The Competitive λ-Dynamics

We plan to explore the λ-dynamics (LD) method and its applications to simple Lennard-Jones (LJ) particles and realistic proteins. We hope that investigating a simple model system like the LJ fluid would help to bring up some generalization or improvement over the method, which could afterwards be applied to the more complicated case of protein-ligand interactions.

I. APPLIED TO THE LENNARD-JONES PARTICLES: A MODEL STUDY

A trick of the LD method devised originally by Kong and Brook is that the candidate ligands are bound together to form the hybrid. The components of the hybrid don't interact with each other, but since they are in the same space, the existence of a component ligand somehow sets hurdles to the others in space, even though it has diminished to a negligible fraction (the centers of their atoms can be thought of as some sort of singularities in space which other atoms can never approach).

These singularities are not trivial since they are normally close to the champion ligand which are bound together with others. It remains unclear how their existence will affect the competition dynamics and the sampling of phase space.

In this section we present a model study which doesn't have such a problem. The model is a LJ solute in the LJ solvent. Let's assume that we would like to examine which kind of solute atom is most solvable in a certain solvent. The solute is assumed to be composed of different component of atoms and hence all of the candidate atoms automatically have the same coordinate thus there exists no spatial interference as we mentioned above about the molecular case. With this prototype we can actually check the effectiveness of the LD method.

A. The methodology

Consider a solute in a LJ system. The solvent particles interact among themselves with a LJ potential $\phi_{\epsilon,\sigma}(r)$. The solute particle interacts with its kins a LJ potential $\phi_{\epsilon_l,\sigma_l}(r)$. Now suppose we care about the free energy of solution of different types of atoms, which are actually characterized by different values of ϵ and σ.

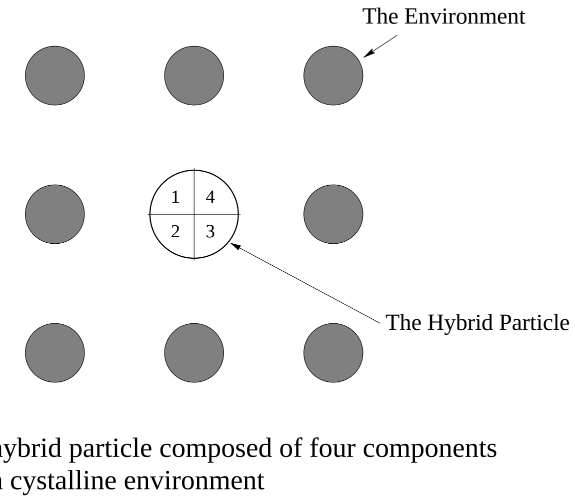

FIG. 1. A four-component hybrid Lennard-Jones particle in a cystalline environment.

We first build a hybrid particle H which consists of a group of different kinds of atoms with weights assigned to each type of atom. Let's suppose we have the following Lagrangean

$$\mathcal{L} = \frac{1}{2}\sum_{l=1}^{L} M_l \dot{\lambda_l}^2 + \frac{1}{2}\sum_{i=1}^{N} m v_i^2 + \frac{1}{2}\sum_{l=1}^{L} \lambda_l^2 m_l v_H^2 - U(\{r_1, \cdots, r_N, r_H\}, \{\lambda_1, \cdots, \lambda_L\}) \tag{1}$$

where m_l are the masses of the real solutes. The third term in the above equation is introduced to take into account the effect of mass on atomic motions. The potential is

$$U(\{r_1,\cdots,r_N,r_H\},\{\lambda_1,\cdots,\lambda_L\}) = U_0 + \sum_{l=1}^{L}[V_l(r_H) - V(r_H)]\lambda_l^2 \tag{2}$$

where

$$U_0 = \frac{1}{2}\sum_{i,j,i\neq j}^{N+1}\phi_{00}(r_{ij}) \tag{3}$$

This potential is obtained by summing ϕ_{00} over all the atoms including the hybrid as if the hybrid were just a solvent atom.

$$V_l(r_H) = \sum_{i=1}^{N}\phi_{l0}(|r_i - r_H|) \tag{4}$$

$$V(r_H) = \sum_{i=1}^{N}\phi_{00}(|r_i - r_H|) \tag{5}$$

$V(r_H)$ can be regarded as a reference potential. The λ coordinates satisfy the normalization condition

$$\sum_{l=1}^{L}\lambda_l^2 = 1 \tag{6}$$

and the λ velocities obey

$$\sum_{l=1}^{L}\lambda_l\dot{\lambda}_l = 0 \tag{7}$$

The Lagrangean equations of motion can be written as

$$M_l\frac{d^2\lambda_l}{dt^2} = -2(V_l - V)\lambda_l + (2\gamma + m_l V_H^2)\lambda_l \tag{8}$$

$$m_i\frac{d^2 r_i}{dt^2} = -\nabla_{r_i}U_0 - \sum_{l=1}^{L}\lambda_l^2(\nabla_{r_i}V_l - \nabla_{r_i}V) \tag{9}$$

where γ are the Lagrange multipliers introduced in order to satisfy the holonomic constraint in eq.(6) (It forces the second order derivative of the holonomic constraint to be zero)

$$\gamma = \frac{\sum_l\left[2(V_l - V)\lambda_l^2 - m_l V_H^2\lambda_l^2 - M_l\dot{\lambda}_l^2\right]}{2\sum_l\lambda_l^2} \tag{10}$$

Note that in eq.(8) the driving force $V_l - V$ is time-dependent. When $V_l < V$, which means that the l-component of the hybrid becomes hydrophilic, the acceleration of λ_l becomes positive. On the contrary, $V_l > V$, the l-component of the hybrid becomes hydrophobic, as a result λ_l is decelerated. Therefore we anticipate that the most hydrophilic component will win out in the competition.

The numerical maintaneance of the holonomic condition eq.(6) constitutes an important part of the LD algorithm. In a simulation it may turn out that the holonomic condition cannot be numerically satisfied using only the second-order derivative constraint eq.(10), namely the Lagrange multiplier γ. In such a case, eqs.(7) and (6) have to be introduced to control the propogation of numerical error. And this kind of control is normally a rescaling scheme, which simply amounts to introducing some sort of heat bath.

The code was first examined by building a hybrid of which five components have actually the same LJ parameters and which is therefore nothing but another solvent particle. The code passed this test—it is observed that the λ coordinates remain exactly constant in the dynamics simulation.

Second, we tested the code by putting a hybrid atom which consists of five components having different ϵ but identical σ and mass into the solvent and run the dynamics at a temperature under the melting point. Again it was observed that the code produced reasonable qualitative results.

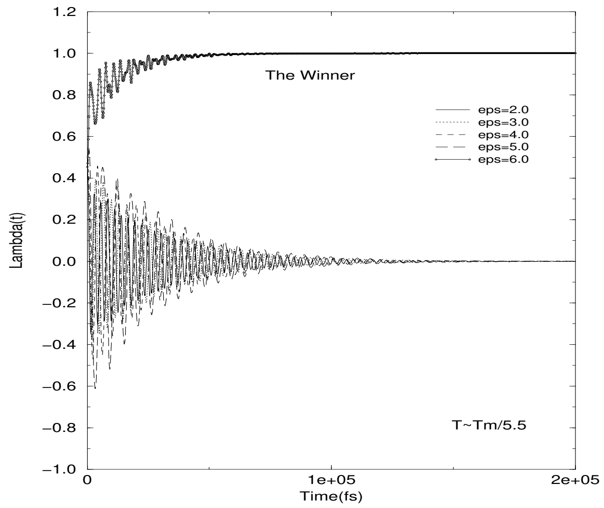

FIG. 2. The trajectories of the λ coordinates show that only the most favored component survives after a long time competition.

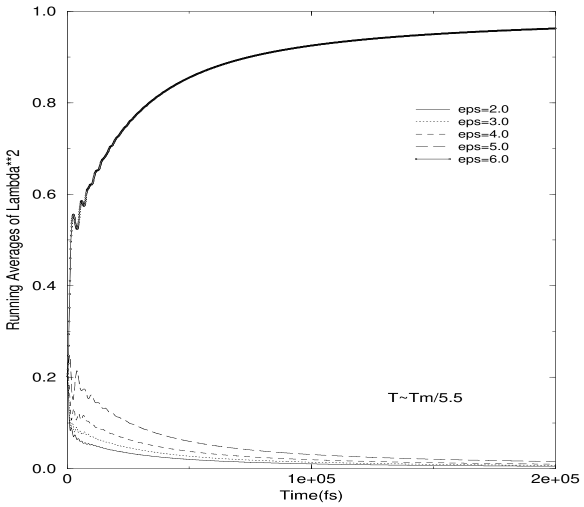

FIG. 3. The running averages of the λ coordinates. The initial composition of the hybrid: $20\%(\epsilon = 2.0)+$ $20\%(\epsilon = 3.0)+$ $20\%(\epsilon = 4.0)+$ $20\%(\epsilon = 5.0)+$ $20\%(\epsilon = 6.0)$.

Third, we built a hybrid composed by five types of atoms having the same ϵ and σ identical to those of the solvent atoms but different masses. In such a limit, we expect from the following analytical formula that the heaviest species will win the competition

$$F_l(T) = -k_B T \ln \left[\int d^N r dr_H \exp(-U/k_B T) \right] + 3Nk_B T \ln \sqrt{h^2/2\pi m k_B T} + 3k_B T \ln \sqrt{h^2/2\pi m_l k_B T} \tag{11}$$

and the free energy difference of two species with different mass reads

$$F_a - F_b = 3k_B T \ln \sqrt{m_b/m_a} \tag{12}$$

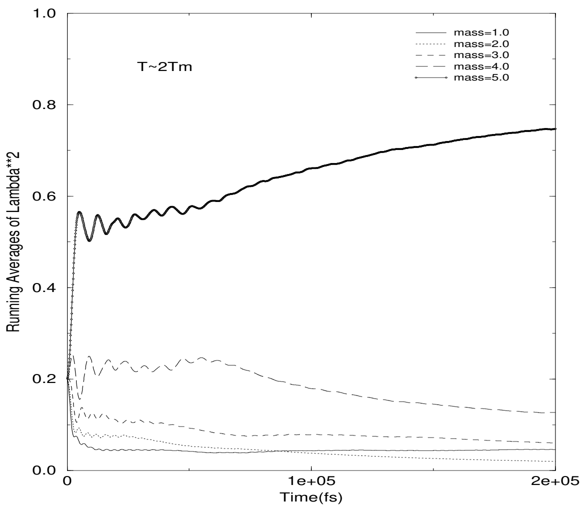

FIG. 4. The running averages of the λ coordinates. The initial composition of the hybrid: 20%($m_1 = 1.0$)+ 20%($m_2 = 2.0$)+ 20%($m_3 = 3.0$)+ 20%($m_4 = 4.0$)+ 20%($m_5 = 5.0$). The LJ parameters of the five components are the same, and identical to those of the solvent atoms.

Fig. 4 shows that our simulation gives qualitatively correct results.

B. Free energy

The partition function of the hybrid system can be written by

$$Z(\{\lambda\}, r_1, \cdots, r_N, r_H) = \int d^N r dr_H \exp\left\{ -\beta \left[U_0 + \sum_{l=1}^{L} (V_l(r_H) - V(r_H)) \lambda_l^2 \right] \right\} \tag{13}$$

The free energy difference between two solutes, a and b, can be obtained by

$$\Delta\Delta F_{ab} = -k_B T \ln \int dr^N dr_H \exp\left[-\beta(U_0 + V_a - V)\right] + k_B T \ln \int dr^N dr_H \exp\left[-\beta(U_0 + V_b - V)\right] \tag{14}$$

It can be further written as

$$\Delta\Delta F_{ab} = -k_B T \ln \frac{Z(\lambda_a = 1, \{\lambda_{l \neq a}\} = 0)}{Z(\lambda_b = 1, \{\lambda_{l \neq b}\} = 0)} \tag{15}$$

A major technical problem is the choice of the fictitious masses. The fictitious masses affect the dynamics of the λ coordinates and therefore that of the atomic coordinates and velocities. It is certain that the system will undergo different dynamics with different choice of fictitious masses. The question is wether or not the system will therefore sample a part of the phase space which is only accessible by a specific choice of masses. This will, if it is the case, result in the mass-dependence of the results, which subsequently pose a challenge to the LD approach.

C. The precision of the λ-dynamics

The following problems have to be addressed. First, we know other methods like the thermodynamical integration method has a prediction accuracy. Namely, the free energy difference shouldn't be smaller than an energy difference normal thermodynamical fluctuations can yield. So how small a free energy difference could the LD method probe? We call this problem the accuracy of the LD method.

Second, the inventors of the LD method claimed that a relatively short time of simulation is required by the approach to find out the best candidate. However, there is no criterion for how long it will take for the λ coordinates to converge, even if there exist no local minima to trap them.

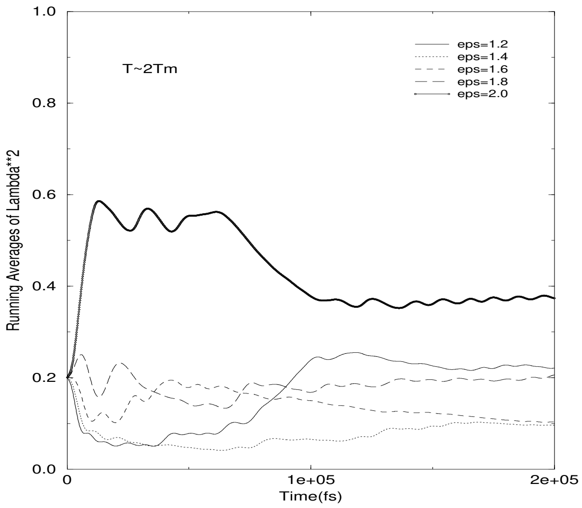

FIG. 5. The running averages of the λ coordinates. The initial composition of the hybrid: 20%($\epsilon = 1.2$)+ 20%($\epsilon = 1.4$)+ 20%($\epsilon = 1.6$)+ 20%($\epsilon = 1.8$)+ 20%($\epsilon = 2.0$). This figure reveals the problems of the LD method.