Editors' Suggestion

Numerical computation of effective thermal equilibria in stochastically switching Langevin systems

Benjamin L. Walker and Katherine A. Newhall*
University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA

(Received 16 September 2021; accepted 4 May 2022; published 14 June 2022)

Stochastically switching force terms appear frequently in models of biological systems under the action of active agents such as proteins. The interaction of switching forces and Brownian motion can create an "effective thermal equilibrium," even though the system does not obey a potential function. In order to extend the field of energy landscape analysis to understand stability and transitions in switching systems, we derive the quasipotential that defines this effective equilibrium for a general overdamped Langevin system with a force switching according to a continuous-time Markov chain process. Combined with the string method for computing most-probable transition paths, we apply our method to an idealized system and show the appearance of previously unreported numerical challenges. We present modifications to the algorithms to overcome these challenges and show validity by demonstrating agreement between our computed quasipotential barrier and asymptotic Monte Carlo transition times in the system.

DOI: 10.1103/PhysRevE.105.064113

I. INTRODUCTION

Biological systems under the influence of microscale active agents such as proteins are frequently modeled using switching forces as the agents shift between different states [1]. Examples include molecular motors [2], crosslinked biopolymer networks [3], and transient antibody crosslinking of antigens to mucus protein networks [4,5]. Protein action also plays a crucial role in the organization of DNA inside the cell nucleus, both in the form of stochastic crosslinking [6–9], and more recently, protein loop extrusion [10–12]. Modeling these active agents in combination with passive diffusion leads to mathematical models with two sources of noise—stochastically switching forces combined with stochastic Brownian motion.

In the case of crosslinking proteins, prior work in the literature on modeling the dynamic organization of DNA with polymer bead-spring models included stochastically switching spring forces between beads representing 5 kbp of DNA [13,14]. These rapidly switching forces are on timescales faster than the time to reach thermal equilibrium; thus the system is in a constant state of disequilibrium. However, in [14] we observed long-lived stable condensed clusters of beads consistent with experimental results, with the stochastic switching rate acting like an effective temperature. Rapid switching produced low-temperature-like stable clusters, slow switching produced high-temperature-like amorphic arrangements, and intermediate switching times allowed for dynamic clusters with beads exchanging between clusters.

To explain the mechanism behind this emergent clustering behavior, we seek an effective thermal equilibrium. Recall that if the forces in a system, $v(x) = -\nabla U(x)$, are the gradient of a potential function $U(x)$, the dynamics governed by the stochastic differential equation (SDE)

$$dX = v(X)dt + \sqrt{2k_BT}dW \tag{1}$$

approach, in the long time limit, a Boltzmann thermal equilibrium distribution given by

$$p(x) \sim \exp\left(-\frac{U(x)}{k_BT}\right). \tag{2}$$

States x that minimize $U(x)$ are long-lived stable configurations at temperatures small enough relative to the energy barriers of $U(x)$ separating such states, defined as the gap ΔU between the energy at the minimum and the energy at the lowest saddle point on the region of attraction of the minimum. The mean transition time between the wells surrounding the energy-minimizing states can be computed asymptotically for vanishing temperature ($k_BT \to 0$): the time τ taken for the system to escape from a potential well under Brownian noise relates asymptotically to the energy barrier of the well following the Arrhenius equation given by

$$\mathbb{E}[\log \tau] \sim \frac{\Delta U}{k_BT}. \tag{3}$$

The most-probable path (MPP) the system traverses as it makes one such transition can also be found asymptotically for vanishing temperature; it is a path that is everywhere parallel to the gradient of the energy landscape $U(x)$.

The study of energy landscapes and large deviations allows for valuable insights to be made by viewing systems through the lens of statistical thermodynamics. In the cases of either stochastic switching of deterministic forces alone [15,16] or nongradient forces with thermal noise [17–19], the derivation of a quasipotential in the small temperature limit successfully predicts equilibrium distributions as well as transition times and paths. We seek an effective equilibrium that takes into account both the thermal fluctuations and the stochasticity

*knewhall@unc.edu

induced by the switching forces in the above-mentioned DNA bead-spring model.

A first thought to find an effective potential function $U(x)$ is to simply time-average the force, thereby removing the switching. However, this approach significantly overestimates the effective strength of a strong force when the switching is not very fast. Considering a thought experiment of an infinitely high potential barrier that is only sometimes on, we can see that no matter the manner of the switching, the barrier remains infinitely high after averaging. However, as the particle could cross the barrier due to diffusion while it is off, transmission is clearly possible, contrary to the expectation of the naive time-averaging.

This time-averaging takes the noise from the stochastic switching to zero first and then considers the effects of thermal noise. Rather, we seek a distinguished limit that takes the switching timescale to zero simultaneously with the thermal noise. To simultaneously consider both sources of randomness while leveraging the power of the energy landscape framework, we compute a quasipotential $W(x)$ whose gradient represents an average force that generalizes the asymptotic properties of the potential function to nongradient systems. We build off the work in Ref. [16] that used a Wentzel-Kramers-Brillouin (WKB) approximation to construct a quasipotential for the Morris-Lecar equation, an ordinary differential equation whose evolution depends on a stochastically changing number of open ion gates. In our case we consider both a switching state modeled by a continuous-time Markov chain and diffusive noise from Brownian motion. We extend the WKB ansatz approach of [16] to form the Hamilton-Jacobi equation for our problem.

The quasipotential $W(x; x_A)$ differs from a global potential function in that it is only defined in the basin of attraction of a particular fixed point x_A of the deterministic system. Since we are interested in using the quasipotential to predict expected transition times, we seek its values along the MPP the system escapes along. As with MPPs for gradient systems, this path connects x_A to a saddle point and is everywhere parallel to the gradient of the quasipotential. To simultaneously find this path and the quasipotential along it we use the string method [20] and its climbing variant [21]. These string methods place a number of copies of the system, or images, along a path in phase space to form a "string." Each image is independently updated via gradient descent, and then the images are reinterpolated along the path to keep them equally spaced in arc length. Thus the path aligns itself with the gradient descent direction, converging to the MPP. In the climbing variant, the final image "climbs" in energy in the direction tangent to the string, and gradient descends in all others. Thus this final image converges to a saddle point, with the remaining images parametrizing the MPP from this saddle point.

In this work, we must couple the evolving string with finding the quasipotential. Therefore we develop a numerical scheme to iteratively solve for the quasipotential along a path, then update the path based on the found quasipotential until it converges to the MPP.

While previous work generally looked at one- to three-dimensional problems [16,22–27], we demonstrate an example for three particles in two dimensions for a total of six spatial dimensions. Under our proposed framework, several aspects of the problem present numerical challenges that were not reported in the previous literature. This numerical instability arises in both the implicit solver for the gradient of the quasipotential and the solver for the MPP. We find that Newton's method does not generally converge from the starting guesses we can make, and so we use a modified version of Newton's method with an additional fallback designed to ensure convergence. This also leads to high sensitivity in the Hessian matrix, which is used in the geometric minimum action method (gMAM) for finding the MPP [16,28]. This explains the choice of using the string method instead, which does not require the Hessian matrix. Still, we see significant high-frequency noise in the string method update, arising for larger numbers of images in the string. We present a set of numerical methods, along with code, that overcomes all of these challenges and demonstrate validity by showing agreement between quasipotential barriers and Monte Carlo escape time asymptotics.

In Sec. II, we derive the form of the Hamiltonian for our problem and describe the numerical methods we apply to solve for the gradient of the quasipotential and most-probable transition paths. In Sec. III we validate the formulation of the Hamiltonian on a one-dimensional problem in which the string method is not necessary. In Sec. IV we then apply the full method to compute transition paths and asymptotic escape times for a system of three particles moving in two dimensions, revealing an important physical principle, that the interaction of stochastic switching and Brownian noise leads to a weaker effective force than would be expected by simple averaging. Finally, in Sec. V we review our contributions and note future directions of research.

II. FORMULATION

In this section we detail the steps required to compute the gradient of the quasipotential ∇W along transition paths. We start by introducing a general model system of the form of Eq. (1) but with a deterministic force term that switches based on a continuous-time Markov chain. Then in Sec. II B we derive the Hamiltonian and the associated Hamilton-Jacobi equation that defines the gradient of the quasipotential ∇W for our problem. In Sec. II C we describe the algorithm that we use to address the additional numerical challenges in this version of the problem and solve the Hamilton-Jacobi equation for ∇W. In Sec. II D we describe the numerical procedure for finding the most-probable transition paths which minimize the action. In the following sections we apply our method to two example systems.

A. Model framework

Taking Eq. (1) and introducing a parameter $\epsilon = k_B T$, we modify the forces $v(x)$ to switch between different states to arrive at the general mathematical form of the equation for the dynamics. The configuration of the system is represented as the combination of a position $X_t \in \mathbb{R}^m$ and a switching state index $s_t \in \mathbb{Z}_1^n$ for some dimensionality m and number of possible switching states n. The time evolution of the position follows an overdamped Langevin equation given by

$$dX_t = v(X_t; s_t)dt + \sqrt{2\epsilon}dW, \tag{4}$$

where the switching state $s_t \in \{1, ..., n\}$ affects the drift term. The time evolution of the state s_t follows a continuous-time Markov Chain (CTMC) process whose transition rate matrix $\frac{1}{\epsilon}S(X_t)$ depends on position. In this way the processes governing position and switching state are coupled.

We chose the scaling of $\frac{1}{\epsilon}$ in front of S to produce interaction between the CTMC switching noise and the $\epsilon = k_B T$ noise from Brownian motion at lowest order asymptotically, thus preparing to take a distinguished limit of small noise. This is also a physically relevant scaling for biological systems where the timescale of these switching forces is likely temperature dependent. We compare our results to the case of faster scaling of the CTMC [for example, $\frac{1}{\epsilon^2}S(X_t)$], finding this limit significantly overestimates the effective energy barrier.

B. Deriving the Hamilton-Jacobi equation

The Hamilton-Jacobi equation that defines the quasipotential arises from making a WKB-like quasi-steady-state assumption for the solution to a system of Fokker-Planck equations that are coupled by the transitions between switching states. We start by defining these coupled Fokker-Planck equations. Let the force (drift) on the ith position coordinate under switching configuration s be represented as v_i^s. Recall the CTMC process transition matrix elements S_{jk} hold the transition rate into state j from state k. Finally, let $p_s(x,t)$ represent the joint probability function between the discrete variable s and the continuous position variables of each bead,

$$p_s(x,t) = \rho(x,t|s_t = s)\mathrm{P}(s_t = s) \quad \text{for } s = 1, 2, \ldots n, \quad (5)$$

where $\rho(x,t|s_t = s)$ is the conditional density for the process X_t at time t given that the force state is currently in state s.

Each individual p_s will follow a Fokker-Planck equation associated with that state's drift term from Eq. (4), with an additional coupling term to represent transitions between states. These coupled Fokker-Planck equations take the form

$$\frac{\partial p_s}{\partial t} = -\sum_{i=1}^{m} \frac{\partial}{\partial x_i}\left[v_i^s p_s\right] + \epsilon \sum_{i=1}^{m} \frac{\partial^2}{\partial x_i^2}[p_s] + \frac{1}{\epsilon}\sum_{k=1}^{n} S_{sk} p_k, \quad (6)$$

with the associated steady-state equation

$$0 = -\sum_{i=1}^{m} \frac{\partial}{\partial x_i}\left[v_i^s p_s\right] + \epsilon \sum_{i=1}^{m} \frac{\partial^2}{\partial x_i^2}[p_s] + \frac{1}{\epsilon}\sum_{k=1}^{n} S_{sk} p_k \quad (7)$$

for $s = 1, 2, \ldots n$.

If the drift term v did not depend on the switching process s_t, the Langevin process for X_t would have a potential function constructed from a path integral of v. However, in our formulation in which the drift function v exhibits random switching, it is no longer the gradient of a potential function, and so such a U cannot be found. Intuitively, in the small-noise limit, $\epsilon \to 0$, in which the magnitude of the Brownian noise goes to zero as the rate of the stochastic switching goes to infinity, there is no diffusion and the forces are extremely rapidly switching, creating an effective force given by the average of the switching forces v^s over the steady-state distribution of the switching matrix.

The above informs our choice of the WKB-like ansatz for the steady-state distribution $p_s(x)$ in the form

$$p_s(x) = r_s(x)\exp\left(-\frac{1}{\epsilon}W(x)\right) \quad (8)$$

for $s = 1 \ldots n$ as $\epsilon \to 0$, similar to the one employed in [16] for purely stochastically switching forces (no diffusion). We see here that $W(x)$ takes the place of the potential $U(x)$ in Eq. (2); $W(x)$ is the quasipotential. The preexponential term r_s superimposes the different states s.

We then plug Eq. (8) into Eq. (7), seeking equations for $r_s(x)$ and $W(x)$ given by the order $\frac{1}{\epsilon}$ terms, which are the lowest order in ϵ. In this way, Eq. (8) differs from a typical WKB expansion in which there would be no preexponential term present at lowest order. This r_s term only encapsulates the relationship between states—the full preexponential term would emerge at higher order in ϵ.

The resulting order $\frac{1}{\epsilon}$ equation has the form

$$M(x, \nabla W)r(x) = 0, \quad (9)$$

where the matrix M depends on both the position x and the gradient of the quasipotential ∇W. The vector $r(x) = (r_1, r_2, \ldots, r_n)^T$ has components $r_s(x)$ for each state s. Details are shown in Appendix B 1. The $n \times n$ matrix M is given by

$$M(x, \nabla W) = D(\nabla W) + A(x, \nabla W) + S(x), \quad (10)$$

the sum of three matrices corresponding to the three terms in Eq. (7): D, the diffusion matrix; A, the advection matrix; and S, the switching matrix. Note that S is unchanged from its original definition as the CTMC transition rate matrix and serves the purpose of coupling the different states, while D and A are diagonal matrices with diagonal elements for each switching state $s = 1 \ldots n$ given by

$$D_{ss} = \sum_{i=1}^{m}\left(\frac{\partial W}{\partial x_i}\right)^2 \quad \text{and} \quad A_{ss} = \sum_{i=1}^{m} v_i^s \frac{\partial W}{\partial x_i}. \quad (11)$$

We can observe that in order to have nontrivial solutions to the system (9) we must have that $\det M(x, \nabla W) = 0$. Choosing the Hamiltonian as the greatest eigenvalue of M, $\mathcal{H}(x,p) = \max \lambda$ s.t. $M(x,p)u = \lambda u$, we see that having nontrivial solutions to the system (9) is equivalent to the Hamilton-Jacobi equation, $\mathcal{H}(x,p) = 0$. Note that we have introduced a new variable p as the second argument by analogy to the typical form of the Hamilton-Jacobi equation. This means that our solution for ∇W will be given by the value of p that solves the Hamilton-Jacobi equation.

Combining the above with an additional curl-zero constraint, we obtain the pair of equations

$$\mathcal{H}(x, \nabla W(x)) = 0, \quad (12)$$

$$\nabla \times \nabla W(x) = 0, \quad (13)$$

that uniquely defines the gradient of the quasipotential. In this work we will compute the quasipotential along MPPs parameterized as $\phi(s) : [0,1] \to \mathbb{R}^m$. On the interior of the path $0 < s < 1$, Eq. (12) will define a convex surface, and the curl constraint that defines the unique solution can be replaced

with a constraint evaluated only on the path given by

$$\nabla_p \mathcal{H}(x,p)|_{p=\nabla W(x)} \parallel \frac{d\phi}{ds}. \tag{14}$$

We will require that the endpoints of this path are fixed points of the "deterministic dynamics," defined by the limit $\epsilon \to 0$ of Eq. (4). Recall that in this limit the Brownian noise term vanishes, and the CTMC switching rates go to infinity such that the system always exists in a superposition of states consistent with the steady state of the switching matrix $S(x)$ at its current position. These dynamics can be expressed as

$$\begin{aligned} \frac{dx_i}{dt} &= \sum_{k=1}^{n} v_i^k r_k \text{ for } i = 1 \ldots m \\ r &= \text{null } S \qquad \sum_{k=1}^{n} r_k = 1, \end{aligned} \tag{15}$$

where v_i^s defines the force on coordinate i when in switching state s, and r is a null vector of $S(x)$ appropriately normalized so that it represents the steady-state distribution of the CTMC at fixed position x. While $\mathcal{H}(x,0) = 0$ everywhere, we show in Appendix B3 that at fixed points of the deterministic dynamics, $\mathcal{H}(x,0) = 0$ is a minimum of $\mathcal{H}$, and therefore the unique solution (note that $\mathcal{H}$ is a convex function). Thus, at the endpoints we necessarily have $\nabla W = 0$, and on the interior ∇W is defined by the simultaneous solution of Eqs. (12) and (14).

Solving Eq. (14) for the MPP ϕ can be done using variational methods such as the string method [20]. However, in the nongradient case, as the quasipotential can be thought of as propagating along caustics, we do not know the true value of the quasipotential except on an MPP. This creates the need for an iterative algorithm that alternates between estimating ∇W using the current value of ϕ and using the computed values of ∇W to update the path ϕ. We discuss this procedure in more detail next.

C. Solving for ∇W

We will now describe the numerical optimization procedure by which we obtain ∇W, the gradient of the quasipotential, through which we obtain the final quasipotential by numerical integration. This will require formulating the simultaneous solution of Eqs. (12) and (14) as an optimization problem using Lagrange multipliers and then plugging in the Hamiltonian constructed in Sec. II B.

Our approach for computing ∇W is based on the Newton's method equations in [16], which are presented for an arbitrary Hamiltonian $\mathcal{H}$, and also require the gradient and Hessian of $\mathcal{H}$ with respect to the momentum variables p [see Eqs. (C1) and (C2)]. Using the Hamiltonian derived in Sec. II B along with the differentiation formulas shown in Appendix B2, we are able to apply these equations to our problem.

In practice, we have observed that direct application of these Newton's method equations often fails to converge for the problems we have attempted. To address this, we have added a fallback scheme that is designed to make guaranteed iterative improvements until the region in which Newton's method converges is reached.

This fallback method reframes Eqs. (12) and (14) as the problem of maximizing the dot product with $\frac{d\phi}{ds}$ on the convex surface $\mathcal{H}(x,p) = 0$. The fallback method moves the current guess a small distance in the direction of $\frac{d\phi}{ds}$ projected onto the normal of the surface $\mathcal{H}(x,p) = 0$, and then applies Newton's method to return to a solution of Eq. (12), $\mathcal{H}(x,p) = 0$. Note that the application of Newton's method in the fallback is only on Eq. (12), as opposed to the outer Newton's method which seeks a simultaneous solution to Eqs. (12) and (14).

If the outer Newton's step fails to improve the quality of the solution, this fallback is used instead, which is guaranteed to produce a better solution. In practice, the fallback is able to quickly bring the guess close enough to the true solution for Newton's method to begin to converge quadratically. Additional details of this routine are written in Appendix C 1.

D. Computing minimum action paths

Previous work on computing quasipotentials has either computed the quasipotential along MPPs (for example, [15,16]) or on a grid using an upwind scheme [16,26]. As we are interested in transition asymptotics, we restrict this work to considering computing quasipotentials along MPPs, and so the routine for computing the quasipotential must simultaneously search for an MPP.

An MPP, $\phi(s)$, is a curve in the configuration space of the system parameterized by s, typically taken to be the normalized arc length along the curve. It connects two points x_a, x_b such that $\phi(0) = x_a$, $\phi(1) = x_b$. Numerically, the path $\phi(s)$ is discretized into a sequence of "images," $\phi_1, \phi_2, \phi_3, \ldots, \phi_N$, representing the state of the physical system along the transition path. These images are chosen so that the arc length is constant between images, e.g., $||\phi_k - \phi_{k-1}|| = \text{const}$.

To compute the MPP, we make an initial guess of the transition path (typically linearly interpolated) and then apply the string method [20], in which each image along the path is moved a small distance opposite to the direction of ∇W and then interpolated along the path to remain equally spaced in arc length. Compared to typical applications of the string method, there is an additional step of applying the implicit solver to find the value of ∇W at each image, which itself depends on the current direction of the string.

Because we are looking for escape trajectories from a well, we allow the final image of the string to move according to the climbing string method [21]. The final image moves in a different direction: letting ∇W_N refer to the gradient of the quasipotential at the final image and $d\phi_N$ refer to the tangent direction of the string at the final image, the direction for moving the final image is given by

$$\nabla W_N - (1+\alpha)\text{proj}_{d\phi_N} \nabla W_N \tag{16}$$

for some $\alpha > 0$. This can be interpreted as descending in the directions orthogonal to $d\phi_N$, but climbing in the direction parallel to $d\phi_N$, as the projection term inverts that component. Details of the algorithm are presented in Appendix C 2.

The more general method for finding transition paths in nongradient systems is gMAM [16,28]. This method also time-evolves a discretized path but requires additionally computing the Hessian of the Hamiltonian, in our case defined as the largest eigenvalue of the matrix M. In practice, we

noticed the interaction between the largest and second-largest eigenvalues led to high sensitivity of the Hamiltonian on the state of the system and corresponding instability in the gMAM iteration. For this reason we chose to work with the string method to find MPPs.

Once the MPP $\phi(s)$ is found, together with ∇W along this path, the quasipotential barrier is defined via integration,

$$\Delta W = \int_0^1 \nabla W[\phi(s)]|\phi'(s)|ds.$$

We compute this numerically using the trapezoid scheme with centered differences for computing $\phi'(s)$.

III. ONE-DIMENSIONAL CASE

To demonstrate our formulation of the Hamiltonian (Appendix B1) and correspondingly, the quasipotential, we begin by applying our method to the one-dimensional case of a single bead subject to a constant excluded volume force and an on-off switching attractive force pulling it towards the origin. We note that the presence of only a single spatial dimension means the transition path is trivially known; this frees us from the additional optimization step of computing the most-probable transition path out of the minimum.

In this case there are only two states of the switching force corresponding to whether the attractive force is switched off or on—we label these forces v^1 and v^2, respectively, in line with the notation introduced in Sec. II B. They are given by

$$v^1(x) = a_{ev}x\exp\left(-\frac{x^2}{c_{ev}}\right),$$

$$v^2(x) = a_{ev}x\exp\left(-\frac{x^2}{c_{ev}}\right) - kx.$$

In this example we use parameters $k = 5$, $a_{ev} = 3$, $c_{ev} = 0.5$. We note that the method is independent of the choice of these parameters, and the choice is simply motivated to create a well with a basin of attraction extending to $|x| \approx 1$.

Correspondingly, the transition rate matrix takes the form

$$\frac{1}{\epsilon}S = \frac{1}{\epsilon}\begin{pmatrix} -a(x) & c \\ a(x) & -c \end{pmatrix}, \tag{17}$$

with $a(x)$ being some decreasing function of distance from the origin, and c a constant that is here taken to be 0.5. Note the $1/\epsilon$ scaling of the CTMC to couple the switching rates with the magnitude of the Brownian motion, thus allowing us to take the limit as both fluctuations go to zero with $\epsilon \to 0$.

We consider three different choices of $a(x)$, chosen to create quasipotential barriers of different heights, to allow us to show agreement in three different cases. These functions $a(x)$ are

$$a_1(x) = 2e^{-3|x|^2}, \tag{18}$$

$$a_2(x) = \frac{2}{1+e^{20(|x|-0.75)}}, \tag{19}$$

$$a_3(x) = \frac{4}{1+e^{20(|x|-0.75)}}. \tag{20}$$

Because the most-probable transition path can only move in the single dimension of the problem, we simply compute the quasipotential along some interval $(0, x_0)$ such that x_0 is to the right of the "saddle point" of the quasipotential (which in one dimension is, in fact, a maximum). For the above choices, the maximum is in the vicinity of $x = 1$, so we compute values through $x_0 = 2$.

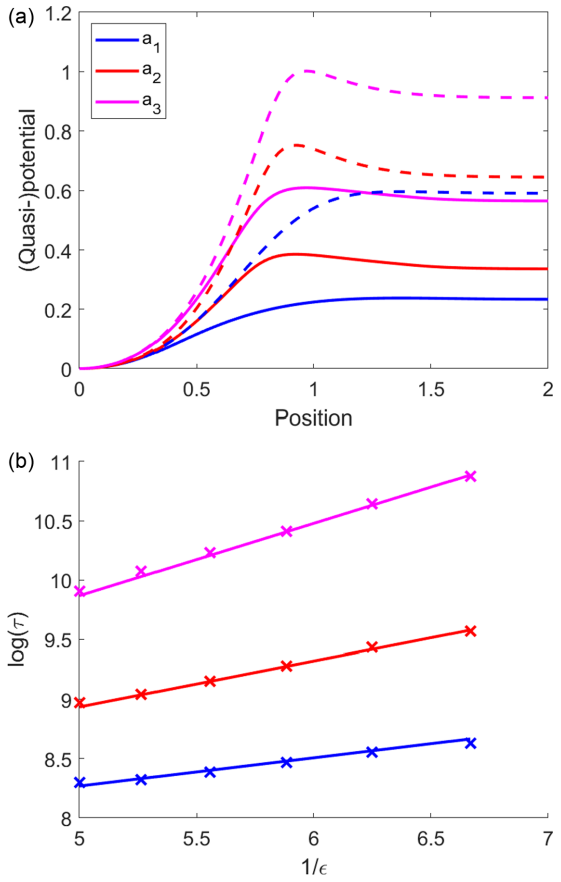

FIG. 1. Comparison of quasipotential and escape time asymptotics for the three different affinity functions given by Eqs. (18)–(20). (a) Comparison of quasipotential (solid) along the string with the deterministic average (dashed) that time averages the switching force, illustrating that the deterministic average would not agree with Monte Carlo statistics. (b) Average Monte Carlo escape times (points) overlaid on a solid line whose slope is given by the quasipotential barrier, showing full agreement between Monte Carlo statistics and asymptotic prediction. Logarithms are taken base e.

In order to validate the size of the computed quasipotential barrier, we compare to escape times, computed as described in Appendix A, using the modified Euler-Maruyama method. Note that the transition rate from state $s = 1$ to $s = 2$ depends on the changing variable x, and so waiting times are resampled each time step; the rate c is constant, and so the waiting time from state $s = 2$ to $s = 1$ can be preserved until it is reached.

Figure 1(a) shows the results of computing the quasipotential for each of the three affinity functions, along with the deterministic energy computed by numerically integrating the deterministic force from Eq. (15). This shows that while the two agree on the location of the highest point of the barrier, they disagree significantly on its height. Figure 1(b) shows the average Monte Carlo escape times for different values of ϵ, compared to lines whose slope is given by the

height of the quasipotential barrier. The quasipotential barrier height clearly predicts well the slope followed by the escape times, confirming that our formulation of the quasipotential is consistent with the theory and preparing us to add the string descent element in higher-dimensional problems. We note that the theoretical intercept of these lines would follow from higher-order asymptotics. Here, it is chosen by hand to fit the data.

We also recall the thought experiment of the infinitely high switching barrier from the introduction. Figure 1 confirms that the interaction between Brownian noise and stochastic switching leads the deterministic average to significantly overestimate the effective barrier, in this case by a factor of roughly two. This reinforces the need to properly consider the distinguished asymptotic limit in both sources of noise.

IV. TWO-DIMENSIONAL CASE

In this section we detail the model of three beads moving in two dimensions that exhibits the same mixing behavior that originally inspired this work. We characterize the four stable states of the system and show that the system spends most of its time in those states in which a pair of the beads is clustered together. We compute a most-probable escape path from one of these states and show that it agrees with Monte Carlo simulated escape times for transitions between two clustered states.

A. Model

Our interest in stochastically switching Langevin systems is originally motivated by prior work on modeling dynamics of chromosomes inside a yeast nucleus [14]. We consider here a toy model using the same functional forms of forces, while noting that the methods apply in general to any type of stochastically switching force between particles.

We consider three beads, affected by three forces: bonding, repulsion (excluded volume), and a global confinement force. These are collectively sufficient to produce multiple stable equilibria in the system. The confinement force is given by

$$f_c^i(x_i) = -\eta x_i, \tag{21}$$

the excluded volume force is given by

$$f_{\mathrm{EV}}^i(\{x_j\}_j) = \sum_{j\neq i} a_{ev}(x_i - x_j)\exp\left[-\frac{(x_i - x_j)^2}{c_{ev}}\right], \tag{22}$$

and the binding attraction force is given by

$$f_{\mathrm{bond}}^i(\{x_j\}_j) = \sum_{j\neq i} k b_{ij}(x_j - x_i), \tag{23}$$

where $b_{ij} = 1$ if beads i and j are presently bonded and zero otherwise. We use parameter values $k = 5, a_{ev} = 2, c_{ev} = 0.5$, and $\eta = 1$. With this notation, the entries b_{ij} stochastically switch as beads bind and unbind. Thus only the binding force f_{bond} is stochastic; the others are deterministic.

This gives the following SDE for the position of bead i:

$$dX_i = \left(f_c^i + f_{\mathrm{EV}}^i + f_{\mathrm{bond}}^i\right)dt + \sqrt{2\epsilon}\,dW, \tag{24}$$

where ϵ is a small positive parameter that controls the amount of stochasticity in the system and also appears in the switching, as described below.

The stochastic switching of the bonding term models the crosslinking proteins that bind two nearby beads. Each bead can be either unbound or bound to a single other bead. Bonds are symmetric. If a bond is formed between two beads, assume that the lifetime of the bond is an exponentially distributed random variable with rate c, meaning it has an expected lifetime of $\frac{1}{c}$. To simulate it, one can simply draw such an exponentially distributed random variable and use it as the lifetime.

It is natural to assume that crosslinking proteins would be more likely to bind beads that are closer together. Therefore, by analogy to the form of Eq. (17) we include an "affinity function" $a(r)$ dependent on pairwise distances for the binding rate. Specifically, $a(r)$ gives the (exponential process) rate at which a bond forms between two currently unbound beads i, j with positions x_i, x_j that are separated by a distance $r = |x_i - x_j|$. We note that this rate is only meaningful until a bond forms—a rate of $a(r)$ is equivalent to stating that in an infinitesimally short time dt, there is a probability $\frac{dt}{a(r)}$ that a bond forms. However, note that as the system moves in time, these probabilities will change accordingly. In this section we will use $a = a_2$, given in Eq. (19), which is

$$a(x) = \frac{2}{1 + e^{20(|x|-0.75)}}.$$

Following the framework of Sec. II A, the CTMC switches between the different binding configurations. These binding configurations corresponding to the states of the Markov chain are enumerated as all three beads unbound ($s = 1$), bead 1 bound to 2 ($s = 2$), bead 1 bound to 3 ($s = 3$), bead 2 bound to 3 ($s = 4$). The corresponding transition rate matrix S takes the form

$$\frac{1}{\epsilon}S = \frac{1}{\epsilon}\begin{pmatrix} b & c & c & c \\ a(x_1 - x_2) & -c & 0 & 0 \\ a(x_1 - x_3) & 0 & -c & 0 \\ a(x_2 - x_3) & 0 & 0 & -c \end{pmatrix}, \tag{25}$$

with $b = -a(x_1 - x_2) - a(x_1 - x_3) - a(x_2 - x_3)$. Recall again the fixed rate $c = 0.5$ describes bonds breaking.

To illustrate the qualitative behavior of this constructed model system, Fig. 2 demonstrates a sample simulated trajectory. In particular, we can see that this model replicates the "mixing" property that motivated this research, with rapid switching between which pair of beads (analogous to a cluster) are currently close. This motivates the further investigation into the behavior of this system, and the stability of these cluster states.

B. Computing most-probable escape paths

We now proceed to compute MPPs out of a basin of attraction using the quasipotential climbing string method. We begin by initializing a climbing string to search from the two-bead cluster minimum, with two beads in the same location and the third further away.

Compared to previous work utilizing a string method, our problem exhibits a greater level of numerical instability. In

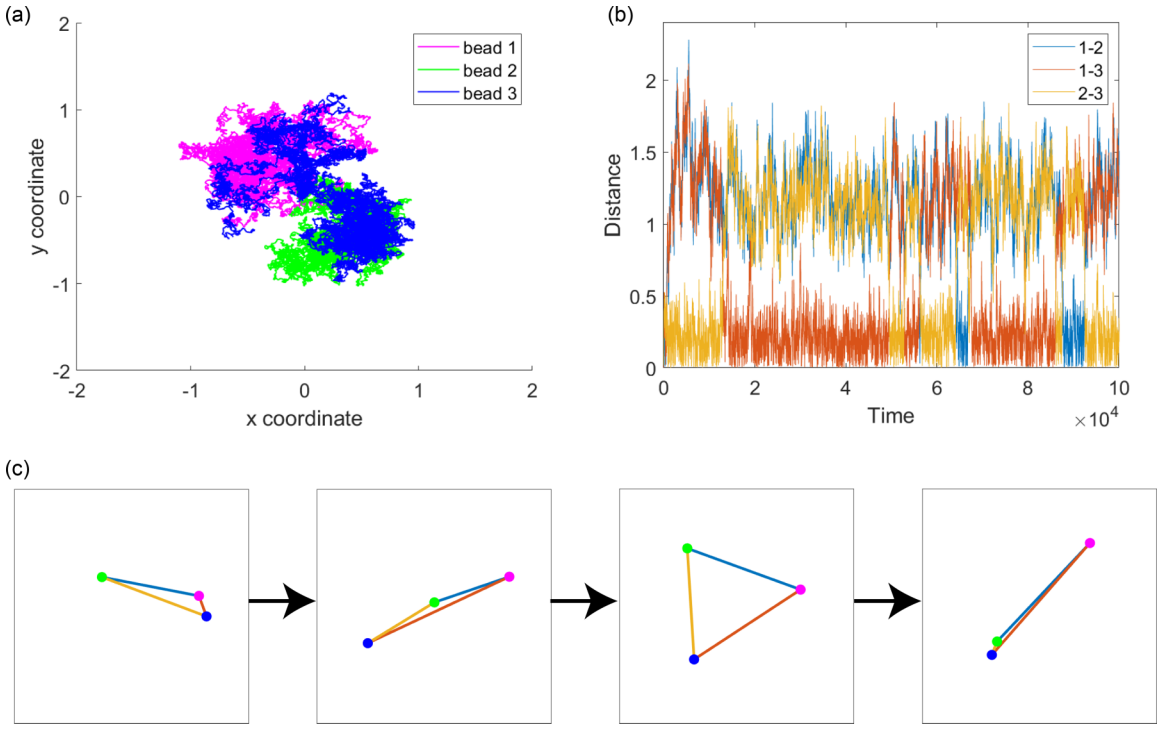

FIG. 2. (a) Trajectories taken by three beads over the first 20 000 time steps of a simulation. (b) Plot of pairwise distances between the three beads over time. Observe that at any time, one distance is small (between the two currently clustered beads) and the other two distances are large, with which pair is bound rapidly switching. (c) Snapshots of the simulation illustrating possible transitions between two clustered states. The nearby beads separate, arranging into a line. Another possible intermediate state is a triangle. Finally, two beads again approach and enter a cluster state.

particular, we notice high-frequency error along the string with each image zigzagging back and forth. This coincides with a lack of convergence, as shown in Fig. 3(a), where we plot the total change over all images since the previous iteration for different numbers of images along the string. Only the string with 10 images shows convergence. While reducing the time-step size for evolving the string helped with convergence, we noticed that reducing the number of images had a greater effect on convergence. Further evidence for the lack of convergence for strings with more than 10 images appears in Fig. 3(b). This figure shows that the computed barrier height of the quasipotential continues to evolve with more iterations, except for the string with 10 images.

Figure 4(a) shows the computed escape trajectory from the two-bead cluster minimum to the saddle point of three colinear beads, using a string with 10 images along it. This is consistent with the observation in Fig. 2(c), that transitions out of the two-bead cluster state pass through an intermediate state in which the three beads are colinear. In fact, the beads remain colinear along the entire transition path; therefore in Fig. 4(b), we plot only the y coordinate of the beads along the MPP, as the x coordinate remains unchanged. Note that our system is rotationally symmetric, so any transition from the cluster to the colinear state is simply a rotation of this transition taking place along the y axis.

Figure 4(c) shows a different computed escape trajectory from the two-bead cluster to a saddle point shown by the black circles with two beads still on top of each other, but closer to the third bead, before separating into the stable triangle state. This spatially longer transition path was computed using a string with 30 images along it. We point out that this placed about 10 images between the starting minimum and saddle point, which covers about the same spatial distance as the MPP in Fig. 4(a). This hints at a possible relationship between the physical distance between images and the time step to evolve the images to ensure convergence, similar to some finite difference methods for evolving partial differential equations. Further investigation is required to determine the exact relationship between convergence and the number of images along a string.

C. Monte Carlo escape statistics

We validate our computation of the quasipotential barrier by comparing it to Monte Carlo simulations of the model test system in Eq. (24). As mentioned in Sec. IV A, our test system exhibits switching between permutations of the stable cluster state, with two beads close together and one far away; see Fig. 2(b). By design, it spends minimal time in other stable configurations such as the triangle, as we are particularly interested in transitions between the stable cluster states.

To compute the Monte Carlo escape times, we initialize the simulation with beads 1 and 2 close and bound together, and measure the time until a different pair of beads form a cluster by a criterion that the newly clustered beads should be separated by a distance under 0.3 and the original pair is separated by a distance of at least 1. While this criterion represents entering a new cluster state, as opposed to simply

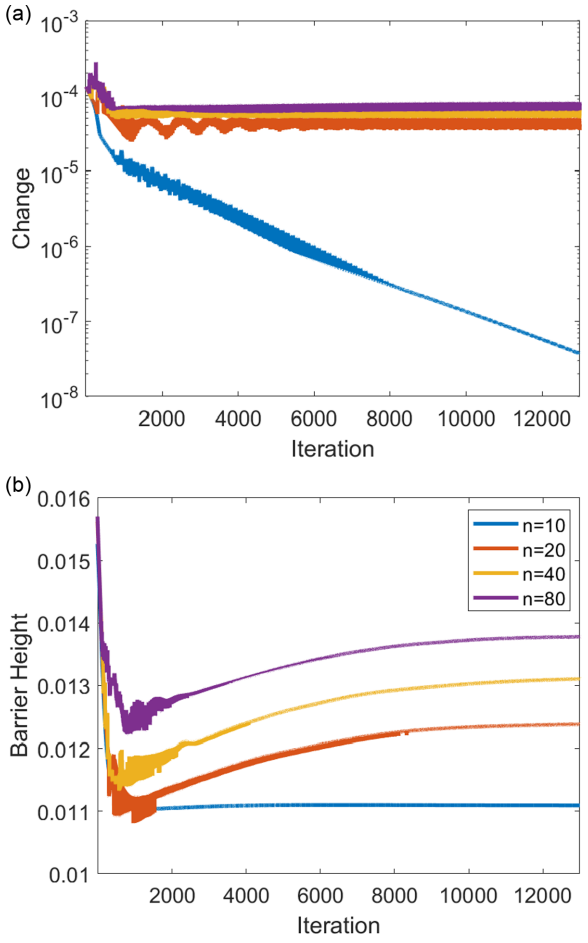

FIG. 3. (a) Maximum change over all images from previous iteration, showing convergence only in the case of 10 images. (b) Quasipotential barrier height over iterations, showing that the $n = 10$ converges to a barrier height of approximately 0.011.

reaching the edge of the basin of attraction, these two events take the same amount of time asymptotically to the lowest (logarithmic) order. Due to the multidimensionality of the system combined with the stochastic switching force, directly determining the edge of the basin of attraction is not trivial. Further details on the numerical simulations and estimate of the mean escape time τ are given in Appendix A.

In Fig. 5(a) we compute the quasipotential (solid line) along the MPP shown in Fig. 4(a), connecting the two-bead cluster to the saddle point of the quasipotential. For comparison, we also include the deterministic average potential (dashed line) computed by averaging the stochastically switching force over the steady-state distribution of the Markov chain, thereby taking the switching noise to zero first before considering the effect of the Brownian noise. Note that in both these cases this effective potential is in the six-dimensional configuration space of the three-bead system; it is not a pairwise effective potential between two beads. We find the quasipotential barrier ΔW to be equal to approximately 0.011, whereas the deterministic average is an order of magnitude greater at approximately 0.096. This illustrates the

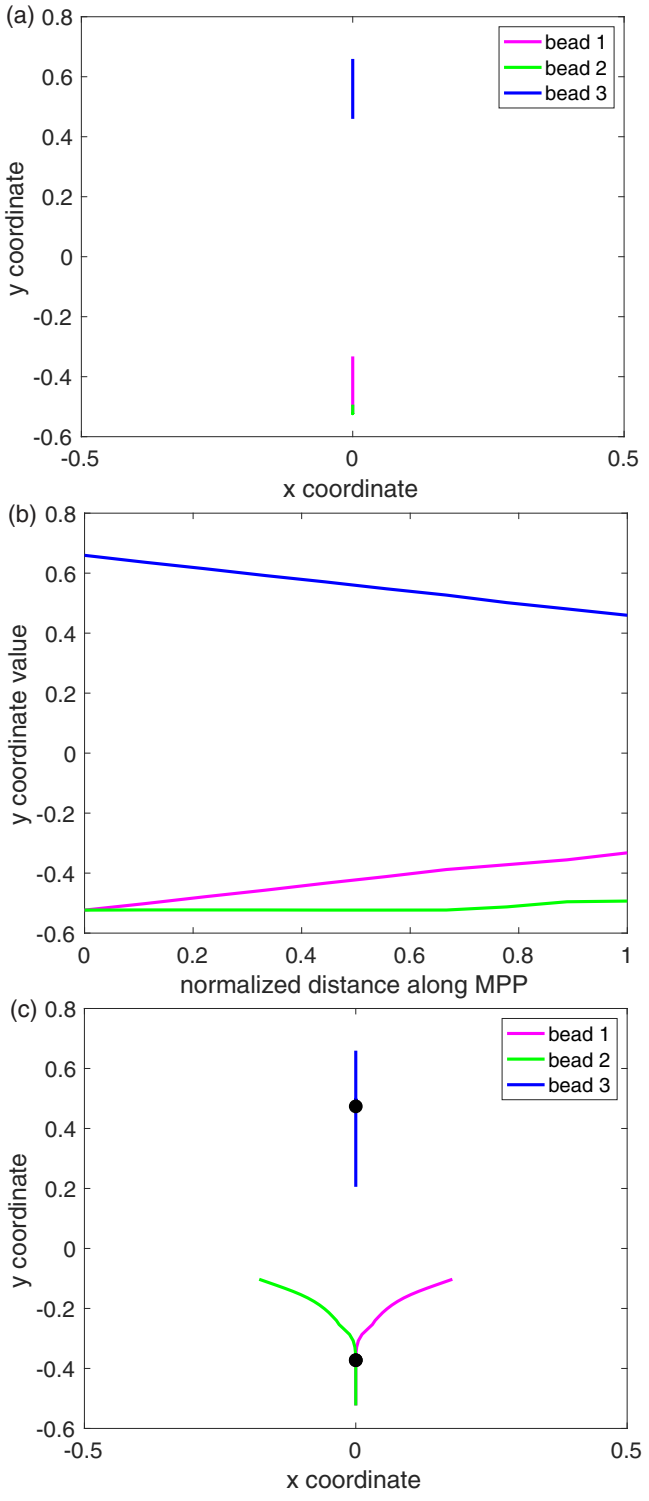

FIG. 4. Visualization of escape paths computed using quasipotential string descent. (a) Path connecting the two-bead cluster state [Fig. 2(c), first panel] to the saddle point of three colinear beads [Fig. 2(c), second panel]. (b) The y coordinate only of the transition in (a) as the x coordinate remains unchanged. Note that the behavior is not simply linear, especially visible in the case of bead 2 (green). (c) Path connecting the two-bead cluster state [Fig. 2(c), first panel] through a different saddle point indicated by the black circles to the triangle state [Fig. 2(c), third panel].

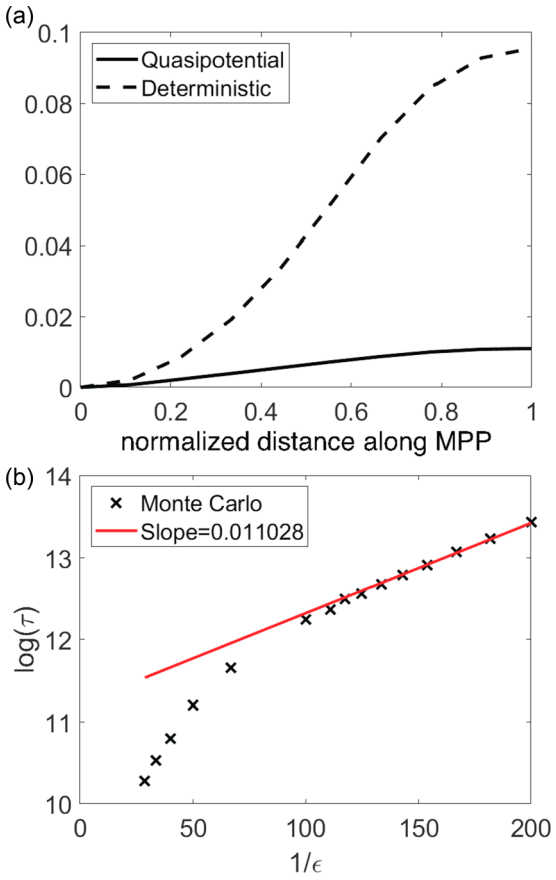

FIG. 5. (a) Quasipotential (solid) and deterministic average (dashed) along transition path from Fig. 4(a). Quasipotential barrier height is approximately 0.011. (b) Comparison of asymptotic escape times computed via Monte Carlo simulation (points) to slope taken from quasipotential barrier height (red line).

disagreement between the quasipotential and the deterministic average even more strikingly than in the one-dimensional case.

In Fig. 5(b) we test if the Monte Carlo escape times are well described by the quasipotential theory that $\log \tau \sim \Delta W/\epsilon$. We run simulations for a collection of values of $\epsilon \in [0.005, 0.035]$ and compute the mean escape time τ for each ϵ as described in Appendix A. For values of $1/\epsilon \geqslant 100$ ($\epsilon < 0.01$), the escape times are linear in the log-log plot, as predicted by the asymptotic relation, with a best-fit slope of approximately 0.118. We then compare this to the quasipotential barrier from Fig. 5(a), drawing a red line whose slope is given by the theory and fitting the y intercept. The slope of the Monte Carlo points agrees essentially perfectly with the line, demonstrating the validity of our quasipotential. We importantly note that this slope of approximately 0.011 clearly disagrees both with the deterministic average (approximately 0.096) and the quasipotential that would be derived along a linear trajectory between minimum and saddle (approximately 0.015), as opposed to our nonlinear path shown in Fig. 4(b). Thus we are confident that our numerical method converged to the correct MPP and quasipotential and that this asymptotic limit is the correct way to capture the dynamics of metastable switching systems under Brownian noise.

The main source of numerical error in our computation comes from integrating the gradient of the quasipotential along the string. This error decreases with more images, but more images cause the convergence issues mentioned previously, including that the computed barrier continues to increases as more iterations are performed. To get a sense of the error, we also computed the same quasipotential using 6 through 12 images along the string, finding the quasipotential barrier to vary from 0.0108 (6 images) to 0.0115 (12 images). These all round to the value of 0.011 reported above.

V. DISCUSSION

In this paper we extended the theory of quasipotentials to systems of Langevin equations with an additional source of stochasticity arising from a continuous-time Markov process switching the force term, a natural mathematical framework for modeling systems in biology under the effect of protein binding and unbinding mechanics. We demonstrated that under the interaction of switching forces and Brownian motion, an effective potential is generated that is weaker than predicted by a simple time averaging of the switching force. We further demonstrated that the quasipotential represents an effective thermal equilibrium in the low-noise regime and can be used to predict asymptotic transition times between metastable states created by such switching forces. This sets the stage for further analysis of switching models of biological systems, representing states created by thermal action of proteins as minima in an effective energy landscape given by the quasipotential and transitions as most-probable paths in this landscape, only made possible by understanding the interaction between both sources of stochasticity.

By taking a distinguished limit simultaneously in both sources of noise, we derived the Hamiltonian for this problem and demonstrated the numerical problems that arise when using the standard methods in the literature. We developed modifications to these methods that allow us to compute most-probable transition paths and quasipotential barriers along these paths. These quasipotential barriers accurately predict escape times, reinforcing that this simultaneous treatment of both noise sources is required over a time average of only the stochastic switching to produce a single deterministic force before considering the effects of thermal noise. This deterministic average significantly overestimates the effective energy barrier and thus the stability of a pseudoequilibrium state such as the gene clusters observed in [14]. Thus we have shown that an effective thermal equilibrium can be constructed by simultaneously incorporating both the stochastic switching that pushes the system out of equilibrium and the thermal noise.

The inclusion of Brownian motion in our framework contrasts with previous work in which the only source of noise comes from random switching between deterministic differential equations [16,22]. With the correct choice of scaling between the thermal noise and the stochastic switching rates, we were able to demonstrate that a similar WKB ansatz approach can be applied to derive the Hamilton-Jacobi equation for this framework.

Our example six-dimensional system produced new numerical issues not reported by others computing quasipo-

tentials for one- to three-dimensional systems in previous literature [16,22–27]. These numerical issues arose in two separate places in our algorithm: the implicit solve step for ∇W, as discussed in Sec. II C, and the optimization method for finding the most-probable transition path, as discussed in Sec. IV B. In each case we showed how to modify the algorithm to overcome these issues. As solving for ∇W is a convex optimization problem, it is expected that there should exist a good solver, and our modified Newton's method seems to fit this role, with guaranteed convergence under a short number of iterations from practical considerations. However, there is room for future work in further understanding the error in the MPP computation step, as the current remedy of reducing the number of images leaves desired the ability to compute paths on a finer resolution. We reiterate that the most commonly observed alternative to the string method, gMAM, does not alleviate this problem, as it runs into instability due to high sensitivity in the Hessian matrix. However, it is conceivable that continued work on the method used for this optimization step could produce a more stable algorithm in higher numbers of images.

Applying our methods to our idealized model, we showed how a stochastically switching pairwise crosslinking force could create effective bound states, with the system exhibiting switching between which pair of beads is in the effective bound state. We note that this is not simply switching between which pair of beads is currently bound by the switching crosslinking force. When beads are in this so-called effective bound state, they are switching back and forth between being bound and unbound, but on average are bound enough of the time to remain in close proximity until a rare event through a combination of Brownian noise and switching of the force allows them to separate and a new pair of beads to enter the effective bound state. We showed the ability to accurately predict the asymptotic timescales on which this escape occurs, showing strong agreement between the slope of Monte Carlo simulations of the timescale and the height of the quasipotential barrier, in line with the Arrhenius law of escape times. This validates our methods and code and sets the stage for use in further applications.

Our work here was originally motivated by observations in [14] in which we observed particles representing beads in a polymer model of the yeast genome associating into clusters as though on an effective energy landscape, even though the stochastically switching forces meant the system was never truly in equilibrium. In this work we were able to mathematically demonstrate that the observed behavior can be explained by an effective energy landscape through the analysis of a reduced model. Having derived the Hamilton for a general overdamped Langevin system and the numerical methods to compute transition paths, we would like to apply our methods in future work to analyze the clustering states observed in this model as fixed points of the associated deterministic dynamics, and then compute the quasipotential barrier to escape from these states to understand the stability of each clustering state in terms of system parameters.

One of the challenges facing scaling up this method to more particles is the size of the matrix M, which increases factorially with the number of particles. Our previous work in [14] showed a collection of approximately 380 beads condensing into clusters dynamic clusters of approximately 5–10 beads in suitable parameter regimes. In future work we would like to apply our methods to larger numbers of beads using forces directly taken from biological modeling. The largest challenge to this will be addressing the state space of the switching process, which increases exponentially with the number of beads. The algorithm does not use the matrix $M(x, p)$ in full, only the derived (scalar) Hamiltonian $\mathcal{H}(x, p)$ arising from the greatest eigenvalue. Since x and p only scale up in dimension linearly with the number of particles, this suggests an alternative approach in which we construct an approximation either for the map $(x, p) \to \mathcal{H}(x, p)$, which is used to implicitly solve for p, or even directly for the map $x \to p$ (noting that the direction $\frac{d\phi}{ds}$ of the string would also need to be included). Recently machine learning methods have been used to develop approximations to challenging-to-compute mathematical functions, and so this is an approach we would like to incorporate into our method in future work.

The WKB ansatz approach also allows for the computation of additional terms by following through the asymptotic expansion to higher order. In particular, the next term in the series would give the full preexponential term that would allow for computation of the intercept of the escape time asymptotics as shown in Figs. 1 and 5. While the current work focused on developing the machinery to compute most-probable transition paths under our framework, future work could expand on this and compute this preexponential coefficient.

The full code to compute transition paths and quasipotential barriers for these systems and replicate all results in this work is available on GitHub [29].

ACKNOWLEDGMENTS

We would like to thank Jay Newby for insightful conversations and advice on the computation of quasipotentials for switching systems. This work was partially supported by the National Science Foundation under Grants No. DMS-1816630 and No. DMS-1816394.

APPENDIX A: MONTE CARLO SIMULATION

To validate our quasipotential, we compare it to Monte Carlo simulations of the system. These simulations are taken by numerically evaluating Eq. (24) using a modified Euler-Maruyama method, to incorporate both the SDE and the stochastic switching. We track the current switching state s in addition to the position x and update the position in each step as

$$x_{\text{new}} = x + \Delta t F(x; s) + \sqrt{2\epsilon}\Delta B, \tag{A1}$$

where $F(x; s)$ is the drift function at position x in state s, and $\Delta B \sim \mathcal{N}(0, \Delta t)$ is the increment of Brownian noise. Then we compute any changes to the state variable s. The specifics of this update depend on the nature of the switching but in general involve checking whether any changes occur within the time step Δt by drawing waiting times from exponential distributions. Due to the memoryless property of exponential random variables, we may safely redraw in future time steps any waiting times that do not correspond to a transition

within the current time step, in case the transition rate changes between time steps. In this work we make a first-order approximation and treat all switches as occurring at the beginning of the time step.

To empirically estimate the mean escape time associated with a particular ϵ, we initialize a Monte Carlo simulation in a state corresponding to a stable equilibrium of the deterministic dynamics, as obtained through descent of the deterministic force. We proceed to simulate Eq. (A1) until a termination condition is reached indicating escape from the basin of attraction that the simulation began in. The escape time is then recorded as τ_k. The simulation may also reach maximum time T without exiting the original basin of attraction, in which we take $\tau_k = T$. We assume escape times follow an exponential distribution with mean μ and use the maximum likelihood estimator (MLE) for the mean of an exponential distribution given samples capped at a maximum time, given by [30]:

$$\tau = \frac{\sum_k \tau_k}{\sum_k I(\tau_k < T)}. \tag{A2}$$

Note that as $T \to \infty$, this reduces to the simple mean of the samples τ_k.

We then repeat the above process for a sequence of values of ϵ, computing a relationship $\tau(\epsilon)$. We expect a linear relationship between $\frac{1}{\epsilon}$ and $\log \tau$,

$$\log \tau = a + \frac{b}{\epsilon}, \tag{A3}$$

where the value b is corresponds theoretically with the quasipotential barrier height, and a would be given by the higher-order preexponential term.

APPENDIX B: DERIVATION DETAILS

1. Derivation of the Hamiltonian

Here we will present the computation arising from plugging in the WKB ansatz Eq. (8),

$$p_s(x) = r_s \exp\left(-\frac{1}{\epsilon}W(x)\right),$$

to the steady-state equation, Eq. (7),

$$0 = -\sum_i \frac{\partial}{\partial x_i}\left[v_i^s p_s\right] + \epsilon \sum_i \frac{\partial^2}{\partial x_i^2}[p_s] + \frac{1}{\epsilon}[Sp]_s,$$

(both reproduced for clarity) and collecting lowest order terms in ϵ, which will be the $\frac{1}{\epsilon}$ terms.

The spatial derivatives of the WKB ansatz are given by

$$\frac{\partial}{\partial x_i}\exp\left(-\frac{W(x)}{\epsilon}\right) = \exp\left(-\frac{W(x)}{\epsilon}\right)\left[-\frac{1}{\epsilon}\frac{\partial W}{\partial x_i}\right]$$

and

$$\frac{\partial^2}{\partial x_i^2}\exp\left(-\frac{W(x)}{\epsilon}\right) = \exp\left(-\frac{W(x)}{\epsilon}\right)\left[-\frac{1}{\epsilon}\frac{\partial W}{\partial^2 x_i^2}\right] + \exp\left(-\frac{W(x)}{\epsilon}\right)\left[\frac{1}{\epsilon}\frac{\partial W}{\partial x_i}\right]^2.$$

For the drift term, we obtain a $\frac{1}{\epsilon}$ term by differentiating once the exponential, thus giving us

$$-\sum_i \frac{\partial}{\partial x_i}\left[v_i^s p_s\right] \sim \frac{1}{\epsilon} r_s \exp\left(-\frac{W(x)}{\epsilon}\right)\sum_i v_i^s \frac{\partial W}{\partial x_i}.$$

For the diffusion term we obtain a $\frac{1}{\epsilon}$ term by differentiating the exponential twice and combining with the ϵ prefactor, giving us

$$\epsilon \sum_i \frac{\partial^2}{\partial x_i^2}[p_s] \sim \frac{1}{\epsilon} r_s \exp\left(-\frac{W(x)}{\epsilon}\right)\sum_i \left(\frac{\partial W}{\partial x_i}\right)^2.$$

Finally, the switching term itself is a $\frac{1}{\epsilon}$ term, given by

$$\frac{1}{\epsilon}[Sp]_s = \frac{1}{\epsilon}\exp\left(-\frac{W(x)}{\epsilon}\right)[Sr]_s.$$

Together, Eq. (7) reduces to, at lowest order in ϵ,

$$\begin{aligned} 0 = {} & \frac{1}{\epsilon} r_s \exp\left(-\frac{W(x)}{\epsilon}\right)\sum_i v_i^s \frac{\partial W}{\partial x_i} \\ & + \frac{1}{\epsilon} r_s \exp\left(-\frac{W(x)}{\epsilon}\right)\sum_i \left(\frac{\partial W}{\partial x_i}\right)^2 \\ & + \frac{1}{\epsilon}\exp\left(-\frac{W(x)}{\epsilon}\right)[Sr]_s. \end{aligned} \tag{B1}$$

We can cancel out the common term of $\frac{1}{\epsilon}\exp(-\frac{W(x)}{\epsilon})$ and then rewrite the resulting equation as a matrix equation by defining vector r with components r_s, $s = 1 \ldots n$ and turning the advection (drift), diffusion, and switching terms into the matrices A, D, and S, respectively:

$$[A + D + S]r = Mr = 0, \tag{B2}$$

with the diffusion matrix D defined as

$$D = \left[\sum_i \left(\frac{\partial W}{\partial x_i}\right)^2\right]\mathbb{I}, \tag{B3}$$

where $\mathbb{I}$ represents the identity matrix, and the advection matrix A is defined as

$$A = \mathrm{diag}(V\nabla W) \tag{B4}$$

(note that diag here indicates the mapping of a vector to the square matrix with it on the diagonal), and the matrix S unchanged from its original definition. Note also that D and A are diagonal matrices, so the only off-diagonal contributions come from S, which is not dependent on W.

2. Derivatives of the Hamiltonian

In order to apply Newton's method for finding critical points, we will need to compute the gradient and Hessian of $\mathcal{H}(x, p)$ with respect to the momentum variable p. To do so we will use the formulas of [31] in differentiating eigenvalues of a matrix with respect to the entries of that matrix. Consider a real, square matrix $M = M_0 + dM$ such that $M_0 u_0 = \lambda_0 u_0$ and $v_0 M_0 = \lambda_0 v_0$, and then consider the function $\lambda(dM)$ s.t.

$\lambda(0) = \lambda_0$. Using the superscript $+$ to refer to the Moore-Penrose inverse, we have

$$d\lambda = \frac{v_0^T dM u_0}{v_0^T u_0},$$
$$d^2\lambda = \frac{2v_0^T (dM) K_0 (\lambda_0 I - M_0)^+ K_0 (dM) u_0}{v_0^T u_0},$$
$$K_0 = I - \frac{u_0 v_0^T}{v_0^T u_0}.$$

We now consider the case that M is in fact a function of the momentum variable p, centered at a value p_0. In this case we have $M(p_0) = M_0$. Applying the chain rule, we get the formulas

$$\frac{d\lambda}{dp_i} = \frac{v_0^T M_{p_i} u_0}{v_0^T u_0}, \tag{B5}$$
$$\frac{d^2\lambda}{dp_i dp_j} = \frac{2v_0^T M_{p_i} K_0 (\lambda_0 I - M_0)^+ K_0 M_{p_j} u_0}{v_0^T u_0} + \delta_{ij} \frac{v_0^T M_{p_i,p_i} u_0}{v_0^T u_0}, \tag{B6}$$

where $p_i = \partial W / \partial x_i$ for $i = 1 \ldots m$.

Because $\mathcal{H}(x, p)$ is simply the largest eigenvalue of the matrix M, we can compute the gradient and Hessian $\mathcal{H}_p$ and $\mathcal{H}_{pp}$ with direct application of Eqs. (B5) and (B6).

3. Relationship between deterministic dynamics and quasipotential

In this section we provide a proof that fixed points x^d of the deterministic dynamics correspond to points where $\nabla_p \mathcal{H}(x^d, 0) = 0$, showing that $\nabla W = 0$ corresponds to a minimum and therefore unique solution of $\mathcal{H}(x^d, \nabla W) = 0$.

Let x^d be a fixed point of the deterministic dynamics such that $F_{\text{det}}(x^d) = 0$, and let r^d be the probability vector across the states, and let the force term for each switching state v_i^s be represented in a matrix $V_{is} = v_i^s$ such that

$$S(x^d) r^d = 0 \quad \text{because } r^d \text{ is a null vector,}$$
$$V(x^d) r^d = 0 \quad \text{because } x^d \text{ is a fixed point,}$$

Recall the definition $M(x, \nabla W) = A(x, \nabla W) + D(\nabla W) + S(x)$. One can see from the definitions in Eqs. (B3) and (B4) that A and D vanish when $\nabla W = 0$, so that $M(x, 0) = S(x)$. It follows that $M(x, 0) r = S(x) r = 0$, so 0 is an eigenvalue of $M(x, 0)$ and therefore $\mathcal{H}(x, 0) = 0$ for any x.

We apply Eq. (B5) to compute $\frac{d\mathcal{H}}{dp_i}$, which we will show to be 0 for all i,

$$\frac{d\mathcal{H}}{dp_i} = \frac{v_0^T (dM/dp_i) u_0}{v_0^T u_0},$$

where u_0, v_0 are the right and left eigenvectors, respectively, of $M(x_d, 0) = S(x_d)$, corresponding to an eigenvalue $\lambda = 0$. Recalling that S is a CTMC transition rate matrix, the right eigenvector u_0 will naturally be the steady-state vector r^d. Similarly, as a transition rate matrix its columns always sum to 0, and so the ones vector will be a left eigenvector with eigenvalue 0, and thus $v_0 = [1, 1, 1 \ldots 1]^T$.

Noting that

$$\frac{dA}{dp_i} = \text{diag}(V_i),$$
$$\frac{dD}{dp_i} = 2p_i I = 0 \quad (p_i = 0),$$

we get that

$$dM(x^d, 0) u_0 = \text{diag}(V_i) r^d = v_0^T V_i^T r^d = \frac{d}{dt} x^d = 0,$$

as x_d is a fixed point of the deterministic dynamics.

APPENDIX C: ALGORITHM IMPLEMENTATION DETAILS

Note that full code for reproducing all results is available on GitHub [29].

1. Quasipotential implicit solver

The core Newton updates from [16] are given by the iterative update equations:

$$p_{n+1} = p_n + \mathcal{H}_{pp}^{-1} \left[\lambda_n \frac{d\phi}{ds} - \mathcal{H}_p \right], \tag{C1}$$
$$\lambda_n = \sqrt{\frac{\mathcal{H}_p \mathcal{H}_{pp}^{-1} \mathcal{H}_p - 2\mathcal{H}}{d\phi^T \mathcal{H}_{pp}^{-1} d\phi}}, \tag{C2}$$

starting from an initial guess for p_0, where $\mathcal{H}, \mathcal{H}_p, \mathcal{H}_{pp}$ are evaluated at x, p_n, and the derivatives w.r.t p are computed in Appendix B 2. If the starting guess is sufficiently close to the correct value, Eq. (C1) will converge to the value p that satisfies Eqs. (12) and (14), which is the value of ∇W. However, in practice, we find that for our problems the Newton method frequently fails to converge, and so we add a modification based on the structure of the optimization problem to guarantee convergence.

The solution to Eq. (14) under the constraint Eq. (12) can be equivalently posed as the unique solution of

$$\arg\max_p \left[p \cdot \frac{d\phi}{ds} \right] \quad \text{s.t. } \mathcal{H}(x, p) = 0, \tag{C3}$$

as the surface $\mathcal{H}(x, p) = 0$ is convex. In other words, Eq. (14) can be viewed as a maximization problem. Correspondingly, we require that each iteration increase the objective quantity $p \cdot \frac{d\phi}{ds}$. If the Newton's step fails to do so, we move a small distance κ in the direction of $\frac{d\phi}{ds}$ projected onto the normal to the $\mathcal{H}(x, p) = 0$ surface:

$$p_n^* = p_n + \kappa \, \text{proj}_{\perp \mathcal{H}_p} \left(\frac{d\phi}{ds} \right).$$

We then use the root-finding version of Newton's method to find a p_{n+1} such that $\mathcal{H}(x, p_{n+1}) = 0$, starting from p_n^*, until convergence:

$$p_n^* = p_n^* - \frac{\mathcal{H}(x, p_n^*)}{\mathcal{H}_p(x, p_n^*)}. \tag{C4}$$

If this did not increase the objective quantity $p \cdot \frac{d\phi}{ds}$, we reduce κ by a factor of 2 and try again. Because the surface is convex

and we are moving in the objective direction projected onto the surface normal, we are guaranteed an improvement for sufficiently small κ. However, it is desirable that κ not be too much smaller than necessary to reduce the number of iterations until Newton's method begins to converge.

We find that with this modification we always reach a point where Newton's method begins to converge quadratically to the true solution.

2. String method

We apply the climbing string method as described in [21] and summarized here for systems of the form

$$dX = -\nabla U(X)dt + \sqrt{2\epsilon}dW.$$

The string ϕ is a path in the configuration space of X, starting at an energy minimizer x_A of $U(x)$. It is discretized into a number of images, ϕ_i for $i = 1 \ldots N$ with $\phi_1 = x_A$, that are evolved according to

$$\begin{aligned}\phi_i^{n+1} &= \phi_i^n - \nabla U\left(\phi_i^n\right)dt \; i = 2 \ldots N-1,\\ \phi_N^{n+1} &= \phi_N^n - \nabla U\left(\phi_N^n\right)dt + (1+\alpha)\left[\nabla U\left(\phi_N^n\right)\cdot\hat{\tau}\right]\hat{\tau}dt,\end{aligned} \tag{C5}$$

where $\hat{\tau}$ is a unit vector approximating the tangent to the string at the end point

$$\hat{\tau} = \frac{\phi_N^n - \phi_{N-1}^n}{\left|\phi_N^n - \phi_{N-1}^n\right|},$$

and $\alpha > 0$ is a parameter that controls the climbing speed of the last image. After each iteration of Eq. (C5), the intermediate images along the string are interpolated uniformly in arc length to prevent them from bunching up at the energy minimizer.

The only modification is the replacement of the gradient of the potential ∇U with the gradient of the quasipotential ∇W. These differ in that whereas ∇U evaluated at an image ϕ_i would depend only on the value of ϕ_i, ∇W depends on both ϕ_i and the direction of the string $\frac{d\phi}{ds}|_{\phi=\phi_i}$. We estimate this using a centered finite difference approximation,

$$\left.\frac{d\phi}{ds}\right|_{\phi=\phi_i} \approx \frac{\phi_{i+1} - \phi_{i-1}}{2h}, \tag{C6}$$

which turns Eq. (14) into the numerical condition,

$$\left.\frac{d\phi}{ds}\right|_{\phi=\phi_i} \parallel \phi_{i+1} - \phi_{i-1}. \tag{C7}$$

Note that we drop the denominator $\frac{1}{2h}$, as it does not affect the direction and therefore is irrelevant to the parallel condition. This, however, means that the update order affects the algorithm. In our case we compute the directions $\frac{d\phi}{ds}$ at all images first and then perform updates simultaneously.

[1] P. C. Bressloff, Stochastic switching in biology: From genotype to phenotype, J. Phys. A: Math. Theor. **50**, 133001 (2017).

[2] M. A. Welte, Bidirectional transport along microtubules, Curr. Biol. **14**, R525 (2004).

[3] X.-Z. Cao and M. G. Forest, Rheological tuning of entangled polymer networks by transient cross-links, J. Phys. Chem. B **123**, 974 (2019).

[4] M. A. Jensen, Y.-Y. Wang, S. K. Lai, M. G. Forest, and S. A. McKinley, Antibody-mediated immobilization of virions in mucus, Bull. Math. Biol. **81**, 4069 (2019).

[5] H. A. Schroeder, J. Newby, A. Schaefer, B. Subramani, A. Tubbs, M. G. Forest, E. Miao, and S. K. Lai, LPS-binding IgG arrests actively motile Salmonella typhimurium in gastrointestinal mucus, Mucosal Immunology **13**, 814 (2020).

[6] E. Alipour and J. F. Marko, Self-organization of domain structures by DNA-loop-extruding enzymes, Nucleic Acids Res. **40**, 11202 (2012).

[7] V. Ea, M.-O. Baudement, A. Lesne, and T. Forné, Contribution of topological domains and loop formation to 3D chromatin organization, Genes **6**, 734 (2015).

[8] A. Goloborodko, M. V. Imakaev, J. F. Marko, and L. Mirny, Compaction and segregation of sister chromatids via active loop extrusion, eLife **5**, e14864(2016).

[9] S. Schalbetter, A. Goloborodko, G. Fudenberg, J. Belton, C. Miles, M. Yu, J. Dekker, L. Mirny, and J. Baxter, Structural maintenance of chromosome complexes differentially compact mitotic chromosomes according to genomic context, Nat. Cell Biol. **19**, 1071 (2017).

[10] M. Ganji, I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker, Real-time imaging of DNA loop extrusion by condensin, Science **360**, 102 (2018).

[11] Y. He, J. Lawrimore, D. Cook, E. E. Van Gorder, S. C. De Larimat, D. Adalsteinsson, M. G. Forest, and K. Bloom, Statistical mechanics of chromosomes: In vivo and in silico approaches reveal high-level organization and structure arise exclusively through mechanical feedback between loop extruders and chromatin substrate properties, Nucleic Acids Res. **48**, 11284 (2020).

[12] T. Terakawa, S. Bisht, J. M. Eeftens, C. Dekker, C. H. Haering, and E. C. Greene, The condensin complex is a mechanochemical motor that translocates along DNA, Science **358**, 672 (2017).

[13] C. Hult, D. Adalsteinsson, P. A. Vasquez, J. Lawrimore, M. Bennett, A. York, D. Cook, E. Yeh, M. G. Forest, and K. Bloom, Enrichment of dynamic chromosomal crosslinks drive phase separation of the nucleolus, Nucleic Acids Res. **45**, 11159 (2017).

[14] B. Walker, D. Taylor, J. Lawrimore, C. Hult, D. Adalsteinsson, K. Bloom, and M. G. Forest, Transient crosslinking kinetics optimize gene cluster interactions, PLoS Comput. Biol. **15**, e1007124 (2019).

[15] P. C. Bressloff and J. M. Newby, Path integrals and large deviations in stochastic hybrid systems, Phys. Rev. E **89**, 042701 (2014).

[16] J. M. Newby, Spontaneous excitability in the Morris–Lecar model with ion channel noise, SIAM J. Appl. Dyn. Syst. **13**, 1756 (2014).

[17] D. Ryter, The exit problem at weak noise, the two-variable quasipotential, and the kramers problem, J. Stat. Phys. **149**, 1069 (2012).
[18] M. Cameron, Finding the quasipotential for nongradient SDEs, Physica D **241**, 1532 (2012).
[19] M. Cameron and S. Yang, Computing the quasipotential for highly dissipative and chaotic SDEs an application to stochastic Lorenz 63, Commun. Appl. Math. Comput. Sci. **14**, 207 (2019).
[20] W. E, W. Ren, and E. Vanden-Eijnden, String method for the study of rare events, Phys. Rev. B **66**, 052301 (2002).
[21] W. Ren and E. Vanden-Eijnden, A climbing string method for saddle point search, J. Chem. Phys. **138**, 134105 (2013).
[22] P. C. Bressloff and O. Faugeras, On the Hamiltonian structure of large deviations in stochastic hybrid systems, J. Stat. Mech.: Theory Exp. (2017) 033206.
[23] Z. Brzeźniak, S. Cerrai, and M. Freidlin, Quasipotential and exit time for 2D stochastic Navier-Stokes equations driven by space time white noise, Probab. Theory Relat. Fields **162**, 739 (2015).
[24] Z. Chen, J. Zhu, and X. Liu, Non-differentiability of quasipotential and non-smooth dynamics of optimal paths in the stochastic Morris–Lecar model: Type I and II excitability, Nonlinear Dyn. **96**, 2293 (2019).
[25] C. M. Moore, C. R. Stieha, B. C. Nolting, M. K. Cameron, and K. C. Abbott, QPot: An R package for stochastic differential equation quasi-potential analysis, The R Journal **8**, 19 (2016).
[26] S. Yang, S. F. Potter, and M. K. Cameron, Computing the quasipotential for nongradient SDEs in 3D, J. Comput. Phys. **379**, 325 (2019).
[27] J. X. Zhou, M. Aliyu, E. Aurell, and S. Huang, Quasi-potential landscape in complex multi-stable systems, J. R. Soc. Interface **9**, 3539 (2012).
[28] M. Heymann and E. Vanden-Eijnden, The geometric minimum action method: A least action principle on the space of curves, Commun. Pure Appl. Math. **61**, 1052 (2008).
[29] bwalker1, bwalker1/quasi-string-reprod: Quasipotential Strings v1 (v1.0.0), Zenodo (2022), https://doi.org/10.5281/zenodo.6383387.
[30] https://www.itl.nist.gov/div898/handbook/apr/section4/apr412.htm.
[31] J. R. Magnus, On differentiating eigenvalues and eigenvectors, Econometric Theory **1**, 179 (1985).