Annual Review of Physical Chemistry

Variational Path Sampling of Rare Dynamical Events

Aditya N. Singh,[1] Avishek Das,[1,2] and David T. Limmer[1,3,4,5]

[1]Department of Chemistry, University of California, Berkeley, California, USA; email: ansingh@berkeley.edu, avishek_das@berkeley.edu, dlimmer@berkeley.edu

[2]Current affiliation: Fundamental Research on Matter Institute for Atomic and Molecular Physics (AMOLF), Amsterdam, The Netherlands

[3]Kavli Energy Nanoscience Institute, University of California, Berkeley, California, USA

[4]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

[5]Material Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

Annu. Rev. Phys. Chem. 2025. 76:639–62

First published as a Review in Advance on February 19, 2025

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

https://doi.org/10.1146/annurev-physchem-083122-115001

Keywords

molecular simulation, enhanced sampling, nonequilibrium, large deviation theory, rate theory, path sampling

Abstract

This article reviews the concepts and methods of variational path sampling. These methods allow computational studies of rare events in systems driven arbitrarily far from equilibrium. Based upon a statistical mechanics of trajectory space and leveraging the theory of large deviations, they provide a perspective from which dynamical phenomena can be studied with the same types of ensemble reweighting ideas that have been used for static equilibrium properties. Applications to chemical, material, and biophysical systems are highlighted.

1. INTRODUCTION

VPS: variational path sampling

In the past few years, a significant effort has been undertaken to extend the reach of modern molecular simulation tools into the realm of systems driven far from equilibrium (1, 2). This effort is motivated in part by observation. It is increasingly possible to manipulate individual molecules (3), build nanoscale devices subject to extreme driving forces (4), and observe active agents converting energy into directed motion (5). Such experiments drive molecular systems strongly enough that traditional near-equilibrium approaches that form the core theory in physical chemistry are rendered inaccurate (6–8). At the same time, there have been abstract formal developments for nonequilibrium systems (9), including the formulation of stochastic thermodynamics (10) and dynamical large deviation theory (11). Stochastic thermodynamics has ushered in statements of symmetries of dynamical fluctuations (12–14) and generalized response relationships (15–18). Large deviation theory has provided a language for probabilities of time-dependent quantities (19).

Motivated by a desire to bring new theories of nonequilibrium states to bear on detailed molecular systems, we and our coworkers have developed a general computational method for quantifying the likelihoods of dynamical events. The method requires no preconceived notion of the stationary configurational distribution, which is generically unknown when a system is driven from its equilibrium, Boltzmann state. Called variational path sampling (VPS), it leverages the relationship between ensembles of trajectories conditioned on a specific dynamical event and those driven to exhibit that event with high probability through the application of a control force (20, 21). This method generalizes old notions of Brownian bridges (22) and brings them to bear on contemporary molecular problems. It lifts the importance sampling tools available for equilibrium molecular simulations into the study of nonequilibrium systems. The term variational derives from a bound that the control force satisfies, a measure of the distance between the conditioned and driven trajectory ensembles, providing a basis for numerical approximation. Path sampling refers both to the framework of the statistical mechanics of trajectories or paths and to the ability to generate paths that lead to the rare dynamical event directly. We have demonstrated the construction of both approximate controllers and numerically exact ones, opening the way for computational studies of the dynamic pathways of chemical and physical transformations (23–25) and of rare fluctuations in driven and active systems (26–28).

2. RARE EVENTS IN AND AWAY FROM EQUILIBRIUM

Typical fluctuations of molecular systems can be studied straightforwardly by the direct integration of their equation of motion. Through such direct sampling, insights have been made into nonequilibrium systems, and diverse phenomena have been discovered. However, in molecular systems, rare events can be consequential. For example, chemical reactions exhibit a strong separation of timescales—on average a molecule has to wait a long time before a reaction occurs relative to the fundamental timescales of molecular motion, and yet when it occurs, it happens fleetingly, on the timescale of those typical motions (29). Therefore, the likelihood of a reaction occurring at any given time is vanishingly small, yet once it occurs, the properties of the substrate can be dramatically different. Analogously, collections of molecules in one phase can be transformed into another phase by the application of a suitable field. If that transition is between phases of different symmetry, the free energy along a distinguishing order parameter will exhibit two local minima, and the probability distribution along that coordinate will be bimodal. The existence of that second phase, and the coincident high susceptibility to the field, is encoded in the enhancement of probability far away from the typical behavior of the first phase, in states rarely if ever visited (30).

Away from equilibrium, rare fluctuations play important roles in observable molecular behavior, though generically less is understood about their structure. Chemical reactions and phase transitions occur under nonequilibrium driving with more diverse phenomenology than their equilibrium counterparts due to the breaking of time-reversal symmetry. We have found that rates of reactions are generically enhanced (31), and the transitions forbidden at equilibrium can be accessed through the continuous dissipation of heat (25). We have been able to deduce the likelihoods and mechanisms of rare dynamical fluctuations, including dynamic phase transitions (20, 32–34), using the VPS framework.

Action: the negative log likelihood of a trajectory, $\Gamma[\mathbf{X}(\tau)]$

2.1. Path Ensembles and Stochastic Dynamics

The perspective we adopt to study dynamical fluctuations is one in which the fundamental quantity is a stochastic trajectory, or path, of a system over time. By this we mean the time ordered sequence of configurations spanning a defined observation time, τ. This perspective follows from work on dynamical fluctuations by Onsager & Machlup (35) and Ruelle (36) and on transition paths by Chandler and coworkers (37, 38). We denote the configuration of a system at time t as $\mathbf{x}_t$, which could include positions, momenta, or any other dynamical internal variable. A trajectory of a system is denoted as $\mathbf{X}(\tau) = \{\mathbf{x}_t\}_{0\le t\le\tau}$. Molecular systems obey Markovian equations of motion, which admit a factorized formulation of the probability of observing a trajectory, $P[\mathbf{X}(\tau)]$, given by

$$P[\mathbf{X}(\tau)] = \rho(\mathbf{x}_0)P[\mathbf{X}(\tau)|\mathbf{x}_0], \qquad 1.$$

the probability of observing the initial condition $\rho(\mathbf{x}_0)$ times the conditional probability of generating the subsequent trajectory from that initial condition, $P[\mathbf{X}(\tau)|\mathbf{x}_0]$. Moreover, $P[\mathbf{X}(\tau)|\mathbf{x}_0]$ itself can be further factorized over transition probabilities between configurations at arbitrary time displacements.

To obtain a nonequilibrium steady state with persistent injection of energy, a system needs to be in contact with a bath that it can dissipate energy into. If the system is weakly coupled to a bath at a constant temperature, the resulting equation of motion will take the following stochastic form:

$$\dot{\mathbf{x}}_t = \boldsymbol{\mu}\mathbf{F}(\mathbf{x}_t) + \boldsymbol{\eta}_t, \qquad 2.$$

where the dot represents the time derivative, $\boldsymbol{\mu}\mathbf{F}(\mathbf{x}_t)$ is the time-dependent flux of the configurations through phase space, $\boldsymbol{\eta}_t$ is a Gaussian random variable with mean $\langle\boldsymbol{\eta}_t\rangle = 0$, and covariance $\langle\boldsymbol{\eta}_t \otimes \boldsymbol{\eta}_{t'}\rangle = 2k_\mathrm{B}T\boldsymbol{\mu}\delta(t-t')$, where the brackets denote the thermal average while $k_\mathrm{B}T$ is Boltzmann's constant times temperature (39). If $\mathbf{x}$ is a position, then $\boldsymbol{\mu}\mathbf{F}(\mathbf{x}_t)$ is a mobility matrix times a vector force. The force could include a gradient of a potential, $V(\mathbf{x}_t)$, as well as nongradient drifts, $\mathbf{f}(\mathbf{x}_t)$, that take the system away from equilibrium, $\mathbf{F}(\mathbf{x}_t) = -\nabla V(\mathbf{x}_t) + \mathbf{f}(\mathbf{x}_t)$. The probability of a trajectory has an explicit form, written in terms of a stochastic action $\Gamma[\mathbf{X}(\tau)]$ with path integral measure $\mathcal{D}[\mathbf{X}(\tau)]$ satisfying

$$P[\mathbf{X}(\tau)|\mathbf{x}_0] = e^{-\Gamma[\mathbf{X}(\tau)]} \qquad 1 = \int \mathcal{D}[\mathbf{X}(\tau)]e^{-\Gamma[\mathbf{X}(\tau)]}, \qquad 3.$$

a normalization condition. The stochastic action can be written compactly as (9)

$$\Gamma[\mathbf{X}(\tau)] = \frac{1}{4k_\mathrm{B}T}\int_0^\tau \mathrm{d}t\,[\dot{\mathbf{x}}_t - \boldsymbol{\mu}\mathbf{F}(\mathbf{x}_t)]\cdot\boldsymbol{\mu}^{-1}\cdot[\dot{\mathbf{x}}_t - \boldsymbol{\mu}\mathbf{F}(\mathbf{x}_t)], \qquad 4.$$

where the quadratic form follows from the Gaussian noise.[1]

[1]There are subtleties associated with stochastic equations of motion, which arise when the noise depends on the state of the system. While not the case here, it is useful to state that the action above is in the Itô convention, not Stratonovich where the action would include a gradient of the force (40).

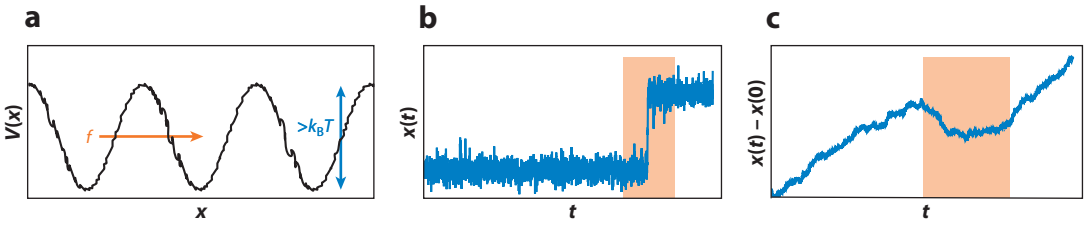

Figure 1

Rare dynamical fluctuations away from equilibrium occur in systems like that shown in panel *a* where a particle evolves under both conservative, $-\nabla V(x)$, and nonconservative f forces. Two examples of rare fluctuations include reactive transitions like the barrier crossing in panel *b* or a particle displacement in the direction opposite the force in panel *c*. Rare dynamical fluctuations are highlighted in orange.

Given a sequence of configurations, the likelihood of a trajectory can thus be exactly computed. While the direct integration of the equation of motion will generate typical trajectories, atypical trajectories exist as well. Take, for example, a particle moving in a periodic potential under an external force like that pictured in **Figure 1*a***. If the amplitude of the potential is large relative to k_BT, and large relative to the external force times the distance between the adjacent minima, the resultant dynamics of barrier crossing will look like that in **Figure 1*b***, where the particle's position will dwell in one state for prolonged periods of time, before fleetingly transitioning over the barrier into a new stable position. As a consequence, the likelihood of a barrier crossing trajectory of finite time is exceptionally low. Analogously, for a constant force propelling a particle in one direction, the second law requires that the particle displaces on average along that direction. If the force is large, displacements against the applied force will be rare, and trajectories like that in **Figure 1*c***, with prolonged domains of displacement against the force, will occur with low likelihoods. Below, we discuss quantifying both kinds of rare events.

2.2. Rare Events in Fixed Time

Transitions between collections of states, like those associated with chemical reactions and phase transitions, are examples of rare events that occur in finite time. To discuss rates of transitions between two sets of states, it is convenient to introduce indicator functions that associate a given configuration to a given set. For example, for two sets of states denoted A and B, we can define h_A and h_B such that

$$h_A[\mathbf{x}_t] = \begin{cases} 1 & \mathbf{x}_t \in A \\ 0 & \mathbf{x}_t \notin A \end{cases} \qquad h_B[\mathbf{x}_t] = \begin{cases} 1 & \mathbf{x}_t \in B \\ 0 & \mathbf{x}_t \notin B \end{cases} \tag{5.}$$

whose action on a configuration at a given time is to return 1 if a configuration is a member of that set and 0 otherwise. These functions act as filters; for example, we can consider the trajectories of a system that start in set A at time $t = 0$, defined as a path partition function Z_A,

$$Z_A = \int \mathrm{d}\mathbf{x}_0 h_A[\mathbf{x}_0]\rho[\mathbf{x}_0] = \langle h_A[\mathbf{x}_0]\rangle, \tag{6.}$$

Path partition function: an integral over a set of trajectories weighted by their probability, Z

which by virtue of the condition acting only on the initial time does not depend on the subsequent path. Notice that Z_A is equivalent to the average of the indicator function, or the probability of being in set A in the steady state. If A is made up of typical configurations, then there will be many paths that start in A, and the partition function will be large. Consider additionally filtering trajectories to be in B at the end of the trajectory. The resultant partition function, $Z_{AB}(\tau)$,

$$Z_{AB}(\tau) = \int \mathrm{d}\mathbf{x}_0 \int \mathcal{D}[\mathbf{X}(\tau)] h_A[\mathbf{x}_0] h_B[\mathbf{x}_\tau] \rho[\mathbf{x}_0] e^{-\Gamma[\mathbf{X}(\tau)]} = \langle h_A[\mathbf{x}_0] h_B[\mathbf{x}_\tau]\rangle \tag{7.}$$

now depends on the duration of the trajectory, τ. The partition function can be recognized as an average, the joint probability that a trajectory starts in set A and ends in set B.

Just as their configurational analogs, ratios of path partition functions are related to physical quantities. The ratio $Z_{AB}(\tau)/Z_A$ has the probabilistic interpretation as the conditional probability of ending in B at time τ given the system started in A initially. The time rate of change of that probability, k_{AB}, is thus

$$k_{AB} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{Z_{AB}(\tau)}{Z_A} \sim \frac{1}{\tau}\frac{Z_{AB}(\tau)}{Z_A}, \qquad 8.$$

equivalent to a microscopic definition of the phenomenological rate constant (29) for transitioning between metastable states (41). Invoking the separation of timescales between the waiting time to transition and the characteristic time for the transition to occur over, it is expected that $Z_{AB}(\tau)$ will increase linearly in time, meaning that the time derivative is well-approximated by dividing by τ.

Rate constant: probability per time of a reaction like A goes to B, k_{AB}

Rate function: the time-intensive negative log likelihood of a time-integrated observable, $I(\mathcal{O}/\tau)$

2.3. Rare Events in the Long Time Limit

In complex systems, relevant dynamic variables are usually collective coordinates, and their fluctuations are encoded along their time series. There are generically two different kinds of time-integrated, time-local observables (19), those that depend on the configuration denoted, $\mathcal{R}$, and those that depend on instantaneous changes to the configuration, $\mathcal{J}$. For trajectories of duration τ, both can be defined as

$$\mathcal{R}[\mathbf{X}(\tau)] = \int_0^\tau \mathrm{d}t\, r[\mathbf{x}_t] \qquad \mathcal{J}[\mathbf{X}(\tau)] = \int_0^\tau \mathrm{d}t\, \mathbf{j}[\mathbf{x}_t]\cdot\dot{\mathbf{x}}_t, \qquad 9.$$

where $r[\mathbf{x}_t]$ and $\mathbf{j}[\mathbf{x}_t]$ are the time-local quantities associated with $\mathcal{R}$ and $\mathcal{J}[\mathbf{x}_t]$. The former is interpreted as an average of a configurational quantity over time. The latter is associated with currents of configurational quantities, and its integral is a generalized displacement.

Just as we considered trajectories filtered for transitioning between two sets of states, we can consider the ensemble of trajectories conditioned on a value of a specific parameter. Take $\mathcal{O}[\mathbf{X}(\tau)]$ to be some generic time-integrated observable. The partition function for trajectories conditioned on a specific value of $\mathcal{O}[\mathbf{X}(\tau)]$, $Z_\mathcal{O}$, is computed from

$$Z_\mathcal{O} = \int \mathrm{d}\mathbf{x}_0 \int \mathcal{D}[\mathbf{X}(\tau)]\delta\left\{\mathcal{O}[\mathbf{X}(\tau)] - \mathcal{O}\right\}\rho[\mathbf{x}_0]e^{-\Gamma[\mathbf{X}(\tau)]} = \left\langle\delta\left\{\mathcal{O}[\mathbf{X}(\tau)] - \mathcal{O}\right\}\right\rangle, \qquad 10.$$

which is recognized as the average of the delta function, or equivalently the probability that trajectories of duration τ exhibit $\mathcal{O}$. In the limit of long observation times for time-integrated quantities, provided a finite correlation time for fluctuations in $\mathcal{O}$, the partition function or probability takes on a simplified form,

$$I(\mathcal{O}/\tau) = \lim_{\tau\to\infty} -\frac{1}{\tau}\ln Z_\mathcal{O}, \qquad 11.$$

where $I(\mathcal{O}/\tau)$ is a so-called rate function from large deviation theory (11), a negative log likelihood independent of the observation time. The rate function characterizes the spectrum of fluctuations of time-integrated observables. Typical fluctuations encoded in $I(\mathcal{O}/\tau)$ portray expectations from the central limit theorem and its concomitant Gaussian fluctuations. Encoded in the extreme fluctuations, those trajectories consistent with $Z_\mathcal{O}$ for $\mathcal{O}$ far away from the mean, are molecular mechanisms for improbable dynamical fluctuations. Signatures of collective effects like nonequilibrium phase transitions can show up as significant enhancements of rare fluctuations relative to priors informed from the central limit theorem, or neglect of correlations. Moreover, the rate function and its dependence on the driving forces contain information about the response of the system to changing driving forces, through generalized fluctuation-dissipation relationships (18, 34, 42–44).

3. PATH ENSEMBLE REWEIGHTING

Path ensemble: collection of trajectories with a common constraint

The scarcity of rare but important events renders them very difficult to study computationally. In equilibrium systems, knowledge of the stationary distribution enables significant simplification. Enhanced sampling tools, from traditional umbrella sampling to current artificial intelligence–augmented methods, have rendered the sampling of rare configurational fluctuations and the evaluation of their likelihoods straightforward in all but the most pernicious cases (45, 46). These methods increase the likelihood of specific configurations by adding an external potential, and account for the change in the probability by leveraging Boltzmann statistics. Optimal ways of reweighting configurational probabilities, decoding the resulting information, and doing so with high accuracy models are still difficult for some systems and important developments are still being made along these lines (47).

The study of dynamical quantities is generally more challenging. Adding potentials to enhance the sampling of rare configurations necessarily alters the dynamics with which a system evolves through those configurations, often in a manner that is difficult to interpret. In some cases, theory can connect configurational fluctuations to dynamical ones, as is done with transition state theory (48). However, these relationships typically require the system to be in equilibrium. Path sampling approaches provide a means of directly importance sampling dynamical quantities (37). While most efficient for the study of transitions accompanying a chemical reaction, path sampling methods have been extended to diffusive events (49–51). The efficiency of traditional path sampling relies on microscopic reversibility, which is violated in systems away from equilibrium. Variations (32, 52, 53) and alternatives (1, 2, 54) have been proposed to evaluate nonequilibrium averages and rates of rare events. These strategies stratify the rare fluctuation into a sequence of less rare events that can be concatenated. Only recently have methods like VPS been developed that importance sample dynamical quantities by directly enhancing their likelihood through a change to their equation of motion. VPS does this through path ensemble reweighting, in close analogy to traditional methods of configurational enhanced sampling.

3.1. Reweighting in Equilibrium

To illustrate how path ensemble reweighting works, it is useful to briefly review importance sampling and reweighting for configurational statistics in equilibrium. In equilibrium, at fixed temperature, volume, and number of particles, the steady-state distribution is given by a Boltzmann form dependent only on the internal energy of the system. If the system's Hamiltonian is $\mathcal{H}_0(\mathbf{x})$, then the resultant distribution is

$$\rho_0(\mathbf{x}) = \frac{1}{Q_0} e^{-\mathcal{H}_0(\mathbf{x})/k_\mathrm{B}T} \qquad Q_0 = \int \mathrm{d}\mathbf{x}\, e^{-\mathcal{H}_0(\mathbf{x})/k_\mathrm{B}T}, \tag{12.}$$

where Q_0 is the canonical partition function, with $\mathrm{d}\mathbf{x}$, a suitably chosen measure, often made dimensionless by introducing factors of Planck's constant. The likelihood of observing a given configuration of the system under a different Hamiltonian, $\mathcal{H}_1(\mathbf{x})$, can be straightforwardly related to the first distribution,

$$\rho_0(\mathbf{x}) = \rho_1(\mathbf{x}) \frac{Q_1}{Q_0} e^{-\Delta\mathcal{H}(\mathbf{x})/k_\mathrm{B}T} \qquad Q_0 = Q_1 \left\langle e^{-\Delta\mathcal{H}(\mathbf{x})/k_\mathrm{B}T} \right\rangle_1, \tag{13.}$$

where $\Delta\mathcal{H} = \mathcal{H}_0 - \mathcal{H}_1$ and the average is taken under the new Hamiltonian. The probability of a configuration in one ensemble, or under a given Hamiltonian, is related to the probability under a different Hamiltonian times a reweighting factor. Given an order parameter, or putative reaction coordinate, $\mathcal{O}$, whose statistics are of interest, the marginal distribution can be computed by averaging Dirac's delta function, $p_0(\mathcal{O}) = \langle \delta[\mathcal{O} - \mathcal{O}(\mathbf{x})] \rangle_0$, in the original ensemble. In most cases

there will be regions of $p_0(\mathcal{O})$ in which the system is unlikely to visit, such as transition states or distinct metastable minima. Within equilibrium importance sampling, one can add a potential that is a function of $\mathcal{O}$, such that $\Delta\mathcal{H} = U(\mathcal{O})$ to enhance the likelihood of a given value of $\mathcal{O}$ (55). Through the reweighting principle above, the distribution of order parameter values generated with the extra potential is related to the original ensemble as

$$p_0(\mathcal{O}) = p_1(\mathcal{O})\frac{Q_1}{Q_0}e^{-U(\mathcal{O})/k_\mathrm{B}T} \qquad Q_0 = Q_1\left\langle e^{-U(\mathcal{O})/k_\mathrm{B}T}\right\rangle_1, \tag{14.}$$

providing access to the free energy, or reversible work, to change the parameter, $\Delta F = -k_\mathrm{B}T\ln p_0(\mathcal{O})$, in the original ensemble.

Reactive path ensemble: collection of trajectories that react

3.2. Conditioned, Tilted, and Driven Ensembles

In analogy to equilibrium, reweighting trajectory ensembles allows for rare fluctuations to be observed with higher probabilities, admitting studies on their mechanisms and origins. Two important distinctions exist for reweighting paths. First, the central limit theorem dictates that distributions of sums of random variables sharpen around their mean. Distributions of observables that are extensive in particle number and integrated over time are consequently very sharply peaked around their typical values, or, equivalently, fluctuations away from the typical value are less likely. This inevitably means that more computational power is needed to sample them than their configurational counterparts. Second, equilibrium configurational properties are independent of the details of the equations of motion used for sampling them. By construction, dynamical properties lose this universality. This can render their interpretation difficult, and care must be taken in order to pose the relevant equation of motion. The mathematics of reweighting Markovian processes largely go back to old work by Doob (56) and Girsanov (57).

We can construct a path ensemble conditioned on exhibiting a dynamical event, $\mathcal{O}$, by manipulating the original path probability

$$P_\mathcal{O}[\mathbf{X}(\tau)] = \frac{1}{Z_\mathcal{O}}\delta\left\{\mathcal{O}[\mathbf{X}(\tau)] - \mathcal{O}\right\}\rho[\mathbf{x}_0]e^{-\Gamma[\mathbf{X}(\tau)]}, \tag{15.}$$

where the delta function provides the filtering of trajectories and $Z_\mathcal{O}$ renormalizes the distribution. Choosing $\mathcal{O}[\mathbf{X}(\tau)] = h_A[\mathbf{x}_0]h_B[\mathbf{x}_\tau]$ generates the reactive path ensemble $P_{AB}[\mathbf{X}(\tau)]$ with transition path partition function $Z_{AB}(\tau)$ (59). For example, in **Figure 2*a***, for a system that exhibits two metastable basins, denoted A and B, most trajectories started in A will remain there, or will only exhibit brief excursions out of that basin, with a high likelihood of returning. Within the conditioned path ensemble, all trajectories react, albeit with varying probabilities. Rather than sampling from $P_\mathcal{O}[\mathbf{X}(\tau)]$, inspired by traditional equilibrium techniques, one could alternatively alter the system in order to make improbable trajectories typical. For example, one could add an external time-dependent potential that drives the system to a rare event, $\Lambda(\mathbf{x}_t, t)$, with an equation of motion:

$$\dot{\mathbf{x}}_t = \boldsymbol{\mu}\mathbf{F}(\mathbf{x}_t) - \boldsymbol{\mu}\nabla\Lambda(\mathbf{x}_t, t) + \boldsymbol{\eta}_t, \tag{16.}$$

as illustrated for a reactive conditioning in **Figure 2*b***. The consequence of the added potential can be understood quantitatively by relating the conditioned path probability to a conditioned path probability under the external driving, denoted $P_{\mathcal{O},\Lambda}[\mathbf{X}(\tau)]$, assuming the initial condition is the same,

$$P_\mathcal{O}[\mathbf{X}(\tau)] = P_{\mathcal{O},\Lambda}[\mathbf{X}(\tau)]\frac{1}{Z_\mathcal{O}(\tau)}e^{-\Delta\Gamma[\mathbf{X}(\tau)]}, \tag{17.}$$

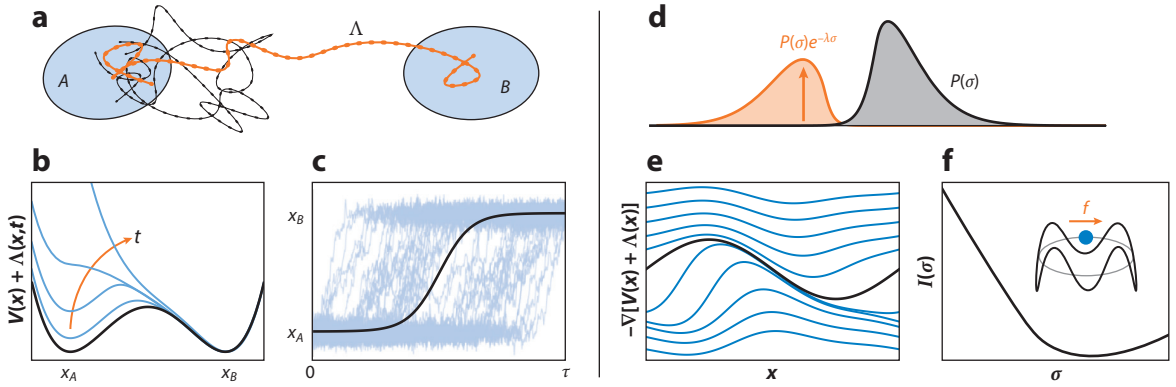

Figure 2

Illustration of conditioned, tilted, and driven trajectory ensembles. (*a*) A typical trajectory (*black*) evolving in a metastable system between states *A* and *B* is rarely reactive (*orange*) unless driven by a control force Λ. (*b*) An example bistable system and time-dependent potential to affect transitions. (*c*) Reactive trajectories (*blue*) and their mean (*black*) with evident time-translational invariance. (*d*) Time-integrated quantities like the entropy production σ exhibit typical fluctuations (*black*) and can be biased to exhibit rare fluctuations by tilting with a parameter λ (*orange*). (*e*) Control forces (*blue*) on top of the conservative force (*black*) that generates rare entropy fluctuations for the system in panel *f* with a particle driven around a ring whose rate function has exponentially small probabilities for negative values due to Gallavotti-Cohen symmetry (58).

where the path partition functions are related by

$$Z_{\mathcal{O}}(\tau) = Z_{\mathcal{O},\Lambda}(\tau)\left\langle e^{-\Delta\Gamma[\mathbf{X}(\tau)]}\right\rangle_{\mathcal{O},\Lambda} \qquad 18.$$

and $\Delta\Gamma$ is the change in stochastic action (31). For the added potential to the stochastic equation of motion above, the change in action is

$$\Delta\Gamma = -\frac{1}{4k_{\mathrm{B}}T}\int_0^\tau \mathrm{d}t\, \nabla\Lambda(\mathbf{x}_t,t)\cdot[2\dot{\mathbf{x}}_t - 2\boldsymbol{\mu}\mathbf{F}(\mathbf{x}_t) + \boldsymbol{\mu}\nabla\Lambda(\mathbf{x}_t,t)], \qquad 19.$$

which depends on the details of the individual trajectories. These reweighting statements are valid for any arbitrary added force. We refer to the collection of trajectories evolved with the added force as the driven path ensemble, and Equation 17 is a relationship between the conditioned path ensemble and the driven one.

Using the previous relation between the rate constant and the ratio of path partition functions (Equation 8), we can define the rate in the driven ensemble analogously,

$$k_{AB,\Lambda} = \frac{1}{\tau}\frac{Z_{AB,\Lambda}(\tau)}{Z_A} \sim \frac{1}{\tau}, \qquad 20.$$

where if the driving force is chosen such that it enforces all trajectories to be reactive, the ratio of partition functions will be one, or the rate in the driven ensemble is just the inverse observation time. Using the reweighting relation, we can relate the rates in the driven and conditioned ensemble in the limit in which the added force ensures reactivity,

$$k_{AB}\tau = \left\langle e^{-\Delta\Gamma[\mathbf{X}(\tau)]}\right\rangle_{\Lambda} \geq e^{-\langle\Delta\Gamma[\mathbf{X}(\tau)]\rangle_{\Lambda}}, \qquad 21.$$

where the first equality is true for all added potentials that enforce reactivity and the inequality follows from the convexity of the exponential and Jensen's inequality (21). An equivalent expression relates the rate to the average change in action in the conditioned ensemble without Λ, $\ln k_{AB}\tau \leq \langle\Delta\Gamma\rangle_{AB}$. While the equality relating the rate constant in the original system to an average in the driven ensemble is exact, it is not of numerical use as exponential averages are

Driven path ensemble: collection of trajectories that react due to an extra, external force

difficult to converge. The inequality is conversely easy to evaluate as it is a simple average, but for an arbitrary potential it may provide a poor estimate of the rate. This inequality forms the basis of the way VPS evaluates rate constants, by variationally optimizing Λ to saturate the bound. The inequality relates the rate in the reference ensemble to the mean change in action, or equivalently since the action is the log likelihood, the rate is given by the Kullback–Leibler divergence between the conditioned and driven ensembles. This observation provides an interpretation for the saturation of the bound, as given by a specific time-dependent potential that renders the subsequent reactive trajectories statistically indistinguishable from those of the conditioned ensemble. In **Figure 2*c***, the time-translational invariance of a steady state would ensure that a reaction would be equally likely to happen at any point over the τ observation time. Driven trajectories should similarly react anywhere in that time window.

Scaled cumulant generating function: negative logarithm of the time-intensive Laplace transform of a probability distribution, $-(\ln Z_\lambda)/\tau$

Alternatives to the conditioned ensembles are so-called tilted ensembles, which have been formulated most often for time-integrated observables (38, 60–65). These ensembles are inspired by the Laplace transform structure of equilibrium statistical mechanics and weight paths exponentially in proportion to the observable of interest times a counting parameter, denoted here as λ. Shown in **Figure 2*d***, the addition of this weighting factor, with specifically chosen values of λ, can enhance the likelihood of otherwise rare fluctuations of $\mathcal{O}$. The resultant distribution of trajectories in the tilted ensemble has probability

$$P_\lambda[\mathbf{X}(\tau)] = \frac{1}{Z_\lambda}\rho[\mathbf{x}_0]e^{-\Gamma[\mathbf{X}(\tau)]}e^{-\lambda\mathcal{O}[\mathbf{X}(\tau)]}, \qquad 22.$$

where in the long time limit the initial condition distribution is irrelevant, and

$$Z_\lambda = \int d\mathbf{x}_0 \int \mathcal{D}[\mathbf{X}(\tau)]\rho[\mathbf{x}_0]e^{-\Gamma[\mathbf{X}(\tau)]}e^{-\lambda\mathcal{O}[\mathbf{X}(\tau)]} = \left\langle e^{-\lambda\mathcal{O}}\right\rangle \qquad 23.$$

is the tilted partition function, equal to the moment generating function, or average exponential of the tilting factor $\lambda\mathcal{O}$. Equivalently, $-(\ln Z_\lambda)/\tau$ is a scaled cumulant generating function. The Laplace structure of the path distribution manifests in the long time limit as a Legendre-transform structure that relates the scaled cumulant generating function to the original rate function,

$$I(\mathcal{O}/\tau) = \frac{1}{\tau}\max_\lambda\left[\lambda\mathcal{O} + \ln Z_\lambda\right], \qquad 24.$$

valid when τ is much longer than the correlation time of $\mathcal{O}$. Working within the tilted ensemble provides equivalent access to the statistics of time-integrated observables.

There are methods to compute $\ln Z_\lambda$ and sample $P_\lambda[\mathbf{X}(\tau)]$, which have been applied to molecular systems (61, 66, 67). Like methods available to compute rates away from equilibrium, these methods stratify the sampling of Z_λ but do not make the fluctuations that contribute most at each value of λ more probable. An alternative is to drive the system directly to the rare value of $\mathcal{O}$. For example, an additional potential, $\mathbf{\Lambda}(\mathbf{x}_t)$, could be added to the equation of motion as in Equation 16, like those drifts shown in **Figure 2*e*** to probe rare particle displacements. With judiciously chosen $\mathbf{\Lambda}(\mathbf{x}_t)$, the same values of $\mathcal{O}$ could be sampled as would be for a given value of λ. The resulting distribution of trajectories, a driven ensemble, is related to the original distribution, $P_\lambda[\mathbf{X}(\tau)]$, as

$$P_\lambda[\mathbf{X}(\tau)] = P_{\lambda,\Lambda}[\mathbf{X}(\tau)]\frac{1}{Z_\lambda}e^{-\Delta\Gamma[\mathbf{X}(\tau)]}, \qquad 25.$$

where the partition function is evaluable as

$$Z_\lambda = \left\langle e^{-\lambda\mathcal{O}[\mathbf{X}(\tau)]-\Delta\Gamma[\mathbf{X}(\tau)]}\right\rangle_\Lambda \geq e^{-\lambda\left\langle\mathcal{O}[\mathbf{X}(\tau)]\right\rangle_\Lambda - \left\langle\Delta\Gamma[\mathbf{X}(\tau)]\right\rangle_\Lambda} \qquad 26.$$

since $P_{\lambda,\Lambda}[\mathbf{X}(\tau)]$ is a normalized distribution so that its partition function is equal to 1 (20). The evaluation of $\ln Z_\lambda$, or equivalently $I(\mathcal{O}/\tau)$, is equivalent to averaging the exponential in the first

equality under the driving dynamics, providing access to rare time-integrated quantities away from equilibrium, an example of which is in **Figure 2*f***. The second inequality acts similar to Equation 21 to provide a variational bound to determine $\ln Z_\lambda$. Analogously, VPS algorithms can be developed to variationally approach the calculation of $I(\mathcal{O}/\tau)$.

Optimal control force: a force that renders rare fluctuations typical in a manner indistinguishable from an unforced spontaneous fluctuation, $-\nabla\Lambda^*$

Committor: the probability of reaching state B before state A starting in configuration $\mathbf{x}$, $\bar{q}_B(\mathbf{x})$

3.3. Relation to Stochastic Control

The variational inequalities in Equations 21 and 26 can be saturated with unique added forces known from previous work in the literature on stochastic processes and have close connections to theories of stochastic control. For this reason, we refer to the added drift that enters into the reweighting expressions as the control force, though it need not be a force in the Newtonian sense. At saturation both equalities follow from a so-called generalized Doob transform (68). For the rate constant problem in Equation 21, the optimal control force, $-\nabla\Lambda^*$, follows from the solution of a Cole–Hopf transformed backward Kolmogorov equation (21),

$$\frac{\partial}{\partial t}\Lambda^* = -\boldsymbol{\mu}\mathbf{F}\cdot\nabla\Lambda^* - k_\mathrm{B}T\nabla\cdot(\boldsymbol{\mu}\nabla\Lambda^*) + \nabla\Lambda^*\cdot(\boldsymbol{\mu}\nabla\Lambda^*)/2 \qquad 27.$$

with boundary conditions applied at the final time, $\Lambda(\mathbf{x}\in B, t=\tau)=0$. The boundary condition enforces that the trajectories are reactive. In the absence of a force in the reference system, this equation is that for a Brownian bridge first solved by Doob (69). Orland and coworkers (70–72) generalized this to barrier crossing events in equilibrium systems and employed analytical approximations to it as a means of generating transition paths, though they did not recognize the connection to the evaluation of rates. Subsequent work (73–75), building off of the transition path theory formalism (76), recognized the stationary solution of the optimal control problem in Equation 27 as isomorphic to that of the committor, or splitting probability, $\bar{q}_B(\mathbf{x}) = \exp[-\Lambda^*(\mathbf{x})\boldsymbol{\mu}^{-1}/2k_\mathrm{B}T]$. We refer to the time-dependent solution of the backward Kolmogorov equation as the time-dependent committor $q_B(\mathbf{x}, t)$ (21, 72), which is related to the optimal controller through $\Lambda^*(\mathbf{x}, t) = -2k_\mathrm{B}T\ln q_B(\mathbf{x}, t)$ (23).

For the evaluation of large deviation functions for observables of the form $\mathcal{O} = \mathcal{J} + \mathcal{R}$, the optimal control force satisfies a Hamilton–Jacobi–Bellman equation (68, 77),

$$\frac{1}{2k_\mathrm{B}T}[\nabla\Lambda^*\cdot(\boldsymbol{\mu}\nabla\Lambda^*) - 2(\boldsymbol{\mu}\mathbf{F})\cdot\nabla\Lambda^*] - \nabla\cdot(\boldsymbol{\mu}\nabla\Lambda^*) \qquad 28.$$
$$= 2\lambda r + 2\psi(\lambda) - 2\lambda^2 k_\mathrm{B}T\mathbf{j}\cdot(\boldsymbol{\mu}\mathbf{j}) - \lambda\Big\{\mathbf{j}\cdot(\boldsymbol{\mu}\nabla\Lambda^*) + (\boldsymbol{\mu}\mathbf{j})\cdot\nabla\Lambda^* - 2(\boldsymbol{\mu}\mathbf{F})\cdot\mathbf{j} - 2k_\mathrm{B}T\nabla\cdot(\boldsymbol{\mu}\mathbf{j})\Big\},$$

where $\psi(\lambda) = (\ln Z_\lambda)/\tau$ is the scaled cumulant generating function. The optimal controller for generating trajectories biased on large deviations, $P_\lambda[\mathbf{X}(\tau)]$, is related to the eigenfunction of a tilted Fokker–Planck equation (78). There is, however, one additional subtlety. If $\mathcal{O}$ includes a current-type variable, $\mathcal{J}$, then the optimal control force cannot be written as a gradient. It contains a gradient part from the solution of the above equation, and an additional nongradient part, such that $\nabla\Lambda^* \to 2k_\mathrm{B}T\lambda\mathbf{j} + \nabla\Lambda^*$. The long time limit in the tilting of the trajectory ensemble renders the optimal control force time-independent. There has been significant work developing the optimal control theory for processes conditioned on large deviations of integrated observables, beginning with its identification (78, 79) and following with a variety of numerical approaches (33, 80, 81). Efficient basis sets like tensor networks have enabled the direct numerical solution of discrete analogies to Equation 28 for lattice models (82–84). Feedback control and alternative variational approaches have also been considered and applied to systems in the continuous space (85, 86).

4. VARIATIONAL PATH SAMPLING

The goal of VPS is to compute the statistics and mechanisms of rare fluctuations by learning the optimal controllers that make the rare events typical. This is done by the parameterization of the control force $-\nabla\Lambda$ through the prescription of an ansatz and minimization of the gradients of the variational cost functions on the right-hand side of Equations 21 and 26 with respect to the parameters. At the heart of VPS is the efficient estimation of these gradients from simulated trajectories given a functional form of Λ. The optimized forces are then used to obtain estimates of the likelihood of the rare fluctuation and can be used to generate the ensemble of trajectories that achieve the fluctuation, from which mechanistic insight can be gained.

4.1. Ansatz for the Control Force

Since the optimal controllers are related to solutions of the Fokker–Planck and backward Kolmogorov equations, they formally span the whole dimensional space of the system. Nevertheless, the parameterization of forces along a low-rank set of descriptors can be both efficient and accurate. Once a set of descriptors is selected, one has to choose an ansatz for parameterization of the forces. This choice strongly depends on the physical symmetries of the system and the nature of the rare fluctuation. Since the optimal force generally respects the highest symmetries common to both the system dynamics and observable, particular advantage can be gained by choosing an ansatz that also obeys the same symmetries. For example, in a spatially periodic system, sampling rare values of a current that does not break that symmetry is achieved with an optimal controller that is also periodic. In many-particle systems with a spatial translation symmetry, sampling rare values of dynamical activity without reference to absolute positions corresponds to a controller that is also translation invariant. The importance of these symmetries has analogously been a guiding principle for the development of neural network–based force fields for molecular dynamics (87, 88).

Generally, linear basis sets, i.e., an ansatz with linear dependence on the variational parameters, are desirable due to the ease in the evaluation of the gradients, as well as the reduced computational cost of performing simulations with them. The choice of the basis sets can be easily tuned to the symmetries of the system with examples illustrated in **Figure 3*a*** including Fourier series for periodic potentials, Laguerre polynomials for long-range pairwise interactions, or Jacobi polynomials to represent time dependence (20, 21, 26). For problems that involve collective dynamics where the dynamically relevant descriptors are challenging to approximate a priori, physics-informed neural networks offer an alternative ansatz for representing the forces and have been used recently for solving a multitude of different problems within physics that are grounded in stochastic optimal control (24, 85, 89–91). Deep models can be advantageous for constructing many-body controllers that can be adapted to the symmetries of the system; however, they can be computationally expensive depending on the size of the system and method of implementation.

For systems with sets of fast and slow variables, such as an underdamped system evolving with a large friction coefficient, methods from homogenization theory (92) can be utilized to marginalize over the fast variables and express the controller solely in terms of the slow variables (24, 93). Finally, for the case of finite-time conditioning, learning the optimal controller requires learning both its spatial and temporal features (21, 26). The time dependence can be parameterized based on knowledge from exactly solvable models, which show that the potential varies away from the observation time over timescales associated with its intrinsic relaxation time. For a reactive conditioning this means that the potential is time independent when $k_{AB}(\tau - t)$ is large, for $t < \tau$ (23).

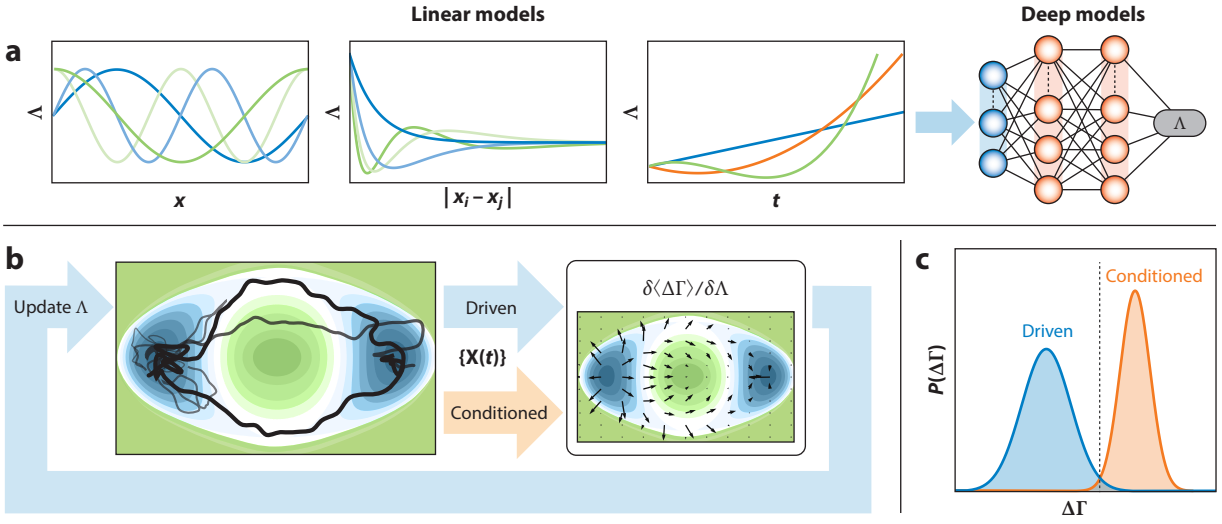

Figure 3

Schematic of the variational path sampling algorithm. (*a*) A representation for the control force is determined by defining descriptors in space and time. These are spanned by basis functions, for example, Fourier (*left*), Laguerre (*middle*), or Jacobi (*right*) functions. The force can be represented by linear combinations of these descriptors or nonlinear models like deep feed-forward neural networks. (*b*) Optimization of the control force follows from either one-shot fitting (*orange*) or iterative learning procedures (*blue*). The rare event being sampled is a barrier crossing transition from the left to the right well on the potential energy surface. The learning procedure requires several iterations to obtain both the rare trajectories and the optimization of the control force. The black lines represent reactive trajectories and the increasing opaqueness highlights the decrease in Kullback–Leibler divergence as the optimization proceeds. (*c*) Estimators can be evaluated based on either the conditioned or driven ensembles, and their accuracy depends on the spread of the distribution of $\Delta\Gamma$.

4.2. Optimization

Once the ansatz has been chosen, the next step involves the minimization of the variational loss functions to obtain the optimal controller. Similar to the construction of the ansatz, the most effective method of optimization depends on the specific type of problem being studied. In the past, we have adopted two distinct approaches that are best referred to as a direct fitting approach, or a reinforcement learning approach, both illustrated in **Figure 3*b***.

When a conditioned ensemble is accessible via other importance sampling methods, a variational loss can be constructed by reweighting Equations 21 and 26 back into the undriven conditioned trajectory ensembles (24, 94). Since the control force then only enters into the relative action, not the trajectory probability, this form of optimization can be done directly. The relative action is fit with either a linear or nonlinear basis set. For linear ansatzes, the optimization can be accomplished by a simple matrix inversion, and for a nonlinear model, the optimization can be performed through gradient descent. In terms of applications, this method of optimization is appropriate for finite-time processes where the computation and analysis of the commitment probability are of interest (95–101), and this method is amenable to problems in equilibrium where the conditioned path ensemble can be accessed with relative ease by leveraging state-of-the-art methods for sampling the reactive path ensembles (98). Within nonequilibrium steady states, however, the applicability of this fitting procedure is diminished by the lack of efficient sampling tools that afford easy access to the conditioned path ensemble.

For cases where a conditioned ensemble is not available directly, VPS can iteratively generate rare trajectories and evaluate their likelihoods by learning the control force on the fly. The

iterative optimization algorithm consists of the choice of initial values of variational parameters, simulation of an ensemble of driven trajectories to compute the variational estimate and its gradients, and updating the parameters for the next iteration with the gradients. At convergence, the optimized force simultaneously gives an estimate of the rate as well as generates rare trajectories as typical ones. Particular advantage can be gained if the force is expressed as a linear combination of a complete set of basis functions. This allows the easy differentiation of the force with respect to the variational parameters. This approach efficiently combines the evaluation of forces at every time step with the accumulation of the gradient estimators for variational optimization. For infinite-time conditioning, temporal self-averaging in the steady state can be used to improve the estimation of the gradients of the force with respect to the variational parameters. This is accomplished with the use of so-called Malliavin weights (102), which are auxiliary variables propagated along with the real dynamical variables in the system. Each Malliavin weight keeps the gradient of the trajectory probability with respect to a variational parameter. At the end of the steady-state trajectory, the gradient of the cost function is directly estimated from the correlation function of the Malliavin weights with the observable.

Tilted path ensemble: collection of trajectories reweighted with a linear bias

For complex, interacting systems where the basis sets span more than a few dimensions, the combinatorial explosion of the number of spatiotemporal variational parameters in learning implies that the statistical fluctuations in each cost-function gradient are too large to learn efficiently with a naive gradient descent algorithm. Specialized reinforcement learning tools have been developed for this purpose that learn not only the force but also a value function for the expected future cost starting from any point in space and time. This approach reduces the variance of the gradient estimate drastically by subtracting a zero-mean high-variance baseline. Such specialized reinforcement learning algorithms such as Monte Carlo, with a value baseline and the actor-critic algorithm, can learn an optimal controller sufficiently accurately that the reaction rate, scaled cumulant generating function, and reactive trajectory ensemble are quantitatively estimated (21, 26). Alternative optimization methods also exist for the construction of the effective approximate controllers. These include feedback algorithms (80, 81), the brute force enumeration of variational parameters (103), or evolutionary algorithms (104). Recently, the exact controller for many-particle systems in one and two dimensions has been constructed with tensor network basis sets (82, 83). This approach has still been limited to spatially discrete systems but holds promise in its treatment of fluctuations that are many body in origin.

4.3. Estimation of Rare Event Statistics

The optimal force is, in general, many-bodied, but a low-rank ansatz is often a tractable approximation. For example, a control force might be truncated at one- and two-bodied forces. We have found that such an approximate force is able to describe the principal mechanism leading to rare fluctuations, as discussed in the following sections. In such cases where the variational bound is nearly saturated, the approximation error from a limited basis set can be corrected for by estimating the exponential averages in Equations 21 and 26 either directly or through a cumulant expansion (20). In the case that both the driven and conditioned or tilted path ensembles are available, and the driving force approximation is accurate enough that there is sufficient overlap between the two distributions, as shown in **Figure 3*c***, the likelihood of the rare fluctuation can be evaluated using a generalization of the Bennett acceptance ratio (21, 23, 55).

There are cases where the bound is sufficiently unsaturated that converging the correction is not numerically possible. In this case, the calculation of the exponential averages can also be performed iteratively in small exponential increments with a cloning algorithm (33, 105). Using optimized forces from VPS speeds up the calculation by many orders of magnitude and has been

shown to outperform stratification-based sampling algorithms such as cloning and forward flux sampling (20, 21). Future work is needed on developing such hybrid, efficient techniques in finite time for calculating reaction rates.

5. APPLICATIONS TO RATE CONSTANTS

VPS has been used to estimate rate constants and characterize rare fluctuations, enabling mechanistic studies of reactive events both in and out of equilibrium. The algorithms in Section 4 offer various means for the computation of reactive rates from both conditioned and driven reactive trajectory ensembles, offering robust means for the estimation of reactive probabilities. Working in the conditioned ensemble enables the computation of the many-body reaction coordinates and, through the framework of stochastic optimal control, enables simple quantitative metrics for distilling rare fluctuations on a microscopic level of detail.

5.1. Quantifying Mechanisms

A long-standing goal in theoretical chemistry has been to develop methods to distill dynamics on a high-dimensional landscape into a few physically meaningful descriptors (48). Within the framework of transition path sampling (106) and transition path theory (76), the committor has been identified as the ideal reaction coordinate. However, as the solution of the backward Kolmogorov equation, the committor formally lives in the full dimensional space of the system.

Within this context, the application of theories of stochastic optimal control admits a rather alternative perspective into the analysis of these highly nonlinear functions, which can be summarized as follows: (*a*) Any conditioned Markov process can be mapped onto an unconditioned Markov process with an external drift that encapsulates all the natural fluctuations that induce the conditioning; (*b*) this external drift can be related to the conditional probability itself; and (*c*) the natural fluctuations that induce the conditioning can be quantified by the computation of the change of measure between systems with and without this external drift and can generally be decomposed into effective contributions from different degrees of freedom. As such, the formalism of VPS (20, 21, 23–25, 94) and of similar methods based on stochastic optimal control (85, 89) can be utilized to visualize and quantify rare fluctuations on a molecular level of detail.

5.1.1. Colloidal cluster dissociation. The simple relation between the optimal time-dependent controller that solves the generalized bridge problem and the splitting probability was employed in Reference 23 to study the dissociation of a particle from a heptameric cluster, originally inspired by Reference 41 and used in a range of different studies (107–109). The variational form in References 95, 97, and 100 was used to construct the splitting probability, and the optimal controller was used to generate reactive trajectories. **Figure 4*a*** shows a snapshot of a representative reactive trajectory with the optimized control forces. These control forces encapsulate the natural fluctuations of the system that make a trajectory reactive, and this system offers an illustrative example of how the formalism of VPS can be employed to visualize many-body fluctuations of rare events that involve collective motion. What emerged in this case is reminiscent of a T1 transition (110), a characteristic deformation pattern for switching neighbors in dense media. Utilizing this formalism for studying collective rare events such as nucleation and developing Monte Carlo moves for importance sampling of glassy systems offer promising avenues of future research.

5.1.2. Alanine dipeptide isomerization. For cases where the high dimensionality of the system restricts visualization of the forces, one can utilize the fact that the stochastic action due to the control forces $\langle \Delta\Gamma \rangle$ can be decomposed into effective contributions from different degrees of freedom to infer mechanistic insight. This framework affords a direct quantification of rare fluctuations

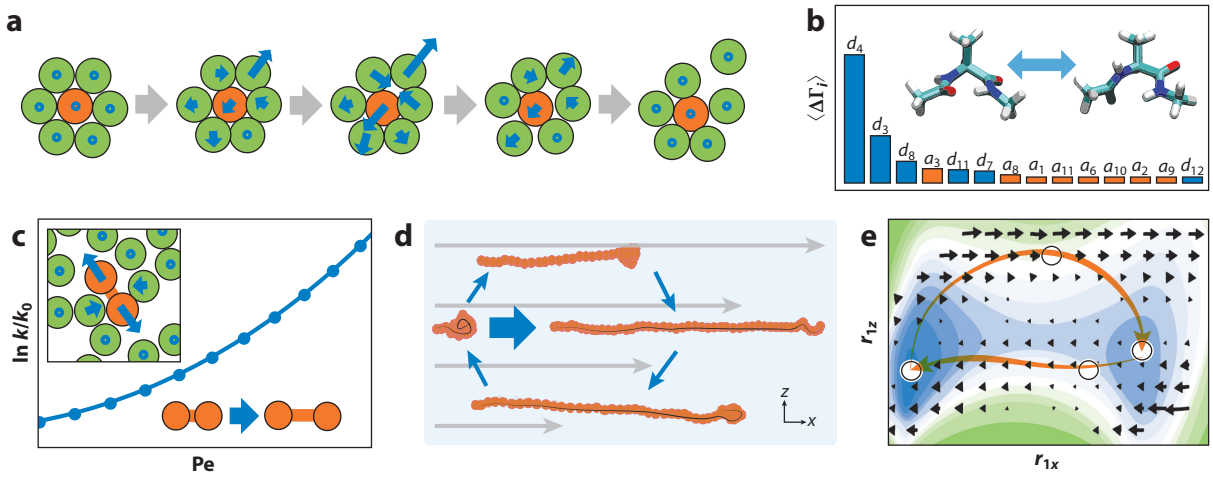

Figure 4

Mechanisms of reactive events clarified with variational path sampling. (*a*) The optimal control forces (*blue arrows*) for the dissociation (or association) of a self-assembled colloidal cluster demonstrate a T1-like mechanism. (*b*) The relative contributions of each dihedral (d_i) and angle (a_i) to the isomerization of alanine dipeptide. (*c*) Rate relative to its equilibrium value, k/k_0, with increasing activity [Péclet (Pe) number] for the conformational change of a passive dimer in an active fluid with optimal control forces in the inset (*blue arrows*). (*d*) The shear-induced unfolding of a polymer proceeds through distinct mechanisms in the folding and unfolding pathways due to persistent probability currents, as illustrated in panel *e* by the projection of the mean velocity (*black arrows*) and log likelihood (contour plot) onto two Rouse modes along the shear gradient r_{1z} and direction of the shear r_{1x}. Typical unfolding (*top*) and unfolding (*bottom*) pathways are shown in orange.

along different descriptors and can be transformed across different representations of the system, allowing simple ways to perform dynamically consistent coarse graining along physically relevant descriptors. This was shown in the study of the isomerization of alanine dipeptide that serves as a standard model for the exploration of unimolecular reactions in interacting systems (106). The reaction was explored by fitting the time-dependent committor (24) from trajectories obtained from transition path sampling, and a bond-angle-torsion transformation of the system was used to represent the set of descriptors. **Figure 4*b*** shows the conditioned action along the leading descriptors that were identified by this method. The most important descriptor was found to be the Ramachandran angle ϕ (d_4), consistent with previous studies on the system (106), and the analysis additionally highlighted that the relative motion of the alkyl bond was activated during the transition.

5.2. Rate Enhancement and Stability

A key aspect of the discussion in previous sections was that all the methods of analysis do not invoke detailed balance and can be employed to characterize reactive events in nonequilibrium steady states. In this section, we explore the application of VPS to two examples of reactions far from equilibrium that highlight features of nonequilibrium steady states that are strongly relevant to the endeavors of physical chemistry.

5.2.1. Dimer in an active solvent. A paradigmatic model for nonequilibrium soft matter is active matter, where energy is continuously injected to keep particles autonomously motile. Due to an intrinsic self-propulsion, active matter often exhibits unusual dynamics such as ballistic motion at short timescales (111), the formation of living clusters (112), swarming (113), and motile topological defects (114). Such dynamics are common across a range of length scales in biological

systems such as the cytoskeleton in cells, colonies of bacteria, and flocks of birds (115). Besides being useful models for living systems, active matter can be a tool for designing self-organized materials that show novel structure and function. A general quest has been to find ways to extract useful work from an active system. This is challenging because the directions of self-propulsion in neighboring active particles are not always aligned; thus a large portion of the self-propulsion energy is lost as dissipation through interparticle collisions. A new approach to solve the problem of harvesting energy from active systems has been to design passive solute particles or obstacles that are then driven by their active environment (116–118). It is in this spirit that VPS was recently applied to quantify the rate and mechanism of the isomerization of a colloidal dimer in a dense active fluid (20). As shown in **Figure 4*c***, the colloidal dimer has an extended and collapsed state. When it is placed in a bath of active Brownian particles and the self-propulsion of the bath is gradually increased, the rate of the reaction of the dimer from the collapsed to the extended state is enhanced quadratically by a factor of 20. The mechanism of the enhancement was revealed by the most probable reactive trajectories extracted from those driven with the optimized force. The mechanism is shown in the inset of **Figure 4*c*** and illustrates that a long-lived collective polarization fluctuation of the active fluid is responsible for sterically driving the extension reaction. This observation is in line with prior experiments on active particles adhering to the passive surface depending on the surface geometry (118) and demonstrated the ability of VPS to estimate rates and quantify many-body reaction mechanisms in nonequilibrium matter.

5.2.2. Attractive polymer in a shear flow. The emergence of the metastability of thermodynamically unfavorable states is often observed in systems that are driven out of equilibrium. The exploration of these systems is an active area of research in active matter (114, 119), self-assembly (117, 120, 121), and biophysics (122, 123). In the case where the emergent state is metastable, inference into the mechanism of the reaction is challenging due to the existence of persistent probability currents that can mediate reactions (108, 124). Standard approaches have leveraged effective equilibrium-based theories (125, 126), extensions of transition path theory to time-irreversible systems (127), and the computation of minimum action paths (128, 129) to infer mechanisms into reactions in nonequilibrium steady states.

An example of this phenomenology is a grafted polymer with strong attractive interactions under shear, which has been employed to gain insight into the role of the von Willebrand cofactor in the process of blood clotting (122, 130). These polymers have been observed to exhibit a globule-stretch transition at a critical shear, as shown in **Figure 4*d***, similar to the unfolding of the von Willebrand cofactor into thin fibers under physiologically relevant shear rates. Using the formalism of VPS and leveraging the transferability of the metric for discovering relevant dynamical coordinates, the leading normal modes r were identified as the best set of descriptors that were activated during globule-stretch transitions. Since the stretched state was thermodynamically unstable in equilibrium, it was not surprising that the nonequilibrium reaction coordinates were found to be orthogonal to the equilibrium reaction coordinate. Moreover, due to the breakdown of detailed balance, the forward and backward transitions in the nonequilibrium steady state were found to be distinct from each other, with the pathways passing through different intermediates, as shown in **Figure 4*d***. Both of these features can be visualized in **Figure 4*e***, which shows the negative log likelihood and net flux at finite shear along r_{1x} and r_{1z}, Rouse modes that correspond to the equilibrium and one of the nonequilibrium reaction coordinates, respectively, along with representative folding and unfolding reactive pathways. The system illustrates how probability currents can mediate reactions. Instead of a standard barrier crossing transition, the unfolding reaction involves coupling to strong unidirectional currents observed in **Figure 4*e***. Similarly, the folding reaction has a low free energy barrier that can be overcome through thermal fluctuations

and requires escape from the strong cyclic currents that confine the polymer within the extended state.

6. APPLICATIONS TO LARGE DEVIATION FUNCTIONS

Both the rate function and scaled cumulant generating function, as defined through Equations 11 and 23, emerge as the solutions to several variational principles. Variational principles exist because large deviations are achieved in the least rare of all rare pathways. This idea is called the contraction principle for large deviations (68). The contraction principle underpins the expressions for large deviation functions in terms of control forces that can achieve a given large deviation, such as for the scaled cumulant generating function in Equation 26. The saturation of Jensen's inequality with the control force $-\nabla\Lambda^*$, the solution of the Hamilton–Jacobi–Bellman equation (Equation 28), provides the mechanism of this least rare set of fluctuations. Solving for the optimal force thus gives us quantitative estimates for the rate functions and mechanism of the rare event.

6.1. Dynamic Phase Transitions

When the rare event involves the coexistence of distinct macroscopic dynamics for an extensive amount of time, such as the formation of a macroscopic cluster starting from a dispersed phase, it is said to be in a different dynamical phase. Transitions between dynamical phases with different values of a dynamical observable are associated with bistability in $I(\mathcal{O}/\tau)$ or singularities in derivatives of $\ln Z_\lambda$ (79, 131–133). For example, a discontinuity in the first derivative of the $\ln Z_\lambda$ for the current denotes the coexistence of distinct phases that have different mean values of the current.

A classic example of a dynamic phase transition occurs in active matter. Among the diverse phenomenology exhibited by active matter (115), the emergence of stable phase separation in purely repulsive particles, as shown in **Figure 5*a***, garnered significant theoretical attention. The phase diagram describing the so-called motility-induced phase separation was worked out as a function of the Péclet number, a measure of active to diffusive motions, using theoretical arguments and direct simulations (119). The entropy production was found to be a useful order parameter distinguishing the condensed from dilute phase, and controlling the entropy production could trigger phase separation (6, 86), at least for linear theories (134). In Reference 28, a general lower bound on distributions of entropy production for active matter systems was derived using the variational principles underpinning VPS. An approximate control force was used to sample entropy production fluctuations consisting of the exact solution to the single particle problem and using a linearization of a hydrodynamic model motivated from work on hard rods (135). It was found that for systems far away from dynamical phase transitions, the bound was tight, and its weakening signaled the importance of long-range correlations. This observation linked the locations of phase transitions to enhanced entropy production fluctuations as probed by a tilted ensemble, shown in **Figure 5*a***. Large deviation theory and finite size scaling of the rate function were used to motivate a generalization of a Maxwell construction between the two phases (**Figure 5*b***).

6.2. Long-Range Forces

The steady state of many-particle systems out of equilibrium, in general, has long-range correlations (136). Even for dynamics that do not lead to long-range correlations in their typical behavior, rare fluctuations of size-extensive variables generally result from long-range correlations between particles. The corresponding optimal control forces are long-range, often in the form of generalized Coulombic interactions (137). While analytical approximations for the optimal control forces can be obtained for small biases in the hydrodynamic limit using macroscopic fluctuation

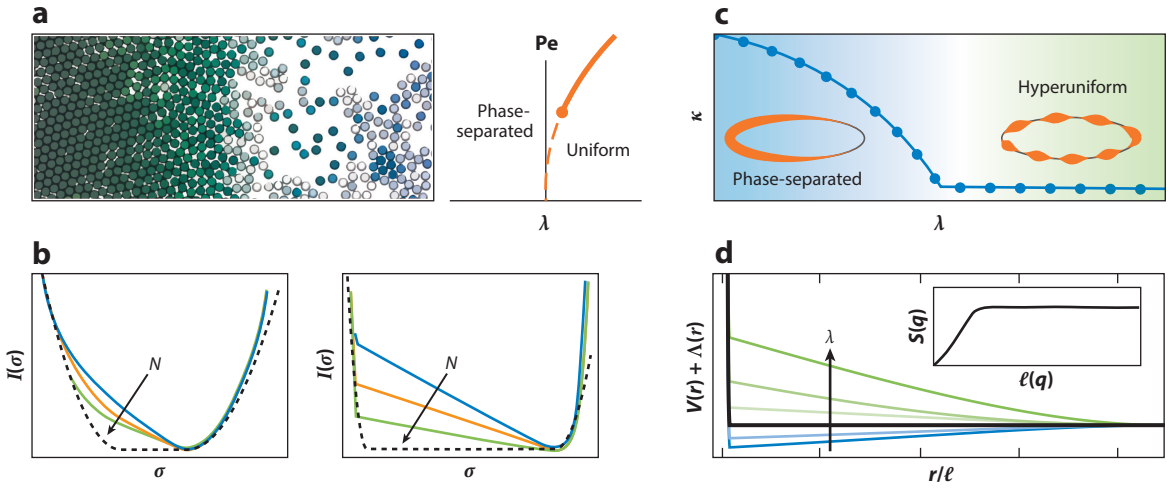

Figure 5

Large deviation function calculations with variational path sampling (VPS) include studies on (*a*) phase separation in active fluids, whose diagram depends on activity [Péclet (Pe) number] and the tilting parameter (λ). (*b*) Finite size scaling analysis of the entropy production rate functions with particle number N were computed from VPS (*solid lines*), with theoretical calculations yielding an asymptotic form (*dashed lines*) at low (*left*) and high (*right*) Pe numbers. (*c*) Soft Brownian particles on a line exhibit a dynamical phase transition as a function of the tilting parameter between phase-separated and hyperuniform states, stabilized by control forces shown in panel *d* whose range was much longer than the particle size (ℓ). (*d*, *inset*) Hyperuniformity is defined as the suppression of density fluctuations, evident in the behavior of the structure factor $S(q)$ at low wave vectors q.

theory (138), they are not available for complicated interparticle interactions. VPS can be used to numerically learn the optimal forces provided that a long-range ansatz is used in the basis functions.

VPS was applied to a model of overdamped repulsive particles in one dimension to characterize fluctuations of dynamical activity (κ). Dynamical activity measures the time-integrated propensity of particles to move and explore the space around them (139). Activity has been used as a dynamical order parameter to study the glassiness of materials (61), to evaluate the frequency of dissipative transitions in a model of trapped ions (140), and to characterize fluids that show long-range correlations despite disorder (79, 141). The last example, called a hyperuniform phase, arises in the one-dimensional model for low magnitudes of dynamical activity fluctuations. Similar models have been shown to have a dynamical phase transition at near-zero tilting, $\lambda \approx 0$, crossing between a hyperuniform phase and a clustered state with macroscopic phase separation (79, 135). Using macroscopic fluctuation theory, the hyperuniform phase can be shown to exhibit pairwise long-range repulsive interactions that push the particles away from each other so that collisions are minimized.

VPS enabled the quantification of the two phases outside of the regimes of the validity of approximate theory (20). When the ansatz for learning the optimal force allows for long-range interactions via basis functions of scaled Laguerre polynomials, VPS learns $(\ln Z_\lambda)/\tau$ and the correct rare trajectories leading to high (phase-separated) and low (hyperuniform) values of activity, as shown in **Figure 5*c***. The clustered phase in this case has a quadratic scaling of the activity with respect to system size, as the most long-lived clustered configurations are achieved by minimizing collisions on the surface of the cluster as opposed to in its center, where particles have collisions with both neighboring particles at once. In contrast, the hyperuniform phase consists of particles colliding only pairwise at a time and has a linearly scaling activity. The optimal force here

consists of a long-range repulsion that increases strongly with the bias and leads to a suppression of long-wavelength density fluctuations [the structure factor $S(q) \to 0$ as $q \to 0$], as shown in **Figure 5*d***.

7. FUTURE DIRECTIONS

The preceding sections describe applications from several branches of chemical, material, and biological physics. The approaches we review here allow for the calculation of rates and the evaluation of large deviation functions for high-dimensional, complex systems evolving arbitrarily away from equilibrium. However, there is still much work to be done to understand the structure of the theory underpinning VPS. The structure of the variational landscape in terms of control forces and its relation to the relaxation timescales of the system may affect the convergence of the variational algorithm, yet this behavior is not formally well-understood. Further, the application of VPS to some systems is no doubt still difficult owing to large statistical uncertainties in the formally computed parameter gradients. Understanding the best way to construct control forces spanned by low-dimensional collective coordinates is unclear. Improvements in their accuracy and more knowledge of the best statistical estimators to use in these cases would help in extending VPS to processes that are not spatially local like nucleation. Hybrid approaches in which optimized control forces are added to transition path sampling or the cloning algorithm (20, 21) have shown promise and warrant continued study. Extending these approaches to other methods like nonequilibrium instantons (129) or nonequilibrium umbrella sampling (2) would likely also be fruitful. Finally, extending VPS to the evaluation of other observables like stationary distribution functions or multitime functions would increase the scope of problems tractable within its application. Early work along the former direction has just begun (94). It is our hope that others' experience and diverse perspectives will lead to advances in these areas and unanticipated directions.

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 work reviewed in this article was supported largely by National Science Foundation grant CHE1954580 and the US Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, and Office of Basic Energy Sciences via the Scientific Discovery through Advanced Computing program. Several members of the Limmer group and other colleagues have contributed significantly to the development of the ideas described in this review, including Phillip Geissler, Juan Garrahan, Garnet Chan, Robert Jack, Benjamin Rotenberg, Dominic Rose, Ben Kuznets-Speck, Chloe Gao, Trevor GrandPre, Anthony Poggioli, Ushnish Ray, Jorge Rosa-Raíces, and Eric R. Heller.

LITERATURE CITED

1. Allen RJ, Valeriani C, Ten Wolde PR. 2009. Forward flux sampling for rare event simulations. *J. Phys. Condens. Matter* 21(46):463102
2. Dickson A, Dinner AR. 2010. Enhanced sampling of nonequilibrium steady states. *Annu. Rev. Phys. Chem.* 61:441–59
3. Bustamante C, Bryant Z, Smith SB. 2003. Ten years of tension: single-molecule DNA mechanics. *Nature* 421(6921):423–27
4. Bocquet L, Charlaix E. 2010. Nanofluidics, from bulk to interfaces. *Chem. Soc. Rev.* 39(3):1073–95

5. Needleman D, Dogic Z. 2017. Active matter at the interface between materials science and cell biology. *Nat. Rev. Mater.* 2(9):17048
6. Fodor É, Jack RL, Cates ME. 2022. Irreversibility and biased ensembles in active matter: insights from stochastic thermodynamics. *Annu. Rev. Condens. Matter Phys.* 13:215–38
7. Limmer DT, Gao CY, Poggioli AR. 2021. A large deviation theory perspective on nanoscale transport phenomena. *Eur. Phys. J. B* 94:145
8. Ciliberto S. 2017. Experiments in stochastic thermodynamics: short history and perspectives. *Phys. Rev. X* 7(2):021051
9. Limmer DT. 2024. *Statistical Mechanics and Stochastic Thermodynamics*. Oxford, UK: Oxford Univ. Press
10. Seifert U. 2019. From stochastic thermodynamics to thermodynamic inference. *Annu. Rev. Condens. Matter Phys.* 10:171–92
11. Touchette H. 2018. Introduction to dynamical large deviations of Markov processes. *Phys. A Stat. Mech. Appl.* 504:5–19
12. Evans DJ, Searles DJ. 2002. The fluctuation theorem. *Adv. Phys.* 51(7):1529–85
13. Crooks GE. 1999. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* 60(3):2721
14. Jarzynski C. 1997. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* 78(14):2690
15. Horowitz JM, Gingrich TR. 2020. Thermodynamic uncertainty relations constrain non-equilibrium fluctuations. *Nat. Phys.* 16(1):15–20
16. Baiesi M, Maes C. 2013. An update on the nonequilibrium linear response. *New J. Phys.* 15(1):013004
17. Speck T, Seifert U. 2006. Restoring a fluctuation-dissipation theorem in a nonequilibrium steady state. *Europhys. Lett.* 74(3):391
18. Gao CY, Limmer DT. 2019. Nonlinear transport coefficients from large deviation functions. *J. Chem. Phys.* 151(1):014101
19. Touchette H. 2009. The large deviation approach to statistical mechanics. *Phys. Rep.* 478(1–3):1–69
20. Das A, Limmer DT. 2019. Variational control forces for enhanced sampling of nonequilibrium molecular dynamics simulations. *J. Chem. Phys.* 151(24):244123
21. Das A, Kuznets-Speck B, Limmer DT. 2022. Direct evaluation of rare events in active matter from variational path sampling. *Phys. Rev. Lett.* 128(2):028005
22. Doob JL. 1942. The Brownian movement and stochastic equations. *Ann. Math.* 43(2):351–69
23. Singh AN, Limmer DT. 2024. Splitting probabilities as optimal controllers of rare reactive events. *J. Chem. Phys.* 161(5):054113
24. Singh AN, Limmer DT. 2023. Variational deep learning of equilibrium transition path ensembles. *J. Chem. Phys.* 159(2):024124
25. Singh AN, Limmer DT. 2025. Reactive path ensembles within nonequilibrium steady-states. arXiv:2501.19233 [physics.chem-ph]
26. Das A, Rose DC, Garrahan JP, Limmer DT. 2021. Reinforcement learning of rare diffusive dynamics. *J. Chem. Phys.* 155(13):134105
27. GrandPre T, Limmer DT. 2018. Current fluctuations of interacting active Brownian particles. *Phys. Rev. E* 98(6):060601
28. GrandPre T, Klymko K, Mandadapu KK, Limmer DT. 2021. Entropy production fluctuations encode collective behavior in active matter. *Phys. Rev. E* 103(1):012613
29. Chandler D. 1978. Statistical mechanics of isomerization dynamics in liquids and the transition state approximation. *J. Chem. Phys.* 68(6):2959–70
30. Binder K, Vollmayr K, Deutsch HP, Reger JD, Scheucher M, Landau DP. 1992. Monte Carlo methods for first order phase transitions: some recent progress. *Int. J. Mod. Phys. C* 3(05):1025–58
31. Kuznets-Speck B, Limmer DT. 2021. Dissipation bounds the amplification of transition rates far from equilibrium. *PNAS* 118(8):e2020863118
32. Ray U, Chan GK, Limmer DT. 2018. Importance sampling large deviations in nonequilibrium steady states. I. *J. Chem. Phys.* 148(12):124120
33. Ray U, Chan GKL, Limmer DT. 2018. Exact fluctuations of nonequilibrium steady states from approximate auxiliary dynamics. *Phys. Rev. Lett.* 120(21):210602

34. Ray U, Limmer DT. 2019. Heat current fluctuations and anomalous transport in low-dimensional carbon lattices. *Phys. Rev. B* 100(24):241409
35. Onsager L, Machlup S. 1953. Fluctuations and irreversible processes. *Phys. Rev.* 91(6):1505
36. Ruelle D. 1999. Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. *J. Stat. Phys.* 95:393–468
37. Bolhuis PG, Chandler D, Dellago C, Geissler PL. 2002. Transition path sampling: throwing ropes over rough mountain passes, in the dark. *Annu. Rev. Phys. Chem.* 53:291–318
38. Chandler D, Garrahan JP. 2010. Dynamics on the way to forming glass: bubbles in space-time. *Annu. Rev. Phys. Chem.* 61:191–217
39. Zwanzig R. 2001. *Nonequilibrium Statistical Mechanics*. Oxford, UK: Oxford Univ. Press
40. De Pirey TA, Cugliandolo LF, Lecomte V, Van Wijland F. 2022. Path integrals and stochastic calculus. *Adv. Phys.* 71(1–2):1–85
41. 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
42. Speck T. 2016. Thermodynamic formalism and linear response theory for nonequilibrium steady states. *Phys. Rev. E* 94(2):022131
43. Lesnicki D, Gao CY, Rotenberg B, Limmer DT. 2020. Field-dependent ionic conductivities from generalized fluctuation-dissipation relations. *Phys. Rev. Lett.* 124(20):206001
44. Gao CY, Limmer DT. 2017. Transport coefficients from large deviation functions. *Entropy* 19(11):571
45. Binder K, Ceperley DM, Hansen JP, Kalos M, Landau D, et al. 2012. *Monte Carlo Methods in Statistical Physics*, Vol. 7. Berlin: Springer
46. Mehdi S, Smith Z, Herron L, Zou Z, Tiwary P. 2024. Enhanced sampling with machine learning. *Annu. Rev. Phys. Chem.* 75:347–70
47. Hénin J, Lelièvre T, Shirts M, Valsson O, Delemotte L. 2022. Enhanced sampling methods for molecular dynamics simulations. *Liv. J. Comput. Mol. Sci.* 4(1):1583
48. Peters B. 2017. *Reaction Rate Theory and Rare Events*. Amsterdam: Elsevier
49. Guttenberg N, Dinner AR, Weare J. 2012. Steered transition path sampling. *J. Chem. Phys.* 136(23):234103
50. Gingrich TR, Geissler PL. 2015. Preserving correlations between trajectories for efficient path sampling. *J. Chem. Phys.* 142(23):234104
51. Grünwald M, Dellago C, Geissler PL. 2008. Precision shooting: sampling long transition pathways. *J. Chem. Phys.* 129(19):194101
52. Crooks GE, Chandler D. 2001. Efficient transition path sampling for nonequilibrium stochastic dynamics. *Phys. Rev. E* 64(2):026109
53. Buijsman P, Bolhuis P. 2020. Transition path sampling for non-equilibrium dynamics without predefined reaction coordinates. *J. Chem. Phys.* 152(4):044108
54. Zuckerman DM, Chong LT. 2017. Weighted ensemble simulation: review of methodology, applications, and software. *Annu. Rev. Biophys.* 46:43–57
55. Frenkel D, Smit B. 2023. *Understanding Molecular Simulation: From Algorithms to Applications*. Amsterdam: Elsevier
56. Doob JL. 1984. *Classical Potential Theory and Its Probabilistic Counterpart*. New York: Springer
57. Girsanov IV. 1960. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory Probab. Appl.* 5(3):285–301
58. Gallavotti G, Cohen EGD. 1995. Dynamical ensembles in stationary states. *J. Stat. Phys.* 80:931–70
59. Dellago C, Bolhuis PG, Chandler D. 1999. On the calculation of reaction rate constants in the transition path ensemble. *J. Chem. Phys.* 110(14):6617–25
60. Garrahan JP, Jack RL, Lecomte V, Pitard E, van Duijvendijk K, van Wijland F. 2009. First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories. *J. Phys. A Math. Theor.* 42(7):075007
61. Hedges LO, Jack RL, Garrahan JP, Chandler D. 2009. Dynamic order-disorder in atomistic models of structural glass formers. *Science* 323(5919):1309–13
62. Hurtado PI, Garrido PL. 2011. Spontaneous symmetry breaking at the fluctuating level. *Phys. Rev. Lett.* 107(18):180601

63. Baek Y, Kafri Y, Lecomte V. 2017. Dynamical symmetry breaking and phase transitions in driven diffusive systems. *Phys. Rev. Lett.* 118(3):030604
64. Vaikuntanathan S, Gingrich TR, Geissler PL. 2014. Dynamic phase transitions in simple driven kinetic networks. *Phys. Rev. E* 89(6):062108
65. Garrahan JP, Lesanovsky I. 2010. Thermodynamics of quantum jump trajectories. *Phys. Rev. Lett.* 104(16):160601
66. Jack RL, Garrahan JP, Chandler D. 2006. Space-time thermodynamics and subsystem observables in a kinetically constrained model of glassy materials. *J. Chem. Phys.* 125(18):184509
67. Giardina C, Kurchan J, Lecomte V, Tailleur J. 2011. Simulating rare events in dynamical processes. *J. Stat. Phys.* 145:787–811
68. Chetrite R, Touchette H. 2015. Variational and optimal control representations of conditioned and driven processes. *J. Stat. Mech. Theory Exp.* 2015(12):P12001
69. Doob JL. 1957. Conditional Brownian motion and the boundary limits of harmonic functions. *Bull. Soc. Math. Fr.* 85:431–58
70. Orland H. 2011. Generating transition paths by Langevin bridges. *J. Chem. Phys.* 134(17):174114
71. Koehl P, Orland H. 2022. Sampling constrained stochastic trajectories using Brownian bridges. *J. Chem. Phys.* 157(5):054105
72. Majumdar SN, Orland H. 2015. Effective Langevin equations for constrained stochastic processes. *J. Stat. Mech. Theory Exp.* 2015(6):P06039
73. Hartmann C, Schütte C. 2012. Efficient rare event simulation by optimal nonequilibrium forcing. *J. Stat. Mech. Theory Exp.* 2012(11):P11004
74. Hartmann C, Banisch R, Sarich M, Badowski T, Schütte C. 2013. Characterization of rare events in molecular dynamics. *Entropy* 16(1):350–76
75. Zhang W, Wang H, Hartmann C, Weber M, Schütte C. 2014. Applications of the cross-entropy method to importance sampling and optimal control of diffusions. *SIAM J. Sci. Comput.* 36(6):A2654–72
76. 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
77. Rose DC, Mair JF, Garrahan JP. 2021. A reinforcement learning approach to rare trajectory sampling. *New J. Phys.* 23(1):013013
78. Chetrite R, Touchette H. 2015. Nonequilibrium Markov processes conditioned on large deviations. *Ann. Henri Poincaré* 16:2005–57
79. Jack RL, Sollich P. 2015. Effective interactions and large deviations in stochastic processes. *Eur. Phys. J. Spec. Top.* 224(12):2351–67
80. Nemoto T, Sasa Si. 2014. Computation of large deviation statistics via iterative measurement-and-feedback procedure. *Phys. Rev. Lett.* 112(9):090602
81. Nemoto T, Bouchet F, Jack RL, Lecomte V. 2016. Population-dynamics method with a multicanonical feedback control. *Phys. Rev. E* 93(6):062123
82. Causer L, Bañuls MC, Garrahan JP. 2023. Optimal sampling of dynamical large deviations in two dimensions via tensor networks. *Phys. Rev. Lett.* 130(14):147401
83. Banuls MC, Garrahan JP. 2019. Using matrix product states to study the dynamical large deviations of kinetically constrained models. *Phys. Rev. Lett.* 123(20):200601
84. Helms P, Ray U, Chan GKL. 2019. Dynamical phase behavior of the single-and multi-lane asymmetric simple exclusion process via matrix product states. *Phys. Rev. E* 100(2):022101
85. Yan J, Touchette H, Rotskoff GM, et al. 2022. Learning nonequilibrium control forces to characterize dynamical phase transitions. *Phys. Rev. E* 105(2):024115
86. Nemoto T, Fodor É, Cates ME, Jack RL, Tailleur J. 2019. Optimizing active work: dynamical phase transitions, collective motion, and jamming. *Phys. Rev. E* 99(2):022605
87. Kocer E, Ko TW, Behler J. 2022. Neural network potentials: a concise overview of methods. *Annu. Rev. Phys. Chem.* 73:163–86
88. Keith JA, Vassilev-Galindo V, Cheng B, Chmiela S, Gastegger M, et al. 2021. Combining machine learning and computational chemistry for predictive insights into chemical systems. *Chem. Rev.* 121(16):9816–72

89. Boffi NM, Vanden-Eijnden E. 2024. Deep learning probability flows and entropy production rates in active matter. *PNAS* 121(25):e2318106121
90. Holdijk L, Du Y, Hooft F, Jaini P, Ensing B, Welling M. 2023. Stochastic optimal control for collective variable free sampling of molecular transition paths. In *Advances in Neural Information Processing Systems*, Vol. 36, ed. A Oh, T Naumann, A Globerson, K Saenko, M Hardt, S Levine, pp. 79540–56. Red Hook, NY: Curran Assoc.
91. Kang P, Trizio E, Parrinello M. 2024. Computing the committor with the committor to study the transition state ensemble. *Nat. Comput. Sci.* 4:451–60
92. Pavliotis GA. 2014. *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. New York: Springer
93. Hartmann C, Latorre JC, Zhang W, Pavliotis GA. 2014. Optimal control of multiscale systems using reduced-order models. *J. Comput. Dyn.* 1(2):279–306
94. Rosa-Races JL, Limmer DT. 2024. Variational time reversal for free-energy estimation in nonequilibrium steady states. *Phys. Rev. E* 110(2):024120
95. Rotskoff GM, Mitchell AR, Vanden-Eijnden E. 2022. Active importance sampling for variational objectives dominated by rare events: consequences for optimization and generalization. *Proc. Mach. Learn. Res.* 145:757–80
96. Thiede EH, Giannakis D, Dinner AR, Weare J. 2019. Galerkin approximation of dynamical quantities using trajectory data. *J. Chem. Phys.* 150(24):244111
97. Hasyim MR, Batton CH, Mandadapu KK. 2022. Supervised learning and the finite-temperature string method for computing committor functions and reaction rates. *J. Chem. Phys.* 157(18):184111
98. 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.* 3:334–45
99. Strahan J, Finkel J, Dinner AR, Weare J. 2023. Predicting rare events using neural networks and short-trajectory data. *J. Comput. Phys.* 488:112152
100. Khoo Y, Lu J, Ying L. 2019. Solving for high-dimensional committor functions using artificial neural networks. *Res. Math. Sci.* 6:1
101. Liang S, Singh AN, Zhu Y, Limmer DT, Yang C. 2023. Probing reaction channels via reinforcement learning. *Mach. Learn. Sci. Technol.* 4(4):045003
102. Warren PB, Allen RJ. 2012. Malliavin weight sampling for computing sensitivity coefficients in Brownian dynamics simulations. *Phys. Rev. Lett.* 109(25):250601
103. Jacobson D, Whitelam S. 2019. Direct evaluation of dynamical large-deviation rate functions using a variational ansatz. *Phys. Rev. E* 100(5):052139
104. Whitelam S, Jacobson D, Tamblyn I. 2020. Evolutionary reinforcement learning of dynamical large deviations. *J. Chem. Phys.* 153(4):044113
105. Giardina C, Kurchan J, Peliti L. 2006. Direct evaluation of large-deviation functions. *Phys. Rev. Lett.* 96(12):120603
106. Bolhuis PG, Dellago C, Chandler D. 2000. Reaction coordinates of biomolecular isomerization. *PNAS* 97(11):5877–82
107. Das A, Limmer DT. 2021. Variational design principles for nonequilibrium colloidal assembly. *J. Chem. Phys.* 154(1):014107
108. Das A, Limmer DT. 2023. Nonequilibrium design strategies for functional colloidal assemblies. *PNAS* 120(40):e2217242120
109. Yuan J, Shah A, Bentz C, Cameron M. 2024. Optimal control for sampling the transition path process and estimating rates. *Commun. Nonlinear Sci. Numer. Simul.* 129:107701
110. Hasyim MR, Mandadapu KK. 2021. A theory of localized excitations in supercooled liquids. *J. Chem. Phys.* 155(4):044504
111. Redner GS, Hagan MF, Baskaran A. 2013. Structure and dynamics of a phase-separating active colloidal fluid. *Biophys. J.* 104(2):640a (Abstr.)
112. Mognetti BM, Šarić A, Angioletti-Uberti S, Cacciuto A, Valeriani C, Frenkel D. 2013. Living clusters and crystals from low-density suspensions of active colloids. *Phys. Rev. Lett.* 111(24):245702
113. Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O. 1995. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* 75(6):1226

114. Marchetti MC, Joanny JF, Ramaswamy S, Liverpool TB, Prost J, et al. 2013. Hydrodynamics of soft active matter. *Rev. Mod. Phys.* 85(3):1143
115. Ramaswamy S. 2010. The mechanics and statistics of active matter. *Annu. Rev. Condens. Matter Phys.* 1:323–45
116. Grünwald M, Tricard S, Whitesides GM, Geissler PL. 2016. Exploiting non-equilibrium phase separation for self-assembly. *Soft Matter* 12(5):1517–24
117. Mallory SA, Valeriani C, Cacciuto A. 2018. An active approach to colloidal self-assembly. *Annu. Rev. Phys. Chem.* 69:59–79
118. Sokolov A, Apodaca MM, Grzybowski BA, Aranson IS. 2010. Swimming bacteria power microscopic gears. *PNAS* 107(3):969–74
119. Cates ME, Tailleur J. 2015. Motility-induced phase separation. *Annu. Rev. Condens. Matter Phys.* 6:219–44
120. Nguyen M, Qiu Y, Vaikuntanathan S. 2021. Organization and self-assembly away from equilibrium: toward thermodynamic design principles. *Annu. Rev. Condens. Matter Phys.* 12:273–90
121. Palacci J, Sacanna S, Steinberg AP, Pine DJ, Chaikin PM. 2013. Living crystals of light-activated colloidal surfers. *Science* 339(6122):936–40
122. Alexander-Katz A, Schneider M, Schneider S, Wixforth A, Netz R. 2006. Shear-flow-induced unfolding of polymeric globules. *Phys. Rev. Lett.* 97(13):138101
123. Tan TH, Mietke A, Li J, Chen Y, Higinbotham H, et al. 2022. Odd dynamics of living chiral crystals. *Nature* 607(7918):287–93
124. Fang X, Kruse K, Lu T, Wang J. 2019. Nonequilibrium physics in biology. *Rev. Mod. Phys.* 91(4):045004
125. Redner GS, Wagner CG, Baskaran A, Hagan MF. 2016. Classical nucleation theory description of active colloid assembly. *Phys. Rev. Lett.* 117(14):148002
126. Cates M, Nardini C. 2023. Classical nucleation theory for active fluid phase separation. *Phys. Rev. Lett.* 130(9):098203
127. Helfmann L, Ribera Borrell E, Schütte C, Koltai P. 2020. Extending transition path theory: periodically driven and finite-time dynamics. *J. Nonlinear Sci.* 30(6):3321–66
128. Zakine R, Vanden-Eijnden E. 2023. Minimum-action method for nonequilibrium phase transitions. *Phys. Rev. X* 13(4):041044
129. Heller ER, Limmer DT. 2024. Evaluation of transition rates from nonequilibrium instantons. *Phys. Rev. Res.* 6:043110
130. Schneider S, Nuschele S, Wixforth A, Gorzelanny C, Alexander-Katz A, et al. 2007. Shear-induced unfolding triggers adhesion of von Willebrand factor fibers. *PNAS* 104(19):7899–903
131. Buča B, Garrahan JP, Prosen T, Vanicat M. 2019. Exact large deviation statistics and trajectory phase transition of a deterministic boundary driven cellular automaton. *Phys. Rev. E* 100(2):020103
132. Garrahan JP, Jack RL, Lecomte V, Pitard E, van Duijvendijk K, van Wijland F. 2007. Dynamical first-order phase transition in kinetically constrained models of glasses. *Phys. Rev. Lett.* 98(19):195702
133. Espigares CP, Garrido PL, Hurtado PI. 2013. Dynamical phase transition for current statistics in a simple driven diffusive system. *Phys. Rev. E* 87(3):032115
134. Tociu L, Fodor É, Nemoto T, Vaikuntanathan S. 2019. How dissipation constrains fluctuations in nonequilibrium liquids: diffusion, structure, and biased interactions. *Phys. Rev. X* 9(4):041026
135. Dolezal J, Jack RL. 2019. Large deviations and optimal control forces for hard particles in one dimension. *J. Stat. Mech. Theory Exp.* 2019(12):123208
136. Spohn H. 1983. Long range correlations for stochastic lattice gases in a non-equilibrium steady state. *J. Phys. A Math. Gen.* 16(18):4275
137. Popkov V, Schütz GM, Simon D. 2010. Asep on a ring conditioned on enhanced flux. *J. Stat. Mech. Theory Exp.* 2010(10):P10007
138. Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G, Landim C. 2015. Macroscopic fluctuation theory. *Rev. Mod. Phys.* 87(2):593
139. Pitard E, Lecomte V, Van Wijland F. 2011. Dynamic transition in an atomic glass former: a molecular-dynamics evidence. *Europhys. Lett.* 96(5):56002
140. Ates C, Olmos B, Garrahan JP, Lesanovsky I. 2012. Dynamical phases and intermittency of the dissipative quantum Ising model. *Phys. Rev. A* 85(4):043620
141. Torquato S. 2018. Hyperuniform states of matter. *Phys. Rep.* 745:1–95