Equilibrium sampling by reweighting nonequilibrium simulation trajectories

Cheng Yang,[1] Biao Wan,[1] Shun Xu,[1] Yanting Wang,[2] and Xin Zhou[1,*]

[1]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[2]*State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100086, China*

(Received 29 August 2015; revised manuscript received 31 December 2015; published 21 March 2016)

Based on equilibrium molecular simulations, it is usually difficult to efficiently visit the whole conformational space of complex systems, which are separated into some metastable regions by high free energy barriers. Nonequilibrium simulations could enhance transitions among these metastable regions and then be applied to sample equilibrium distributions in complex systems, since the associated nonequilibrium effects can be removed by employing the Jarzynski equality (JE). Here we present such a systematical method, named reweighted nonequilibrium ensemble dynamics (RNED), to efficiently sample equilibrium conformations. The RNED is a combination of the JE and our previous reweighted ensemble dynamics (RED) method. The original JE reproduces equilibrium from lots of nonequilibrium trajectories but requires that the initial distribution of these trajectories is equilibrium. The RED reweights many equilibrium trajectories from an arbitrary initial distribution to get the equilibrium distribution, whereas the RNED has both advantages of the two methods, reproducing equilibrium from lots of *nonequilibrium* simulation trajectories with an *arbitrary initial* conformational distribution. We illustrated the application of the RNED in a toy model and in a Lennard-Jones fluid to detect its liquid-solid phase coexistence. The results indicate that the RNED sufficiently extends the application of both the original JE and the RED in equilibrium sampling of complex systems.

DOI: 10.1103/PhysRevE.93.033309

I. INTRODUCTION

Molecular dynamics (MD) simulation is an important tool for investigating the macroscopic physical properties of complex molecular systems by looking at their microscopic interactions. However, for many complex systems, such as biopolymers, glasses, phase-coexistence systems, etc., the potential energy surface is very rugged, with lots of energy basins, in the high-dimensional conformational space. Thus the usual equilibrium simulation within a finite time is usually trapped in a local conformational region nearby its initial conformation. The conformational space could be reduced to several metastable regions separated by high free energy barriers. The local equilibrium inside each of the metastable states can be reached within the normal simulation time scale, but the global equilibrium may need a far longer time to generate sufficiently the rare interstate transition events. In the past decades, many enhanced sampling simulation techniques have been invented to circumvent that problem and great successes have been achieved in many cases [1–20]. Nevertheless, more endeavours have continuously focused on improving the simulation efficiency, especially in complex systems.

Among several newly developed techniques fulfilling this purpose, one is the reweighted ensemble dynamics (RED) method [18]. The RED generates many independent short simulation trajectories obeying the same dynamics but starting from (arbitrarily) dispersed initial conformations. Due to the short simulation time, each of those trajectories can visit only a limited conformational region. When lots of dispersed initial conformations are applied, the whole set of trajectories could cover the whole equilibrium conformational space, but with a biased conformational distribution from the global equilibrium distribution, because these trajectories do not completely lose their memory of the initial conformations. The RED method extracts the relations among these trajectories to establish a linear equation whose solution provides the weights of these trajectories for reproducing the global equilibrium [18,21,22]. Practically, ones can use many independent computers to generate these trajectories simultaneously to shorten the waiting time.

In the RED, if each single simulation trajectory can visit a larger conformational region, fewer trajectories are required to cover the whole conformational space. Most of the existing enhanced sampling techniques which bias the potential energy surface to improve the visiting efficiency of a single trajectory [14,15,17,19,20] can be employed in the RED.

Nonequilibrium simulations under a time-dependent Hamiltonian can also be applied in the RED scheme, since the nonequilibrium effects can be removed by employing the Jarzynski equality (JE) [23,24], while the transitions among metastable states could be efficiently enhanced. The JE relates the works of nonequilibrium trajectories to the equilibrium free energy difference of the initial and final systems of the nonequilibrium process. It is widely applied to estimate free energy in simulation and in single molecular experiments [25,26]. However, in the JE, the initial conformational distribution of nonequilibrium trajectories is required to be equilibrium [27]; the requirement usually limits its application in many complex cases where the initial equilibrium distribution itself is hard to get.

In this paper, we combine the RED and the JE to present an efficient sampling method, named as the reweighted nonequilibrium ensemble dynamics (RNED) for generating equilibrium distributions in complex systems. The RNED method enhances the transitions among metastable states by a nonequilibrium process and applies the JE to remove the

*Author to whom correspondence should be addressed: xzhou@ucas.ac.cn

associated nonequilibrium effects. Therefore, RNED improves the efficiency of RED and extends the application range of JE at the same time.

This paper is organized as follows. The basic theory of the RNED is established in Sec. II, the simulations and results are introduced in Sec. III, and a short conclusion is given in Sec. IV.

II. THEORY

Let us consider an ensemble of simulation trajectories $\{q_i(\tau)\}, i = 1,...,N$, starting from different initial conformations $\{q_i(0)\}$ under the same Hamiltonian $H(q;\lambda)$ with the time-dependent parameter λ. We assume $H(q,\lambda)$ equals $H_0(q)$ while $0 \leqslant \tau \leqslant t_1$ and $t_2 \leqslant \tau \leqslant t$, i.e., two segments of equilibrium processes, but is time dependent between the two equilibrium segments, i.e., a nonequilibrium process in (t_1,t_2). In the RED, we construct the equilibrium distribution by reweighting the trajectories in a segment of equilibrium simulation, such as in the first equilibrium segment, $0 \leqslant \tau \leqslant t_1$:

$$P_w = \frac{1}{\sum_k w_k} \sum_j w_j P_j(q) \to P_{eq}(q), \tag{1}$$

with $N \to \infty$. Here w_k is the weight of the kth trajectory, and $P_k(q)$ is the distribution of its sample in conformational space q. The weight corresponds to the deviation of the initial distribution $P(q,0) = \frac{1}{N}\sum_j \delta(q - q_j(0))$ from the equilibrium distribution,

$$w_k = \frac{P_{eq}(q)}{P(q;0)}|_{q=q_k(0)} \approx \left\langle \frac{P_{eq}(q)}{P(q;0^+)} \right\rangle_{P_k(q,0^+)}, \tag{2}$$

where $\delta(\cdot)$ is the Dirac δ function. $\langle \cdots \rangle_{P(q)}$ means the ensemble average under the distribution $P(q)$, which is estimated from the corresponding sample of the $P(q)$ in practice. Here we replace a single initial configuration at $\tau = 0$ by a short initial segment of the trajectory ($\tau \in [0,0^+]$) to depress possible statistical errors [18]. It is easy to know $\sum_k w_k = N$ since $P(q;0^+) = \frac{1}{N}\sum_k P_k(q,0^+)$.

Substituting Eq. (1) into Eq. (2) leads to a linear equation of weights $\{w_j\}$,

$$\sum_j G_{ij} w_j = 0, \tag{3}$$

where $G_{ij} = \Lambda_{ij} - \delta_{ij}$ with $\Lambda_{ij} = \frac{1}{N}\langle \frac{P_j(q)}{P(q,0^+)} \rangle_{P_i(q,0^+)}$, and δ_{ij} is the Kronecker δ symbol. By applying a complete set of orthonormalized basis functions $\{A^\mu(q)\}$ [21], we have [18,22]

$$\frac{P_j(q)}{P(q,0^+)} = \sum_\mu A^\mu(q)\, a_j^\mu, \tag{4}$$

with the expanded coefficient $a_j^\mu = \langle A^\mu(q) \rangle_{P_j(q)}$. Here the $\{A^\mu(q)\}$ is orthonormalized by a standard orthogonal normalization process from lots of (arbitrarily) chosen basis functions [21], i.e., $\langle A^\mu(q) A^\nu(q) \rangle_{P(q,0^+)} = \delta_{\mu\nu}$. Consequently,

$$\Lambda_{ij} = \frac{1}{N} \sum_\mu \langle A^\mu(q) \rangle_{P_i(q,0^+)} \langle A^\mu(q) \rangle_{P_j(q)}. \tag{5}$$

Equation (3) is the key result of the RED [18], whose matrix elements are estimated from simulation data and whose solution provides $\{w_i\}$ for reproducing equilibrium properties. However, when the conformational space consists of multiple metastable regions separated by very high free energy barriers, simulation trajectories may be trapped in the local regions without crossing the barriers. In this case, Eq. (3) has multiple solutions and then the global equilibrium distribution cannot be uniquely determined, which is the main limitation in application of the RED [18].

In the RNED method, we design the nonequilibrium segment in (t_1,t_2) to overcome the very high free energy barriers and then to promote the transition events between metastable states. We then build a connection between the equilibrium simulations before and after the nonequilibrium segment. The solution degeneration problem in Eq. (3), due to inadequate interstate transitions, is then effectively overcome. We describe the remaining details of the RNED below.

Similar to Eq. (1), we can also construct the equilibrium distribution from the trajectories sampled in the second equilibrium segment, $t_2 \leqslant \tau \leqslant t$:

$$P_w^{(2)} = \frac{1}{\sum_k w_k^{(2)}} \sum_j w_j^{(2)} P_j^{(2)}(q) \to P_{eq}(q). \tag{6}$$

We can relate the weights $w_k^{(2)}$ to w_k by the nonequilibrium work of the kth trajectory in the nonequilibrium interval (t_1,t_2) [24,28],

$$w_k^{(2)} = w_k \Omega_k, \tag{7}$$

where $\Omega_k = \exp\{-W[q_k(\tau)]\}$, and the nonequilibrium work is

$$W[q_k(\tau)] = \int_{t_1}^{t_2} \frac{\partial H(q_k(\tau);\lambda)}{\partial \lambda} \frac{d\lambda}{d\tau} d\tau. \tag{8}$$

We can then combine the two weighted samples together to form the equilibrium distributions. For example, we estimate the equilibrium distribution by

$$P_{eq} \approx \frac{P_w(q) + \gamma P_w^{(2)}(q)}{1+\gamma}. \tag{9}$$

Here

$$\gamma = \frac{M_{\text{eff}}^{(2)}}{M_{\text{eff}}^{(1)}}, \tag{10}$$

where $M_{\text{eff}}^{(2)}$ and $M_{\text{eff}}^{(1)}$ are the effective sizes of the weighted sample in the second equilibrium segment and the first equilibrium segment, respectively. The effective size of a weighted sample is usually smaller than its real size. The exact formula may be dependent on these weights besides the real size of the sample, and thus we might write [20]

$$M_{\text{eff}} = \frac{(\sum w_i)^2}{\sum (w_i)^2} = M \frac{1}{1+\sigma^2}. \tag{11}$$

Here M is the real size of the sample, and σ is the fluctuation of the normalized weights, $\hat{w}_i = M \frac{w_i}{\sum w_i}$. Therefore, for an unweighted sample, the effective size of the sample is equal

to its real size, but for a weighted sample, the effective size is $1+\sigma^2$ times smaller than the real size M.

While both the effective sizes of the two samples are sufficiently large, any one of the two weighted samples itself gives a good estimate of the equilibrium distribution, and thus any γ could be applied in Eq. (9). In other words, the estimate is insensitive to the value of γ, although there exists a best γ for the accuracy of the estimate. We show that the selection of γ is not sensitive to the reproduced equilibrium distribution in the RNED. Therefore, the exact formula of the effective size of the sample and its weights is not important for our current purpose.

We have

$$P_{eq} \approx \frac{1}{N(1+\gamma)} \sum_j w_j \left[P_j(q) + \frac{\gamma}{c} \Omega_j P_j^{(2)}(q) \right], \tag{12}$$

where $c \equiv \frac{\sum w_i \Omega_i}{\sum w_k}$ should ideally be unity according to the JE [23], but practically it might slightly differ from unity due to the statistical error caused by a finite value of N. We can then obtain the linear equation $\sum_j G_{ij} w_j = 0$ with

$$G_{ij} = \frac{1}{1+\gamma} \left[(\Lambda_{ij} - \delta_{ij}) + \gamma \left(\frac{\Omega_j}{c} \Lambda_{ij}^{(2)} - \delta_{ij} \right) \right], \tag{13}$$

where $\Lambda_{ij}^{(2)} = \frac{1}{N} \langle \frac{P_j^{(2)}(q)}{P(q,0^+)} \rangle_{P_i(q,0^+)}$, the same as the definition of Λ_{ij} in Eq. (3), except that $P_j(q)$ is replaced by $P_j^{(2)}(q)$. This equation can be rewritten as

$$\sum_j \tilde{G}_{ij} w_j = 0, \tag{14}$$

where

$$\tilde{G}_{ij} = G_{ij} - \bar{G}_j = \frac{1}{1+\gamma} \left(\tilde{G}_{ij}^{(1)} + \gamma \tilde{G}_{ij}^{(2)} \right), \tag{15}$$

with $\tilde{G}_{ij}^{(1)} = \Lambda_{ij} - \delta_{ij}$ and $\tilde{G}_{ij}^{(2)} = \frac{\Omega_j}{c} (\Lambda_{ij}^{(2)} - \frac{1}{N}) - (\delta_{ij} - \frac{1}{N})$. Here $\bar{G}_j \equiv \frac{1}{N} \sum_i G_{ij}$. For convenience, we usually construct a symmetric matrix $\mathbf{H} = \tilde{G}^T \tilde{G}$ (i.e., $H_{jk} = \sum_i \tilde{G}_{ij} \tilde{G}_{ik}$) and calculate the ground state of $\mathbf{H}$ with the weight $\mathbf{w} = (w_1, \ldots, w_N)^T$, or equivalently, solve the equation

$$\mathbf{H}\mathbf{w} = 0, \tag{16}$$

to obtain the weight vector. In Eq. (16), $\tilde{G}_{ij}$ depends on the parameters γ and c. From the JE, $c = \frac{1}{N} \sum_i w_i \Omega_i$ is approximately equal to unity.

III. RESULTS

A. One-dimensional potential

We first employ a simple model with a single particle moving in a one-dimensional potential to illustrate the validity of the RNED method and what we should pay attention to when using this method. A particle moving in a one-dimensional potential $U(x)$ obeys the overdamped Langevin equation

$$\frac{dx}{dt} = -\frac{dU}{dx} + \sqrt{2T}\xi(t), \tag{17}$$

where $T = 0.2$ is the simulation temperature, and $\xi(t)$ is a Gaussian white noise with zero mean and obeys the correlation function $\langle \xi(t)\xi(t') \rangle = \delta(t-t')$. The integration time step is 0.001, and

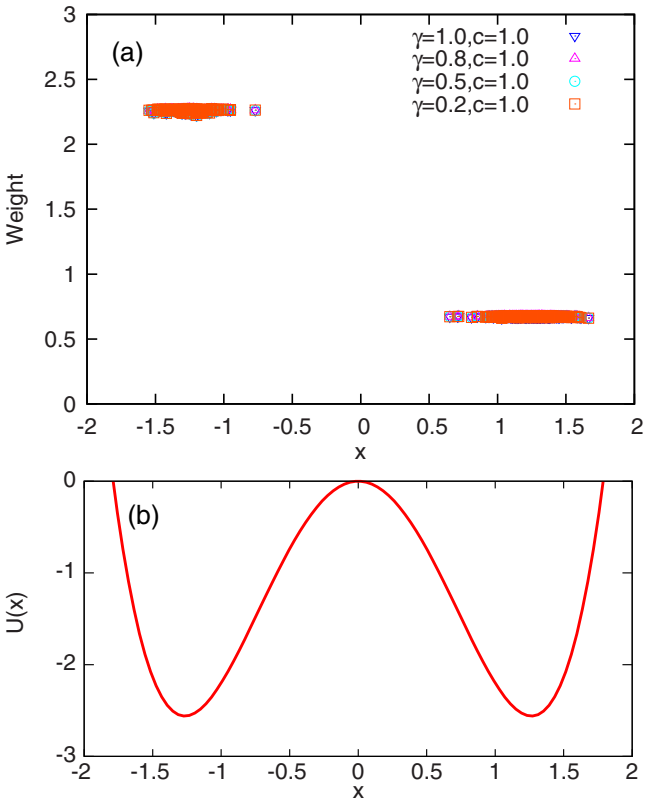

FIG. 1. (a) The relationship between the initial conformation of each trajectory and its weight. The difference between the weights with different γ values is very small. (b) The potential energy surface of $U(x)$.

$$U(x) = x^4 - kx^2, \tag{18}$$

where k is a controllable variable which determines the energy surface. In our case we choose $k = 3.2$ as the initial value and then change it over time to implement a nonequilibrium process. The two potential wells are located at $x = \pm\sqrt{\frac{k}{2}}$ and the height of the energy barrier is $\frac{k^2}{4}$. We deployed 1000 trajectories, with about 800 starting from the right well and the others from the left. The set

$$A^\mu(x) = \begin{cases} 1, & x_\mu < x < x_{\mu+1}, \\ 0, & \text{others} \end{cases} \tag{19}$$

was selected as the basis functions, where $x_\mu = -1.6 + 0.05\mu$, with $\mu = 1,\ldots,64$ covering all the important regions of the conformational space. Each bin of $\{x_\mu\}$ can be combined with its neighbors if it contains too few samples.

During the simulation, the value of k is changed with time as

$$k = \begin{cases} 3.2, & 0 \leqslant t < 100, \\ 3.2 - \Delta k \mathrm{int}(\frac{t-100}{0.02}), & 100 \leqslant t < 103, \\ 2.0, & 103 \leqslant t < 153, \\ 2.0 + \Delta k \mathrm{int}(\frac{t-153}{0.02}), & 153 \leqslant t < 156, \\ 3.2, & 156 \leqslant t \leqslant 256, \end{cases} \tag{20}$$

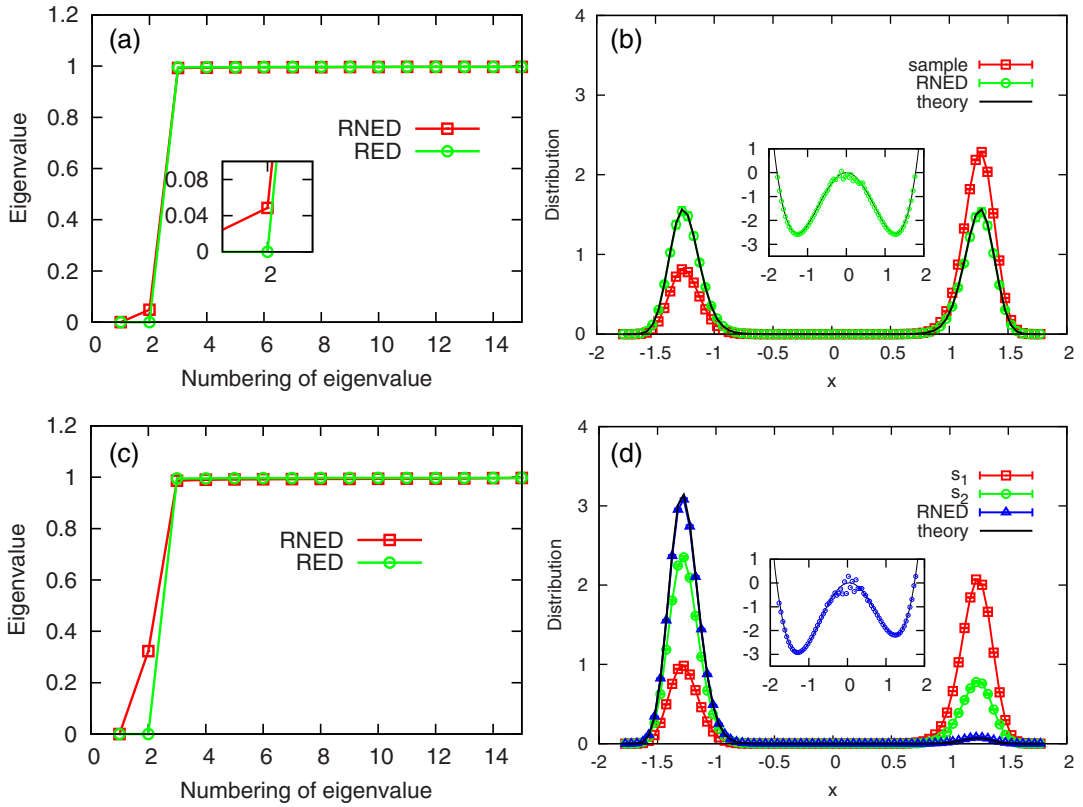

FIG. 2. (a) The first 15 smallest eigenvalues of RED and RNED when the potential energy is $U(x)$. The RED has two zero eigenvalues, which means there are few trajectories across the free energy barrier, so we cannot obtain the right weights via the RED method in this case. (b) Compare the sampled distribution (red line) and weighted distribution (green line) of RNED with the theoretical distribution of $U(x)$ (black line). The weighted distribution is almost the same as the theoretical distribution, showing that the RNED method is effective. The inset shows the free energy surfaces of theory and those obtained by RNED. They are almost the same except on the free energy barrier for rare samples. (c) The first 15 smallest eigenvalues of RED and RNED when the potential energy is $U_b(x)$. The RED method cannot obtain weights of trajectories for there are two zero eigenvalues. (d) The black line is a theoretical distribution of $U_b(x)$. The red line is the sampled distribution of the first segment and the green line is the sampled distribution of the second segment. Both sampled distributions are not equilibrium. The weighted distribution is consistent with the theoretical distribution. The inset shows the energy surfaces of theory and those obtained by the RNED. They are almost the same, except on the free energy barrier for rare samples.

where $\Delta k = 0.008$ and the function $\mathrm{int}(x)$ determines the largest integer smaller or equal to x. This nonequilibrium process is designed to first decrease the free energy barrier and then increase back to its original value. Samples are taken when $t \in [0,100]$ and $[156,256]$ with the interval of $\Delta t = 0.1$. We choose $\gamma = 1.0, 0.8, 0.5, 0.2$ and set $c = 1.0$ to calculate the weights of trajectories by the RNED, which are shown in Fig. 1(a). We can see that the weights are independent of γ, so we fix $\gamma = 1.0$ for simplicity in later calculations. Moreover, the trajectories in the same metastable state have similar weights, consistent with the RED method [18]. The free energy surface is shown in Fig. 1(b).

We have also performed a set of standard MD simulations with $k = 3.2$. The initial conformations and total simulation time of each trajectory of this ensemble are the same as in the RNED method. The MD trajectories have been analyzed with the RED method. The first 15 smallest eigenvalues of the RED are shown in Fig. 2(a) with green line. Since the ground state is degenerate due to the fact that very few trajectories cross the free energy barrier, the weights cannot be determined uniquely. In the RNED, decreasing the free energy barrier helps the trajectories to transit between the two metastable states, so the ground state is nondegenerate (red line) and the elements of the corresponding eigenvector are the weights of the trajectories. Figure 2(b) shows the distributions obtained by the RNED. The sampled distribution (red line) and the theoretical distribution (black line) are different, but the weighted distribution (green line) is almost the same with the latter, demonstrating the effectiveness of the RNED. The inset shows the free energy surface of the theoretical (black line) and that obtained by the RNED (green circle). They are almost the same, except on the free energy barrier for rare samples. In order to further explain how the RNED method works, we employ a asymmetric potential energy $U_b(x) = x^4 - kx^2 + 0.3x$ as an additional example. The nonequilibrium process is the same as Eq. (20), and standard MD simulations are also performed as a comparison. Figure 2(c) shows the 15 smallest eigenvalues of RED (green) and RNED (red). The RED method gets two zero eigenvalues while the RNED method is nondegenerate. The samples in RNED can be divided into two segments: the first segment is sampled before the nonequilibrium process and the second one is sampled after the nonequilibrium process. These two sampled distributions are shown in Fig. 2(d) (red and green), and both deviate from the theoretical distribution

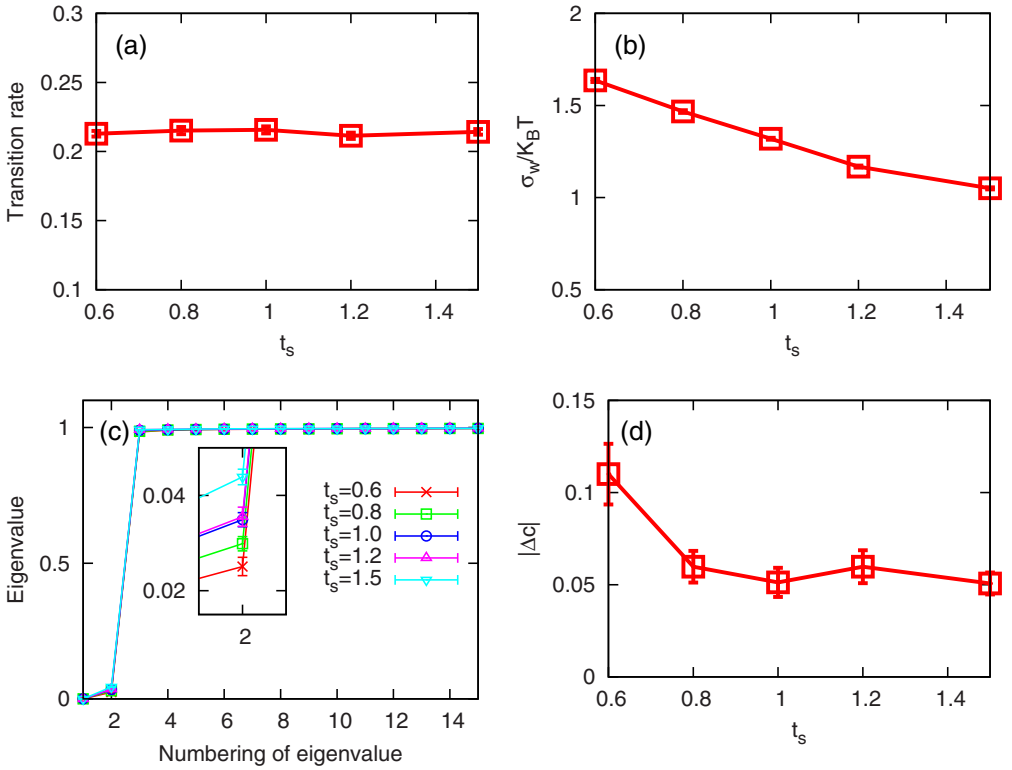

FIG. 3. Several parameters changed with t_s. (a) Transition rate between the two metastable states. (b) Standard deviation of the work. (c) The first 15 smallest eigenvalues of the RNED. (d) The deviation of parameter c from unity. In (a), (b), and (c) the statistical uncertainties are smaller than the symbols.

of $U_b(x)$ (black). The weighted distribution is the blue line, which is the same as the theoretical distribution. The inset shows the free energy surface of theory (black line) and that obtained by the RNED method (blue circle).

We then demonstrate what factors impact the results of the RNED. All the simulations are done under potential energy $U(x)$. Let $\Delta k = \frac{3.2-2.0}{m}$, where m is a controllable parameter. $t_s = 0.02m$ is the duration for k to decrease from 3.2 to 2.0, and t_m is the time for the system to stay at $k = 2.0$. We first vary the switching time t_s, representing the speed of changing k, by choosing a different m and keep $t_m = 50$. The results are shown in Fig. 3. As t_s increases, the transition rate (the ratio between the number of transition trajectories and total trajectories) does not change [see Fig. 3(a)]. However, the standard deviation of the accumulated work decreases [see Fig. 3(b)]. Correspondingly, the second smallest eigenvalue apparently rises when t_s increases [see Fig. 3(c)]. We employ $|\Delta c| = |c - 1.0|$ to describe the deviation of c from unity. It can be seen from Fig. 3(d) that the deviation decreases with t_s. Next, we vary t_m, the system evolving time with $k = 2.0$, and keep $m = 150$. The results are shown in Fig. 4. The transition rate increases along with t_m [see Fig. 4(a)], but the standard deviation of the accumulated work does not change [see Fig. 4(b)]. The second smallest eigenvalue becomes larger when t_m increases [see Fig. 4(c)]. In addition, the deviation of c from unity is shown in Fig. 4(d), which shows that $|\Delta c|$ dose not change with t_m.

The second smallest eigenvalue and $|\Delta c|$ are two indicators of the precision of the RNED. Weights are more precise when the second smallest eigenvalue deviates from zero more obviously and $|\Delta c|$ is closer to zero. The deviation between the second smallest eigenvalue and zero implies the ground state of the RNED is nondegenerate, and $|\Delta c|$ being close to zero suggests the JE is suitable for our example [29–31]. Therefore, the transition rate and standard deviation of the work are two main factors affecting the precision of the RNED. The RNED method can give a reasonable estimation of the equilibrium distribution only if both requirements are met.

B. Lennard-Jones fluids

Next we apply the RNED to a more complex system with the Lennard-Jones (L-J) potential. Our MD simulations of the system with the L-J potential are performed under the NVT ensemble by using the LAMMPS simulation package [32]. This system consists of 256 particles and has a box size of 22.58 Å × 22.58 Å × 22.58 Å with the periodic boundary condition applied. The potential parameters of L-J are $\epsilon/k_B = 119.8$ K and $\sigma = 3.405$ Å, and the cutoff is 8.5 Å. The properties of this model are similar to argon [33]. As shown in Fig. 5(a), there is a hysteresis loop in the potential energy U and the temperature T space while cooling and heating the system. Here we simulate 4 ns at each temperature. The system is located in one of the two metastable states (liquid and solid) between 50 K and 80 K, depending on its history. The higher energy branch corresponds to the liquid state, while the lower energy branch corresponds to solid state. The dashed line represents 68 K, where we will reconstruct the equilibrium distribution by the RNED, and the triangle is the obtained equilibrium energy due to the RNED reconstruction.

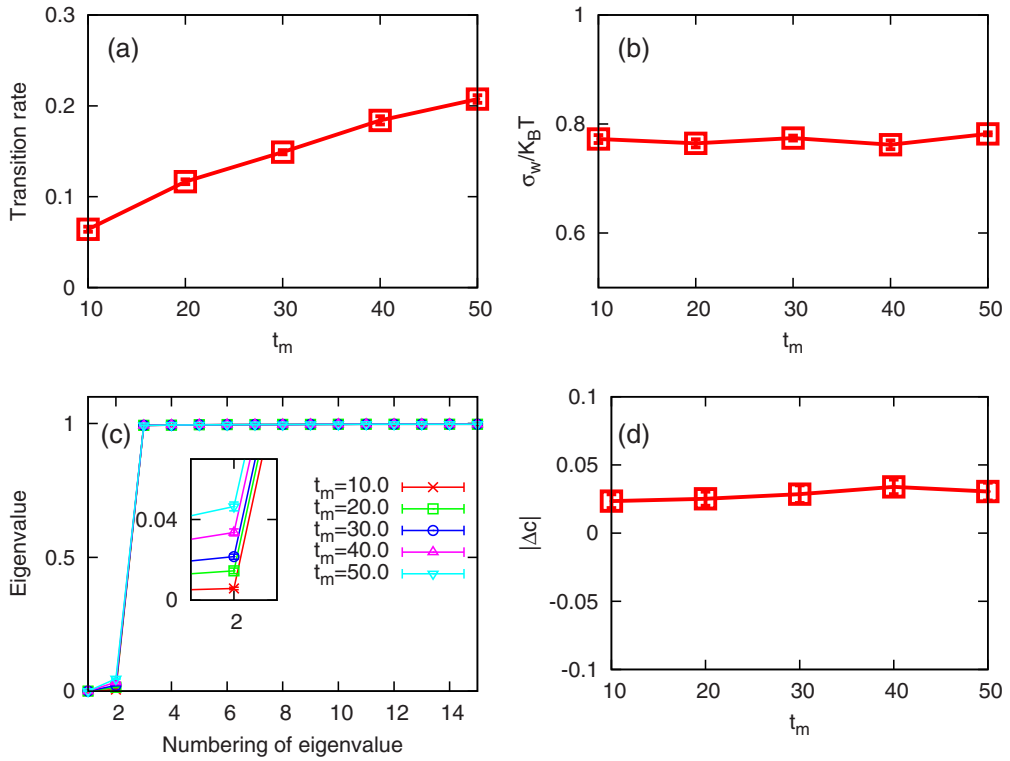

FIG. 4. Several parameters changed with t_m. (a) Transition rate between the two metastable states. (b) Standard deviation of the work. (c) The first 15 smallest eigenvalues of the RNED. (d) The deviation of c from unity. All statistical uncertainties are smaller than the symbols.

The L-J system is described by two order parameters. The first one is the potential energy U and the second is the average local bond order parameter $Q_6 = \langle Q_6(i)\rangle$ [34–36], where $\langle..\rangle$ denotes averaging over all particles. $Q_6(i)$ is defined as

$$Q_6(i) = \sqrt{4\pi/13}|\hat{q}_6(i)|, \tag{21}$$

where

$$\hat{q}_{6m}(i) = \frac{1}{N_b(i)}\sum_{k=0}^{N_b(i)} q_{6m}(k), \tag{22}$$

with $m = -6, -5 \ldots 5,6$ and $N_b(i)$ the number of first neighbors around particle i, and

$$q_{6m}(k) = \frac{1}{N_b(k)}\sum_{j=1}^{N_b(k)} Y_{6m}(\hat{r}_{kj}). \tag{23}$$

Here Y_{6m} is the spherical harmonic function, and $\hat{r}_{kj}$ is the normalized vector from particle k to particle j. The two metastable states of the L-J system in the (U, Q_6) map are shown in Fig. 5(b); the red points are solid conformations and green points are liquid conformations [35,36].

The functions of U and Q_6,

$$A^{\mu}(U,Q_6) = \begin{cases} 1, & U^l < U < U^{l+1}, Q_6^k < Q_6 < Q_6^{k+1}, \\ 0, & \text{others}, \end{cases} \tag{24}$$

are selected as the basis functions, where $U^l = -360 + l, l = 1,..,30$ and $Q_6^k = 0.1 + 0.02k, k = 1,..,20$, which cover the most important part of the conformational space.

We have simulated 1000 nonequilibrium trajectories and each trajectory lasts 8.0 ns. There are 500 trajectories starting from solid conformations and others starting from liquid conformations. The nonequilibrium process was implemented by changing the potential energy to $U_{\text{eff}} = U + \frac{\alpha(t)}{2\beta}(U - U_0)^2$, where U is the physical potential energy of the L-J system, $\beta = 1/k_B T$ with k_B the Boltzmann factor, and α is a controllable parameter. When $\alpha = 0.0$, it degenerates into a standard MD simulation. U_0 is chosen to approach the position of the free energy barrier between the liquid and solid, about $U_0 = -345.0$ in our case. The system evolves under the new potential energy U_{eff}, and thus the free energy surface is changed, leading to a higher transition rate between the liquid and solid states. α changes with time as

$$\alpha = \begin{cases} 0.0, & 0.0 \leqslant t < 0.5, \\ 0.0 + 0.0025 \times \text{int}\left(\frac{t-0.5}{0.005}\right), & 0.5 \leqslant t < 1.5, \\ 0.5, & 1.5 \leqslant t < 6.5, \\ 0.5 - 0.0025 \times \text{int}\left(\frac{t-6.5}{0.005}\right), & 6.5 \leqslant t < 7.5, \\ 0.0, & 7.5 \leqslant t \leqslant 8.0. \end{cases} \tag{25}$$

Samples are taken when $t \in [0,0.5]$ and $[7.5,8.0]$. We choose $\gamma = 1.0, 0.8, 0.5, 0.2$ and set $c = 1.0$ to calculate the weights of the nonequilibrium ensemble with the RNED method, as shown in Fig. 5(c). Since the weights are independent of the specific choice of γ, we set $\gamma = 1.0$ for simplicity. The 15 smallest eigenvalues of RNED are shown in Fig. 5(d) (green line). We also simulated 1000 equilibrium trajectories as a comparison. Each trajectory of the comparison ensemble lasted 8.0 ns, too. Then we used the RED method to analyze this ensemble. The first 15 smallest eigenvalues are shown in

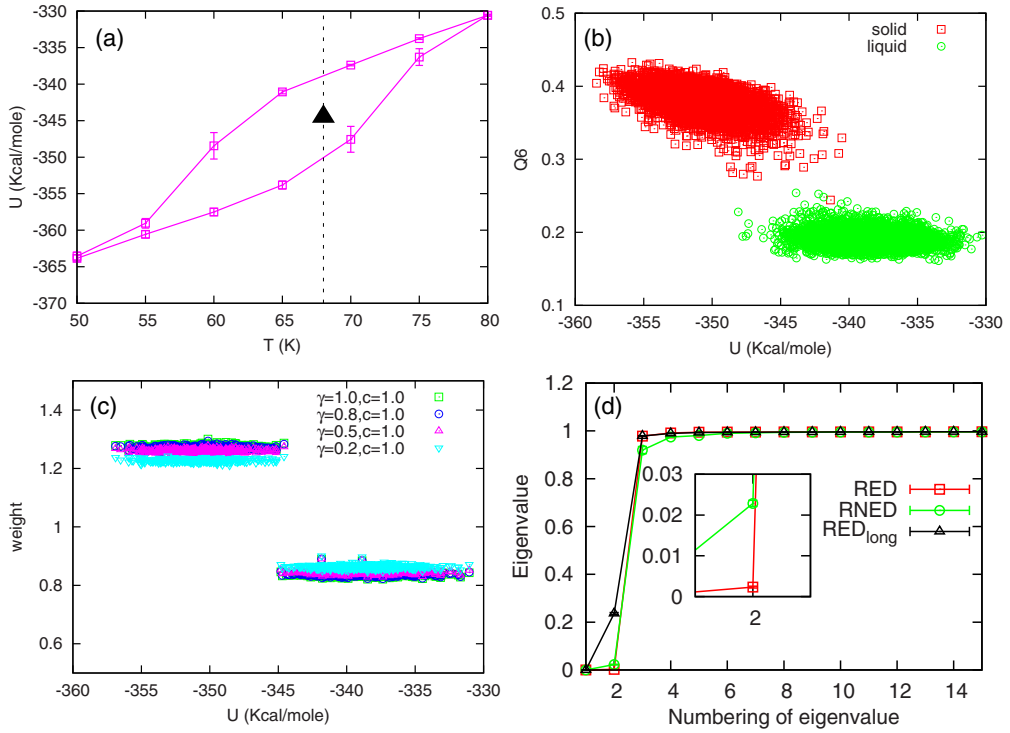

FIG. 5. (a) The hysteresis loop in the potential energy U and temperature T space. The dashed line is the temperature point at which we will reconstruct the equilibrium distribution. The triangle is the equilibrium potential energy of the L-J system we obtained by the RNED method. (b) (U, Q_6) map for the metastable states of the L-J system; the red points are solid and the green points are liquid. (c) The relationship between the initial conformation of each trajectory and its weight. The differences between the weights with different γ values are very small. (d) The 15 smallest eigenvalues of RED and RNED. The black line is the result of RED for a 200-ns-length ensemble. The red line is the result of RED for an 8-ns-length ensemble. The second eigenvalue is very close to zero, implying there are few trajectories across the free energy barrier. The green line is the result of RNED for the nonequilibrium ensemble. The second eigenvalue obviously deviates from zero when compared with the red line.

Fig. 5(d) (red line). The second eigenvalue is very closed to zero when compared with the RNED's, for there are few trajectories across the free energy barrier. So the precise weights of the 8.0-ns-length ensemble cannot be obtained by the RED method. As an addition, we simulated 200 equilibrium trajectories, with 100 trajectories starting from the solid state and the others starting from the liquid state. Each trajectory lasted 200 ns. We analyzed this 200-ns-length ensemble with the RED method. The first 15 smallest eigenvalues are shown in Fig. 5(d) (black line). The second eigenvalue obviously deviates from zero.

We discarded 100 trajectories starting from the liquid state of the nonequilibrium ensemble randomly and the others constitute a new ensemble which we call simulation1. In the same way, we discarded 200 trajectories starting from the solid state of the nonequilibrium ensemble and we call the rest of the nonequilibrium ensemble simulation2. The sampled distributions of these two ensembles in different order parameter spaces are shown in Figs. 6(a) and 6(b). Then we used the RNED method to analyze these ensembles. The weighted distributions of the two ensembles in parameter Q_6 space are shown in Fig. 6(c) and those in parameter U space are shown in Fig. 6(d). The weighted distributions of the two ensembles are very similar in both order parameter spaces. The black lines in Figs. 6(c) and 6(d) are weighted distributions of a 200-ns-length ensemble obtained by the RED method. The RNED method can give the same weighted distribution from different initial conformations, and the weighted distribution is also consistent with the weighted distribution of a longer MD ensemble analyzed by the RED method.

IV. CONCLUSIONS

In this work, we generalized the reweighted ensemble dynamics (RED) scheme to the reweighted nonequilibrium ensemble dynamics (RNED) method with the help of the Jarzynski equality. The RNED is especially useful when the free energy barriers between metastable states are so high that the transitions overcoming these barriers cannot adequately happen in the usual equilibrium simulations within reasonable time scales. In the RNED, a well-designed nonequilibrium process can greatly enhance the transition rates between different free energy basins, and weights for nonequilibrium trajectories can be calculated from the simulation data itself to reconstruct the global equilibrium. This novel method has been successfully applied to two systems here. For the one-dimensional system, we compared RED and RNED with the same simulation time. The results show that the RNED method is efficient and accurate, while the fluctuation of nonequilibrium works is limited in a few $k_B T$, and the fraction

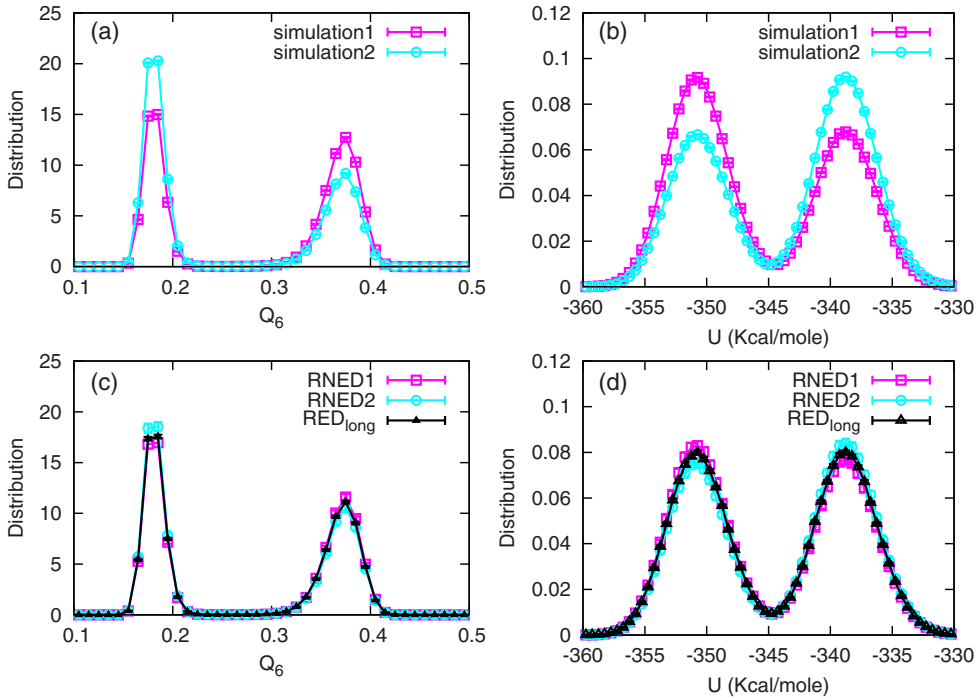

FIG. 6. (a) The sampled distributions of simulation1 and simulation2 in parameter U space. (b) The sampled distributions of simulation1 and simulation2 in Q_6 space. (c, d) The weighted distributions of these two nonequilibrium ensembles. Black lines are the weighted distributions of a 200-ns-length ensemble obtained by the RED method.

of transition trajectories is significantly enhanced. In the L-J system, from different initial distributions, we show that the obtained equilibrium distributions by the RNED are consistent each other, and also in good agreement with the results of the RED with longer trajectories. The RNED is advantageous in the sense that it does not require much *a priori* knowledge about the simulated system, for example, the foreknown information about metastable states. The initial conformations of RNED are almost arbitrary rather than the equilibrium in the original JE, and they can be generated in many different ways, such as simulations under high or low temperatures, the usual nonequilibrium simulations, etc. The RNED method widely extends the application of the original JE and a very recent improvement where the JE was extended to a matrix equality to calculate the free energies of foreknown metastable states [37].

ACKNOWLEDGMENTS

This work was supported by the NSFC under Grant No. 11175250 and the Open Project from the State Key Laboratory of Theoretical Physics. The authors are thankful for fruitful discussions with D. P. Landau. X.Z. thanks the Hundred Talent Program of the Chinese Academy of Sciences for financial support.

[1] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[2] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[3] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).
[4] D. D. Frantz, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **93**, 2769 (1990).
[5] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).
[6] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[7] H. Grubmuller, Phys. Rev. E **52**, 2893 (1995).
[8] A. F. Voter, Phys. Rev. Lett. **78**, 3908 (1997).
[9] X. Wu and S. Wang, J. Chem. Phys. **110**, 9401 (1999).
[10] Q. Yan and J. J. de Pablo, J. Chem. Phys. **111**, 9509 (1999).
[11] A. Mitsutake, Y. Sugita, and Y. Okamoto, Biopolymers (Peptide Sci.) **60**, 96 (2001).
[12] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. USA **99**, 12562 (2002).
[13] I. Andricioaei, A. R. Dinner, and M. Karplus, J. Chem. Phys. **118**, 1074 (2003).
[14] D. Hamelberg, T.-Y. Shen, and J. A. McCammon, J. Chem. Phys. **122**, 241103 (2005).
[15] X. Zhou, Y. Jiang, K. Kremer, H. Ziock, and S. Rasmussen, Phys. Rev. E **74**, 035701(R) (2006).
[16] A. Barducci, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **100**, 020603 (2008).
[17] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[18] L. Gong and X. Zhou, Phys. Rev. E **80**, 026707 (2009).
[19] C. Zhang and J. Ma, J. Chem. Phys. **130**, 194112 (2009).
[20] S. Xu, X. Zhou, Y. Jiang, and Y. Wang, Sci. China: Phys., Mech. Astron. **58**, 090501 (2015).

[21] Z. Chuan-Biao, L. Ming, and Z. Xin, Chin. Phys. B **24**, 120202 (2015).
[22] L. Gong, X. Zhou, and Z.-C. OuYang, Chin. Phys. B **24**, 060202 (2015).
[23] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[24] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. USA **98**, 3658 (2001).
[25] S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, J. Chem. Phys. **119**, 3559 (2003).
[26] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, Science **296**, 1832 (2002).
[27] Z. Gong and H. T. Quan, Phys. Rev. E **92**, 012131 (2015).
[28] G. Hummer and A. Szabo, Acc. Chem. Res. **38**, 504 (2005).
[29] D. M. Zuckerman and T. B. Woolf, Phys. Rev. Lett. **89**, 180602 (2002).
[30] J. Gore, F. Ritort, and C. Bustamante, Proc. Natl. Acad. Sci. USA **100**, 12564 (2003).
[31] S. Park and K. Schulten, J. Chem. Phys. **120**, 5946 (2004).
[32] See http://lammps.sandia.gov/ for information about the LAMMPS molecular dynamics simulator.
[33] J.-P. Hansen and L. Verlet, Phys. Rev. **184**, 151 (1969).
[34] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983).
[35] W. Lechner and C. Dellago, J. Chem. Phys. **129**, 114707 (2008).
[36] J. Russo and H. Tanaka, Sci. Rep. **2**, 505 (2012).
[37] B. Wan, C. Yang, Y. Wang, and X. Zhou, arXiv:1512.08631.