Final Project

Due: Wednesday, 12/11/19 at 12pm

Objective: To use a molecular simulation to investigate some physical phenomenon of interest, and to present numerical results alongside written and visual explanations.

Topic selection and problem formulation

For this project, you will select a model of an atomic or molecular system that will allow you to explore its physical properties and behavior through simulation. You are encouraged to select a system that is of interest to you or to your research. Alternatively, you may select a system designed to address or explain recent results in the broad scientific literature, such as those that might be found in a Science or Nature article.

It is very highly suggested that you select a fairly simple model for exploring the behavior of interest. While detailed all-atom models are widely used in the literature, implementations of these can require more effort than what is reasonable in the time frame of the course. Instead, choose a simple or coarse-grained model that contains the very basic driving forces that you expect to be relevant to the physical phenomena of interest, i.e., models and systems with:

- **only a few atom types and kinds of terms in the potential energy function**, and hence only a few parameters to tune
- **no long-range interactions** (e.g., Coulombic). These would require special methods to compute the energies (e.g., the Ewald summation). You may, however, choose highly approximate electrostatic methods, such as screened Coulomb interactions, that are simple to use and short-ranged.
- **a reasonable number of atoms.** Remember that simulation expense scales as N^2. Atomic systems with approximately $N > 500$ will likely require long simulation times.
- **no strong order-to-disorder transitions.** It is very challenging to simulate transitions from disordered to ordered phases (e.g., crystallization, micellization) because the time scales can be quite long. One might simulate a crystal, given a known initial structure, but a simulation of crystallization from liquid would be very difficult.

The selection of your model should be motivated by one or two physical questions of interest that you will investigate with simulation. Some sample questions:

- What are the relevant scaling laws or functional dependence of properties with model parameters or thermodynamic state?
- How do properties change with molecular architecture (e.g., polymer length) or composition (e.g., fraction of hydrophobic groups)?
- What are the characteristic time or length scales in the system?
- What is the microscopic mechanism for an observed behavior?
- What are the dominant driving forces behind an observed behavior?
- To what extent can the behavior be described by simple theoretical expressions (e.g., mean field)?
- To what extent do finite-size effects play a role? At what system sizes do macroscopic descriptions start to break down?

If you are at a loss for creativity, here are a few ideas to help get your thinking started:

- binary mixtures of Lennard-Jones, soft-spheres, or other types of particles
- interactions of various kinds of molecules with surfaces and interfaces (e.g., relative binding affinities, adsorption isotherms, density profiles)
- behavior of systems under confinement between rigid walls, relative to bulk
- simple polymer models of proteins (e.g., "Go-models") or other biomolecules
- behavior of "functionalized" particles (e.g., particles with "sticky" spots on them)

Don't be afraid to *develop* a simple model that might shed light on some real system of interest. This may be easier than searching the literature for existing simulation models.

Methodology requirement

Your project needs to implement at least one advanced simulation method or topic discussed:

- distributions of molecular order parameters, such as radial distribution functions $g(r)$, orientational correlation functions, polymer metrics (e.g. Rg), fluid density as a function of distance from a wall, etc. [*doesn't count for advanced track*]
- time correlation functions and correlation times [*doesn't count for advanced track*]

- MC simulations using the isothermal-isobaric, grand canonical, Gibbs, or other ensemble besides the canonical one
- advanced MC moves other than displacement: orientational moves, biased moves, lattice polymer moves, hybrid MC/MD moves
- biased sampling methods: umbrella, multicanonical, or Wang-Landau sampling
- free energy calculations for thermodynamic free energies or potentials of mean force (PMFs) using any of a variety of methods: simple reweighting, weighted multiple histogram analysis or MBAR, Bennet's method, thermodynamic integration, Widom insertion, alchemical transforms, transition matrix methods
- phase equilibrium calculations
- other ideas – consult the instructor

Can I use an off-the-shelf simulation package for complex systems?

I strongly recommend that you pick a simple enough system so that you can write your own simulation code. However, if you are dead-set on studying a system that is relevant to your research and it would require advanced techniques (e.g., an Ewald summation) or many-parameter models (e.g., proteins, DNA, etc.), there is the possibility that you may use a package like AMBER, CHARMM, GROMACS, or LAMMPS. However, the analysis codes must be written by you *from scratch* (i.e., not taken from existing code outside of this class) and must include a sufficiently advanced methodology from the above list.

Topic pre-approval

Please pre-approve your topic with the instructor. This is to ensure that the project is feasible within the allotted time frame and satisfies the requirements above. The instructor is available to discuss options with you during the course office hours (preferred) or by appointment.

Some ideas for analysis

Your project should be designed to interrogate one or more physical questions of interest. The results, therefore, should provide a convincing argument for whatever conclusions or hypotheses you propose. It will be key to perform analyses that support your arguments. There are many thermodynamic, kinetic, and structural properties you might examine:

- thermodynamic averages for U, P, ρ, C_V and manipulations thereof (derivatives, etc.)

- free energies and entropies, and potentials of mean force
- conditions for phase equilibrium (e.g., phase diagrams)
- transport coefficients like the self-diffusivity D
- molecular structure metrics like $g(r)$, R_g, or coordination numbers (the average number of molecules within a first coordination shell of a central molecule)

You can study the behavior of these as a function of many different variables:

- thermodynamic state (e.g., T and P or ρ)
- system size (N or V)
- composition, for mixtures
- parameters in the force field (e.g., well depth or particle diameter, if multiple are defined)

Error bars

In any simulation study, it is always critical to report data of high statistical quality, with a clear understanding of the errors involved. In your project, you should provide error estimates on simulation averages (error bars in graphs). This is most easily accomplished by performing multiple simulation trials with different initial random number seeds.

What to turn in

For this project, you will put together two components:

You will write a short technical report describing the project and the results (~4-5 single spaced pages, no smaller than 11 point font with 1 inch margins). Your report should include:

1. A descriptive title of your project, your name, and the date.
2. A short abstract that summarizes the relevance of your system and the major results found.
3. A discussion that addresses the following:
    a. What is the problem being studied and why is it of interest?
    b. What is the model and its assumptions? What are the parameters, force field equations, etc.?

c. What is the simulation algorithm? What exact protocols and settings did you use? Enough detail should be provided so that your results could be reproduced.

d. What results were found? How should they be interpreted?

e. What conclusions or hypotheses can one generate about the behavior of the system from the results?

f. What might be potential deficiencies of the model and/or technique used, and how might your approach be improved (i.e., next steps)?

4. An annotated schematic / molecular image of the system under study
5. Graphs of important results found
6. A descriptive caption for the movie (below)

In addition, you will create a representative movie of one of your simulations. This can be developed by outputting positions periodically to pdb files (atomwrite.py is available on the course website) and importing the pdb trajectory file into one of a number of possible visualization packages. Perhaps the easiest program for this purpose is the freely-available program UCSF Chimera.

You will upload your project results to a course Box folder. Please create a subfolder with your first initial and last name (e.g., SShell). You will need to add to the Box subfolder:

1. your report in PDF format
2. a .zip or .gz compressed archive of your source code (and any relevant input files)
3. your movie in a format suitable for viewing on both Windows and Mac systems (without specialized software)

All of the projects for the course will be included in a virtual website gallery of simulation results on the course webpage that will be made available to the public. Links to your report, source code, and movie will be created.

Voting

Once the website is assembled, members of the class will be able to vote on their top three projects and winners will be announced.