A replica exchange transition interface sampling method with multiple interface sets for investigating networks of rare events

Swenson, D.W.H.; Bolhuis, P.G.

DOI
10.1063/1.4890037

Publication date
2014

Document Version
Final published version

Published in
Journal of Chemical Physics

A replica exchange transition interface sampling method with multiple interface sets for investigating networks of rare events

David W. H. Swenson and Peter G. Bolhuis[a)]
van 't Hoff Institute for Molecular Sciences, Universiteit van Amsterdam, PO Box 94157, 1090 GD Amsterdam, The Netherlands

(Received 23 April 2014; accepted 30 June 2014; published online 22 July 2014)

The multiple state transition interface sampling (TIS) framework in principle allows the simulation of a large network of complex rare event transitions, but in practice suffers from convergence problems. To improve convergence, we combine multiple state TIS [J. Rogal and P. G. Bolhuis, J. Chem. Phys. **129**, 224107 (2008)] with replica exchange TIS [T. S. van Erp, Phys. Rev. Lett. **98**, 268301 (2007)]. In addition, we introduce multiple interface sets, which allow more than one order parameter to be defined for each state. We illustrate the methodology on a model system of multiple independent dimers, each with two states. For reaction networks with up to 64 microstates, we determine the kinetics in the microcanonical ensemble, and discuss the convergence properties of the sampling scheme. For this model, we find that the kinetics depend on the instantaneous composition of the system. We explain this dependence in terms of the system's potential and kinetic energy. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4890037]

I. INTRODUCTION

Rare events, in which transitions between stable regions of phase space are extremely infrequent due to the presence of large free energy barriers, are ubiquitous in chemical physics. Examples range from nucleation, to chemical reactions, to conformational changes in biomolecules. Usually the rare event occurring in the system is (considered) independent from possible other events occurring. For instance, in protein folding, the folding of a single molecule is considered not to be influenced by the state of another protein.[1] Similarly, isomerization reactions of molecules in the gas phase are supposed to be independent of each other. The kinetics of each rare event are therefore treated individually. Still, there are occasions when the environment does influence the kinetics of the rare event. One example is high pressure or density where molecules are (sterically) hindering each other, changing their kinetics. Another more complex example is the enhancement of conformational changes in amyloidogenic or prion proteins by other such proteins.[2] A fundamental question is therefore: how are the kinetics of rare events influenced by their direct environment, consisting of identical molecules under going the same rare transitions? Of course, it is well-understood that the same reaction can have a very different rate in the condensed phase and in vacuum. However, the rates can vary significantly as a function of reaction extent even at a single state point. This effect is usually not taken into account. In particular, when we have a system that contains many particles which can undergo rare transitions, what happens when some of those particles are in different states? How separable is the state of a given rare event system from the global state of its environment?

To study this problem theoretically, molecular dynamics (MD) simulations are the tool of choice. However, as all-atom simulations require time steps on the scale of femtoseconds, the reactions occur with a frequencies rare enough that direct simulation ranges from impractical to impossible with straightforward MD. Therefore, many techniques to study such processes have been developed, including umbrella sampling,[3] blue moon sampling,[4] local elevation,[5] flooding,[6] hyper-dynamics,[7] meta-dynamics,[8] adaptive bias force,[9] replica exchange,[10] simulated tempering,[11] integrated sampling,[12] orthogonal space sampling,[13] and many others. Such methods enhance the sampling of the rare event by biasing along an *a priori* known collective variable. These methods do alter the natural dynamics, and hence do not directly yield kinetic information. This shortcoming can be ameliorated by applying the Bennett-Chandler approach[14,15] which gives precise kinetic information, but often requires *a priori* knowledge of the reaction coordinate, which might be difficult to establish. This notion drove the development of so-called two-ended methods such as nudged elastic band,[16] action minimization,[17] the finite temperature string method,[18] and transition path sampling.[19] These techniques aim to find pathways between two predefined endpoints. Transition path sampling (TPS) results in an ensemble of dynamically unbiased trajectories between two specified states: the initial and final state of a rare event process. Applying the path sampling framework yields mechanisms, rates, and even insight in the reaction coordinates of complex processes.[20] Many advances have been made in the last decade, such as transition interface sampling (TIS),[21] which introduces a set of interfaces to allow efficient rate constant calculations. TIS is related to forward flux sampling,[22] the adaptive multilevel splitting method,[23,24] stochastic process rare event sampling,[25] and the restart method.[26] These methods create a rare event path ensemble by splitting trajectories, and select the ones that make

[a)]Electronic mail: p.g.bolhuis@uva.nl

progress along an collective variable, whereas the path sampling techniques are based on a Monte Carlo importance sampling of (rare) trajectories.

The current path sampling methodology cannot directly be applied to the problem of many molecules undergoing transitions individually, because it is not known in advance which molecule is making the transitions. Therefore, TPS and TIS have been extended to simultaneously study the networks of rare events which arise when a system has more than two stable states.[27] While this multiple-state approach achieves sampling advantages by studying all transitions simultaneously instead of individually, convergence of the path ensemble turned out difficult.[28] Coupling TIS with replica exchange substantially enhances the sampling for two state systems,[29,30] and was also proposed for the multiple state system. Recently, one of us developed a single replica method to explore the trajectory space in a multiple state system.[31] In this paper, we focus on applying the full replica exchange TIS approach to a multiple state system.

We choose to work with a very simply model system: a 2D fluid of multiple independent dimers that can undergo a transition from compressed to extended. A similar system was also studied in Ref. 27. Using this simple model we can study the influence of the composition of the system on the rate constant at a single state point. The question then becomes whether the rates of the transitions are dependent on the state of the entire system. While this dimer model seems simple, the number of possible states grows exponentially with the number of constituents: two to the power of the number of the dimers. The complex system kinetics between such states could be analyzed by computing a complete Markov state model, a.k.a. an equilibrium kinetic network,[32] but this would be very cumbersome, as the number of possible transitions is the square of the number of states. Therefore, often one artificially reduces the number of states by lumping similar microstates into a macrostate.[32] Moreover, the improbability of having more than one reaction simultaneously reduces the number of transitions drastically.

A challenge with the multiple state formalism is that a good order parameter for one transition might not be a good order parameter for another transition exiting the same initial state. This can be particularly problematic if one of the transitions is much less frequent than the other — indeed, the grouping of states into a macrostate in Ref. 33 was primarily to avoid this problem, not to reduce the total number of transitions. In the present work, we introduce a new version of TIS which allows different order parameters for each transition.

The aim of this paper is fourfold: (1) to investigate and improve the convergence of replica exchange multiple state transition interface sampling methods, (2) to develop and apply grouping methods to lump micro-states into macro-states, (3) to develop and apply different order parameters for each transition, and (4) to investigate the effect of the environment (as a function of composition) on the kinetics.

The rest of the paper is organized as follows. Section II introduces the simple model system we study, as well as the order parameters and state/macrostate definitions. This clarifies the need for the multiple interface set TIS approach, which is introduced in Sec. III, after a brief review of TIS in general. Section IV describes the Monte Carlo moves used for sampling path space in this new approach, and introduces the new multiple set minus move. Section V shows the results of this approach, and Sec. VI summarizes and concludes.

II. MODEL SYSTEM

The double-well dimer in a bath of Weeks-Chandler-Anderson (WCA) particles is a widely used model system for testing and developing rare event methods.[21,34] In order to study the correlation between rare transitions, we extend that model to a system with an arbitrary number of such dimers.

The dimer intramolecular potential is given by

$$V_{\text{intra}}(r) = h\left(1 - \left(\frac{r - r_0 - w}{w}\right)^2\right)^2 \tag{1}$$

and the intermolecular potentials are pairwise interactions on the atomic sites given by the WCA potential

$$V_{\text{inter}}(r) = \begin{cases} 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) + \epsilon & \text{if } r \le 2^{1/6}\sigma, \\ 0 & \text{if } r > 2^{1/6}\sigma, \end{cases} \tag{2}$$

where σ is the diameter of the particles, and ϵ the strength of attraction. The unit of time is $(m\sigma^2/\epsilon)^{1/2}$, with m as the mass of the particles. In this work, we have set $r_0 = 2^{1/6}\sigma$, $h = 10.0\epsilon$, and $w = 0.3\sigma$. In the remainder of the paper, we consider reduced dimensionless variables by setting $\epsilon = 1$, and $\sigma = 1$. We study this system in a two-dimensional square box with side lengths $L_x = L_y = 6.5$. The box always contains 25 total particles, although the number of dimers N_d ranges between $N_d = 1$ and $N_d = 6$ (so the number of monomers ranges from $N_m = 13$ to $N_m = 23$). Note that for $N_d = 1$, this reduces to a model system which has been previously studied with several parameter sets for the double-well.[21,34]

Each dimer can switch between condensed and extended conformations. This means that systems with the number of extended dimers N_{ex} in the range $0 \le N_{ex} \le N_d$ can make a transition either to extend a dimer or to condense a dimer. The processes involved are fundamentally different for the two directions, so instead of trying to use a single order parameter to define interfaces for all transitions out of a given state, we will provide a different order parameter for each transition, and each order parameter will have its own set of interfaces. This approach, which we call "multiple interface set transition interface sampling," is described in detail in Sec. III.

Since our system has N_d dimers, each of which can be extended or condensed, we have 2^{N_d} different stable microstates, with $2N_d\,2^{N_d-1}$ possible transitions between microstates with $\Delta N_{ex} = \pm 1$. However, the dimers are identical, so we can lump these microstates into macrostates (hereafter denoted just "states") according to N_{ex}. This means that there are $N_d + 1$ such states, and there are $2N_d$ transitions between them with $\Delta N_{ex} = \pm 1$.

It is also possible for more than one dimer to make a transition simultaneously, but this is an extremely rare event and we do not attempt to study it here (we only observe it at very high energies, and even then very infrequently). This can

lead to transitions with $\Delta N_{ex} = \pm 2$ or more, or can involve a transition to the same macrostate with a different microstate. In the following, we only consider transitions with $\Delta N_{ex} = \pm 1$.

We define state boundaries by setting a maximum distance r_{sh} for a condensed dimer and a minimum distance r_{ex} for an extended dimer. If r_k, the distance between the particles in the kth dimer, is between r_{sh} and r_{ex}, then that dimer is in the transition region. We define the system as being in a state if and only if no dimers are in the transition region. Mathematically, this means we give the configuration a state label $I = N_{ex}$ if $\sum_{k=1}^{N_d}[\Theta(r_{sh} - r_k) + \Theta(r_k - r_{ex})] = N_d$ (where $\Theta(x)$ is the Heaviside function); otherwise, the state label is undefined.

Each interface set α has its own order parameter, but for convenience we can group them into "forward" (increasing N_{ex}) and "backward" (decreasing N_{ex}). We must also account for the fact that we want the order parameter to be independent of the specific initial microstate or final "target" microstate. To accomplish this, we define a local distance order parameter d_k^α for each dimer, where $1 < k < N_d$ is a label on the dimer, and α is a label on the interface set. This is given by

$$d_k^\alpha = \begin{cases} r_k - r_{sh} & \text{if } r_k^0 < r_{sh} \text{ and } \alpha \in \{\text{fwd}\}, \\ r_{ex} - r_k & \text{if } r_k^0 > r_{ex} \text{ and } \alpha \in \{\text{bkwd}\}, \\ \text{undefined} & \text{in all other cases,} \end{cases} \quad (3)$$

where r_k^0 is the dimer distance r_k at the start of the trajectory, and the collection {fwd} (resp. {bkwd}) is the collection of all forward (backward) interface sets. These collections are defined purely for convenience for this model. Each member of one of these collections will be associated with a different initial state I: i.e., there is a "forward" interface set associated with state $N_{ex} = 0$, and another associated with $N_{ex} = 1$: these are two elements of {fwd}. For $N_{ex} = 1$, there is also a different interface set for the "backward" transition, which would be a member of {bkwd}.

The overall order parameter λ_α for the entire system in a given interface set α is given by the maximum of the local order parameters

$$\lambda_\alpha = \max_{k \in \text{dimers}} \left(d_k^\alpha\right). \quad (4)$$

Note that the same configuration can be associated with different order parameters, depending on which interface set α it is associated with.

The cutoffs r_{sh} and r_{ex} define the width of the state: in the following, we take $r_{sh} = 1.20$ and $r_{ex} = 1.64$ for all states. We place 8 regular "in-set" interfaces along the order parameters in both "forward" and "backward" directions, starting at $\lambda = 0.000$ and separated by 0.025, up to an outer (regular) interface at $\lambda = 0.175$ for each interface set. We also set a multiple-state interface at $\lambda = 0.220$, which is the midpoint between the two state boundaries (as such, it represents the same physical interface whether it associates with a "forward" or a "backward" interface).

It is important to observe that these order parameter definitions actually provide *different* order parameters for each interface set. For example, the order parameter for a "forward" interface going from $N_{ex} = 0$ to $N_{ex} = 1$ is undefined if the path starts in $N_{ex} = 2$. Despite the similarities in the definitions as given above, each interface set formally has its own order parameter.

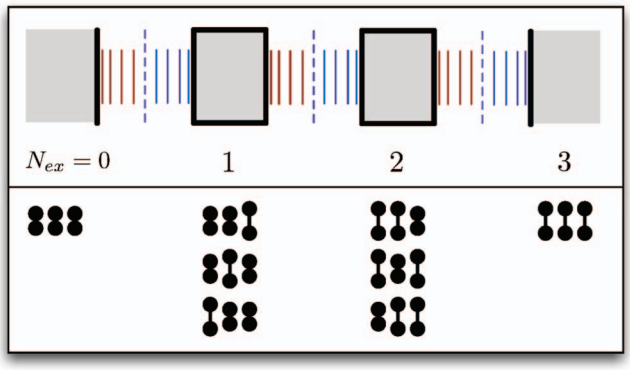

FIG. 1. State topology and microstates for $N_d = 3$. Top panel: state and interface topology, showing 4 normal interfaces (short solid lines) per interface set, 1 multiple state interface (dashed line) per transition, and minus/multiple set minus interfaces (thick lines/boxes) at each state. Extending interfaces are shown in red, condensing interfaces in blue, and multistate interfaces in purple. The minus/multiple set minus interfaces are shown in black. Bottom panel: microstates associated with each state, shown as a triplet of (extended or condensed) dimers.

The top panel of Fig. 1 shows a schematic representation of the states and interfaces for the $N_d = 3$ system. Only four in-set interfaces are shown per transition. The bottom panel enumerates the microstates for each state.

III. MULTIPLE INTERFACE SET TRANSITION INTERFACE SAMPLING

A. Multiple state TIS

Transition interface sampling is a path sampling approach to rare events which is particularly efficient at calculating rates. The fundamental equation of TIS calculates the rate as the product of the flux $\phi_{0,A}$ through an innermost interface of state A and the product of all the n successive crossing probabilities $P(\lambda_{(j+1),A}|\lambda_{j,A})$[21]

$$k_{AB} = \phi_{0,A} P(\lambda_{0,B}|\lambda_{n,A}) \prod_{j=0}^{n-1} P(\lambda_{(j+1),A}|\lambda_{j,A}), \quad (5)$$

where $\phi_{0,A}$ is the effective positive flux out of the stable state A through the first interface $\lambda_{0,A}$. "Effective positive" means we count only a crossing when the trajectory came directly from A (no other crossings of $\lambda_{0,A}$ in between). $P(\lambda_{i,A}|\lambda_{j,A})$ gives the probability that a trajectory that crosses $\lambda_{j,A}$ while coming directly from state A also crosses $\lambda_{i,A}$. In this format, TIS connects only two states, A and B. More recently, TIS (and transition path sampling) have been extended to the case of multiple states connected by a network of transitions.[27,28,33] When there are multiple states, the path ensemble is extended to allow transitions between *any* two states. In this case, the equation for the rate is the

straightforward extension of Eq. (5)

$$k_{IJ} = \phi_{0,I} P(\lambda_{0,J}|\lambda_{n_I,I}) \prod_{j=0}^{n_I-1} P(\lambda_{(j+1),I}|\lambda_{j,I}), \tag{6}$$

where n_I is now the number of interfaces associated with state I.

One important point about TIS in general (and which holds for MS-TIS) is that the results are formally independent of the choice of order parameter λ. However, a poorly chosen form for λ can, indeed, fail in practice due to poor convergence.

B. Multiple interface set TIS

In this work, we introduce the multiple interface set TIS formalism, which generalizes the multiple state formalism. The key idea is that, when multiple transitions can come from the same initial state, a single order parameter may not be the best way to bias all the transitions. Previous versions of MS-TIS had only one set of interfaces leaving each state. As such, the state label and the interface set label were equivalent. In the formalism developed below, we distinguish between the two. Instead of each interface i belonging to a given state I, it belongs to a given interface set α. Each interface set then has a unique initial state $S(\alpha) = I$ (although many different interface sets can be associated with any particular state I).

1. The in-set interface path ensemble

This work will involve several types of interfaces, each with its own path ensemble. The primary type is the in-set interface. The path ensemble for an in-set interface i belonging to interface set α (associated with an initial state given by $S(\alpha)$) is

$$\mathcal{P}_{\alpha,i}[\mathbf{x}(L)] = \frac{1}{Z_{\alpha,i}} \mathcal{P}[\mathbf{x}(L)] h_{S(\alpha)}(x_0) \prod_J \bar{h}_J[\mathbf{x}(L)] \times \hat{h}_{\alpha,i}[\mathbf{x}(L)] \sum_J h_J(x_L), \tag{7}$$

where $\mathcal{P}[\mathbf{x}(L)]$ is the natural probability to observe the dynamical path $\mathbf{x}(L)$. The trajectory $\mathbf{x}(L) = \{x_0, x_1 \ldots x_L\}$ contains $L+1$ phase points $x = \{r^N, p^N\}$ with r^N and p^N, respectively, the positions and momenta of all the particles at a given time. These phase points are separated by a time $\Delta\tau = \mathcal{T}/L$, where $\mathcal{T}$ is the total time of the trajectory. $Z_{\alpha,i}$ is a normalization factor similar to a partition function, such that the path integral over trajectories of all lengths is $\int \mathcal{D}\mathbf{x}(L)\mathcal{P}_{\alpha,i}[\mathbf{x}(L)] = 1$. The characteristic function $h_{S(\alpha)}(x_0)$ is unity if the configuration x_0 is in state $S(\alpha) = I$, and zero otherwise, enforcing that the paths start in the initial state I associated with interface set α. Similarly, the sum over $h_J(x_L)$ means that the path must end in a state, but that it can end in any other state, including state I. The $\hat{h}_{\alpha,i}[\mathbf{x}(L)]$ path function is one if the path crosses interface $\lambda_{\alpha,i}$, and is zero otherwise. The function $\bar{h}_J[\mathbf{x}(L)]$ is zero if the path ever enters state J except at its endpoints, and is one otherwise. It can be represented mathematically in terms of $h_J(x)$ as

$$\bar{h}_J[\mathbf{x}(L)] \equiv \prod_{i=1}^{L-1} (1 - h_J(x_i)). \tag{8}$$

Including a product of this function over all states ensures that our paths are only in states at their endpoints. In summary, this path ensemble includes paths that start in I, end in any other state J, and cross interface $\lambda_{\alpha,i}$, without visiting any other state in between.

2. The multiple state path ensemble

In contrast to the in-set path ensembles which require each path to start in the state associated with it, the multiple state path ensemble allows paths to start (and end) in different states. For each multiple state interface, we define a group $\mathcal{M}$ of allowed sets (and therefore initial states). In the present work, we limit the members of $\mathcal{M}$ to the "forward" and "backward" versions of each transition, e.g., the extending set from initial state $N_{ex} = 0$ and the condensing set from initial state $N_{ex} = 1$ form such a pair. The multiple state path ensemble then becomes

$$\mathcal{P}_{\mathcal{M}}[\mathbf{x}(L)] = \frac{1}{Z_{\mathcal{M}}} \mathcal{P}[\mathbf{x}(L)] \prod_J \bar{h}_J[\mathbf{x}(L)] \times \sum_{\alpha \in \mathcal{M}} h_{S(\alpha)}(x_0) \hat{h}_{\alpha,m}[\mathbf{x}(L)] \times \sum_J h_J(x_L), \tag{9}$$

where the mth interface of set α is understood to be the outermost interface of the set. $Z_{\mathcal{M}}$ is again a normalizing factor. Both the in-set and the multiple state interfaces in this formalism reduce to the previous work on multiple state TIS in the appropriate limits. If each state has only one interface set, our sum over interface sets is equivalent to the sum over states used in previous work.[27] If we also include all interface sets in $\mathcal{M}$, we regain the old definition of the multiple state outer interface path ensemble.

3. The multiple set minus interface path ensemble

Finally, we introduce the multiple set minus interface path ensemble, which allows the exchange of replicas between different interface sets associated with the same initial state. This path ensemble is an extension of the minus interface path ensemble introduced by van Erp in Ref. 29. Normal (in-set) interface path ensembles begin and end in a state, and cross an interface (a "positive" crossing). The minus interface uses paths with "negative" crossings: paths which start at the interface boundary, and go into the state instead of away from it. In Ref. 29, the minus interface is used for two purposes. First, the decorrelation of the path inside the state means that different exits from the state are sampled, which, combined with replica exchange, leads to more efficient sampling of the path space. We show a similar effect by demonstrating the improved sampling of the dimer–dimer distance when the minus interface is included. The second aim of the minus interface

in Ref. 29 is that the flux can be calculated from the "plus" and "minus" versions of the same interface. While we did not implement that in the present work, the multiple set minus interface can also be used in that way.

Paths in the multiple set minus interface must begin and end with segments that cross the innermost interface of one of the interface sets coming out of a given initial state. The path should be long enough that these two segments should be decorrelated. As the dimers in the simple model exhibit "ringing" (exiting the same interface several times) until perturbed by interaction with another particle in the system, instead of taking two consecutive segments that cross first interfaces, we take a path with N_l such segments. We choose N_l such that the segments are decorrelated, and find that this corresponds with the number of harmonic periods in the mean free path time. Mathematically, we express the path ensemble for the multiple set minus interface as

$$\mathcal{P}_I^{N_l}[\mathbf{x}(L)] = \frac{1}{Z_I^{N_l}} \mathcal{P}[\mathbf{x}(L)] \check{h}_I^{N_l}[\mathbf{x}(L)] \prod_{J \neq I} \bar{h}_J[\mathbf{x}(L)], \quad (10)$$

where the indicator function $\check{h}_I^{N_l}[\mathbf{x}(L)]$ is given by

$$\check{h}_I^{N_l}[\mathbf{x}(L)] = \prod_{k=1}^{N_l} h_I(x_{i_k}) h_I(x_{f_k}) \sum_{\{\alpha | S(\alpha)=I\}} \hat{h}_{\alpha,0}[x_{i_k}, \ldots, x_{f_k}], \quad (11)$$

with i_k and f_k the initial and final time slices of the segment k, such that $i_0 = 0$, $f_{N_l} = L$, and $\forall k$, $i_{k+1} > f_k > i_k$. Thus, this indicator function is only unity if the path includes N_l segments that cross the first interface of any set α while coming directly from I and returning to I, and zero otherwise. This defines a longer trajectory $\mathbf{x}(L)$ which includes N_l (disjoint) segments. Each of these segments must itself be a member of the path ensemble for the innermost interface of at least one of the interface sets α belonging to state I. This means that, in the limit where $N_l = 1$, this path ensemble just combines all the innermost interface path ensembles for the interface sets starting in this state. If we only have one interface set per state, then we get an ensemble identical to the innermost interface ensemble.

IV. MOVES IN THE MULTIPLE INTERFACE SET ENSEMBLES

In order to efficiently sample the transition interface path ensemble, we use several different types of Monte Carlo moves in path space. In addition to the normal interfaces and their associated shooting moves[21,35] and replica exchange moves,[29,36] we implement a version of van Erp's "minus interface move,"[29] and an extension thereof which we call the "multiple set minus interface move."

A. Shooting moves

In order to preserve the desired path distribution, the shooting move acceptance probability in each ensemble is slightly different. For an in-set interface i in interface set α, the acceptance probability is

$$P_{\alpha,i}^{\text{acc}}(\mathbf{x}^{(o)} \to \mathbf{x}^{(n)}) = h_{S(\alpha)}(x_0^{(n)}) \prod_J \bar{h}_J[\mathbf{x}^{(n)}(L^{(n)})] \times \hat{h}_{\alpha,i}[\mathbf{x}^{(n)}(L^{(n)})] \sum_J h_J(x_L^{(n)}) \times \min\left(1, \frac{L^{(o)}}{L^{(n)}}\right), \quad (12)$$

where the superscript (o) indicates a property associated with the "old" path, and (n) indicates a property associated with the "new" (trial) path. The terms in this expression require that the path only be in a state at its endpoints, that it cross the desired interface, that the initial state be the correct state for this interface set, that the final state be any other state, and the probability is then controlled by the ratio of the selection probabilities for each shooting point (i.e., the ratio of the path lengths). This last part is necessary to maintain detailed balance, since the probability of selecting a given phase point as the shooting point is inversely proportional to the number of time slices in the path. If the trial path exceeds the maximum allowed length (here 10^4 slices) it is rejected.

The multiple state interface shooting move has the acceptance probability

$$P_{\mathcal{M}}^{\text{acc}}(\mathbf{x}^{(o)} \to \mathbf{x}^{(n)}) = \prod_J \bar{h}_J[\mathbf{x}^{(n)}(L^{(n)})] \times \sum_{\alpha \in \mathcal{M}} h_{S(\alpha)}(x_0^{(n)}) \hat{h}_{\alpha,m}[\mathbf{x}^{(n)}(L^{(n)})] \times \sum_J h_J(x_L^{(n)}) \min\left(1, \frac{L^{(o)}}{L^{(n)}}\right). \quad (13)$$

Both the multiple state interface and the in-set interfaces invoke the multiple state TIS formalism, in that they are both allowed to *end* in any state. The primary difference between them is that the in-set interfaces for set α must start in state $S(\alpha)$, whereas the multiple state interfaces for group $\mathcal{M}$ can start in any of the states for the interface sets in that group. This is essential for the multiple state interface to act as a bridge between different interface sets, such as between "extending" and "condensing" transitions connecting the same pairs of states.

B. Replica exchange moves

In addition to improvements to the underlying formalism, TIS has seen several improvements to its implementation. Perhaps the most notable of these in the inclusion of replica exchange between interfaces (RETIS), which can lead to a significantly faster convergence.[29,30,36] The basic RETIS technique can be directly applied to our situation. We only perform swaps between "neighboring" interfaces: i.e., within a set, only interfaces with successive values of λ_i can swap with each other, only the innermost interfaces from each set can swap with the minus or multiple set minus interfaces, and only the outermost interfaces from each set can swap with the multiple state interfaces.

The acceptance rules are simply whether both paths are members of the appropriate interface ensembles, described in Sec. III. However, getting replicas to cross from one side of the multiple state interface to the other (changing interfaces sets in the process) requires that the path change its initial state: a path that came to the multiple state interface from the $N_{ex} \rightarrow N_{ex} + 1$ interface set cannot then be swapped into an interface in the $N_{ex} + 1 \rightarrow N_{ex}$ interface set, because it does not have the correct initial state. Whereas previous implementations of multiple state TIS have used a separate "path reversal" move, we accomplish the same thing by preceding swaps from the multiple state outer interface with a 50% chance of reversing the multiple state path. This is also essential for sampling different initial microstates of a given state.

C. Multiple set minus interface move

In our system, all states except $N_{ex} = 0$ and $N_{ex} = N_d$ are associated with two interface sets. In order to have replica exchange flow across a state, we need to introduce a new type of move that forms a bridge between the "forward" and "backward" interface sets associated with that state. We achieve this by defining a path ensemble that stores two segments from a longer trajectory, where each segment exits state I and crosses the innermost interface of at least one of the interface sets associated with initial state I. In our dimer application, there can be two possible innermost interfaces for each state; one for the "forward" interface set and one for the "backward" interface set.

The Monte Carlo move for the multiple set minus interface combines a replica exchange with a trajectory extension. This move uses the following algorithm, illustrated in Fig. 2:

1. At random, select either the first or last trajectory segment described by Eq. (11) ($k = 1$ or $k = N_l$) for this multiple set minus interface. By definition, this segment will satisfy the requirements for *at least one* of the sets for this multiple set interface. We refer to the selected trajectory segment as A. Also randomly select one of the interface sets (labeled α, with innermost interface $(\alpha, 0)$ with the path in that interface labeled B).
2. If trajectory segment A from the multiple set minus interface satisfies the requirements for the interface $(\alpha, 0)$ (i.e., has a non-zero probability according to Eq. (7) for that in-set interface ensemble), swap paths A and B. Now A is in interface $(\alpha, 0)$, and B is in the multiple set minus interface. If the requirements are not met, abort and reject the move.
3. Extend B (randomly choosing whether to extend forward or backward in time) until it satisfies the multiple set minus interface path ensemble—i.e., until it consists of N_l segments crossing the innermost interface of at least one interface set, as described in Eqs. (10) and (11). Stop the extension and abort the move if the path length grows beyond some pre-set maximum. In that case, restore A and B to the old identity.

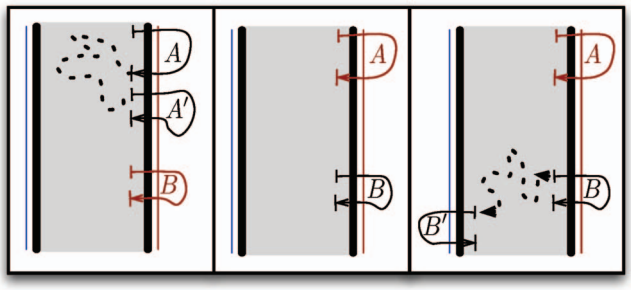

FIG. 2. Cartoon of the multiple state minus move algorithm. When a path is in the multiple set minus interface, it is shown in black. Paths in the "red" interface set to the right are shown in red. The dotted line indicates the long decorrelating trajectory between the two saved segments of the multiple set minus interface. For clarity, only the first and last exit of the multiple set minus path are shown (so this cartoon corresponds to the $N_l = 2$ case). The thick black lines are the state boundary, and the red (resp. blue) vertical lines indicate the innermost interface for the red (blue) interface set. The red interface shown corresponds to $(\alpha, 0)$ described in the text. The first panel shows the system before the move. The path segments A and A' are in the multiple set minus interface, and path B is in the $(\alpha, 0)$ interface. The second panel shows the system after the replica exchange: Path A is now in the $(\alpha, 0)$ in-set interface, and path B is in the multiple set minus interface. However, B does not satisfy the multiple set minus path ensemble until after the extension, as shown in the third panel.

The move is aborted if the replica exchange swap is not allowed (at very little computational cost) or if the extension is unsuccessful, which only happens if its path length grows beyond its maximum value (for our model system, this maximum value is 10^4 time steps, which is 20 time units). For efficient exchange, it is important that the fluxes through the innermost interface of all sets α are not too different from each other, so that a trajectory segment from a given extension step has around the same probability to satisfy any of the interface sets.

In the limit of only one interface set per state, the set selection becomes redundant and this reduces to an implementation of the minus interface. For this reason, we refer to this as the multiple set minus interface move. We have presented this move as a way to bridge different interface sets during replica exchange, which was not part of the purpose of the original minus interface.[29,30] But since our multiple set minus interface amounts to an extension of the original minus interface, we expect to also obtain some of the benefits of the original. In particular, the minus interface improves sampling by yielding significantly different trajectory segments, due to longer in-state decorrelation times. In our system, diffusion of the dimers is very slow with respect to the transition timescale. The (multiple set) minus interfaces allow for dimer diffusion during the (much longer) propagation between saved trajectory segments.

Another important aspect of the (multiple set) minus move is that we expect the two saved trajectory segments to be separated by enough time that they are essentially independent. For a system such as the one here, especially when simulated microcanonically, consecutive excursions from within the state can be part of the "ringing" of the free dimer, and are thus highly correlated. We obtained better results by saving only the fifth consecutive eligible segment ($N_l = 5$). This result matches with the observation that, at our lowest energy

$E = 25$, the mean free path is around 5 times the period of a harmonic approximation to the minima of the double well potential $V_{\text{intra}}(r)$. As a result, the (multiple set) minus interface moves tend to be the most computationally expensive part of the calculation. In order to prevent this computation from going to waste, we define the multiple set minus interface such that crossing it is equivalent to crossing the innermost interfaces for all interface sets. While, formally, any location of the multiple set minus interface that is not outside the innermost interface is allowed, for this particular choice each saved segment of the multiple set minus interface can be directly swapped with paths in the corresponding innermost interfaces. With a maximum path length of 10^4 timesteps, we find that the acceptance probability once an extension is made is above 90%, with rejections only due to path lengths growing too long. The extension of the segment only takes place if needed (if the move did not abort after attempting the replica exchange). To improve the replica flow, we reduce the probability of selecting a (multiple set) minus move such that the total probability of selecting the move and accepting it is similar to the probability of selecting and accepting a normal replica exchange move. This also reduces the impact of the higher cost of the multiple set minus move, since it is used less frequently.

V. RESULTS AND DISCUSSION

A. Rate calculations

Rates were calculated both directly and using RETIS. The flux for the RETIS rate was calculated as a byproduct of the direct rate calculation. At $E = 30$, rates were calculated for numbers of dimers from $N_d = 1$ to $N_d = 6$. In the $N_d = 2$ case, we also calculated the rate at $E = 25$, $E = 40$, and $E = 50$.

Errors in the crossing probabilities are estimated by dividing the full data set into 20 blocks with equal numbers of Monte Carlo move attempts, and calculating the standard deviation of the crossing probabilities for those blocks. Fluxes and direct rates are determined by running four MD trajectories for each interface set (giving 16-48 trajectories per data point). Errors in the direct rates are determined by logging the residence time for each transition in those trajectories, splitting the list for each transition into 20 blocks with equal numbers of transitions, and taking the standard deviation of the rates in the blocks. Errors in the fluxes are determined by taking the standard deviation of the fluxes from each of those trajectories (weighted by total residence time for the state). The RETIS error is determined by assuming the flux and crossing probability give independent standard deviations. Errors bars shown are ± 1 standard deviation. Error bars are not shown if they are larger than 50% of the rate.

Fig. 3 compares the multiple interface set RETIS rates to the directly calculated rates. It is important to note that the errors are just to give an idea of the accuracy of the method: they do *not* represent equal computational times. The effort for both the RETIS and the direct rate calculations scaled linearly with the number of states, but the direct rate calculations took about 10 times more computer wall time and still have

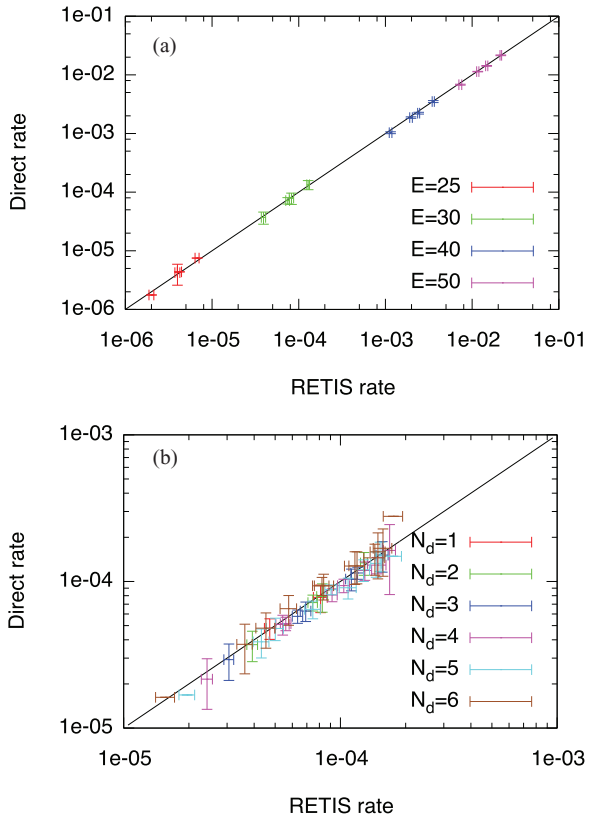

FIG. 3. Direct rate vs. TIS rate: (a) $N_d = 2$, various values of total energy; (b) $E = 30$, various numbers of dimers. Multiple interface set TIS rates match the direct rate calculations.

significantly larger error bars, except at the highest energies. Of course, this speedup factor will be largest for systems with lower energy, or for systems with higher barriers.

In Fig. 3(a), we consider the rates for a system of two dimers over a range of energies. This demonstrates the accuracy of the method over 5 orders of magnitude in the rates. Fig. 3(b) looks at fixed energy of $E = 30$, and compares the rates for different numbers of dimers in the system. The multiple interface set RETIS approach accurately reproduces the direct rate results.

B. Convergence of the sampling

One of the primary purposes of the multiple set minus interface move was to enable replicas to travel through the whole system. Only when a replica is able to visit all interfaces of all states, the system can be ergodically sampled. When the replica exchange network has a linear topology, as ours does, one good measure for this is the replica flow.[37] The replica flow, f_i, for interface i, is defined by

$$f_i = \frac{n_i^{\uparrow}}{n_i^{\uparrow} + n_i^{\downarrow}}. \tag{14}$$

The values $n_i^{\uparrow}$ and $n_i^{\downarrow}$ are obtained by labeling each replica as "up" when it visits the "bottom" interface (here, the minus interface where $N_{ex} = 0$) and labeling it "down" when it visits the "top" interface (here, the minus interface where $N_{ex} = N_d$).

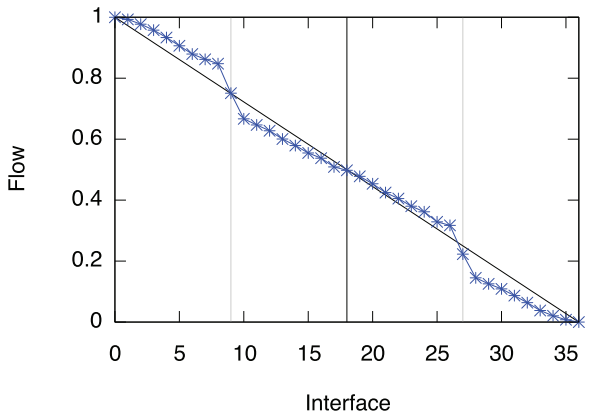

FIG. 4. Replica flow for $N_d = 2$ and $E = 30$. The black vertical line marks the multiple set minus interface; the grey vertical lines mark the multiple state interfaces. We see good flow through the system, with the only significant jumps from imperfect optimization of the acceptance ratio at the multiple state interfaces.

Then $n_i^{\uparrow}$ and $n_i^{\downarrow}$ are simply the number of "up" and "down" replicas (respectively) that visit interface i.

By definition, f_i therefore starts at 1 and ends at 0: at the "bottom" (leftmost) interface, all replicas are marked "up," so $n_{\text{bottom}}^{\downarrow} = 0$ and $f_{\text{bottom}} = 1$. At the "top" (rightmost) interface, all replicas are marked "down," so $n_{\text{top}}^{\uparrow} = 0$ and $f_{\text{top}} = 0$. Ideally, the flow should be linear as a function of the replica index. The flow for our system with two dimers and energy $E = 30$ is shown in Fig. 4. Overall the flow is good, even through the multiple set minus interface marked by the black vertical line. The only divergence from excellent flow is at the multiple state interfaces, marked by grey lines. This is a result of imperfectly optimized acceptance ratios at these multiple state interfaces: whereas the other interfaces typically have acceptance probabilities around 25%-40%, the multiple state interface acceptance probabilities are closer to 10%-15%. The difference in acceptance probability creates a jump in the flow.

Since our approach specifically samples transitions between macrostates, it is important to show that we are also sampling all relevant microstate transitions. In Fig. 5, we consider the distribution of Monte Carlo steps required for a given

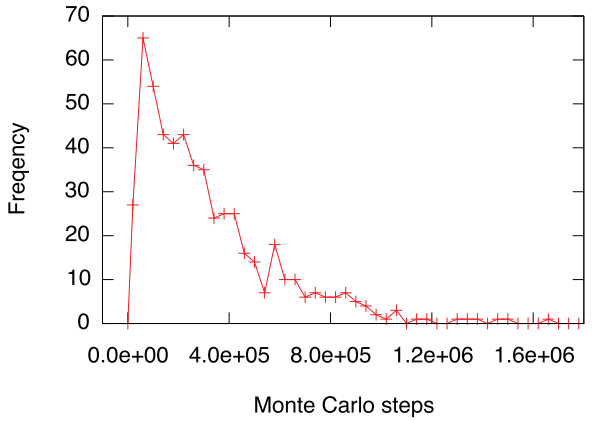

FIG. 5. Histogram of Monte Carlo steps for a given replica switching from one initial microstate to the other in the $N_{ex} = 1$ state, with $N_d = 2$ and $E = 30$. Switching almost always occurs within 10^6 Monte Carlo moves.

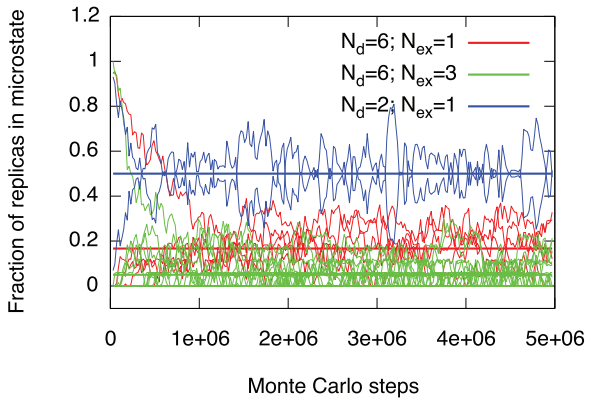

FIG. 6. Populations of various initial microstates as a function of Monte Carlo step. The horizontal lines show the expected average values. Although all states begin with only one microstate populated, the microstate populations relax to fluctuate around the expected values.

replica to switch between the two microstates associated with the $N_{ex} = 1$ in the case $N_d = 2$ and $E = 30$. Nearly all switches occur in less than 10^6 Monte Carlo steps, well within our total simulation time.

Starting a simulation from a specific microstate for each macrostate, Fig. 6 shows the evolution of the fraction of replicas in the microstates as a function of Monte Carlo step. The number of microstates range from 2 to 20, but after initial equilibration of around 10^6 Monte Carlo steps, all microstate populations are fluctuating around their expected values.

An important role of the minus and multiple set minus interfaces is that they improve sampling of the total phase space. The $N_d = 2$ example gives a simple illustration: a good sampler would efficiently visit the entire distribution of dimer center of mass distances. In this system, the timescale of diffusion is much longer than the timescale of a transition path; therefore, it can be difficult to sample the center of mass distribution with path sampling techniques. Fig. 7 shows that including the minus moves improves this sampling by comparing the autocorrelation function of a replica's dimer-dimer

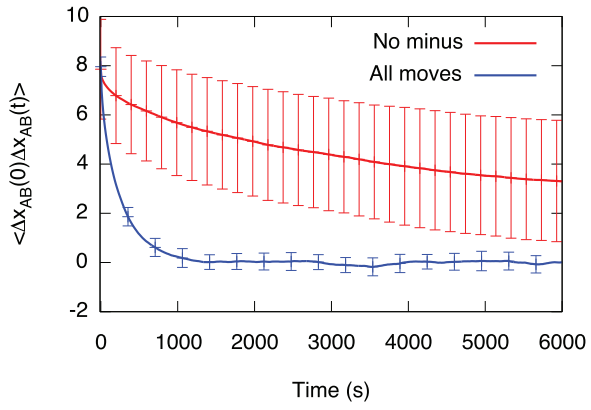

FIG. 7. Autocorrelation function of the dimer-dimer center of mass vector when $N_d = 2$ and $E = 30$, as a function of wall time. Error bars are placed every 1000 shooting moves: since the (multiple set) minus interface is computationally expensive, including it means fewer shooting moves are attempted within a fixed wall time. However, this is more than compensated by the improved sampling. Including the minus moves leads to much faster decorrelation, indicating better sampling of the configurational space.

center of mass vector in $N_d = 2$; $E = 30$. When the minus and multiple-set minus interfaces are included, the correlation decays much more quickly than when they are not, indicating that we are finding new center of mass configurations more quickly.

C. Effect of the environment on the kinetics

The rates shown in Figs. 3(a) and 3(b) are the macrostate-to-macrostate rates. We can get more detailed information about the effect of the environment (specifically, whether other rare transitions have already occurred) by assuming that all microstate-to-microstate transitions within a given macrostate-to-macrostate transition have the same rates (as would be expected in this simple system). We can then put these transitions on an equal footing by dividing the macrostate-macrostate rate by the number of microstate-to-microstate transitions. When increasing the number of dimers, this factor is $N_{sh} = N_d - N_{ex}$; when decreasing the number of dimers, it is N_{ex}. These "microstate-normalized" rates are shown in Fig. 8. Clearly, these rates decrease as a function of N_{ex}, showing that even these normalized rates are dependent on the reaction extent.

Considering the two parts of the RETIS calculation (the flux and the crossing probability), it is clear that the microstate normalization primarily applies to the flux part of the calculation. So we also compare the microstate normalized fluxes. The microstate-normalized flux and the total crossing probabilities are shown in Fig. 9. The microstate-normalized flux for all transitions, extending and condensing, decreases as the number of dimers increases. However, as the number of extended dimers increases within a given total number of dimers, the flux in the extending direction increases, whereas the flux in the condensing direction decreases.

On the other hand, the trends in the flux are overwhelmed by the trend in the crossing probability when the two are combined to give the total rate. All rates (and crossing probabilities) show a downward trend as the number of dimers increases and as the number of extended dimers increases. Both of these trends are manifestations of increasing average potential energy, specifically increasing the average potential energy at the barrier and therefore decreasing the kinetic energy available at the barrier, $E - E^{\ddagger}$. To show that the crossing probability depends on the average kinetic energy, in Fig. 10 we show the results for $N_d = 4$ at different energies and densities. Of course, decreasing the total energy decreases

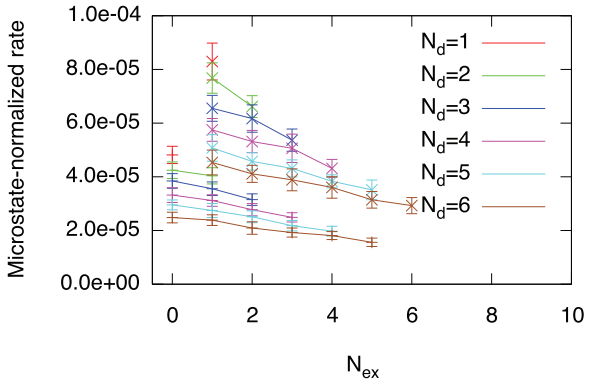

FIG. 8. Microstate-normalized rates for extending and condensing transitions at $E = 30$. Extending transitions are marked with "+," condensing transitions with "×." For both extending and condensing processes, the rate decreases as more dimers are initially extended.

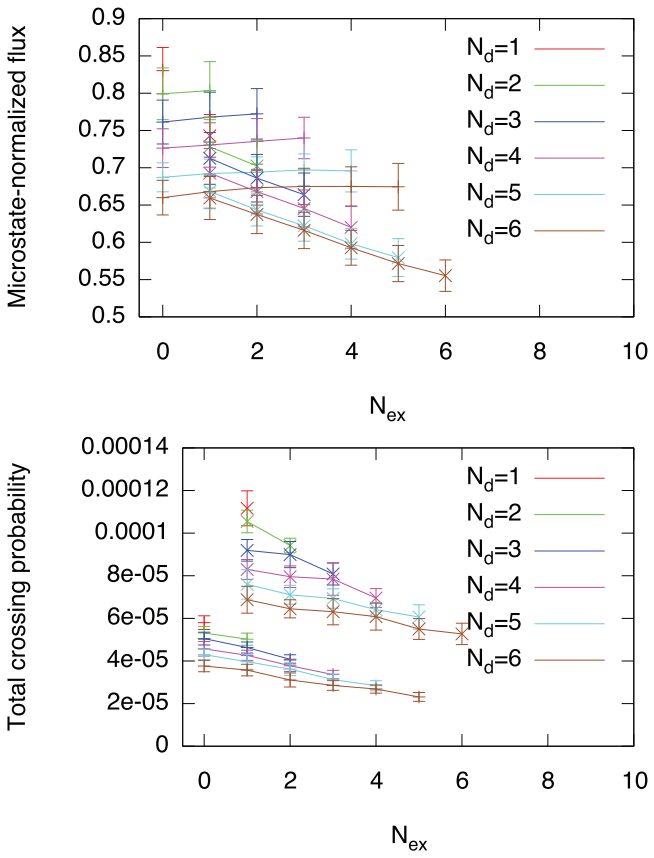

FIG. 9. Components of the microstate-normalized rates at $E = 30$: (a) microstate-normalized flux and (b) total crossing probability. In both plots, extending transitions are marked with "+," and condensing transitions are marked with "×." The microstate-normalized flux shows opposite trend directions for the extending and condensing processes as the number of initially extended dimers is increased, but that is overwhelmed by the trends in the crossing probability.

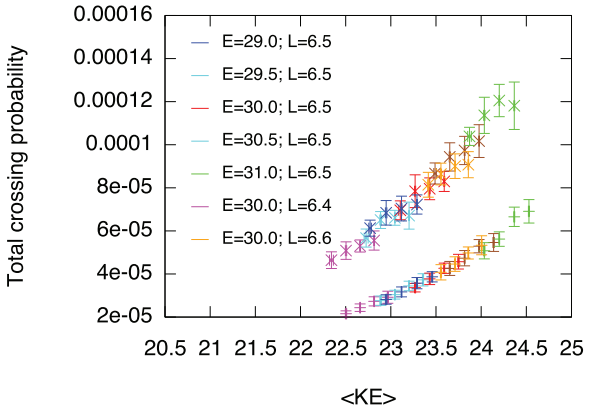

FIG. 10. Crossing probability as a function of average kinetic energy in the initial state. Results are shown for the 4-dimer system at different energies and different densities (square box length given by L). Regardless of energy or density, there is a single curve for the extending process (points marked with "+") and a single curve for the condensing process (points marked with "×").

the average kinetic energy, as does increasing the density (since the purely repulsive intermolecular interactions will lead to a larger average potential energy at higher densities). What Fig. 10 demonstrates is that the crossing probability is determined by the average kinetic energy in the initial state for both the forward and backward transitions, whether we change the total energy or whether we change the density of the system. The results that come from changing the total energy align with the results that come from changing the density. For example, the higher two crossing probabilities of the high density ($E = 30$; $L = 6.4$), have nearly the same average kinetic energy and crossing probabilities as the lower two crossing probabilities of the same system at our standard density ($E = 30$; $L = 6.5$), for both the extending and condensing processes. However, these correspond to different transitions: in the extending process, the higher two crossing probabilities are $N_{ex} = 0 \rightarrow 1$ and $N_{ex} = 1 \rightarrow 2$. The lower two crossing probabilities are $N_{ex} = 2 \rightarrow 3$ and $N_{ex} = 3 \rightarrow 4$. So by changing the density of the system, we change the average kinetic energy such that the high density $N_{ex} = 0 \rightarrow 1$ transition has about the same crossing probability as the standard density $N_{ex} = 2 \rightarrow 3$ crossing probability. This indicates that the average kinetic energy is the primary factor determining the crossing probability, as one might expect from microcanonical transition state theory.

The global state of the system affects the local rate: the average potential energy is affected by the global state, and since total energy is conserved, it also affects the average kinetic energy at the barrier, which in turn determines the crossing probability and the rate. To have different transitions show the same crossing probability, we must change system parameters such as density or total energy.

VI. CONCLUSION

In order to sample the kinetics of a large network of rare events, we have developed and employed a combination of replica exchange MSTIS with multiple interface set definitions. This allows us to define different order parameters for different transitions leaving the same initial state. Efficient use of the multiple interface set approach also requires the multiple interface set minus move, a novel Monte Carlo move for RETIS simulations. To keep the computations tractable, we reduced the size of the network by combining transitions into "macrostate" transitions.

Applying these tools to a simple system of dimers, we were able to study environmental effects on the rare transition, demonstrating that even in very simple systems, the rate for single type of rare transition can depend on its environment — specifically, what other transitions may have preceded it. Although the condensing and extended transitions in the present example have rates on the same order of magnitude, the multiple interface set approach may prove particularly useful when different order parameters can be used to distinguish between two transitions with very different rates.

The method can be easily parallelized by assigning each replica to a different computer node. Of course, as the paths can have different length, the replica exchange move has to wait until all shooting moves are finished. However, the use of a queuing system can avoid unwanted idle time. Alternatively, an asynchronous replica exchange technique could be developed.[38]

More generally, the methodology introduced in this work aims to sample paths in a complex landscape of metastable states, using order parameters to bias the selection of trajectories, but not biasing the dynamics. Our approach therefore combines positive features from both the rare event (free) energy landscape methodology, as well as from the path-based methodology.

We believe the tools developed here can be applied to a wide variety of more complicated systems, including protein folding, nucleation, and crystal transformations, opening up the way for studying large networks for rare events.

ACKNOWLEDGMENTS

This work is part of the research programme VICI 700.58.442, which is financed by the Netherlands Organization for Scientific Research (NWO).

[1] A. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (W.H. Freeman, New York, 1999).
[2] F. Chiti and C. M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006).
[3] G. M. Torrie and J. P. Valleau, Chem. Phys. Lett. **28**, 578 (1974).
[4] E. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral, Chem. Phys. Lett. **156**, 472 (1989).
[5] T. Huber, A. Torda, and W. van Gunsteren, J. Comput. Aided Mol. Des. **8**, 695 (1994).
[6] H. Grubmüller, Phys. Rev. E **52**, 2893 (1995).
[7] A. F. Voter, J. Chem. Phys. **106**, 4665 (1997).
[8] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002).
[9] E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).
[10] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[11] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[12] L. Zheng, M. Chen, and W. Yang, Proc. Natl. Acad. Sci. U.S.A. **105**, 20227 (2008).
[13] Y. Q. Gao, J. Chem. Phys. **128**, 064105 (2008).
[14] C. H. Bennett, in *Algorithms for Chemical Computations*, ACS Symposium Series Vol. 46, edited by R. Christofferson (American Chemical Society, Washington, DC, 1977).
[15] D. Chandler, J. Chem. Phys. **68**, 2959 (1978).
[16] G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9978 (2000).
[17] R. Elber, A. Cardenas, A. Ghosh, and H. A. Stern, Adv. Chem. Phys. **126**, 93 (2003).
[18] L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, J. Chem. Phys. **125**, 024106 (2006).
[19] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[20] C. Dellago and P. G. Bolhuis, Adv. Polym. Sci. **221**, 167 (2009).
[21] T. S. van Erp, D. Moroni, and P. G. Bolhuis, J. Chem. Phys. **118**, 7762 (2003).
[22] R. Allen, D. Frenkel, and P. ten Wolde, J. Chem. Phys. **124**, 024102 (2006).
[23] F. Cerou, Stoch. Anal. Appl. **25**, 417 (2007).
[24] F. Cerou, A. Guyader, T. Lelievre, and D. Pommier, J. Chem. Phys. **134**, 054108 (2011).
[25] J. T. Berryman and T. Schilling, J. Chem. Phys. **133**, 244101 (2010).
[26] M. Villen-Altamirano and J. Villen-Altamirano, Eur. Trans. Telecommun. **13**, 373 (2002).
[27] J. Rogal and P. G. Bolhuis, J. Chem. Phys. **129**, 224107 (2008).
[28] J. Rogal and P. G. Bolhuis, J. Chem. Phys. **133**, 034101 (2010).
[29] T. S. van Erp, Phys. Rev. Lett. **98**, 268301 (2007).
[30] P. G. Bolhuis, J. Chem. Phys. **129**, 114108 (2008).
[31] W.-N. Du and P. G. Bolhuis, J. Chem. Phys. **139**, 044105 (2013).
[32] 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).

[33]W.-N. Du, K. A. Marino, and P. G. Bolhuis, J. Chem. Phys. **135**, 145102 (2011).
[34]C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **110**, 6617 (1999).
[35]P. G. Bolhuis, C. Dellago, and D. Chandler, Faraday Discuss. **110**, 421 (1998).
[36]P. G. Bolhuis and C. Dellago, *Reviews of Computational Chemistry* (Wiley-VCH, Hoboken, 2009).
[37]H. G. Katzgraber, S. Trebst, D. A. Huse, and M. Troyer, J. Stat. Mech. (**2006**) P03018.
[38]E. Gallicchio, R. M. Levy, and M. Parashar, J. Comput. Chem. **29**, 788 (2008).