UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Dynamical Reweighting for Biased Rare Event Simulations

Keller, B.G.; Bolhuis, P.G.

DOI
10.1146/annurev-physchem-083122-124538
Publication date
2024
Document Version
Final published version
Published in
Annual Review of Physical Chemistry

License
CC BY

Link to publication

Citation for published version (APA):
Keller, B. G., & Bolhuis, P. G. (2024). Dynamical Reweighting for Biased Rare Event Simulations. *Annual Review of Physical Chemistry*, *75*, 137-162. https://doi.org/10.1146/annurev-physchem-083122-124538

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:09 Sep 2024

ANNUAL REVIEWS

Annual Review of Physical Chemistry

Dynamical Reweighting for Biased Rare Event Simulations

Bettina G. Keller[1] and Peter G. Bolhuis[2]

[1]Department of Biology, Chemistry and Pharmacy, Freie Universität Berlin, Berlin, Germany; email: bettina.keller@fu-berlin.de

[2]Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, The Netherlands

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Phys. Chem. 2024. 75:137–62

The *Annual Review of Physical Chemistry* is online at physchem.annualreviews.org

https://doi.org/10.1146/annurev-physchem-083122-124538

Copyright © 2024 by the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.

OPEN ACCESS CC BY

Keywords

molecular simulation, dynamical reweighting, path reweighting, enhanced sampling, rare events, potential energy surface

Abstract

Dynamical reweighting techniques aim to recover the correct molecular dynamics from a simulation at a modified potential energy surface. They are important for unbiasing enhanced sampling simulations of molecular rare events. Here, we review the theoretical frameworks of dynamical reweighting for modified potentials. Based on an overview of kinetic models with increasing level of detail, we discuss techniques to reweight two-state dynamics, multistate dynamics, and path integrals. We explore the natural link to transition path sampling and how the effect of nonequilibrium forces can be reweighted. We end by providing an outlook on how dynamical reweighting integrates with techniques for optimizing collective variables and with modern potential energy surfaces.

1. INTRODUCTION

1.1. Rare Events in Molecular Systems

Rare events: occur infrequently, such that, on average, the system has ample time to explore the current state before the event occurs

Collective variable (CV): any useful function of the coordinates; it does not need to follow the committor

Reaction coordinate: a CV that accurately describes the progress of the reactive process; it should follow the committor closely

Understanding rare events is crucial in molecular systems. Most chemical reactions, even catalyzed ones, are rare events on the microscopic scale. Likewise, conformational transitions, molecular associations, binding processes, phase transitions, and nucleation in materials involve rare events. Transition state theory (TST) (1) and Kramers theory (2) demonstrate that the mean first passage time, which indicates the waiting time, exponentially increases with energy barrier height between the initial and final states of a rare event. Simulating rare events is challenging because the timescale of molecular rare events are often well beyond timescales accessible by direct molecular dynamics (MD) simulations. Even if occasional rare event transitions are observed, thermodynamic or kinetic properties estimates may lack statistical significance. Furthermore, the dynamics in the fast degrees of freedom influence free energies and diffusion profiles of rare event transitions, and rare events can comprise multiple separate transition paths with various intermediate and transition states. In short, sampling all relevant pathways that contribute to a rare event is vital.

Enhanced sampling techniques have been developed mostly in the context of empirical potential energy functions, with applications evolving around biomolecular processes, structural changes in materials, and transport processes. Recent years have witnessed dramatic advances in the development of molecular potential energy functions. Examples are reactive and polarizable force fields (3), advances in quantum-chemical potential energies (4), and, most notably, neural network potentials that use deep neural networks to parameterize potential energies with quantum-chemical accuracy (5, 6). With these more accurate and more versatile potential energy functions, one can now sample chemical reactions and the influence of the environment on these reactions, as well as processes at high temperature and pressure, where the approximations of empirical force fields break down.

However, reactions are associated with extremely high free-energy barriers, often one to two orders of magnitude higher than the thermal energy. Thus, direct simulations of these processes are numerically prohibitive. Instead, one relies on enhanced sampling techniques that might distort the dynamics; therefore, the simulations need to be corrected by dynamical reweighting. Whether one can harvest the full benefit of modern molecular potentials depends critically on the accuracy and the usability of the dynamical reweighting method.

1.2. Enhanced Sampling and Reweighting

Enhanced sampling techniques increase the frequency of rare events in a simulation by modifying the system's dynamics and can be categorized into three main groups: (*a*) temperature changes, (*b*) modifications to the potential energy function, and (*c*) adjustments to the initial simulation state (for a comprehensive review of these different approaches, see 7). Enhanced sampling methods often make use of collective variables (CVs). If the reaction coordinate of the rare event is poorly represented by the CVs, one might severely misestimate the rare event rate, and even arrive at a completely incorrect understanding of the rare event's mechanism (8, 9) (see the **Supplemental Material**). In this review, we assume that the reaction coordinate is reasonably well represented by the chosen CVs and that we do not have to worry about this problem.

Supplemental Material >

Stationary reweighting techniques that recover free energies from enhanced sampling simulations of the first two types have been well established (7, 10, 11). With enhanced sampling and stationary reweighting, the statistical certainty of thermodynamic estimates is drastically increased in comparison to direct simulations. Therefore, thermodynamic properties are now routinely calculated using this approach.

In contrast, retrieving dynamical properties from enhanced sampling simulations is much more challenging and is a very active field of research (12–14). Dynamical reweighting has been best developed for enhanced sampling simulations of the third type that adjust the initial simulation state. Such methods preserve the underlying true dynamics but alter the probability of observing a rare event trajectory. Several classes of trajectory sampling techniques exist. In so-called splitting methods, an existing trajectory is extended forward in time, creating multiple trajectories using stochastic dynamics. Subsequent selection or pruning then enhances the probability of the event of interest. After each split, the daughter trajectories are reweighted to mimic the probability of occurrence in the original ensemble. Examples of such methods include weighted ensemble (15, 16), forward flux sampling (17), and adaptive multilevel splitting (18). Another class of methods that adjust the initial state can be regarded as importance sampling of trajectories. Transition path sampling (TPS) (19, 20) performs a random walk through trajectory space by creating trial paths and accepting or rejecting them using Metropolis–Hastings Monte Carlo. The resulting collection of trajectories converges to the path distribution of interest. Adding constraints on the path distribution yields the probabilities of observing reactive trajectories. Examples of such methods include TPS (20, 21) and the related TIS (transition interface sampling) method (22).

TPS and splitting methods are best suited for rare transitions between two well-defined and a priori known states. The presence of multiple states, unknown states, or hierarchies of free-energy barriers requires different techniques that can sample rare event transitions and explore state space simultaneously. In this sense, enhanced sampling by modifying the potential is complementary to adjusting the initial simulation state, as one not only increases rare event transitions but also enhances state-space exploration. A combination of modified potentials and adjusting the initial state could tackle high free-energy barriers and complex dynamics in emerging applications of MD simulations. To do so, we need dynamical reweighting techniques for enhanced sampling simulations of the second type, that is, with modified potentials. These techniques are the topic of this review.

2. MODELS OF RARE EVENT TRANSITIONS

Many resources (9, 21, 23–25) discuss rate theories in detail. In this section, we provide an overview of the kinetic models used as the foundation for dynamical reweighting methods.

2.1. Collective Variables and Free Energy

Consider a system with N atoms, for which $x \in \Omega \subset \mathbb{R}^{3N}$ is the position vector and Ω is the position or configuration space. $V(x)$ is a potential energy function whose gradient determines the deterministic forces in the system. We assume that the system samples the configurational Boltzmann density:

$$\pi(x) = \frac{1}{Z} e^{-\beta V(x)}. \qquad 1.$$

Here, $\beta = (k_B T)^{-1}$, where k_B is the Boltzmann constant and T is the temperature. $Z = \int_\Omega e^{-\beta V(x)}\, dx$ is the configurational partition function.

Due to the high dimensionality of the full state space, it is more practical to analyze the dynamics using low-dimensional CVs. A CV vector is a (possibly nonlinear) function,

$$q : \Omega \to \Omega_{CV} \subset \mathbb{R}^m, \qquad 2.$$

which maps each position $x \in \Omega$ onto a vector $q \in \Omega_{CV} \subset \mathbb{R}^m$, where $m \ll 3N$. Ω_{CV} is the CV space. The stationary density $\pi(q)$ is the configurational Boltzmann density $\pi(x)$ marginalized to

the CV space:

$$\pi(q) = \frac{1}{Z}\int_{\Omega} \mathrm{d}x\, e^{-\beta V(x)}\delta\left(q(x) - q\right) = \frac{1}{Z}e^{-\beta F(q)}, \tag{3.}$$

where $\delta\left(q(x) - q\right)$ is the Dirac delta function. We define the free energy as

$$F(q) = -\frac{1}{\beta}\ln \pi(q) + C, \tag{4.}$$

where $C = (-1/\beta)\ln Z$ depends on V but is a constant with respect to q. The free energy serves as an effective potential for the CVs. A time-dependent probability density in CV space is denoted $p(q, t)$.

Rate and rate constant: not the same thing (Equation 5 defines their relation); rate is closely related to the concept of flux

2.2. Rare Events and Transition Rates

We define two regions in position space, $A \subset \Omega$ and $B \subset \Omega$, and a one-dimensional reaction coordinate $q : \Omega \to \mathbb{R}$, which connects them. The reaction $A \xrightarrow{k_{AB}} B$ is modeled by a unimolecular (phenomenological) rate equation:

$$\mathrm{rate}(t) = -\frac{\mathrm{d}}{\mathrm{d}t}p_A(t) = k_{AB}\, p_A(t). \tag{5.}$$

Here, $p_A(t) = \int_A p(q, t)\mathrm{d}q$ is the relative population in state A at time t. The rate of the reaction is the decrease in p_A per infinitesimally small time interval $\mathrm{d}t$, which is proportional to the current relative population in A. The proportionality constant k_{AB} is the rate constant.

The rate constant can be calculated as $k_{AB} = \langle \tau_{AB}\rangle^{-1}$, where the mean first passage time $\langle \tau_{AB}\rangle$ can be estimated from a (large) set of N_{fpt} first passage times as

$$\langle \tau_{AB}\rangle = \lim_{N_{\mathrm{fpt}}\to\infty}\frac{1}{N_{\mathrm{fpt}}}\sum_{i=1}^{N_{\mathrm{fpt}}}\tau_{AB,i}. \tag{6.}$$

A state A is referred to as metastable if the time the system spends on average in A is long enough to fully sample $\pi(q)$ within A. A rare event (i.e., a transition between two metastable states A and B) arises if the two states are separated by a large free-energy barrier, leading to a timescale separation between the barrier crossing time and the residence time in the states. The maximum of this free-energy barrier is called the transition state.

2.3. Transition State Theory and Flux Through a Dividing Surface

TST provides a formula for k_{AB}. Define an indicator function of state A:

$$h_A(q) = \begin{cases} 1 & \text{if } q \in A \\ 0 & \text{otherwise.} \end{cases} \tag{7.}$$

The indicator function of state B, $h_B(q)$, is defined analogously. We further assume that state A and B are adjacent and that the dividing surface $q(x) = q^{\ddagger}$ between the two states is close to the transition state. The directed probability flux is the amount of probability density per time interval $\mathrm{d}t$ that passes from A to B through $q^{\ddagger}$ (**Figure 1*a***). For a two-state system, any decrease in p_A is caused by this flux across $q^{\ddagger}$. In this case, flux and rate are equal ($\mathrm{flux}_{AB}(t) = \mathrm{rate}(t)$), and

$$k_{AB} = \frac{\mathrm{flux}_{AB}(t)}{p_A(t)}. \tag{8.}$$

Since we assume that k_{AB} does not depend on the specific probability density $p(q, t)$ at time t, we can discuss Equation 8 for the case of thermal equilibrium, that is, for $p(q, t) = \pi(q)$. In this

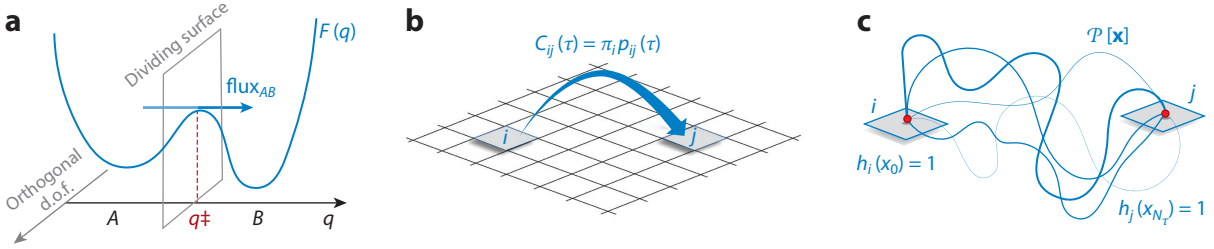

Figure 1

Models of rare event transitions. (*a*) Transition state theory as flux$_{AB}$ through a dividing surface. (*b*) Absolute and conditional transition probabilities in discretized state space. (*c*) Paths **x** of length τ from Ω_i to Ω_j. Line thickness represents the path probability $\mathcal{P}[\mathbf{x}]$. The starting points for dynamic reweighting are flux$_{AB}$, $C_{ij}(\tau) = \pi p_{ij}(\tau)$, and $\mathcal{P}[\mathbf{x}]$. Figure adapted from Reference 14.

case, one can derive explicit expressions for flux$_{AB}$ and $p_A(t)$ to obtain an equation for k_{AB} within the TST framework:

$$k_{AB}^{\mathrm{TST}} = \frac{1}{2}\langle|\dot{q}|\rangle_{q=q^\ddagger} \cdot \frac{Z^\ddagger}{Z_A}, \qquad 9.$$

where $\langle\ldots\rangle_{q=q^\ddagger}$ denotes an ensemble average with respect to the stationary density on the transition state $\pi^\ddagger = \delta(q - q^\ddagger)e^{-\beta F(q)}/Z^\ddagger$. Z_A and $Z^\ddagger$ are the partition functions of state A and the transition state (for a derivation, see the **Supplemental Material**).

Supplemental Material >

Equation 9 is closely related to the Eyring–Polanyi equation, in which the rate is derived by assuming an equilibrium between A and the transition state, with an equilibrium constant $K^\ddagger = Z^\ddagger/Z_A$. The rate constant is then $k_{AB}^{\mathrm{TST}} = \kappa\nu^\ddagger K^\ddagger$, where $\nu^\ddagger$ is the decay rate of the transition state. In Equation 9 the decay rate is represented by the factor $\langle|\dot{q}|\rangle_{q=q^\ddagger}$, whereas in the Eyring–Polanyi theory $\nu^\ddagger$ is approximated by the universal frequency factor $\nu^\ddagger \approx (\beta h)^{-1}$, with Planck constant h. κ is the transmission coefficient, which accounts for the fact that not all systems that reach the transition state proceed to B (some instead return to A). The original formulation of the Eyring–Polanyi equation was developed for gas-phase reactions, and one assumed that any system that reaches the transition state would proceed to B; thus, κ is set to unity. In thermal equilibrium, equally many systems move from A to B as in the opposite direction, leading to $\kappa = 1/2$. Many real systems exhibit recrossings; that is, a system crosses $q^\ddagger$ but returns to A after a brief excursion into B (without committing to B). By including these recrossings in the flux, one overestimates the rate constant. To compensate, one can set κ to values below 1/2. In fact, κ can become low if the CV is not aligned with the reaction coordinate.

2.4. Multistate Dynamics and Markov State Models

While many systems can be accurately modeled by a single rare event transition between two states, most molecular systems exhibit a rugged potential energy surface with multiple metastable states and a hierarchy of (free-)energy barriers between them. These systems are best modeled by considering effective dynamics in the CV space. Under certain assumptions, which are not too restrictive (25), the time-dependent probability density $p(q, t)$ evolves according to a Fokker–Planck equation in the CV space:

$$\frac{\partial}{\partial t}p(q,t) = \mathcal{Q}p(q,t), \qquad 10.$$

where $\mathcal{Q}$ is a Fokker–Planck operator. Metastable states and the transition rate constants between them can be obtained by analyzing the eigenfunctions and eigenvalues of $\mathcal{Q}$ (24, 26).

Computationally, the eigenfunctions and eigenvectors are obtained by discretizing Equation 10. One discretizes the CV space Ω_{CV} into n nonoverlapping states Ω_i. The grid

can be highly nonregular. Projecting the time-dependent probability density $p(q, t)$ onto this grid yields a vector $\mathbf{p}(t)$: $p_i(t) = \int_{\Omega_i} p(q,t)\,\mathrm{d}q$. This vector evolves in time according to a master equation (compare with Equation 5):

Row-normalized rate matrices: Equation 13 defines the rate matrix **K** as row normalized.

$$\frac{\mathrm{d}}{\mathrm{d}t}p_i(t) = \sum_{j\neq i} k_{j\to i}\,p_j(t) - k_{i\to j}\,p_i(t), \qquad 11.$$

where $k_{i\to j} = k_{ij}$ is the transition rate constant from state Ω_i to state Ω_j. In matrix vector notation, Equation 11 reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{p}^\top(t) = \mathbf{p}^\top(t)\mathbf{K}, \qquad 12.$$

where $\mathbf{p}^\top(t)$ denotes the transpose of $\mathbf{p}(t)$ and the rate matrix **K** has elements

$$k_{ij} = \begin{cases} k_{ij} & \text{if } i \neq j, \text{ and } \Omega_i \text{ adjacent to } \Omega_j, \\ 0 & \text{if } i \neq j, \text{ and } \Omega_i \text{ not adjacent to } \Omega_j, \\ -\sum_{j\neq i} k_{ij} & \text{if } i = j. \end{cases} \qquad 13.$$

K is the discretized Fokker–Planck operator $\mathcal{Q}$.

Formally integrating Equation 12 yields the propagator of the dynamics:

$$\mathbf{p}^\top(t+\tau) = \mathbf{p}^\top(t)\,\exp(\tau\mathbf{K}) = \mathbf{p}^\top(t)\,\mathbf{P}(\tau), \qquad 14.$$

where $\mathbf{P}(\tau)$ is the (row-normalized) transition matrix. Its elements $p_{ij}(\tau) = \mathbb{P}[q(t+\tau) \in \Omega_j | q(t) \in \Omega_i]$ represent the conditional probability of finding the system in state Ω_j at time $t+\tau$, given that it has been in state Ω_i at time t. τ is called the lag time. The transition probabilities are normalized for all rows i as

$$\sum_{j=1}^{n} p_{ij}(\tau) = 1 \qquad \forall i. \qquad 15.$$

Kinetic models that construct either a rate matrix **K** or a transition matrix $\mathbf{P}(\tau)$ are called Markov state models (MSMs) (24, 26–28). Because the underlying grid is defined in a multidimensional space Ω_{CV}, MSMs can accurately model multistate dynamics. Because of Equation 14, **K** and $\mathbf{P}(\tau)$ share the same eigenvectors, and their eigenvalues are connected by a similarly simple relation. These eigenvectors and eigenvalues are approximations of the eigenfunctions and eigenvalues of the Fokker–Planck operator $\mathcal{Q}$. Specifically, the stationary distribution $\boldsymbol{\pi}$ is given by $\boldsymbol{\pi}^\top\mathbf{K} = 0$ or, equivalently, $\boldsymbol{\pi}^\top\mathbf{P}(\tau) = \boldsymbol{\pi}^\top$. The elements of the stationary distribution vector are $\pi_i = \int_{\Omega_i} \pi(q)\,\mathrm{d}q$.

In thermal equilibrium, the absolute probability for an Ω_i-to-Ω_j transition within τ is (**Figure 1*b***)

$$C_{ij}(\tau) = \pi_i p_{ij}(\tau). \qquad 16.$$

Furthermore, no net flux between any pair of states occurs. Therefore,

$$\pi_i p_{ij}(\tau) = \pi_j p_{ji}(\tau) \qquad \forall\, i, j. \qquad 17.$$

Equation 17 is referred to as the detailed-balance condition.

The parameters of an MSM are the transition rate constants k_{ij} or the transition probabilities $p_{ij}(\tau)$. There are two ways to determine these parameters. First, one can assume that the system evolves according to overdamped Langevin dynamics and derive a formula for k_{ij} by analyzing the associated Fokker–Planck equation. Second, one can simulate the system and estimate the parameters from the observed transitions between grid cells. We introduce both approaches in the following subsections.

2.4.1. Square-root approximation of the Smoluchowski diffusion equation. We assume that the system moves according to overdamped Langevin dynamics in a free-energy surface $F(q)$. A position-dependent diffusion profile $D(q)$ accounts for the equilibration time in the orthogonal degrees of freedom and the curvature of q. The Fokker–Planck equation associated with these dynamics is called the Smoluchowski diffusion equation. To analyze this equation, one discretizes Ω_{CV} into grid cells that are so small that the free energy is essentially constant within each cell. The rate constant k_{ij} between adjacent cells Ω_i and Ω_j is then modeled as the density flux through the cell boundary, which allows for the use of the divergence theorem (Gauss's theorem). This is analogous to **Figure 1*a***, except that one calculates flux_{ij} through the cell boundary rather than through the transition state surface. One thus obtains the expression

$$k_{ij} = \frac{D_{ij}}{d_{ij}} \frac{\mathcal{S}_{ij}}{\mathcal{V}_i} \cdot \sqrt{\frac{\pi(q_j)}{\pi(q_i)}} = \frac{D_{ij}}{d_{ij}} \frac{\mathcal{S}_{ij}}{\mathcal{V}_i} \cdot e^{-\beta(F(q_j)-F(q_i))/2}, \qquad 18.$$

where q_i is the center of grid cell Ω_i with volume $\mathcal{V}_i$. d_{ij} is the distance between two cell centers q_i and q_j, and $\mathcal{S}_{ij}$ is the area of the intersecting hypersurface between two adjacent grid cells. D_{ij} represents the diffusion profile on the intersecting surface, which is approximated to be constant and is calculated as $D_{ij} = (D(q_i) + D(q_j))/2$. $\pi(q_i)$ and $\pi(q_j)$ are the stationary densities at the cell centers, and the second equality in Equation 18 arises from inserting Equation 3.

Several variants of Equation 18 have been derived. For regular grids in a one-dimensional CV space, Equation 18 appears in, for example, References 29 and 30. In this case, the prefactor simplifies as $(D_{ij}/d_{ij})(\mathcal{S}_{ij}/\mathcal{V}_i) = D_{ij}/d_{ij}^2$ (for a detailed derivation, see the supplement of 31). Equation 18 has also been derived and analyzed for arbitrary grids in high-dimensional spaces (32–34). Because of the square-root factor in Equation 18, we call this approach the square-root approximation (SqRA) of the Fokker–Planck equation. Maximum caliber (MaxCal) is another way to derive Equation 18 (35). Given a set of states, their population, and a number of external constraints such as state averages and fluxes, the caliber $\sum_{ij} \pi_i k_{ij} \ln k_{ij}$ can be optimized to yield a rate matrix with k_{ij} analogous to Equation 18 (for more details on the MaxCal approach, see Section 5.3).

2.4.2. Maximum likelihood estimate of Markov models. To estimate the transition probabilities $p_{ij}(\tau)$ from a simulation at potential V, one counts the number of transitions $c_{ij}(\tau) = c_{ij}$ from Ω_i to Ω_j within lag time τ in this trajectory. "Transition within lag time τ" means that the system starts in Ω_i at a time t and ends in Ω_j at time $t + \tau$. When exactly the transition occurs, or whether other grid cells are visited in the time window $[t, t + \tau]$, does not play a role. In particular, Ω_i and Ω_j do not have to be adjacent.

The likelihood of an MSM given the simulation data is

$$\mathcal{L}_{\text{MSM}} = \prod_i \prod_j p_{ij}^{c_{ij}}, \qquad 19.$$

where the transition probabilities $p_{ij}(\tau) = p_{ij}$ are the parameters of the model and $c_{ij}(\tau) = c_{ij}$ are the data. We drop the dependence on τ to declutter the notation. The transition probabilities can be estimated by maximizing the log likelihood $\ln \mathcal{L}_{\text{MSM}}$ under the constraint that Equation 15 is fulfilled. Using the method of Lagrange multipliers, one obtains the Lagrange function

$$L_{\text{MSM}} = \sum_i \sum_j c_{ij} \cdot \ln p_{ij} + \lambda_i \left(1 - \sum_{j=1}^{n} p_{ij}\right), \qquad 20.$$

$\mathcal{S}_{ij}$: the area of the intersecting hypersurface; not to be confused with the entropy production ΔS_{ij} in Equation 46

Maximum caliber (MaxCal) approach: an instance of the maximum entropy principle that is defined on trajectory space

Supplemental Material >

where λ_i are Lagrange multipliers. Maximizing L_{MSM} yields the maximum likelihood estimator for the MSM transition probabilities (for the full derivation, see the **Supplemental Material**):

$$p_{ij}(\tau) = \frac{c_{ij}(\tau)}{\sum_{j=1}^{n} c_{ij}(\tau)}. \qquad 21.$$

The denominator in Equation 21 estimates the equilibrium population in state Ω_i, $\pi_i = (1/N_\mathrm{t}) \sum_{j=1}^{n} c_{ij}(\tau)$, where N_t is the number of frames in the trajectory. The numerator estimates the absolute transition probability $C_{ij}(\tau) = (1/N_\mathrm{t}) c_{ij}(\tau)$ (see Equation 16). For transition rate constants k_{ij}, one can derive a similar maximum likelihood estimator (28).

The assumptions in the sampling approach for MSMs are less severe than in the SqRA. Instead of assuming a specific type of dynamics for the system as in the SqRA, the dynamical model is given implicitly by the MD simulation and can therefore be very complex. Instead of assuming that the free energy is constant within each cell Ω_i, one assumes that the system has sufficiently sampled the local stationary density within Ω_i (local equilibrium within Ω_i), allowing for larger grid cells. However, because the transitions are sampled by a simulation in the full state space of the system, the estimators can be very hard to converge. Additionally, convergence is difficult to test and control.

2.5. Path Integrals and Correlation Functions

In estimated MSMs, the underlying dynamics are not included explicitly in the model. Instead, they are included implicitly by estimating the transition probabilities from simulations in the full-dimensional state space Ω. Path integrals offer a way to connect the MSM transition probabilities with the dynamical model of the simulation.

The absolute transition probability between two grid cells Ω_i and Ω_j within lag time τ can be interpreted as a correlation function:

$$C_{ij}(\tau) = \langle h_i(x_0) h_j(x_\tau) \rangle = \int_{x_0 \in \Omega} \int_{x_\tau \in \Omega} h_i(x_0) \pi(x_0) p(x_0|x_\tau) h_j(x_\tau) \, \mathrm{d}x_0 \mathrm{d}x_\tau, \qquad 22.$$

where $\pi(x_0)$ is given by Equation 1 and $p(x_0|x_\tau)$ is the conditional transition probability density of finding the system in $x_\tau \mathrm{d}x_\tau$ at time $t = \tau$, given that it was in state x_0 at time $t = 0$. Usually no closed-form expression for $p(x_0|x_\tau)$ can be found for time-continuous dynamics, but it can be related to the time-discretized dynamics of the simulation.

The lag time is $\tau = N_\tau \Delta t$, where Δt is the integration time step of the MD simulation and $N_\tau + 1$ is the number of states that make up a simulated path, $\mathbf{x} = (x_0, x_1, \ldots, x_{N_\tau})$, with initial state x_0. Such a time-discretized path $\mathbf{x}$ is an element of the path space $\mathcal{S} = \Omega^{N_\tau + 1}$. By construction, most MD integrators implement a Markov process; in other words, the state x_{k+1} depends only on the current state x_k and is independent of any of the previous states. Thus, each MD integration step is associated with a single-step transition probability $p(x_{k+1}|x_k)$, and the (normalized) path probability is

$$\mathcal{P}[\mathbf{x}] = p(x_0) \cdot \mathcal{P}[x_1 \ldots x_{N_\tau}|x_0] = p(x_0) \cdot \prod_{k=0}^{N_\tau - 1} p(x_{k+1}|x_k). \qquad 23.$$

The transition probability density $p(x_0|x_\tau)$ for a lag time τ is expressed by integrating over all intermediate time steps between x_0 and x_{N_τ}:

$$p(x_\tau|x_0) = p(x_0) \int_{x_1} \cdots \int_{x_{N_\tau - 1}} \prod_{k=0}^{N_\tau - 1} p(x_{k+1}|x_k). \qquad 24.$$

Thus, Equation 22, which additionally contains the integral over the initial and the final state, becomes a path integral (**Figure 1*c***):

$$C_{ij}(\tau) = \int_{\mathcal{S}} \mathcal{D}\mathbf{x}\, h_i(x_0)\, \mathcal{P}[\mathbf{x}]\, h_j(x_{N_\tau}), \qquad 25.$$

where $\int_{\mathcal{S}} \mathcal{D}\mathbf{x} \ldots$ represents an integral over all possible time-discretized paths $\mathbf{x}$ of length τ.

Equation 25 is formally the starting point for the computation of reactive trajectory ensembles between metastable states A and B using TPS. In this case, states $i = A$ and $j = B$ are not arbitrary cells within Ω_{CV} but instead represent free-energy minima separated by a transition state region. The transition itself is assumed to be a rare event. One can use path sampling to collect an ensemble of rare event trajectories that can be utilized to compute the rate constant k_{AB} and to obtain mechanistic insight (20, 21) (**Figure 1*b***). The rate constant k_{AB}, then, is the time derivative of the ratio of the correlation function C_{AB} to the unrestricted correlation function C_A:

$$k_{AB} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{C_{AB}(\tau)}{C_A} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\int_{\mathcal{S}} \mathcal{D}\mathbf{x}\, h_A(x_0)\, \mathcal{P}[\mathbf{x}]\, h_B(x_{N_\tau})}{\int_{\mathcal{S}} \mathcal{D}\mathbf{x}\, h_A(x_0)\, \mathcal{P}[\mathbf{x}]} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\langle h_A(x_0) h_B(x_{N_\tau})\rangle}{\langle h_A\rangle}, \qquad 26.$$

where the last equality defines a path ensemble. Path sampling can evaluate this correlation function (21) by using, for instance, transition interface sampling (22, 36). Note the direct relationship to Equation 22. Finally, note that in order to evaluate these integrals, one uses reweighting methods such as the weighted histogram analysis method (WHAM) (37), which is essentially a likelihood maximization method.

3. DYNAMICAL REWEIGHTING: GENERAL SETUP

The simulation potential at which the actual simulation is carried out is denoted V. All properties associated with this potential are indicated without a tilde: k_{AB}, τ_{AB}, and so forth. The objective of dynamical reweighting is to estimate kinetic properties at a modified potential $\widetilde{V}$ on the basis of simulations conducted at V. We refer to $\widetilde{V}$ as the target potential, and all properties related to $\widetilde{V}$ are denoted with a tilde: $\widetilde{k}_{AB}$, $\widetilde{\tau}_{AB}$, and so on. The potentials are related by

$$V = \widetilde{V} + U. \qquad 27.$$

Here, U is the differential potential, which can also be time dependent: $U(t)$.

Two different sign conventions exist in the literature on dynamical reweighting. In Equation 27, we take the viewpoint of enhanced sampling simulation, where a bias U is added to the target potential to obtain the simulation potential. Simulations at different bias potentials U^α can be reweighted to a single target potential $\widetilde{V}$ (**Figure 2*a***). Alternatively, one can take the viewpoint of

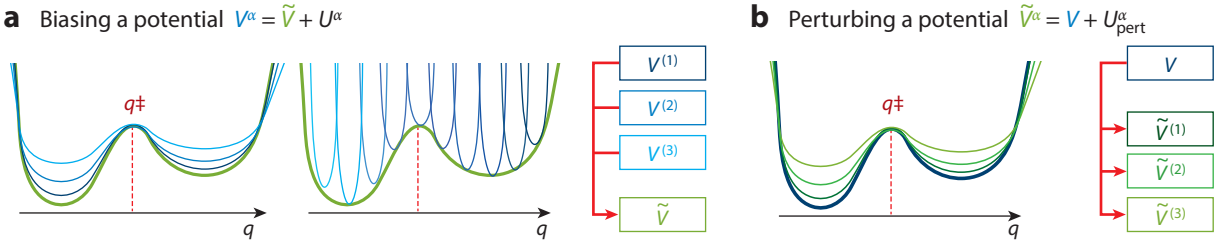

Figure 2

Dynamical reweighting. (*a*) Viewpoint of biasing a potential. Several biased simulation potentials V^α are reweighted to one target potential $\widetilde{V}$. (*Left*) Single bias potential (metadynamics). (*Right*) Series of potentials (umbrella sampling). (*b*) Viewpoint of perturbing a potential. One simulation potential V is reweighted to several perturbed target potentials $\widetilde{V}^\alpha$.

perturbing a potential, where a simulation conducted for a single potential V is reweighted to a range of target potentials $\widetilde{V}^{\alpha}$ (**Figure 2*b***). This viewpoint is intuitive when optimizing potential energy functions, and one defines $\widetilde{V} = V + U_{\text{pert}}$, where U_{pert} represents the perturbation. Setting $U = -U_{\text{pert}}$ recovers Equation 27. Here, we present all equations in line with the sign convention used in Equation 27. However, note the sign convention used in the original literature when comparing equations. We discuss all methods for a single bias U, because extension to multiple biases U^{α} is usually straightforward.

4. TWO-STATE DYNAMICS

4.1. Do Not Reweight at All

Supplemental Material >

Often, the focus is not on absolute rate constants but rather on relative rate constants for a series of similar systems. If one can safely assume that the differential potential U does not perturb this relative order, one can directly compare rate constants obtained at V instead of reweighting them to $\widetilde{V}$. This viewpoint is embraced by τ–random acceleration molecular dynamics (τRAMD) (38, 39), a method designed to gauge the relative unbinding rates in protein–ligand complexes (for more details, see the **Supplemental Material**). Targeted MD (40), a method closely related to τRAMD, allows one to estimate the vector fields for underdamped Langevin dynamics. By using underdamped Langevin as a computationally efficient surrogate model, one can resimulate and analyze the system's dynamics (41).

4.2. Reweighting the Transition State Theory Rate

Methods that accelerate dynamics by adding bias potentials to the currently sampled minimum A have been proposed since the 1990s (42–44); among them, metadynamics (45) is probably the most prominent. Assuming that the bias has been deposited only within A and not on the transition state region, one can derive an acceleration or boost factor α for the TST rate (Equation 9) (43, 44, 46):

$$\alpha = \frac{\widetilde{\tau}_{AB}}{\tau_{AB}} \approx \langle e^{+\beta U} \rangle_A. \tag{28.}$$

Here, τ_{AB} and $\widetilde{\tau}_{AB}$ are the first passage times at the simulation potential V and at the target potential $\widetilde{V}$, respectively, and $\langle \ldots \rangle_A$ denotes an ensemble average at V, where the domain of the integral is restricted to state A. If no bias is deposited on the transition state, then $\langle |\dot{q}| \rangle_{q=q^{\ddagger}}$ and $Z^{\ddagger}$ in Equation 9 are not affected by U and cancel in Equation 28. One is left with the fraction $\widetilde{Z}_A/Z_A$, which can be evaluated using standard reweighting techniques for stationary properties leading to the approximation in Equation 28.

The reweighted first passage time follows from the simulated first passage time as

$$\widetilde{\tau}_{AB} = \alpha \tau_{AB} \approx \left[\frac{1}{n} \sum_{t=1}^{n} e^{+\beta U(q_t)} \right] \tau_{AB} = \Delta t \sum_{t=1}^{n} e^{+\beta U(q_t)}, \tag{29.}$$

where $\tau_{AB} = n\Delta t$ and Δt is the simulation time step. Reference 46 extends this approach to time-dependent biases, that is, to the buildup phase of metadynamics:

$$\widetilde{\tau}_{AB} = \Delta t \sum_{t=1}^{n} e^{+\beta U(q_t;t)}, \tag{30.}$$

where $U(q_t; t)$ is the bias at position q_t at time t.

From a set of reweighted transition times, one can calculate the rate at $\widetilde{V}$ via Equation 6 and $\widetilde{k}_{AB} = \langle \widetilde{\tau}_{AB} \rangle^{-1}$. A more robust estimator is obtained (47) by fitting the cumulative distribution

function (CDF) of $\widetilde{\tau}_{AB}$ to the CDF of a Poisson process, where the mean first passage time at the target potential $\langle\widetilde{\tau}_{AB}\rangle$ is the fit parameter. Reweighting the TST rate in this manner assumes that:

1. In the transition state region, the potential is not modified throughout the simulation.
2. The simulation is in local equilibrium within state A before transitioning across the barrier.

Cumulative distribution function (CDF): the probability of observing at least one transition by time t

Several protocols to ensure the first assumption have been proposed (46, 48). Nonetheless, if the biased coordinate and the reaction coordinate are not optimally aligned, some bias will be deposited at the transition region and, consequently, the barrier height will be overestimated. Reference 49 offers a way to correct for this overestimation, using ideas from single-molecule force spectroscopy experiments (50). This process leads to a time-dependent acceleration factor,

$$\alpha(t) = \frac{k_{AB}(t)}{\widetilde{k}_{AB}} = e^{+\gamma\beta U_{\mathrm{MB}}(t)}, \qquad 31.$$

where $k_{AB}(t)$ is the time-dependent rate, $U_{\mathrm{MB}}(t)$ is the time-dependent maximum bias averaged over several runs with bias buildup, and $\widetilde{k}_{AB}$ is the rate at the target potential. The parameter $\gamma \in [0, 1]$ represents the fraction of $U_{\mathrm{MB}}(t)$ that actually contributes to the acceleration. Any time interval τ_{AB} measured at the time-dependent simulation potential $V(t) = \widetilde{V} + U(t)$ can then be reweighted by

$$\widetilde{\tau}_{AB} = \int_{t=0}^{\tau_{AB}} \alpha(t)\,\mathrm{d}t \approx \Delta t \sum_{t=1}^{n} e^{+\gamma\beta U_{\mathrm{MB}}(q_t,\,t)}. \qquad 32.$$

Equations 30 and 32 differ by the scaling factor γ and by the way the bias at time t is defined. γ and the $\widetilde{k}_{AB}$ are jointly determined by maximizing a likelihood function of the Poisson process (49). If the biased coordinate is well aligned with the reaction coordinate, one expects that 100% of the bias will contribute to the acceleration and, thus, $\gamma = 1$; otherwise, γ falls noticeably below 1. Therefore, γ indicates whether the biased coordinate has been chosen wisely. With specialized biasing protocols and an effective correction method, the first assumption can be adequately controlled.

However, the validity of the second assumption is a subject of controversy (47, 51, 52), as introducing a bias weakens this very assumption. The stronger the acceleration is, the shorter τ_{AB} will be, making it less likely that the system will fully explore region A before leaving it. For time-dependent biases $U(t)$, one even needs to assume that the system fully samples A between increments of $U(t)$; thus, the accuracy of the approximation depends on the deposition rate of the bias. In some cases, violations of the second assumption can be identified by testing whether the reweighted transition times are indeed consistent with Poisson statistics (47). This condition is necessary but not sufficient. Hidden barriers within A are a particular concern in this regard (51, 52), because at a biased potential the transition across the hidden barrier can become slow compared with the transition into B.

Figure 3 illustrates that a system that is well described by two-state dynamics at the target potential $\widetilde{V}$ may exhibit multistate dynamics at the simulation potential V. The distribution of transition times may then exhibit Poisson-like behavior, even though the hidden minima have not been sufficiently explored, leading to inaccurate rate estimates. Considering that most molecular potentials exhibit some degree of ruggedness, this scenario is far from hypothetical.

However, if the two assumptions are met, reweighting according to the TST rate is a powerful technique. Applications range from protein–ligand unbinding (53) to dissociation from metal surfaces (54) to chemical reactions (55, 56). Finally, researchers have proposed several methods

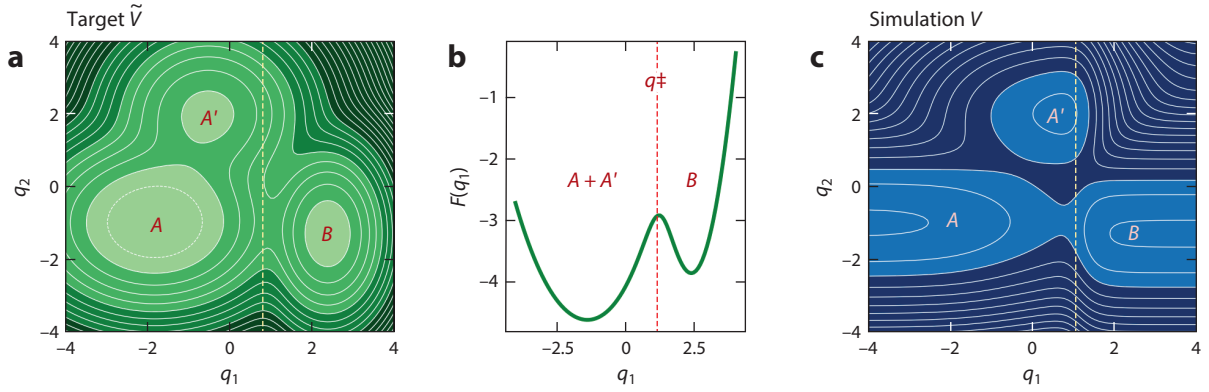

Figure 3

(*a*) Target potential $\widetilde{V}(q_1, q_2)$, with state $A + A'$ exhibiting a barrier that is not visible in the free-energy profile along q_1 shown in panel *b*. (*c*) Simulation potential $V(q_1, q_2) = \widetilde{V}(q_1, q_2) + U(q_1)$ with $U(q_1) = -F(q_1)$. Because of the bias, the barrier between A and B is low in comparison to the hidden barrier between A and A', and exploration of A' becomes a rare event on the timescale of the biased simulation.

(57, 58) in which the bias is scaled by a parameter a and the rate $k_{AB}(a)$ is determined at various bias strengths aU. Extrapolation schemes to the target potential $\widetilde{V}$ at $a = 0$ have been derived on the basis of the Kramers rate equation. Another, closely related method is ligand Gaussian accelerated molecular dynamics (59).

5. MULTISTATE DYNAMICS

Systems with multistate dynamics, at $\widetilde{V}$ or at V (**Figure 3**), can be modeled by an MSM (Equations 11–14). Methods for reweighting MSM transition rates k_{ij} or probabilities $p_{ij}(\tau)$ involve either explicitly defining the effective dynamics in the discretized state space or inferring the dynamics from simulations conducted in the full-dimensional state space Ω.

5.1. Reweighting a Markov Model Using the Square-Root Approximation

The SqRA of an MSM (Equation 18) assumes that the system's dynamics are accurately described by overdamped Langevin dynamics on a low-dimensional free-energy surface $\widetilde{F}(q)$ associated with $\widetilde{V}(q)$ via Equation 4. Equation 18 can be directly evaluated by calculating $\widetilde{F}(q)$ (60, 61) and $\widetilde{D}(q)$ (62–64), usually from umbrella sampling simulations, making the SqRA an intermediate between direct MSM estimation and MSM reweighting.

Equation 18 can also be used to reweight transition rates sampled at the free energy $F(q)$ associated with $V(x) = \widetilde{V}(x) + U(q(x))$. The resulting acceleration factor is

$$\alpha_{ij} = \frac{k_{ij}}{\widetilde{k}_{ij}} = e^{-\beta(U(q_j)-U(q_i))/2}, \qquad 33.$$

where $U(q)$ is the differential potential. Equation 33 can be viewed as reweighting flux_{ij} through the cell boundary (analogous to **Figure 1*a***). In addition to the derivation via the SqRA (65), Equation 33 can be derived via the dynamic histogram analysis method (DHAM) (66), where α_{ij} is used to reweight transition probabilities (Equation 21) as $\widetilde{p}_{ij}(\tau) = \alpha_{ij} p_{ij}(\tau)$. When using this reweighting method, one makes the following assumptions:

1. The system's dynamics are accurately described by overdamped Langevin dynamics at $\widetilde{F}(q)$ as well as at $F(q)$.

2. The diffusion profile is not influenced by the $U(q)$, such that $\widetilde{D}(q) = D(q)$.
3. $\widetilde{F}(q)$ and $F(q)$ are approximately constant within each grid cell.

Badaoui et al. (67) used this approach to model the membrane permeability of a series of drug molecules using the position along the membrane and the orientation of the drug molecule as CVs. The reweighted MSM in this two-dimensional space yielded insights into how the chemical structure of the drug molecule shapes the transition pathway across the membrane. Furthermore, Tiwary & Berne (68) used this approach to identify reaction coordinates from metadynamics simulations by maximizing timescale separation at the target potential $\widetilde{V}$.

5.2. Reweighting a Markov Model Using the Maximum Likelihood Function

Because of assumptions 1–3 above, dynamical reweighting via the SqRA is limited to systems accurately described by a few CVs. For higher-dimensional spaces, or when overdamped Langevin dynamics is not applicable, reweighting the MSM likelihood function (Equation 19) is an alternative. Inspired by methods like WHAM (60) and the multistate Bennett acceptance ratio (MBAR) (61), researchers have proposed protocols for reweighting the MSM likelihood function. Analogous to WHAM and MBAR, no closed-form expression of the maximum likelihood estimator can be found; instead, a system of coupled nonlinear equations arises and is then solved by either fixed-point iteration or gradient-based methods.

5.2.1. Dynamic histogram analysis method. In DHAM (66), one directly reweights the transition probabilities (**Figure 1*b***) as

$$f_i \alpha_{ij} \widetilde{p}_{ij} = p_{ij} \qquad \forall i, \tag{34.}$$

where $\widetilde{p}_{ij}$ are normalized as $\sum_j \widetilde{p}_{ij} = 1 \ \forall i$. Inserting Equation 34 into the MSM Lagrange function (Equation 20), and adding the reweighted normalization as an additional multiplier, yields the DHAM Lagrange function:

$$L_{\mathrm{DHAM}} = \sum_i \sum_j c_{ij} \cdot \ln f_i \alpha_{ij} \widetilde{p}_{ij} + \sum_i \lambda_i \left(1 - \sum_j f_i \alpha_{ij} \widetilde{p}_{ij}\right) + \sum_i \mu_i \left(1 - \sum_j \widetilde{p}_{ij}\right). \tag{35.}$$

Here, c_{ij} are obtained from the simulation at V. The optimization parameters are f_i, α_{ij}, and μ_i, making the system underdetermined.

Building on the SqRA, the authors (66) therefore assume that $\alpha_{ij} = e^{-\beta w U(q_j)} / e^{-\beta(1-w)U(q_i)}$, where $U(q_i)$ is the differential potential at the center of grid cell Ω_i. w can be scaled between 1/2 and 1, where $w = 1/2$ corresponds to an extremely short lag time τ and $w = 1$ corresponds to a lag time τ that is longer than the global equilibration time of the system. By optimizing L_{DHAM} with respect to f_i and μ_i, one obtains the reweighted MSM as a function of w. However, the authors report that the optimization tends to converge to inaccurate results, because detailed balance is not enforced in Equation 35 (66). Setting $w = 1/2$, and normalizing the transition probabilities after reweighting, reduces this approach to reweighting via the SqRA.

5.2.2. Dynamic histogram analysis method extended to detailed balance. The first insight obtained from DHAM is that reweighting a transition probability directly is practical only when a specific dynamical model, such as overdamped Langevin dynamics in the SqRA, is assumed. The second insight is the importance of enforcing detailed balance during the reweighting process. The dynamic histogram analysis method extended to detailed balance (DHAMed) (69, 70) incorporates these lessons in a new ansatz. The MSM likelihood function is reformulated such that detailed

balance and row normalization are guaranteed:

$$\mathcal{L}_{\mathrm{MSM,db}} = \prod_{i>j}\left[p_{ij}^{c_{ij}} \cdot p_{ji}^{c_{ji}}\right] \cdot \prod_i p_{ii}^{c_{ii}}$$
$$= \prod_{i>j}\left[\left(\frac{C_{ij}}{\pi_i}\right)^{c_{ij}} \cdot \left(\frac{C_{ij}}{\pi_j}\right)^{c_{ji}}\right] \cdot \prod_i \exp\left(-c_{ii}\sum_{j\neq i}\frac{C_{ij}}{\pi_i}\right). \quad 36.$$

Here, $C_{ij} = C_{ij}(\tau)$ are absolute transition probabilities (Equation 16) and $C_{ij}(\tau) = C_{ji}(\tau)$ due to detailed balance (Equation 17). The last factor enforces row normalization (69, eq. 31).

Note that detailed balance and normalization are enforced not by Lagrange multipliers but rather by adjusting the functional form of $\mathcal{L}_{\mathrm{MSM}}$. Equation 36 now contains the stationary density $\boldsymbol{\pi}$ at the simulation potential V explicitly. One can reweight $\boldsymbol{\pi}$ by using Equation 3:

$$\pi_i \propto \widetilde{\pi}_i e^{-\beta U(q_i)}, \quad 37.$$

where one assumes that the differential potential $U(q_i)$ is approximately constant within each cell, such that $\pi_i = \text{const.}$ and $\widetilde{\pi}_i = \text{const.}$ Taking the logarithm of reweighted Equation 36 yields the DHAMed log likelihood (see the **Supplemental Material**). By maximizing it and solving the resulting set of equations, one obtains estimates of the absolute transition probability C_{ij} at the simulation potential V and of the stationary density $\widetilde{\boldsymbol{\pi}}$ at the target potential $\widetilde{V}$. If only simulation data from V are available, DHAMed can be viewed as an improvement on WHAM, as it incorporates dynamical information into the estimate of the target stationary density $\widetilde{\boldsymbol{\pi}}$. The estimate of the free-energy surface calculated by DHAMed is more accurate and converges faster than the corresponding WHAM estimate (69).

Supplemental Material >

To obtain the MSM transition matrix at the target potential $\widetilde{V}$, one must also conduct simulations at $\widetilde{V}$. The data then consist of the transition counts $\widetilde{c}_{ij}$ sampled at $\widetilde{V}$ and the transition counts c_{ij} sampled at V. The likelihood is extended as follows:

$$\mathcal{L}_{\mathrm{DHAMed}} = \mathcal{L}_{\mathrm{MSM,db}} \cdot \widetilde{\mathcal{L}}_{\mathrm{MSM,db}}, \quad 38.$$

where $\widetilde{\mathcal{L}}_{\mathrm{MSM}}$ is the MSM likelihood function for $\widetilde{V}$. Equation 38 is then reweighted and maximized as above. This process also leads to maximum likelihood estimates for the absolute transition probabilities $\widetilde{C}_{ij}(\tau)$ at $\widetilde{V}$, which can be converted into transition probabilities as $\widetilde{p}_{ij}(\tau) = \widetilde{C}_{ij}(\tau)/\widetilde{\pi}_i$.

5.2.3. Transition-based reweighting analysis method. Like DHAMed, the transition-based reweighting analysis method (TRAM) (71) recognizes that likelihood-based dynamical reweighting requires simulation data at both V and $\widetilde{V}$. While DHAMed assumes a constant differential potential U within each grid cell when reweighting the stationary density, TRAM relaxes this assumption and instead assumes that the sampled states in the simulation originate from a locally equilibrated stationary density. Local equilibrium within grid cell Ω_i is defined as

$$\pi_i(q) = \frac{1}{Z_i}e^{-\beta V}h_i(q) = \frac{Z}{Z_i}\pi(q)h_i(q), \quad 39.$$

where $Z_i = \int_\Omega e^{-\beta V}h_i(q)\,\mathrm{d}q$ is the local partition function for cell Ω_i and $h_i(q) = h_i(q(x))$ is the indicator function for Ω_i (Equation 7). $\widetilde{\pi}_i(q)$ and $\widetilde{Z}_i$ at the target potential $\widetilde{V}$ are defined analogously.

The likelihood that a trajectory $\mathbf{q} = (q_0, q_1, \ldots, q_{N_t})$ has been sampled from a locally equilibrated stationary density is

$$\mathcal{L}_{\mathrm{LEQ}}(\mathbf{q}) = \prod_{i=1}^{n}\prod_{k=0}^{N_t}\pi_i(q_k) = \prod_{i=1}^{n}\prod_{k=0}^{N_t}\frac{Z}{Z_i}\pi(q_k)h_i(q_k), \quad 40.$$

where n is the number of grid cells and N_t is the number of frames in the trajectory. The TRAM likelihood then combines $\mathcal{L}_{\mathrm{LEQ}}$ and the MSM likelihood (Equation 19) for both data sets:

$$\mathcal{L}_{\mathrm{TRAM}}(\mathbf{q}, \widetilde{\mathbf{q}}) = \mathcal{L}_{\mathrm{MSM}}(\mathbf{q}) \cdot \mathcal{L}_{\mathrm{LEQ}}(\mathbf{q}) \cdot \widetilde{\mathcal{L}}_{\mathrm{MSM}}(\widetilde{\mathbf{q}}) \cdot \widetilde{\mathcal{L}}_{\mathrm{LEQ}}(\widetilde{\mathbf{q}}). \quad 41.$$

Maximum entropy (MaxEnt) approach: an instance of the maximum entropy principle that is used for ensembles in configuration or phase space

The likelihood is reweighted by reweighting $\mathcal{L}_{\mathrm{LEQ}}$ via the Boltzmann reweighting factor $g(q)$ on the stationary probability

$$\frac{Z}{\widetilde{Z}} g(q) = \frac{\widetilde{\pi}(q)}{\pi(q)} = \frac{Z}{\widetilde{Z}} e^{+\beta U(q)}. \quad 42.$$

The reweighted stationary probabilities $\widetilde{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}$ are enforced onto their respective MSMs using a Lagrange multiplier for the detailed-balance condition (Equation 17). Further Lagrange multipliers ensure that both MSM transition matrices are row normalized and that $\widetilde{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}$ are normalized. The resulting system of coupled nonlinear equations for the optimized Lagrange function yields the transition probabilities $\widetilde{p}_{ij}(\tau)$ at $\widetilde{V}$.

Earlier versions of this method (72, 73) were published prior to TRAM; they in turn profited from research on temperature reweighting for MSMs (74, 75). Reference 69 contains a detailed analytical and numerical comparison of DHAMed and TRAM. While the two methods are similar in terms of accuracy, the iterative solution of the DHAMed equations converges faster than those of the TRAM equations. The convergence for the TRAM equations was recently improved through the introduction of a stochastic approximator to the iterative solution (76). Furthermore, Reference 77 presents a method closely related to TRAM and DHAM that is derived via the MaxCal approach (see Section 5.3). Reference 78 benchmarks this method against TRAM.

TRAM has been applied to the modeling of protein and peptide folding (79), binding processes to proteins (80), and ion transport through a membrane channel (81). Both DHAMed and TRAM sample trajectories at the target potential $\widetilde{V}$ and augment these data by simulations at a biased potential $V = \widetilde{V} + U$. Reference 82 uses a sensitivity analysis to determine which MSM transitions are undersampled and chooses V accordingly.

5.3. Reweighting a Markov Model Using the Maximum Caliber–Based Approach

MaxCal is an extension of the maximum entropy (MaxEnt) principle to trajectories proposed by Jaynes (83). By taking X to be a discrete state variable space, one can define the entropy connected to an ensemble of trajectories Γ in that space as a sum over all possible sequences of states, or trajectories (84–86):

$$S = -\sum_{\Gamma} p_{\Gamma} \ln \frac{p_{\Gamma}}{q_{\Gamma}}. \quad 43.$$

Here, $\Gamma = \{X_0, X_1, \ldots, X_L\}$ is a time-ordered sequence of states, or trajectory, where the indices refer to the time. p_{Γ} denotes the probability of this not-necessarily-Markovian trajectory, and q_{Γ} is a prior or reference distribution. For any dynamical feature $s(\Gamma)$, one can compute its average over the ensemble of trajectories: $\langle s \rangle = \sum_{\Gamma} p_{\Gamma} s(\Gamma)$.

The MaxCal approach aims to find the optimal distribution p_{Γ} for which this average is constrained to the desired value s_0 by maximizing the path entropy in Equation 43. By use of the method of Lagrange multipliers, the optimization function, or caliber, then becomes

$$L_{\mathrm{MaxCal}} = -\sum_{\Gamma} p_{\Gamma} \ln \frac{p_{\Gamma}}{q_{\Gamma}} - \eta \left(\sum_{\Gamma} p_{\Gamma} s(\Gamma) - s_0 \right) - \nu \left(\sum_{\Gamma} p_{\Gamma} - 1 \right), \quad 44.$$

where η is a Lagrange multiplier that imposes the ensemble average $\langle s(\Gamma)\rangle$ constraint and ν ensures normalization. This caliber function, completely analogously to the MaxEnt or likelihood maximization approach, can be maximized by setting $\partial\mathcal{L}/\partial p_\Gamma = 0$ and solving for p_Γ, yielding $p_\Gamma = e^{-\eta s(\Gamma)}/\mathcal{Z}$. Here, $\mathcal{Z}$ is a normalization constant analogous to the partition function $\mathcal{Z} = \sum_\Gamma e^{-\eta s(\Gamma)}$. Taking the derivative of $\mathcal{Z}$ with respect to η returns the path ensemble average of s, $\partial \ln \mathcal{Z}/\partial\eta = -\langle s_i\rangle$, which allows evaluation of the value of the Lagrange multiplier (for more details, see 84, 85).

An application of the MaxCal approach that is pertinent to this review is the optimization of transition rate matrices in MSMs, for both equilibrium and nonequilibrium steady states (87, 88). For this application, the path probability is approximated as a discrete Markov chain $p_\Gamma \approx \pi_0 \prod_{t=0}^{T-1} p_{i_t i_{t+1}}$, where $p_{i_t i_{t+1}}$ denotes the transition probability from state i_t to the state at the next time step i_{t+1}. By summing over all path realizations, one can approximate the path entropy for this discrete Markov chain as

$$S \approx -N_\mathrm{t} \sum_{ij} \pi_i p_{ij} \ln \frac{p_{ij}}{q_{ij}}, \qquad 45.$$

where the sum is now over the discrete state space. This much simpler equation can be used to optimize a transition probability matrix p_{ij} (or, equivalently, a rate matrix k_{ij}).

Aside from its equilibrium applications, the MaxCal approach is very useful for analyzing nonequilibrium steady states. Bause et al. (87) employed the MaxCal method to analyze MSMs out of equilibrium, for instance, by applying an (active) nonconservative force. The most straightforward approach is to add a constraint involving the local entropy production at each transition $\Delta S_{ij} = \ln(p_{ij}/p_{ji})$, which leads to the caliber (87)

$$C_{\Delta S_{ij}} = \sum_{ij} \pi_i p_{ij} \ln \frac{p_{ij}}{q_{ij}} + \sum_j \nu_j \left(\sum_i \pi_i p_{ij} - \pi_j \right) + \sum_{ij} \pi_i \mu_{ij} \left(\ln \frac{p_{ij}}{p_{ji}} - \Delta S_{ij} \right). \qquad 46.$$

Optimizing this caliber function for the transition matrix requires the entropy production ΔS_{ij} for each transition, which is approximated by integrating the work performed as a result of the nonconservative force that drives the system out of equilibrium (87, 88; for comments on temperature reweighting for MSMs, see the **Supplemental Material**).

Supplemental Material >

6. REWEIGHTING DYNAMICAL PATHS

Girsanov reweighting directly addresses the problem of reweighting the conditional transition probability $p_{ij}(\tau)$ when only data at the simulation potential V are available. Rather than assuming effective dynamics in Ω_CV, this method analyzes the dynamics of the MD simulation in the full molecular state space Ω. The key realization is that the trajectories are generated by a well-known MD integration algorithm in Ω, and it is ultimately these dynamics that require reweighting.

6.1. Girsanov Reweighting for Path Integrals

The starting point for Girsanov reweighting is the path integral formulation of the absolute transition probability $C_{ij}(\tau)$ (Equation 25), which is reweighted as

$$\widetilde{C}_{ij}(\tau) = \int_{\mathcal{S}} \mathcal{D}\mathbf{x}\, h_i(x_0)\, W[\mathbf{x}] \mathcal{P}[\mathbf{x}]\, h_j(x_{N_\tau}). \qquad 47.$$

Here, $\widetilde{C}_{ij}(\tau)$ is the absolute transition probability at the target potential $\widetilde{V}$, and $\widetilde{\mathcal{P}}[\mathbf{x}] = W[\mathbf{x}]\mathcal{P}[\mathbf{x}]$ is the reweighted path probability density at $\widetilde{V}$. The relative path probability is

$$W[\mathbf{x}] = \frac{\widetilde{\mathcal{P}}[\mathbf{x}]}{\mathcal{P}[\mathbf{x}]} = \frac{\widetilde{p}(x_0)}{p(x_0)} \cdot \prod_{k=0}^{N_\tau - 1} \frac{\widetilde{p}(x_{k+1}|x_k)}{p(x_{k+1}|x_k)} = g(x_0) \cdot M[\mathbf{x}|x_0], \qquad 48.$$

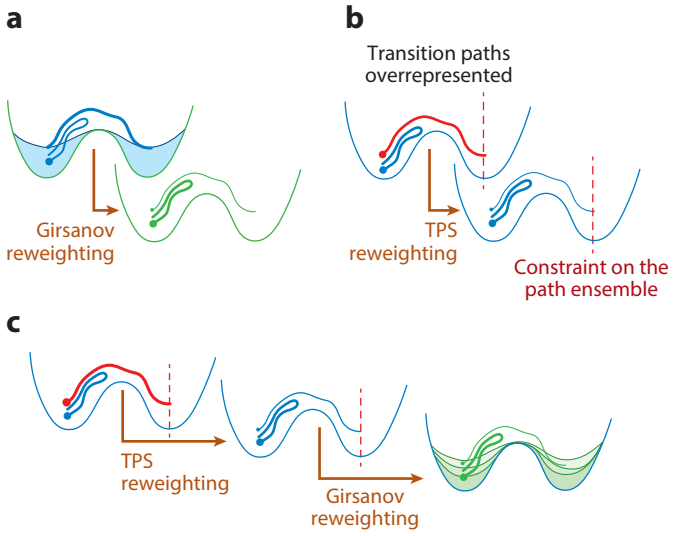

Figure 4

(*a*) Girsanov reweighting samples paths on modified potential energy V (*blue*) and reweights them to the target potential $\widetilde{V}$ (*green*). (*b*) TPS samples at V (*blue*) but overrepresents the transition paths by putting a constraint on the path ensemble; reweighting recovers the correct path ensemble at V. The potential is not modified. (*c*) Reference 89 combines both approaches with the goal of optimizing the force field. TPS is run on the current potential energy surface V (*blue*). The TPS data are reweighted to obtain kinetic models at modified potential energy surfaces $\widetilde{V}$ (*green*). Abbreviation: TPS, transition path sampling.

where $g(x_0)$ is the Boltzmann reweighting factor for the initial state of the path (analogous to Equation 42). The path reweighting factor $M[\mathbf{x}|x_0]$ and, more specifically, the single-step transition probabilities $\widetilde{p}(x_{k+1}|x_k)$ and $p(x_{k+1}|x_k)$ depend on the MD integrator. $W[\mathbf{x}] = g(x_0)\cdot M[\mathbf{x}\,|\,x_0]$ represents the relative statistical weight of a certain path $\mathbf{x}$ at $\widetilde{V}$ in comparison to V (**Figures 1*c*** and **4*a***).

This reweighting approach builds on research by Onsager & Machlup (90) and Girsanov (91), who analyzed relative path probabilities in continuous stochastic dynamics. The Girsanov theorem ensures the preservation of relative path probabilities in overdamped Langevin dynamics as long as the forces due to U remain bounded (91, 92). The well-established expression for $M[\mathbf{x}|x_0]$ in the Euler–Maruyama algorithm enables the reweighting of overdamped Langevin dynamics (92–94). Girsanov reweighting, or dynamical importance sampling, has proven effective in unbiasing overdamped Langevin dynamics for both model potentials (93, 95, 96) and small molecular systems (89, 97–99). However, overdamped dynamics are limited to the mesoscopic regime, as they lack accuracy in capturing atomistic conformational transitions because they suppress intramolecular fluctuations.

By contrast, underdamped Langevin dynamics, or Langevin thermostats, are an accurate and frequently used equation of motion for atomistic MD (100–104). In Langevin MD integrators, the stochasticity of the dynamics is modeled by Gaussian random numbers, and the path probability density $\mathcal{P}[\mathbf{x}]$ can be related to the probability of these random numbers (74). Let $\boldsymbol{\eta}_l = (\eta_{l,0}, \eta_{l,1}, \ldots, \eta_{l,N_\tau-1})$ be a sequence of random numbers drawn for the lth degree of freedom to generate the path $\mathbf{x} = (x_0, x_1, \ldots, x_{N_\tau})$ from the initial state x_0. It follows that

$$\mathcal{P}[\mathbf{x}] = p(x_0)\prod_{l=1}^{3N}\mathcal{P}[\boldsymbol{\eta}_l] = p(x_0)\prod_{l=1}^{3N}\mathcal{N}\exp\left(-\frac{1}{2}\sum_{k=0}^{N_\tau-1}\eta_{l,k}^2\right), \qquad 49.$$

where $\mathcal{N}$ is the normalization constant of the multivariate Gaussian. Equation 49 can be generalized to Langevin thermostats that draw multiple random numbers per integration step (105).

An effective way to implement Girsanov reweighting for an MD simulation is to use on-the-fly reweighting (106, 107), where the path reweighting factor per degree of freedom is expressed as

$$M_l[\mathbf{x}|x_0] = \frac{\mathcal{P}[\widetilde{\boldsymbol{\eta}}_l]}{\mathcal{P}[\boldsymbol{\eta}_l]} = \exp\left(-\sum_{k=0}^{N_\tau-1} \eta_{l,k}\Delta\eta_{l,k}\right) \cdot \exp\left(-\frac{1}{2}\sum_{k=0}^{N_\tau-1} \Delta\eta_{l,k}^2\right). \quad 50.$$

Here, $M[\mathbf{x}|x_0] = \prod_{l=1}^{3N} M_l[\mathbf{x}|x_0]$, and $\widetilde{\eta}_{l,k} = \eta_{l,k} + \Delta\eta_{l,k}$, where $\Delta\eta_{l,k}$ is the random number difference. Equation 50 offers the following intuition. Given a path $\mathbf{x}$ that was generated at the simulation potential V using random number sequences $\boldsymbol{\eta}_l$, we can ask: Which random number sequences $\widetilde{\boldsymbol{\eta}}_l$ are needed to generate the same path at the target potential $\widetilde{V}$? Comparing the probabilities of these two sets of random number sequences yields the path reweighting factor $M[\mathbf{x}|x_0]$. It is important for the implementation that each random number at V can be related to the corresponding random number at $\widetilde{V}$ via $\widetilde{\eta}_{l,k} = \eta_{l,k} + \Delta\eta_{l,k}$, because $\eta_{l,k}$ can be recorded during the simulation at V. $\Delta\eta_{l,k}$ needs to be derived for each stochastic MD integrator.

We recently developed a systematic approach (94, 105) to deriving $\Delta\eta_{l,k}$ and $M[\mathbf{x}|x_0]$ for integrators of underdamped Langevin dynamics. We find that the random number difference is proportional to the force due to the differential potential:

$$\Delta\eta_{l,k} \propto -\frac{\partial}{\partial x_l} U(q(x_{l,k})). \quad 51.$$

Since the $U(q)$ is a function of a low-dimensional CV q, most of these derivatives vanish, and $M_l[\mathbf{x}|x_0] = 1$ for most degrees of freedom. Thus, only those degrees of freedom that contribute to q also contribute to the path reweighting factor $M[\mathbf{x}|x_0] = \prod_{l=1}^{3N} M_l[\mathbf{x}|x_0]$. The proportionality factor in Equation 51 balances dissipation and fluctuation and is a function of T, the mass m associated with the lth degree of freedom, the time step of the integrator Δt, and the friction coefficient of the underdamped Langevin dynamics. The exact functional form depends on the Langevin integrator (for the complete equations, see 105).

Equations 49–51 allow for a simple implementation of Girsanov reweighting: $\eta_{k,l}$ and $U(q(x_k))$ can be recorded during the simulation at V. The parameters for the proportionality factor in Equation 51 are input parameters to the simulation. With this information, one can readily calculate Equation 50 and use it to reweight the path integral in Equation 47.

The Girsanov reweighting equation does not rely on any assumptions, since it reweights the dynamics at the level of the MD integrator. However, several caveats need to be respected for it to be practical:

1. The relative path probability needs to exist.
2. The paths $\mathbf{x}$ should be short.
3. The strength of U should be moderate.

Some Langevin integrators violate the first caveat by sampling different paths at V and $\widetilde{V}$ (105). In this case, Girsanov reweighting cannot be performed without making additional assumptions. The overlap of $\mathcal{P}[\mathbf{x}]$ and $\widetilde{\mathcal{P}}[\mathbf{x}]$ diminishes as the path lengths increase, which makes reweighting long paths numerically inefficient. To address this issue, one can either restrict the path length to the MSM lag time when reweighting MSMs (99, 106) or focus on reweighting the transition path ensemble of rare events, that is, paths that connect A and B (89). Strong U weakens the overlap of the path probability density. An analysis of the optimal biasing strength for Girsanov reweighting is currently lacking.

Girsanov reweighting has been applied to molecular model systems (94, 106, 107) and used in metadynamics simulations of β-hairpin peptide–folding equilibrium (107). Reference 65 compares Girsanov reweighting of MSMs with SqRA reweighting. Girsanov reweighting has also been incorporated into machine learning of CVs (108) and force-field optimization based on kinetic constraints (89) in combination with MaxCal.

6.2. Reweighting in the Continuum Path Ensemble Maximum Caliber Approach

With path integrals, one can extend MaxCal (see Section 5.3) to continuum path (CoPE MaxCal) ensembles (109, 110), where "continuum" means that paths are continuous in space but are still time discretized. The CoPE Lagrange function, or caliber, reads as follows:

$$L_{\mathrm{MaxCal}} = -\int \mathcal{D}\mathbf{x}\widetilde{\mathcal{P}}[\mathbf{x}] \ln \frac{\widetilde{\mathcal{P}}[\mathbf{x}]}{\mathcal{P}[\mathbf{x}]} - \nu\left(\int \mathcal{D}\mathbf{x}\widetilde{\mathcal{P}}[\mathbf{x}] - 1\right) - \mu\left(\int \mathcal{D}\mathbf{x}\widetilde{\mathcal{P}}[\mathbf{x}]s[\mathbf{x}] - s_0\right). \quad 52.$$

Here, the first term is the continuous version of the path entropy (Equation 43), where $\mathcal{P}[\mathbf{x}]$ denotes the prior (reference) distribution and $\widetilde{\mathcal{P}}[\mathbf{x}]$ is the target path probability density (both are defined in Equation 23). Just as in the discrete MaxCal case, the optimal distribution $\widetilde{\mathcal{P}}[\mathbf{x}]$ is given by the maximization of the caliber function. The constraint in the last term is on a specific value s_0 of the path observable $s[\mathbf{x}]$, which can be a time-dependent quantity $s[\mathbf{x}] = s(x_0)s(x_t)$, including the kinetic rate constant (for more details, see 109, 110). Taking the functional derivative of L_{MaxCal} with respect to the path distribution, and setting it to zero, yields for the posterior distribution

$$\widetilde{\mathcal{P}}^{MC}[\mathbf{x}] \propto e^{-\mu s[\mathbf{x}]}\mathcal{P}[\mathbf{x}]. \quad 53.$$

The procedure allows one to find the optimized reweighted path ensembles between (multiple) states, given imposed dynamical constraints such as rate constants (109, 110).

With Girsanov reweighting, one can reweight Equation 52. Doing so enables the evaluation and optimization of L_{MaxCal} at a target potential based on simulation data at V. An intriguing application is to optimize force fields to reproduce accurate experimental kinetic rate constants (89). On the basis of simulations at the current force field V, force-field parameters $\mathbf{a} = (a_1, a_2, \ldots)$ are tuned to reproduce certain dynamical properties while keeping the difference from the current force field (as measured via the path entropy) minimal.

The rate constant predicted by an ensemble of paths starting in initial state A is given by Equation 26. Altering the force field from V to $\tilde{V} = V - U(\mathbf{a})$ changes this prediction to

$$\widetilde{k}_{AB} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\int_{\mathcal{S}} \mathcal{D}\mathbf{x}\, h_A(x_0)\, W[\mathbf{x},\mathbf{a}]\mathcal{P}[\mathbf{x}]\, h_B(x_{N_\tau})}{\int_{\mathcal{S}} \mathcal{D}\mathbf{x}\, h_A(x_0)\, W[\mathbf{x},\mathbf{a}]\mathcal{P}[\mathbf{x}]} = \frac{\mathrm{d}}{\mathrm{d}\tau}\langle h_A(x_0)h_B(x_{N_\tau})\rangle_W. \quad 54.$$

Here, $W[\mathbf{x},\mathbf{a}]$ is the relative path probability (Equation 48), which now depends explicitly on the force-field parameters $\mathbf{a}$. In the second equality, we use the path ensemble average notation, where the subscript denotes that the path average is done in the reweighted ensemble. This path ensemble can be obtained by using, for example, TPS.

Defining the path ensemble average as

$$\langle s[\mathbf{x}]\rangle_W = \frac{Z}{\widetilde{Z}(\mathbf{a})}\int \mathcal{D}\mathbf{x}\, W[\mathbf{x};\mathbf{a}]\, \mathcal{P}[\mathbf{x}]s[\mathbf{x}], \quad 55.$$

and inserting the rate constant expression (Equation 54) into Equation 52, yields for the caliber

$$L_{\mathrm{MaxCal}}(\mathbf{a}) = \langle \ln W[\mathbf{x};\mathbf{a}]\rangle_W - \ln\frac{\widetilde{Z}(\mathbf{a})}{Z} - \mu\left(\langle h_A(x_0)\dot{h}_B(x_{N_\tau})\rangle_W - k_{AB}^{\mathrm{exp}}\right). \quad 56.$$

This equation can be maximized to find the optimal potential parameters $\mathbf{a}$, which reproduce the imposed kinetics and at the same time provide a minimal perturbation with respect to the reference

Supplemental Material >

path distribution. To evaluate Equations 54–56, we combine TPS and Girsanov reweighting (for examples, see 89) (**Figure 4*c***). Another interesting approach is variational path sampling (briefly discussed in the **Supplemental Material**).

7. OUTLOOK

7.1. Enhanced (Biased) Sampling in Path Space

Since an accurate estimate of the kinetics requires that all relevant pathways be sampled, it seems natural to enhance the sampling by applying the bias directly in path space. (Note that applying a bias in path space does not include TPS, which uses a selection/constraint rather than a bias.) A step in this direction is CORE-MD (111), which derives a bias for the underlying dynamics from a path integral. Another method involves representing the path as a polymer with harmonic bonds by using the Onsager–Machlup action and performing enhanced sampling on this polymer representation (112). The individual conformations in the path act as beads, the progress in conformational space represents the bond length, and the expected progress from deterministic forces serves as the reference bond length (see also 19). A challenge in this approach is the high dimensionality of the path space, namely $\mathcal{S} = \Omega^{N_\tau+1}$, where Ω is the system's state space and $N_\tau + 1$ is the number of conformations in the path.

7.2. Machine Learning for Collective Variables

Fitting nonlinear functions in high-dimensional spaces has become possible with deep neural networks. Since CVs fall exactly into this class of functions, a wide range of machine learning methods to discover and improve CVs have been proposed in recent years (113–116). Machine learning techniques are notoriously data hungry. Therefore, enhanced sampling techniques must be applied in order to generate data with enough rare event transitions. As mentioned in Section 1, the optimal reaction coordinate for a two-state process is the committor function, which is an inherently dynamical property that has been used in several transition path–based approaches (117–119). An important future application of dynamical reweighting is to reweight dynamical properties that serve as a loss function for machine learning CVs. The Girsanov reweighting enhanced sampling technique (108) integrates enhanced sampling and CV optimization, employing neural networks to represent the dominant MSM eigenfunctions. These networks maximize the dominant eigenvalues by training on Girsanov-reweighted data from metadynamics simulations along the CVs. The spectral gap optimization of order parameters technique (68) ranks the spectral gap for a SqRA-reweighted MSM for a range of (conventionally represented) CVs. Additionally, force-field optimization techniques, discussed above, offer an alternative machine learning approach to fine-tune parameters, but for nonequilibrium forces or even the potential model itself (89, 120).

7.3. Multistate Chemical Reactions

The key applications of modern potential energy functions are chemical reaction networks with multiple intermediates and reaction pathways. These multistate processes are associated with free-energy barriers that can only be overcome with enhanced sampling simulations. It is likely that modern potential energy functions will not be sampled directly; instead, an iterative approach of enhanced sampling, reweighting, and CV learning will become an integral part of the simulation. **Figure 5*a*** shows what such a computer experiment might look like. Applications ranging from drug unbinding (121) to enzyme catalysis (122) to protonation dynamics at solid–water interfaces (123) have already succeeded in using such integrative approaches. However, the accuracy of the

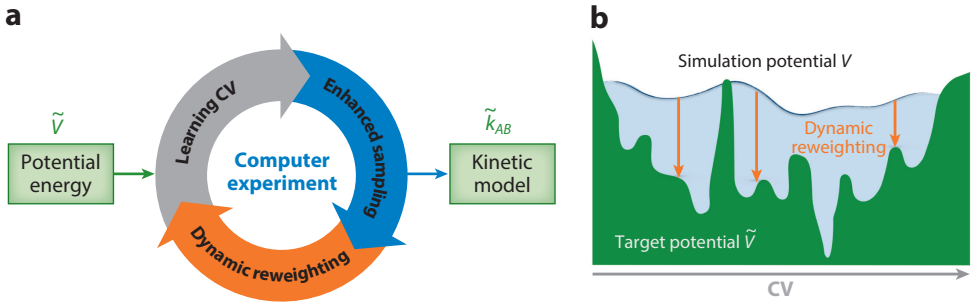

Figure 5

(*a*) Workflow of an integrated simulation approach. Enhanced sampling simulations are dynamically reweighted to obtain a kinetic model, from which an improved CV is machine learned. This optimized CV initiates a new cycle of sampling, reweighting, and machine learning. After convergence, a kinetic model at $\tilde{V}$ is obtained. (*b*) Since enhanced sampling at V is used in panel *a*, dynamic reweighting is crucial for retaining the details of the actual molecular dynamics at $\tilde{V}$. Abbreviation: CV, collective variable.

reweighting will determine whether we can profit from the (quantum-mechanical) accuracy of the potential energy function (**Figure 5*b***). All reweighting methods discussed in this review reach their limit when the bias forces are very strong; on the other hand, they allow for agnostic state-space exploration. Therefore, it is likely that biasing will be combined with path sampling methods to model potential energy surfaces with a maze of high barriers (**Figure 4*c***).

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors thank Simon Ghysbrecht for proofreading the manuscript. This research was funded by Deutsche Forschungsgemeinschaft through grant CRC 1114 Scaling Cascades in Complex Systems, project number 235221301, project B05.

LITERATURE CITED

1. Eyring H. 1935. The activated complex in chemical reactions. *J. Chem. Phys.* 3(2):107–15
2. Kramers HA. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7(4):284–304
3. Bedrov D, Piquemal JP, Borodin O, MacKerell AD Jr., Roux B, Schröder C. 2019. Molecular dynamics simulations of ionic liquids and electrolytes using polarizable force fields. *Chem. Rev.* 119(13):7940–95
4. Bannwarth C, Caldeweyher E, Ehlert S, Hansen A, Pracht P, et al. 2021. Extended tight-binding quantum chemistry methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 11(2):e1493
5. Kocer E, Ko TW, Behler J. 2022. Neural network potentials: a concise overview of methods. *Annu. Rev. Phys. Chem.* 73:163–86
6. Noé F, Tkatchenko A, Müller KR, Clementi C. 2020. Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.* 71:361–90
7. Hénin J, Lelièvre T, Shirts MR, Valsson O, Delemotte L. 2022. Enhanced sampling methods for molecular dynamics simulations [article v1.0]. *Living J. Comput. Mol. Sci.* 4:1583
8. Peters B. 2016. Reaction coordinates and mechanistic hypothesis tests. *Annu. Rev. Phys. Chem.* 67:669–90
9. Peters B. 2017. *Reaction Rate Theory and Rare Events*. Amsterdam: Elsevier

10. Kamenik AS, Linker SM, Riniker S. 2022. Enhanced sampling without borders: on global biasing functions and how to reweight them. *Phys. Chem. Chem. Phys.* 24(3):1225–36
11. Chen H, Chipot C. 2022. Enhancing sampling with free-energy calculations. *Curr. Opin. Struct. Biol.* 77:102497
12. Camilloni C, Pietrucci F. 2018. Advanced simulation techniques for the thermodynamic and kinetic characterization of biological systems. *Adv. Phys. X* 3(1):1477531
13. Linker SM, Weiß RG, Riniker S. 2020. Connecting dynamic reweighting algorithms: derivation of the dynamic reweighting family tree. *J. Chem. Phys.* 153:234106
14. Kieninger S, Donati L, Keller BG. 2020. Dynamical reweighting methods for Markov models. *Curr. Opin. Struct. Biol.* 61:124–31
15. Huber G, Kim S. 1996. Weighted-ensemble Brownian dynamics simulations for protein association reactions. *Biophys. J.* 70(1):97–110
16. Zhang BW, Jasnow D, Zuckerman DM. 2010. The "weighted ensemble" path sampling method is statistically exact for a broad class of stochastic processes and binning procedures. *J. Chem. Phys.* 132:054107
17. Allen RJ, Warren PB, ten Wolde PR. 2005. Sampling rare switching events in biochemical networks. *Phys. Rev. Lett.* 94:018104
18. Cerou F, Guyader A, Lelievre T, Pommier D. 2011. A multiple replica approach to simulate reactive trajectories. *J. Chem. Phys.* 134:054108
19. Dellago C, Bolhuis PG, Csajka FS, Chandler D. 1998. Transition path sampling and the calculation of rate constants. *J. Chem. Phys.* 108(5):1964–77
20. Bolhuis PG, Chandler D, Dellago C, Geissler PL. 2002. Transition path sampling: throwing ropes. *Annu. Rev. Phys. Chem.* 53:291–318
21. Dellago C, Bolhuis PG. 2009. Transition path sampling and other advanced simulation techniques for rare events. In *Advanced Computer Simulation Approaches for Soft Matter Sciences III*, ed. C Holm, K Kremer, pp. 167–233. Berlin: Springer
22. Cabriolu R, Refsnes KMS, Bolhuis PG, van Erp TS. 2017. Foundations and latest advances in replica exchange transition interface sampling. *J. Chem. Phys.* 147:152722
23. E W, Vanden-Eijnden E. 2010. Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* 61:391–420
24. Bowman GR, Pande VS, Noé F, eds. 2013. *An Introduction to Markov State Models and Their Application to Long-Timescale Molecular Simulation*. Berlin: Springer
25. Zwanzig R. 2001. *Nonequilibrium Statistical Mechanics*. Oxford, UK: Oxford Univ. Press
26. Prinz JH, Wu H, Sarich M, Keller B, Senne M, et al. 2011. Markov models of molecular kinetics: generation and validation. *J. Chem. Phys.* 134:174105
27. Swope WC, Pitera JW, Suits F. 2004. Describing protein folding kinetics by molecular dynamics simulations. 1. Theory. *J. Phys. Chem. B* 108(21):6571–81
28. Buchete NV, Hummer G. 2008. Coarse master equations for peptide folding dynamics. *J. Phys. Chem. B* 112(19):6057–69
29. Bicout D, Szabo A. 1998. Electron transfer reaction dynamics in non-Debye solvents. *J. Chem. Phys.* 109(6):2325–38
30. Berneche S, Roux B. 2003. A microscopic view of ion conduction through the K^+ channel. *PNAS* 100(15):8644–48
31. Krämer A, Ghysels A, Wang E, Venable RM, Klauda JB, et al. 2020. Membrane permeability of small molecules from unbiased molecular dynamics simulations. *J. Chem. Phys.* 153:124107
32. Lie HC, Fackeldey K, Weber M. 2013. A square root approximation of transition rates for a Markov state model. *SIAM J. Matrix Anal. Appl.* 34(2):738–56
33. Donati L, Heida M, Keller BG, Weber M. 2018. Estimation of the infinitesimal generator by square-root approximation. *J. Condens. Matter Phys.* 30:425201
34. Donati L, Weber M, Keller BG. 2021. Markov models from the square root approximation of the Fokker–Planck equation: calculating the grid-dependent flux. *J. Condens. Matter Phys.* 33:115902
35. Dixit PD, Jain A, Stock G, Dill KA. 2015. Inferring transition rates of networks from populations in continuous-time Markov processes. *J. Chem. Theory Comput.* 11:5464–72

36. van Erp TS, Moroni D, Bolhuis PG. 2003. A novel path sampling method for the calculation of rate constants. *J. Chem. Phys.* 118(17):7762–74
37. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* 13(8):1011–21
38. Kokh DB, Amaral M, Bomke J, Grädler U, Musil D, et al. 2018. Estimation of drug-target residence times by τ-random acceleration molecular dynamics simulations. *J. Chem. Theory Comput.* 14(7):3859–69
39. Kokh DB, Doser B, Richter S, Ormersbach F, Cheng X, Wade RC. 2020. A workflow for exploring ligand dissociation from a macromolecule: efficient random acceleration molecular dynamics simulation and interaction fingerprint analysis of ligand trajectories. *J. Chem. Phys.* 153:125102
40. Wolf S, Stock G. 2018. Targeted molecular dynamics calculations of free energy profiles using a nonequilibrium friction correction. *J. Chem. Theory Comput.* 14(12):6175–82
41. Wolf S, Lickert B, Bray S, Stock G. 2020. Multisecond ligand dissociation dynamics from atomistic simulations. *Nat. Commun.* 11:2918
42. Huber T, Torda A, van Gunsteren W. 1994. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *J. Comput.-Aided Mol. Des.* 8:695–708
43. Grubmüller H. 1995. Predicting slow structural transitions in macromolecular systems: conformational flooding. *Phys. Rev. E* 52:2893–906
44. Voter AF. 1997. Hyperdynamics: accelerated molecular dynamics of infrequent events. *Phys. Rev. Lett.* 78:3908–11
45. Laio A, Parrinello M. 2002. Escaping free-energy minima. *PNAS* 99(20):12562–66
46. Tiwary P, Parrinello M. 2013. From metadynamics to dynamics. *Phys. Rev. Lett.* 111:230602
47. Salvalaglio M, Tiwary P, Parrinello M. 2014. Assessing the reliability of the dynamics reconstructed from metadynamics. *J. Chem. Theory Comput.* 10(4):1420–25
48. Ray D, Ansari N, Rizzi V, Invernizzi M, Parrinello M. 2022. Rare event kinetics from adaptive bias enhanced sampling. *J. Chem. Theory Comput.* 18(11):6500–9
49. Palacio-Rodriguez K, Vroylandt H, Stelzl LS, Pietrucci F, Hummer G, Cossio P. 2022. Transition rates and efficiency of collective variables from time-dependent biased simulations. *J. Phys. Chem. Lett.* 13(32):7490–96
50. Hummer G, Szabo A. 2003. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.* 85(1):5–15
51. Dickson BM. 2018. Erroneous rates and false statistical confirmations from infrequent metadynamics and other equivalent violations of the hyperdynamics paradigm. *J. Chem. Theory Comput.* 15(1):78–83
52. Khan SA, Dickson BM, Peters B. 2020. How fluxional reactants limit the accuracy/efficiency of infrequent metadynamics. *J. Chem. Phys.* 153:054125
53. Mahinthichaichan P, Liu R, Vo QN, Ellis CR, Stavitskaya L, Shen J. 2023. Structure–kinetics relationships of opioids from metadynamics and machine learning analysis. *J. Chem. Inf. Model.* 63(7):2196–206
54. Ludwig T, Singh AR, Nørskov JK. 2020. Subsurface nitrogen dissociation kinetics in lithium metal from metadynamics. *J. Phys. Chem. C* 124(48):26368–78
55. Fu CD, He Y, Pfaendtner J. 2019. Diagnosing the impact of external electric fields chemical kinetics: application to toluene oxidation and pyrolysis. *J. Phys. Chem. A* 123(14):3080–89
56. Polino D, Parrinello M. 2019. Kinetics of aqueous media reactions via ab initio enhanced molecular dynamics: the case of urea decomposition. *J. Phys. Chem. B* 123(31):6851–56
57. Doshi U, Hamelberg D. 2011. Extracting realistic kinetics of rare activated processes from accelerated molecular dynamics using Kramers' theory. *J. Chem. Theory Comput.* 7(3):575–81
58. Frank AT, Andricioaei I. 2016. Reaction coordinate-free approach to recovering kinetics from potential-scaled simulations: application of Kramers' rate theory. *J. Phys. Chem. B* 120(33):8600–5
59. Miao Y, Bhattarai A, Wang J. 2020. Ligand Gaussian accelerated molecular dynamics (LiGaMD): characterization of ligand binding thermodynamics and kinetics. *J. Chem. Theory Comput.* 16(9):5526–47
60. Ferrenberg AM, Swendsen RH. 1989. Optimized Monte Carlo data analysis. *Comput. Phys.* 3(5):101–4
61. Shirts MR, Chodera JD. 2008. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* 129:124105

62. Hummer G. 2005. Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J. Phys.* 7(1):34
63. Comer J, Chipot C, González-Nilo FD. 2013. Calculating position-dependent diffusivity in biased molecular dynamics simulations. *J. Chem. Theory Comput.* 9(2):876–82
64. Sicard F, Koskin V, Annibale A, Rosta E. 2021. Position-dependent diffusion from biased simulations and Markov state model analysis. *J. Chem. Theory Comput.* 17(4):2022–33
65. Donati L, Weber M, Keller BG. 2022. A review of Girsanov reweighting and of square root approximation for building molecular Markov state models. *J. Math. Phys.* 63:123306
66. Rosta E, Hummer G. 2015. Free energies from dynamic weighted histogram analysis using unbiased Markov state model. *J. Chem. Theory Comput.* 11(1):276–85
67. Badaoui M, Kells A, Molteni C, Dickson CJ, Hornak V, Rosta E. 2018. Calculating kinetic rates and membrane permeability from biased simulations. *J. Phys. Chem. B* 122(49):11571–78
68. Tiwary P, Berne B. 2016. Spectral gap optimization of order parameters for sampling complex molecular systems. *PNAS* 113(11):2839–44
69. Stelzl LS, Kells A, Rosta E, Hummer G. 2017. Dynamic histogram analysis to determine free energies and rates from biased simulations. *J. Chem. Theory Comput.* 13(12):6328–42
70. Stelzl LS, Hummer G. 2017. Kinetics from replica exchange molecular dynamics simulations. *J. Chem. Theory Comput.* 13(8):3927–35
71. Wu H, Paul F, Wehmeyer C, Noé F. 2016. Multiensemble Markov models of molecular thermodynamics and kinetics. *PNAS* 113(23):E3221–30
72. Wu H, Mey ASJS, Rosta E, Noé F. 2014. Statistically optimal analysis of state-discretized trajectory data from multiple thermodynamic states. *J. Chem. Phys.* 141:214106
73. Mey AS, Wu H, Noé F. 2014. xTRAM: estimating equilibrium expectations from time-correlated simulation data at multiple thermodynamic states. *Phys. Rev. X* 4:041018
74. Chodera JD, Swope WC, Noé F, Prinz JH, Shirts MR, Pande VS. 2011. Dynamical reweighting: improved estimates of dynamical properties from simulations at multiple temperatures. *J. Chem. Phys.* 134:244107
75. Prinz JH, Chodera JD, Pande VS, Swope WC, Smith JC, Noé F. 2011. Optimal use of data in parallel tempering simulations for the construction of discrete-state Markov models of biomolecular dynamics. *J. Chem. Phys.* 134:244108
76. Galama MM, Wu H, Krämer A, Sadeghi M, Noé F. 2023. Stochastic approximation to MBAR and TRAM: batchwise free energy estimation. *J. Chem. Theory Comput.* 19(3):758–66
77. Wan H, Zhou G, Voelz VA. 2016. A maximum-caliber approach to predicting perturbed folding kinetics due to mutations. *J. Chem. Theory Comput.* 12(12):5768–76
78. Ge Y, Voelz VA. 2022. Estimation of binding rates and affinities from multiensemble Markov models and ligand decoupling. *J. Chem. Phys.* 156:134115
79. Knoch F, Speck T. 2018. Unfolding dynamics of small peptides biased by constant mechanical forces. *Mol. Syst. Des. Eng.* 3(1):204–13
80. Dodd T, Botto M, Paul F, Fernandez-Leiro R, Lamers MH, Ivanov I. 2020. Polymerization and editing modes of a high-fidelity DNA polymerase are linked by a well-defined path. *Nat. Commun.* 11:5379
81. Hu X, Wang Y, Hunkele A, Provasi D, Pasternak GW, Filizola M. 2019. Kinetic and thermodynamic insights into sodium ion translocation through the μ-opioid receptor from molecular dynamics and machine learning analysis. *PLOS Comput. Biol.* 15(1):e1006689
82. He Z, Paul F, Roux B. 2021. A critical perspective on Markov state model treatments of protein–protein association using coarse-grained simulations. *J. Chem. Phys.* 154:084101
83. Jaynes ET. 1980. The minimum entropy production principle. *Annu. Rev. Phys. Chem.* 31:579–601
84. Dixit PD, Wagoner J, Weistuch C, Pressé S, Ghosh K, Dill KA. 2018. Perspective: Maximum caliber is a general variational principle for dynamical systems. *J. Chem. Phys.* 148:010901
85. Pressé S, Ghosh K, Lee J, Dill KA. 2013. Principles of maximum entropy and maximum caliber in statistical physics. *Rev. Mod. Phys.* 85(3):1115–41
86. Ghosh K, Dixit PD, Agozzino L, Dill KA. 2020. The maximum caliber variational principle for nonequilibria. *Annu. Rev. Phys. Chem.* 71:213–38

87. Bause M, Wittenstein T, Kremer K, Bereau T. 2019. Microscopic reweighting for nonequilibrium steady-state dynamics. *Phys. Rev. E* 100:060103
88. Bause M, Bereau T. 2021. Reweighting non-equilibrium steady-state dynamics along collective variables. *J. Chem. Phys.* 154:134105
89. Bolhuis PG, Brotzakis ZF, Keller BG. 2023. Optimizing molecular potential models by imposing kinetic constraints with path reweighting. *J. Chem. Phys.* 159:074102
90. Onsager L, Machlup S. 1953. Fluctuations and irreversible processes. *Phys. Rev.* 91:1505–12
91. Girsanov IV. 1960. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory Probab. Appl.* 5:285–301
92. Øksendal B. 2003. *Stochastic Differential Equations: An Introduction with Applications*. Berlin: Springer. 6th ed.
93. Adib AB. 2008. Stochastic actions for diffusive dynamics: reweighting, sampling, and minimization. *J. Phys. Chem. B* 112(19):5910–16
94. Kieninger S, Keller BG. 2021. Path probability ratios for Langevin dynamics—exact and approximate. *J. Chem. Phys.* 154:094102
95. Zuckerman DM, Woolf TB. 1999. Dynamic reaction paths and rates through importance-sampled stochastic dynamics. *J. Chem. Phys.* 111(21):9475–84
96. Zuckerman DM, Woolf TB. 2000. Efficient dynamic importance sampling of rare events in one dimension. *Phys. Rev. E* 63:016702
97. Athenes M. 2004. A path-sampling scheme for computing thermodynamic properties of a many-body system in a generalized ensemble. *Eur. Phys. J. B* 38:651–63
98. Xing C, Andricioaei I. 2006. On the calculation of time correlation functions by potential scaling. *J. Chem. Phys.* 124:034110
99. Schütte C, Nielsen A, Weber M. 2015. Markov state models and molecular alchemy. *Mol. Phys.* 113(1):69–78
100. Hünenberger PH. 2005. Thermostat algorithms for molecular dynamics simulations. In *Advanced Computer Simulation Approaches for Soft Matter Sciences I*, ed. C Holm, K Kremer, pp. 105–49. Berlin: Springer
101. Bussi G, Parrinello M. 2007. Accurate sampling using Langevin dynamics. *Phys. Rev. E* 75:056707
102. Leimkuhler B, Matthews C. 2013. Robust and efficient configurational molecular sampling via Langevin dynamics. *J. Chem. Phys.* 138:174102
103. Sivak DA, Chodera JD, Crooks GE. 2013. Using nonequilibrium fluctuation theorems to understand and correct errors in equilibrium and nonequilibrium simulations of discrete Langevin dynamics. *Phys. Rev. X* 3:011007
104. Kieninger S, Keller BG. 2022. GROMACS stochastic dynamics and BAOAB are equivalent configurational sampling algorithms. *J. Chem. Theory Comput.* 18(10):5792–98
105. Kieninger S, Ghysbrecht S, Keller BG. 2023. Girsanov reweighting for simulations of underdamped Langevin dynamics. Theory. arXiv:2303.14696 [cond-mat.stat-mech]
106. Donati L, Hartmann C, Keller BG. 2017. Girsanov reweighting for path ensembles and Markov state models. *J. Chem. Phys.* 146:244112
107. Donati L, Keller BG. 2018. Girsanov reweighting for metadynamics simulations. *J. Chem. Phys.* 149:072335
108. Shmilovich K, Ferguson AL. 2023. Girsanov reweighting enhanced sampling technique (GREST): on-the-fly data-driven discovery of and enhanced sampling in slow collective variables. *Phys. Chem. A* 127(15):3497–517
109. Brotzakis ZF, Vendruscolo M, Bolhuis PG. 2020. A method of incorporating rate constants as kinetic constraints in molecular dynamics simulations. *PNAS* 119(2):e2012423118
110. Bolhuis PG, Brotzakis ZF, Vendruscolo M. 2021. A maximum caliber approach for continuum path ensembles. *Eur. Phys. J. B* 94(9):188
111. Peter EK, Manstein DJ, Shea JE, Schug A. 2021. Core-MD II: a fast, adaptive, and accurate enhanced sampling method. *J. Chem. Phys.* 155:104114
112. Mandelli D, Hirshberg B, Parrinello M. 2020. Metadynamics of paths. *Phys. Rev. Lett.* 125:026001

113. Ketkaew R, Luber S. 2022. DeepCV: a deep learning framework for blind search of collective variables in expanded configurational space. *J. Chem. Inf. Model.* 62(24):6352–64
114. Sidky H, Chen W, Ferguson AL. 2020. Machine learning for collective variable discovery and enhanced sampling in biomolecular simulation. *Mol. Phys.* 118(5):e1737742
115. Chen M. 2021. Collective variable–based enhanced sampling and machine learning. *Eur. Phys. J. B* 94(10):211
116. Hooft F, de Alba Ortíz AP, Ensing B. 2021. Discovering collective variables of molecular transitions via genetic algorithms and neural networks. *J. Chem. Theory Comput.* 17(4):2294–306
117. Ma A, Dinner AR. 2005. Automatic method for identifying reaction coordinates in complex systems. *J. Phys. Chem. B* 109(14):6769–79
118. Frassek M, Arjun A, Bolhuis PG. 2021. An extended autoencoder model for reaction coordinate discovery in rare event molecular dynamics datasets. *J. Chem. Phys.* 155:064103
119. Jung H, Covino R, Arjun A, Leitold C, Dellago C, et al. 2023. Machine-guided path sampling to discover mechanisms of molecular self-organization. *Nat. Comput. Sci.* 3(4):334–45
120. Das A, Kuznets-Speck B, Limmer DT. 2022. Direct evaluation of rare events in active matter from variational path sampling. *Phys. Rev. Lett.* 128:028005
121. Lamim Ribeiro JM, Provasi D, Filizola M. 2020. A combination of machine learning and infrequent metadynamics to efficiently predict kinetic rates, transition states, and molecular determinants of drug dissociation from G protein-coupled receptors. *J. Chem. Phys.* 153:124105
122. Paul TK, Taraphder S. 2022. Nonlinear reaction coordinate of an enzyme catalyzed proton transfer reaction. *J. Phys. Chem. B* 126(7):1413–25
123. Zeng Z, Wodaczek F, Liu K, Stein F, Hutter J, et al. 2023. Mechanistic insight on water dissociation on pristine low-index TiO_2 surfaces from machine learning molecular dynamics simulations. *Nat. Commun.* 14:6131