Reflection on "Coarse-Grained Simulation of DNA using LAMMPS,
An implementation of oxDNA model and its applications"

This paper mainly introduced the process of using Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) to simulate the molecule dynamics of oxDNA. In the final part, it also presented the results on the behavior and relevant applications.

We have covered the information about molecule dynamics and have experience in LAMMPS simulation from the class homework #4. In our homework #5, we also simulated the behaviors of a gold nanowire under the simple tension test in MATLAB. We defined the crystal structure of gold in 0K which is FCC packing. Then we calculate the interactive force of two atoms by Lennard-Jones potential equation which is only function of the distance between two atoms. By summing the interactive force between one atom and all other atoms. The net force of an atom in 3 directions can be calculated. Next, we can get new coordinate system for all atoms and repeat the force calculation. This is the basic logic for our simulation. In this paper, the setting for most parts are similar with what we did in our homework. But it still has some difference.

Firstly, this paper have discussed about the oxDNA model. The structure of oxDNA model is not simple FCC as what we did in homework. The oxDNA has complex stacking which will generate more bonding and interactive forces. The overall structure is shown in Fig. 1. It contained some bonding or interactions in oxDNA such as phosphate backbone connectivity and excluded volume, hydrogen-bonding, stacking, cross-stacking and coaxial stacking interaction. These will have impact on the set up for LAMMPS simulation. In LAMMPS code library, the oxDNA model with specific sequence interaction has already created. In this paper, the authors directly used them as the input file.

For the force and torque calculation, the means of quaternions are used instead of the Euler angles. The pair potential energy can be determined which is function of $r, \hat{a},$ and $\hat{b}$ where r is the norm of the relative distance vector between two rigid bodies A and B, $\hat{a}$ and $\hat{b}$ are the principal axes of these two rigid bodies. Then the force and torque can be straightforwardly calculated from potential energy. In this simulation, the potential energy used is Lennard-Jones potential which is exactly what we used in homework.

The authors also gave some specific information about the input file, data file, output and visualization. In this simulation, the model structures are imported from the library and the coefficient of bonding and interaction for each pair are also defined in their input file. In data file, they did the similar thing as we did in MATLAB simulation set up. They defined the dimensions of the simulation box and arranged the atoms in a certain order within the simulation box. They also defined the coordinate positions of ellipsoids and set initial velocity all equal to 0. The last thing is to define the bond information by connecting the atom-ID defined before. To simplify the setup procedure, a setup tool in USER-CGDNA package is used as well. For data output, the coordinate position x, y and z of atoms are written in a text-file. Additionally, the velocity, force and torque are also printed out in the output file. The rotational, kinetic and potential energy are also included in this output file. Since LAMMPS doesn't contain a direct visualization toolkit. They are using VMD (Visual Molecular Dynamics) to generate the motion history and some images for the behaviors of oxDNA.

In the next section, the authors introduced the Langevin-type rigid body integrators. Since only a limited choice of suitable Langevin integrators can be selected in LAMMPS. The DOT

integrator is an alternative choice instead of standard LAMMPS NVE integrator for aspherical particle. This setting can be changed in the input file. This integrator is useful for analyzing the accuracy of the pair interactions at a given timestep. Fig. 2 shows the total energy comparison between the aspherical integrator and DOT integrator. The aspherical integrator is based on the Richardson iteration in the update of the quaternion degrees of freedom while DOT used a rotation sequence. From the figure, in a certain timestep size, the DOT integrator is not accurate and stable than standard integrator. However, with a longer LJ time unit length, the difference will tend to disappear. For the Langevin dynamics, we used an estimate based on the average kinetic, rotational, potential and total energy of the benchmark. The results are shown in Fig. 3. Typically, the error when the timestep size $\Delta t = 2 * 10^{-2}$ is still low for DOT-C integrator. This timestep size is about 4 times larger than the maximum timestep size for standard LAMMPS Langevin integrator. By using DOT-C integrator, we can increase the computational efficiency by 400%.

In this project, it performed several LAMMPS implementation with a few simple benchmarks with different density. The low-density benchmarks was formed by a 10*10 array of duplexes and it's giving a total of 60 kbp. The high-density one was formed by a 40*40 array and it's giving 960 kbp totally. The results of parallel speedup for both benchmarks relative to the single node performance with 24 MPI-task are shown in Fig. 4. From the figure we can tell, the low-density benchmark performed well up to about 128 MPI-tasks with a high parallel efficiency. With the increase of MPI-tasks, the performance is becoming worse and worse, and the parallel efficiency is deceasing as well. The authors believed that the degradation happened because of the limit ratio between ghost atoms and local atoms. LAMMPS is known to require at least a few hundred local atoms or more for a good parallel performance. In this section, authors also spent time on discussing the performance of 3SPN.2 as well.

In the last section of this paper, it listed some applications from structural properties of DNA. Specifically, it presented the results obtained from Langevin dynamic simulation of the two ssDNA sequence at a certain conditions. By using the linear DNA molecules under a certain conditions, the system will converge from a random initial conditions to a new steady state. When the steady state is reaching, the gyration radius R_g and the number of base-pairs N_C will tend to be a constant. Fig. 5 shows the gyration radius of λ-DNA and poly-A and poly-T strands which are different linear DNA molecules of the same length. From the figure we can tell, when the temperature increasing, the radius of gyration became same for all sequences. However, with the decrease of temperature, the difference of gyration radius between A nucleotides and T nucleotides became larger and larger. This is because the stacking strength of these two strands are different. But the average should be very close to the final gyration radius which is shown in Fig 5 as well. In next part, authors presents some images of the λ-ssDNA sequence structures under different temperatures and discussed the different behaviors of that λ-ssDNA. Additionally, a plasmid cloning vector pUC19 is also introduced in this section. Since it's only a two pages summary, I will leave the remaining part and stop from here.

Generally, this is a great paper that demonstrated the implementation of the oxDNA model by using LAMMPS. It gave some details about the setup in LAMMPS even how to change the setting conditions by using the USER-package. The overall strategy used in this LAMMPS simulation is pretty similar with what we did in the homework. I think it helped me reinforce the stuff learned relative to the LAMMPS.

All figures are from: “Coarse-Grained Simulation of DNA using LAMMPS:An implementation of the oxDNA model and its applications,” by Oliver Henrich, Yair Augusto,Gutierrez Fosado, Tine Curk,and Thomas E. Ouldridge; arXiv:1802.07145v2 [cond-mat.soft] 7 May 2018.

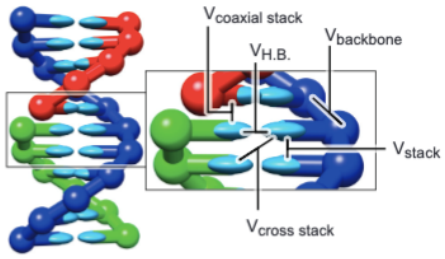

Fig. 1 Overview of oxDNA structure

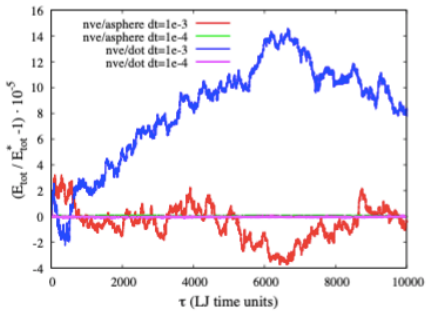

Fig. 2 Relative normalized accuracy between aspherical particle integrator and DOT integrator

	Δt	E_{kin}	E_{rot}	E_{pot}	E_{tot}	standard error of E_{tot} fit
fix nve/asphere & fix langevin						
	10^{-4}	2.3999	2.4001	-21.4512	-16.6513	± 0.00377 (0.0227%)
	10^{-3}	2.4015	2.4021	-21.5564	-16.7582	± 0.00349 (0.0208%)
	$5 \cdot 10^{-3}$	2.4012	2.3999	-21.6352	-16.8315	± 0.00322 (0.0191%)
nve/dotc/langevin						
	10^{-4}	2.3989	2.3997	-21.5278	-16.7292	± 0.00362 (0.0216%)
	10^{-3}	2.3998	2.4008	-21.6631	-16.8624	± 0.00335 (0.0199%)
	10^{-2}	2.3959	2.3941	-21.6151	-16.8251	± 0.00318 (0.0189%)
	$2 \cdot 10^{-2}$	2.3895	2.3752	-21.6266	-16.8619	± 0.00313 (0.0185%)

Fig. 3 Results of energy for different integrators

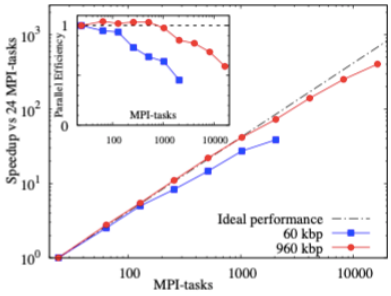

Fig. 4 Speedup of low and high density benchmarks

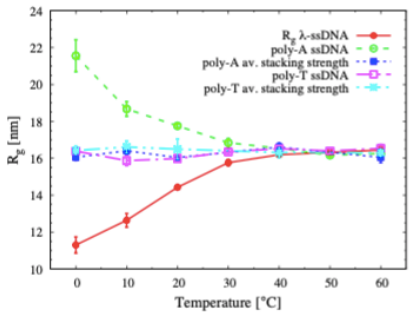

Fig. 5 Response of the radius of gyration