Adaptive single replica multiple state transition interface sampling

Wei-Na Du and Peter G. Bolhuis

Citation: The Journal of Chemical Physics **139**, 044105 (2013); doi: 10.1063/1.4813777
View online: http://dx.doi.org/10.1063/1.4813777
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/139/4?ver=pdfcov
Published by the AIP Publishing

Adaptive single replica multiple state transition interface sampling

Wei-Na Du and Peter G. Bolhuis
Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, The Netherlands

(Received 29 April 2013; accepted 28 June 2013; published online 24 July 2013)

The multiple state transition path sampling method allows sampling of rare transitions between many metastable states, but has the drawback that switching between qualitatively different pathways is difficult. Combination with replica exchange transition interface sampling can in principle alleviate this problem, but requires a large number of simultaneous replicas. Here we remove these drawbacks by introducing a single replica sampling algorithm that samples only one interface at a time, while efficiently walking through the entire path space using a Wang-Landau approach or, alternatively, a fixed bias. We illustrate the method on several model systems: a particle diffusing in a simple 2D potential, isomerization in a small Lennard Jones cluster, and isomerization of the alanine dipeptide in explicit water. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4813777]

I. INTRODUCTION

Straightforward molecular dynamics simulations of rare events in complex systems usually take much longer than the available computer time. These long time scales are in many cases caused by high free energy barriers separating the metastable states. Examples include chemical and enzymatic reactions, the nucleation of crystal phases, and conformational transitions in biomolecular systems. To sample the configuration space many different techniques have been developed, such as umbrella sampling,[1] blue moon sampling,[2] local elevation,[3] flooding,[4] hyper-dynamics,[5] meta-dynamics,[6] adaptive bias force,[7] replica exchange,[8] simulated tempering,[9] integrated sampling,[10] orthogonal space sampling,[11] and many others. Such methods bias the Boltzmann distribution usually by biasing the dynamics itself, and hence do not directly yield kinetic information. The Bennett-Chandler approach[12,13] in principle corrects for this, but often requires *a priori* knowledge of the reaction coordinate, which might be limited or not available. Two-ended methods such as nudged elastic band,[14] action minimization,[15] the finite temperature string method,[16] or transition path sampling[17] avoid this problem. Transition path sampling (TPS) collects an ensemble of dynamically unbiased trajectories between two specified endpoints: the initial and final states of a rare event process. The path sampling methodology allows for evaluation of mechanisms, rates, and reaction coordinates, and as such yields insight in complex processes.[18] Since the initial development of TPS many advances have been made in the field, notably the introduction of the transition interface sampling (TIS),[19] in which a set of interfaces is introduced to enable efficient rate constant calculations. Other trajectory based (but not two-ended) methods include milestoning,[20] the forward flux sampling methodology,[21] adaptive multilevel splitting,[22] non-equilibrium umbrella sampling,[23] stochastic process rare event sampling,[24] and the Weighted Ensemble.[25]

Even though path sampling is very powerful in principle, in practice there are two challenges: (1) There is a dependence on the initial path which precludes a sampling of multiple parallel channels. (2) The presence of meta-stable intermediate states leads to excessively long paths. The development of the replica exchange transition interface sampling (RETIS)[26,27] solves the first problem to a large extent, while the latter problem is solved by the multiple state TPS and TIS (MSTPS/MSTIS).[28] The MSTPS/MSTIS method aims to construct a Markov state model in which a complex rare event process such as protein folding or conformational change is viewed as a series of transitions between meta-stable states, but suffers from slow switching between different types of pathways. Biasing the path types using a Wang-Landau algorithm turned out to be not very efficient.[29] A natural way of solving the low efficiency problem for sampling different types of pathways is to combine MSTIS with the replica exchange approach. In this method for every state a set of TIS interfaces is defined. All TIS runs are performed simultaneously, and are allowed to exchange pathways. In this way the entire path space is sampled. While we pursue this replica exchange MSTIS approach in Ref. 30 it is obvious that for large systems a replica exchange of many interfaces is not very practical to implement, due to the large number of interfaces that have to be treated simultaneously, and the disparity in length of the pathways, which makes a parallel implementation inefficient.

Based on ideas from simulated tempering,[9,31] we propose here to solve the low efficiency problem by employing a single replica that samples all interfaces in the MSTIS framework. Starting from an initial path in an initial interface, a Monte Carlo swapping move enables switching to a neighboring replica. When the path fulfills the criteria for the interface it will be accepted. Shooting moves will create new paths. Naturally, this approach would strongly favor the stable states, and will most likely not result in barrier crossings. To enforce such crossings, we can bias the distribution, e.g., by invoking the power of the Wang-Landau approach to create a *density of paths* at each interface. The Wang-Landau density of paths (DOP) will bias the paths away from the stable states and over to other basins of attraction. A multiple state

version of this approach will then allow sampling of all predefined metastable states. The DOP fills up all interfaces until the "histogram is flat." In the spirit of the Wang-Landau algorithm the increment of the DOP will be decreased and the sampling continues until convergence. This approach also enables an adaptive scheme in which stable states are added as the sampling of path space progresses. Because the dynamics of the trajectories themselves will be still unchanged, the resulting path ensembles when combined in a reweighting scheme will represent an (in principle exponential) acceleration of the dynamical trajectories, and yield the rate, free energy, and mechanism (see, e.g., Ref. 32). Convergence of the Wang-Landau algorithm is known to be slow, and therefore we also propose to use it only to build up the DOP, and then switch to a fixed bias based on the crossing probabilities. We note that the Wang-Landau methodology in combination with TPS has been proposed before.[29,33]

The paper is organized as follows. In Sec. II we give a brief review of the MSTIS methodology, and introduce the single replica algorithms – one based on the Wang-Landau approach, the second on the fixed bias – as well as an adaptive sampling scheme. In Sec. III these algorithms are applied to several model systems, a particle diffusing in a simple 2D potential, isomerization in small Lennard Jones clusters, and isomerization of the alanine dipeptide in explicit water. We end with concluding remarks in Sec. IV.

II. THEORETICAL BACKGROUND

A. Multiple state transition interface sampling

We discretize a dynamical trajectory (or path) $\mathbf{x}^L \equiv \{\mathbf{x}_0, \mathbf{x}_1, \ldots \mathbf{x}_L\}$ as a series of consecutive phase space points $\mathbf{x} = \{\mathbf{r}^N, \mathbf{p}^N\}$, with $\mathbf{r}$ and $\mathbf{p}$ the coordinates and momenta of the N-particle system, separated by a time Δt, such that the total time duration of the path is $\mathcal{T} = L\Delta t$. An unbiased trajectory $\mathbf{x}^L$ occurs with a probability that is proportional to[17]

$$\pi[\mathbf{x}^L] = \rho(\mathbf{x}_0) \prod_{i=0}^{L-1} p(\mathbf{x}_i \to \mathbf{x}_{i+1}). \tag{1}$$

Here, $\rho(\mathbf{x})$ is the steady state distribution (e.g., the canonical one) for a phase point $\mathbf{x}$. The Markov probability $p(\mathbf{x} \to \mathbf{y})$ for going from a state $\mathbf{x}$ to $\mathbf{y}$ within one time step depends on the underlying dynamics.[17] To normalize this probability we introduce a factor akin to a partition function

$$Z \equiv \int \mathcal{D}\mathbf{x}^L \, \pi[\mathbf{x}^L], \tag{2}$$

where the integration runs over all possible paths of all lengths, so that the normalized path probability becomes

$$\mathcal{P}[\mathbf{x}^L] = \pi[\mathbf{x}^L]/Z. \tag{3}$$

We consider a set of M states $\mathbf{S} = \{0, 1, \ldots I, \ldots M\}$ each with its own state definition. We denote the definition for state I as $\{\mathbf{x} : \lambda_I(\mathbf{x}) < \lambda_{0I}\}$, where $\lambda(\mathbf{x})$ is a function of the phase point $\mathbf{x}$, and the parameter λ_{0I} defines the boundary of state I. Typically, λ acts as a metric defining the distance of a point $\mathbf{x}$ to a reference point characterizing the stable state. The transition interface sampling (TIS)[19] method introduces for each state I $m_I + 1$ non-intersecting hyper-surfaces defined by $\{x : \lambda(\mathbf{x}) = \lambda_{iI}\}$. For convenience these interfaces are also described by the function $\lambda(\mathbf{x})$, although this is strictly not necessary. The $m_I + 1$ interfaces are defined by the ordered sequence $\lambda_{0I}, \lambda_{1I}, \ldots, \lambda_{mI}$ where the first interface λ_{0I} equals the boundary of state I, and the last one λ_{mI} is labeled the outermost interface. The TIS path ensemble for interface i of state I contains paths that start in I, end in any of the M states in $\mathbf{S}$ (including state I), and also cross interface λ_i. If we define the region of phase space beyond interface i by $\Lambda_{iI}^+ = \{\mathbf{x} : \lambda(\mathbf{x}) > \lambda_{iI}\}$, the TIS path probability for state I is

$$\mathcal{P}_{I\Lambda_{iI}}[\mathbf{x}^L] = \tilde{h}_i^I[\mathbf{x}^L]\pi[\mathbf{x}^L]/Z_{I\Lambda_{iI}}, \tag{4}$$

where the normalizing factor $Z_{I\Lambda_i}$ is again defined by $\int \mathcal{D}\mathbf{x}^L \mathcal{P}_{I\Lambda_i}[\mathbf{x}^L] = 1$. The indicator function $\tilde{h}_i^I[\mathbf{x}^L]$ for paths that begin in I, end in $\mathbf{S}$, and cross λ_i is

$$\tilde{h}_i^I[\mathbf{x}^L] = \begin{cases} 1 & \text{if } \mathbf{x}_0 \in I \ \wedge \ \mathbf{x}_L \in \mathbf{S} \wedge \\ & \forall\{j|0 < j < L\} : \mathbf{x}_j \notin \mathbf{S} \wedge \\ & \exists\{j|0 < j < L\} : \mathbf{x}_j \in \Lambda_{Im}^+ \\ 0 & \text{otherwise} \end{cases}. \tag{5}$$

This function is only unity if the path starts in I, ends in any stable state, does not visit another state in between, and crosses the interface. Since all paths have to start in I, the TIS path ensembles for each state are seemingly disconnected. They can be connected by introducing the multiple state path ensemble comprising all trajectories from all states $I \in \mathbf{S}$ that cross the outermost interface of I

$$\mathcal{P}_{\text{MSTIS}}[\mathbf{x}^L] = \sum_I^M \tilde{h}_m^I[\mathbf{x}^L]\pi[\mathbf{x}^L]/Z. \tag{6}$$

The path ensembles (4) and (6) can be sampled using the regular shooting algorithm.[19] The Metropolis acceptance criterion for the shooting algorithm is

$$P_{acc}[\mathbf{x}^L(o) \to \mathbf{x}^L(n)] = \tilde{h}_i[\mathbf{x}^L(n)] \min\left[1, \frac{L^{(o)}}{L^{(n)}}\right], \tag{7}$$

where o and n denote the old and new paths, respectively. An important quantity in the TIS framework is the crossing probability $P(\lambda|\lambda_{iI})$, which, for $\lambda > \lambda_{iI}$, is given by

$$P_I(\lambda|\lambda_{iI}) = \int \mathcal{D}\mathbf{x}\mathcal{P}_{I\Lambda_i}[\mathbf{x}^L]\theta(\lambda_{I,max}[\mathbf{x}^L] - \lambda), \tag{8}$$

where $\theta(\lambda)$ is the Heaviside step function. All crossing probability histograms $P_I(\lambda|\lambda_{iI})$ belonging to state I can be joined using the weighted histogram analysis method (WHAM),[34,35] to give a single crossing probability $P_I(\lambda|\lambda_{1I})$. Note again that for pathways connecting I with a different state, all interfaces are crossed. This results in a automatic saturation of the crossing probability to a plateau for arbitrary large λ values, rather than a decay to zero.

B. Single replica MSTIS

Employing a combination of the replica exchange formalism with the TIS algorithm (RETIS) improves the sampling efficiency dramatically.[26,27] In the RETIS scheme an

additional replica is needed to sample the stable states themselves. This so-called *minus*-interface ensemble is defined by

$$\mathcal{P}^-_{I\Lambda_1}[\mathbf{x}^L] = \tilde{h}^I_-[\mathbf{x}^L]\pi[\mathbf{x}^L]/Z^-_{I\Lambda_1}, \tag{9}$$

$$\tilde{h}^I_-[\mathbf{x}(L)] = \begin{cases} 1 & \text{if } x_0 \in \Lambda_1^+ \wedge x_L \in \Lambda_1^+ \wedge \\ & \quad \forall\{j|0<j<L\} : x_j \in I \\ 0 & \text{otherwise} \end{cases}. \tag{10}$$

Reweighting the path ensemble will then allow a complete description of the processes, and yield the rates, free energy landscape, and the mechanisms.[29] A drawback of the RETIS method for multiple states is the large number of interfaces that simultaneously need to be sampled. For instance, for 5 states, each with 10 interfaces (9 regular +1 minus interface), a total of 51 interfaces need to be sampled. This becomes cumbersome from a practical viewpoint, not only because of the large computational effort, but also because a parallel implementation has the complication of being asynchronous as all pathways have different lengths.

To avoid such problems, we can adopt a single replica approach.[9,31] In this approach a single replica walks along the interface space, while sampling paths by shooting and reversal moves. The sampling is interspersed with exchange moves between interfaces, resulting in a "flat" random walk across the interface space. To obey detailed balance, such an exchange move should be accepted with

$$P_{acc}(\mathbf{x}^L; \lambda_{iI} \to \lambda_{jI}]) = \tilde{h}^I_j[\mathbf{x}^L] \min\left[\frac{g(\lambda_{iI})}{g(\lambda_{jI})}\right], \tag{11}$$

where $g(\lambda_{jI})$ is the correct *density of paths* for each interface. This DOP on the interfaces will not be equal for the different interfaces but high close to stable states, and low close to the transition state region. In fact, as it turns out, the correct DOP is proportional to the crossing probability $g(\lambda_{iI}) \propto P_I(\lambda_{iI}|\lambda_{1I})$. This can be seen as follows. While an exchange to a lower interface is always possible, an exchange opportunity to a higher interface occurs with the naturally occurring probability for pathways at the higher interfaces, which, in fact, is the crossing probability. To obtain an equal population the exchange acceptance should therefore be biased with the ratio of the crossing probabilities. As an exchange between two interfaces belonging to the same state I are governed by the same crossing probability, the proportionality factor cancels.

Besides exchanges between interfaces belonging to the same state, we can attempt exchanging between states when the current replica is in the outermost interface, the multiple state ensemble. This state swapping can only be accepted when the path connects different states. After the swap the path is reversed in order to have it start in the correct initial state. The interface set, with the attached DOP, is also changed to the new state. We note that here the proportionality factor does not cancel, and that a direct identification between the crossing probabilities and the DOP will not result in a flat histogram sampling of the different states.

As in regular single replica exchange approaches developed for sampling phase space[9,31] the sampling is only correct when the biasing function is already known in advance. However, we can obtain the correct DOP in two ways: (1) build up the density of states using a Wang-Landau approach and (2) use a fixed bias based on the estimated crossing probability from a previous path sampling.

The Wang-Landau approach is useful because it allows to initialize the replicas from scratch. In Secs. II C and II D we therefore introduce the Wang-Landau approach first, followed by Sec. II E on fixed biases.

C. Wang-Landau biasing in single replica MSTIS

The original Wang-Landau scheme[36,37] builds up an estimate of the density of states $g(E)$ of a system via a Monte Carlo algorithm that aims for a flat sampling in energy. To achieve this the Monte Carlo random walk is performed with a probability proportional to the reciprocal of the density of states $1/g(E)$. The metropolis move for a symmetric move that changes the energy from $E^{(o)}$ to $E^{(n)}$ is then

$$P_{\text{acc}}\left[E^{(o)} \to E^{(n)}\right] = \min\left[1, \frac{g(E^{(o)})}{g(E^{(n)})}\right]. \tag{12}$$

Each time an energy E is visited a histogram $H(E)$ is recorded and the corresponding density of states is multiplied with a modification factor $f > 1.0$. This results in the histogram $H(E)$ becoming flat within a certain accuracy. The modification factor is then reduced, e.g., $f_{i+1} = \sqrt{f_i}$ and the histogram is reset. The sampling continues to update the density of states $g(E)$ until the new histogram is again flat. This procedure is continued until the modification factor f is smaller than the desired accuracy.

As was shown in Ref. 29 the basic idea of the Wang-Landau algorithm can also be applied to improved sampling of different types of pathways in the multiple state path ensemble. Here we extend that idea to bias instead of the path types, the DOP on the interfaces. A single replica dubbed the current interface λ_i performs a Monte Carlo walk across the interface space. A move from interface i to j is accepted with the criterion (11). The (logarithm of) density of paths on every interface visited is updated with f, and a histogram of visited interfaces is recorded. When this histogram is almost flat, the histogram is reset and the factor f reduced with a factor 2. In the outermost interface one can either exchange with the previous interface $m-1$ or a switch to another state. If the swap to a previously not visited state is accepted, the DOP for that set of interface is set to one (ln DOP to zero).

D. The Wang-Landau MSTIS algorithm

The Wang-Landau MSTIS algorithm consists of an initialization phase followed by a main production loop.

1. Initialization phase

1. Define the database of stable states I.
2. Define each stable state I by an interface λ_{0I}, and a set of TIS interfaces λ_{iI}.

3. Select one of the stable state configurations as an initial configuration.
4. Create an initial path using the minus interface move and set the current interface to 0 (the minus interface) or read in an existing path from the previous run and set the current interface accordingly.
5. Set all the DOPs to 1 (ln DOP to zero) or read in DOPs from a previous Wang-Landau run.

2. Main loop

Repeat one of the following moves in random order with specified frequency:

1. *A regular shooting move*: Select a random slice from the path, change its momenta, and integrate forward and backward in time until a stable state is found in both directions, or the path exceeds a maximum length. The path is accepted with the regular TIS Metropolis criterion $P_{acc} = \min[1, L^{(o)}/L^{(n)}]$. If rejected the old path is retained.
2. *A reversal move*: A time reversal is executed by reversing the order of the slices in the path, and reversing all velocities of all particles. This move can only be accepted for an I to I path.
3. *An interface exchange move*: Multiply the DOP of the current interface i by f (or add $\ln f$ to the logarithmic density). Attempt to move the current interface from i to j (usually the neighboring interface $j = i \pm 1$). For $1 \rightarrow 0$, perform a minus interface move by extending one of the paths termini, until the path exits the state again, and accept the move. For $0 \rightarrow 1$, extend one of the paths termini, until the path enters the state again, and accept the move. Otherwise, if the current path crosses the new interface the trial move is accepted via a Metropolis criterion $P_{acc} = \min[1, \frac{g(\lambda_{iI})}{g(\lambda_{jI})}]$.
4. *A state swap*: When the current interface is the outermost interface m there is a substantial chance for an I to J pathway. If so, an exchange to a J to I s is attempted. This exchange is accepted with the Metropolis criterion $P_{acc} = \min[1, \frac{g(\lambda_{mI})}{g(\lambda_{mJ})}]$. When accepted, the initial state changes from I to J. This means that the set of interfaces and the DOP of state J will replace that of I. The current interface will be the outermost interface m_J.

The Wang-Landau MSTIS algorithm is schematically illustrated in Fig. 1. While one is free to change the relative attempt frequency of each move, we found that an uniform attempt frequency gives good results. However, attempting more interface exchanges than shootings is computationally not expensive, and might result in (slightly) better diffusion along the interfaces.

E. Fixed bias in single replica MSTIS

A drawback of the Wang-Landau approach is the slow convergence, requiring longer and longer runs when f is reduced. Therefore, after the first round of sampling, it might be advantageous to switch to a fixed bias, which is based on the crossing probability itself rather than a (slowly) converging DOP. The crossing probability can be extracted from an initial round of Wang-Landau MSTIS sampling using WHAM.[34,35] The proportionality factors for the multiple state TIS ensembles can be extracted from the Wang-Landau histogram and be fixed. Since these proportionality factors only attempt to give equal weight to the different sets of paths starting in each state, a slight deviation from the correct value will only result in unequal numbers of paths for each state. This will not affect the crossing probabilities, the rates, or the mechanisms.

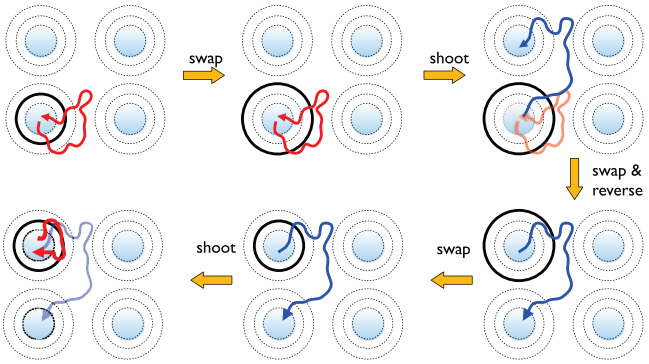

FIG. 1. Cartoon of the exchange of interfaces and state swaps for a 4-state system, where each state has two interfaces. The initial path in the left top corner crosses not only the current (black) interface, but also the outermost interface so an exchange is allowed. A shooting move creates a path toward another state, so that state swap is possible. The current interface is then the outermost interface of the new state. Another interface exchange and a subsequent shot creates the red path in the lower left corner cartoon.

F. Adaptive sampling

Our approach naturally allows for an adaptive version of the Wang-Landau MSTIS algorithm. When a state is discovered that is not in the current database of stable states, it can be added and sampled (see Fig. 2). We start by sampling pathways on the set of interfaces of an initial state using Wang-Landau MSTIS as described in Secs. II A–II E. At some point the pathways will escape the basin of attraction of this initial state, and will explore another stable state. At this moment the paths become very long, because the algorithm has not yet recognized that it has landed in a different stable state. The trajectory will be rejected because the maximum path length has been reached. At that point, we can analyze the trajectories for stability, for instance, by minimizing the potential energy to obtain a local minimum, or by performing a cluster analysis. This local minimum is again a (meta) stable state, characterized by an order parameter, such as the distance to the configuration of the minimum. We can add this state to the set of allowed states, and define a set of a allowed interfaces with it. The sampling is continued, adding more and more states, until convergence.

III. RESULTS AND DISCUSSION

A. Langevin dynamics in a simple 2D potential

In this section we apply the Wang-Landau MSTIS methodology to a simple 2D potential with multiple states.

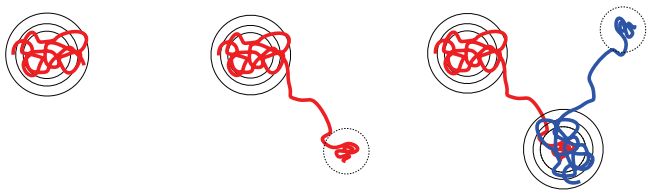

FIG. 2. Concept of the adaptive Wang-Landau MSTIS scheme. Sampling the outermost interface of the initial state (left) creates an activated path that finds another basin, which results in a very long trial trajectory (middle). This path is analyzed, and if the new state is stable it is treated as a new state with its own set of interfaces. This process is repeated, and new states are added to the database as needed (right).

The potential is defined as

$$V(x, y) = 10^{-5}(x^6 + y^6) + 3\sum_{i=1}^{6} e^{-2((x-x_i)^2-(y-y_i)^2)}, \quad (13)$$

where $\{x_i, y_i\}$ the location of the set of minima acting as attractors for the dynamics. The (randomly chosen) set of minima is $\{x_i, y_i\} = \{(0.852944, -1.95778), (-1.69449, -4.34425), (4.88473, 4.00666), (1.84771, 4.51368), (3.09019, -2.69672), (-4.64209, -2.80314)\}$. A plot of this potential is shown in Fig. 3.

A single particle moves in this potential according to full Langevin dynamics as detailed in Ref. 17, with a time step $\Delta t = 0.05$, and a friction of $\gamma = 1$. The inverse temperature $\beta = 1/k_B T$, with T the reduced (artificial) temperature and k_B the Boltzmann constant, determines how rare transition is between minima. As this is only a proof of principle we fix the temperature $\beta = 4$, to ensure a low rate constant.

Stable states are defined by a circle around the minima and comprise all points defined by $\{(x, y)|((x - x_i)^2 + (y - y_i)^2) < 0.1^2\}$. For each state I, the interfaces are defined by the metric $\lambda(x, y) = \sqrt{(x - x_I)^2 + (y - y_I)^2}$. In total 20 interfaces are put at $\lambda_i = \{0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1.0, 1.05, 1.1, 1.15\}$.

Wang-Landau MSTIS simulations were performed with an equal amount of shooting, interface exchange, reversal and state swap move attempts. We used the constrained one-way shooting move, which integrates only forward from the first crossing point of the path with the interface λ_i.[27] This integration is stopped when a stable state or the maximum path length is reached. This shooting move has an acceptance ratio of 1. Decorrelation of the paths is achieved via the interface exchange move. When the visiting histogram is flat within 10%, the Wang-Landau factor is decreased, and the histogram is reset. In addition, a minus interface is sampled to improve decorrelation in the stable states. The sampling is continued until convergence, i.e., a Wang-Landau factor $f < 1 \times 10^{-6}$, which required a total number of 1×10^8 shooting moves.

In Fig. 3 a random sample of paths is shown. Note that all states are visited. We stress that the dynamics is untouched, and the system is completely sampled according to the original dynamics.

The Wang-Landau MSTIS sampling decorrelates quickly. In Fig. 4 we plot the state index as a function of time. State swapping is relatively fast, since each couple of hundreds shots a different state is visited. For the converged system, within a few thousand shooting moves, all replicas are visited. When starting from a zero DOP it takes slightly longer, as the DOP has to be built up first. We plot in the same figure the distributions of the time interval (in shooting moves) between visits to the minus interface. The distribution has a Poisson shape, with a maximum around a few hundred shots. The inset shows that the distribution of intervals between state visits (without having visited the minus interface) is peaked at a much small number of shots, as expected.

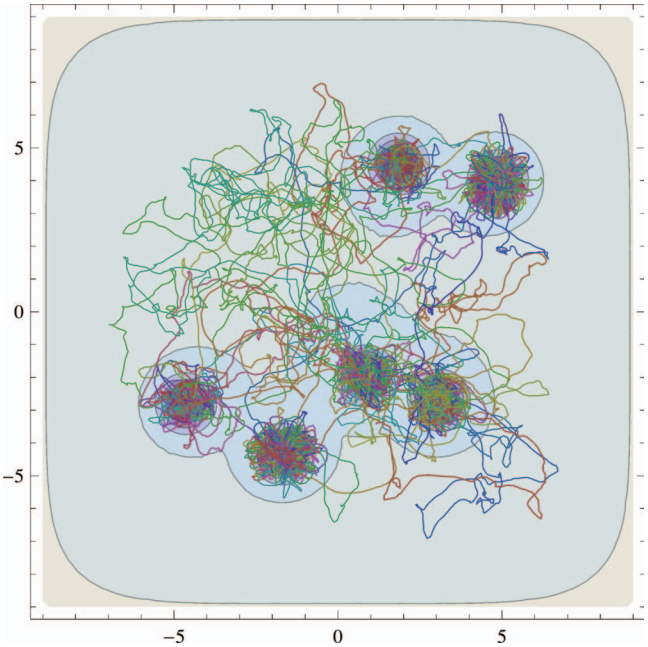

FIG. 3. The 2D potential with the 6 minima plotted in the x,y plane. Samples of the Wang-Landau MSTIS paths at $\beta = 4$ are drawn on top. The dynamically unbiased paths connect all states.

Fig. 5(a) shows the convergence of the DOP for all interfaces. To test whether our method is correct we conducted a RETIS simulation using the exact same settings and simulation duration. Fig. 5(b) plots the crossing probabilities for each state as a function of λ for both the RETIS and Wang-Landau MSTIS methods for $\beta = 8$. We conclude that both methods give the same statistics, as expected. Further confirmation of convergence can be obtained by comparing the DOP with the crossing probability. The DOP converges to the proper weight of the pathways in unbiased ensemble and hence should be equal to the crossing probability. In Fig. 5(c) we plot the comparison of the two, which indicates that indeed our sampling has converged.

Using the crossing probabilities we can compute the reweighted path ensemble (RPE),[29] and reconstruct the free energy (FE) landscape, as well as the committor surface for a single state (in this case state 3), both shown in Fig. 6. Note that the FE landscape naturally is equal to the potential as given by Eq. (13).

We can compute the rate constant matrix using the expression $k_{IJ} = \phi_{1I} P_I(\lambda_{mI}|\lambda_{1I}) P(\lambda_{0J}|\lambda_{mI})$. The first term, the

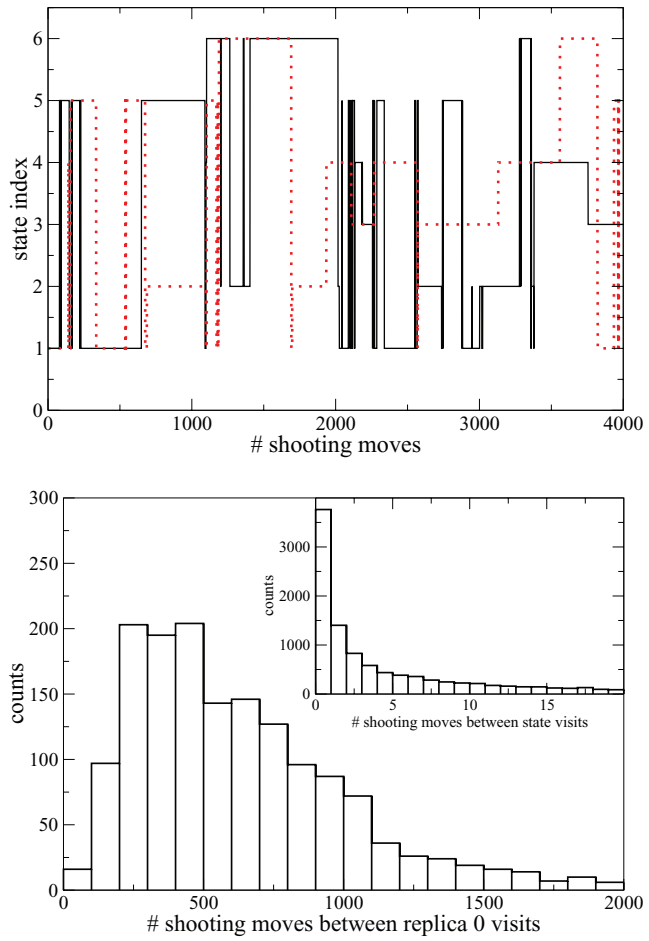

FIG. 4. Decorrelation in the Wang-Landau MSTIS sampling. Top: The state index as a function of simulation time measured in number of shooting moves. The dotted line shows the results when the DOP is initialized to zero and solid line shows the result for a converged DOP. Note that the decorrelation is almost equal. Bottom: The decorrelation between stable state minima. The main window shows the histogram of interval lengths between stable state visits (replica 0). The inset shows the histogram of interval lengths between stable state visits for the whole ensemble. Note that this interval is naturally much shorter than the one between replica 0, as the replica has to diffuse through the whole interface space.

flux $\phi_{1I} = \langle \tau_0 + \tau_1 \rangle^{-1}$, follows from the average sum of the path lengths in the minus interface and the first interface. The second factor is the crossing probability at the outermost interfaces, and the last term is given by the transition matrix, computed at the outermost interface. Using the rate matrix we can construct the time dependent population of the system using the expression $p_i = p_0 \exp(\mathbf{K}t)$, where p_0 denotes the (arbitrary) initial population distribution, and the diagonal terms of $\mathbf{K}$ are set to $k_{ii} = -\sum_{j \neq i} k_{ij}$. Taking the limit $t \to \infty$ we obtain the equilibrium population p_{eq}. This population is given by $p_{eq} = \{0.223, 0.168, 0.104, 0.154, 0.2046, 0.147\}$, which is close to the minima of the potential $p_{eq} = \{0.218, 0.1676, 0.107, 0.155, 0.208, 0.144\}$. The equilibrium population is dependent on the accuracy of the rate matrix, and hence can be used to assess the convergence properties of the method in a different way. In Fig. 7(b) we plot the equilibrium populations computed from the rate matrix as a function of simulation time, starting from the very initialization of the system. The Wang-Landau algorithm converges rather slowly, and needs to reduce the factor f almost fully to zero, before the populations are converged. As mentioned in Sec. II, a better convergence might be obtained if we do not rely on the Wang-Landau algorithm, but use the cumulative crossing probabilities as a fixed biasing function. As is shown in Fig. 7(a) the fixed biased approach converges much quicker.

B. Adaptive sampling on the LJ7 cluster

To illustrate the adaptive sampling method we choose a system with only a limited amount of states to prove that all states can be found automatically. A cluster of 7 LJ particles is ideal for this, since it has just four potential energy minima, given in Fig. 8.[38] At low temperature these minima are also the stable states. As an order parameter we can use the sum of the squared distances to a reference structure. Define for each particle i the sum of the distances in a structure x by

$$d_i(x) = \sum_{j \neq i}^{N} (r_i - r_j)^2, \tag{14}$$

where N is the number of particles, and r_i is the position of particle i. Then the squared distance to the reference structure x_I of stable state I is defined as

$$\Delta(x, x_I) = \sum_{i}^{N} |d_i(x) - d_i(x_I)|. \tag{15}$$

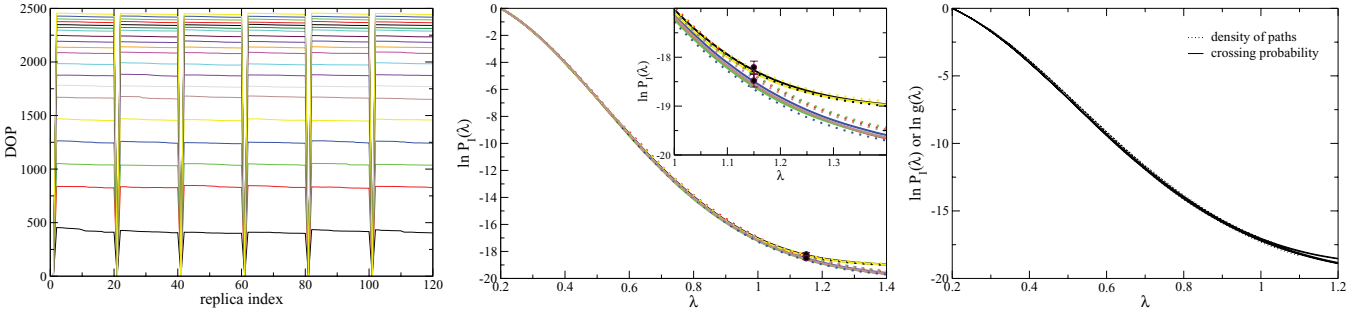

FIG. 5. Left: Convergence of the DOP in the course of the Wang-Landau MSTIS simulation. The replica index is defined as $Im + i$. Middle: Comparison of the logarithmic crossing probability $\ln P(\lambda)$ for the six states as a function of λ between the replica exchange MSTIS (dotted curves) and Wang-Landau MSTIS (solid curves) simulations. For each state the two approaches result in an identical crossing probability within the error bar (shown for 2 points). Right: The density of paths in the Wang-Landau algorithm converges to crossing probability.

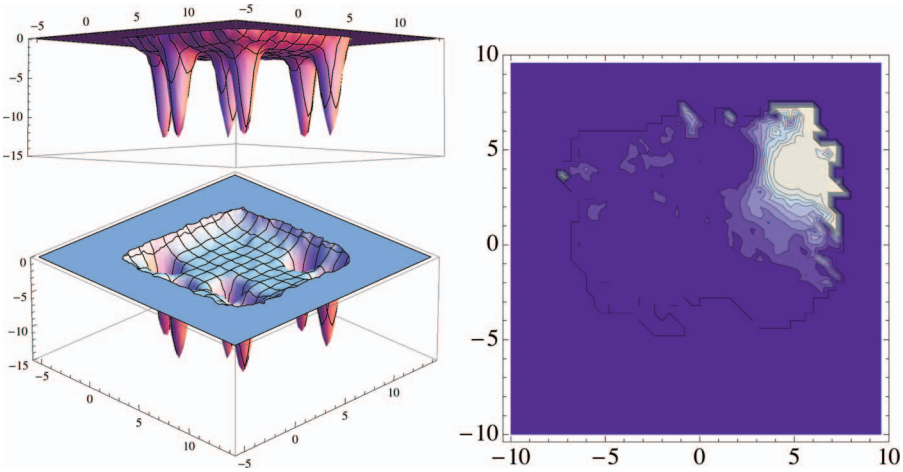

FIG. 6. The RPE contains both information about the free energy and the kinetics. Left: FE projections of the RPE in the x, y plane. Note that this projection is naturally equal to the potential $V(x, y)$. Right: Projection of the committor p_3 in the x,y plane. Purple denotes $p_3 = 0$ and white denotes $p_3 = 1$.

This metric is zero if the configuration x is identical to the reference structure of stable state I, and positive for deviations from it. The metric has the advantage that it is simpler to compute than, for instance, the root mean square deviation. However, this order parameter also does not allow for permutations of labeled particles in the structures, so that all cluster permutations are identified as a separate state. This leads to $7! = 5040$ different configurations per potential energy minimum. While in principle it is possible to sample all configurations, it seems unnecessary, since all of these states are degenerate. We avoid the permutation problem by sorting the sum of the distances for each particle of the reference structure, and of the current configuration, before comparing the two. Thus, the order parameter becomes

$$\Delta(x, x_I) = \sum_i^N |d_i^*(x) - d_i^*(x_I)|, \tag{16}$$

where the asterisk indicates that the distances have been sorted. The sorted distances for the four state minima are given in Table I.

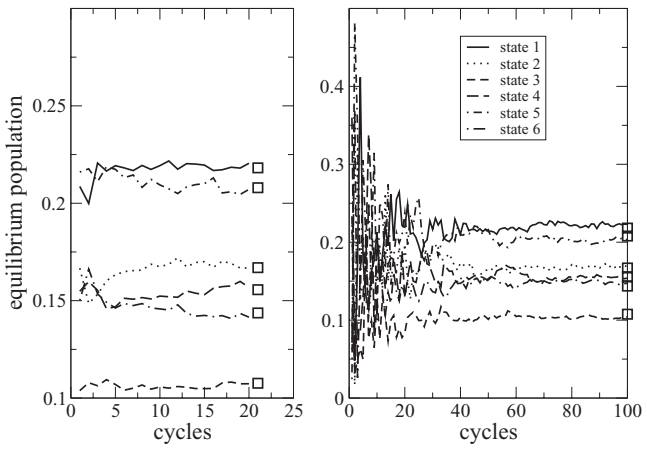

FIG. 7. Convergence of the equilibrium populations as a function of simulation time. Left: Fixed bias. Right: Wang-Landau MSTIS scheme.

We evolve the equation of motions using the Nose-Hoover-Langevin integrator,[40] with a time step $dt = 0.01$ and a friction $\gamma = 1$. A quick MD run of the ground states shows that at low temperature of $\beta = 20$ the state is very stable, and well defined by $\Delta(x, x_I) < 3$. We define a set of 15 interfaces for the ground state $\lambda_{i1} = \{3, 3, 4, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5\}$. For simplicity we used the stable state definition and this interface sets also for the other states, although this is not strictly necessary. During the path sampling we attempt shooting, exchange, reversal, and state swap moves with equal probability.

Starting from the ground state (the pentagonal bipyramid) the Wang-Landau MSTIS algorithm explores the interfaces by building up the DOP. At some point by coincidence an activated path at a high interface enters another stable state. As this state is not yet recognized, a very long trial trajectory of the maximal allowed length (in this case 10^4 time steps) is the result. The end point of this trajectory is a potential candidate for a stable state. The configuration is minimized and checked for stability, i.e., it does not relax back to another state in the maximum allowed time. If it is stable the configuration is added to the database of states. After a while another state is detected and the procedure is repeated, until all states are found. The build-up for the DOP for the four states is shown in Fig. 9. Here the DOP is plotted every 1000 trial shots. Note that in a several rounds all states have been found.

The Wang-Landau scheme can be continued until the DOP of all replicas of all states have converged until $f = 5 \times 10^{-6}$, which took of the order 10^8 moves. The resulting crossing probabilities and density of paths are plotted in

FIG. 8. The configurations corresponding to the stable states in the LJ7 cluster. From left to right: the pentagonal bipiramid, the capped octahedron, bicapped trigonal bipyramid, and tricapped tetrahedron.[38,39]

TABLE I. Sorted squared distance sum for the four minima.

Index	1	2	3	4
1	7.5356	8.7401	7.52625	7.52302
2	7.5356	8.74011	9.60047	9.59817
3	11.6306	8.74012	9.60047	9.59817
4	11.6306	11.2149	11.7586	9.59817
5	11.6306	11.215	11.7586	13.9461
6	11.6306	11.215	15.4297	13.9462
7	11.6306	14.9534	15.4298	13.9464

Fig. 10. Employing the fixed biasing function scheme resulted in a converged DOP after only 10^6 shots.

Note that the requirement to sample all states during the simulation is not strictly necessary. State swap moves are optional, and can be left out. However, each state will then require its own separate single replica MSTIS simulation.

From the single replica MSTIS results we can extract the rate using the expression $k_{IJ} = \phi_{1I} P_I(\lambda_{mI}|\lambda_{1I}) P(\lambda_{0J}|\lambda_{mI})$. The first term, the flux $\phi_{1I} = \langle \tau_0 + \tau_1 \rangle^{-1}$, follows from the length of the paths in the minus and first interfaces, τ_0 and τ_1, respectively. The second factor is the crossing probability at the outermost interface, or alternatively, the ratio of the DOP for the outermost and the first interfaces. The last factor is the fraction of paths in the outermost interface ensemble that make it from I to J. The three factors are summarized in Table II and Fig. 10, and the final rate matrix is given in Table III. As expected the transitions out of the ground state are very rare, whereas those leaving the excited states are much more frequent.

As the unbiased dynamical reactive pathways are available in the simulations, analysis of these pathways can lead to improved insight in the mechanism of the rare event. We illustrate here how such reactive pathways can be visualized. For each pathway the distance squared order parameter $\Delta(x, x_I)$ with respect to each state I is available. This information could be plotted in a four dimensional space, or, alternatively, a lower dimensional projection. Here we make use of the fact that four dimensional distance data can be visualized in a tetrahedral projection, in which the corners of the tetrahedron denote a perfect overlap with one of the structures,

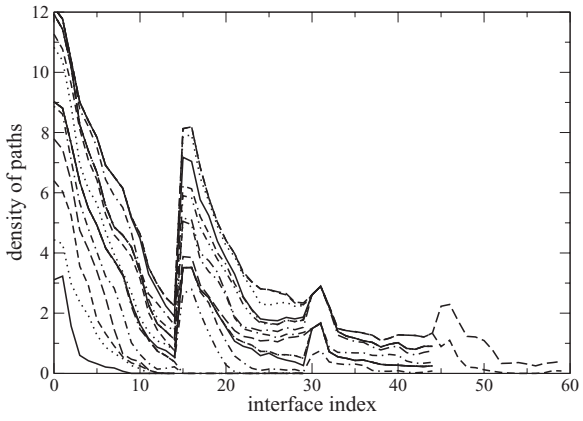

FIG. 9. Buildup of the density of paths for the LJ7 cluster using the adaptive Wang-Landau MSTIS scheme. The index is defined as $Im + i$.

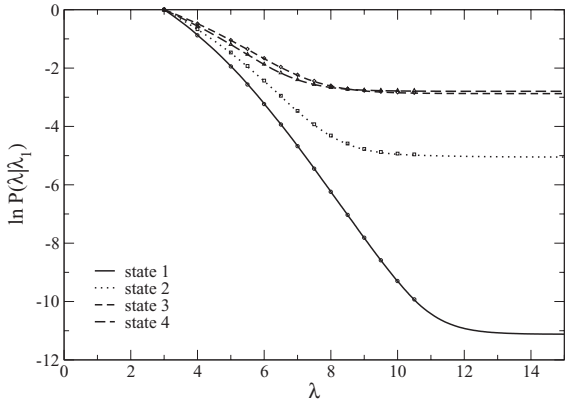

FIG. 10. Converged crossing probability (curves) and density of paths (symbols) for the LJ7 cluster.

and the center of the tetrahedron corresponds to structures that are far from any of the minima. This is achieved as follows. First for each configuration x along the path the inverse of the square order parameter $\Delta(x, x_I)$ is taken as a projected 4D coordinate, and normalized such that the sum equals 1. This results in a coordinate close to unity if the configuration is near the target, and 0.25 for all if the configuration is far from all targets. The inverted 4D normalized coordinate is now projected to the tetrahedron using a simple geometrical formula also used for constructing four component phase diagrams.[41] Fig. 11 shows several pathways between the stable states. From this picture it is clear the paths from state 1 are relatively rare, since it is much more likely to return from the outermost interface to state 1. When it succeeds it almost always goes to state 2. From there transitions to states 3 and 4 can occur, which are plentiful due to much lower stability of the excited states. While this projection is for illustrative purposes, it highlights the fact that all rare unbiased reactive transitions are available in the simulation.

We note that optimizing the location of the interfaces can make the amount of paths in the outermost transition matrix equivalent.[42] We did not pursue that strategy here.

TABLE II. (a) Outermost interface transition matrix for the four states. Subscripts denote the error in the last digits. Columns are leaving states and rows are arriving states. (b) Flux through the first interface for the four states.

(a)

Index	1	2	3	4
1	0.69_1	0.599_8	0.0348_8	0.023_1
2	0.29_1	0.065_1	0.9336_9	0.9505_7
3	0.0034_4	0.187_5	0.0171_4	0.0185_7
4	0.0020_1	0.148_2	0.0143_5	0.0070_6

(b)

	ϕ_{1I}
1	0.00419_2
2	0.00357_2
3	0.00343_1
4	0.003323_6

TABLE III. Rate matrix for the four states. Subscripts denote the error in the last digits. Columns are leaving states and rows are arriving states.

	1	2	3	4
1	$1.41_8 \times 10^{-07}$	$1.48_4 \times 10^{-05}$	$7.0_2 \times 10^{-06}$	$4.8_3 \times 10^{-06}$
2	$5.9_3 \times 10^{-08}$	$1.62_3 \times 10^{-06}$	$1.88_1 \times 10^{-04}$	$1.93_3 \times 10^{-04}$
3	$7.0_7 \times 10^{-10}$	$4.6_2 \times 10^{-06}$	$3.5_1 \times 10^{-06}$	$3.8_1 \times 10^{-06}$
4	$4.0_3 \times 10^{-10}$	$3.6_1 \times 10^{-06}$	$2.9_1 \times 10^{-06}$	$1.4_1 \times 10^{-06}$

C. Alanine dipeptide

As a final test system we consider alanine dipeptide in explicit solvent (Fig. 12). This is a standard model system in the biomolecular simulation field, for which we have applied MSTIS previously.[43]

Energy minimization and molecular dynamics (MD) simulations were performed with the Gromacs package (version 4.0.5),[44–47] employing the AMBER96[48] force field. The system was prepared by solvating a single alanine dipeptide molecule (Ace-Ale-NME) with 620 TIP3P water molecules[49] in a truncated octahedral periodic box. The non-bonded van der Waals cutoff radius was 1.1 nm. The SHAKE algorithm handles bond constraints, and fast Particle-Mesh Ewald method treated long-distance electrostatic interactions.[50,51] After minimization and equilibration at ambient conditions (see Ref. 43), we fixed the box size to $d = 2.92160$Å, which yielded a average pressure of 1 atm. All MD simulations are performed in the NVT ensemble with a time step of 2 fs. The temperature was kept constant using the v-rescale thermostat.[52]

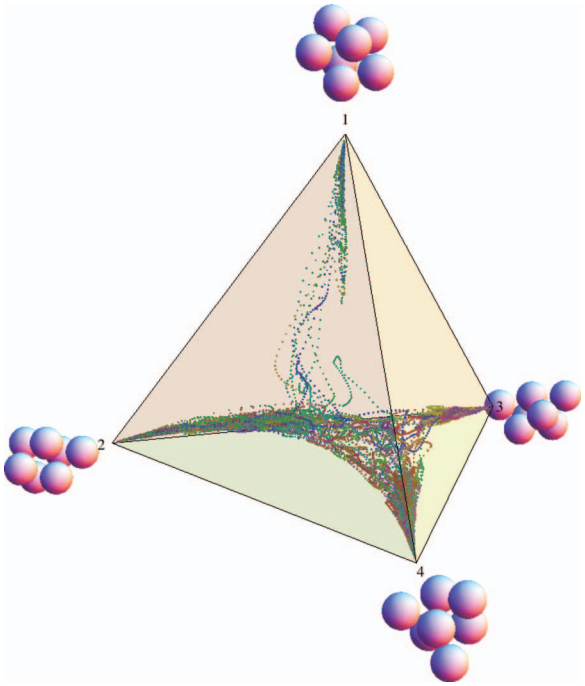

FIG. 11. Tetrahedral projection of the trajectories consisting of the 4 square distance sum order parameters. The projection is such that the corners represent exact correspondence to the target structure, whereas the center is infinitely away from all structures.

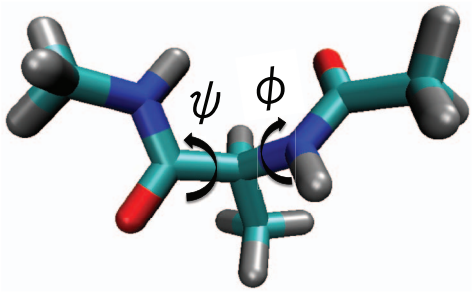

FIG. 12. The molecular structure of alanine dipeptide rendered in licorice. Carbons in cyan, oxygen in red, nitrogen in blue, and hydrogen in white. The two order parameters describing the metastable states are the dihedral angles ϕ and ψ.

Fig. 13 shows the previously computed[43] free energy surface plotted in the plane of dihedral angles ϕ and ψ. There are 6 minima with their central points located at (−150, 150), (−70, 135), (−150, −65), (−70, −50), (50, −100), and (40, 65) (in degrees). Due to low barriers between the basins A and B, and between C and D, we combine these into joint states β and α, respectively, using the combination rule.[43] The state I is defined by the set $\{x : \lambda_{iI} < \lambda_{0I}\}$ where the distance $\lambda_I = \sqrt{(\phi-\phi_I^{\text{ref}})^2+(\psi-\psi_I^{\text{ref}})^2}$ is measured from the central point of the state $(\phi_I^{\text{ref}}, \psi_I^{\text{ref}})$. All stable state boundaries are set identical ($\lambda_{0\beta} = \lambda_{0\alpha} = \lambda_{0E} = \lambda_{0F} = 10$). Four interfaces are defined for each state, as shown in Table IV.

We performed the single replica MSTIS approach with the number of interface exchange moves to number of shooting moves being roughly 4:1. We used the one-way shooting algorithm with a uniform random choice of shooting point. When an exchange with interface 0 is attempted, the minus interface move is invoked. The exchange move is governed by a fixed bias DOP, given by the crossing probability based on the sampled paths. This crossing probability and thus the bias function is updated after every 5000 exchange moves, i.e., around 1250 shooting moves.

After 50 000 exchange moves the DOP is converged. We plot the (logarithm of) the last three bias functions and the final crossing probability for each state to evaluate the convergence in Fig. 14. We observe indeed proper convergence.

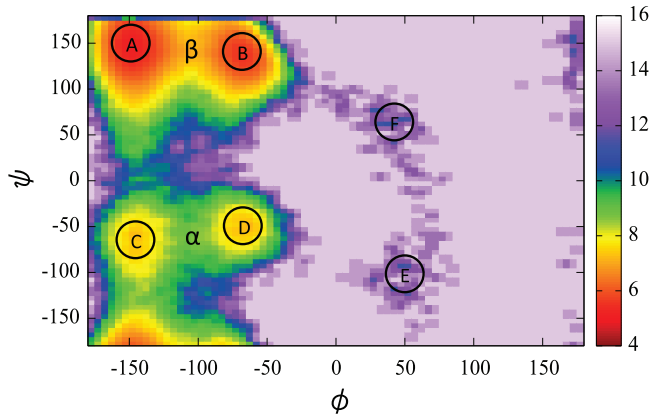

FIG. 13. The free energy surface from a replica exchange MD simulation obtained by projecting the logarithm of the probability to find a (ϕ, ψ) pair.[43]

TABLE IV. Definition of interfaces for states β, α, E, and F.

Interface	β	α	E	F
λ_0	10	10	10	10
λ_1	20	20	20	20
λ_2	40	33	40	40
λ_3	60	46	60	60
λ_4	75	60	80	80

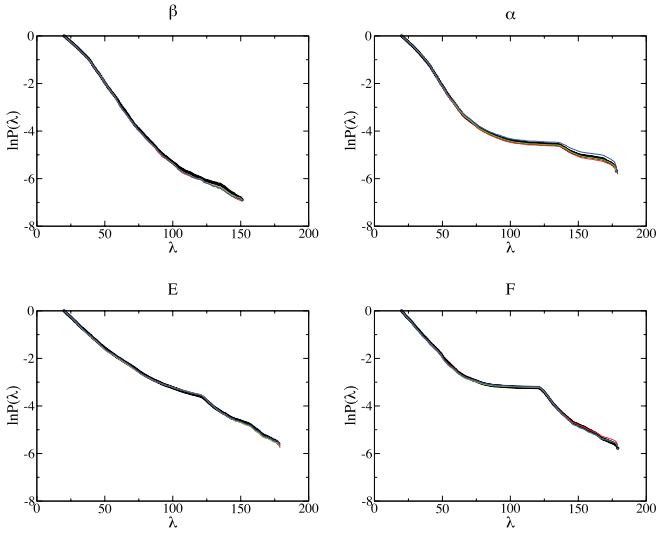

FIG. 14. Logarithm of crossing probability against interface for states β, α, E, and F. The red, green, and blue lines are the last three DOP biasing functions, while the black line with circles is the final logarithm of the crossing probability.

TABLE V. Rate constant data. (a) Flux through first interface, crossing probability from the first to the last interface, and flux at the outermost interface for each state. (b) Transition probability for states β, α, E, and F at their respective outer interfaces. The rows denote the leaving state and the columns denote the arriving state. (c) Rate constant matrix from single replica MSTIS for states β, α, E, and F.

(a)

State	ϕ_{1i} (ps^{-1})	$P_i(\lambda_{mi}\|\lambda_{1i})$	ϕ_{mi} (ps^{-1})
β	2.15240	0.02005	0.04316
α	1.98760	0.05385	0.10703
E	1.43730	0.06570	0.09443
F	2.44490	0.04694	0.11476

(b)

	β	α	E	F
β	0.92819	0.07181	0.00000	0.00000
α	0.18381	0.81552	0.00067	0.00000
E	0.01155	0.19401	0.71095	0.08350
F	0.72675	0.00173	0.10397	0.16755

(c)

	β	α	E	F
β		0.00310	0.00000	0.00000
α	0.01967		0.00007	0.00000
E	0.00109	0.01832		0.00788
F	0.08340	0.00020	0.01193	

In Table V the fluxes, crossing probabilities, multiple state transition matrix for the outer interfaces, and the final rate constants are given. The results correspond to the rate matrix given in Ref. 43, with the exception of the rate from state F to state β, which is higher.

IV. CONCLUSIONS

We have introduced a novel single replica method to efficiently sample the path space for rare events in complex environment. The method uses a density of paths as a bias to achieve a flat sampling along the interfaces. The density of paths can be built up using a Wang-Landau scheme or fixed to the crossing probability, which is the correct bias.

The single replica MSTIS algorithm has the great advantage over the RETIS approach that it provides an asynchronous sampling scheme, in contrast to the replica exchange approach, that has to be conducted in a parallel way. The single replica sampling can still be conducted in parallel, by independent walkers, and thus is almost trivially parallelizable. The only data needed for all replicas are the density of paths and the visiting histogram. In contrast, the parallel RETIS algorithm needs information about all paths at all times, and hence needs to be synchronized.

The single replica MSTIS algorithm becomes of course more computationally demanding with the number of stable states. When all states are connected with all other states, and all transitions are similarly populated, the computational effort will scale quadratically with the number of states. However, when the network of transition is sparse (e.g., a protein folding network) the approach will scale roughly linear with the number of states. In practice, the scaling will probably somewhere in between these limits.

As shown in this work, the single replica approach also allows an adaptive sampling scheme in which the database of states is slowly built up. We believe this adaptive scheme will be useful in the exploration of complex (biomolecular) systems, where states are initially largely unknown. In this adaptive scheme the recognition of new states is a crucial part. Currently the recognition of new states is made in an ad hoc manner. Systematic ways of recognizing the states, e.g., using the Markov State Model methodology might be beneficial.[53]

ACKNOWLEDGMENTS

The authors thank David Swenson for a careful reading of the manuscript. This work is part of the research programme VICI 700.58.442, which is financed by the Netherlands Organization for Scientific Research (NWO).

[1]G. M. Torrie and J. P. Valleau, Chem. Phys. Lett. **28**, 578 (1974).
[2]E. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral, Chem. Phys. Lett. **156**, 472 (1989).
[3]T. Huber, A. Torda, and W. van Gunsteren, J. Comput.-Aided Mol. Des. **8**, 695 (1994).
[4]H. Grubmüller, Phys. Rev. E **52**, 2893 (1995).
[5]A. F. Voter, J. Chem. Phys. **106**, 4665 (1997).
[6]A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002).
[7]E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).

[8]Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[9]E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[10]L. Zheng, M. Chen, and W. Yang, Proc. Natl. Acad. Sci. U.S.A. **105**, 20227 (2008).
[11]Y. Q. Gao, J. Chem. Phys. **128**, 064105 (2008).
[12]C. H. Bennett, in *Algorithms for Chemical Computations*, ACS Symposium Series Vol. 46, edited by R. Christofferson (American Chemical Society, Washington, DC, 1977).
[13]D. Chandler, J. Chem. Phys. **68**, 2959 (1978).
[14]G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9978 (2000).
[15]R. Elber, A. Cardenas, A. Ghosh, and H. A. Stern, Adv. Chem. Phys. **126**, 93 (2003).
[16]L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, J. Chem. Phys. **125**, 024106 (2006).
[17]C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[18]C. Dellago and P. G. Bolhuis, Adv. Polym. Sci. **221**, 167 (2009).
[19]T. S. van Erp, D. Moroni, and P. G. Bolhuis, J. Chem. Phys. **118**, 7762 (2003).
[20]A. K. Faradjian and R. Elber, J. Chem. Phys. **120**, 10880 (2004).
[21]R. J. Allen, D. Frenkel, and P. R. ten Wolde, J. Chem. Phys. **124**, 024102 (2006).
[22]F. Cérou, A. Guyader, T. Lelièvre, and D. Pommier, J. Chem. Phys. **134**, 054108 (2011).
[23]A. Warmflash, P. Bhimalapuram, and A. R. Dinner, J. Chem. Phys. **127**, 154112 (2007).
[24]J. T. Berryman and T. Schilling, J. Chem. Phys. **133**, 244101 (2010).
[25]G. Huber and S. Kim, Biophys. J. **70**, 97 (1996).
[26]T. S. van Erp, Phys. Rev. Lett. **98**, 268301 (2007).
[27]P. G. Bolhuis, J. Chem. Phys. **129**, 114108 (2008).
[28]J. Rogal and P. G. Bolhuis, J. Chem. Phys. **129**, 224107 (2008).
[29]J. Rogal, W. Lechner, J. Juraszek, B. Ensing, and P. G. Bolhuis, J. Chem. Phys. **133**, 174109 (2010).
[30]D. Swendson and P. G. Bolhuis, "Multiple state replica exchange transition interface sampling," J. Chem. Phys., (to be published).
[31]M. Hagen, B. Kim, P. Liu, R. A. Friesner, and B. J. Berne, J. Phys. Chem. B **111**, 1416 (2007).
[32]W. Lechner, C. Dellago, and P. G. Bolhuis, J. Chem. Phys. **135**, 154110 (2011).
[33]E. E. Borrero and C. Dellago, J. Chem. Phys. **133**, 134112 (2010).
[34]A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
[35]S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, J. Comput. Chem. **13**, 1011 (1992).
[36]F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[37]F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[38]D. J. Wales and R. S. Berry, J. Chem. Phys. **92**, 4283 (1990).
[39]D. Moroni, "Efficient sampling of rare event pathways," Ph.D. thesis (Universiteit van Amsterdam, 2005).
[40]B. Leimkuhler, E. Noorizadeh, and F. Theil, J. Stat. Phys. **135**, 261 (2009).
[41]F. Spear, Am. Mineral. **65**, 1291 (1980).
[42]E. E. Borrero, M. Weinwurm, and C. Dellago, J. Chem. Phys. **134**, 244118 (2011).
[43]W.-N. Du, K. A. Marino, and P. G. Bolhuis, J. Chem. Phys. **135**, 145102 (2011).
[44]B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008).
[45]D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, J. Comput. Chem. **26**, 1701 (2005).
[46]E. Lindahl, B. Hess, and D. van der Spoel, J. Mol. Model. **7**, 306 (2001).
[47]H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, Comput. Phys. Commun. **91**, 43 (1995).
[48]W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[49]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[50]T. Darden, D. York, and L. G. Pedersen, J. Comput. Chem. **98**, 10089 (1993).
[51]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[52]G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).
[53]J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé, J. Chem. Phys. **134**, 174105 (2011).