MENG 255 PSET 2:
NVE and NVT Molecular Dynamics

Noah Dohrmann

(Dated: 25 February 2022)

Please see the corresponding code (simulation2_main.py) and the example Jupyer notebook which go along with this report.

I. INTRODUCTION

This problem set makes use of phase space conserving integration techniques in NVE and NVT Molecular Dynamics simulation for the prediction of physical observables of a trial system.

A. Theory

The underlying potential governing the system is the shifted and truncated Lennard-Jones potential:

$$V_{\mathrm{LJTS}}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] - V_{\mathrm{LJ}}(r_c) \tag{1}$$

so that the total Hamiltonian is $H = T + V_{\mathrm{LJTS}}$. A cutoff radius of $r_c = 2.5\sigma$ is used. The integration scheme used for the NVE simulation is the Velocity-Verlet scheme in which velocity is updated in two half steps $1/2\delta t$ while the spatial coordinates are updated in one step. They are given by Allen and Tildesley[1] as

$$\boldsymbol{v}\left(t + \frac{1}{2}\delta t\right) = \boldsymbol{v}(t) + \frac{1}{2}\delta t \boldsymbol{a}(t) \tag{2}$$

$$\boldsymbol{r}(t + \delta t) = \boldsymbol{r}(t) + \delta t\, \boldsymbol{v}\left(t + \frac{1}{2}\delta t\right) \tag{3}$$

$$\boldsymbol{v}(t + \delta t) = \boldsymbol{v}\left(t + \frac{1}{2}\delta t\right) + \frac{1}{2}\delta t\, \boldsymbol{a}(t + \delta t) \tag{4}$$

For the NVT simulations via the Nosé-Hoover Thermostat, an additional coordinate s with an effective mass Q is considered. The total Hamiltonian of this system is written (using the thermodynamic friction coefficient ξ) as:[2]

$$H_{\mathrm{Nosé}} = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m_i} + U\left(\boldsymbol{r}^N\right) + \frac{\xi^2 Q}{2} + L\frac{\ln s}{\beta} \tag{5}$$

Before attempting either the NVE or NVT scheme, the particles are placed on a cubic lattice at the specified density in the problem statement (usually $\rho = 0.9$ [particle/vol]). This is manually verified by inspection of the spatial coordinates. The following are the results for each of the questions asked in the problem statement.

II. QUESTIONS 1 AND 2

In the first question, we write functions to find the neighbor list from a cell list via periodic boundary conditions. Additionally, a force calculation method is also employed via finding the gradient of the shifted-truncated Lennard-Jones potential, ie. $\boldsymbol{f} = -\nabla V$. In practice it is easier to treat the Lennard-Jones potential as the function of a scalar r, so that is may be differentiated with respect to this variable only. Information on the direction of the resulting force vector is found via the "nearest-neighbor" vector, the components of which are defined as

$$x_{ij,\mathrm{nearest}} = x_{ij} - L_x\ \mathrm{nint}(x_{ij}/L_x) \tag{6}$$

with the nearest integer function appearing on the right hand side. Then, the force vector in my code is updated as $\delta\boldsymbol{f} = -f_{\mathrm{mag}}\boldsymbol{r}_{\mathrm{nn}}$, where the RHS vector is the normalized vector pointing to the nearest neighbor. The magnitude of the force is also adjusted accordingly so that it falls off to zero continuously when the cutoff radius is reached[3]. These $\delta\boldsymbol{f}$ are then summed for every non-truncated interaction that the particle experiences.

In terms of constructing the neighbor list itself, a cell list is first made by dividing the simulation volume into a series of cubic cells and recording which particles reside in which cell at the current time step. Then, each particle's own neighbor list is found by adding together the sets of other particles which are located in the first particle's cell as well as those that lie adjacent (this also includes cells on the other end of the total box if the center cell is located on a boundary.) Then, this set is further refined by comparing the distance from particle i to the nearest image of particle j via the above neighbor formula. If this distance is greater than r_c, the interaction will not be considered.

To determine if the neighbor list should be updated, the two largest displacements of any particles i and j with respect to the geometry at the time of the last neighbor list update is considered[1]. If this distance is larger than a buffer radius parameter r_{buff}, then the neighbor list will be updated at the start of the next iteration.

Lastly, the time step corresponding to each integration iteration is $\delta t = 0.001$ in reduced units for most NVE simulations, and 0.01 for most NVT simulations, unless noted otherwise.

A. Trial NVE Integration

Following the Verlet procedure from above, a sample integration of 120 iterations was run as a proof of concept for

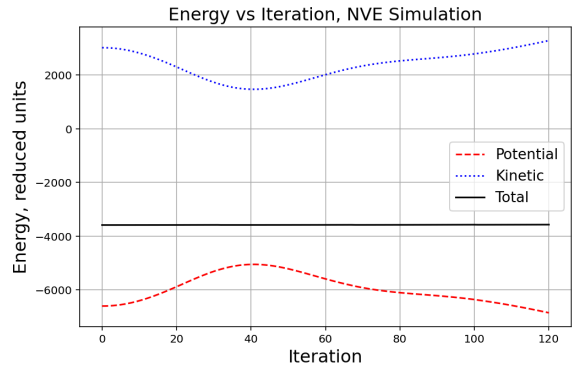

FIG. 1. The potential, kinetic, and total energy sums for a trial NVE integration at $N = 2000$, $\rho = 0.9$, and reduced $T = 1.0$

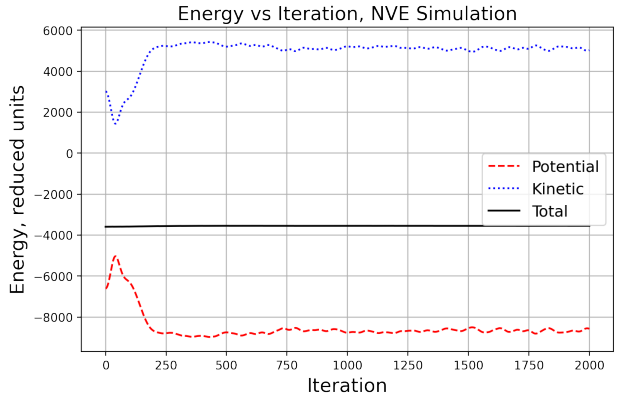

FIG. 2. (For a somewhat longer simulation) The potential, kinetic, and total energy sums for a longer NVE integration at $N = 2000$, $\rho = 0.9$, and reduced $T = 1.0$

the NVE integrator. The results are found in Fig. 1 As expected, there is little seen total energy drift, though K and U do not settle, as the duration of the simulation is too short. A longer duration simulation of 2000 iterations is also reported with Fig. 2. Both trials succeed in the goal of demonstrating energy conservation with the NVE simulation ensemble.

B. Trial NVT Simulation

There are more parameters involved in the NVT simulation, namely those stemming from the introduction of the thermal masses ξ_k which act as thermal reservoirs for the system. The full implementation of an integration step is found in the posted guide on NVT implementation as well as the resulting code. There are two parameters Q_1 and Q_2 that are initially guessed and later tuned to improve results; they appear in the definitions of two bookkeeping quantities as follows:

$$G_1 = \frac{1}{Q_1}\left(\sum_i m_i v_i^2 - L/\beta\right) \tag{7}$$

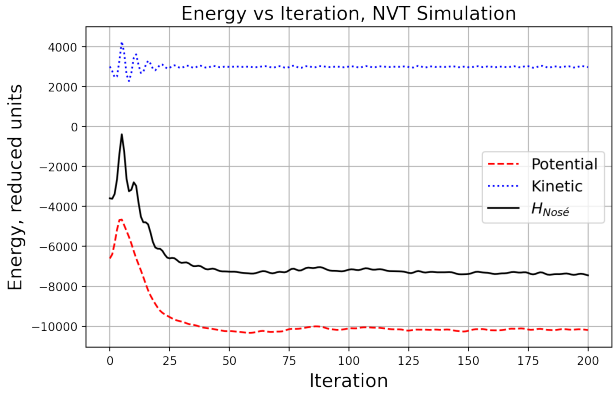

FIG. 3. Potential, Kinetic, and $H_{\text{Nosé}}$ energies for a trial NVT run. $N = 2000$, $\rho = 0.9$, $T = 1.0$.

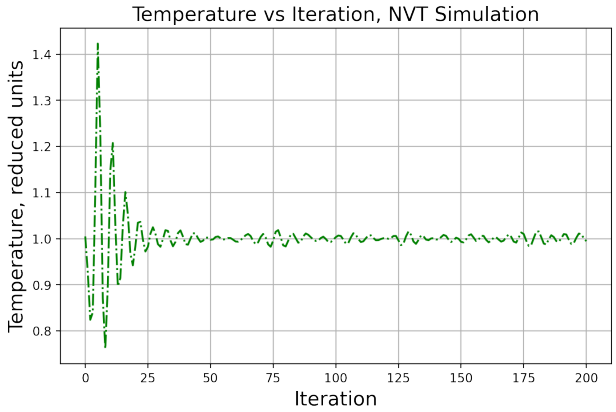

FIG. 4. Temperature log for a trial NVT run. $N = 2000$, $\rho = 0.9$, $T = 1.0$.

$$G_2 = \frac{1}{Q_2}\left(Q_1 v_{\xi_1}^2 - k_B T\right) \tag{8}$$

which are used when rescaling velocities to match the heat bath. For the following results, assume $Q_1 = Q_2 = 1.0$ unless otherwise stated. It should be noted that the quantity L in the definition of G_1 *is not* the length of the simulation box, but instead equal to $3N$. In addition to tracking the total U, K, and T of the system, we are also interested in monitoring the conserved Hamiltonian $H_{\text{Nosé}}$ that has been defined above. Please see the corresponding code in "NVT_full" for the implementation of tracking the Nosé-Hoover term. Ideally, the data for $H_{\text{Nosé}}$ should not see any substantial fluctuations.

The preliminary results display a convergence for both energy (U, K, and the terms added in $H_{\text{Nosé}}$) and temperature. The resulting value around which temperature settles is approximately a the desired $k_B T = 1.0$, which suggests there are no serious numerical flaws in the NVT integrator (aside from the the tracking of $H_{\text{Nosé}}$) for the time scale this simulation was run (200 steps at $\delta t = 0.01$). An energy log is found in Fig. 3 and a temperature log is found in Fig. 4.

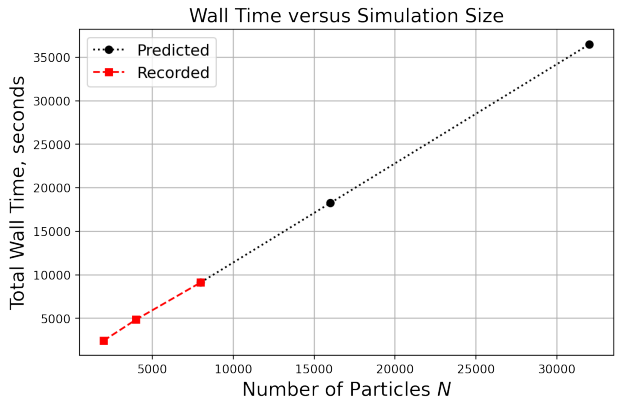

FIG. 5. Wall time results for a cell list enabled simulation at various N over 1000 iterations at $\delta t = 0.01$. Observed and predicted (estimated slope of 2) time shown.

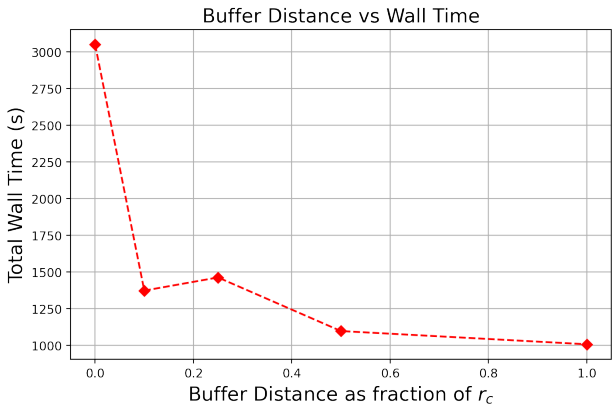

FIG. 6. Buffer distance as a fraction of r_c versus required computation time. $N = 2000$, $T = 1.0$, $\rho = 0.9$, iterations $= 500$, $\delta t = 0.01$.

III. QUESTION 3

For this question we run a series of tests on the NVT integrator and track its performance for a range of temperatures, particle counts N, and buffer distances r_{buff}. The effect of the neighbor list on the total computation time will also be recorded. However, in order to keep total wall-time spent waiting on the simulations reasonable (while still producing results more quantitative than "> 10 minutes"), it was decided to reduce the total time steps for the simulations from fifty thousand to one thousand.

A. Simulation Wall Time

The first round of simulations conducted are displayed in Fig. 5 for the conditions $T = 1.0$, $\rho = 0.9$, and $N = 2, 4, 8, 16$, and $32 * 10^3$ with the cell list enabled at a buffer distance of $r_{\text{buff}} = 0.5 r_c$ for 1000 iterations. As might be expected, there is roughly a two fold increase in total simulation time from each N to the next. This is indicative of the success of the cell list in reducing the computation time needed from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$. The last two data points still would require a very long simulation time, so they are extrapolated from the measured times rather than recorded directly.

The second round of calculations (without a cell or neighbor list) is run for the same parameters as above. However, due to an extreme increase in computation time associated with the neglect of optimization techniques such as the cell/neighbor list, only much more limited data is available. For a system of $N = 2000$ particles, the run time for 1000 time steps is: 62 768.77 seconds. It was deemed impractical to continue waiting on results past this.

B. Optimizing the Buffer Distance r_{buff}

In order to investigate the impact of buffer radii on the total wall time of a simulation, a series of tests will be run for $r_{\text{buff}} = (0, 0.1, 0.25, 0.5, 1.0) * r_c$. Again, to keep the times reasonable the simulation parameters were adapted so that $N = 2000$ and the total number of iterations is 500. Other parameters ($T = 1.0$, $\rho = 0.9$) were left unchanged. See Fig. 6 for data. Surprisingly, there is a slight increase in computation time from $r_{\text{buff}} = 0.1 r_c$ to $0.25 r_c$, although the shortest simulation of the set is with $r_{\text{buff}} = 1.0 r_c$. This is caused by, in the way I have formulated the update criteria for the neighbor list, the buffer length scales negatively with the frequency of the neighbor list updates. A higher frequency in updates will lead to a slower computation time.

IV. QUESTION 4

In this question, we return to the study of MSD's, given by

$$\text{MSD} = \frac{1}{N} \sum_{i=1}^{N} [\vec{r}_i(t) - \vec{r}_i(0)]^2 \tag{9}$$

for a given geometry $C(t)$ as compared to its original geometry $C(0)$. As with the last problem set, diffusion constants from MSD are calculated with the *real space* displacement of the particles, not their virtual image positions at time t.

First, for an NVE simulation, the initial temperature is set at $T = 1.0$, and the other parameters are left at what their standards have been thus far ($N = 2000$, $\rho = 0.9$). Over the course of the run (1000 iterations at $\delta t = 0.001$), the MSD along with the new velocity-averaged T at the new time. From previous NVE runs, it is expected that T will converge to a stable value after a certain number of iterations when the system reaches equilibrium U and K values. The corresponding results are in Fig. 7 and Fig. 8. Interestingly, the derivative of MSD with respect to iteration appears to decrease, while remaining positive, for later iterations. The temperature for the same simulation, once settling after about 200 iterations, settles on an average value of approximately $T = 1.699$ in reduced units. In order to see if these trends persisted, a longer series of trials was conducted with the number of iterations being 10,000. In these subsequent results, (Fig. 9 and Fig. 10) it is confirmed

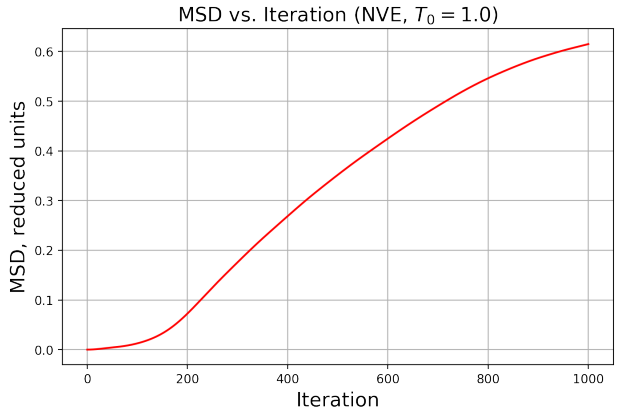

FIG. 7. The system total MSD as a function of iteration for an NVE simulation at the standard system parameters.

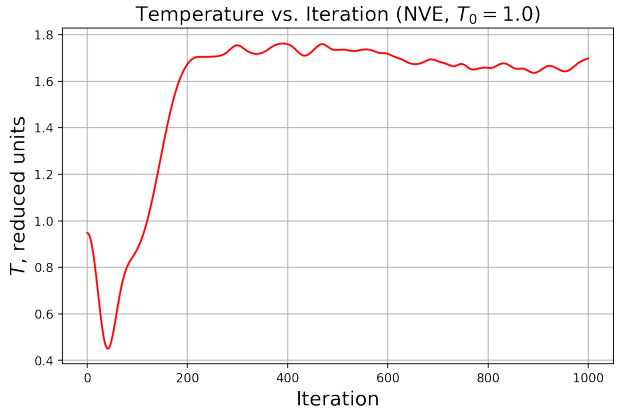

FIG. 8. The corresponding temperature for the previous figure as a function of iteration for an NVE simulation at the standard system parameters.

that MSD will rise with the number of iterations in a linear fashion rather than reaching a plateau as suggested by the first run. The equilibrium average of the 10,000 iteration run temperature was found to be $T = 1.714$.

Next, similar tests were run with NVT simulations, this time at the set temperature of $T = 1.70$ to match the previous NVE run. Moreover, several values of the friction parameter ξ (applied on both ξ_1 and ξ_2) were tested ($\xi \in \{0.1, 0.5, 1.0\}$). The output of these trials is found Fig. 11 and Fig. 12. There is not a large effect on the overall MSD path of the system for the three values of ξ tested, though this could be due to the fact that the sampled values only range one order of magnitude.

V. QUESTION 5

Having now verified both the NVE and NVT integrators as well as their MSD results, this next question will focus on other physical and observable properties of the sytem, namely the heat capacity C_v of the fluid. This will be determined via

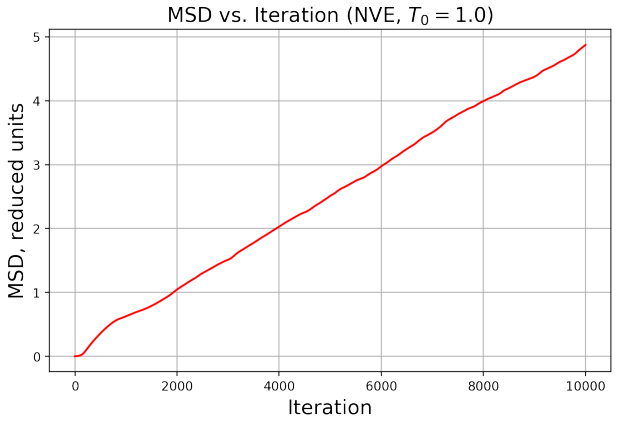

FIG. 9. (Longer duration simulation, $\delta t = 0.001$) The system total MSD as a function of iteration for an NVE simulation at the standard system parameters.

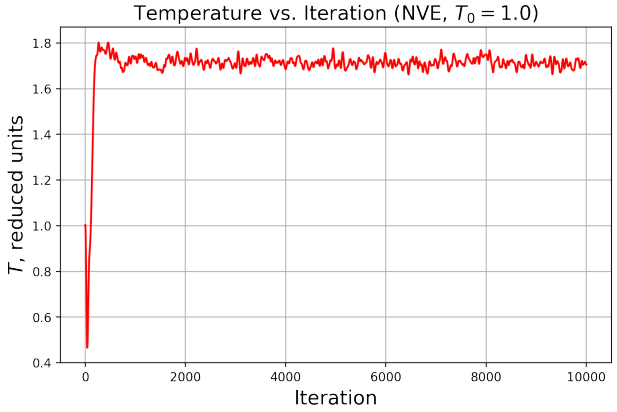

FIG. 10. (Longer duration simulation, $\delta t = 0.001$). The corresponding temperature for the previous figure as a function of iteration for an NVE simulation at the standard system parameters.

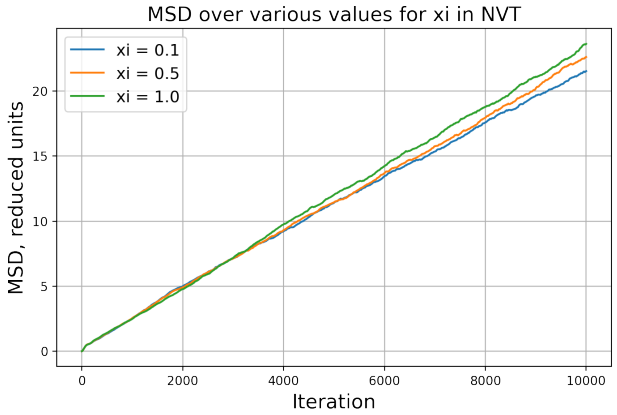

FIG. 11. MSD results for NVT simulation at varying ξ's at $T = 1.70$.

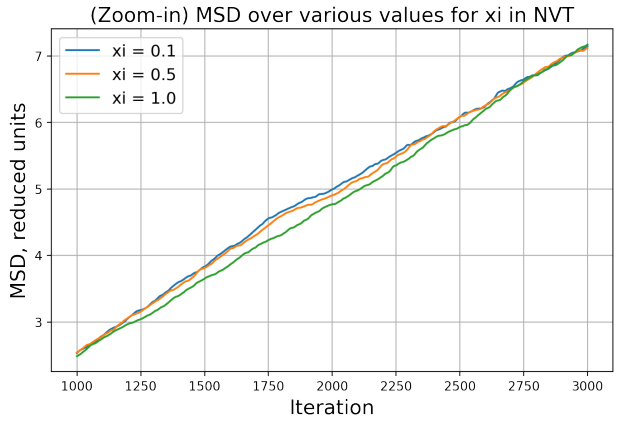

FIG. 12. A zoomed in version of the previous figure for an iteration range of 1000 to 3000.

TABLE I. NVT Specific Heat Capacity Results (reduced units)

Temperature	Heat Capacity
1.0	2031.6
2.0	2583.3

the relation

$$C_v k_B T^2 = \langle \delta E^2 \rangle \tag{10}$$

which will be used to verify heat capacity with several different starting temperatures for NVT simulations. For each case, the variance in E will be measured after the system has reached equilibrium, which usually occurs on the order of 10^2 or 10^3 time steps ($\delta t = 0.01$ for both in this case). The method adopted for NVE simulations must be formulated somewhat differently, as by definition there will be a very low variance in total E for an NVE simulation. Accordingly, the familiar physical interpretation of heat capacity is used (i.e. the amount of energy increase associated with an increase in temperature), written as

$$C_v = \frac{Q}{\Delta T} \tag{11}$$

In this context, Q is interpreted as the difference in total energy $U_2 - U_1$ of the $N = 2000$, $\rho = 0.9$ system between different temperatures.

Trial 1 for the NVT integrator was run at the standard $N = 2000, \rho = 0.9$, and $T = 1.0$ for 1000 iterations. From the energetic results, the equilibrium values of U and K are reached at the latest by the 200th epoch. As such, the variance in the energy for the last 800 epochs is recorded, and the heat capacity is found. The second NVT integration run was conducted at the same conditions, excepting $T = 2.0$. Again, after equilibrium was reached the resulting variance in total energy is recorded. Limited results are available due to high computation time, but it appears that there is an increase in C_v with an increase in T. These are found in Table I.

Now, for the NVE tests (at the same specifications), several different T_0 are selected: $T_0 \in \{1.0, 2.0, 2.5, 5.0\}$. The average

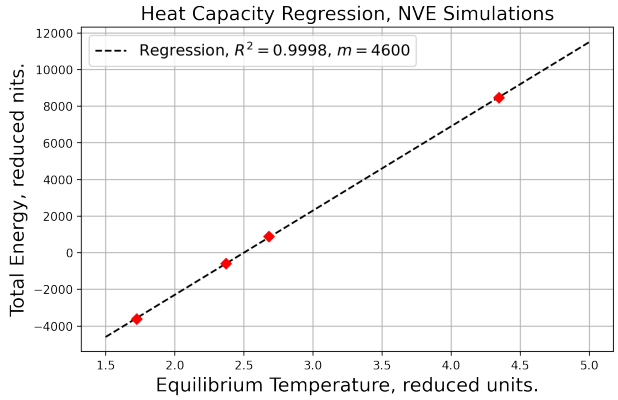

FIG. 13. The results of several NVE simulations at the above equilibrium temperatures. The change in total or internal energy between the systems (which each started at the same $N = 2000$, $\rho = 0.9$ lattice) is used as a measure of heat capacity C_v.

energy and average equilibrium temperatures are recorded and compared in Fig. 13. A regression of these points gives the slope

$$m = C_v = \frac{\Delta U}{\Delta T} = 4600.9$$

as the resulting heat capacity. Interestingly, this quantity lies around twice the value as the C_v obtained for the NVT method. If the two methods are supposed to produce similar results, this may be due to numerical errors in either of the integration methods.

VI. QUESTION 6

Lastly, we calculate the radial distribution function $g(r)$ for the NVT integrator for a series of temperatures and densities. At a given geometry, the radial distribution function is calculated as[1]:

$$g(r) = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i} \delta(r - r_{ij}) \right\rangle \tag{12}$$

It is noted that for practical reasons, the delta function should be replaced with a small range of acceptable values instead of a pure Dirac delta spike. This is implemented in the "rad_dist_2" function in the code.

For the initial crystal-like structure that emerges when the particles are placed on a cubic lattice, the distance from one particle to another will take on only certain values, and a continuous distribution is not expected. Rather, one might anticipate that there are sharp peaks around certain distances, eg $\approx 1, \sqrt{2}, 2, \ldots$ (for $N = 2000$, $\rho = 0.95$). This is precisely what is found for $g(r)$ when it is calculated for the initial geography, present in Fig. 14. At equilibrium (Fig. 15, there are no longer sharp peaks, but there is rather a series of jagged oscillations over the recorded range of r. Qualitatively, this corresponds to the mixing of particles across the spatial coordinates of the box. Below a reduced r of one, there is very

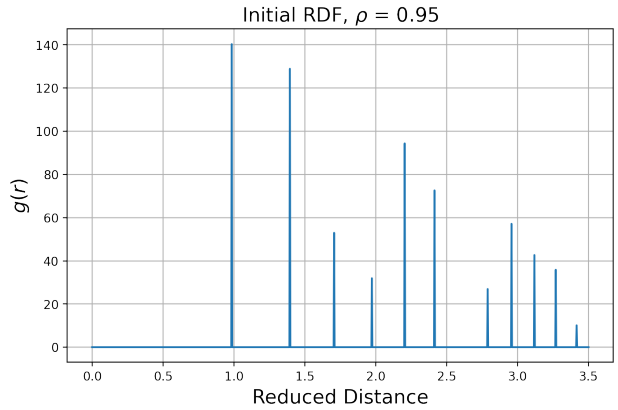

FIG. 14. Radial distribution function $g(r)$ for the initial cubic lattice. The geometry of the lattice at $N = 2000$ and $\rho = 0.95$ is such that the distance from particle $P_{i,j,k}$ to $P_{i+1,j,k}$ is ≈ 1. (if i, j, and k correspond to the x, y, and z indices of an element in a 3D array.)

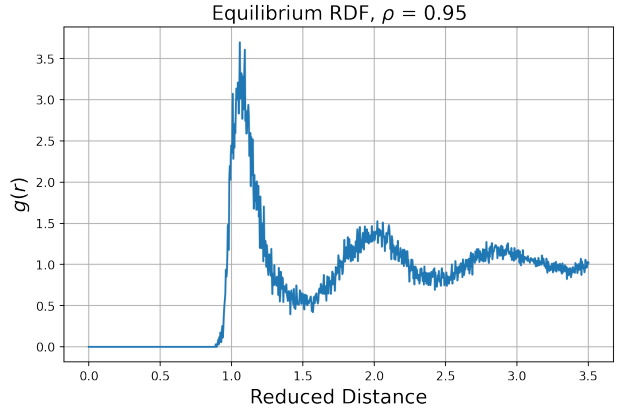

FIG. 15. Radial distribution function $g(r)$ for the equilibrated structure. The geometry of the lattice is $N = 2000$ and $\rho = 0.95$, with $T = 1.0$.

little particle density seen, as this region corresponds to a extremely high energies as defined from the Lennard-Jones potential. At larger r, $g(r)$ varies around 1.0 as the surveyed density across the larger region becomes closer to the bulk density of the fluid.

When similar tests are run for higher density configurations, $\rho = 1.05$, qualitatively $g(r)$ retains its shape, but the initial peaks are seen to be somewhat higher due to the increased density (see Fig. 16 for equilibrium at $T = 2.0$ and Fig. 17 for equilibrium at $T = 0.75$).

[1]M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 2nd ed. (Oxford University Press, 2017).

[2]D. Frenkel and B. Smit, "Chapter 4 - molecular dynamics," in *Understanding Molecular Simulation (Second Edition)*, edited by D. Frenkel and B. Smit (Academic Press, San Diego, 2002) second edition ed., pp. 63–105.

[3]S. Toxvaerd and J. C. Dyre, "Communication: Shifted forces in molecular dynamics," The Journal of Chemical Physics **134**, 081102 (2011), https://doi.org/10.1063/1.3558787.

[4]Z. Yao, J.-S. Wang, G.-R. Liu, and M. Cheng, "Improved neighbor list algorithm in molecular simulations using cell decomposition and data sorting method," (2003), 10.1016/j.cpc.2004.04.004, arXiv:physics/0311055.

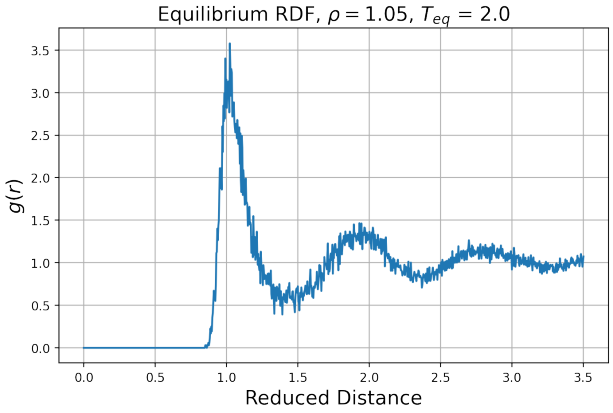

FIG. 16. Radial distribution function $g(r)$ for the equilibrated structure. The system parameters are $N = 2000$ and $\rho = 1.05$, with $T_{\text{eq}} = 2.0$.

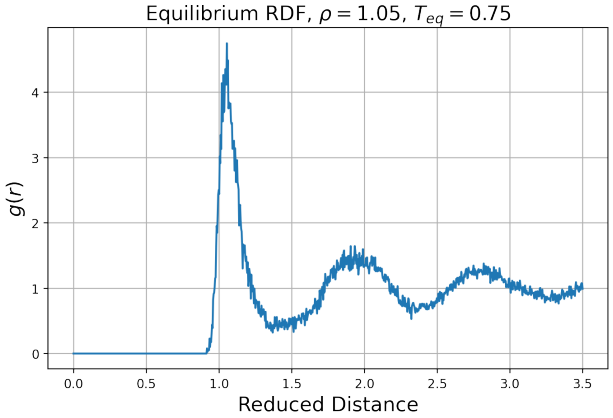

FIG. 17. Radial distribution function $g(r)$ for the equilibrated structure. The system parameters are $N = 2000$ and $\rho = 1.05$, with $T_{\text{eq}} = 0.75$.

[5]G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, "Explicit reversible integrators for extended systems dynamics," Molecular Physics **87**, 1117–1157 (1996), https://doi.org/10.1080/00268979600100761.