Irreversible Langevin samplers and variance reduction: a large deviations approach

2015 Nonlinearity 28 2081

(http://iopscience.iop.org/0951-7715/28/7/2081)

IOP Publishing | London Mathematical Society Nonlinearity

Nonlinearity **28** (2015) 2081–2103 doi:10.1088/0951-7715/28/7/2081

Irreversible Langevin samplers and variance reduction: a large deviations approach

Luc Rey-Bellet[1] **and Konstantinos Spiliopoulos**[2]

[1] Department of Mathematics and Statistics, University of Massachusetts Amherst, Amherst, MA 01003, USA
[2] Department of Mathematics and Statistics, Boston University, Boston, MA 02215, USA

E-mail: luc@math.umass.edu and kspiliop@math.bu.edu

Received 27 August 2014, revised 17 February 2015
Accepted for publication 22 April 2015
Published 19 May 2015

CrossMark

Recommended by Dmitry Dolgopyat

Abstract

In order to sample from a given target distribution (often of Gibbs type), the Monte Carlo Markov chain method consists of constructing an ergodic Markov process whose invariant measure is the target distribution. By sampling the Markov process one can then compute, approximately, expectations of observables with respect to the target distribution. Often the Markov processes used in practice are time-reversible (i.e. they satisfy detailed balance), but our main goal here is to assess and quantify how the addition of a non-reversible part to the process can be used to improve the sampling properties. We focus on the diffusion setting (overdamped Langevin equations) where the drift consists of a gradient vector field as well as another drift which breaks the reversibility of the process but is chosen to preserve the Gibbs measure. In this paper we use the large deviation rate function for the empirical measure as a tool to analyze the speed of convergence to the invariant measure. We show that the addition of an irreversible drift leads to a larger rate function and it strictly improves the speed of convergence of ergodic average for (generic smooth) observables. We also deduce from this result that the asymptotic variance decreases under the addition of the irreversible drift and we give an explicit characterization of the observables whose variance is not reduced reduced, in terms of a nonlinear Poisson equation. Our theoretical results are illustrated and supplemented by numerical simulations.

Keywords: non reversible Monte Carlo Markov chains, large deviations, variance reduction, Langevin equation
Mathematics Subject Classification: 60F05, 60F10, 60J25, 60J60, 65C05, 82B80

(Some figures may appear in colour only in the online journal)

1. Introduction

In a wide range of applications it is often of interest to sample from a given high-dimensional distribution. However, often, the target distribution, say $\bar{\pi}(\mathrm{d}x)$, is known only up to normalizing constants and then one has to rely on approximations. In practice, one often relies on approximations using Markov processes that have the particular target distributions as their invariant measure, as for example in Monte Carlo Markov Chain methods. Closely related, in steady-state simulations one is often interested in quantities of the form $\int_E f(x)\bar{\pi}(\mathrm{d}x)$, where E is the state space and f is a given function. When closed-form evaluation of such integrals is prohibitive, one considers a Markov process X_t which has $\bar{\pi}$ as its invariant distribution and under the assumption that X_t is positive recurrent, the ergodic theorem gives

$$\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s \to \int_E f(x)\bar{\pi}(\mathrm{d}x), \text{ a.s. as } t \to \infty, \tag{1.1}$$

for all $f \in L^1(\bar{\pi})$. Hence, the estimator $f_t \equiv \frac{1}{t}\int_0^t f(X_s)\mathrm{d}s$ can be used to approximate the expectation $\bar{f} \equiv \int_E f(x)\bar{\pi}(\mathrm{d}x)$.

Standard criteria to analyze the degree of efficiency of a simulation method relies on the ergodic properties of the Markov process. The spectral gap of the semigroup in $L^2(\pi)$ (or in other functional settings), which provides a bound for the distance between the distribution of X_t and π, as well as the asymptotic variance of f_t are commonly used, see for example [1, 3, 5, 6, 9, 10, 15–17, 19–21, 23, 26–30, 31, 33]. A couple of years ago, in [13, 14], the theory of large deviations, specifically the rate function for the empirical measure, has been proposed as a comparison tool to assess Monte-Carlo methods and used to analyze the swapping algorithm. In this paper we use this criterium as a guide to design and analyze non-reversible Markov processes and compare them with reversible ones. We show that the rate function increases under the addition of an irreversible drift. This is shown to improve the convergence properties of the ergodic average f_t for generic (smooth) observables. We prove as well that a fine analysis of the large deviation rate function allows us to show that the asymptotic variance for generic smooth observables decreases.

In this paper, we specialize to the diffusion setting: to sample the Gibbs measure $\bar{\pi}$ on the set E with density

$$\frac{\mathrm{e}^{-2U(x)}}{\int_E \mathrm{e}^{-2U(x)}\mathrm{d}x},$$

one can consider the (time-reversible) Langevin equation

$$\mathrm{d}X_t = -\nabla U(X_t)\mathrm{d}t + \mathrm{d}W_t, \tag{1.2}$$

whose invariant measure is $\bar{\pi}$. There are however many other stochastic differential equations with the same invariant measure, for example the family of equations

$$\mathrm{d}X_t = [-\nabla U(X_t) + C(X_t)]\,\mathrm{d}t + \mathrm{d}W_t, \tag{1.3}$$

where the vector field $C(x)$ satisfies the condition

$$\mathrm{div}(C\mathrm{e}^{-2U}) = 0.$$

This constraint ensures that $\bar{\pi}$ remains the unique invariant measure, but then the Markov process is time-reversible only if $C = 0$. There are many possible choices for the vector field $C(x)$. Indeed, since $\mathrm{div}(C\mathrm{e}^{-2U}) = 0$ is equivalent to

$$\mathrm{div}(C) = 2C\nabla U,$$

we can choose for example C to be both divergence free and orthogonal to ∇U. In any dimension one can for example set $C = S\nabla U$ where S is an (arbitrary) anti-symmetric matrix S. More generally, by theorem 5.3 of [4] any divergence free vector field in dimension d can be written, locally, as the exterior (or wedge) product $C = \nabla V_1 \wedge \cdots \wedge \nabla V_{n-1}$ for some $V_i \in \mathcal{C}^1(E)$. Therefore for our purpose we can pick C of the form

$$C = \nabla U \wedge \nabla V_2 \cdots \wedge \nabla V_{n-1}.$$

for arbitrary $V_2, \cdots, V_{n-1} \in \mathcal{C}^1(E)$, and this guarantees that $C\nabla U = 0$ by the properties of the exterior product.

The main result in [21] is that the absolute value of the second largest eigenvalue of the Markov semigroup in $L^2(\bar{\pi})$ strictly decreases under a natural non-degeneracy condition on C (the corresponding eigenspace should not be invariant under the action of the added drift C). More detailed results on the spectral gap are in [7, 15] where the authors consider diffusions on compact manifolds with $U = 0$ and a one-parameter families of perturbations $C = \delta C_0$ for $\delta \in \mathbb{R}$ and C_0 is some divergence vector field. In these papers the behavior of the spectral gap is related to the ergodic properties of the flow generated by C (for example if the flow is weak-mixing then the second largest eigenvalue tends to 0 as $\delta \to \infty$). Further, a detailed analysis of linear diffusion processes with $U(x) = \frac{1}{2}x^T Ax$ and $C = JAx$ for a antisymmetric J can be found in [20, 23] where the optimal choice of J is determined.

We consider here the same class of problems but we take the large deviations rate function as a measure of the speed of convergence to equilibrium and deduce from it results on the asymptotic variance for a given observable. While the spectral gap measures the distance of the distribution of X_t compared to the invariant distribution, from a practical Monte-Carlo point of view one is often more interested in the distribution of the ergodic average $t^{-1}\int_0^t f(X_s)\,\mathrm{d}s$ and how likely it is that this average differs from the average $\int f \mathrm{d}\bar{\pi}$. It will be useful to consider in a first step the empirical measure

$$\pi_t \equiv \frac{1}{t}\int_0^t \delta_{X_s}\,\mathrm{d}s, \tag{1.4}$$

which converges to $\bar{\pi}$ almost surely. Let us assume that we have a large deviation principle for the family of measures π_t, which we write, symbolically as

$$\mathbb{P}\{\pi_t \approx \mu\} \asymp \mathrm{e}^{-tI_C(\mu)}.$$

Here $\asymp$ denotes logarithmic equivalence (the formal definition is given in definition 2.1). Then, the rate function $I_C(\mu)$ which is non-negative and vanishes if and only if $\mu = \bar{\pi}$ quantifies the exponential rate at which the random measure π_t converges to $\bar{\pi}$. Clearly, the larger I_C is, the faster the convergence occurs.

Breaking detailed balance has been shown to accelerate convergence to equilibrium for Markov chains by increasing spectral gap and/or decreasing asymptotic variance and for diffusions by increasing spectral gap, e.g. [6, 9, 10, 15–17, 20, 21, 26–28]. The novelty of the present paper lies in that (a): we use large deviations theory in a novel way to characterize convergence to equilibrium, (b): we prove that asymptotic variance is also decreased when breaking detailed balance for diffusions, and (c): we derive a Poisson equation which characterizes when irreversible perturbations lead to strict improvement in performance.

Our first key result here is that if $\mu(\mathrm{d}x) = p(x)\mathrm{d}x$ has a smooth density p and satisfies the non-degeneracy condition $\mathrm{div}(pC) \neq 0$, the large deviation rate function strictly increases, $I_C(\mu) > I_0(\mu)$, when one adds a non-zero appropriate drift $C(x)$ to make the process X_t irreversible, see theorem 2.2. Moreover, specializing to perturbations of the form $C(x) = \delta C_0(x)$ for appropriate $C_0(x)$ and $\delta \in \mathbb{R}$, we find that the rate function for the empirical measure is quadratic in $\delta \in \mathbb{R}$, see theorem 2.3.

Our second key result is that the information in $I_C(\mu)$ can be used to study specific observable: from the large deviation for the empirical measure we have a large deviation for principle for observables $f \in \mathcal{C}(E;\mathbb{R})$,

$$\mathbb{P}\left\{\frac{1}{t}\int_0^t f(X_s)\,\mathrm{d}s \approx \ell\right\} \asymp \mathrm{e}^{-t\tilde{I}_{f,C}(\ell)},$$

and we show that $\tilde{I}_{f,C}(\ell) > \tilde{I}_{f,0}(\ell)$ unless f and ℓ satisfy the non degeneracy condition (in form of a Poisson equation) given in theorem 2.4, see also remarks 2.5 and 2.6.

Moreover, one can deduce information about asymptotic variances from the large deviations rate function, since the second derivative of the rate function $\tilde{I}_{f,C}(\ell)$ evaluated at $\ell = \bar{f}$ is inversely proportional to the asymptotic variance of the estimator, denoted by $\sigma^2_{f,C}$. Based on this relation, we show that the asymptotic variance strictly decreases $\sigma^2_{f,C} < \sigma^2_{f,0}$, for generic observables.

The paper is organized as follows. In section 2 we recall some well-known results about large deviations due to Donsker–Varadhan and Gärtner and we present our main results. Proofs of statements related to the rate function for the empirical measure are in section 3. In particular, we prove theorems 2.2 and 2.3 by using a representation of the rate function $I(\mu)$ due to Gärtner [18]. Proofs related to the rate function for a given observable and the results for variance reduction are in section 4. In particular, we use the results of section 3 to deduce the results on the rate function and asymptotic variance for observables, i.e. theorems 2.4 and 2.7. In section 5 we present a few simulation results to illustrate the theoretical findings.

2. Main results

Let us first recall the definition of the large deviations principle for a family of empirical measures π_t. Let E be a Polish space, i.e. a complete and separable metric space. Denoting by $\mathcal{P}(E)$ the space of all probability measures on E, we equip $\mathcal{P}(E)$ with the topology of weak convergence, which makes $\mathcal{P}(E)$ metrizable and a Polish space.

Definition 2.1. *Consider a sequence of random probability measures $\{\pi_t\}$. The family $\{\pi_t\}$ is said to satisfy a large deviations principle (LDP) with rate function (equivalently action functional) $I : \mathcal{P}(E) \mapsto [0,\infty]$ if the following conditions hold:*

- *For all open sets $O \subset \mathcal{P}(E)$, we have*
$$\liminf_{t\to\infty} \frac{1}{t}\log \mathbb{P}\{\pi_t \in O\} \geqslant -\inf_{\mu\in O} I(\mu).$$
- *For all closed sets $F \subset \mathcal{P}(E)$, we have*
$$\limsup_{t\to\infty} \frac{1}{t}\log \mathbb{P}\{\pi_t \in F\} \leqslant -\inf_{\mu\in F} I(\mu).$$
- *The level sets $\{\mu : I(\mu) \leqslant M\}$ are compact in $\mathcal{P}(E)$ for all $M < \infty$.*

If the random measures π_t are the empirical measures of an ergodic Markov process X_t (see 1.4) with invariant distribution $\bar{\pi}$ then $I(\mu)$ is a nonnegative convex function with $I(\bar{\pi}) = 0$ and thus $I(\mu)$ controls the rate at which the random measure π_t concentrates to $\bar{\pi}$.

For convenience we will assume that the diffusion process X_t which solves the SDE 1.3 takes values in a compact space and that the vector fields are sufficiently smooth. We fully expect, though, our result to still hold in $\mathbb{R}^d$ under suitable confining assumptions on the potential U to ensure a large deviation principle. Throughout the rest of the paper we assume that

(H) The state space E is a connected, compact, d-dimensional smooth Riemann manifold without boundary, and there exists an $\alpha \in (0,1)$ such that the potential $U \in \mathcal{C}^{(2+\alpha)}(E)$ and the vector field $C \in \mathcal{C}^{(1+\alpha)}(E)$. Moreover, we assume that $\mathrm{div}(C\mathrm{e}^{-2U}) = 0$ so that the measure $\bar{\pi}$ is invariant.

From the work of Gärtner and Donsker–Varadhan, [11, 18], under condition (**H**), the empirical measures π_t satisfy a large deviation principle which is uniform in the initial condition, i.e. the rate function is independent of the distribution of $X_0 \sim \mu_0$. Let us denote by $\mathcal{L}$ the infinitesimal generator of the Markov process X_t and by $\mathcal{D}$ its domain of definition. The rate function $I(\mu)$ (usually referred to as the Donsker–Varadhan functional) takes the form

$$I(\mu) = -\inf_{u \in \{u \in \mathcal{D}, u>0\}} \int_E \frac{\mathcal{L}u}{u} \mathrm{d}\mu.$$

An alternative formula for $I(\mu)$, more useful in the context of this paper, is given in terms of the Legendre transform

$$I(\mu) = \sup_{f \in \mathcal{C}(E)} \left\{ \int f \,\mathrm{d}\mu - \lambda(f) \right\},$$

where $\lambda(f)$ is the maximal eigenvalue of the Feyman–Kac semigroup $T_t^f h(x) = \mathbb{E}_x[\mathrm{e}^{\int_0^t f(X_s)\mathrm{d}s} h(X_t)]$ acting on the Banach space $\mathcal{C}(E;\mathbb{R})$. As shown in [18] for nice μ this formula can be used to derive a useful, more explicit, formula for $I(\mu)$ which will be central in our analysis (see theorem 3.1 below).

In the sequel and in order to emphasize the dependence on C of the rate function we will use the notation $I_C(\mu)$. Our first two results show that adding an irreversible drift C increases the Donsker–Varadhan rate function pointwise.

Theorem 2.2. *Assume that* $C \neq 0$ *is as in Assumption* (**H**). *For any* $\mu \in \mathcal{P}(E)$ *we have* $I_C(\mu) \geqslant I_0(\mu)$. *Let* $\mu(\mathrm{d}x) = p(x)\mathrm{d}x$ *be a probability measure with positive density* $p \in \mathcal{C}^{(2+\alpha)}(E)$ *for some* $\alpha > 0$ *and* $\mu \neq \bar{\pi}$. *Then, we have*

$$I_C(\mu) = I_0(\mu) + \frac{1}{2}\int_E |\nabla \psi_C(x) - \nabla U(x)|^2 \,\mathrm{d}\mu(x).$$

where ψ_C *is the unique solution (up to a constant) of the elliptic equation*

$$\mathrm{div}\,[p\,(-\nabla U + C + \nabla \psi_C)] = 0.$$

Moreover, we have $I_C(\mu) = I_0(\mu)$ *if and only if the positive density* $p(x)$ *satisfies* $\mathrm{div}\,(p(x)C(x)) = 0$. *Equivalently such* p *have the form* $p(x) = \mathrm{e}^{2G(x)}$ *where* G *is such that* $G + U$ *is an invariant for the vector field* C *(i.e.* $C\nabla(G+U) = 0$*).*

To obtain a slightly more quantitative result let us consider a one-parameter family $C(x) = \delta C_0(X)$ where $\delta \in \mathbb{R}$ and C_0. We show that for any fixed measure μ the functional $I_{\delta C_0}(\mu)$ is quadratic in $\delta \in \mathbb{R}$.

Theorem 2.3. *Assume that* $C = \delta C_0 \neq 0$ *is as in Assumption* (**H**) *and consider the measure* $\mu(\mathrm{d}x) = p(x)\mathrm{d}x$ *with positive density* $p \in \mathcal{C}^{(2+\alpha)}(E)$ *for some* $\alpha > 0$. *Then we have*

$$I_{\delta C_0}(\mu) = I_0(\mu) + \delta^2 K(\mu),$$

where the functional $K(\mu)$ *is strictly positive if and only if* $\mathrm{div}\,(p(x)C_0(x)) \neq 0$. *Moreover, the functional* $K(\mu)$ *takes the explicit form*

$$K(\mu) = \frac{1}{2}\int_E |\nabla \xi(x)|^2 \,\mathrm{d}\mu(x),$$

where ξ *is the unique solution (up to a constant) of the elliptic equation*

$$\mathrm{div}\,[p\,(C_0 + \nabla \xi)] = 0.$$

For $f \in \mathcal{C}(E)$ the contraction principle implies that the ergodic average $\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s$ satisfies a large deviation principle with the rate function

$$\tilde{I}_{f,C}(\ell) = \inf_{\mu\in\mathcal{P}(E)} \{I_C(\mu) : \langle f, \mu\rangle = \ell\}.$$

Note that $\tilde{I}_{f,C}(\ell)$ can also be expressed in terms of a Legendre transform

$$\tilde{I}_{f,C}(\ell) = \sup_{\beta\in\mathbb{R}} \{\beta\ell - \lambda(\beta f)\},$$

where

$$\lambda(\beta f) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left[\mathrm{e}^{\int_0^t \beta f(X_s)\mathrm{d}s}\right].$$

The eigenvalue $\lambda(\beta f)$ is a smooth strictly convex function of β so that if ℓ belongs to the range of f we have

$$\tilde{I}_{f,C}(\ell) = \widehat{\beta}\ell - \lambda(\widehat{\beta} f), \quad \text{with } \widehat{\beta} \text{ given by } \ell = \frac{\mathrm{d}}{\mathrm{d}\beta}\lambda(\widehat{\beta} f).$$

In fact, if $f \in \mathcal{C}^{(\alpha)}(E)$, then by proposition 4.1 there is $\mu_C^*(\mathrm{d}x) = p_C(x)\mathrm{d}x$, with $p_C(x) > 0$ and $p_C \in \mathcal{C}^{(2+\alpha)}(E)$ such that $\tilde{I}_{f,C}(\ell) = I_C(\mu_C^*)$. Then, theorem 2.2 and proposition 4.1 give theorem 2.4. Theorem 2.4 shows that the rate function for observables increases pointwise under a non-degeneracy condition.

Theorem 2.4. *Assume that $C \neq 0$ is as in Assumption* (**H**). *Consider $f \in \mathcal{C}^{(\alpha)}(E)$ and $\ell \in (\min_x f(x), \max_x f(x))$ with $\ell \neq \int f d\bar{\pi}$. Then we have*

$$\tilde{I}_{f,C}(\ell) \geqslant \tilde{I}_{f,0}(\ell).$$

Moreover if there exists ℓ_0 such that for the vector field C, $\tilde{I}_{f,C}(\ell_0) = \tilde{I}_{f,0}(\ell_0)$ then we must have

$$\widehat{\beta}(\ell_0) f = \frac{1}{2}\Delta(G+U) + \frac{1}{2}|\nabla G|^2 - \frac{1}{2}|\nabla U|^2, \tag{2.1}$$

where G is such that $G + U$ is invariant under the particular vector field C.

The following remarks are of interest.

Remark 2.5. Letting $\mathcal{L}_0$ denote the infinitesimal generator of the reversible process X_t defined in (1.2), we get that (2.1) can be rewritten as a nonlinear Poisson equation of the form

$$\widehat{\beta}(\ell_0) f = \mathcal{H}(G+U),$$

where

$$\mathcal{H}(G+U) = \mathrm{e}^{-(G+U)}\mathcal{L}_0\mathrm{e}^{G+U} = \frac{1}{2}\Delta(G+U) + \frac{1}{2}|\nabla G|^2 - \frac{1}{2}|\nabla U|^2.$$

Recalling theorem 2.2 (see the proof of theorem 2.4), an alternative condition that gives $\tilde{I}_{f,C}(\ell_0) = \tilde{I}_{f,0}(\ell_0)$ is as follows. By proposition 4.1 there is $\mu_C^*(\mathrm{d}x; \ell_0) = p_C(x; \ell_0)\mathrm{d}x$, with $p_C > 0$ and $p_C \in \mathcal{C}^{(2+\alpha)}(E)$ such that $\tilde{I}_{f,C}(\ell) = I_C(\mu_C^*(\cdot; \ell_0))$. Then, the condition $\mathrm{div}(p_C(x; \ell_0)C(x)) = 0$, implies that $\tilde{I}_{f,C}(\ell_0) = \tilde{I}_{f,0}(\ell_0)$.

Remark 2.6. In is interesting to note here that the Poisson equation (2.1) is reminiscent of Poisson equations that have appeared in the literature in the analysis of MCMC algorithms, see for example chapter 17 of [25]. In this paper, we see that the particular Poisson equation can be also used to characterize when irreversible perturbations do actually strictly improve convergence to equilibrium.

A standard measure of efficiency of a sampling method for an observable f is to use the asymptotic variance. Under our assumptions the central limit theorem holds for the ergodic average f_t and we have

$$t^{1/2}\left(\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s - \int f\mathrm{d}\bar{\pi}\right) \Rightarrow N(0,\sigma^2_{f,C}), \tag{2.2}$$

where the asymptotic variance $\sigma^2_{f,C}$ is given in terms of the integrated autocorrelation function, see e.g. proposition 4.1.3 in [2],

$$\sigma^2_{f,C} = 2\int_0^\infty \mathbb{E}_{\bar{\pi}}\left[\left(f(X_0)-\bar{f}\right)\left(f(X_t)-\bar{f}\right)\right]\mathrm{d}t.$$

This is a convenient quantity from a practical point of view since there exists easily implementable estimators for $\sigma^2_{f,C}$. On the other hand the asymptotic variance $\sigma^2_{f,C}$ is related to the curvature of the rate function $I_{f,C}(\ell)$ around the mean $\ell = \bar{f}$ (e.g. see [8]): we have

$$\tilde{I}''_{f,C}(\bar{f}) = \frac{1}{2\sigma^2_{f,C}}.$$

From theorem 2.4 it follows immediately that $\sigma^2_{f,C} \leqslant \sigma^2_{f,0}$ but in fact the addition of an appropriate irreversible drift strictly decreases the asymptotic variance.

Theorem 2.7. *Assume that $C \neq 0$ is a vector field as in assumption* (**H**) *and let $f \in \mathcal{C}^{(\alpha)}(E)$ such that for some $\epsilon > 0$ and $\ell \in (\bar{f}-\epsilon,\bar{f}+\epsilon)\setminus\{\bar{f}\}$ we have $\tilde{I}_{f,C}(\ell) > \tilde{I}_{f,0}(\ell)$. Then we have*

$$\sigma^2_{f,C} < \sigma^2_{f,0}.$$

Remark 2.8. An examination of the proof of theorem 2.7 shows that a less restrictive condition is needed for the strict decrease in variance to hold. In particular, it is enough to assume that

$$\operatorname{div}\left(\frac{\partial p_C(x)}{\partial \ell}|_{\ell=\bar{f}}C(x)\right) \neq 0,$$

where $p_C(x) = p_C(x;\ell)$ is the strictly positive invariant density of $\mu^*_C(\mathrm{d}x) = \mu^*_C(\mathrm{d}x;\ell)$ such that $\tilde{I}_{f,C}(\ell) = I_C(\mu^*_C)$.

Let us conclude this section with an example demonstrating that adding irreversibility in the dynamics does not always result in a increase of the spectral gap, even though the variance of the estimator decreases. The key point is that the imaginary part of complex eigenvalues of the generator for irreversible processes creates oscillations in the autocorrelation function which can dramatically reduce the value of its integral. A related discussion regarding comparison of convergence criteria can be also found in [14]. Related computations for the asymptotic behavior of the mean-square displacement of tracers can be found in [24]. The purpose of this example is to demonstrate that spectral gap as a criterium of convergence may not be tight enough to assess improvement in performance when breaking irreversibility. On the other hand, the large deviations rate function and the asymptotic variance both reflect the improved convergence properties due to the irreversible perturbation.

Example 2.9. *Let us consider the family of diffusions*

$$\mathrm{d}X_t = \delta\mathrm{d}t + \mathrm{d}W_t$$

on the circle S^1 with generator

$$\mathcal{L}_\delta = \Delta + \delta\nabla.$$

For any $\delta \in \mathbb{R}$ the Lebesgue measure on S^1 is invariant, but $\mathcal{L}_\delta$ is self-adjoint on $L^2(\mathrm{d}x)$ and thus X_t is reversible if and only if $\delta = 0$. A simple computation (using for example lemma 3.2) shows that for a measure $\mu(\mathrm{d}x)$ that has positive and sufficiently smooth density $p(x)$ we have

$$I(\mu) = \frac{1}{8}\int_{S^1}\left|\frac{p'(x)}{p(x)}\right|^2 p(x)\mathrm{d}x + \delta^2\frac{1}{2}\left[1 - \frac{1}{\int_{S^1}\frac{1}{p(x)}\mathrm{d}x}\right],$$

and in this case $I(\mu)$ strictly increases unless $\mu(\mathrm{d}x) = \mathrm{d}x$. The eigenvalues of $\mathcal{L}_\delta$ are $\lambda_n = -n^2 + in\delta, \quad n \in \mathbb{Z}$ with eigenfunction $\mathrm{e}^{\mathrm{inx}}$ and thus the spectral gap is -1 for any $\delta \in \mathbb{R}$. However for any real-valued function f the asymptotic variance decreases: for f with $\int_{S^1} f\mathrm{d}x = 0$ with Fourier coefficients c_n we have

$$\sigma_f^2(\delta) = \int_0^\infty \langle \mathrm{e}^{t\mathcal{L}} f(x), f(x)\rangle_{L^2(\mathrm{d}x)}\,\mathrm{d}t = \sum_{n\in\mathbb{Z}, n\neq 0}\frac{|c_n|^2}{n^2 + in\delta} = \sum_{n=1}^{\infty}\frac{2|c_n|^2}{n^2+\delta^2}.$$

In this example, even though the spectral gap does not increase at all, the variance not only decreases, but it can be made as small as we want by increasing δ^2. The latter is in agreement with both theorems 2.3 and 2.7 and illustrates how irreversibility improves sampling.

3. The Donsker–Varadhan functional

A standard trick in the theory of large deviations, when computing the probability of an unlikely event, is to perform a change of measure to make the unlikely event typical. In the context of SDE's, this takes of the form of changing the drift of the SDE's itself. This is the idea behind the proof of the following result due to Gärtner, [18].

Theorem 3.1 (Theorem 3.2 in [18]). *Consider the SDE*

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_t$$

on E with $b \in \mathcal{C}^{(1+\alpha)}(E)$ and with generator

$$\mathcal{L} = \Delta + b\nabla.$$

Let $\mu \in \mathcal{P}(E)$, where $\mu(\mathrm{d}x) = p(x)\mathrm{d}x$ is a measure with positive density $p \in \mathcal{C}^{(2+\alpha)}(E)$ for some $\alpha > 0$. The Donsker–Varadhan rate function $I(\mu)$ takes the form

$$I(\mu) = \frac{1}{2}\int_E |\nabla\phi(x)|^2\,\mathrm{d}\mu(x), \tag{3.1}$$

where ϕ is the unique (up to constant) solution of the equation

$$\Delta\phi + \frac{1}{p}(\nabla p, \nabla\phi) = \frac{1}{p}\mathcal{L}^* p, \tag{3.2}$$

and $\mathcal{L}^* = \Delta - \nabla b$ is the formal adjoint of L in $L^2(\mathrm{d}x)$.

In the special case where $b = -\nabla U$ is a gradient, then up to an additive constant $\phi(x) = \frac{1}{2}\log p(x) + U(X)$, and we get

$$I(\mu) = \frac{1}{2}\int_E \left|\frac{1}{2}\frac{\nabla p(x)}{p(x)} + \nabla U(x)\right|^2 \mathrm{d}\mu(x), \tag{3.3}$$

which is the usual explicit formula for the rate function in the reversible case.

It will be useful to rewrite $I(\mu)$ in a different form.

Lemma 3.2. *Under the conditions of theorem 3.1, we have*

$$I(\mu)=\frac{1}{8}\int_E\left|\frac{\nabla p(x)}{p(x)}\right|^2\mathrm{d}\mu(x)+\frac{1}{2}\int_E|\nabla\psi(x)|^2\,\mathrm{d}\mu(x)-\frac{1}{2}\int_E\frac{b\nabla p}{p}\mathrm{d}\mu(x),$$

where ψ is the unique (up to constant) solution of the elliptic equation

$$\operatorname{div}\left[p\left(b+\nabla\psi\right)\right]=0.$$

Proof. Motivated by the solution in gradient case, let us write $\phi(x)=\frac{1}{2}\log p(x)+\psi(x)$. By plugging $\phi(x)=\frac{1}{2}\log p(x)+\psi(x)$ in (3.1), we get

$$\begin{aligned}
I(\mu)&=\frac{1}{2}\int_E\left|\frac{1}{2}\frac{\nabla p(x)}{p(x)}+\nabla\psi(x)\right|^2\mathrm{d}\mu(x)\\
&=\frac{1}{8}\int_E\left|\frac{\nabla p(x)}{p(x)}\right|^2\mathrm{d}\mu(x)+\frac{1}{2}\int_E|\nabla\psi(x)|^2\,\mathrm{d}\mu(x)+\frac{1}{2}\int_E\frac{\nabla\psi\nabla p}{p}\mathrm{d}\mu(x)\\
&=\left[\frac{1}{8}\int_E\left|\frac{\nabla p(x)}{p(x)}\right|^2\mathrm{d}\mu(x)+\frac{1}{2}\int_E|\nabla\psi(x)|^2\,\mathrm{d}\mu(x)-\frac{1}{2}\int_E\frac{b\nabla p}{p}\mathrm{d}\mu(x)\right]\\
&\quad+\frac{1}{2}\int_E\left[\left(b+\nabla\psi\right)\nabla p\right]\mathrm{d}x\\
&=I(\mu,1)+I(\mu,2).
\end{aligned}$$

We claim that $I(\mu,2)=0$. Indeed, using $\phi(x)=\frac{1}{2}\log p(x)+\psi(x)$, the constraint (3.2) gives the following chain of equalities

$$\begin{aligned}
&\Delta\phi+\frac{1}{p}\left(\nabla p,\nabla\phi\right)=\frac{1}{p}\mathcal{L}^*p\Rightarrow\\
&\frac{\Delta p}{2p}-\frac{|\nabla p|^2}{2p^2}+\Delta\psi+\frac{|\nabla p|^2}{2p^2}+\frac{1}{p}\left(\nabla p,\nabla\psi\right)=\frac{\Delta p}{2p}-\frac{1}{p}\operatorname{div}(bp)\Rightarrow\\
&\Delta\psi+\frac{1}{p}\left(\nabla p,\nabla\psi\right)=-\frac{1}{p}\operatorname{div}(bp)\Rightarrow\\
&p\Delta\psi+\left(\nabla p,\nabla\psi\right)+\operatorname{div}(bp)=0\Rightarrow\\
&\nabla\cdot\left[p\left(b+\nabla\psi\right)\right]=0.
\end{aligned}$$

The weak formulation of the latter statement reads as follows

$$\int_E\left(b(x)+\nabla\psi(x)\right)\nabla g(x)p(x)\mathrm{d}x=0,\qquad\forall g\in\mathcal{C}^1(E).$$

Choosing $g=\log p$, we obtain

$$\int_E\left(b(x)+\nabla\psi(x)\right)\nabla p(x)\mathrm{d}x=0,$$

which is precisely the statement $I(\mu,2)=0$. So we have indeed proven the claim. □

With the representation of $I_C(\mu)$ we can now prove theorem 2.2.

Proof of theorem 2.2: since $b(x)=-\nabla U(x)+C(x)$, using lemma 3.2, $I_C(\mu)$ becomes

$$\begin{aligned}
I_C(\mu)=&\frac{1}{8}\int_E\left|\frac{\nabla p(x)}{p(x)}\right|^2\mathrm{d}\mu(x)+\frac{1}{2}\int_E|\nabla\psi_C(x)|^2\,\mathrm{d}\mu(x)\\
&+\frac{1}{2}\int_E\frac{\nabla U(x)\nabla p(x)}{p(x)}\mathrm{d}\mu(x)-\frac{1}{2}\int_E\frac{C(x)\nabla p(x)}{p(x)}\mathrm{d}\mu(x),
\end{aligned}\qquad(3.4)$$

where ψ_C is the unique (up to constant) solution of the equation

$$\operatorname{div}\left[p\left(-\nabla U + C + \nabla\psi_C\right)\right] = 0.$$

The proof of lemma 3.2 shows that $\psi_C(x) = \phi(x) - \frac{1}{2}\log p(x)$ where ϕ is the unique solution (up to constants) of the equation (3.2) with $\mathcal{L} = \mathcal{L}_0 + C\nabla$.

Using the explicit formula 3.3 for the reversible case we obtain for the difference $J_C(\mu) = I_C(\mu) - I_0(\mu)$

$$\begin{aligned} J_C(\mu) = I_C(\mu) - I_0(\mu) &= \frac{1}{2}\int_E \left[|\nabla\psi_C(x)|^2 - |\nabla U(x)|^2\right]\mathrm{d}\mu(x) \\ &\quad -\frac{1}{2}\int_E \frac{C(x)\nabla p(x)}{p(x)}\mathrm{d}\mu(x). \end{aligned}$$

The condition $\operatorname{div}\left(C(x)\mathrm{e}^{-2U(x)}\right) = 0$ can be rewritten as

$$\operatorname{div} C(x) = 2C(x)\nabla U(x).$$

Integration by parts gives for the last term in $J_C(\mu)$

$$\begin{aligned} \int_E \frac{C(x)\nabla p(x)}{p(x)}\mathrm{d}\mu(x) &= \int_E C(x)\nabla p(x)\mathrm{d}x = -\int_E \operatorname{div} C(x)p(x)\mathrm{d}x = -\int_E \operatorname{div} C(x)\mathrm{d}\mu(x) \\ &= -\int_E 2C(x)\nabla U(x)\mathrm{d}\mu(x). \end{aligned}$$

Hence, we obtain

$$J_C(\mu) = \frac{1}{2}\int_E \left[|\nabla\psi_C(x)|^2 - |\nabla U(x)|^2 + 2C(x)\nabla U(x)\right]\mathrm{d}\mu(x).$$

Using the constraint in its weak form

$$\int_E \left[\nabla\psi_C(x) - \nabla U(x) + C(x)\right]\nabla g(x)\mathrm{d}\mu(x) = 0, \quad \text{for every } g \in \mathcal{C}^1(E) \tag{3.5}$$

we can pick freely $g \in \mathcal{C}^1(E)$. If we first choose $g = \psi_C + U$, then, (3.5) gives

$$\int_E \left[|\nabla\psi_C(x)|^2 - |\nabla U(x)|^2\right]\mathrm{d}\mu(x) = -\int_E C(x)\left(\nabla\psi_C(x) + \nabla U(x)\right)\mathrm{d}\mu(x)$$

and thus

$$J_C(\mu) = \frac{1}{2}\int_E C(x)\left(\nabla U(x) - \nabla\psi_C(x)\right)\mathrm{d}\mu(x). \tag{3.6}$$

Choosing $g = \psi_C - U$ and we get from (3.5)

$$\int_E |\nabla\psi_C(x) - \nabla U(x)|^2\,\mathrm{d}\mu(x) = \int_E C(x)\left(\nabla U(x) - \nabla\psi_C(x)\right)\mathrm{d}\mu(x).$$

Plugging this in (3.6) we obtain

$$J_C(\mu) = \frac{1}{2}\int_E |\nabla\psi_C(x) - \nabla U(x)|^2\,\mathrm{d}\mu(x).$$

Clearly $J_C(\mu) \geqslant 0$. If μ possesses a strictly positive density, it is clear that $J_C(\mu) = 0$ if and only if $\operatorname{div}(pC) = 0$. In other words, $J_C(\mu) > 0$ if and only if $\operatorname{div}(pC) \neq 0$.

Finally let us write the positive density as $p(x) = \mathrm{e}^{2G(x)}$, since we have $\operatorname{div}(C\mathrm{e}^{-2U}) = 0$ and $\operatorname{div}(C\mathrm{e}^{2G}) = 0$ we have

$$\operatorname{div} C = -2C\nabla U = 2C\nabla G$$

and thus $C\nabla(G + U) = 0$, i.e. $(G + U)$ is a conserved quantity under the flow $\frac{\mathrm{d}x}{\mathrm{d}t} = C(x)$. $\square$

We now consider the one-parameter family $C(x) = \delta C_0(x)$ and prove theorem 2.3.

Proof of theorem 2.3: for notational convenience let us write $J_\delta(\mu)$ instead of $J_{\delta C_0}(\mu)$ and let us set $\varphi_\delta(x) = \psi_{\delta C_0}(x) - U(x)$. From theorem 2.2 we have

$$J_\delta(\mu) = \frac{1}{2}\int_E |\nabla\varphi_\delta(x)|^2\, \mathrm{d}\mu(x), \tag{3.7}$$

where φ_δ is the unique (up to constant) solution of the equation

$$\int_E (\delta C_0(x) + \nabla\varphi_\delta(x))\, \nabla g(x)\mu(\mathrm{d}x) = 0, \quad \forall g \in \mathcal{C}^1(E). \tag{3.8}$$

Let us define $\xi_\delta(x) = \delta^{-1}\varphi_\delta(x)$. Then,

$$J_\delta(\mu) = \delta^2 \frac{1}{2}\int_E |\nabla\xi_\delta(x)|^2\, \mathrm{d}\mu(x)$$

and because $\delta \neq 0$, ξ_δ is the unique (up to constant) solution of the equation

$$\int_E (C_0(x) + \nabla\xi_\delta(x))\, \nabla g(x)\mu(\mathrm{d}x) = 0, \quad \forall g \in \mathcal{C}^1(E).$$

The last equation makes it clear that, modulo an additive constant, $\xi_\delta(x)$ is in fact independent of δ. Thus, there exists a functional $K(\mu) \geqslant 0$ such that

$$J_\delta(\mu) = \delta^2 K(\mu).$$

Clearly, if $\mu(\mathrm{d}x) = p(x)\mathrm{d}x$ with $\mathrm{div}(pC_0) = 0$ then $K(\mu) = 0$, otherwise $K(\mu) > 0$. □

4. Large deviation for observables and the asymptotic variance

Let us consider a function $f \in \mathcal{C}(E)$ with mean $\bar{f} = \int_E f(x)\mathrm{d}\bar{\pi}(x)$. Let us set

$$f_t = \langle f, \pi_t \rangle = \int_E f(x)\mathrm{d}\pi_t(x) = \frac{1}{t}\int_0^t f(X_s)\mathrm{d}s.$$

By the contraction principle f_t satisfies a large deviation principle with action functional given by

$$\tilde{I}_{f,C}(\ell) = \inf_{\mu\in\mathcal{P}(E)} \{ I_C(\mu) : \langle f, \mu \rangle = \ell \}, \tag{4.1}$$

where $\ell \in \mathbb{R}$ and $I_C(\mu)$ is the Donsker–Varadhan action functional for the empirical measure π_t.

In subsection 4.1 we prove theorem 2.4, whereas in subsection 4.2 we prove theorem 2.7.

4.1. Large deviation for observables

Theorem 2.4 is a fairly immediate consequence of theorem 2.2 and proposition 4.1.

Proposition 4.1. *Let $f \in \mathcal{C}^{(\alpha)}(E)$, and $\ell \in (\min_x f(x), \max_x f(x))$. Then there exists $\mu^*(\mathrm{d}x) = p(x)\mathrm{d}x$ with $p(x) > 0$ and $p(x) \in \mathcal{C}^{(2+\alpha)}(E)$ such that*

$$\tilde{I}_{f,C}(\ell) = I_C(\mu^*).$$

Proof. As discussed in Gärtner [18], the semigroup $T_t h(x) = \mathbb{E}_x\,[h(X_t)]$ is strong-Feller and the strong-Feller property is inherited by the Feynman–Kac semigroup

$$T_t^f h(x) = \mathbb{E}_x \left[\mathrm{e}^{\int_0^t f(X_s)\mathrm{d}s} h(X_t) \right],$$

if $f \in \mathcal{C}(E)$. Moreover the semigroups T_t^f are quasi-compact on the Banach space $\mathcal{C}(E)$ and by a Perron–Frobenius argument the semigroup T_t^f has a dominant simple positive eigenvalue

$e^{\lambda(f)t}$ with a corresponding strictly positive eigenvector $u(f) = e^{\phi(f)}$. We write $\lambda(f)$ and $u(f)$ instead of λ, u in order to emphasize their dependence on the observable f.

For any $f, g \in \mathcal{C}(E)$, $T_t^{f+\gamma g}$ is a bounded perturbation of T_t^f. By analytic perturbation theory (see for example chapter 8 of [22]) and the simplicity of the eigenvalue $\lambda(f)$ this implies that the maps $\gamma \mapsto \lambda(f + \gamma g)$ and $\gamma \mapsto u(f + \gamma g)$ are real-analytic functions. If we require, in addition, that $f \in \mathcal{C}^{(\alpha)}(E)$, then the bounded linear operator $(\mathcal{L}_C + f)$ that maps $\mathcal{C}^{(2+\alpha)}(E)$ to $\mathcal{C}^{(\alpha)}(E)$ is invertible with compact inverse. Hence, the relation

$$(\mathcal{L}_C + f)u(f) = \lambda(f)u(f).$$

implies that $\lambda(f) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left[e^{\int_0^t f(X_s)ds}\right]$ is a simple eigenvalue of the operator $(\mathcal{L}_C + f)$ in $\mathcal{C}^{(\alpha)}(E)$ and that the solution $u(f)$ is in $\mathcal{C}^{(2+\alpha)}(E)$ (see [12]). This implies

$$\nabla\phi(f) = \nabla \log u(f) \in \mathcal{C}^{(1+\alpha)}(E).$$

The rate function $I_C(\mu)$ can be written as

$$I_C(\mu) = \sup_{f\in\mathcal{C}^{(\alpha)}(E)} \{\mu(f) - \lambda(f)\}.$$

If we pick $\mu(dx) = p(x)dx$ with $p(x) > 0$ and $p \in C^{(2+\alpha)}(E)$ then it is shown in [18] that the supremum is attained when f is chosen such that μ is the invariant measure for the SDE with infinitesimal generator

$$\mathcal{L}_C + \nabla\phi(f)\nabla = \mathcal{L}_{C+\nabla\phi}.$$

Turning now to the rate function for observables we note first that if $\ell \in (\min_x f(x), \max_x f(x))$ then $I_{f,C}(\ell)$ is finite. Indeed simply pick any measure μ with a $\mathcal{C}^{(2+\alpha)}(E)$ strictly positive density such that $\int f d\mu = \ell$, then $I_{f,C}(\ell) \leqslant I_C(\mu)$ which is finite by theorem 3.1. Besides the representation 4.1 we can also represent the rate function $\tilde{I}_{f,C}$ as the Legendre transform of the moment generating function of $\bar{f}_t$

$$\tilde{I}_{f,C}(\ell) = \sup_{\beta\in\mathbb{R}} \{\ell \cdot \beta - \lambda(\beta f)\}$$

where

$$\lambda(\beta f) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left[e^{\int_0^t \beta f(X_s)ds}\right].$$

Due to the relation

$$(\mathcal{L}_C + \beta f)u(\beta f) = \lambda(\beta f)u(\beta f), \tag{4.2}$$

$\lambda(\beta f)$ is a simple eigenvalue of $\mathcal{L}_C + \beta f$ in $\mathcal{C}^{(\alpha)}(E)$ and as mentioned before $u(\beta f)$ is in $\mathcal{C}^{(2+\alpha)}(E)$. We can then compute $\tilde{I}_{f,C}(\ell)$ by calculus and the sup is attained if $\hat{\beta}$ is chosen such that $\ell = \frac{\partial}{\partial\beta}\lambda(\hat{\beta} f)$. With $u(\beta f) = e^{\phi(\beta f)}$, the eigenvalue equation (4.2) can be equivalently written as

$$\mathcal{L}_C\phi(\beta f) + \frac{1}{2}|\nabla\phi(\beta f)|^2 = \lambda(\beta f) - \beta f. \tag{4.3}$$

Differentiating 4.3 with respect to β and setting $\psi(\beta f) = \frac{\partial\phi}{\partial\beta}(\beta f)$ we see that $\psi(\beta f)$ satisfies the equation

$$\mathcal{L}_C\psi(\beta f) + (\nabla\phi(\beta f), \nabla\psi(\beta f)) = \frac{d}{d\beta}\lambda(\beta f) - f,$$

or equivalently

$$\mathcal{L}_{C+\nabla\phi(\beta f)}\psi = \frac{d}{d\beta}\lambda(\beta f) - f.$$

Thus, the constraint $\langle f, \mu\rangle = \ell$, implies that in order to have $\ell = \frac{\mathrm{d}}{\mathrm{d}\beta}\lambda(\hat{\beta} f)$ for some $\hat{\beta}$, $\mu_{\hat{\beta}}$ should be the invariant measure for the process with generator $\mathcal{L}_{C+\nabla\phi(\hat{\beta} f)}$. Since $\nabla\phi \in \mathcal{C}^{(1+\alpha)}(E)$ the corresponding invariant measure $\mu_{\hat{\beta}}$ is strictly positive and has a density $p(x) \in \mathcal{C}^{(2+\alpha)}(E)$.

To conclude the proof of the proposition, by [18] we have $I_C(\mu_{\hat{\beta}}) = \mu(\hat{\beta} f) - \lambda(\hat{\beta} f)$. But since $\mu(f) = \ell$ this is also equal to $I_{f,C}(\ell)$. □

Completion of the proof of theorem 2.4: let ℓ be such that $\ell \neq \int f d\bar{\pi}$. By proposition 4.1, there exists measures μ_0^* and μ_C^*, both with strictly positive densities p_0, $p_C \in \mathcal{C}^{(2+\alpha)}(E)$ such that $\tilde{I}_{f,C}(\ell) = I_C(\mu_C^*)$ and $\tilde{I}_{f,0}(\ell) = I_0(\mu_0^*)$.

Let us first assume that $\mathrm{div}(p_C C) \neq 0$. Since $I_C(\mu) > I_0(\mu)$ for any μ with strictly positive densities $p \in C^{(2+\alpha)}$ such that $\mathrm{div}(pC) \neq 0$, this implies that $\tilde{I}_{f,0}(\ell) \leqslant \tilde{I}_{f,C}(\ell)$.

By contradiction let us now assume that

$$\tilde{I}_{f,0}(\ell) = \tilde{I}_{f,C}(\ell).$$

Let us first assume that $\mu_0^* \neq \mu_C^*$. Since $\mathrm{div}(p_C C) \neq 0$, we have

$$I_0(\mu_0^*) = I_C(\mu_C^*) > I_0(\mu_C^*).$$

which contradicts $\tilde{I}_{f,0}(\ell) = I_0(\mu_0^*)$. Now if $\mu_0^* = \mu_C^*$ then we have

$$I_0(\mu_C^*) = I_0(\mu_0^*) = I_C(\mu_C^*).$$

However, this contradicts the fact that we always have $I_C(\mu_C^*) > I_0(\mu_C^*)$ for $\mu_C^*(\mathrm{d}x) = p_C(x)\mathrm{d}x$ such that $\mathrm{div}(p_C C) \neq 0$. This proves that $\tilde{I}_{f,0}(\ell) < \tilde{I}_{f,C}(\ell)$.

If $\mathrm{div}(p_C C) = 0$ then with $p = \mathrm{e}^{2G}$ we must have $C\nabla(G+U) = 0$. As in the proof of proposition 4.1, the density p_C is an invariant measure for the SDE with added drift ϕ_C, i.e.

$$\mathcal{L}^*_{C+\nabla\phi_C} p_C = 0.$$

but since $\mathrm{div}(p_C C) = 0$ we have in fact

$$\mathcal{L}^*_{\nabla\phi_C} p_C = 0.$$

Also $\mathcal{L}_{\nabla\phi}$ is the generator of a reversible ergodic Markov process and thus $p_C = \mathrm{e}^{2(\phi-U)}$ from which we see that

$$\phi = G + U.$$

On the other hand e^{ϕ} is the solution of the eigenvalue equation

$$(\mathcal{L}_C + \widehat{\beta} f)\mathrm{e}^{\phi} = \lambda(f)\mathrm{e}^{\phi}.$$

Since $C\nabla(G+U) = 0$, we have that $C\nabla\mathrm{e}^{\phi} = C\nabla\mathrm{e}^{G+U} = 0$. Thus, the last display reduces to

$$(\mathcal{L}_0 + \widehat{\beta} f)\mathrm{e}^{\phi} = \lambda(f)\mathrm{e}^{\phi}.$$

We also note that changing f into $f + c$ leaves ϕ unchanged but changes $\lambda(f)$ to $\lambda(f) + \widehat{\beta} c$. So, for some constant c, we must have

$$\widehat{\beta} f = \mathrm{e}^{-(G+U)}\mathcal{L}_0\mathrm{e}^{(G+U)} + c = \frac{1}{2}\Delta(G+U) + \frac{1}{2}|\nabla G|^2 - |\nabla G|^2 + c. \qquad \square$$

4.2. Asymptotic variance

In this subsection, we prove that adding irreversibility results in reducing the asymptotic variance of the estimator. The existence of the central limit theorem, see (2.2), of the second derivative $\tilde{I}''_f(\bar{f})$ and of the relation $\sigma_f^2 = \frac{1}{2\tilde{I}''_f(\bar{f})}$ implies that it is enough to prove that for $C \neq 0$ and $f \in \mathcal{C}^{(\alpha)}(E)$

$$\tilde{I}''_{f,C}(\bar{f}) - \tilde{I}''_{f,0}(\bar{f}) > 0.$$

We recall that by (3),

$$J_C(\mu) = I_C(\mu) - I_0(\mu) = \frac{1}{2}\int_E |\nabla\psi_C(x) - \nabla U(x)|^2 \, \mathrm{d}\mu(x).$$

By proposition 4.1, it is enough to consider measures that have a strictly positive density in $\mathcal{C}^{(2+\alpha)}(E)$. We start by computing the first and second order Gâteaux directional derivatives of $J_C(\mu)$ for $\mu(\mathrm{d}x) = p(x)\mathrm{d}x$ with $p(x) \in \mathcal{C}^{(2+\alpha)}(E)$. For notational convenience we shall often write $J_C(p)$ instead of $J_C(\mu)$. Let $\gamma \in \mathbb{R}$ and let us define

$$\tilde{J}_C(\gamma; p, q) = J_C(p + \gamma q), \quad \text{for } p, q \in \mathcal{C}^{(2+\alpha)}(E). \tag{4.4}$$

In subsubsection 4.2.1 we compute first order Gâteaux directional derivative, whereas in subsubsection 4.2.1 we compute second order Gâteaux directional derivative. Then, in subsection 4.3 we put things together proving theorem 2.7.

4.2.1. First order Gâteaux directional derivative Let $p(x), q(x) \in \mathcal{C}^{(2+\alpha)}(E)$ and notice that

$$\begin{aligned}
&\frac{1}{\gamma}\left[J_C(p+\gamma q) - J_C(p)\right] \\
&\quad = \frac{1}{\gamma}\left[\frac{1}{2}\int_E \left|\nabla\psi_C^{p+\gamma q}(x) - \nabla U(x)\right|^2 (p(x) + \gamma q(x))\,\mathrm{d}x \right.\\
&\quad \left. -\frac{1}{2}\int_E \left|\nabla\psi_C^{p}(x) - \nabla U(x)\right|^2 p(x)\mathrm{d}x\right] \\
&\quad = \frac{1}{2\gamma}\left[\int_E \left(\nabla\psi_C^{p+\gamma q}(x) - \nabla\psi_C^{p}(x)\right)\left(\nabla\psi_C^{p+\gamma q}(x) + \nabla\psi_C^{p}(x) - 2\nabla U(x)\right)\right.\\
&\qquad \left. \times p(x)\mathrm{d}x + \gamma\int_E \left|\nabla\psi_C^{p+\gamma q}(x) - \nabla U(x)\right|^2 q(x)\mathrm{d}x\right] \\
&\quad = \frac{1}{2}\left[\int_E \frac{\nabla\psi_C^{p+\gamma q}(x) - \nabla\psi_C^{p}(x)}{\gamma}\left(\nabla\psi_C^{p+\gamma q}(x) + \nabla\psi_C^{p}(x) - 2\nabla U(x)\right)p(x)\mathrm{d}x \right.\\
&\qquad \left. + \int_E \left|\nabla\psi_C^{p+\gamma q}(x) - \nabla U(x)\right|^2 q(x)\mathrm{d}x\right].
\end{aligned}$$

For every $g \in \mathcal{C}^1(E)$, we notice that $\frac{\nabla\psi_C^{p+\gamma q}(x) - \nabla\psi_C^{p}(x)}{\gamma}$ satisfies

$$\begin{aligned}
0 &= \frac{1}{\gamma}\int_E \left[\left(-\nabla U(x) + C(x) + \nabla\psi_C^{p+\gamma q}(x)\right)\nabla g(x)(p(x) + \gamma q(x))\right.\\
&\qquad \left. - \left(-\nabla U(x) + C(x) + \nabla\psi_C^{p}(x)\right)\nabla g(x) p(x)\right]\mathrm{d}x \\
&= \int_E \frac{\nabla\psi_C^{p+\gamma q}(x) - \nabla\psi_C^{p}(x)}{\gamma}\nabla g(x) p(x)\mathrm{d}x \\
&\qquad + \int_E \left(-\nabla U(x) + C(x) + \nabla\psi_C^{p+\gamma q}(x)\right)\nabla g(x) q(x)\mathrm{d}x.
\end{aligned}$$

Since $p, q \in \mathcal{C}^{(2+\alpha)}(E)$, it follows (as in section 3 of [18]) that there is a $\hat{\psi}_C^{p,q} \in \mathcal{C}^{(2+\alpha)}(E)$ such that

$$\psi_C^{p+\gamma q}(x) = \psi_C^p(x) + \gamma \hat{\psi}_C^{p,q}(x) + o_1(\gamma),$$

where $\|o_1(\gamma)\|_{(2+\alpha)} \to 0$ as $\gamma \to 0$. Then $\forall g \in \mathcal{C}^1(E)$, $\nabla\hat{\psi}_C^{p,q}(x)$ satisfies

$$\int_E \left[\nabla\hat{\psi}_C^{p,q}(x)p(x) + \left(-\nabla U(x) + C(x) + \nabla\psi_C^p(x)\right)q(x)\right]\nabla g(x)\mathrm{d}x = 0. \tag{4.5}$$

Let us then denote

$$\mathrm{d}J_C(p;q) = \lim_{\gamma\downarrow 0}\frac{J_C(p+\gamma q) - J_C(p)}{\gamma}.$$

We obtain

$$\mathrm{d}J_C(p;q) = \int_E \nabla\hat{\psi}_C^{p,q}(x)\left(\nabla\psi_C^p(x) - \nabla U(x)\right)p(x)\mathrm{d}x + \frac{1}{2}\int_E \left|\nabla\psi_C^p(x) - \nabla U(x)\right|^2 q(x)\mathrm{d}x.$$

It is clear that if the measure μ is the invariant measure, i.e. $\mu(\mathrm{d}x) = \bar{\pi}(\mathrm{d}x)$, then denoting by $\bar{p}$ the density of $\bar{\pi}(\mathrm{d}x) = \bar{p}(x)\mathrm{d}x$, we have that $\nabla\psi_C^{\bar{p}}(x) = \nabla U(x)$. The latter implies that for any direction q, we get

$$\mathrm{d}J_C(\bar{p};q) = 0,$$

which is of course expected to be true.

4.2.2. Second order Gâteaux directional derivative Next we compute the second order Gâteaux directional derivative. For $p(x), q(x), h(x) \in \mathcal{C}^{(2+\alpha)}(E)$, we get

$$\begin{aligned}
&\frac{1}{\gamma}\left[\mathrm{d}J_C(p+\gamma h;q) - \mathrm{d}J_C(p;q)\right] \\
&= \frac{1}{\gamma}\int_E \nabla\hat{\psi}_C^{p+\gamma h,q}(x)\left(\nabla\psi_C^{p+\gamma h}(x) - \nabla U(x)\right)(p(x)+\gamma h(x))\mathrm{d}x \\
&\quad -\frac{1}{\gamma}\int_E \nabla\hat{\psi}_C^{p,q}(x)\left(\nabla\psi_C^p(x) - \nabla U(x)\right)p(x)\mathrm{d}x \\
&\quad +\frac{1}{2\gamma}\left[\int_E \left|\nabla\psi_C^{p+\gamma h}(x) - \nabla U(x)\right|^2 q(x)\mathrm{d}x - \int_E \left|\nabla\psi_C^p(x) - \nabla U(x)\right|^2 q(x)\mathrm{d}x\right] \\
&= \int_E \nabla\hat{\psi}_C^{p+\gamma h,q}(x)\left(\nabla\psi_C^{p+\gamma h}(x) - \nabla U(x)\right)h(x)\mathrm{d}x \\
&\quad + \int_E \nabla\hat{\psi}_C^{p+\gamma h,q}(x)\frac{\nabla\psi_C^{p+\gamma h}(x) - \nabla\psi_C^p(x)}{\gamma}p(x)\mathrm{d}x \\
&\quad + \int_E \frac{\nabla\hat{\psi}_C^{p+\gamma h,q}(x) - \nabla\hat{\psi}_C^{p,q}(x)}{\gamma}\left(\nabla\psi_C^p(x) - \nabla U(x)\right)p(x)\mathrm{d}x \\
&\quad +\frac{1}{2}\int_E \frac{\nabla\psi_C^{p+\gamma h}(x) - \nabla\psi_C^p(x)}{\gamma}\left(\nabla\psi_C^{p+\gamma h}(x) + \nabla\psi_C^p(x) - 2\nabla U(x)\right)q(x)\mathrm{d}x.
\end{aligned}$$

As it was done for the computation of the first order directional derivative, we next notice that for every $g \in \mathcal{C}^1(E)$, $\frac{\nabla\hat{\psi}_C^{p+\gamma h,q}(x) - \nabla\hat{\psi}_C^{p,q}(x)}{\gamma}$ satisfies

$$\begin{aligned}
0 = &\int_E \left[\frac{\nabla\hat{\psi}_C^{p+\gamma h,q}(x) - \nabla\hat{\psi}_C^{p,q}(x)}{\gamma}p(x) + \nabla\hat{\psi}_C^{p+\gamma h,q}(x)h(x)\right]\nabla g(x)\mathrm{d}x \\
&+ \int_E \frac{\nabla\psi_C^{p+\gamma h}(x) - \nabla\psi_C^p(x)}{\gamma}q(x)\nabla g(x)\mathrm{d}x.
\end{aligned}$$

As in section 3 of [18], it follows then that there is a $\hat{\hat{\psi}}_C^{p,q,h}(x) \in \mathcal{C}^{(1+\alpha)}(E)$ such that

$$\hat{\psi}_C^{p+\gamma h,q}(x) = \hat{\psi}_C^{p,q}(x) + \gamma \hat{\hat{\psi}}_C^{p,q,h}(x) + o_2(\gamma),$$

where $\|o_2(\gamma)\|_{(1+\alpha)} \to 0$ as $\gamma \to 0$. Then, for every $g \in \mathcal{C}^1(E)$, $\nabla\hat{\hat{\psi}}^{p,q,h}(x)$ satisfies

$$\int_E \left[\nabla\hat{\hat{\psi}}_C^{p,q,h}(x)p(x) + \nabla\hat{\psi}_C^{p,q}(x)h(x) + \nabla\hat{\psi}_C^{p,h}(x)q(x)\right]\nabla g(x)\mathrm{d}x = 0. \tag{4.6}$$

Let us then denote

$$\mathrm{d}^2 J_C(p; q, h) = \lim_{\gamma\downarrow 0} \frac{dJ_C(p+\gamma h; q) - dJ_C(p; q)}{\gamma}.$$

We get

$$\begin{aligned}
\mathrm{d}^2 J_C(p; q, h) &= \int_E \left(\nabla\psi_C^{p,q}(x) - \nabla U(x)\right)\nabla\hat{\psi}_C^{p,q}(x)h(x)\mathrm{d}x \\
&\quad + \int_E \nabla\hat{\psi}_C^{p,q}(x)\nabla\hat{\psi}_C^{p,h}(x)p(x)\mathrm{d}x \\
&\quad + \int_E \nabla\hat{\hat{\psi}}_C^{p,h,q}(x)\left(\nabla\psi_C^{p}(x) - \nabla U(x)\right)p(x)\mathrm{d}x \\
&\quad + \int_E \nabla\hat{\psi}_C^{p,h}(x)\left(\nabla\psi_C^{p}(x) - \nabla U(x)\right)q(x)\mathrm{d}x \\
&= \int_E \left(\nabla\psi_C^{p}(x) - \nabla U(x)\right) \\
&\quad \times \left(\nabla\hat{\psi}^{p,q}(x)h(x) + \nabla\hat{\psi}^{p,h}(x)q(x) + \nabla\hat{\hat{\psi}}^{p,q,h}(x)p(x)\right)\mathrm{d}x \\
&\quad + \int_E \nabla\hat{\psi}^{p,q}(x)\nabla\hat{\psi}^{p,h}(x)p(x)\mathrm{d}x.
\end{aligned} \tag{4.7}$$

Using the constraint (4.6) with the test function $g(x) = \psi_C^p(x) - U(x)$, we then obtain

$$\mathrm{d}^2 J_C(p; q, h) = \int_E \nabla\hat{\psi}^{p,q}(x)\nabla\hat{\psi}^{p,h}(x)p(x)\mathrm{d}x.$$

Recall that for every $g \in \mathcal{C}^1(E)$, $\nabla\hat{\psi}^{p,q}(x)$ satisfies (4.5) and similarly for $\nabla\hat{\psi}^{p,h}(x)$. Thus, selecting $h(x) = q(x)$, we get

$$\mathrm{d}^2 J_C(p; q, q) = \int_E \left|\nabla\hat{\psi}_C^{p,q}(x)\right|^2 p(x)\mathrm{d}x. \tag{4.8}$$

Relation (4.8) implies that pointwise in p and for non-zero directions $q(x)$ the second order directional derivative of $I_C(p)$ increases when adding an appropriately non-zero irreversible drift C, i.e.

$$\mathrm{d}^2 J_C(p; q, q) \geqslant 0.$$

Of course, this is expected to be true due to convexity. Let us next investigate what happens at the law of large numbers limit $\mu = \bar{\pi}$. So, let us choose $\mu(\mathrm{d}x)$ to be the invariant measure $\bar{\pi}(\mathrm{d}x)$ and let us denote its density by $\bar{p}(x)$. Then, we notice that in this case $\nabla\psi_C^{\bar{p}}(x) = \nabla U(x)$. So, (4.8) becomes

$$\mathrm{d}^2 J_C(\bar{p}; q, q) = \int_E \left|\nabla\hat{\psi}_C^{\bar{p},q}(x)\right|^2 \bar{\pi}(\mathrm{d}x) \geqslant 0 \tag{4.9}$$

where, $\forall g \in \mathcal{C}^1(E)$, $\nabla\hat{\psi}^{\bar{p},q}(x)$ satisfies

$$\int_E \left[\nabla\hat{\psi}_C^{\bar{p},q}(x)\bar{p}(x) + C(x)q(x)\right]\nabla g(x)\mathrm{d}x = 0. \tag{4.10}$$

In fact, we get for q, C such that $\mathrm{div}(qC) \neq 0$ that $\nabla\hat{\psi}_C^{p,q}(x) \neq 0$. Then, by (4.10) and (4.9) we have

$$\mathrm{d}^2 J_C(\bar{p}; q, q) > 0. \tag{4.11}$$

In addition, (4.10) shows that if $C = 0$, or if q, C are such that $\mathrm{div}(qC) = 0$, then

$$\mathrm{d}^2 J_0(\bar{p}; q, q) = 0.$$

4.3. Completion of the proof of theorem 2.7

Let $C \neq 0$ and $f \in \mathcal{C}^{(\alpha)}(E)$ be such that $\tilde{I}_{f,C}(\ell) > \tilde{I}_{f,0}(\ell)$ for every $\ell \neq \bar{f}$. Then, we want to prove

$$\tilde{I}''_{f,C}(\bar{f}) - \tilde{I}''_{f,0}(\bar{f}) > 0.$$

We know by proposition 4.1 that there exist measures, say $\mu_C(\mathrm{d}x;\ell)$ and $\mu_0(\mathrm{d}x;\ell)$, that have a strictly positive densities in $\mathcal{C}^{(2+\alpha)}(E)$ such that

$$\tilde{I}_{f,C}(\ell) = I_C(\mu_C(\cdot;\ell)), \quad \text{and} \quad \tilde{I}_{f,0}(\ell) = I_0(\mu_0(\cdot;\ell))$$

By convexity and the definitions of $\mu_C(\mathrm{d}x;\ell) = p_{C,\ell}(x)\mathrm{d}x$ and $\mu_0(\mathrm{d}x;\ell) = p_{0,\ell}(x)\mathrm{d}x$, we have that for all $\ell \in (\min_x f(x), \max_x f(x))$

$$\begin{aligned}
\tilde{I}''_{f,C}(\ell) - \tilde{I}''_{f,0}(\ell) &= \frac{\partial^2}{\partial\ell^2}\left[I_C(\mu_C(\cdot;\ell)) - I_0(\mu_0(\cdot;\ell))\right] \\
&= \frac{\partial^2}{\partial\ell^2}\left[(I_C(\mu_C(\cdot;\ell)) - I_0(\mu_C(\cdot;\ell))) + (I_0(\mu_C(\cdot;\ell)) - I_0(\mu_0(\cdot;\ell)))\right] \\
&\geqslant \frac{\partial^2}{\partial\ell^2}\left[I_C(\mu_C(\cdot;\ell)) - I_0(\mu_C(\cdot;\ell))\right] \\
&= \frac{\partial^2}{\partial\ell^2} J_C(\mu_C(\cdot;\ell)).
\end{aligned}$$

Then, (4.9) implies that when evaluated at the law of large numbers $\ell = \bar{f}$,

$$\frac{\partial^2}{\partial\ell^2} J_C(p_{C,\ell})\big|_{\ell=\bar{f}} = \int_E \left|\nabla\hat{\psi}_C^{\bar{p},\bar{q}}(x)\right|^2 \bar{\pi}(\mathrm{d}x), \tag{4.12}$$

such that (4.10) holds with $q(x) = \bar{q}(x) = \frac{\partial}{\partial\ell} p_C(x;\ell)|_{\ell=\bar{f}}$, i.e. $\nabla\hat{\psi}^{\bar{p},q}(x)$ satisfies

$$\int_E \left[\nabla\hat{\psi}_C^{\bar{p},\bar{q}}(x)\bar{p}(x) + C(x)\bar{q}(x)\right]\nabla g(x)\mathrm{d}x = 0, \quad \forall g \in \mathcal{C}^1(E). \tag{4.13}$$

Then, (4.11) implies

$$\frac{\partial^2}{\partial\ell^2} J_C(p_{C,\ell})|_{\ell=\bar{f}} > 0,$$

as long as $\mathrm{div}\,(\bar{q}C) \neq 0$. This concludes the proof of theorem 2.7. □

Table 1. Estimated variance values for different pairs (δ, t).

$\delta \mid t$	25	100	160	220	295
0	0.22	0.08	0.038	0.029	0.011
10	0.19	0.01	0.007	0.005	0.002
100	0.09	0.001	$3e-04$	$2.8e-04$	$1.3e-04$

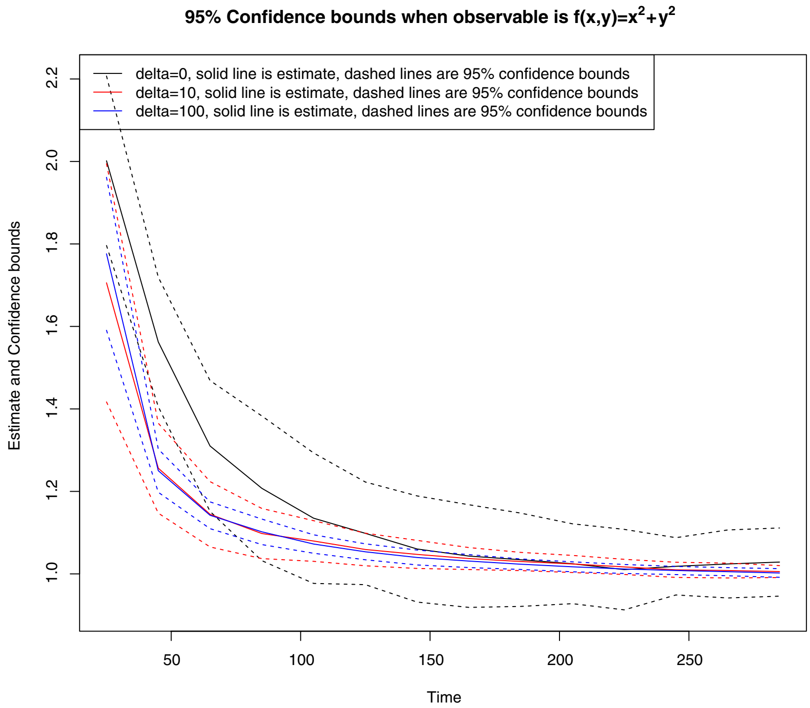

Figure 1. Convergence of $\hat{\bar{f}}_1(t)$ to $\bar{f}_1$.

5. Simulations

In this section we present some numerical results to illustrate the theoretical findings. We study numerically the effect that adding irreversibility has on the speed of convergence to the equilibrium. Consider the SDE in 2 dimensions

$$\mathrm{d}Z_t = [-\nabla U(Z_t) + C(Z_t)]\,\mathrm{d}t + \sqrt{2D}\mathrm{d}W_t, \quad Z_0 = 0$$

where $D = 0.1$ and, for $z = (x, y)$, $C(x, y) = \delta C_0(x, y)$ with $C_0(x, y) = J\nabla U(x, y)$. Here, $\delta \in \mathbb{R}$, I is the 2×2 identity matrix and J is the standard 2×2 antisymmetric matrix, i.e. $J_{12} = 1$, $J_{21} = -1$ and $J_{11} = J_{22} = 0$.

Clearly, in the case $\delta = 0$ we have reversible dynamics, whereas for $\delta \neq 0$ the dynamics is irreversible. Notice that for any $\delta \in \mathbb{R}$, the invariant measure is

$$\bar{\pi}(\mathrm{d}x\mathrm{d}y) = \frac{\mathrm{e}^{-\frac{U(x,y)}{D}}}{\int_{\mathbb{R}^2} \mathrm{e}^{-\frac{U(x,y)}{D}}\mathrm{d}x\mathrm{d}y}\mathrm{d}x\mathrm{d}y$$

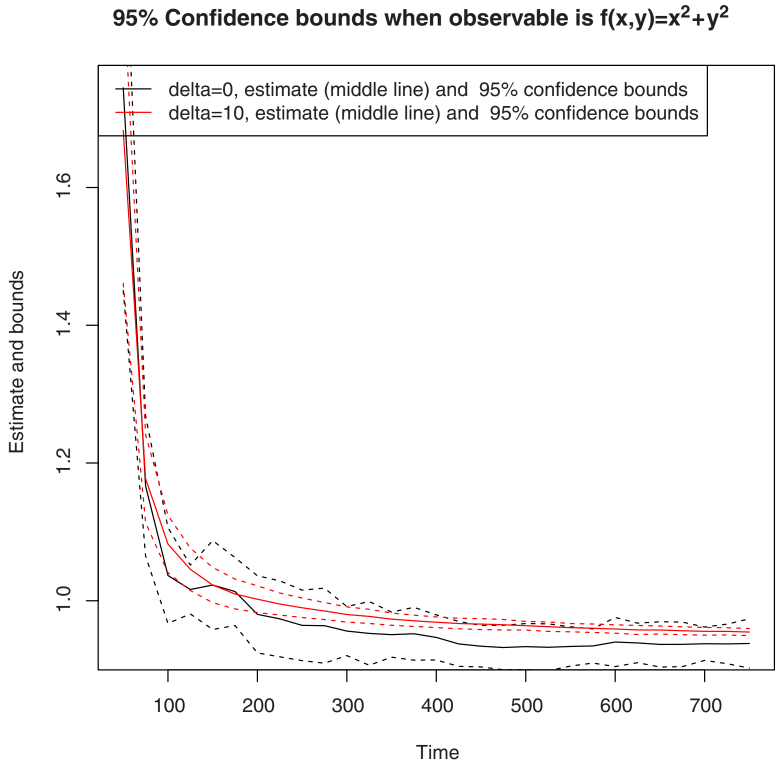

Figure 2. Estimate and 95% Confidence bounds when $U(x, y) = (x^2 - 1)^2 + \frac{1}{2}(3y + x^2 - 1)^2$ and $f(x, y) = x^2 + y^2$.

Table 2. Estimated variance values for different pairs (δ, t).

$\delta \mid t$	100	200	300	400	500	600	700
0	0.01	0.006	0.002	0.002	0.002	0.003	0.002
10	0.003	0.0007	0.0002	0.0001	$7e - 05$	$6e - 05$	$6e - 05$

Let us suppose that we are given an observable $f(x, y)$ and we want to compute

$$\bar{f} = \int_{\mathbb{R}^2} f(x, y)\, \bar{\pi}(\mathrm{d}x\mathrm{d}y).$$

It is known that an estimator for $\bar{f}$ is given by

$$\hat{\bar{f}}(t) = \frac{1}{t - v} \int_v^t f(X_s, Y_s)\, \mathrm{d}s$$

where v is some burn-in period that is used with the hope that the bias has been significantly reduced by time v. This estimate is based on simulating a very long trajectory $Z_s = (X_s, Y_s)$.

For illustration purposes, we present in table 1, variance estimates for different values of δ and time horizons t in the set-up of figure 1. It is noteworthy that the variance reduction for this particular example is about two orders of magnitude.

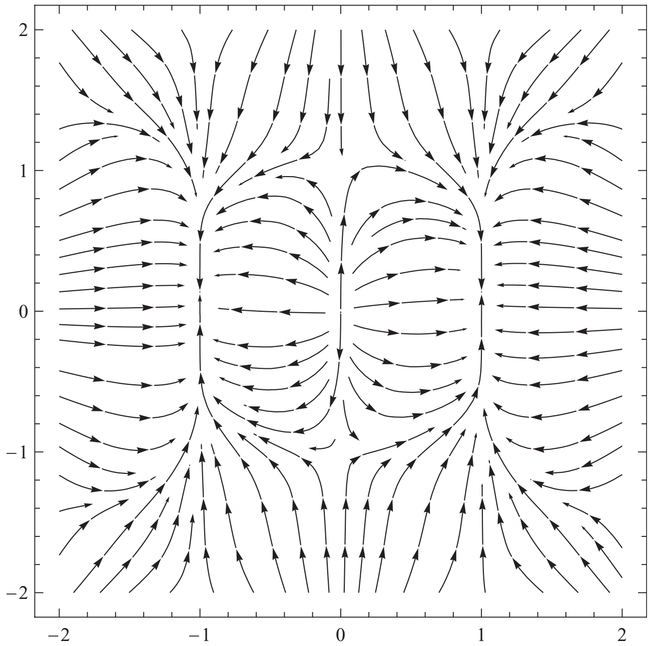

Figure 3. Phase portrait of $U(x, y) = \frac{1}{4}\left[(x^2-1)^2((y^2-2)^2+1)+2y^2-y/8\right] + e^{-8x^2-4y^2}$.

In the second example we pick again a bimodal potential $U(x, y) = (x^2-1)^2 + \frac{1}{2}(3y + x^2-1)^2$ and the observable $f(x, y) = x^2 + y^2$. In figure 2 we see 95% confidence bounds for $\hat{\bar{f}}(t)$. In table 2, we present numerical data for the variance estimates that are illustrated in figure 2. Again, we see variance reduction and it is at the order of about two magnitudes.

In the third example we pick the potential

$$U(x, y) = \frac{1}{4}\left[(x^2-1)^2((y^2-2)^2+1)+2y^2-y/8\right] + e^{-8x^2-4y^2}.$$

Due to the somewhat complex form of $U(x, y)$, we have also plotted in figure 3 its phase portrait. We see that it has two local minima at $(\pm 1.00051, 0.125314)$, two saddle points at $(0, -1.00711)$ and at $(0, 1.08849)$ and a local maximum at $(0, -0.0139)$.

We consider again the observable $f(x, y) = x^2 + y^2$. In figure 4 we see 95% confidence bounds for $\hat{\bar{f}}(t)$. In table 3, we present numerical data for the variance estimates that are illustrated in figure 4. Again, we see variance reduction and it is at the order of about one magnitude when the irreversible parameter is $\delta = 10$.

We conclude this section with a remark on the optimal choice of irreversibility. Theorem 2.3 suggests that in the generic situation, perturbations of the form $C(\cdot) = \delta C_0(\cdot)$ yield better results as the parameter δ increases. However, in practice the higher the δ is, the smaller the discretization step in the simulation algorithm should be, i.e. there is a trade-off to consider here. Thus it makes sense to look for the optimal perturbation $C(x)$ and this could be formulated as a solution to a variational problem that involves minimizing the asymptotic variance of the estimator. Since, the asymptotic variance is inversely proportional to the second derivative of the rate function of the observable evaluated at $\bar{f}$, the variational problem to consider is basically maximization over vector fields C that satisfy condition (**H**)

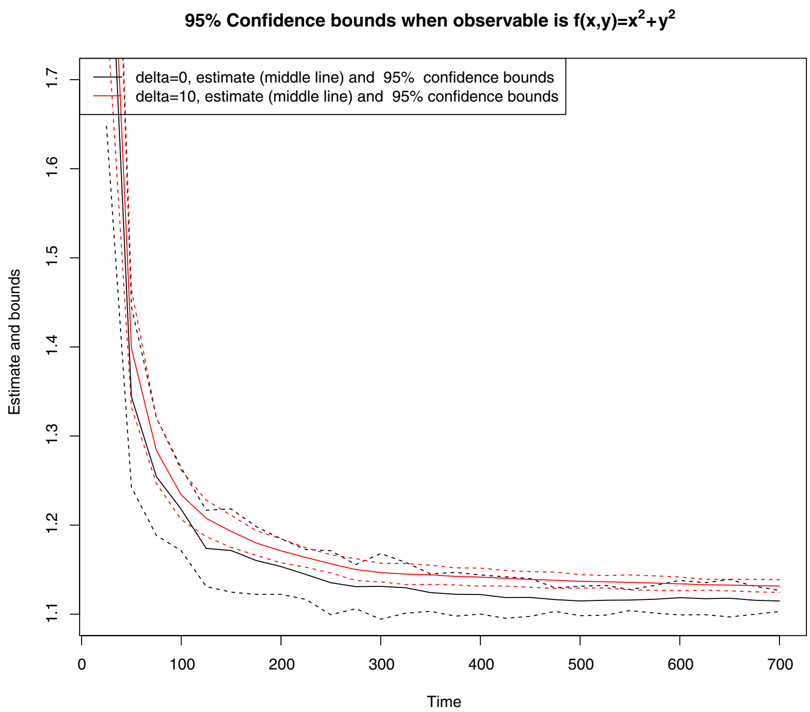

Figure 4. Estimate and 95% Confidence bounds when $U(x, y) = \frac{1}{4}\left[(x^2 - 1)^2((y^2 - 2)^2 + 1) + 2y^2 - y/8\right] + e^{-8x^2-4y^2}$ and $f(x, y) = x^2 + y^2$.

Table 3. Estimated variance values for different pairs (δ, t).

$\delta \mid t$	100	200	300	400	500	600	700
0	0.004	0.002	0.002	0.001	0.001	0.001	0.001
10	0.001	0.0003	0.0002	0.0001	0.0001	0.0001	0.0001

of the quantity (4.12) under the constraint (4.13). We plan to investigate this question in a future work.

6. Conclusions

In this article we have considered the problem of estimating the expected value of a functional of interest using as estimator the long time average of a process that has as its invariant distribution the target measure. We have argued using large deviations theory, both theoretically and numerically, that adding an appropriate drift to the dynamics of a reversible Langevin equation, results in smaller asymptotic variance for the time average estimator. We characterize when observables do not see their variance reduced in terms of a precise non-linear Poisson equation.

Acknowledgments

KS was partially supported by the National Science Foundation (DMS 1312124). LRB was partially supported by the National Science Foundation (DMS 1109316) and the Department of Energy-ASCR (ER 26161) and he also thanks N Pillai for bringing his attention to the issue of non-reversible Monte Carlo methods.

References

[1] Amit Y and Grenander U 1991 Comparing sweeping strategies for stochastic relaxation *J. Multivariate Anal.* **37** 197–222
[2] Asmussen S and Glynn P W 2007 *Stochastic Simulation* (Berlin: Springer)
[3] Athreya K A, Doss H and Sethuraman J 1996 On the convergence of the Markov chain simulation method *Ann. Stat.* **24** 69–100
[4] Barbarosie C 2011 Representation of divergence-free vector fields *Q. Appl. Math.* **69** 309–16
[5] Bedard M and Rosenthal J S 2008 Optimal scaling of metropolis algorithms: heading towards general target distributions *Can. J. Stat.* **36** 483–503
[6] Bierkens J 2014 Non-reversible Metropolis-Hastings. Preprint (arxiv.org/abs/1401.8087)
[7] Constantin P, Kiselev A, Ryshik L and Zlatos A 2008 Diffusion and mixing in fluid flow *Ann. Math.* **168** 643–74
[8] Den Hollander F 2000 *Large Deviations* (Providence, RI: American Mathematical Society)
[9] Diaconis P, Holmes S and Neal R 2010 Analysis of a nonreversible Markov chain sampler *Ann. Appl. Probab.* **10** 726–52
[10] Diaconis P and Miclo L 2013 On the spectral analysis of second-order Markov chains *Ann. Fac. Sci. Toulouse Math.* **3** 573–621
[11] Donsker M D and Varadhan S R S 1975 Asymptotic evaluation of certain Markov process expectations for large times, I *Commun. Pure Appl. Math.* **28** 1–47
Donsker M D and Varadhan S R S 1975 Asymptotic evaluation of certain Markov process expectations for large times, II *Commun. Pure Appl. Math.* **28** 279–301
Donsker M D and Varadhan S R S 1976 Asymptotic evaluation of certain Markov process expectations for large times, III *Commun. Pure Appl. Math.* **29** 389–461
[12] Douglis A and Nirenberg L 1975 Interior estimates for elliptic systems of partial differential equations *Commun. Pure Appl. Math.* **8** 503–38
[13] Plattner N, Doll J D, Dupuis P, Wang H, Liu Y and Gubernatis J E 2011 An infinite swapping approach to the rare-event sampling problem *J. Chem. Phys.* **135** 134111
[14] Dupuis P, Liu Y, Plattner N and Doll J D 2012 On the infinite swapping limit for parallel tempering *SIAM Multiscale Modeling Simul.* **10** 986–1022
[15] Franke B, Hwang C-R, Pai H-M and Sheu S-J 2010 The behavior of the spectral gap under growing drift *Trans. Am. Math. Soc.* **362** 1325–50
[16] Frigessi A, Hwang C R, Sheu S J and Di Stefano P 1993 Convergence rates of the Gibbs sampler, the Metropolis algorithm, and their single-site updating dynamics *J. R. Stat. Soc.* B **55** 205–19
[17] Frigessi A, Hwang C R and Younes L 1992 Optimal spectral structures of reversible stochastic matrices, Monte Carlo methods and the simulation of Markov random fields *Ann. Appl. Probab.* **2** 610–28
[18] Gärtner J 1977 On large deviations from the invariant measure *Theory Probab. Appl.* **22** 24–39
[19] Gilks W R and Roberts G O 1996 Strategies for improving MCMC *Monte Carlo Markov Chain in Practice* (London: Chapman and Hall) pp 89–114
[20] Hwang C R, Hwang S Y and Sheu S J 1993 Accelerating gaussian diffusions *Ann. Appl. Probab.* **3** 897–913
[21] Hwang C R, Hwang S Y and Sheu S J 2005 Accelerating diffusions *Ann. Appl. Probab.* **15** 1433–44
[22] Kato T 1969 *Perturbation Theory of Linear Operators* (New York: Gordon and Breach)
[23] Lelievre T, Nier F and Pavliotis G A 2013 Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion *J. Stat. Phys.* **152** 237–74
[24] Majda A J and Kramer P R 1999 Simplified models for turbulent diffusion: theory, numerical modelling and physical phenomena *Phys. Rep.* **314** 237–574
[25] Meyn S and Tweedie R L 2009 *Markov Chains and Stochastic Stability* 2nd edn (Cambridge: Cambridge University Press)
[26] Mira A 2001 Ordering and improving the performance of Monte Carlo Markov chains *Stat. Sci.* **16** 340–50

[27] Mira A and Geyer C J 2000 On non-reversible Markov chains *Monte Carlo methods* (*Fields Institute Communications* vol 26) (Providence, RI: American Mathematical Society) pp 95–110

[28] Neal R M 2004 Improving asymptotic variance of MCMC estimators: non-reversible chains are better *Technical Report* No. 0406, Department of Statistics, University of Toronto

[29] Mergessen K L and Tweedie R L 1996 Rates of convergence of the Hastings and Metropolis algorithms *Ann. Stat.* **24** 101–21

[30] Peskun P H 1973 Optimum Monte-Carlo sampling using Markov Chains *Biometrika* **60** 607–12

[31] Roberts G O and Rosenthal J S 2004 General state space Markov Chain and MCMC algorithms *Probab. Surv.* **1** 20–71

[32] Rockafellar R T 1970 *Convex Analysis* (*Princeton Mathematical Series*) (Princeton, NJ: Princeton University Press)

[33] Tierney L 1998 A Note on Metropolis-Hastings kernels for general state spaces *Ann. Appl. Probab.* **8** 1–9