String method in collective variables: Minimum free energy paths and isocommittor surfaces

Luca Maragliano, Alexander Fischer, and Eric Vanden-EijndenGiovanni Ciccotti

Citation: The Journal of Chemical Physics **125**, 024106 (2006); doi: 10.1063/1.2212942
View online: http://dx.doi.org/10.1063/1.2212942
View Table of Contents: http://aip.scitation.org/toc/jcp/125/2
Published by the American Institute of Physics

String method in collective variables: Minimum free energy paths and isocommittor surfaces

Luca Maragliano,[a] Alexander Fischer,[b] and Eric Vanden-Eijnden[c]
Courant Institute of Mathematical Sciences, New York University, New York, New York 10012

Giovanni Ciccotti[d]
INFM and Dipartimento di Fisica, Università di Roma La Sapienza, Piazzale Aldo Moro, 2, 00185 Roma, Italy

(Received 13 February 2006; accepted 18 May 2006; published online 14 July 2006)

A computational technique is proposed which combines the string method with a sampling technique to determine minimum free energy paths. The technique only requires to compute the mean force and another conditional expectation locally along the string, and therefore can be applied even if the number of collective variables kept in the free energy calculation is large. This is in contrast with other free energy sampling techniques which aim at mapping the full free energy landscape and whose cost increases exponentially with the number of collective variables kept in the free energy. Provided that the number of collective variables is large enough, the new technique captures the mechanism of transition in that it allows to determine the committor function for the reaction and, in particular, the transition state region. The new technique is illustrated on the example of alanine dipeptide, in which we compute the minimum free energy path for the isomerization transition using either two or four dihedral angles as collective variables. It is shown that the mechanism of transition can be captured using the four dihedral angles, but it cannot be captured using only two of them. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2212942]

I. INTRODUCTION

In recent years, many techniques, such as adaptive biasing force[1] (ABF), adiabatic molecular dynamics,[2] metadynamics,[3] etc., have been introduced to map free energy landscapes in a few collective variables of interest for a given reaction. Unfortunately, these techniques become rapidly inefficient as the number of collective variables increases. This is unfortunate since it is often the case that the choice of variables that describe a reaction is speculative, and the smaller their number, the higher the risk to miss some important effects in the reaction. At the other end of the spectrum, several techniques, such as transition path sampling (TPS),[4] the boundary value formulation of Elber *et al.*,[5] or the string method,[6] have been recently designed to determine reaction pathways without making any *a priori* assumption about its mechanism. Unfortunately, at present day these techniques do not allow to investigate very large systems, especially if the reaction is dominated by entropic effects and many different pathways exist, due, e.g., to symmetry by permutation of certain degrees of freedom (such as the one associated with the solvent).

The purpose of this paper is to propose a method which borrows from both classes of approaches listed above and eliminates some of their restrictions. Specifically, we propose to combine the string method with some biased sampling method to determine minimum free energy paths in some selected variables. This is worthwhile to do since minimum free energy paths (MFEPs), and the free energy along them, are often the quantities of most interest. By restricting ourselves to the determination of minimum free energy paths rather than the full free landscape itself, we are able to lift the restriction on the number of collective variables that can be investigated simultaneously. Indeed, the string method is very robust and efficient at determining minimum (free) energy paths in a given landscape, and as an input it only requires the calculation of the mean force. This calculation can be performed locally, which means that the cost of the calculation scales linearly with the number of points along the discretized string, but is independent of the dimension. This is in marked contrast with the free energy mapping techniques listed above whose cost increases exponentially with the number of collective variables used to describe the reaction. At the same time, by eliminating many degrees of freedom that are believed to be irrelevant for the reaction, we eliminate the potential sources of difficulty for the string method when it is applied in the original state space of the system. Working in free energy space is also easier because the free energy landscape is usually much smoother than the original potential energy landscape of the system.

Provided that the number of collective variables is large enough, the approach that we propose here captures the mechanism of the reaction in the following sense. As is well known by now, the best reaction coordinate to describe the reaction, in fact, *the* reaction coordinate, is the committor function which, at any point in phase space, gives the probability that a trajectory initiated at this point will reach first the product state rather than the reactant state. As was shown

[a]Electronic mail: maraglia@cims.nyu.edu
[b]Electronic mail: fischer@cims.nyu.edu
[c]Electronic mail: eve2@cims.nyu.edu
[d]Electronic mail: giovanni.ciccotti@roma1.infn.it

in Ref. 7 (see also Refs. 8–10) the committor function allows one to compute the probability density of reactive trajectories, their probability current, and the rate of the reaction. In general, the committor function is a very complicated function of the positions and momenta of all the atoms in the system. Suppose, however, that this function can be approximated, at least locally (in a sense specified below), by some (possibly large) set of suitable collective variables function only of the positions. Then the isocommittor surfaces (i.e., the hypersurfaces in the original state space in which the committor function is constant) can be locally recovered from the MFEP: they simply are the lift up in state space of hyperplanes defined in the collective variables space which make a specific angle with the MFEP. Among other things, the knowledge of the isocommittor surfaces allows one to identify *a priori* (i.e., without identifying actual reactive trajectories beforehand) the transition region defined as the region in the committor $\frac{1}{2}$ surface with significant probability. This is promising, since extracting the same information from the TPS path ensemble is not straightforward.[11,12]

In the following, we first recall in Sec. II the definition of MFEPs on the free energy landscape in some collective variables. The definition of the MFEP is quite simple and natural, but it is somewhat more complicated to give a full justification of why the MFEP is relevant to understand the mechanism of a reaction and, in particular, to determine the isocommittor surfaces. So, we defer this justification to Sec. III. In Sec. IV we discuss the implementation of the string method in the present context, by first recalling the method itself, then detailing the way the mean force is locally computed. In this work, we use a specific type of biased sampling for the mean force calculation, but this can be straightforwardly generalized, say, to blue moon sampling.[13] In Sec. V we illustrate the technique on the example of alanine dipeptide in vacuum. This system has been extensively used as a benchmark for free energy and conformational sampling methods.[12,14–21] In this example, we show how to determine the mean free energy paths in two and four dihedral angles simultaneously and we discuss consequences of these results in terms of the mechanism of isomerization of alanine dipeptide. In particular, we show that the mechanism of transition cannot be captured by two dihedral angles only but, rather, requires four dihedral angles. These four angles allow one to identify *a priori* the transition region for the reaction, and this can be checked *a posteriori* by computing the committor distribution of this region and verifying that it is indeed peaked at $\frac{1}{2}$. Finally, some concluding remarks are given in Sec. VI.

II. MINIMUM FREE ENERGY PATH

Consider a system in the *NVT* ensemble whose statistics is described by the Boltzmann-Gibbs probability density function

$$\rho(x) = Z^{-1} e^{-\beta V(x)}, \tag{1}$$

where $Z=\int_{\mathbb{R}^n} e^{-\beta V(x)} dx$, $\beta = 1/k_B T$ is the inverse temperature, and $x \in \mathbb{R}^n$ are the Cartesian coordinates. Thus we assume that we have already integrated out the part of the density arising from the momenta and we focus on its configurational component. We also assume for simplicity that there are no molecular constraints on the system; these could be straightforwardly accounted for by modifying (1) appropriately.

Suppose that one introduces N collective variables,

$$\theta(x) = (\theta_1(x), \ldots, \theta_N(x)), \tag{2}$$

which are functions of x and are good candidates to describe some reaction of interest in the system. The free energy associated with $\theta(x)$ is the function depending on $z = (z_1, \ldots, z_N)$ defined as

$$F(z) = -k_B T \ln\left(Z^{-1} \int_{\mathbb{R}^n} e^{-\beta V(x)} \times \delta(z_1 - \theta_1(x)) \cdots \delta(z_N - \theta_N(x)) dx \right). \tag{3}$$

In Eq. (8) below, we give the equation which determines the MFEP on $F(z)$. As mentioned earlier, a full justification of this equation and, in fact, of why the MFEP is relevant at all, is somewhat more involved, so we defer this justification to Sec. III. But Eq. (8) for the MFEP is rather natural and it can also be introduced in a simple way which we explain now.

To begin, it is useful to recall first what a minimum energy path (MEP) is in the original landscape $V(x)$ in terms of the Cartesian coordinates x. A MEP is a path which connects two minima of $V(x)$ via a saddle point and corresponds to the steepest descent path on $V(x)$ from this saddle point. In other words, a MEP is composed of two heteroclinic orbits connecting a saddle point to two minima by the steepest descent dynamics $\dot{x}(t) = -\nabla V(x(t))$. Since the two relevant solutions of this equation take doubly infinite time to go out of the saddle point and in the minima, it is useful to change perspective, to look at the MEP as a curve in configuration space and represent it using a different parametrization than time. Let us represent the MEP by the curve $x(\alpha)$, where $\alpha \in [0,1]$ is the parameter used to parametrize the curve. Since by definition the force $-\nabla V$ must be everywhere tangent to the MEP, we must have

$$\frac{dx_k(\alpha)}{d\alpha} \text{ parallel to } \frac{\partial V(x(\alpha))}{\partial x_k}. \tag{4}$$

If the MEP connects the two minima of $V(x)$ located at x_a and x_b, Eq. (4) must be supplemented by the boundary conditions $x(0)=x_a$ and $x(1)=x_b$. Notice that Eq. (4) implies that the component of the force perpendicular to the curve $x(\alpha)$ is zero everywhere along the curve, which can also be written shorthand as $0=(\nabla V)^{\perp}$.

Let us now consider things in the space of $z=\theta(x)$ and, for simplicity, suppose first that $N=n$, i.e., the map $x \mapsto z = \theta(x)$ is simply a change of coordinates. Letting $z(\alpha) = \theta(x(\alpha))$ and $V(x) = U(\theta(x))$, Eq. (4) implies that

$$\frac{dz_i(\alpha)}{d\alpha} = \sum_{k=1}^{n} \frac{\partial \theta_i(x(\alpha))}{\partial x_k} \frac{dx_k(\alpha)}{d\alpha}$$

$$\text{parallel to } \sum_{k=1}^{n} \frac{\partial \theta_i(x(\alpha))}{\partial x_k} \frac{\partial V(x(\alpha))}{\partial x_k}$$

$$= \sum_{j,k=1}^{n} \frac{\partial \theta_i(x(\alpha))}{\partial x_k} \frac{\partial \theta_j(x(\alpha))}{\partial x_k} \frac{\partial U(z(\alpha))}{\partial z_j}. \tag{5}$$

When $N<n$, some extra averaging is required due to the projection step from the original Cartesian space with n dimensions to the space of $z=\theta(x)$ with N dimensions. This is the step which cannot be justified satisfactorily within the present line of argument and requires the more elaborate explanation given in Sec. III. It is shown there that it amounts to replacing $U(z)$ by $F(z)$ and the tensor $\Sigma_k(\partial\theta_i/\partial x_k)\times(\partial\theta_j/\partial x_k)$ by its average M_{ij} defined as

$$M_{ij}(z) = Z^{-1} e^{\beta F(z)} \int_{\mathbb{R}^n} \sum_{k=1}^{n} \frac{\partial \theta_i(x)}{\partial x_k} \frac{\partial \theta_j(x)}{\partial x_k} e^{-\beta V(x)} \times \delta(z_1 - \theta_1(x)) \cdots \delta(z_N - \theta_N(x)) dx. \tag{6}$$

Making these substitutions in Eq. (5), this equation becomes

$$\frac{dz_i(\alpha)}{d\alpha} \text{ parallel to } \sum_{j=1}^{N} M_{ij}(z(\alpha)) \frac{\partial F(z(\alpha))}{\partial z_j}. \tag{7}$$

This equation implies that the component of $M\nabla_z F$ perpendicular to the path $z(\alpha)$ vanishes identically, i.e., $0=(M\nabla_z F)^{\perp}$, which is a shorthand notation for

$$0 = \sum_{j,k=1}^{N} P_{ij}(\alpha) M_{jk}(z(\alpha)) \frac{\partial F(z(\alpha))}{\partial z_k}, \tag{8}$$

where $P_{ij}(\alpha)$ is the projector on the plane perpendicular to the path at $z(\alpha)$,

$$P_{ij}(\alpha) = \delta_{ij} - \hat{t}_i(\alpha)\hat{t}_j(\alpha), \quad \hat{t}_i(\alpha) = \frac{\partial z_i/\partial\alpha}{|\partial z_i/\partial\alpha|}. \tag{9}$$

If one is interested in finding the MFEP joining two minima of the free energy located at z_a and z_b, Eq. (8) must be solved with the boundary conditions $z_a=z(0)$ and $z_b=z(1)$. How to find this solution in practice is explained in Sec. IV.

III. THEORETICAL BACKGROUND

The free energy in (3) has a well-defined probabilistic meaning: $e^{-\beta F(z)}$ gives the equilibrium probability density function of the variables $z=\theta(x)$. On the other hand, often one has in mind to extract from $F(z)$ not only equilibrium properties but also dynamical ones: for instance, saddle points on the free energy surface are often identified with the transition states of a reaction and the minimum free energy paths with the most likely reaction paths. While appealing, such identifications are by no means straightforward to justify. The purpose of this section is to elucidate the conditions when (and the sense in which) the MFEP is indeed the most likely reaction path, and discuss the underlying assumptions this entails about the collective coordinate chosen to describe the reaction. The reader more interested in practical implementation issues and applications may skip this section in first reading.

A. Reaction coordinate and committor function

Consider a system governed by

$$\begin{cases} \dot{x}_i(t) = v_i(t), \\ \dot{v}_i(t) = -\dfrac{\partial V(x(t))}{\partial x_i} - \gamma m_i^{-1} v_i(t) + \sqrt{2\gamma\beta^{-1}m_i^{-1}}\,\eta_i(t). \end{cases} \tag{10}$$

Equation (10) is written in mass-weighted coordinates and in this equation, γ is the friction coefficient, m_i is the mass of the particle with coordinate x_i, and $\eta_i(t)$ a white noise satisfying $\langle \eta_i(t)\eta_j(t')\rangle = \delta_{ij}\delta(t-t')$. The dynamics in Eq. (10) is consistent with the Boltzmann-Gibbs probability density in Eq. (1).

Suppose that one is interested in understanding the mechanism of reaction between two sets $A \in \mathbb{R}^{2n}$ and $B \in \mathbb{R}^{2n}$ in phase space which may, for instance, be the regions in phase space associated with the reactant and product states of a reaction. It has been shown in Ref. 7 (see also Refs. 8–10) that the best reaction coordinate to describe the mechanism of reaction between A and B, in fact, *the* reaction coordinate, is the committor function $q(x,v)$. $q(x,v)$ gives the probability that the trajectory solution of Eq. (10) initiated at the point (x,v) will reach (i.e., get committed to) B first rather than A. As shown in Ref. 7, $q(x,v)$ allows one to compute many statistical quantities of interest regarding the reactive trajectories by which the reaction from A to B occurs like, in particular, the rate constant of this reaction. It is also well known that $q(x,v)$ satisfies the following partial differential equation, referred to as the backward Kolmogorov equation in the probability literature:[22]

$$\begin{cases} 0 = Lq \equiv \displaystyle\sum_{i=1}^{n} \left(v_i \frac{\partial q}{\partial x_i} - \frac{\partial V(x)}{\partial x_i} \frac{\partial q}{\partial v_i} \right) \\ \qquad + \gamma \displaystyle\sum_{i=1}^{n} m_i^{-1} \left(-v_i \frac{\partial q}{\partial v_i} + \beta^{-1} \frac{\partial^2 q}{\partial v_i^2} \right), \\ q|_{(x,v)\in A} = 0, \quad q|_{(x,v)\in B} = 1. \end{cases} \tag{11}$$

Equation (11) is a partial differential equation in $2n$ dimensions which, as such, is way too complex to be solved by traditional numerical methods. Next we simplify this equation by assuming that $q(x,v)$ does not depend on v and depends on x only via the collective variables chosen to describe the reaction.

B. Variational formulation and collective variables

Consistent with the assumption that the collective variables $(\theta_1(x),\ldots,\theta_N(x))$ are a good set of variables to describe the reaction at hand, we will suppose that the function $q(x,v)$ can be approximated by a function depending only on $(\theta_1(x),\ldots,\theta_N(x))$, i.e.,

$$q(x,v) \approx f(\theta_1(x), \ldots, \theta_N(x)), \tag{12}$$

for some f to be determined. In order to implement the approximation (12) in (11), we use the following variational formulation of this equation:

$$I = \int_{\mathbb{R}^n \times \mathbb{R}^n} e^{-\beta H} |Lq|^2 dx dv, \tag{13}$$

where $H(x,v) = \frac{1}{2}|v|^2 + V(x)$ is the Hamiltonian. Since $I \geqslant 0$ and $I=0$ if q is the solution of (11), this solution minimizes I over all test functions q satisfying $q|_{(x,v)\in A}=0$, $q|_{(x,v)\in B}=1$. Consistent with the approximation (12) we can then look for the minimum of (13) over $q(x,v)$ depending only on $(\theta_1(x), \ldots, \theta_N(x))$ as in (12) and thereby determine the optimal f. Inserting (12) into (13) and integrating out the momenta, we are left with

$$\begin{aligned} I &= C \int_{\mathbb{R}^n} e^{-\beta V} \sum_{i,j=1}^{N} \frac{\partial f}{\partial z_i} \nabla \theta_i \cdot \nabla \theta_j \frac{\partial f}{\partial z_j} dx \\ &= C \int_{\mathbb{R}^N} \int_{\mathbb{R}^n} e^{-\beta V} \sum_{i,j=1}^{N} \frac{\partial f}{\partial z_i} \nabla \theta_i \cdot \nabla \theta_j \frac{\partial f}{\partial z_j} \\ &\quad \times \delta(z_1 - \theta_1(x)) \cdots \delta(z_N - \theta_N(x)) dx dz \\ &= CZ \int_{\mathbb{R}^N} e^{-\beta F} \sum_{i,j=1}^{N} \frac{\partial f}{\partial z_i} M_{ij}(z) \frac{\partial f}{\partial z_j} dz, \end{aligned} \tag{14}$$

where C is a constant whose precise value is irrelevant and F and M are, respectively, the free energy (3) and the tensor defined in (6). We are thus left to minimize

$$\bar{I} = \int_{\mathbb{R}^N} e^{-\beta F} \sum_{i,j=1}^{N} \frac{\partial f}{\partial z_i} M_{ij}(z) \frac{\partial f}{\partial z_j} dz \tag{15}$$

over all functions $f(z)$ subject to $f|_{z\in a}=0$, $f|_{z\in b}=1$, where a and b are the representation in z space of the sets A and B, respectively.

The Euler-Lagrange equation associated with the minimization of (15) is

$$\begin{cases} 0 = \displaystyle\sum_{i,j=1}^{N} \frac{\partial}{\partial z_i} \left(e^{-\beta F(z)} M_{ij}(z) \frac{\partial f}{\partial z_j} \right), \\ f|_{z\in a} = 0, \quad f|_{z\in b} = 1. \end{cases} \tag{16}$$

Thus, we are left with analyzing this equation to determine $f(z)$ and thereby $q(x,v)$ via (12).

C. MFEP as maximum likelihood path

Equation (16) is the backward Kolmogorov equation associated with the stochastic differential equation[22]

$$\begin{aligned} \dot{z}_i(\tau) &= -\sum_{j=1}^{N} \left(M_{ij}(z(\tau)) \frac{\partial F(z(\tau))}{\partial z_j} - \beta^{-1} \frac{\partial M_{ij}(z(\tau))}{\partial z_j} \right) \\ &\quad + \sqrt{2\beta^{-1}} \sum_{j=1}^{N} M_{ij}^{1/2}(z(\tau)) \eta_j(\tau), \end{aligned} \tag{17}$$

where $\eta_i(\tau)$ is a white noise satisfying $\langle \eta_i(\tau) \eta_j(\tau') \rangle = \delta_{ij} \delta(\tau - \tau')$ and τ is an artificial time. Since the argument above holds for any sets A and B, we are thus led to the following conclusion: if the collective variables $(\theta_1(x), \ldots, \theta_N(x))$ are good variables to describe the mechanism of the reaction, then this mechanism can also be understood by looking at reactions in the system governed by (17). [Notice that we are talking about the mechanism of the reaction and not its timing or rate: this is why τ in (17) is an artificial time whose relation with the physical time t in (10) is unspecified by our argument.]

By assuming that the collective variables $(\theta_1(x), \ldots, \theta_N(x))$ are good variables to describe the reaction, we must assume that the relevant metastable sets in phase space correspond to regions around the minima of $F(z)$ and that the free energy barriers between these minima are much larger than $k_B T$. Then the transition from one local minima of $F(z)$ to another is a rare event. This means that, when one of these rare events happens, it happens almost always in the same predictable way because any other way is even much more unlikely (this is the essence of the Freidlin-Wentzell theory of large deviations, see Ref. 23). Thus we are left to determine the trajectory of maximum likelihood that connects local minima of $F(z)$ by the dynamics in (17). How to determine this trajectory is somewhat technical, so we defer this explanation to Appendix A. It is shown there that the trajectory of maximum likelihood can be constructed out of specific solutions of the following ordinary differential equation obtained from (17) in the limit as $\beta \to \infty$:

$$\dot{z}_i(\tau) = -\sum_{j=1}^{N} M_{ij}(z(\tau)) \frac{\partial F(z(\tau))}{\partial z_j}. \tag{18}$$

More precisely, the trajectory of maximum likelihood by which the solution of (17) hops from the vicinity of the free energy minimum located at $z_a \in a$ to the vicinity of the one located at $z_b \in b$ is made of the heteroclinic orbits of (18), i.e., those solutions of (18) which connect a saddle point of $F(z)$ to two minima of $F(z)$ in (doubly) infinite time (and time needs to be reversed on one of the branch to allow the solution to go uphill). But as we already know from Sec. II, the MFEP defined by (8) is a curve in z space whose graph coincides with (i.e., whose geometric location in z space is the same as) the one of the heteroclinic orbit of (18). Thus the MFEP is the most likely path of transition between z_a and z_b.

D. Isocommittor surfaces around the MFEP

From the argument in Sec. III C and Appendix A, it follows that transitions are most likely to occur around the MFEP. Here we determine the committor function $f(z)$ [and hence $q(x,v)$ via (12)] around this MFEP. To this end, let

$z(\alpha)$, $\alpha \in [0,1]$ be the MFEP solution of (8) and denote by $g(\alpha)=f(z(\alpha))$ the value of the committor function $f(z)$ along this MFEP. Consider the family of isocommittor surfaces $S(\alpha)$ where $f(z)$ takes the value $f(z(\alpha))$, i.e., $S(\alpha)=\{z:f(z)=f(z(\alpha))\}$, and let $\Omega(\alpha)$ be the region in z space between a, the metastable set around z_a at the boundary of which $f(z)=0$ [see the boundary condition in (16)], and $S(\alpha)$.

Integrating (16) over $\Omega(\alpha)$, using the divergence theorem, we deduce that

$$\begin{aligned} 0 &= \int_{\Omega(\alpha)} \sum_{i,j=1}^{N} \frac{\partial}{\partial z_i}\left(e^{-\beta F(z)} M_{ij}(z) \frac{\partial f}{\partial z_j}\right) dz \\ &= -\int_{\partial a} \sum_{i,j=1}^{N} \hat{n}_{\partial a,i}(z) M_{ij}(z) \frac{\partial f}{\partial z_j} e^{-\beta F(z)} d\sigma_{\partial a}(z) \\ &\quad + \int_{S(\alpha)} \sum_{i,j=1}^{N} \hat{n}_{S(\alpha),i}(z) M_{ij}(z) \frac{\partial f}{\partial z_j} e^{-\beta F(z)} d\sigma_{S(\alpha)}(z), \end{aligned} \tag{19}$$

where $\hat{n}_{\partial a,i}(z)$ is the outward pointing unit normal on ∂a and $d\sigma_{\partial a}(z)$ is the surface element on ∂a and similarly for $\hat{n}_{S(\alpha),i}(z)$ and $d\sigma_{S(\alpha)}(z)$ on $S(\alpha)$. The surface integral on ∂a in (19) is a constant in α which we shall denote by C.

Now, since all the action takes place near the MFEP, it must be the case that $e^{-\beta F(z)}$ restricted on $S(\alpha)$ is strongly peaked around $z(\alpha)$. We will assume that this allows one to approximate the surface integral on $S(\alpha)$ in (19) by a surface integral where all the factors in the integrand but $e^{-\beta F(z)}$ are evaluated at $z(\alpha)$ and hence can be pulled out of the integral, i.e., we will assume that (19) can be approximated by

$$C \approx \sum_{i,j=1}^{N} \hat{n}_{S(\alpha),i}(z(\alpha)) M_{ij}(z(\alpha)) \frac{\partial f(z(\alpha))}{\partial z_j} Q(\alpha), \tag{20}$$

where

$$Q(\alpha) = \int_{S(\alpha)} e^{-\beta F(z)} d\sigma_{S(\alpha)}(z). \tag{21}$$

In fact, if $e^{-\beta F(z)}$ is strongly peaked in $S(\alpha)$, we can even assume that the integral in (21) can be approximated by the surface integral on $P(\alpha)$, the hyperplane tangent to $S(\alpha)$ at $z(\alpha)$. In other words, denoting by $\hat{n}(\alpha)=\hat{n}_{S(\alpha),i}(z(\alpha))$ the unit normal to $P(\alpha)$, we have

$$Q(\alpha) \approx \int_{P(\alpha)} e^{-\beta F(z)} d\sigma_{P(\alpha)}(z). \tag{22}$$

Remarkably, the argument above already indicates the orientation of the planes $P(\alpha)$ [and hence of the isocommittor surfaces $S(\alpha)$, at least locally] with respect to the MFEP $z(\alpha)$. Indeed, since $e^{-\beta F(z)}$ is strongly peaked at $z(\alpha)$ in $P(\alpha)$, $z(\alpha)$ must be the location of the minimum of $F(z)$ in $P(\alpha)$, which implies that the gradient of $F(z)$ has zero component in the plane and hence

$$\hat{n}_i(\alpha) \text{ parallel to } \frac{\partial F(z(\alpha))}{\partial z_i}. \tag{23}$$

Combining this equation with Eq. (7) for the MFEP, we deduce that

$$\sum_{j=1}^{N} M_{ij}(z(\alpha)) \hat{n}_j(\alpha) \text{ parallel to } \frac{dz_i(\alpha)}{d\alpha}, \tag{24}$$

which, knowing $z(\alpha)$, fixes $\hat{n}(\alpha)$. Thus the only thing which remains to be determined is the value of $f(z)$ on $P(\alpha)$ near $z(\alpha)$, i.e., the value of $f(z(\alpha))$.

To determine the latter, we follow an argument similar to that given in Sec. 9.3.3 of Ref. 24 in a different context. First, notice that

$$\frac{df(z(\alpha))}{d\alpha} = \sum_{i=1}^{N} \frac{dz_i(\alpha)}{d\alpha} \frac{\partial f(z(\alpha))}{\partial z_i}. \tag{25}$$

Combining this expression with (24) we deduce that

$$\frac{df(z(\alpha))}{d\alpha} = R(\alpha) \sum_{i,j=1}^{N} M_{ij}(z(\alpha)) \hat{n}_j(\alpha) \frac{\partial f(z(\alpha))}{\partial z_i}, \tag{26}$$

where

$$R(\alpha) = \frac{|dz(\alpha)/d\alpha|}{|M(z(\alpha))\hat{n}(\alpha)|}. \tag{27}$$

Here $|dz(\alpha)/d\alpha|$ denotes the Euclidean norm of the vector with components $dz_i(\alpha)/d\alpha$ and $|M(z(\alpha))\hat{n}(\alpha)|$ of the one with components $\Sigma_{j=1}^{N} M_{ij}(z(\alpha))\hat{n}_j(\alpha)$. Using (26) in (20), and solving the resulting equation for $df(z(\alpha))/d\alpha$ we arrive at

$$\frac{df(z(\alpha))}{d\alpha} = \frac{CR(\alpha)}{Q(\alpha)}. \tag{28}$$

Solving this equation with the boundary conditions $f(z(0))=0$ and $f(z(1))=1$ determines both $f(z(\alpha))$ and C:

$$f(z(\alpha)) = \frac{\int_0^\alpha R(\alpha')/Q(\alpha')d\alpha'}{\int_0^1 R(\alpha')/Q(\alpha')d\alpha'}, \tag{29}$$

where the denominator is $1/C$.

Expression (29) can be estimated by evaluating the integral in (22) defining $Q(\alpha)$ by Laplace method. Indeed since the factor $e^{-\beta F(z)}$ in this integral is strongly peaked at $z(\alpha)$, the integral can be approximated by

$$Q(\alpha) \approx G(\alpha) e^{-\beta F(z(\alpha))}, \tag{30}$$

where $G(\alpha)$ is a prefactor term which accounts for the curvature of $F(z)$ in $P(\alpha)$ around $z(\alpha)$. Inserting (30) in (29) and evaluating this integral again by Laplace method, we deduce that the integral at the numerator will be much smaller than the one at the denominator if $\alpha<\alpha_s$, where α_s is the value where $F(z(\alpha))$ reaches its maximum for $\alpha \in [0,1]$; the two integrals will be of the same order if $\alpha>\alpha_s$, and the integral at the numerator will be approximately half of the one at the numerator if $\alpha=\alpha_s$. Summarizing,

$$f(z(\alpha)) \approx \begin{cases} 0 & \text{if } \alpha < \alpha_s \\ \frac{1}{2} & \text{if } \alpha = \alpha_s \\ 1 & \text{if } \alpha > \alpha_s. \end{cases} \tag{31}$$

This expression indicates that the committor $\frac{1}{2}$ surface is approximately the plane $P(\alpha_s)$ where $F(z(\alpha))$ is maximum. This is the estimate that we will use in Sec. V to verify *a posteriori* that the plane $P(\alpha_s)$ identified by the string

method is indeed a good approximation of the isocommittor $\frac{1}{2}$ surface (which, at the same time, is a check whether the collective variables that we use are a good set or not).

We note that estimates for $f(z(\alpha))$ more accurate than (31) can be obtained by more elaborate sampling in $P(\alpha)$ of the distribution proportional to $e^{-\beta F(z)}$ to obtain a more accurate estimate of $Q(\alpha)$ to be used in (29).

IV. STRING METHOD AND MEAN FORCE CALCULATION

Now we explain how to solve (8) and identify the minimum free energy path $z(\alpha)$ by combining the string method with a specific sampling technique to evaluate locally the mean force $\nabla_z F(z)$ and the tensor $M(z)$.

A. String method

Let us first assume that $\nabla_z F(z)$ and $M(z)$ are known. We will come back to the issue of determining them in Sec. IV B. Then, (8) can be solved by a straightforward extension of the string method.[6] The string method constructs an evolution equation for a parametrized curve, which we refer to as a string, which is such that any initial guess for the string converges to a solution of (8) as time evolves. Let us denote the string by $z(\alpha,t)$, where $\alpha \in [0,1]$ is used to parametrize the string. The simplest way to evolve $z(\alpha,t)$ so that it converges to a MFEP is to use

$$\frac{\partial z_i(\alpha,t)}{\partial t} = -\sum_{j,k=1}^{N} P_{ij}(\alpha,t) M_{jk}(z(\alpha,t)) \frac{\partial F(z(\alpha,t))}{\partial z_k}, \tag{32}$$

where $P_{ij}(\alpha,t)$ is defined as in (9) with $z(\alpha)=z(\alpha,t)$. If z_a and z_b are the locations of the two minima of the free energy and (32) is solved with the boundary conditions $z(0,t)=z_a$ and $z(1,t)=z_b$, then as $t\to\infty$ the solution of (32) converges to a MFEP connecting z_a and z_b since (8) is a stable fixed point of (32). Notice that if the locations of the two minima of the free energy, z_a and z_b, are not known beforehand, then one possibility is to let the end point of $z(\alpha,t)$ free to evolve by steepest descent in the potential, i.e., take

$$\frac{\partial z_i(0,t)}{\partial t} = -\frac{\partial F(z(0,t))}{\partial z_i} \text{ and } \frac{\partial z_i(1,t)}{\partial t} = -\frac{\partial F(z(1,t))}{\partial z_i}. \tag{33}$$

The equations in (33) for the end points can be solved in parallel with (32), or they can be integrated independently first to locate the two minima of the free energy which are then kept fixed as boundary conditions for (32). In either case, provided only that the end points of the initial guess are sufficiently close to z_a and z_b, as time evolves the string will converge to a mean free energy path solution of (8) with $z(0)=z_a$ and $z(1)=z_b$.

Equation (32), however, may lead to a practical difficulty, namely, that it does not constrain the parametrization of the curve $z(\alpha,t)$ during the evolution. This may lead to stability problems in the numerical scheme to integrate (32), for instance, if the points along the discretized string drift away from one another, leaving the string under-resolved in certain portions along it. However, this difficulty is very easy to solve. At any given time we can reparametrize $z(\alpha,t)$ in a way which is convenient, then evolve this reparametrized curve via (32). This operation of reparametrization does not change the geometric location of the curve and hence it does not interfere with (32), bringing $z(\alpha,t)$ towards the MFEP as $t\to\infty$. However, this reparametrization step, if done often enough, prevents any numerical instability (32) may lead to if $z(\alpha,t)$ is let free to evolve and never reparametrized. How to implement this reparametrization step in practice is very simple and will be explained in Sec. IV C.

B. Mean force calculation

The integration of (32) requires to evaluate the mean force $\nabla_z F(z)$ and the tensor $M(z)$ locally along the string. Since the tensor $M(z)$ is given in (6) in terms of a conditional expectation on $\theta(x)=z=cst$ and $\nabla_z F$ can also be expressed in terms of such a conditional expectation, we could compute both via constrained dynamical simulation on $\theta(x)=z=cst$ using, e.g., the blue moon sampling method.[13] In this paper, however, we will use a different approach, not because it is more efficient but because it is simpler to implement.

Suppose that we aim at computing $\nabla_z F(z)$ and the tensor $M(z)$ at a given value of z, for instance, the current value of an image along the string. Consider the equation

$$\dot{x}_i(t) = -\frac{\partial V(x(t))}{\partial x_i} - k\sum_{j=1}^{N} (\theta_j(x(t)) - z_j)\frac{\partial \theta_j(x(t))}{\partial x_i} + \sqrt{2k_B T}\eta_i(t), \tag{34}$$

where $k>0$ is a parameter and $\eta(t)$ is a white noise, i.e., a Gaussian process with mean zero and covariance $\langle \eta_i(t)\eta_j(t')\rangle = \delta_{ij}\delta(t-t')$. Equation (34) is the overdamped dynamics in the extended potential

$$U_{k,z}(x) = V(x) + \frac{k}{2}\sum_{j=1}^{N} (\theta_j(x) - z_j)^2. \tag{35}$$

As a result, the equilibrium probability density function for (34) is the Boltzmann-Gibbs density

$$\rho_{k,z}(x) = Z_{k,z}^{-1} e^{-\beta U_{k,z}(x)}, \tag{36}$$

where $Z_{k,z}=\int_{\mathbb{R}^n} e^{-\beta U_{k,z}(x)}dx$. Consider the estimator

$$\frac{k}{T}\int_0^T (z_j - \theta_j(x(t)))dt. \tag{37}$$

By ergodicity, we have

$$\lim_{T\to\infty} \frac{k}{T}\int_0^T (z_j - \theta_j(x(t)))dt = k\int_{\mathbb{R}^n} (z_j - \theta_j(x))\rho_{k,z}(x)dx \equiv \frac{\partial F_k(z)}{\partial z_j}, \tag{38}$$

where

$$F_k(z) = -k_B T \ln\left(Z^{-1}\int_{\mathbb{R}^n} e^{-\beta U_{k,z}(x)}dx\right). \tag{39}$$

[The constant $Z=\int_{\mathbb{R}^n} e^{-\beta V(x)}dx$ in this expression is added for dimensional consistency but is otherwise arbitrary.] The key

property of $\rho_{k,z}(x)$ that we use next is the following: given any test function $f(x)$, we have

$$\lim_{k\to\infty}\int_{\mathbb{R}^n} f(x)\rho_{k,z}(x)dx = Z^{-1}e^{\beta F(z)}\int_{\mathbb{R}^n} f(x)e^{-\beta V(x)} \times \delta(z_1-\theta_1(x))\cdots\delta(z_N-\theta_N(x))dx. \quad (40)$$

As a result, for k large enough, (34) allows us to approximate expectations conditional on $\theta(x)=z$. In particular, it follows from (39) and (40) that $\lim_{k\to\infty}\nabla_z F_k(z)=\nabla_z F(z)$, and hence, combining (37), (38), and (40) that, for k and T large enough,

$$\frac{k}{T}\int_0^T (z_j-\theta_j(x(t)))dt \approx \frac{\partial F(z)}{\partial z_j}. \quad (41)$$

Similarly, from (6),

$$\frac{1}{T}\sum_{k=1}^{n}\int_0^T \frac{\partial\theta_i(x(t))}{\partial x_k}\frac{\partial\theta_j(x(t))}{\partial x_k}dt \approx M_{ij}(z). \quad (42)$$

The quality of the approximations in (41) and (42) depends on k and T. In Appendix B, it is shown that the error can be estimated as a contribution of order $O(1/\sqrt{T})$ due to finite sampling effects, and one of order $O(1/k)$ due to finiteness of k. In practice, the statistical errors dominate, which means that k does not have to be taken too big. This is fortunate since big values of k would make (34) stiff.

Notice that using the overdamped dynamics (34) is not essential, and other types of dynamics, such as Langevin, Nosé-Hoover,[25] etc., can be used as well, provided only that they lead to the equilibrium probability density function (36). For this to be the case, the key is simply to use the extended potential (35) rather than the bare one $V(x)$. For instance, in the application to the alanine dipeptide presented in Sec. V, Nosé-Hoover dynamics was used.

Finally, notice that the free energy along the string can be straightforwardly computed since

$$\frac{dF(z(\alpha))}{d\alpha} = \sum_{i=1}^{N}\frac{dz_i(\alpha)}{d\alpha}\frac{\partial F(z(\alpha))}{\partial z_i}. \quad (43)$$

Therefore

$$F(z(\alpha))-F(z(0)) = \int_0^\alpha \sum_{i=1}^{N}\frac{dz_i(\alpha')}{d\alpha'}\frac{\partial F(z(\alpha'))}{\partial z_i}d\alpha'. \quad (44)$$

C. Implementation

In practice, the evolution of the string toward the MFEP can be simulated as follows. The string is discretized using R discretization points (or images), for instance, as $z^m(t)=z(m\Delta\alpha,t)$ with $\Delta\alpha=1/(R-1)$. Suppose that the current state of these images is known at time t with the images evenly distributed along the string, i.e.,

$$|z^{m+1}(t)-z^m(t)| = |z^{m'+1}(t)-z^{m'}(t)| \quad (45)$$

for any $m,m'=1,\ldots,R-1$. To update the discretized string, we then proceed in four steps:

(1) Calculation of the mean force $\nabla_z F(z)$ and the tensor $M(z)$,

(2) evolution of the string,

(3) smoothing of the string and

(4) reparametrization of the string.

The details of each of these steps is as follows.

Step 1. To determine the mean force $\nabla_z F(z)$ and the tensor $M(z)$ at each $z^m(t)$, simulate R replicas of (34) [or any evolution equation where the potential is modified as in (35) so that it has (36) as equilibrium density] with $z=z^m(t)$, $m=2,\ldots,R-1$, and use the estimators (41) and (42).

Once $\nabla_z F(z^m(t))$ and $M(z^m(t))$ have been estimated at every $z^m(t)$, each of these images is evolved by one time step using the following multistep scheme for (32), which amounts to going through steps 2–4.

Step 2. Evolve the images using, e.g., forward Euler with time step Δt on the discretized version of (32). This leads to the following updating rule:

$$z_i^{m,*} = z_i^m(t) - \Delta t\sum_{j,k=1}^{N} P_{ij}^m M_{jk}(z^m(t))\frac{\partial F(z^m(t))}{\partial z_k}, \quad (46)$$

in which P_{ij}^m is a discrete approximation of the projector in (9):

$$P_{ij}^m = \delta_{ij} - \frac{z_i^{m+\sigma}(t)-z_i^m(t)}{|z^{m+\sigma}(t)-z^m(t)|}\frac{z_j^{m+\sigma}(t)-z_j^m(t)}{|z^{m+\sigma}(t)-z^m(t)|}, \quad (47)$$

where $\sigma=+1$ if $(z^{m+1}(t)-z^m(t))\cdot M(z^m(t))\nabla_z F(z^m(t))\geqslant 0$ and $\sigma=-1$ otherwise (using either forward or backward difference is required for numerical stability, see Refs. 6 and 26 for details).

Step 3. The main issue in the present scheme turns out to be the control of the statistical errors which impede strict convergence in time of the string toward the minimum free energy path. Indeed, since $\nabla_z F(z)$ and $M(z)$ keep fluctuating at every time step made by the string, the latter may never settle down completely. In practice, this issue is resolved by estimating the statistical error and terminating the calculation as soon as the remaining motion of the string is solely due to statistical fluctuations. It is also convenient to smooth the string to damp these statistical fluctuations, and this is what step 3 amounts to. A simple way to smooth the string is to use for $m=2,\ldots,M-1$

$$z^{m,**} = (1-s)z^{m,*} + \frac{s}{2}(z^{m-1,*}+z^{m+1,*}), \quad (48)$$

where $s\in[0,1]$ is a parameter whose amplitude controls how aggressive the smoothing is. Equation (48) is a simple convex interpolation of z^*. It is important to make sure that (48) does not impede on the accuracy at which the continu-

ous string is resolved at discrete level. Since this accuracy is $O(\Delta\alpha)$ because of the way we discretize P_{ij} in (47), this means that we must have $s=O(\Delta\alpha)$. This also implies that the statistical error must be below this threshold; in other words, consistent with the error estimate derived in Appendix B, in step 1 we must take $T=O(\Delta\alpha^{-2})$ in the estimators (41) and (42) to compute $\nabla_z F(z)$ and $M(z)$. We must also take $k=O(\Delta\alpha^{-1})$ to not introduce a bias above accuracy threshold due to the finiteness of k.

Step 4. Reparametrize the string. At discrete level, this amounts to interpolating a curve through the images $z_i^{m,**}$, then distributing new images along this interpolated curve. In the present implementation, it suffices to use piecewise linear interpolation: this is first order accurate in $\Delta\alpha$, which is also the accuracy at which we compute the projector in (47). The reparametrization step goes as follows. Denote by $L(k)$ the length of the string up to image k, i.e., $L(0)=1$ and

$$L(k) = \sum_{m=1}^{k} |z^{m,**} - z^{m-1,**}|, \quad k = 2, \ldots, R, \tag{49}$$

and for $m=2,\ldots,R$ let $s(m)=(m-1)L(R)/(R-1)$. Then, set $z^1(t+\Delta t)=z^{1,**}$, $z^R(t+\Delta t)=z^{R,**}$, and for $m=2,\ldots R-1$ take

$$z^m(t+\Delta t) = z^{k-1,**} + (s(m) - L(k-1))\frac{z^{k,**} - z^{k-1,**}}{|z^{k,**} - z^{k-1,**}|}, \tag{50}$$

where $k=2,\ldots,R$ is such that $L(k-1)<s(m)\leqslant L(k)$. It can be readily checked that the $z^m(t+\Delta t)$ are then the points along the piecewise linear interpolated path which are at equal distance along this path. Notice that the new images $z^m(t+\Delta t)$ only satisfy (45) to order $O(\Delta\alpha^2)$: this is consistent with the overall order of accuracy of our scheme, but, if needed, arbitrarily high order of accuracy on (45) can be obtained by iterating (50).

Steps 1–4 leave us with a new update of the string, $z^m(t+\Delta t)$, and these steps can be repeated until convergence to the MFEP.

More sophisticated methods to discretize, evolve, smooth, and reparametrize the string can be used to achieve higher order of accuracy (see e.g., Refs. 6, 14, and 27), but in Sec. V we will simply use the scheme discussed above. Finally, let us note that we use the string method here because it is simple, robust, and efficient, but other techniques such as the nudged elastic band could be used as well.[28]

V. EXAMPLE: ALANINE DIPEPTIDE

As an example to illustrate the procedure that we propose, we analyze the isomerization transition of alanine dipeptide molecule at 300 K in vacuum. We have studied the transition between the two metastable conformers usually named $C_{7\mathrm{eq}}$ and $C_{7\mathrm{ax}}$. This transition has been extensively studied in the literature with different methods.[12,14–20] The two conformers are usually defined as local minima in the space of the two dihedral angles ϕ and ψ. Figure 1 shows a pictorial description of the molecule and of the two conformers. All the dihedral angles here used as collective variables are also shown, and the atoms involved in their definitions

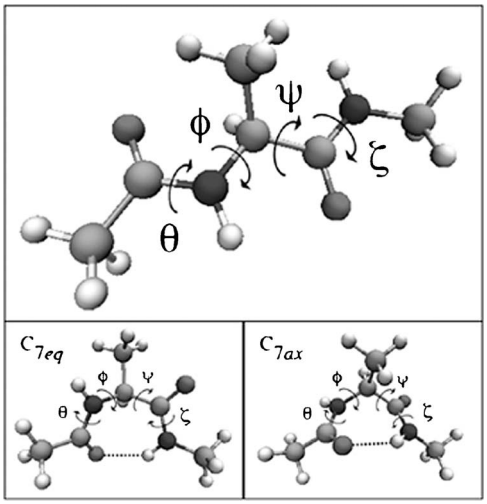

FIG. 1. Ball-and-stick representation of the alanine dipeptide molecule: $(CH_3)-(CO)-(NH)-(C_\alpha HCH_3)-(CO)-(NH)-(CH_3)$. The central carbon atom is referred to as C_α. The dihedral angles used in the calculations are showed. They are defined through the following quadruples of atoms: (O,C,N,C_α) for θ, (C,N,C_α,C) for ϕ, (N,C_α,C,N) for ψ, and (C_α,C,N,H) for ζ. In the bottom row are shown the two metastable conformers of the molecule in vacuum. The dashed line represents a hydrogen bond.

are listed in the caption. For what follows, we recall that the central carbon atom in the figure is usually called C_α. We have looked first for minimum free energy paths using the two angles ϕ and ψ as collective variables [i.e., $(z_1,z_2)=(\phi,\psi)$], then the four angles ϕ, ψ, θ, and ζ as collective variables [i.e., $(z_1,z_2,z_3,z_4)=(\phi,\psi,\theta,\zeta)$]. The angles ϕ and ψ describe the rotations around the $N-C_\alpha$ and $C_\alpha-C$ bonds, respectively, while θ and ζ describe the rotations around the $C-N$ bonds (i.e., the peptide bonds). The atoms O, C, N, and H are usually collectively referred to as peptide group. In the alanine dipeptide molecule, two such groups are present, one on the left and one on the right of the C_α atom in Fig. 1. Along simulated trajectories of the molecule, these atoms stay in average on a plane. Different torsion angles can be used to monitor the conservation of this structure.[29] Among these, we decided to work with θ and ζ because the two molecular isomers $C_{7\mathrm{eq}}$ and $C_{7\mathrm{ax}}$ are both characterized by the presence of a hydrogen bond between the O and H atoms involved in the definition of these angles (see Fig. 1).

The simulations using the two angles ϕ and ψ as collective variables were done mainly to benchmark the method, since in this case the full free energy profile in ϕ and ψ can also be computed by umbrella sampling. On the other hand, the results reported below clearly indicate that using only ϕ and ψ as collective variables is not enough to describe the mechanism of transition in alanine dipeptide in vacuum. However, using the four angles ϕ, ψ, θ, and ζ as collective variables is sufficient: in particular, it will be shown that working with these angles permits to determine the isocommittor $\frac{1}{2}$ surface, whereas working with the two angles ϕ and ψ does not.

In all simulations, we used the full-atom representation of the molecule in the CHARMM force field.[30] For the dynamics, the Nosé-Hoover integrator[25] in the CHARMM code[31] was used. All chemical bonds in the system were represented

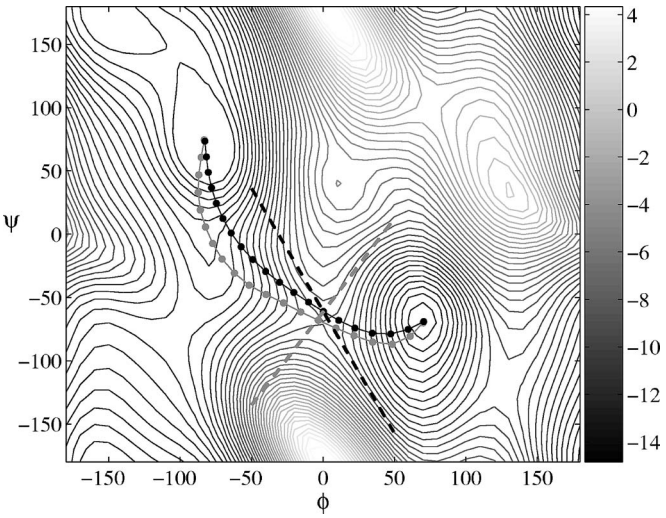

FIG. 2. Minimum free energy paths obtained by the string method using the two dihedral angles (ϕ,ψ) (gray curve) and the four angles (ϕ,ψ,θ,ζ) (black curve). The latter is the projection into the (ϕ,ψ) plane. The dots along the curves represent the images of the discretized strings. Notice that even though the MFEPs look close in this projected view, the MFEP in (ϕ,ψ,θ,ζ) is essentially four dimensional and cannot be represented using the variables ϕ and ψ alone—see Fig. 3. Notice also that the hyperplanes defined by (51) are very different in the two cases: the gray dashed line is the one associated with the MFEP in two angle (ϕ,ψ) space going through the image with maximum free energy along the MFEP (See Fig. 4), whereas the black dashed line is the intersection between the (ϕ,ψ) plane and the hyperplane associated with the MFEP in four angle (ϕ,ψ,θ,ζ) space going through the image with maximum free energy along the MFEP (see Fig. 4).

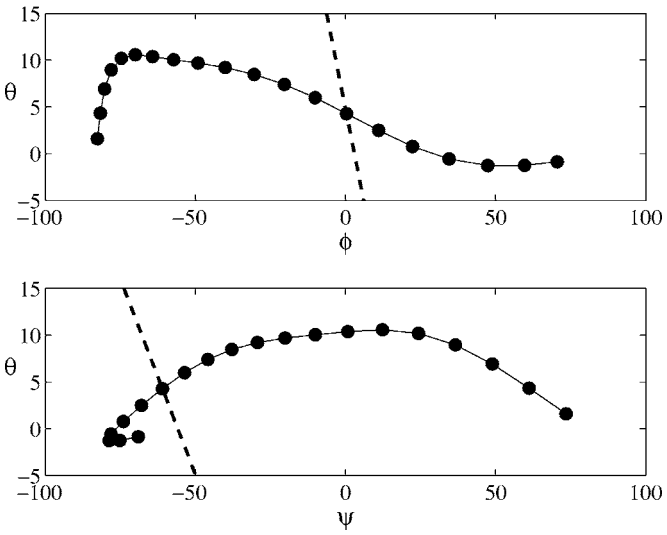

FIG. 3. Minimum free energy path obtained by the string method using the four dihedral angles (ϕ,ψ,θ,ζ) projected into the (ϕ,θ) plane (top panel) and the (ψ,θ) plane (bottom panel). The intersections of the hyperplane associated to the image with maximum free energy along the MFEP (see Fig. 4) with the planes (ϕ,θ) and (ψ,θ) are shown as dashed lines.

with springs. For the calculation of the mean force, harmonic potentials were added involving the dihedral angles, as requested from Eq. (35), with force constants $k=1000$ kcal/(mol rad^2). The CHARMM force subroutine was modified in order to obtain the quantities needed for the computation of the tensor defined in Eq. (6).

To start the computation of the MFEP with the string method, an initial condition is needed. First, we computed the end points by minimization using (33). The end points so obtained are $(\phi,\psi)_{C_{7\mathrm{eq}}}=(-83.2,74.5)$ and $(\phi,\psi)_{C_{7\mathrm{ax}}}=(70,-70)$ for the two angle simulation, and $(\phi,\psi,\theta,\zeta)_{C_{7\mathrm{eq}}}=(-82.7,73.5,1.6,-4.3)$, $(\phi,\psi,\theta,\zeta)_{C_{7\mathrm{ax}}}=(70.5,-69.1,-0.8,5.7)$ in the four angle case. We keep the names $C_{7\mathrm{eq}}$ and $C_{7\mathrm{ax}}$ for these end points, no matter the dimension of the space in which they are defined. Once the end points are identified, we take as starting condition for the string 20 images along a linear path connecting these end points, and then we follow the procedure described in Sec. IV C to update this string till convergence to the MFEP. To compute the mean force $\nabla_z F$ and the tensor M (step 1) we use the Nosé-Hoover integrator in CHARMM with a time step of 0.5 fs for a total of 500 000 steps. From the resulting trajectories, $\nabla_z F$ and M are calculated for each image using (41) and (42). The string is then updated using (46) with $\Delta t=0.02$ (step 2). For the smoothing step (step 3), we use (48) with $s=0.1$. Finally, the reparametrization step is performed as described in step 4 at every update. In order for the string to converge to the MFEP, about 100 updates were needed in two angles and about 250 in four angles. At the last update, simulations of up to 10^6 steps in the Nosé-Hoover integrator are performed to minimize the statistical error in $\nabla_z F$ and M.

Figure 2 shows the result of our MFEP calculation using (ϕ,ψ) and (ϕ,ψ,θ,ζ) as collective variables. In the latter case, the MFEP identified by the string is a curve in the four dimensional space $(\phi,\psi,\theta,\zeta)\in[0,2\pi]^4$. The projection of this MFEP in the space (ϕ,ψ) is shown in the figure. For graphical purposes, the paths are overimposed on the so-called *adiabatic energy* landscape of alanine dipeptide for the (ϕ,ψ) variables. This is defined as the surface of minimum potential energy of the system $V(x)$ at fixed values of ϕ and ψ. The two angles path goes through the major saddle point in the landscape. It also passes through a flat region located around $(\phi,\psi)=(-60,-40)$. As for the four angles path, it can be seen that its projection on the (ϕ,ψ) plane differs from the MFEP obtained by working with the variables (ϕ,ψ) only. This indicates that the MFEP in (ϕ,ψ,θ,ζ) is curved in the θ and ζ directions. In order to visualize the variation in θ space, Fig. 3 shows the projections of the four dimensional MFEP in the space of (ϕ,θ) and (ψ,θ) variables. Also shown in Figs. 2 and 3 are the hyperplanes $P(\alpha)$ (see Sec. III D) associated with the images with maximum free energy along the MFEPs. These planes are defined by the equation

$$\sum_{j=1}^{N}\hat{n}_j(\alpha)(z_j-z_j(\alpha))=0, \tag{51}$$

meaning that they are lines when two angles are used and $(z_1,z_2)=(\phi,\psi)$, and three dimensional hyperplanes when four angles are used and $(z_1,z_2,z_3,z_4)=(\phi,\psi,\theta,\zeta)$ [in the latter case, what is shown in the figures is the intersection of the hyperplane with the planes ϕ-ψ (Fig. 2) or ϕ-θ and ϕ-ζ (Fig. 3)]. Figure 2 shows that the plane associated with the MFEP in two angles is very different from the one associated with the MFEP in four angles, and this already indicates that (ϕ,ψ) is not a set of collective variables large enough to describe the mechanism of the reaction [of course, this does

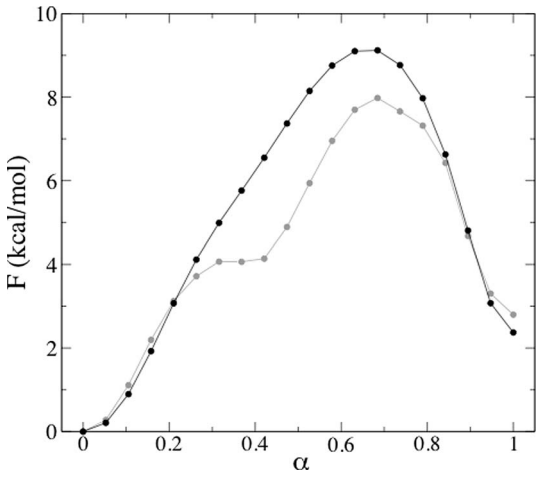

FIG. 4. Free energy profiles along the MFEPs in the variables (ϕ,ψ) (gray curve) and (ϕ,ψ,θ,ζ) (black curve). The profile of the free energy along the MFEP in (ϕ,ψ) indicates that this MFEP goes by a flat region of the free energy, consistent with what is seen in Fig. 2. However, there is no such region along the MFEP in (ϕ,ψ,θ,ζ). The free energy along the MFEP in (ϕ,ψ,θ,ζ) is more similar to the one obtained in Ref. 14 using the string method in the full Cartesian space, which is already an indication that the collective variables ϕ, ψ, θ, and ζ are a set large enough to describe the transition between $C_{7\mathrm{eq}}$ and $C_{7\mathrm{ax}}$.

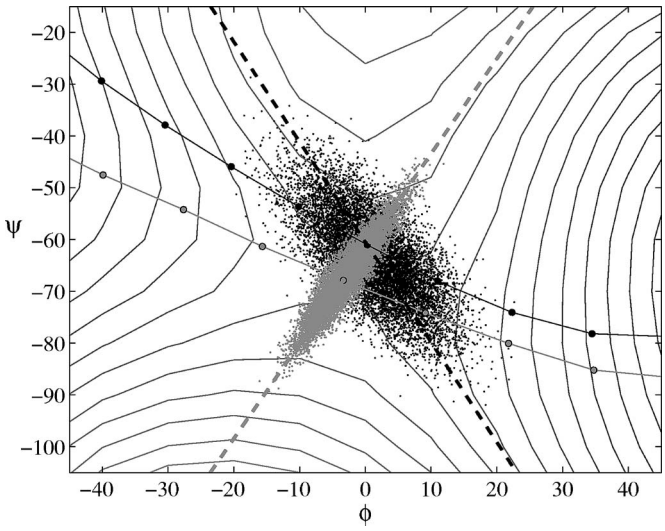

FIG. 5. The projections on the (ϕ,ψ) plane of points sampled in the transition ensemble [that is, the ensemble of points on the isocommittor $\frac{1}{2}$ surface in state space (x,v) equipped with the Boltzmann-Gibbs probability density function restricted to this surface]. The gray cloud is slightly scattered around the gray dashed line due to the sampling procedure used; the black cloud is more scattered because the isocommittor $\frac{1}{2}$ surface associated with the MFEP in (ϕ,ψ,θ,ζ) bends in the direction of θ and ζ. This reveals the importance of these variables in describing the mechanism of the transition.

not says that (ϕ,ψ,θ,ζ) is a large enough set, though this is indeed the case as will be verified below].

The free energy profile along the MFEP can be calculated from (44). Figure 4 shows the results obtained for the MFEPs in (ϕ,ψ) and (ϕ,ψ,θ,ζ). It can be seen that the free energy along the MFEP in (ϕ,ψ,θ,ζ) is different from the one along the MFEP in (ϕ,ψ). In the two angles case, the profile shows the flat region through which the string goes (see Fig. 2). Free energy differences calculated between points along the two angles profile are in very good agreement with results previously obtained with different methods, all of which used ϕ and ψ only as collective variables.[15–19] Moreover, it is interesting to note that the free energy profile along the MFEP in (ϕ,ψ,θ,ζ) is in fact closer to the one calculated in Ref. 14 with the string method using all the variables in Cartesian space.

To further prove the importance of the θ and ζ variables, we generated configurations from the transition ensemble. By definition, the transition ensemble is the ensemble of points on the isocommittor $\frac{1}{2}$ surface in state space (x,v) equipped with the Boltzmann-Gibbs probability density function restricted to this surface. From the results in Sec. III D, we know that if one assumes that the collective variables used are a good set, then the isocommittor $\frac{1}{2}$ surface is locally approximated by the lift up in state space (x,v) of the plane defined by (51) associated with the point of maximum free energy along the MFEP. Denoting by α_s the point along the MFEP where $F(z(\alpha))$ is maximum, i.e., $F(z(\alpha_s))=\max_{\alpha\in[0,1]}F(z(\alpha))$, the equation for this hypersurface is

$$\sum_{j=1}^{N}\hat{n}_j(\alpha_s)(\theta_j(x)-z_j(\alpha_s))=0. \tag{52}$$

In order to sample points on the hypersurface defined by (52) distributed according to the Boltzmann-Gibbs probability density function restricted to this surface, we added the following expression to the potential used in the CHARMM code:

$$V_{\alpha,k_p}(x)=\frac{k_p}{2}\left(\sum_{j=1}^{N}\hat{n}_j(\alpha)(\theta_j(x)-z_j(\alpha))\right)^2, \tag{53}$$

and used the Nosé-Hoover integrator. The constant k_p was chosen to be 1000 kcal/(mol rad^2). Figure 5 shows an enlarged view of the major saddle point region of Fig. 2, together with the projections on the (ϕ,ψ) plane of the points in the transition ensemble. The gray cloud are the points obtained by working in the two angles (ϕ,ψ) only. [In this case, in this projected view, the transition ensemble should lie on the gray line—the latter corresponds to (51) with $(z_1,z_2)=(\phi,\psi)$ and $\alpha=\alpha_s$; the finite width of this cloud away from the line is due to the finiteness of k_p in (53)]. The black cloud contains the points from the transition ensemble obtained by working in the four angles (ϕ,ψ,θ,ζ). The fact that the black cloud covers a more extended region than the gray cloud reveals that the transition ensemble obtained by working with (ϕ,ψ,θ,ζ) is actually bended in the θ and ζ variables, which is a clear indication that these two variables are important for the transition under study.

As a conclusive and definitive test on the role of the full set of collective variables, we computed committor values distributions of the transition ensembles on the hypersurface (52), since this surface is supposed to be a local approximation of the isocommittor $\frac{1}{2}$ surface. We first calculated the committor distribution associated with the surface (52) when only the angles ϕ and ψ were used. One hundred different configurations were extracted from the transition ensemble on this surface, and their committor value was computed from 200 trajectories generated from each of these configu-

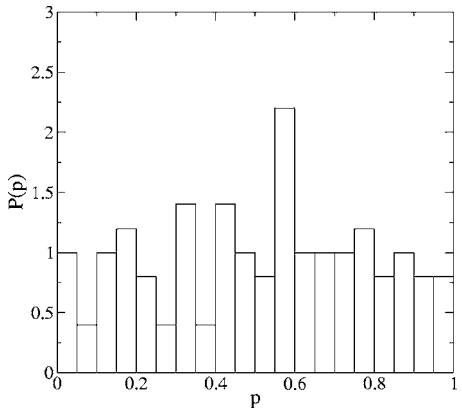

FIG. 6. Committor distribution of the transition ensemble in the hypersurface in the two angles MFEP with putative committor value $\frac{1}{2}$. The flatness of this distribution indicates that this surface is not a good approximation of the isocommittor $\frac{1}{2}$ surface, i.e., (ϕ,ψ) are not good collective variables to describe the reaction.

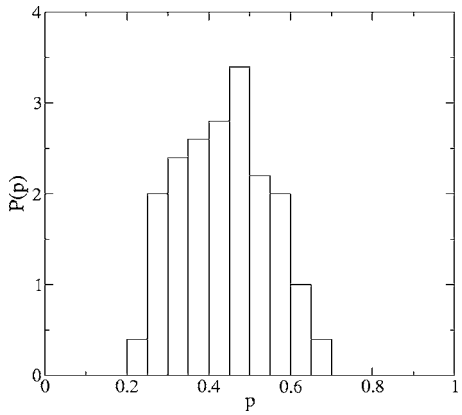

FIG. 8. Committor distribution for the interpolated hypersurface with committor value $\frac{1}{2}$ along the MFEP in four angles. The fact that this distribution is peaked at $\frac{1}{2}$ indicates that the set (ϕ,ψ,θ,ζ) is a good set of collective variables to describe the reaction.

rations by assigning initial random velocities. The resulting distribution is reported in Fig. 6. The distribution is almost uniform, indicating that the corresponding surface is not a good approximation of the isocommittor $\frac{1}{2}$ surface, i.e., ϕ and ψ alone are not a good set of collective variables.

The situation improves markedly if one uses the four dihedral angles, ϕ, ψ, θ, and ζ. Figure 7 shows the committor distributions calculated for the hypersurfaces (52) corresponding to images number 13 and 14 along the string, corresponding to $\alpha=0.63$ and $\alpha=0.68$, respectively. One hundred points were extracted from the ensemble restricted on these hypersurfaces, and 200 trajectories were launched from each of these points. The resulting distributions are peaked, which is an indication that the set of collective variables properly describes the reaction since it shows that the hypersurfaces (52) with $\alpha=0.63$ and $\alpha=0.68$ do indeed approximate the isocommittor surfaces. Figure 8 shows the distribution obtained for the transition ensemble on the hypersurface labeled by α_s, with $\alpha_s=0.66$, corresponding to committor value $\frac{1}{2}$, obtained by linear interpolation between those at $\alpha=0.63$ and $\alpha=0.68$. The distribution is peaked at $\frac{1}{2}$, indicating that we have chosen the correct surface.

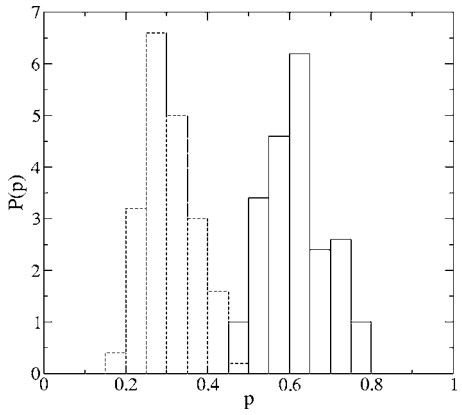

FIG. 7. Committor distributions for the hypersurfaces at the images along the string with $\alpha=0.63$ (dashed line) and $\alpha=0.68$ (solid line), which are the closest to the maximum of the free energy along the MFEP in the four angle MFEP.

The results above—namely, that the set (ϕ,ψ) is not sufficient to describe the reaction but the enlarged set (ϕ,ψ,θ,ζ) is—are consistent with those in Refs. 12 and 20. Two points are worth noting, however. First, the results in Refs. 12 and 20 were obtained by identifying a set of reactive trajectories first, then analyzing these trajectories to identify the transition state regions. This second step is quite a tedious one in practice, and by our technique we avoid it completely since we identify the isocommittor surfaces directly (i.e., without running reactive trajectories beforehand). Even though the validity of these surfaces (i.e., the validity of the collective variables chosen) must then be assessed *a posteriori*, the method that we propose is still substantially cheaper than the one in Refs. 12 and 20. The second point worth noting is a difference with Refs. 12 and 20 in terms of the protocol used to compute the committor distributions. In Ref. 20, these were calculated for ensembles constrained in a point of the collective variables space, while here we only restrict the system to be on the hypersurface defined in (52). This leads to a more severe test than the one used in Ref. 20, but also a more appropriate one, since the isocommittor surface is a dividing surface in the original state space which therefore cannot be reduced to a point in the collective variables space (it should be a surface of dimension $N-1$ in this space if the number of collective variables is N).

VI. CONCLUDING REMARKS

Let us summarize the highlights of this paper. We have shown that the minimum free energy path (MFEP) in a given set of collective variables is relevant to describe a reaction: the MFEP is the reaction pathway of maximum likelihood in the space of these collective variables. We have also shown how to compute the MFEP by combining the string method with a restrained sampling technique to calculate the mean force and the metric tensor [see the definition of the MFEPs in (7)].

The main advantage of this computational approach is that the update of the string involves a local calculation along the path. As a result, the cost of the technique presented here scales linearly with the number of images along

the discretized string, and it is independent of the number of collective variables used. In addition, since the computation of the estimators at one image is independent from those at the others, the method is straightforwardly parallel in the number of images: one update of the whole string is obtained at the cost of one computation of the mean force. As a result, we can compute the MFEP in the space of a large number of collective variables. This is the key feature that differentiates our technique from other methods[1–3] which aim at mapping the whole free energy landscape (and hence become prohibitively expensive when the number of collective variables is large). It also makes our technique promising in tackling problems where intuition might be not sufficient to identify *a priori* the mechanism of the reaction via a small number of collective variables.

Regarding this last point, it should be stressed that the relevance of the MFEP depends on the collective variables that are used to describe the reaction. The approach described here does not tell how to choose these collective variables. Clearly, this may be a difficult step for which some *a priori* insight about the reaction may be needed. However, the choice is made simpler by the fact that many collective variables can be used simultaneously since the computation cost scales well with their number. In addition, the choice of collective variables can be tested *a posteriori*. If the variables chosen are insufficient, the committor distribution of the isocommittor $\frac{1}{2}$ surface will not be peaked at $\frac{1}{2}$. In this case, it will be necessary to add some additional collective variables to describe the reaction. If, on the contrary, they are redundant, the committor distribution will be peaked at $\frac{1}{2}$, but the normal vector to the isocommittor $\frac{1}{2}$ surface will only have nonzero components in a subset of the collective variables. In this case, these variables in which the components of the normal vector are zero are irrelevant for the reaction (since the reaction pathway is perpendicular to them) and can be safely removed from the set of collective variables.

It should also be stressed that, even if the right collective variables are used, there may be more than one reaction channel. (There may be also intermediate metastable states along these channels but this is not a problem since the string is parametrized by arclength, not time, and hence it can go through these intermediate states without slowing down.) If there are multiple reaction channels, in order to identify the different MFEPs associated with them it will be necessary to vary the initial condition for the string, a step which may again require some *a priori* insight. In these situations, the relative probability of the different channels can be deduced from the free energy barrier along the associated MFEPs (the MFEP with the lowest free energy barrier being the most probable one).

Proceeding this way, we can envision to use the method proposed here to systematically analyze the mechanism of complicated reactions and arrive at a description of this mechanism in terms of a few, not necessarily intuitive collective variables.

ACKNOWLEDGMENTS

We thank Weinan E and Weiqing Ren for many useful discussions and comments. This work is partially supported by NSF Grant Nos. DMS02-09959 and DMS02-39625, and by ONR Grant No. N00014-04-1-0565. One of the authors (L.M.) is grateful to COST for partial support under its Action P13. This work was performed while one of the authors (E.V.-E.) was visiting UC Berkeley on a Visiting Professorship position sponsored by the Miller Institute.

APPENDIX A: PATH INTEGRALS AND MINIMUM FREE ENERGY PATH

Introducing the following shorthand notation for (17):

$$\dot{z} = -M\nabla_z F + \beta^{-1}\,\mathrm{div}\,M + \sqrt{2\beta^{-1}}M^{1/2}\eta, \tag{A1}$$

the transition probability density $p(z_1,\tau|z_0)$ that the solution be at z_1 at time τ conditional on it being at z_0 at initial time $\tau=0$ can be formally expressed as the path integral

$$p(z_1,\tau|z_0) = C^{-1}\int_{\text{paths}} \exp(-\beta S_\tau^\beta) \prod_{0<\tau'<\tau} dz(\tau'). \tag{A2}$$

Here C is a normalization constant,

$$S_\tau^\beta = \int_0^\tau \|\dot{z} + M\nabla_z F - \beta^{-1}\,\mathrm{div}\,M\|_M^2 d\tau', \tag{A3}$$

where for any $a=(a_1,\ldots,a_N)$, $\|a\|_M^2=\Sigma_{i,j=1}^N a_i M_{ij}^{-1} a_j$ and the path integral in (A2) is supposed to be taken over all path $z(\tau')$, $0\leqslant\tau'\leqslant\tau$, such that $z(0)=z_0$ and $z(\tau)=z_1$. Equation (A2) is formal because the constant C is actually infinite and $\Pi_{0<\tau'<\tau}dz(\tau')$ is ill defined. The correct way to give a precise mathematical meaning to (A2) is via Girsanov formula in probability theory.[22,32] We shall not dwell on this issue here but simply note that another way to make (A3) practical is to discretize in time the process $z(\tau)$ in (A2) and (A4) [a very nice derivation of the rigorous equivalent of (A2) along these lines is given in Sec. IV of Kac's book[33]]. Once the paths $z(\tau)$ are discretized, e.g., as $z(m\Delta\tau)=z^m$, $m=1\ldots,R$ with $\Delta\tau=\tau/(R-1)$ and $z^1=z_0$, $z^R=z_1$, (A2) becomes a regular $R-2$ dimensional integral over $z_2,\ldots,z_{R-1}$. This last approach is what is done, e.g., in transition path sampling,[4] where (A2) is used to design a Metropolis Monte-Carlo scheme to sample the solutions of (17).

Equation (A2) is also a useful tool to formally derive rigorous asymptotic expansion. Suppose, in particular, that the temperature is small compared with the main free energy barriers in the system. Then, much the same way as what happens in finite dimension when one evaluates an integral by Laplace method, to leading order in β^{-1}, the path integral will be dominated by the trajectory which minimizes

$$S_\tau^\infty = \int_0^\tau \|\dot{z} + M\nabla_z F\|_M^2 d\tau', \tag{A4}$$

subject to the appropriate boundary conditions $z(0)=z_0$ and $z(\tau)=z_1$. Equation (A4) is the action functional in the Freidlin-Wentzell theory of large deviations: the minimizer of (A4)—the minimum action path—subject to $z(0)=z_0$ and $z(\tau)=z_1$ is the path of maximum likelihood by which the solution of (17) goes from z_0 to z_1 in time τ. Next, we show the following: if τ in (A3) is taken large enough (i.e., if one does not constraint the time over which the reaction occurs

and let it happen on its natural time scale), then the graph of the minimum action path connecting two minima of the free energy is the MFEP.

To see this, let z_a and z_b be the location of two minima of the free energy, and let z_s be the location of the saddle point in between [that is, z_s is the location of the minimum of the free energy on the separatrix which separates the basins of attraction of z_a and z_b by the deterministic dynamics in (18)]. We wish to determine the trajectory which minimizes the integral (A3) subject to $z(0)=z_a$ and $z(t)=z_b$. To identify this path, let us assume that when the path crosses the separatrix between z_a and z_b at time $t_s \in (0,t)$, it does so by z_s (this assumption will be verified afterwards). Then we can split the integral in (A4) as

$$\int_0^\tau \|\dot{z}+M\nabla_z F\|_M^2 d\tau' = \int_0^{\tau_s} \|\dot{z}+M\nabla_z F\|_M^2 d\tau' + \int_{\tau_s}^{\tau} \|\dot{z}+M\nabla_z F\|_M^2 d\tau'. \tag{A5}$$

Since $z(\tau_s)=z_s$ by assumption, the second integral from τ_s to τ can be made arbitrarily small by letting $\tau-\tau_s\to\infty$ and choosing for $z(\tau)$ the solution of (18) which connects z_s to z_b in doubly infinite time. As for the first integral from 0 to τ_s, let us write it as

$$\begin{aligned}
&\int_0^{\tau_s} \|\dot{z}+M\nabla_z F\|_M^2 d\tau' \\
&= \int_0^{\tau_s} \|\dot{z}-M\nabla_z F+2M\nabla_z F\|_M^2 d\tau' \\
&= \int_0^{\tau_s} \|\dot{z}-M\nabla_z F\|_M^2 d\tau' + 4\int_0^{\tau_s} \dot{z}\cdot\nabla_z F d\tau' \\
&= \int_0^{\tau_s} \|\dot{z}-M\nabla_z F\|_M^2 d\tau' + 4(F(z_s)-F(z_a)) \\
&= \int_{-\tau_s}^{0} \|\dot{\bar{z}}+M\nabla_z F\|_M^2 d\tau' + 4(F(z_s)-F(z_a)),
\end{aligned} \tag{A6}$$

where in the third step we simply expanded the square, and in the last step we inverted time and defined $\bar{z}(\tau)=z(-\tau)$. The boundary term $4(F(z_s)-F(z_a))$ is the smallest possible since to go from z_a to z_b, the trajectory must cross the separatrix between these minima, and z_s is the point where the free energy is minimum on this separatrix. On the other hand, the integral term can be made arbitrarily small by letting $\tau_s\to\infty$ and choosing for $\bar{z}(\tau)$ the solution of (18) which connects z_s to z_a in doubly infinite time. Putting everything together, we conclude that the minimum action path between z_a and z_b is the patch of the two heteroclinic orbits of (18) which join z_s to z_a and z_s to z_b. In other words, the graph of the minimum action path is the MFEP, as claimed, which also means that the reactive trajectories by which transitions between z_a and z_b occur are most likely to follow the MFEP.

APPENDIX B: SOME ERROR ESTIMATES

Consider the following expectation:

$$A(z) = Z^{-1}e^{\beta F(z)}\int_{\mathbb{R}^n} a(x)e^{-\beta V(x)} \times \delta(z_1-\theta_1(x))\cdots\delta(z_N-\theta_N(x))dx, \tag{B1}$$

where $a(x)$ is some function defined on Cartesian space $\mathbb{R}^n$. Both $\nabla_z F$ and $M(z)$ are expectations of this type, and we wish to estimate the error introduced by using the following estimator for $A(z)$:

$$A_{T,k}(z) = \frac{1}{T}\int_0^T a(x(t))dt, \tag{B2}$$

where $x(t)$ is a trajectory satisfying (34). As $T\to\infty$, we know from the law of large numbers that

$$A_{T,k}(z) \to A_k(z) \equiv \int_{\mathbb{R}^n} a(x)\rho_{k,z}(x)dx, \tag{B3}$$

where $\rho_{k,z}(x)$ is the equilibrium probability density function (36). By the central limit theorem, we also know that, as $T\to\infty$,

$$\sqrt{T}(A_{T,k}(z)-A_k(z)) \to \sqrt{\tau_k(z)}\xi_k(z), \tag{B4}$$

where $\xi_k(z)$ is a Gaussian variable with mean zero and variance

$$\text{var}\,\xi_k(z) = \int_{\mathbb{R}^n} (a(x)-A_k(z))^2\rho_{k,z}(x)dx, \tag{B5}$$

and $\tau_k(z)$ is the following correlation time:

$$\tau_k(z) = (\text{var}\,\xi_k(z))^{-1}\int_0^\infty\int_{\mathbb{R}^n} \mathbf{E}(a(x(t))-A_k(z)) \times (a(x)-A_k(z))\rho_{k,z}(x)dxdt. \tag{B6}$$

Here $x(t)$ is the solution of (34) starting at $x(t=0)=x$, and $\mathbf{E}$ denotes expectation with respect to the white noise in (34). Equation (B4) means that

$$A_{T,k}(z) = A_k(z) + \sqrt{\frac{\tau_k(z)}{T}}\xi_k(z) + \cdots, \tag{B7}$$

where $(\cdots)$ stands for higher order terms in $1/T$. This gives half of the error estimate. The other half involves the error due to the finiteness of k. To estimate this part, given any function $a(x)$ consider the functions $\langle a\rangle_k(z)$ and $\langle a\rangle(z)$ defined as

$$\langle a\rangle_k(z) = \int_{\mathbb{R}^n} a(x)e^{-\beta U_{k,z}(x)}dx \tag{B8}$$

and

$$\langle a\rangle(z) = \int_{\mathbb{R}^n} a(x)e^{-\beta V(x)} \times \delta(z_1-\theta_1(x))\cdots\delta(z_N-\theta_N(x))dx. \tag{B9}$$

Notice the absence of the normalization factors $Z_{k,z}$ and $Ze^{-\beta F(z)}$, respectively, in these expressions. In fact,

$$A_k(z) = \frac{\langle a\rangle_k(z)}{\langle 1\rangle_k(z)},$$

$$A(z) = \frac{\langle a\rangle(z)}{\langle 1\rangle(z)} \equiv Z^{-1}e^{\beta F(z)}\langle a\rangle(z). \tag{B10}$$

Now, let

$$\langle \hat{a}\rangle_k(\eta) = \int_{\mathbb{R}^N} e^{i\eta\cdot z}\langle a\rangle_k(z)dz \tag{B11}$$

be the Fourier transform of (B8) (here $\eta\cdot z=\Sigma_{j=1}^N \eta_j z_j$), and define similarly $\langle \hat{a}\rangle(\eta)$, the Fourier transform of (B9). Using

$$\begin{aligned}
&\int_{\mathbb{R}^N} e^{i\eta\cdot z-\beta U_{k,z}(x)}dz \\
&\quad = \left(\frac{2\pi}{\beta k}\right)^{N/2}\exp\left(i\eta\cdot\theta(x)-\beta V(x)-\frac{|\eta|^2}{2\beta k}\right) \\
&\quad = \left(\frac{2\pi}{\beta k}\right)^{N/2}\exp(i\eta\cdot\theta(x)-\beta V(x)) \\
&\qquad \times\left(1-\frac{|\eta|^2}{2\beta k}+O(k^{-2})\right)
\end{aligned} \tag{B12}$$

where $|\eta|^2=\Sigma_{j=1}^N \eta_j^2$, we deduce that

$$\begin{aligned}
\langle \hat{a}\rangle_k(\eta) &= \left(\frac{2\pi}{\beta k}\right)^{N/2}\int_{\mathbb{R}^n} g(x) \\
&\qquad \times\exp\left(i\eta\cdot\theta(x)-\beta V(x)-\frac{|\eta|^2}{2\beta k}\right)dx \\
&= \left(\frac{2\pi}{\beta k}\right)^{N/2}\left(1-\frac{|\eta|^2}{2\beta k}+O(k^{-2})\right) \\
&\qquad \times\int_{\mathbb{R}^n} g(x)\exp(i\eta\cdot\theta(x)-\beta V(x))dx \\
&= \left(\frac{2\pi}{\beta k}\right)^{N/2}\left(1-\frac{|\eta|^2}{2\beta k}+O(k^{-2})\right)\langle \hat{a}\rangle(\eta).
\end{aligned} \tag{B13}$$

Going back to z space, (B13) implies that $\langle a\rangle_k(z)$ and $\langle a\rangle(z)$ are related as

$$\langle a\rangle_k(z) = \left(\frac{2\pi}{\beta k}\right)^{N/2}\left(\langle a\rangle(z)+\frac{1}{2\beta k}\sum_{j=1}^N \frac{\partial^2\langle a\rangle(z)}{\partial z_j^2}+O(k^{-2})\right). \tag{B14}$$

Inserting (B10) into this expression and Taylor series expanding in $1/k$ then gives

$$\begin{aligned}
A_k(z) &= A(z)-\frac{1}{2\beta k}\sum_{j=1}^N \frac{\partial^2 A(z)}{\partial z_j^2}+\frac{1}{\beta}\sum_{j=1}^N \frac{\partial A(z)}{\partial z_j}\frac{\partial F(z)}{\partial z_j} \\
&\quad +O(k^{-2}).
\end{aligned} \tag{B15}$$

By a similar argument, we can show that

$$\operatorname{var}\xi_k(z) = \operatorname{var}\xi(z)+O(k^{-1}), \tag{B16}$$

where

$$\begin{aligned}
\operatorname{var}\xi(z) &= Z^{-1}e^{\beta F(z)}\int_{\mathbb{R}^n}(a(x)-A(z))^2 e^{-\beta V(x)} \\
&\qquad \times\delta(z_1-\theta_1(x))\cdots\delta(z_N-\theta_N(x))dx.
\end{aligned} \tag{B17}$$

This means that the Gaussian variable $\xi_k(z)$ can be represented as

$$\xi_k(z) = \xi(z)+O(k^{-1/2}), \tag{B18}$$

where $\xi(z)$ is a Gaussian variable with mean zero and variance (B17). Similarly, it can be shown that $\tau_k(z)=\tau(z)+O(k^{-1})$ where $\tau(z)$ is the correlation time of the process defined by (34) in the limit as $k\to\infty$. As discussed in Ref. 34, this limiting process satisfies

$$\begin{aligned}
\dot{x}_i(t) &= -\sum_{j=1}^n \hat{P}_{ij}(x(t))\frac{\partial V(x(t))}{\partial x_j}+\sqrt{2k_BT}\sum_{j=1}^n \hat{P}_{ij}(x(t))\,\eta_j(t) \\
&\quad +k_BT v_i(x(t)).
\end{aligned} \tag{B19}$$

Here $\hat{P}_{ij}(x)$ is the orthogonal projector into $\theta(x(t))=z=cst$, which can be expressed, e.g., as

$$\hat{P}_{ij} = \delta_{ij}-\sum_{k,l=1}^N G_{kl}^{-1}(x)\frac{\partial\theta_k(x)}{\partial x_i}\frac{\partial\theta_l(x)}{\partial x_j} \tag{B20}$$

where

$$G_{kl}(x) = \sum_{i=1}^n \frac{\partial\theta_k(x)}{\partial x_i}\frac{\partial\theta_l(x)}{\partial x_i} \tag{B21}$$

and $v_i(x)=\Sigma_{j=1}^n \partial\hat{P}_{ij}/\partial x_j$. [The equilibrium probability density function of (B19) is

$$Z^{-1}e^{\beta F(z)}e^{-\beta V(x)}\delta(z_1-\theta_1(x))\cdots\delta(z_N-\theta_N(x)), \tag{B22}$$

and therefore this equation could be used to eliminate the error in k altogether. This possibility will be exploited elsewhere.]

Inserting (B15) and (B18) into (B7), we finally arrive at

$$\begin{aligned}
A_{T,k}(z) &= A(z)-\frac{1}{2\beta k}\sum_{j=1}^N \frac{\partial^2 A(z)}{\partial z_j^2}+\frac{1}{\beta}\sum_{j=1}^N \frac{\partial A(z)}{\partial z_j}\frac{\partial F(z)}{\partial z_j} \\
&\quad +\sqrt{\frac{\tau(z)}{T}}\xi(z)+\cdots,
\end{aligned} \tag{B23}$$

where $(\cdots)$ stands for higher order terms. Equation (B23) shows that the leading order terms in the error using the estimator (B2) for $A(z)$ are of orders $1/\sqrt{T}$ and $1/k$, as claimed in Sec. IV B.

[1] E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).
[2] L. Rosso, P. Minárý, S. Shou, and M. E. Tuckerman, J. Chem. Phys. **116**, 4389 (2001).
[3] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002).
[4] P. G. Bolhuis, C. Dellago, P. L. Geissler, and D. Chandler, Annu. Rev. Phys. Chem. **54**, 20 (2002).
[5] R. Elber, A. Ghosh, and A. Cardenas, Acc. Chem. Res. **35**, 396 (2002).
[6] W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 52301 (2002).
[7] W. E and E. Vanden-Eijnden, J. Stat. Phys. (in press).
[8] W. E, W. Ren, and E. Vanden-Eijnden, Chem. Phys. Lett. **413**, 242 (2005).
[9] W. E and E. Vanden-Eijnden, in *Multiscale Modelling and Simulation*,

edited by S. Attinger and P. Koumoutsakos (Springer-Verlag, Berlin, 2004).
[10] E. Vanden-Eijnden, in *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer-Verlag, Berlin, in press).
[11] G. Hummer, J. Chem. Phys. **120**, 516 (2004).
[12] A. Ma and A. R. Dinner, J. Phys. Chem. B **109**, 6769 (2005).
[13] E. A. Carter, G. Ciccotti, J. T. Hynes, and R. Kapral, Chem. Phys. Lett. **156**, 472 (1989).
[14] W. Ren, E. Vanden-Eijnden, P. Maragakis, and W. E, J. Chem. Phys. **123**, 134109 (2005).
[15] B. M. Pettitt and M. Karplus, Chem. Phys. Lett. **121**, 194 (1985).
[16] R. Czerminski and R. Elber, J. Chem. Phys. **92**, 5580 (1990).
[17] T. Lazaridis, D. J. Tobias, C. L. Brooks III, and M. E. Paulaitis, J. Chem. Phys. **95**, 7612 (1991).
[18] D. J. Tobias and C. L. Brooks III, J. Phys. Chem. **96**, 3864 (1992).
[19] C. L. Brooks III and D. A. Case, Chem. Rev. (Washington, D.C.) **93**, 2487 (1993).
[20] P. G. Bolhuis, C. Dellago, and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **97**, 5877 (2000).
[21] G. Hummer and I. G. Kevrekidis, J. Chem. Phys. **118**, 10762 (2003).
[22] R. Durrett, *Stochastic Calculus* (CRC, Washington, DC, 1996).
[23] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, New York/Verlag, Berlin, 1984).
[24] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer, New York, 1997).
[25] S. Nosé, Mol. Phys. **57**, 187 (1986).
[26] G. Henkelman and H. Jonsson, J. Chem. Phys. **113**, 9978 (2000).
[27] W. Ren, Commun. Math. Sci. **1**, 5877 (2003).
[28] H. Jonsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker, (World Scientific, Singapore, 1998).
[29] M. Buck and M. Karplus, J. Am. Chem. Soc. **121**, 9645 (1999).
[30] A. S. MacKerell, Jr., D. Bashford, M. Bellott *et al.*, J. Phys. Chem. B **102**, 3586 (1998).
[31] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[32] B. Oksendal, *Stochastic Differential Equations* (Springer, New York/Verlag, Berlin, 1984).
[33] M. Kac, *Probability and Related Topics in Physical Sciences. Lectures in Applied Mathematics* (Interscience, New York, 1959).
[34] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden, ChemPhysChem **6**, 1809 (2005).