Accepted: 2 February 2022

DOI: 10.1002/cpa.22177

RESEARCH ARTICLE

Sharp asymptotic estimates for expectations, probabilities, and mean first passage times in stochastic systems with small noise

Tobias Grafke[1] | **Tobias Schäfer**[2,3] | **Eric Vanden-Eijnden**[4]

[1]Mathematics Institute, University of Warwick, Coventry, UK

[2]Department of Mathematics, College of Staten Island, New York, USA

[3]Physics Program, CUNY Graduate Center, New York, USA

[4]Courant Institute, New York University, New York, USA

Correspondence
Eric Vanden-Eijnden, Courant Institute, New York University, 251 Mercer Street, New York, NY, USA.
Email: eve2@cims.nyu.edu

Funding information
EPSRC, Grant/Award Numbers: EP/T011866/1, EP/V013319/1; NSF, Grant/Award Numbers: DMS-1522737, DMS-2012548, DMS-1522767; Materials Research Science and Engineering Center Program, Grant/Award Number: DMR-1420073; Simons Collaboration on Wave Turbulence, Grant/Award Number: No. 617006

Abstract

Freidlin-Wentzell theory of large deviations can be used to compute the likelihood of extreme or rare events in stochastic dynamical systems via the solution of an optimization problem. The approach gives exponential estimates that often need to be refined via calculation of a prefactor. Here it is shown how to perform these computations in practice. Specifically, sharp asymptotic estimates are derived for expectations, probabilities, and mean first passage times in a form that is geared towards numerical purposes: they require solving well-posed matrix Riccati equations involving the minimizer of the Freidlin-Wentzell action as input, either forward or backward in time with appropriate initial or final conditions tailored to the estimate at hand. The usefulness of our approach is illustrated on several examples. In particular, invariant measure probabilities and mean first passage times are calculated in models involving stochastic partial differential equations of reaction-advection-diffusion type.

CONTENTS

1 | INTRODUCTION

Rare events in stochastic dynamical systems tend to cluster around their most likely realization. As a result they have predictable features that can be calculated via some optimization problem. This profound observation has been made in numerous fields, and used for example, to explain phase transitions in statistical mechanics [18], derive Arrhenius' law in chemical kinetics [1], or use semiclassical trajectories in quantum field theory [49]. Large deviation theory (LDT) [52] gives a mathematical justification to these results and provide us with an action, or rate function, to minimize in order to calculate paths of maximum likelihood, also known as *instantons*. The theory also gives exponential asymptotic estimates of rare event probabilities. While this information is already useful in many cases, more refined estimates are often desirable. These 'prefactor' calculations attempt to quantify the effect of Gaussian fluctuations around the instanton, a notion that has also been separately rediscovered in the literature through various means [5]. For example, in the context of chemical reaction rates, next order refinements of the exponential reaction rate are known as the Eyring-Kramers law [19, 38]. Similarly in quantum field theory, perturbing around the semiclassical trajectory, the second order variations leads to a Gaussian path-integral, which ultimately results in an additional contribution in the form of a ratio of functional determinants [49].

Over the last 2 decades, several computational methods have been developed to calculate instantons. Among others, we refer to the string method in the context of gradient flows [15, 17], the minimum action method [16, 22], the adaptive minimum action method (aMAM) [53] and the geometric minimum action method (gMAM) [34, 35, 50]. These methods are now efficient enough to be used in the context high-dimensional systems, including stochastically driven partial differential equations arising in fluid dynamics [28, 30].

In contrast, surprisingly little work has been done on the numerical side of prefactor calculations (see however [45]). The main objective of this paper is to show how to extend methods such as MAM or gMAM to efficiently estimate prefactors in the context of the calculations of expectations, probabilities, and mean first passage times.

1.1 | LDT and instantons

Consider a family of stochastic differential equations (SDEs) for $X_t^\epsilon \in \mathbb{R}^n$, with drift vector field $b : \mathbb{R}^n \to \mathbb{R}^n$ and diffusion matrix $a = \sigma\sigma^\top$, where $\sigma \in \mathbb{R}^{n\times n}$,

$$dX_t^\epsilon = b\left(X_t^\epsilon\right) dt + \sqrt{\epsilon}\sigma\, dW_t. \tag{1.1}$$

Here, W_t is an n-dimensional Wiener process, and we have introduced a small parameter $\epsilon > 0$ to characterize the strength of the noise. For simplicity we will assume that the diffusion matrix $a \in \mathbb{R}^{n\times n}$ is positive-definite (hence invertible) and constant (but not necessarily diagonal), that is the case of additive Gaussian noise—the generalization of the methods presented below to a covariance matrix a that depends on x, that is multiplicative Gaussian noise, is straightforward. Large deviations theory [24, 52] indicates that, in the limit as $\epsilon \to 0$, the solutions to (1.1) that contribute most to the probability of an event or the value of an expectation are likely to be close to the minimizer of the Freidlin-Wentzell rate function S_T subject to appropriate boundary conditions. This action functional is given by

$$S_T(\phi) = \int_0^T L(\phi, \dot\phi)\, dt \tag{1.2}$$

with the Lagrangian

$$L(\phi, \dot\phi) = \tfrac{1}{2}\langle \dot\phi - b(\phi), a^{-1}(\dot\phi - b(\phi))\rangle \equiv \tfrac{1}{2}|\dot\phi - b(\phi)|_a^2. \tag{1.3}$$

Here $\langle x, y\rangle$ stands for the Euclidean scalar product between the vectors x and y and we introduced the norm induced by a, $|x|_a^2 = \langle x, a^{-1}x\rangle \equiv \sum_{i,j=1}^n x_i a^{-1}_{i,j} x_j$. If the diffusion matrix a is the identity, this norm becomes simply the Euclidean length.

The minimizer of the action (1.2) is referred to as path of maximum likelihood or *instanton*, and it can be found by solving the corresponding Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot\phi} = \frac{\partial L}{\partial \phi} \tag{1.4}$$

with boundary conditions appropriate to the event under consideration. Thus, in the context of LDT, the leading order estimation of probabilities or expectations can be reduced to the solution of the deterministic system (1.4).

Alternatively, there is a Hamiltonian formulation to the problem. Taking the Legendre transform of the Lagrangian and introducing the momentum $\theta = \partial L/\partial\dot\phi$, we obtain the Hamiltonian

$$H(\phi, \theta) = \langle b(\phi), \theta\rangle + \tfrac{1}{2}\langle \theta, a\theta\rangle, \tag{1.5}$$

and the Euler-Lagrange equations for the instanton become

$$\dot\phi = b(\phi) + a\theta, \qquad \dot\theta = -(\nabla b(\phi))^\top \theta. \tag{1.6}$$

1.2 | Prefactor estimates

The instanton gives the leading contribution to the exponential decay of the probability for observing a rare event. In order to obtain sharp estimates, one needs to furthermore consider prefactor contributions.

Intuitively these prefactors can be calculated by accounting for the effects of the fluctuations around the instanton ϕ, which can be done by linearizing the solution of the SDE (1.1) around ϕ and considering

$$dZ_t = \nabla b(\phi(t))Z_t\,dt + \sigma\,dW_t. \tag{1.7}$$

The solution to this equation defines a Gaussian process, and the prefactor contribution to expectations, probabilities, or mean first passage times can be calculated as specific expectations over this process. In turns, these expectations are ratios between determinants of specific positive-definite matrices or operators that can be expressed in terms of the solutions of deterministic Riccati equations, as can be intuited by analogy with results from optimal control theory. Our objectives here are to: (i) formulate these Riccati equations, including their boundary conditions, in the specific cases of the calculation of expectations, probabilities, and mean first passage time; and (ii) develop efficient numerical methods for their solution.

1.3 | Related works

From a theoretical point of view, our approach builds on a large corpus of works dealing with expansion beyond the exponential estimate of LDT and the evaluation of quadratic path integrals or Wiener integrals, as initiated by Schilder [44]. Results from probability theory and stochastic analysis in this direction include for example the pioneering works by Kifer [37] and Azencott [2]. On the analytical side, formal asymptotic expansions were used for example in [40, 41], the WKB expansion in [21], optimal control theory in [3], and more recently potential theory in [5, 12]. This last approach allows one to establish rigorously the Eyring-Kramers law for reversible system [11], and has also been extended to other problems, such as lattice gas models [13]. Similarly, prefactor calculations recently included non-reversible systems [9]. From a mathematical perspective it is much harder to treat the infinite dimensional case, even though recent breakthroughs have been made at least for systems in detailed balance [4, 6, 7, 20].

From a computational viewpoint, none of the references above deal with explicit calculation of the instanton or the prefactors, and the equations they derive for these objects, while often very general, do not lend themselves automatically to numerical implementation. Explicit (especially numerical) computation of prefactors is usually confined to quantum field theory, where the evaluation of Gaussian path integrals is a classical result [23, 26, 39, 43], see also [14] for a recent review. However, the Lagrangian in quantum mechanics is usually assumed to take a special form with no first order time derivative in the Euler-Lagrange equation (corresponding to a stochastic process in detailed balance in the stochastic interpretation). The processes we are interested in are considerably more general from that perspective, and require the more general methods we develop here. These methods built on formulas derived by asymptotic analysis of backward Kolmogorov equations (BKEs), and are complementary to those recently proposed in [45] using a path integral approach.

1.4 | Main contributions

Our main results can be summarized as follows: (1) We provide formal but short and simple proofs of propositions establishing sharp estimates for expectations, probability densities, probabilities,

exit probabilities, and mean first passage times. While most of these results can be found in some form in the literature, they are scattered in many different papers, and we believe it is useful to collect and summarize them in one place. The methods used in these proofs can also be extended to more general situations not covered here. (2) We phrase each of our theoretical statements in a way geared towards numerical implementation, unlike what is usually found in the literature on the topic. (3) We use the geometric approach from gMAM to provide statements that are valid at infinite time (i.e. on the invariant measure of the process), by explicitly computing this infinite time limit via mapping of $t \in (-\infty, 0]$ onto the normalized arc-length $s \in (0, 1]$ along the instantons involved. (4) We formally generalize our results to the infinite-dimensional setup, with applications to stochastic partial equations of reaction-advection-diffusion type. (5) We also generalize our result to examples driven by non-Gaussian noise, specifically Markov jump processes (MJPs) in specific limits. (6) We illustrate the applicability of our method through tests cases in finite and infinite dimension, in which we discuss how to perform the numerical calculation involved.

In terms of limitations, our work focuses on situations where the stochastic system at hand has a single attracting point in the limit of vanishing noise. This setup is of interest in several situations, but it excludes the important problem of analyzing rare transitions between metastable states. Some of the tools we develop here, in particular the Riccati equations whose solutions enter the expressions for the prefactors, may be useful to analyze these transitions, but we will not consider them here.

1.5 | Assumptions and organization

As stated before, we are interested in obtaining sharp asymptotic estimates for expectations, probabilities, exit probabilities, and mean first passage times. We will do so under the generic assumption that the LDT optimization problem associated with each of these questions, that is the minimization of the action in (1.2) subject to appropriate constraints and/or boundary conditions, is strictly convex. This simplifying assumption guarantees that the solution of the equations presented below exists and is unique, and therefore allows us to avoid dealing with local minimizers, flat minima, etc. that may require to generalize/amend some of the statements below. While this is not necessary difficult to do, at least formally, it leads to a zoology of subcases that we want to avoid listing.

To make (1.1) have unique solutions, we will also make:

Assumption 1.1. The vector field b is $C^2(\mathbb{R}^n)$ and such that:

$$\exists \alpha, \beta > 0 \ : \ \langle b(x), x\rangle \leq \alpha - \beta|x|^2 \quad \forall x \in \mathbb{R}^n;$$

and the matrix a is such that:

$$\exists \gamma, \Gamma \text{ with } 0 < \gamma < \Gamma < \infty \ : \ \gamma|x|^2 \leq \langle x, ax\rangle < \Gamma|x|^2 \quad \forall x \in \mathbb{R}^n.$$

This assumption guarantees [42] that the solution to the SDE (1.1) exists for all times and is ergodic with respect to a unique invariant measure with a probability density function $\rho_\varepsilon : \mathbb{R}^n \to (0, \infty)$. This density is the unique solution to

$$0 = -\nabla \cdot (b\rho_\varepsilon) + \tfrac{1}{2}\varepsilon a : \nabla\nabla\rho_\varepsilon, \qquad \rho_\varepsilon \geq 0, \qquad \int_{\mathbb{R}^n} \rho_\varepsilon(x)dx = 1, \tag{1.8}$$

where here and below the colon denotes the trace, that is $a : \nabla\nabla = \mathrm{tr}(a[\nabla \otimes \nabla])$. We will also make a stronger assumption:

Assumption 1.2. The ODE $\dot{x} = b(x)$ has a single fixed point located at x_* (i.e. x_* is the only solution to $b(x) = 0$), which is linearly stable locally (i.e. the real part of all eigenvalues of the matrix $\nabla b(x_*)$ are strictly negative), and globally attracting (i.e. any solution to $\dot{x} = b(x)$ approaches x_* asymptotically).

This assumption implies that ρ_ε becomes atomic on $x = x_*$ as $\varepsilon \to 0$, a property we will need when looking at expectations or probability on the invariant measure.

The remainder of this paper is organized as follows: In Section 2 we will first consider the problem of calculating sharp estimates of expectations, both at finite time (Sections. 2.1 and 2.2) and on the invariant measure of (1.1) (Section 2.3). In Section 3, we will show how to calculate sharp estimates of probabilities densities at finite (Section 3.1) and infinite times (Section 3.2). In Section 4 we then build on these results to calculate probabilities at finite times (Section 4.1) and on the invariant measure of the process (Section 3.2). Finally, in Section 5 we consider the problem of mean exit time calculation. These results are illustrated on finite dimensional examples throughout the paper, and on infinite dimensional examples in Section 6, where we consider linear and nonlinear reaction-advection-diffusion equations with random forcing. The results in this paper are mostly for diffusions, but they can be generalized to other set-ups, in particular MJPs: this is discussed in Section 7. We end the paper with some conclusions in Section 8 and defer some technical results to Appendices.

2 | EXPECTATIONS

2.1 | Finite time expectations

Given the observable $f : \mathbb{R}^n \to \mathbb{R}$ with $f \in C^2(\mathbb{R}^n)$, consider

$$A_\varepsilon(T, x) = \mathbb{E}^x \exp\left(\varepsilon^{-1} f(X_T^\varepsilon)\right), \tag{2.1}$$

where the expectation is taken over samples of the SDE (1.1) conditioned on $X_0^\varepsilon = x$ and evaluated at the final time $t = T < \infty$. Expectations of the form (2.1) are an important first step to compute all other probabilistic variables, as they correspond to the moment generating function of the probability distribution of the random variable $f(X_T)$. Thus, they represent a problem dual to the one of finding the associated probability densities. This will become clear in Section 3, where we use the results presented here to derive similar sharp limits for probability densities.

We have the following proposition:

Proposition 2.1. *Let $(\phi_x(t), \theta_x(t))$ solve the instanton equations*

$$\begin{aligned} \dot{\phi}_x &= a\theta_x + b(\phi_x), & \phi_x(0) &= x, \\ \dot{\theta}_x &= -(\nabla b(\phi_x))^\top \theta, & \theta_x(T) &= \nabla f(\phi_x(T)), \end{aligned} \tag{2.2}$$

and $W_x(t)$ be the solution to the Riccati equation

$$\dot{W}_x = -\nabla\nabla\langle b(\phi_x), \theta_x\rangle - (\nabla b(\phi_x))^\top W_x - W_x(\nabla b(\phi_x)) - W_x a W_x, \qquad W_x(T) = \nabla\nabla f(\phi_x(T)), \tag{2.3}$$

integrated backwards in time from $t = T$ to $t = 0$ along the instanton $\phi_x(t)$. Then the expectation (2.1) satisfies

$$\lim_{\varepsilon\to 0} \frac{A_\varepsilon(T,x)}{\bar{A}_\varepsilon(T,x)} = 1, \tag{2.4}$$

where

$$\bar{A}_\varepsilon(T,x) = R(T,x)\exp\left(\varepsilon^{-1}\left(f(\phi_x(T)) - \frac{1}{2}\int_0^T \langle\theta_x(t), a\theta_x(t)\rangle\, dt\right)\right), \tag{2.5}$$

with

$$R(T,x) = \exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(aW_x(t))\, dt\right). \tag{2.6}$$

Remark 2.1. The function $R(T,x)$ is typically referred to as the prefactor. LDT gives a rougher estimate

$$\lim_{\varepsilon\to 0} \frac{\log A_\varepsilon(T,x)}{\log \bar{A}_\varepsilon(T,x)} = 1, \tag{2.7}$$

which would be unaffected if we were to neglect the prefactor $R(T,x)$ in $\bar{A}_\varepsilon(T,x)$. Of course, this prefactor is key to get the more refined estimate in (2.4).

Remark 2.2. If the SDE (1.1) is modified into

$$dX_t^\epsilon = b\left(X_t^\epsilon\right) dt + \varepsilon\tilde{b}\left(X_t^\epsilon\right) dt + \sqrt{\epsilon}\sigma\, dW_t. \tag{2.8}$$

with $\tilde{b} : \mathbb{R}^n \to \mathbb{R}^n$ is $C^1(\mathbb{R}^n)$ with bounded derivatives, Proposition 2.1 can be generalized by replacing (2.6) with

$$R(T,x) = \exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(aW_x(t))\, dt + \int_0^T \langle\theta_x(t), \tilde{b}(\phi_x(t))\rangle dt\right), \tag{2.9}$$

while leaving unchanged all the other equations in the proposition. The statements in the propositions below can similarly be straightforwardly amended to apply to (2.8), but for the sake of brevity we will stick to (1.1).

Proof of Proposition 2.1. Let

$$u_\varepsilon(T-t,x) = \mathbb{E}^x \exp\left(\varepsilon^{-1} f\left(X_t^\varepsilon\right)\right) \tag{2.10}$$

so that

$$u_\varepsilon(0,x) = A_\varepsilon(T,x). \tag{2.11}$$

It is well-known that u_ε satisfies the BKE

$$\partial_t u_\varepsilon + L_\varepsilon u_\varepsilon = 0, \qquad u_\varepsilon(T,x) = \exp\left(\varepsilon^{-1} f(x)\right), \tag{2.12}$$

where L_ε is the generator of the process (1.1):

$$L_\varepsilon = b(x)\cdot\nabla + \tfrac{1}{2}\varepsilon a \,:\, \nabla\nabla \tag{2.13}$$

Look for a solution of (2.12) of the form

$$u_\varepsilon(t,x) = Z_\varepsilon(t,x)\, \exp\left(\varepsilon^{-1} S(t,x)\right), \tag{2.14}$$

where $S(t,x)$ satisfies the Hamilton-Jacobi equation

$$\partial_t S + b(x)\cdot\nabla S + \tfrac{1}{2}\langle \nabla S, a\nabla S\rangle = \partial_t S + H(x,\nabla S) = 0, \qquad S(T,x) = f(x). \tag{2.15}$$

If $(\phi_x(t),\theta_x(t))$ solves the instanton equations in (2.2), we have $\theta_x(t) = \nabla S(t,\phi_x(t))$ and a direct calculation shows that

$$\frac{d}{dt} S(t,\phi_x(t)) = \partial_t S + \dot\phi_x\cdot\nabla S = \tfrac{1}{2}\langle \theta_x(t), a\theta_x(t)\rangle, \tag{2.16}$$

implying

$$S(T,\phi(T)) - S(0,\phi(0)) = \frac{1}{2}\int_0^T \langle \theta, a\theta\rangle\, dt. \tag{2.17}$$

Since $S(T,\phi(T)) = f(\phi(T))$, we have

$$\exp(\varepsilon^{-1} S(0,\phi(0))) = \exp\left(\varepsilon^{-1}\left(f(\phi(T)) - \frac{1}{2}\int_0^T \langle \theta, a\theta\rangle\, dt\right)\right). \tag{2.18}$$

As a result, to show that (2.4) holds, it remains to establish that the factor $Z_\varepsilon(t,x)$ has a limit as $\varepsilon\to 0$ with $\lim_{\varepsilon\to 0} Z_\varepsilon(0,x) = R(T,x)$. To this end, notice that $Z_\varepsilon(t,x)$ satisfies

$$\partial_t Z_\varepsilon + (b + a\nabla S)\cdot\nabla Z_\varepsilon + \tfrac{1}{2} Z_\varepsilon a \,:\, \nabla\nabla S + \tfrac{1}{2}\varepsilon a \,:\, \nabla\nabla Z_\varepsilon = 0, \qquad Z_\varepsilon(T,x) = 1. \tag{2.19}$$

Formally taking the limit as $\varepsilon\to 0$ on this equation, we deduce that $\lim_{\varepsilon\to 0} Z_\varepsilon(t,x) = Z(t,x)$, where $Z(t,x)$ solves

$$\partial_t Z + (b + a\nabla S)\cdot\nabla Z + \tfrac{1}{2} Z a \,:\, \nabla\nabla S = 0 \qquad Z(T,x) = 1. \tag{2.20}$$

Setting $G_x(t) = Z(t, \phi_x(t))$ on the instanton path, so that $G_x(0) = Z(0, \phi_x(0)) = Z(0, x)$, and using the instanton equation $\dot{\phi}_x = b(\phi_x) + a\theta_x$, we find as evolution equation for G_x

$$\dot{G}_x = -\frac{1}{2} G_x \operatorname{tr}(aW_x), \qquad G_x(T) = 1, \tag{2.21}$$

where the matrix $W_x(t)$ is defined as the Hessian of $S(t, x)$ evaluated along the characteristics (i.e. the instanton path $\phi_x(t)$): $W_x(t) = \nabla\nabla S(t, \phi_x(t))$. In order to solve the equation (2.21) for $G_x(t)$, we need an equation for $W_x(t)$. Differentiating the Hamilton-Jacobi equation (2.15) twice with respect to x and evaluating the result at $x = \phi_x(t)$, it is easy to show that W_x solves the Riccati equation in (2.3). Therefore, by integrating (2.21), we deduce

$$G_x(0) = Z(0, x) = \exp\left(\frac{1}{2}\int_0^T \operatorname{tr}(aW_x(t))\, dt\right) \equiv R(T, x), \tag{2.22}$$

which terminates the proof. □

2.2 | Expectations via Girsanov theorem

An alternative justification of the Riccati equation (2.3) can be given by introducing a stochastic process that samples the Gaussian fluctuations around the instanton. This approach opens up the possibility to alternatively compute the prefactor contribution as an expectations via sampling techniques. This result is well-known (see e.g. [24]) and can be phrased as:

Proposition 2.2. *The prefactor (2.6) satisfies*

$$R(T, x) = \mathbb{E}^0 \exp\left(\frac{1}{2}\int_0^T \nabla\nabla\langle b(\phi_x(t)), \theta_x(t)\rangle : Z_t Z_t\, dt + \frac{1}{2}\nabla\nabla f(\phi_x(T)) : Z_T Z_T\right), \tag{2.23}$$

where $(\phi_x(t), \theta_x(t))$ are defined as in Proposition 2.1, Z_t solves

$$dZ_t = \nabla b(\phi_x(t)) Z_t\, dt + \sigma\, dW_t, \tag{2.24}$$

and the expectation $\mathbb{E}^0$ is taken over realizations of (2.24) conditioned on $Z_0 = 0$.

Proof. Let $Y_t^\varepsilon \in \mathbb{R}^n$ satisfy

$$dY_t^\varepsilon = \dot{\phi}_x(t)\, dt + \sqrt{\varepsilon}\sigma\, dW_t, \tag{2.25}$$

where $\phi_x(t)$ is the instanton solution to (2.2). By invoking Girsanov's theorem, we can write $A_\varepsilon(x)$ as

$$A_\varepsilon(T, x) = \mathbb{E}^x M_T^\varepsilon \exp\left(\varepsilon^{-1} f\left(Y_T^\varepsilon\right)\right) \tag{2.26}$$

where M_T^ε is the Radon-Nikodym density

$$M_T^\varepsilon = \exp\left(-\frac{1}{2\varepsilon}\int_0^T \left|\dot\phi_x - b\left(Y_t^\varepsilon\right)\right|_a^2 dt - \frac{1}{\sqrt{\varepsilon}}\int_0^T \left\langle \sigma^{-1}\left(\dot\phi_x - b\left(Y_t^\varepsilon\right)\right), dW_t\right\rangle \right). \tag{2.27}$$

Since $Y_t^\varepsilon = \phi_x(t) + \sqrt{\varepsilon}\sigma W_t$, after expanding both $b\left(Y_t^\varepsilon\right)$ and $f\left(Y_t^\varepsilon\right)$ in ϵ, it is easy to see that the leading order contribution is precisely given by

$$\begin{aligned}&\exp\left(-\frac{1}{2\varepsilon}\int_0^T |\dot\phi_x - b(\phi_x)|_a^2\, dt + \varepsilon^{-1} f(\phi_x(T))\right)\\ &= \exp\left(-\frac{1}{2\varepsilon}\int_0^T \langle \theta_x(t), a\theta_x(t)\rangle\, dt + \varepsilon^{-1} f(\phi_x(T))\right)\end{aligned}$$

The next order vanishes due to the criticality of the minimizer. The first correction term in the exponential is therefore $O(\epsilon^0)$, that is it gives the prefactor $R(T,x)$, and reads

$$\begin{aligned}R(T,x) = \mathbb{E}\exp\Bigg(&-\frac{1}{2}\int_0^T |\nabla b(\phi_x(t))U_t|_a^2\, dt + \frac{1}{2}\int_0^T \nabla\nabla\langle b(\phi_x),\theta_x\rangle : U_t U_t\, dt\\ &+ \int_0^T \langle \nabla b(\phi_x)U_t, dU_t\rangle_a + \tfrac{1}{2}\nabla\nabla f(\phi_x(T)) : U_T U_T\Bigg),\end{aligned} \tag{2.28}$$

where $U_t = \sigma W_t$. Noticing that the term

$$\exp\left(-\frac{1}{2}\int_0^T |\nabla b(\phi_x)U_t|_a^2\, dt + \int_0^T \langle \nabla b(\phi_x)U_t, dU_t\rangle_a\right) \tag{2.29}$$

is a itself Radon-Nikodym density for the change of measure from the random process Z_t defined in (2.24) to $U_t = \sigma W_t$, we can therefore alternatively write the right hand-side of (2.28) as in (2.23). □

Note that the formula (2.23) immediately tells us that the prefactor, defined as the limit as $\varepsilon \to 0$ of the ratio between $A_\varepsilon(T,x)$ and $\bar A^\varepsilon(T,x)$, is unity when both the drift b and the observable f are linear. Note also that computing the expectation by using the change of measure from the original process to one representing fluctuations around the instanton can be seen as an approximated way (up to terms of order $O(\varepsilon)$) of performing importance sampling via Monte-Carlo method: how to do this importance sampling exactly is harder in general, as discussed for example in [51].

Finally, note that another, more direct, proof of Proposition 2.2 goes as follows. It is easy to see that

$$\mathbb{E}^0\exp\left(\frac{1}{2}\int_0^T \nabla\nabla\langle b(\phi_x),\theta_x\rangle : Z_t Z_t\, dt + \tfrac{1}{2}\nabla\nabla f(\phi_x(T)) : Z_T Z_T\right) = v(0,0) \tag{2.30}$$

where $v(t,z)$ solves

$$\partial_t v + \langle \nabla b(\phi_x) z, \nabla v\rangle + \frac{1}{2} a : \nabla\nabla v + q(t,z) v = 0, \quad v(T,z) = \exp\left(\frac{1}{2}\nabla\nabla f(\phi_x(T)) : zz\right), \tag{2.31}$$

with

$$q(t,z) = \frac{1}{2}\nabla\nabla\langle b(\phi_x(t)), \theta_x\rangle : zz. \tag{2.32}$$

The solution of (2.31) can be expressed as

$$v(t,z) = G_x(t)\exp\left(\frac{1}{2}\langle z, W_x(t) z\rangle\right), \tag{2.33}$$

where $W_x(t)$ and $G_x(t)$ solve (2.3) and (2.21), respectively. Therefore $v(0,0) = G_x(0)$, consistent with the statement in Proposition 2.1. In Appendix B we list a few more expressions that relate the solution to Riccati equations like (2.3) to expectations over the solution of a linear SDE like (2.24). These expressions are useful as they give a possible route to solve the Riccati equation (2.3) via sampling, which is less accurate in general but simpler than solving the Riccati equation itself.

2.3 | Expectations on the invariant measure

Given the observable $f : \mathbb{R}^n \to \mathbb{R}$ with $f \in C^2(\mathbb{R}^n)$, consider the expectation

$$B_\varepsilon = \int_{\mathbb{R}^n} \exp\left(\varepsilon^{-1} f(y)\right)\rho_\varepsilon(y) dy \tag{2.34}$$

where $\rho_\varepsilon(y)$ denotes the invariant density solution to (1.8). (2.34) can also be written as

$$B_\varepsilon = \lim_{T\to\infty} A_\varepsilon(T,x) \tag{2.35}$$

Assumption 1.1 guarantees that this limit exists for appropriate f (e.g. if $|f|$ is bounded), and is independent on x.

The next proposition shows how to calculate a sharp estimate for (2.34) using the procedure of gMAM that allows us to compactify the physical time interval $[0,\infty)$ onto [0,1].

Proposition 2.3. *Let $(\hat\phi(s), \hat\theta(s))$ solve the geometric instanton equations*

$$\begin{aligned} \lambda\hat\phi' &= a\hat\theta + b(\hat\phi), & \hat\phi(0) &= x_*, \\ \lambda\hat\theta' &= -(\nabla b(\hat\phi))^\top\hat\theta, & \hat\theta(1) &= \nabla f(\hat\phi(1)), \end{aligned} \tag{2.36}$$

with $\lambda = |b(\hat\phi)|_a/|\hat\phi'|_a$, and $\hat W(s)$ be the solution to the Riccati equation

$$\lambda\hat W' = -\nabla\nabla\langle b(\hat\phi),\hat\theta\rangle - (\nabla b(\hat\phi))^\top\hat W - \hat W(\nabla b(\hat\phi)) - \hat W a\hat W, \qquad \hat W(1) = \nabla\nabla f(\hat\phi(1)), \tag{2.37}$$

integrated backwards in time from $s = 1$ to $s = 0$ along the instanton $\hat{\phi}(s)$. Then the expectation (2.34) satisfies

$$\lim_{\varepsilon \to 0} \frac{B_\varepsilon}{\bar{B}_\varepsilon} = 1, \tag{2.38}$$

where

$$\bar{B}_\varepsilon = \hat{R} \exp\left(\varepsilon^{-1} \left(f(\hat{\phi}(1)) - \frac{1}{2} \int_0^1 \lambda^{-1}(s) \langle \hat{\theta}(s), a\hat{\theta}(s) \rangle \, ds \right) \right), \tag{2.39}$$

with

$$\hat{R} = \exp\left(\frac{1}{2} \int_0^1 \lambda^{-1}(s) \mathrm{tr}(a\hat{W}(s)) \, ds \right). \tag{2.40}$$

For more details about the geometric instanton equations we refer the reader to [29, 31, 34] as well as Appendix A. Note that the parametrization of $\hat{\phi}$ can be chosen arbitrarily: In the calculations, it is convenient to use normalized arc-length, that is impose that $|\hat{\phi}'(s)|_a = L$, where the constant L is the length of the instanton.

Proof. In order to speak meaningfully about the limit $T \to \infty$, we introduce a reparametrization of time, $t(s) : [0,1] \to \mathbb{R}^+$, to compactify the infinite "physical" time interval. In particular, we choose $t(s)$ such that

$$\frac{d}{ds}(\phi \circ t) = \left(\int_0^T |\dot{\phi}| \, dt \right)^{-1}, \tag{2.41}$$

and denote $\hat{\phi}(s) = (\phi \circ t)(s)$. We then have

$$\hat{\phi}'(s) = \frac{d}{ds}(\phi \circ t)(s) = \frac{dt}{ds} \dot{\phi}\big|_{t=t(s)} \equiv \lambda^{-1}(s) \dot{\phi}\big|_{t=t(s)} \tag{2.42}$$

where we introduced $\lambda(s) = ds/dt$. As a consequence, $\hat{\phi}(s)$ fulfills the geometric instanton equations (2.36). These equations can be used to show that, in the limit as $T \to \infty$, the initial condition becomes irrelevant, and we can consider $\hat{\phi}(0) = x_*$. First, from (1.2) we know that $S_T = \frac{1}{2} \int_0^T |\dot{x} - b(x)|^2 \, dt$, so that in the limit $T \to \infty$, we must have $\dot{x} = b(x)$ for an infinite amount of time if the action is to remain finite. As a consequence, for $T \to \infty$, the global minimizer will decay towards the unique fixed point x_*, as all initial points are attracted towards it.

In order to find the global minimum, we can consider the two separate problems of first approaching the fixed point, and subsequently leaving again. For this purpose, consider the trajectory $\eta(t)$, and corresponding reparametrized trajectory $\hat{\eta}(s)$ with $\dot{\eta} = \lambda \hat{\eta}'$, $s \in [-1, 1]$, and

$$\hat{\eta}(s) = \begin{cases} \hat{\eta}_1(s) & \text{for} \quad s \in [-1, 0], \\ \hat{\eta}_2(s) = \hat{\phi}(s) & \text{for} \quad s \in [0, 1], \end{cases} \tag{2.43}$$

where

$$\lambda\hat{\eta}_1' = b(\hat{\eta}_1), \quad \hat{\eta}_1(-1) = x, \quad \hat{\eta}_1(0) = x_*. \tag{2.44}$$

It follows that $\hat{\eta}(s)$ corresponds to the trajectory that deterministically decays into the fixed point x_* starting from x, and then corresponds to the minimizer $\hat{\phi}$, solution to equations (2.36) from then on.

Since x_* is the unique fixed point and all $x \in \mathbb{R}^n$ are attracted to it, such an $\hat{\eta}_1$ exists and is unique for all x. On the other hand, since $(\hat{\eta}_1, \hat{\theta}_{\eta_1})$ fulfill instanton equations on $s \in [-1, 0]$, and $\lambda\hat{\eta}_1' = b(\hat{\eta}_1)$, we have that $\hat{\theta}_{\eta_1} = 0$ and the corresponding action vanishes on the interval $s \in [-1, 0]$. Therefore, the action associated with $\hat{\eta}(s)$ is equal to the action associated with $\hat{\phi}(s)$. The problem of finding a global minimizer starting at x reduces to the problem of solving (2.36).

Another consequence of taking the limit $T \to \infty$ is that the Hamiltonian is zero along the instanton, that is $H(\hat{\phi}(s), \hat{\theta}(s)) = 0 \; \forall s \in [0, 1]$. As a result

$$|a\hat{\theta} + b(\hat{\phi})|_a^2 = \langle\hat{\theta}, a\hat{\theta}\rangle + 2\langle\hat{\theta}, b(\hat{\phi})\rangle + |b(\hat{\phi})|_a^2 = 2H(\hat{\phi}, \hat{\theta}) + |b(\hat{\phi})|_a^2 = |b(\hat{\phi})|_a^2. \tag{2.45}$$

Since $\dot{\phi}|_{t(s)} = \lambda(s)\hat{\phi}'(s)$, this allows us to deduce the following expression for λ:

$$\lambda(s) = \frac{|\dot{\phi}(t(s))|_a}{|\hat{\phi}'(s)|_a} = \frac{|a\hat{\theta} + b(\hat{\phi})|_a}{|\hat{\phi}'|_a} = \frac{|b(\hat{\phi})|_a}{|\hat{\phi}'|_a}. \tag{2.46}$$

This is the expression stated in the proposition.

It remains to be shown that the reparametrization of the Riccati equation, (2.37), is well posed. To this end, notice that as $s \to 0$, we have $\hat{\phi}(s) \to x_*$ and $\hat{\theta}(s) \to 0$. Therefore,

$$\lambda\hat{W}' = -(\nabla b(x_*))^\top\hat{W} - \hat{W}(\nabla b(x_*)) - \hat{W}a\hat{W} \quad \text{for} \quad 0 < s \ll 1, \tag{2.47}$$

which leads to the conclusion that, since $\nabla b(x_*)$ is negative definite by definition, we have $\hat{W}(s) \to 0$ as $s \to 0$. □

2.3.1 | Example: Gradient system

An easy example that can be computed explicitly is the case of diffusion in a potential landscape,

$$dX_t^\varepsilon = -\nabla U(X_t^\varepsilon)\, dt + \sqrt{2\varepsilon} dW_t, \tag{2.48}$$

where $X \in \mathbb{R}^n$, and $U : \mathbb{R}^n \to \mathbb{R}$ is a potential with a unique minimum x_*, such that $\int_{\mathbb{R}^n} e^{-\varepsilon^{-1}U(x)}dx < \infty$ for all $\varepsilon > 0$—we can set $U(x_*) = 0$ without loss of generality. Given an observable $f : \mathbb{R}^n \to \mathbb{R}$ such that $f(y)$ grows sufficiently slower than $U(y)$ as $|y| \to \infty$, we want to obtain a sharp estimate of the expectation

$$B_\varepsilon = \int_{\mathbb{R}^n} \exp(\varepsilon^{-1}f(y))\rho_\varepsilon(y)dy, \tag{2.49}$$

where $\rho_\varepsilon(y)$ is the density of the invariant measure of (2.48), given by

$$\rho_\varepsilon(y) = Z_\varepsilon^{-1} \exp(-\varepsilon^{-1}U(y)) \qquad \text{with} \qquad Z_\varepsilon = \int_{\mathbb{R}^n} \exp(-\varepsilon^{-1}U(y))dy. \tag{2.50}$$

Combining (2.49) and (2.50) we see that B_ε is given by

$$B_\varepsilon = \frac{\int_{\mathbb{R}^n} \exp(\varepsilon^{-1}(f(y) - U(y)))dy}{\int_{\mathbb{R}^n} \exp(-\varepsilon^{-1}U(y))dy}, \tag{2.51}$$

and both integrals can be estimated by Laplace's method. The result is that $\lim_{\varepsilon\to 0} B_\varepsilon/\bar{B}_\varepsilon = 1$ with $\bar{B}_\varepsilon$ given by

$$\bar{B}_\varepsilon = \left(\frac{\det H(x_*)}{\det\left(H(x_f) - \nabla\nabla f(x_f)\right)}\right)^{1/2} \exp\left(\varepsilon^{-1}\big(f(x_f) - U(x_f)\big)\right), \tag{2.52}$$

where $H(x) = \nabla\nabla U(x)$ is the Hessian of the potential and the point $x_f \in \mathbb{R}^n$ is the solution of the optimization problem

$$x_f = \operatorname*{argmin}_{y\in\mathbb{R}^n} (U(y) - f(y)). \tag{2.53}$$

We will now show that Proposition 2.3 yields the same result.

First, we know explicitly that the instanton fulfills

$$\lambda\hat{\phi}'(s) = -\nabla U(\hat{\phi}(s)) \quad \text{and} \quad \hat{\theta}(s) = \nabla U(\hat{\phi}(s)), \tag{2.54}$$

where primes again denote derivatives with respect to arc-length. Further, the endpoints of the instanton are $x_* = \hat{\phi}(0)$ and $x_f = \hat{\phi}(1)$: the first is by definition, and the second since the final condition for the equation for $\hat{\theta}(s)$ gives

$$\hat{\theta}(1) = \nabla f(\hat{\phi}(1)) = \nabla U(\hat{\phi}(1)), \tag{2.55}$$

where the second equality follows from the second equation in (2.54). Since (2.55) is also the Euler-Lagrange equation for the minimization problem in (2.53), we deduce $x_f = \hat{\phi}(1)$. Therefore, using

$$\lambda^{-1}(s)\langle\hat{\theta},\hat{\theta}\rangle = \langle\nabla U(\hat{\phi}(s)),\hat{\phi}'(s)\rangle = \frac{d}{ds}\nabla U(\hat{\phi}(s)) \tag{2.56}$$

we deduce that the exponential term in the expectation is given by

$$\exp\left(\varepsilon^{-1}\left(f(x_f) - \int_0^1 \lambda^{-1}(s)|\hat{\theta}(s)|^2\, ds\right)\right) = \exp\left(\varepsilon^{-1}\big(f(x_f) - U(_f)\big)\right). \tag{2.57}$$

Second, we have

$$\nabla\nabla\langle b,\hat{\theta}\rangle = -\nabla\nabla\nabla U(\hat{\phi})\nabla U(\hat{\phi}) = -\nabla\nabla\nabla U(\hat{\phi})\lambda\hat{\phi}' = \lambda H'. \tag{2.58}$$

This means that the Riccati equation (2.37) is here given by

$$\lambda\hat{W}' = \lambda H' + H\hat{W} + \hat{W}H + 2\hat{W}^2, \qquad \hat{W}(1) = \nabla\nabla f(x_f). \tag{2.59}$$

Let us look for a solution of the form

$$\hat{W} = \lambda D^{-1}D', \tag{2.60}$$

for a matrix $D \in \mathbb{R}^{n\times n}$, with $D(1) = \mathrm{Id}$, and $\lambda(1)D'(1) = \nabla\nabla f(x_f)$ to fulfill the boundary conditions for $\hat{W}$. Equation (2.60) can be written as $D' = \lambda^{-1}D\hat{W}$ which, from Liouville's formula, $\Psi'(s) = A(s)\Psi(s)$ implies $\det\Psi(s) = \det\Psi(0)\exp(\int_0^s \mathrm{tr}A(\tau)d\tau)$ for $A, \Psi \in \mathbb{R}^{n\times n}$, yields

$$\det D(0) = \exp\left(\int_0^1 \lambda^{-1}(s)\mathrm{tr}\hat{W}(s)\,ds\right). \tag{2.61}$$

From Equation (2.60) it also follows that

$$\lambda\hat{W}' = \lambda D^{-1}(\lambda D')' - \lambda^2 D^{-1}D'D^{-1}D', \tag{2.62}$$

which we can use in conjunction with the Riccati equation to get

$$\lambda(\lambda D')' = \lambda(DH)' + \lambda(DH - \lambda D')D^{-1}D' \tag{2.63}$$

or equivalently

$$\lambda(\lambda D' - DH)' = -(\lambda D' - DH)W. \tag{2.64}$$

Using again Liouville's formula (this time for $\Psi = \lambda D' - DH$) yields the relation

$$\det\big(\lambda(0)D'(0) - D(0)H(x_*)\big) = \det\big(\lambda(1)D'(1) - D(1)H(x_f)\big)\exp\left(-\int_0^1 \lambda^{-1}(s)\mathrm{tr}\hat{W}(s)\,ds\right), \tag{2.65}$$

into which we can insert the boundary conditions $\lambda(1)D'(1) = \nabla\nabla f(x_f)$, $D(1) = \mathrm{Id}$ and $\lambda(0) = 0$, as well as $\det D(0)$ from Equation (2.61) to obtain

$$\det H(x_*) = \det\big(H(x_f) - \nabla\nabla f(x_f)\big)\exp\left(-2\int_0^1 \lambda^{-1}(s)\mathrm{tr}\hat{W}(s)\,ds\right). \tag{2.66}$$

This gives eventually the prefactor contribution,

$$\hat{R} = \exp\left(\int_0^1 \lambda^{-1}(s)\mathrm{tr}\hat{W}(s)\,ds\right) = \left(\frac{\det H(x_*)}{\det\big(H(x_f) - \nabla\nabla f(x_f)\big)}\right)^{1/2}. \tag{2.67}$$

Therefore we recover $\bar{B}_\varepsilon$ given in (2.52), as needed.

Note that the above results can be generalized to a diffusion in a potential landscape with mobility matrix, that is systems of the form

$$dX = -M\nabla U(X)\,dt + \sqrt{2\varepsilon}M^{1/2}\,dW, \tag{2.68}$$

where $M \in \mathbb{R}^{n\times n}$ is the symmetric, positive definite mobility matrix, and $M^{1/2} \in \mathbb{R}^{n\times n}$ is a symmetric matrix with $M^{1/2}M^{1/2} = M$. The procedure above can be repeated by replacing $\hat{W}$ with $M^{-1/2}\hat{W}M^{-1/2}$ and H with $M^{-1/2}HM^{-1/2}$. The result (2.52) is unchanged.

3 | PROBABILITY DENSITIES

3.1 | Probability densities at finite time

Here we estimate the probability density function of X_t^ε in the limit of small ε. Denote this density by $\rho_\varepsilon^x(t,y)$, so that, given any suitable $F : \mathbb{R}^n \to \mathbb{R}$,

$$\int_{\mathbb{R}^n} F(y)\rho_\varepsilon^x(T,y)dx = \mathbb{E}^x F\left(X_T^\varepsilon\right) \tag{3.1}$$

The next proposition shows how to get a sharp estimate of $\rho_\varepsilon^x(t,y)$ at any y by purely local considerations:

Proposition 3.1. *Let $(\phi_{x,y}(t), \theta_{x,y}(t))$ solve the instanton equations*

$$\begin{aligned} \dot{\phi}_{x,y} &= a\theta_{x,y} + b(\phi_{x,y}), \qquad & \phi_{x,y}(0) &= x, \\ \dot{\theta}_{x,y} &= -(\nabla b(\phi_{x,y}))^\top\theta, \qquad & \phi_{x,y}(T) &= y, \end{aligned} \tag{3.2}$$

and let $Q_{x,y}(t)$ be the solution to the forward Riccati equation

$$\dot{Q}_{x,y} = Q_{x,y}K_{x,y}Q_{x,y} + Q_{x,y}(\nabla b(\phi_{x,y}))^\top + (\nabla b(\phi_{x,y}))Q_{x,y} + a, \quad Q_{x,y}(0) = 0. \tag{3.3}$$

where we denote

$$K_{x,y}(t) = \nabla\nabla\langle b(\phi_{x,y}(t)), \theta_{x,y}(t)\rangle, \tag{3.4}$$

In addition let $I_x(T,y)$ be defined as

$$I_x(T,y) = \frac{1}{2}\int_0^T \langle\theta_{x,y}(t), a\theta_{x,y}(t)\rangle\,dt \geq 0. \tag{3.5}$$

Then the probability density function $\rho_\varepsilon^x(T,y)$ of satisfies

$$\lim_{\varepsilon\to 0}\frac{\rho_\varepsilon^x(T,y)}{\bar{\rho}_\varepsilon^x(T,y)} = 1 \qquad \textit{pointwise in } y, \tag{3.6}$$

where $\bar\rho^x_\varepsilon(T,y)$ *is given by*

$$\bar\rho^x_\varepsilon(T,y) = (2\pi\varepsilon)^{-n/2}|\det Q_{x,y}(T)|^{-1/2}\exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(K_{x,y}(t)Q_{x,y}(t))\,dt - \frac{1}{2\varepsilon}I_x(T,y)\right). \tag{3.7}$$

Remark 3.1. Equation (3.3) for $Q_{x,y}$ is structurally equivalent to the equation (2.3) for $W_{x,y} = Q_{x,y}^{-1}$ except for the boundary conditions. The boundary conditions necessitate solving $Q_{x,y}$ forward in time. Notably, this direction of integration is also the one in which the equation for $Q_{x,y}$ is numerically stable, as can be intuited for example by considering the Ornstein-Uhlenbeck process $b(x) = -\gamma x$, for $\gamma > 0$ (see Section 3.1.1 below). This feature will become even more apparent in the context of stochastic partial differential equations, see the discussion after Proposition 6.1 in Section 6.

Remark 3.2. In Appendix B we discuss how to solve (3.3) via sampling of the solution of a nonlinear (in the sense of McKean) SDE.

Intuitively, the local estimation of of $\rho^x_\varepsilon(t,y)$ implied by Proposition 3.1 is possible because in the limit $\varepsilon \to 0$, this probability density is dominated by the instanton, and the prefactor contributions can again be estimated from the Gaussian fluctuations around this instanton. More concretely, we will see in the proof of Proposition 3.1 below (see (3.23)) that $Q_{x,y}(T)$ is the inverse of the Hessian of the action $I_x(T,y)$, that is

$$Q_{x,y}(T) = [\nabla_y\nabla_y I_x(T,y)]^{-1} \tag{3.8}$$

This implies that we can also write (3.7) as

$$\bar\rho^x_\varepsilon(T,y) = (2\pi\varepsilon)^{-n/2}|\det \nabla_y\nabla_y I_x(T,y)|^{1/2}\exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(K_{x,y}(t)Q_{x,y}(t))\,dt - \frac{1}{2\varepsilon}I_x(T,y)\right). \tag{3.9}$$

This form shows that $\bar\rho^x_\varepsilon(T,y)$ is normalized in the limit as $\varepsilon \to 0$, a result we state as

Lemma 3.2. *The density* $\bar\rho^x_\varepsilon(T,y)$ *defined in (3.7) is asymptotically normalized, that is*

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\bar\rho^x_\varepsilon(T,y)dy = 1 \tag{3.10}$$

Proof. Starting from expression (3.9) for $\bar\rho^x_\varepsilon(T,y)$, which as we show in the proof of Proposition 3.1 is equivalent to (3.7), let us evaluate the integral $\int_{\mathbb{R}^n}\bar\rho^x_\varepsilon(T,y)dy$ by Laplace's method. To this end, we must first identify the point where $I_x(T,y)$ is minimal. It is easy to see that this point is $y = y_x(T)$, where $y_x(t)$ is the solution to the ODE $\dot y_x = b(y_x)$, $y_x(0) = x$. Indeed, $y_x(t)$ is also the instanton obtained when $\theta_{x,y_x(T)}(t) = 0$, and from (3.5) it gives $I_x(T,y_x(T)) = 0$, which implies that the minimum at y_x is necessarily a global minimum. Similarly, $\theta_{x,y_x(T)}(t) = 0$ implies that $K_{x,y_x(T)}(t) = 0$. Therefore

$$\begin{aligned}\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\bar\rho^x_\varepsilon(T,y)dy &= \lim_{\varepsilon\to 0}(2\pi\varepsilon)^{-n/2}|\det\nabla_y\nabla_y I_x(T,y_x(T))|^{1/2}\int_{\mathbb{R}^n}\exp\left(-\frac{1}{2\varepsilon}I_x(T,y)\right)dy\\ &= (2\pi)^{-n/2}|\det\nabla_y\nabla_y I_x(T,y_x(T))|^{1/2}\int_{\mathbb{R}^n}\exp\left(-\frac{1}{2}\langle u,\nabla_y\nabla_y I_x(T,y_x(T))u\rangle\right)du = 1\end{aligned} \tag{3.11}$$

where to get the second equality we used $u = (z - y_x(T))/\sqrt{\varepsilon}$ as new integration variable and only keep the zeroth order term in ε that contributes to the limit. □

In the proof of this lemma, only the behavior of $\bar{\rho}^x_\varepsilon(T,y)$ near its maximum at $y_x(T)$ matters, but let us stress that the expression for $\bar{\rho}^x_\varepsilon(T,y)$ given in (3.7) offers a much finer approximation of this density valid arbitrary far away from $y_x(T)$. This will allow us to use $\bar{\rho}^x_\varepsilon(T,y)$ to estimate expectations dominated by tail events, as shown in Section 3.3.

Proof of Proposition 3.1. Consider the expectation

$$C^x_\varepsilon(\eta) = \mathbb{E}^x \exp(\varepsilon^{-1}\langle \eta, X^\varepsilon_T\rangle), \qquad \eta \in \mathbb{R}^n. \tag{3.12}$$

We can alternatively express this expectation as

$$C^x_\varepsilon(\eta) = \int_{\mathbb{R}^n} \exp\left(\varepsilon^{-1}\langle \eta, y\rangle\right)\rho^x_\varepsilon(T,y)\,dy, \tag{3.13}$$

We will show that $\bar{\rho}^x_\varepsilon(T,y)$ allows us to estimate this expectation for all $\eta \in \mathbb{R}^d$. Assuming that $\rho^x_\varepsilon(T,y)$ is of the form

$$\rho^x_\varepsilon(T,y) = (2\pi\varepsilon)^{-n/2}\left(R^x_T(y) + \mathcal{O}(\varepsilon)\right)\exp(-\varepsilon^{-1}I_x(T,y)), \tag{3.14}$$

we can evaluate the integral in (3.13) by Laplace's method to obtain

$$C^x_\varepsilon(\eta) = \left(R^x_T(y) + \mathcal{O}(\varepsilon)\right)|\det \nabla_y\nabla_y I_x(T,y)|^{-1/2}\exp(\varepsilon^{-1}(\langle \eta, y\rangle - I_x(T,y))), \tag{3.15}$$

where

$$y = \operatorname*{argmax}_{z\in\mathbb{R}^n}(\langle \eta, z\rangle - I_x(T,z)). \tag{3.16}$$

From Proposition 2.1 we also know that

$$C^x_\varepsilon(\eta) = \left(\exp\left(\frac{1}{2}\int_0^T \operatorname{tr}(aW_x(t))\,dt\right) + \mathcal{O}(\varepsilon)\right)\exp\left(\varepsilon^{-1}\left(\langle \eta, \phi_x(T)\rangle - \frac{1}{2}\int_0^T \langle \theta_x, a\theta_x\rangle\,dt\right)\right), \tag{3.17}$$

where $(\phi_x(t), \theta_x(t))$ solve the instanton equations (2.2) and $W_x(t)$ solves the Riccati equation (2.3); note that, since $f(x) = \langle \eta, x\rangle$ here, the boundary conditions at $t = T$ reduce to $\theta_x(T) = \eta$ and $W_x(T) = 0$. Comparison between (3.15) and (3.17) implies that $\phi_x(T) = y$, that is the instanton equations for $(\phi_x(t), \theta_x(t)) \equiv (\phi_{x,y}(t), \theta_{x,y}(t))$ reduce to the form in (3.2). In addition, $I_x(T,y)$ is given by (3.5) and

$$R^x_T(y) = |\det \nabla_y\nabla_y I_x(T,y)|^{1/2}\exp\left(\frac{1}{2}\int_0^T \operatorname{tr}(aW_x(t))\,dt\right). \tag{3.18}$$

Using Proposition C.1 with $K(t) = K_{x,y}(t)$, we can identify $W(t) = W_x(t)$, and $Q(t) = Q_{x,y}(t)$ to deduce that

$$\int_0^T \mathrm{tr}(aW_x(t))dt = \int_0^T \mathrm{tr}(K_{x,y}(t)Q_{x,y}(t))dt \tag{3.19}$$

since $W_x(T) = 0$ and $Q_{x,y}(0) = 0$ and so $\det(\mathrm{Id} - W_x(T)Q_{x,y}(T)) = \det(\mathrm{Id} - W_x(0)Q_{x,y}(0)) = 1$. Therefore, to finish the proof it remains to evaluate $\det[\nabla_y \nabla_y I_x(T,y)]$. To this end, notice first that, since $I_x(T,y)$ is given by (3.5), it satisfies

$$\partial_t I_x(t,y) + H(y, \nabla_y I_x(t,y)) = 0, \qquad \lim_{t\to 0} tI_x(t,y) = \tfrac{1}{2}\langle (y-x), a^{-1}(y-x)\rangle. \tag{3.20}$$

Indeed, the solution to this equation can be expressed as in (3.5), and the boundary condition follows from the fact that, for small T, to leading order the solution to the instanton equations (3.2) reads

$$\phi_{x,y}(t) = x + (y-x)t/T + O(T), \qquad \theta_{x,y}(t) = a^{-1}(y-x)/T + O(T), \qquad t \in [0,T]. \tag{3.21}$$

This implies that

$$I_x(T,y) = \frac{1}{2}\int_0^T \langle \theta_{x,y}(t), a\theta_{x,y}(t)\rangle\, dt = \tfrac{1}{2}T^{-1}\langle (y-x), a^{-1}(y-x)\rangle + O(1) \quad \text{for} \quad T \ll 1, \tag{3.22}$$

consistent with the boundary condition in (3.20). Therefore, if we introduce

$$[\nabla_y \nabla_y I_x(t, \phi_{x,y}(t))]^{-1} = Q_{x,y}(t) \tag{3.23}$$

so that $\nabla_y \nabla_y I_x(T,y) = Q_{x,y}^{-1}(T)$, then an equation for $Q_{x,y}(t)$ can be derived by (i) differentiating (3.20) twice with respect to y and evaluating the result at $y = \phi_{x,y}(t)$ to obtain an equation for $\nabla_y \nabla_y I_x(t, \phi_{x,y}(t))$, and (ii) using this equation to derive one for $Q_{x,y}(t)$. The result is the Riccati equation (3.3), in which the boundary condition follows from $\lim_{t\to 0} t\nabla_y \nabla_y I_x(t,y) = \frac{1}{2}a^{-1}$, that is $[\nabla_y \nabla_y I_x(t,y)]^{-1} = O(t)$ and hence $Q_{x,y}(t) = [\nabla_y \nabla_y I_x(t, \phi_{x,y}(t))]^{-1} = O(t)$. □

3.1.1 | Example: The one-dimensional Ornstein-Uhlenbeck process

Consider Equation (1.1) with $n = 1$, $\sigma = 1$, and $b(x) = -x$, that is

$$dX_t^\varepsilon = -X_t^\varepsilon\, dt + \sqrt{\varepsilon} dW_t, \quad X_0^\varepsilon = x. \tag{3.24}$$

This one-dimensional and linear case is the easiest possible non-trivial scenario, and in particular we know explicitly that

$$\rho_\varepsilon^x(T,y) = \sqrt{\frac{1}{\pi\varepsilon(1-e^{-2T})}} \exp\left(-\frac{|y - xe^{-t}|^2}{\varepsilon(1-e^{-2T})}\right). \tag{3.25}$$

We want to compare this analytical result to the approximation $\bar{\rho}_\varepsilon^x(T,y)$ given in Equation (3.7). In fact it turns out that $\bar{\rho}_\varepsilon^x(T,y)$ of proposition 3.1 is exact in this case, since there is no higher order contribution in ε in the prefactor. To show this, we have the instanton equations

$$\begin{cases} \dot{\phi} = -\phi + \theta, & \phi(0) = x, \\ \dot{\theta} = \theta, & \phi(T) = y, \end{cases} \tag{3.26}$$

which are solved by

$$\begin{cases} \phi(t) = \dfrac{y - xe^{-T}}{1 - e^{-2T}}\left(e^{t-T} - e^{-t-T}\right) \\ \theta(t) = \dfrac{2\left(y - xe^{-T}\right)}{1 - e^{-2T}} e^{t-T}. \end{cases} \tag{3.27}$$

The exponential estimate therefore yields

$$\exp\left(-\frac{1}{2\varepsilon}\int_0^T \theta^2(t)\,dt\right) = \exp\left(-\frac{\left|y - xe^{-T}\right|^2}{\varepsilon(1 - e^{-2T})}\right). \tag{3.28}$$

For the prefactor, we have the Riccati equation

$$\dot{Q} = -2Q + 1, \qquad Q(0) = 0, \tag{3.29}$$

which is solved by

$$Q(t) = \frac{1}{2}(1 - e^{-2t}). \tag{3.30}$$

Therefore

$$|Q(T)|^{-1/2} = \left(\frac{1}{2}(1 - e^{-2T})\right)^{-1/2}, \tag{3.31}$$

and since $\nabla\nabla b(x) = 0$ in this example, we obtain

$$\bar{\rho}_\varepsilon^x(T,y) = \sqrt{\frac{1}{\pi\varepsilon(1 - e^{-2T})}} \exp\left(-\frac{\left|y - xe^{-T}\right|^2}{\varepsilon(1 - e^{-2T})}\right), \tag{3.32}$$

which is precisely the analytical result.

3.2 | Probability density of the invariant measure

We can generalize Proposition 3.1 to calculate the probability density of the invariant measure, using again the procedure of gMAM to compactify physical time:

Proposition 3.3. *Let $(\hat\phi_y(s),\hat\theta_y(s))$ solve the geometric instanton equations*

$$\begin{aligned}\lambda\hat\phi_y' &= a\hat\theta_y + b(\hat\phi_y), &\qquad \hat\phi_y(0) &= x_*,\\ \lambda\hat\theta_y' &= -(\nabla b(\hat\phi_y))^\top\theta_y, &\qquad \hat\phi_y(1) &= y.\end{aligned} \tag{3.33}$$

where $\lambda = |a\hat\theta_y + b(\hat\phi_y)|/|\phi_y'|$. In addition let Q_* be the solution to the Lyapunov equation

$$0 = Q_*(\nabla b(x_*))^\top + (\nabla b(x_*))Q_* + a, \tag{3.34}$$

and let $\hat Q_y(s)$ be the solution to the forward Riccati equation

$$\lambda\hat Q_y' = \hat Q_y\hat K_y\hat Q_y + \hat Q_y(\nabla b(\hat\phi_y))^\top + (\nabla b(\hat\phi_y))\hat Q_y + a, \qquad \hat Q_y(0) = Q_*, \tag{3.35}$$

where we denote

$$\hat K_y(s) = \nabla\nabla\langle b(\hat\phi_y(s)),\hat\theta_y(s)\rangle, \tag{3.36}$$

Finally, let $\hat I(y)$ be given by

$$\hat I(y) = \frac{1}{2\varepsilon}\int_0^1 \lambda^{-1}(s)\langle\hat\theta_y(s), a\hat\theta_y(s)\rangle ds, \tag{3.37}$$

Then the probability density function $\rho_\varepsilon(y)$ of satisfies

$$\lim_{\varepsilon\to 0}\frac{\rho_\varepsilon(y)}{\bar\rho_\varepsilon(y)} = 1, \tag{3.38}$$

in which $\bar\rho_\varepsilon(y)$ is given by either one of the two equivalent expressions

$$\begin{aligned}\bar\rho_\varepsilon(y) &= (2\pi\varepsilon)^{-n/2}|\det\hat Q_y(1)|^{-1/2}\exp\left(\frac{1}{2}\int_0^1\lambda^{-1}(s)\mathrm{tr}(\hat K_y(s)\hat Q_y(s))\,ds - \frac{1}{2\varepsilon}\hat I(y)\right)\\ &= (2\pi\varepsilon)^{-n/2}|\det Q_*|^{-1/2}\exp\left(-\int_0^1\lambda^{-1}(s)\Big(\nabla\cdot b(\hat\phi_y(s)) + \tfrac{1}{2}\mathrm{tr}(a\hat Q_y^{-1}(s))\Big)\,ds - \frac{1}{2\varepsilon}\hat I(y)\right).\end{aligned} \tag{3.39}$$

The function $\hat I(y)$ defined in (3.37) is important: it is the quasipotential of y with respect to x_*. Interestingly $\hat Q_y(s) = [\nabla\nabla\hat I(\hat\phi_y(s))]^{-1}$ and so (3.7) can also be written as

$$\begin{aligned}\bar\rho_\varepsilon(y) &= (2\pi\varepsilon)^{-n/2}|\det\nabla\nabla\hat I(y)|^{1/2}\exp\left(\frac{1}{2}\int_0^1\lambda^{-1}(s)\mathrm{tr}(\hat K_y(s)\hat Q_y(s))\,ds - \frac{1}{2\varepsilon}\hat I(y)\right)\\ &= (2\pi\varepsilon)^{-n/2}|\det\nabla\nabla\hat I(x_*)|^{1/2}\exp\left(-\int_0^1\lambda^{-1}(s)\hat L(y,s)\,ds - \frac{1}{2\varepsilon}\hat I(y)\right),\end{aligned} \tag{3.40}$$

where we defined

$$\hat{L}(s,y) = \nabla \cdot b(\hat{\phi}_y(s)) + \tfrac{1}{2}\mathrm{tr}(a\nabla\nabla\hat{I}(\hat{\phi}_y(s))) \tag{3.41}$$

The second expression in (3.40) is consistent with the one derived in [9]. Note that we can find yet another illuminating form to write (3.40) by use of the *transverse decomposition* that decomposes the drift $b(x)$ into a gradient of the quasipotential, and a transverse portion $l(x)$, defined as

$$l(x) = b(x) + \tfrac{1}{2}a\nabla\hat{I}(x) \qquad \text{with} \qquad \langle \nabla I(x), l(x)\rangle = 0. \tag{3.42}$$

Such a decomposition is guaranteed to exist under the assumption that the quasipotential is continuously differentiable [24, chap. 3.4]. In terms of $l(x)$ (3.41) reduces to $\hat{L}(s,y) = \nabla \cdot l(\hat{\phi}_y(s))$. Note that for a reversible diffusion, that is a gradient flow, we have $U(x) = 2\hat{I}(x)$, and so $l(x) = 0$ and $\hat{L}(s,y)$ vanishes. More generally, in flows that have $\nabla \cdot l(x) = 0$, the invariant measure remains the Gibbs-measure, $\rho_\varepsilon(y) = (2\pi\varepsilon)^{-n/2}|\det \nabla\nabla\hat{I}(x_*)|^{1/2}\exp(-\frac{1}{2\varepsilon}\hat{I}(y))$, leading the authors of [9] to call the quantity $\exp(-\int_0^1 \lambda^{-1}(s)\nabla \cdot l(x)\,ds)$ the *non-Gibbsianness* of the flow. It is important to remark, though, that this simplification is only available in the infinite time case on the invariant measure, as for finite times the corresponding quantity $\nabla_y\nabla_y I_x(t,y)$ becomes singular at $t = 0$. Further, since $\nabla \cdot l$ is generally not available explicitly, the first form of (3.40) is the one most easily used for numerical computation.

We can use (3.40) to show that $\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\bar{\rho}_\varepsilon(y)dy = 1$, that is the equivalent of Lemma 3.2 holds in the infinite time limit using the geometric expressions above. Basically, this is because the dominating point in this integral is $y = x_*$, for which $\hat{\theta}_{x_*}(s) = 0$ and $\hat{Q}_{x_*}(s) = Q_* = \nabla\nabla\hat{I}^{-1}(x_*)$. This also shows that, $O(\sqrt{\varepsilon})$ away from x_*, $\rho_\varepsilon(y)$ can be approximated by the Gaussian density with mean x_* and covariance εQ_* given by

$$(2\pi\varepsilon)^{-n/2}|\det Q_*|^{-1/2}\exp\left(-\frac{1}{2\varepsilon}\langle (y-x_*), Q_*^{-1}(y-x_*)\rangle\right), \tag{3.43}$$

and this density is normalized. However, for locations y that are $O(1)$ away from x_*, this Gaussian approximation is no longer valid, and the full expression (3.39) must be used as estimate of the probability density on the invariant measure, as shown in Section 3.3.

Proof of Proposition 3.3. To get the first equality in (3.39), we can mimic the steps in the proof of Proposition 3.1. The only difference is that we need to show the validity of the boundary conditions for $\hat{Q}(0)$ given in equation (3.34). It was established before that, regardless of the initial conditions x, the instanton $\hat{\eta}(s)$ first decays into the unique fixed point x_* on $s \in [-1, 0]$. For $s = 0$, we then have $\hat{\phi}(0) = x_*$, $\hat{\theta}(0) = 0$, and $\lambda(0) = |b(x_*)|_a/|\hat{\phi}'(0)|_a = 0$, so that we deduce from the arc-length reparametrization of Equation (3.3), given in (3.35), that

$$0 = \hat{Q}(0)(\nabla b(x_*))^\top + (\nabla b(x_*))\hat{Q}(0) + a. \tag{3.44}$$

This shows that $\hat{Q}_y(0) = Q_*$ with Q_* solution to (3.34).

To get the second equality in (3.39), start from (3.35) for $\hat{Q}_y$ and use Liouville's formula to deduce that

$$\begin{aligned}\lambda\frac{d}{ds}\log\det\hat{Q}_y &= \mathrm{tr}\big[\hat{Q}_y^{-1}(\hat{Q}_y\hat{K}_y\hat{Q}_y + \hat{Q}_y(\nabla b(\hat{\phi}_y))^\top + (\nabla b(\hat{\phi}_y))\hat{Q}_y + a)\big]\\ &= \mathrm{tr}(\hat{K}_y\hat{Q}_y) + 2\nabla\cdot b + \mathrm{tr}(a\hat{Q}_y^{-1})\end{aligned} \tag{3.45}$$

Integrating this equation on $s\in[0,1]$ using $\hat{Q}_y(0)=Q_*$ gives

$$\det\hat{Q}_y(1) = \det Q_*\exp\left(\int_0^1\lambda^{-1}(s)\big(\mathrm{tr}(\hat{K}_y(s)\hat{Q}_y) + 2\nabla\cdot b(\hat{\phi}_y(s)) + \mathrm{tr}(a\hat{Q}_y^{-1})\big)\right) \tag{3.46}$$

that is

$$\begin{aligned}&|\det\hat{Q}_y(1)|^{-1/2}\exp\left(\frac{1}{2}\int_0^1\lambda^{-1}(s)\mathrm{tr}(\hat{K}_y(s)\hat{Q}_y)\right)\\ &\quad= |\det Q_*|^{-1/2}\exp\left(-\int_0^1\lambda^{-1}(s)\Big(\nabla\cdot b(\hat{\phi}_y(s)) + \tfrac{1}{2}\mathrm{tr}(a\hat{Q}_y^{-1})\Big)\right)\end{aligned} \tag{3.47}$$

This shows that the two expressions for $\bar{\rho}_\varepsilon(y)$ in (3.39) are identical. □

Remark 3.3. Note that this is consistent with the intuitive interpretation that $\hat{Q}$ quantifies the covariance of fluctuations around the instanton. For $T\to\infty$ the fluctuations will "thermalize" around the fixed point, which corresponds to considering the linearized (Ornstein-Uhlenbeck) dynamics around x_*, the covariance of which solves the Lyapunov equation (3.34).

Remark 3.4. We seemingly encounter a practical problem with the forward Riccati equation (3.35) at $s=0$, as also discussed in [10]: Since the instanton remains at the fixed-point x_* for an infinite amount of time, we have $\lambda(0)=0$, as well as $\hat{\theta}_y(0)=0$, and thus the forward Riccati equation (3.35) reads $0=0$ at $s=0$—a similar issue arises with the equation for $\hat{\phi}_y$ in (3.33) and is discussed in Appendix A. This problem is only apparent however and the limit of $\hat{Q}_y'(s)$ as $s\to 0$ can be straightforwardly obtained by writing (3.35) as

$$\hat{Q}_y' = \lambda^{-1}\big(\hat{Q}_y\hat{K}_y\hat{Q}_y + \hat{Q}_y(\nabla b(\hat{\phi}_y))^\top + (\nabla b(\hat{\phi}_y))\hat{Q}_y + a\big), \tag{3.48}$$

and sending $s\to 0$ at both sides using l'Hôpital rule to compute the limit of the right hand side. Accounting for the fact that $\hat{Q}_y(0)=Q_*$ and $\hat{K}_y(0)=0$ since $\hat{\theta}_y(0)=0$, the result is the following equation for $\hat{Q}_y'(0)$:

$$\begin{aligned}\hat{Q}_y'(0) = [\lambda'(0)]^{-1}\Big(&Q_*(\nabla\nabla b(x_*)\hat{\theta}_y'(0))Q_* + \hat{Q}_y'(0)(\nabla b(x_*))^\top + (\nabla b(x_*))\hat{Q}_y'(0)\\ &+ Q_*(\nabla\nabla b(x_*)\hat{\phi}_y'(0))^\top + (\nabla\nabla b(x_*)\hat{\phi}_y'(0))Q_*\Big).\end{aligned} \tag{3.49}$$

All the quantities in this equation except $\hat{Q}_y'(0)$ are available from the solution of the geometric instanton equation (3.33) as well as the Lyapunov equation (3.34) for Q_*, and (3.49) can also be

written as a Lyapunov equation

$$\mathfrak{C}\,\hat{Q}_y'(0) + \hat{Q}_y'(0)\,\mathfrak{C}^\top = \mathfrak{K}, \tag{3.50}$$

where we defined

$$\mathfrak{C} = \tfrac{1}{2}\lambda'(0)\mathrm{Id} - \nabla b(x_*)\mathfrak{K} = Q_*(\nabla\nabla b(x_*)\hat{\theta}_y'(0))Q_* + Q_*(\nabla\nabla b(x_*)\hat{\phi}_y'(0))^\top + (\nabla\nabla b(x_*)\hat{\phi}_y'(0))Q_*. \tag{3.51}$$

Since $\lambda'(0) > 0$ and $-\nabla b(x_*)$ is positive-definite as x_* is a stable fixed point of $\dot{x} = b(x)$ by assumption, the matrix $\mathfrak{C}$ is positive-definite and invertible, meaning that (3.50) has a unique solution $\hat{Q}_y'(0)$. Once $\hat{Q}_y'(0)$ has been calculated, one can initialize the integration of $\hat{Q}_y(s)$ with arc-length stepsize $\Delta s > 0$ using for example

$$\hat{Q}_y(\Delta s) = Q_* + \Delta s\hat{Q}_y'(0) + O(\Delta s^2). \tag{3.52}$$

For the examples presented in Sections 3.2.2 or 4.2.1, it turns out that $\nabla\nabla b(x_*) = 0$ and thus $Q_y'(0) = 0$. Therefore, approximating $\hat{Q}_y(\Delta s)$ by Q_* is correct to $\mathcal{O}(\Delta s^2)$, which is precise enough for the numerical scheme we employ, and is therefore what we used in practice. However, we note that it is straighforward to go to higher order if necessary: for example, the derivation to obtain $\hat{Q}_y''(0)$ when $\hat{Q}_y'(0) = 0$ is given in appendix A.2.

3.2.1 | Example: Invariant measure of gradient diffusions

Similar to the expectations for gradient diffusion, discussed in Section 2.3.1, we can also derive a formula for the small ε approximation of the invariant density of the gradient diffusion process (2.48). The result is known to be

$$\bar{\rho}_\varepsilon(y) = (2\pi\varepsilon)^{-n/2}(\det H(x_*))^{1/2}\exp(-\varepsilon^{-1}U(y)), \tag{3.53}$$

where the only approximation made is on the prefactor Z_ε that can be evaluated by Laplace's method (using $U(x_*) = 0$): $Z_\varepsilon = (2\pi\varepsilon)^{n/2}((\det H(x_*))^{-1/2} + \mathcal{O}(\varepsilon))$. Let us show that the small ε approximation of Proposition 3.3 recovers this result, including the normalization factor.

First, the instanton contribution yields

$$\begin{aligned} \lambda\hat{\phi}_y' &= 2\hat{\theta}_y - \nabla U(\hat{\phi}_y), & \hat{\phi}_y(0) &= x_*, \\ \lambda\hat{\theta}_y' &= H(\hat{\phi}_y)\theta_y, & \hat{\phi}_y(1) &= y, \end{aligned} \tag{3.54}$$

which is solved by

$$\lambda\hat{\phi}_y'(s) = \nabla U(\hat{\phi}_y(s)) = \hat{\theta}_y(s) \tag{3.55}$$

so that the exponential large deviation contribution is given by

$$\hat{I}(y) = \int_0^1 \lambda^{-1}(s)|\hat{\theta}_y|^2(s)\,ds = \int_0^1 \langle\nabla U(\hat{\phi}(s)), \hat{\phi}'(s)\rangle\,ds = U(y), \tag{3.56}$$

which recovers the exponential part of $\bar{\rho}_\varepsilon(y)$. If we additionally notice that for a gradient system we know explicitly the transverse decomposition to be trivial, $l(x) = 0$ and thus $\hat{L}(s, y) = 0$, the second form of (3.39) in proposition 3.3 immediately yields the limiting density of (3.53).

It is instructive to recover the same result by explicitly solving the Riccati equation. Notice that $\hat{K}_y(s) = \nabla\nabla\langle b(\hat{\phi}_y(s), \hat{\theta}_y(s)\rangle = -\lambda(s)H'(\hat{\phi}_y(s))$, so that forward Riccati equation (3.35) reduces to

$$\lambda\hat{Q}_y' = -\lambda\hat{Q}_y H'\hat{Q}_y - H\hat{Q}_y - \hat{Q}_y H + 2\mathrm{Id}, \qquad \hat{Q}_y(0) = Q_*. \tag{3.57}$$

We can directly solve the Lyapunov equation (3.34) for Q_*,

$$Q_* = \hat{Q}_y(0) = H(x_*)^{-1}. \tag{3.58}$$

Given this fact, note that Equation (3.57) is solved by

$$\hat{Q}_y(s) = H^{-1}(\hat{\phi}_y(s)). \tag{3.59}$$

Therefore, we can read off

$$|\det \hat{Q}_y(1))|^{-1/2} = |\det H(y))|^{1/2}. \tag{3.60}$$

and

$$\begin{aligned}&\exp\left(\frac{1}{2}\int_0^1 \lambda^{-1}(s)\mathrm{tr}(\hat{K}_y(s)\hat{Q}_y(s))ds\right)\\ &= \exp\left(-\frac{1}{2}\int_0^1 \mathrm{tr}(H^{-1}(\hat{\phi}_y(s))H'(\hat{\phi}_y(s))ds\right) = \left|\frac{\det \hat{Q}_y(0)}{\det \hat{Q}_y(1)}\right|^{1/2} = \left|\frac{\det H(y)}{\det H_*}\right|^{1/2}\end{aligned} \tag{3.61}$$

Combining these results we indeed recover exactly the limiting density given in (3.53).

3.2.2 | Example: Invariant measure of a nonlinear, irreversible process in $\mathbb{R}^2$

For all above examples, analytical results were available from case-specific calculations, since the systems are very simple. The easiest example of a system where no analytical results can be easily derived is a system which is nonlinear, and further the drift is not given by a gradient of a potential. Since the latter is always the case in 1D, we need to consider a state space of at least dimension two. In this case, in order to compute expectations, probability densities, or probabilities with our approach, we need to solve the corresponding equations, both instanton equation and Riccati equation, by numerical means.

As a concrete example, consider the non-gradient drift given by

$$b(x_1, x_2) = \left(-\alpha x_1 - \gamma x_1^3 + \beta x_2, -\alpha x_2 - \gamma x_2^3 - \beta x_1\right). \tag{3.62}$$

For positive $\alpha, \gamma \in \mathbb{R}$, this drift corresponds to a nonlinear attractive force towards the fixed-point $x_* = (0, 0)$. The parameter $\beta \neq 0$ adds a swirl on the drift, so that the total dynamics are

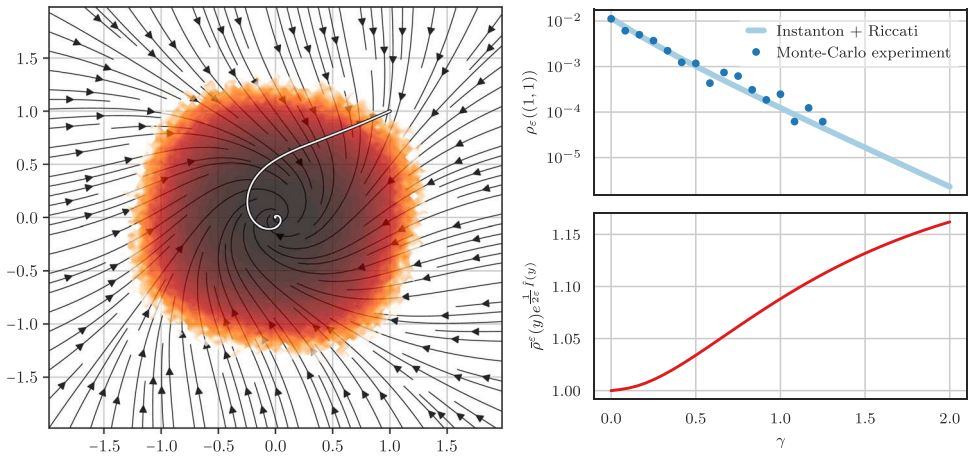

FIGURE 1 Left: Instanton (solid curve) connecting the fixed point $x_* = (0,0)$ to the point (1,1). The shaded contours indicate the invariant measure, obtained from sampling the process for long times. The flow field depicts the drift $b(x)$. Here, $\alpha = 0.5, \beta = 1, \gamma = 1, \varepsilon = 0.25$. Top right: Comparison of the pdf $\rho_\varepsilon((1,1))$ between the computation via proposition 3.3 (light blue solid) and random sampling of the stochastic process (dark blue dots) for varying $\gamma \in [0,2]$. The pdf is reproduced at every point from instanton and fluctuations, including its normalization factor. Bottom right: Prefactor of the invariant density, $\bar{\rho}^\varepsilon(y)e^{\frac{1}{2}\varepsilon^{-1}\hat{I}(y)}$, at $y = (1,1)$ as a function of the nonlinearity parameter γ, obtained by numerically integrating the Riccati equation along the instanton using proposition 3.3. Clearly, the strength of the nonlinearity influences the prefactor. The other parameters are again $\alpha = 0.5, \beta = 1$.

no longer gradient. The behavior of this drift can be seen by the streamlines in Figure 1 (left). Note in particular that the system is not rotationally symmetric in the presence of cubic terms (i.e. $\gamma \neq 0$).

In order to estimate numerically the density $\bar{\rho}^\varepsilon(y)$ defined in (3.39), we need to:

(1) compute the instanton $\hat{\phi}_y(s)$ connecting (0,0) to y, as well as the parametrization $\lambda(s)$, using gMAM; and
(2) solve the forward Riccati equation (3.35).

The results of this procedure are shown in Figure 1. For a given arbitrary endpoint (1,1), we can obtain the invariant density by sampling the process for long times, and approximating the density by a normalized histogram. The result is depicted as shaded contours on the left. Note that for this procedure, not only do we need many samples hitting close to the point (1,1), but furthermore statistics everywhere else in state space, because these determine the normalization constant. The instanton connecting (0,0) to (1,1) is depicted here as a solid line, while the drift is given by the flowlines.

We now want to establish how the prefactor contribution, $\bar{\rho}^\varepsilon(y)e^{\frac{1}{2}\varepsilon^{-1}\hat{I}(y)}$, changes when changing the parameter γ, which determines the strength of the nonlinear term. In particular, for the linear case $\gamma = 0$, we know that $\bar{\rho}^\varepsilon(y)e^{\frac{1}{2}\varepsilon^{-1}\hat{I}(y)} = 1 \; \forall \, y \in \mathbb{R}^2$, since the invariant measure is Gaussian, and its covariance can be computed by solving a Lyapunov equation. For the nonlinear case, $\gamma > 0$, this is no longer possible, and we need to resort to numerical results instead. In the right panel of Figure 1 we show the computation of the prefactor through proposition 3.3, computing the

instanton and solving the Riccati equation. For $\gamma = 0$, this reproduces the known linear result, but we obtain values for finite values of γ as well. These results are in agreement with results from sampling the whole process, and looking at the normalized histogram at the point (1,1) for different $\gamma \in [0, 2]$. The numerical parameters here are $\alpha = 0.5, \beta = 1, \gamma \in [0, 2], \varepsilon = 0.25$, and the histogram binning is done with square bins of side-lengths $\Delta x = 0.02$ and with 10^7 samples per value of γ.

3.3 | Calculation of expectations revisited

We can use Proposition 3.1 to give an expression alternative to that in Proposition 2.1 for finite-time expectations:

Proposition 3.4. *Let $(\phi_x(t), \theta_x(t))$ solve the instanton equations in (2.2), and $Q_x(t)$ be the solution to the forward Riccati equation*

$$\dot{Q}_x = Q_x K_x Q_x + Q_x(\nabla b(\phi_x))^\top + (\nabla b(\phi_x))Q_x + a, \qquad Q_x(0) = 0, \tag{3.63}$$

where we denote $K_x(t) = \nabla\nabla\langle b(\phi_x(t)), \theta_x(t)\rangle$. Then the factor $R(T, x)$ defined in (2.6) of Proposition 2.1 can also be expressed as

$$R(T, x) = |\det(Id - \nabla\nabla f(\phi_x(T))Q_x(T))|^{1/2} \exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(K_x(t)Q_x(t))\,dt\right). \tag{3.64}$$

Proof. We can arrive at expression (3.64) for $R(T, x)$ in two ways. We can use Proposition C.1 with $W_x(t)$ solution to (2.3) and $Q_x(t)$ solution to (3.63) to deduce from (C.2) that

$$|\det(\mathrm{Id} - \nabla\nabla f(\phi_x(T))Q_x(T))|^{1/2} \exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(K_x(t)Q_x(t))\,dt\right) = \exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(aW_x(t)\,dt\right) \tag{3.65}$$

The left hand side is (3.64) and the right hand side is (2.6), showing that these two equations contain identical expressions. Alternatively, we can evaluate

$$\int_{\mathbb{R}^n} \exp(\varepsilon^{-1} f(x))\bar{\rho}^\varepsilon(T, x)dx \tag{3.66}$$

by Laplace's method, using the expression in (3.7) for $\bar{\rho}^\varepsilon(T, x)$. This calculation show that this integral is asymptotically equivalent to $\bar{A}^\varepsilon(T, x)$ in (2.5) with $R(T, x)$ given by (3.64) instead of (3.64), thereby establishing again that these two equations give the same $R(T, x)$ since this factor is independent on ε. □

Similarly, we can use Proposition 3.3 to give an expression alternative to that in Proposition 2.3 for expectations on the invariant measure:

Proposition 3.5. *Let $(\hat{\phi}(s), \hat{\theta}(s))$ solve the instanton equations in (2.36), and $\hat{Q}(s)$ be the solution to the forward Riccati equation*

$$\lambda(s)\hat{Q}' = \hat{Q}\hat{K}\hat{Q} + \hat{Q}(\nabla b(\hat{\phi}))^\top + (\nabla b(\hat{\phi}))\hat{Q} + a, \qquad \hat{Q}(0) = Q_*. \tag{3.67}$$

where Q_* solves (3.34) and we denote $\hat{K}(s) = \nabla\nabla\langle b(\hat{\phi}(s)), \hat{\theta}(s)\rangle$. Then the factor $\hat{R}$ defined in (2.40) of Proposition 2.3, can also be expressed as

$$\hat{R} = |\det(Id - \nabla\nabla f(\hat{\phi}(1))\hat{Q}(1))|^{1/2} \exp\left(\frac{1}{2}\int_0^1 \lambda^{-1}(s)\mathrm{tr}(\hat{K}(s)\hat{Q}(s))\, ds\right). \tag{3.68}$$

The proof of this proposition is similar to that of Proposition 3.4.

4 | PROBABILITIES

4.1 | Probabilities at finite time

Having access to pointwise estimates of the probability density allows us to estimate probabilities by integration. For example, let $f : \mathbb{R}^n \to \mathbb{R}$ be some observable, and assume we want to compute the probability that $f(X_T^\varepsilon)$ exceeds some value $a \in \mathbb{R}$. This probability can be expressed as

$$\mathbb{P}^x(f(X_T^\varepsilon) \geq a) = \int_{f(y)\geq a} \rho_\varepsilon^x(T, y)\, dy, \tag{4.1}$$

and calculating for various values of a gives the complementary cumulative distribution function (aka tail distribution) of the random variable $f(X_T^\varepsilon)$. For concreteness, we will focus on the calculation of this tail probability for one value of a which, without loss of generality, we can set to zero. In this case, (4.1) can also be interpreted as the probability that the stochastic process (1.1) hits the set

$$A = \{z \in \mathbb{R}^n | f(z) \geq 0\}, \tag{4.2}$$

at time $t = T$. This is a problem interesting in its own right, and we will denote this probability as

$$P_\varepsilon^A(T, x) = \mathbb{P}^x(X_T^\varepsilon \in A) = \int_A \rho_\varepsilon^x(T, y)\, dy. \tag{4.3}$$

To make the problem interesting, we will work under:

Assumption 4.1. The function $f : \mathbb{R}^n \to \mathbb{R}$ is in C^2, $x \notin A$ (i.e. $f(x) < 0$), and the vector field b points outward A everywhere on ∂A (i.e. $\langle \nabla f(z), b(z)\rangle < 0 \; \forall \; z \in \partial A = \{z \in \mathbb{R}^n | f(z) = 0\}$).

This assumption implies that the event X_T^ε is noise-driven, and as a result $P_\varepsilon^A(T, x) \to 0$ as $\varepsilon \to 0$, which is the nontrivial case. We also need

Assumption 4.2. The point on ∂A with minimal $I_x(T,z)$,

$$y = \operatorname*{argmin}_{z\in\partial A} \int_0^T \langle \theta_{x,z}(t), a\theta_{x,z}(t)\rangle\, dt \tag{4.4}$$

is unique, and so is the instanton leading to it. In addition y is not a critical point of f.

Intuitively we demand that the set A does not admit multiple distinct points that can be reached by competing instantons of identical action. Then, we have:

Proposition 4.3. *Let $(\phi_{x,y}(t), \theta_{x,y}(t))$ solve the instanton equations in (2.2) with y specified as*

$$y = \operatorname*{argmin}_{z\in\partial A} \int_0^T \langle \theta_{x,z}(t), a\theta_{x,z}(t)\rangle\, dt, \tag{4.5}$$

and let $Q_{x,y}(t)$ solve the forward Riccati equation in (3.3). Let also $F(T,x)$ and $V(T,x)$ be defined as

$$\begin{aligned} F(T,x) &= Id - |\theta_{x,y}(T)||\nabla f(y)|^{-1} Q_{x,y}(T)\nabla\nabla f(y), \\ V(T,x) &= \langle \theta_{x,y}(T), Q_{x,y}(T)\theta_{x,y}(T)\rangle^{-1/2} \exp\left(\frac{1}{2}\int_0^T \operatorname{tr}(K_{x,y}(t)Q_{x,y}(t))\, dt\right), \end{aligned} \tag{4.6}$$

where $K_{x,y}(t) = \nabla\nabla\langle b(\phi_{x,y}(t)), \theta_{x,y}(t)\rangle$. Finally denote by $\hat{n} = \nabla f(y)/|\nabla f(y)|$ the inward pointing surface normal of A at y. Then the probability $P_\varepsilon^A(T,x)$ satisfies

$$\lim_{\varepsilon\to 0} \frac{P_\varepsilon^A(T,x)}{\bar{P}_\varepsilon^A(T,x)} = 1, \tag{4.7}$$

in which

$$\bar{P}_\varepsilon^A(T,x) = (2\pi)^{-1/2}\varepsilon^{1/2}\left(\frac{\langle \hat{n}, F(T,x)\hat{n}\rangle}{\det F(T,x)}\right)^{1/2} V(T,x)\exp\left(-\frac{1}{2\varepsilon}\int_0^T \langle \theta_{x,y}(t), a\theta_{x,y}(t)\rangle\, dt\right). \tag{4.8}$$

In addition , the unit normal $\hat{n}$ can be expressed as $\hat{n} = \theta_{x,y}(T)/|\theta_{x,y}(T)|$.

The situation of Proposition 4.3 is depicted in Figure 2.

Remark 4.1. As we will see in the proof, the factor involving $F(T,x)$ describes the effect of the geometry (specifically, the curvature) of the set A around the maximum likelihood hitting point y: If the set is planar, then $\nabla\nabla f(y) = 0$ and the factor evaluates to unity. Equivalently, this factor disappears for linear observables f. As will also be clear from the proof, in the definition of $F(T,x)$ in (4.6) we can replace $|\nabla f(y)|^{-1}\nabla\nabla f(y)$ by $\nabla\hat{n}(y)$, that is in (4.8) we can replace $F(T,x)$ with

$$F_\perp(T,x) = \mathrm{Id} - |\theta_{x,y}(T)| Q_{x,y}(T)\nabla\hat{n}(y) \tag{4.9}$$

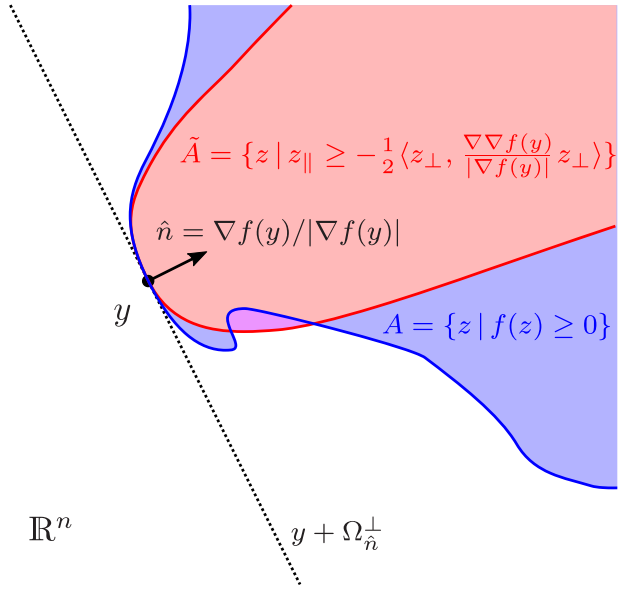

FIGURE 2 Schematic representation of the situation in proposition 4.3. To obtain the probability to hit the set A, we replace the integral of the density over A with an integral over the paraboloid $\tilde{A}$, which is tangential A at the point of maximum likelihood y and shares its curvature.

This is because $\nabla\hat{n} = |\nabla f|^{-1}\nabla\nabla f + |\nabla f)|^{-1}\hat{n}(\nabla\nabla f\hat{n})^T$ and the extra factor $|\nabla f|^{-1}\hat{n}(\nabla\nabla f\hat{n})^T$ does not contribute to the ratio $\langle\hat{n}, F(T,x)\hat{n}\rangle|\det F(T,x)|^{-1}$ in (4.8). This shows that this expression is intrinsic to the set A, that is it does not depend on the way we parametrize its boundary by the zero level-set of f, as it should be. For computational purposes, using the parametrized version in (4.6) is more convenient, however.

Proof. Let us evaluate the integral in (4.3) by Laplace's method using the ansatz (3.14) for $\rho_{\varepsilon}^{x}(T,y)$. To this end notice that the point y specified in (4.5) is also given by

$$y = \operatorname*{argmin}_{z\in\partial A} I_x(T,z) \in \partial A. \tag{4.10}$$

Notice also that $\theta_{x,y}(T) = \nabla_y I_x(T,y)$, and that $\hat{n} = \theta_{x,y}(T)/|\theta_{x,y}(T)|$ is the inward pointing unit normal to ∂A at y. Therefore, to perform the integral

$$\bar{P}_{\varepsilon}^{A}(T,x) = \int_A \bar{\rho}_{\varepsilon}^{x}(T,z)\,dz, \tag{4.11}$$

we split the integration variable into components parallel to $\hat{n}$ and perpendicular to $\hat{n}$, as

$$z - y = \varepsilon z_{\|}\hat{n} + \sqrt{\varepsilon}z_{\perp}, \tag{4.12}$$

where $z_{\|} \in \mathbb{R}$ and $z_{\perp} \in \Omega_{\hat{n}}^{\perp}$ with

$$\Omega_{\hat{n}}^{\perp} = \{z \in \mathbb{R}^d \,:\, \langle\hat{n}, z\rangle = 0\}. \tag{4.13}$$

For any $z \in A$ we have $f(z) \geq 0$. Further using $f(y) = 0$, we can expand f around y to obtain

$$f(z) = f(y) + \sqrt{\varepsilon}\langle z_\perp, \nabla f(y)\rangle + \varepsilon z_\parallel \langle \hat{n}, \nabla f(y)\rangle + \frac{\varepsilon}{2}\langle z_\perp, \nabla\nabla f(y) z_\perp\rangle + \mathcal{O}(\varepsilon^{3/2}) \geq 0. \tag{4.14}$$

Since $\langle z_\perp, \nabla f(y)\rangle = 0$ by definition, we have for points in A around y that

$$z_\parallel \geq -\frac{1}{2}\langle z_\perp, \frac{\nabla\nabla f(y)}{|\nabla f(y)|} z_\perp\rangle. \tag{4.15}$$

Effectively, to the relevant order in ε, A can be approximated by a paraboloid

$$\tilde{A} = \{z \in \mathbb{R}^n \,|\, z_\parallel \geq -\frac{1}{2}\langle z_\perp, |\nabla f(y)|^{-1}\nabla\nabla f(y) z_\perp\rangle\}, \tag{4.16}$$

where $\nabla\nabla f(y)/|\nabla f(y)|$ is the (normalized) curvature of the set A at y.

We can now evaluate the integral in (4.3) by Laplace's method,

$$\begin{aligned}
\bar{P}_\varepsilon^A(T,x) &= (2\pi\varepsilon)^{-n/2} R_\varepsilon(y) \exp(-\varepsilon^{-1} I_x(T,y))\, \varepsilon^{(n-1)/2} \int_{\Omega_{\hat{n}}^\perp} e^{-\frac{1}{2}\langle z_\perp, \nabla\nabla I_x(T,y) z_\perp\rangle} \\
&\quad \times \varepsilon \int_{-\frac{1}{2}\langle z_\perp, |\nabla f(y)|^{-1}\nabla\nabla f(y) z_\perp\rangle}^{\infty} e^{-z_\parallel |\nabla I_x(T,y)|} dz_\parallel\, dz_\perp \\
&= (2\pi)^{-1/2}\varepsilon^{1/2} R_\varepsilon(y) \exp(-\varepsilon^{-1} I_x(T,y)) |\nabla I_x(T,Y)|^{-1} \\
&\quad \times \int_{\Omega_{\hat{n}}^\perp} e^{-\frac{1}{2}\langle z_\perp, \nabla\nabla I_x(T,y) z_\perp\rangle + \frac{1}{2}\langle z_\perp, |\nabla f(y)|^{-1} |\nabla I_x(T,y)| \nabla\nabla f(y) z_\perp\rangle} dz_\perp \\
&= (2\pi)^{-1/2}\varepsilon^{1/2} R_\varepsilon(y) |D(T,y)|^{-1/2} |\theta_{x,y}(T)|^{-1} \exp(-\varepsilon^{-1} I_x(T,y)),
\end{aligned} \tag{4.17}$$

Here

$$\begin{aligned}
R_\varepsilon(y) &= \left(|\det Q_{x,y}(T)|^{-1/2} \exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(K_{x,y}(t) Q_{x,y}(t)))\, dt \right) + \mathcal{O}(\varepsilon) \right) \\
&= \left(|\det \nabla\nabla I_x(T,y)|^{1/2} \exp\left(\frac{1}{2}\int_0^T \mathrm{tr}(K_{x,y}(t) Q_{x,y}(t)))\, dt \right) + \mathcal{O}(\varepsilon) \right)
\end{aligned} \tag{4.18}$$

and

$$D(T,y) = \det{}_\perp \left(\nabla\nabla I_x(T,y) - |\nabla f(y)|^{-1} |\nabla I_x(T,y)| \nabla\nabla f(y) \right) \tag{4.19}$$

where $\det_\perp$ is defined as follows: Given an invertible, positive definite, symmetric $H \in \mathbb{R}^{n\times n}$ and a unit vector $\hat{n}$ we define $\det_\perp H$ via

$$(2\pi)^{(n-1)/2} |\det{}_\perp H|^{-1/2} = \int_{\Omega_{\hat{n}}^\perp} e^{-\frac{1}{2}\langle y, Hy\rangle} dy. \tag{4.20}$$

In other words, $\det_\perp$ is the determinant evaluated only in the space $\Omega^\perp_{\hat{n}}$ perpendicular to a vector $\hat{n}$. As shown in Appendix E we have

$$\det_\perp H = \langle \hat{n}, H^{-1}\hat{n}\rangle \det H, \tag{4.21}$$

so that

$$\begin{aligned} &|\det \nabla\nabla I_x(T,y)|^{1/2}|D(T,y)|^{-1/2} \\ &= \left|\det_\perp \left(1 - |\nabla f(y)|^{-1}|\nabla I_x(T,y)|Q_{x,y}(T)\nabla\nabla f(y)\right)\right|^{-1/2} \langle \hat{n}, Q_{x,y}(T)\hat{n}\rangle^{-1/2} \end{aligned} \tag{4.22}$$

and thus, inserting R_ε in (4.17), this equation gives (4.7). Finally, note that by definition of $\det_\perp$ we can replace $|\nabla f(y)|^{-1}\nabla\nabla f(y)$ by $\nabla\hat{n}(y)$ in (4.19), which justifies the alternative expression in (4.9) for $F(T,x)$. □

4.2 | Probabilities on the invariant measure

The equivalent of (4.3) at infinite time is

$$P^A_\varepsilon = \int_A \rho_\varepsilon(y)\, dy. \tag{4.23}$$

where $\rho_\varepsilon(y)$ is the invariant density defined by (1.8) and the set A is defined as before. Then we have:

Proposition 4.4. *Let $(\hat{\phi}_y(s), \hat{\theta}_y(s))$ solve the instanton equations in (3.33) with y specified as*

$$y = \operatorname*{argmin}_{\hat{\phi}_y(1)\in\partial A} \int_0^1 \lambda^{-1}(s)\langle \hat{\theta}_y(s), a\hat{\theta}_y(s)\rangle\, ds; \tag{4.24}$$

and let $\hat{Q}_y(s)$ solve the forward Riccati equation in (3.35). Let also $\hat{F}$ and $\hat{V}$ be given by

$$\begin{aligned} \hat{F} &= Id - |\hat{\theta}_y(1)||\nabla f(y)|^{-1}\hat{Q}_y(1)\nabla\nabla f(y), \\ \hat{V} &= \langle \hat{\theta}_y(1), \hat{Q}_y(1)\hat{\theta}_y(1)\rangle^{-1/2} \exp\left(\frac{1}{2}\int_0^1 \lambda^{-1}(s)\mathrm{tr}(\hat{K}_y(s)\hat{Q}_y(s))\, dt\right), \end{aligned} \tag{4.25}$$

where $\hat{K}_y(s) = \nabla\nabla\langle b(\hat{\phi}_y(s)), \hat{\theta}_y(s)\rangle$. Finally denote by $\hat{n} = \nabla f(y)/|\nabla f(y)|$ the inward pointing surface normal of A at y. Then the probability P^A_ε satisfies

$$\lim_{\varepsilon\to 0} \frac{P^A_\varepsilon}{\bar{P}^A_\varepsilon} = 1, \tag{4.26}$$

where

$$\bar{P}^A_\varepsilon = (2\pi)^{-1/2}\varepsilon^{1/2}\left(\frac{\langle \hat{n}, \hat{F}\hat{n}\rangle}{\det \hat{F}}\right)^{1/2} \hat{V} \exp\left(-\frac{1}{2\varepsilon}\int_0^1 \lambda^{-1}(s)\langle \hat{\theta}_y(s), a\hat{\theta}_y(s)\rangle\, ds\right). \tag{4.27}$$

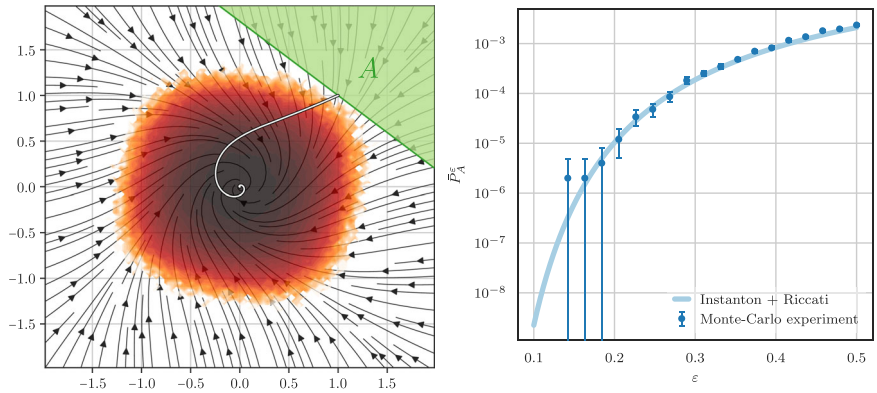

FIGURE 3 Left: Instanton (solid curve) connecting the fixed point $x_* = (0,0)$ to the set A, with minimal action at boundary point (1,1). The shaded contours indicate the invariant measure, obtained from sampling the process for long times, and the set A is depicted in green. The flow field depicts the drift $b(x)$. Here, $\varepsilon = 0.25$. Right: Comparison of the probability to hit the set A between the computation via proposition 4.4 (light blue solid) and random sampling of the stochastic process (dark blue dots) for varying ε. The probability is reproduced at every point from instanton and fluctuations, including its normalization factor. The parameters are $\alpha = 0.5, \beta = 1, \gamma = 0.5$.

In addition , the unit normal $\hat{n}$ can be expressed as $\hat{n} = \hat{\theta}_y(1)/|\hat{\theta}_y(1)|$.

The proof is similar to that of Proposition 4.3.

4.2.1 | Example: Nonlinear, irreversible process in $\mathbb{R}^2$ revisited

To show the applicability of these propositions to an actual system, we re-use the nonlinear, irreversible process in $\mathbb{R}^2$, as defined in Equation (3.62), as a simple example for which the solution is not readily accessible by analytical considerations. Instead of computing the invariant density, as in Section 3.2.2 we use proposition 4.4 to compute the probability of hitting the set A on the invariant measure, where A is the half-space defined by

$$A = \left\{ x \in \mathbb{R}^2 | \langle \hat{n}, x - (1,1) \rangle \geq 0 \right\}, \tag{4.28}$$

with $\hat{n} \approx (0.6304, 0.7762)$ chosen specifically such that the point $y = (1,1)$ is the maximum likelihood point on ∂A. Because of this, the solutions to the instanton equations and the Riccati equation do not need to be re-computed, and we can insert their results into equation (4.27) to obtain the instanton and prefactor estimate for the hitting probability on the invariant measure.

The results of this experiment are shown in Figure 3: While the dynamics and instanton in Figure 3 (left) look identical to the original problem in Section 3.2.2, we are now trying to estimate the probability of hitting the set A denoted by the light green shading. The result shown in the right panel of Figure 3 confirms that the asymptotic prediction of proposition 4.4 agrees with the Monte-Carlo simulations for different values of ε. The parameters are $\alpha = 0.5, \beta = 1, \gamma = 0.5$, and we are taking $N_{\text{samples}} = 10^6$ samples each value of ε.

5 | EXIT PROBABILITIES AND MEAN FIRST PASSAGE TIMES

Let $A^c \subset \mathbb{R}^n$ be the complement of the set A defined in (4.2) that is

$$A^c = \{z \in \mathbb{R}^n \mid f(z) < 0\} \tag{5.1}$$

Under Assumption 4.1, we know from the argument given after Proposition 3.3 that

$$\lim_{\varepsilon \to 0} \int_{A^c} \rho_\varepsilon(y) dy = \lim_{\varepsilon \to 0} \int_{A^c} \bar{\rho}_\varepsilon(y) dy = 1 \tag{5.2}$$

We wish to estimate the exit probability

$$P_\varepsilon^{\partial A} = \int_{\partial A} \rho_\varepsilon(y) d\sigma(y) \tag{5.3}$$

in the limit as $\varepsilon \to 0$, since this surface integral enters the limiting expression for the mean first passage time (MFPT) of the process to set A. We have

Proposition 5.1. *Let $(\hat{\phi}_y(s), \hat{\theta}_y(s))$ solve to the instanton equations in (3.33) with y specified as*

$$y = \underset{z \in \partial B}{\operatorname{argmin}} \int_0^1 \lambda^{-1}(s) \langle \hat{\theta}_z(s), a\hat{\theta}_z(s) \rangle \, ds; \tag{5.4}$$

and let $\hat{F}$ and $\hat{V}$ be given by (4.25). Denote by $\hat{n} = \nabla f(y)/|\nabla f(y)|$ the inward pointing unit normal vector on ∂A at y. Then the exit probability $P_\varepsilon^{\partial A}$ satisfies

$$\lim_{\varepsilon \to 0} \frac{P_\varepsilon^{\partial A}}{\bar{P}_\varepsilon^{\partial A}} = 1, \tag{5.5}$$

where

$$\bar{P}_\varepsilon^{\partial A} = (2\pi\varepsilon)^{-1/2} |\hat{\theta}_y(1)| \left(\frac{\langle \hat{n}, \hat{F}\hat{n} \rangle}{\det \hat{F}} \right)^{1/2} \hat{V} \exp\left(-\frac{1}{2\varepsilon} \int_0^1 \lambda^{-1}(s) \langle \hat{\theta}_y(s), a\hat{\theta}_y(s) \rangle \, ds \right) \tag{5.6}$$

Proof. The proof has the same ingredients as the proof of proposition 4.4, except instead of integrating over the paraboloid approximation $\tilde{A}$, we integrate over its boundary $\partial\tilde{A}$. No normal integral is performed and therefore the scaling in ε and $|\hat{\theta}_y(1)|$ differs. □

With this knowledge, we can now compute exit times. Consider the entrance into the domain A for a particle starting at $x \in A^c$. The mean exit time $T_\varepsilon : A^c \to \mathbb{R}^+$ is given by the solution of the problem

$$\begin{aligned} L_\varepsilon T_\varepsilon &= -1, \qquad \text{if } x \in A^c, \\ T_\varepsilon &= 0, \qquad \text{if } x \in \partial A, \end{aligned} \tag{5.7}$$

where L_ε is the generator of our process defined in (2.13). In the following, we quickly review the standard approach to use a boundary layer expansion in order to approximate the exit time T_ϵ. Multiplying the first equation by the invariant measure ρ and integrating we find—after applying Green's theorem and making use of the fact that $L_\varepsilon^* \rho_\varepsilon = 0$ [25]:

$$\epsilon \int_{\partial A} \rho_\varepsilon \langle \hat{n}, \nabla T_\epsilon \rangle \, d\sigma = - \int_{A^c} \rho_\varepsilon \, dx. \tag{5.8}$$

Here, $\hat{n}(z) = \nabla f(z)/|\nabla f(z)|$ denotes the inward pointing normal unit vector of the surface ∂A at $z \in \partial A$. An approximation of $\langle \hat{n}(z), \nabla T_\varepsilon(z) \rangle$ for $z \in \partial A$ can be obtained by boundary layer analysis. For this purpose, we first expand

$$T_\epsilon(x) = e^{C/\epsilon} \tau(x) \tag{5.9}$$

such that we find from (5.7) the corresponding equation for τ

$$L_\varepsilon \tau = e^{-C/\epsilon} \approx 0. \tag{5.10}$$

Since we assumed that $\langle \hat{n}(z), b(z) \rangle < 0$ for all $z \in \partial A$, the appropriate scaling of the boundary layer is $x = z - \epsilon \eta \hat{n}$. In the scaled variables, the equation for τ becomes

$$-\langle b(z), \hat{n}(z) \rangle \tau_\eta + \tau_{\eta\eta} = 0 \tag{5.11}$$

with the solution

$$\tau = \tilde{C} \left(1 - e^{\langle \hat{n}(z), b(z) \rangle \eta} \right) \tag{5.12}$$

This means that we obtain for the exit time T_ϵ the expression

$$T_\epsilon = \tilde{C} \, e^{C/\epsilon} \left(1 - e^{\langle \hat{n}(z), b(z) \rangle \eta} \right) \tag{5.13}$$

and therefore

$$\nabla T_\epsilon = \frac{\tilde{C}}{\epsilon} \, e^{C/\epsilon} \, \langle \hat{n}(z), b(z) \rangle \hat{n}(u) \tag{5.14}$$

which we can use in the solvability condition (5.8). From there, for x away from the boundary layer, we obtain

$$T_\epsilon(x) = \tilde{C} \, e^{C/\epsilon} = \frac{- \int_{A^c} \rho_\varepsilon(z) \, dz}{\int_{\partial A} \rho_\varepsilon(z) \langle \hat{n}(z), b(u) \rangle d\sigma(z)} \tag{5.15}$$

The following proposition shows how to estimate $T_\epsilon(x)$ in the limit as $\varepsilon \to 0$:

Proposition 5.2. *Let $(\hat{\phi}_y(s), \hat{\theta}_y(s))$ solve the instanton equations in (3.33) with y as specified in (5.4), and let $\hat{F}$ and $\hat{V}$ be given by (4.25). Denote by $\hat{n} = \nabla f(y)/|\nabla f(y)|$ the inward pointing unit normal*

vector on ∂A at y. Then for any $x \in A^c$, the mean first passage time $T_\varepsilon(x)$ satisfies

$$\lim_{\varepsilon\to 0} \frac{T_\varepsilon(x)}{\bar{T}_\varepsilon} = 1 \tag{5.16}$$

where

$$\bar{T}_\varepsilon = (2\pi\varepsilon)^{1/2} |\langle \hat{n}, b(y)\rangle|^{-1} |\hat{\theta}_y(1)|^{-1} \left(\frac{\langle \hat{n}, \hat{F}\hat{n}\rangle}{\det \hat{F}}\right)^{-1/2} \hat{V}^{-1} \exp\left(\frac{1}{2\varepsilon}\int_0^1 \lambda^{-1}(s)\langle \hat{\theta}_y(s), a\hat{\theta}_y(s)\rangle\, ds\right). \tag{5.17}$$

Proof. From (5.15) it follows with the use of proposition 5.1 that

$$\bar{T}_\varepsilon = -\left(\langle \hat{n}(y), b(y)\rangle \bar{P}_\varepsilon^{\partial B}\right)^{-1}, \tag{5.18}$$

where we additionally used that $\lim_{\varepsilon\to 0}\int_{A^c} \rho_\varepsilon(z)\, dz = 1$. □

Remark 5.1. Another interesting case is when we demand $\langle \hat{n}(z), b(z)\rangle = 0$ everywhere on ∂A, such that A^c corresponds exactly to the basin of attraction of the process. In this case, the situation is more complicated. Generally, one expects a different scaling of the form

$$T_\varepsilon(x) \sim C\varepsilon^{-1/2}\left(\bar{P}_\varepsilon^{\partial A}\right)^{-1}. \tag{5.19}$$

The underlying assumptions need to make sure that the quasipotential is twice differentiable at the exit point, which is true for gradient systems, but not generally true for an arbitrary drift. We refer to [9, 40] for details.

5.0.1 | Example: Ornstein-Uhlenbeck process

For the 1D Ornstein-Uhlenbeck process,

$$dX_t^\varepsilon = -\gamma X_t^\varepsilon\, dt + \sqrt{\epsilon}\, dW_t, \qquad X_0 = 0, \tag{5.20}$$

$X_t \in \mathbb{R}$, the formula for the MFPT is known exactly [48]. Concretely, the expected time for the process (5.20) to leave the set $A^c = [-\infty, z]$ is given by

$$T_\varepsilon = \frac{1}{\gamma}\sqrt{\frac{\pi}{2}}\int_0^{z\sqrt{2\gamma/\epsilon}} \left(1 + \operatorname{erf}\left(\frac{t}{\sqrt{2}}\right)\right)\exp\left(\frac{t^2}{2}\right) dt. \tag{5.21}$$

In the limit $\varepsilon \to 0$, truncating the prefactor at $\mathcal{O}(\varepsilon^{3/2})$, this yields the limiting result

$$\tilde{T}_\epsilon = \left(\frac{1}{z}\sqrt{\frac{\pi\epsilon}{\gamma^3}} + O(\epsilon^{3/2})\right) e^{\epsilon^{-1}\gamma z^2}. \tag{5.22}$$

Since necessarily $y = z$, and there is no perpendicular direction in 1D, Proposition 5.2 tells us that $\bar{T}_\varepsilon = C\exp(\varepsilon^{-1}\gamma z^2)$ with C given by

$$C = (2\pi\varepsilon)^{-1}(|b(y)||\hat{\theta}_y(1)|\hat{V}|)^{-1}, \tag{5.23}$$

Using

$$\hat{\theta}_y(1) = 2\gamma z, \qquad |b(y)| = \gamma z$$
$$\hat{Q}^* = (2\gamma)^{-1}, \qquad \hat{Q}_y(t) = (2\gamma)^{-1}$$
$$\hat{V} = \left(\hat{\theta}_y(1)^2\hat{Q}_y(1)\right)^{-1/2} = (2\gamma z^2)^{-1/2}$$

we obtain

$$C = \frac{1}{z}\sqrt{\frac{\pi\varepsilon}{\gamma^3}},$$

in agreement with the analytical result in (5.22).

5.0.2 | Example: Exit from a displaced circle

Consider instead the situation of $X_t \in \mathbb{R}^2$, but still

$$dX_t^\epsilon = -\gamma X_t^\epsilon\, dt + \sqrt{\epsilon}\, dW_t, \qquad X_0 = 0. \tag{5.24}$$

We are interested in the expected time to enter the complement of the translated circle [33] of radius r around $(z - r, 0)$, that is exit

$$A_r^c = \left\{x \in \mathbb{R}^2 \mid |x - (z - r, 0)| \le r\right\}. \tag{5.25}$$

We want to compare this result against the translated half-plane,

$$A_|^c = \left\{x \in \mathbb{R}^2 \mid x_1 \le z\right\}, \tag{5.26}$$

where x_1 is the first component of $x = (x_1, x_2)$. The most likely exit point, located at $y = (z, 0)$, is identical in both cases, as is the instanton, but we expect to find different prefactors due to the difference in curvature at y between the two sets. In the left panel of Figure 4 we illustrate the problem setup. We will focus on the case $r > z$ when the instanton lies on the x-axis: at $r = z$ the instanton becomes degenerate and we obtain *caustics*, that is due to rotational symmetry every point on the circle is equally likely to be the exit point.

When $r > z$, we have, basically identically to Section 5.0.1,

$$\hat{\theta}_y(1) = 2\gamma z, \qquad |\langle\hat{n}, b(y)\rangle| = \gamma z$$
$$\hat{Q}^* = (2\gamma)^{-1}\mathrm{Id}, \qquad \hat{Q}_y(t) = (2\gamma)^{-1}\mathrm{Id}$$
$$\hat{V} = (2\gamma z^2)^{-1/2}.$$

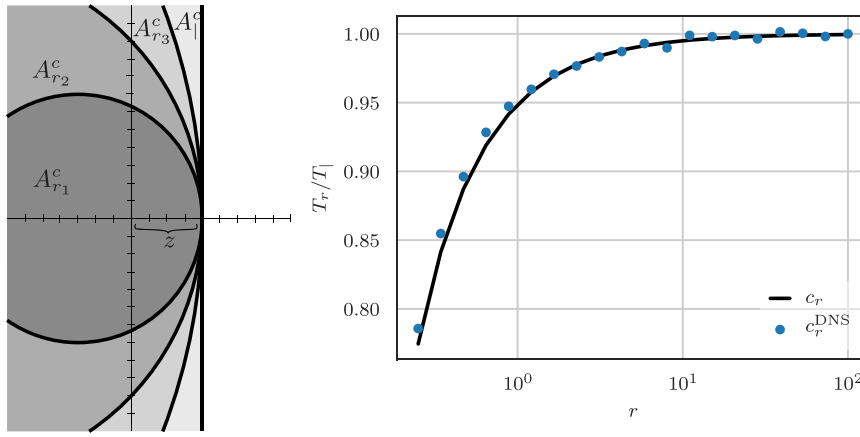

FIGURE 4 Left: Schematic representation of the curvature experiment. An Ornstein-Uhlenbeck process is started at the origin (center), and we are interested in the time T_r at which it leaves the circular set A^c_r, or time $T_|$ at which it leaves the half-plane $A^c_|$. The most likely exit point is always $(z, 0)$, but the curvature of the set boundary differs. Right: Ratios of T_r^{DNS} to $T_|^{\mathrm{DNS}}$ for different r, in comparison to the theoretical prediction (5.28). For $r \to \infty$, this ratio should converge to 1, but for small radii the curvature has a measurable effect. For example, at a radius $r = \frac{1}{4}$, the measured exit time is roughly 25% smaller than for the planar exit.

The only difference between the two cases (5.25) and (5.26) are the curvature contributions $\hat{F}_r$ and $\hat{F}_|$, respectively. For the half-space $A^c_|$, the curvature at y is 0, and thus $\hat{F}_| = \mathrm{Id}$. For the circle, instead, we can choose

$$f(x) = |x - (z - r, 0)|^2 - r^2$$

so that A^c_r is the zero level-set of f. Then,

$$|\nabla f(y)| = 2r$$

and

$$\nabla\nabla f(y) = 2\mathrm{Id},$$

and thus

$$\begin{aligned}\hat{F}_r &= \mathrm{Id} - \hat{Q}_y(1)|\hat{\theta}_y(1)||\nabla f(y)|^{-1}\nabla\nabla f(y)\\ &= \left(1 - \frac{z}{r}\right)\mathrm{Id}.\end{aligned}$$

Defining the scalar contribution of the curvature to the prefactor $c = (\det_\perp \hat{F})^{-1/2}$ as $c_|$ for the planar case and c_r for the case of a circle with radius r, respectively, we obtain

$$c_| = 1, \qquad \text{and} \qquad c_r = \left(\det{}_\perp(1 - \frac{z}{r})\mathrm{Id}\right)^{1/2} = \sqrt{\frac{r - z}{r}} \tag{5.27}$$

In order to measure this curvature prefactor experimentally, we performed the following numerical experiment: For different radii r from $r = 0.25$ to $r = 100$, we performed $N = 2 \cdot 10^5$ Monte-Carlo simulations each, and measured the mean time T_r^{DNS} to exit the set A_r^c. We performed further $2 \cdot 10^5$ Monte-Carlo simulations in the planar case $A_|^c$ to obtain $T_|^{\mathrm{DNS}}$. The ratio of the measured passage times of the planar case to the circular case yields a numerical estimate of the curvature component only:

$$c_r^{\mathrm{DNS}} = \frac{T_r^{\mathrm{DNS}}}{T_|^{\mathrm{DNS}}}$$

which can be compared against the analytical prediction

$$c_r = \sqrt{(r-z)/r}. \tag{5.28}$$

This comparison is shown in Figure 4. For $r \to \infty$, c_r indeed converges to 1, but for small radii the curvature has a measurable effect, and the measured discrepancy to the planar prediction agrees with the correction term given in (5.28). The other parameters are $\gamma = 1$, $z = 0.1$, and $\varepsilon = 5 \cdot 10^{-3}$.

6 | INFINITE-DIMENSIONAL EXAMPLES

In the previous sections, we derived prefactor estimates and sharp limits in finite dimension. All of these estimates have a counterpart in infinite dimension, that is when applied to stochastic partial differential equations.

For concreteness we will focus on reaction-advection-diffusion equations of the type (for more general equations see Appendix F)

$$\partial_t c + v(x) \cdot \nabla c = \nabla \cdot (D(x)\nabla c) + r(x)c + f(c) + \sqrt{\varepsilon}\eta, \quad c(0) = c_0, \tag{6.1}$$

Here $t \in [0, \infty)$, $x \in \Omega \subset \mathbb{R}^d$, with Ω compact, and $c : [0, \infty) \times \Omega \to \mathbb{R}$; the vector field $v : \Omega \to \mathbb{R}^d$ is in $C^2(\Omega)$; the diffusion tensor $D : \Omega \to \mathbb{R}^d \times \mathbb{R}^d$ is in $C^2(\Omega)$, symmetric, $D^T(x) = D(x)$, and positive-definite for all $x \in \Omega$; $r : \Omega \to \mathbb{R}$ is in $C^2(\Omega)$; the reaction term $f : \mathbb{R} \to \mathbb{R}$, is in $C^2(\Omega)$, nonlinear in general; and the noise η is white-in-time Gaussian with covariance

$$\mathbb{E}\eta(t,x)\eta(t',x') = \delta(t-t')\chi(x,x'). \tag{6.2}$$

where $\chi : \Omega \times \Omega \to \mathbb{R}$ given and in $C^2(\Omega \times \Omega)$. If Ω is a rectangular domain in $\mathbb{R}^d$, we can impose periodic boundary condition on c, assuming that all the other functions in (6.1) are also periodic; otherwise, denoting by $\hat{n}(x)$ the unit normal to $\partial\Omega$, we impose that $r(x) = 0$ and $\hat{n}(x) \cdot v(x) = 0$ on $\partial\Omega$, $\chi(x,y) = 0$ for all $x \in \partial\Omega$ and $y \in \Omega$ or $y \in \partial\Omega$ and $x \in \Omega$, and also that

$$\hat{n}(x) \cdot D(x)\nabla c(t,x) = 0 \qquad \text{for all } x \in \partial\Omega \text{ and } t \geq 0. \tag{6.3}$$

In this infinite-dimensional setting, the determinants must be replaced by functional determinants, the instanton equations become PDEs, and the Riccati equation must be replaced by a functional variant as well. While these changes require a whole new set of techniques to rigorously prove validity, our results and methods remain formally applicable in infinite dimension.

For example, considering probabilities for a linear observable, that is

$$P_\varepsilon^{c_0}(T,z) = \mathbb{P}^{c_0}\left(\int_\Omega \phi(x)c(T,x)\,dx \geq z\right), \tag{6.4}$$

for some test function $\phi : \Omega \to \mathbb{R}$, we obtain a proposition analogous to Proposition 4.3:

Proposition 6.1 (Probabilities for SPDEs). *Let the fields $c(t,x)$, $\theta(t,x)$ solve the instanton equations*

$$\begin{cases} \partial_t c = \nabla\cdot(D(x)\nabla c) - v(x)\cdot\nabla c + r(x)c + f(c) + \displaystyle\int_\Omega \chi(x,y)\theta(t,y)\,dy, & c(0) = c_0, \\ \partial_t \theta = -\nabla\cdot(D(x)\nabla\theta) - \nabla\cdot(v(x)\theta) - r(x)\theta - f'(c)\theta, & \theta(T) = \phi, \end{cases} \tag{6.5}$$

with either periodic boundary conditions, or

$$\hat{n}(x)\cdot D(x)\nabla c(t,x) = \hat{n}(x)\cdot D(x)\nabla\theta(t,x) = 0 \qquad \text{for all } \ x\in\partial\Omega \ \text{ and } \ t\in[0,T]; \tag{6.6}$$

Let also the field $Q(t,x,y) = Q(t,y,x)$ solve

$$\begin{aligned} \partial_t Q = {} & \nabla_x\cdot(D(x)\nabla_x Q) + \nabla_y\cdot\big(D(y)\nabla_y Q\big) - v(x)\cdot\nabla_x Q - v(y)\cdot\nabla_y Q + r(x)Q + r(y)Q \\ & + f'(c(t,x))Q + f'(c(t,y))Q + \int_\Omega Q(t,x,z)Q(t,z,y)f''(c(t,z))\theta(t,z)\,dz + \chi(x,y), \end{aligned} \tag{6.7}$$

with initial condition $Q(0) = 0$, and either periodic boundary conditions in x and y, or

$$\begin{aligned} &\hat{n}(x)\cdot D(x)\nabla_x Q(t,x,y) = 0 && \text{for all } x\in\partial\Omega,\ y\in\Omega,\ \text{and } t\in[0,T], \\ &\hat{n}(y)\cdot D(y)\nabla_y Q(t,x,y) = 0 && \text{for all } y\in\partial\Omega,\ x\in\Omega,\ \text{and } t\in[0,T]. \end{aligned} \tag{6.8}$$

Then the probability $P_\varepsilon^{c_0}(T,z)$ in (6.4) satisfies

$$\lim_{\varepsilon\to 0}\frac{P_\varepsilon^{c_0}(T,z)}{\bar{P}_\varepsilon^{c_0}(T,z)} = 1, \tag{6.9}$$

where

$$\bar{P}_\varepsilon^{c_0}(T,z) = (2\pi)^{-1/2}\varepsilon^{1/2}\mathcal{V}(T,c_0)\exp\left(-\frac{1}{2\varepsilon}\int_0^T\int_{\Omega^2}\theta(t,x)\chi(x,y)\theta(t,y)\,dx\,dy\,dt\right), \tag{6.10}$$

with

$$\begin{aligned} \mathcal{V}(T,c_0) = {} & \left(\int_{\Omega^2}\theta(T,x)Q(T,x,y)\theta(T,y)\,dx\,dy\right)^{-1/2} \\ & \times\exp\left(\frac{1}{2}\int_0^T\int_\Omega f''(c(t,x))\theta(t,x)Q(t,x,x)\,dx\,dt\right). \end{aligned} \tag{6.11}$$

We will omit the proof of this proposition since it essentially amounts to translating the equations in Proposition 4.3 to the infinite dimensional setting, and using the fact that the term involving the tensor $F(T,x)$ disappears since the equivalent of $\nabla\nabla f$ is zero for a linear observable as in (6.4). Similar reformulations of our other propositions are straightforward as well, and for the sake of brevity we will therefore omit writing them down.

Note the instanton equation for $c(t,x)$ in (6.5) as well as the Riccati equation in (6.7) for $Q(t,x,y)$ are well-posed as formulated, that is forward in time; similarly the instanton equation for $\theta(t,x)$ in (6.5) is well-posed as formulated, that is backward in time.

In Sections 6.1 and 6.2, we confirm numerically that Proposition 6.1, or its counterparts for other quantities, produces results that agree with those obtained via direct sampling of the SPDE in (6.1). As example problems, we consider two special cases of (6.1): First, in Section 6.1 we study a linear advection-reaction-diffusion equation with spatially non-homogeneous forcing, for which some analytical results can be derived. The probability that the concentration exceeds a threshold at a given location can then be estimated by a mixed analytical-numerical approach. Second, in Section 6.2, we study a reaction-advection-diffusion equation with cubic nonlinearity, where we need to resort to fully solving numerically all involved instanton and Riccati equations.

6.1 | Linear reaction-advection-diffusion equation with non-local forcing

Here we consider the stochastic one-dimensional advection-diffusion-reaction equation,

$$\partial_t c = \kappa\partial_x^2 c - \partial_x(v(x)c) - \alpha c + \sqrt{\varepsilon}\eta. \tag{6.12}$$

where $x \in [0,1]$ periodic. This a special case of (6.1) with $D = \kappa$, $r(x) = -\partial_x v$, and $f(c) = -\alpha c$. For the covariance of the white-in-time forcing $\eta(t,x)$ we take

$$\chi(x,x') = \psi^\delta_{x_1}(x)\psi^\delta_{x_1}(x'). \tag{6.13}$$

where $\psi^\delta_{x_1}(x)$ is a mollifier of length $\delta < 1$ concentrated around $x_1 \in [0,1]$. For the observable, we take

$$\mathbb{P}\left(\int_0^1 c(x)\psi^\delta_{x_2}(x)\,dx \geq z\right), \tag{6.14}$$

where $x_2 \neq x_1$ and the expectation is taken on the invariant measure of the solution to (6.12). Intuitively, the scenario we are investigating is therefore that of a pollutant, the density of which is described by $c(t,x)$, along a one-dimensional periodic channel. The pollutant is randomly emitted into the environment at a spatial location x_1, and gets advected and diffused conservatively, but decays over time with rate α. We are interested in measuring extreme concentrations of the pollutant around the location x_2 somewhere else in the channel.

6.1.1 | Finite-dimensional analogous case

Equation (6.12) is a linear SPDE, an infinite-dimensional generalization of the (non-normal, n-dimensional) Ornstein-Uhlenbeck process,

$$dX_t^\varepsilon = -\Gamma X_t^\varepsilon\, dt + \sqrt{2\varepsilon}\sigma\, dW_t, \quad t \geq 0, \quad X_t^\varepsilon \in \mathbb{R}^n, \tag{6.15}$$

for $\Gamma \in \mathbb{R}^{n\times n}$, where $\Gamma \neq \Gamma^\top$ (non-symmetric), $\Gamma\Gamma^\top \neq \Gamma^\top\Gamma$ (non-normal), W is an n-dimensional Wiener process, σ not necessarily invertible, and $a = \sigma\sigma^\top$. In analogy to the above scenario, we can ask for probabilities on the invariant measure of the form

$$P_\varepsilon^{A_z} = \mathbb{P}(X^\varepsilon \in A_z), \quad \text{where} \quad A_z = \{x \in \mathbb{R}^n | \langle k, x\rangle \geq z\}. \tag{6.16}$$

Intuitively, we want to estimate the probability that the process reaches the far-side of a plane with normal k, distance z away from the origin.

If we define the symmetric, positive semi-definite matrix C as the solution of the Lyapunov equation

$$\Gamma C + C\Gamma^\top = 2a, \tag{6.17}$$

the quasi-potential of (6.15) is given by

$$V(x) = \langle x, C^{-1}x\rangle. \tag{6.18}$$

Since C is not necessarily invertible, we interpret $w = C^{-1}v$ to be the solution of $v = Cw$ if it exists, and otherwise the quasi-potential is set to infinity. The final point of the geometric instanton $\hat{\phi}(s)$ for observable value z must be given by

$$y = \underset{x\in\partial A_z}{\operatorname{argmin}}(V(x)) = \frac{z}{\langle k, Ck\rangle}Ck \tag{6.19}$$

Here, we must assume that k is not in the kernel of C, which is the same as saying that k is in the support of the invariant measure of (6.15). The action is

$$I(z) = \frac{z^2}{\langle k, Ck\rangle}, \tag{6.20}$$

which follows from evaluating $V(x)$ at $x = y$ given by (6.19).

To estimate the prefactor, we need to solve the Q-equation with appropriate boundary condition. Because the Equation (6.15) is linear, $\hat{Q}(s) = \frac{1}{2}C$. As a result

$$\lim_{\varepsilon\to 0}\frac{P_\varepsilon^{A_z}}{\bar{P}_\varepsilon^{A_z}} = 1 \qquad \text{with} \qquad \bar{P}_\varepsilon^{A_z} = \sqrt{\frac{\varepsilon\langle k, Ck\rangle}{4\pi z^2}}\exp\left(-\frac{z^2}{\varepsilon}\langle k, Ck\rangle^{-1}\right) \tag{6.21}$$

where we used

$$\hat{V} = \langle\hat{\theta}_y(1), \hat{Q}_y(1)\hat{\theta}_y(1)\rangle^{-1/2} = \frac{1}{\sqrt{2}z}\langle k, Ck\rangle^{1/2} \tag{6.22}$$

since

$$\hat{\theta}(s) = 2C^{-1}\hat{\phi}(s), \quad \text{so that} \quad \hat{\theta}(1) = 2C^{-1}y = \frac{2z}{\langle k, Ck\rangle}k. \tag{6.23}$$

Due to the linearity of the system, only the end location of the instanton at $s = 1$ plays a role.

6.1.2 | Infinite dimensional setting

Coming back to the infinite-dimensional case, we can proceed similarly, noticing that (6.12) can be written as

$$\partial_t c = -\mathcal{G}c + \eta, \tag{6.24}$$

where

$$\mathcal{G} = -\kappa\partial_x^2 + v(x)\partial_x + (\partial_x v) + \alpha \tag{6.25}$$

is a linear differential operator, acting on functions in L^2. Notably, $\mathcal{G}$ is not normal, and we need to solve the Lyapunov equation

$$\mathcal{G}C + C\mathcal{G}^\top = 2\psi^\delta_{x_1}(x)\psi^\delta_{x_1}(y), \tag{6.26}$$

where $C(x, y)$ is symmetric and positive semi-definite in L^2 and $C\mathcal{G}^\top = (\mathcal{G}C)^\top$, that is it is the differential operator acting on the second variable of $C(x, y)$. Explicitly (6.26) reads

$$-\kappa\left(\partial_x^2 + \partial_y^2\right)C - (v\partial_x + v\partial_y)C - (\partial_x v + \partial_y v)C + 2\alpha C = 2\psi^\delta_{x_1}(x)\psi^\delta_{x_1}(y). \tag{6.27}$$

By extending the argument for the finite dimensional case to the functional setting, we also deduce that the endpoint at $s = 1$ of the geometric instanton $\hat{\phi}(s, x)$ for hitting the set $\mathcal{A}_z$ is given by

$$\hat{\phi}(1, x) = \operatorname*{argmin}_{\xi\in\partial\mathcal{A}_z} V(\xi) = \frac{z\int_0^1 C(x, y)\psi^\delta_{x_2}(y)\,dy}{\int_0^1\int_0^1 \psi^\delta_{x_2}(x)C(x, y)\psi^\delta_{x_2}(y)\,dx\,dy}, \tag{6.28}$$

in analogy to Equation (6.19) by replacing the inner product with the L^2 inner product. The corresponding action is (compare (6.20))

$$I(z) = z^2\left(\int_0^1\int_0^1 \psi^\delta_{x_2}(x)C(x, y)\psi^\delta_{x_2}(y)\,dx\,dy\right)^{-1}. \tag{6.29}$$

Concerning the prefactor, we have (compare (6.23))

$$\theta(t = 0, x) = 2z\left(\int_0^1\int_0^1 \psi^\delta_{x_2}(x)C(x, y)\psi^\delta_{x_2}(y)\,dx\,dy\right)^{-1}\psi^\delta_{x_2}(x) \tag{6.30}$$

and thus (compare (6.22))

$$\hat{\mathcal{V}} = \frac{1}{2z}\left(\int_0^1 \int_0^1 \psi^\delta_{x_2}(x) C(x,y) \psi^\delta_{x_2}(y)\, dx\, dy\right)^{1/2}, \tag{6.31}$$

so that in total,

$$\lim_{\varepsilon\to 0} \frac{\mathbb{P}\left(\int_0^1 c(x)\psi^\delta_{x_2}(x)\, dx \geq z\right)}{P_\varepsilon(z)} = 1 \tag{6.32}$$

with

$$P_\varepsilon(z) = \sqrt{\frac{\varepsilon}{4\pi z^2}}\left(\int_0^1 \int_0^1 \psi^\delta_{x_2}(x) C(x,y) \psi^\delta_{x_2}(y)\, dx\, dy\right)^{1/2} \exp\left(-\varepsilon^{-1} I(z)\right) \tag{6.33}$$

as in Equation (6.21). Clearly, this probability is Gaussian, as expected when we consider a linear system with Gaussian input.

In order to obtain C, we numerically solve the operator Lyapunov equation (6.26) for a finite difference representation of $\mathcal{G}$. As an example, we take the advection-diffusion-reaction equation (6.12), with a velocity field

$$v(x) = A\sin(2\pi x) \qquad (A > 0), \tag{6.34}$$

that is with negative divergence at the center of the domain, $x = 1/2$. We expect (if we were not to condition on any outcome, and would force homogeneously in space) that in a typical configuration the concentration has higher variance at $x = 1/2$ and lower around $x = 0$, where it is depleted by the velocity. We also choose $x_1 = 1/32$ and $x_2 = 1/4$, that is the forcing is localized on the very left of the domain, where concentration is released randomly, and we are sensing concentration further down the channel, where it is transported by advection and diffusion. We also check the results above numerically by comparing the instanton prediction derived here with the result of a direct simulation of the advection reaction diffusion equation (6.12) with stochastic forcing localized according to (6.13), and counting the number of times the observable exceeds a given threshold.

In the left panel of Figure 5 we show the corresponding instanton: It has a non-trivial dependence on both the location of the forcing as well as the location of the observation. Additionally shown are the localized functions defining the forcing and the observation, as well as the velocity field. In the right panel of Figure 5 we show a comparison against direct simulation results of the SPDE. In particular we note that the mean of all observed events within $\mathcal{A}_z$ resembles the instanton. Further, the highest variance in the direct simulation is clearly observed where noise is injected, as well as close to $x = 1/2$, at the sink of the velocity field.

Depending on z and ε, the probability can be computed via (6.32). In Figure 6 we show a comparison of these probabilities against numerics, obtained by integrating the SPDE (6.12) for a long time, and observing how often each threshold z was exceeded. We note that the prediction indeed captures not only the scaling, but also the correct prefactor.

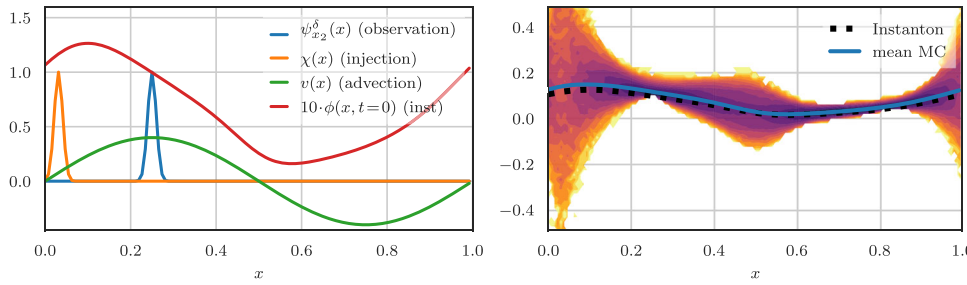

FIGURE 5 *Left*: Endpoint of the instanton configuration (rescaled, red), that is solution of the minimization problem (6.28) for the advection-reaction-diffusion equation (6.12) with velocity field $v(x)$ (green), conditioning on observing a high concentration in the region specified by $\psi^{\delta}_{x_2}$ (blue), where concentration is released randomly around location x_1 (yellow). Here, $\kappa = 10^{-2}, \alpha = 1, v(x) = 0.4\sin(2\pi x), \varepsilon = 5\cdot 10^{-2}, z = 0.0025, x_1 = 1/32$, $x_2 = 1/4$, and $\delta = 10^{-2}$, as well as $N_x = 128$. *Right*: Comparison between instanton and the result from direct numerical simulations, conditioning on hitting the set $\mathcal{A}_z$. Notably, the mean realization recovers the instanton, but variances are very high around $x = 1/32$, where the concentration fluctuations are inserted, and also high around $x = 1/2$, where the velocity field compresses the fluctuations of the concentration field.

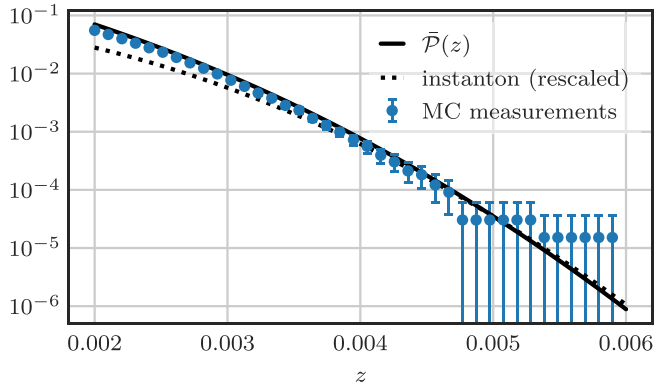

FIGURE 6 Probabilities for the advection-diffusion-reaction problem with localized forcing obtained by direct numerical simulations (blue), from formula (6.32) (black). Also shown is the instanton prediction (black dotted), that is only the exponential contribution from the large deviation scaling. It is rescaled by a factor 10^{-2} to obtain the best fit in the tail. The difference between the instanton prediction and $\bar{\mathcal{P}}(z)$ is noticeable.

6.2 | Nonlinear reaction-advection-diffusion equation

For the linear reaction-advection-diffusion equation of Section 6.1, we were able to harness linearity to avoid explicitly using Proposition 6.1 and numerically solving the instanton and Riccati equations. This is no longer possible if we choose the system to be nonlinear, for example by taking a cubic reaction term. In this case, we no longer have access analytically to the invariant measure, the instanton trajectory, or the solution to the prefactor terms.

Concretely, we consider the nonlinear stochastic partial differential equation

$$\partial_t c = \kappa\partial_x^2 c - v(x)\partial_x c - \alpha c - \gamma c^3 + \sqrt{2\varepsilon}\eta, \tag{6.35}$$

where the spatial variable x is periodic on the domain $[-L/2, L/2]$ and $t \in [-T, 0]$. This equation is a special case of (6.1) with $D(x) = \kappa$, $r(x) = 0$, and $f(c) = -\alpha c - \gamma c^3$: the parameters α and γ control the linear and nonlinear part of the reaction term, respectively. As velocity field we pick

$$v(x) = 4 + 2\sin(4\pi x/L), \tag{6.36}$$

which always advects to the right, but with spatially varying speed. As a consequence, the whole equation is no longer translation invariant. We will assume that the noise η is white in space and time, with covariance

$$\mathbb{E}(\eta(t,x)\eta(t',x')) = \delta(t-t')\delta(x-x'). \tag{6.37}$$

This is a rougher noise than the one in the SPDE (6.5), which is allowed because (6.35) is well-posed in 1D with this forcing [20].

We are interested in the probability that a sample on the invariant measure of (6.35) exceeds the threshold z at the location $x = 0$, that is

$$\mathbb{P}(c(x=0) \geq z), \tag{6.38}$$

Note that this problem can either be viewed as a nonlinear version of the reaction-advection-diffusion equation of Section 6.1, or as an infinite dimensional version of the $\mathbb{R}^2$ process given in Section 4.2.1.

To apply our method we need to first solve the geometric instanton equations (2.36) using the appropriate boundary conditions. Since we are focusing in this example on the limit $T \to \infty$, an efficient way of numerically solving these equations using an iterative scheme and arc-length parametrization has been discussed in previous work [29]. Since for Equation (6.35) the Hamiltonian is

$$H(c,\theta) = \int_{-L/2}^{L/2} \left(\left(\kappa \partial_x^2 \hat{c} - v(x)\partial_x \hat{c} - \alpha \hat{c} - \gamma \hat{c}^3 \right) \hat{\theta}(x) + |\hat{\theta}(x)|^2 \right) dx \tag{6.39}$$

we immediately obtain the (geometric) instanton equations in this case as

$$\begin{cases} \lambda \partial_s \hat{c} = \kappa \partial_x^2 \hat{c} - v(x)\partial_x \hat{c} - \alpha \hat{c} - \gamma \hat{c}^3 + 2\hat{\theta}, & \hat{c}(0) = 0 \\ \lambda \partial_s \hat{\theta} = -\kappa \partial_x^2 \hat{\theta} - \partial_x (v(x)\hat{\theta}) + \alpha \hat{\theta} + 2\gamma \hat{c}^2 \hat{\theta}, & \hat{\theta}(1) = \delta(x). \end{cases} \tag{6.40}$$

In this case, the boundary conditions $\hat{c}(0) = 0$ and $\hat{\theta}(1) = \delta(x)$ allow direct application of the previously developed geometric techniques [29] to solve the above system iteratively in order to find the instanton solution.

Once the instanton is found, we need to solve the corresponding (geometric variant of the) forward Riccati equation for $\hat{Q}(s,x,y)$ given by (6.7).

For the specific SPDE (6.35), the equation for $\hat{Q}(s,x,y) = \hat{Q}(s,y,x)$ is

$$\begin{aligned} \lambda \partial_s \hat{Q} = {} & \kappa \partial_x^2 \hat{Q} + \kappa \partial_y^2 \hat{Q} - v(x)\partial_x \hat{Q} - v(y)\partial_y \hat{Q} - 2\alpha \hat{Q} - 3\gamma \left(\hat{c}^2(s,x) + \hat{c}^2(s,y) \right) \hat{Q} \\ & - 6 \int_{-L/2}^{L/2} \hat{Q}(s,x,z)\hat{c}(s,z)\hat{\theta}(s,z)\hat{Q}(s,z,y)\, dz + 2\delta(x-y) \end{aligned} \tag{6.41}$$

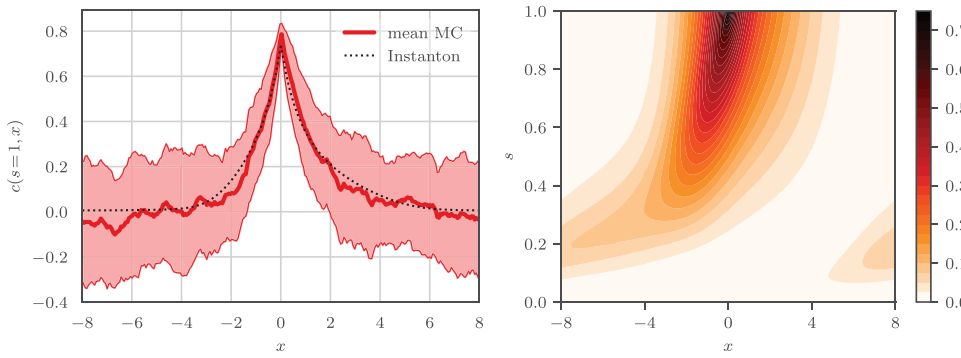

FIGURE 7 Left: Comparison of the numerically computed instanton (black dots) with the filtered solution from direct numerical simulations (red) realizing a high amplitude event $z = 0.7$ at the origin $x = 0$, plus/minus one standard deviation (light red shaded region), for the nonlinear stochastic partial differential equation (6.35). The spatially dependent positive velocity field $v(x)$ leads to an asymmetry if the typical final configuration exceeding z at $x = 0$, which is visible both in the instanton and the direct simulation results. Right: Instanton trajectory along the arc-length parameter s, showing the temporal evolution of the instanton into the large amplitude configuration at $s = 1$. The positive velocity leads to a traveling wave instanton solution that slowly amplifies over time.

to be solved with the initial condition, $\hat{Q}(0) = \hat{Q}_*$, solution to the Lyapunov equation

$$-\frac{1}{2}\kappa\left(\partial_x^2\hat{Q}_* + \partial_y^2\hat{Q}_*\right) + \alpha\hat{Q}_* + \frac{1}{2}\left(v(x)\partial_x\hat{Q}_* + v(y)\partial_y\hat{Q}_*\right) = \delta(x-y). \tag{6.42}$$

One can numerically integrate the instanton equations (6.40) to obtain the most likely configuration that achieves a given amplitude z at $x = 0$ at final time $t = 0$. Fourier transforms were used to calculate the spatial derivatives. When solving the associated Riccati equation for Q, however, we chose an equally distant point grid to discretize the time interval $[-T, 0]$ for stability reasons. Here, the time T can be found from the geometric parametrization [34] and the instanton solution needs to be interpolated onto this new point grid. For the solution of the Riccati equation on the equidistant grid, exponential time differencing [36] was employed and the diffusive term $\kappa(\partial_x^2 + \partial_y^2)Q$ can be treated for numerical efficiency using the two-dimensional fast Fourier transform.

In the left panel of Figure 7 we compare this numerically computed instanton with the filtered instanton from direct simulations of the SPDE (6.35) using the method described in [28]. As can be seen, the mean realization with $z = 0.7$ at $x = 0$ in the SPDE is very similar to the instanton. In particular, both instanton and the typical sample from the SPDE show an asymmetry around $x = 0$ coming from the spatially inhomogeneous velocity field (6.36). Also shown is one standard deviation of the fluctuations around the instanton as shaded red region. In the right panel of Figure 7 we show the whole evolution of the instanton in arc-length parameter s that compactifies the infinite time interval $t \in [0, \infty)$ into $s \in [0, 1]$. Clearly visible is the (inhomogeneous) movement of the peak as it is advected with the positive velocity $v(x)$ given in Equation (6.36).

In the left panel of Figure 8 shows the conjugate momentum $\hat{\theta}(s, x)$ along the arc-length parameter s, and in the right panel we show the quantity $f''(\hat{c}(s, x))\hat{\theta}(s, x)\hat{Q}(s, x, x)$: the exponential of its integral along x and s enters the expression in (6.11) for the prefactor component $\mathcal{V}(T, c_0)$.

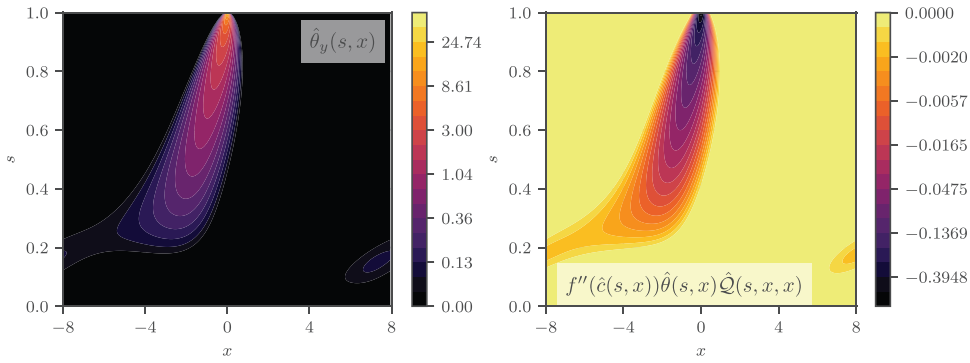

FIGURE 8 Left: Conjugate momentum $\hat{\theta}(s,x)$ along the arc-length parameter s. Right: The quantity $f''(\hat{c}(s,x))\hat{\theta}(s,x)\hat{Q}(s,x,x)$. The exponential of its integral along x and s enters the expression in (6.11) for the prefactor component $\mathcal{V}(T,c_0)$.

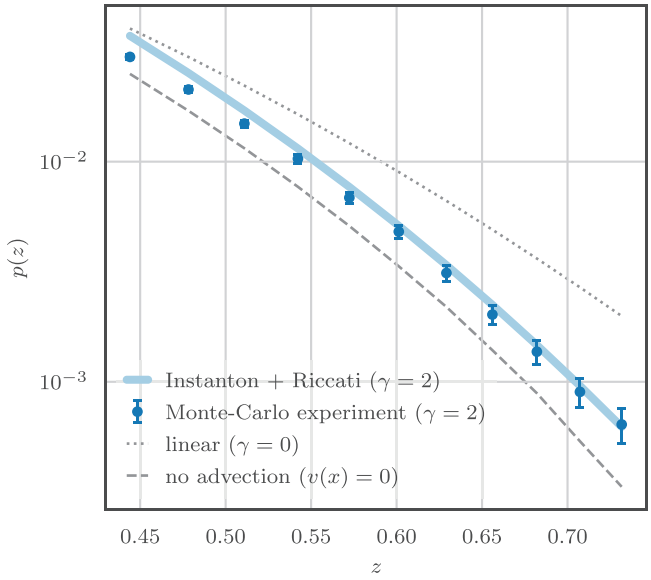

FIGURE 9 Comparison of the predicted probabilities $p(z)$ of exceeding the threshold z. The dark blue dots shows the probability p estimated via direct numerical simulations for the nonlinear case ($\gamma = 2$), while the light blue line is the theoretical prediction from instanton and Riccati equation. Clearly, the instantons, together with the corresponding prefactors, approximate these probabilities well. For comparison, we show the analytical prediction that can be obtained for the linear case ($\gamma = 0$), highlighting the fact that the nonlinearity indeed plays a role for the tail probabilities in particular. Similarly, we show the situation without advection ($v(x) = 0$), demonstrating that the advection term modifies the probabilities.

The combination of both results inserted into Proposition 6.1 allows us to determine the prefactor. In Figure 9 we compare the result from direct simulations of the SPDE (6.35) to the predictions of the instanton equations together with the prefactor, showing clear agreement especially for large values of z, as expected. A comparison is further made to the analytical result available for the linear case, $\gamma = 0$, which clearly shows that the nonlinear term affects the probability. Similarly we compare against the case without advection ($v(x) = 0$), which is gradient, demonstrating

that the advective term also has a considerable (opposite) effect on the probabilities. In this numerical example, we chose $N_x = 256$ grid points for the discretization in space for a domain of size $L = 16$, and $N_t = 5000$ grid points in time for the direct simulations for a temporal domain of size $T = 25$. The instanton is computed with $N_s = 2000$ discretization points in the arc-length parameter. Additional parameters are $\varkappa = 1$, $\alpha = 0.6$, and $\gamma = 2$. The noise level is set to $\varepsilon = 0.1$, and we collected $N_{\text{samples}} = 10^5$ samples in the direct simulations.

7 | GENERALIZATION TO PROCESSES DRIVEN BY A NON-GAUSSIAN NOISE

It is possible to generalize the above results to continuous time MJPs that cannot be represented by an SDE with additive Gaussian noise like in (1.1). The intuition is similar: considering a WKB approximation to the BKE allows us to obtain a Hamilton-Jacobi equation that defines the behavior of the instanton. If we furthermore keep track of the prefactor in the WKB, we will obtain an additional equation for it on the next order, which will yield a generalization of the Riccati equations to obtain sharp prefactor estimates.

Concretely, consider a continuous-time MJP on the state space $\mathbb{R}^d$ with the generator

$$L_\varepsilon f(x) = \frac{1}{\varepsilon} \sum_{r=1}^{R} a_r(x)(f(x + \nu_r \varepsilon) - f(x)) \tag{7.1}$$

which encodes a reaction network for $d \in \mathbb{N}$ species with $R \in \mathbb{N}$ reactions, each reaction r leading to a change in species defined by the vector $\nu_r \in \mathbb{R}^d$, and happening with rate $a_r(x) \in \mathbb{R}_+$ possibly depending on the state $x \in \mathbb{R}^d$. Depending on the microscopic model at hand, we can expand $a_r(x)$ in terms of orders of ε,

$$a_r(x) = a_r^{(0)}(x) + \varepsilon a_r^{(1)}(x) + \mathcal{O}(\varepsilon^2). \tag{7.2}$$

If we insert the WKB ansatz $f(t,x) = Z(t,x)\exp(\varepsilon^{-1}S(t,x))$ in the BKE

$$\partial_t f = L_\varepsilon f, \tag{7.3}$$

and collect terms of successive orders in ε, we obtain

$$\mathcal{O}(\varepsilon^{-1}): \qquad \partial_t S = -\sum_{r=1}^{R} a_r^{(0)}(x)(e^{\nu_r \cdot \nabla S} - 1) = -H(x, \nabla S(x)) \tag{7.4}$$

$$\mathcal{O}(\varepsilon^{0}): \quad \partial_t Z = -\sum_{r=1}^{R} \left(a_r^{(0)}(x) e^{\nu_r \cdot \nabla S} \left(\nu_r \cdot \nabla Z + \frac{1}{2} Z \nu_r \cdot \nabla\nabla S \nu_r \right) + a_r^{(1)}(x) Z (e^{\nu_r \cdot \nabla S} - 1) \right). \tag{7.5}$$

The instanton ϕ solves the Hamilton's equations

$$\begin{cases} \dot{\phi} = \nabla_\theta H(\phi, \theta), & \phi(0) = x \\ \dot{\theta} = -\nabla_\phi H(\phi, \theta), \end{cases} \tag{7.6}$$

where we additionally get a boundary condition $\phi(T) = y$ or $\theta(T) = \nabla f(\phi(T))$ depending on the scenario under consideration. For the generator (7.1), the Hamiltonian is given by

$$H(\phi, \theta) = \sum_{r=1}^{R} a_r^{(0)}(\phi)\left(e^{\nu_r \cdot \theta} - 1\right), \tag{7.7}$$

but an arbitrary Hamiltonian is possible in general. If we evaluate $Z(t, x)$ along the instanton, $G(t) = Z(t, \phi(t))$, we have

$$\dot{G}(t) = \dot{Z} + \nabla Z \cdot \dot{\phi} = \dot{Z} + \sum_{r=1}^{R} \nu_r \cdot \nabla Z a_r^{(0)}(\phi) e^{\nu_r \cdot \nabla S}, \tag{7.8}$$

and therefore, along the instanton, Equation (7.5) becomes

$$\dot{G} = -\frac{1}{2} G \sum_{r=1}^{R} \left(e^{\nu_r \cdot \nabla S} a_r^{(0)}(\phi) \nu_r \cdot \nabla \nabla S \nu_r + 2 a_r^{(1)}(\phi)(e^{\nu_r \cdot \nabla S} - 1) \right), \quad G(T) = 1. \tag{7.9}$$

Written in terms of the Hamiltonian and the additional $\mathcal{O}(\varepsilon)$ drift term, this becomes

$$\dot{G} = -\frac{1}{2} G \left(\operatorname{tr}(H_{\theta\theta} W) + A^{(1)} \right), \quad G(T) = 1, \tag{7.10}$$

with $A^{(1)} = 2 \sum_r a_r^{(1)}(\phi)(e^{\nu_r \cdot \nabla S} - 1)$. Note that, similar to the derivations for Gaussian SDEs, $G(t)$ plays the role of prefactor along the instanton (as $G(t) = Z(t, \phi(t))$ and is used to compute the eventual prefactor correction (which is merely $G(0)$).

As before, we obtain an evolution equation for $W = \nabla \nabla S$ by differencing the HJB equation twice,

$$\dot{W} = -H_{\phi\phi} - H_{\phi\theta} W - W H_{\phi\theta}^{\top} - W H_{\theta\theta} W, \tag{7.11}$$

to be solved with boundary conditions that depend on the scenario, and where subscripts of the Hamiltonian denote differentiation. We can also derive an equation for Q,

$$\dot{Q} = Q H_{\phi\phi} Q + Q H_{\phi\theta} + H_{\phi\theta}^{\top} Q + H_{\theta\theta}. \tag{7.12}$$

This allows one to compute sharp estimates for expectations, probability densities, hitting probabilities and exit times in a similar way as before, replacing the instanton equations (1.6) and its variants with (7.6), and the forward and backward Riccati equations with (7.11) and (7.12).

7.0.1 | Example: Continuous time MJP

Consider the following continuous-time MJP, inspired by [8], in which a particle hops on a grid with spacing ε , $X \in \varepsilon \mathbb{Z}$ (see Figure 10), that is the spatial coordinate becomes continuous for

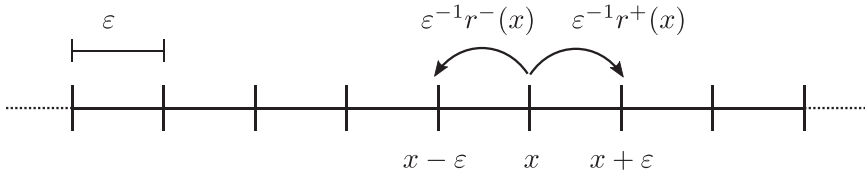

FIGURE 10 Schematic depiction of the MJP with generator (7.14): A particle at point $x \in \varepsilon\mathbb{Z}$ jumps with rates $r^{\pm}$ to the left and right, respectively. The rates are given in Equation (7.13). For $\varepsilon \to 0$, we compute the probability $P_T(z)$ of excursions larger than $z \in \mathbb{R}$ at time $t = T$.

$\varepsilon \to 0$. Left and right jumps happen with a rate

$$r^{\varepsilon}_{\pm}(x) = \exp\big(-\varepsilon^{-1}(E(x \pm \varepsilon) - E(x))\big), \tag{7.13}$$

and the generator is given by

$$L_\varepsilon f = \varepsilon^{-1}(r_+(x)(f(x+\varepsilon) - f(x)) + r_-(x)(f(x-\varepsilon) - f(x))). \tag{7.14}$$

By construction, this process is in detailed balance with respect to the Gibbs distribution

$$\mu_\infty(x) = Z^{-1}e^{-2\varepsilon^{-1}E(x)}, \quad \text{where} \quad Z = \sum_{x\in\varepsilon\mathbb{Z}} e^{-2\varepsilon^{-1}E(x)} \tag{7.15}$$

that is $E(x)$ plays the role of the free energy and $\frac{1}{2}\varepsilon$ that of the temperature.

In the continuum limit, $\varepsilon \to 0$, the rates (7.13) can be expanded as

$$r^{\varepsilon}_{\pm}(x) = e^{\mp E'(x)}\Big(1 + \tfrac{1}{2}\varepsilon E''(x) + \mathcal{O}(\varepsilon^2)\Big) = r^{(0)}_{\pm}(x) + \varepsilon r^{(1)}_{\pm} + \mathcal{O}(\varepsilon^2) \tag{7.16}$$

with

$$r^{(0)}_{\pm}(x) = e^{\mp E'(x)} \qquad \text{and} \qquad r^{(1)}_{\pm}(x) = \frac{1}{2}E''(x)e^{\mp E'(x)}. \tag{7.17}$$

Correspondingly, the LDT Hamiltonian takes the form

$$H(\phi,\theta) = r^{(0)}_+(\phi)\big(e^{\theta} - 1\big) + r^{(0)}_-(\phi)\big(e^{-\theta} - 1\big). \tag{7.18}$$

We are interested in estimating the probability to observe a large excursion $z \in \mathbb{R}$ at time $t = T$,

$$P_T(z) = \mathbb{P}(X_T > z | X_0 = 0) \tag{7.19}$$

for the MJP with generator (7.1), with the particle starting at $X_0 = 0$ at $t = 0$.

Solving this problem numerically by our approach amounts to performing the following steps:

(i) Solve the instanton equations

$$\begin{cases} \dot\phi = H_\theta(\phi,\theta) = e^{-E'+\theta} - e^{E'-\theta}, & \phi(0) = 0, \\ \dot\theta = -H_\phi(\phi,\theta) = -E''e^{-E'}\big(e^{\theta} - 1\big) + E''e^{E'}\big(e^{-\theta} - 1\big), & \phi(T) = z, \end{cases} \tag{7.20}$$

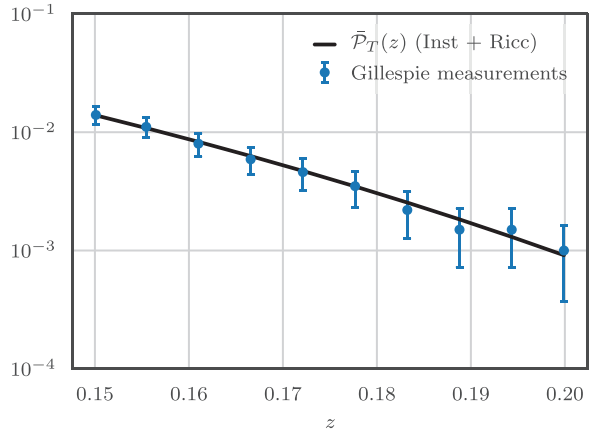

FIGURE 11 Comparison between numerical estimate of $\bar{P}_T(z)$ and sampling estimate obtained from the Gillespie algorithm [27] for the original MJP for small ε. Here, the parameters are $E(x) = \frac{1}{4}x^4$, $T = 5.12$, while for the instanton, $N_t = 512$, $\Delta t = 10^{-2}$, and for the Gillespie algorithm $N = 2 \cdot 10^3$, $N_{\text{samples}} = 10^4$.

(ii) Solve the Riccati equation

$$\begin{aligned}\dot{Q} &= H_{\phi\phi}Q^2 + 2H_{\phi\theta}Q + H_{\theta\theta} \\ &= ((-E''' + (E'')^2)e^{-E'}(e^{\theta} - 1) + (E''' + (E'')^2)e^{E'}(e^{-\theta} - 1))Q^2 \\ &\quad - E''(e^{-E'+\theta} + e^{E'-\theta})Q + (e^{-E'+\theta} + e^{E'-\theta}), \qquad Q(0) = 0 \end{aligned} \tag{7.21}$$

forward in time,

(iii) Assemble the full estimate as

$$\begin{aligned}\bar{P}_T(z) = &\left(\frac{\varepsilon}{2\pi Q(T)\theta^2(T)}\right)^{1/2} \exp\left(\frac{1}{2}\int_0^T \left(H_{\phi\phi}(\phi(t),\theta(t))Q(t) + A^{(1)}(\phi(t))\right) dt\right) \\ &\times \exp\left(-\varepsilon^{-1}\left(\int \theta d\phi - H(\theta,\phi)T\right)\right)\end{aligned} \tag{7.22}$$

This procedure can be carried out for various $z \in \mathbb{R}$ to, for example, investigate the probability $P_T(z)$ for rare events, that is large z. Displayed as black solid line in Figure 11 is the estimate $\bar{P}_T(z)$ at $T = 5.12$. Here, we choose $E(x) = \frac{1}{4}x^4$. For the numerical computation of the instanton, we employed a simple forward Euler scheme integrating the instanton equations (7.6) forward and backward multiple times until convergence, with $N_t = 512$ time discretization points, and $\Delta t = 10^{-2}$ temporal resolution. The blue markers compare the instanton prediction against a numerical computation of the actual stochastic process for finite (but small) $\varepsilon \ll 1$ by using the Gillespie algorithm [27]. Numerical parameters are $\varepsilon = 5 \cdot 10^{-4}$, and we took $N_{\text{samples}} = 10^4$ samples to estimate $P_T(z)$.

8 | CONCLUSIONS

In this paper, we have proposed explicit formulas to calculate the prefactor contribution of expectations, probabilities, probability densities, exit probabilities, and mean first passage times for

stochastic processes, both at finite time and over the invariant measure. The approach gives sharp estimates in the small noise limit, corresponding to the next order (pre-exponential) term to the Freidlin-Wentzell large deviation limit. This allows us to compute these probabilistic quantities in absolute terms, that is including normalization constants, for finite noise values, instead of merely producing the exponential scaling. This feature makes the approach valuable whenever full quantitative estimates of probabilities are required, as is the case in almost all applications, for example in physics, chemistry, biology, and engineering.

From a physical point of view the prefactor formulas given above represent an explicit evaluation of the fluctuation determinant. In principle, these ideas have been formulated multiple times in various contexts (compare Section 1.3), from quantum field theory over linear-quadratic control to calculus of variations. As noted by Schulman [46]:

> *"Methods for handling the quadratic Lagrangian are legion and have been well developed since the earliest work on path integrals. Oddly enough, papers on the subject continue to appear and may give some historian of science material for a case history on the nondiffusion of knowledge."*

Importantly, though, numerical methods for the calculation of prefactors in the general setup we consider here remain, to the best of our knowledge, mostly nonexistent. Concretely, for the prefactor terms we (i) formulate algorithms suitable for their explicit numerical computation, (ii) phrase them for the more general class of Lagrangians encountered in LDT, and (iii) treat the case of the invariant measure and infinite time horizon. Computations on stochastic partial differential equations highlight the fact that our results produce correct results even in the irreversible infinite dimensional case, and are efficient enough to be used for quantitative estimates of probabilistic quantities in regimes where direct sampling is inaccessible.

ACKNOWLEDGMENTS

TG wants to thank Timo Schorlepp and Rainer Grauer for helpful discussions. The work of TG was supported by the EPSRC projects EP/T011866/1 and EP/V013319/1. The work of TS was supported by the NSF grants DMS-1522737 and DMS-2012548 as well as the PSC-CUNY grant TRADB-51-281. The work of EVE was supported by National Science Foundation (NSF) Materials Research Science and Engineering Center Program grant DMR-1420073, NSF grant DMS-1522767, and the Simons Collaboration on Wave Turbulence, Grant No. 617006.

REFERENCES

1. S. Arrhenius, *Ueber die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Saeuren*, Z. Phys. Chem. **4** (1889), 226–248.
2. R. Azencott, *Petites perturbations aléatoires des systèmes dynamiques: développements asymptotiques*, Bull. Sci. Math. **109** (1985), no. 3, 253–308.
3. G. Barles and B. Perthame, *Exit time problems in optimal control and vanishing viscosity method*, SIAM J. Control Optim. **26** (1988), no. 5, 1133–1148.
4. F. Barret, *Sharp asymptotics of metastable transition times for one dimensional SPDEs*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 1, 129–166.
5. N. Berglund, *Kramers' law: Validity, derivations and generalisations*, Markov Process. Related Fields **19** (2013), no. 3, 459–490.
6. N. Berglund, G. Di Gesu, and H. Weber, *An Eyring-Kramers law for the stochastic Allen-Cahn equation in dimension two*, arXiv:1604.05742 [math-ph], 2016.
7. N. Berglund and B. Gentz, *Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers' law and beyond*, Electron. J. Probab. **18** (2013).

8. G. A. Bonaschi and M. A. Peletier, *Quadratic and rate-independent limits for a large-deviations functional*, Contin. Mech. Thermodyn. **28** (2016), no. 4, 1191–1219.
9. F. Bouchet and J. Reygner, *Generalisation of the Eyring–Kramers transition rate formula to irreversible diffusion processes*, Ann. Henri Poincaré **17** (2016), no. 12, 3499–3532.
10. F. Bouchet and J. Reygner, *Path integral derivation and numerical computation of large deviation prefactors for non-equilibrium dynamics through matrix Riccati equations*, arXiv:2108.06916 [cond-mat], 2021.
11. A. Bovier and F. den Hollander, *Metastability*, Grundlehren der mathematischen Wissenschaften, vol. 351, Springer International Publishing, Cham, 2015.
12. A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, *Metastability in reversible diffusion processes I: Sharp asymptotics for capacities and exit times*, J. Eur. Math. Soc. **6** (2004), no. 4, 399–424.
13. F. den Hollander, *Metastability under stochastic dynamics*, Stochastic Process. Appl. **114** (2004), no. 1, 1–26.
14. G. V. Dunne, *Functional determinants in quantum field theory*, J. Phys. A **41** (2008), no. 30, 304006.
15. W. E , W. Ren, and E. Vanden-Eijnden, *String method for the study of rare events*, Phys. Rev. B **66** (2002), no. 5, 052301.
16. W. E , W. Ren, and E. Vanden-Eijnden, *Minimum action method for the study of rare events*, Comm. Pure Appl. Math. **57** (2004), no. 5, 637–656.
17. W. E , W. Ren, and E. Vanden-Eijnden, *Simplified and improved string method for computing the minimum energy paths in barrier-crossing events*, J. Chem. Phys. **126** (2007), 164103.
18. R. S. Ellis, *Entropy, large deviations, and statistical mechanics*, Springer-Verlag, Berlin, Heidelberg, 2006.
19. H. Eyring, *The activated complex in chemical reactions*, J. Chem. Phys. **3** (1935), no. 2, 107–115.
20. W. G. Faris and G. Jona-Lasinio, *Large fluctuations for a nonlinear heat equation with noise*, J. Phys. A **15** (1982), no. 10, 3025.
21. W. H. Fleming and M. R. James, *Asymptotic series and exit time probabilities*, Ann. Probab. **20** (1992), no. 3, 1369–1384.
22. H. C. Fogedby and W. Ren, *Minimum action method for the Kardar-Parisi-Zhang equation*, Phys. Rev. E **80** (2009), no. 4, 041116.
23. R. Forman, *Functional determinants and geometry*, Invent. Math. **88** (1987), no. 3, 447–493.
24. M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, Springer, Heidelberg, 2012.
25. C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, Springer, Berlin, 2009.
26. I. M. Gel'fand and A. M. Yaglom, *Integration in functional spaces and its applications in quantum physics*, J. Math. Phys. **1** (1960), no. 1, 48–69.
27. D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem. **81** (1977), no. 25, 2340–2361.
28. T. Grafke, R. Grauer, and T. Schäfer, *The instanton method and its numerical implementation in fluid mechanics*, J. Phys. A **48** (2015), no. 33, 333001.
29. T. Grafke, R. Grauer, T. Schäfer, and E. Vanden-Eijnden, *Arclength parametrized hamilton's equations for the calculation of instantons*, Multiscale Model. Simul. **12** (2014), no. 2, 566–580.
30. T. Grafke, R. Grauer, and S. Schindel, *Efficient computation of instantons for multi-dimensional turbulent flows with large scale forcing*, Commun. Comput. Phys. **18** (2015), no. 3, 577–592.
31. T. Grafke, T. Schäfer, and E. Vanden-Eijnden, Long term effects of small random perturbations on dynamical systems: Theoretical and computational tools, In Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science, Fields Institute Communications, Springer, New York, NY, 2017, pp. 17–55.
32. T. Grafke and E. Vanden-Eijnden, *Numerical computation of rare events via large deviation theory*, Chaos **29** (2019), no. 6, 063118.
33. J. Grasman and O. van Herwaarden, *Asymptotic Methods for the Fokker-Planck Equation and the Exit Problem in Applications*, Springer, Berlin, 1999.
34. M. Heymann and E. Vanden-Eijnden, *The geometric minimum action method: A least action principle on the space of curves*, Comm. Pure Appl. Math. **61** (2008), no. 8, 1052–1117.
35. M. Heymann and E. Vanden-Eijnden, *Pathways of maximum likelihood for rare events in nonequilibrium systems: Application to nucleation in the presence of shear*, Phys. Rev. Lett. **100** (2008), no. 14, 140601.
36. A. Kassam and L. Trefethen, *Fourth-order time-stepping for stiff PDEs*, SIAM J. Sci. Comput. **26** (2005), no. 4, 1214–1233.

37. Y. I. Kifer, *On the asymptotics of the transition density of processes with small diffusion*, Theory Probab. Appl. **21** (1977), no. 3, 513–522.
38. H. A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica **7** (1940), no. 4, 284–304.
39. S. Levit and U. Smilansky, *A new approach to Gaussian path integrals and the evaluation of the semiclassical propagator*, Ann. Phys. **103** (1977), no. 1, 198–207.
40. R. Maier and D. Stein, *Limiting exit location distributions in the stochastic exit problem*, SIAM J. Appl. Math. **57** (1997), no. 3, 752–790.
41. B. J. Matkowsky and Z. Schuss, *The exit problem for randomly perturbed dynamical systems*, SIAM J. Appl. Math. **33** (1977), no. 2, 365–382.
42. J. C. Mattingly, A. M. Stuart, and D. J. Higham, *Ergodicity for sdes and approximations: locally lipschitz vector fields and degenerate noise*, Stochastic Process. Appl. **101** (2002), no. 2, 185–232.
43. E. W. Montroll, *Markoff chains, Wiener integrals, and quantum theory*, Comm. Pure Appl. Math. **5** (1952), no. 4, 415–453.
44. M. Schilder, *Some asymptotic formulas for wiener integrals*, Trans. Amer. Math. Soc. **125** (1966), no. 1, 63–85.
45. T. Schorlepp, T. Grafke, and R. Grauer, *Gel'fand-Yaglom type equations for calculating fluctuations around Instantons in stochastic systems*, J. Phys. A: Math. Theor. **54** (2021), no. 23, 235003.
46. L. S. Schulman, *Techniques and Applications of Path Integration*, John Wiley & Sons, 1981.
47. A.-S. Sznitman, Topics in propagation of chaos, In P.-L. Hennequin, (ed.) Ecole d'Eté de Probabilités de Saint-Flour XIX—1989, Springer Berlin Heidelberg, Berlin, Heidelberg, 1991, pp. 165–251.
48. M. U. Thomas, *Some mean first-passage time approximations for the Ornstein-Uhlenbeck process*, J. Appl. Probab. **12** (1975), no. 3, 600–604.
49. A. Vainshtein, V. I. Zakharov, V. Novikov, and M. A. Shifman, *ABC of instantons*, Sov. Phys. Usp. **25** (1982), 195.
50. E. Vanden-Eijnden and M. Heymann, *The geometric minimum action method for computing minimum energy paths*, Jour. Chem. Phys. **128** (2008), 061103.
51. E. Vanden-Eijnden and J. Weare, *Rare event simulation of small noise diffusions*, Comm. Pure Appl. Math. **65** (2012), no. 12, 1770–1803.
52. S. R. S. Varadhan, *Large deviations*, Ann. Prob. **36** (2008), no. 2, 397–419.
53. X. Zhou, W. Ren, and W. E , *Adaptive minimum action method for the study of rare events*, J. Chem. Phys. **128** (2008), 104111.

APPENDIX A: THE GEOMETRIC INSTANTON AND RICCATI EQUATIONS

A.1 | Geometric instanton equations

Here we give some additional information about the geometric instanton equations (2.36), referring the reader to [29, 31, 32, 34] for more information. Recall that these equations read

$$\begin{cases} \lambda\hat{\phi}' = b(\hat{\phi}) + a\hat{\theta} \\ \lambda\hat{\theta}' = -(\nabla b(\hat{\phi}))^\top\hat{\theta} \end{cases} \tag{A.1}$$

with boundary conditions $\hat{\phi}(0) = x_*$, $\hat{\theta}(1) = \nabla f(\hat{\phi}(1))$ as in (2.36), or $\hat{\phi}(0) = x_*$, $\hat{\phi}(1) = y$ as in (3.33). Here $\lambda(s) = ds/dt$ is the reparametrization factor from physical time $t \in [-\infty, 0]$ to normalized arc-length $s \in [0, 1]$, and primes denote derivatives with respect to s. In practice, these equations can be solved globally by a relaxation method [34]. This amounts to solving the first equation in $a\hat{\theta}$ and inserting the result in the second equation multiplied by a to deduce

$$0 = \lambda^2\hat{\phi}'' + \lambda\lambda'\hat{\phi}'' - \lambda a\big(a^{-1}\nabla b(\hat{\phi}) - (a^{-1}\nabla b(\hat{\phi}))^\top\big)\hat{\phi}' - a(\nabla b(\hat{\phi}))^\top a^{-1} b(\hat{\phi}) \tag{A.2}$$

with boundary conditions $\hat{\phi}(0) = x_*$, $a\hat{\phi}'(1) = \nabla f(\hat{\phi}(1))$ or $\hat{\phi}(0) = x_*$, $\hat{\phi}(1) = y$. We can now find the stable fixed points of this equation by introducing an artificial relaxation time τ and evolving

$\{\phi(s) : s \in (0,1)\}$ globally from some initial guess via

$$\partial_\tau \hat{\phi} = \lambda^2 \hat{\phi}'' + \lambda\lambda' \hat{\phi}'' - \lambda a\big(a^{-1}\nabla b(\hat{\phi}) - (a^{-1}\nabla b(\hat{\phi}))^\top\big)\hat{\phi}' - a(\nabla b(\hat{\phi}))^\top a^{-1} b(\hat{\phi}). \tag{A.3}$$

while enforcing $\hat{\phi}(0) = x_*$, $a\hat{\phi}'(1) = \nabla f(\hat{\phi}(1))$ or $\hat{\phi}(0) = x_*$, $\hat{\phi}(1) = y$ for all $\tau \geq 0$. This is the procedure that was used in this paper to solve the geometric instanton equations.

Alternatively, in situations where we want to solve (A.2) with boundary condition $\hat{\phi}(0) = x_*$, $\hat{\theta}(1) = \nabla f(\hat{\phi}(1))$ we can also iteratively solve the equation for $\hat{\phi}$ forward in s, then that for $\hat{\theta}$ backward in time forward-integrate the $\hat{\phi}$-equation [29, 31]. In this case, we encounter a similar problem as with the equation for $\hat{Q}$ in (3.3): at $s = 0$, corresponding to physical time $t = -\infty$ when the instanton starts at the fixed-point $x_* = \hat{\phi}(0)$, one has $b(\hat{\phi}(0)) = b(x_*) = 0$ and $\hat{\theta}(0) = 0$, as well as $\lambda(0) = 0$. This means that the equation for $\hat{\phi}$ in (A.1) reads $0 = 0$ at $s = 0$ and it is *a priori* unclear how to start the integration. Here too, this problem is only apparent since the equation for $\hat{\phi}$ can be written as

$$\hat{\phi}'(s) = \lambda^{-1}(s)(b(\hat{\phi}(s)) + a\hat{\theta}(s)) \tag{A.4}$$

and a direct application of l'Hôpital rule shows that $\hat{\phi}'(0) = \lim_{s\to 0}\phi'(s)$ satisfies

$$\hat{\phi}'(0) = [\lambda'(0)]^{-1}(\nabla b(x_*)\hat{\phi}'(0) + a\hat{\theta}'(0)) \tag{A.5}$$

where we used $\hat{\phi}(0) = x_*$. Noticing that the matrix $(\lambda'(0)\mathrm{Id} - \nabla b(x_*))$ is invertible since $\lambda'(0) > 0$ and $-\nabla b(x_*)$ is positive-definite as x_* is a stable fixed point of $\dot{x} = b(x)$ by assumption, this equation has the unique solution

$$\hat{\phi}'(0) = (\lambda'(0)\mathrm{Id} - \nabla b(x_*))^{-1} a\hat{\theta}'(0). \tag{A.6}$$

This solution can be computed numerically since we have access to $\lambda'(0)$ and $\hat{\theta}'(0)$ through our knowledge of $\lambda(s)$ and $\hat{\theta}(s)$ from the last backward integration. In particular, since $\lambda(0) = 0$ and $\hat{\theta}(0) = 0$, we can write, for arc-length stepsize $\Delta s > 0$,

$$\hat{\theta}'(0) = \hat{\theta}(\Delta s)/\Delta s + O(\Delta s), \qquad \lambda'(0) = \lambda(\Delta s)/\Delta s + O(\Delta s). \tag{A.7}$$

Once $\hat{\phi}'(0)$ has been computed, it can be used together with $\hat{\phi}(0) = x_*$ to start off the numerical integration of the $\hat{\phi}$-equation forward in time using

$$\hat{\phi}(\Delta s) = x_* + \hat{\phi}'(0)\Delta s + O(\Delta s^2) = x_* + (\lambda(\Delta s)\mathrm{Id} - \Delta s\nabla b(x_*))^{-1} a\hat{\theta}(0) + O(\Delta s^2). \tag{A.8}$$

A.2 | Geometric Riccati equations

As discussed in remark 3.4, we can similarly derive the initial arc-length derivative $\hat{Q}_y'(0)$ by solving an additional Lyapunov equation. Concretely, let us start from

$$\hat{Q}_y' = \lambda^{-1}\big(\hat{Q}_y K_y \hat{Q}_y + \hat{Q}_y(\nabla b(\hat{\phi}_y))^\top + (\nabla b(\hat{\phi}_y))\hat{Q}_y + a\big) \equiv \lambda^{-1}\mathfrak{R}. \tag{A.9}$$

Since $\mathfrak{R}(0) = 0$ and $\lambda(0) = 0$, via l'Hôpital's rule it follows that $\hat{Q}_y'(0) = \lambda'(0)^{-1}\mathfrak{R}'(0)$., which can be written explicitly as

$$\mathfrak{C}\,\hat{Q}_y'(0) + \hat{Q}_y'(0)\,\mathfrak{C}^\top = \mathfrak{K}. \tag{A.10}$$

where we defined

$$\mathfrak{C} = \frac{1}{2}\lambda'(0)\mathrm{Id} - \nabla b(x_*)\mathfrak{K} = Q_*(\nabla\nabla b(x_*)\hat{\theta}'_y(0))Q_* + Q_*(\nabla\nabla b(x_*)\hat{\phi}'_y(0))^\top + (\nabla\nabla b(x_*)\hat{\phi}'_y(0))Q_*. \tag{A.11}$$

If $\nabla\nabla b(x_*) = 0$, we have $\mathfrak{K} = 0$ and hence one obtains $\hat{Q}'_y(0) = 0$—this is the case for example in the examples presented in Sections 3.2.2 or 4.2.1. In this situation, the approximation $\hat{Q}_y(\Delta s) = Q_*$ is correct up to order $\mathcal{O}(\Delta s^2)$. If a more accurate approximation is needed, we can consider the next order: Taking the derivative of (A.9) we deduce that

$$\hat{Q}''_y(s) = \lambda(s)^{-1}\mathfrak{R}'(s) - \lambda(s)^{-2}\mathfrak{R}(s)\lambda'(s), \tag{A.12}$$

and hence, using l'Hôpital's rule again together with $\mathfrak{R}(0) = \mathfrak{R}'(0) = 0$ and $\lambda(0) = 0$, we obtain

$$\hat{Q}''_y(0) = \frac{1}{2}\lambda'(0)^{-1}\mathfrak{R}''(0). \tag{A.13}$$

For brevity, we will refrain to write down this equation explicitly, but we note that it is again a Lyapunov equation, this time for $\hat{Q}''_y(0)$, where all other terms are known. Knowledge of $\hat{Q}''_y(0)$ in situations where $\hat{Q}'_y(0) = 0$ allows one to use

$$\hat{Q}_y(\Delta s) = Q_* + \frac{1}{2}\Delta s^2\hat{Q}''_y(0) + \mathcal{O}(\Delta s)^3. \tag{A.14}$$

APPENDIX B: EXPRESSIONS FOR THE SOLUTION OF RICCATI EQUATIONS AS EXPECTATIONS

The following two propositions give ways to express the solution of backward and forward Riccati equations of the type considered in text in terms of expectations over the solution of some SDE:

Proposition B.1. *Given $A : [0,T] \to \mathbb{R}^{n\times n}$, $B : [0,T] \to \mathbb{R}^{n\times n}$, $\eta : [0,T] \to \mathbb{R}^n$, all in $C^1([0,T])$ and with A symmetric, as well as $C \in \mathbb{R}^{n\times n}$, with C symmetric, and $\xi \in \mathbb{R}^n$, the following equality holds*

$$\begin{aligned}\mathbb{E}^z \exp\left(\int_0^T \left(\frac{1}{2}A(t) : Z_tZ_t + \eta(t)\cdot Z_t\right)dt + \frac{1}{2}C : Z_TZ_T + \xi\cdot Z_T\right)\\ = G(0)\exp\left(\frac{1}{2}\langle z, W(0)z\rangle + r(0)\cdot z\right),\end{aligned} \tag{B.1}$$

where Z_t solves the linear SDE

$$dZ_t = B(t)Z_t dt + \sigma dW_t\,; \tag{B.2}$$

and $W : [0,T] \to \mathbb{R}^{n\times n}$, $r : [0,T] \to \mathbb{R}^n$, $G : [0,T] \to \mathbb{R}$, solve

$$\begin{cases}\dot{W} + B^T(t)W + WB(t) + WaW + A(t) = 0, & W(T) = C,\\ \dot{r} + B^T(t)r + War + \eta(t) = 0, & r(T) = \xi,\\ \dot{G} + \frac{1}{2}\mathrm{tr}(aW)G + \frac{1}{2}a : rrG = 0, & G(T) = 1.\end{cases} \tag{B.3}$$

Note that the solution to the equation for $G(t)$ in (B.3) can be expressed as

$$G(t) = \exp\left(\frac{1}{2}\int_t^T (\mathrm{tr}(aW(s)) + a : r(s)r(s))ds\right). \tag{B.4}$$

Proof. Let $v : [0,T]\times\mathbb{R}^n \to \mathbb{R}$ solve

$$\partial_t v + \langle B(t)z, \nabla v\rangle + \frac{1}{2}a : \nabla\nabla v + \left(\frac{1}{2}A(t) : zz + \eta(t)\cdot z\right)v = 0, \tag{B.5}$$

for the final condition

$$v(T,z) = \exp\left(\frac{1}{2}C : zz + \xi\cdot z\right). \tag{B.6}$$

Then: (i) computing $d(v(t,Z_t)\exp(\int_0^t(\frac{1}{2}A(s) : Z_sZ_s + \eta(s)\cdot Z_s)\,ds))$ via Ito's formula, taking expectation, and integrating on $t\in[0,T]$ shows that the solution to this equation at time $t=0$ can be expressed as the expectation in (B.1); and (ii) substituting $G(t)\exp(\frac{1}{2}\langle z, W(t)z\rangle + r(t)\cdot z)$ in (B.5) shows that this expression satisfies this equation as well as (B.6) if $W(t)$, $r(t)$, and $G(t)$ satisfy (B.3). □

Proposition B.2. *Using the same notations as in Proposition B.1, let $Q : [0,T]\to\mathbb{R}^{n\times n}$ solve*

$$\dot{Q} = B(t)Q + QB^T(t) + QA(t)Q + a, \qquad Q(0) = Q_0, \tag{B.7}$$

for some $Q_0 = Q_0^T$, positive semidefinite (possibly zero). Then

$$Q(t) = \mathbb{E}Z_t^Q(Z_t^Q)^T, \tag{B.8}$$

where Z_t^Q solves the nonlinear (in the sense of McKean) SDE

$$dZ_t^Q = B(t)Z_t^Q dt + \frac{1}{2}QA(t)Z_t^Q dt + \sigma dW_t, \tag{B.9}$$

and the expectation in (B.8) is taken over solutions to (B.9) with initial conditions drawn from a Gaussian distribution with mean zero and covariance Q_0.

Proof. Application of Ito's formula shows that

$$\frac{d}{dt}\mathbb{E}\left[Z_t^Q\left(Z_t^Q\right)^T\right] = B(t)\mathbb{E}\left[Z_t^Q\left(Z_t^Q\right)^T\right] + \mathbb{E}\left[Z_t^Q\left(Z_t^Q\right)^T\right]B^T(t)$$

$$+\frac{1}{2}QA(t)\mathbb{E}\left[Z_t^Q\left(Z_t^Q\right)^T\right] + \frac{1}{2}\mathbb{E}\left[Z_t^Q\left(Z_t^Q\right)^T\right]A(t)Q + a. \tag{B.10}$$

Since $\mathbb{E}\left[Z_0^Q\left(Z_0^Q\right)^T\right] = Q_0 = Q(0)$ initially, this equation shows that $\mathbb{E}\left[Z_t^Q\left(Z_t^Q\right)^T\right] = Q(t)$ for $t\geq 0$. □

The following proposition offers a practical way to simulate (B.9):

Proposition B.3. *Let* $\{Z_t^i\}_{i=1}^n$ *solve*

$$dZ_t^i = B(t)Z_t^i dt + \frac{1}{2}n^{-1}\sum_{j=1}^{n}\left\langle Z_t^i, A(t)Z_t^j\right\rangle Z_t^j dt + \sigma dW_t^i, \qquad i = 1,\dots,n, \tag{B.11}$$

where $\{W_t^i\}_{i=1}^n$ *is a set of independent Wiener processes. Then if we drawn the initial conditions for (B.11) independently from a Gaussian distribution with zero mean and covariance* Q_0*, we have*

$$\frac{1}{n}\sum_{i=1}^{n} Z_t^i \left(Z_t^i\right)^T \to Q(t) \qquad \textit{almost surely as } n \to \infty \tag{B.12}$$

Proof. The proposition is a direct consequence of a 'propagation of chaos' argument (see e.g. [47]) applied to (B.11). □

APPENDIX C: LINK BETWEEN THE FORWARD AND BACKWARD RICCATI EQUATIONS

We have:

Proposition C.1. *Using the same notations as in Proposition B.1, let* $W : [0,T] \to \mathbb{R}^{n\times n}$ *and* $Q : [0,T] \to \mathbb{R}^{n\times n}$ *solve*

$$\begin{cases} \dot{W} + B^T(t)W + WB(t) + WaW + A(t) = 0, & W(T) = W_T, \\ \dot{Q} = QB^T(t) + B(t)Q + a + QA(t)Q = 0, & Q(0) = Q_0, \end{cases} \tag{C.1}$$

where $W_T \in \mathbb{R}^{n\times n}$ *and* $Q_0 \in \mathbb{R}^{n\times n}$*, both symmetric. Then the following identity holds*

$$\det\left(Id - W_T Q(T)\right) = \det\left(Id - W(0)Q_0\right)\exp\left(\int_0^T \operatorname{tr}(A(t)Q(t) - W(t)a)dt\right) \tag{C.2}$$

Proof. From (C.1), it is easy to see that the matrix $W(t)Q(t)$ satisfies

$$\frac{d}{dt}(WQ) = (B^T(t) + Wa)(\mathrm{Id} - WQ) - (\mathrm{Id} - WQ)(B^T(t) + A(t)Q). \tag{C.3}$$

As a result, using the Jacobi formula, we deduce that

$$\frac{d}{dt}\log\det(\mathrm{Id} - WQ) = -\mathrm{tr}\big((\mathrm{Id} - WQ)^{-1}(d/dt)(WQ)\big) = \mathrm{tr}(A(t)Q - Wa). \tag{C.4}$$

Integrating both side on $t \in [0,T]$ and taking the exponential of the result gives (C.2). □

APPENDIX D: THE RADON'S LEMMA FOR THE RICCATI EQUATION

It is well-known that a matrix-Riccati equation can be equivalently represented by a linear matrix equation. Sometimes, this transformation is called Radon's Lemma. Consider a differential

equation of the form

$$\frac{d}{dt}\begin{pmatrix}\Phi\\ \Theta\end{pmatrix}=\begin{pmatrix}M_{11} & M_{12}\\ M_{21} & M_{22}\end{pmatrix}\begin{pmatrix}\Phi\\ \Theta\end{pmatrix}, \tag{D.1}$$

and set $W=\Theta\Phi^{-1}$. Then

$$\dot{W}=\dot{\Theta}\Phi^{-1}+\Theta\dot{\Phi}^{-1}=\dot{\Theta}\Phi^{-1}-\Theta\Phi^{-1}\dot{\Phi}\Phi^{-1}=(M_{21}\Phi+M_{22}\Theta)\Phi^{-1}-\Theta\Phi^{-1}(M_{11}\Phi+M_{12}\Theta)\Phi^{-1}$$
$$=M_{21}+M_{22}W-WM_{11}-WM_{12}W.$$

We can apply this to the instanton matrix-Riccati equation by choosing $M_{21}=-\langle\nabla\nabla b,\theta\rangle$, $M_{22}=-(\nabla b)^T$, $M_{11}=\nabla b$, and $M_{12}=a$ to obtain:

$$\frac{d}{dt}\begin{pmatrix}\Phi\\ \Theta\end{pmatrix}=\begin{pmatrix}\nabla b & a\\ -\langle\nabla\nabla b,\theta\rangle & -(\nabla b)^T\end{pmatrix}\begin{pmatrix}\Phi\\ \Theta\end{pmatrix}, \tag{D.2}$$

and as final conditions we can choose $\Theta(T)=W(T)$ and $\Phi(T)=\mathrm{Id}$. While it seems appealing to solve the Riccati equation this way, in practice the issue is that the equation for Φ is well-posed forward in time, whereas that for Θ is well-posed backward in time. This means that the system (D.2) has to be solved iteratively, and the final condition are not simple to impose. This is why we did not use (D.2) in this paper.

APPENDIX E: DERIVATION OF $\det_\perp \boldsymbol{H}=(\hat{\boldsymbol{n}}^\top \boldsymbol{H}^{-1}\hat{\boldsymbol{n}})\det \boldsymbol{H}$

Given an invertible, positive definite $H=H^\top\in\mathbb{R}^{n\times n}$ and a unit vector $\hat{n}$ we define $\det_\perp H$ via

$$(2\pi)^{(n-1)/2}|\det{}_\perp H|^{-1/2}=\int_P e^{-\frac{1}{2}\langle y,Hy\rangle}d\sigma(y)=:A \tag{E.1}$$

where $P=\{y:\langle\hat{n},y\rangle=0\}$. For $m>0$, let

$$H_m=H+m\hat{n}\hat{n}^\top \tag{E.2}$$

Clearly

$$A=\int_P e^{-\frac{1}{2}\langle y,H_m y\rangle}d\sigma(y) \tag{E.3}$$

since $\langle\hat{n},y\rangle=0$ in P. At the same time we have

$$(2\pi)^{n/2}|\det H_m|^{-1/2}=\int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle u,H_m u\rangle}du=(\hat{t}\cdot\hat{n})\int_P\int_{\mathbb{R}} e^{-\frac{1}{2}\langle(y+s\hat{t}),H_m(y+s\hat{t})\rangle}dsd\sigma(y)$$
$$=(\hat{t}\cdot\hat{n})\int_P e^{-\frac{1}{2}\langle y,H_m y\rangle}d\sigma(y)\int_{\mathbb{R}} e^{-\frac{1}{2}s^2\langle\hat{t},H_m\hat{t}\rangle}ds \tag{E.4}$$

where we used $u = y + s\hat{t}$ to change integration variable with

$$\hat{t} = \frac{H^{-1}\hat{n}}{|H^{-1}\hat{n}|} \quad \Leftrightarrow \quad \hat{n} = \frac{H\hat{t}}{|H\hat{t}|} \tag{E.5}$$

Comparing (E.1) and (E.5) we deduce

$$|\det{}_{\perp} H| = (\hat{n}\cdot\hat{t})^2 |\hat{t}^{\top} H_m \hat{t}|^{-1} |\det H_m| \tag{E.6}$$

Since

$$t^{\top} H_m \hat{t} = \hat{t}^{\top}(H + m\,\hat{n}\,\hat{n}^{\top})\hat{t} = \hat{t}^{\top} H \hat{t} + m(\hat{n}\cdot\hat{t})^2 \tag{E.7}$$

we have

$$(\hat{n}\cdot\hat{t})^2 |t^{\top} H_m \hat{t}|^{-1} = \frac{\hat{n}\cdot\hat{t}}{|H\hat{t}| + m(\hat{n}\cdot\hat{t})} \tag{E.8}$$

and (E.6) can be written as

$$\det{}_{\perp} H = \frac{\hat{n}\cdot\hat{t}}{|H\hat{t}| + m(\hat{n}\cdot\hat{t})} \det H_m \tag{E.9}$$

Next, write (E.2) as

$$H_m = H\big(\mathrm{Id} + mH^{-1}\hat{n}\,\hat{n}^{\top}\big) = H\big(\mathrm{Id} + m|H^{-1}\hat{n}|\,\hat{t}\,\hat{n}^{\top}\big) \tag{E.10}$$

so that

$$\det H_m = \det H \det\big(\mathrm{Id} + m|H^{-1}\hat{n}|\,\hat{t}\,\hat{n}^{\top}\big) \tag{E.11}$$

The matrix $\mathrm{Id} + m|H^{-1}\hat{n}|\hat{t}\hat{n}^{\top}$ has $n-1$ eigenvectors perpendicular to $\hat{n}$, each with eigenvalue 1, and one eigenvector $\hat{t}$ with eigenvalue $1 + m|H^{-1}\hat{n}|(\hat{n}\cdot\hat{t})$ since

$$(\mathrm{Id} + m|H^{-1}\hat{n}|\,\hat{t}\,\hat{n}^{\top})\hat{t} = \big(1 + m|H^{-1}\hat{n}|(\hat{n}\cdot\hat{t})\big)\hat{t} \tag{E.12}$$

Therefore

$$\det{}_{\perp} H = \frac{(\hat{n}\cdot\hat{t})\big(1 + m|H^{-1}\hat{n}|(\hat{n}\cdot\hat{t})\big)}{|H\hat{t}| + m(\hat{n}\cdot\hat{t})} \det H \tag{E.13}$$

Since $|H^{-1}\hat{n}| = |H\hat{t}|^{-1}$ this can be written as

$$\det{}_{\perp} H = \frac{(\hat{n}\cdot\hat{t})}{|H\hat{t}|} |\det H| = (\hat{n}^{\top} H^{-1}\hat{n})\,\det H \tag{E.14}$$

APPENDIX F: GENERAL FORM OF THE INSTANTON AND RICCATI EQUATIONS FOR SPDEs

Let us generalize (6.1) into

$$\partial_t u = \mathcal{B}[u] + \sqrt{\varepsilon}\eta, \quad u(0) = u_0, \tag{F.1}$$

for $t \in [0, \infty)$, $x \in \Omega \subseteq \mathbb{R}^d$, and $u : [0, \infty) \times \Omega \to \mathbb{R}$, and where $\mathcal{B}[u]$ is a (possibly nonlinear) differential operator in the spatial variable x and the noise η is white-in-time Gaussian with covariance

$$\mathbb{E}\eta(t, x)\eta(t', x') = \delta(t - t')\chi(x, x'). \tag{F.2}$$

If we consider again probabilities that a linear observable exceeds a certain threshold,

$$P_\varepsilon^{u_0}(T, z) = \mathbb{P}^{u_0}\left(\int_\Omega \phi(x)u(T, x)\, dx \geq z\right), \tag{F.3}$$

we formally obtain a proposition analogous to Proposition 6.1:

Proposition F.1 (Probabilities for SPDEs – general case). *Let the fields $u(t, x)$, $\theta(t, x)$ solve the instanton equations*

$$\begin{cases} \partial_t u = \mathcal{B}[u] + \displaystyle\int_\Omega \chi(x, y)\theta(t, y)\, dy, & u(0) = u_0, \\ \partial_t \theta = -\displaystyle\int_\Omega \frac{\delta \mathcal{B}[u](t, y)}{\delta u(x)}\theta(t, y)\, dy, & \theta(T) = \phi, \end{cases} \tag{F.4}$$

and let $Q(t, x, y)$ solve

$$\begin{aligned} \partial_t Q = &\int_{\Omega^3} Q(t, x, z_1)\mathcal{K}(t, z_1, z_2)Q(t, z_2, y)\, dz_1\, dz_2\, dz_3 \\ &+ \int_\Omega \frac{\delta \mathcal{B}[u](t, x)}{\delta u(t, z)}Q(t, y, z)\, dz + \int_\Omega \frac{\delta \mathcal{B}[u](t, y)}{\delta u(t, z)}Q(t, z, x)\, dz + \chi(x, y), \end{aligned} \tag{F.5}$$

with $Q(0) = 0$ and where we denote

$$\mathcal{K}(t, x, y) = \int_\Omega \frac{\delta \mathcal{B}[u](t, z)}{\delta u(t, x)\delta u(t, y)}\theta(t, z)\, dz. \tag{F.6}$$

Then the probability $P_\varepsilon^{u_0}(T, z)$ in (F.3) satisfies

$$\lim_{\varepsilon \to 0} \frac{P_\varepsilon^{u_0}(T, z)}{\bar{P}_\varepsilon^{u_0}(T, z)} = 1, \tag{F.7}$$

where

$$P_\varepsilon^{u_0}(T,z) = (2\pi)^{-1/2}\varepsilon^{1/2}\mathcal{V}(T,u_0)\exp\left(-\frac{1}{2\varepsilon}\int_0^T\int_{\Omega^2}\theta(t,x)\chi(x,y)\theta(t,y)\,dx\,dy\,dt\right) \tag{F.8}$$

with

$$\mathcal{V}(T,u_0) = \left(\int_{\Omega^2}\theta(T,x)\mathcal{Q}(T,x,y)\theta(T,y)\,dx\,dy\right)^{1/2}\exp\left(\frac{1}{2}\int_0^T\int_\Omega \mathcal{K}(t,x,x)\mathcal{Q}(t,x,x)\,dx\,dt\right). \tag{F.9}$$