

A General Long-Time Molecular Dynamics Scheme in Atomistic Systems: Hyperdynamics in Entropy Dominated Systems

Xin Zhou and Yi Jiang

Los Alamos National Laboratory, Los Alamos, NM 87545, USA
xzhou@lanl.gov

Abstract. We extend the hyperdynamics method developed for low-dimensional energy-dominated systems, to simulate slow dynamics in more general atomistic systems. We show that a few functionals of the pair distribution function forms a low-dimensional collective space, which is a good approximation to distinguish stable and transitional conformations. A bias potential that raises the energy in stable regions, where the system is at local equilibrium, is constructed in the pair-correlation space on the fly. Thus a new MD scheme is present to study any time-scale dynamics with atomic details. We examine the slow gas-liquid transition of Lennard-Jones systems and show that this method can generate correct long-time dynamics and focus on the transition conformations without prior knowledge of the systems. We also discuss the application and possible improvement of the method for more complex systems.

1 Introduction

The atomistic molecular dynamics (MD) simulations are typically limited to a time scale of less than a microsecond, so many interesting slow processes in chemistry, physics, biology and materials science cannot be simulated directly. Many new methods, such as kinetic monte carlo [1], transition path ensemble methods [2], minimal action/time methods [3] have been developed to study the slow processes (for a review see [4]). They all require prior knowledge of the system, which is often hard to obtain, and they can only deal with a few special processes inside a small part of configurational space of the system.

In many systems, the interesting slow dynamics are governed by the infrequent, fast transitions between (meta-) stable regions; yet the systems spend most of their time in the stable regions, whose dynamics can be well described by some time-average properties. Hence an alternative approach to describing the long-time dynamic trajectory would be some suitable time propagator entering in/out the stable regions as well as the real short transition trajectory among the regions. In other words, we would coarse-grain the stable regions while keeping the needed details outside the stable regions. This natural coarse-graining technique is different from the usual method that average some degrees of freedom. The averaging process in the latter method usually changes the dynamics, although the static properties of systems might remain correct.

Hyperdynamics, originally developed by Voter [5], is an example of such a coarse-graining method. The hyperdynamics method treats the potential basins as the stable configurational regions that are separated by saddle surfaces. A bias potential is designed to lift the potential energy of the system in the basins, while keeping the saddle surfaces intact. Dynamics on the biased system lead to accelerated evolution from one stable region to another. Based on transition state theory (TST), the realistic waiting time t_{real} before a transition from the basins can be re-produced statistically:

$$t_{real} = \Delta t \sum_i \exp(\beta \Delta V(r_i)), \quad (1)$$

where Δt is MD time step, and $\Delta V(r_i)$ is the bias potential at the conformation r_i , $\beta = 1/k_B T$, k_B is the Boltzmann constant and T is the temperature of the system; r refers to the $3N$ -dimensional conformation vector throughout this paper. The method has been applied successfully to systems in which the relevant states correspond to deep wells in the potential energy surface (PES), with dividing surfaces at the energy ridge tops separating these states. This is typical of solid-state diffusion systems [6]. However, in systems where entropy is not negligible, the basins of PES do not completely correspond to the long-time stable regions. Hyperdynamics cannot readily apply. An extreme example is hard sphere systems: all physical permitted conformations have the same zero potential energy, but some conformations, which correspond to transition regions among stable regions, are rarely visited. A complication occurs even when applying hyperdynamics in solids with fairly clearly defined stable regions: after applying the bias potential, the energy landscape becomes much flatter and the system can start to have entropic-like characteristics. These effects limit the improvement in the simulation rate that can be achieved by the hyperdynamics method over the direct MD approach [5]. Thus, although there have been some attempts [7,8] to apply hyperdynamics to enhance conformational sampling in biological systems, generally, accurate slow dynamics or kinetics can only be expected for relatively simple or low-dimensional systems.

Here, we derive a more general hyperdynamics method that can be used to access very long (possibly all) time scale in fluids where both entropy and potential energy are important. A description of part of this method has appeared in [9]. We first present explicit conditions for applying this method without using TST. Then we give some possible collective variables for identifying the stable and transition conformations to design suitable bias potential. We then examine the performance of this method in simple fluids. Finally we discuss the further development and application of this method in more general (and complex) systems.

2 Theory

We begin this process by introducing a time-compressing transformation,

$$d\tau = a(r)dt, \quad (2)$$

where $d\tau$ is a pseudo-time step, dt is the real time step, and the local dimensionless compression factor is given by a conformational function $a(r)$, which is ≤ 1 . Thus the trajectory $r(t)$ can be rewritten as $r(\tau) = r(\tau(t))$ in a shorter pseudo-time interval τ , and we have,

$$\tau = \int_0^t dt' a(r(t')) = t \int dr D(r; r(t); t) a(r), \quad (3)$$

where $D(r; r(t); t)$ is the density probability of $r(t)$ in the conformational space in interval $[0, t]$,

$$D(r; r(t); t) = \frac{1}{t} \int_0^t dt' \delta(r - r(t')) \quad (4)$$

The compressed trajectory $r(\tau)$ will satisfy a new equation of motion [9]. If we use the usual Langevin equation to simulate the evolution of system in NVT ensemble, the new equation is,

$$\begin{aligned} \frac{d}{d\tau} R_i &= \frac{\mathcal{P}_i}{M_i} \\ \frac{d}{d\tau} \mathcal{P}_i &= -\frac{\partial U(r)}{\partial R_i} - \zeta'(r) \mathcal{P}_i + f_i(\tau) - \sum_j \mathcal{L}_{ij} \frac{\partial \Delta V(r)}{\partial R_j}, \end{aligned} \quad (5)$$

where R_i is the i th component of conformational vector r , $M_i = m_i a^2(r)$ is the pseudo-mass of the particle with the real mass m_i . It is an equation of motion of particles with smaller conformation-dependent mass $M(r)$ under new potential $U(r) = V(r) + \Delta V(r)$ where $\Delta V(r) = -k_B T \ln a(r)$, as well as new friction coefficient, if we neglect the zero ensemble-average term

$$\mathcal{L}_{ij} = \frac{1}{k_B T} \frac{\mathcal{P}_i \mathcal{P}_j}{M_j} - \delta_{ij},$$

where δ_{ij} is the kroneck δ -symbol. Similarly, the zero-mean white noise friction force $f(\tau)$ satisfies the fluctuation-dissipation theorem of the new system,

$$\langle f_i(\tau) f_j(\tau') \rangle = 2k_B T M_i \zeta' \delta_{ij} \delta(\tau - \tau'). \quad (6)$$

where $\zeta'(r) = \zeta/a(r)$ and ζ is the friction coefficient of the original system.

At the first glance, it would not appear to be advantageous over directly generate $r(\tau)$ from (5), as very short time steps are necessary due to the small mass M . However, if we only focus on the long-time dynamics, we can use a smoother pseudo-trajectory $\mathcal{R}(\tau)$ to replace $r(\tau)$, provided that the reproduced time from $\mathcal{R}(\tau)$ is the same as that from $r(\tau)$. Thus, a sufficient condition to replace $r(\tau)$ with $\mathcal{R}(\tau)$ is that their density probability are the same. Actually, if the compressed factor is defined as function of some collective variables, notated as X , rather than that of r , a simpler condition is,

$$D(X; \mathcal{R}(\tau); \tau) = D(X; r(\tau); \tau), \quad (7)$$

where $D(X; \mathcal{R}(\tau); \tau)$ and $D(X; r(\tau); \tau)$ are the density probability of the trajectory $\mathcal{R}(\tau)$ and $r(\tau)$ in the X space, respectively.

In the time-consuming regions (stable regions), similar conformations would be visited repeatedly many times even during a finite simulation time. Thus we can assume the distribution can be approximated by $D(r; r(t); t) \propto \exp(-\beta V(r))$. Many methods might be used to generate $\mathcal{R}(\tau)$ with the required distribution. One of them is to use a realistic trajectory corresponding to the local equilibrium of a new potential $U(r) = V(r) - k_B T \ln a(r)$. Actually, if we select $a(r) < 1$ only in the potential wells, we effectively have the hyperdynamics presented by Voter [5].

In many cases, some collective variables X can be used to identify transition conformations. The transition conformations related to slow dynamics always locate in some regions of X space, so we can select $a(r) = a(X(r)) < 1$ outside these regions that involve transition conformations, thus the new potential is $U(r) = V(r) - k_B T \ln a(X(r))$. For example, in polymers, the transition regions of slow conformational transitions can be identified in the soft torsion angles space. We require that the density probability of pseudo-trajectory in the collective-variable space equals to that of the compressed trajectory. Thus, if we use a bias potential to generate the pseudo-trajectory, a simple design of the bias potential is

$$\Delta V(r) = k_B T f_+(\ln(D(X(r))/D_c)), \quad (8)$$

where $f_+(y) = y\Theta(y)$, $\Theta(y)$ is the step function. $D(X)$ is the density probability of a segment of trajectory, and D_c is a pre-selected threshold value. The design means we compress the trajectory to make the density probability reach D_c if the original density is larger than the value. We can repeat the biasing process: simulating a segment of trajectory to get a bias potential, then simulating another segment of trajectory under the designed bias potentials to get another bias potential. The sum of all bias potentials forms the total bias potential to reach (almost) any time scale. In practice, we also add a correction the definition of $f_+(y)$ near $y = 0$ to get continuous bias force.

The key to apply hyperdynamics successfully is to design suitable bias potentials $\Delta V(r)$. Obviously, $\Delta V(r)$ should have the same symmetry as $V(r)$. For a simple case of N identical particles, the conformational vector $\{\mathbf{R}_j\}$ ($j = 1, \dots, N$) can be rewritten as a density field, $\hat{\rho}(\mathbf{x}) = \sum_j \delta(\mathbf{x} - \mathbf{R}_j)$. Here both \mathbf{x} and \mathbf{R} are the normal 3-dimension spatial vectors. Since the neighboring conformations are equivalent in the viewpoint of slow dynamics, $\hat{\rho}(\mathbf{x})$ can be averaged to get a smooth function, $\bar{\rho}(\mathbf{x})$, for example, by using a Gaussian function to replace the Dirac- δ function. If the width of the Gaussian function is small, $\bar{\rho}(\mathbf{x})$ can be used to identify different conformations. Another similar description is the k -space density field, $\hat{\rho}(\mathbf{k}) = \sum_i \exp(i\mathbf{k} \cdot \mathbf{R}_i)$. If neglecting the effects of multi-body correlations and directional correlations, we can approximate the density field $\bar{\rho}(\mathbf{x})$ to the radial pair distribution function $g(z)$, which is defined as,

$$g(z) = \frac{1}{4\pi\rho z^2 N} \sum_i \sum_{j \neq i} \delta(r_{ij} - z), \quad (9)$$

where $r_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$, or more exactly, some bin-average values of $g(z)$ along z ,

$$g_p = 2\pi\rho \int g(z)h_p(z)z^2 dz, \quad (10)$$

Where $h_p(z)$ is a two-step function, unity in a special z range (a_p, b_p) and zero otherwise. Thus, each conformation corresponds to a group of g_p (g vector), the spatial neighborhood of the conformation and their symmetric companions also correspond to the same g vector. If we select enough g_p , all conformations with the same $\{g_p\}$ are thought to be identical in the viewpoint of slow dynamics. Therefore, we can define the bias potential in the low-dimension g space, $\Delta V(\mathbf{R}^N) = \Delta V(\{g_p(\mathbf{R}^N)\})$. To better identify conformations with not-too-small bin size, we can add some important dynamics-related physical variables, such as, potential $V(\mathbf{R}^N)$, to the collective variable group. Another important variable is the two-body entropy s_2 , presented first by H. S. Green [10], defined as,

$$s_2 = -2\pi\rho \int [g(z) \ln g(z) - g(z) + 1]z^2 dz. \quad (11)$$

The two-body entropy, which forms the main part (85% – 95%) of macroscopic excess entropy, has been studied widely [11,12]. Actually, both g_p and s_2 are functional of $g(z)$, similarly, it is also possible to use some another functional of $g(z)$ to form the collective variables. In some special systems, it may be also useful to add some special order parameters O_q to take into account possible multi-body correlations. Finally, we have a group (of order 10) of general collective variables $X = \{X^p\}$, which might involve V , s_2 , some $\{g_p\}$ and some possible $\{O_q\}$, to identify conformations and form an appropriate bias potential $\Delta V(X(\mathbf{R}^N)) = k_B T f_+(\ln D(X)/D_c)$. The corresponded bias force on each particle can be calculated by the chain rule of differentiation,

$$\Delta \mathbf{f}_i = - \sum_p \frac{\partial \Delta V}{\partial X^p} \frac{\partial X^p}{\partial \mathbf{R}_i}. \quad (12)$$

For example, from the (9), we have,

$$\frac{\partial g(z)}{\partial \mathbf{R}_k} = \frac{1}{2\pi\rho z^2 N} \sum_{j \neq k} \hat{r}_{kj} \frac{\partial}{\partial z} \delta(r_{kj} - z) \quad (13)$$

where \hat{r}_{kj} is the corresponding unit vector of the distance vector $\mathbf{r}_{kj} = \mathbf{R}_j - \mathbf{R}_k$, and $r_{kj} = r_{jk}$ is the length of \mathbf{r}_{kj} . Thus, for any functional of $g(z)$, for example, s_2 and g_p , we can easily calculate its derivative, and hence the bias force.

In practice, in order to get continuous bias forces, we replace the Dirac- δ function by some smooth functions, for example, Epanechnikov kernel function,

$$K_e(z) = \frac{3}{4\sqrt{5}\epsilon} \left(1 - \frac{z^2}{5\epsilon^2}\right), \quad (14)$$

if $-\sqrt{5}\epsilon \leq z \leq \sqrt{5}\epsilon$, otherwise, $K_e(z) = 0$. While $\epsilon \rightarrow 0$, the $K_e(z) \rightarrow \delta(z)$. Under the replacement, we redefine

$$g_p = \frac{1}{4\pi\rho\delta N z_p^2} \sum_i \sum_{j \neq i} \int_{z_p - \delta/2}^{z_p + \delta/2} K_e(z - r_{ij}) dz, \quad (15)$$

where δ and z_p is the size and the center of the p^{th} bin, respectively. It is easy to know, while $\epsilon \rightarrow 0$, eq.(15) is same as the normal formula of pair correlation function in bins. The s_2 is redefined as,

$$s_2 = -2\pi\rho\delta \sum_{p=1}^{p_m} z_p^2 (g_p \ln g_p - g_p + 1), \quad (16)$$

where p_m is the maximal index of the bins and we have already set a cutoff along z , g_p is the bin-average value of $g(z)$ at the p^{th} bin, defined by the (15). Therefore, the derivative of the s_2 and g_p is continuous. For example,

$$\frac{\partial s_2}{\partial \mathbf{R}_k} = -\frac{1}{N} \sum_{j \neq k} \hat{r}_{kj} c(r_{kj}), \quad (17)$$

$$c(z) = \sum_p \ln g_p [K_e(z_p + \delta/2 - z) - K_e(z_p - \delta/2 - z)].$$

Here $c(z) \rightarrow -\frac{d}{dz} \ln g(z)$ under the limit $\epsilon \rightarrow 0$.

Besides choosing some functionals of pair distribution function $g(z)$ as the collective variables, another selection is possible, and may be better in some special systems. For more complex systems, for example, multi-component mixtures or macromolecular systems, etc., we should identify different kinds of atoms and calculate different pair distribution function $g_{AB}(z)$, where A and B are the types of atoms. An alternative method is to use energy distribution function. For any pair-interactive potential $E = u(z)$, such as Lennard-Jones interaction, coulomb interaction, the distance r_{ij} of atom pairs can be replace by the interactive energy $u(r_{ij})$, thus we can define the pair distribution function in energy space,

$$G(E) \propto \sum_i \sum_{j \neq i} \delta(E - u(r_{ij})). \quad (18)$$

Actually, $G(E)$ is a transformation of $g(z)$. Since the interactive energy is more directly related to the dynamics, using $G(E)$ replace $g(z)$ might be a good idea in identifying conformations. In addition, while existing many kinds of interaction between atoms, we can define the $G(E)$ for different kinds of interactive energy, for example $G(E_{bond})$, $G(E_{angle})$, $G(E_{torsion})$, $G(E_{lj})$, $G(E_{coul})$ etc., thus it is possible to take into account higher-order correlation. In the $G(E)$ cases, the bias force can be calculated similarly as that in the $g(z)$ cases. We are testing this idea in water condensation, details will appear elsewhere.

We have examined the general hyperdynamics method in a simple system of N identical Lennard-Jones (LJ) particles and studied the slow gas/liquid transition in the NVT ensemble. The main results and simulation details are published in ref. ([9]). In such a system, the potential energy of transitional conformations (liquid drops with critical size) is lower than that of the stable gas phase. Thus, simple bias potential based on potential energy [13] cannot work at all. In general, for entropy-important systems, using the potential alone is not sufficient to identify the transitional conformations and then form the bias

potential. We also found that with only two functionals of $g(z)$, the potential energy V and the two-body entropy s_2 , as the collective variables, we were able to correctly produce the slow dynamics. By designing suitable bias potentials, we could reach almost any slow timescale gas-liquid transition in reasonable time frame: we used a 10^6 time boost factor to find the transition in the system with very small saturation shown in the Fig. (1) [9].

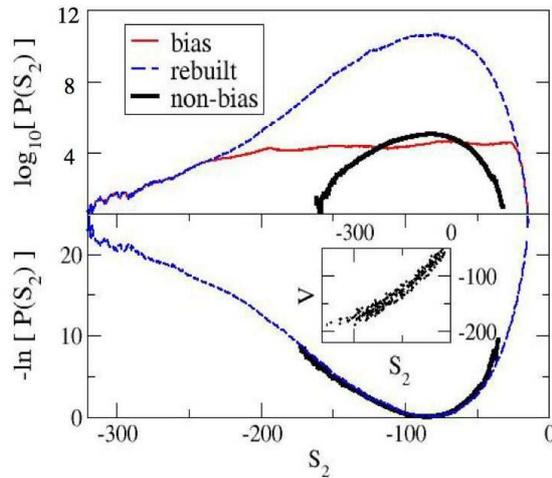


Fig. 1. Top: the distributions of two body entropy S_2 from non-biased and biased simulations, and the rebuilt distribution of the bias simulation. Here, the simulated system starts from gas phase with very small saturation. Bottom: the free energy profiles from the non-biased and biased simulations are compared. The inset shows the simulated samples in the (S_2, V) space. The observed liquid phase does not show here.

To summarize, we have extended the hyperdynamics method to more general cases by inhomogeneously compressing time. Our approach directly generates an explicit method to design the bias potential. In simple systems, two-body entropy s_2 as a functional of the pair distribution function provide a good approximation of the density field in identifying the important conformations and for constructing the bias potential without prior knowledge of the conformational space. The method can be applied in complex fluids, such as glass transition and liquid/solid transition of single or multi component Lennard-Jones fluids, water etc., and s_2 should be still the leading collective variable in the complex systems. For more complex cases, for example, polymers, biological systems, in whole conformational space, it is possible that we will need too many collective variables in identifying transitions, so that it is very difficult to estimate the density probability in the higher-dimension collective variables. A possible improvement is to divide the whole conformational space into some small parts, and use a few collective variables locally in each part.

Acknowledgments

This work was supported by the US DOE under contract No. DE-AC52-06NA25396. We are grateful to K. Kremer, H. Ziock, S. Rasmussen and A. F. Voter for stimulating discussions, comments and suggestions.

References

1. Bortz, A. B. and Kalos, M. H. and Lebowitz, J. L.: A new algorithm for Monte Carlo simulation of Ising spin systems. *J. Comp. Phys.* **17**, (1975) 10-18.
2. Bolhuis, P. G. and Chandler, D. and Dellago, C. and Geissler, P. L.: Transition Path Sampling: Throwing Ropes over Mountain Passes in the Dark. *Ann. Rev. Phys. Chem.* **53**, (2002) 291-318.
3. Olender, R. and Elber, R.: Calculation of classical trajectories with a very large time step: Formalism and numerical examples. *J. Chem. Phys.* **105**, (1996) 9299-9315.
4. Elber, R.: Long-timescale simulation methods. *Curr. Opin. Struct. Bio.* **15**, (2005) 151-156 .
5. Voter, A. F.: A method for accelerating the molecular dynamics simulation of infrequent events. *J. Chem. Phys.* **106**, (1997) 4665-4677.
6. Miron, R. A. and Fichthorn, K. A.: Accelerated molecular dynamics with the bond-boost method. *J. Chem. Phys.* **119**, (2003) 6210-6216.
7. Rahman, J. A. and Tully, J. C.: Puddle-skimming: An efficient sampling of multi-dimensional configuration space. *J. Chem. Phys.* **116**, (2002) 8750-8760.
8. Hamelberg, D. and Shen, T.-Y. and McCammon, J. A.: Phosphorylation effects on cis/trans isomerization and the backbone conformation of serine-proline Motifs: accelerated molecular dynamics analysis. *J. Am. Chem. Soc.* **127**, (2005) 1969-1974.
9. Zhou, X. and Jiang, Y. and Kremer, K. and Ziock, H. and Rasmussen, S.: Hyperdynamics for entropic systems: Time-space compression and pair correlation function approximation. *Phys. Rev. E* **74**, (2006) R035701.
10. Green, H. S.: *The Molecular Theory of Fluids*. (North-Holland, Amsterdam, 1952);
11. Baranyai, A. and Evans, D. J.: Direct entropy calculation from computer simulation of liquids. *Phys. Rev. A* **40**, (1989) 3817-3822.
12. Giaquinta, P. V. and Giunta, G. and PrestipinoGiarritta, S.: Entropy and the freezing of simple liquids., *Phys. Rev. A* **45** (1992) R6966 - R6968.
13. Steiner, M. M. and Genilloud, P.-A and Wilkins, J. W.: Simple bias potential for boosting molecular dynamics with the hyperdynamics scheme. *Phys. Rev. B* **57**, (1998) 10236-10239.