Computing Large Deviation Rate Functions of Entropy Production for Diffusion Processes in the Vanishing-Noise Limit and High Dimensions by an Interacting Particle Method

Zhizhang Wu[a], Renaud Raquépas[b], Jack Xin[c], Zhiwen Zhang[a]

[a] *Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China.*
[b] *Courant Institute, New York University, New York, NY 10012, United States.*
[c] *Department of Mathematics, University of California at Irvine, Irvine, CA 92697, United States.*

Abstract

We study an interacting particle method (IPM) for computing the large deviation rate function of entropy production for diffusion processes, with emphasis on the vanishing-noise limit and high dimensions. The crucial ingredient to obtain the rate function is the computation of the principal eigenvalue λ of elliptic, non-self-adjoint operators. We show that this principal eigenvalue can be approximated in terms of the spectral radius of a discretized evolution operator obtained from an operator splitting scheme and an Euler–Maruyama scheme with a small time step size, and we show that this spectral radius can be accessed through a large number of iterations of this discretized semigroup, suitable for the IPM. The IPM applies naturally to problems in unbounded domains, scales easily to high dimensions, and adapts to singular behaviors in the vanishing-noise limit. We show numerical examples in dimensions up to 16. The numerical results show that our numerical approximation of λ converges to the analytical vanishing-noise limit with a fixed number of particles and a fixed time step size. Our paper appears to be the first one to obtain numerical results of principal eigenvalue problems for non-self-adjoint operators in such high dimensions.

Keywords interacting particle methods, principal eigenvalues, large deviation rate functions, entropy production, vanishing-noise limits, high dimensions.

AMS subject classifications 37M25, 47D08, 60F10, 82C31

1. Introduction

The problem we are interested in concerns the time reversibility of diffusion processes, as famously studied by Kolmogorov as early as 1937 [39]. He found among other things that, with V a smooth potential function and b a non-conservative smooth vector field, stochastic differential equations (SDEs) in $\mathbb{R}^d$ of the form

$$\begin{cases} \mathrm{d}X_t = -\nabla V(X_t)\,\mathrm{d}t + b(X_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}B_t, \\ X_0 \sim \mu \end{cases} \tag{1.1}$$

Email addresses: wuzz@hku.hk (Zhizhang Wu), rr4374@nyu.edu (Renaud Raquépas), jack.xin@uci.edu (Jack Xin), zhangzw@hku.hk (Zhiwen Zhang)

are invariant under time reversal only when $b = 0$ and the density of the initial measure μ is proportional to $\exp(-\varepsilon^{-1}V)$; see Section 2 for precise statements of the assumptions on V and b we will work with. While time-reversed models have a long history of applications to fields such as signal processing [48, 55] and electric circuit theories [2, 3], and have been adopted in recent years as a way to generate high-quality images in computer vision [56, 57], we will focus on questions from stochastic thermodynamics. When the time reversal of a diffusion process is still a diffusion process [1, 35], a natural question is how distinguishable the two processes are, i.e. how irreversible the original diffusion is. One classical way to quantify irreversibility is to compute an observable called entropy production. In the large-time limit or the steady-state regime, the entropy production for (1.1) can be computed through the Clausius-like entropy (Stratonovich) integral

$$S_t^\varepsilon = \frac{1}{\varepsilon} \int_0^t \langle\, b(X_s), \circ\, \mathrm{d}X_s \rangle, \tag{1.2}$$

which in the language of statistical thermodynamics is the work done by the non-conservative part of the drift force in (1.1), rescaled by temperature [41, 44]. Here, the definition and physical interpretation of the entropy production (1.2) for (1.1) rely on the interpretation as a small-mass approximation (a.k.a. Kramers–Smoluchowski limit); it should be adapted in a natural way in the presence of momentum variables, which should change sign under time reversal; see e.g. [20, 37, 44]. We refer the readers to [19] for a discussion of other decompositions of the drift force and to [37, 54] for a rigorous comparison with other measures of irreversibility, including the point of view of hypothesis testing of the arrow of time. The study of these different notions of entropy production — and more precisely of their large deviations — has driven important theoretical progress in non-equilibrium statistical physics since the 1990s; see e.g. [11, 22, 23, 29, 41, 44, 64]. One key feature of the theory of entropy production is that the positivity of the mean entropy production rate is considered as a key signature of steady non-equilibrium phenomena.

Let $\mathrm{Prob}^{\mu,\varepsilon}$ refer to the law for the solution of (1.1) starting from an initial measure μ, which we assume for simplicity to have a smooth, positive, rapidly decaying density with respect to the Lebesgue measure on $\mathbb{R}^d$. The large deviation rate function $I^\varepsilon : \mathbb{R} \to [0, \infty]$ in this problem is the function that gives the exponential rate of decay in t of fluctuations of order t in S_t^ε,

$$\mathrm{Prob}^{\mu,\varepsilon}\{t^{-1}S_t^\varepsilon \approx s\} \asymp \exp\left(-tI^\varepsilon(s)\right) \tag{1.3}$$

as $t \to \infty$; see Section 2 for a more precise formulation of the large deviation principle. We are interested in an efficient way of numerically computing this rate function.

Before we discuss numerical considerations, let us briefly explain how the rate function is related to an eigenvalue computation. The moment-generating function of S_t^ε with respect to $\mathrm{Prob}^{\mu,\varepsilon}$ is

$$\chi_t^\varepsilon(\alpha) = \int_{C_t} \exp(-\alpha S_t^\varepsilon)\, \mathrm{dProb}^{\mu,\varepsilon}, \tag{1.4}$$

where $\alpha \in \mathbb{R}$ and C_t is the space $C([0,t];\mathbb{R}^d)$ of continuous paths in $\mathbb{R}^d$ over the time interval $[0,t]$. Under our assumptions, the following Feynman–Kac representation of the moment-generating function $\chi_t^\varepsilon(\alpha)$ holds:

$$\chi_t^\varepsilon(\alpha) = \int_{\mathbb{R}^d} \left(\exp(t\,\mathscr{A}^{\varepsilon,\alpha})\mathbb{1}\right)(\xi)\,\mathrm{d}\mu(\xi)\,, \tag{1.5}$$

where the operator $\mathscr{A}^{\varepsilon,\alpha}$ is a second-order differential operator that is elliptic but not self-adjoint. Such a representation dates at least back to [41, 44] and relies on Girsanov's theorem and the Feynman–Kac formula; we refer to [7, 54] for rigorous proofs that cover our hypotheses. With $\lambda^{\varepsilon,\alpha}$ the principal eigenvalue (the one with the largest real part) of $\mathscr{A}^{\varepsilon,\alpha}$, the identity

$$\lim_{t\to\infty} \frac{1}{t} \log \chi_t^\varepsilon(\alpha) = \lambda^{\varepsilon,\alpha} \tag{1.6}$$

provides a spectral-theoretic point of view on the large-t behavior of the moment-generating function, which is instrumental in the study of large deviations. Hence, the moment-generating function is of course convex in α and it is symmetric about $\alpha = \frac{1}{2}$. The spectral-theoretic point of view provides tools for showing smoothness in α. The Legendre transform of $\lambda^{\varepsilon,\alpha}$ in the variable α is the large deviation rate function I^ε in (1.3):

$$I^\varepsilon(s) = \sup_\alpha \left(-\alpha s - \lambda^{\varepsilon,\alpha}\right). \tag{1.7}$$

The symmetry about $\alpha = \frac{1}{2}$ gives rise to the Gallavotti–Cohen symmetry $I^\varepsilon(-s) = I^\varepsilon(s) + s$. In sufficiently regular situations, many statistical properties of the family $(S_t^\varepsilon)_{t>0}$ can be equivalently read off the limiting cumulant-generating function $\lambda^{\varepsilon,\alpha}$ or off the rate function $I^\varepsilon(s)$. For example, the asymptotic mean entropy production per unit time is both $-\partial_\alpha \lambda^{\varepsilon,\alpha}|_{\alpha=0}$ and the zero of I^ε. Again, we refer to [7, 37, 54] for proofs and more thorough theoretical discussions.

There are several motivations for seeking novel numerical methods for accessing I^ε via $\lambda^{\varepsilon,\alpha}$. First, trying to probe the large deviations of S_t^ε from direct simulations of (1.1) and computation of (1.2) is not realistic since these large deviations are events with exponentially small probabilities. In most cases where rigorous theorems on entropy production are proved, $\lambda^{\varepsilon,\alpha}$ is the only available access to the rate function I^ε, but admits no closed-form formula. Second, the assumptions for these theorems are relatively stringent — most significantly by the non-degeneracy assumption on the noise — and we are looking for ways to explore the large deviations in situations where no rigorous results are available. We will be particularly interested in the small-noise regime $0 < \varepsilon \ll 1$ since, under additional assumptions at the critical points of V, [54] provides explicit formulas for the limits $\lambda^{0,\alpha} = \lim_{\varepsilon\to 0^+} \lambda^{\varepsilon,\alpha}$ for α in an interval of the form $(-\delta, 1+\delta)$ and $I^0(s) = \lim_{\varepsilon\to 0^+} I^\varepsilon(s)$ for s in an interval around the mean entropy production rate, allowing us to compare our numerical results. In the presence of momentum variables, the vanishing-noise limit has attracted independent interest in the physics literature since [42], due to its relation to deterministic systems; it still does to this day [8, 54, 51]. We will come back to this point in Section 2.

In this paper, we study an interacting particle method (IPM) [13, 18, 25, 33, 46] for numerically computing $\lambda^{\varepsilon,\alpha}$ — and thus $I^{\varepsilon}(s)$ — at $0 < \varepsilon \ll 1$. More precisely, we consider an α- and ε-dependent, discrete-time semigroup obtained from an operator splitting and an Euler–Maruyama scheme with a time step size Δt, and then show that the spectral radius associated with this discrete-time semigroup has the following two properties:

- on the one hand, it is accessible through large iterates of the semigroup and lends itself to the IPM, thanks to suitable stability properties [24];
- on the other hand, it provides a good approximation of $\lambda^{\alpha,\varepsilon}$ for small Δt, thanks to different results from (non-self-adjoint) perturbation theory [5, 38, 63].

We also discuss techniques for setting the measure of initial conditions to obtain faster approximations of this spectral radius.

To put things into perspective, let us briefly discuss the computational difficulties. The following three issues pose great challenges to traditional mesh-based numerical methods such as finite element methods [58] and finite difference methods [9, 43].

1. *Unboundedness of the physical domain*: Since the stochastic dynamics (1.1) is defined in all of $\mathbb{R}^d$, truncation of the domain is usually needed in mesh-based methods [34], and this may introduce numerical errors.
2. *High dimensionality*: In having in mind applications to stochastic thermodynamics in which the dimension d of X_t in (1.1) is proportional to the number of particles, we would like to be able to handle situations where d is large, but most mesh-based methods suffer from the curse of dimensionality.
3. *Singularities in the vanishing-noise limit*: With $\psi^{\varepsilon,\alpha}$ the normalized principal eigenfunction, it is known from [26] that $\varepsilon \log \psi^{\varepsilon,\alpha}$ has a nontrivial limit as $\varepsilon \to 0^+$ under certain conditions. This implies that $\psi^{\varepsilon,\alpha}$ is asymptotically proportional to $\exp(-\varepsilon^{-1}\Psi^{\alpha})$ for some function Ψ^{α} and thus admits singularities in the vanishing-noise limit. For mesh-based methods, finer grids are needed in order to capture the singularity.

On the other hand, the IPM provides an alternative to the computation of $\lambda^{\varepsilon,\alpha}$ from the perspective of Feynman–Kac semigroups, which has already been applied to the computation of ground state energies of Schrödinger operators using Diffusion Monte Carlo [4, 10, 28, 31], to the computation of effective diffusivity [49, 66, 67, 68] and KPP front speeds [50, 70], and to non-linear filtering problems [12, 14, 15], to mention only a few. Since the IPM is based on simulations of an SDE, it naturally applies to unbounded domains and it is independent of whether the operator whose principal eigenvalue is sought is self-adjoint. In addition, the IPM is essentially a Monte Carlo scheme for approximating a Feynman–Kac semigroup, and together with resampling as a way of controlling variance it is expected to be free of the curse of dimensionality [47]. Also, the IPM scales easily to high dimensions in terms of coding.

In numerical examples with different values of d up to 16, the numerical approximation of $\lambda^{\varepsilon,\alpha}$ converges to its predicted vanishing-noise limit with a fixed number of particles and a fixed time step size, which shows the scalability and robustness of our method for large d and small ε. In these examples, the computational time spent grows linearly with respect to d and does not significantly change with respect to ε when other numerical parameters are fixed, including the number of particles, time step size, and final time. Moreover, the empirical density of the particles we obtain at the final time (after resampling) accurately captures singularities of the vanishing-noise limit, which is compatible with [26]. We point out here that Feynman–Kac semigroups have a long history in large deviation theory; see e.g. [16, 17, 40, 65, 69, 62] for theoretical literature and [30, 36, 45, 52, 53, 59] for numerical literature. However, there are few theoretical results on the properties of the numerical method and few numerical results of the challenging high-dimensional case in the existing works.

The rest of this paper is organized as follows. In Section 2, we present the Feynman–Kac semigroup formulation of the principal eigenvalue problem and the formulation of the large deviation principle. In Section 3, we introduce the discrete-time semigroup at the heart of our numerical approximation and present our theoretical results on the corresponding spectral radius. In Section 4, we present the interacting particle algorithm and techniques for setting the initial measure. In Section 5, we show numerical examples in dimensions up to 16 and compare them with the analytical vanishing-noise limits to find excellent agreement with explicit theoretical predictions. Numerical experiments also allow us to probe situations for which we are not aware of explicit theoretical predictions for the vanishing-noise limit. Finally, we give some concluding remarks in Section 6.

Notation. Let $\mathcal{P}(\mathbb{R}^d)$ be the space of all probability measures over $\mathbb{R}^d$. For a measure μ with finite mass, let $(\mu, \varphi) = \int_{\mathbb{R}^d} \varphi \, \mathrm{d}\mu$ for any $\varphi \in L^\infty(\mathbb{R}^d)$. We use C_0 for the space of continuous real-valued functions on $\mathbb{R}^d$ that vanish at infinity, and given a function $W : \mathbb{R}^d \to [1, +\infty)$, we use the notation

$$L^\infty_W(\mathbb{R}^d) = \left\{ \varphi \in L^\infty_{\mathrm{loc}}(\mathbb{R}^d) : \left\| \frac{\varphi}{W} \right\|_{L^\infty(\mathbb{R}^d)} < +\infty \right\}. \tag{1.8}$$

We use $|\cdot|$ for the Euclidean norm on $\mathbb{R}^d$ and $\|\cdot\|$ for the operator norm it induces on d-by-d matrices.

2. Continuous-time Feynman–Kac semigroups and large deviations

This section, together with Section 3, serves to show that the principal eigenvalue $\lambda^{\alpha,\varepsilon}$ of $\mathscr{A}^{\alpha,\varepsilon}$ in (1.4)–(1.6) can be approximated using a discretized procedure, which will then be combined with resampling to yield our IPM in Section 4. Before proceeding further, we make some assumptions on V and b, trying to strike a balance between optimality and readability.

Assumption 2.1. *We assume that (1) $V \in C^\infty(\mathbb{R}^d)$; (2) there exists a positive-definite matrix H_0 such that $\langle x, H_0 \nabla V(x)\rangle \geq |x|^2$ whenever $|x|$ is large enough; (3) $\|D^2 V(x)\| = o(|\nabla V(x)|)$ as $|x| \to \infty$.*

Assumption 2.2. *We assume that (1) $b \in (C^\infty(\mathbb{R}^d))^d$; (2) $\|b\|_{C^1(\mathbb{R}^d)} < \infty$; (3) $\langle b, \nabla V\rangle \leq c|\nabla V|^2$ for some constant $0 < c < \frac{1}{2}$.*

Recall that the formulas (1.5)–(1.6) from the introduction appealed to an elliptic operator $\mathscr{A}^{\alpha,\varepsilon}$ as the generator of a semigroup. On a suitable function space, this semigroup is compact and irreducible and this is what guarantees that the principal eigenvalue $\lambda^{\alpha,\varepsilon}$ of $\mathscr{A}^{\alpha,\varepsilon}$ appropriately captures the large-t behavior of the moment-generating function $\chi_t^\varepsilon(\alpha)$; see e.g. [7, 54]. For the analysis of the present paper, we will instead work with the spectrally equivalent operator

$$\begin{aligned}\mathcal{A}^{\alpha,\varepsilon} f &:= \exp((-2\varepsilon)^{-1}V)\mathscr{A}^{\alpha,\varepsilon}(\exp((2\varepsilon)^{-1}V)) \\ &= \varepsilon\Delta f + \langle (1-2\alpha)b, \nabla f\rangle \\ &\quad - \frac{1}{4\varepsilon}|\nabla V|^2 f + \frac{1}{2\varepsilon}\langle b, \nabla V\rangle f - \frac{\alpha(1-\alpha)}{\varepsilon}|b|^2 f + \frac{1}{2}(\Delta V) f - \alpha(\nabla\cdot b) f, \end{aligned} \tag{2.1}$$

which is also associated with a semigroup, which we now take the time to describe. Define the operator $\mathcal{L}^{\alpha,\varepsilon}$ by

$$\mathcal{L}^{\alpha,\varepsilon} f = \varepsilon\Delta f + \langle (1-2\alpha)b, \nabla f\rangle, \tag{2.2}$$

on sufficiently regular functions and let

$$U^{\alpha,\varepsilon} = -\frac{1}{4\varepsilon}|\nabla V|^2 + \frac{1}{2\varepsilon}\langle b, \nabla V\rangle - \frac{\alpha(1-\alpha)}{\varepsilon}|b|^2 + \frac{1}{2}(\Delta V) - \alpha(\nabla\cdot b). \tag{2.3}$$

For readability, let us fix α and ε and omit the dependence on α and ε from the notation for the time being. Consider the SDE with infinitesimal generator $\mathcal{L}$, that is

$$\mathrm{d}X_t = (1-2\alpha)b\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}B_t, \tag{2.4}$$

where B_t is a d-dimensional Brownian motion, and the evolution operator P_t^U defined by

$$P_t^U\varphi(x) = \mathbb{E}\left[\varphi(X_t)\exp\left(\int_0^t U(X_s)\mathrm{d}s\right)\Big| X_0 = x\right], \tag{2.5}$$

where $\mathbb{E}$ is the expectation over all realizations of (2.4) and φ is a function in a suitable space. With natural choices of domain and space, $\mathcal{A} = \mathcal{L} + U$ is indeed the generator of the positivity-preserving semigroup $(P_t^U)_{t>0}$ with the same desirable properties as that generated by $\mathscr{A}$ — albeit on a different space. While these properties and their consequences can be obtained in many different ways, we present a result that foreshadows our upcoming analysis of the discrete semigroups behind our IPM.

Theorem 2.3. *Let $W(x) = \mathrm{e}^{\theta|x|^2}$ and suppose that Assumptions 2.1–2.2 hold. For $\theta > 0$ small enough, there exists a unique measure $\mu_U^\star \in \mathcal{P}(\mathbb{R}^d)$ with $(\mu_U^\star, W) < +\infty$ and a constant $\kappa > 0$ with the following property: for any initial measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ with $(\mu, W) < +\infty$, there exists a constant $C_\mu > 0$ such that*

$$\left| \frac{(\mu, P_t^U \varphi)}{(\mu, P_t^U \mathbb{1})} - (\mu_U^\star, \varphi) \right| \le C_\mu \mathrm{e}^{-\kappa t} \|\varphi\|_{L_W^\infty} \tag{2.6}$$

for all $\varphi \in L_W^\infty(\mathbb{R}^d)$ and $t > 0$. Moreover,

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left[\exp\left(\int_0^t U(X_s) \mathrm{d}s \right) \middle| X_0 \sim \mu \right]. \tag{2.7}$$

Proof sketch. We follow Section 2.3 in [24]. Picking θ small enough such that $32\theta < \|H_0\|^{-2}$ in Assumption 2.1, one can show that the growth bounds in Assumptions 2.1–2.2 imply that W is a Lyapunov function. The regularity properties in Assumptions 2.1–2.2 can be used to show that the semigroup satisfies a Deoblin-type minorization property, an irreducibility property, and a local regularity property that then suffice to deduce (2.6). Moreover, one can show that, for any fixed $t_0 > 0$, the spectral radius of $P_{t_0}^U$ — which equals $\mathrm{e}^{t_0\lambda}$ by the spectral mapping theorem [21] — admits a positive eigenvector h with $\|h\|_{L_W^\infty} = 1$, and that no other eigenvalue admits a positive eigenvector. Taking $\varphi = h$ in (2.6) at times of the form $t = kt_0$, one can deduce that

$$t_0\lambda = \lim_{k\to\infty} \frac{1}{k} \log(\mu, P_{kt_0}^U \mathbb{1}),$$

and then pass to (2.7) using standard arguments. □

We now reintroduce the dependence on α and ε in the notation. We also note that it follows from standard perturbation-theory arguments that the limiting function $\alpha \mapsto \lambda^{\alpha,\varepsilon}$ is real-analytic. Hence, by (1.6) and the Gärtner–Ellis theorem, the following large deviation principle holds: with I^ε the Legendre transform of the function $\alpha \mapsto \lambda^{\alpha,\varepsilon}$, we have

$$\begin{aligned} -\inf_{s\in\operatorname{int} E} I^\varepsilon(s) &\le \liminf_{t\to\infty} \frac{1}{t} \log \mathrm{Prob}^{\mu,\varepsilon}\left\{ \tfrac{1}{t} S_t^\varepsilon \in E \right\} \\ &\le \limsup_{t\to\infty} \frac{1}{t} \log \mathrm{Prob}^{\mu,\varepsilon}\left\{ \tfrac{1}{t} S_t^\varepsilon \in E \right\} \le -\inf_{s\in\operatorname{cl} E} I^\varepsilon(s) \end{aligned}$$

for every Borel set $E \subseteq \mathbb{R}$; again see [7, 54]. It was shown in [54] that, locally and under additional conditions at the critical points of V, easily accessible formulas can be given in the subsequent limit $\varepsilon \to 0^+$, without any rescaling of $\lambda^{\alpha,\varepsilon}$ nor I^ε. Roughly speaking, this means that we get easy access to a limiting rate function I^0 such that

$$\mathrm{Prob}^{\mu,\varepsilon}\{t^{-1} S_t^\varepsilon \approx s\} \asymp \exp\left(-tI^0(s)\right) \tag{2.8}$$

for $t \gg \varepsilon^{-1} \gg 1$ and s near the mean entropy production rate. These additional conditions will be met here if we further assume that

$$\det D^2 V|_{x_j} \neq 0 \tag{2.9}$$

at each of the finitely many critical points $\{x_j\}_{j=1}^J$ of V, and that

$$|b(x_j + \xi)| = O(|D^2 V|_{x_j} \xi|) \tag{2.10}$$

there as well. These extra conditions force the deterministic dynamics obtained by plainly putting $\varepsilon = 0$ in (1.1) to have only very simple invariant structures. The limiting $\lambda^{0,\alpha}$ turns out to be the principal eigenvalue for a quadratic approximation of $\mathcal{A}$ at some α-dependent choice of critical point of V, in such a way that the limiting I^0 is the convex envelope of different rate functions that would arise from linear diffusions approximating (1.1) near critical points of V.

Suppose that on the contrary, that (2.9)–(2.10) fail, say because V has a whole critical circle to which b is tangent as in Section 5 of [7]. Then, we expect to see, as $\varepsilon \to 0^+$, the principal eigenvalue $\lambda^{\alpha,\varepsilon}$ diverge for $\alpha \notin [0,1]$. In such situations, one can consider the rescaling of [7, 8] to obtain further information on the behaviour of those divergences and their relations to the deterministic dynamics and Freidlin–Wentzell theory. We will explore this numerically in Section 5.

3. Numerical discretization using discrete-time semigroups

We again fix α and ε and omit keeping track of them in the notation. To compute the principal eigenvalue λ, we consider a discretization of the operator semigroup $(P_t^U)_{t>0}$, which consists of two steps: an operator splitting scheme and an Euler–Maruyama scheme for the SDE (2.4).

With a time step size $\Delta t > 0$, define an evolution operator $\widetilde{P}_{\Delta t}^U$ by

$$\widetilde{P}_{\Delta t}^U \varphi(x) = \exp(\Delta t U(x)) \mathbb{E}\left[\varphi(X_{\Delta t}) | X_0 = x\right], \tag{3.1}$$

where $X_{\Delta t}$ is the solution to (2.4) at time Δt and φ is a function in a suitable space. Note that if we define an operator P_t by

$$P_t \varphi(x) = \mathbb{E}\left[\varphi(X_t) | X_0 = x\right], \tag{3.2}$$

then $P_t = \exp(t\mathcal{L})$ on a suitable space. Hence, $\widetilde{P}_{\Delta t}^U = \exp(\Delta t U) \exp(\Delta t \mathcal{L})$ can be seen as an approximation of $P_{\Delta t}^U$ using an operator splitting scheme. One can show using the Krein–Rutman theorem that, just like in the case of $P_{\Delta t}^U$, the spectral radius $\widetilde{\Lambda}_{\Delta t}$ of $\widetilde{P}_{\Delta t}^U$ admits a positive eigenvector and that no other eigenvalue admits a positive eigenvector. It should be expected that, for $\Delta t \ll 1$, we have

$$\frac{1}{\Delta t} \log \widetilde{\Lambda}_{\Delta t} \approx \lambda.$$

We will come back to this point at the end of this section.

We now further discretize $\widetilde{P}^U_{\Delta t}$ by considering an Euler–Maruyama scheme for (2.4) with the time step size Δt, which reads

$$\begin{cases} \widehat{X}_{n+1} = \widehat{X}_n + (1-2\alpha)b(\widehat{X}_n)\Delta t + \sqrt{2\varepsilon\Delta t}G_n, \\ \widehat{X}_0 \sim \mu, \end{cases} \tag{3.3}$$

where G_n is a d-dimensional standard Gaussian random variable. Recall that we are considering $\Delta t \ll 1$. We define the evolution operator $\widehat{P}_{\Delta t}$ by

$$\widehat{P}_{\Delta t}\varphi(x) = \mathbb{E}[\varphi(\widehat{X}_{n+1})|\widehat{X}_n = x], \tag{3.4}$$

and define $\widehat{P}^U_{\Delta t}$ by

$$\widehat{P}^U_{\Delta t}\varphi(x) = \exp(U(x)\Delta t)\widehat{P}_{\Delta t}\varphi(x). \tag{3.5}$$

In view of the good convergence properties of the Euler–Mayurama scheme in the total variation distance [6] and the growth of U inherited from Assumptions 2.1–2.2, we expect the spectral radius $\widehat{\Lambda}_{\Delta t}$ of $\widehat{P}^U_{\Delta t}$ to satisfy

$$\frac{1}{\Delta t}\log\widehat{\Lambda}_{\Delta t} \approx \frac{1}{\Delta t}\log\widetilde{\Lambda}_{\Delta t}$$

for $\Delta t \ll 1$. We define a normalized, discrete-time, dual Feynman–Kac semigroup associated with $\widehat{P}^U_{\Delta t}$ by

$$(\Phi_{k,\Delta t}\mu, \varphi) = \frac{(\mu, (\widehat{P}^U_{\Delta t})^k\varphi)}{(\mu, (\widehat{P}^U_{\Delta t})^k\mathbb{1})} = \frac{\mathbb{E}\left[\varphi(\widehat{X}_k)\exp\left(\Delta t\sum_{j=0}^{k-1}U(\widehat{X}_j)\right)\middle|\widehat{X}_0\sim\mu\right]}{\mathbb{E}\left[\exp\left(\Delta t\sum_{j=0}^{k-1}U(\widehat{X}_j)\right)\middle|\widehat{X}_0\sim\mu\right]} \tag{3.6}$$

for any initial measure μ and any bounded measurable function φ. The following theorem establishes, following [24], desirable stability properties of $\Phi_{k,\Delta t}$ for the purpose of numerically accessing the spectral radius $\widehat{\Lambda}_{\Delta t}$.

Theorem 3.1. *Suppose that Assumptions 2.1–2.2 hold. Then, there exists a measure $\widehat{\mu}^\star_{U,\Delta t} \in \mathcal{P}(\mathbb{R}^d)$ with $\widehat{\Lambda}_{\Delta t} = (\widehat{\mu}^\star_{U,\Delta t}, \widehat{P}^U_{\Delta t}\mathbb{1})$ and a constant $\widehat{\beta} \in (0,1)$ with the following property: for any initial measure $\mu \in \mathcal{P}(\mathbb{R}^d)$, there is a constant C_μ for which*

$$|(\Phi_{k,\Delta t}\mu, \varphi) - (\widehat{\mu}^\star_{U,\Delta t}, \varphi)| \le C_\mu\widehat{\beta}^k||\varphi||_{L^\infty} \tag{3.7}$$

for all $\varphi \in L^\infty(\mathbb{R}^d)$ and $k \ge 1$. Moreover,

$$\log\widehat{\Lambda}_{\Delta t} = \lim_{k\to\infty}\frac{1}{k}\log\mathbb{E}\left[\exp\left(\Delta t\sum_{j=0}^{k-1}U(\widehat{X}_j)\right)\middle|\widehat{X}_0\sim\mu\right]. \tag{3.8}$$

Proof sketch. We follow Section 2.2 in [24]. The constant function $\mathbb{1}$ is a Lyapunov function for $\widehat{P}^U_{\Delta t}$. To see this, note that the action of the first operator in the splitting leaves $\mathbb{1}$ invariant and that the action of the second operator is such that

$$\exp(U\Delta t)\mathbb{1} \leq \left(\sup_{|y|>R} \exp(\Delta t U(y))\right)\mathbb{1} + \left(\sup_{|y|\leq R} \exp(\Delta t U(y))\right)\mathbb{1}_{\{y:|y|\leq R\}}, \tag{3.9}$$

and then take $R \to \infty$ using the growth bounds in Assumptions 2.1–2.2. The regularity properties in Assumptions 2.1–2.2 can be used to show that $\widehat{P}^U_{\Delta t}$ satisfies a Deoblin-type minorization property, an irreducibility property, and a local regularity property that then suffice to deduce (3.7) for some uniquely determined probability measure $\widehat{\mu}^\star_{\Delta t}$ satisfying

$$\Phi_{1,\Delta t}\widehat{\mu}^\star_{\Delta t} = \widehat{\mu}^\star_{\Delta t}. \tag{3.10}$$

Moreover, one can show that the spectral radius $\widehat{\Lambda}_{\Delta t}$ for $\widehat{P}^U_{\Delta t}$ admits a positive eigenvector $\widehat{h}$ with $\|\widehat{h}\|_{L^\infty} = 1$, and that no other eigenvalue admits a positive eigenvector. Taking $\varphi = \widehat{h}$ in (3.7), one can deduce that

$$\log \widehat{\Lambda}_{\Delta t} = \lim_{k\to\infty} \frac{1}{k} \log(\mu, (\widehat{P}^U_{\Delta t})^k \mathbb{1})$$

holds. Finally, since (3.10) implies in particular that $(\Phi_{1,\Delta t}\widehat{\mu}^\star_{\Delta t}, \widehat{h}) = (\widehat{\mu}^\star_{\Delta t}, \widehat{h})$, it follows from the eigenvalue equation for $\widehat{h}$ and the definition of $\Phi_{1,\Delta t}$ that $\widehat{\Lambda}_{\Delta t} = (\widehat{\mu}^\star_{U,\Delta t}, \widehat{P}^U_{\Delta t}\mathbb{1})$. □

The aforementioned intuition that

$$\frac{1}{\Delta t}\log \widehat{\Lambda}_{\Delta t} \approx \frac{1}{\Delta t}\log \widetilde{\Lambda}_{\Delta t} \approx \lambda.$$

can indeed be turned into a soft convergence result. This is the content of our next result.

Proposition 3.2. *Under Assumptions 2.1–2.2, for every $T > 0$, we have*

$$\lim_{n\to\infty} n \log \widehat{\Lambda}_{Tn^{-1}} = \lim_{n\to\infty} n \log \widetilde{\Lambda}_{Tn^{-1}} = T\lambda.$$

Proof sketch. Fix $T > 0$ and set

$$\widetilde{\Pi}_n := (\widetilde{P}^U_{Tn^{-1}})^n \text{ and } \widehat{\Pi}_n := (\widehat{P}^U_{Tn^{-1}})^n.$$

We show in four steps that the spectral radii of $\widetilde{\Pi}_n$ and $\widehat{\Pi}_n$ both converge to that of P^U_T as operators on C_0 equipped with the L^∞-norm.

Step 1. *Operator norm convergence* $\|\widehat{\Pi}_n - \widetilde{\Pi}_n\| \to 0$. The growth conditions on V in Assumption 2.1 and the control on b in Assumption 2.2 imply that $\exp(TU)$ is bounded by $K := \exp(T\|U_+\|_{L^\infty})$ and satisfies the following decay property: for every $\delta > 0$, there exists R_δ such that

$$\sup_{|x|>R_\delta} \exp(TU(x)) < \delta.$$

The boundedness in Assumption 2.2 allows for the application of a classical martingale argument that shows that, for every $\eta > 0$, there exists ρ_η such that

$$\sup_x \mathbb{P}^x \left\{ \sup_{t\in[0,T]} |X_t^x - x| \geq \rho_\eta \right\} < \eta. \tag{3.11}$$

One can show that $|\widehat{\Pi}_n\varphi(x) - \widetilde{\Pi}_n\varphi(x)|$ can be made arbitrarily small with large n, uniformly in x and φ with $\|\varphi\|_{L^\infty} = 1$ as follows. Choose δ and η small enough, then R_δ and ρ_η accordingly, and then consider separately the cases $|x| > R_\delta + \rho_\eta$ and $|x| \leq R_\delta + \rho_\eta$. The former will be small as is, and as for the latter, take n large to leverage results for the Euler–Mayurama scheme in total variation norm [6].

Step 2. *Strong convergence* $\widetilde{\Pi}_n - P_T^U \xrightarrow{s} 0$. Since we already know that $\mathcal{L} + U$ generates a strongly continuous semigroup on C_0, this is a direct consequence of Trotter's product formula for the semigroups generated by $\mathcal{L}$ and U on that same space [63].

Step 3. *Collective compactness of* $(\widetilde{\Pi}_n)_{n=1}^\infty$. We want to show that

$$S := \{\widetilde{\Pi}_n\varphi : n \in \mathbb{N}, \varphi \in C_0, \|\varphi\|_{L^\infty} \leq 1\}$$

is precompact in C_0. To do this we need to show three properties: boundedness, uniform vanishing at infinity, and equicontinuity.

3a. Pointwise, it follows from the definition of $\widetilde{\Pi}_n$ and the assumption that $\|\varphi\|_{L^\infty} \leq 1$ that

$$|(\widetilde{\Pi}_n\varphi)(x)| \leq \mathbf{E}\left[\exp\left(\sum_{k=0}^{n-1} \frac{T}{n} U(X_{kTn^{-1}}) \right) \middle| X_0 = x \right] \tag{3.12}$$

This shows among other things that the family S is bounded in norm by K.

3b. Using (3.11) in conjunction with properties of U discussed in Step 1, we see that the expectation on the right-hand side of (3.12) is arbitrarily small, simultaneously for all n and φ, as soon as x is outside of a sufficiently large ball. Hence, the family S does uniformly vanish at infinity.

3c. To show equicontinuity, we consider the differences

$$\begin{aligned}&\widetilde{\Pi}_n\varphi(x) - \widetilde{\Pi}_n\varphi(y)\\ &\quad = \mathbf{E}^x\left[\varphi(X_T) \exp\left(\sum_{k=0}^{n-1} \frac{T}{n} U(X_{kTn^{-1}}) \right) \right] - \mathbf{E}^y\left[\varphi(Y_T) \exp\left(\sum_{k=0}^{n-1} \frac{T}{n} U(Y_{kTn^{-1}}) \right) \right]\end{aligned}$$

with y close to some fixed x — we require $|y - x| < 1$ to begin. Note that this difference of expectations can be computed by realizing the two processes $(X_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ on a common probability space as we see fit — this is the classical coupling method; see e.g. [61] and historical references therein.

By continuity, the difference

$$\delta_2 := |U(x) - U(y)|$$

can be made arbitrarily small by taking y close enough to x. In view of the ultra-Feller property [32], given any $\tau > 0$, the difference

$$\delta_3(\tau) := \|P_\tau^* \delta_x - P_\tau^* \delta_y\|_{\mathrm{TV}}$$

can be made arbitrarily small by taking y close enough to x. Hence, once such a τ is given, we can realize the two processes $(X_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ on a common probability space so that

$$\mathbb{P}^{(x,y)}\{X_t \neq Y_t \text{ for some } t \in [\tau, T]\} \leq \delta_3(\tau).$$

All in all, we have

$$\begin{aligned}
&|\widetilde{\Pi}_n \varphi(x) - \widetilde{\Pi}_n \varphi(y)| \\
&\quad \leq \mathbf{E}^{(x,y)} \left| \varphi(X_T) \exp\left(\sum_{k=0}^{n-1} \frac{T}{n} U(X_{kTn^{-1}}) \right) - \varphi(Y_T) \exp\left(\sum_{k=0}^{n-1} \frac{T}{n} U(Y_{kTn^{-1}}) \right) \right| \\
&\quad \leq 2K\eta + 2K\delta_3(\tau) + KT\delta_2 + 4K\tau \sup_{|z-x|<1+\rho_\eta} |U(z)|.
\end{aligned}$$

This can be made arbitrarily small, uniformly in φ and n, as follows. First, we choose η so that the first term is as small as desired, and then we fix ρ_η accordingly. Next, we take τ small enough that the last term is as small as desired. Finally, once τ is fixed, we can choose a coupling to compute the expectation, and the second and third terms will be as small as desired as long as y is close enough to x.

Step 4. *Spectral theory.* On the one hand, Step 1 and the fact that both sequences of operators are uniformly bounded by K ensures that $|\operatorname{spr}(\widetilde{\Pi}_n) - \operatorname{spr}(\widehat{\Pi}_n)| \to 0$ by classical perturbation theory arguments; see e.g. [38]. On the other hand, thanks to the spectral analysis of [5] for collectively compact sequences of operators that converge strongly, Steps 2 and 3 show that $|\operatorname{spr}(\widetilde{\Pi}_n) - \operatorname{spr}(P_T^U)| \to 0$.

Clearly, if $\widetilde{\Lambda}_{Tn^{-1}}$ is an eigenvalue of $\widetilde{P}^U_{Tn^{-1}}$ with a positive eigenvector, then $(\widetilde{\Lambda}_{Tn^{-1}})^n$ is an eigenvalue of $\widetilde{\Pi}_n = (\widetilde{P}^U_{Tn^{-1}})^n$ with a positive eigenvector. Hence, the identity $\operatorname{spr}(\widetilde{\Pi}_n) = (\widetilde{\Lambda}_{Tn^{-1}})^n$ follows from the fact that the spectral radius is the only eigenvalue admitting a positive eigenvector. Similarly, $\operatorname{spr}(\widehat{\Pi}_n) = (\widehat{\Lambda}_{Tn^{-1}})^n$. Finally, the fact that $\operatorname{spr}(P_T^U) = \exp(\lambda T)$ is a consequence of the spectral mapping theorem [21], so the proof is completed. □

4. Interacting particle methods

Sections 2 and 3 show that the principal eigenvalue $\lambda^{\varepsilon,\alpha}$ can be approximated in terms of the logarithmic spectral radius (3.8) that is accessible through large iterates of a discrete-time semigroup with good stability properties. This strongly suggests that it can be efficiently accessed using a discrete-time IPM. Given an ensemble of particles, the IPM proceeds within each time interval as follows. The particles evolve according to the dynamics of $\widehat{P}_{\Delta t}$ with an importance weight assigned to each particle, and then to control variance [25] (or to avoid weight degeneracy [46]) the particles are resampled according to the multinomial distribution associated with their respective weights. The logarithmic spectral radius (3.8) is accessed using the particles at each time step. The complete algorithm of the IPM for computing $\lambda^{\varepsilon,\alpha}$ is given in Algorithm 1, where we only emphasize the dependence of the final approximation $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ on ε and α. Note that the particles $\{\mathbf{q}^{n,m}\}_{m=1}^{M}$ are no longer independent as soon as $n \geq 1$ but still exchangeable.

Algorithm 1 The interacting particle method for computing $\lambda^{\varepsilon,\alpha}$

Input: α, noise level ε, velocity field b, potential V, number of particles M, initial measure μ, final time T, time step size $\Delta t = \frac{T}{N}$.

1: Generate M independent and μ-distributed particles $\{\mathbf{q}^{0,m}\}_{m=1}^{M}$.
2: **for** n = 1:N **do**
3: Compute each $\widetilde{\mathbf{q}}^{n,m}$ using the Euler–Maruyama scheme (3.3) with $\mathbf{q}^{n-1,m}$ the initial value.
4: Compute each weight $w^{n-1,m} = \exp(\Delta t U(\mathbf{q}^{n-1,m}))$ according to (2.3).
5: Compute the quantities $P^{n-1} = \sum_{m=1}^{M} w^{n-1,m}$ and $\widehat{\lambda}^{n-1} = \log(P^{n-1}/M)$.
6: Compute the probabilities $p^{n-1,m} = w^{n-1,m}/P^{n-1}$ and sample M non-negative integers $(K_m)_{m=1}^{M}$ summing to M according to the multinomial law

$$\mathrm{Prob}\{K_1 = k_1, \ldots, K_M = k_M\} = \frac{M!}{\prod_{m=1}^{M} k_m!} \prod_{m=1}^{M} (p^{n-1,m})^{k_m}.$$

7: Set $(\mathbf{q}^{n,m})_{m=1}^{M}$ to contain K_m copies of $\widetilde{\mathbf{q}}^{n,m}$.
8: **end for**
9: Compute the approximation

$$\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t} = \frac{1}{T} \sum_{n=0}^{N-1} \widehat{\lambda}^{n}$$

of the principal eigenvalue.

Output: the approximation of the principal eigenvalue $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$.

4.1. The empirical measure of particles at the final time

The empirical measure of particles at the final time T (equivalently after the N-th step) is a random measure that is thought of as an approximation to the Feynman–Kac semigroup

$\Phi_{N,\Delta t}\mu$, defined in (3.6), and thus to the invariant measure $\widehat{\mu}^{\star}_{U,\Delta t}$ for N large by (3.7). To justify this, let us first consider the one-step evolution $\Phi_{1,\Delta t}\mu$. First of all, by the Glivenko–Cantelli theorem or a variant thereof (see e.g. [27, 60]), the empirical measure

$$\widehat{\mu}^{0,+}_{U,\Delta t,M} := \frac{1}{M}\sum_{m=1}^{M}\delta_{\mathbf{q}^{0,m}}$$

of the particles $\{\mathbf{q}^{0,m}\}_{m=1}^{M}$ approximates μ well provided that M is large. Then, on the one hand, by (3.6) and the definitions in Algorithm 1, the measure $\Phi_{1,\Delta t}\mu$ can be approximated by the weighted empirical measure

$$\widehat{\mu}^{1,-}_{U,\Delta t,M} := \Phi_{1,\Delta t}\widehat{\mu}^{0}_{U,\Delta t,M} = \sum_{m=1}^{M} p^{0,m}\delta_{\tilde{\mathbf{q}}^{1,m}}.$$

On the other hand, since the multinomial law used in Algorithm 1 satisfies

$$\sum_{k_1,\dots,k_M} k_m \,\mathrm{Prob}\{K_1 = k_1,\dots,K_M = k_M\} = Mp^{0,m},$$

we have that, for any test function φ,

$$\begin{aligned}
&\sum_{k_1,\dots,k_M}\left(\frac{1}{M}\sum_{m=1}^{M}\varphi(\mathbf{q}^{1,m})\right)\mathrm{Prob}\{K_1 = k_1,\dots,K_M = k_M\}\\
&\qquad= \sum_{k_1,\dots,k_M}\left(\frac{1}{M}\sum_{m=1}^{M}k_m\varphi(\tilde{\mathbf{q}}^{1,m})\right)\mathrm{Prob}\{K_1 = k_1,\dots,K_M = k_M\}\\
&\qquad= \sum_{m=1}^{M}\left(\frac{1}{M}\sum_{k_1,\dots,k_M} k_m\,\mathrm{Prob}\{K_1 = k_1,\dots,K_M = k_M\}\right)\varphi(\tilde{\mathbf{q}}^{1,m})\\
&\qquad= \sum_{m=1}^{M} p^{0,m}\varphi(\tilde{\mathbf{q}}^{1,m}).
\end{aligned}$$

Hence, the resampled empirical measure $\widehat{\mu}^{1,+}_{U,\Delta t,M}$ of $\{\mathbf{q}^{1,m}\}_{m=1}^{M}$ yields, once the randomness in the resampling process is averaged out, the exact same expectations as the weighted empirical measure $\widehat{\mu}^{1,-}_{U,\Delta t,M}$. In particular, this holds when $\varphi = \widehat{P}^{U}_{\Delta t}\exp(\Delta t U)$, which is relevant at the next step for carrying on with our approximation of the principal eigenvalue. In fact, it is expected that, for that purpose and when M is large, the empirical measure $\widehat{\mu}^{1,+}_{U,\Delta t,M}$ is a numerically sounder choice as it gives more importance to the regions where $\exp(\Delta t U)$ is large. We refer the readers to e.g. [13] and [46] for a more thorough discussion.

Iterating this argument, the measure $\Phi_{N,\Delta t}\mu$ should indeed be well approximated by the resampled empirical measure $\widehat{\mu}^{N,+}_{U,\Delta t,M}$ of $\{\mathbf{q}^{N,m}\}_{m=1}^{M}$. In our numerical examples in Section 5, as $\varepsilon \to 0^{+}$, the asymptotic behavior of the empirical density of particles at T is consistent with the theory in [26].

4.2. Choice of the initial measure

The IPM involves the choice of an initial measure μ for the particles. Recall from Theorem 3.1 that

$$\widehat{\lambda}_{\Delta t} = \frac{1}{\Delta t} \log(\widehat{\mu}^{\star}_{U,\Delta t}, \widehat{P}^{U}_{\Delta t} \mathbb{1}),$$

and that

$$\widehat{\lambda}_{\Delta t} = \lim_{k \to \infty} \frac{1}{k \Delta t} \log(\mu, (\widehat{P}^{U}_{\Delta t})^{k} \mathbb{1}).$$

This suggests that we can obtain faster convergence in (3.8) with respect to k if we can set μ to be close to $\widehat{\mu}^{\star}_{U,\Delta t}$ in some sense. We introduce here two techniques for choosing the initial measure μ in Algorithm 1 based on this intuition.

The first technique is the so-called burn-in procedure, in which we start computing $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ only using results from times later than some $t > 0$, i.e. compute $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ using

$$\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t} = \frac{1}{T - t} \sum_{n=\left\lceil \frac{t}{\Delta t} \right\rceil}^{N-1} \widehat{\lambda}^{n}.$$

This is equivalent to setting the initial measure to be the empirical measure of particles at t, which should be closer to $\widehat{\mu}^{\star}_{U,\Delta t}$ than μ in any sense.

The second technique applies when computing $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ for ε_1 and ε_2 with $\varepsilon_1 > \varepsilon_2$. Let $\mu^{\star}_{U,\varepsilon}$ be the invariant measure in Theorem 2.3 for each ε and $p^{\star}_{U,\varepsilon}$ be the corresponding density. The results in [26] show that under certain conditions $\varepsilon \log p^{\star}_{U,\varepsilon}$ has a nontrivial asymptotic limit as $\varepsilon \to 0^{+}$. Hence it is intuitively reasonable that given the invariant measure at ε_1 we can use it as the initial measure for ε_2 in order to obtain faster computation of (3.8).

5. Numerical examples

5.1. Computational setup

The following computation is performed on a high-performance computing cluster with 2 Intel Xeon Gold 6226R (16 Core) CPUs and 96GB RAM. We consider the computation of $\lambda^{\varepsilon,\alpha}$ for certain values of ε and α. In particular, we choose $\varepsilon = 0.1, 0.01, 0.001$. For each fixed ε, we let $\alpha \in \left[-\frac{1}{10}, \frac{11}{10}\right]$ and compute $\widehat{\lambda}^{\varepsilon,\alpha}$ for $\alpha = -\frac{1}{10} + \frac{j}{31}\frac{12}{10}$ with $j = 0, 1, \ldots, 31$. The computation of $\widehat{\lambda}^{\varepsilon,\alpha}$ for each ε with 32 different values of α is performed at the same time in parallel on the 32 cores of the CPUs. For the numerical discretization of our method, we choose $M = 500\,000$ and $\Delta t = 2^{-8}$ in Algorithm 1. Also, unless specified, the initial measure of the particles is chosen to be the standard multivariate Gaussian distribution.

5.2. The principal eigenvalue and the rate function in the vanishing-noise limit

Example 5.1. Consider

$$V(x_1, x_2) = \frac{x_1^2 + x_2^2}{2} + \frac{x_1^4 + x_2^4}{8},$$

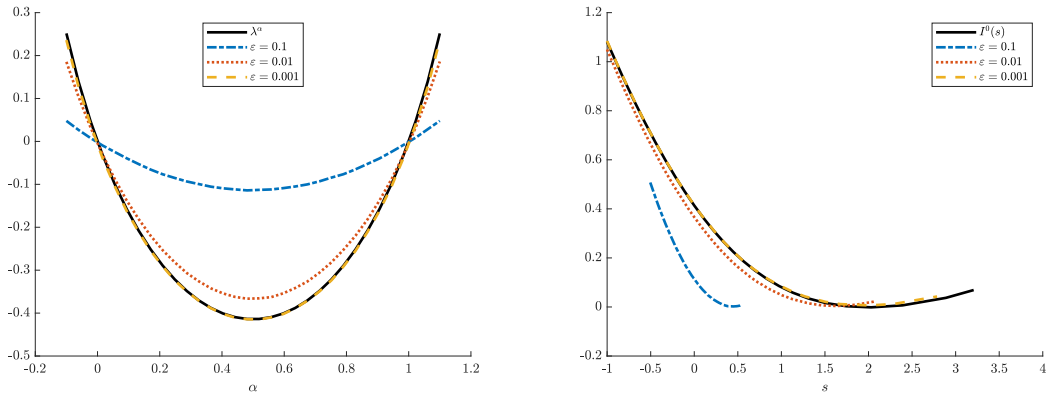
and

$$b(x_1, x_2) = \pi^{-1}(\cos(\pi x_1)\sin(\pi x_2), -\sin(\pi x_1)\cos(\pi x_2)).$$

Note that V has a global minimum point at $(0, 0)$ and no other critical points. It can be shown [54] that

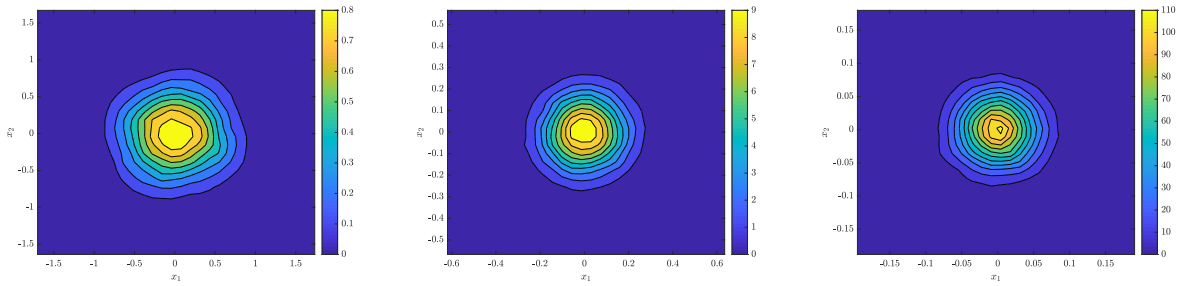
$$\lambda^{0,\alpha} = \lim_{\varepsilon \to 0^+} \lambda^{\varepsilon,\alpha} = 1 - \sqrt{1 + 4\alpha(1-\alpha)}. \tag{5.1}$$

We choose $T = 1024$. We show the numerical eigenvalue $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ in Figure 1a. In addition, the numerical rate function $\widehat{I}^{\varepsilon}_{\Delta t}(s)$ obtained by the Legendre transform of $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ is shown in Figure 1b. Moreover, the empirical density of particles at T with $\alpha \approx 0.6742$ is shown in Figure 2. It can be seen from Figure 2 that the particles get more localized around the global minimum point $(0, 0)$ of V as $\varepsilon \to 0^+$.



(a) $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ (b) $\widehat{I}^{\varepsilon}_{\Delta t}(s)$

Figure 1: In the context of Example 5.1, we plot our numerical approximation $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ of the principal eigenvalue $\lambda^{\varepsilon,\alpha}$ and the resulting approximation $\widehat{I}^{\varepsilon}_{\Delta t}(s)$ of the rate function $I^{\varepsilon}(s)$, compared respectively to the limit $\lambda^{0,\varepsilon}$ in (5.1) and its Legendre transform $I^0(s)$. Note the consistency of the symmetries mentioned in the Introduction. Also note that the restriction of $\widehat{I}^{\varepsilon}_{\Delta t}(s)$ to certain values of s is due to our restriction of $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ and how it interacts with the derivatives.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 2: In the context of Example 5.1, we plot the empirical density of particles at T with $\alpha \approx 0.6742$. As $\varepsilon \to 0^+$, the particles gets more localized around $(0, 0)$.

Example 5.2. Consider

$$V(x_1, x_2) = x_1^4 - 2x_1^2 + (1 + a(x_1 - 1)^2)x_2^2 + x_2^4$$

with $a = 0.4$, and

$$b(x_1, x_2) = \widetilde{b}\pi^{-1}(\cos(\pi x_1)\sin(\pi x_2), -\sin(\pi x_1)\cos(\pi x_2))$$

with $\widetilde{b} = 1$. Note that V has two local minima at $(-1, 0)$ and $(1, 0)$, as well as a saddle point at $(0, 0)$. It can be shown [54] that

$$\lambda^{0,\alpha} = \widetilde{\lambda}^{\alpha}(a, \widetilde{b}) = \lim_{\varepsilon \to 0^+} \lambda^{\varepsilon,\alpha} = \max(\lambda_+^{\alpha}, \lambda_-^{\alpha}), \tag{5.2}$$

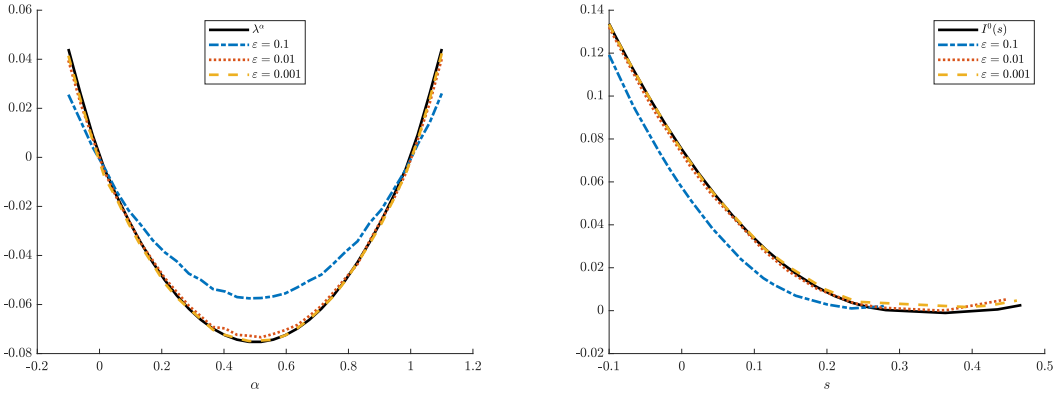
where

$$\lambda_+^{\alpha} = -\operatorname{Tr} X_+(\alpha) + \frac{1}{2}\operatorname{Tr} D^2V|_{(1,0)}, \quad \lambda_-^{\alpha} = -\operatorname{Tr} X_-(\alpha) + \frac{1}{2}\operatorname{Tr} D^2V|_{(-1,0)},$$

with $X_+(\alpha)$ and $X_-(\alpha)$ satisfying the following algebraic Riccati equations

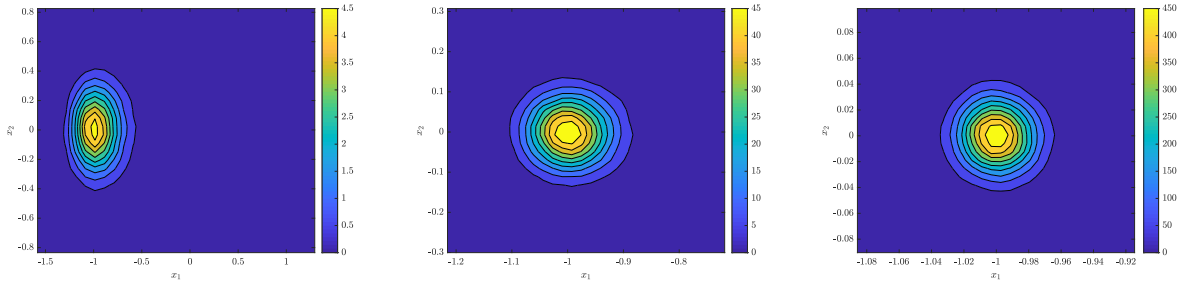
$$\begin{aligned}
&X_+(\alpha)^2 - \frac{1-2\alpha}{2}(\nabla b^T X_+(\alpha) + X_+(\alpha)^T \nabla b) \\
&- \frac{1}{4}D^2V|_{(1,0)}D^2V|_{(1,0)} + \frac{1}{4}(\nabla b^T D^2V|_{(1,0)} + D^2V|_{(1,0)}\nabla b) - \alpha(1-\alpha)\nabla b^T \nabla b = 0, \\
&X_-(\alpha)^2 - \frac{1-2\alpha}{2}(\nabla b^T X_-(\alpha) + X_-(\alpha)^T \nabla b) \\
&- \frac{1}{4}D^2V|_{(-1,0)}D^2V|_{(-1,0)} + \frac{1}{4}(\nabla b^T D^2V|_{(-1,0)} + D^2V|_{(-1,0)}\nabla b) - \alpha(1-\alpha)\nabla b^T \nabla b = 0.
\end{aligned}$$

We choose $T = 2048$. We use the burn-in procedure, in which we start computing the eigenvalue from $t = 1024$. We show $\widehat{\lambda}_{\Delta t}^{\varepsilon,\alpha}$ in Figure 3a and $\widehat{I}_{\Delta t}^{\varepsilon}(s)$ in Figure 3b. The empirical density of particles at T with $\alpha \approx 0.5968$ is shown in Figure 4 and that with $\alpha \approx 1.0613$ is shown in Figure 5. We can see from Figures 4 and 5 that the particles are localized around different local minimum points of V for different values of α.



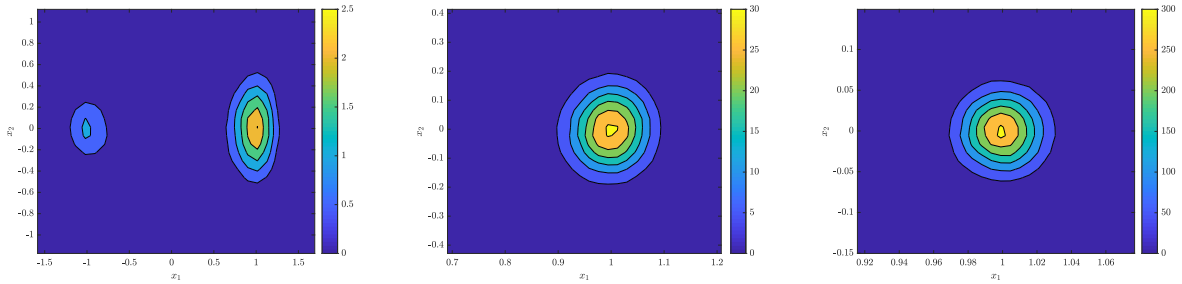
(a) $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ (b) $\widehat{I}^{\varepsilon}_{\Delta t}(s)$

Figure 3: In the context of Example 5.2, we plot our numerical approximation $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ of the principal eigenvalue $\lambda^{\varepsilon,\alpha}$ and the resulting approximation $\widehat{I}^{\varepsilon}_{\Delta t}(s)$ of the rate function $I^{\varepsilon}(s)$, compared respectively to the limit $\lambda^{0,\varepsilon}$ in (5.2) and its Legendre transform $I^0(s)$. Note that the maximum in (5.2) causes a discontinuity of the derivative of the limit of the eigenvalue in $\alpha = 0$ and $\alpha = 1$, in turn causing flat regions in the limit of the rate function.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 4: In the context of Example 5.2, we plot the empirical density of particles at T with $\alpha \approx 0.5968$. As $\varepsilon \to 0^+$, the particles get more localized around $(-1, 0)$.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 5: In the context of Example 5.2, we plot the empirical density of particles at T with $\alpha \approx 1.0613$. As $\varepsilon \to 0^+$, the particles get more localized around $(1, 0)$.

Example 5.3. Consider

$$V(x_1,\dots,x_8)=\sum_{j=1}^{2}\left(\frac{x_{2j-1}^2+x_{2j}^2}{2}+\frac{x_{2j-1}^4+x_{2j}^4}{8}\right)$$
$$+\sum_{j=3}^{4}\left(x_{2j-1}^4-2x_{2j-1}^2+(1+a(x_{2j-1}-1)^2)x_{2j}^2+x_{2j}^4\right)$$

with $a=0.3$, and

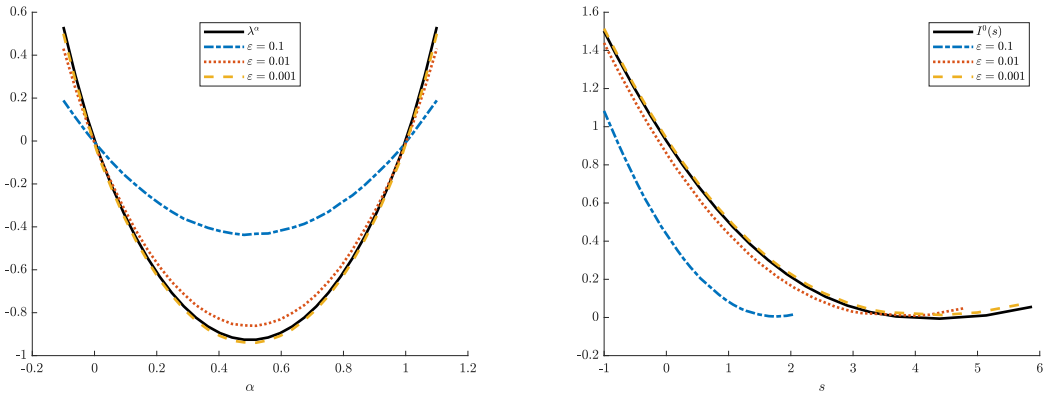
$$b(x_1,\dots,x_8)=\pi^{-1}\bigoplus_{j=1}^{4}(\widetilde{b}_j\cos(\pi x_{2j-1})\sin(\pi x_{2j}),-\widetilde{b}_j\sin(\pi x_{2j-1})\cos(\pi x_{2j})),$$

with $\widetilde{b}_1=1,\widetilde{b}_2=0.5,\widetilde{b}_3=1$ and $\widetilde{b}_4=2$. It can be shown [54] that

$$\lambda^{0,\alpha}=\lim_{\varepsilon\to 0^+}\lambda^{\varepsilon,\alpha}=1-\sqrt{1+4\alpha(1-\alpha)}+1-\sqrt{1+\alpha(1-\alpha)}+\lambda_3^\alpha+\lambda_4^\alpha, \tag{5.3}$$

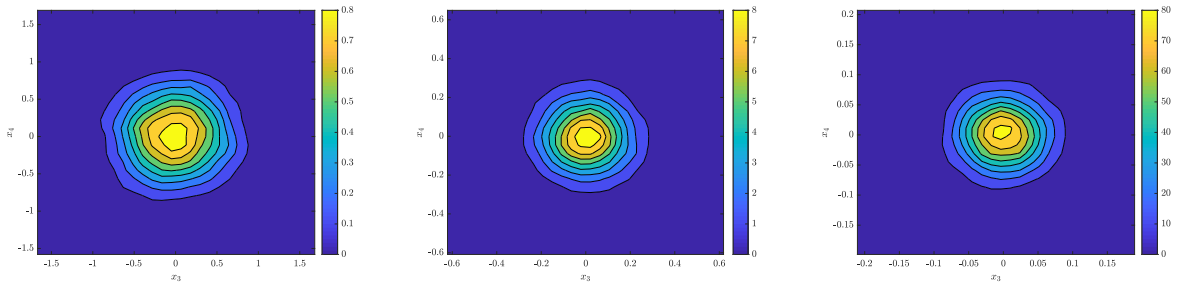
where $\lambda_3^\alpha=\widetilde{\lambda}^\alpha(a,\widetilde{b}_3)$, $\lambda_4^\alpha=\widetilde{\lambda}^\alpha(a,\widetilde{b}_4)$, with $\widetilde{\lambda}^\alpha(a,\widetilde{b})$ given in (5.2).

We choose $T=2048$. We use the burn-in procedure for $\varepsilon=0.1,0.01$, in which we start computing the eigenvalue from $t=1024$. For $\varepsilon=0.001$, we use the empirical measure of particles at T obtained at $\varepsilon=0.01$ as the initial measure. We show $\widehat{\lambda}_{\Delta t}^{\varepsilon,\alpha}$ in Figure 6a and $\widehat{I}_{\Delta t}^{\varepsilon}(s)$ in Figure 6b. The 2-dimensional marginal empirical densities of particles at T with $\alpha\approx 0.3645$ are shown in Figure 7 and 8.



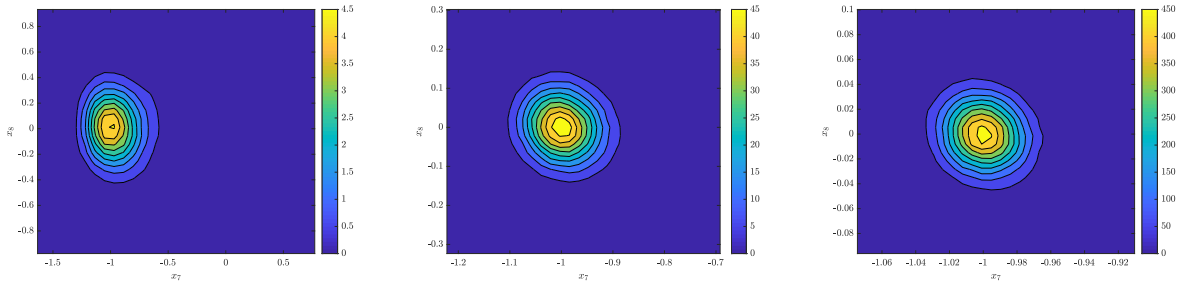
(a) $\widehat{\lambda}_{\Delta t}^{\varepsilon,\alpha}$ (b) $\widehat{I}_{\Delta t}^{\varepsilon}(s)$

Figure 6: In the context of Example 5.3, we plot our numerical approximation $\widehat{\lambda}_{\Delta t}^{\varepsilon,\alpha}$ of the principal eigenvalue $\lambda^{\varepsilon,\alpha}$ and the resulting approximation $\widehat{I}_{\Delta t}^{\varepsilon}(s)$ of the rate function $I^\varepsilon(s)$, compared respectively to the limit $\lambda^{0,\varepsilon}$ in (5.3) and its Legendre transform $I^0(s)$.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 7: In the context of Example 5.3, we plot the 2D marginal empirical density of (x_3, x_4) of particles at T with $\alpha \approx 0.3645$.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 8: In the context of Example 5.3, we plot the 2D marginal empirical density of (x_7, x_8) of particles at T with $\alpha \approx 0.3645$.

Example 5.4. Consider

$$\begin{aligned} V(x_1, \ldots, x_{16}) = &\frac{x_1^2 + x_2^2}{2} + \frac{x_1^4 + x_2^4}{8} + \frac{x_3^2 + x_4^2}{2} + \frac{x_3^4 + x_4^4}{8} \\ &+ x_5^4 - 2x_5^2 + (1 + a_1(x_5 - 1)^2)x_6^2 + x_6^4 \\ &+ x_7^4 - 2x_7^2 + (1 + a_1(x_7 - 1)^2)x_8^2 + x_8^4 \\ &+ x_9^4 - 2x_9^2 + (1 + a_2(x_9 - 1)^2)x_{10}^2 + x_{10}^4 \\ &+ x_{11}^4 - 2x_{11}^2 + (1 + a_2(x_{11} - 1)^2)x_{12}^2 + x_{12}^4 \\ &+ \frac{x_{13}^2 + x_{14}^2}{2} + \frac{x_{13}^4 + x_{14}^4}{8} \\ &+ x_{15}^4 - 2x_{15}^2 + (1 + a_3(x_{15} - 1)^2)x_{16}^2 + x_{16}^4 \end{aligned}$$

with $a_1 = 0.2, a_2 = 0.7, a_3 = 0.5$, and

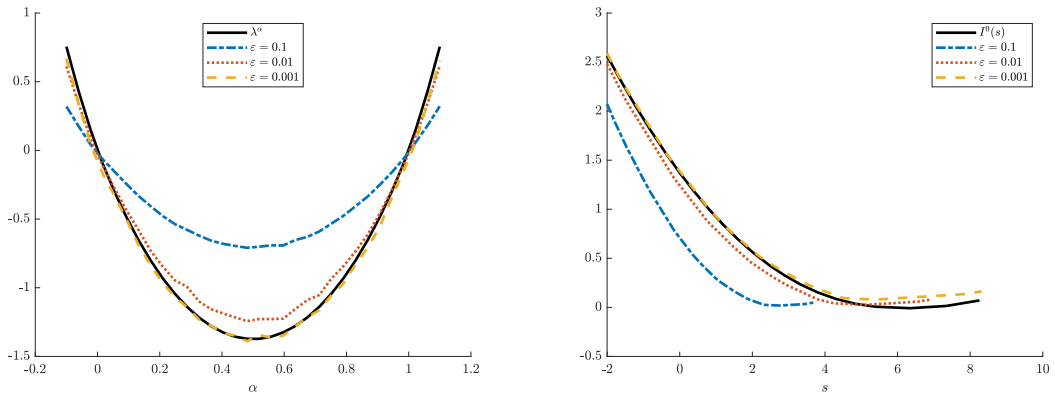
$$b(x_1, \ldots, x_{16}) = \pi^{-1} \bigoplus_{j=1}^{6} (\widetilde{b}_j \cos(\pi x_{2j-1}) \sin(\pi x_{2j}), -\widetilde{b}_j \sin(\pi x_{2j-1}) \cos(\pi x_{2j})) \bigoplus (0, 0, 0, 0),$$

with $\widetilde{b}_1 = 1, \widetilde{b}_2 = 0.5, \widetilde{b}_3 = 1, \widetilde{b}_4 = 2, \widetilde{b}_5 = 1, \widetilde{b}_6 = 2$. It can be shown [54] that

$$\lambda^{0,\alpha} = \lim_{\varepsilon \to 0^+} \lambda^{\varepsilon,\alpha} = 1 - \sqrt{1 + 4\alpha(1-\alpha)} + 1 - \sqrt{1 + \alpha(1-\alpha)} + \lambda_3^\alpha + \lambda_4^\alpha + \lambda_5^\alpha + \lambda_6^\alpha, \quad (5.4)$$

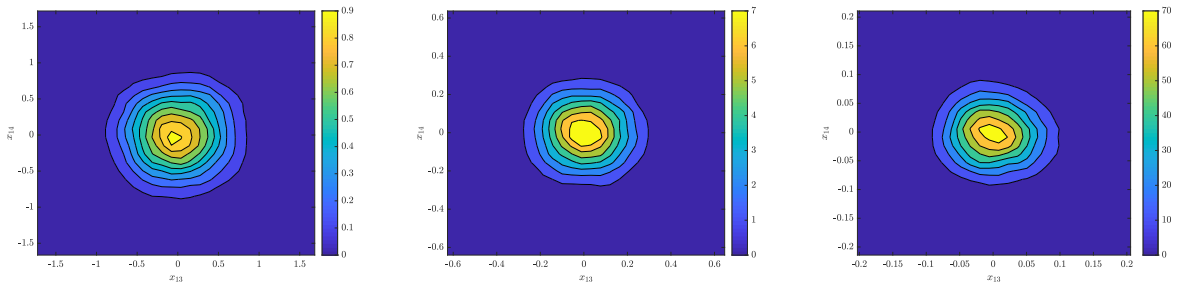
where $\lambda_3^\alpha = \widetilde{\lambda}^\alpha(a_1, \widetilde{b}_3), \lambda_4^\alpha = \widetilde{\lambda}^\alpha(a_1, \widetilde{b}_4), \lambda_5^\alpha = \widetilde{\lambda}^\alpha(a_2, \widetilde{b}_5), \lambda_6^\alpha = \widetilde{\lambda}^\alpha(a_2, \widetilde{b}_6),$, with $\widetilde{\lambda}^\alpha(a, \widetilde{b})$ given in (5.2).

We choose $T = 2048$. We use the burn-in procedure for $\varepsilon = 0.1$, in which we start computing the eigenvalue from $t = 1024$. For $\varepsilon = 0.01, 0.001$, we use the empirical measure of particles at T obtained at $\varepsilon = 0.1$ as the initial measure. We show $\widehat{\lambda}_{\Delta t}^{\varepsilon,\alpha}$ in Figure 9a and $\widehat{I}_{\Delta t}^{\varepsilon}(s)$ in Figure 9b. The 2-dimensional marginal empirical densities of particles at T with $\alpha \approx 0.2097$ are shown in Figure 10 and 11.



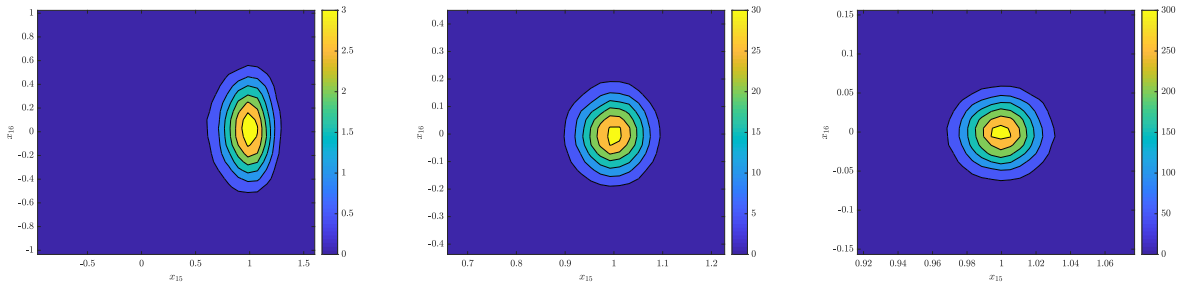
(a) $\widehat{\lambda}_{\Delta t}^{\varepsilon,\alpha}$ (b) $\widehat{I}_{\Delta t}^{\varepsilon}(s)$

Figure 9: In the context of Example 5.4, we plot our numerical approximation $\widehat{\lambda}_{\Delta t}^{\varepsilon,\alpha}$ of the principal eigenvalue $\lambda^{\varepsilon,\alpha}$ and the resulting approximation $\widehat{I}_{\Delta t}^{\varepsilon}(s)$ of the rate function $I^\varepsilon(s)$, compared respectively to the limit $\lambda^{0,\varepsilon}$ in (5.4) and its Legendre transform $I^0(s)$.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 10: In the context of Example 5.4, we plot the 2D marginal empirical density of (x_{13}, x_{14}) of particles at T with $\alpha \approx 0.2097$.



(a) $\varepsilon = 0.1$ (b) $\varepsilon = 0.01$ (c) $\varepsilon = 0.001$

Figure 11: In the context of Example 5.4, we plot the 2D marginal empirical density of (x_{15}, x_{16}) of particles at T with $\alpha \approx 0.2097$.

From all the above examples with different values of d, we can observe the convergence of both the numerical principal eigenvalue $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ and the numerical rate function $\widehat{I}^{\varepsilon}_{\Delta t}(s)$ to their respective analytical vanishing-noise limits $\lambda^{0,\varepsilon}$ and $I^0(s)$, with a fixed number of particles and a fixed time step size. Furthermore, the maximum of the 2-dimensional (marginal) empirical density of particles at T is proportional to ε^{-1}. We know from [26] that $\varepsilon \log p^{\star}_{U,\varepsilon}$ has a nontrivial asymptotic limit as $\varepsilon \to 0^+$ under certain conditions with $p^{\star}_{U,\varepsilon}$ the invariant density, which shows that $p^{\star}_{U,\varepsilon}$ is asymptotically proportional to $\exp(-\varepsilon^{-1}\Phi)$ for some function Φ and that the normalizing constant of $p^{\star}_{U,\varepsilon}$ is asymptotically $\mathcal{O}(\varepsilon^{-d/2})$. Hence the asymptotic behavior of the empirical density of particles at T as $\varepsilon \to 0^+$ is consistent with the theory in [26]. In addition, we show in Figure 12 the computational time in minutes versus d based on the computational times of Examples 5.2–5.4 as well as that of a 4-dimensional example (which we do not show here). The computational setups of these 4 examples have the same value of $M, \Delta t, T$ and only differ in the value of d. Here, the computational time is the maximum computational time over all α for each fixed ε. We can see that the computational time grows linearly with respect to d and does not change significantly as ε varies.

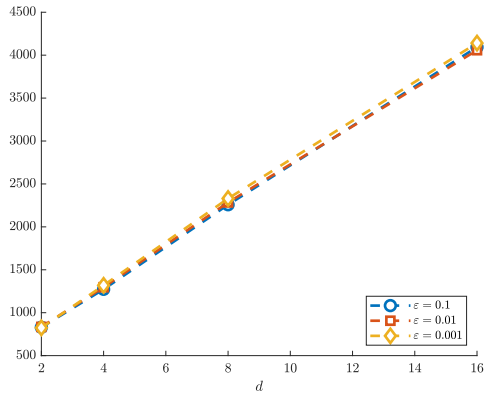


Figure 12: Computational time in minutes versus d.

Finally, we discuss an example where the assumptions (2.9)–(2.10) fail, preventing us

from appealing to the proof of [54] for convergence in the limit $\varepsilon \to 0^+$. In such situations, it is possible for $\lambda^{\varepsilon,\alpha}$ to diverge as $\varepsilon \to 0^+$.

Example 5.5. Consider

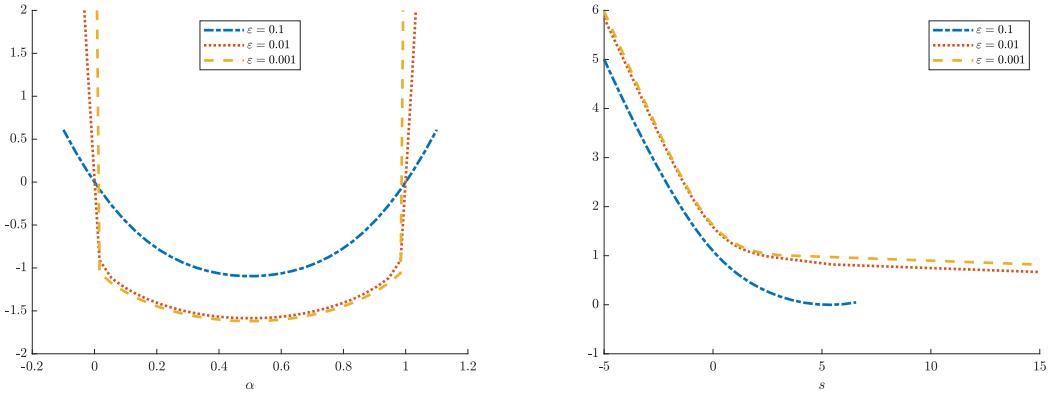
$$V(x_1, x_2) = -\frac{x_1^2 + x_2^2}{4} + \frac{x_1^4 + 2x_1^2 x_2^2 + x_2^4}{8},$$

and

$$b(x_1, x_2) = (\cos(x_1)\sin(x_2), -\sin(x_1)\cos(x_2)).$$

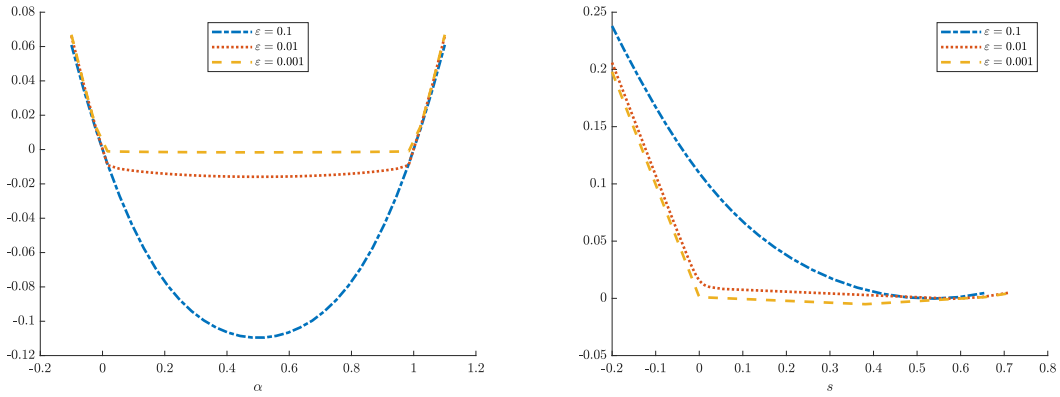
Note that the $\nabla V(x) = 0$ for all x on the circle $\{(x_1, x_2) : x_1^2 + x_2^2 = 1\}$, whereas b acts nontrivially along that circle. In Figures 13a and 14a, we see that the eigenvalue is of different orders in ε depending on whether $\alpha \in [0, 1]$ or $\alpha \notin [0, 1]$.

For this example, we use $T = 1024$. We show $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ in Figure 13a and its Legendre transform $\widehat{I}^{\varepsilon}_{\Delta t}(s)$ in Figure 13b. In particular, the zero of $\widehat{I}^{\varepsilon}_{\Delta t}$ — which is the mean entropy production rate for that value of ε — seems to diverge as $\varepsilon \to 0^+$, as expected due to the inverse power of ε in the definition of the entropy production and the periodic orbit of the deterministic dynamics along which the work done by b per unit time is nonzero. We also show $\varepsilon\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ in Figure 14a and its Legendre transform $\varepsilon\widehat{I}^{\varepsilon}_{\Delta t}(\varepsilon^{-1}s)$ in Figure 14b, as studied in [7, 8]. In particular, a key feature discussed in Section 5 of [7] is emerging as $\varepsilon \to 0^+$: a kink in $\varepsilon\widehat{I}^{\varepsilon}_{\Delta t}(\varepsilon^{-1}s)$ at $s = 0$, where two flat regions meet at an angle compatible with the Gallavotti–Cohen symmetry. The example also confirms that in some (but not all) scenarios, the limits of I^{ε} and $\varepsilon I^{\varepsilon}(\varepsilon^{-1}\,\cdot\,)$ provide complementary, nontrivial information on the fluctuations of S^{ε}_t.



(a) $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ (b) $\widehat{I}^{\varepsilon}_{\Delta t}(s)$

Figure 13: In the context of Example 5.5, we plot our numerical approximation $\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ of the principal eigenvalue $\lambda^{\varepsilon,\alpha}$ and the resulting approximation $\widehat{I}^{\varepsilon}_{\Delta t}(s)$ of the rate function $I^{\varepsilon}(s)$.



(a) $\varepsilon\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ (b) $\varepsilon\widehat{I}^{\varepsilon}_{\Delta t}(\varepsilon^{-1}s)$

Figure 14: In the context of Example 5.5, we plot our numerical approximation $\varepsilon\widehat{\lambda}^{\varepsilon,\alpha}_{\Delta t}$ of the rescaled principal eigenvalue $\varepsilon\lambda^{\varepsilon,\alpha}$ and the resulting approximation $\varepsilon\widehat{I}^{\varepsilon}_{\Delta t}(\varepsilon^{-1}s)$ of the rescaled rate function $\varepsilon I^{\varepsilon}(\varepsilon^{-1}s)$.

6. Conclusion

We study an interacting particle method for the computation of rate functions I^{ε} for the large deviations of entropy production in the context of diffusion processes by equivalently computing the principal eigenvalue for a family of non-self-adjoint elliptic operators. We are particularly interested in the high-dimensional and vanishing-noise case, which is challenging to traditional numerical methods. We show that the principal eigenvalue can be well approximated in terms of the spectral radius of a discretized semigroup, making it suitable for an IPM. Moreover, we discuss two techniques for setting the initial measure in the IPM for faster computation. We present numerical examples in dimensions up to 16. The numerical results provide evidence that the numerical principal eigenvalue converges to the analytical vanishing-noise limit with a fixed number of particles and a fixed time step size. Furthermore, the asymptotic behavior of the empirical density at the final time in the vanishing-noise limit is consistent with the theory in [26]. Our paper appears to be the first one to obtain numerical results of principal eigenvalue problems in such high dimensions. Our method also allows us to probe the rate function I^{ε} in situations where no explicit formulas are available, as well as to explore the gap between the theoretical works on different scalings for the vanishing-noise limit $\varepsilon \to 0^{+}$.

In the future, it would be interesting to investigate the error estimate of the IPM with respect to the numerical parameters of the method. Furthermore, the method should also be used to study the large deviation rate functions in situations that go beyond the scope of the theoretical works [7, 8, 37, 54], e.g. combining non-linearity of the vector field with the degeneracy of the noise.

Acknowledgements

RR was partially funded by the *Fonds de recherche du Québec — Nature et technologies* (FRQNT) and by the Natural Sciences and Engineering Research Council of Canada (NSERC). JX was partially supported by NSF grant DMS-2309520. ZZ was supported by the National Natural Science Foundation of China (Project 12171406), Hong Kong RGC grant (Project 17307921), Seed Funding Programme for Basic Research (HKU), the Outstanding Young Researcher Award of HKU (2020–21), and Seed Funding for Strategic Interdisciplinary Research Scheme 2021/22 (HKU), and an R&D Funding Scheme from the HKU-SCF FinTech Academy.

The project was initiated at Courant Institute, New York University where JX was visiting in the Fall of 2022. The authors would like to thank Professors R. Caflisch, R. Kohn, D. McLaughlin, C. Peskin, S. R. S. Varadhan, and L.-S. Young for helpful conversations and scientific activities that made our collaboration possible.

The computations were performed using research computing facilities provided by Information Technology Services, the University of Hong Kong.

References

[1] B. Anderson, *Reverse-time diffusion equation models*, Stoch. Process. Appl., 12(3) (1982), pp. 313–326.

[2] B. Anderson and T. Kailath, *Forwards, backwards, and dynamically reversible Markovian models of second-order processes*, IEEE Trans. Circuits Syst., 26 (1979), pp. 956–965.

[3] B. Anderson and T. Kailath, *Passive network synthesis via dual spectral factorization*, IEEE Trans. Circuits Syst., 26 (1979), pp. 866–873.

[4] J. B. Anderson, *A random-walk simulation of the Schrödinger equation: H+ 3*, J. Chem. Phys., 63 (1975), pp. 1499–1503.

[5] P. Anselone and T. Palmer, *Spectral analysis of collectively compact, strongly convergent operator sequences*, Pacific J. Math., 25 (1968), pp. 423–431.

[6] O. Bencheikh and B. Jourdain, *Convergence in total variation of the Euler–Maruyama scheme applied to diffusion processes with measurable drift coefficient and additive noise*, SIAM J. Numer. Anal., 60 (2022), pp. 1701–1740.

[7] L. Bertini and G. Di Gesù, *Small noise asymptotic of the Gallavotti–Cohen functional for diffusion processes*, ALEA, Lat. Am. J. Probab. Math. Stat., 12 (2015), pp. 743–763.

[8] L. Bertini, D. Gabrielli, and C. Landim, *Large deviations for diffusions: Donsker–Varadhan meet Freidlin–Wentzell*, arXiv preprint arXiv:2211.02593, (2022).

[9] A. Carasso, *Finite-difference methods and the eigenvalue problem for nonselfadjoint Sturm–Liouville operators*, Math. Comp., 23 (1969), pp. 717–729.

[10] D. M. Ceperley and B. J. Alder, *Ground state of the electron gas by a stochastic method*, Phys. Rev. Lett., 45 (1980), p. 566.

[11] G. E. Crooks, *Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences*, Phys. Rev. E, 60 (1999), p. 2721.

[12] P. Del Moral, *Nonlinear filtering: Interacting particle resolution*, C. R. Acad. Sci. Paris, 325 (1997), pp. 653–658.

[13] ——, *Feynman–Kac formulae*, Springer, 2004.

[14] P. Del Moral and A. Guionnet, *Central limit theorem for nonlinear filtering and interacting particle systems*, Ann. Appl. Probab., (1999), pp. 275–297.

[15] P. Del Moral and L. Miclo, *Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to non-linear filtering*, in Séminaire de Probabilités XXXIV, J. Azéma, M. Ledoux, M. Émery, and M. Yor, eds., Springer Berlin Heidelberg, 2000, pp. 1–145.

[16] F. den Hollander, *Large deviations*, vol. 14 of Fields Institute monographs, American Mathematical Society, 2000.

[17] M. D. Donsker and S. S. Varadhan, *On a variational formula for the principal eigenvalue for operators with maximum principle*, Proc. Natl. Acad. Sci., 72 (1975), pp. 780–783.

[18] A. Doucet, N. De Freitas, N. J. Gordon, et al., *Sequential Monte Carlo methods in practice*, Springer, 2001.

[19] M. H. Duong and J. Zimmer, *On decompositions of non-reversible processes*, J. Phys. Conf. Ser., 2514 (2023), p. 012007.

[20] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Entropy production in nonlinear, thermally driven hamiltonian systems*, J. Stat. Phys., 95 (1999), pp. 305–331.

[21] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, vol. 194 of Graduate Texts in Mathematics, Springer, 2000.

[22] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, *Probability of second law violations in shearing steady states*, Phys. Rev. Lett., 71 (1993), p. 2401.

[23] D. J. Evans and D. J. Searles, *Equilibrium microstates which generate second law violating steady states*, Phys. Rev. E, 50 (1994), p. 1645.

[24] G. Ferré, M. Rousset, and G. Stoltz, *More on the long time stability of Feynman–Kac semigroups*, Stoch. Partial Differ. Equ. Anal. Comput., 9 (2021), pp. 630–673.

[25] G. Ferré and G. Stoltz, *Error estimates on ergodic properties of discretized Feynman–Kac semigroups*, Numer. Math., 143 (2019), pp. 261–313.

[26] W. H. Fleming and S.-J. Sheu, *Asymptotics for the principal eigenvalue and eigenfunction of a nearly first-order operator with large potential*, Ann. Probab., 25 (1997), pp. 1953–1994.

[27] R. Fortet and E. Mourier, *Convergence de la répartition empirique vers la répartition théorique*, Ann. Sci. École Norm. Sup. (3), 70 (1953), pp. 267–285.

[28] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, *Quantum Monte Carlo simulations of solids*, Rev. Mod. Phys., 73 (2001), p. 33.

[29] G. Gallavotti and E. G. D. Cohen, *Dynamical ensembles in nonequilibrium statistical mechanics*, Phys. Rev. Lett., 74 (1995), p. 2694.

[30] C. Giardina, J. Kurchan, and L. Peliti, *Direct evaluation of large-deviation functions*, Phys. Rev. Lett., 96 (2006), p. 120603.

[31] R. Grimm and R. Storer, *Monte-Carlo solution of Schrödinger's equation*, J. Comput. Phys., 7 (1971), pp. 134–156.

[32] M. Hairer, *Ergodic properties of a class of non-Markovian processes*, in Trends in Stochastic Analysis, vol. 353 of (London Math. Soc. Lecture Note Ser.), Cambridge University Press, 2009.

[33] M. Hairer and J. Weare, *Improved diffusion Monte Carlo*, Commun. Pure Appl. Math., 67 (2014), pp. 1995–2021.

[34] H. Han and X. Wu, *Artificial boundary method*, Springer Science & Business Media, 2013.

[35] U. G. Haussmann and E. Pardoux, *Time reversal of diffusions*, Ann. Probab., (1986), pp. 1188–1205.

[36] E. G. Hidalgo, T. Nemoto, and V. Lecomte, *Finite-time and finite-size scalings in the evaluation of large-deviation functions: Numerical approach in continuous time*, Phys. Rev. E, 95 (2017), p. 062134.

[37] V. Jakšić, C.-A. Pillet, and A. Shirikyan, *Entropic fluctuations in thermally driven harmonic networks*, J. Stat. Phys., 166 (2017), pp. 926–1015.

[38] T. Kato, *Perturbation theory for linear operators*, vol. 132 of Grundlehren der mathematischen Wissenschaften, Springer, second ed., 1995.

[39] A. N. Kolmogorov, *Zur Umkehrbarkeit der statistischen Naturgesetze (Engl.: On the reversibility of the statistical laws of nature)*, Math. Ann., 113 (1937), pp. 766–772.

[40] I. Kontoyiannis and S. Meyn, *Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes*, Electron. J. Probab., 10 (2005), pp. 61–123.

[41] J. Kurchan, *Fluctuation theorem for stochastic dynamics*, J. Phys. A, 31 (1998), p. 3719.

[42] ——, *Gallavotti–Cohen theorem, chaotic hypothesis and the zero-noise limit*, J. Stat. Phys., 128 (2007), pp. 1307–1320.

[43] J. R. Kuttler, *Finite difference approximations for eigenvalues of uniformly elliptic operators*, SIAM J. Numer. Anal., 7 (1970), pp. 206–232.

[44] J. L. Lebowitz and H. Spohn, *A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics*, J. Stat. Phys., 95 (1999), pp. 333–365.

[45] V. Lecomte and J. Tailleur, *A numerical approach to large deviations in continuous time*, J. Stat. Mech. Theory Exp., 2007 (2007), p. P03004.

[46] T. Lelièvre, M. Rousset, and G. Stoltz, *Free energy computations: a mathematical perspective*, World Scientific, 2010.

[47] L.-H. Lim and J. Weare, *Fast randomized iteration: Diffusion Monte Carlo through the lens of numerical linear algebra*, SIAM Rev., 59 (2017), pp. 547–587.

[48] L. Ljung and T. Kailath, *Backwards Markovian models for second-order stochastic processes (corresp.)*, IEEE Trans. Inform. Theory, 22 (1976), pp. 488–491.

[49] J. Lyu, Z. Wang, J. Xin, and Z. Zhang, *Convergence analysis of stochastic structure-preserving schemes for computing effective diffusivity in random flows*, SIAM J. Numer. Anal., 58 (2020), pp. 3040–3067.

[50] ——, *A convergent interacting particle method and computation of KPP front speeds in chaotic flows*, SIAM J. Numer. Anal., 60 (2022), pp. 1136–1167.

[51] C. Monthus, *Large deviations for trajectory observables of diffusion processes in dimension $d > 1$ in the double limit of large time and small diffusion coefficient*, J. Stat. Mech.: Theory Exp., (2024), p. 013205.

[52] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, *Population-dynamics method with a multicanonical feedback control*, Phys. Rev. E, 93 (2016), p. 062123.

[53] T. Nemoto, E. G. Hidalgo, and V. Lecomte, *Finite-time and finite-size scalings in the evaluation of large-deviation functions: Analytical study using a birth-death process*, Phys. Rev. E, 95 (2017), p. 012102.

[54] R. Raquépas, *The large-time and vanishing-noise limits for entropy production in non-degenerate diffusions*, Ann. Inst. Henri Poincaré B: Probab. Stat., 60 (2024), pp. 431–462.

[55] G. Sidhu and U. Desai, *New smoothing algorithms based on reversed-time lumped models*, IEEE Trans. Automat. Control, 21 (1976), pp. 538–541.

[56] J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, and S. Ganguli, *Deep unsupervised learning using nonequilibrium thermodynamics*, International Conference on Machine Learning, (2015).

[57] J. Song, C. Meng, and S. Ermon, *Denoising diffusion implicit models*, ICLR, (2021).

[58] J. Sun and A. Zhou, *Finite element methods for eigenvalue problems*, CRC Press, 2016.

[59] J. Tailleur and V. Lecomte, *Simulation of large deviation functions using population dynamics*, in AIP Conference Proceedings, vol. 1091, American Institute of Physics, 2009, pp. 212–219.

[60] M. Talagrand, *The Glivenko–Cantelli problem*, Ann. Probab., 15 (1987), pp. 837–870.

[61] H. Thorisson, *Coupling methods in probability theory*, Scandinavian J. Stat., (1995), pp. 159–182.

[62] H. Touchette, *The large deviation approach to statistical mechanics*, Phys. Rep., 478 (2009), pp. 1–69.

[63] H. F. Trotter, *On the product of semi-groups of operators*, Proc. Amer. Math. Soc., 10 (1959), pp. 545–551.

[64] R. van Zon and E. G. D. Cohen, *Extension of the fluctuation theorem*, Phys. Rev. Lett., 91 (2003), p. 110601.

[65] S. R. S. Varadhan, *Large deviations and applications*, SIAM, 1984.

[66] Z. Wang, J. Xin, and Z. Zhang, *Computing effective diffusivity of chaotic and stochastic flows using structure-preserving schemes*, SIAM J. Numer. Anal., 56 (2018), pp. 2322–2344.

[67] Z. Wang, J. Xin, and Z. Zhang, *Sharp uniform in time error estimate on a stochastic structure-preserving Lagrangian method and computation of effective diffusivity in 3D chaotic flows*, Multiscale Model. Simul., 19 (2021), pp. 1167–1189.

[68] Z. Wang, J. Xin, and Z. Zhang, *Computing effective diffusivities in 3D time-dependent chaotic flows with a convergent Lagrangian numerical method*, ESAIM: Math. Model. Numer. Anal., 56 (2022), pp. 1521–1544.

[69] L. Wu, *Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems*, Stochastic Process. Appl., 91 (2001), pp. 205–238.

[70] T. Zhang, Z. Wang, J. Xin, and Z. Zhang, *A convergent interacting particle method for computing KPP front speeds in random flows*, arXiv preprint arXiv:2308.14479, (2023).