

## Capacities and efficient computation of first-passage probabilities

Jackson Loper\*

*Data Science Institute, Columbia University, 10027 New York, New York, USA*Guangyao Zhou<sup>\*,†</sup> and Stuart Geman*Division of Applied Mathematics, Brown University, Providence, 02912 Rhode Island, USA*

(Received 14 May 2019; revised 8 May 2020; accepted 12 June 2020; published 10 August 2020)

A reversible diffusion process is initialized at position  $x_0$  and run until it first hits any of several targets. What is the probability that it terminates at a particular target? We propose a computationally efficient approach for estimating this probability, focused on those situations in which it takes a long time to hit any target. In these cases, direct simulation of the hitting probabilities becomes prohibitively expensive. On the other hand, if the timescales are sufficiently long, then the system will essentially “forget” its initial condition before it encounters a target. In these cases the hitting probabilities can be accurately approximated using only local simulations around each target, obviating the need for direct simulations. In empirical tests, we find that these local estimates can be computed in the same time it would take to compute a single direct simulation, but that they achieve an accuracy that would require thousands of direct simulation runs.

DOI: [10.1103/PhysRevE.102.023304](https://doi.org/10.1103/PhysRevE.102.023304)

### I. INTRODUCTION

Reversible diffusions play a key role in a wide variety of physical systems. For example, the folding of macromolecules into their native configurations is often posed as a diffusion (either directly through simulations of atomic dynamics or indirectly through mesoscopic models [1–3]), the fluctuation of chemical species in solution can be modeled as diffusions (“chemical Langevin equations” are one classic example of this approach [4,5]), the motion of particles through membranes can be posed as a diffusion [6], and so on. In all cases, the physical state of the system at time  $t$  is represented through a variable  $X_t$ , confined to a finite region  $\Omega$  and evolving according to a stochastic differential equation.

In this paper we seek to estimate first-passage probabilities (sometimes also called “splitting probabilities” [7] or simply “hitting probabilities”) of such diffusions: given an initial condition  $x_0$  and a collection of targets, what is the probability that the process first hits any particular target before hitting any of the other targets? Targets might represent, for example, multiple exit locations from a region of the state space in a statistical mechanics problem or the establishment of new intramolecular bonds in a folding problem. Where will the system first exit or what is the next step in folding? This paper develops an alternative algorithm for the approximation of first-passage probabilities.

To keep things simple, most of the discussion will be in the context of just two targets,  $A$  and  $B$ ,  $A, B \subset \Omega$ , although all of the results apply, with trivial notational changes, to any finite set of targets. In this context, “hitting probability” (also known as “first-passage probability”) refers to the specific probability that  $X_t$  hits target  $A$  before target  $B$ . The algorithm is designed for cases where the diffusion “forgets” its initial condition before encountering a target. Precise definitions are in Sec. II, but to get the main idea consider a region  $M$  in  $\Omega$  that lies outside of the targets  $A$  and  $B$ ,  $M \subset \Omega \setminus (A \cup B)$ , and assume that the diffusion starts in  $M$  ( $x_0 \in M$ ). Recall that reversible diffusions are *ergodic*: as time goes on, it gets harder and harder to guess the initial condition from the current configuration. As this happens, the process is said to become “mixed,” and the speed with which the initial condition is forgotten is called the “mixing rate.” If the first-passage time out of  $M$  is long relative to this mixing rate, then the hitting probability will be nearly independent of  $x_0$ . This is what we mean by “forgets” its initial condition. Our main result (Theorem 1) asserts that if the hitting probability varies by at most  $\varepsilon$  among all initial conditions in  $M$ , then the proposed algorithm’s approximation error is less than  $\varepsilon + \sqrt{\varepsilon/2}$ . Figure 1 gives a rough sketch of this idea.

When would these conditions hold? We will have more to say about this in Sec. III, but for now we note that the narrow-escape literature gives at least one class of examples. In these examples  $X_t$  is a Brownian motion trapped inside a set by reflecting boundaries, and the targets are very small windows in boundary. These models have been successfully applied to a variety of mesoscopic biosystems (cf. Ref. [8] and the references therein). These kinds of diffusions may also be relevant for modeling the folding of large molecules. According to McLeish, “folding rates are controlled by the rate of diffusion of pieces of the open coil in their search for favorable

\*These two authors contributed equally to this work.

†Corresponding author: [tczhouguangyao@gmail.com](mailto:tczhouguangyao@gmail.com)

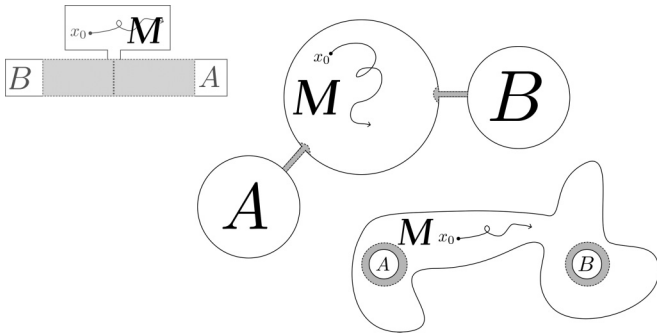


FIG. 1. First-passage capacities allow efficient computation of hitting probabilities. We consider a reversible stochastic differential equation in an  $n$ -dimensional state space, initialized at some configuration  $x_0$  inside a set  $M$ . What is the probability that the diffusion will reach region  $A$  before  $B$ ? We assume that the hitting probability is nearly the same for every initial condition  $x_0 \in M$ —differing by at most  $\varepsilon$  across all initial conditions. (See Theorem 2 for some sufficient conditions, involving small targets and/or high dimensions.) Theorem 1 establishes an approximation to the hitting probability that is accurate to within  $\varepsilon + \sqrt{\varepsilon/2}$  and uses only “first-passage capacities.” These capacities can be computed from simulations that take place locally, meaning in the neighborhoods of the targets and entirely outside of  $M$ . Here we depict several examples. In each case, the first-passage capacities can be computed efficiently using simulations confined only to the areas shaded in gray.

contacts” [9]. There is also some experimental evidence that the exploration of large regions is the rate-limiting step for a variety of processes [10–13]. If these results bear out, then to the extent that exit times are long it will be possible to obtain good approximations for hitting probabilities using our approach.

Many standard methods for calculating hitting probabilities fail for diffusions of this kind. When the escape time from  $M$  is large, the associated partial differential equations often include large Lipschitz constants, and direct Monte Carlo simulation requires a prohibitive number of time steps (cf. Refs. [14–16]). Some authors take a pessimistic view of this subject: “If these processes are intrinsically slow, i.e., require an extensive sampling of state space, not much can be done to speed up their simulation without destroying the dynamics of the system” [17]. But in fact diffusions of this kind enjoy certain properties which can actually make analyses easier. The literature includes a variety of tools which use such properties to advantage. Here we review a few of these tools:

*Analytic analyses.* When targets are very small, the first-passage times (as opposed to probabilities) are often well approximated by analytic formulas. The narrow-escape literature has developed many techniques in this direction. Much of this literature focuses on making these formulas as accurate as possible for specific geometric configurations, such as the case of  $N$  circular targets (cf. Ref. [18]), the case that one target lies at the end of a long tube (cf. Ref. [19]), or the case that the motion is trapped inside a symmetric domain (cf. Refs. [20–23]). There is also some work on efficient numerical methods for low-dimensional problems using insights from these analytic results (cf. Ref. [24]). Note that this literature mostly focuses on hitting times, whereas in this paper we

focus on hitting probabilities. Nevertheless, the first-passage probabilities are nearly inversely proportional to the mean first-passage times for many cases, so these techniques can often be used to estimate first-passage probabilities. A helpful overview of many of the main ideas in this literature can be found in Ref. [25].

*Markov state models.* Markov state models (MSM) begin by partitioning the state space of the diffusion into  $n$  sets (“states”). The diffusion can be understood at a coarse level by looking at which state the process is in at any given time (cf. Refs. [26–28]). In some cases the time it takes to move between states is long relative to the rate of mixing inside the states. This separation of timescales ensures that the coarse process is approximately Markovian. We can therefore approximately simulate the discrete process as long as we know the distribution on the exit times and the probability of transitioning to each possible state. The amount of computational time required to simulate from this approximate process does not depend upon how long the process actually spends in each state.

*Site-localizing functions.* Site-localizing functions define a kind of continuous version of the MSMs: instead of creating a hard partitioning of the space, one constructs basis functions  $g_1(x) \cdots g_n(x)$  such that  $\sum_i g_i(x) = 1$ . In many cases the action of the relevant diffusion operators can be well approximated by their action on the  $n$ -dimensional space spanned by these basis functions. For example, under a separation-of-timescales assumption, Morro [29] shows a way to design basis functions which faithfully represent the diffusion’s behavior on the slow timescale.

*Milestoning.* In some cases it is possible to construct a low-dimensional reaction coordinate which measures the distance from the targets in some suitable metric. In some cases the movement along this reaction coordinate is slow relative to the mixing rate along all other directions. This separation of timescales ensures that the dynamics of the diffusion along this coordinate are approximately Markovian. If the reaction coordinate is carefully chosen, many properties of the original diffusion are maintained in this low-dimensional projection [7]. Simulating the low-dimensional approximate diffusion can yield useful insights into the behavior of the original system [30].

Here we will investigate a different way to make use of separated timescales. Let  $h_{A,B}(x)$  denote the probability of hitting  $A$  before  $B$  given that the diffusion is initialized at  $x \in M$ . It is well known that  $h_{A,B}(x)$  can be represented as the solution of a variational problem. (A detailed account of this and related facts is included in Appendix A.) Under the special condition that  $h_{A,B}(x)$  is a constant, say  $h_{A,B}(x) = \bar{h}$  for all  $x \in M$ , the solution of the variational problem, and hence  $\bar{h}$ , can be written in terms of certain integrals over local regions surrounding the targets (indicated by the gray-shaded areas in the examples depicted in Fig. 1). We will refer to these integrals as “first-passage capacities”; see Sec. II for the formal definition. If, on the other hand,  $h_{A,B}(x)$  is not exactly constant, then  $\bar{h}$  becomes an approximation rather than an exact formula: for all  $x \in M$ ,  $|\bar{h} - h(x)| \leq \varepsilon + \sqrt{\varepsilon/2}$ , where  $\varepsilon$  is the difference between the maximum and minimum hitting probabilities over  $M$ . This is the main theoretical result, Theorem Sec. 1, which reduces the first-passage problem to

one of evaluating local integrals. For the purpose of evaluating these integrals, we introduce a Monte Carlo approach, that we call the shell method, and demonstrate its accuracy and computational efficiency.

The key advantage of our approach is that it allows us to get good estimates while completely ignoring all dynamics inside  $M$ . This is particularly valuable if  $M$  is a large, high-dimensional set. In this case it is unclear how one would partition  $M$  or formulate a meaningful reaction coordinate, calling into question the utility of Markov state models and milestoneing under these conditions. Furthermore, integrating over all of  $M$  is computationally burdensome, making site-localizing functions also difficult to implement. Asymptotic techniques from the narrow escape and reaction-controlled diffusion literature may still be applied, but only if the problem falls into one of a collection of very specific cases, most of which lie in two or three dimensions, feature regular targets, and assume that the diffusion has no energetic potential. By contrast, first-passage capacities can be calculated efficiently for a diffusion with an arbitrary energetic potential, target shape, and ambient dimension. Moreover, first-passage capacities are fundamentally nonasymptotic quantities; for any problem the approximation error is bounded explicitly in terms of the extent to which the hitting probability varies across  $M$ . On the other hand, the strengths of the capacity-based approach are accompanied by a significant limitation: the approach gives no information about first-passage *times*. Indeed, the temporal process is explicitly eliminated. In general, there is an overall proportionality constant which connects first-passage hitting probabilities to first-passage hitting times, but this relationship cannot be recovered from the first-passage capacities. We discuss this limitation further in Sec. VI.

In the following section (Sec. II “Preliminaries”) we introduce notation, define the diffusion and first-passage probabilities, enumerate the main assumptions, and finally define the first-passage capacities that are at the heart of this approach. In Sec. III we show that these capacities can be used to accurately estimate first-passage probabilities, and by way of examples, we give some sufficient conditions that guarantee that the main assumptions are satisfied. In Sec. IV, we develop the “shell method,” a numerical approach to computing first-passage capacities. And in Sec. V we investigate the speed and accuracy of the capacity-based approach in a variety of computational experiments.

**II. PRELIMINARIES**

The results in this paper are about diffusion processes  $X$  confined to an open bounded set  $\Omega \subset \mathbb{R}^n$  with reflecting smooth boundary  $\partial\Omega$  for  $n \geq 3$ . We assume  $X$  is driven by an  $n$ -dimensional standard Brownian motion  $W$ ,

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t, \tag{1}$$

where  $b: \Omega \rightarrow \mathbb{R}^n$  and  $\sigma: \Omega \rightarrow \mathbb{R}^{n \times n}$  are continuously differentiable vector-valued and matrix-valued functions. We further assume that  $a(x) = \sigma(x)\sigma(x)^T$  is uniformly elliptic on  $\Omega$ , i.e., the smallest eigenvalues of  $a$  are bounded away from zero. Let  $\bar{\Omega}$  denote the closure of  $\Omega$  (and in general let  $\bar{S}$  denote the closure of any set  $S \subset \bar{\Omega}$ ). For the precise definition of the reflected process, we adopt the framework

developed by Lions and Sznitman [31]: Let  $n = n(x)$  denote the outward normal of  $\partial\Omega$  and  $v: \partial\Omega \rightarrow \mathbb{R}^n$  a smooth vector field satisfying  $n^T v \geq c > 0$ , and assume that  $x_0 \in \Omega$ . Then there is a unique pathwise continuous and  $W$ -adapted strong Markov process  $X_t \in \bar{\Omega}$ , and (random) measure  $L$ , such that

$$X_t = x_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s - \int_0^t v(X_s) L(ds) \tag{2}$$

and  $L(\{t: X_t \notin \partial\Omega\}) = 0$ . For convenience, we will refer to  $X$  by simply saying “the reflected diffusion process (1).” We further assume that  $X$  is a reversible process, with equilibrium distribution given by

$$\rho(x) \doteq \frac{1}{Z} e^{-U(x)}, \quad Z = \int_{x \in \Omega} e^{-U(x)} dx,$$

where  $U: \bar{\Omega} \rightarrow \mathbb{R}$  is continuously differentiable. As shown by Chen [32], to ensure that  $X$  is reversible it is sufficient that

$$b_i(x) = \frac{1}{2} \sum_j \partial a_{ij}(x) / \partial x_j - \frac{1}{2} \sum_j a_{ij}(x) \partial U(x) / \partial x_j, \tag{3}$$

$$v(x) = a(x)n(x),$$

where  $a(x) = \sigma(x)\sigma(x)^T$  is uniformly elliptic. When the conditions in (3) are in force we will say that  $X$  satisfies the reversibility conditions relative to  $U$ . We assume these conditions throughout the paper.

The main goal of this paper is to estimate first-passage probabilities for  $X$ . We here give a formal definition for these hitting probabilities for the two-target case, although all of our results extend, straightforwardly, to finite collections of targets:

*Definition 1.* (First passage probabilities). Fix two disjoint sets  $A, B \subset \Omega$ . The first-passage probability function  $h_{A,B}(x)$  is the probability that the process  $X$  visits  $A$  before  $B$  if it is initialized at  $X_0 = x$ . Formally,

$$h_{A,B}(x) \triangleq \mathbb{P}(X_{\tau_{A \cup B}} \in A | X_0 = x),$$

where  $\tau_{A \cup B}$  indicates the first-passage time to  $A \cup B$ , i.e.,  $\tau_{A \cup B} \triangleq \inf\{t \geq 0: X_t \in A \cup B\}$ . Throughout this paper we will use  $\tau_S$  to denote the first-passage time to a set  $S$  and  $h_{S,S'}$  to denote the hitting probability function for targets  $S, S'$ .

As discussed in the introduction, these hitting probabilities are much easier to estimate if the process forgets its initial condition to such an extent that the hitting probabilities are nearly constant for any initial condition inside a set  $M$ . Here we give a formal definition for this property:

*Definition 2.* (The  $\varepsilon$ -flatness condition) A hitting probability function  $h_{A,B}(x)$  is said to be “ $\varepsilon$ -flat relative to  $M$ ” whenever

$$\sup_{x,y \in M} |h_{A,B}(x) - h_{A,B}(y)| < \varepsilon.$$

As we shall see in Theorem 1, this condition is exactly what we need to show that the hitting probabilities within  $M$  can be well approximated using “first-passage capacities.” These capacities are the last piece we must define:

*Definition 3.* (Capacity) Let  $S \subset \bar{S} \subset \Omega$  be open sets and let  $X$  be a diffusion governed by the SDE in Eq. (2) and

satisfying the reversibility conditions in Eq. (3). The first-passage capacity  $\text{cap}(S, \tilde{S})$  for  $X$  is defined as

$$\text{cap}(S, \tilde{S}) \triangleq \int_{\tilde{S} \setminus S} \|\sigma(x) \nabla h_{S, \tilde{S}^c}(x)\|^2 e^{-U(x)} dx.$$

We refer the reader to Appendix A for more details on capacity and the related concept of Dirichlet form. Note that there are several related definitions of ‘‘capacity’’ in the probabilistic potential theory literature, all slightly different. For example, the harmonic capacity arises by taking the above definition in the special case when the diffusion is a simple Brownian motion. The definition used here most closely follows the work of Bovier [33,34]. Throughout this work, the term is used only in the sense of the above definition.

### III. MAIN THEORETICAL RESULTS

Our main theorem shows that the first-passage capacities give accurate approximations of the hitting probabilities when the hitting probabilities are themselves  $\varepsilon$ -flat in a region outside a neighborhood of the targets.

*Theorem 1.* Let  $A \subset \tilde{A}, B \subset \tilde{B}$  be open sets and assume that  $h_{A,B}(x)$  is  $\varepsilon$ -flat relative to  $\Omega \setminus (\tilde{A} \cup \tilde{B})$ . Then the first-passage probabilities are well-approximated by the first-passage capacities:

$$\sup_{x \in \tilde{A}, \tilde{B}} \left| h_{A,B}(x) - \frac{\text{cap}(A, \tilde{A})}{\text{cap}(A, \tilde{A}) + \text{cap}(B, \tilde{B})} \right| \leq \varepsilon + \sqrt{\varepsilon/2}.$$

We defer the proof to Appendix C. Note that the generalization to multiple targets is straightforward due to the additive property of capacities (Proposition 2 in Appendix A). In general, the hitting probability is approximately proportional to the corresponding capacity.

When might the  $\varepsilon$ -flatness condition hold? In the introduction, we gave an intuitive explanation for when this might be expected to happen, namely, whenever the diffusion forgets its initial condition before it hits the targets. To prove the  $\varepsilon$ -flatness condition for a particular problem, one must make this idea rigorous. Here are two examples of how this might be done:

*Theorem 2.* (Examples of  $\varepsilon$ -flatness). Let  $\mathcal{B}(x, r) = \{y: \|x - y\| < r\}$  denote the ball of radius  $r$  centered at  $x$ . Let us say that a diffusion governed by Eq. (1) ‘‘behaves like Brownian motion’’ on a set if  $\sigma(x) = I$  and  $\nabla U(x) = 0$  for every  $x$  in that set.

(1) (*Small targets*) Fix  $r, \varepsilon > 0$ . We can then find  $r' > 0$  with the following property. For any bounded convex set  $\Omega$  with diameter less than or equal to 1, any  $x_A, x_B \in \Omega$ , any reversible stationary diffusion trapped inside  $\Omega$  and behaving like Brownian motion for all  $x \notin \mathcal{B}(x_A, r') \cup \mathcal{B}(x_B, r')$ , and any  $A \subset \mathcal{B}(x_A, r'), B \subset \mathcal{B}(x_B, r')$ , we have that  $h_{A,B}$  is  $\varepsilon$ -flat on  $\Omega \setminus (\mathcal{B}(x_A, r) \cup \mathcal{B}(x_B, r))$ .

(2) (*High dimensions*) Fix  $r, \varepsilon > 0$  and  $r' < r$ . We can then find  $n$  with the following property. For any convex set  $\Omega \in \mathbb{R}^n$  with diameter less than or equal to 1, any  $x_A, x_B$  such that  $\mathcal{B}(x_A, r), \mathcal{B}(x_B, r) \subset \Omega$ , any reversible stationary diffusion trapped inside  $\Omega$  and behaving like Brownian motion for all  $x \notin \mathcal{B}(x_A, r') \cup \mathcal{B}(x_B, r')$ , and any  $A \subset \mathcal{B}(x_A, r'), B \subset$

$\mathcal{B}(x_B, r')$ , we have that  $h_{A,B}$  is  $\varepsilon$ -flat on  $\Omega \setminus [\mathcal{B}(x_A, r) \cup \mathcal{B}(x_B, r)]$ .

A proof can be found in Appendix B. The first example shows how hitting probabilities become flat as the targets become small. The second example notes that even if we fix the radius of the targets, the hitting probabilities becomes flat if the ambient dimension  $n$  is high. In both cases, the results are proved by showing that the rate of mixing inside  $M$  is fast relative to the time it takes to exit  $M$ .

### IV. CAPACITY ESTIMATION VIA THE SHELL METHOD

To make use of Theorem 1 in practice, we must be able to compute capacities. The calculation is local, in that  $\text{cap}(A, \tilde{A})$  is an integral on  $\tilde{A} \setminus A$ . We will propose here a Monte Carlo approach to evaluating the integral, using a combination of analytic reductions and local sampling.

We begin by using an alternative formulation of the capacity:

*Proposition 1.* For any regions  $G$  and  $\tilde{G}$  having smooth boundaries and such that  $A \subset G \subset \tilde{G} \subset \tilde{A}$ ,  $\text{cap}(A, \tilde{A})$  can be expressed as a flux leaving  $\tilde{G} \setminus G$ :

$$\text{cap}(A, \tilde{A}) = \int_{\partial(\tilde{G} \setminus G)} h_{A, \tilde{A}^c}(x) \mathbf{n}(x)^T a(x) \nabla h_{G, \tilde{G}^c}(x) \times e^{-U(x)} \mathcal{H}(dx), \tag{4}$$

where  $a(x) = \sigma(x)\sigma(x)^T$  is the diffusion matrix,  $\mathcal{H}(dx)$  is the  $(n - 1)$ -dimensional Hausdorff measure, and  $\mathbf{n}$  represents the outward-facing (relative to the set  $\tilde{G} \setminus G$ ) normal vector on  $\partial(\tilde{G} \setminus G)$ .

(Results like this are well known, though perhaps not in exactly this form. In any case, we have included a formal proof, which can be found in Appendix D.)

There is a great deal of freedom in choosing  $G$  and  $\tilde{G}$ ; the idea is to choose them so as to make the surface integrals as simple as possible. Each of these surface integrals can be viewed as an expectation. Let  $\mathcal{H}(dx)$  denote the  $n - 1$ -dimensional Hausdorff measure. We can define a probability measure on  $\partial G$  by

$$\mathcal{P}(dx) \doteq \frac{1}{Z} e^{-U(x)} \mathcal{H}(dx) \text{ where } Z = \int_{\partial G} e^{-U(x)} \mathcal{H}(dx).$$

We can define  $\tilde{\mathcal{P}}(dx)$  and  $\tilde{Z}$  analogously, on  $\partial \tilde{G}$  rather than  $\partial G$ . Then

$$\begin{aligned} \text{cap}(A, \tilde{A}) &= \int_{\partial \tilde{G}} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{G, \tilde{G}^c} e^{-U} \mathcal{H}(dx) \\ &\quad - \int_{\partial G} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{G, \tilde{G}^c} e^{-U} \mathcal{H}(dx) \\ &= \tilde{Z} \int_{\partial \tilde{G}} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{G, \tilde{G}^c} \tilde{\mathcal{P}}(dx) - Z \\ &\quad \int_{\partial G} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{G, \tilde{G}^c} \mathcal{P}(dx), \end{aligned} \tag{5}$$

where, in these integrals, the normal,  $\mathbf{n}$ , points outward from both  $G$  and  $\tilde{G}$ . Let  $y_1, y_2, \dots, y_m \sim \text{iid } \mathcal{P}(dx)$ , so that

$$\frac{1}{m} \sum_{i=1}^m h_{A, \tilde{A}^c}(y_i) \mathbf{n}^T(y_i) a(y_i) \nabla h_{G, \tilde{G}^c}(y_i) \xrightarrow{m \rightarrow \infty} \int_{\partial G} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{G, \tilde{G}^c} \mathcal{P}(dx)$$

and

$$\frac{1}{m} \sum_{i=1}^m e^{U(y_i)} \xrightarrow{m \rightarrow \infty} \int_{\partial G} e^U \mathcal{P}(dx) = \frac{|\partial G|}{Z},$$

where  $|\partial G|$  is the surface area of  $G$ . Putting these together, we get the large  $n$  approximation

$$Z \int_{\partial G} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{G, \tilde{G}^c} \mathcal{P}(dx) \approx |\partial G| \frac{\sum_{i=1}^m h_{A, \tilde{A}^c}(y_i) \mathbf{n}^T(y_i) a(y_i) \nabla h_{G, \tilde{G}^c}(y_i)}{\sum_{i=1}^m e^{U(y_i)}}.$$

If we now extend all of this to  $\partial \tilde{G}$ , with  $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_n \sim \text{iid } \tilde{\mathcal{P}}(dx)$ , and put the approximations into (5), then for large  $n$  and  $m$ ,

$$\begin{aligned} \text{cap}(A, \tilde{A}) &\approx |\partial \tilde{G}| \frac{\sum_{i=1}^n h_{A, \tilde{A}^c}(\tilde{y}_i) \mathbf{n}^T(\tilde{y}_i) a(\tilde{y}_i) \nabla h_{G, \tilde{G}^c}(\tilde{y}_i)}{\sum_{i=1}^n e^{U(\tilde{y}_i)}} \\ &\quad - |\partial G| \frac{\sum_{i=1}^m h_{A, \tilde{A}^c}(y_i) \mathbf{n}^T(y_i) a(y_i) \nabla h_{G, \tilde{G}^c}(y_i)}{\sum_{i=1}^m e^{U(y_i)}}. \end{aligned} \quad (6)$$

In summary: we can use a Monte Carlo technique to compute the capacity as long as we can (1) sample from  $\mathcal{P}(dx)$  and  $\tilde{\mathcal{P}}(dx)$ , (2) compute the surface areas  $|\partial G|$  and  $|\partial \tilde{G}|$ , (3) compute the first-passage probability  $h_{A, \tilde{A}^c}$ , and (4) compute the gradient  $\nabla h_{G, \tilde{G}^c}$ . We can often make tasks (1) and (2) straightforward by a judicious choice of  $G, \tilde{G}$ . However, tasks (3) and (4) are more difficult.

Thus, to compute the capacity using this approach, the most difficult challenge is to compute the first-passage probabilities along  $G, \tilde{G}$ . Broadly speaking there are two approaches for this kind of problem. First-passage probabilities satisfy an elliptic PDE related to the infinitesimal generator [see Appendix A, Eq. (A1)], and we could therefore choose from a selection of numerical solvers. Here, in a different direction, we exploit the connection between first-passage probabilities and the underlying random walk in order to develop Monte Carlo tools suitable for estimating both  $h_{A, \tilde{A}^c}$  and  $\nabla h_{G, \tilde{G}^c}$  on the surfaces  $\partial G$  and  $\partial \tilde{G}$ . These tools are based on what we will call the ‘‘shell method,’’ which we describe briefly in the following paragraphs and in full detail in Appendix F.

Generically, given two simply connected regions  $R$  and  $\tilde{R}$ , with  $R \subset \tilde{R}$ , and a set  $S$  such that  $R \subset S \subset \tilde{R}$ , we seek an approximation to the function  $h_{R, \tilde{R}^c}$  on the surface  $\partial S$ . In principle, we could begin with a fine-grained partitioning of  $\partial S$  into simply connected ‘‘cells,’’ and for each cell simulate the diffusion many times, recording whether or not the path first exits  $\partial(\tilde{R} \setminus R)$  at  $\partial R$ . The fraction of paths that first exit at  $\partial R$  constitutes an estimate of  $h_{R, \tilde{R}^c}(x)$  for any  $x$  in the current

cell. But this is wasteful and likely infeasible in all but the simplest of settings. Much of the waste stems from the fact that the ensemble of all paths generated from all cells will likely include many near collisions of paths scattered throughout  $\partial(\tilde{R} \setminus R)$ . An alternative, divide-and-conquer approach, is to introduce multiple sets,  $S_0, S_1, \dots, S_n$  such that

$$R = S_0 \subset \dots \subset S_{m-1} \subset S_m = S \subset S_{m+1} \subset \dots \subset S_n = \tilde{R}$$

and use sample paths from  $X$ , *locally*, to estimate the transition probability matrices from each cell within each ‘‘shell’’  $\partial S_k$  to each cell of its neighboring shells,  $\partial S_{k-1}$  and  $\partial S_{k+1}$ . Equipped with these transition matrices, the first-passage probability for a given  $x \in S$  is computed algebraically, without further approximation.

$S$  must have been chosen not only to satisfy  $R \subset S \subset \tilde{R}$  but also in such a way as to make it feasible to sample from  $\partial S$  under the probability measure  $\frac{1}{Z} e^{-U} \mathcal{H}(dx)$ . After that,  $S_k$   $k = 1, \dots, n-1$  are chosen so that the shells nest and are in close proximity; the hitting times starting from a sample in  $\partial S_k$  and ending at  $\partial S_{k-1} \cup \partial S_{k+1}$  must be short enough to encourage many repeated runs. The output is a set of samples,  $z_1, \dots, z_N \sim \frac{1}{Z} e^{-U} \mathcal{H}(dx)$  on  $\partial S$  together with the approximate value of  $h_{R, \tilde{R}^c}(x)$  at each sample  $x = z_i$ . [In fact, though the main purpose is to estimate  $h_{R, \tilde{R}^c}$  on  $\partial S$ , a byproduct is a sample from  $\frac{1}{Z} e^{-U} \mathcal{H}(dx)$  on all of the shells  $\partial S_k$ , along with an estimate of  $h_{R, \tilde{R}^c}$  at every sample.] With the choice of  $A$  for  $R$  and  $\tilde{A}$  for  $\tilde{R}$ , the algorithm becomes directly applicable to the estimation of  $h_{A, \tilde{A}^c}$  on  $\partial G$  and  $\partial \tilde{G}$ , taking  $S = G$  in the former case and  $S = \tilde{G}$  in the latter.

The shell method is closely related to milestoneing [30,35,36] and Markov state models [26–28], though more tailored to the problem at hand. In particular, our interest here is in computing the first-passage probabilities rather than in approximating the underlying process. Also, the discretizations of the shells are *adaptive*, in that they are based on clusters that are derived from an ensemble of samples, as opposed to being crafted for a particular landscape; see Appendix F.

As for the required gradients, these are generally harder to estimate. Nevertheless, for the particular gradient  $\nabla h_{G, \tilde{G}^c}$ , the problem is substantially mitigated by noting that we are interested only in its evaluation on  $\partial G$  and  $\partial \tilde{G}$ , each of which is a level set of  $h_{G, \tilde{G}^c}$  ( $h_{G, \tilde{G}^c} = 1$  on  $G$  and 0 on  $\tilde{G}$ ). Consequently, on each surface the gradient is in the normal direction, and we need only estimate its magnitude. For this purpose it is enough to know the values of  $h_{G, \tilde{G}^c}$  on a surface close to  $G$  and interior to  $\tilde{G} \setminus G$  (for estimating  $\nabla h_{G, \tilde{G}^c}$  on  $G$ ) and on another surface close to  $\tilde{G}$  and also interior to  $\tilde{G} \setminus G$  (for estimating  $\nabla h_{G, \tilde{G}^c}$  on  $\tilde{G}$ ). Two such surfaces would be  $\partial S_1$  and  $\partial S_{n-1}$ , were we to apply the shell method with  $R = G$  and  $\tilde{R} = \tilde{G}$ , since, as already noted, a by-product of the method is an estimate of  $h_{R, \tilde{R}^c}$  on all of the shells. Alternatively, in the interest of better accuracy, the method could be run twice, once with  $S = S_1$ , a well-chosen outer approximation of  $G$ , and then again with  $S = S_{n-1}$ , a well-chosen inner approximation of  $\tilde{G}$ .

## V. NUMERICAL EXPERIMENTS

We use numerical experiments to test the practical relevance of our theoretical results and capacity-estimation algorithm [37]. There several questions we would like to address empirically: We used a series of experiments to investigate the accuracy and computational efficiency of the proposed approach to estimating first-passage probabilities. Three questions were addressed empirically, in each of several settings:

*Is the first-passage probability function  $\varepsilon$ -flat?* The capacity approach to first-passage probabilities (Theorem 1) is predicated on the assumption that  $h_{A,B}(x)$  is  $\varepsilon$ -flat over a region that is sufficiently removed from the immediate neighborhoods of the targets. Theorem 2 gives asymptotic (small-target or high-dimension) sufficient conditions. How quickly do these asymptotic limits come into play? We estimated, numerically, the extent of  $\varepsilon$ -flatness in each setting.

*How tight are the bounds in Theorem 1?* The theorem guarantees that the first-passage probability function is within  $\varepsilon + \sqrt{\varepsilon/2}$  of the probability determined by the target-capacity ratios. We estimated the tightness of the bound in each setting.

*Is the shell method accurate and computationally efficient?* Target capacities can rarely be computed analytically, and hence the applicability of the capacity-based approach to first-passage probabilities depends upon the accuracy and computational efficiency of the shell method (Sec. IV). We examined both in each of many settings.

We looked at problems of the following form. In every case the configuration space,  $\Omega$ , was the unit ball in  $\mathbb{R}^5$ . In other words,  $\Omega = \mathcal{B}(0, 1)$ , where  $\mathcal{B}(x, r) \triangleq \{y: \|y - x\| < r\}$ . The diffusion was confined to  $\Omega$  by a reflecting boundary at  $\partial\Omega$ , and within  $\Omega$  the dynamics were assumed to obey the first-order (high-viscosity) Langevin equation,

$$dX_t = -\nabla U(X_t) dt + dW_t, \quad (7)$$

where  $W_t$  is an  $n$ -dimensional Brownian motion and  $U$  is a potential energy. Although the method applies to multiple targets (as noted earlier), the experiments involved only two,

$$A \triangleq \mathcal{B}(x_A, r_A),$$

$$B \triangleq \mathcal{B}(x_B, r_B).$$

Given an initial condition  $X_0 \in \Omega \setminus (A \cup B)$ , we are interested in knowing whether the dynamics carry the system into  $A$  or  $B$  first. We further assumed that  $U$  is constant beyond the immediate neighborhoods of  $A$  and  $B$ : letting  $\tilde{A} = \mathcal{B}(x_A, r_{\tilde{A}})$  and  $\tilde{A} = \mathcal{B}(x_A, r_{\tilde{A}})$  such that  $r_A < r_{\tilde{A}} < r_{\tilde{A}}$ , and letting  $\tilde{B} = \mathcal{B}(x_B, r_{\tilde{B}})$  and  $\tilde{B} = \mathcal{B}(x_B, r_{\tilde{B}})$  such that  $r_B < r_{\tilde{B}} < r_{\tilde{B}}$ , we assumed that  $\nabla U(x) = 0$  for all  $x \in \Omega \setminus (\tilde{A} \cup \tilde{B})$ . Our interest is in estimating  $h_{A,B}(x)$  over the region  $x = X_0 \in \Omega \setminus \tilde{A} \cup \tilde{B}$ .

In order to establish good estimates for ground truth, we ran a total of 400 000 simulations of (7) until the first passage to  $A$  or  $B$ , in each of six experiments. This was greatly facilitated by the assumption that  $\nabla U(x) = 0$  outside of  $\tilde{A} \cup \tilde{B}$ , since we could use the “walk-on-spheres” method (cf. Ref. [38]) for the corresponding portions of the trajectory.

We considered two broadly different variations of the setup: one (Brownian diffusion) in which the *exact* target capacities could be computed, thereby allowing for a direct assessment of the accuracy of the shell method, and the other

involving complex energy landscapes in the neighborhoods of the targets. In the latter case, the accuracy of the capacity estimates could only be inferred from the accuracy of the resulting estimates of first-passage probabilities.

### A. Brownian diffusion

The first set of experiments tested the capacity-based approach in the simplest possible case:  $U(x) = 0$  for all  $x$  outside of the targets. The target centers were fixed at  $x_A = (0.5, 0.6, 0.0, 0, 0, 0.0)$  and  $x_B = (-0.7, 0.0, 0.0, 0, 0, 0.0)$ , the radii  $r_{\tilde{A}}$  and  $r_{\tilde{B}}$  were fixed at 0.2, and the radii  $r_A$  and  $r_B$  were varied across three conditions. These were defined as “small targets” ( $r_A = 0.02, r_B = 0.04$ ), “medium targets” ( $r_A = 0.05, r_B = 0.075$ ), and “large targets” ( $r_A = r_B = 0.15$ ). The three cases are illustrated, along with the corresponding results, in Fig. 2. (In light of the assumption  $U(x) = 0$ , the particular values of  $r_{\tilde{A}}$  and  $r_{\tilde{B}}$  are irrelevant in this set of experiments, though not in the next set; see Sec. VB.)

Each of the three questions raised earlier (Is the first-passage probability function  $\varepsilon$ -flat? How tight are the bounds in Theorem 1? Is the shell method accurate and computationally efficient?) was explored for each of the three target sizes.

#### 1. Is the first-passage probability function $\varepsilon$ -flat?

Theorem 2, although an asymptotic result, suggests that the variability of the first-passage probabilities will decrease with target size. To investigate this relationship we conducted 2000 diffusion simulations at each of 100 randomly selected initial conditions,  $X_0 \in \Omega \setminus \tilde{A} \cup \tilde{B}$ . These simulations yielded 100 well-estimated hitting probabilities. The histograms of these probabilities are displayed in Fig. 2. As the targets become smaller, the histograms become more peaked. A crude estimate of the flatness of  $h_{A,B}(x)$  is the spread of the 100 estimated probabilities:

- (1)  $\varepsilon \approx 0.0295$  for the small targets
- (2)  $\varepsilon \approx 0.0420$  for the medium targets
- (3)  $\varepsilon \approx 0.1060$  for the large targets.

In other words, the dependency of the hitting probability on the initial condition decreases with target size, as anticipated.

#### 2. How tight are the bounds in Theorem 1?

If we were to take these estimates of  $\varepsilon$ -flatness and apply Theorem 1, as though they were precise measurements, then we would conclude that

$$\sup_{x \notin \tilde{A}, \tilde{B}} |h_{A,B}(x) - p_A| \leq \varepsilon + \sqrt{\varepsilon/2}, \quad (8)$$

where

$$p_A \triangleq \frac{\text{cap}(A, \tilde{A})}{\text{cap}(A, \tilde{A}) + \text{cap}(B, \tilde{B})} \quad (9)$$

and  $\varepsilon + \sqrt{\varepsilon/2} = 0.0295 + \sqrt{0.0295/2} \approx 0.15$  for the small targets, and 0.19 and 0.34 for the medium and large targets, respectively.

In this special case, in which the diffusion is just a Brownian motion outside of the targets, the capacities can be

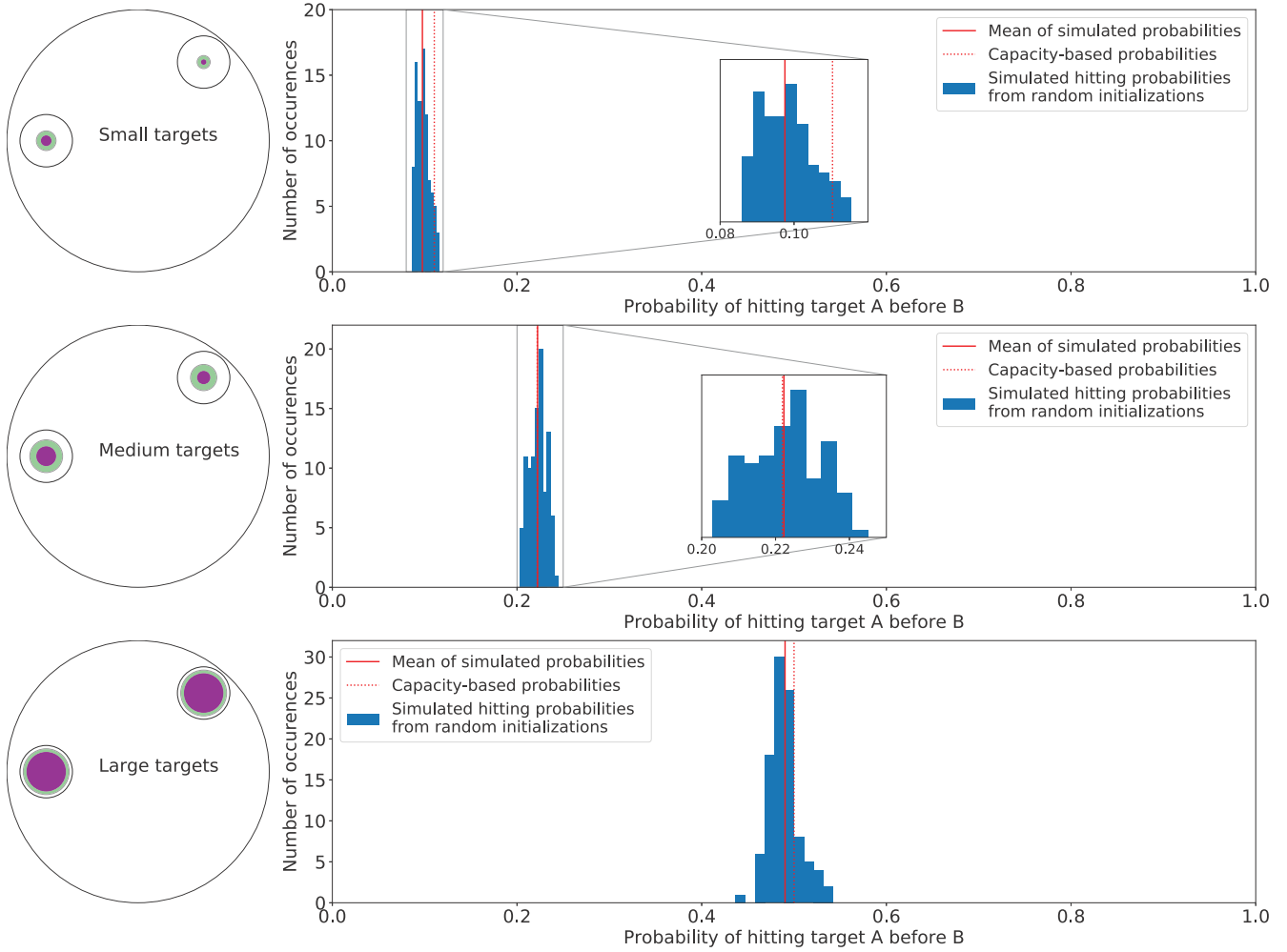


FIG. 2. Capacities and first-passage probabilities: Brownian motion. We consider three different cases (small, medium, and large), as illustrated on the left side. In each case we have two targets and are interested in the probability that a Brownian motion will first encounter target  $A$  before target  $B$ . We consider 100 random initial locations, confined to be outside of the small black circles. For each location, we use 2000 runs of the diffusion to estimate the probability of hitting target  $A$  before  $B$ . We plot these 100 estimates as a histogram. The histograms reveal that the hitting probabilities fall within a narrow range (0.0295, 0.0420, 0.1060 for small, medium, and large targets, respectively), especially when the targets are small. Moreover, we see that the location of this narrow range can be accurately predicted by the capacity-based approximation in Theorem 1 [ $p_A$  from Eq. (9), shown as a dotted red line, 0.1104, 0.2219, 0.5000 for small, medium, and large targets, respectively]; the capacity-based approximation, derived from purely *local* computations within the neighborhoods of the targets, always falls inside the span of the histogram. We can further compare the capacity-based value to an estimate of the mean hitting probability for a process initialized at a random location, which is just the relative frequency of hitting  $A$  before  $B$  in 200 000 additional simulations from random initial conditions (shown as a solid red line, 0.0975, 0.2223, 0.4903 for small, medium, and large targets, respectively.)

computed exactly via the formulas

$$\text{cap}(A, \tilde{A}) = \frac{6\pi^{\frac{5}{2}}}{\Gamma(\frac{5}{2})(r_A^{-3} - r_{\tilde{A}}^{-3})}, \quad \text{cap}(B, \tilde{B}) = \frac{6\pi^{\frac{5}{2}}}{\Gamma(\frac{5}{2})(r_B^{-3} - r_{\tilde{B}}^{-3})}, \quad (10)$$

which are easily derived using the representation in equation (4), established in Proposition 1. Together with (9), we then get the following formula for the corresponding probabilities:

$$p_A = \frac{\frac{1}{r_A^{-3} - r_{\tilde{A}}^{-3}}}{\frac{1}{r_A^{-3} - r_{\tilde{A}}^{-3}} + \frac{1}{r_B^{-3} - r_{\tilde{B}}^{-3}}},$$

which evaluates to 0.1104, 0.2219, and 0.5, respectively for the small, medium and large targets. These capacity-based estimates are indicated in Fig. 2 by the dotted red lines, which are superimposed on the corresponding histograms of the 100 empirically derived probabilities.

To the extent that the range of the 100 well-estimated hitting probabilities (2000 samples of the diffusion starting at each of the 100 randomly selected starting locations) is a reasonable estimate of the  $\varepsilon$ -flatness, for a given pair of targets, we are now in a position to evaluate the accuracy of the bound guaranteed by Theorem 1, i.e., the bound in equation (8). In particular, in each of the three examples depicted in Fig. 2 the value of  $p_A$  (dotted red line) lies within the range of the corresponding histogram. Therefore, in each of the three

TABLE I. Accuracy of the shell method: Brownian motion. There are three pairs of targets, labeled small, medium, and large. The target capacities, which can be calculated exactly for Brownian motion, are entered in the first row. The second row contains the corresponding estimates derived by the shell method (cf. Sec. IV). According to Theorem 1,  $p_A$  (9) is an estimate of the probability that the first passage occurs at  $A$ , accurate to within an error of  $\varepsilon + \sqrt{\varepsilon/2}$ . The exact values of  $p_A$  are entered in the third row, and the estimated values, derived from the estimated capacities, are in the fourth row.

Targets	Small	Medium	Large
$(\text{cap}(A, \tilde{A}), \text{cap}(B, \tilde{B}))$	(0.000632, 0.005094)	(0.010026, 0.035164)	(0.460937, 0.460937)
$(\widehat{\text{cap}}(A, \tilde{A}), \widehat{\text{cap}}(B, \tilde{B}))$	(0.000549, 0.004699)	(0.009308, 0.033625)	(0.452441, 0.451537)
$p_A$	0.1104	0.2219	0.5000
$\hat{p}_A$	0.1091	0.2168	0.5005

examples, the maximum error is smaller than our estimate of  $\varepsilon$  itself. For these examples, the theoretical bound appears to be fairly loose—in each case  $\varepsilon + \sqrt{\varepsilon/2}$  could be replaced by  $\varepsilon$ , which comes to 0.0295 instead of 0.15 for the small targets, and 0.0420 instead of 0.19 and 0.1060 instead of 0.34 for the medium and large targets. Possibly, the *asymptotic* bound ( $\varepsilon \rightarrow 0$ ), which is not addressed by Theorem 1, is dominated by  $\varepsilon$  rather than  $\sqrt{\varepsilon}$ .

### 3. Is the shell method accurate and computationally efficient?

As noted in Sec. VA 2, the special assumptions in these examples allow us to compute the exact capacities, according to Eq. (10). Hence we can directly measure the accuracy of the approximations computed with the shell method.<sup>1</sup> Table I summarizes the results for each of the three target sizes. By Theorem 1, the capacities enter into an approximation of first-passage probabilities through  $p_A$ , as defined in Eq. (9). The third and fourth rows of the table give the exact values of  $p_A$  and the corresponding values derived from the shell-estimated capacities, respectively. The differences are small relative to the spread of first-passage probabilities and, in particular, each of the estimated values lies comfortably within the range of estimated probabilities (cf. Fig. 2). The shell method does not introduce any significant additional error for these problems.

Turning now to the question of computational efficiency, how fast is a capacity-estimation approach when compared against a simulation-based approach? There are of course many variables that might affect the comparison, including the dimension and the details of the energies, not to mention the implementation details. In our experiments with direct simulations, we used the “walk-on-spheres” method to simulate trajectories in the flat region [38], JIT compilation to remove loop overhead, multi-CPU parallelization, and the coarsest time step that yielded accurate results. (In that the walk-on-spheres method requires a flat energy landscape and Theorem 1 does not, our inferences about computational efficiency are, in this regard, somewhat tilted in favor of the simulation approach.) As for capacity estimation, we made no effort to adjust the number of samples or the discretization parameters.

In both the Brownian-motion experiments discussed in this section and the experiments with nontrivial landscapes

discussed in the next section, we observed that a single run of direct simulation took about as long as estimating two capacities using the shell method, i.e., as long as it takes to compute the capacity-based estimate of the first-passage probability. With this benchmark in hand, the comparison of the two approaches comes down to the question of accuracy: How many direct simulations would be needed to compute the first-passage probability to within an accuracy comparable to what was obtained using capacities?

A single simulation run produces a single Bernoulli variable—one if target  $A$  is encountered before target  $B$ , and zero otherwise. The total number of ones, then, in a sequence of independent simulations will have a binomial distribution. Hence, given any probability  $p$  of the Bernoulli event “one,” and any desired accuracy  $\delta$ , the number of independent simulations needed to achieve a confidence interval of size  $\delta$ , with 95% confidence, is about  $n(\delta) = 4 \frac{p(1-p)}{\delta^2}$ .<sup>2</sup> Since a single simulation run takes about as long as our capacity-based estimation for  $p_A$ , we can assert that in this sense the capacity-based estimation is approximately  $n = n(\delta)$  times more efficient.

What is a reasonable value of  $\delta$ ? Here are two points of view, leading to similar estimates of  $n$ :

(1) Let  $\hat{p}_A$  be the capacity-based estimate of first encountering  $A$  after starting at a given location  $x \in \Omega \setminus \tilde{A} \cup \tilde{B}$  (indicated, for each of the three experiments, by the dotted red line in Fig. 2), and recall that  $h_{A,B}(x)$  represents the actual probability, which in general depends on the starting location,  $x$ . Keep in mind that the capacity-based estimator is independent of  $x$ . If the starting location were chosen randomly, then an unbiased estimator of the mean absolute error,  $\mathbb{E}[|h_{A,B}(X) - \hat{p}_A|]$ , for a given value of  $\hat{p}_A$ , is obtained by taking expectation with respect to the empirical probability distribution, which is displayed in the figure by a histogram for each target size. The resulting three estimates, for the small, medium, and large targets, of the mean absolute errors are 0.0127, 0.0080, and 0.0167, respectively. From this point of view (i.e., using these three values for  $\delta$ ), the number of simulation runs from a given location  $x$  needed to achieve an error approximately equal to the mean absolute error of

<sup>1</sup>With reference to Appendix F, the following parameters were used to implement the shell method:  $m = 2$ ,  $n = 4$ ,  $N_p = 100$ ,  $N_b = 3$ ,  $N_s = 1000$ , and a time step of  $10^{-7}$ .

<sup>2</sup>In principle, the value of  $p$  depends on the starting location. But in light of the tight clustering of first-passage probabilities ( $\varepsilon$ -flatness), we used the mean value obtained from 200 000 samples at random starting locations, which is designated “mean of simulated probabilities” in the figures, for all calculations of sample sizes.



the capacity method, and hence the computational advantage of the capacity method, is about 2000 for the small targets, 10 000 for the medium targets, and 3500 for the large targets.

(2) Suppose, instead, that the goal is to estimate the mean first-passage probability, i.e., the average over all starting locations  $x \in \Omega \setminus \tilde{A} \cup \tilde{B}$ . For each of the three target sizes, we used 200 000 simulation runs, each one initiated at an independent and randomly chosen starting location, and recorded the relative frequencies with which target  $A$  was reached before target  $B$ . The resulting estimates of the mean first-passage probabilities are reported in the figure and indicated by the solid red lines in the histograms. The difference between these means and the capacity-based probabilities (the distance between the red and dotted-red lines) is an approximation of the accuracy of the capacity-based approach in estimating the mean first-passage probabilities. We used these differences as  $\delta$ , but slightly corrected (made larger) to account for the binomial approximation error associated with a sample size of 200 000 (i.e., corrected for the estimation of the mean in the first place). In this way we estimate a computational advantage for the capacity-based method to be about 2000-fold for the small targets, 400 000 for the medium targets, and 8000 for the large targets.

## B. Nontrivial landscape

In the second set of experiments the targets  $A$  and  $B$  were only partially contained in  $\Omega$ , and the energy landscape  $U$  was nontrivial. We fixed the target centers at  $x_A = (0.6402, 0.7682, 0.0, 0, 0, 0.0)$  and  $x_B = (-1.0, 0.0, 0.0, 0, 0, 0.0)$ , and the outer radii at  $r_{\tilde{A}} = r_{\tilde{B}} = 0.5$ . We investigated three different cases with targets of different sizes: small (where  $r_A = 0.02$ ,  $r_{\tilde{A}} = 0.05$ ,  $r_B = 0.04$ , and  $r_{\tilde{B}} = 0.075$ ), medium (where  $r_A = 0.1$ ,  $r_{\tilde{A}} = 0.15$ ,  $r_B = 0.2$ , and  $r_{\tilde{B}} = 0.25$ ), and large (where  $r_A = r_B = 0.45$  and  $r_{\tilde{A}} = r_{\tilde{B}} = 0.475$ ). The specifications of the energy functions can be found in Appendix G. Illustrations of the three set ups and corresponding results are shown in Fig. 3. We will now revisit the three issues discussed in the previous examples, but in the context of the more complex landscapes surrounding the targets.

### 1. Is the first-passage probability function $\varepsilon$ -flat?

Following the same procedures used in Sec. VA 1, we estimated  $\varepsilon$ -flatness from the spreads of the probability histograms:

- (1)  $\varepsilon \approx 0.0530$  when the targets are small
- (2)  $\varepsilon \approx 0.0495$  when the targets are medium
- (3)  $\varepsilon > 0.75$  when the targets are large.

In the third example, the first-passage probability function  $h_{A,B}(x)$  can hardly be called flat. Evidently, the large targets are not subject to the asymptotic guarantees offered in Theorem 2. In this circumstance, Theorem 1 is irrelevant, and therefore we will discuss the remaining questions only in the context of the small and medium targets.

### 2. How tight are the bounds in Theorem 1?

In other words, how large is  $\sup_x |h_{A,B}(x) - p_A|$  on  $\Omega \setminus \tilde{A} \cup \tilde{B}$  compared to  $\varepsilon + \sqrt{\varepsilon/2}$ ? For Brownian motion (Sec. VA),

the capacities were derivable analytically [Eq. (10)], which led to a closed-form expression for  $p_A$ . Here, we followed the same reasoning as in Sec. VA 2 to get an approximation of the tightness of the theorem's bound, but replaced  $p_A$  by  $\hat{p}_A$ , the estimate computed from the shell method. The resulting value of  $\hat{p}_A$ , for each target size, is displayed in Fig. 3 as a dotted red line superimposed on the target-probability histogram. For both the small and medium targets, the estimated hitting probability lies within the range of the 100 sampled probabilities, as it did in the case of Brownian motion. Reasoning as we did there, these observations suggest that the true error may be better approximated by  $\varepsilon$  rather than the  $\varepsilon + \sqrt{\varepsilon/2}$  guarantee given in the theorem. Specifically, the magnitudes of the ranges, 0.0530 and 0.0495, respectively, serve as ballpark upper bounds on the estimation errors for the first-passage probabilities for the small and medium targets.

### 3. Is the shell method accurate and computationally efficient?

How accurate is the shell method in these cases? In the absence of a closed-form expression for the capacities, we are forced to rely on an indirect measure. In Sec. VB 2 we used the shell method to estimate capacities and used those capacities to get hitting-probability estimates. These were in good agreement with the results obtained from 200 000 runs of simulated diffusions (2000 from each of 100 random starting points). This, then, constitutes indirect evidence for the accuracy of the shell method, but with the important proviso that the estimated hitting probabilities depend only on the *ratio* of the (estimated) capacities, and not on the capacities themselves.

To assess computational efficiency, we repeated the approach used for Brownian motion, step by step as laid out in Sec. VA 3. As noted there, one simulation run is nearly the same as using the shell method to estimate  $p_A$ , in terms of computational cost. Therefore, efficiency comparisons between the shell method (capacity-based) approach and direct simulation come down to computing the number of simulation runs required to achieve comparable accuracy in estimating first-passage probabilities. In Sec. VA 3 we explored two interpretations of accuracy, which we repeat here, but using instead the results reported in Fig. 3:

(1) For each target size (small and medium) and corresponding value of  $\hat{p}_A$ , we used the probability histogram to estimate  $\mathbb{E}[|h_{A,B}(X) - \hat{p}_A|]$ , the mean absolute error of the capacity-based method. The resulting accuracy measures came to 0.0078 for the small targets and 0.0095 for the medium targets. Then, using the respective probabilities of 0.8217 and 0.3893 (solid red lines), we conclude that in each case about 10 000 simulation runs would be needed to achieve an accuracy comparable to capacity-based approach. In this sense, the capacity approach, based on shell-method estimates, is about 10 000-fold more efficient than direct simulation.

(2) In the second interpretation, we consider the problem of estimating the mean first-passage probability for a randomly chosen starting position (approximately 0.8154 for the small targets and 0.3893 for the medium targets, based on 200 000 independent random samples). Errors in the capacity-based approach were about 0.0072 and 0.0089, which come

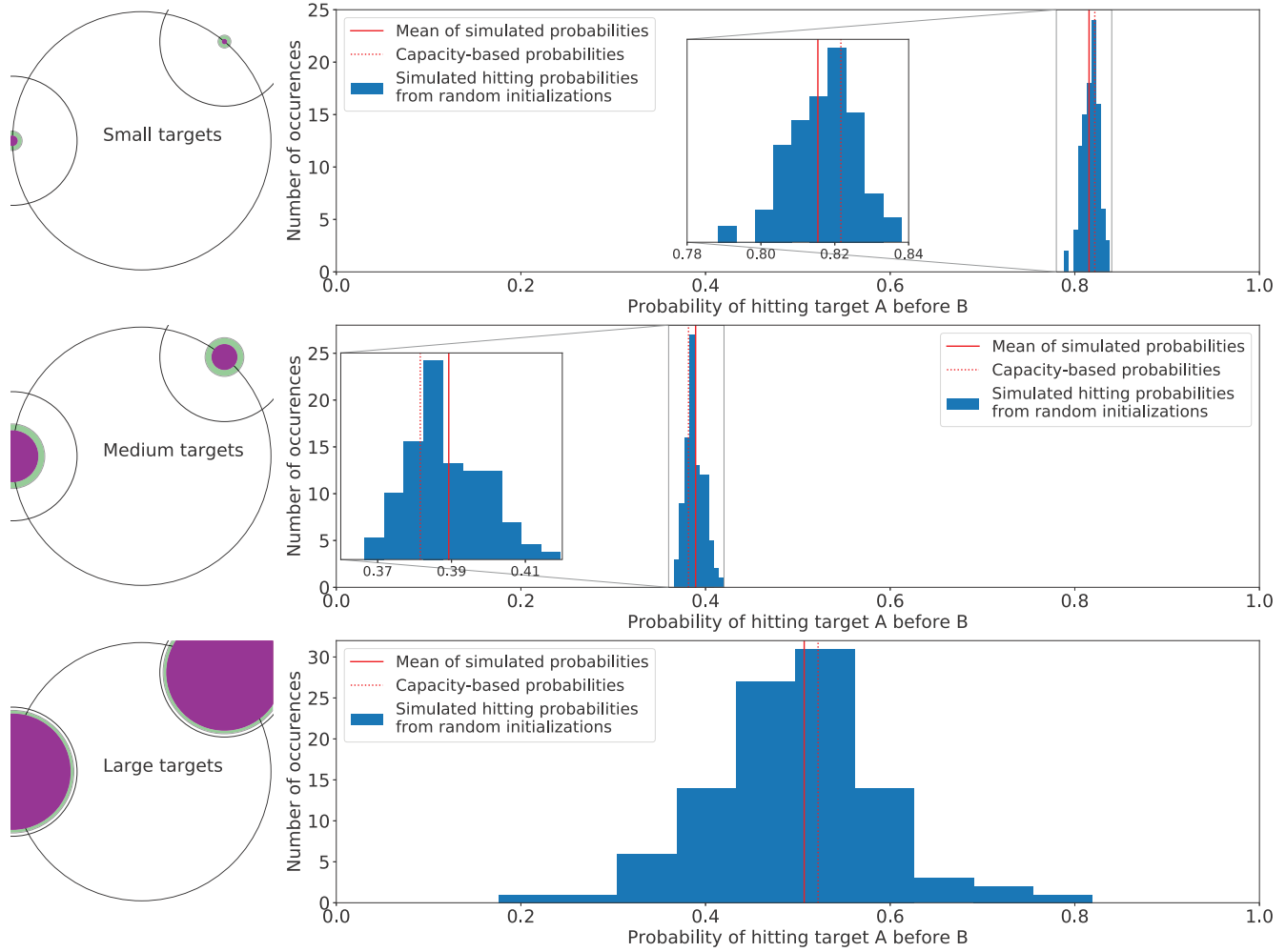


FIG. 3. Capacities and first-passage probabilities: nontrivial energy landscapes. See Fig. 2. The setup here is the same except that the targets  $A$  and  $B$  are only partially contained in  $\Omega$ , and the energy landscapes become complicated near the targets. The histograms show that the first-passage probabilities are largely independent of starting location for the small and medium targets, as they were in the previous set of experiments. In contrast, the first-passage probabilities for the large targets depend strongly on starting location. This is expected, since these targets are substantially larger than the corresponding targets from the previous experiments. (The size difference is even bigger than it might appear, given that these figures represent two-dimensional slices of a five-dimensional region.) In contrast to the Brownian motion examples, the complex landscape in these examples precludes an exact evaluation of  $p_A$  [Eq. (9)]. Hence the dotted red line represents the *estimated* value,  $\hat{p}_A$ , computed using the shell method (cf. Fig. 2). Nonetheless, these estimates fall within the span of the respective histograms in all three examples. For small, medium, and large targets, the ranges of the 100 hitting probabilities are 0.0495, 0.0530, 0.6430, respectively; the capacity-based hitting probabilities (dotted lines) are 0.8217, 0.3815, 0.5219, respectively; the mean hitting probabilities (solid lines), estimated using 200 000 additional simulations from random initial conditions, are 0.8154, 0.3893, 0.5069, respectively.

from the differences in the positions of the red and dotted-red lines, but inflated to account for the expected sampling error when using a sample of size  $n = 200\,000$ . Comparable accuracies would require about 12 000 simulation runs, for each of the small and medium targets. Hence, from this point of view, we would estimate that the capacity-based approach is about 12 000-fold more efficient than direct simulation.

### VI. SUMMARY AND DISCUSSION

If a first-passage probability is nearly the same for every initial condition in a region  $M$ , then it can be well-approximated in terms of certain integrals. We dub these integrals the “first-passage capacities.” These capacities can be approximately computed using a milestoning-like tech-

nique, as described in Sec. IV. In Theorem 1 we show that the error introduced by this approximation is small whenever hitting probabilities are sufficiently similar across different initial conditions inside a large region,  $M$ . When this condition holds, we say that the hitting probabilities are “ $\varepsilon$ -flat,” where the coefficient  $\varepsilon$  bounds how much the hitting probabilities vary over the initial conditions. Under what circumstances are hitting probabilities similar for all initial conditions in a large region,  $M$ ? Theorem 2 provides some sufficient conditions, though the concept is more general and we would expect some form of  $\varepsilon$ -flatness to hold anytime the mixing rate of the diffusion is fast relative to the typical escape time from  $M$ .

The capacity of a given target is local, in the sense that the capacity integral is over a region in the neighborhood

of the target. We have devised a Monte Carlo approach to estimating capacities that we call the shell method (Sec. IV). In Sec. V we provide evidence for the accuracy and computational efficiency of using estimated capacities and Theorem 1 to compute first-passage probabilities. In a series of computational experiments we compared our capacity-based approach to the direct estimation of hitting probabilities using repeated runs of the diffusion. In each of these experiments, repeated simulation required at least three orders of magnitude more computation to achieve comparable accuracy.

We were motivated by the narrow-escape problem, and in particular the potential connection to modeling the folding of large molecules. To the extent that folding can be thought of as a series of discrete steps, each requiring a rate-limiting search for new connections in a large conformational space [9–13], it might be possible to sidestep the exploratory phases by computing first-passage probabilities—one for each potential connection. An example might be the sequence of stem formations in the folding of structural RNA molecules. The idea would be to model the creation of each new stem as the outcome of a first-passage problem, defined by the already established stems. Not only does the search, as opposed to the stem creation, have to be the rate-limiting step, but it must be slow enough to allow for the hitting probabilities to be nearly independent of the starting configuration. But the larger problem would be the possibility of the dissolution of an existing stem. RNA folding is not necessarily monotonic; stems come and go. Which will come first, the undoing of an existing stem or the creation of a new one? Unless the undoing of a stem could be formulated as itself a target or some kind of narrow escape, we would be forced to confront the problem of estimating first-passage *times* in addition to probabilities.

In general, it may turn out that estimating first-passage times is profoundly more difficult than estimating first-passage probabilities. In this work we showed that the first-passage probabilities are sometimes invariant to the diffusion dynamics on the interior of a large set  $M$ , but the first-passing times will admit no such invariance. If the dynamics inside of  $M$  are genuinely intricate, it appears that estimating hitting times will always require a correspondingly intricate analysis.

#### ACKNOWLEDGMENT

S.G. was partially supported by the Office of Naval Research under Contracts No. ONR N000141613168 and No. ONR N000141512267.

#### APPENDIX A: THREE PERSPECTIVES ON HITTING PROBABILITIES

Many of our results are based on the fact that hitting probabilities can actually be seen from three distinct perspectives. To prepare for these results, we take a moment to review these three perspectives here. Let  $A, B \subset \Omega$ , disjoint, open with smooth boundary.

(1) Hitting probabilities: Let  $\tau_S \triangleq \inf\{t: X_t \in S\}$  for any set  $S$  and  $h_{A,B}(x) \triangleq \mathbb{P}(X_{\tau_{A \cup B}} \in \partial A | X_0 = x)$ .

(2) Elliptic equation: Let  $h_{A,B}^{\text{dir}}(x)$  denote the solution to the partial differential equation

$$\begin{aligned} 0 &= \sum_{i=1}^n b_i(x) \frac{\partial u(x)}{\partial x_i} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij}(x) \frac{\partial^2 u(x)}{\partial x_i \partial x_j} \quad x \notin \bar{A}, \bar{B}, \\ 1 &= u(x), x \in A, \\ 0 &= u(x), x \in B, \\ 0 &= \mathbf{n}(x)^T a(x) \nabla u(x), x \in \partial \Omega. \end{aligned} \quad (\text{A1})$$

This solution is unique and smooth [39]. What's more, it is equal to the hitting probability function:  $h_{A,B}^{\text{dir}}(x) = h_{A,B}(x)$  (cf. Sec. 6.7 of Chen [40]).

(3) Variational form: For any open set  $S \subset \Omega$  let  $\mathcal{E}_S(f, g) \triangleq \int_S \nabla f(x)^T a(x) \nabla g(x) \rho(dx)$  denote the ‘‘Dirichlet form’’ of  $f, g$  on the domain  $S$ . Let  $\mathcal{L}^2(S, \rho)$  denote the Hilbert space of functions on  $S$  which are square-integrable with respect to  $\rho$ . Let  $\mathcal{H}^1(S, \rho) = W^{1,2}(S) \subset \mathcal{L}^2(S)$  denote the corresponding once weakly differentiable Hilbert-Sobolev space. We define  $h_{A,B}^{\text{var}}(x)$  as the solution to

$$\begin{aligned} \min_{u \in \mathcal{H}^1(S)} \quad & \mathcal{E}_S(u, u) \\ \text{subject to} \quad & u(x) = 1, x \in \partial A \\ & u(x) = 0, x \in \partial B, \end{aligned}$$

where  $S = \Omega \setminus (A \cup B)$ . This solution is unique and equal to  $h_{A,B}^{\text{dir}}$  on  $S$  (cf. Sec. 4 of Dret [41]). This variational perspective leads us to the notion of the ‘‘condenser capacity’’ associated with  $h_{A,B}$ . It is defined as

$$\text{cap}(A, \Omega \setminus B) \triangleq \mathcal{E}_S(h_{A,B}, h_{A,B}),$$

where again  $S = \Omega \setminus (A \cup B) = (\Omega \setminus B) \setminus A$ .

We will use all three of these perspectives to show our results. For example, consider how the hitting probability perspective helps us show a result about capacities:

*Proposition 2.* Let  $A \subset \bar{A}, B \subset \bar{B}$  with  $\bar{A}, \bar{B}$  disjoint. Then  $\text{cap}(A \cup B, \bar{A} \cup \bar{B}) = \text{cap}(A, \bar{A}) + \text{cap}(B, \bar{B})$ .

*Proof.* This result is certainly known, but we include a proof here because we were unable to find a proof in the literature. Since  $\bar{A}, \bar{B}$  are disjoint and  $X$  is continuous, the process cannot cross from one to the other without hitting the boundary. Thus we have  $\tau_{\partial \bar{A} \cup \partial \bar{B} \cup \partial A \cup \partial B} = \tau_{\partial \bar{A} \cup \partial A}$  as long as  $X_0 \in \bar{A}$ . We get a symmetric result if  $X_0 \in \bar{B}$ . It follows that

$$h_{A \cup B, (\bar{A} \cup \bar{B})^c}(x) = \begin{cases} h_{A, \bar{A}^c}(x) & \text{if } x \in \bar{A} \\ h_{B, \bar{B}^c}(x) & \text{if } x \in \bar{B} \end{cases}$$

We can now use this probabilistic perspective to help us understand the capacity by articulating it as the Dirichlet form on the relevant hitting probability functions

$$\begin{aligned} \text{cap}(A \cup B, \bar{A} \cup \bar{B}) &= \int_{\bar{A} \cup \bar{B} \setminus (A \cup B)} \|\sigma \nabla h_{A \cup B, (\bar{A} \cup \bar{B})^c}\|^2 \rho(dx) \\ &= \int_{\bar{A} \setminus A} \|\sigma \nabla h_{A, \bar{A}^c}\|^2 \rho(dx) \\ &\quad + \int_{\bar{B} \setminus B} \|\sigma \nabla h_{B, \bar{B}^c}\|^2 \rho(dx) \\ &= \text{cap}(A, \bar{A}) + \text{cap}(B, \bar{B}) \end{aligned}$$

as desired.  $\blacksquare$

**APPENDIX B: PROOF OF THEOREM 2**

*Theorem 2.* (Examples of  $\varepsilon$ -flatness). Let  $\mathcal{B}(x, r) = \{y: \|x - y\| < r\}$  denote the ball of radius  $r$  centered at  $x$ . Let us say that a diffusion governed by Eq. (1) “behaves like Brownian motion” on a set if  $\sigma(x) = I$  and  $\nabla U(x) = 0$  for every  $x$  in that set.

(1) (*Small targets*) Fix  $r, \varepsilon > 0$ . We can then find  $r' > 0$  with the following property. For any bounded convex set  $\Omega$  with diameter less than or equal to 1, any  $x_A, x_B \in \Omega$ , any reversible stationary diffusion trapped inside  $\Omega$  and behaving like Brownian motion for all  $x \notin \mathcal{B}(x_A, r') \cup \mathcal{B}(x_B, r')$ , and any  $A \subset \mathcal{B}(x_A, r'), B \subset \mathcal{B}(x_B, r')$ , we have that  $h_{A,B}$  is  $\varepsilon$ -flat on  $\Omega \setminus (\mathcal{B}(x_A, r) \cup \mathcal{B}(x_B, r))$ .

(2) (*High dimensions*) Fix  $r, \varepsilon > 0$  and  $r' < r$ . We can then find  $n$  with the following property. For any convex set  $\Omega \subset \mathbb{R}^n$  with diameter less than or equal to 1, any  $x_A, x_B$  such that  $\mathcal{B}(x_A, r), \mathcal{B}(x_B, r) \subset \Omega$ , any reversible stationary diffusion trapped inside  $\Omega$  and behaving like Brownian motion for all  $x \notin \mathcal{B}(x_A, r') \cup \mathcal{B}(x_B, r')$ , and any  $A \subset \mathcal{B}(x_A, r'), B \subset \mathcal{B}(x_B, r')$ , we have that  $h_{A,B}$  is  $\varepsilon$ -flat on  $\Omega \setminus [\mathcal{B}(x_A, r) \cup \mathcal{B}(x_B, r)]$ .

*Proof.* Let  $\dot{A} = \mathcal{B}(x_A, r'), \dot{B} = \mathcal{B}(x_B, r')$ . Let  $\{Z_t\}_t$  denote a Brownian motion trapped inside  $\Omega$  and coupled to  $X$  such that  $X_t = Z_t$  for all  $t \leq \tau = \{\inf t: X_t \in \dot{A} \cup \dot{B}\}$  (we can do this because we assumed that  $X$  behaves like Brownian motion outside of  $\dot{A}, \dot{B}$ ). Let  $Z_\infty$  denote the stationary distribution of  $Z$ , i.e., the uniform distribution on  $\Omega$ . Using Lemma 1, we have that

$$|\mathbb{E}[h_{A,B}(Z_t) - h_{A,B}(Z_\infty)|Z_0 = x]| \leq \frac{1}{t} \quad \forall x.$$

On the other hand, Dynkin’s formula gives that

$$h_{A,B}(x) = \mathbb{E}[h_{A,B}(Z_{\tau \wedge t})|Z_0 = x] \quad \forall t, x.$$

Putting those two facts together:

$$\begin{aligned} & |h_{A,B}(x) - \mathbb{E}[h_{A,B}(Z_\infty)]| \\ &= |\mathbb{E}[h_{A,B}(Z_{\tau \wedge t}) - h_{A,B}(Z_\infty)|Z_0 = x]| \\ &\leq |\mathbb{E}[h_{A,B}(Z_{\tau \wedge t}) - h_{A,B}(Z_t)|Z_0 = x]| \\ &\quad + |\mathbb{E}[h_{A,B}(Z_t) - h_{A,B}(Z_\infty)|Z_0 = x]| \\ &\leq \mathbb{P}(\tau \leq t|Z_0 = x) + \frac{1}{t}. \end{aligned}$$

To make this small, it thus suffices to get  $t$  large but keep  $\mathbb{P}(\tau \leq t)$  small. In short, it suffices to show that  $\tau$  is usually big. We prove this differently in the two different cases:

(1) The first example follows because the targets are vanishing into nothing, so it is no surprise that the time to hit the targets will get longer and longer. However, note that this intuition is valid only because the ambient dimension  $n$  is at least three (which we have assumed throughout). A formal proof of this argument can be found in Lemma 2.

(2) The second example is similar; even though the targets maintain the same diameter, their relative volume is vanishing because the ambient dimension  $n$  is increasing. Thus the targets are effectively becoming smaller, so the hitting time increases. This matter is a bit trickier, and so for this second case note we have made additional requirements that

$\mathcal{B}(x_A, r), \mathcal{B}(x_B, r) \subset \Omega$ . A formal proof of this argument can be found in Lemma 3. ■

*Lemma 1.* (Uniform ergodicity) Let  $\Omega \subset \mathbb{R}^d$  be a convex set with diameter  $\xi$  and let  $M$  denote a Brownian motion trapped by reflecting boundaries inside  $\Omega$ . The distribution of  $M_t$  converges uniformly to the uniform distribution, in the sense that

$$\sup_{f: \Omega \rightarrow [0,1]} |\mathbb{E}[f(M_t) - f(Z)|M_0 = x]| \leq \xi^2/4t, \quad \forall x \in \Omega, t > 0,$$

where  $Z$  is uniformly distributed on  $\Omega$ .

*Proof.* Per Loper [42],

$$\sup_{f: \Omega \rightarrow [0,1]} |\mathbb{E}[f(M_t) - f(Z)|M_0 = x]| \leq \mathbb{P}(\tau > 4t),$$

where  $\tau$  is the first exit time of a one-dimensional Brownian motion from the interval  $[-\xi, \xi]$  when initialized at the origin. Dynkin’s formula gives that  $\mathbb{E}[\tau] = \xi^2$ . Hence, by the Markov inequality,  $\mathbb{P}(\tau > 4t) \leq \xi^2/4t$ . ■

*Lemma 2.* Let  $Z$  denote a Brownian motion trapped inside a convex bounded open set  $\Omega \subset \mathbb{R}^n$  with smooth boundary. We further assume that  $n \geq 3$ . Fix  $x \in \mathbb{R}^n$  and  $r, \varepsilon > 0$ . Let  $\tau_A$  denote  $\inf\{t: Z_t \in A\}$  for any set  $A$ . We can always find  $r_1 > 0$  such that  $\mathbb{P}(\tau_A < 1/\varepsilon|Z_0 = z) < \varepsilon$  for all  $z \notin \mathcal{B}(x, r_3)$  and all  $A \subset \mathcal{B}(x, r_1)$ .

*Proof.* First consider the case that  $x \in \Omega$ . To measure the time it takes to hit  $\mathcal{B}(x, r_1)$ , we will make use of two additional concentric spheres around  $x$ . First, find  $r_3$  such that  $\mathcal{B}(x, r_3) \subset \Omega$ . Second, let  $r_2 = \sqrt{r_1 r_3}$ , an intermediate radius. We will use these spheres to break the trajectory of  $Z$  into smaller manageable pieces. If  $\|Z_0 - x\| > r_3$ , the first part of the trajectory must carry it to a radius of  $r_2$  before it reaches  $r_1$ . The next part of the trajectory will then do one of two things: either it will carry on to  $r_1$ , or it will first exit the ball of radius  $r_3$ . Assuming it exits the ball of radius  $r_3$ , the strong Markov property can be used to show that we are essentially in the same place we started. Under this latter condition, the total time will be the the time we’ve spent so far plus a new variable which has the same properties as the original time. We can use this recursive relation to put a lower bound on the overall time.

Let us make this rigorous. To measure the tails of  $\tau$ , we will be interested in

$$g(z) = \mathbb{E}[e^{-\frac{1}{2}\tau}|Z_0 = z].$$

To get a lower bound on the moments of  $\tau$ , we need to get an upper bound on this object. To do so, we will use two objects of interest:

$$\begin{aligned} g_3 &= \sup_{z: \|z-x\|=r_3} g(z), \\ g_2 &= \sup_{z: \|z-x\|=r_2} g(z), \end{aligned}$$

using two stopping times of interest:

$$\begin{aligned} T_2 &= \inf\{t: \|Z_t - x\| = r_2\}, \\ T_3 &= \inf\{t > T_1: \|Z_t - x\| = r_3\}. \end{aligned}$$

Applying the strong Markov property, we observe that

$$\begin{aligned} g_2 &= \sup_{z: \|z-x\|=r_2} \mathbb{E}[\mathbb{E}[e^{-\frac{1}{2}\tau} | Z_{T_3 \wedge \tau} | Z_0 = z]] \\ &= \sup_{z: \|z-x\|=r_2} \mathbb{E}[e^{-\frac{1}{2}T_3 \wedge \tau} g(Z_{T_3 \wedge \tau}) | Z_0 = z] \\ &\leq \sup_{z: \|z-x\|=r_2} \mathbb{E}[e^{-\frac{1}{2}T_3} g_3 \mathbb{1}_{T_3 < \tau} + e^{-\frac{1}{2}\tau} \mathbb{1}_{T_3 > \tau} | Z_0 = z]. \end{aligned}$$

On the other hand, applying the continuity of Brownian motion we have that  $T_2 < \tau$  whenever  $Z_0$  lies outside the ball of radius  $r_2$ , so

$$g_3 = \sup_{z: \|z-x\|=r_3} \mathbb{E}[e^{-\frac{1}{2}T_2} g(T_2) | Z_0 = z] \leq g_2.$$

So in fact

$$\begin{aligned} g_3 &\leq \sup_{z: \|z-x\|=r_2} \mathbb{E}[e^{-\frac{1}{2}T_3} \mathbb{1}_{T_3 < \tau} g_3 + e^{-\frac{1}{2}\tau} \mathbb{1}_{T_3 > \tau} | Z_0 = z], \\ g_3 &\leq \frac{\sup_{z: \|z-x\|=r_2} \mathbb{E}[e^{-\frac{1}{2}\tau} \mathbb{1}_{T_3 > \tau} | Z_0 = z]}{1 - \sup_{z: \|z-x\|=r_2} \mathbb{E}[e^{-\frac{1}{2}T_3} \mathbb{1}_{T_3 < \tau} | Z_0 = z]} = L. \end{aligned}$$

To calculate this upper-bound  $L$ , defined above, Wendel gives explicit formulas (cf. Ref. [43]). We caution that in this document a somewhat nonstandard notation is used, namely Wendel uses  $\mathbb{E}[A; B]$  to indicate  $\mathbb{E}[A \times \mathbb{1}_B]$ . For any  $z$  with  $\|z - x\| = r_2$ , Wendel uses the symmetry of the problem to show that

$$\begin{aligned} \mathbb{E}[e^{-\frac{1}{2}\tau} \mathbb{1}_{T_3 > \tau} | Z_0 = z] &= \left(\frac{r_1}{r_2}\right)^h \frac{I(r_3)K(r_2) - I(r_2)K(r_3)}{I(r_3)K(r_1) - I(r_1)K(r_3)}, \\ \mathbb{E}[e^{-\frac{1}{2}\tau} \mathbb{1}_{T_3 < \tau} | Z_0 = z] &= \left(\frac{r_3}{r_2}\right)^h \frac{I(r_1)K(r_2) - I(r_2)K(r_1)}{I(r_1)K(r_3) - I(r_3)K(r_1)}, \end{aligned}$$

where  $h = (n - 2)/2$  and  $I, K$  denote the modified Bessel functions of order  $h$ . This leads to the following explicit formula for  $L$ :

$$L = \frac{\left(\frac{r_1}{r_2}\right)^h (I_3 K_2 - I_2 K_3)}{I_2 K_1 \left[\frac{I_3}{I_2} - \left(\frac{r_3}{r_2}\right)^h\right] + I_1 K_2 \left[\left(\frac{r_3}{r_2}\right)^h - \frac{K_3}{K_2}\right]}, \tag{B1}$$

where we denote  $I_1 = I(r_1), K_1 = K(r_1), I_2 = I(r_2)$ , and so on. To prove our result, it thus suffices to show we can drive  $L$  to zero by taking  $n$  sufficiently large.

The numerator of  $L$  in Eq. (B1) converges to zero as  $r_1 \rightarrow 0$ , because  $\frac{r_1}{r_2} \rightarrow 0, h > 0$  and the other terms are constant. On the other hand, the denominator explodes, because as  $r_1 \rightarrow 0$  we have

$$\begin{aligned} \left[\frac{I_3}{I_2} - \left(\frac{r_3}{r_2}\right)^h\right] K_1 I_2 &\rightarrow +\infty, \\ \left[\left(\frac{r_3}{r_1}\right)^h - \frac{K_3}{K_2}\right] I_1 K_2 &\rightarrow 0. \end{aligned}$$

These limits follow immediately from three properties of Bessel functions:

- (1)  $K(x) \rightarrow \infty, I(x) \rightarrow 0$  as  $x \rightarrow 0$  for  $h > 0$ ,
- (2)  $K(x), I(x) > 0$  for  $x > 0, h > 0$ ,
- (3)  $\frac{I(y)}{I(x)} > \left(\frac{y}{x}\right)^h$  for  $y > x$  and  $h > 0$ .

The first two properties are well known and can be found in DLMF (cf. Ref. [44]); the second can be found in Baricz (cf. Refs. [44,45]). In conclusion, since the numerator vanishes and the denominator explodes, we have that overall  $L$  vanishes. Applying a Chernoff bound, we have our result for the case that  $x \in \Omega$ .

What happens if  $x$  is not in the closure of  $\Omega$ ? Then by taking  $r_1$  sufficiently small we can ensure that  $A \cap \Omega = \emptyset$ , so  $\tau = \infty$ . Thus the result is proved automatically.

What if  $x$  is precisely on the boundary of  $\Omega$ ? Here we must use the smoothness of the boundary of  $\Omega$ . By picking  $r_3$  sufficiently small, the smoothness guarantees that  $\Omega$  is arbitrarily well approximated by a half-plane when restricted to a ball  $\mathcal{B}(x, r_3)$ . In this half-plane case, identical arguments to the ones above can be used to argue that we can always take a  $r_1$  that is yet smaller than  $r_3$  and ensure that the hitting time is arbitrarily long. Finally, to account for the slight discrepancy between the half-plane case and the actual case for some finite  $r_3$ , we appeal to the continuity of the Poisson equation which governs  $g$  with respect to boundary condition (cf. Ref. [46]). ■

*Lemma 3.* Fix  $0 < \epsilon$  and  $0 < r_1 < r_3$ . Then we can always find an  $n$  with the following property. Pick any open set  $\Omega \in \mathbb{R}^n$ , any  $x$  satisfying  $\mathcal{B}(x, r_3) \subset \Omega$ , and any  $x_0 \notin \mathcal{B}(x, r_3)$ . Consider a Brownian motion  $Z$  trapped inside  $\Omega$  by reflecting boundaries and let  $\tau$  denote the first hitting time of the process to  $\mathcal{B}(x, r_1)$ . Our choice of  $n$  guarantees that

$$\mathbb{P}(\tau < 1/\epsilon | Z_t = x_0) \leq \epsilon.$$

*Proof.* Following the same arguments found in Lemma 2, we again obtain that  $\mathbb{E}[e^{-\tau/2} | Z_0 = x_0]$  is bounded by

$$L = \frac{\left(\frac{r_1}{r_2}\right)^h (I_3 K_2 - I_2 K_3)}{I_2 K_1 \left[\frac{I_3}{I_2} - \left(\frac{r_3}{r_2}\right)^h\right] + I_1 K_2 \left[\left(\frac{r_3}{r_2}\right)^h - \frac{K_3}{K_2}\right]}. \tag{B2}$$

To show our proof, it thus suffices to show we can drive this quantity to zero and then apply a Chernoff bound.

Let us first look at the numerator of  $L$  in Eq. (B1). Asymptotics from the DLMF give that as  $h \rightarrow \infty$  we have

$$I(x) \sim \frac{x^h}{2^h \Gamma(h + 1)}, \quad K(x) \sim \frac{2^h \Gamma(h + 1)}{(2h)x^h}.$$

Here by  $f_1(h) \sim f_2(h)$  we mean ‘‘asymptotic equivalence,’’ i.e.,  $\lim_{h \rightarrow \infty} f_1(h)/f_2(h) = 1$ . Plugging these in, we see that both terms in the numerator are asymptotically vanishing.

Now let us turn to the denominator. Using the asymptotic formula above, we first note that

$$\begin{aligned} \left(\frac{r_3}{r_2}\right)^h I_1 K_2 &\sim \frac{1}{2h} \left(\frac{r_3 r_1}{r_2^2}\right)^h = \frac{1}{2h} \rightarrow 0, \\ -\frac{K_3}{K_2} I_1 K_2 &\sim -\frac{1}{2h} \left(\frac{r_1}{r_3}\right)^h \rightarrow 0. \end{aligned}$$

So those terms are negligible. However, the other two terms of the denominator are in fact exploding: one to positive infinity and one to negative infinity. To understand this delicate balance, we these we turn to Lemma 4. Applying this Lemma

and the asymptotics of the DLMF, we obtain that

$$\begin{aligned} \left[ \frac{I_3}{I_2} - \left( \frac{r_3}{r_2} \right)^h \right] K_1 I_2 &\geq \frac{r_3^h}{I_2} \times \frac{r_3^2 - r_2^2}{2^{h+2}\Gamma(h+2)} K_1 I_2 \\ &\sim r_3^h \times \frac{r_3^2 - r_2^2}{2^{h+2}\Gamma(h+2)} \frac{2^h \Gamma(h+1)}{2hr_1^h} \\ &= \left( \frac{r_3}{r_1} \right)^h \left[ \frac{r_3^2 - r_2^2}{8h(h+1)} \right], \end{aligned}$$

which is indeed exploding to positive infinity as  $h = (n - 2)/2 \rightarrow \infty$ .

Thus, since the numerator vanishes and the denominator explodes to positive infinity, we have that  $L$  vanishes as  $n \rightarrow \infty$ . ■

*Lemma 4.* Let  $I$  denote the modified bessel function of the first kind of order  $h$ . If  $a > b$  then

$$\frac{I(a)}{I(b)} - \left( \frac{a}{b} \right)^h \geq \frac{a^h}{I(b)} \times \frac{a^2 - b^2}{2^{h+2}\Gamma(h+2)}.$$

*Proof.* Recall that  $I$  may be defined as

$$I(x) = \sum_{m=0}^{\infty} \frac{x^{h+2m}}{2^{h+2m}\Gamma(m+h+1)\Gamma(m+1)}.$$

Thus

$$\begin{aligned} \frac{I(a)}{I(b)} - \left( \frac{a}{b} \right)^h &= \frac{I(a)b^h - a^h I(b)}{b^h I(b)} \\ &= \frac{a^h b^h \sum_{m=0}^{\infty} \frac{a^{2m} - b^{2m}}{2^{h+2m}\Gamma(m+h+1)\Gamma(m+1)}}{b^h I(b)}. \end{aligned}$$

Since  $a > b$ , we have that  $a^{2m} - b^{2m}$  is always positive. Thus we can get a lower bound by simply taking one of the terms. Choosing  $m = 1$ , we get our final result. ■

**APPENDIX C: PROOF OF THEOREM 1**

Let  $A \subset \tilde{A} \subset \Omega, B \subset \tilde{B} \subset \Omega$ . Let  $\tilde{A}, \tilde{B}$  be disjoint and  $h_{A,B}(x)$   $\varepsilon$ -flat with respect to  $\Omega \setminus (\tilde{A} \cup \tilde{B})$ . We assume the set boundaries are all smooth.

Under these conditions, we will show we can use local capacities to get good approximations for  $h_{A,B}(x)$  when  $x \notin \tilde{A}, \tilde{B}$ . To do so, our key idea is to uncover upper and lower bounds on the value of the Dirichlet form applied to this function,  $\mathcal{E}(h_{A,B}, h_{A,B})$ . We will see that these bounds can be understood in terms of local capacities, and the resulting inequalities will then yield our main result in the form of Theorem 1.

*Lemma 5.* Let  $S = \Omega \setminus (A \cup B)$ . The Dirichlet form of  $h_{A,B}$  on  $S$  can be upper-bounded in terms of the capacities:

$$\mathcal{E}_S(h_{A,B}, h_{A,B}) \leq \frac{\text{cap}(A, \tilde{A}) \text{cap}(B, \tilde{B})}{\text{cap}(A, \tilde{A}) + \text{cap}(B, \tilde{B})}.$$

*Proof.* We recall from Appendix A that

$$\mathcal{E}_S(h_{A,B}, h_{A,B}) = \text{cap}(A, \Omega \setminus B) \leq \mathcal{E}_S(u, u)$$

for any  $u$  with  $u(\partial A) = 1, u(\partial B) = 0$ . Thus, to prove an upper bound it suffices to find any such function for which we can calculate  $\mathcal{E}(u, u)$ . To this end, consider

$$u_c(x) \triangleq \begin{cases} (1-c)h_{A,\tilde{A}^c}(x) + c & \text{if } x \in \tilde{A} \\ c[1 - h_{B,\tilde{B}^c}(x)] & \text{if } x \in \tilde{B} \\ c & \text{otherwise.} \end{cases}$$

These functions are well suited to giving us upper bounds on  $\mathcal{E}_S(h_{A,B}, h_{A,B})$ . Indeed:

(1)  $u_c(\partial A) = 1, u_c(\partial B) = 0$ . In fact,  $u_c$  takes a constant value  $c$  outside of  $\tilde{A}, \tilde{B}$ , drops smoothly in  $\tilde{B}$  to achieve 0 on  $\partial B$ , and rises smoothly in  $\tilde{A}$  to achieve 1 on  $\partial \tilde{A}$ .

(2) Noting that  $u_c$  is written as a piecewise combination of hitting probability functions, we see that its Dirichlet form can be calculated in terms of capacities on local regions:  $\mathcal{E}_S(u_c, u_c) = (1-c)^2 \text{cap}(A, \tilde{A}) + c^2 \text{cap}(B, \tilde{B})$ .

Thus the  $u_c$  functions give us a practical way to calculate upper bounds:

$$\mathcal{E}_S(h_{A,B}, h_{A,B}) \leq (1-c)^2 \text{cap}(A, \tilde{A}) + c^2 \text{cap}(B, \tilde{B}).$$

This inequality holds for any value of  $c$ . To get the best bound, we can take derivatives to minimize the right hand side with respect to  $c$ . The result is

$$c^* = \frac{\text{cap}(A, \tilde{A})}{\text{cap}(A, \tilde{A}) + \text{cap}(B, \tilde{B})}.$$

Plugging this into the previous equation, we obtain our final result. ■

*Lemma 6.* Let  $S = \Omega \setminus (A \cup B)$ . Let  $m = \frac{1}{2} \{ \sup_{x \notin \tilde{A}, \tilde{B}} h_{A,B}(x) + \inf_{x \notin \tilde{A}, \tilde{B}} [h_{A,B}(x)] \}$ . The Dirichlet form of  $h_{A,B}$  can be lower bounded in terms of  $m$  and the capacities:

$$\mathcal{E}_S(h_{A,B}, h_{A,B}) \geq \left( 1 - m - \frac{\varepsilon}{2} \right)^2 \text{cap}(A, \tilde{A}) \mathbb{I}_{m \leq 1 - \frac{\varepsilon}{2}} + \left( m - \frac{\varepsilon}{2} \right)^2 \text{cap}(B, \tilde{B}) \mathbb{I}_{m \geq \frac{\varepsilon}{2}}.$$

*Proof.* Recall that  $\mathcal{E}_S(h_{A,B}, h_{A,B})$  can be expressed as an integral over  $S$ . We decompose this into three integrals: one over  $\tilde{A}$ , one over  $\Omega \setminus \tilde{A}, \tilde{B}$ , and one over  $\tilde{B}$ :

$$\mathcal{E}_S(h_{A,B}, h_{A,B}) = \int_{\tilde{A} \setminus A} \|\sigma \nabla h_{A,B}\|^2 \rho(dx) + \int_{\tilde{B} \setminus B} \|\sigma \nabla h_{A,B}\|^2 \rho(dx) + \int_{\Omega \setminus \tilde{A}, \tilde{B}} \|\sigma \nabla h_{A,B}\|^2 \rho(dx).$$

Since the integrand is always positive, we can get a lower bound by simply ignoring the integral over  $\Omega \setminus \tilde{A}, \tilde{B}$  and focusing on the integrals over  $\tilde{A}, \tilde{B}$ . The  $\tilde{A}, \tilde{B}$  integrals can be lower bounded using capacities.

For example, let us focus on the  $A$  integral. There are two different possibilities we must consider:

- (1) If  $m > 1 - \varepsilon/2$  we will simply note that the integral over the  $\tilde{A}$  region is non-negative.
- (2) If  $m \leq 1 - \varepsilon/2$ , then we define

$$u_A(x) \triangleq \frac{h_{A,B}(x) - m - \frac{\varepsilon}{2}}{1 - m - \frac{\varepsilon}{2}}.$$

Note that  $h_{A,B}(x) = 1$  for  $x \in \partial A$  and the  $\varepsilon$ -flatness condition shows that  $h_{A,B}(x) \leq m + \frac{\varepsilon}{2}$  for  $x \in \partial \tilde{A}$ . Thus  $u_A(\partial A) \geq 1, u_A(\partial \tilde{A}) \leq 0$ . Lemma 8 from Appendix E may then be applied to yield that  $\mathcal{E}_{\tilde{A} \setminus A}(u_A, u_A) \geq \text{cap}(A, \tilde{A})$ . We can thus obtain the bound

$$\begin{aligned} \int_{\tilde{A} \setminus A} \|\sigma \nabla h_{A,B}\|^2 \rho(dx) &= \left(1 - m - \frac{\varepsilon}{2}\right)^2 \mathcal{E}_{\tilde{A} \setminus A}(u_A, u_A) \\ &\geq \left(1 - m - \frac{\varepsilon}{2}\right)^2 \text{cap}(A, \tilde{A}). \end{aligned}$$

Putting these two possibilities together, we obtain

$$\int_{\tilde{A} \setminus A} \|\sigma \nabla h_{A,B}\|^2 \rho(dx) \geq \left(1 - m - \frac{\varepsilon}{2}\right)^2 \text{cap}(A, \tilde{A}) \mathbb{I}_{m \leq 1 - \frac{\varepsilon}{2}}.$$

Applying the same ideas to the integral over  $\tilde{B}$ , we obtain our result. ■

We are now in a position to prove Theorem 1:

*Theorem 1.* Assume that  $h_{A,B}(x)$  is  $\varepsilon$ -flat relative to  $\Omega \setminus (\tilde{A} \cup \tilde{B})$ . Then the first-passage probabilities can be well approximated in terms of the target capacities:

$$\sup_{x \notin \tilde{A}, \tilde{B}} \left| h_{A,B}(x) - \frac{\text{cap}(A, \tilde{A})}{\text{cap}(A, \tilde{A}) + \text{cap}(B, \tilde{B})} \right| \leq \varepsilon + \sqrt{\varepsilon/2}.$$

*Proof.* To simplify notation, let  $\kappa_A = \text{cap}(A, \tilde{A})$  and  $\kappa_B = \text{cap}(B, \tilde{B})$ . Applying the previous two lemmas together, we obtain the inequality

$$\frac{\kappa_A \kappa_B}{\kappa_A + \kappa_B} \geq \mathcal{E}(h_{A,B}, h_{A,B}) \geq \left(1 - m - \frac{\varepsilon}{2}\right)^2 \kappa_A \mathbb{I}_{m \leq 1 - \frac{\varepsilon}{2}} + \left(m - \frac{\varepsilon}{2}\right)^2 \kappa_B \mathbb{I}_{m \geq \frac{\varepsilon}{2}},$$

where  $m = \frac{1}{2}(\sup_{x \notin \tilde{A}, \tilde{B}} h_{A,B}(x) + \inf_{x \notin \tilde{A}, \tilde{B}} h_{A,B}(x))$ . In analyzing this inequality, there are three possibilities to consider.

If  $m \in (\varepsilon/2, 1 - \varepsilon/2)$ , the quadratic formula yields

$$\begin{aligned} m &\geq \frac{\kappa_A}{\kappa_A + \kappa_B} + \frac{\frac{\varepsilon}{2}(\kappa_B - \kappa_A) - \sqrt{\kappa_A \kappa_B \varepsilon(2 - \varepsilon)}}{\kappa_A + \kappa_B}, \\ m &\leq \frac{\kappa_A}{\kappa_A + \kappa_B} + \frac{\frac{\varepsilon}{2}(\kappa_B - \kappa_A) + \sqrt{\kappa_A \kappa_B \varepsilon(2 - \varepsilon)}}{\kappa_A + \kappa_B}. \end{aligned}$$

Applying  $\left| \frac{\kappa_B - \kappa_A}{\kappa_A + \kappa_B} \right| \leq 1$  and the fact that the geometric mean  $\sqrt{\kappa_A \kappa_B}$  never exceeds the arithmetic mean  $(\kappa_A + \kappa_B)/2$ , it follows that

$$\left| m - \frac{\kappa_A}{\kappa_A + \kappa_B} \right| \leq \frac{\varepsilon + \sqrt{\varepsilon(2 - \varepsilon)}}{2}.$$

Applying the fact that  $m$  was designed so that  $|h_{A,B}(x) - m| < \varepsilon/2$  for all  $x \notin \tilde{A}, \tilde{B}$ , we obtain

$$\left| h_{A,B}(x) - \frac{\kappa_A}{\kappa_A + \kappa_B} \right| \leq \frac{2\varepsilon + \sqrt{\varepsilon(2 - \varepsilon)}}{2}.$$

If  $m < \varepsilon/2$ , our equations become

$$\frac{\kappa_A \kappa_B}{\kappa_A + \kappa_B} \geq \left(1 - m - \frac{\varepsilon}{2}\right)^2 \kappa_A.$$

Our assumption that  $m \leq \varepsilon/2$  indicates that  $(1 - m - \varepsilon/2)^2 \geq (1 - \varepsilon)^2$ , and thus in fact we have

$$\frac{\kappa_B}{\kappa_A + \kappa_B} \geq (1 - \varepsilon)^2 = 1 + \varepsilon^2 - 2\varepsilon,$$

which means that  $\kappa_A/(\kappa_A + \kappa_B) \leq 2\varepsilon - \varepsilon^2 \leq 2\varepsilon$ . Thus we assumed  $m \in [0, \varepsilon/2]$  and showed that  $\kappa_A/(\kappa_A + \kappa_B) \in [0, 2\varepsilon - \varepsilon^2]$ , so it follows that

$$\left| m - \frac{\kappa_A}{\kappa_A + \kappa_B} \right| \leq 2\varepsilon - \varepsilon^2,$$

and so for any  $x \notin \tilde{A}, \tilde{B}$ , we have

$$\left| h_{A,B}(x) - \frac{\kappa_A}{\kappa_A + \kappa_B} \right| \leq 2.5\varepsilon - \varepsilon^2.$$

If  $m > 1 - \varepsilon/2$ , the same bound can be achieved by arguments which are symmetric to those employed in  $m < \varepsilon/2$ :

$$\left| h_{A,B}(x) - \frac{\kappa_A}{\kappa_A + \kappa_B} \right| \leq 2.5\varepsilon - \varepsilon^2.$$

Our final result is found by noting that all these bounds are upper bounded by  $\varepsilon + \sqrt{\varepsilon/2}$ . ■

#### APPENDIX D: PROOF OF PROPOSITION 1

We first establish a lemma, using Green’s first identity and some properties of the stationary SDE (1), under the reversibility conditions (3), relative to  $U$ :

*Lemma 7.* Fix some  $S \subset \Omega$  with smooth boundary. Then for any smooth function  $g$  that satisfies  $\mathcal{L}g = 0$  and smooth function  $f$

$$\int_S \nabla f(x)^T a(x) \nabla g(x) e^{-U(x)} dx = \int_{\partial S} f(x) \mathbf{n}(x)^T a(x) \nabla g(x) e^{-U(x)} \mathcal{H}(dx),$$

where  $\mathbf{n}$  are the normal vectors facing out of the set  $S$  and  $\mathcal{H}(dx)$  is the integral with respect to the  $(n - 1)$ -dimensional Hausdorff measure and

$$(\mathcal{L}f)(x) \triangleq \sum_i b_i(x) \frac{\partial f(x)}{\partial x_i} + \frac{1}{2} \sum_{ij} a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j}.$$

*Proof.* This result is essentially a direct corollary of Green’s identities. Apply Green’s first identity to get

$$\int_S \nabla f^T a \nabla g e^{-U} dx = \int_{\partial S} f \mathbf{n}^T a \nabla g e^{-U} \mathcal{H}(dx) - \int_S f \nabla \cdot (a \nabla g e^{-U}) dx,$$

where

$$\nabla \cdot (a \nabla g e^{-U}) = \sum_i \frac{\partial}{\partial x_i} \left[ \sum_j e^{-U} a_{ij} \frac{\partial g}{\partial x_j} \right].$$

Next, using the reversibility constraint on  $b$  from Equation (3), it’s not hard to verify that

$$\nabla \cdot (a \nabla g e^{-U}) = 2e^{-U} \mathcal{L}g = 0.$$

This gives us the desired result. ■

*Proposition 1.* For any regions  $G$  and  $\tilde{G}$  having smooth boundaries and such that  $A \subset G \subset \tilde{G} \subset \tilde{A}$ ,  $\text{cap}(A, \tilde{A})$  can be expressed as a flux leaving  $\tilde{G} \setminus G$ :

$$\text{cap}(A, \tilde{A}) = \int_{\partial(\tilde{G} \setminus G)} h_{A, \tilde{A}^c}(x) \mathbf{n}(x)^T a(x) \nabla h_{G, \tilde{G}^c}(x) e^{-U(x)} \mathcal{H}(dx),$$

where  $a(x) = \sigma(x)\sigma(x)^T$  is the diffusion matrix,  $\mathcal{H}(dx)$  is the  $(n - 1)$ -dimensional Hausdorff measure, and  $\mathbf{n}$  represents the outward-facing (relative to  $\tilde{G} \setminus G$ ) normal vector on  $\partial(\tilde{G} \setminus G)$ .

*Proof.* This result is certainly known, but we include a proof here because we were unable to find a proof in the literature. First, recall that  $\mathcal{L}h_{A, \tilde{A}^c} = 0$  and

$$\text{cap}(A, \tilde{A}) = \int_{\tilde{A} \setminus A} \nabla h_{A, \tilde{A}^c}(x)^T a(x) \nabla h_{A, \tilde{A}^c}(x) e^{-U(x)} dx.$$



Together with Lemma 7, this yields that

$$\text{cap}(A, \tilde{A}) = \int_{\partial(\tilde{A} \setminus A)} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx) = - \int_{\partial A} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx), \tag{D1}$$

where in the last step we used  $h_{A, \tilde{A}^c}(x) = 1, x \in \partial A, h_{A, \tilde{A}^c}(x) = 0, x \in \partial \tilde{A}$ . Also, note that the normal vector on the right-hand side is pointing *out* of the set  $A$ , as is our convention. Hence then negative sign.

Next we apply Lemma 7 again to get

$$0 = \int_{G \setminus A} \nabla(\mathcal{L}) a \nabla h_{A, \tilde{A}^c}^T e^{-U} dx = \int_{\partial(G \setminus A)} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx).$$

Combining this with Eq. (D1) gives us

$$\text{cap}(A, \tilde{A}) = - \int_{\partial A} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx) = \int_{\partial G} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx). \tag{D2}$$

Using the facts that  $h_{G, \tilde{G}^c}(x) = 1, x \in \partial G, h_{G, \tilde{G}^c}(x) = 0, x \in \partial \tilde{G}$ , and  $\mathcal{L}h_{G, \tilde{G}^c} = 0$ , we apply Lemma 7 two more times to obtain

$$\begin{aligned} \text{cap}(A, \tilde{A}) &= \int_{\partial G} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx) = \int_{\partial G} h_{G, \tilde{G}^c} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx) \\ &= \int_{\partial(\tilde{G} \setminus G)} h_{G, \tilde{G}^c} \mathbf{n}^T a \nabla h_{A, \tilde{A}^c} e^{-U} \mathcal{H}(dx) = \int_{\tilde{G} \setminus G} \nabla h_{G, \tilde{G}^c} a \nabla h_{A, \tilde{A}^c} e^{-U} dx \\ &= \int_{\partial(\tilde{G} \setminus G)} h_{A, \tilde{A}^c} \mathbf{n}^T a \nabla h_{G, \tilde{G}^c} e^{-U} \mathcal{H}(dx). \end{aligned}$$

*Corollary.* For any region  $S$  having smooth boundary  $\partial S$ , and such that  $A \subset S \subset \tilde{A}$ ,  $\text{cap}(A, \tilde{A})$  can be expressed as a flux leaving  $S$ :

$$\text{cap}(A, \tilde{A}) = \int_{\partial S} \mathbf{n}(x)^T a(x) \nabla h_{A, \tilde{A}^c}(x) e^{-U(x)} \mathcal{H}(dx),$$

where  $a$  and  $\mathcal{H}(dx)$  are as defined in the proposition, and  $\mathbf{n}$  is the outward-facing normal on  $\partial S$ .

*Proof.* Put  $G = S$  in Eq. (D2).

### APPENDIX E: INEQUALITY BOUNDARY CONDITIONS FOR THE VARIATIONAL FORM

Recall that  $h_{A,B}(x)$  may be defined variationally. We have let

$$\mathcal{E}(f, g) \triangleq \int_{\Omega} \nabla f(x)^T a(x) \nabla g(x) \rho(dx)$$

denote the ‘‘Dirichlet form.’’ Let  $\Omega \subset \mathbb{R}^n$  compact and open with smooth boundary. Let  $\mathcal{L}^2(\bar{\Omega}, \rho)$  denote the Hilbert space of functions on  $\bar{\Omega}$  which are square-integrable with respect to a continuous positive measure  $\rho(dx) = e^{-U} dx$ . Let  $\mathcal{H}^1(\bar{\Omega}, \rho) = W^{1,2}(\bar{\Omega}, \rho) \subset \mathcal{L}^2(\bar{\Omega}, \rho)$  denote the corresponding once weakly differentiable Hilbert-Sobolev space. Let  $A, B \subset \Omega$ , open, disjoint, with smooth boundary, and define  $\text{cap}(A, \Omega \setminus B) \in \mathbb{R}$  as the minimizing value of the problem

$$\begin{aligned} &\min_{u \in \mathcal{H}^1} \mathcal{E}(u, u) \\ &\text{subject to } u(x) = 1, x \in A \\ &\quad u(x) = 0, x \in B. \end{aligned}$$

It is natural to consider an apparently different problem, where the equality boundary conditions are replaced with inequalities. Here we show that it is not possible to get lower than  $\text{cap}(A, \Omega \setminus B)$  by such a relaxation.

*Lemma 8.* Let  $\tilde{h}$  satisfy  $\tilde{h}(x) \geq 1$  on  $A$  and  $\tilde{h}(x) \leq 0$  on  $B$ . Then  $\mathcal{E}(\tilde{h}, \tilde{h}) \geq \text{cap}(A, \Omega \setminus B)$ .

*Proof.* This result is certainly known, but we include a proof here because we were unable to find a proof in the literature. Let  $k = \text{clamp}(\tilde{h}, 0, 1)$ , i.e.,

$$k(x) = \begin{cases} \tilde{h}(x) & \tilde{h}(x) \in [0, 1] \\ 0 & \tilde{h}(x) \leq 0 \\ 1 & \tilde{h}(x) \geq 1 \end{cases}.$$

Note that  $k \in \mathcal{H}^1$  and satisfies the equality boundary conditions. Thus, by definition,  $\text{cap}(A, \Omega \setminus B) \leq \mathcal{E}(k, k)$ . This immediately yields our result:

$$\text{cap}(A, \Omega \setminus B) \leq \mathcal{E}(k, k) = \int \|\sigma \nabla k\|^2 \rho(dx) = \int_{x: \tilde{h}(x) \in [0,1]} \|\sigma \nabla \tilde{h}\|^2 \rho(dx) \leq \int \|\sigma \nabla \tilde{h}\|^2 \rho(dx) = \mathcal{E}(\tilde{h}, \tilde{h}).$$

■

**APPENDIX F: SHELL METHOD**

The algorithm for estimating local hitting probabilities is outlined as follows:

*Algorithm 1.* Estimating  $h_{R, \tilde{R}^c}(x)$  for many values of  $x$  on a shell  $\partial S$

*Input.*  $R \subset S \subset \tilde{R} \subset \Omega$  and a stationary reversible diffusion process  $\{X_t\}_{t \geq 0}$  in  $\Omega$  with invariant measure  $\mu = e^{-U(x)} dx$ . We also require a series of subsets

$$R = S_0 \subset S_1 \subset \dots \subset S_{m-1} \subset S_m = S \subset S_{m+1} \subset \dots \subset S_n = \tilde{R},$$

which indicate a kind of reaction coordinate.

*Output.* A collection of points  $z_1, \dots, z_{N_p}$  on  $\partial S$  sampled from the invariant measure  $\mu = e^{-U(x)} dx$  restricted on  $\partial S$ , along with estimates of  $h_{R, \tilde{R}^c}(z_i)$  for each point.

(1) Discretize the space.

(a) Generate an ensemble of samples  $z_1, \dots, z_{N_p}$  on  $\partial S$  according to the invariant measure  $\mu = e^{-U(x)} dx$  restricted to  $\partial S$ .

(b) Evolve the ensemble on  $\partial S$ , by repeatedly sampling an initial location from the uniform distribution on  $\{z_1, \dots, z_{N_p}\}$  and carry out a local simulation following the dynamics of  $\{X_t\}_{t \geq 0}$  until the trajectory hits either  $\partial S_{m-1}$  or  $\partial S_{m+1}$ . Record the hitting locations on  $\partial S_{m-1}$  and  $\partial S_{m+1}$  until we have  $N_p$  points on both  $\partial S_{m-1}$  and  $\partial S_{m+1}$ . In most cases, the process is more likely to hit one of  $S_{m-1}, S_{m+1}$  than the other, and we need to run more than  $2N_p$  local simulations to get at least  $N_p$  samples on both shells. We store the results of the redundant local simulations for future estimation of transition probabilities.

(c) Repeat the above process to sequentially evolve the ensembles on  $\partial S_{m-1}, \dots, \partial S_2$  and on  $\partial S_{m+1}, \dots, \partial S_{n-2}$ , to get  $N_p$  samples on all of the intermediate shells  $\partial S_1, \dots, \partial S_{n-1}$ . Store the results of the redundant local simulations for future estimation of transition probabilities.

(d) For each one of the shells  $\partial S_1, \dots, \partial S_{n-1}$ , cluster the  $N_p$  samples on that shell into  $N_b$  states. In our implementation, we use  $k$ -means, and represent the  $N_b$  states by the  $N_b$  centroids we get from the algorithm.

The result of this step is a partitioning of each shell  $\partial S_i$  into  $N_b$  regions, representing an adaptive discretization of the shells. For a point on a shell  $\partial S_i$ , we identify its corresponding discrete state by finding the nearest centroid.

(2) Estimate the transition probabilities between these discrete states by running an additional  $N_s$  local simulations for each one of the  $N_b$  states on each shell. The result of this step is an estimate of the probability of transitioning from state  $k$  on  $\partial S_i$  to state  $l$  on  $\partial S_j$ , which we denote by  $P_{k,l}^{(i,j)}$ , where  $k, l \in \{1, \dots, N_b\}$  and  $i, j \in \{1, \dots, n-1\}$  with  $|i-j|=1$ .

(3) Use the transition probabilities to get an estimate of the hitting probabilities for the  $N_b$  states on  $\partial S$ . In line with related works on Markov state models [26–28], we approximate the continuous dynamics using closed-form calculations from the discrete Markov chain we have developed in the previous two steps. In particular, we estimate overall hitting probabilities using the standard “one-step analysis.” For any  $k \in \{1, \dots, N_b\}$  and  $i \in \{1, \dots, n-1\}$ , let  $u_k^{(i)}$  denote the probability of hitting  $\partial R = \partial S_0$  before hitting  $\partial \tilde{R} = \partial S_n$  if we start the discretized process at state  $k$  on  $\partial S_i$ . We can calculate our object of interest by solving the matrix difference equation

$$u^{(i)} = P^{(i,i+1)} u^{(i+1)} + P^{(i,i-1)} u^{(i-1)}, \quad i = 1, \dots, n-1$$

with boundary conditions  $u^{(0)} = 1, u^{(n)} = 0$ , where 0 and 1 are vectors of all 0’s and 1’s. This gives the estimated hitting probability for each discrete state. We then estimate the hitting probability of each point  $z_i$  by

$$h_{R, \tilde{R}^c}(z_i) = u_k^{(m)}, \quad z_i \in \text{state } k \text{ on } \partial S. \tag{F1}$$

**APPENDIX G: DETAILS ON ENERGY FUNCTION**

For the energy function, we hand designed two different kinds of landscape: random well energy, which we use for the region around target  $A$ , and random crater energy, which we use for the region around target  $B$ . The basic components of these energy functions are the well component, given by

$$F_w(x|d_w, r) = -\frac{d_w}{r^4} (||x - x_A||_2^4 - 2r^2 ||x - x_A||_2^2) - d_w, \tag{G1}$$

where  $d_w$  gives the depth of the well; the crater component, given by

$$F_c(x|d_c, h, r) = \frac{d_c}{3b^2 r^4 - r^6} (2||x - x_B||_2^2 - 3(b^2 + r^2) ||x - x_B||_2^4 + 6b^2 r^2 ||x - x_B||_2^2) - d_c, \tag{G2}$$

where  $d_c$  and  $h$  give the depth and the height of the crater, respectively, and

$$b^2 = -\frac{1}{3d} \left( -3d_c r^2 + C + \frac{\Delta_0}{C} \right) \quad (\text{G3})$$

with

$$C = 3r^2 \sqrt[3]{d_c h [d_c + \sqrt{d_c(d_c + h)}]}, \quad \Delta_0 = -9d_c h r^4;$$

and finally a random component, given by

$$F_r(x|\mu, \sigma) = \sum_{i=1}^m \prod_{j=1}^d \exp \left[ -\frac{(x_j - \mu_{ij})^2}{2\sigma_{ij}^2} \right], \quad (\text{G4})$$

where  $\mu = (\mu_{ij})_{m \times d}$  and  $\sigma = (\sigma_{ij})_{m \times d}$ , with  $\mu_i = (\mu_{i1}, \dots, \mu_{id})$ ,  $i = 1, \dots, m$  being the locations of  $m$  Gaussian random bumps in the region around the targets, and  $\sigma_{ij}$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, d$  gives the corresponding standard deviations.

To make sure the energy function is continuous, and the different components of the energy function are balanced, we introduce a mollifier, given by

$$F_m(x|x_0, r) = \exp \left( -\frac{r}{r - \|x - x_0\|_2^{20}} \right), \quad (\text{G5})$$

where  $x_0 = x_A$ ,  $r = r_A$  or  $x_0 = x_B$ ,  $r = r_B$ , depending on which target we are working with, and a rescaling of the random component, which is given by  $0.1 \times d_w$  if we are perturbing the well component, and  $0.1(d_c + h)$  if we are perturbing the crater component.

Intuitively, for the well component, we use a fourth-order polynomial to get a well-like energy landscape around the target that is continuous and differentiable at the boundary. Similarly, for the crater component, we use a sixth-order polynomial to get a crater-like energy landscape around the target that is also continuous and differentiable at the boundary. For the random component, we are essentially placing a number of Gaussian bumps around the target. And for the mollifier, we are designing the function such that it's almost exactly 1 around the target, until it comes to the outer boundary, when it transitions smoothly and swiftly to 0. To summarize, given parameters  $d_w, d_c, h$  and random bumps  $\mu_A, \mu_B$  with  $\mu_i^A \in \dot{A} \setminus A$ ,  $i = 1, \dots, m_A$ ,  $\mu_i^B \in \dot{B} \setminus B$ ,  $i = 1, \dots, m_B$ , and the corresponding standard deviations  $\sigma^A, \sigma^B$  with  $\sigma_{ij}^A$ ,  $i = 1, \dots, m_A$ ,  $j = 1, \dots, d$ ,  $\sigma_{ij}^B$ ,  $i = 1, \dots, m_B$ ,  $j = 1, \dots, d$ , the energy function we used in the experiments is given by

$$F(x) = F_w(x|d_w, r_A) + 0.1 \times d_w \times F_m(x|x_A, r_A) + F_r(x|\mu^A, \sigma^A), \quad \forall x \in \dot{A} \setminus A, \quad (\text{G6})$$

$$F(x) = F_c(x|d_c, h, r_B) + 0.1 \times (d_c + h) \times F_m(x|x_B, r_B) + F_r(x|\mu^B, \sigma^B), \quad \forall x \in \dot{B} \setminus B. \quad (\text{G7})$$

In our actual experiments, we used

$$d_w = 10.0, \quad d_c = 6.0, \quad h = 1.0, \quad \sigma_{ij}^A = \sigma_{kl}^B = 0.01, \quad \forall i, j, k, l$$

and

$$\mu^A = \begin{pmatrix} 0.512 & 0.583 & -0.013 & 0.013 & -0.001 \\ 0.464 & 0.575 & -0.001 & 0.019 & -0.014 \\ 0.503 & 0.611 & -0.012 & -0.024 & 0.023 \\ 0.5 & 0.601 & -0.024 & 0.034 & 0.011 \\ 0.486 & 0.586 & 0.006 & 0.01 & 0.001 \\ 0.489 & 0.588 & -0.017 & 0.002 & 0.027 \\ 0.493 & 0.585 & 0.015 & -0.001 & -0.032 \\ 0.516 & 0.596 & 0.027 & -0.026 & 0.022 \\ 0.514 & 0.624 & 0.01 & 0.01 & -0.002 \\ 0.5 & 0.605 & 0.017 & -0.016 & 0.004 \end{pmatrix}, \quad \mu^B = \begin{pmatrix} -0.696 & -0.006 & 0.023 & -0.041 & 0.019 \\ -0.731 & 0.021 & -0.033 & -0.014 & 0.017 \\ -0.694 & -0.034 & -0.009 & 0.031 & 0.019 \\ -0.666 & -0.013 & 0.002 & 0.017 & 0.009 \\ -0.68 & 0.058 & 0.007 & -0.011 & -0.008 \\ -0.704 & -0.022 & 0.034 & 0.003 & 0.026 \\ -0.714 & -0.015 & 0.017 & 0.027 & 0.028 \\ -0.681 & 0.017 & -0.046 & -0.04 & -0.002 \\ -0.648 & -0.009 & 0.002 & -0.012 & -0.022 \\ -0.664 & -0.04 & 0.05 & -0.012 & -0.002 \end{pmatrix}.$$

[1] H. A. Scheraga, M. Khalili, and A. Liwo, *Annu. Rev. Phys. Chem.* **58**, 57 (2007).

[2] A. Hospital, J. R. Gofii, M. Orozco, and J. L. Gelpi, *Adv. Appl. Bioinform. Chem.* **8**, 37 (2015).

[3] H. Lei, B. Caswell, and G. E. Karniadakis, *Phys. Rev. E* **81**, 026704 (2010).

[4] V. Sotiropoulos and Y. N. Kaznessis, *Chem. Eng. Sci.* **66**, 268 (2011).

- [5] D. T. Gillespie, *J. Chem. Phys.* **113**, 297 (2000).
- [6] D. Holcman and Z. Schuss, *J. Stat. Phys.* **117**, 975 (2004).
- [7] W. E. and E. Vanden-Eijnden, *J. Stat. Phys.* **123**, 503 (2006).
- [8] Z. Schuss, A. Singer, and D. Holcman, *Proc. Natl. Acad. Sci. USA* **104**, 16098 (2007).
- [9] T. C. B. McLeish, *Biophys. J.* **88**, 172 (2005).
- [10] J. M. Goldberg and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **96**, 2019 (1999).
- [11] M. Jacob, M. Geeves, G. Holtermann, and F. X. Schmid, *Nat. Struct. Biol.* **6**, 923 (1999).
- [12] K. W. Plaxco and D. Baker, *Proc. Natl. Acad. Sci. USA* **95**, 13591 (1998).
- [13] D. J. Wales, *Int. Rev. Phys. Chem.* **25**, 237 (2006).
- [14] E. B. Baum, *Phys. Rev. Lett.* **57**, 2764 (1986).
- [15] L. T. Wille, *Phys. Rev. Lett.* **59**, 372 (1987).
- [16] J. Machta, *Phys. Rev. E* **80**, 056706 (2009).
- [17] M. Christen and W. F. van Gunsteren, *J. Comput. Chem. Computer Simul.* **29**, 157 (2008).
- [18] A. F. Cheviakov, M. J. Ward, and R. Straube, *Multiscale Model. Simul.* **8**, 836 (2010).
- [19] X. Li, *J. Phys. A: Math. Theor.* **47**, 505202 (2014).
- [20] C. Chevalier, O. Bénichou, B. Meyer, and R. Voituriez, *J. Phys. A: Math. Theor.* **44**, 025002 (2010).
- [21] S. Condamin and O. Bénichou, *J. Chem. Phys.* **124**, 206103 (2006).
- [22] D. Coombs, R. Straube, and M. Ward, *SIAM J. Appl. Math.* **70**, 302 (2009).
- [23] A. E. Lindsay, A. J. Bernoff, and M. J. Ward, *Multiscale Model. Simul.* **15**, 74 (2017).
- [24] J. Kaye and L. Greengard, *J. Comput. Phys.: X* **5**, 100047 (2020).
- [25] O. Bénichou and R. Voituriez, *Phys. Rep.* **539**, 225 (2014).
- [26] V. S. Pande, K. Beauchamp, and G. R. Bowman, *Methods* **52**, 99 (2010).
- [27] J. D. Chodera and F. Noé, *Curr. Opin. Struct. Biol.* **25**, 135 (2014).
- [28] B. E. Husic and V. S. Pande, *J. Am. Chem. Soc.* **140**, 2386 (2018).
- [29] G. J. Moro, *J. Chem. Phys.* **103**, 7514 (1995).
- [30] J. M. Bello-Rivas and R. Elber, *J. Chem. Phys.* **142**, 094102 (2015).
- [31] P.-L. Lions and A.-S. Sznitman, *Commun. Pure Appl. Math.* **37**, 511 (1984).
- [32] Z.-Q. Chen, *Probab. Theory Relat. Fields* **94**, 281 (1993).
- [33] A. Bovier and F. den Hollander, *Metastability: A Potential-Theoretic Approach*, Grundlehren der mathematischen Wissenschaften Vol. 351 (Springer International Publishing Cham, Switzerland, 2016).
- [34] A. Bovier, M. Eckhoff, V. Gaynard, and M. Klein, *J. Eur. Math. Soc.* **6**, 399 (2004).
- [35] A. M. A. West, R. Elber, and D. Shalloway, *J. Chem. Phys.* **126**, 145104 (2007).
- [36] D. Aristoff, J. Bello-Rivas, and R. Elber, *Multiscale Model. Simul.* **14**, 301 (2016).
- [37] All the experimental results can be reproduced or easily modified from open-source code, which can found, along with detailed instructions, at [https://github.com/StannisZhou/entropic\\_barrier](https://github.com/StannisZhou/entropic_barrier).
- [38] N. H. Bingham, *Z. Wahrscheinlichkeitstheorie verwandte Gebiete* **22**, 169 (1972).
- [39] G. M. Lieberman, *J. Math. Anal. Appl.* **113**, 422 (1986).
- [40] Z.-Q. Chen and M. Fukushima, *Symmetric Markov Processes, Time Change, and Boundary Theory*, London Mathematical Society Monographs Vol. 35 (Princeton University Press, Princeton, 2012).
- [41] H. Dret and B. Lucquin, *Partial Differential Equations: Modeling, Analysis and Numerical Approximation*, International Series of Numerical Mathematics Vol. 168 (Birkhäuser, Cham, Switzerland, 2016).
- [42] J. Loper, *J. Theor. Probab.* **33**, 22 (2018).
- [43] J. G. Wendel, *Ann. Probab.* **8**, 164 (1980).
- [44] NIST digital library of mathematical functions, <http://dlmf.nist.gov/>.
- [45] Á. Baricz, *Proc. Edinburgh Math. Soc.* **53**, 575 (2010).
- [46] G. Strang, *Numer. Math.* **19**, 81 (1972).