Reweighted ensemble dynamics simulations: Theory, improvement, and application*

Gong Lin-Chen(龚麟宸)[a], Zhou Xin(周 昕)[b]†, and Ouyang Zhong-Can(欧阳钟灿)[a,c]

[a] *Center for Advanced Study, Tsinghua University, Beijing* 100084, *China*
[b] *School of Physics, University of Chinese Academy of Sciences, Beijing* 100049, *China*
[c] *Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing* 100190, *China*

(Received 20 November 2014; revised manuscript received 19 January 2015; published online 10 April 2015)

Based on multiple parallel short molecular dynamics simulation trajectories, we designed the reweighted ensemble dynamics (RED) method to more efficiently sample complex (biopolymer) systems, and to explore their hierarchical metastable states. Here we further present an improvement to depress statistical errors of the RED and we discuss a few keys in practical application of the RED, provide schemes on selection of basis functions, and determination of the free parameter in the RED. We illustrate the application of the improvements in two toy models and in the solvated alanine dipeptide. The results show the RED enables us to capture the topology of multiple-state transition networks, to detect the diffusion-like dynamical behavior in an entropy-dominated system, and to identify solvent effects in the solvated peptides. The illustrations serve as general applications of the RED in more complex biopolymer systems.

Keywords: ensemble dynamics simulations, peptides, enhanced sampling, metastable states

PACS: 02.70.Ns, 87.15.A–, 82.20.Wt **DOI:** 10.1088/1674-1056/24/6/060202

1. Introduction

Exploring for the high-dimensional conformational space of complex molecular systems has long been the focus of computer simulation studies in recent years. Two challenges exist under this topic, i.e., how to thoroughly sample the conformational space, and how to understand the equilibrium and dynamic properties based on lots of simulation data in hand. Although standard Monte-Carlo (MC) and molecular-dynamics (MD) simulation methods have been widely applied and achieved great successes, for many complex systems, such as glassy systems, phase-transition systems, and solvated biopolymers, sufficient sampling is by no means an easy task. The rugged energy landscapes in the systems frequently trap simulation trajectories in a small conformational region, disallowing the thorough exploration of the whole conformational space within limited simulation time. Due to those, lots of various advanced simulation techniques, such as umbrella sampling,[1–3] J-walk,[4] multicanonical ensembles,[5] simulated tempering,[6,7] conformational flooding,[8] hyperdynamics,[9–11] parallel tempering,[12,13] Wang–Landau sampling[14] and its derivatives,[15,16] metadynamics,[17] and multidimensional generalized-ensemble methods,[18] have been developed and applied to overcome the difficulties. New hardware[19] and parallel computational technique[20] were also designed to extend MD simulations to reach microsecond scale in protein systems. On the other hand, due to the growing power of massively distributed computation, the ensemble dynamics (ED),[21–23] which independently runs multiple simulations in distributed computers, provides an alternative way for bridging the gap between simulation power and practical interest.

Besides efficiently sampling, proper analysis and understanding systems from the simulation data are also not a trivial task. The major purpose in data analyses, is to search for a coarse-grained description of the conformational space. In the traditional way, a few (usually only one or two) knowledge-based reaction coordinates (or order parameters) are first defined, then the equilibrium (and dynamic) properties of the system are achieved by mapping sampled configurations in the low-dimensional reaction-coordinate space. However, this method may suffer from the arbitrariness in the selection of reaction coordinates, and give an oversimplified description of the system.[24] Recently, a systematic metastable-state-network-based view brought prominent improvement.[25–34] Due to being theoretically transparent and inclusive of nearly all the useful information, the network model has received much attention.

In a previous work,[35] we have proposed a method, which was previously denoted as the weighted ensemble dynamics, now renamed as the reweighted ensemble dynamics (RED), to enhance sampling and to form a metastable-state transition network.[36] The RED is following the spirit of ED, which generated multiple short simulation trajectories from dispersively-chosen initial conformations, then applies a network-based analysis to get equilibrium and dynamics properties from all of these trajectories. In the present pa-

*Project supported by the National Natural Science Foundation of China (Grant No. 11175250).
†Corresponding author. E-mail: xzhou@ucas.ac.cn
© 2015 Chinese Physical Society and IOP Publishing Ltd

per, we further improve and validate the RED, including: (i) we present a scheme to select basis functions and judge their completeness; (ii) we discuss the unique free parameter in the RED; (iii) we provide an improvement to depress statistical errors. We illustrate the improved RED in a few different systems, give the topology structure of a multiple-state transition network, and predict the diffusion-like dynamical property in entropy-dominated systems. More importantly, we apply the RED to analyze the effects of different aspects (degrees of freedom, interactions, etc.) separately, such as the solvent effects in the solvated dipeptide.

The article is organized as follows. In Section 2, we briefly represent the main formula of the RED, then we discuss a few keys in the RED, such as selection of basis functions, the effects of unique parameters, and the improvement of statistical errors. In Section 3, we illustrate the application of the RED in two-dimensional multiple-well potentials and in solved dipeptides. A short conclusion is given in Section 4. The simulation and analysis details are presented in Appendix A.

2. Theory and method

In the RED, multiple MD (or MC) trajectories, $\{q_i(t)\}$, $i = 1,\ldots,p, t \in [0,\tau]$, are independently generated from initial conformations, denoted as $\{q_i(0) = q_{i,0}\}$, $i = 1,\ldots,p$, under the same simulation condition, respectively. Each trajectory involves m conformations, $\{q_i(t = j\Delta) = q_{i,j}\}$, $j = 0,\ldots,m-1$, thus total $m \times p$ conformations are sampled. Here q is a simple notation of configurational coordinates. Δ is the time interval of the samples, and $m\Delta = \tau$. We denote the distribution of samples in a single trajectory as $P_i(q) = \frac{1}{m}\sum_{j=1}^{m}\delta(q-q_{i,j})$, the initial distribution of all the trajectories as $P_{\rm init}(q) = \lim_{p\to\infty}\frac{1}{p}\sum_{i=1}^{p}\delta(q-q_{i,0})$.

2.1. Theory of the RED

Usually, while τ is shorter than the equilibrium time of systems, a single trajectory does not reach the equilibrium, i.e., $P_i(q) \neq P_{eq}(q)$, but we can use all of these trajectories to reproduce the equilibrium distribution,

$$P_{\rm eq}(q) = \frac{\sum_i w_i P_i(q)}{\sum_i w_i}, \tag{1}$$

where w_i is the weight of the i-th trajectory, which is due to the deviation of the initial distribution from the equilibrium distribution. A possible selection of $\{w_i\}$ is,

$$w_i = \left.\frac{P_{\rm eq}(q)}{P_{\rm init}(q)}\right|_{q=q_{i,0}}. \tag{2}$$

Thus, we have a linear equation about $\{w_i\}$,

$$w_i = \frac{1}{p}\sum_j \Gamma_{ij} w_j, \tag{3}$$

where

$$\Gamma_{ij} = \left.\frac{P_j(q)}{P_{\rm init}(q)}\right|_{q=q_{i,0}},$$

and $\sum_j w_j = p$.

Due to the finite size of samples $\{q_{i,j}\}$, equation (1) should be understood as

$$\langle A(q)\rangle_{\rm eq} = \frac{1}{p}\sum_i w_i \langle A(q)\rangle_i$$

for any $A(q)$, while $\langle A(q)\rangle_i$ is estimated as $\frac{1}{m}\sum_{j=0}^{m-1} A(q_{i,j})$. Similarly, the ratio of two distributions is understood as

$$\left\langle \frac{P_j(q)}{P_{\rm init}(q)} A(q)\right\rangle_{\rm init} = \langle A(q)\rangle_j,$$

while $\langle\ldots\rangle_\alpha$ is estimated by average in the corresponding conformation samples. By expanding on a complete set of conformational functions (named as basis functions) $\{A^\mu(q)\}, \mu = 1,2,\ldots$, we have,

$$\Gamma_{ij} = \sum_{\mu,\nu} g_{\mu\nu} A^\mu(q_{i,0}) \langle A^\nu(q)\rangle_j. \tag{4}$$

Here $g_{\mu\nu}$ is the inverse of $g^{\mu\nu} \equiv \langle A^\mu(q) A^\nu(q)\rangle_{\rm init}$ which is estimated by the average in the initial sample.

In practical application, instead of $A(q_{i,0})$ in estimation of Γ_{ij}, we use the average of $A(q)$ in a short $\tilde{t}$-length initial segment of trajectory (the first m_0 configurations with $m_0 \ll m$ or $\tilde{t} \ll \tau$), i.e.,

$$A(q_{i,0}) \leftarrow \langle A(q)\rangle_{i^+} = \frac{1}{m_0}\sum_{j=1}^{m_0} A(q_{i,j}),$$

then

$$\Gamma_{ij} = \frac{1}{p}\sum_{\mu,\nu} \langle \hat{A}^\mu(q)\rangle_{i^+} \langle \hat{A}^\nu(q)\rangle_j. \tag{5}$$

Here $\{\hat{A}^\mu(q)\}$ is the orthonormalized basis functions satisfying $\langle \hat{A}^\mu(q)\hat{A}_\nu(q)\rangle_{\rm init+} = \delta_{\mu\nu}$ with

$$P_{\rm init+}(q) = \frac{1}{p}\sum_i P_{i^+}(q).$$

Thus, equation (3) can be rewritten as

$$H\boldsymbol{w} = 0. \tag{6}$$

Here $H \equiv G^{\rm T}G$ is the variance–covariance matrix of the trajectory-mapped vectors

$$\boldsymbol{G}^i = (G_{i1},\ldots,G_{i,p})^{\rm T}, \tag{7}$$

where $G_{ij} \equiv \Gamma_{ij} - \delta_{ij}$ with $i = 1,\ldots,p$, and $\boldsymbol{w} = (w_1,\ldots,w_p)^{\rm T}$. Equation (6) is the main formula of the RED. The $\boldsymbol{w}$ corresponds to the ground state of matrix H whose number of degenerated ground states gives the number of disconnected metastable states under the timescale of the individual trajectory τ, the existence of a small fraction of transition trajectories among states can slightly split the ground state of H. More details are shown in our previous work.[35]

2.2. Practical schemes of the RED

The selection of basis functions is a key in the practical application of RED. In our previous works,[35,36] we briefly discussed the selection of basis functions, and used some triangle functions as basis functions like Fourier expansion. Here we more generally discuss the key point of the RED.

2.2.1. Selection of basis functions

Although the expansion in the RED required a complete set of basis functions in principle, it is usually not necessary to include too many basis functions in practice. Usually, it is easy to choose the initial configurations so that its distribution $P_{\text{init+}}(q)$ is already local equilibrium, thus only slow processes need to be reweighted by the RED. Therefore $\Omega(q)$ in the all-atomic q space can be replaced by $\Omega(x) = P_{\text{eq}}(x)/[P_{\text{init+}}(x)]$ in a coarse-grained x space, where x represents the slow (coarse-grained) degrees of freedom, $P(x)$ is the corresponding distribution function in the x space. For example, in proteins, if we disregard the first few nanoseconds of relaxation simulation and reset the zero point of time, it is safe to suppose only some slow degrees of freedom, such as torsion angles of backbone, positions of side-chain atomic groups, be x, thus only functions in the x space should be applied as basis functions. In many cases, systems might reach equilibrium within the τ time (length of individual trajectory) except for only a few collective variables, thus we even apply the RED in a low-dimensional space. In addition, we usually suppose there are many potential-energy basins (or super-basins) separated by high (free) energy barriers in the conformational space; each (super-)basin corresponds to a metastable state wherein a τ-length trajectory is sufficient to reach local equilibrium. Thus, $\Omega(x)$ looks like a step-like function, which is almost constant in each metastable state, variations only happening at boundaries between states. We approximately have

$$\Omega(x) \approx \sum_{\alpha} \Theta_{\alpha}(x)\omega_{\alpha}, \tag{8}$$

where $\Theta_{\alpha}(x)$ is character function of the α metastable state, which is unity if x is inside the state, zero otherwise. The set of character functions of states $\{\Theta_{\alpha}(x)\}$ provides a natural set of basis functions to expand $\Omega(x)$, thus the required number of basis functions is equal to the number of states. Since the boundaries of these states are unknown *a priori*, we usually use far more basis functions to linearly combine the step-like $\Omega(x)$ function. We may use some usual collective variables with clear physical meaning as basis functions. In biological macromolecules, dihedral angles and their analytical functions, pair distances between some key atomic groups, potential energy or its parts, solved area of proteins, radius of gyration, native contact number, number of hydrogen, etc., can be applied as basis functions. The variance–covariance matrix $g^{\mu\nu}$ will ensure the consistent consideration of different classes of basis functions, and provide the measure of their relative importance by the orthonormalization process. In addition, the number of independent basis functions is also limited by the finite size of init+, thus more selected basis functions will be discarded after considering the matrix $g_{\mu\nu}$, then do not significantly contribute to final results. Another possible scheme about the selection of basis functions is to use local functions in the conformational space. For example, for a set of sampled configurations, $\{x_i\}, i = 1, \ldots, M$, one may group them into some subsets by clustering algorithms based on their geometric similarity, and use the character functions of clusters as basis functions. Thus the number of basis functions is the number of clusters, $n_{\text{cl}} = M/m$. Here M and m are the total number of sampled conformations and the (average) number of conformations in each cluster. The idea is actually already applied in the Markov State Model (MSM) to divide the whole conformational space into metastable states.[24,30–34] In this paper, we use physically meaningful collective variables as basis functions, the application of the sample-based discrete cluster functions will appear elsewhere.

2.2.2. Completeness of basis functions

The completeness of selected basis functions can be directly test-based on sampled configurations in the RED. For a set of basis functions $\{A^{\mu}(x)\}$, we can estimate the difference between two samples (or the corresponding distributions), such as the initial distribution $P_{\text{init+}}(x)$ and the equilibrium distribution $P_{\text{eq}}(x)$,

$$S^2_{\text{init+,eq}} = \sum_{\mu,\nu} g_{\mu\nu} \langle A^{\mu}(x)\rangle_{\text{eq}} \langle A^{\nu}(x)\rangle_{\text{eq}}. \tag{9}$$

The definition gives a measurement about the similarity of samples, and can be generally applied to compare two samples. For example, we had used the formula to optimize coarse-grained models of the all-atomic water.[37] For assessing the completeness of basis functions in the expansion of $P_{\text{eq}}(x)/[P_{\text{init+}}(x)]$, a reasonable method is to check the convergence of $S^2_{\text{init+,eq}}$ while adding more and more basis functions.

To calculate $S^2_{\text{init+},eq}$, the sample for equilibrium distribution $P_{\text{eq}}(x)$, thus the trajectory weights $\{w_i\}$, are necessary. We propose two methods to bypass this difficulty. First, before reweighting, $P_{\text{eq}}(x)$ is replaced by $\bar{P}(x) = \frac{1}{p}\sum P_i(x)$, and we roughly select basis functions. Second, after reweighting, $S^2_{\text{init+,eq}}$ can be calculated with the resulting trajectory weights. This procedure can be used to test the former selection. If the rigorously calculated $S^2_{\text{init+,eq}}$ has reached its saturation region with previously roughly selected basis functions, the results should be satisfiable. However, this posterior examination is restricted to the cases that the trajectory weights can be uniquely determined. In practice, the distribution $\bar{P}(x)$, has the same local equilibrium characteristics in each metastable

region as $P_{\text{init}+}$ and P_{eq}, thus $\Omega_{\text{init}+,\bar{P}} = \bar{P}(x)/[P_{\text{init}+}(x)]$ is similarly a step function, which is also constant in each meta-stable region, only the constant values are different from those of $\Omega_{\text{init}+,\text{eq}}$. Thus we believe the *a priori* test is already enough, and the following results are all calculated in this way.

Since the statistical errors exist due to the finite sample sizes, S^2 would not strictly stop growing with the number of basis functions, it just increases slowly compared to the initial growing phase with increasing basis functions. Similarly, when a large error occurs, S^2 may abnormally increase within the saturation region. We design a scheme to get rid of the statistically unreliable basis functions. We orthonormalize the functions $\{A^\mu(x)\}$, leading to the explicitly identifiable value and error contribution from the orthonormalized basis functions to S^2. In the present paper, we implement the Gram–Schmidt orthonormalization method to one-by-one derive out the set of orthonormalized basis functions $\{\hat{A}^\mu(x)\}$, thus $S^2 = \sum_\mu s_\mu^2$, with $s_\mu^2 = \langle \hat{A}^\mu(x) \rangle_{\text{eq}}^2$. To estimate the error of S^2, we approximately estimate the error of s_μ, and combine them by

$$\text{Err}(S^2) = \sqrt{\sum_\mu 4 s_\mu^2 \text{Err}^2(s_\mu)}. \tag{10}$$

Here, Err( ) denotes the statistical error of the quantity in brackets. For each orthonormalized basis function, $\hat{A}^\mu$, if $\text{Err}(s_\mu)$ is larger than $|s_\mu|$, i.e., the relative error of s_μ is larger than 1.0, this basis function will not be selected in expansion. One-by-one selection will give out the final basis function set and the estimated error of the completeness measure S^2.

The scheme introduced here provides us with the knowledge about how many basis functions are enough in the RED. After orthonormalization, the contribution of the basis functions to S^2 clearly reflect their relative importance in trajectory reweighting. The basis functions that contribute more are more important for portraying the difference between $P_{\text{init}+}(x)$ and $P_{\text{eq}}(x)$, and the ones that contribute quite little can be linearly expanded by other existing basis functions. Furthermore, various kinds of basis functions can be safely added to investigate the related phenomena. For illustration, we will show that the basis functions of dihedral angles and those of distances between heavy atoms can be consistently treated in calculation, and the effect of a solvent on the conformation of biological macromolecules can be studied by additional basis functions characterising the solute–solvent relationship.

2.2.3. The effects of free parameter $\tilde{t}$

There is only one free parameter in the RED, $\tilde{t}^* = \tilde{t}/\tau$, for collecting the init+ samples in Eq. (5). Large $\tilde{t}^*$ may bring systematical errors, and small $\tilde{t}^*$ will reduce the sample number in init+, leading to a larger statistical uncertainty. $\tilde{t}^* = 0.01 \sim 0.1$ is usually applied in the RED with satisfactory results obtained. We also point out that, for the purpose of discovering the states in conformational space, the results are not sensitive to $\tilde{t}^*$.

The variable $\tilde{t}$ actually reflects the timescale of the process that we are interested in. If it were set to the single trajectory length τ, then the trajectory weights solved by Eq. (6) will be equal to each other. In this case, we are focusing on the processes with a timescale much longer than τ, thus all the trajectories are still in the relaxation to local equilibrium, and should have the same weight. Usually, if a system can be divided into several metastable states under the focused timescale, $\tilde{t}$ should be selected comparable to or smaller than the smallest life time of these states. Given that the kinetic states of the system are frequently unknown, $\tilde{t}$ should be selected relatively short. In practice, it is possible to use shorter and shorter $\tilde{t}$ and check the convergence of the obtained free energy surface. Thus, a proper $\tilde{t}$ may be found as a compromise between correctness and statistical reliability (small fluctuation of trajectory weights). Another practical consideration is that a few percent of τ would be the safe choice for $\tilde{t}$. This is because the transition events with timescales prominently smaller than τ have reached equilibrium in a single trajectory (thus do not need reweighting), we only need to focus on the ones with timescales similar to or slightly smaller than τ. We usually choose $\tilde{t}/\tau$ in the interval of $[0.01, 0.1]$, and the results are satisfactory. In this paper, we will explicitly show that the variable $\tilde{t}^*$ within the range is not sensitive to our results.

2.3. Improvement of the RED

For decreasing statistical errors in the RED, we can generate a few trajectories from each initial configuration, and specify the same weight to the trajectories from the same initial configuration. Suppose there are n_s initial configurations (denoted as $q_{10}, q_{20}, \ldots, q_{n_s 0}$ in the following), and $m_1, m_2, \ldots, m_{n_s}$ trajectories initiated respectively from these initial configurations (the total number of trajectories is still p), we have

$$\sum_{k=1}^{n_s} \left(\sum_j G_{ij} \tilde{T}_{jk} \right) w_k = 0, \tag{11}$$

where w_k is the weight for all the m_k trajectories initiated from q_{k0}, $\tilde{T}$ is a $p \times n_s$ matrix with zero and unity elements. $\tilde{T}_{jk}$ is unity only if the j-th trajectory starts from the k-th initial configuration. Since the number of equations is more than the number of trajectory weights, we estimate Eq. (11) by minimizing $I = |G\tilde{T}\boldsymbol{w}|^2$. The vector of trajectory weights, $\boldsymbol{w}$, corresponds to the ground state of matrix $H \equiv \tilde{T}^{\text{T}} G^{\text{T}} G \tilde{T}$. For the newly defined H matrix, the number of the degenerated ground states still corresponds to the number of disconnected metastable states under the timescale of the individual trajectory. The existence of a small fraction of transition trajectories can slightly split the ground state of H. We name this method SI1.

Alternatively, we can merge the trajectories with the same initial configuration to a single one, and perform the standard RED analysis. Concretely speaking, we can calculate the average of a certain variable in the initial-segment sample set, init+, by,

$$\langle A\rangle_{\text{init}+} = \frac{1}{n_s}\sum_{i=1}^{n_s}\langle A\rangle_{q_{i0}^+}, \tag{12}$$

where $\langle A\rangle_{q_{i0}^+}$ means the average of A over the initial segments of all the trajectories with the initial configuration q_{i0}. By doing so, more samples are included in calculation, and effectively, both the initial region and the whole area in the conformational space which the trajectory explores become statistically more reliable. This method is named SI2. We will illustrate the application of SI1 and SI2 for depressing the large statistical noise in the entropy-dominated systems where the differences between trajectories within a metastable state might be in comparison with that in different states.

3. Results

3.1. Toy models

We first illustrate the RED in two simple toy models. The first one is a two-dimension (2D) system with four identical potential wells, denoted as Two-Dimension-Quad-Well (TDQW), see Fig. 1(a). The second model is a 2D system with a Mexican–Hat potential (MHAT) which has two states, see Fig. 1(b). The outer state is wide (larger entropy) and the relaxation time inside the state is comparable to the transition time between the state and the inner state. Simulation details and the applied basis functions are shown in Appendix A.

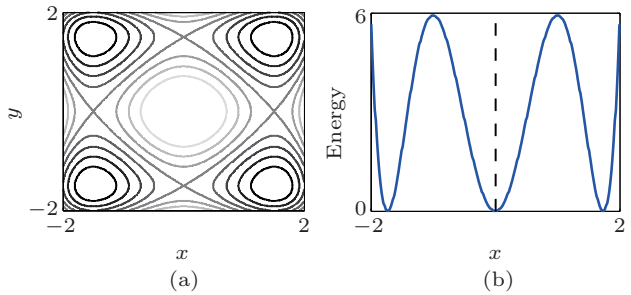


Fig. 1. (color online) The 2D potentials. (a) The TDQW potential shown as contour map. The darker regions correspond to potential wells, and the lighter regions correspond to potential barriers. (b) The MHAT potential projected along one dimension. Rotating around the dashed line reproduces the 2D potential.

3.1.1. Multiple-state transition network

In TDQW system, the eigenvalues of the matrix H in Eq. (6) are plotted in Fig. 2(a). As expected, there are four zero eigenvalues at temperature 0.3, implying the four disconnected states. At higher temperatures, the second to fourth eigenvalues deviate from zero to larger values due to the increasing fraction of the interstate transition trajectories. More illustrative perspective can be obtained from Figs. 2(b), 2(c), and 2(d), where the trajectories are mapped into the three-dimensional space by projecting the trajectory-mapped vectors, $\{\boldsymbol{G}^i\}$ defined in Eq. (7), to the second, third, and fourth eigenvectors of H, denoted as $L_\alpha^i \equiv \boldsymbol{G}^i \cdot \boldsymbol{u}_\alpha$, where $\boldsymbol{u}_\alpha$ is the α-th eigenvector of H. At temperature 0.3, the trajectories agglomerate into four disconnected clusters that locate respectively at the four vertexes of a tetrahedron. Inside each cluster are the trajectories isolated in the same potential well. At temperature 0.85, part of the trajectories can climb over the potential barriers to the other states. As a result, the points corresponding to the transition trajectories are found to gather along the straight lines that shoot out of the vertexes, and the ones corresponding to the non-transition trajectories are still locating on the vertexes. In Fig. 2(c), each vertex only connects to two neighboring vertexes, correctly reflecting the topology of TDQW system. At temperature 1.5, it is relatively easier for the trajectories to transfer between different states. The tetrahedron structure is still preserved, and its interior is filled by multiple-transition trajectories which transit among more than two states.

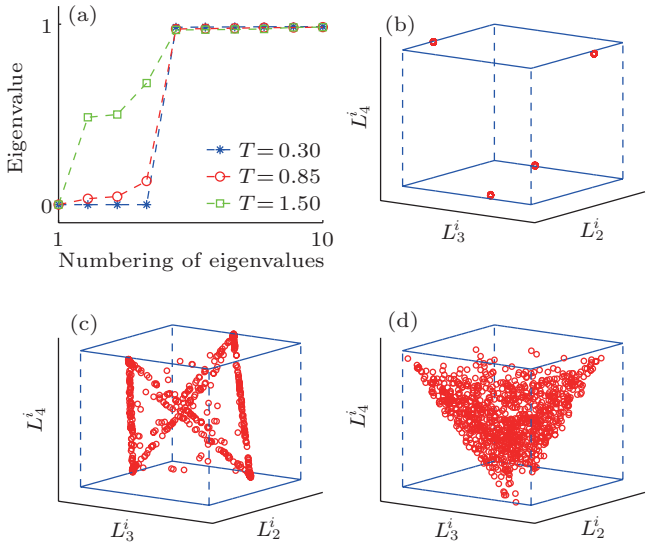


Fig. 2. (color online) The eigenvalues and projection maps for TDQW potential under different temperatures. (a) The first ten eigenvalues of H matrix calculated at different temperatures. For the three temperatures of 0.3 (b), 0.85 (c), and 1.50 (d), the projection of $\{\boldsymbol{G}^i\}$ to the second, third, and fourth eigenvectors of H are also plotted.

3.1.2. Entropy-dominated states

In MHAT system, the 70 smallest eigenvalues of H are shown in Fig. 3(a). Different from TDQW, the eigenvalues of H gradually increase from zero at the low temperature $T = 0.05$, and lots of small eigenvalues exist. The number of the eigenvalues that obviously deviate from 1.0 is always more than the number of physical states in MHAT system, till the temperature is high enough that there is only one zero-eigenvalue apparently different from 1.0. This behavior simply reflects the entropic nature of the outer potential well. At

low temperature, a single trajectory cannot reach local equilibrium in the state, consequently a few small eigenvalues exist, as there are many sub-states in the well. At higher temperatures, the trajectories that start from the outer potential well usually climb over the potential barrier before reaching local equilibrium within the state, thus the inter-trajectory difference will not be extinguished, until the two states in MHAT kinetically become one state at high enough temperature. The projections of $\{G^i\}$ onto the eigenvectors of H are shown in Figs. 3(b), (3c), and 3(d). At both temperatures 0.05 and 0.30, the trajectories are divided into two groups. In one group, the trajectories dispersively locate on the same plane which is perpendicular to the L_2^i axis. These trajectories are found to be isolated in the outer potential well. In the other group, the trajectories concentrate to nearly one point. They are found to be isolated in the inner potential well. These two groups of trajectories are clearly differentiated by the projections along the second eigenvector of H, i.e., the one indicating the most important contribution to ergodicity broken. At $T = 0.05$, each trajectory only explores a very limited region in the outer state, leading to the distribution on the perimeter of a circle. At $T = 0.30$, since each trajectory can explore a larger area, the circle has been filled with points. At $T = 0.85$, the trajectories are found to locate in a cone. The points in the bottom surface and the vertex of the cone correspond to the non-transition trajectories inside the outer and inner potential wells, respectively, and the other points are transition trajectories between the two states. Once again, the trajectory-projected space clearly provides the topologies of this system.

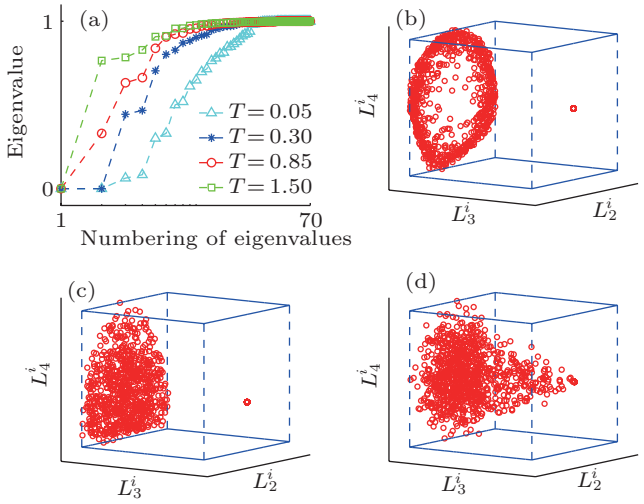


Fig. 3. (color online) The eigenvalues and projection maps for MHAT potential under different temperatures. (a) The first seventy eigenvalues of H matrix calculated at different temperatures. For the three temperatures of 0.05 (b), 0.30 (c), and 0.85 (d), the projection of $\{\boldsymbol{G}^i\}$ to the second, third, and fourth eigenvectors of H are also plotted.

3.1.3. Discussion

In the two 2D models, triangle functions in the 2D space are applied as basis functions, and the details are presented in Appendix A. We did the S^2 analysis mentioned in Subsection 2.2 at $T = 0.85$. For both systems, S^2 initially grows fast with increasing number of basis functions, then the growth slows down and shows a saturation behavior of S^2 as expected [see Figs. 4(a) and 4(b)]. Once S^2 has approximately reached the platform region, more basis functions will not change the calculation results anymore (data not shown). Thus, for TDQW and MHAT systems, about 24 basis functions are sufficient.

We also study the effect of free parameter $\tilde{t}$ in the RED with the simulation data of MHAT at $T = 0.85$. In this case, since there is only one zero eigenvalue of H, the whole free energy surface can be reconstructed. We uniformly divide the whole conformational space into small cells, $k = 1, \ldots, N_p$ (N_p is selected as 1600 here, i.e., 40 bins along each dimension), and the distribution probability in each cell is reconstructed with different $\tilde{t}^* = \tilde{t}/\tau$. We define the function $\delta_{\text{dist}}(\tilde{t}^*)$ as

$$\delta_{\text{dist}}(\tilde{t}^*) = \sqrt{\frac{1}{N_p}\sum_{k=1}^{N_p}(P_k(\tilde{t}^*) - P_k^{\text{theo}})^2} \tag{13}$$

to measure the difference between the reconstructed distribution and the theoretical distribution. Here $P_k(\tilde{t}^*)$ is the distribution in the k-th cell which is reproduced by the RED using $\tilde{t}^*$. P_k^{theo} is the theoretical distribution in the cell. As shown in Fig. 4(c), $\delta_{\text{dist}}(\tilde{t}^*)$ indeed decreases with shorter $\tilde{t}^*$ in calculation. For $\tilde{t}^*$ in the $[0.01, 0.1]$ interval, the $\delta_{\text{dist}}(\tilde{t}^*)$ values are almost the same, and are prominently smaller than the value of the non-weighted distribution of the simulation data. In the non-weighted (all w_i is unity) distribution, the major deviation from theoretical distribution resides in the inner potential well of MHAT, where the distribution is too high for the non-weighted data. In the reweighting process, the trajectories that start from the inner potential well are usually specified lower weights [see Fig. 4(c), inset], thus the over occupation in the inner state can be offset. The other variation in trajectory weights further adjust the intrastate distribution for both the inner and the outer states.

3.1.4. Improvement of statistical errors

We test the statistics improvement methods, SI1 and SI2. In here and below, we always randomly select one-fifth of the initial configurations, then five trajectories are spawned from each configuration to test the improvement. At $T = 0.85$, we apply both SI1 and SI2 to reconstruct the energy surface of MHAT. From Fig. 4(c), it can be seen that both methods correctly reproduce the energy surface with even smaller deviation from the theoretical one, as compared to the standard RED. At $T = 0.05$, the eigenvalues of the H matrix are calculated with SI2 method. As shown in Fig. 4(d), no matter whether one trajectory or multiple trajectories are generated from one initial configuration, the eigenvalues calculated by the standard RED are almost the same. However, analyzing

by the SI2 method significantly alters the structure of eigenvalues. By generating more trajectories from single initial configuration, and combining these trajectories together in analysis, the statistics of the single trajectory are prominently improved, and the difference between trajectories within the outer state are reduced to a large extent. As a result, some of the eigenvalues increase, in particular for the ones deviating but close to 1.0. In other words, the SI2 method has depressed the statistical noise in the spectral analysis of H, leaving the physically relevant states recognized.

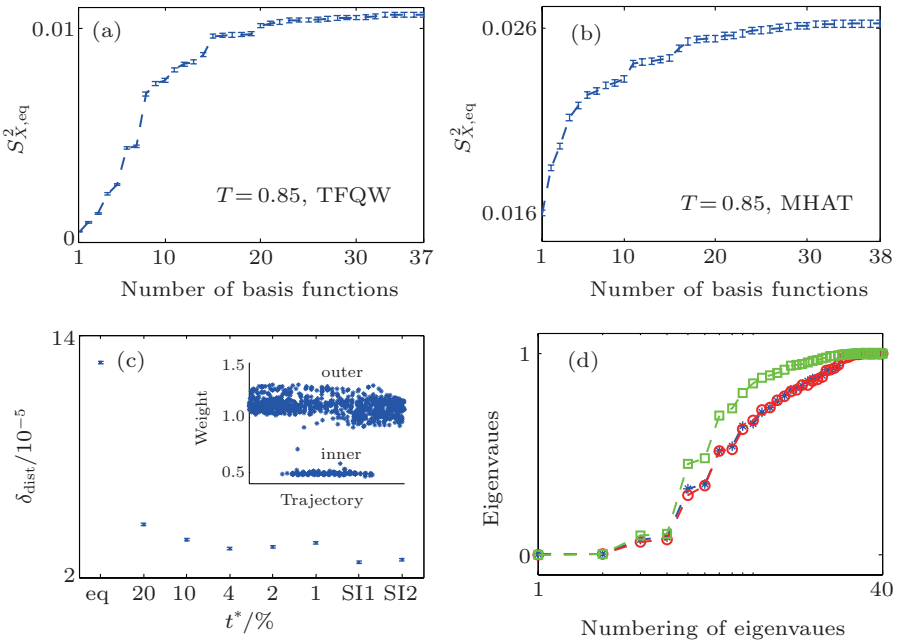


Fig. 4. (color online) The S^2 curve, effect of $\tilde{t}$, and illustration of statistics improvement. The behavior of S^2 with increasing number of basis functions are shown in panel (a) (TDQW) and panel (b) (MHAT). (c) The difference between the reconstructed equilibrium distribution and the theoretical distribution are shown for MHAT potential at $T = 0.85$. The horizontal axis denotes the method for reproducing the equilibrium distribution. 'eq' means the results without reweighting (i.e., $w_i = 1$), numbers mean the selected $\tilde{t}^* = \tilde{t}/\tau$. 'SI1' and 'SI2' label the two methods for statistics improvement, respectively. The trajectory weights calculated with $\tilde{t}^*$ equal to 0.01 are plotted as the inset. In panels (a), (b), and (c), the statistical errors are also shown as error bars. (d) The first 40 eigenvalues of H for MHAT potential at $T = 0.05$. The ones calculated with normal data set (stars), multiple trajectory data set (one-fifth initial configurations and five trajectories for each initial configuration) with the standard RED (circles), multiple trajectory data set with the SI2 (squares) are shown.

3.2. Solvated alanine dipeptide

The simulated system together with the initial configurations of simulation are shown in Fig. 5, where various notations of the atoms, dihedral angles, and free energy wells of this molecule are introduced. More details of simulations are presented in Appendix B. We use two different classes of basis functions to focus on the effects of different physical variables on the observed metastable states.

3.2.1. Dihedral angles

With thorough investigation, the two dihedral angles ϕ and ψ are the major coordinates to characterize the conformational motion of the alanine dipeptide molecule. As in the previous study,[35] we adopted the first-to-fourth order two-dimensional trigonometrical functions of ϕ and ψ as the basis functions. The 10 smallest eigenvalues of H are shown in Fig. 6(a). The smallest zero eigenvalue is intrinsic due to the construction of H. The second eigenvalue corresponds to the kinetic separation between the region with $\phi > 0$ (containing C_7^{ax} and α_{L}) and the region with $\phi < 0$ (containing C_7^{eq} and α_{R}). Consequently, by projecting the trajectory-mapped vectors $\{\boldsymbol{G}^i\}$ to the second eigenvector of H, i.e., through $\{L_2^i\}$, the trajectories can be clearly classified into the abovementioned two regions. The third eigenvalue of H corresponds to the partially kinetic connection between the C_7^{eq} and α_{R} states. Only a fraction of the trajectories in C_7^{eq} can climb over the free energy barriers to the α_{R} state, and *vice versa*. Thus, we may use the $\{L_i^3\}$ to classify the trajectories that are confined in the super-state involving C_7^{eq} and α_{R} into three groups. Since the two potential wells are well defined with their internal relaxation time smaller than the transition timescale, we can manipulate the trajectories to predict the transition times and locate the transition state ensemble.[35] The fourth smallest eigenvalue of H is also smaller than 1.0. Although it no longer indicates a new state in the conformational space and should be taken as the reflection of statistical difference between trajectories, we recently find that, by observing $\{L_i^4\}$, two trajectories that transiently walk into the α_{L} region can be picked out. One of the two trajectories initially located in the α_{L} region, then quickly moved into C_7^{ax}, another just occasionally visited α_{L}. Actually, at the temperature of 300 K, the α_{L} region seems not very stable with the applied force field. The other eigenvalues of H are very close to 1.0 without observable meaningful information.

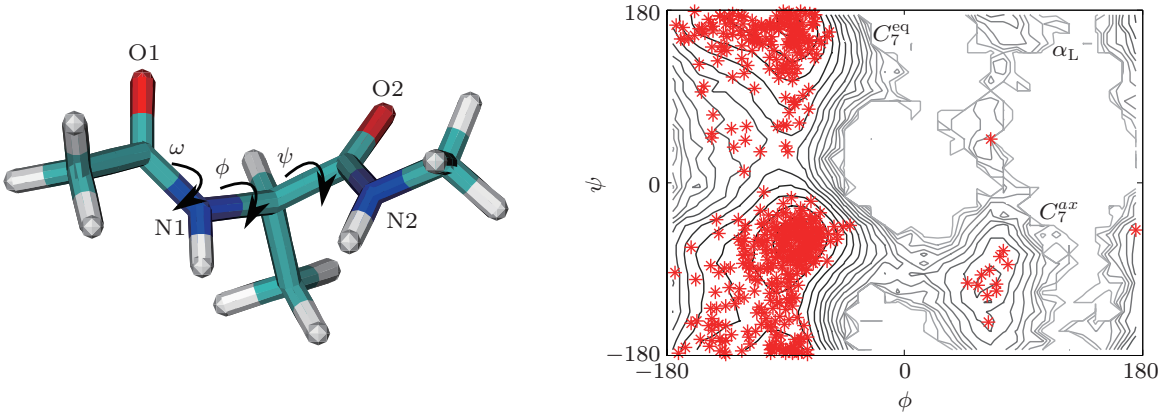


Fig. 5. (color online) Illustration of the alanine dipeptide molecule and the initial conformations of the RED simulation. The alanine dipeptide is shown in the left panel, with some dihedral angles and atoms labeled. The initial conformations projected to the ϕ–ψ plane are shown in the right panel. The background is the free energy surface at $T = 600$ K, with the free energy wells labeled.

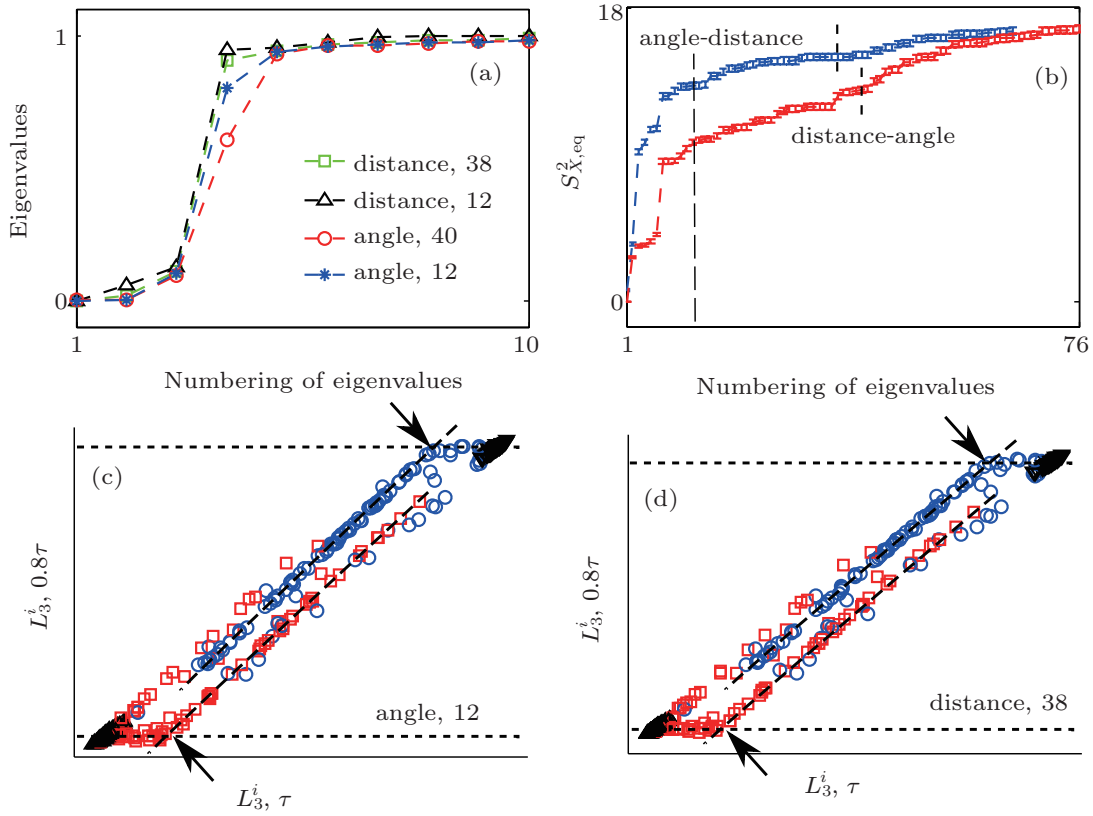


Fig. 6. (color online) RED analysis with the basis functions of dihedral angles and inter-atomic distances. Shown are the results for solvated alanine dipeptide, which is simulated at $T = 300$ K with the modified force field. (a) The eigenvalues of H matrix calculated with different sets of basis functions, including the full set of inter-atomic distances (squares), 12 inter-atomic distances (upward triangles), first-to-fourth order trigonometrical functions of ϕ and ψ (circles), and first-to-second order trigonometrical functions of ϕ and ψ (stars). (b) The behavior of S^2 is shown either with the inter-atomic distances (distance-angle) at first, or with the trigonometrical functions of ϕ and ψ at first (angle-distance). The $\{L_i^3\}$ values calculated with truncated trajectories (80 percent length) versus the ones calculated with full trajectories are shown. The results are obtained either with the first-to-second order trigonometrical functions of ϕ and ψ (c), or with the inter-atomic distances (d). The arrows label the intersection points between the dashed and dotted lines.

3.2.2. Distances between heavy atoms

We also test the effects of distances between atoms in the backbone of dipeptide by using them as basis functions. We compare two aspects, i.e., the eigenvalues of H and the manipulations on $\{L_i^3\}$ for extracting the kinetic information between C_7^{eq} and α_R. There are 10 non-hydrogen heavy atoms in alanine dipeptide. Except for the directly bonded atoms, 36 pairs of atoms exist, corresponding to 36 interatomic distances. Together with the two distances for the possible intramolecular hydrogen bonds (see Table 1), we select the 38 distances as basis functions. The resulting eigenvalues of H are shown in Fig. 6(a). While the three smallest eigenvalues

are almost the same as the results obtained with the functions of dihedral angles, the fourth eigenvalue is prominently lifted to a value closer to 1.0. However, the discrimination of the two α_L-visited trajectories is still possible. We plot the $\{L_3^i\}$ values calculated with the trajectories truncated to 80-percent length (i.e., the 0.2τ ending segment of the trajectories are discarded in analysis) versus the ones calculated with full trajectories[35] in Fig. 6(d). For comparison, the same figure generated with the first-to-second order two-dimensional trigonometrical functions of torsion angles is shown in Fig. 6(c). The two figures have quite similar structure. The points corresponding to the non-transition trajectories in α_R and C_7^{eq} are plotted as upward and downward (black) triangles, respectively. The points corresponding to the transition trajectories that start from α_R and C_7^{eq} are plotted as (red) squares and (blue) circles, respectively. The points along the inclined dashed lines correspond to the trajectories with only one time transition, and the points along the horizontal dotted lines correspond to the trajectories that do not happen to transition within the initial 0.8τ. Thus the intersection between the dotted and dashed lines corresponds to a single-transition trajectory with its transition happening exactly at 0.8τ. These points can be used to predict the times the transitions happen in the single-transition trajectories, consequently the transition state ensemble can be constructed. All the other points correspond to the trajectories with an even number of transitions which all happen within 0.8τ, or with early transitions occurring within $\tilde{t}$, or the multiple transition trajectories with transition happening both within 0.8τ and after 0.8τ. More details about detecting transition time thus transition conformation ensemble from the truncated-trajectory plots are presented in Ref. [35]. The similarity between Fig. 6(c) and Fig. 6(d) suggests both the torsion angles and the inter-atomic distances can be applied as basis functions.

Table 1. Basis function list.

Group	Number	Function
Internal freedom of alanine dipeptide	40	Two-dimensional trigonometric functions of ϕ and ψ
	8	N1–O1, N2–O2, N1H–O1, N2H–O2 and their squares[a]
Interaction between solvent and solute	2	Interaction[b] between peptide and water
	8	Interaction between O1, O2, N1, N2, and water
	2	Water number around peptide (3.3 Å, 5.5 Å)
	1	Solvent accessible surface area
	20	Water number around O1, O2, N1, N2 (3.3 Å, 3.7 Å, 4.1 Å, 4.5 Å, 4.9 Å)
Solvent structure	10	Integration of the OO radial distribution function, $g(r)$, of the bulk water[c], i.e., $\int_0^{r_0} r^2 g(r)\,\mathrm{d}r$ (r^0 is selected to be 2.9 Å, 3.3 Å, 3.7 Å, 4.1 Å, 4.5 Å, 4.9 Å, 5.3 Å, 5.7 Å, 6.1 Å, and 6.5 Å).

[a]The four atoms O1, O2, N1, and N2 are labelled in Fig. 5, left panel. N1H (N2H) is the hydrogen atom bonded to N1 (N2). O1–N1, and O2–N2 may form intra-molecule hydrogen bonds, or may be connected by water bridge.[38]
[b]In here and below, interaction means the electrostatic energy part and the Van de Waals energy part.
[c]These functions reflect the packing of bulk water.

3.2.3. Solvent molecules

For solvated biological macromolecules, the conformational change is sometimes correlated to the surrounding solvent molecules. In the RED, since the solvent-related properties can be selected as basis functions, it becomes possible to systematically one-by-one check and filter out the important collective variables. We study the dipeptide system with 91 basis functions to search for the solvent-related effects. These basis functions are classified into three groups, which respectively reflect the internal freedoms of alanine dipeptide (denoted as P), the solute-solvent relation (denoted as PW), and the structure of bulk solvent (denoted as W). The details of these basis functions are listed in Table 1. The eigenvalues of H are shown in Fig. 7(a). As can be seen, at $T = 300$ K, the inclusion of the solvent-related basis functions does not affect the spectral property of H. Particularly, the eigenvalues that are prominently smaller than 1.0 do not change, reflecting the unchanged physical division of the conformational space. Only a few eigenvalues pretty close to 1.0 are slightly lowered due to the added basis functions. The above results suggest that, for the current system, the solvent-related properties have almost reached equilibrium within the timescale of a single trajectory. Consequently, these properties contribute only a little to the inter-trajectory difference, and are unlikely to be strongly coupled with the dynamic process of the solute molecule under the simulated timescale.

3.2.4. Low temperature

We also apply the RED to study the low-temperature ($T = 150$ K) properties of the dipeptide. We note that, at this temperature, the physical phenomena are quite complicated; the force field and the simulation procedure that we adopt may not reflect the physical reality. Here, we are only meant to illustrate the application of the RED in the extreme condition, instead of providing physical insight into the system.

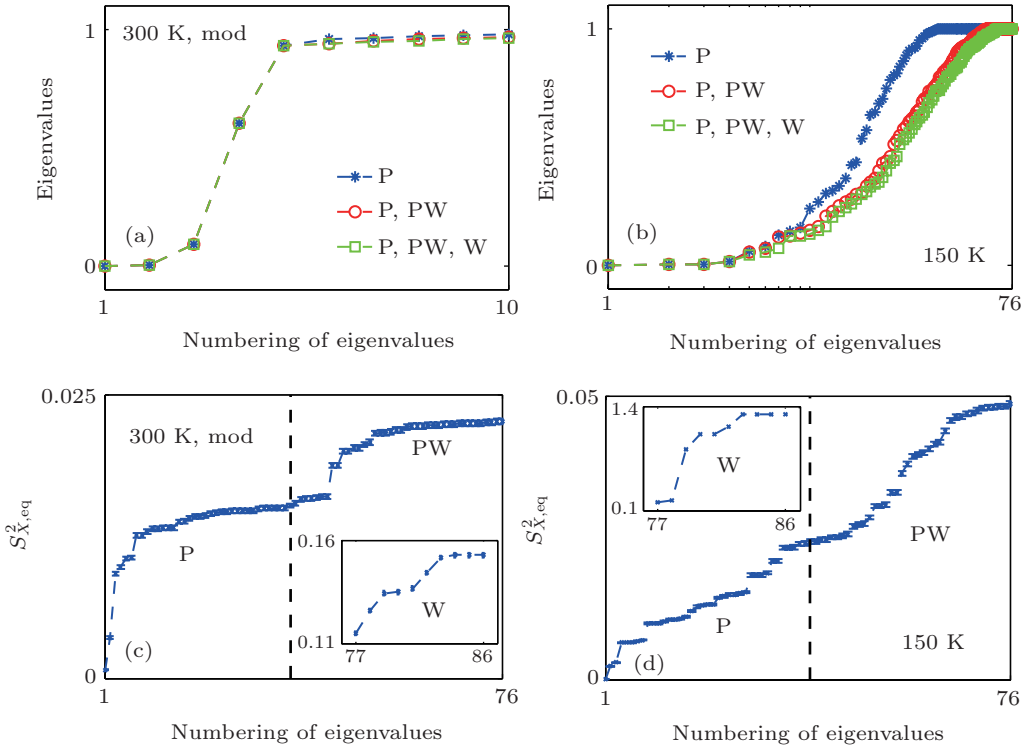


Fig. 7. (color online) The spectral properties of H calculated with the basis functions including solute–solvent relation. Shown are the results for the system of solvated alanine dipeptide. The eigenvalues of H calculated with different groups of basis functions are shown in panel (a) (300-K simulation with modified force field) and panel (b) (150-K simulation with standard forcefield). The functions used in calculation are labeled in the legend. The symbols 'P', 'PW', and 'W' mean the three sets of basis functions included in calculation, the internal degrees of freedom of dipeptide, the solvent–solute interaction, and the bulk water; see the main text for more details. The behavior of S^2 with increasing number of basis functions are shown in panel (c) (300 K with modified force field) and panel (d) (150 K with standard force field). The insets of panels (c) and (d) show the part of S^2 curve that is related to the functions of solvents.

The eigenvalues of the H matrix are calculated with the 91 basis functions listed in Table 1, and are plotted in Fig. 7(b). The solvent-related quantities do affect the eigenvalues of H now, leading to even more small eigenvalues. Correspondingly, the S^2, as shown in Fig. 7(d), almost always increases with the number of basis functions. At such a low temperature, the free energy surface becomes quite rough, and the diffusive dynamics[39] in the conformational space slows down seriously. As a result, the simulation trajectories may only explore a small region in the conformational space. Some degrees of freedom, like the solvent motion around the solute molecule and the conformational motion perpendicular to the ϕ–ψ plane, are no longer equilibrated in a single trajectory. Thus, the inter-trajectory difference becomes larger, more states are detected automatically, and more basis functions are required to expand the rugged distribution of the samples, leading to the abnormal behavior of the eigenvalues of H and S^2.

Manipulating the eigenvectors of H can provide further information of the simulation data. The projections of $\{\boldsymbol{G}^i\}$ to the second and the third eigenvectors of H pick out two completely separated states, C_7^{ax} and the super-state containing α_{R}, and C_7^{eq} [Fig. 8(a), at the current low temperature, the ϕ and ψ angles may not be enough to characterize the conformational motion, thus the notations, C_7^{ax}, α_{R}, and C_7^{eq}, are used mainly for the purpose of illustration]. With the aid of the histogram of $\{L_3^i\}$ [see Fig. 8(a), inset], we can roughly classify the trajectories inside the super-state into three groups, i.e., the trajectories in the leftmost two bins, in the rightmost two bins and the remaining ones. The conformational distribution of trajectories in different groups are shown in Fig. 8(c). The two groups of the leftmost or rightmost bins both possess single peak distribution, either in C_7^{eq} region or α_{R} region, just like the non-transition trajectories identified in the simulations at higher temperature. In contrast, the trajectories in the remaining group have a significant occupation in the intermediate region between C_7^{eq} and α_{R}. Obviously, these trajectories are trapped at the intermediate region due to slow dynamics, and are different from the transition trajectories identified with the similar procedure at 300 K. We plot the $\{L_3^i\}$ calculated with 80 percent trajectory length with the ones with full trajectory length in Fig. 8(c). The behaviors presented in Fig. 6 no longer exist. In terms of kinetics, the transition between the two regions, α_{R} and C_7^{eq}, can neither be approximated as Markovian process nor described with the first-order transition process any more. Consequently, with the current simulation data, it is only possible to reconstruct the intrastate equilibrium distribution in the local regions of conformational space, or to study the more localized kinetic behavior.

Due to the numerous local minimums at the low tem-

perature, the α_R and C_7^{eq} regions are just like two entropic-dominated states. We also test the statistics improvement method in this system. With one-fifth initial conformations and five trajectories from each initial configuration, the inter-trajectory difference, thus the statistical noise, is indeed suppressed as shown in Fig. 8(d).

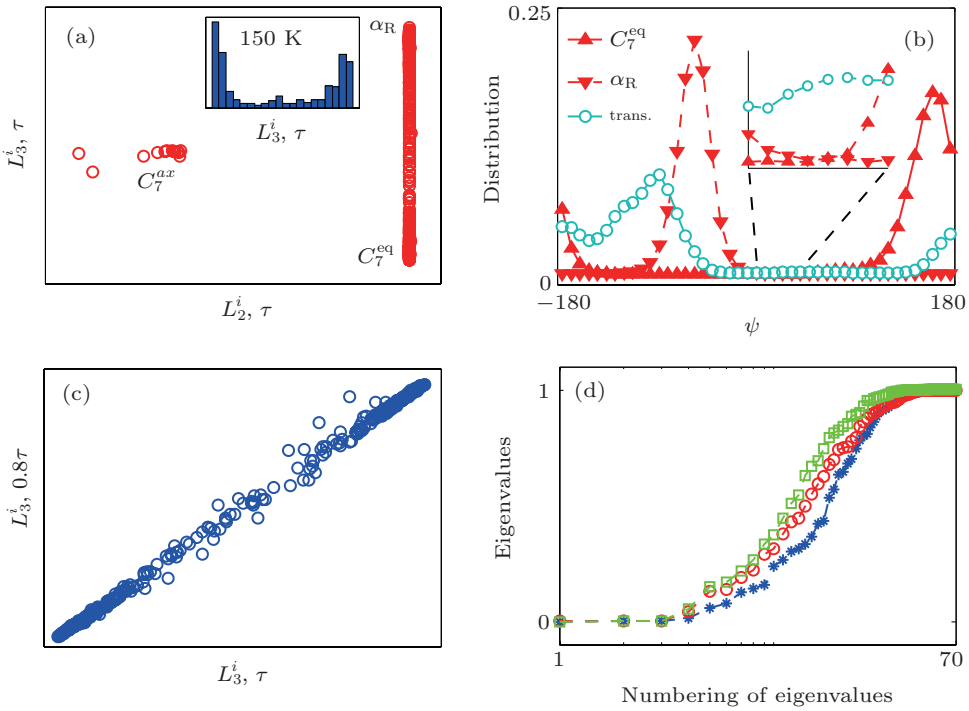


Fig. 8. (color online) Results for the low-temperature simulation of alanine dipeptide. (a) The projection of trajectories into the space of L_2^i and L_3^i. The inset is the histogram of $\{L_3^i\}$ values. The different states, into which the trajectories are roughly classified, are labeled. (b) The distributions along the ψ angle in the groups are shown in panel (a). Shown are the distributions for the trajectories in the leftmost two bins (downward triangles), the rightmost two bins (upward triangles) and the other trajectories (circles). (c) The $\{L_i^3\}$ values calculated with truncated trajectories (80 percent length) versus the ones calculated with full trajectories. (d) The first 70 eigenvalues of H calculated with normal data set (blue stars), multiple trajectory data set (one-fifth initial configurations and five trajectories for each initial configuration) with the standard RED (red circles), multiple trajectory data set with SI2 method (green squares) are shown.

As a summary, in the dipeptide system at $T = 300$ K with modified potential, we find the redundant basis functions (including functions of both the dihedral angles and the interatomic distances) are consistently treated in the RED, and the inclusion of the solvent-related basis functions gives out the expected results. At pretty low temperature of 150 K, the roughness of the free energy surface is reflected in the results of RED. The conformational space can still be roughly divided into parts, however, the kinetics becomes too complicated to be unambiguously characterized.

4. Conclusion

As a systematic method for exploring the conformational space, the RED does not require much *a priori* knowledge and approximation, thus its results can honestly reflect the thermodynamic and kinetic information of the system. For biological macromolecules, the usually expected hierarchical structure can be extracted out step-by-step. For these complex systems, the structure of the conformational space may be much more complicated, these examples studied here become quite relevant for physically interpreting the results. Along with the applications, we design the S^2 quantity for checking the completeness of the selected basis functions, and the schemes for statistics improvement. These tactics are also proved to be helpful for portraying the conformational space.

Acknowledgments

Zhou Xin is grateful for the financial support of the Hundred of Talents Program in the Chinese Academy of Sciences. Gong Lin-Chen and Zhou Xin are also grateful for the support of the Max Planck Society and the Korea Ministry of Education, Science and Technology; part of the work was done while they worked at the Asia Pacific Center for Theoretical Physics, Korea under this support.

Appendix A: Toy models

A1. Simulation details

For the 2D toy models, TDQW and MHAT, simulations are restricted in the square region of $[-2,2] \times [-2,2]$. Reflecting boundary condition is imposed. The over-dampened

Langevin equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{1}{\gamma}\frac{\partial U(x,y)}{\partial x} + \sqrt{\frac{2k_{\mathrm{B}}T}{\gamma}}\xi_1(t),$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{1}{\gamma}\frac{\partial U(x,y)}{\partial y} + \sqrt{\frac{2k_{\mathrm{B}}T}{\gamma}}\xi_2(t) \qquad \text{(A1)}$$

are adopted to generate the dynamical trajectories. Here γ is the frictional coefficient. $\xi_1(t)$ and $\xi_2(t)$ are the white noises satisfying $\langle \xi(t)_i \xi(t')_j \rangle = \delta(t-t')\delta_{ij}$ with $\langle\ \rangle$ denoting the ensemble average of noise. We simply take the Boltzmann constant k_{B} and γ as unity to get the dimension-reduced units for time, position, and temperature.

The potential of TDQW is

$$U(x,y) = \begin{cases} \infty, & |x|>2 \text{ or } |y|>2, \\ 5\left(\frac{4}{9}x^2-1\right)^2 + 5\left(\frac{4}{9}y^2-1\right)^2, & |x|\le 2 \text{ and } |y|\le 2. \end{cases}$$

The potential of MHAT is

$$U(r) = 40\cdot\left(\frac{1}{27}r^6 - \frac{2}{9}r^4 + \frac{1}{3}r^2\right), \quad r^2 = x^2+y^2.$$

For both systems, at each temperature, 900 initial configurations are randomly selected, then the RED simulations are performed, with each trajectory simulated for 50 dimensionless time units. The coordinates are recorded every 0.01 time units. The $\tilde{t}/\tau$ is selected as 0.04, if not further mentioned. In TDQW system, three temperatures of 0.30, 0.85, and 1.50 are discussed. For MHAT system, the RED simulations are performed at four temperatures of 0.05, 0.30, 0.85, and 1.50.

A2. Basis functions

We apply the following two-dimensional triangle functions

$$\begin{aligned} &\sin[(m+n)x], \cos[(m+n)x], \\ &\sin[(m+n)y], \cos[(m+n)y], \\ &m+n>0, \\ &\sin(mx)\sin(ny), \sin(mx)\cos(ny), \\ &\cos(mx)\sin(ny), \cos(mx)\cos(ny), \\ &m\ge 1, n\ge 1 \end{aligned} \qquad \text{(A2)}$$

as basis functions. Here, (x,y) are the 2D Cartesian coordinates multiplied by a transformation factor $\pi/2$, m and n are non-negative integers. We define the summation of m and n in Eq. (A2) as the order of these functions, and adopt the first-to-fourth order functions (in total 40 basis functions) in our analysis.

Appendix B: Solvated alanine dipeptide

For the system of explicitly solvated alanine dipeptide, we generated 500 MD trajectories at $T = 300$ K, each with $\tau =$ 600-ps length in the NAMD package[40] with CHARMM27 force field. In the simulations, a few potential barriers were added on the dihedral energy terms of the standard force field to kinetically separate the conformational space at the room temperature. The detailed simulation settings are the same as those in our previous work.[35] For the low temperature simulations, the 500 configurations selected from a 600-K simulation in the NAMD package are first relaxed by simulating at 150 K for 200 p*s*, then continuously simulate $\tau = 800$ ps, and collect configurations every 0.2 ps for subsequent analysis.

References

[1] Torrie G M and Valleau J P 1977 *J. Comput. Phys.* **23** 187
[2] Bartels C and Karplus M 1998 *J. Phys. Chem. B* **102** 865
[3] Maragakis P, van der Vaart A and Karplus M 2009 *J. Phys. Chem. B* **113** 4664
[4] Frantz D D, Freeman D L and Doll J D 1990 *J. Chem. Phys.* **93** 2769
[5] Berg B A and Neuhaus T 1991 *Phys. Lett. B* **267** 249
[6] Marinari E and Parisi G 1992 *Europhys. Lett.* **19** 451
[7] Lyubartsev A P, Martsinovski A A, Shevkunov S V and Vorontsov-Velyaminov P N 1992 *J. Chem. Phys.* **96** 1776
[8] Grubmüller H 1995 *Phys. Rev. E* **52** 2893
[9] Voter A F 1997 *Phys. Rev. Lett.* **78** 3908
[10] Miron R A and Fichthorn K A 2004 *Phys. Rev. Lett.* **93** 128301
[11] Zhou X, Jiang Y, Kremer K, Ziock H and Rasmussen S 2006 *Phys. Rev. E* **74** 035701
[12] Sugita Y and Okamoto Y 2000 *Chem. Phys. Lett.* **329** 261
[13] Xu S, Zhou X and Ou-Yang Z C 2012 *Commun. Comput. Phys.* **12** 1293
[14] Wang F G and Landau D P 2001 *Phys. Rev. Lett.* **86** 2050
[15] Yan Q and Pablo J J 2003 *Phys. Rev. Lett.* **90** 035701
[16] Kim J, Straub J E and Keyes T 2006 *Phys. Rev. Lett.* **97** 050601
[17] Laio A and Parrinello M 2002 *Proc. Natl. Acad. Sci. USA* **99** 12562
[18] Mitsutake A and Okamoto Y 2009 *J. Chem. Phys.* **130** 214105
[19] Shaw D E, Maragakis P, Lindorff-Larsen K, Piana S, Dror R O, Eastwood M P, Bank J A, Jumper J M, Salmon J K, Shan Y B and Wriggers W 2010 *Science* **330** 341
[20] Wu Y L, Xu X H, Yang X J, Zou S and Ren X G 2014 *Chin. Phys. B* **23** 028903
[21] Voter A F 1998 *Phys. Rev. B* **57** 13985
[22] Shirts M R and Pande V S 2001 *Phys. Rev. Lett.* **86** 4983
[23] Huang X H, Bowman G R and Pande V S 2008 *J. Chem. Phys.* **128** 205106
[24] Noé F and Fischer S 2008 *Curr. Opin. Struct. Biol.* **18** 154
[25] Becker O M and Karplus M 1997 *J. Chem. Phys.* **106** 1495
[26] Wales D J 2001 *Science* **293** 2067
[27] Krivov S V and Karplus M 2004 *Proc. Natl. Acad. Sci. USA* **101** 14766
[28] Rao F and Caflisch A 2004 *J. Mol. Biol.* **342** 299
[29] Gfeller D, Rios P D L, Caflish A and Rao F 2007 *Proc. Natl. Acad. Sci. USA* **104** 1817
[30] Noé F, Horenko I, Schütte C and Smith J C 2007 *J. Chem. Phys.* **126** 155102
[31] Chodera J D, Singhal N, Pande V S, Dill K A and Swope W C 2007 *J. Chem. Phys.* **126** 155101
[32] Prada-Gracia D, Gomez-Gardenes J, Echenique P and Falo F 2009 *PLos Comput. Biol.* **5** e1000415
[33] Bowman G R, Meng L and Huang X H 2013 *J. Chem. Phys.* **139** 121905
[34] Weber J K, Jack R L and Pande V S 2013 *J. Am. Chem. Soc.* **135** 5501
[35] Gong L C and Zhou X 2009 *Phys. Rev. E* **80** 026707
[36] Gong L C and Zhou X 2010 *J. Phys. Chem. B* **114** 10266
[37] Lu S J and Zhou X 2015 *Commun. Theor. Phys.* **63** 10
[38] Drozdov A N, Grossfield A and Pappu R V 2004 *J. Am. Chem. Soc.* **126** 2574
[39] Zwanzig R 1988 *Proc. Natl. Acad. Sci. USA* **85** 2029
[40] Phillips J C, *et al.* 2005 *J. Comput. Chem.* **26** 1781