Extracting the diffusion tensor from molecular dynamics simulation with Milestoning

Mauro L. Mugnai¹ and Ron Elber^{1,2}

¹*Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, USA*

²*The Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas 78712, USA*

(Received 30 May 2014; accepted 9 December 2014; published online 7 January 2015)

We propose an algorithm to extract the diffusion tensor from Molecular Dynamics simulations with Milestoning. A Kramers-Moyal expansion of a discrete master equation, which is the Markovian limit of the Milestoning theory, determines the diffusion tensor. To test the algorithm, we analyze overdamped Langevin trajectories and recover a multidimensional Fokker-Planck equation. The recovery process determines the flux through a mesh and estimates local kinetic parameters. Rate coefficients are converted to the derivatives of the potential of mean force and to coordinate dependent diffusion tensor. We illustrate the computation on simple models and on an atomically detailed system—the diffusion along the backbone torsions of a solvated alanine dipeptide. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4904882>]

I. INTRODUCTION

Many experimental techniques follow the time evolution of a few microscopic quantities to investigate kinetic and thermodynamic properties of biopolymers. Techniques such as FRET (Förster Resonance Energy Transfer)¹ or force spectroscopy^{2,3} can be used to monitor temporal changes of the end-to-end distance of protein or nucleic acid chains. The information extracted by those techniques can be used to study the free energy and kinetics of folding¹⁻⁴ and the dynamics of the protein in the unfolded state.^{5,6}

It is common to describe these complex processes with a reduced set of coordinates called coarse variables (CV). This subset is frequently accessible to experimental measurements and can be computed from atomically detailed simulations. The end-to-end distance is an example of a single coarse variable. It is likely that a single coarse variable is insufficient to describe the process at desired accuracy, and models of higher dimensions are needed. Molecular Dynamics (MD) simulations can provide the parameters and the effective energy landscape for the CV. To model and compute the dynamics in the space of the CV an appropriate theory is required. Theories of this type of dynamics assume stochastic behavior of the CV due to interactions with the remaining degrees of freedom (the “other variables,” or OV). The OV induce diffusion behavior of the CV, which in the present paper is modeled by a Fokker-Planck Equation (FPE).

In atomically detailed simulations, on which our theory of Milestoning is based, all the degrees of freedom are recorded: those that are neglected in a reduced description (OV), and those that are modeled probabilistically in FPE (CV). Therefore, atomically detailed simulations are more informative than the reduced model. However, the reduction is useful for comparison to phenomenological theories and experimental interpretations that follow the FPE or related approaches.

Ideally, theoretical models of the dynamics in coarse space employ an exact projection of the entire dynamics onto a few

degrees of freedom.^{7,8} Exact projection theories are the Generalized Langevin Equation (GLE),⁷ which describes the dynamics of the CV, and the Generalized Master Equation (GME),⁹ which describes the temporal evolution of the probability density in coarse space. To describe the dynamics of the system with GLE, we need to know the friction kernel. Computing the friction kernel from atomistic dynamics is a hard task. Indeed, modeling the friction kernel requires the solution of temporal evolution of variables according to the so-called orthogonal dynamics. An approximate expression for the kernel can be obtained with simplifying assumptions; for example,

- (i) The OV are assumed to be harmonic oscillators.¹⁰
- (ii) The orthogonal dynamics is fast compared to the dynamics of the CV. The friction kernel is obtained from the force-force autocorrelation function.¹¹ In other words, the friction kernel decays faster than the dynamics of the CV. Then a static friction¹² which is a time-independent, space-dependent tensor is computed in the Markov approximation.¹¹
- (iii) The velocity autocorrelation function with a harmonic restraint¹³ is used to estimate the friction kernel at a specific location.

Recently, another solution was found¹⁴ for a linear generalized Langevin equation.¹⁰ Two algorithms compute the friction kernel once the entire MD trajectory is at hand.¹⁴

Here, we consider a different route to extract information on the diffusive behavior, which is based on the GME. In the GME, the time evolution of the probability density in coarse space depends on the transition probabilities per unit time (rates) between positions in coarse space. The rates are complex functions of the OV,⁹ and a possible way to obtain them is with Milestoning. Milestoning was discussed extensively in the literature.^{15,16} In brief (more details are in Sec. II), Milestoning is a method that maps atomically detailed dynamics to kinetics and equilibrium of CV. The space of CVs is

partitioned into cells, and the dividers separating cells are called milestones. MD trajectories are initiated at the milestones and propagated in time until they hit, for the first time, another milestone. The identity of initiating and terminating milestones is recorded and so is the time of termination. The data are used in a non-Markovian theory to compute the kinetics and equilibrium of the CV.¹⁵ Explicit expressions for the free energy and mean first passage time are available.¹⁶ An advantage of using Milestoning is that the trajectories between the milestones are very short (picoseconds for biological molecules) and can be computed on massively parallel systems. Another advantage is that Milestoning speeds up the calculations of activated processes¹⁶ making the extraction of MD information significantly more efficient than in straightforward MD. It is therefore recommended to use Milestoning to connect atomically detailed simulations with CV dynamics. In the present paper, we offer such an algorithm.

The friction kernel does not appear explicitly in the GME. We extract the diffusion tensor, which is related to the friction, for the case in which the dynamics in coarse space is Markovian. In this case, the GME becomes the Master Equation (ME). While this is clearly an approximation, most other approaches we mentioned earlier relied on the same assumption. Milestoning makes it possible to compare the diffusion coefficient to data extracted with alternative approaches, and to assess the accuracy of the procedure. The assessment is possible since Milestoning is derived from atomically detailed trajectories, and can be made exact.⁵² Furthermore, the high numerical efficiency of Milestoning makes it attractive for the construction of CV-based models since statistical convergence at relevant time scales can be achieved. Finally since Milestoning is based on the GME and not the ME, extensions to time dependent friction kernel are possible and are topic for future research.

To derive the diffusion tensor, we exploit the so-called Kramers-Moyal (KM) expansion^{17–19} of a discrete ME which is truncated at the second order to yield the FPE.¹⁸ With the rate coefficients of the ME determined by the Milestoning analysis,¹⁶ it is possible to compute the space-dependent diffusion tensor on each milestone. Einstein relates the diffusion and friction tensors,¹⁰ so computing one is equivalent to computing the other.

Other methods are available to determine the space-dependent diffusion tensor from MD simulations:

- (i) Compute the zero time limit of the average of the position autocorrelation function, which gives the diffusion in the overdamped limit^{20,21}

$$D(q^*) = \lim_{t \rightarrow 0} \langle [q(t) - q(0)]^2 | q(0) = q^* \rangle / t :$$
- (ii) For overdamped dynamics, the space dependent diffusion coefficient in q^* is computed from the first passage time from q^* to q and backwards.²⁰
- (iii) Assuming Langevin dynamics and restraining the system with a harmonic potential (similar to Ref. 13), a local diffusion coefficient can be estimated using the velocity or position autocorrelation functions.^{21,22}
- (iv) Hummer²² proposed a method that shares some similarities with the approach presented in this paper. He also started from the ME. He partitioned the system into cells

and obtained the rates of transitions between them using straightforward MD and Bayesian statistics. A related Bayesian approach was used to derive a formula for the rates in a Markovian formulation of Milestoning that is analogous to the one that we used.²³ Hummer derives the rates using the assumption of detailed balance.²⁴ This approach does not hold in non-equilibrium conditions.²⁵ The formula that we derive does not require detailed balance.

- (v) Others have used Bayesian inference to determine the diffusion tensor. The likelihood used was not the probability of the trajectory given the rates, but the probability of the displacement in the CV space assuming overdamped dynamics.²⁶ Others employed the method in conjunction with adaptive biasing force technique (ABF) to estimate the diffusion tensor.²⁵ This method was recently used to study the permeability through a lipid bilayer.²⁷ Another study considered a small solute permeation through a lipid bilayer using ABF.²⁸ In this case, a two-dimensional reaction coordinate was used: the position and the orientation of the permeant were considered simultaneously. Interestingly, a similar application of Milestoning was already performed, where the Milestoning equations were used to describe the permeation of a small peptide molecule (NATA, N-acetyl-L-tryptophanamide) across a biological membrane.²⁹

The paper is organized as follows: in Sec. II, we illustrate the basic equations. We further show how to use Milestoning to extract the rate coefficients. In Sec. III, we discuss one and two-dimensional test cases. Overdamped trajectories are computed with predetermined diffusion tensors and deterministic forces. The trajectories are analyzed to recover the input forces and diffusion coefficients. We also illustrate an analysis of backbone diffusion of solvated alanine dipeptide based on atomically detailed simulations. Perspectives and conclusions are in Sec. V. The main equations of the calculations of the drift and diffusion coefficients of the FPE are Eqs. (18), (19), (21), (22), and (26).

II. METHOD

Consider a N -dimensional vector $\vec{x}$ of N CV. We assume that the time evolution of $\vec{x}$ is Markovian and that it follows the overdamped equation

$$d\vec{x} = -\beta \hat{D}(\vec{x}) \cdot \vec{\nabla} U(\vec{x}) dt + \vec{\nabla} \cdot \hat{D}(\vec{x}) dt + \sqrt{2} \hat{b}(\vec{x}) d\vec{\xi}(t). \quad (1)$$

In Eq. (1) we defined

1. $\beta = 1/k_B T$, with T temperature and k_B the Boltzmann constant;
2. $U(\vec{x})$ is the potential of mean force for the N CV. We will use the force $\vec{F}(\vec{x}) = -\vec{\nabla} U(\vec{x})$;
3. $\hat{D}(\vec{x})$ is the diffusion tensor. It is a symmetric positive-definite matrix connected to the matrix $\hat{b}(\vec{x})$ by the following relationship:

$$D_{ij}(\vec{x}) = \sum_{k=1}^N b_{ik}(\vec{x}) b_{jk}(\vec{x}). \quad (2)$$

4. $\vec{\xi}(t)$ is a vector of N independent Wiener processes, i.e., $\langle \xi_i(t) \rangle = 0$ and $\langle [\xi_i(t) - \xi_i(s)][\xi_j(t) - \xi_j(s)] \rangle = \delta_{ij}(t - s)$ with $t > s$, and δ_{ij} the Kronecker delta. The product $\hat{b}(\vec{x})d\vec{\xi}(t)$ is integrated according to Ito calculus.³⁰

The probability of finding the system in position $\vec{x}$ at time t is $p(\vec{x}, t)$. If the dynamics of the system follows Eq. (1), the probability $p(\vec{x}, t)$ is obtained solving the Fokker-Planck equation³¹

$$\begin{aligned} \frac{\partial p(\vec{x}, t)}{\partial t} = & - \sum_{i=1}^N \frac{\partial}{\partial x_i} \left[-\beta \sum_{j=1}^N D_{ij}(\vec{x}) \frac{\partial}{\partial x_j} U(\vec{x}) + \sum_{j=1}^N \frac{\partial}{\partial x_j} D_{ij}(\vec{x}) \right] p(\vec{x}, t) \\ & + \sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2}{\partial x_i \partial x_j} D_{ij}(\vec{x}) p(\vec{x}, t). \end{aligned} \quad (3)$$

The equilibrium distribution for this Fokker-Planck equation is the usual Boltzmann distribution $p_{\text{eq}}(\vec{x}) \propto \exp[-\beta U(\vec{x})]$.

Equation (3) is temporally and spatially continuous. We seek a connection between Eq. (3) and the discrete representation of space that we use in Milestoning.¹⁵ A suggestive model that connects continuous and discrete space of Markovian stochastic processes is the ME. The discrete ME is widely used in the chemical and physical sciences. It gained significant popularity recently as a fundamental component of the Markov state model.^{32–35} The continuous ME has also been discussed in the literature and textbooks,³⁶ but is considerably more controversial than the discrete representation of the ME or the continuous FPE. The limit of infinitesimally small time and spatial steps gives rise to singularities in the continuous ME that are not easy to overcome³⁶ and are not trivial to formally connect to Milestoning. Therefore, we restrict the discussion below to the discrete (in space) ME. We show how the rate coefficients of the ME are connected to the FPE, to the diffusion tensor, and to the mean force.

We further comment that the Milestoning theory is non-Markovian and therefore goes beyond the Markovian approximation to stochastic dynamics of coarse variables. The last approximation is used in the FPE and ME. Nevertheless, it is useful to consider this formal and computational language since many experimentalists and theorists are using it to describe their finding. Bridging the results of approximate approaches and exact formulation can be useful to further assess these approximations and to obtain simpler interpretations of complex non-Markovian analysis.

In the analysis of a MD trajectory, we divide the continuous CV into a small number of partitions and look at the transitions between them. For each of the CV (say i), we

introduce n_i states spaced by Δ_i and labeled by an integer index α_i , such that $1 \leq \alpha_i \leq n_i$. For convenience, let us define the square matrix of rank $N\hat{\Delta}$, with entries $\hat{\Delta}_{ij} = \Delta_i \delta_{ij}$, where δ_{ij} is the Kronecker delta. The position of the $\vec{\alpha}$ partition will be given by $\hat{\Delta}\vec{\alpha}$. The dynamics of the system is described by a series of jumps between the different partitions of the CV. The jumps occur at a specific rate, which is connected to the probability and the time scale of the jumps. The time evolution of the probability distribution expresses the balance between gain term and loss term, and is described by a master equation

$$\begin{aligned} \frac{dp(\hat{\Delta}\vec{\alpha}, t)}{dt} = & \sum_{\gamma_1=\alpha_1-\nu_1}^{\alpha_1+\nu_1} \cdots \sum_{\gamma_N=\alpha_N-\nu_N}^{\alpha_N+\nu_N} k(\hat{\Delta}\vec{\gamma}; \hat{\Delta}\vec{\alpha}) p(\hat{\Delta}\vec{\gamma}, t) \\ & - \sum_{\gamma_1=\alpha_1-\nu_1}^{\alpha_1+\nu_1} \cdots \sum_{\gamma_N=\alpha_N-\nu_N}^{\alpha_N+\nu_N} k(\hat{\Delta}\vec{\alpha}; \hat{\Delta}\vec{\gamma}) p(\hat{\Delta}\vec{\alpha}, t), \end{aligned} \quad (4)$$

where $k(\hat{\Delta}\vec{\gamma}; \hat{\Delta}\vec{\alpha})$ is the rate of going from the state $\vec{\gamma}$ to the state $\vec{\alpha}$, and ν_i is such that on the i th CV transitions occur exclusively within the set of states $[\alpha_i - \nu_i; \alpha_i + \nu_i]$. Following the steps described in Ref. 36, a relationship is now developed between the time evolution of the probability density in Eq. (3) and the probability of Eq. (4) at the limit of small spatial displacement Δ . We define the rates as $\bar{k}(\hat{\Delta}\vec{\gamma}; \hat{\Delta}\vec{\alpha})$, which is the rate of being in $\hat{\Delta}\vec{\gamma}$ and performing a jump $\hat{\Delta}\vec{\epsilon}$, where $\vec{\epsilon}$ is such that $\vec{\alpha} = \vec{\gamma} + \vec{\epsilon}$. Now, we can expand in a Taylor series $\bar{k}(\hat{\Delta}\vec{\alpha} - \hat{\Delta}\vec{\epsilon}; \hat{\Delta}\vec{\alpha}) p(\hat{\Delta}\vec{\alpha} - \hat{\Delta}\vec{\epsilon}, t)$ around $\bar{k}(\hat{\Delta}\vec{\alpha}; \hat{\Delta}\vec{\alpha}) p(\hat{\Delta}\vec{\alpha}, t)$. Note that the rate constant is a slowly varying function of the first variable and therefore the expansion is valid. The first term of the Taylor series cancels out with the loss term in Eq. (4). The remaining terms give

$$\begin{aligned} \frac{\partial p(\hat{\Delta}\vec{\alpha}, t)}{\partial t} = & - \sum_{i=1}^N \frac{\partial}{\partial x_i} \left\{ \sum_{\gamma_1=\alpha_1-\nu_1}^{\alpha_1+\nu_1} \cdots \sum_{\gamma_N=\alpha_N-\nu_N}^{\alpha_N+\nu_N} \Delta_i (\gamma_i - \alpha_i) k(\vec{x}; \hat{\Delta}\vec{\gamma}) \right\} p(\vec{x}, t) \Big|_{\vec{x}=\hat{\Delta}\vec{\alpha}} \\ & + \sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2}{\partial x_i \partial x_j} \left\{ \left[\frac{1}{2} \sum_{\gamma_1=\alpha_1-\nu_1}^{\alpha_1+\nu_1} \cdots \sum_{\gamma_N=\alpha_N-\nu_N}^{\alpha_N+\nu_N} \Delta_i \Delta_j (\gamma_i - \alpha_i) (\gamma_j - \alpha_j) k(\vec{x}; \hat{\Delta}\vec{\gamma}) \right] p(\vec{x}, t) \right\} \Big|_{\vec{x}=\hat{\Delta}\vec{\alpha}} + \cdots \end{aligned} \quad (5)$$

The truncation of this expansion at the second order, after appropriate normalization of the probability, yields an approximation to the Fokker-Planck equation (3). This is a Kramers-Moyal expansion^{19,37,38} of the discrete master equation. The properties of this expansion and its truncation for the continuous master equation have been discussed extensively in the past.^{17,19,31,36}

If we truncate Eq. (5) at the second order, we can perform a term-by-term comparison with the Fokker-Planck equation (3). In this way, we derive an expression for the drift and the diffusion terms of the Fokker-Planck equation as a function of the rate coefficients

$$M_i^{(1)}(\vec{x} = \hat{\Delta}\vec{\alpha}) = -\beta \sum_{j=1}^N D_{ij}(\vec{x}) \frac{\partial}{\partial x_j} U(\vec{x}) + \sum_{j=1}^N \frac{\partial}{\partial x_j} D_{ij}(\vec{x})$$

$$= \sum_{\gamma_1=\alpha_1-\nu_1}^{\alpha_1+\nu_1} \cdots \sum_{\gamma_N=\alpha_N-\nu_N}^{\alpha_N+\nu_N} \Delta_i(\gamma_i - \alpha_i) k(\hat{\Delta}\vec{\alpha}; \hat{\Delta}\vec{\gamma}), \quad (6)$$

$$M_{ij}^{(2)}(\vec{x} = \hat{\Delta}\vec{\alpha}) = D_{ij}(\vec{x}) = \frac{1}{2} \sum_{\gamma_1=\alpha_1-\nu_1}^{\alpha_1+\nu_1} \cdots \sum_{\gamma_N=\alpha_N-\nu_N}^{\alpha_N+\nu_N} \Delta_i \Delta_j (\gamma_i - \alpha_i) (\gamma_j - \alpha_j) k(\hat{\Delta}\vec{\alpha}; \hat{\Delta}\vec{\gamma}), \quad (7)$$

where $M_i^{(1)}(\vec{x})$ is the i th component of the vector of first moments of the rates, and $M_{ij}^{(2)}(\vec{x})$ is the ij -th component of the tensor of second moments of the rates.

Equation (7) provides a method to extract the diffusion tensor once the rates are known. We now describe the Milestoning approach to determine the rate coefficient.¹⁵ We partition our space using milestones, as illustrated in Figure 1(a). A two-dimensional system is divided in N_{MS} milestones defined by lines parallel to the y -axis and crossing the x -axis at positions x , $x + \Delta_x$, $x + 2\Delta_x, \dots$. In this two-dimensional space, the x -coordinate is the CV. The milestones are denoted by Greek letters and are used to label fragments of a trajectory.

FIG. 1. The red, green, and blue lines are the milestones. Every time that the trajectory crosses a milestone it is assigned to that milestone. In the figure, the trajectory is labeled with the color of the last milestone that was crossed. Panel (a): One dimensional milestones. Panel (b): Milestones adopted to compute Eqs. (13), (14), and (17); All the milestones are parallel to the y -axis, but in contrast to (a), they are not lines. Instead, they are segments whose ends are rhomboidal arrows. Panel (c): Milestones adopted to compute Eqs. (15)–(17). Same trajectory as (a) and (b), but this time discretized to show transitions between milestones parallel to the x -axis. As in (b), the milestones are segments whose ends are represented as rhomboidal arrows. Figures A, B, and C have previously been published in the Ph.D. dissertation of M. L. Mugnai.⁵¹

Every time a trajectory crosses a milestone it is assigned to that milestone until it crosses another milestone. In this way, the trajectory is coarse grained to hops between a series of milestones. Instead of $(x(t), y(t))$ —the full coordinate set as a function of time—we have $\alpha(t_\alpha \leq t < t_\gamma), \gamma(t_\gamma \leq t < t_\eta), \eta(t_\eta \leq t < t_k), \dots$ where t_μ is the time in which the μ milestone was crossed by the trajectory $(x(t), y(t))$. Alternatively, we can run multiple trajectories starting from each milestone and terminate them when they cross a neighboring milestone. Both approaches yield the kinetic information that we need, namely, the local transition probabilities and local transition times between nearby milestones. The use of trajectory fragments instead of a single long trajectory is significantly more efficient. The calculations with trajectory fragments can be approximate, or with the investment of additional computational resources and iterations can be made exact (see Ref. 52).

To simplify the notation, we write $f(\hat{\Delta}\vec{\alpha}) = f(\vec{\alpha})$ in the rest of the paper. We record the average amount of time, $\tau(\alpha)$, that it takes to cross another milestone given that at time zero the trajectory crossed milestone α . We also record the probability that, given that the trajectory starts from milestone α , the next milestone crossed is $\gamma - p(\alpha, \gamma)$. The rate coefficient of going from milestone α to milestone γ is¹⁶

$$k(\alpha; \gamma) = \frac{p(\alpha, \gamma)}{\tau(\alpha)}. \quad (8)$$

The validity of the Milestoning procedure was discussed extensively in the past (see for instance Ref. 39) and we refer the reader to the literature for further discussions.

Once we extract the rate coefficients, we can calculate the force and the space-dependent diffusion using Eqs. (6) and (7). For clarity we consider the one-dimensional case first, extension to higher dimensions are discussed later in the article. The drift and diffusion coefficients that we need are

$$M^{(1)}(\alpha) = \beta D(\alpha) F(\alpha) + \frac{d}{dx} D(x) \Big|_{x=\Delta\alpha}$$

$$= \sum_{\gamma=\alpha-\nu}^{\alpha+\nu} \Delta(\gamma - \alpha) k(\alpha; \gamma), \quad (9)$$

$$M^{(2)}(\alpha) = D(\alpha) = \frac{1}{2} \sum_{\gamma=\alpha-\nu}^{\alpha+\nu} \Delta^2(\gamma - \alpha)^2 k(\alpha; \gamma). \quad (10)$$

In one dimension, the only accessible milestones to α are milestones $\gamma = \alpha \pm 1$. Consider two rates per milestone: the rate of going forward, $k^+(\alpha) = k(\alpha; \alpha + 1)$ and the rate of going backward, $k^-(\alpha) = k(\alpha; \alpha - 1)$. Therefore, the drift and diffusion coefficients are

$$M^{(1)}(\alpha) = \beta D(\alpha) F(\alpha) + \left. \frac{d}{dx} D(x) \right|_{x=\Delta\alpha} = \Delta [k^+(\alpha) - k^-(\alpha)], \quad (11)$$

$$M^{(2)}(\alpha) = D(\alpha) = \Delta^2 \frac{k^+(\alpha) + k^-(\alpha)}{2}. \quad (12)$$

The drift term (11) is the difference between the forward and backward rate, while the diffusion term (12) is proportional to the average of the rates.

This paper computes the diffusion coefficient as a function of space, i.e., of the position of the milestone. To compute the diffusion coefficient Eq. (12) is enough. On the other hand, we could compute the mean force using Eq. (11). As we will show in Sec. III, this computation often turns out to be very noisy.

Fortunately, the mean force can be extracted efficiently with other means such as umbrella sampling,⁴⁰ so in principle it is not necessary to solve Eq. (11). Nevertheless, if we compute the mean force from the ME, our method can be tested and compared to others. To do so, we use Eq. (11) and solve for $\beta F(\alpha)$

$$M^{(1)}(\alpha) \leftarrow M^{(1)}(\alpha) - \frac{D(\alpha+1) - D(\alpha-1)}{2\Delta},$$

$$\beta F(\alpha) = \frac{M^{(1)}(\alpha)}{D(\alpha)},$$

where the derivative of the diffusion coefficient is estimated numerically.

The one-dimensional case is the simplest problem. Often though, we need to account for multiple coarse variables. These variables may not be independent, and the analysis of their correlation is of interest. Let us consider a two-dimensional case, where the two coarse variables are (x, y) . To simplify the notation, we identify the two components of the vector $\vec{v}$ as (v_x, v_y) . We need to compute the drift and diffusion coefficients

$$M_x^{(1)}(\vec{\alpha}) = \beta D_{xx}(\vec{\alpha}) F_x(\vec{\alpha}) + \beta D_{xy}(\vec{\alpha}) F_y(\vec{\alpha}) + \frac{\partial}{\partial x} D_{xx}(\vec{\alpha}) + \frac{\partial}{\partial y} D_{yx}(\vec{\alpha})$$

$$= \sum_{\gamma_x=\alpha_x-\nu_x}^{\alpha_x+\nu_x} \sum_{\gamma_y=\alpha_y-\nu_y}^{\alpha_y+\nu_y} \Delta_x (\gamma_x - \alpha_x) k(\vec{\alpha}; \vec{\gamma}), \quad (13)$$

$$M_{xx}^{(2)}(\vec{\alpha}) = D_{xx}(\vec{\alpha}) = \frac{1}{2} \sum_{\gamma_x=\alpha_x-\nu_x}^{\alpha_x+\nu_x} \sum_{\gamma_y=\alpha_y-\nu_y}^{\alpha_y+\nu_y} \Delta_x^2 (\gamma_x - \alpha_x)^2 k(\vec{\alpha}; \vec{\gamma}), \quad (14)$$

$$M_y^{(1)}(\vec{\alpha}) = \beta D_{yx}(\vec{\alpha}) F_x(\vec{\alpha}) + \beta D_{yy}(\vec{\alpha}) F_y(\vec{\alpha}) + \frac{\partial}{\partial x} D_{xy}(\vec{\alpha}) + \frac{\partial}{\partial y} D_{yy}(\vec{\alpha})$$

$$= \sum_{\gamma_x=\alpha_x-\nu_x}^{\alpha_x+\nu_x} \sum_{\gamma_y=\alpha_y-\nu_y}^{\alpha_y+\nu_y} \Delta_y (\gamma_y - \alpha_y) k(\vec{\alpha}; \vec{\gamma}), \quad (15)$$

$$M_{yy}^{(2)}(\vec{\alpha}) = D_{yy}(\vec{\alpha}) = \frac{1}{2} \sum_{\gamma_x=\alpha_x-\nu_x}^{\alpha_x+\nu_x} \sum_{\gamma_y=\alpha_y-\nu_y}^{\alpha_y+\nu_y} \Delta_y^2 (\gamma_y - \alpha_y)^2 k(\vec{\alpha}; \vec{\gamma}), \quad (16)$$

$$M_{xy}^{(2)}(\vec{\alpha}) = D_{xy}(\vec{\alpha}) = \frac{1}{2} \sum_{\gamma_x=\alpha_x-\nu_x}^{\alpha_x+\nu_x} \sum_{\gamma_y=\alpha_y-\nu_y}^{\alpha_y+\nu_y} \Delta_x \Delta_y (\gamma_x - \alpha_x) (\gamma_y - \alpha_y) k(\vec{\alpha}; \vec{\gamma}). \quad (17)$$

The drift and diffusion coefficients in Eqs. (13)–(17) can be classified into three types. The first two (Eqs. (13) and (14)) are the rate of jump first and second moments in $\Delta_x(\gamma_x - \alpha_x)$. The next two (Eqs. (15) and (16)) are the first and second moments in $\Delta_y(\gamma_y - \alpha_y)$. Finally, the last one (Eq. (17)) is the moment in $\Delta_x(\gamma_x - \alpha_x)\Delta_y(\gamma_y - \alpha_y)$.

What is the best choice of milestones for this two-dimensional system? When building a lattice in coarse space to estimate trajectory crossing through milestones, care must be used in the selection of the milestones. There are several considerations. First, the milestones must be sufficiently close to each

other such that the truncation of the expansion of the ME will be sound. Second, the milestones should not be too close to each other to allow for reasonably long trajectory between the milestones and some loss of memory, which is required to satisfy the Markov assumption. Third, the implementation of the hypersurfaces should be simple. The simplicity makes it possible to conduct efficient and accurate numerical calculations of the drift and diffusion coefficients. These requirements are not all compatible and some compromises must be made. We discuss below a few options and implementations for Milestoning implementation in more than one dimension.

The first practical paper examining Milestoning in more than one dimension used Voronoi tessellation.⁴¹ The computational implementation is simple and attractive. However, the distance between milestones is not a simple function and can be as short as zero, not leaving sufficient time for memory loss. The memory loss was addressed in a later paper⁴² in which the milestones depend on the direction of crossing and a spatial shift of the boundaries of the Voronoi cells was imposed. However, the use of this type of milestones makes the calculation of the drift and diffusion coefficients complicated. This is because the milestones are not parallel to each other. As a result, the displacements in the x and y -directions between the starting point on milestone $\vec{\alpha}$ and the arrival point on milestone $\vec{\beta}$ are not $\Delta_x(\gamma_x - \alpha_x)$ and $\Delta_y(\gamma_y - \alpha_y)$, as it appears in Eqs. (13)–(17), but depend on the actual starting and arrival points on the milestone. An adequate choice of milestones would reduce the uncertainty in the distance.

Here, we adopt two different sets of milestones to compute the functions (13)–(17). The first two are simplified if the transitions that we count occur between milestones parallel to the y -axis and at a fixed distance along the x -axis. With this choice, we get that in Eqs. (13) and (14) $\Delta_x(\gamma_x - \alpha_x) = \pm\Delta_x$ or $\Delta_x(\gamma_x - \alpha_x) = 0$, and it is identical for all points of milestone $\vec{\alpha}$ and $\vec{\gamma}$. Similarly to compute Eqs. (15) and (16), we propose to use a set of milestones parallel to the x -axis and at fixed distance along the y -axis. With this choice, we get in these equations $\Delta_y(\gamma_y - \alpha_y) = \pm\Delta_y$ or $\Delta_y(\gamma_y - \alpha_y) = 0$. Again, these differences do not depend on specific points that

were crossed on milestone $\vec{\alpha}$ and $\vec{\gamma}$, but are only function of the two milestones. A related reasoning is not possible for the calculation of the off-diagonal term of the diffusion tensor (17). In that case, we use a combination of the rates derived in the previous two cases and the displacements $\Delta_x(\gamma_x - \alpha_x)$ and $\Delta_y(\gamma_y - \alpha_y)$ are the displacements of the center of milestone $\vec{\gamma}$ from the center of milestone $\vec{\alpha}$.

We proceed to explain in detail each of these three different cases.

To compute the moments in Eqs. (13) and (14), we discretize the trajectory shown in Figure 1(a) as illustrated in Figure 1(b).

Each of the N_{MSX} milestones of Figure 1(a) is now divided into N_{MSY} milestones (represented in Figure 1(b) by continuous, dashed, and dotted lines) of size Δ_y . The idea of partitioning the milestones in this way was originally put forward by Hawk.⁴³ We do not adopt a clustering procedure to partition the milestones, but we divide them *a priori*. The N_{MSY} milestones resulting from the partitioning are labeled with increasing integers, from 1 to N_{MSY} , from bottom to top. Following Figure 1(b), each milestone is defined as the set of points (x, y) such that $\{x = \Delta_x\alpha_x, \Delta_y\alpha_y - \frac{\Delta_y}{2} \leq y < \Delta_y\alpha_y + \frac{\Delta_y}{2}\}$. The Milestoning analysis of the trajectory determines the rate of going from any point on a milestone $\vec{\alpha}$ to any point on a neighbor milestone $\vec{\gamma}$. Note that each milestone has $3N_{MSY} - 1$ neighbors belonging to the set $\{(\gamma_x, \gamma_y) | \alpha_x - 1 \leq \gamma_x \leq \alpha_x + 1 \text{ and } (\gamma_x, \gamma_y) \neq (\alpha_x, \alpha_y)\}$. The rates computed with Milestoning are labeled with the superscript (x) . The resulting drift and diffusion equations are

$$\begin{aligned} M_x^{(1)}(\vec{\alpha}) &= \beta D_{xx}(\vec{\alpha}) F_x(\vec{\alpha}) + \beta D_{xy}(\vec{\alpha}) F_y(\vec{\alpha}) + \frac{\partial D_{xx}(\vec{\alpha})}{\partial x} + \frac{\partial D_{xy}(\vec{\alpha})}{\partial y} \\ &= \sum_{\gamma_y=1}^{N_{MSY}} \Delta_x \left[k^{(x)}(\alpha_x, \alpha_y; \alpha_x + 1, \gamma_y) - k^{(x)}(\alpha_x, \alpha_y; \alpha_x - 1, \gamma_y) \right] \\ &= \Delta_x \left[k^{(x)+}(\vec{\alpha}) - k^{(x)-}(\vec{\alpha}) \right], \end{aligned} \quad (18)$$

$$\begin{aligned} M_{xx}^{(2)}(\vec{\alpha}) &= D_{xx}(\vec{\alpha}) \\ &= \frac{1}{2} \sum_{\gamma_y=1}^{N_{MSY}} \Delta_x^2 \left[k^{(x)}(\alpha_x, \alpha_y; \alpha_x + 1, \gamma_y) + k^{(x)}(\alpha_x, \alpha_y; \alpha_x - 1, \gamma_y) \right] \\ &= \Delta_x^2 \frac{k^{(x)+}(\vec{\alpha}) + k^{(x)-}(\vec{\alpha})}{2}, \end{aligned} \quad (19)$$

where

$$\begin{aligned} k^{(x)+}(\vec{\alpha}) &= \sum_{\gamma_y=1}^{N_{MSY}} k^{(x)}(\alpha_x, \alpha_y; \alpha_x + 1, \gamma_y), \\ k^{(x)-}(\vec{\alpha}) &= \sum_{\gamma_y=1}^{N_{MSY}} k^{(x)}(\alpha_x, \alpha_y; \alpha_x - 1, \gamma_y). \end{aligned} \quad (20)$$

As in the one-dimensional case, the drift term (18) is given by the difference between the forward and backward rates. The diffusion term is the average of the two rates.

A similar strategy can be devised to compute the drift and diffusion coefficients in Eqs. (15) and (16), and the same trajectory is re-analyzed with a new set of milestones parallel to the x axis as in Figure 1(c). In this case, the N_{MSY} milestones parallel to the x -axis are separated by Δ_y . Each of these milestones is then divided into N_{MSX} milestones of size Δ_x , numbered with increasing labels from left to right. A milestone $\vec{\alpha} = (\alpha_x, \alpha_y)$ is identified by the set of points (x, y) such that $\{\Delta_x\alpha_x - \frac{\Delta_x}{2} \leq x < \Delta_x\alpha_x + \frac{\Delta_x}{2}, y = \Delta_y\alpha_y\}$. As before, the $3N_{MSX} - 1$ milestones that are neighbor to milestone $\vec{\alpha}$ are those

belonging to the set $\{(\gamma_x, \gamma_y) \mid \alpha_y - 1 \leq \gamma_y \leq \alpha_y + 1 \text{ and } (\gamma_x, \gamma_y) \neq (\alpha_x, \alpha_y)\}$. The rates obtained with the scheme in Figure

1(c) are labeled by a superscript (y) . The resulting drift and diffusion coefficients are

$$\begin{aligned} M_y^{(1)}(\vec{\alpha}) &= \beta D_{yx}(\vec{\alpha}) F_x(\vec{\alpha}) + \beta D_{yy}(\vec{\alpha}) F_y(\vec{\alpha}) + \frac{\partial D_{yx}(\vec{\alpha})}{\partial x} + \frac{\partial D_{yy}(\vec{\alpha})}{\partial y} \\ &= \sum_{\gamma_x=1}^{N_{MSX}} \Delta_y \left[k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y + 1) - k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y - 1) \right] \\ &= \Delta_y \left[k^{(y)+}(\vec{\alpha}) - k^{(y)-}(\vec{\alpha}) \right], \end{aligned} \quad (21)$$

$$\begin{aligned} M_{yy}^{(2)}(\vec{\alpha}) &= D_{yy}(\vec{\alpha}) \\ &= \frac{1}{2} \sum_{\gamma_x=1}^{N_{MSX}} \Delta_y^2 \left[k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y + 1) + k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y - 1) \right] \\ &= \Delta_y^2 \frac{k^{(y)+}(\vec{\alpha}) + k^{(y)-}(\vec{\alpha})}{2}, \end{aligned} \quad (22)$$

where, as before

$$\begin{aligned} k^{(y)+}(\vec{\alpha}) &= \sum_{\gamma_x=1}^{N_{MSX}} k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y + 1), \\ k^{(y)-}(\vec{\alpha}) &= \sum_{\gamma_x=1}^{N_{MSX}} k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y - 1). \end{aligned} \quad (23)$$

As in the one-dimensional case, the drift is given by the difference of the upward and downward rates, the diffusion term by their average.

We are left with the last case: the calculation of the off-diagonal term of the diffusion tensor in Eq. (17). To do so, we plug the rates obtained from the first scheme $k^{(x)}(\vec{\alpha}, \vec{\gamma})$ in (17). We obtain

$$\begin{aligned} M_{xy}^{(2)(x)}(\alpha_x, \alpha_y) &= \frac{1}{2} \sum_{\gamma_y=1}^{N_{MSY}} (\gamma_y - \alpha_y) \Delta_y \Delta_x \\ &\quad \times \left[k^{(x)}(\alpha_x, \alpha_y; \alpha_x + 1, \gamma_y) \right. \\ &\quad \left. - k^{(x)}(\alpha_x, \alpha_y; \alpha_x - 1, \gamma_y) \right]. \end{aligned} \quad (24)$$

Note that the term $(\gamma_y - \alpha_y) \Delta_x \Delta_y$ is the product of the displacement on the x and y-direction that occurs in going from the milestone $\vec{\alpha}$ to $\vec{\gamma}$. While the displacement on the x-direction ($\pm \Delta_x$) is the same for every starting and arrival point, the displacement in the y-direction $((\gamma_y - \alpha_y) \Delta_y)$ is the displacement between the centers of the milestones (see Figure 2(a)). This is an approximation: the displacement on the y-direction depends on the initial and final points on the milestones. When the size of the milestone is small compared to the length scale in which the rate coefficient changes, the approximation is accurate.

The same procedure can be carried out with the second Milestoning scheme, i.e., we can plug $k^{(y)}(\vec{\alpha}; \vec{\gamma})$ in Eq. (17) to get

$$\begin{aligned} M_{xy}^{(2)(y)}(\alpha_x, \alpha_y) &= \frac{1}{2} \sum_{\gamma_x=1}^{N_{MSX}} (\gamma_x - \alpha_x) \Delta_x \Delta_y \\ &\quad \times \left[k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y + 1) \right. \\ &\quad \left. - k^{(y)}(\alpha_x, \alpha_y; \gamma_x, \alpha_y - 1) \right]. \end{aligned} \quad (25)$$

Here, the situation is inverted. The displacement between two milestones on the y-direction ($\pm \Delta_y$) is the same for every point in the milestone, while the displacement in the x-direction $((\gamma_x - \alpha_x) \Delta_x)$ depends on the actual crossing point on the starting and arrival milestone (see Figure 2(b)).

The two Eqs. (24) and (25) can be used to compute exactly the same quantity. There is no reason to choose one or the other approaches to compute the off-diagonal elements *a priori*. So, we choose to average them, which also ensures symmetry in the treatment of the x and y-coordinate, and define a new equation for the off-diagonal term of the diffusion tensor

$$D_{xy}(\alpha_x, \alpha_y) = \frac{1}{2} \left[M_{xy}^{(2)(x)}(\alpha_x, \alpha_y) + M_{xy}^{(2)(y)}(\alpha_x, \alpha_y) \right]. \quad (26)$$

This is the equation that we use in Sec. III. It is interesting to note that we have tested the difference of the average (Eq. (26)) compared to the individual elements (Eqs. (24) and (25)) and found it to be negligible in the examples we considered (see Figures 6(d) and 10(b)–10(e)).

To perform the analysis we adopt the following protocol:

- (1) Run the simulation;
- (2) Perform the Milestoning analysis with the first scheme and extract the rates $k^{(x)}(\vec{\alpha}; \vec{\gamma})$.
- (3) Perform the Milestoning analysis with the second scheme and extract the rates $k^{(y)}(\vec{\alpha}; \vec{\gamma})$.
- (4) Extract the different elements of the diffusion tensor using Eqs. (19), (22), and (26). This completes the calculation of the diffusion tensor.
- (5) Extract the two elements of the drift vector using Eqs. (18) and (21).
- (6) Extract the force vector using the following steps:

$$\begin{aligned}
M_x^{(1)}(\alpha_x, \alpha_y) &\leftarrow M_x^{(1)}(\alpha_x, \alpha_y) - \frac{D_{xx}(\alpha_x + 1, \alpha_y) - D_{xx}(\alpha_x - 1, \alpha_y)}{2\Delta_x} - \frac{D_{xy}(\alpha_x, \alpha_y + 1) - D_{xy}(\alpha_x, \alpha_y - 1)}{2\Delta_y} \\
M_y^{(1)}(\alpha_x, \alpha_y) &\leftarrow M_y^{(1)}(\alpha_x, \alpha_y) - \frac{D_{yy}(\alpha_x, \alpha_y + 1) - D_{yy}(\alpha_x, \alpha_y - 1)}{2\Delta_y} - \frac{D_{xy}(\alpha_x + 1, \alpha_y) - D_{xy}(\alpha_x - 1, \alpha_y)}{2\Delta_x} \\
\begin{pmatrix} \beta F_x(\alpha_x, \alpha_y) \\ \beta F_y(\alpha_x, \alpha_y) \end{pmatrix} &= \begin{pmatrix} D_{xx}(\alpha_x, \alpha_y) & D_{xy}(\alpha_x, \alpha_y) \\ D_{xy}(\alpha_x, \alpha_y) & D_{yy}(\alpha_x, \alpha_y) \end{pmatrix}^{-1} \begin{pmatrix} M_x^{(1)}(\alpha_x, \alpha_y) \\ M_y^{(1)}(\alpha_x, \alpha_y) \end{pmatrix}.
\end{aligned}$$

Steps (1), (2), and (3) of the algorithm are the only ones that are computationally costly and involve calculations of trajectories. The other steps are performed on the reduced space of the milestones, so they are not expensive.

As discussed earlier, the main aim of this paper is to extract the diffusion tensor, not necessarily the derivatives of the potential of mean force. The estimates of the derivatives of the potential of mean force that we obtain with this method are noisy. Also, there is no assurance that the numerically

derived force is a gradient of a potential function. Nevertheless, since the components of the force are computed essentially for free, it is a good practice to estimate them and compare them with forces estimated by other methods. For example, the mean force can be calculated by a numerical differentiation of the potential of mean force (PMF). The PMF can be estimated with umbrella sampling,⁴⁰ the log of the probability density extracted from MD simulations, Milestoning,¹⁶ and more. The current calculation is deemed more reliable if the force computed with it matches another estimate.

The choice of milestones made for this paper relies on termination of trajectories in only one dimension, the one perpendicular to the milestones (x for the first scheme discussed, y for the second). As we discussed above, the advantage is that calculations of drift and diffusion coefficients are straightforward. Also, the memory loss requirement is trivial

FIG. 2. (a) Milestoning scheme to compute the off-diagonal element of the diffusion tensor adopting the same scheme as in Figure 1(b). The trajectories start from the milestones highlighted by the green dot and terminate in any one of the other milestones. The relative position in space of the arrival milestone compared to the starting milestone is reported in parenthesis. This figure has previously been published in the Ph.D. dissertation of M. L. Mugnai.⁵¹ (b) Same as (a), but the milestone scheme adopted here is the one in Figure 1(c).

FIG. 3. Force (top) and the diffusion coefficient (bottom) as a function of coordinate x . The red line shows the input force and the diffusion coefficient, the green dots are the results of simulations. The error bars are smaller than the size of the symbols.

FIG. 4. Force, x-component. (a) The thin meshed surface represents the force given in input, the thicker lines interpolate the values computed at the 11×11 milestones. (b)–(e) Different cross-sections of the 3D plot. The red line is the expected profile. The green points are the result of the simulation. (b) The force as a function of x with $y = -0.6$. (c) The force as a function of x with $y = 0.6$. (d) The force as a function of y with $x = -0.6$. (e) The force as a function of y with $x = 0.6$.

to satisfy. The disadvantage is that the termination times in the direction parallel to the milestones (y for the first scheme, x for the second) are not precise since termination positions in this direction have errors of order of the milestone spacing. Another concern is that in the direction parallel to the milestone the diffusion will not be local. The last concern, however, is never realized in practice since the locality is determined by the short termination time of the diffusion along the direction perpendicular to the milestones. We attempted another discretization based on directional Milestoning implementation of the Voronoi tessellation,⁴² and obtained worse results that are not reported here.

III. RESULTS

As an illustration of the algorithm presented in Sec. II, we conducted several test cases for which we know the exact answers. All the tests consist of the following steps:

- Choose a functional form for the potential $U(\vec{x})$ and the matrix $\hat{b}(\vec{x})$ or $\hat{D}(\vec{x})$.
- Run a Brownian dynamics simulation using Ermak and McCammon formula⁴⁴ that follows the Euler-Maruyama algorithm⁴⁵

$$\vec{x}_{n+1} = \vec{x}_n + \beta \hat{D}(\vec{x}_n) \cdot \vec{F}(\vec{x}_n) dt + \vec{\nabla} \cdot \hat{D}(\vec{x}_n) dt + \sqrt{2dt} \hat{b}(\vec{x}_n) \cdot \vec{N}(0,1), \quad (27)$$

FIG. 5. Force, y-component and the diffusion coefficient. (a) The thin meshed surface represents the force given in input. The thicker lines interpolate the values computed at the 11×11 milestones. (b)–(c) Different cross-sections of the 3D plot. The red line is the expected profile. The green points are the result of the simulation. (b) The force as a function of x with $y = 0.0$. (c) The force as a function of y with $x = 0.0$. (d) Diffusion tensor, x - x -component. The thin meshed surface represents the function given in input. The thicker lines interpolate the values computed at the 11×11 milestones. (e) A cross-section corresponding to $y = 0.0$ is shown. The error bars are smaller than the size of the symbols. Panels (d) and (e) have previously been published in the Ph.D. dissertation of M. L. Mugnai.⁵¹

FIG. 6. Diffusion tensor, yy ((a) and (b) panels) and xy components ((c) and (d) panels). (a) The thin meshed surface represents the function given in input. The thicker lines interpolate the values computed at the 11×11 milestones. (b) A cross section corresponding to $y = 0.0$ is shown. (c) The lines interpolate the values computed at the 11×11 milestones. The expected result is 0. (d) A cross-section corresponding to $y = 0.0$ is shown. The error bars are smaller than the size of the symbols. The green dots are obtained using Eq. (26), the blue squares using Eq. (24), and the pink triangles using Eq. (25). Note that the deviations from zero are an order of magnitude smaller than the other elements of the diffusion tensor. Figures (a) and (b) have previously been published in the Ph.D. dissertation of M. L. Mugnai.⁵¹

FIG. 7. Force, second 2D test case, x-component. (a) The thin meshed surface represents the force given in input, the thicker lines interpolate the values computed at the 11×11 milestones. (b)–(e) Different cross-sections of the 3D plot. The red line is the expected profile, the green points represent the result of the simulation. (b) The force as a function of x with $y = -0.6$. (c) The force as a function of x with $y = 0.6$. (d) The force as a function of y with $x = -0.6$. (e) The force as a function of y with $x = 0.6$.

FIG. 8. Force, second 2D test case, y-component. (a) The thin meshed surface represents the force given in input, the thicker lines interpolate the values computed at the center of the 11×11 milestones. (b)–(c) Different cross-sections of the 3D plot. The red line is the expected profile, the green points represent the result of the simulation. (b) The force as a function of x with $y = 0.0$. (c) The force as a function of y with $x = 0.0$.

FIG. 9. Diffusion tensor, second 2D test case, xx ((a) and (b) panels) and yy ((c) and (d) panels) components. (a) and (c) The thin meshed surface represents the function given in input, the thicker lines interpolate the values computed at the 11×11 milestones. (b) and (d) A cross-section corresponding to $y = 0.0$ is shown. The red line is the expected profile. The green dots represent the result of the simulation. The error bars are smaller than the size of the symbols. Panels (a)–(d) have previously been published in the Ph.D. dissertation of M. L. Mugnai.⁵¹

FIG. 10. Diffusion tensor, second test case, xy -component. (a) The thin meshed surface represents the function given in input, the thicker lines interpolate the values computed at the 11×11 milestones. (B)–(E) Four cross-sections of the 2D surface: (b) cross-section corresponding to $y = -0.6$; (c) cross-section corresponding to $y = 0.6$; (d) cross-section corresponding to $x = -0.6$; (e) cross-section corresponding to $x = 0.6$. The red line is the expected profile. Green dots, blue squares, and pink triangles represent the result of the simulation. The green dots are obtained using Eq. (26), the blue squares using Eq. (24), and the pink triangles using Eq. (25). The statistical error bars are smaller than the size of the symbols.

FIG. 11. Analysis of the numerical errors in the calculations of the diffusion tensor. The green line is the exact result. (a) The red squares are computed for distances between two milestones $\Delta_x = 0.2$ and variable size Δ_y of a milestone. (b) The blue dots are simulations carried out with $\Delta_x = 0.2$ and $\Delta_y = 0.2$ using the time steps reported on the x -axis, while the red squares were obtained from simulations performed after changing the spacing between milestones to $\Delta_x = 0.01$. Figures 11(a) and 11(b) have previously been published in the Ph.D. dissertation of M. L. Mugnai.⁵¹

FIG. 12. ϕ (panel (a)) and ψ (panel (c)) components of the force vector extracted using the algorithm described in Sec. II. The values reported are only those corresponding to regions of dihedral space that are visited with probability $p > 0.0001$. Panels (b) and (d) are the differences between the estimates of the force obtained with the algorithm described in Sec. II and the estimates obtained from numerical differentiation of the free energy. Note that the scale for the error is around 1/3 of the scale of the signal.

where $\vec{N}(0,1)$ is a vector of independent random numbers sampled from a Gaussian distribution of average 0 and variance 1, and $\vec{F}(\vec{x}) = -\vec{\nabla}U(\vec{x})$ is the force.

- iii. Perform a Milestoning analysis to compute the rate coefficients from the time traces.
- iv. Use the algorithm presented before to extract the diffusion tensor and the force vector.
- v. Compare the input force vector and diffusion tensor with the results generated by our analysis.

A. Test 1

We first consider a one-dimensional test case. We examine the example used in Ref. 22 where the potential is $\beta U(x) = -\cos(2x)$ and the diffusion constant is $D(x) = [0.2 + 0.1 \sin(x)] \text{rad}^2 / \text{ps}$. We divide the system into 24 milestones. We run 30 000 trajectories starting from each milestone and terminate the trajectories when the following or previous milestone is reached. The equations are integrated with a time step of 10^{-5} ps. The results for the potential and the diffusion are reported in Figure 3.

The variances of the estimates of the force and the diffusion coefficients are obtained by propagating the error estimated for the rates. The statistical error for the evaluation of the rate constant is computed as

$$\sigma^2[k(\alpha, \gamma)] = \frac{k^2(\alpha, \gamma)}{N_{\alpha\gamma}} \times \left[\frac{p(\alpha, \gamma)[1 - p(\alpha, \gamma)]}{p(\alpha, \gamma)^2} + \frac{\langle \tau_\alpha^2 \rangle - \langle \tau_\alpha \rangle^2}{\langle \tau_\alpha \rangle^2} \right], \quad (28)$$

where $N_{\alpha\gamma}$ is the number of trajectories that start from milestone α and reach milestone γ before any other milestone.

In Eq. (28), we assume that $p(\alpha, \gamma)$ is binomial. The errors are then propagated to the diffusion and the force. A normal distribution of the errors was assumed in an earlier study⁴⁶ which is an approximation to the expression derived here.

The force is recovered with high accuracy (Figure 3 top). The estimate of the space-dependent diffusion coefficient slightly overestimate the extrema of the function (Figure 3 bottom).

B. Test 2

We also simulated two-dimensional models. For each model, we run one long trajectory of 10^{11} steps with a time step of $10^{-5}\tau$ (τ is an arbitrary unit of time, which is set here to one), saving the configurations every 10 steps. The analysis is then performed using 11 milestones that are used to assign the long trajectory to the milestones as a function of time and to estimate the rate coefficients. In both cases, the length of the

periodic box is of 2.2λ (λ is an arbitrary unit of length which is set here to one).

In the first case, the following two dimensional diffusion tensor and potential energy are used:

$$\begin{aligned}\beta U(x, y) &= \beta U_0 \cos\left[\frac{2\pi}{L}\left(x - \frac{L}{2}\right)\right] \sin\left[\frac{2\pi}{L}\left(y - \frac{L}{2}\right)\right], \\ \hat{D}(x, y) &= \left\{ D_0 + D_1 \cos\left[\frac{2\pi}{L}\left(x - \frac{L}{2}\right)\right] \right\} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.\end{aligned}\quad (29)$$

Here, the off-diagonal terms of the diffusion tensor are zero, and the diagonal ones are equal. The parameters βU_0 , D_0 , and D_1 are set equal to 0.5, $0.03\lambda^2/\tau$, and $0.01\lambda^2/\tau$. The results are reported in Figures 4–6. The statistical errors are derived from the rate coefficients (Eq. (28)). There is also an error in the estimate of the displacement in the x - or y -direction in the definition of the off-diagonal term of the diffusion tensor. In Eq. (24), the contribution coming from the displacement in x is exact. However, the displacement in y lies in the interval $[(\gamma_y - \alpha_y - 1)\Delta_y, (\gamma_y - \alpha_y + 1)\Delta_y]$, as it depends on the actual crossing points on milestones $\vec{\alpha}$ and $\vec{\gamma}$. The resulting error was computed as if the distribution of distances is uniform, and it is $(\Delta_y/2)^2/3N_{\alpha\beta}$. A similar reasoning holds for Eq. (25).

Figures 4 and 5 show a comparison between the input forces and the result of the analysis to recover them. The statistical errors are roughly the size of the dots. Figures 5 and 6 show the different components of the diffusion tensor. The results are less noisy than those in Figures 4 and 5(a)–5(c). However, the simulated space-dependent diffusion coefficient is slightly under-estimated. The possible causes of this discrepancy are discussed in the end of Sec. III. The shapes are recovered quite well, and the off-diagonal term, which is expected to be zero, is found to be more than an order of magnitude smaller than the diagonal terms.

C. Test 3

In the second two-dimensional test case, we use the same potential as in the first case (Eq. (29)) but a different diffusion tensor. The off-diagonal terms of the diffusion tensor are similar to the diagonal terms. We use a $\hat{b}(x, y)$ tensor with the following components:

$$\begin{aligned}b_{xx}(x, y) &= \sqrt{D_0 + D_1 \sin\left[\frac{2\pi}{L}\left(x - \frac{L}{2}\right)\right]}, \\ b_{yy}(x, y) &= \sqrt{D_0 + D_1 \sin\left[\frac{2\pi}{L}\left(y - \frac{L}{2}\right)\right]}, \\ b_{xy}(x, y) &= b_{yx}(x, y) \\ &= \sqrt{\alpha_0 D_0 + \alpha_1 D_1 \sin\left[\frac{2\pi}{L}\left(x - \frac{L}{2}\right)\right]},\end{aligned}\quad (30)$$

with the parameters α_0 , α_1 , D_0 , and D_1 set equal to 0.25, 0.125, $0.03\lambda^2/\tau$, and $0.01\lambda^2/\tau$. The diffusion tensor $\hat{D}(x, y)$ is obtained from $\hat{b}(x, y)$ as in Eq. (2). The results are reported in Figures 7–10.

FIG. 13. Comparison between the forces obtained following the algorithm described in Sec. II (blue dots, with error bars smaller than the size of the dot) and the forces obtained by numerical differentiation of the free energy (red line). (a) The forces are reported as a function of ψ , with $\phi = -115^\circ$. Top: ϕ -component of the force. Middle: ψ -component of the force. Bottom: Probability of being at $\phi = -115^\circ$ as a function of ψ . (b) Same as 13(a). The forces are reported as a function of ϕ , with $\psi = -65^\circ$. (c) Same as 13(a). The forces are reported as a function of ϕ , with $\psi = 125^\circ$.

Figures 7 and 8 show the two components of the force vector. The quality of the expected value is the same as in Figures 4 and 5, the recovery of the two-dimensional surface is noisier, but the results are again within the error bars. Figures 9 and 10 show the different components of the diffusion

FIG. 14. (a) $\phi\phi$ -component of the diffusion tensor extracted using the algorithm described in Sec. II. The values reported are only those corresponding to regions of dihedral space that are visited with probability $p > 0.0001$. (b) Standard deviation of the estimate of the $\phi\phi$ -component of the diffusion tensor. Note that the scale for the error is around 1/10 of the scale of the average.

tensor. The shape of each of the elements of the diffusion tensor is recovered even though, as highlighted from the one-dimensional sections of the two-dimensional surfaces (Figures 9(b), 9(d) and 10(b)–10(e)). The deviations at the maxima and the minima exceed the statistical error.

The discrepancies between the expected diffusion tensor and the result of our method might be due, at least partially, to the accuracy of the integrator and to the spacing between the milestones. To test this hypothesis, we evaluated and reported in Figure 11 the dependence of the error in the evaluation of $D_{xx}(x, y)$ at $x = -0.6$ and $y = 0.0$ from

- The size of the milestone in the y direction, Δ_y (Figure 11(a)).
- The distance between two milestones in the x direction, Δ_x (Figure 11(b)).
- The time step (Figure 11(b)).

FIG. 15. (a) $\psi\psi$ -component of the diffusion tensor extracted using the algorithm described in Sec. II. The values reported are only those corresponding to regions of dihedral space that are visited with probability $p > 0.0001$. (b) Standard deviation of the estimate of the $\psi\psi$ -component of the diffusion tensor. Note that the scale for the error is around 1/10 of the scale of the average.

The results in Figure 17 are obtained by starting the trajectories at the milestone $\{(x, y) | x = 0.6 \text{ and } -\frac{\Delta_y}{2} \leq y < \frac{\Delta_y}{2}\}$. The starting point in y is randomly drawn from a uniform distribution. Each point is computed running 10^4 trajectories, with the exception of the first and second points in Figure 11(a), that are obtained from 3×10^4 and 10^5 runs, respectively. As seen in Figure 11(a), the dependence on the size of the milestone Δ_y is weak. Reducing the time step does not give a significant improvement (Figure 11(b), blue dots). Finally, reducing the spacing between the milestones Δ_x and the time step (Figure 11(b), red squares) are closer to the exact value. Hence, it is possible to obtain highly accurate results if significantly larger computational resources are used. We believe that the difficulties we experience in reaching high-level accuracy are associated with errors in the algorithm for Brownian dynamics. The results are better for a molecular simulation, a much more complex system.

FIG. 16. (a) $\phi\psi$ -component of the diffusion tensor extracted using the algorithm described in Sec. II. The values reported are only those corresponding to regions of dihedral space that are visited with probability >0.0001 . (b) Standard deviation of the estimate of the $\phi\psi$ -component of the diffusion tensor. Note that the scale for the error is around 1/10 of the scale of the average.

IV. MOLECULAR DYNAMICS APPLICATION: ALANINE DIPEPTIDE

In this section, we illustrate the method presented in this paper for the case of alanine dipeptide. We ran four MD trajectories of a solvated alanine dipeptide of 60 ns each. The periodic box is of size 20 Å, the total number of water molecules is 248. The simulations are in the canonical ensemble, fixing the temperature by velocity-rescaling.⁴⁷ The force field is united atom OPLS (Optimized Potential for Liquid Simulations)⁴⁸ for the peptide, Transferable Intermolecular Potential 3P⁴⁹ for water.

During each of the four runs, we collected the values of the ϕ - and ψ -dihedral angles. From the time traces of the dihedral angles, we extracted the components of the force in the ϕ - and ψ -directions, and the three components of the diffusion tensor ($\phi\phi$, $\phi\psi$, and $\psi\psi$) using the algorithm presented in Sec. II. The milestones were placed every 10° , for a total of 36 milestones

per dihedral angle. We also computed the probability distribution in ϕ, ψ -space by computing the probability of being in each of the 36^2 cells, each of size 10° .

The results from the four simulations were averaged using as a weight the inverse of the standard deviation computed within each simulation. If a milestone (ϕ_α, ψ_α) was not visited in a simulation, its contribution to the average was neglected. If a milestone (ϕ_α, ψ_α) was visited only once in a simulation, an error to the estimate of an observable on that milestone could not be computed. In that case, an error was assigned based on the other simulations. This means that if in simulation i milestone (ϕ_α, ψ_α) was visited only once, we computed the largest difference between the value of the observable on the milestone (ϕ_α, ψ_α) for simulation i and all the other simulations $j \neq i$, plus (or minus) the standard deviation computed for simulation j (if available). The largest distance was used as error.

The average probability distribution $P(\phi, \psi)$ for alanine dipeptide is computed directly from 60 ns MD trajectories. We use the distribution to extract the ϕ and ψ components of the force, $\beta F_\phi^B(\phi, \psi)$ and $\beta F_\psi^B(\phi, \psi)$, by numerical differentiation of the free energy $\beta \mathcal{F}(\phi, \psi) = -\ln P(\phi, \psi)$. We compared this to the estimate of the ϕ and ψ components of the forces extracted using the algorithm presented in Sec. II to verify the Milestoning analysis. The forces are shown in Figure 12.

In Figure 12(a), it is possible to see a white path connecting the three minima that closely resembles a minimum free energy path. In Figure 12(c), the three minima are clearly highlighted between red (positive) and blue (negative) regions.

In Figures 12(b) and 12(d), we report the difference between the average of the data extracted using the algorithm described in Sec. II and the force obtained by numerical differentiation of the free energy. To be clearly visible, the difference is reported on a scale that is around 1/3 of the scale of the signal. The largest deviations are concentrated at the boundary of the reported region. Those are the regions that are visited less frequently.

Figures 13(a)–13(c) show one-dimensional profiles of the forces as a function of one of the dihedral angles. Figure 13(a) shows the profile as a function of ψ with ϕ fixed at 125° , highlighting the forces in the proximity of the minimum free energy path that connects the three minima of alanine dipeptide. Figures 13(b) and 13(c) instead are shown as a function of ϕ , with ψ fixed at a position of the two deepest minima in alanine dipeptide dihedral free energy map. In all of these circumstances, the agreement between the two estimates of the forces is excellent.

In Figures 14–16 the three independent components of the diffusion tensor are reported.

From Figures 14(a) and 15(a) it is clear that the diagonal components of the diffusion tensor do not change significantly as a function of the coordinates, aside from the boundaries of the reported region, where the error tends to be larger (see Figures 14(b) and 15(b)). On the other hand, the off-diagonal element, reported in Figure 16(a), is smaller in magnitude, but shows more substantial coordinate dependence, going from positive to negative values as ϕ increases. Even in this case, the larger error is concentrated at the boundary of the region (see Figure 16(b)).

FIG. 17. In each panel: comparison between different components of the diffusion tensor. As red dots we show $D_{\phi\phi}$, as green squares $D_{\psi\psi}$, and as blue triangles $D_{\phi\psi}$. Bottom: Equilibrium probability distribution of the corresponding angle. Top left: Component as a function of ϕ , with $\psi = 125^\circ$. Top right: With a fixed $\psi = -65^\circ$. Bottom: With a fixed $\psi = 125^\circ$ as a function of ϕ . The error bars are smaller than the size of the symbols.

As we did for the different components of the forces, we show three one-dimensional profiles of the three components of the diffusion tensor.

Figure 17 top left panel shows the profile of the different components of the diffusion tensor as a function of ψ , with $\phi = -115^\circ$, around the minimum free energy path for alanine dipeptide. In red we report $D_{\phi\phi}$, in green $D_{\psi\psi}$, and in blue the off-diagonal term $D_{\phi\psi}$. The $\phi\phi$ -component (red circles) is almost completely flat throughout this region. The $\psi\psi$ -component (green squares) shows two shallow minima around the peak of the probability distribution (see bottom of the panels of Figure 17). The off-diagonal term (blue triangles) changes more significantly in this region, it is mostly negative, and its absolute value is less than half of the value of the diagonal components. Note also that the values of the $\psi\psi$ -component of the diffusion tensor obtained here are similar to those reported by Hummer,²² who computed the diffusion for alanine dipeptide as a function of ψ in one-dimension with a different potential. The diagonal terms of the diffusion tensor as a function of ψ around the two deepest maxima of alanine dipeptide free energy map (Figure 17) are almost flat, with a sharp increase around the boundary of the domain we investigated. This might be an effect of poor sampling at the edge. The off-diagonal term changes sign from positive to negative when ϕ increases. Aside for the boundaries of the reported region, its

absolute value is smaller than the values of the diagonal terms of the diffusion tensor.

V. CONCLUSIONS

In the present paper, we report a novel method to extract the space dependent diffusion tensor from the time traces of MD simulations either directly or by Milestoning. The method is based on the connection between the master equation and the Fokker-Planck equation provided by the Kramers-Moyal expansion of a discrete master equation. Using this expansion, we connect the rates in the master equation with the drift vector and diffusion tensor in the Fokker-Planck equation. A Milestoning-type data analysis was carried out to extract the rates from a Brownian dynamics trajectory. The results showed that the method is capable of reproducing accurately both one-dimensional and two-dimensional test cases. The method was also tested on a MD simulation of alanine dipeptide. The forces extracted with this method closely reproduced those obtained from the numerical derivatives of the free energy (log of the probability). The diagonal elements of the diffusion tensor were found to be almost coordinate independent. The off-diagonal element, which is smaller than diagonal elements, changes significantly as a function of both the dihedral angles.

The use of Milestoning has a significant promise for future applications for significantly more complex systems as Milestoning can be applied efficiently to systems that are hard to sample with straightforward MD.

The method that we developed is based upon the following hypotheses:

1. We assumed that the dynamics of the CV is overdamped.
2. The expansion and truncation to the second order of the master equation provide the connection between the rates and the elements of the diffusion tensor. We assume that these approximations are accurate.
3. We assumed that Milestoning provides a reliable estimate of the rates. In the low dimension test cases, the use of Milestoning is exact (exploiting a long trajectory in two dimensions). The use of approximate Milestoning can be advantageous since it is significantly more efficient.¹⁶ Considerable previous investigations formulated the conditions under which Milestoning is accurate, even if it is not formally exact.^{16,50} Recently, an extension to Milestoning was introduced that makes this method formally exact (provided that the statistics is sufficient).⁵²
4. We assume that the approaches to place milestones in one and higher dimensions allow a good estimate of the components of diffusion tensor.

In the recovery of the diffusion tensor from overdamped trajectories, we meet condition (1) within the accuracy of the Euler-Maruyama algorithm. The accuracy of the method was shown to depend on the time step and the spacing between milestones (see Figure 11(b)). In the case of MD simulations of alanine dipeptide, we found that an overdamped description of the dynamics in dihedral space was accurate. Indeed, the numerical derivatives of the potential of mean force, computed by other means, are remarkably similar to the mean force we estimated. This observation suggests that the estimate of the element of the diffusion tensor is accurate as well.

We provide a numerical illustration of the feasibility of the method in different cases, showing that hypotheses (2), (3), and (4) are met in concrete examples.

The availability of the method hereby presented allows new and interesting avenues of analysis of MD trajectories with Milestoning. In particular, we are in the process of using this method to extract the different components of the diffusion tensor when we follow two or more degrees of freedom in a coarse description of a biopolymer's dynamics. In some circumstances, clear evidence of the need of higher dimensionality reaction coordinate was reported.²⁹ The off-diagonal terms of the diffusion tensor will provide us with a description of the coupling between important coarse variables at the overdamped limit.

The reliability of this method in the analysis of a trajectory resulting from an MD simulation of a large biopolymer is yet to be tested. In particular, hypothesis (1) depends on a proper choice of variables of interest. A general discussion of what would be variables that meet this hypothesis is likely to be system dependent and is beyond the scope of this paper.

ACKNOWLEDGMENTS

Discussions with the lab-members Juan M. Bello-Rivas and Michele di Pierro are gratefully acknowledged. This research was supported by NIH Grant No. GM59796 and Welch Grant No. F-1783 to R.E.

- ¹B. Schuler and W. A. Eaton, *Curr. Opin. Struct. Biol.* **18**(1), 16–26 (2008).
- ²J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, *Science* **296**(5574), 1832–1835 (2002).
- ³J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante, *Science* **292**(5517), 733–737 (2001).
- ⁴C. Bustamante, Y. R. Chemla, N. R. Forde, and D. Izhaky, *Annu. Rev. Biochem.* **73**, 705–748 (2004).
- ⁵A. Soranno, B. Buchli, D. Nettels, R. R. Cheng, S. Muller-Spath, S. H. Pfeil, A. Hoffmann, E. A. Lipman, D. E. Makarov, and B. Schuler, *Proc. Natl. Acad. Sci. U. S. A.* **109**(44), 17800–17806 (2012).
- ⁶H. Lannon, J. S. Haghighpanah, J. K. Montclare, E. Vanden-Eijnden, and J. Brucic, *Phys. Rev. Lett.* **110**(12), 128301 (2013).
- ⁷R. Zwanzig, *Phys. Rev.* **124**(4), 983 (1961).
- ⁸H. Mori, *Prog. Theor. Phys.* **33**(3), 423 (1965).
- ⁹R. Zwanzig, *J. Stat. Phys.* **30**(2), 255–262 (1983).
- ¹⁰R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford, New York, 2001).
- ¹¹H. Yamakawa, G. Tanaka, and W. H. Stockmayer, *J. Chem. Phys.* **61**(11), 4535–4539 (1974).
- ¹²B. J. Berne, M. E. Tuckerman, J. E. Straub, and A. L. R. Bug, *J. Chem. Phys.* **93**(7), 5084–5095 (1990).
- ¹³J. E. Straub, M. Borkovec, and B. J. Berne, *J. Phys. Chem.* **91**(19), 4995–4998 (1987).
- ¹⁴A. Carof, R. Vuilleumier, and B. Rotenberg, *J. Chem. Phys.* **140**(12), 124103 (2014).
- ¹⁵A. K. Faradjian and R. Elber, *J. Chem. Phys.* **120**(23), 10880–10889 (2004).
- ¹⁶A. M. A. West, R. Elber, and D. Shalloway, *J. Chem. Phys.* **126**(14), 145104 (2007).
- ¹⁷R. F. Pawula, *Phys. Rev.* **162**(1), 186–188 (1967).
- ¹⁸H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed. (Springer-Verlag, New York, 1996).
- ¹⁹P. Hanggi and H. Thomas, *Phys. Rep.* **88**(4), 207–319 (1982).
- ²⁰M. Hinczewski, Y. von Hansen, J. Dzubiella, and R. R. Netz, *J. Chem. Phys.* **132**(24), 245103 (2010).
- ²¹T. B. Woolf and B. Roux, *J. Am. Chem. Soc.* **116**(13), 5916–5926 (1994).
- ²²G. Hummer, *New J. Phys.* **7**, 34 (2005).
- ²³C. Schutte, F. Noe, J. F. Lu, M. Sarich, and E. Vanden-Eijnden, *J. Chem. Phys.* **134**(20), 204105 (2011).
- ²⁴D. J. Bicout and A. Szabo, *J. Chem. Phys.* **109**(6), 2325–2338 (1998).
- ²⁵J. Comer, C. Chipot, and F. D. Gonzalez-Nilo, *J. Chem. Theory Comput.* **9**(2), 876–882 (2013).
- ²⁶S. Turkcan, A. Alexandrou, and J. B. Masson, *Biophys. J.* **102**(10), 2288–2298 (2012).
- ²⁷J. Comer, K. Schulten, and C. Chipot, *J. Chem. Theory Comput.* **10**(2), 554–564 (2014).
- ²⁸J. Comer, K. Schulten, and C. Chipot, *J. Chem. Theory Comput.* **10**(7), 2710–2718 (2014).
- ²⁹A. E. Cardenas and R. Elber, *Mol. Phys.* **111**(22–23), 3565–3578 (2013).
- ³⁰Z. Schuss, *Theory and Applications of Stochastic Processes: An Analytical Approach* (Springer, New York, 2009).
- ³¹C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, 3rd ed. (Springer-Verlag, Berlin, New York, 2004).
- ³²M. Sarich, F. Noe, and C. Schutte, *Multiscale Model. Simul.* **8**(4), 1154–1177 (2010).
- ³³F. Noe, C. Schutte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl, *Proc. Natl. Acad. Sci. U. S. A.* **106**(45), 19011–19016 (2009).
- ³⁴F. Noe, I. Horenko, C. Schutte, and J. C. Smith, *J. Chem. Phys.* **126**(15), 155102 (2007).
- ³⁵J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope, *J. Chem. Phys.* **126**(15), 155101 (2007).
- ³⁶N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, Rev. and enl. ed. (North-Holland, Amsterdam, New York, 1992).
- ³⁷H. A. Kramers, *Physica* **7**, 284–304 (1940).

- ³⁸J. E. Moyal, J. R. Statist. Soc. B **11**(2), 150–210 (1949).
- ³⁹S. Kirmizialtin and R. Elber, *J. Phys. Chem. A* **115**(23), 6137–6148 (2011).
- ⁴⁰G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**(2), 187–199 (1977).
- ⁴¹E. Vanden-Eijnden and M. Venturoli, *J. Chem. Phys.* **130**(19), 194101 (2009).
- ⁴²P. Májek and R. Elber, *J. Chem. Theory Comput.* **6**(6), 1805–1817 (2010).
- ⁴³A. T. Hawk, *J. Chem. Phys.* **138**(15), 154105 (2013).
- ⁴⁴D. L. Ermak and J. A. McCammon, *J. Chem. Phys.* **69**(4), 1352–1360 (1978).
- ⁴⁵P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, New York, 1992).
- ⁴⁶S. Kreuzer and R. Elber, *J. Chem. Phys.* **139**, 121902 (2013).
- ⁴⁷D. M. Heyes, *Chem. Phys.* **82**(3), 285–301 (1983).
- ⁴⁸W. L. Jorgensen and J. Tiradorives, *J. Am. Chem. Soc.* **110**(6), 1657–1666 (1988).
- ⁴⁹W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**(2), 926–935 (1983).
- ⁵⁰E. Vanden-Eijnden, M. Venturoli, G. Ciccotti, and R. Elber, *J. Chem. Phys.* **129**(17), 174102 (2008).
- ⁵¹M. L. Mugnai, Ph.D. thesis, The University of Texas at Austin, 2014.
- ⁵²J. M. Bello-Rivas and R. Elber, “Exact milestoning,” *J. Chem. Phys.* (submitted).