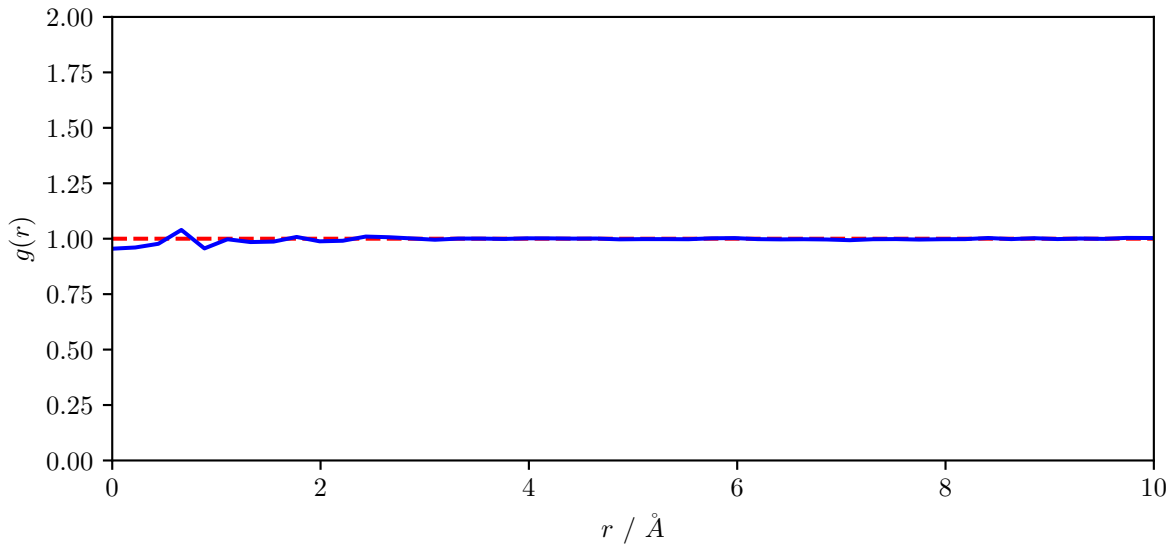


Figure 1: $g(r)$ for ideal gas system, computed from 'ideal.xyz'

As expected for an ideal gas, the RDF is (approximately) equal to 1 for all values of r.

Typically for RDFs, $v_{\text{shell}}(r, \Delta r) = 4\pi r^2 \Delta r$. For my code, I used the more accurate formula $v_{\text{shell}}(r, \Delta r) = 4\pi((r + \Delta r)^3 - r^3)/3$. The only difference to the plots is the behaviour right at $r = 0$, which is fairly unimportant in most cases (see water/colloidal systems).

2.2 RDF of bulk liquid water

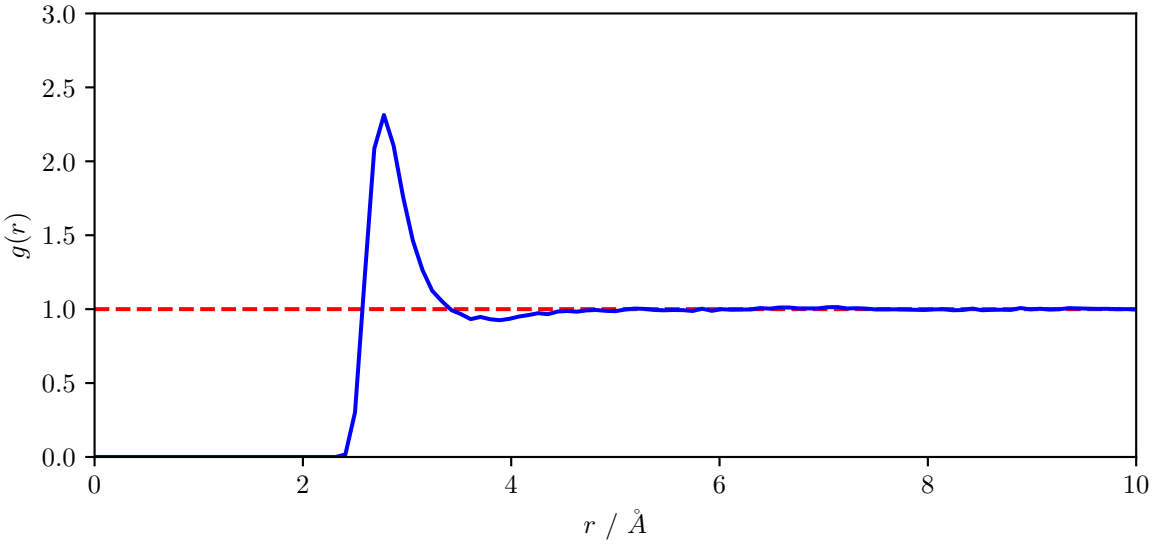


Figure 2: $g_{\text{OO}}(r)$ for water, computed from 'water.xyz'

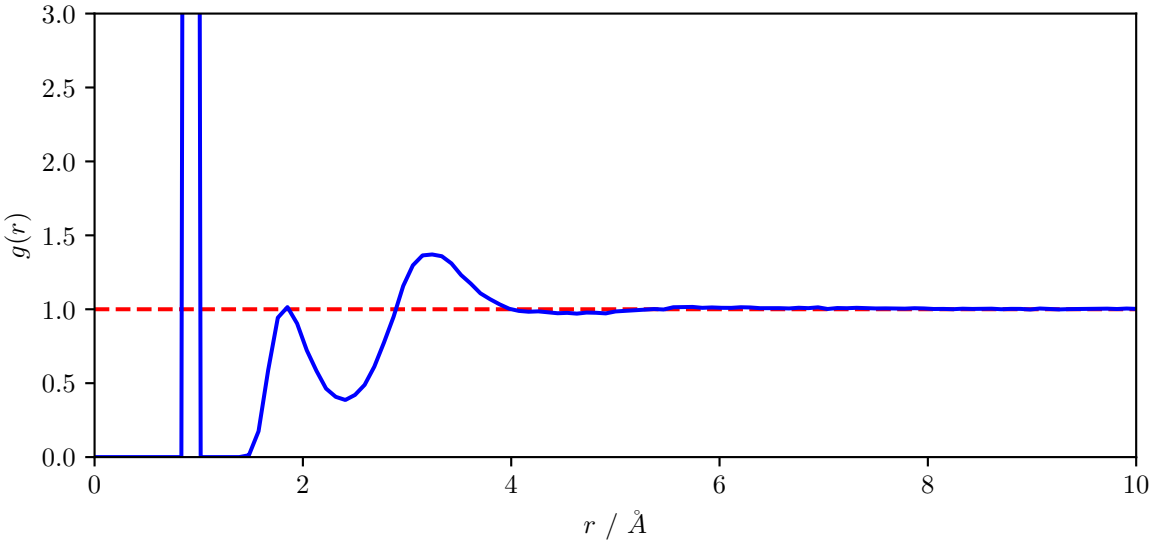


Figure 3: $g_{\text{OH}}(r)$ for water, computed from 'water.xyz'

It is clear that for the oxygen-hydrogen RDF, the sharp peak at around 1Å is due to the two hydrogen atoms on the same molecule as a given oxygen atom.

2.3 Characterising different crystal and liquid phases through their RDFs

Fig. 4 shows a colloidal system in the liquid phase, which is clear due to the RDF's smoother shape. Fig. 5 and 6 show crystal phase colloidal systems.

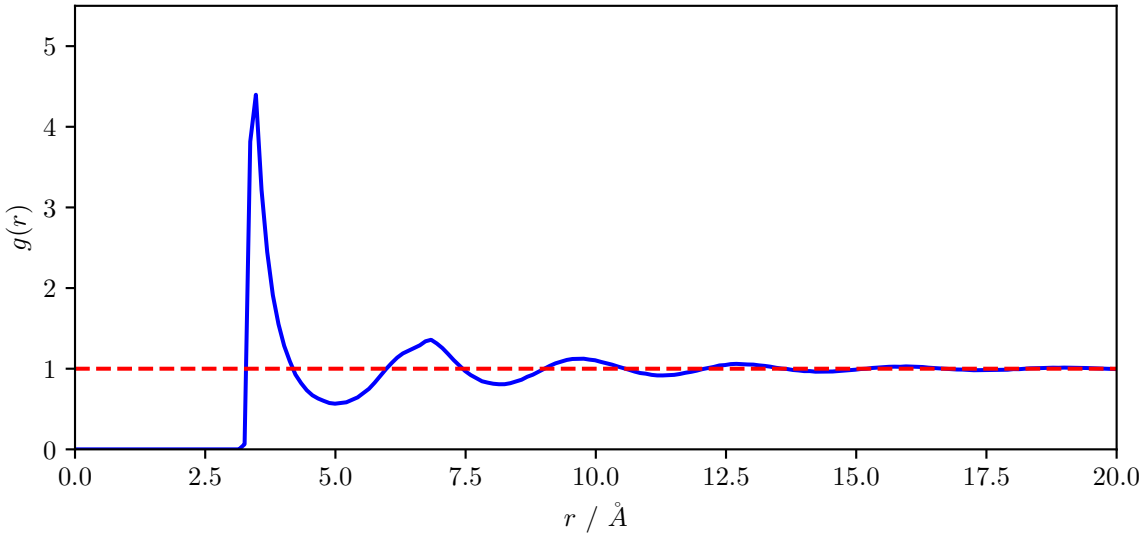


Figure 4: $g(r)$ for a colloidal system, computed from 'set1.xyz'

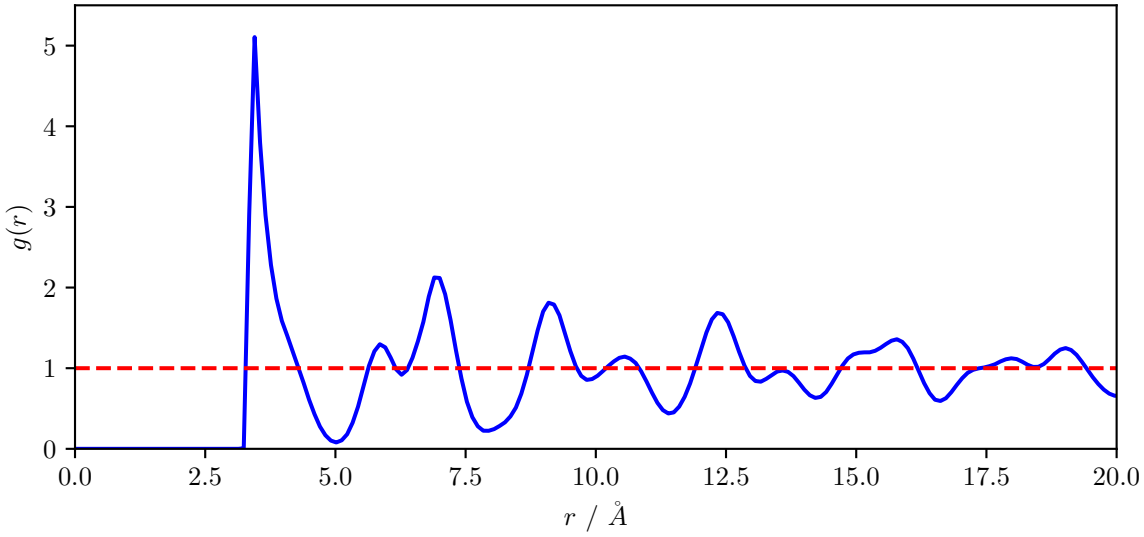


Figure 5: $g(r)$ for a colloidal system, computed from 'set2.xyz'

The coordination number around an atom at a distance r_c can be calculated as

$$z(r_c) = 4\pi\rho \int_0^{r_c} \mathrm{d}r\, r^2 g(r) \tag{18}$$

This can be computed by performing numerical integration on the RDF up to a specific point r_c. Computing the coordination number for Fig. 5 for its first shell ($r_c \approx 5$Å) gives $z(r_1) = 14$.

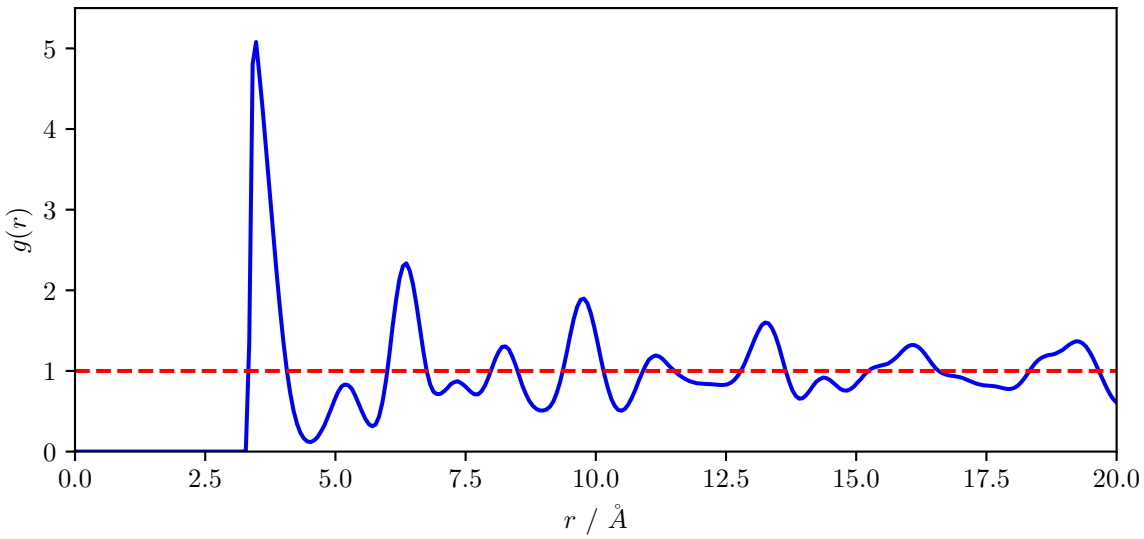


Figure 6: $g(r)$ for a colloidal system, computed from 'set3.xyz'

Computing the coordination number for Fig. 6 for its first and second shells gives $z(r_1) = 12$ and $z(r_2) = 18$ respectively. This suggests that the crystalline colloidal system in Fig. 6 is face-centred cubic (FCC).

3 Potential energy, kinetic energy and total energy of a given configuration

3.1 Potential energy

The three potentials used in this exercise are:

$$E_{\text{LJ}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{19}$$

with $\sigma = 3.405$Å, $\epsilon/k_B = 119.87$K, with a cut-off at $r = 3\sigma$.

$$E_{\text{PHS}} = \begin{cases} \lambda_r\left(\frac{\lambda_r}{\lambda_a}\right)^{\lambda_a}\epsilon_R\left[\left(\frac{\sigma}{r}\right)^{\lambda_r} - \left(\frac{\sigma}{r}\right)^{\lambda_a}\right] + \epsilon_R & r < \left(\frac{\lambda_r}{\lambda_a}\right)\sigma \\ 0 & r \geq \left(\frac{\lambda_r}{\lambda_a}\right)\sigma \end{cases} \tag{20}$$

with $\lambda_a = 49$, $\lambda_r = 50$, $\sigma = 3.405$Å, $\epsilon/k_B = 119.87$K and a cut-off at $r = 1.5\sigma$.

$$E_{\text{HS}} = \begin{cases} \infty & r < \sigma \\ 0 & r \geq \sigma \end{cases} \tag{21}$$

$$E_{\text{Yukawa}} = \pm\epsilon\frac{\sigma}{r}\mathrm{e}^{-\kappa(r-\sigma)} \tag{22}$$

with $\sigma = 3.05\text{Å}$, $\epsilon/k_B = 119.87\text{K}$ and $\kappa = 5\text{Å}^{-1}$. The following potential energies were obtained by looping over every pair of particles and summing the interaction energies:

$$U(\mathbf{r}^N) = \sum_{i=1}^{N}\sum_{j\neq i}^{N} E(r_{ij}) \tag{23}$$

- Lennard-Jones: $U(\mathbf{r}^N)/k_B = -3432512.480910064\text{K}$
- Pseudo-Hard-Sphere: $U(\mathbf{r}^N)/k_B = 210847.14669559113\text{K}$
- Hard sphere Yukawa-Debye-Huckel: $U(\mathbf{r}^N)/k_B = -27584.575122335023\text{K}$

3.2 Kinetic energy

The Maxwell-Boltzmann distribution can be factorised into its cartesian components:

$$\begin{aligned} f(\mathbf{p}) &= \left(\frac{\beta}{2\pi m}\right)^{3/2} \exp\left(-\frac{\beta|\mathbf{p}|^2}{2m}\right) \\ &= \left(\frac{\beta}{2\pi m}\right)^{3/2} \exp\left(-\frac{\beta p_x^2}{2m}\right)\exp\left(-\frac{\beta p_y^2}{2m}\right)\exp\left(-\frac{\beta p_z^2}{2m}\right) \end{aligned} \tag{24}$$

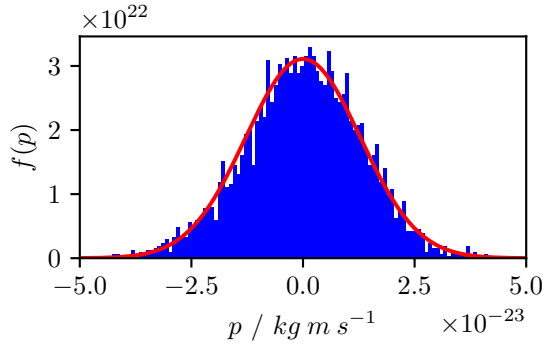
giving individual distributions for the three cartesian coordinates, e.g.

$$f(p_x) = \left(\frac{\beta}{2\pi m}\right)^{1/2} \exp\left(-\frac{\beta p_x^2}{2m}\right) \tag{25}$$

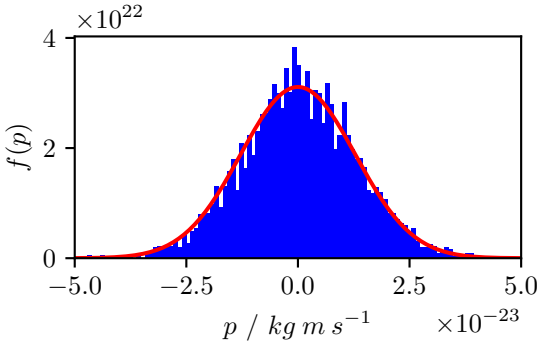
This exactly resembles the probability density function for a normal distribution. Therefore, each component of momentum is normally distributed as such:

$$P_x \sim N(0, \sqrt{m/\beta}) \tag{26}$$

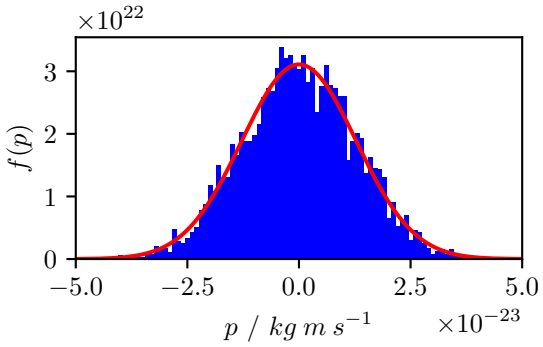
Assigning momenta to each particle in the initial configuration from 'conf.xyz' gave the following histograms for each component (overlayed with $f(p_i)$) with $m = 6.63\times10^{-26}\text{kg}$ and $T = 179.81\text{K}$:



(a) Distribution of p_x



(b) Distribution of p_y



(c) Distribution of p_z

Figure 7: Maxwell-Boltzmann momentum distributions

- Lennard-Jones: $K(\mathbf{p}^N)/k_B = 1089743.0386029673\text{K}$
- Pseudo-Hard-Sphere: $K(\mathbf{p}^N)/k_B = 1050327.8047292393\text{K}$
- Hard sphere Yukawa-Debye-Huckel: $K(\mathbf{p}^N)/k_B = 1069137.2914479263\text{K}$

Since all three potentials do not affect the momentum distribution, it is expected that all three kinetic energies will be similar (not exactly equal since momenta are normally distributed).

3.3 Total energy

The classical Hamiltonian can be written as

$$H(\mathbf{r}^N, \mathbf{p}^N) = U(\mathbf{r}^N) + K(\mathbf{p}^N) \tag{27}$$

The following total energies for the initial configuration in 'conf.xyz' were computed as:

- Lennard-Jones: $H(\mathbf{r}^N, \mathbf{p}^N)/k_B = -2342769.442307097\text{K}$
- Pseudo-Hard-Sphere: $H(\mathbf{r}^N, \mathbf{p}^N)/k_B = 1261174.9514248304\text{K}$
- Hard sphere Yukawa-Debye-Huckel: $H(\mathbf{r}^N, \mathbf{p}^N)/k_B = 1041552.7163255913\text{K}$

The average kinetic energy follows the following relation in the thermodynamic limit for three dimensions:

$$\left\langle \frac{1}{2} m |\mathbf{v}|^2 \right\rangle = \frac{3}{2} k_B T \tag{28}$$

- Lennard-Jones: $T = 181.62\text{K}$
- Pseudo-Hard-Sphere: $T = 175.05\text{K}$
- Hard sphere Yukawa-Debye-Huckel: $T = 178.19\text{K}$

4 Calculation of the isotropic instantaneous pressure

4.1 Pressure through the virial expansion

The virial expression for the pressure of a system of interacting particles is given by

$$P = \frac{1}{3V}\left\langle \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{m} + \mathbf{r}_i \cdot \mathbf{f}_i \right\rangle \tag{29}$$

where $\mathbf{r}_i$ and $\mathbf{p}_i$ are the position and momentum of particle i, and $\mathbf{f}_i$ is the total force on particle i from all other particles. Assuming pairwise additive forces, $\mathbf{f}_i$ can be rewritten as

$$\mathbf{f}_i = \sum_{j \neq i} \mathbf{f}_{ij} \tag{30}$$

The pressure is therefore given by

$$\begin{aligned} P &= \frac{1}{3V}\left\langle \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{m} \right\rangle + \frac{1}{3V}\left\langle \sum_{i=1}^{N} \sum_{j \neq i} \mathbf{r}_i \cdot \mathbf{f}_{ij} \right\rangle \\ &= \frac{1}{3V}\left\langle \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{m} \right\rangle + \frac{1}{6V}\left\langle \sum_{j=1}^{N} \sum_{i \neq j} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \right\rangle \end{aligned} \tag{31}$$

where the second line uses a swap of dummy indices $i \leftrightarrow j$ and Newton's third law ($\mathbf{f}_{ij} = -\mathbf{f}_{ji}$). Now the pressure depends only on the momenta, interparticle distances and pairwise forces. The following plots were obtained for the Lennard-Jones and Pseudo-Hard-Sphere potentials:

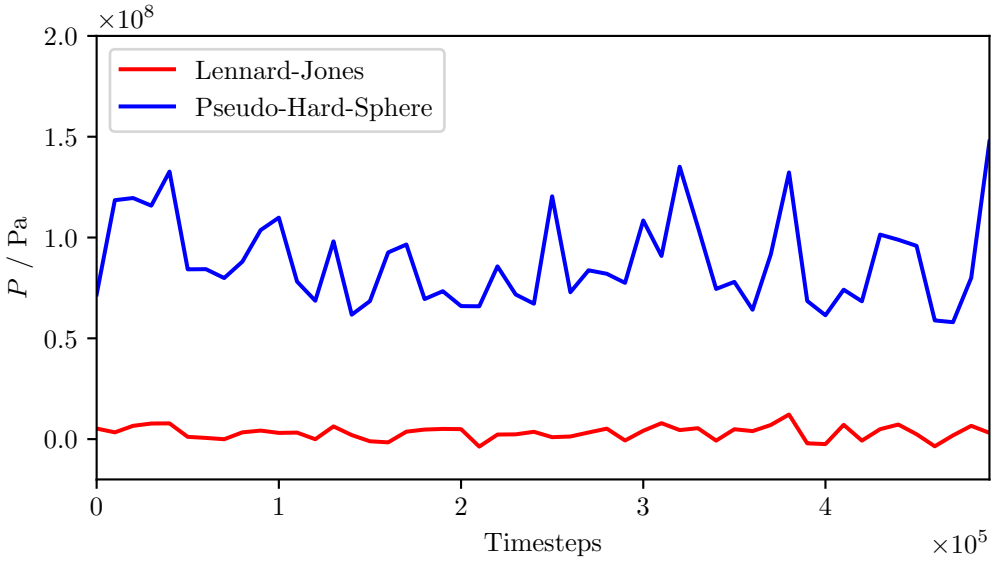


Figure 8: Instantaneous pressures over time using two different potentials, computed using data from 'pres.xyz'

4.2 Time evolution of enthalpy

The enthalpy is given by

$$H = U + PV \tag{32}$$

where U is the internal energy, P is the pressure, and V is the volume of the system. The following plot for the time evolution of the enthalpy was obtained:

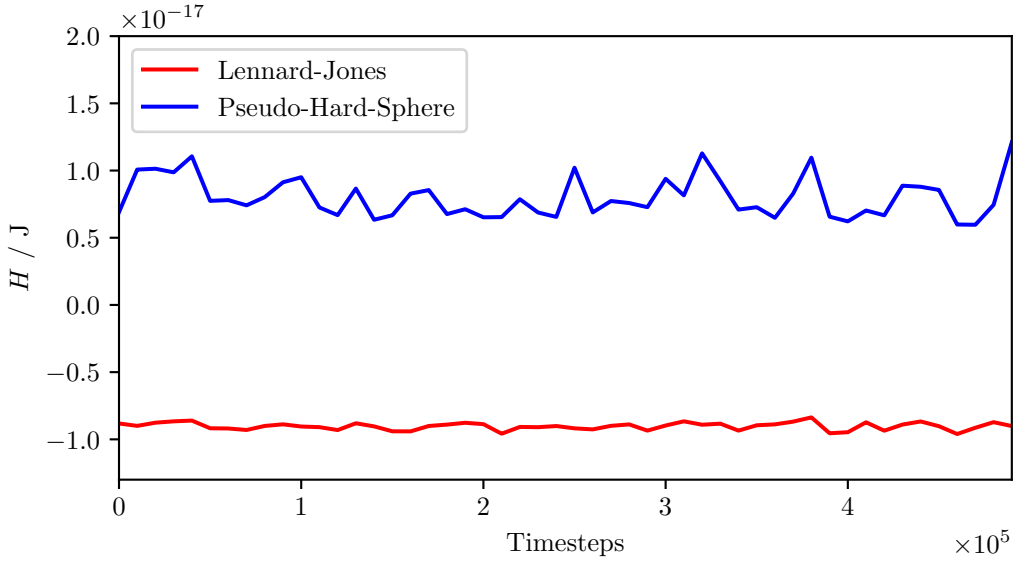


Figure 9: Instantaneous enthalpies over time using two different potentials, computed using data from 'pres.xyz'

4.3 Statistical analysis of pressure and enthalpy

The variance of an observable A can be estimated from the block average method as follows:

$$\sigma^2(A_{t,b}) = \sum_{b=1}^{n_b} \frac{(A_{\tau,b} - \langle A_t \rangle)^2}{n_b} \tag{33}$$

where n_b is the number of blocks, t is the length of the trajectory, τ is the length of a block, and $A_{\tau,b}$ is the average of the observable A within block b. The following results for the pressure were obtained for two different block lengths:

Blocks of 5 samples:

- Lennard-Jones: $P = 3.175 \pm 1.452 \times 10^6$ Pa
- Pseudo-Hard-Sphere: $P = 8.804 \pm 1.094 \times 10^7$ Pa

Blocks of 10 samples:

- Lennard-Jones: $P = 3.175 \pm 1.023 \times 10^6$ Pa

- Pseudo-Hard-Sphere: $P = 8.804 \pm 0.795 \times 10^{7}$ Pa

Here, the error was taken as the standard deviation from the block average method. The following results were obtained for the enthalpy:

Blocks of 5 samples:

- Lennard-Jones: $H = -9.025 \pm 0.115 \times 10^{-18}$ J
- Pseudo-Hard-Sphere: $H = 8.007 \pm 0.744 \times 10^{-18}$ J

Blocks of 10 samples:

- Lennard-Jones: $H = -9.025 \pm 0.083 \times 10^{-18}$ J
- Pseudo-Hard-Sphere: $H = 8.007 \pm 0.543 \times 10^{-18}$ J

5 Dynamic properties: Diffusion coefficient

The mean square displacement of a system of N particles (accounting for centre-of-mass drift) is given by

$$\langle r^2(t) \rangle = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{r}_i(t) - \mathbf{r}_i(0) - (\mathbf{r}_{\text{cm}}(t) - \mathbf{r}_{\text{cm}}(0))|^2 \tag{34}$$

where $\mathbf{r}_i(t)$ is the position of particle i at time t, and $\mathbf{r}_{\text{cm}}(t)$ is the position of the centre of mass of all the particles. The following plots were obtained from two sets of MD trajectories:

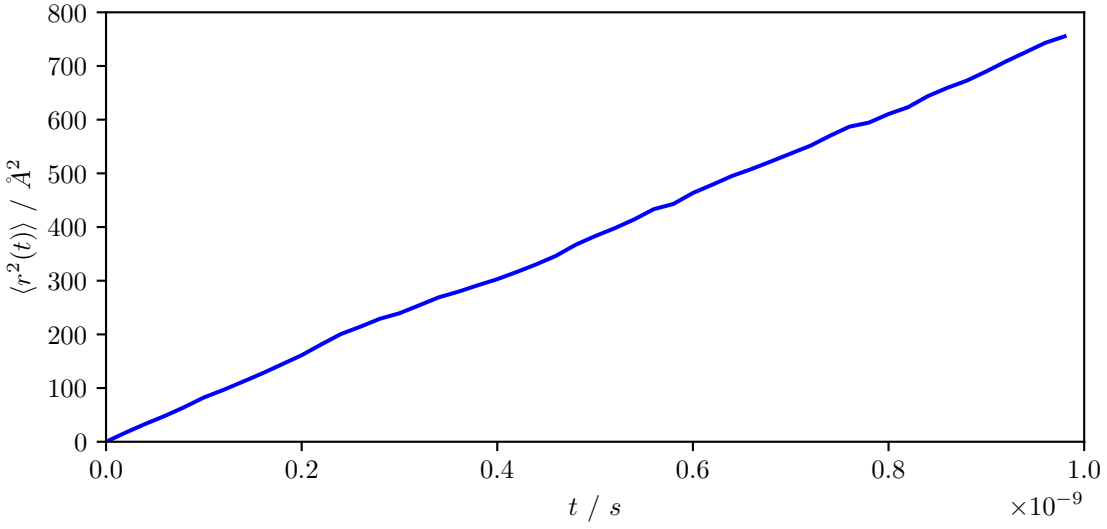


Figure 10: Plot of MSD over time, computed using data from 'diffA.xyz'

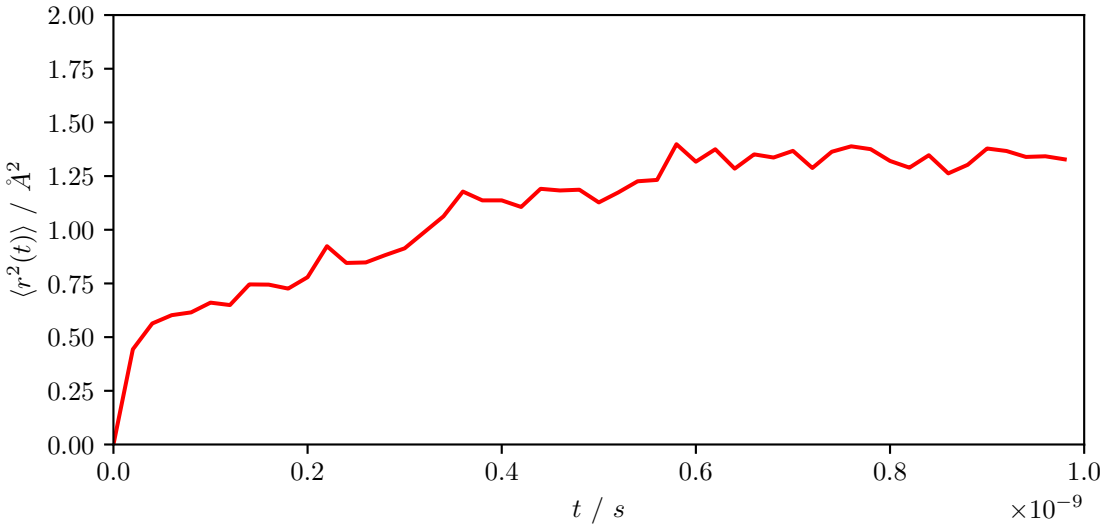


Figure 11: Plot of MSD over time, computed using data from 'diffB.xyz'

The diffusion coefficient, D, obeys the following relation

$$\langle r^2(t)\rangle = 6Dt \tag{35}$$

The computed diffusion coefficients in reduced units of $\sigma^2\tau^{-1}$ ($\sigma = 3.405\mathring{A}$ and $\tau = \sqrt{m\sigma^2/\epsilon}$, with $m = 6.63 \times 10^{-26}$ kg and $\epsilon = 0.24$ kcal/mol) were as follows:

- Set A: $D = 2.38 \times 10^{-2}\ \sigma^2\tau^{-1}$
- Set B: $D = 4.18 \times 10^{-5}\ \sigma^2\tau^{-1}$

Note how set B diffuses much slower than set A, which is evident from Figs 12 and 11.

6 Local order parameters and phase diagram

6.1 Number of contacts through a local order parameter

The average number of polymer contacts for a set of 90 40-bead polymers was computed with the following expression:

$$\langle N_c\rangle = \frac{1}{MN}\sum_{j=1}^{M}\sum_{i=1}^{N} n_i(r < r_c) \tag{36}$$

The following plot was obtained for $\langle N_c\rangle$ as a function of temperature:

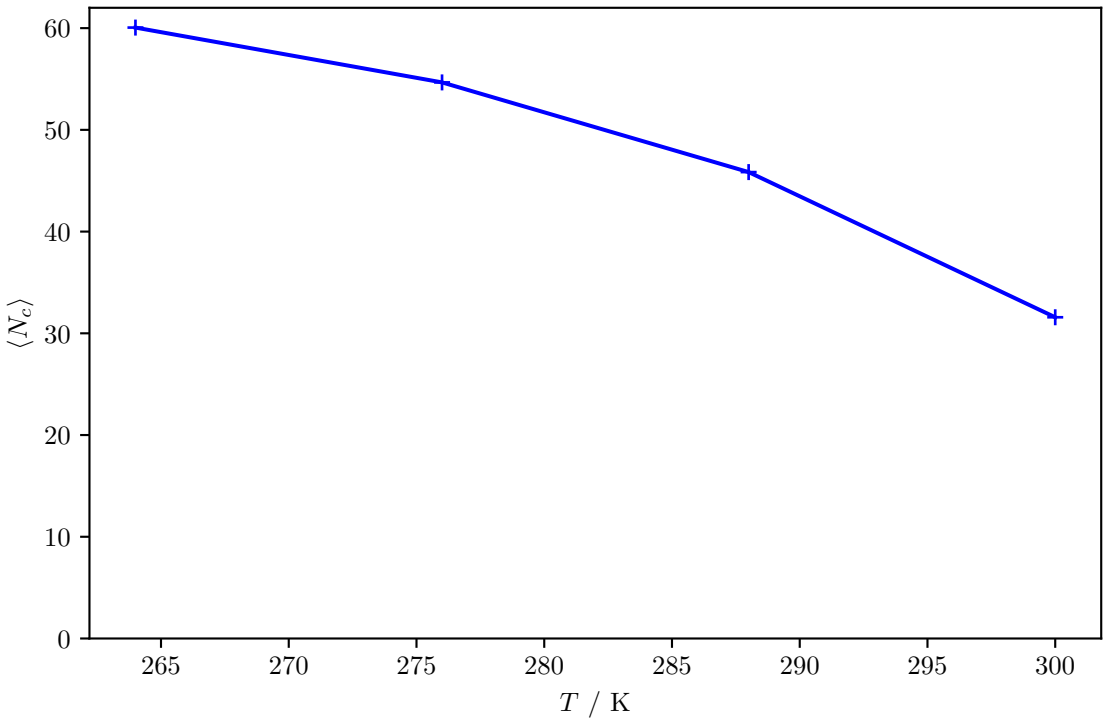


Figure 12: Plot of $\langle N_c \rangle$ vs T from MD trajectories of polymers

6.2 Computing the phase diagram

To determine the phase diagram for this system of polymers, a series of plots were produced for the density along the z axis for four different temperatures. These were used to examine the densities of the two coexistent phases:

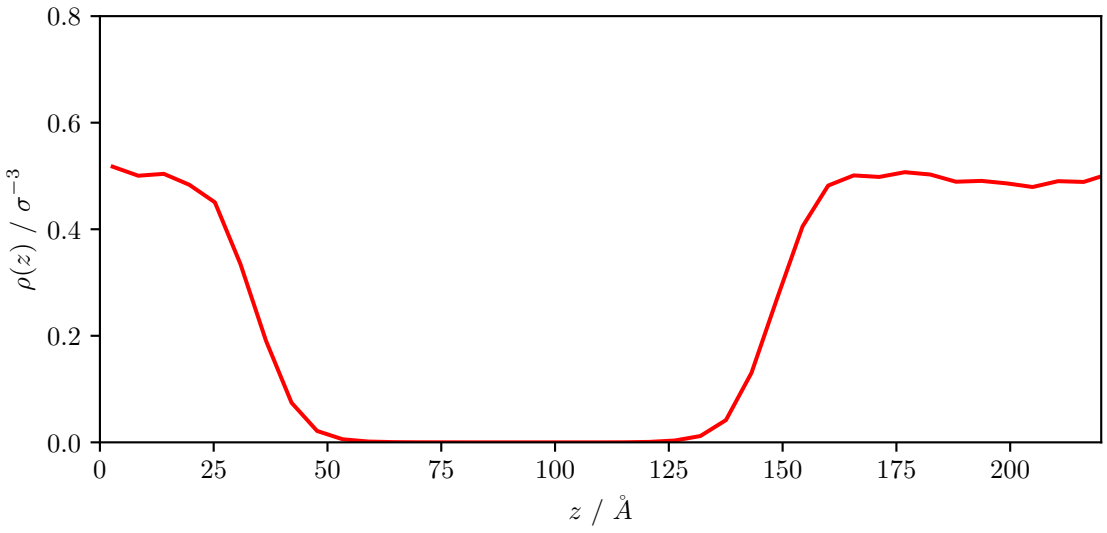


Figure 13: Density plot, computed using data from '264.xyz'

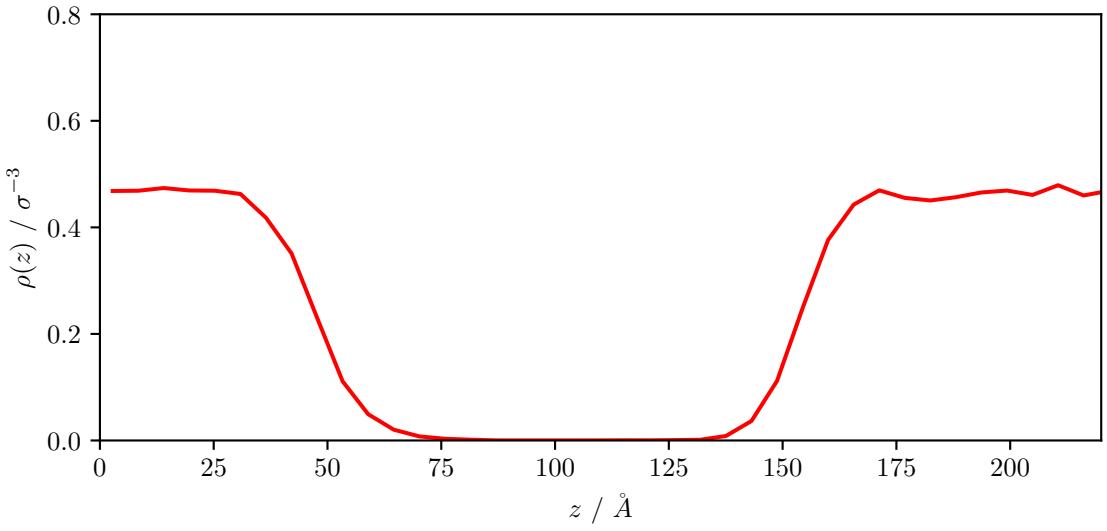


Figure 14: Density plot, computed using data from '276.xyz'

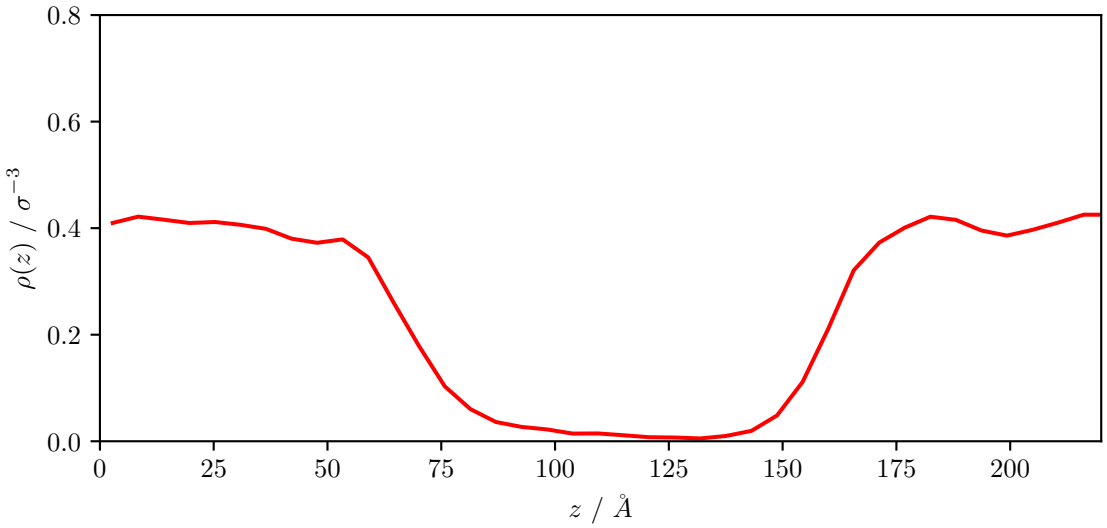


Figure 15: Density plot, computed using data from '288.xyz'

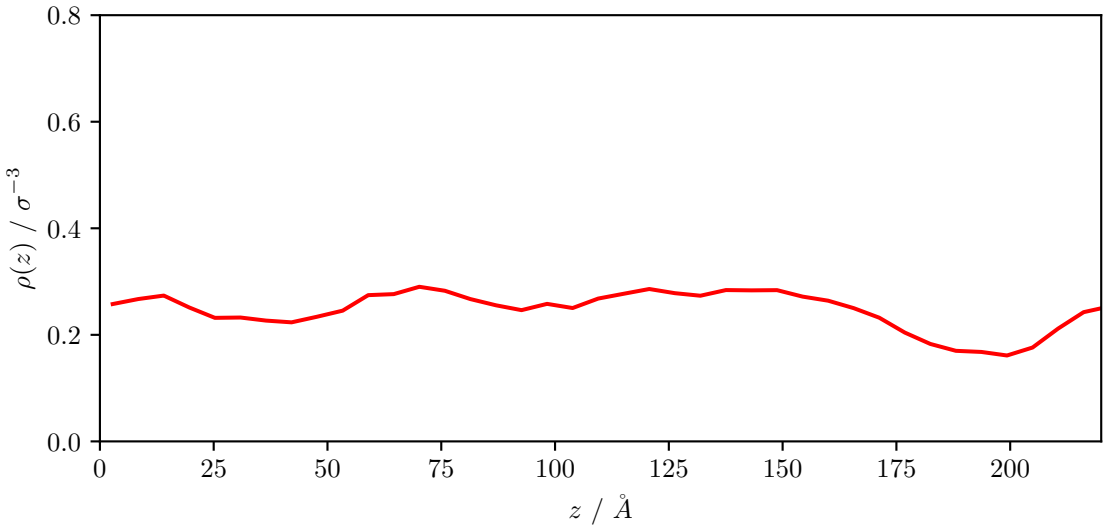


Figure 16: Density plot, computed using data from '300.xyz'

The following phase diagram was produced from these plots:

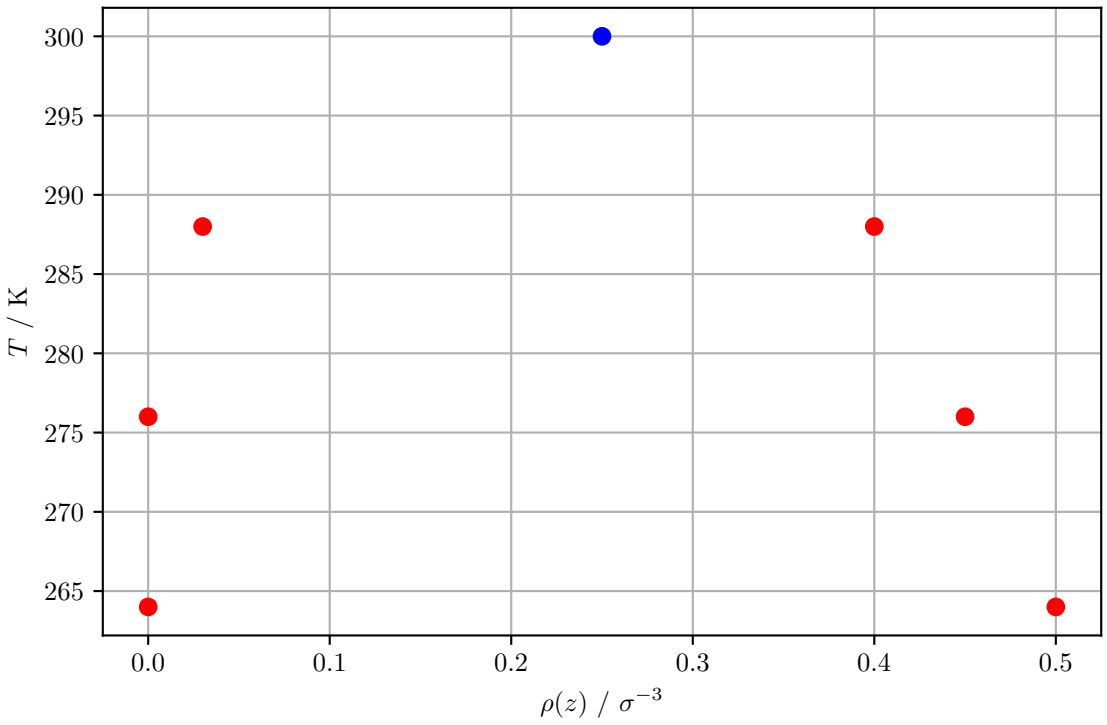


Figure 17: Phase diagram for the polymer system

7 Monte Carlo Code

7.1 Photon Gas

The total energy of the photon gas is given by

$$U = \sum_{j=1}^{N} n_j \omega_j \hbar = \sum_{j=1}^{N} n_j \epsilon_j \tag{37}$$

The canonical partion function is defined as

$$Q = \sum_{i} \mathrm{e}^{-\beta E_i} \tag{38}$$

where we sum over all microstates i, and $E_i = \sum_{j=1}^{N} n_j \epsilon_j$ is the energy of microstate i for a given set of values for the occupancy numbers $n_j \in [0, 1, 2, \ldots, \infty]$. Using the equation for the total energy of the photon gas, the exponent in Equation (38) can be written as

$$\mathrm{e}^{-\beta E_i} = \prod_{j=1}^{N} \mathrm{e}^{-\beta n_j \epsilon_j} \tag{39}$$

Summing over all microstates is equivalent to summing over all possible values of n_j for each j. The partition function therefore becomes

$$\begin{aligned} Q &= \sum_{n_1=0}^{\infty} \mathrm{e}^{-\beta n_1 \epsilon_1} \sum_{n_2=0}^{\infty} \mathrm{e}^{-\beta n_2 \epsilon_2} \dots \sum_{n_N=0}^{\infty} \mathrm{e}^{-\beta n_N \epsilon_N} \\ &= \boxed{\prod_{j=1}^{N} \frac{1}{1-\mathrm{e}^{-\beta\epsilon_j}}} \end{aligned} \tag{40}$$

where the final result is obtained using the following identity for $|x| < 1$:

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x} \tag{41}$$

To evaluate the average occupancy number for mode j, we write

$$\langle n_j \rangle = \frac{1}{Q} \sum_i n_j \mathrm{e}^{-\beta E_i} \tag{42}$$

Using Equation (39) it is clear to see that

$$\frac{\partial}{\partial(-\beta\epsilon_j)} \left(\mathrm{e}^{-\beta E_i}\right) = n_j \mathrm{e}^{-\beta E_i} \tag{43}$$

Substituting this relation into Equation (42) gives the result:

$$\langle n_j \rangle = \frac{1}{Q} \frac{\partial Q}{\partial(-\beta\epsilon_j)} = \frac{\partial \ln Q}{\partial(-\beta\epsilon_j)} \tag{44}$$

Evaluating this for the expression for Q yields

$$\begin{aligned} \langle n_j \rangle = \frac{\partial \ln Q}{\partial(-\beta\epsilon_j)} &= \frac{\partial}{\partial(-\beta\epsilon_j)} \left[\ln \left(\prod_{j=1}^{N} \frac{1}{1-\mathrm{e}^{-\beta\epsilon_j}} \right) \right] \\ &= \frac{\partial}{\partial(-\beta\epsilon_j)} \left[\ln \left(\frac{1}{1-\mathrm{e}^{-\beta\epsilon_1}} \right) + \ldots + \ln \left(\frac{1}{1-\mathrm{e}^{-\beta\epsilon_N}} \right) \right] \\ &= \frac{\partial}{\partial(-\beta\epsilon_j)} \left[-\ln\left(1-\mathrm{e}^{-\beta\epsilon_1}\right) - \ldots - \ln\left(1-\mathrm{e}^{-\beta\epsilon_N}\right) \right] \\ &= \boxed{\frac{1}{\mathrm{e}^{\beta\epsilon_j}-1}} \end{aligned} \tag{45}$$

The following plots were obtained for $\langle n_j \rangle$ using the Metropolis Monte Carlo algorithm:

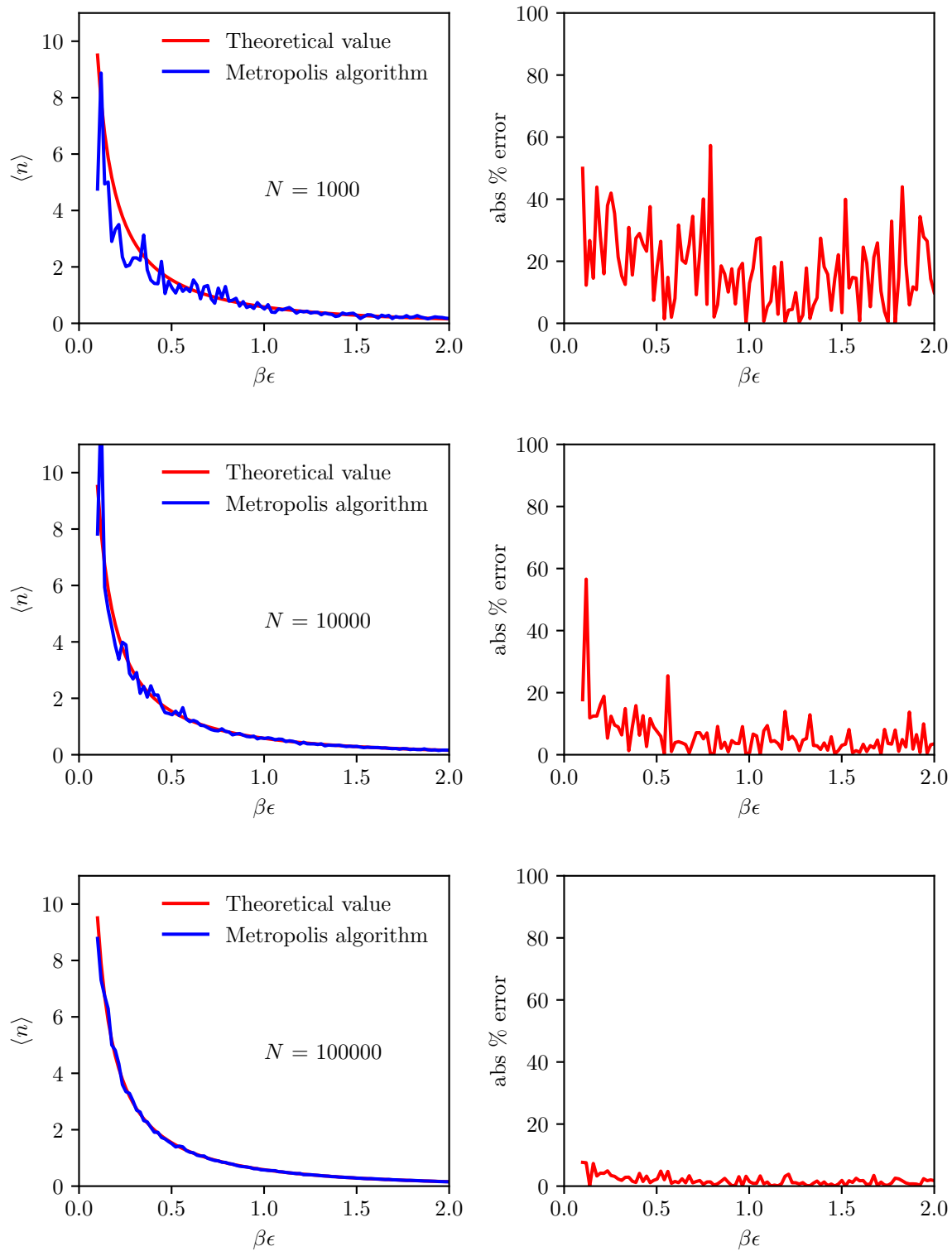


Figure 18: Plots of $\langle n_j \rangle$ against $\beta\epsilon$ computed for different numbers of Monte Carlo steps, N.

As is expected, the Metropolis algorithm computes $\langle n_j \rangle$ with decreasing error with increasing N. By only updating the averages when the trial move is accepted, the following result was obtained for $N = 100000$:

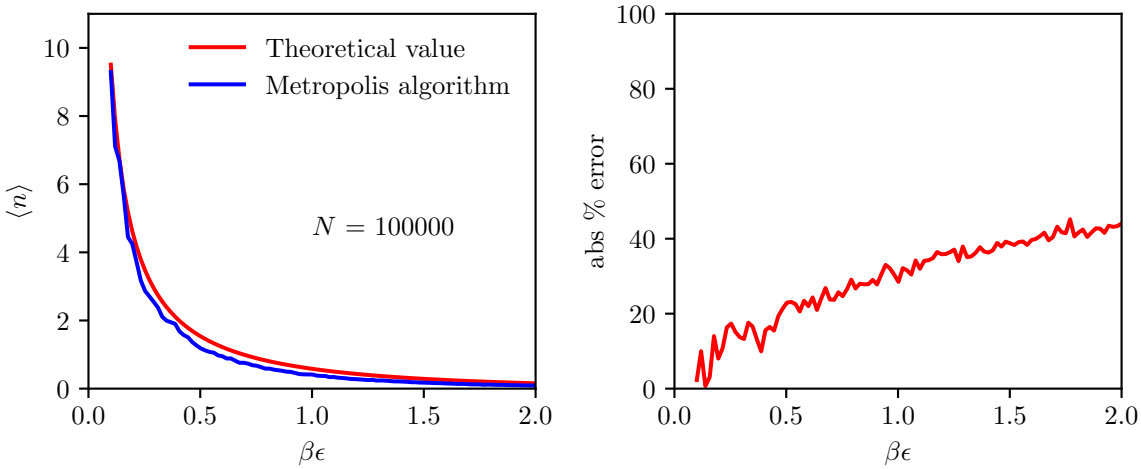


Figure 19: $\langle n_j \rangle$ computed without updating the average if the trial move was rejected.

Here it is clear to see that as $\beta\epsilon$ increases, the relative error increases. It is also evident that this method of computing $\langle n_j \rangle$ is consistently underestimating the theoretical value. The error is more pronounced for higher $\beta\epsilon$ values since this corresponds to an exponentially decreasing probability of a trial move being accepted, so an increasing number of steps are not included in the average.

8 Importance Sampling

The integral to be evaluated in this exercise is the following:

$$f(x) = 3x^2 \qquad I = \int_0^1 f(x)\mathrm{d}x = 1 \tag{46}$$

The Monte Carlo importance sampling method evaluates this integral in the following way:

$$I = \int_a^b f(x)\,\mathrm{d}x = \int_a^b \frac{f(x)}{g(x)} g(x)\,\mathrm{d}x = \mathbb{E}\left[\frac{f(x)}{g(x)}\right]_g = \frac{1}{N}\sum_{i=1}^{N} \frac{f(x_{i|g})}{g(x_{i|g})} \tag{47}$$

for some weight function $g(x)$. The notation on the right-hand side of Equation (47) means that one takes an average of the quantity $f(x)/g(x)$ for a set of values $x_{i|g} \sim G$ (i.e. follows the distribution $g(x)$ defined over the region $[a, b]$). The following results were obtained for the Monte Carlo evaluation of the integral in Equation (46) using three different weight functions. Results from the trapezoidal rule are also shown:

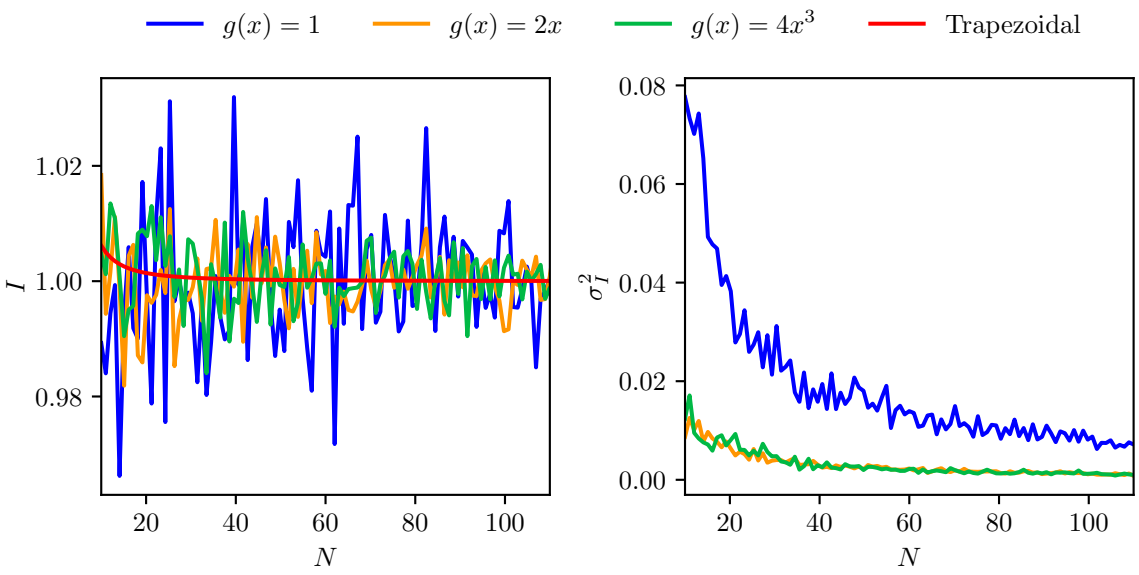


Figure 20: Monte Carlo numerical integration of the function $f(x) = 3x^2$. The left-hand panel shows the average value of the integral for 100 runs of the Monte Carlo code for each weight function, as a function of the number of sampling points, N. The right-hand panel displays the variance scaling of each weight function with respect to N.

The data clearly shows that the naive uniform weight function has the highest error, and therefore the poorest estimation of I. The two other weight functions skew sampling towards the high end of the range $[0, 1]$, which both give comparable accuracy for the integral. The accuracy for these two weight functions is better due to the nature of $f(x)$ having a larger proportion of its integral at higher values in the range. For this simple one-dimensional integral, it is clear that the trapezoidal rule is far superior in accuracy. However, the error in the trapezoidal rule scales as $N^{-3/d}$ where d is the dimensionality of the integral, so for multi-dimensional integrals over some hypervolume, Monte Carlo quickly becomes a more useful method (error scales as $N^{-1/2}$).

The best weight function for this integral is the function itself, i.e. $g(x) = f(x)/\langle f \rangle$. In this case $\sigma_I^2 \to 0$. This requires knowledge of the integral in the first place, so in practice one would not know exactly what the best $g(x)$ is.