HW. 5 Ch121 Atomic Simulations
Winter 2017
TA: Daniel Brooks (daniel.brooks@caltech.edu)

Purpose Of This Homework

In this homework assignment we are going to become familiar with a common open source MD code known as LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator). We will use LAMMPS to run molecular dynamics on a box of water which will give us a trajectory. Next we will utilize a code written in the Goddard group known as 2PT (Two Phase Thermodynamics) to calculate the thermodynamic properties of our system, from the trajectory.

As a challenge students will be given a polymer system which they will need to run dynamics on and then analyze.

This homework assignment should be thought of as the 'midterm' for the class, you will need to give a full lab report of your findings.

The water box calculation will be done to show you how the code works, you will then do your own analysis on the polymer system. Only the polymer system will be considered in the final write-up.

Materials To Be Reviewed Before The Lab Session

1. 2PT User Manual
2. Module Handout
3. Briefly Review the original 2pt publication 'The Two Phase Model For Calculating Thermodynamic Properties ...' by S-T. Lin, M. Blanco, W. Goddard (just read the basic introduction to get an understanding of the problem do not worry about the theory).

1 Building Our Water Box

To begin we need a box of water that we can run molecular dynamics on.

1. Open VMD (on your personal computer) and open the TK console (located in Extensions).

2. Once the console loads type the following
   % package require solvate

3. For our calculation lets study a system containing about 500 water atoms. To do this lets assume a density of $1 g/cm^3$ and type the following into the TK consoule.
   % solvate -minmax { { 0 0 0 } { a a a } }.
   Here a should be the side length of your box which will give about 500 atoms (do not forget the two sets of brackets, and the spaces between the 0's and a's).

4. Now you should see a box of water in your VMD window please record the actual number of atoms, and the number of water molecules within your system.

5. Now we need to save our file to send to the cluster. In the VMD main window (that has the file, molecule, graphics..tab) click on your water box and then right-click your system and select save coordinates. Set the selected atoms to all and choose bgf as the file type, therefore the file should be called waterbox.bgf

6. Now that we have constructed a system we need to send it to the cluster to perform our calculations. Please use whatever software you want to send this file to your account on atom (mobaxterm, winscp, filezilla ....). As always make a new directory for this homework assignment.

7. Inside your 2pt homework directory make a directory called structure and place your bgf file there. When VMD prepares the bgf format it makes some atom typing mistakes that we need to fix for running a LAMMPS calculation (if you try submitting the original file you will get an error). To fix these run the script in /ul/sflynn/scripts directory called fixbgf.pl, the path to this script is /net/hulk/home3/sflynn/scripts/fixbgf.pl Therefore you could type
   % /ul/sflynn/scripts/fixbgf.pl waterbox.bgf
   Notice you should now have 2 files, the file with the Tilda is the original input file, and the corrected input file will now be called waterbox.bgf

8. We are now going to make the input file for our LAMMPS calculation. Make a new directory (at the same level as the structure) called lammps. Copy the corrected waterbox.bgf structure to this lammps directory. Running Lammps is a very common task in the group, therefore a script already exists to make the proper input files, written by Tod Pascal (/ul/tpascal). Type the following command into the terminal to make the lammps input file. This should all be one command, use tab complete to help you in typing this out and then press enter.

   % /ul/tpascal/scripts/createLammpsInput.pl -b waterbox.bgf -s waterbox -f "/ul/tpascal/ff/WAT/spcew.ff"

   As the script runs it will will write some notes to terminal about what it's doing, there should be something like 7 steps that complete. This script should give you multiple files: data.waterbox, in.waterbox_singlepoint, waterbox.lammps.pbs, in.waterbox.

9. Now that we have a LAMMPS file we are going to make some changes to set up our specific trajectory calculation (please feel free to ask what these specific changes do). For all of these changes be sure to keep the same spacing format as the remainder of the in.waterbox file. Open the in.waterbox file and type
   % log ${sname}_npt.log
   This should go just before the
   print
   print ======================
   print "NPT dynamics with an isotropic pressure of 1 atm"
   print ======================
   print

10. Inside the NPT dynamics section make the following changes as well:
    dump 1 all custom ~~5000~~ 500 $sname.npt.lammps id type xu yu zu vx vy vz
    run ~~7500000~~ 50000 # run for ~~15~~ 0.1 ns

11. Finally at the end of the file add the following lines:
    kspace_style ewald 1e-04
    compute atomPE all pe/atom
    compute atomKE all ke/atom
    variable atomEng atom c_atomPE+c_atomKE
    dump 4 all custom 1 ${sname}_2pt.atom.eng id v_atomEng
    run 0
    undump 4
    uncompute atomPE
    uncompute atomKE

12. We have now completed our input lammps file to run a molecular dynamics calculation on this waterbox. All that is left is to submit this calculation to the cluster. To do this lets make another directory called lammp_pbs. Inside this directory copy over the data.waterbox file, and the in.waterbox files to this new directory. The other files (waterbox.bgf, in.waterbox_singlepoint, and waterbox.lammps.pbs) are not needed for the calculation we are running and can be removed if you want.

13. In the lammps_pbs directory make a new file called waterbox.pbs

    In this file write the following lines:
    #!/bin/csh
    #PBS -l walltime=500:00:00
    #PBS -l nodes=1:ppn=2
    #PBS -l pmem=4GB
    #PBS -N waterbox
    #PBS -q workq
    cd $PBS_O_WORKDIR
    module load use.own
    module load saladimodules
    module load openmpi-x86_64
    module load lammps
    mpirun -np 2 lmp -in in.waterbox

    Your lammps_pbs directory should now have a data.waterbox, in.waterbox, and waterbox.pbs.

    At this time if you have not loaded the modules Shyam maintains, please go do this (the calculation will not work otherwise). To be clear all you need to do is return to your main directory (type cd). Then load use.own, and the private modules, once you do these two things you are all set!

    Submit your calculation and wait for the results.

    The first line of the submission script tells the computer what language should be used, our cluster uses a csh environment (.cshrc). Line 2 defines how long the job should be run for, it is good form to give a wall time that is reasonably close to how long your job should take, however, we will ignore good form. Line 3 defines how many compute nodes we want to use for this calculation (in this case 1 node) and ppn is the number of processors (here we use 2). These lines can be changed depending on the work you are doing, lammps allows for parallel computing (using multiple nodes/processors) which can greatly decrease the time it takes for a calculation to run. Line 4 sets the memory to be allocated to this job. Line 5 sets the name for the job, this name will appear when you use qstat after submitting the job to the cluster. Line 6 specifies which queue to submit the job to (our cluster only has one Q so this is not necessary for our work). Line 7 Sets the directory to your current directory (therefore the calculation will output all of its results to the directory where you wrote the qsub file). The four module lines specify the lammps module that Shyam Saladi has created. Lammps is an opensource code meaning it changes periodically, a module allows one person to maintain the software and everyone else just chooses to run whatever software the module has. This allows us to avoid worrying about having each person compile their own lammps. Finally the last line says to use mpirun, a software that allows lammps to use parallel computing, it then says how many processors mpi should use, it calls on the lamps package (lmp) and then specifies the lammps "in file".

2 2PT Calculations On Water

Now that we have a lammps trajectory we may want to analyze it. There are various ways and methods to do this, depending on what you are trying to understand. For this assignment we are going to use a code and methodology written in the group known as 2PT.

1. Inside of your lammps_pbs directory make a new directory called 2pt, and then copy over your data.waterbox file, and the waterbox.2pt.lammps trajectory.
   2pt is a stand alone code that can be used to derive thermodynamic data from a molecular dynamics trajectory (which we calculated in LAMMPS). The software was written in the Goddard group and used extensively by /ul/tpascal. We will use the code Tod has compiled on his account to run our calculation.
   To run our 2pt calculation we will need an in file for 2pt (waterbox_2pt.in), a lammps trajectory (waterbox.2pt.lammps), the lammps data file (data.waterbox), and a submission file (2pt.pbs).

2. We need to make an input file for the analysis, make a file called waterbox_2pt.in
   Inside of the file add the following lines.
   IN_LMPDATA data.waterbox
   IN_LMPTRJ waterbox.2pt
   ANALYSIS_FRAME_INITIAL 1
   ANALYSIS_FRAME_FINAL 0
   ANALYSIS_FRAME_STEP 1
   ANALYSIS_VAC_CORLENGTH 0.5
   ANALYSIS_VAC_MEMORYMB 5000
   ANALYSIS_VAC_2PT 3
   ANALYSIS_OUT waterbox_2pt
   ANALYSIS_LMP_TSTEP 0.001
   ANALYSIS_VAC_LINEAR_MOL 0
   ANALYSIS_VAC_ROTN_SYMMETRY 2
   ANALYSIS_VAC_FIXED_DF N

   Save the file. Here the last line N is the number of atoms in your system (look at the beginning of your data file). Please look at the User Manual provided by Tod to see what commands are available for this software.

3. Now make your submission file, call it 2pt.pbs. Write the following lines inside this file.
   #! /bin/csh
   #PBS -l nodes=1:ppn=1,walltime=72:00:00
   #PBS -q workq
   #PBS -N wbox_2pt
   cd $PBS_O_WORKDIR
   tpascal/codes/bin/md_analysis waterbox_2pt.in

4. We are ready to submit this thermodynamics analysis, submit your job to the cluster (it should take a few minutes to calculate).

3 2PT Analysis Of Water

After running this calculation please take a look at each file that is generated (use the manual from Tod to help you understand what is happening). Specifically look at the thermodynamics file and make sure you understand everything that we can learn about the system from doing a 2pt calculation.

This is the end of the formal lab session material, the remainder of the homework assignment should be completed and discussed in your write-up.

4 Homework: Polymer Calculations

Now that we have completed a detailed example of how to analyze an MD system using 2PT, I would like for you to run a 2PT calculation on a polymer system. I will provide you with a lammps in file and associated data file for this calculation. It is your job to run a lammps calculation on this system to get a trajectory for the polymer chains. Hint: get a trajectory size similar to the water case, don't run it for longer. Another hint: never try to open a trajectory with vi, it is too large, instead use the less command (if this does not make sense ask or google). Once you have your trajectory determine the entropy of the entire polymer system using 2pt just as we did for the water.

Do you think it would be reasonable to assume that each polymer chain has about the same entropy? To test this run another calculation where you define each polymer chain as a group (to do this you will need to refer to the 2PT user manual).

A final comment do not make multiple copies of your trajectory file, your accounts have a limited amount of space. Instead simply move the trajectory file to wherever you want to run your calculation.

So you need to get a trajectory for your polymer system, then analyze this trajectory using 2pt. Once you do this analysis set-up another 2pt calculation (using the same trajectory) where you define each polymer chain as a group and determine the thermodynamics of each chain).

Things to Consider

What atoms are in your polymer, what is this a polymer of? How many chains are making up the polymer you are studying in this simulation, how long are the chains? How do you cap the chains (i.e. how does a chain start and end). What is the charge on the atoms in your polymer, are they reasonable? Do some research, why should we care about polymers? What are the differences between a polymer electrolyte and a liquid electrolyte in the context of batteries?

5 Homework Polymer Write-Up

This is the final homework assignment for the class, the remainder of the course will be spent working on your individual research projects. Therefore it is crucial that you have identified a project you would like to work on, and get in contact with people in the group who will be able to assist you with the project. Remember Bill is a key resource for all of this, he will be able to provide advice, and to help put you in contact with people in the group. If for some reason you are unable to get in touch with group members Bill suggests, send Bill and email about it!

For this homework assignment I am expecting a formal write-up of your results, think of this as a lab report. Be warned this should be considered the "midterm" for the class, and should take you longer to complete than any of the previous reports. The final for the class will be a formal write-up of your original research, this assignment will be used to make sure you understand the type of information that is necessary for presenting (and potentially publishing) research in computational chemistry.

This report should consist of:

1. Abstract
2. Introduction and Methods
3. Results and Conclusions

Below are some general comments meant to be suggestions, please make the report your own!

Abstract

A brief review of the problem, the system, and the methods of analysis. Then provide key results. Make it short and concise.

Introduction and Methods

This is meant to be a brief review about polymers, not a summary of the 2PT formalism! You are welcome to give a few equations and some comments about the formalism, however, this is already published work and should be briefly summarized and cited (be warned it would take some non-equilibrium statistical mechanics and solid state physics to truly understand the methodology). Talk about the system we are studying, why it may be an important system, and why these types of calculations would be useful.

Results and Conclusions

Make sure your system and model are well defined for the reader. Interpret your calculations and give some analysis of why the results are such. Make this your own, do some research on polymer batteries, and the useful application this type of research has in regards to a "greener" energy supply.