Classroom Tutorial : Analysis of Simulation Results

Dr. Madhav Ranganathan

Department of Chemistry,
IIT Kanpur

November 8, 2010

Outline

- Programming Tips
- Analyzing simulation results: Static properties
- Analysis of simulation results: Dynamic properties

Reference: Frenkel and Smit: Appendix D and Appendix L

Programming Tips

- Preexisting code vs. self-written code/modified code
- Ensuring that the program is correct !!
- Inspection/Analysis of the results: Visual vs. Numerical
- Compare to experiments/other simulations

Whether you are using a pre-existing code or a self-written code, it is most likely that you will be writing some analysis tools of your own. For example, you might want to track one particle of interest in your simulation. Or calculate some higher order correlation functions.

Programming Tips

- **Design (Algorithm), Implementation (Programming), Testing (Analysis)**
- **Implementation (Programming)**
- **Testing (Analysis)**

Only when these three aspects are complete can one use the program for generating results.
Be Suspicious and Thorough
Simulations always give results for a model system

Modifying/Testing existing programs

- Understand every line in the code
- Comments/Notes on proposed modifications
- Test against limiting cases and *known results*- Ex. High temperature/Low Temperature limits/ Turn off certain effects

Testing Programs

- NVE Ensemble: Check Energy conservation, Energy fluctuations
- Monte Carlo Simulations: Calculate total energy at each step
- Reduced units - values should be of the order of 1.0
- Monte Carlo Simulations: Try different *pseudo* Random Number generators – VERY IMPORTANT FOR LONG SIMULATIONS – Mersenne Twister algorithm

Analysis of Simulation Results: Data

Simulation *DATA*

Data is characterized by Statistical Errors – Need ways to characterize these statistical errors

Most(all) properties of interest are related to averages.

- If you are given a bunch of data for a property X that are statistically independent, thenthe best estimate for the value of X is the

$$X_{bestest} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

- However, it is crucial that the values of X are statistically independent - Independent, Idetntically distributed Random numbers
- We need to generate independent, identically distributed random numbers - MC vs MD

Ergodicity and Simulations

- Ergodic Hypothesis

$$A_{expt} = \langle A \rangle = Lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau A(\tau)$$

  Not very useful for simulations since we cannot take very large time

- Fundamental Theorem of Computer Simulations

$$\langle A \rangle = \left(\frac{1}{t_2 - t_1} \right) \int_{t1}^{t2} A(t) = \frac{1}{N_{conf}} \sum_i^{configurations} A_i$$

Thus we have to make sure that our system has equilibrated and then calculate averages

Static Properties - Averages

$$A_\tau = \frac{1}{\tau}\int_0^\tau A(t)dt$$

What should be the value of τ over which we average ?
Average should be over numbers that span a time range much larger than the correlation time because

$$\sigma^2(A) = \frac{1}{\tau^2}\int_0^\tau\int_0^\tau C(t,t')dtdt'$$

If t_c is the correlation time then τ should be chosen to be much larger than the correlation time in order for the RMS error to be small.

Static Properties - Pair Correlation Functions

Pair correlation functions give distance dependence in our equilibrium properties.

$$\langle \rho(r)\rho(r')\rangle = \rho^2 g(r - r')$$

How can one calculate $g(r)$ from computer simulation data ?
We have a table of values of r_i of all the particles. Use this to calculat $g(r)$. Do on Blackboard.

Pick a value of dr

1. Consider each particle you have in turn. Count all particles that are a distance between r and r + dr away from the particle you're considering.
2. Divide your total count by N, the total number of particles in your data.
3. Divide this number by $4\pi r^2 dr$, the volume of the spherical shell
4. Divide this by the particle number density. This ensures that g(r)=1 for data with no structure.

Dynamic Properties - Time Correlation

Given a whole trajectory $\mathbf{R_i}(t), \mathbf{P_i}(t)$, how can you calculate the time correlation function ? Block Averaging.Do on Blackboard