Zuse Institute Berlin

Takustr. 7
14195 Berlin
Germany

CARSTEN HARTMANN[1], LORENZ RICHTER[1],
CHRISTOF SCHÜTTE[2,3], WEI ZHANG[3]

[1] *Brandenburgische Technische Universität Cottbus-Senftenberg, D-03046 Cottbus, Germany*

[2] *Department of Mathematics and Computer Science, Freie Universität Berlin, D-14195 Berlin, Germany*

[3] *Zuse Institute Berlin, D-14195 Berlin, Germany*

VARIATIONAL CHARACTERIZATION OF FREE ENERGY: THEORY AND ALGORITHMS

ZIB Report 17-52 (September 2017)

Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany

Telephone: +49 30-84185-0
Telefax: +49 30-84185-125

E-mail: bibliothek@zib.de
URL: http://www.zib.de

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Article

VARIATIONAL CHARACTERIZATION OF FREE ENERGY: THEORY AND ALGORITHMS

Carsten Hartmann [1]*, Lorenz Richter [1], Christof Schütte [2,3] and Wei Zhang [3]

[1] Brandenburgische Technische Universität Cottbus-Senftenberg, D-03046 Cottbus, Germany
[2] Freie Universität Berlin, D-14195 Berlin, Germany
[3] Zuse Institute Berlin, D-14195 Berlin, Germany
* Correspondence: schuette@zib.de

Abstract: The article surveys and extends variational formulations of the thermodynamic free energy and discusses their information-theoretic content from the perspective of mathematical statistics. We revisit the well-known Jarzynski equality for nonequilibrium free energy sampling within the framework of importance sampling and Girsanov change-of-measure transformations. The implications of the different variational formulations for designing efficient stochastic optimization and nonequilibrium simulation algorithms for computing free energies are discussed and illustrated.

Keywords: Importance sampling, Donsker-Varadhan principle, thermodynamic free energy, nonequilibrium molecular dynamics, stochastic approximation, cross-entropy method

MSC: 37A60, 60J60, 82C05

---

1. Introduction

It is one of the standard problems in statistical physics and its computational applications, e.g. in molecular dynamics, that one desires to compute expected values of an observable f with respect to a given (equilibrium) probability density π,

$$\mathbf{E}_\pi[f] = \int f(x)\pi(x)dx$$

Even if samples from the density π are available, the simplest Monte Carlo estimator, the mean value, may suffer from a large variance (compared to the quantity that one tries to estimate), such that the accurate estimation of $\mathbf{E}_\pi[f]$ requires an unreasonably large sample size. Various approaches to circumvent this problem and to reduce the variance of an estimator are available, one of the most prominent representatives being importance sampling where samples are drawn from another probability density ρ and reweighted with the likelihood ratio π/ρ [20,37]. It is well-known that theoretically—and under certain assumptions—there exists an *optimal* importance sampling density ρ^* such that the resulting estimator has *variance zero*. By a clever choice of the importance sampling proposal density, it is thus possible to completely remove the stochasticity from the problem and to obtain what is sometimes called *certainty equivalence*. Yet, drawing samples from the optimal density (or an approximation of it) is a difficult problem in itself, so that the striking variance-reduction due to importance sampling often does not pay off in practice.

The zero variance property of importance sampling and the challenge to utilize it algorithmically is the starting point of this article where the focus is on its generalization to path sampling problems and its algorithmic realization. Regarding the former, we will show that the Donsker-Varadhan variational

principle, a well-known measure-theoretic characterization of the cumulant generating functions [7] that gives rise to a variational characterization of the thermodynamic free energy [6,10] permits several stunning utilizations of the importance sampling framework for path sampling problems; examples involve trajectory-dependent expectations like expected hitting times or free energy differences [4,21]. We will see that finding the optimal change of measure in path space is equivalent to solving an optimal control problem for the underlying dynamical system in which the dynamics is controlled by external driving forces and thus driven out of equilibrium [18,22]. By this connection, optimized importance sampling is shown to share some features with the famous non-equilibrium fluctuation relations for the thermodynamic free energy by Jarzynski and Crooks [5,27]. We will elaborate on how to explore this connection to devise better non-equilibrium free energy algorithms, and hopefully obtain a better understanding of Jarzynski-based estimators; cf. [35].

Regarding the algorithmic realization, the theoretical insight into the relation between (adaptive) importance sampling and optimal control leads to novel algorithms that aim at utilizing the zero variance property without having to sample from the optimal importance sampling density. We will demonstrate how this can be achieved by discretizing the optimal control problem using ideas from stochastic approximation and stochastic optimization [22,44]; see [14,15,42] for an alternative approach using ideas from the theory of large deviations.

Outline

The article is organized as follows: Firstly, in Section 2 we review certainty equivalence and the zero variance property of optimized importance sampling in state space, starting from the Donsker-Varadhan principle and its relation to importance sampling, and comment on some algorithmic issues. Then, in Section 3, we consider the generalization to path space, discuss the relation to stochastic optimal control and revisit Jazynski-based estimators for thermodynamic free energies. Section 4 surveys and discusses novel algorithms that are exploiting the theoretical properties of the control-based importance sampling scheme. We briefly discuss some of these algorithms with simple toy examples in Section 5, before the article concludes in Section 6 with a brief summary and a discussion of open issues. The article contains three appendices that record various technical identities, a brief derivation of Girsanov's change of measure formula, and the proof of the main theorem.

2. Certainty equivalence

In mathematical finance, the guaranteed payment that an investor would accept instead of a potentially higher, but uncertain return on an asset is called a *certainty equivalent*. In physics, certainty equivalence amounts to finding a deterministic surrogate system that reproduces averages of certain fluctuating thermodynamic quantities with probability one. One such example is the thermodynamic free energy difference between two equilibrium states that can be either computed by an exponential average over the fluctuating nonequilibrium work done on the system or by measuring the work of an adiabatic transformation between these states.

2.1. Donsker-Varadhan variational principle

Before getting into the technical details, we briefly review the classical Donsker-Varadhan variational principle for the cumulant generating function of a random variable. To this end, let X be a real-valued, n-dimensional random variable with smooth probability density π and call

$$\mathbf{E}_\pi[f(X)] = \int_{\mathbb{R}^n} f(x)\pi(x)dx \tag{1}$$

the expectation with respect to π for any integrable function $f\colon \mathbb{R}^n \to \mathbb{R}$.

Definition 1. *Let $W\colon \mathbb{R}^n \to \mathbb{R}$ be a bounded random variable. The quantity*

$$\gamma\colon \mathscr{B}(\mathbb{R}^n) \to \mathbb{R}\,, \quad W \mapsto -\log \mathbf{E}_\pi[\exp(-W)] \tag{2}$$

is called the free energy of the random variable $W = W(X)$ with respect to π, where $\mathscr{B}(\mathbb{R}^n)$ is the set of bounded and measurable, real-valued functions on $\mathbb{R}^n$.[1]

Definition 2. *Let ρ be another probability density on $\mathbb{R}^n$. Then*

$$D(\rho|\pi) = \int_{\mathbb{R}^n} \log \frac{\rho(x)}{\pi(x)}\, \rho(x) dx \tag{3}$$

is called the relative entropy of ρ with respect to π (or: Kullback-Leibler divergence), provided that $\pi(x) = 0$ implies that $\rho(x) = 0$ for every $x \in \mathbb{R}^n$. Otherwise we set $D(\rho|\pi) = +\infty$.

The requirement that $\pi(x)$ must not be zero without $\rho(x)$ being zero is known as absolute continuity and guarantees that the likelihood ratio $L = \rho/\pi$ is well defined. In what follows, we may assume without loss of generality that $\pi > 0$. (Otherwise we may exclude those states $x \in \mathbb{R}^n$ for which $\pi(x) = 0$.)

A well-known thermodynamic principle states that the free energy is the Legendre transform of the entropy. The following variant of this principle is due to Donsker and Varadhan and says that (e.g., see [7] and the references therein)

$$-\log \mathbf{E}_\pi[\exp(-W)] = \min_{\rho \geq 0} \left\{ \mathbf{E}_\rho[W] + D(\rho|\pi) \right\}\,, \tag{4}$$

where the minimum is over all probability density functions ρ on $\mathbb{R}^n$. The last equality easily follows from Jensen's inequality by noting that

$$\begin{aligned}
-\log \int_{\mathbb{R}^n} \exp(-W)\pi\, dx &= -\log \int_{\mathbb{R}^n} \exp(-W) \frac{\pi}{\rho}\, \rho\, dx \\
&= -\log \int_{\mathbb{R}^n} \exp\Big(-W - \log \frac{\rho}{\pi}\Big)\, \rho\, dx \\
&\leq \int_{\mathbb{R}^n} \Big(W + \log \frac{\rho}{\pi}\Big)\, \rho\, dx \\
&= \int_{\mathbb{R}^n} W\rho\, dx + \int_{\mathbb{R}^n} \log \frac{\rho}{\pi}\, \rho\, dx\,.
\end{aligned}$$

And it can be readily seen that equality is attained if and only if

$$\rho^* = \exp(\gamma - W)\pi\,, \tag{5}$$

which defines a probability measure with γ given in (2) .

Importance sampling

The relevance of (4)–(5) lies in the fact that, by sampling X from the probability distribution with density ρ^*, one removes the stochasticity from the problem, since the random variable

$$U^* = W(X) + \log \frac{\rho^*(X)}{\pi(X)}$$

[1] "If you can write it down, it's measurable!" (S.R.S. Varadhan).

is almost surely (a.s.) constant. As a consequence, the Monte Carlo scheme for computing the free energy on the left hand side of (4) based on the empirical mean of the independent draws of $U^* = U^*(X)$ with $X \sim \rho^*$, will have a zero variance. This *zero-variance property* is a consequence of Jensen's inequality and the strict concavity of the logarithmic function which implies that equality is attained if and only if the random variable inside the expectation is almost surely constant. The next statement makes this precise.

Theorem 1 (Optimal importance sampling)**.** *Let ρ^* be the probability density given in (5). Then the random variable $Z = \exp(-W)\pi/\rho^*$ has zero variance under ρ^* and we have*

$$Z = \mathbf{E}_\pi[\exp(-W)]\,, \quad \rho^*\text{-a.s.}$$

Proof. We need to show that $\mathrm{Var}_{\rho^*}(Z) = \mathbf{E}_{\rho^*}[Z^2] - (\mathbf{E}_{\rho^*}[Z])^2 = 0$. Using (5) and noting that $\rho^* > 0$ since W is bounded and $\pi > 0$, it follows that Z has finite second moment and

$$\begin{aligned}\mathbf{E}_{\rho^*}[Z^2] - (\mathbf{E}_{\rho^*}[Z])^2 &= \mathbf{E}_\pi[\exp(-2W)\pi/\rho^*] - (\mathbf{E}_\pi[\exp(-W)])^2 \\ &= \exp(-\gamma)\mathbf{E}_\pi[\exp(-W)] - \exp(-2\gamma) \\ &= 0\,,\end{aligned}$$

where we have used that $\exp(-\gamma) = \mathbf{E}_\pi[\exp(-W)]$. □

The above theorem asserts that ρ^*-almost surely (ρ^*-a.s.)

$$Z = \mathbf{E}_{\rho^*}[Z]\,, \tag{6}$$

which means that the importance sampling scheme based on estimating $\mathbf{E}_{\rho^*}[Z]$ using draws from the density ρ^* is a zero-variance estimator of $\mathbf{E}_\pi[\exp(-W)]$. We will discuss the problem of drawing from an approximation of the optimal distribution ρ^* later on in Section 4.

Remark 1. *Equation (4) furnishes the famous relation $F = U - TS$ for the Helmholtz free energy F, with U being the internal energy, T the temperature and S denoting the Gibbs entropy. If we modify the previous assumptions by setting $\pi \equiv 1$ and assuming that $W = \beta E$ where $\beta = (k_B T)^{-1}$ with $k_B > 0$ being Boltzmann's constant and E denoting a smooth potential energy function that is bounded from below and growing at infinity, then*

$$\underbrace{-\beta^{-1}\log\int \exp(-\beta E)\,dx}_{=F} = \min_{\rho>0}\left\{\underbrace{\int E\rho\,dx}_{=U} + \underbrace{\beta^{-1}\int \rho\log\rho\,dx}_{=-TS}\right\}\,,$$

with the unique minimizer being the Gibbs-Boltzmann density $\rho^* = \exp(-\beta E)/Z$ with normalization constant $Z = \exp(-\beta F)$. In the language of statistics, ρ^* is a probability distribution from the exponential family with sufficient statistic $E(X)$ and parameter $\beta > 0$.

2.2. Yet another certainty equivalence

A similar variational characterization of expectation values based on convexity arguments and Jensen's inequality can be formulated for non-negative random variables $W = W(X) \geq 0$. For simplicity and as before, we assume W to be bounded and measurable and $\pi > 0$. Then, for all $p \in [1, \infty)$, it holds that

$$(\mathbf{E}_\pi[W^p])^{1/p} = \max_{\eta\geq 0}\mathbf{E}_\eta\left[W\left(\frac{\eta}{\pi}\right)^{-1/p}\right]\,, \tag{7}$$

where η is any non-negative probability density. If we exclude the somewhat pathological case $W \equiv 0$ a.s., it follows that $\mathbf{E}_\pi[W^p] > 0$ and the supremum is attained for

$$\eta^* = \frac{W^p}{\mathbf{E}_\pi[W^p]}\,\pi\,. \tag{8}$$

The proof is along the lines of the proof of the Donsker-Varadhan principle (4). Indeed, applying Jensen's inequality and noting that η/π is non-zero η-a.s. it readily follows that

$$(\mathbf{E}_\pi\,[W^p])^{1/p} \geq \mathbf{E}_\eta\left[W\left(\frac{\eta}{\pi}\right)^{-1/p}\right],$$

and it is easy to verify that the supremum in (7) is attained at $\eta = \eta^*$ given by (8).

Importance sampling, cont'd

Similar as in Subsection 2.1, the above discussions can be applied to study the importance sampling schemes for the pth moment of random variable W. We have

Theorem 2 (Optimal importance sampling, cont'd)**.** *Let $p > 1$ and η^* be defined in (8). Then the random variable $Y = W^p(\eta^*/\pi)^{-1}$ has zero variance under η^* and we have*

$$Y = \mathbf{E}_\pi[W^p]\,, \quad \eta^*\text{-a.s.}$$

Again, Theorem 2 implies that drawing random variable X from η^* and then estimating the reweighted expectation $\mathbf{E}_{\eta^*}[Y]$ provides a zero variance estimator for the quantity $\mathbf{E}_\pi[W^p]$.

Remark 2. *When $0 < p \leq 1$ the function $f(u) = u^p$ is concave for $u \geq 0$ and the variational principle (7) need to be modified as (see [6])*

$$(\mathbf{E}_\pi\,[W^p])^{1/p} = \min_{\eta \geq 0} \mathbf{E}_\eta\left[W\left(\frac{\eta}{\pi}\right)^{-1/p}\right], \tag{9}$$

where the minimizer η^* is given by (8). If $W > 0$ a.s., then η^* belongs to the exponential family with sufficient statistic $S(X) = p\log W(X)$ and reference density π.

2.3. Computational issues

In practice, the above result is of limited use because the optimal importance sampling distribution is only known up to the normalizing constant C where the latter is just the sought quantity $C = \exp(-\gamma)$.

We recall that one possible way to sample the probability density π is based on the Euler-Maruyama discretization of the stochastic differential equation (SDE)

$$dX_s = -\nabla V(X_s)ds + \sqrt{2}dB_s\,, \quad X_0 = x\,. \tag{10}$$

Here $V\colon \mathbb{R}^n \to \mathbb{R}$ is a smooth potential that is bounded from below and at least quadratically growing at infinity, so that the process $(X_s)_{s\geq 0}$ is ergodic with respect to the probability distribution with density $\pi = \exp(-V)$. Depending on the desired accuracy and the discretization time step one may use a metropolized variant of the Euler-Maruyama scheme [40]. The importance sampling density ρ^* can be generated in the same way by simulating

$$dY_s = -\nabla\Phi(Y_s)ds + \sqrt{2}dB_s\,, \quad Y_0 = x, \tag{11}$$

with the modified potential

$$\Phi(y) = V(y) + W(y)\,. \tag{12}$$

(The density η^* can be generated likewise.) Clearly, we can draw from ρ^* in the sense that, by the law of large numbers (ergodic theorem) for ergodic diffusions such as (11), it holds

$$\mathbf{E}_{\rho^*}[f] = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(Y_t)\,dt\,, \quad \text{a.s.} \tag{13}$$

for almost all initial conditions $Y_0 = x$ and any integrable function $f\colon \mathbb{R}^n \to \mathbb{R}$. The discrete form of (13) then reads

$$\mathbf{E}_{\hat{\rho}^*}[f] = \lim_{N\to\infty} \frac{1}{N}\sum_{i=0}^{N-1} f(\hat{Y}_i)\,, \quad \text{a.s.} \tag{14}$$

where $(\hat{Y}_i)_{i\geq 0}$ is the time discretization of $(Y_s)_{s\geq 0}$ and $\hat{\rho}^*$ is the invariant distribution of the corresponding Markov chain that clearly depends on the time discretization, but that we assume to be a reasonable approximation of the exact density ρ^*.

Sources of bias

In the situation at hand, we wish to apply either (13) or (14) to estimate $\mathbf{E}_{\rho^*}(Z)$ in (6), where $Z = \exp(-W)\pi/\rho^*$ is given in Theorem 1. However, the problem is that the likelihood ratio π/ρ^* is only known up to the normalizing factor. In this case, the self-normalized importance sampling estimator must be used (see, e.g. [19]),

$$\hat{C}_N = \frac{\sum_{i=0}^{N-1} \exp(-W(\hat{Y}_i))\exp(W(\hat{Y}_i))}{\sum_{i=0}^{N-1} \exp(W(\hat{Y}_i))} = \frac{N}{\sum_{i=0}^{N-1} \exp(W(\hat{Y}_i))}\,, \tag{15}$$

which is a consistent estimator for $C = \exp(-\gamma)$. Note that unlike in the case of the importance sampling estimators with known likelihood ratio, the self-normalized estimator is only asymptotically unbiased—even if we can draw exactly from ρ^* (See Appendix A for details.)

To avoid the bias due to the self-normalization, it is helpful to note that $\exp(\gamma) = \mathbf{E}_{\rho^*}[\exp(W)]$ holds ρ^*-a.s. As a consequence,

$$\hat{C}_N^{-1} = \frac{\sum_{i=0}^{N-1} \exp(W(\hat{Y}_i))}{N} \tag{16}$$

is an unbiased estimator of $C^{-1} = \exp(\gamma)$, provided that we can generate i.i.d. samples from ρ^*. Taking the logarithm, it follows that

$$\hat{\gamma}_N = -\log N + \log\left(\sum_{i=0}^{N-1} \exp(W(\hat{Y}_i))\right)\,, \tag{17}$$

is a consistent estimator for γ, which by Jensen's inequality and the strict concavity of the logarithm will again be only asymptotically unbiased.

Comparison with the standard Monte Carlo estimator

In most cases, the samples from ρ^* will be generated by Markov Chain Monte Carlo or by simulating an SDE like (11). If we disregard the possible time disretization bias and consider the advantages of $\hat{\gamma}_N$ as compared to the plain vanilla Monte Carlo estimator

$$\tilde{\gamma}_N = \log N - \log \left(\sum_{i=0}^{N-1} \exp(-W(\hat{X}_i)) \right) , \tag{18}$$

with $\hat{X}_i$ being a discretization of (10), there are two aspects that will influence the efficiency of (17) relative to (18), namely

(a) the speed of convergence towards the stationary distribution and
(b) the (asymptotic) variance of the estimator.

By construction, the asymptotic variance of the importance sampling estimator is zero or close to zero if we take numerical discretization errors into account, hence the efficiency of the estimator (17) is solely determined by the speed of convergence for SDE (11) to its stationary distribution ρ^* which, depending on the problem at hand, may be larger or smaller than the speed of convergence of the original process (10) to π. It may even happen that π is unimodal, whereas $\rho^* \propto e^{-W}\pi$ is multimodal and hence difficult to generate, for example when π is the standard Gaussian density and $W = (x^2 - d)^2$ with $d \gg 0$ is a bistable (energy) function. We refrain from going into details here, and instead refer to the review article [32] for an in-depth discussion of the asymptotic properties of reversible diffusions.

3. Certainty equivalence in path space

The previous considerations nicely generalize from the case of real-valued random variables to time-dependent problems and path functionals.

3.1. Donsker-Varadhan variational principle in path space

Let $(X_s)_{s\geq 0}$ with $X_0 = x \in \mathbb{R}^n$ be the solution of the SDE

$$dX_s = b(X_s, s)ds + \sigma(X_s)dB_s\,, \quad X_0 = x\,, \tag{19}$$

where $b\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}^n$ is a smooth, possibly time-dependent vector field, $\sigma\colon \mathbb{R}^n \to \mathbb{R}^{n\times m}$ is a smooth matrix field and B is an m-dimensional Brownian motion. Our standard example will be an SDE with $b(x,s) = -\nabla V(x)$ for a smooth potential energy function V and $\sigma(x) = \sqrt{2}I_{n\times n}$, so that X_s satisfies a gradient dynamics (10). We assume throughout this paper that the functions b, σ, V are such that either (19) or (10) have unique strong solutions for all $s \geq 0$.

Now suppose that we want to compute the free energy (2) where W is now considered to be a functional of the paths $X = \{X_s : 0 \leq s \leq \tau\}$ for some bounded stopping time τ:

$$W_\tau(X) = \int_0^\tau f(X_s, s)\, ds + g(X_\tau)\,, \tag{20}$$

for some bounded and sufficiently smooth, real valued functions f, g. We assume throughout the rest of the paper that f, g are bounded from below and that W is integrable.

We define P to be the probability measure on the space $\Omega = C([0,\infty), \mathbb{R}^n)$ of continuous trajectories that is induced by the Brownian motion $(B_s)_{s\geq 0}$ that drives the SDE (19). We call P a *path space measure*, and we denote the expectation with respect to P by $\mathbf{E}_P[\cdot]$.

Definition 3 (Path space free energy)**.** *Let $(X_s)_{s\geq 0}$ be the solution of (19) and $W_\tau = W_\tau(X) \geq 0$ be integrable and defined by (20). The quantity*

$$\gamma = -\log \mathbf{E}_P\left[\exp(-W_\tau)\right] = -\log \mathbf{E}_P\left[\exp\left(-\int_0^\tau f(X_s,s)\,ds - g(X_\tau)\right)\right] \tag{21}$$

is called the free energy of W_τ with respect to the path space measure P.

Note that (21) simply is the path space version of (2) which now implicitly depends on the initial condition $X_0 = x$. The Donsker-Varadhan variational principle now reads

$$\gamma = \inf_{Q\ll P}\left\{\mathbf{E}_Q\left[\int_0^\tau f(X_s,s)\,ds + g(X_\tau)\right] + D(Q|P)\right\}, \tag{22}$$

where $Q \ll P$ stands for *absolute continuity* of Q with respect to P, which means that $P(E) = 0$ implies that $Q(E) = 0$ for any measurable set $E \subset \Omega$, as a consequence of which

$$D(Q|P) = \int_\Omega \log\frac{dQ}{dP}(\omega)\,dQ(\omega)\,, \tag{23}$$

exists. Note that (23) is just the generalization of the relative entropy (3) from probability densities on $\mathbb{R}^n$ to probability measures on the measurable space $(\Omega,\mathcal{E})$, with $\mathcal{E}$ being a σ-algebra containing measurable subsets of Ω, where we again declare that $D(Q|P) = \infty$ when Q is not absolutely continuous with respect to P. Therefore it is sufficient that the infimum in (22) is taken over all path space measures $Q \ll P$.

If $W_\tau \geq 0$, it is again a simple convexity argument (see, e.g. [6]) which shows that the minimum in (22) is attained at Q^* given by

$$\left.\frac{dQ^*}{dP}\right|_{[0,\tau]} = \exp(\gamma - W_\tau)\,, \tag{24}$$

where $\varphi|_{[0,\tau]}$ denotes the restriction of the path space density $\varphi(X) = (dQ^*/dP)(X)$ to trajectories $X = (X_s)_{s\geq 0}$ of length τ.[2]

Even though (24) is the direct analogue of (5), this result is not particularly useful if we do not know how to sample from Q^*. So let us first characterise the admissible path space measures $Q \ll P$ and discuss the practical implications later on.

Likelihood ratio of path space measures

It turns out that the only admissible change of measure from P to Q such that $D(Q|P) < \infty$ results in a change of the drift in (19). Let $(u_s)_{s\geq 0}$ be an $\mathbb{R}^m$-valued stochastic process that is adapted, in that u_t depends only on the Brownian motion B_s up to time $s \leq t$, and that satisfies the *Novikov condition* (see e.g. [36])

$$\mathbf{E}_P\left[\exp\left(\frac{1}{2}\int_0^\tau |u_s|^2\,ds\right)\right] < \infty\,. \tag{25}$$

Now define the auxiliary process

$$B_t^u = B_t - \int_0^t u_s\,ds\,.$$

[2] More precisely, $\varphi|_{[0,\tau]}$ is understood as the restriction of the measure Q^* defined by $dQ^* = \varphi dP$ to the σ-algebra $\mathcal{F}_\tau$ that contains all measurable sets $E \in \mathcal{E}$, with the property that for every $t \geq 0$ the set $E \cap \{\tau \leq t\}$ is an element of the σ-algebra $\mathcal{F}_t = \sigma(X_s : 0 \leq s \leq t)$ that is generated by all trajectories $(X_s)_{0\leq s\leq t}$ of length t. In other words, $\mathcal{F}_\tau \subset \mathcal{E}$ is a σ-algebra that contains the history of the trajectories of (the random) length τ.

Using the definition of B_t^u, we may write (19) as

$$dX_s = (b(X_s, s) + \sigma(X_s)u_s)\, ds + \sigma(X_s)dB_s^u\,, \quad X_0 = x\,. \tag{26}$$

By construction, $(B_s^u)_{s\geq 0}$ is not a Brownian motion under P, because the expectation of B_s^u with respect to P is not zero in general. On the other hand, $(B_s)_{s\geq 0}$ is a Brownian motion under the measure P, and our aim is to find a measure $Q \ll P$ under which $(B_s^u)_{s\geq 0}$ is a Brownian motion. To this end, let $(Z_s^u)_{s\geq 0}$ be the process defined by

$$Z_t^u = \int_0^t u_s \cdot dB_s - \frac{1}{2}\int_0^t |u_s|^2\, ds\,, \tag{27}$$

or, equivalently,

$$Z_t^u = \int_0^t u_s \cdot dB_s^u + \frac{1}{2}\int_0^t |u_s|^2\, ds\,. \tag{28}$$

Girsanov's theorem (see e.g. [36, Thm. 8.6.4] or Appendix B) now states that $(B_s^u)_{0\leq s\leq \tau}$ is a standard Brownian motion under the probability measure Q with likelihood ratio

$$\left.\frac{dQ}{dP}\right|_{[0,\tau]} = \exp(Z_\tau^u) \tag{29}$$

with respect to P where the Novikov condition (25) guarantees that $\mathbf{E}_P[\exp(Z_\tau^u)] = 1$, i.e. that Q is a probability measure. Inserting (28)–(29) into the Donsker-Varadhan formula (22), using that B_s^u is Brownian motion with respect to Q, it follows that the term in Z_τ^u in the expression of the relative entropy which is linear in u drops out, and what remains is (cf. [4,6])

$$\gamma = \inf_u \mathbf{E}_Q\left[\int_0^\tau \left(f(X_s, s) + \frac{1}{2}|u_s|^2\right) ds + g(X_\tau)\right]\,,$$

with X_s being the solution of (26). Since the distribution of B^u under Q is the same as the distribution of B under P, an equivalent representation of the last equation is

$$\gamma = \inf_u \mathbf{E}_P\left[\int_0^\tau \left(f(X_s^u, s) + \frac{1}{2}|u_s|^2\right) ds + g(X_\tau^u)\right]\,. \tag{30}$$

where X_s^u is the solution of the controlled SDE

$$dX_s^u = (b(X_s^u, s) + \sigma(X_s^u)u_s)\, ds + \sigma(X_s^u)dB_s\,, \quad X_0^u = x\,, \tag{31}$$

with B_s being our standard, m-dimensional Brownian motion (under P). See Appendix B for a sketch of derivation of Girsanov's formula.

Importance sampling in path space

Similarly to the finite dimensional case considered in the last section, we can derive optimal importance sampling strategies from the Donsker-Varadhan principle. To this end, we consider the case that τ is a random stopping time, which is a case that is often relevant in applications (e.g. when computing transition rates or committor functions [21]), but that is rarely considered in the importance sampling literature. Let $T > 0$ and $O \subset \mathbb{R}^n$ be an open and bounded set with smooth boundary ∂O. We define

$$\tau_O = \inf\{s > 0\colon X_s^u \notin O\}\,,$$

as the first exit time of the set O and define the stopping time

$$\tau = \tau_O \wedge T \tag{32}$$

to be the minimum of τ_O and T, i.e. the exit from the set O or the end of the maximum time interval, whatever comes first.[3] Here X_s^u satisfies the controlled SDE (31).

We will argue that the optimal Q^* which yields zero variance in the reweighting scheme

$$\mathbf{E}_Q[Y_\tau^u] = \mathbf{E}_P[\exp(-W_\tau)]$$

via $Y_\tau^u = \exp(-Z_\tau^u - W_\tau^u)$, can be generated by a feedback control of the form

$$u_s^* = c(X_s^{u^*}, s)\,,$$

with a suitable function $c\colon \mathbb{R}^n \times [0,\infty) \supset D \to \mathbb{R}^m$. Finding u^* turns the Donsker-Varadhan variational principle (22) into an optimal control problem by virtue of (30)–(31). The following statement characterises the optimal control by which the infimum in (22) is attained and which, as a consequence, provides a zero variance reweighting scheme (or: change of measure).

Theorem 3. *Let*

$$\Psi(x,t) = \mathbf{E}_P\left[\exp\left(-\int_t^\tau f(X_s,s)\,ds - g(X_\tau)\right)\middle| X_t = x\right] \tag{33}$$

be the exponential of the negative free energy, considered as a function of the initial condition $X_t = x$ with $0 \le t \le \tau \le T$. Then the path space measure Q^* induced by the feedback control

$$u_s^* = \sigma(X_s^{u^*})^T \nabla_x \log \Psi(X_s^{u^*}, s) \tag{34}$$

yields a zero variance estimator, i.e.

$$Y_\tau^{u^*} = \Psi(x,0)\,, \quad Q^*\text{-a.s.} \tag{35}$$

Proof. See Appendix C. □

Remark 3. *Under the above conditions, if $f, g \ge 0$, excluding again the case $W \equiv 0$, Theorem 3 readily carries over to the case considered in (7), with Ψ replaced by*

$$\tilde{\Psi}(x,t) = \mathbf{E}_P[(W_\tau)^p \mid X_t = x]\,,$$

and the optimal control given by

$$\tilde{u}_s^* = \sigma(X_s^{u^*})^T \nabla_x \log \tilde{\Psi}(X_s^{u^*}, s)\,.$$

Remark 4. *We should mention that Theorem 3 covers also the special cases that either $\tau = T$ is a deterministic stopping time (see e.g. [29, Prop. 5.4.4]) or, by sending $T \to \infty$, that $\tau < \infty$ is an a.s. finite hitting time or first exit time.*

3.2. Revisiting Jarzynski's identity

The Donsker-Varadhan variational principle shares some features with the nonequilibrium free energy formulae of Jarzynski [27] and Crooks [5], and, in fact, the variational form makes these formulae amenable to the analysis of the previous paragraphs, with the aim of improving the quality of the corresponding statistical estimators. It turns out that the resulting importance sampling scheme shares some features with available adaptive numerical sampling schemes, such as the adaptive biasing

[3] For the ease of notation, we will use the same symbol τ to denote the stopping time under the controlled or uncontrolled process (i.e. for $u = 0$) throughout the article. Unless otherwise noted, it should be clear from the context whether τ is understood with respect to X^u or $X = X^{u=0}$.

force method [8,30,31], and we will discuss possible implications of this observation later on in Section 4.

Jarzynski's identity relates the Helmholtz equilibrium free energy to averages that are taken over an ensemble of non-equilibrium trajectories generated by forcing the dynamics. Let $(V_\lambda)_{0\leq\lambda\leq 1}$ be a parametric family of smooth potential energy functions $V_\lambda\colon \mathbb{R}^n \to \mathbb{R}$ and define the free energy difference between the two equilibrium densities $\pi_0 \propto \exp(-V_0)$ and $\pi_1 \propto \exp(-V_1)$ (also called: *thermodynamic states*) as the log-ratio

$$\Delta F = -\log \frac{\int_{\mathbb{R}^n} \exp(-V_1(x))dx}{\int_{\mathbb{R}^n} \exp(-V_0(x))dx}\,. \tag{36}$$

Defining the energy difference $V_{\text{diff}} = V_1 - V_0$ and the equilibrium probability density

$$\pi_0(x) = \frac{\exp(-V_0(x))}{\int_{\mathbb{R}^n} \exp(-V_0(x))dx}\,,$$

the Helmholtz free energy is seen to be an exponential average of the familiar form (2):

$$\Delta F = -\log \mathbf{E}_{\pi_0}[\exp(-V_{\text{diff}})]\,. \tag{37}$$

Jarzynski's formula [27] states that the last equation can be represented as an exponential average over a non-stationary realizations of a parameter-dependent process $X^\lambda = (X_s^\lambda)_{0\leq s\leq T}$. Specifically, letting $W_T^\lambda = W_T^\lambda(X^\lambda)$ denote the nonequilibrium work done on the system by varying the parameter from $\lambda = 0$ to $\lambda = 1$ within time $[0, T]$, Jarzynski's equality states that

$$\Delta F = -\log \mathbf{E}[\exp(-W_T^\lambda)]\,, \tag{38}$$

where W_T^λ will be specified below. In the last equation the expectation is taken over all realizations of X^λ, with initial conditions distributed according to the equilibrium density π_0. To be specific, we assume that the parametric process X_s^λ is the solution of the SDE

$$dX_s^\lambda = -(1-\lambda_s)\nabla V_0(X_s^\lambda)ds - \lambda_s \nabla V_1(X_s^\lambda)ds + \sqrt{2}dB_s\,, \tag{39}$$

with $(\lambda_s)_{0\leq s\leq T}$ being a differentiable parameter process (called: *protocol*) that interpolates between $\lambda_0 = 0$ and $\lambda_T = 1$. Further let the work exerted by the protocol be given by

$$W_T^\lambda = \int_0^T V_{\text{diff}}(X_s^\lambda)\dot{\lambda}_s\, ds\,, \tag{40}$$

where $\dot{\lambda}_s = d\lambda_s/ds$ denotes the time derivative of λ_s. Note that W_τ^λ is a path functional of the standard form (20), with bounded deterministic stopping time $\tau = T$ and cost functions

$$f(X_s, s) = V_{\text{diff}}(X_s^\lambda)\dot{\lambda}_s\,, \quad g \equiv 0\,.$$

Letting now P denote the path space measure that is generated by the Brownian motion $(B_s)_{s\geq 0}$ in the parameter dependent SDE (39), we can express Jarzynski's equality (38) by

$$\Delta F = -\log \mathbf{E}[\exp(-W_T^\lambda)] = -\log \int_{\mathbb{R}^n} \mathbf{E}_P\left[\exp(-W_T^\lambda)\right] \pi_0(x)dx\,, \tag{41}$$

where the (conditional) expectation $\mathbf{E}_P[\exp(-W_T^\lambda)] = \mathbf{E}_P[\exp(-W_T^\lambda(X^\lambda))|X_0^\lambda = x]$ is understood over all realizations of (39) with initial condition $X_0^\lambda = x$.

Optimized protocols by adaptive importance sampling

The applicability of Jarzynski's formula heavily depends on the choice of the protocol λ_s. The observation that an uneducated choice of a protocol may render the corresponding statistical estimator virtually useless because of a dramatic increase of its variance is in accordance with what one observes in importance sampling. An attempt to optimize the protocol by minimizing the variance of the estimator has been carried out in [35], but here we shall follow an alternative route, exploiting the fact that Jarzysnki's formula has the familiar exponential form considered in this paper.

Having this said and recalling Theorem 3, it is plausible that there exists a zero variance estimator for $\mathbf{E}_P[\exp(-W_T^\lambda)]$ which appeared in the integrand of Jarzynski's equality (41), under certain assumptions on the functional W_T^λ. For simplicity, we confine the following considerations to the above example of a diffusion process of the form (39) with a deterministic protocol $(\lambda_s)_{s\in[0,T]}$. To make the idea of optimizing the protocol more precise, we introduce the shorthand $Y_s = X_s^\lambda$ for the solution of (39) and define

$$\gamma(x,t) = \min_v \mathbf{E}_P\left[\int_t^T \left(f(Y_s^v,s) + \frac{1}{2}|v_s|^2\right) ds \,\middle|\, Y_t^v = x\right], \tag{42}$$

with $f(x,s) = V_{\text{diff}}(x)\dot{\lambda}_s$ and the expectation taken over all realizations of Y_s^v. The process Y_s^v solves a controlled variant of the SDE (39), specifically,

$$dY_s^v = \left(b_\lambda(Y_s^v,s) + \sqrt{2}v_s\right) ds + \sqrt{2}dB_s\,, \quad Y_t^v = x\,. \tag{43}$$

Here, we have used the shorthand $b_\lambda(x,s) = -(1-\lambda_s)\nabla V_0(x) - \lambda_s \nabla V_1(x)$. Theorem 3 that specifies the zero-variance importance sampling estimator in terms of a feedback control policy can be adapted to our situation (see e.g. [21,42]) by letting $O \uparrow \mathbb{R}^n$ so that $\tau = \tau_O \wedge T \to T$ a.s. The zero-variance estimator is generated by the feedback control

$$v_s^* = -\sqrt{2}\nabla_y \gamma(Y_s^{v^*},s)\,,$$

with $\gamma(x,t)$ given by (42), and thus by the SDE

$$dY_s^{v^*} = \left(b_\lambda(Y_s^{v^*},s) - 2\nabla_y\gamma(Y_s^{v^*},s)\right) ds + \sqrt{2}dB_s\,, \quad Y_0^{v^*} = x\,. \tag{44}$$

Specifically, given N independent draws $x_1,\ldots,x_N \sim \pi_0$ from the equilibrium distribution and corresponding N independent trajectories $(Y_s^{v^*})_{s\in[0,T]}$ of (44) with initial conditions $Y_0^{v^*} = x_i$, an asymptotically unbiased, minimum variance estimator of the free energy is given by

$$\widehat{\Delta F}_N = -\log \frac{1}{N}\sum_{i=1}^N G_T^*(x_i) \tag{45}$$

where $G_T^* = \exp(-Z_T^{v^*} - W_T^\lambda(Y^{v^*}))$ with $Z_T^{u=v^*}$ given by (27) and $W_T^\lambda(Y^{v^*})$ being the nonequilibrium work (40) under the controlled process (44).

Remark 5. *Note that the estimator (45) is not a zero-variance estimator because we have minimized only the conditional estimator (for fixed initial condition). Moreover the estimator is only asymptotically unbiased by Jensen's inequality and the strict concavity of the logarithm.*

Further notice that the estimator hinges on the availability of $\gamma(x,t)$ which is typically difficult to compute. An idea, inspired by the adaptive biasing force (ABF) algorithm [8,31] is to estimate γ on the fly and then iteratively refine the estimate in the course of the simulation using a suitable parametric representation [24,43]. If good collective variables or reaction coordinates are known, it is further possible to choose a representation that depends only on these variables and still obtain low variance estimators [23,25].

4. Algorithms: gradient descent, cross entropy minimization and beyond

According to Theorem 3 designing reweighting (importance sampling) schemes on path space that feature zero variance estimators comes at the price of solving an optimal control problem of the following form: minimize the cost functional

$$J(u) = \mathbf{E}_P\left[\int_0^\tau \left(f(X_s^u, s) + \frac{1}{2}|u_s|^2\right) ds + g(X_\tau^u)\right] \tag{46}$$

over all admissible controls and subject to the dynamics

$$dX_s^u = (b(X_s^u, s) + \sigma(X_s^u)u_s)\, ds + \sigma(X_s^u) dB_s\,. \tag{47}$$

Here admissible controls are Markovian feedback controls $u_s = c(X_s^u, s)$ such that (47) has a unique strong solution. Leaving all technical details aside (see [18, Sec. IV.3]), it can be shown that the value function (or: *optimal cost-to-go*)

$$\gamma(x,t) = \mathbf{E}_P\left[\int_t^\tau \left(f(X_s^{u^*}, s) + \frac{1}{2}|u_s^*|^2\right) ds + g(X_\tau^{u^*})\middle| X_t^{u^*} = x\right] \tag{48}$$

with u^* being the unique optimal control given by (34) is the solution of a nonlinear partial differential equation of Hamilton-Jacobi-Bellman type. Solving this equation numerically is typically even more difficult than solving the original sampling problem by brute-force Monte Carlo (especially when the state space dimension n is large).

Note that (46)–(47) is simply the concrete form of the Donsker-Varadhan principle when the path space measure is generated by a diffusion. Therefore the equivocation with the path space free energy (21) or (42) is not a coincidence, because by definition the value function *is* the free energy, considered as a function of the initial conditions. In other words and in view of Theorem 3, there is no need for further sampling once the value function is known.

We will now discuss concrete numerical algorithms to minimize (46)–(47) without resorting to the associated Hamilton-Jacobi-Bellman equation.

4.1. Gradient descent

The fact that solving the optimal control problem can be as difficult as solving the sampling problem suggests to combine the two in an iterative fashion using a parametric representation of the value function (or: free energy). To this end, notice that the optimal control is essentially a gradient force that can be approximated by

$$\hat{u}_s = -\sigma(X_s^u)^T \sum_{i=1}^N \alpha_i \nabla_x \phi_i(X_s^u, s)\,, \tag{49}$$

based on a finite-dimensional approximation

$$\hat{\gamma}(x,t) = \sum_{i=1}^N \alpha_i \phi_i(X_s^u, s) \tag{50}$$

of the value function with suitable smooth basis functions $\{\phi_i : \overline{D} \to \mathbb{R}^m : i = 1, \ldots, N\}$ that span an N-dimensional subspace of the space $C^{2,1}(D) \cap C(\overline{D})$ of classical solutions of the associated Hamilton-Jacobi-Bellman equation. Plugging the above representation into (46)–(47) yields the following finite-dimensional optimization problem: minimize

$$J(\hat{u}) = \mathbf{E}_P\left[\int_0^\tau \left(f(X_s^{\hat{u}}, s) + \frac{1}{2}|\hat{u}_s|^2\right) ds + g(X_\tau^{\hat{u}})\right] \tag{51}$$

over the controls $\hat{u}$ where $X^{\hat{u}}$ is the solution of the SDE (31) with control $u = \hat{u}$.

Let us define $\hat{J}(\boldsymbol{\alpha}) = J(\hat{u}(\boldsymbol{\alpha}))$, with the shorthand $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^T \in \mathbb{R}^N$. Because of the dependence of the process $X^{\boldsymbol{\alpha}}$ and the random stopping time $\tau = \tau^{\boldsymbol{\alpha}}$ on the parameter $\boldsymbol{\alpha}$, the functional $\hat{J}$ is not quadratic in $\boldsymbol{\alpha}$, but it has been shown [33] that it is strongly convex if the basis functions ϕ_i are non-overlapping. In this case $\hat{J}$ has a unique minimum, which suggests to do a gradient descent in the parameter $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha}^{(m+1)} = \boldsymbol{\alpha}^{(m)} - h_m \nabla \hat{J}\left(\boldsymbol{\alpha}^{(m)}\right) . \tag{52}$$

Here $(h_m)_{m\geq 0}$ is a sequence of step sizes that goes to zero as $m \to \infty$, and the gradient $\nabla \hat{J}(\boldsymbol{\alpha})$ must be interpreted in the sense of a functional derivative

$$\delta J(\hat{u}) \cdot \xi = \frac{d}{d\epsilon} J(\hat{u} + \epsilon\xi)\Big|_{\epsilon=0} , \tag{53}$$

for suitable test functions $\xi \in \mathscr{V}$ (i.e. square-integrable and adapted to the Brownian motion). Then the gradient $\nabla \hat{J}(\boldsymbol{\alpha})$ has the components

$$\frac{\partial \hat{J}}{\partial \alpha_k} = -\delta J(\hat{u}(\boldsymbol{\alpha})) \cdot (\sigma^T \nabla_x \phi_k) . \tag{54}$$

Introducing the shorthand

$$\ell(X^{\hat{u}}, \hat{u}) = \int_0^\tau \left(f(X_s^{\hat{u}}, s) + \frac{1}{2}|\hat{u}_s|^2 \right) ds + g(X_\tau^{\hat{u}})$$

for the cost and the convention $\mathbf{E}[\cdot] = \mathbf{E}_P[\cdot]$ for the expectation with respect to P, the derivative (53) can again be found by means of Girsanov's formula: there exists a measure Q^ϵ that is absolutely continuous with respect to the reference measure P, such that

$$\frac{d}{d\epsilon} J(\hat{u} + \epsilon\xi)\Big|_{\epsilon=0} = \frac{d}{d\epsilon} \mathbf{E}\left[\ell(X^{\hat{u}+\epsilon\xi}, \hat{u} + \epsilon\xi)\right]\Big|_{\epsilon=0} = \frac{d}{d\epsilon} \mathbf{E}\left[\ell(X, \hat{u} + \epsilon\xi) \frac{dQ^\epsilon}{dP}\right]\Big|_{\epsilon=0} , \tag{55}$$

with the likelihood ratio

$$\frac{dQ^\epsilon}{dP}\Big|_{[0,\tau]} = \exp(Z_\tau^{u+\epsilon\xi}) .$$

Assuming that the derivative and the expectation in (55) commute, we can differentiate inside the expectation $\mathbf{E}[\cdot]$ which is independent of the parameter ϵ and then switch back to the controlled process X^u under the reference measure P, by which we obtain (see [33])

$$\delta J(\hat{u}) \cdot \xi = \mathbf{E}\left[\ell(X^{\hat{u}}, \hat{u}) \int_0^\tau \xi_s \cdot dB_s + \int_0^\tau \hat{u}_s \cdot \xi_s \, ds\right] . \tag{56}$$

Hence, using (54), we find

$$\frac{\partial \hat{J}}{\partial \alpha_k} = -\mathbf{E}\left[\ell(X^{\hat{u}}, \hat{u}) \int_0^\tau (\sigma^T \nabla_x \phi_k)(X_s^{\hat{u}}, s) \cdot dB_s + \int_0^\tau \hat{u}_s \cdot (\sigma^T \nabla_x \phi_k)(X_s^{\hat{u}}, s) ds\right] . \tag{57}$$

where the last expression can be estimated by Monte Carlo, possibly in combination with variance minimizing strategies to improve the convergence of the gradient estimation in the course of the gradient descent [33,39]. Before we conclude, we shall briefly explain why the gradient vanishes when the variance is zero.

Lemma 4. *Under the optimal control u^*, it holds that*

$$\delta J(u^*) \cdot \xi = 0 \quad \forall \xi \in \mathscr{V}\,.$$

Proof. By the Itô isometry [36, Cor. 3.1.7], we can recast (56) as

$$\begin{aligned}\delta J(u^*) \cdot \xi &= \mathbf{E}\left[\left(\ell(X^{u^*}, u^*) + \int_0^\tau u_s^* \cdot dB_s\right) \int_0^\tau \xi_s \cdot dB_s\right] \\ &= \mathbf{E}\left[\left(W(X^{u^*}) + Z_\tau^{u^*}\right) \int_0^\tau \xi_s \cdot dB_s\right] \\ &= \left(W(X^{u^*}) + Z_\tau^{u^*}\right) \mathbf{E}\left[\int_0^\tau \xi_s \cdot dB_s\right],\end{aligned} \tag{58}$$

where in the last equality we have used that $W + Z$ is a.s. constant under the optimal control. Since B is a Brownian motion under P, the expectation is zero and it follows that

$$\delta J(u^*) \cdot \xi = 0 \quad \forall \xi \in \mathscr{V}\,,$$

and hence the assertion is proved. □

We summarize the above considerations in Algorithm 1 below.

Algorithm 1 Gradient descent

Set maximum no. of iteration M_{axit} and $\epsilon > 0$
Initialize $m = 0$, $\boldsymbol{\alpha}^{(0)} \in \mathbb{R}^N$ and $h_0 > 0$
Evaluate $G_0 = \nabla \hat{J}(\boldsymbol{\alpha}^{(0)})$
while $m < M_{\text{axit}}$ & $h_m > \epsilon$ **do**
  $\boldsymbol{\alpha}^{(m+1)} = \boldsymbol{\alpha}^{(m)} - h_m G_m$
  Evaluate $G_{m+1} = \nabla \hat{J}(\boldsymbol{\alpha}^{(m+1)})$
  Evaluate step size, e.g. $h_{m+1} = \frac{\left(\boldsymbol{\alpha}^{(m+1)} - \boldsymbol{\alpha}^{(m)}\right) \cdot \left(G_{m+1} - G_m\right)}{\left|G_{m+1} - G_m\right|^2}$
  $m \mapsto m + 1$
end while

Remark 6. *The step size control in Algorithm 1 follows the Barzilai–Borwein procedure that guarantees convergence as $m \to \infty$ when the funtional is convex. Another alternative is to do a line search after each iterate in the descent direction and then determine h_{m+1} so that it satisfies the Wolfe condition; see [34] for further details.*

Remark 7. *In practice, it may be advantageous to pick the basis functions that are not be explicitly time-dependent (e.g. Gaussians, Chebyshev polynomials or the alike). If the associated control problem is stationary, as is for example the case when the SDE is homogeneous and the stopping time is a hitting time, the value function will be stationary too and, as a consequence, the control policy will be stationary. If, however, the problem is explicitly time-dependent, one may change the ansatz (50) to have stationary basis functions, but time-dependent coefficients α_i, where the time-dependence is mediated by the initial data; see [24] for a discussion.*

4.2. Cross-entropy minimization

Another algorithm for minimizing $J(\hat{u})$ is based on an entropy representation of $J(u)$, namely,

$$J(u) = J(u^*) + D(Q|Q^*) \tag{59}$$

where u is any admissible control for (46)–(47), u^* is the optimal control, and $Q = Q(u)$ and $Q^* = Q(u^*)$ are the corresponding path space measures. Equation (59) is a consequence of the zero-variance property of the optimal change of measure, since (35) implies that

$$\exp(-W)\frac{dP}{dQ^*} = \psi(x,0) = \exp(-J(u^*)) \tag{60}$$

and hence

$$W - \log\left(\frac{dP}{dQ}\frac{dQ}{dQ^*}\right) = J(u^*)\,. \tag{61}$$

Taking the expectation with respect to Q and using that both Q and Q^* are absolutely continuous with respect to P and vice versa yields (59).

The idea now is to seek a minimizer of $D(Q|Q^*)$ in the set of probability measures $Q \in \hat{M}$ that are generated by the discretized controls $\hat{u}$, i.e. one would like to minimize

$$\hat{I}(\boldsymbol{\alpha}) = D(Q(\hat{u}(\boldsymbol{\alpha}))|Q^*) \tag{62}$$

over $\boldsymbol{\alpha} \in \mathbb{R}^N$, such that $\hat{Q} = Q(\hat{u}(\boldsymbol{\alpha}))$ is absolutely continuous with respect to Q^*. By (24) the optimal change of measure is only known up to the normalizing factor $\exp(\gamma)$, so unfortunately minimizing $\hat{I}$ is not possible without knowledge of γ or, equivalently, $J(u^*)$.[4] With a little trick, however, we can turn the unfeasible minimization of (62) into a feasible minimization problem, simply by flipping the arguments. To this end, we define

$$\hat{H}(\boldsymbol{\alpha}) = D(Q^*|Q(\hat{u}(\boldsymbol{\alpha})))\,. \tag{63}$$

Clearly (59) does not hold with arguments in the KL divergence term reversed, since $D(\cdot|\cdot)$ is not symmetric, nevertheless it holds that

$$\hat{I}(\boldsymbol{\alpha}) \geq 0\,,\ \hat{H}(\boldsymbol{\alpha}) \geq 0 \quad \text{and} \quad \hat{I}(\boldsymbol{\alpha}) = 0 \text{ if and only if } \hat{H}(\boldsymbol{\alpha}) = 0\,, \tag{64}$$

where the minimum is attained if and only if $\hat{Q} = Q^*$. Hence, by continuity of the relative entropy, we may expect that by minimizing the "wrong" functional $\hat{H}$ we get close to the optimal change of measure, provided that the optimal Q^* can be approximated by our parametric family $\hat{Q}$. We have the following handy result (see [44]).

Lemma 5 (Cross-entropy minimization)**.** *The minimization of (63) is equivalent to the minimization of the cross-entropy functional*

$$CE(\boldsymbol{\alpha}) = -\mathbf{E}\left[\log \varphi(\hat{u})e^{-W(X)}\right] \tag{65}$$

where the log likelihood ratio $\log\varphi = \log(dQ/dP)$ between controlled and uncontrolled trajectories is quadratic in the unknown $\boldsymbol{\alpha}$ and can be computed via Girsanov's theorem.

Proof. By definition of the KL divergence, we have

$$\hat{H}(\boldsymbol{\alpha}) = \int \log\left(\frac{dQ^*}{dP}\right)\frac{dQ^*}{dP}dP - \int \log\left(\frac{d\hat{Q}}{dP}\right)\frac{dQ^*}{dP}dP\,,$$

[4] We call $\exp(\gamma)$ or $\exp(J(u^*))$ a normalizing *factor*, even though it is clearly a function of the initial conditions (x,t) or $(x,0)$.

since all measures are mutually absolutely continuous. The first term in the last equation is independent of $\boldsymbol{\alpha}$ and the second term is proportional to the cross-entropy functional $-\mathbf{E}[\log \varphi \exp(-W)]$ up to the unknown normalizing factor $\exp(\gamma)$. $\square$

The fact that the cross-entropy functional is quadratic in $\boldsymbol{\alpha}$ implies that the necessary optimality condition is of the form

$$S\boldsymbol{\alpha} = b\,, \tag{66}$$

where $S = (S_{ij})_{1\le i,j\le N}$ and $b = (b_i)_{1\le i\le N}$ are given by

$$\begin{aligned} S_{ij} &= \mathbf{E}\left[e^{-W(X)} \int_0^\tau (\sigma^T \nabla_x \phi_i)(X_s, s)(\sigma^T \nabla_x \phi_j)(X_s, s)\, ds\right] \\ b_i &= -\mathbf{E}\left[e^{-W(X)} \int_0^\tau (\sigma^T \nabla_x \phi_i)(X_s, s) \cdot dB_s\right]\,. \end{aligned} \tag{67}$$

Note that the average in (67) is over the uncontrolled realizations X. It is easy to see that the matrix S is positive definite if the basis functions ϕ_i are linearly independent, which implies that (66) has a unique solution and our necessary condition is in fact sufficient. Nevertheless it may happen in practice that the coefficient matrix S is badly conditioned, in which case it may be advisable to evaluate the coefficients using importance sampling or a suitable annealing strategy; see [24,44] for further details.

A simple, iterative variant of the cross-entropy algorithm is summarized in Algorithm 2.

Algorithm 2 Simple cross-entropy method

Set maximum no. of iteration M_{axit} and $\boldsymbol{\alpha}^{(0)} = 0$
Evaluate $S = S^{(0)}$ and $b = b^{(0)}$ according to (67)
for $m = 0$ to M_{axit} **do**
  Solve linear system of equations $S^{(m)}\boldsymbol{\alpha}^{(m+1)} = b^{(m)}$
  Evaluate $S^{(m+1)}$ and $b^{(m+1)}$ by importance sampling using realizations of $X^{\boldsymbol{\alpha}^{(m+1)}}$
end for

4.3. Other Monte-Carlo-based methods

We refrain from listing all possibilities to compute the optimal change of measure or the optimal control, and mention only two more possilities that are functional in situations in which grid-based discretization methods (e.g. for solving the nonlinear Hamilton-Jacobi-Bellman equation) are unfeasible.

Approximate policy iteration

The first option that is based on successive linearization of the Hamilton-Jacobi-Bellman equation of the underlying optimal control problem. The idea is the following: Given any admissible control $u_s = c(X_s^u, s)$, the Feynman-Kac Theorem [36, Thm. 8.2.1] states that the cost functional $J(u)$, considered as a function of the initial data (x,t) of the controlled process $X^u = (X_s^u)_{s\ge t}$ with $X_t^u = x$, solves a linear boundary value problem of the form

$$\mathcal{A}(c)\Theta(c) = \ell(x,c)\,, \tag{68}$$

where $\mathcal{A}(c)$ is a linear differential operator that depends on the chosen control policy and which precise form (e.g. parabolic, elliptic or hypoelliptic) depends on the problem at hand. Clearly, $\gamma(x,t) = \min_c \Theta(c;x,t)$ is the value function (or free energy), i.e. the solution we seek. For an arbitrary initial

choice of a control policy $c_0 \neq c^*$ we have $\gamma < \Theta(c_0)$, and a successive improvement of the policy can be obtained by iterating

$$c_{n+1}(x,s) = -\sigma(x,s)^T \nabla_x \Theta(c_n; x, t)\,, \quad n \geq 0\,. \tag{69}$$

Under suitable assumptions on the drift and diffusion coefficients, iteration of (68)–(69) yields a convergent series of control policies c_n that converges to the unique optimal control, hence the name of the method is *policy iteration*. Clearly, solving the linear partial differential equation (68) by any grid-based method will be unfeasible if the state space dimension is larger than, say, 3 or 4. In this case, it is possible to approximate the infinitesimal generators $\mathcal{A}(c)$ by a suitable grid-free Markov chain approximation of the underlying dynamics $X^u = X^{u(c)}$; see e.g. [1]. In this case, one speaks of an *approximate policy iteration*. For further details on approximate policy iteration algorithms we refer to the article [3] and the references therein.

Least-squares Monte Carlo

If $\tau = T$ is a finite stopping time, another alternative is to exploit that the value function of the control problem (46)–(47) can be computed as the solution to a forward-backward stochastic differential equation (FBSDE) of the following form

$$\begin{aligned} dX_s &= b(X_s,s)ds + \sigma(X_s)dB_s\,,\ X_t = x \\ dY_s &= -f(X_s,s)ds + \frac{1}{2}|Z_s|^2 + Z_s \cdot dB_s\,,\ Y_T = g(X_T)\,. \end{aligned} \tag{70}$$

where $t \leq s \leq T$ and the second equation must be interpreted as an equation that runs backwards in time. A solution of the FBSDE (70) is a triplet (X_s, Y_s, Z_s), with the property that $Y_{s'}$ and $Z_{s'}$ at time $s' \in [t,T]$ depend only on the history of the forward process $(X_s)_{t \leq s \leq s'}$ up to time s'. In particular, since $X_t = x$, the backward process Y_t is a deterministic function of the initial data (x,t) only, and it holds that (e.g. [17])

$$\gamma(x,t) = Y_t\,. \tag{71}$$

The specific structure of the control problem (46)–(47) implies that the forward equation is decoupled from the backward equation and that the backward process (Y_s, Z_s) can be expressed by

$$Y_s = \gamma(X_s,s)\,, \quad Z_s = -\sigma(X_s)^T \nabla_x \gamma(X_s,s)\,,$$

where X_s is the uncontrolled forward process. Since we can simulate the forward process and we know the functional dependence of (Y_s, Z_s) on X_s, the idea here is again to use the representation (50) of γ in terms of a finite basis. It turns out that the coefficient vector $\boldsymbol{\alpha} \in \mathbb{R}^N$ can be computed by solving a least-squares problem in every time step of the time-discretized backward SDE, which is why methods for solving an FBSDE like (70) are termed *least squares Monte Carlo*; for the general approach we refer to [2,16]; details for the situation at hand will be addressed in a forthcoming paper [26].

5. Illustrative examples

From a measure-theoretic viewpoint, changing the drift of an SDE (a.k.a. Girsanov transformation) is an exponential tilting of a Gaussian measure on an infinite-dimensional space. Here, for illustration purposes, we consider a one-dimensional paradigm that is in the spirit of Section 2 and that illustrates the basic features of Gaussian measure changes, Girsanov transformations and the cross-entropy method.

To this end, let $\pi = \mathcal{N}(0,1)$ the density of the standard Gaussian distribution on $\mathbb{R}$, and define an exponential family of "tilted" probability densities by

$$\rho_\alpha(x) = \exp\left(\alpha x - \frac{\alpha^2}{2}\right)\pi(x)\,. \tag{72}$$

It can be readily checked that ρ_α is the density of the normal distribution $\mathcal{N}(\alpha,1)$ with mean α and unit variance, in other words, the exponential tilting results in a shift of the mean, which represents a change of the drift in the case of an SDE.[5]

5.1. Example 1 (moment generating function)

Let $\beta > 0$ and define

$$\psi_\beta = \mathbf{E}_\pi[\exp(-\beta X)]\,. \tag{73}$$

By Jensen's inequality, it follows that

$$-\beta^{-1}\log \mathbf{E}_\pi[\exp(-\beta X)] \leq \mathbf{E}_\alpha\left[X + \beta^{-1}\log\frac{\rho_\alpha}{\pi}\right]\,, \tag{74}$$

where $\mathbf{E}_\alpha[\cdot]$ denotes the expectation with respect to ρ_α. A simple calculation shows that the inequality is sharp where equality is attained for $\alpha = -\beta$, i.e. when $\rho_\alpha = \rho^*$, with

$$\rho^* = \mathcal{N}(-\beta,1)\,. \tag{75}$$

As a consequence, the Donsker-Varadhan variational principle (4) holds when the minimum is taken over the exponential family (72) with sufficient statistic X.

We will now show that ρ^* can be computed by the cross-entropy method. To this end, let

$$J(\alpha) = \mathbf{E}_\alpha\left[X + \beta^{-1}\log\frac{\rho_\alpha}{\pi}\right]\,. \tag{76}$$

As we have just argued, there exists a unique minimizer $\alpha^* = -\beta$ of J that by Theorem 1 has the zero variance property which implies that

$$J(\alpha) = J(\alpha^*) + \beta^{-1}D(\rho_\alpha|\rho^*)\,. \tag{77}$$

The associated cross-entropy functional has the form (see page 16)

$$CE(\alpha) = -\mathbf{E}_\pi\left[\log\frac{\rho_\alpha}{\pi}\,\exp(-\beta X)\right]\,. \tag{78}$$

Using (72), it is easily seen that the cross-entropy funtional is quadratic,

$$CE(\alpha) = \mathbf{E}_\pi\left[\left(\frac{\alpha^2}{2} - \alpha X\right)\exp(-\beta X)\right]\,, \tag{79}$$

with unique minimizer

$$\hat{\alpha} = \frac{\mathbf{E}_\pi[X\exp(-\beta X)]}{\mathbf{E}_\pi[\exp(-\beta X)]} = -\frac{\partial}{\partial\beta}\log\psi_\beta\,, \tag{80}$$

[5] Compare equations (27)–(29).

where the second equality follows from (73), using the fact that the derivative and the expectation commute because π is Gaussian and hence the moment-generating function ψ_β exists for all $\beta \in \mathbb{R}$. Rearranging the terms in the last equation, we obtain

$$-\frac{\partial}{\partial\beta}\log\psi_\beta = \mathbf{E}_\pi\left[X\exp\left(-\beta X - \frac{\beta^2}{2}\right)\right] = \mathbf{E}_{-\beta}[X] = -\beta\,, \tag{81}$$

showing that

$$\hat{\alpha} = \alpha^* = -\beta\,.$$

The above consideration readily generalize to the multidimensional Gaussian case, and hence this simple example illustrates that the cross-entropy method yields the same result as direct minimization of the functional (76)—at least in the finite-dimensional case.

5.2. Example 2 (rare event probabilities)

The following example illustrates that the cross-entropy method can be used and produces meaningful results, even though the Donsker-Varadhan principle does not hold. To this end consider again the case of a real-valued random variable $X \sim P$ with density $\pi = \mathcal{N}(0,1)$ and $W = -\log \mathbf{1}_{\{X>d\}}$ with $d \gg 0$. Then

$$P(X > d) = \mathbf{E}_\pi\left[\exp(-W)\right] \tag{82}$$

is a small probability that is difficult to compute by brute-force Monte Carlo. In this case, a zero-variance change of measure exists, but it is not of the form (72). As a consequence, equality in (74) cannot be attained within the exponential family $\{\rho_\alpha : \alpha \in \mathbb{R}\}$ given by (72). Instead, the optimal density in this case would be the conditional density

$$\rho^*(x) = \frac{\mathbf{1}_{\{x>d\}}}{p}\pi(x)\,, \tag{83}$$

where the normalization constant $p = P(X > d)$ is of course the quantity we want to compute (cf. page 3). Note that this expression formally agrees with the optimal density (5), which was derived under different assumptions though.

The idea now is to minimize the distance between ρ_α and ρ^* in the sense of relative entropy, i.e. we seek a minimizer of the Kullback-Leibler divergence $D(\rho^*|\rho_\alpha)$ in the exponential family $\{\rho_\alpha : \alpha \in \mathbb{R}\}$. The associated cross-entropy functional is given by

$$CE(\alpha) = \mathbf{E}_\pi\left[\left(\frac{\alpha^2}{2} - \alpha X\right)\mathbf{1}_{\{X>d\}}\right]\,, \tag{84}$$

with unique minizer

$$\alpha^* = \frac{\mathbf{E}_\pi[X\mathbf{1}_{\{X>d\}}]}{\mathbf{E}_\pi[\mathbf{1}_{\{X>d\}}]} = \mathbf{E}_\pi[X|X>d]\,. \tag{85}$$

Comparing (85) and (83), we observe that both densities ρ_{α^*} and $\rho^* \neq \rho_{\alpha^*}$ have the same mean (namely α^*), hence the suboptimal density ρ_{α^*} is concentrated around the typical values that the optimal density ρ^* would produce when samples were drawn from it.

Clearly, the optimal tilting parameter (85) is probably as difficult to compute by brute-force Monte Carlo as the probability $p = P(X > d)$ since $\{X > d\}$ is a rare event when $d \gg 0$ is far away from the mean. The strength of both gradient descent and cross-entropy method is, however, that the optimal tilting parameter can be computed iteratively. This is illustrated numerically in Figure 1 for the choice

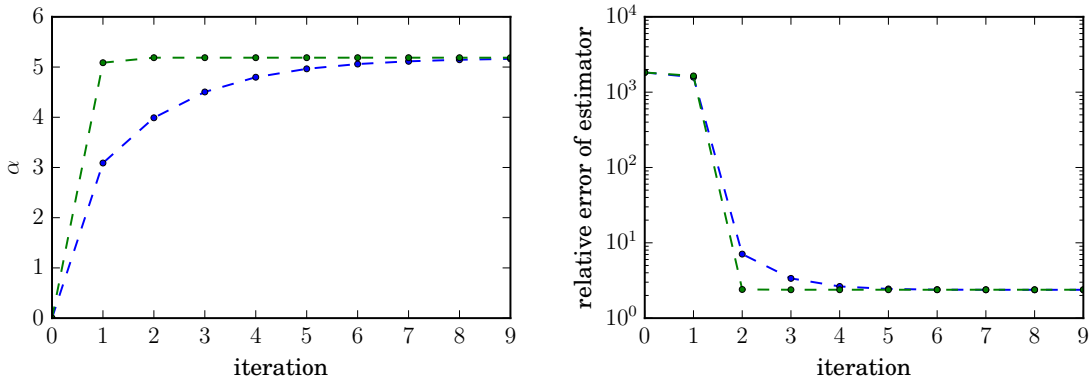

Figure 1. Comparison of the cross-entropy (green) and the gradient descent method (blue) for a rare event with probability $p \approx 2.867 \times 10^{-7}$ for fixed sample size $N = 10^8$. Both algorithms quickly converge to the optimal tilting parameter $\alpha^* \approx 5.187$ for the family $\mathcal{N}(\alpha, 1)$ of importance sampling distributions (left panel) and lead to a drastic reduction of the normalized relative error by a factor of 1000, from about 2000 to 2.38 after few iterations (right panel).

$d = 5$, where we use Algorithm 1 with a constant stepsize and Algorithm 2 as specified. In each iteration m we draw a sample of size $N = 10^8$ from the density $\rho^{\alpha^{(m)}}$, estimate the mean

$$\hat{p} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{\{X_i > d\}} \frac{\pi}{\rho^{\alpha^{(m)}}}(X_i)$$

and the sample variance in each sample. The latter is proportional to the normalized variance $K\,\mathrm{Var}(\hat{p})$ of an estimator that has been estimated K times.

For this (admittedly simple) example both gradient descent and cross-entropy method converge well and lead to a drastic reduction of the normalized relative error $\delta = \sqrt{\mathrm{Var}}(\hat{p})/\hat{p}$ of the estimator by a factor of about 1000, from about 2000 without importance sampling to about $\delta \approx 2.38$ under (suboptimal) importance sampling with exponential tilting, indicating that both methods can handle situations in which the optimal (i.e. $\delta = 0$) change of measure is not available within the set of trial densities.

6. Conclusions

We have presented a method for constructing minimum-variance importance sampling estimators. The method is based on a variational characterization of the thermodynamic free energy and essentially replaces a Monte Carlo sampling problem by a stochastic approximation problem for the optimal importance sampling density. For path sampling, the stochastic approximation problem boils down to a Markov control problem, which again can be solved by stochastic optimization techniques. We have proved that for a large class of path sampling problems that are relevant in e.g. molecular dynamics or rare events simulation the (unique) solution to the optimal control problem can yield zero-variance importance sampling schemes.

The computational gain when replacing the sampling problem by a variational principle is—besides improved convergence due to the variance reduction and often a higher hitting rate of the relevant events—due to the fact that the variational problem can be solved iteratively, which makes it amenable to multilevel approaches. The cross-entropy method as an examples of such an approach has been presented in some detail. A substantial difficulty still is a clever choice of basis functions that is highly problem-specific, and hence future research should address non-parametric approaches

as well as model reduction methods in combination with the stochastic optimization/approximation tools that can be used to solve the underlying variational problems.

Acknowledgments: This work was funded by the Einstein Center for Mathematics (ECMath) and by the Deutsche Forschungsgemeinschaft (DFG) under the grant DFG-SFB 1114 "Scaling Cascades in Complex Systems".

Appendix A Ratio estimators

We shall briefly discuss the properties of the self-normalized importance sampling estimator (15) that is based on estimating a ratio of expectation values

$$q_N = \frac{1}{N}\sum_{i=1}^{N} Q_i\,, \quad p_N = \frac{1}{N}\sum_{i=1}^{N} P_i$$

where Q_i and P_i are i.i.d. random variables living on a joint probability space and having finite variances σ_Q^2 and σ_P^2 and covariance σ_{QP}. Further assume that $q = \mathbf{E}[Q_1] \neq 0$, then, by the strong law of large numbers, the ratio p_N/q_N converges a.s. to p/q where $p = \mathbf{E}[P_1]$.

Appendix A.1 The delta method

We can apply the delta method (e.g. [19, Sec. 4.1]) to analyse the behaviour of the ratio estimator in more detail. Roughly speaking, the delta method says that for a sum $S_N = X_1 + \ldots + X_N$, $N \in \mathbb{N}$ of square-integrable, i.i.d. random variables X_k with mean $\mu \in \mathbb{R}^n$, covariance matrix $\Sigma \in \mathbb{R}^{n\times n}$ and a sufficiently smooth function $\phi\colon \mathbb{R}^n \to \mathbb{R}$ which can be Taylor expanded about μ, the central limit theorem applies. Specifically, using mean value theorem, it is easily seen that

$$\phi(S_N/N) - \phi(\mu) = \nabla\phi(\zeta_N)(S_N/N - \mu)$$

for some $\zeta_N \in \mathbb{R}^n$ lying component-wise in the half open interval between S_N/N and μ. By the continuity of $\nabla\phi$ at μ, the fact that $S_N/N \to \mu$ a.s. as $N \to \infty$ and that $\sqrt{N}(S_N/N - \mu)$ is approximately Gaussian with mean zero and covariance Σ, we have

$$\sqrt{N}\left(\phi(S_N/N) - \phi(\mu)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \nabla\phi(\mu)^T\Sigma\nabla\phi(\mu)\right)\,, \quad N \to \infty\,.$$

Appendix A.2 Asymptotic properties of ratio estimators

Applying the delta method to the function $\phi\colon \mathbb{R}^2 \to \mathbb{R}$, $(u,v) \mapsto u/v$, and assuming that $|v|$ is bounded away from zero, we find that the ratio estimator satisfies a central limit theorem too. Specifically, assuming that $q \neq 0$ so that $|q_N|$ is asymptotically bounded away from zero, the delta method yields

$$\sqrt{N}\left(\frac{p_N}{q_N} - \frac{p}{q}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma^2\right)\,,$$

with variance

$$\sigma^2 = \frac{\mathrm{Var}\left(P_1 - \frac{p}{q}Q_1\right)}{q^2}\,.$$

In particular, the estimator is asymptotically unbiased.

Appendix B Finite-dimensional change of measure formula

We will explain the basic idea behind Girsanov's Theorem and the change of measure formula (29). To keep the presentation easily accessible, we present only a vanilla version of the theorem based on finite-dimensional Gaussian measures, partly following an idea in [38].

Appendix B.1 Gaussian change of measure

Let P be a probability measure on a measurable space $(\Omega, \mathcal{E})$, on which an m-dimensional random variable $B\colon \Omega \to \mathbb{R}^m$ is defined. Further suppose that B has standard Gaussian distribution $P_B = P \circ B^{-1}$. Given a (deterministic) vector $b \in \mathbb{R}^n$ and a matrix $\sigma \in \mathbb{R}^{n\times m}$, we define a new random variable $X\colon \Omega \to \mathbb{R}^n$ by

$$X(\omega) = b + \sigma B(\omega)\,. \tag{A1}$$

The similarity to the SDE (19) is no coincidence. Since B is Gaussian, so is X, with mean b and covariance $C = \sigma\sigma^T$. Now let $u \in \mathbb{R}^n$ and define the shifted Gaussian random variable

$$B^u(\omega) = B(\omega) - u$$

and consider the alternative representation

$$X(\omega) = b^u + \sigma B^u(\omega) \tag{A2}$$

of X that is equivalent to (A1) if and only if

$$\sigma u = b^u - b$$

has a solution (that may not be unique though). Following the line of Section 3.1, we seek a probability measure $Q \ll P$ such that B^u is standard Gaussian under Q, and we claim that such a Q should have the property

$$\frac{dQ}{dP}(\omega) = \exp\left(u \cdot B(\omega) - \frac{1}{2}|u|^2\right) \tag{A3}$$

or, equivalently,

$$\frac{dQ}{dP}(\omega) = \exp\left(u \cdot B^u(\omega) + \frac{1}{2}|u|^2\right)\,, \tag{A4}$$

in accordance with (27)–(29). To show that B^u is indeed standard Gaussian under the above defined measure Q, it is sufficient to check that for any measurable (Borel) set $A \subset \mathbb{R}^m$, the probability $Q(B^u \in A)$ is given by the integral against the standard Gaussian density:

$$Q(B^u \in A) = \frac{1}{(2\pi)^{m/2}} \int_A \exp\left(-\frac{|x|^2}{2}\right) dx\,.$$

Indeed, since B is standard Gaussian under P, it follows that

$$\begin{aligned}
Q(B^u \in A) &= \int_{\{\omega\,:\,B^u(\omega)\in A\}} \exp\left(u \cdot B(\omega) - \frac{1}{2}|u|^2\right) dP(\omega) \\
&= \int_{\{\omega\,:\,B(\omega)-u\in A\}} \exp\left(u \cdot B(\omega) - \frac{1}{2}|u|^2\right) dP(\omega) \\
&= \frac{1}{(2\pi)^{m/2}} \int_{\{x\,:\,x-u\in A\}} \exp\left(u \cdot x - \frac{1}{2}|u|^2 - \frac{1}{2}|x|^2\right) dx \\
&= \frac{1}{(2\pi)^{m/2}} \int_{\{x\,:\,x-u\in A\}} \exp\left(-\frac{|x-u|^2}{2}\right) dx \\
&= \frac{1}{(2\pi)^{m/2}} \int_A \exp\left(-\frac{|y|^2}{2}\right) dy\,,
\end{aligned}$$

showing that B^u has a standard Gaussian distribution under Q.

Appendix B.2 Reweighting

Clearly, by the definition of Q, it holds that

$$\mathbf{E}[f(X)] = \mathbf{E}_Q\left[f(X)\exp\left(-u\cdot B^u(\omega) - \frac{1}{2}|u|^2\right)\right] \tag{A5}$$

for any bounded and measurable function $f\colon \mathbb{R}^n \to \mathbb{R}$, where $\mathbf{E}[\cdot] = \mathbf{E}_P[\cdot]$ denotes the expectation with respect to the reference measure P. Now let

$$X^u(\omega) = b^u + \sigma B(\omega)\,.$$

Since the distribution of the pair (X^u, B) under P is the same as the distribution of the pair (X, B^u) with $X = b^u + \sigma B^u$ under Q, the reweighting identity (A5) entails that

$$\mathbf{E}[f(X)] = \mathbf{E}\left[f(X^u)\exp\left(-u\cdot B(\omega) - \frac{1}{2}|u|^2\right)\right]\,, \tag{A6}$$

with $\mathbf{E}[B] = 0$. Equation (A6) is the finite dimensional analogue of the reweighting identity that has been used to convert the Donsker-Varadhan formula (22) into its final form (30).

Remark 8. *If $\sigma(\cdot)$ in the SDE (26) is square and invertible, then an alternative derivation of Girsanov's Theorem and (29) can be based on the Euler-Maruyama discretization of the SDE and a change of measure for the corresponding Markov chain.*

Appendix C Proof of Theorem 3

The proof is based on the Feynman-Kac formula and Itô's Lemma. Here we give only a sketch of the proof and leave aside all technical details regarding the regularity of solutions of partial differential equations, for which we refer to [18, Sec. VI.5]. Recall the definition

$$\Psi(x,t) = \mathbf{E}_P\left[\exp\left(-\int_t^\tau f(X_s)\,ds - g(X_\tau)\right)\middle| X_t = x\right]\,.$$

By the Feynman-Kac formula, function Ψ solves the parabolic boundary value problem

$$\begin{aligned}(\mathcal{A} - f)\,\Psi &= 0\,, \qquad \text{for } (x,t)\in O\times[0,T)\,,\\ \Psi &= \exp(-g)\,, \quad \text{for } (x,t)\in D^+\,.\end{aligned} \tag{A7}$$

on the domain $D = O\times[0,T)$ where $D^+ = (\partial O\times[0,T))\cup(O\times\{T\})$ denotes the terminal set of the augmented (control-free) process (X_t, t) and

$$\mathcal{A} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma\sigma^T\colon \nabla_x^2 + b\cdot\nabla_x$$

is its infinitesimal generator under the probability measure P. By construction, the stopping time τ is bounded, and we assume that Ψ is of class $C^{2,1}$ on D, and continuous and uniformly bounded away from zero on the closure $\overline{D}$. Now let us define the process

$$\zeta_s^u = -\log\Psi(X_s^u, s)\,,$$

with X_s^u given by (26). Then, using Itô's Lemma (e.g. [36, Thm. 4.2.1]) and introducing the shorthands

$$\Psi_s^u = \Psi(X_s^u, s)\,,\ b_s^u = b(X_s^u, s)\,,\ \sigma_s^u = \sigma(X_s^u)\,,$$

we see that $(\zeta_s^u)_{0\le s<\tau}$ satisfies the SDE

$$\begin{aligned}
d\zeta_s^u &= -\frac{\partial}{\partial t}\log\Psi_s^u\,ds - \nabla_x \log\Psi_s^u\cdot(b_s^u+\sigma_s^u u_s)\,ds \\
&\quad -\frac{1}{2}\sigma_s^u(\sigma_s^u)^T : \nabla_x^2(\log\Psi_s^u)ds - \left((\sigma_s^u)^T\nabla_x\log\Psi_s^u\right)\cdot dB_s^u \\
&= -\left[\frac{\mathcal{A}\Psi_s^u}{\Psi_s^u} + \left((\sigma_s^u)^T\frac{\nabla_x\Psi_s^u}{\Psi_s^u}\right)\cdot u_s - \frac{1}{2}\frac{|(\sigma_s^u)^T\nabla_x\Psi_s^u|^2}{(\Psi_s^u)^2}\right]ds - \left((\sigma_s^u)^T\frac{\nabla_x\Psi_s^u}{\Psi_s^u}\right)\cdot dB_s^u \\
&= -\left[f(X_s^u,s) + \left((\sigma_s^u)^T\frac{\nabla_x\Psi_s^u}{\Psi_s^u}\right)\cdot u_s - \frac{1}{2}\frac{|(\sigma_s^u)^T\nabla_x\Psi_s^u|^2}{(\Psi_s^u)^2}\right]ds - \left((\sigma_s^u)^T\frac{\nabla_x\Psi_s^u}{\Psi_s^u}\right)\cdot dB_s^u\,.
\end{aligned}$$

In the last equation, we have used that the first equation in (A7) holds in the interior of the bounded domain D, i.e. for $s<\tau$. Choosing $u_s = u_s^*$ for $0\le s<\tau$ to be the optimal control

$$u_s^* = \sigma(X_s^{u^*},s)^T\nabla_x\log\Psi(X_s^{u^*},s)$$

as in (34), the last equation can be recast as

$$d\zeta_s^{u^*} = -\left(f(X_s^{u^*},s) + \frac{1}{2}|u_s^*|^2\right)ds - u_s^*\cdot dB_s^{u^*}\,.$$

Similar to (28), if we introduce

$$Z_{s,\tau}^{u^*} = \int_s^\tau u_r^*\cdot dB_r^{u^*} + \frac{1}{2}\int_s^\tau |u_r^*|^2 dr\,,$$

then $Z_{0,\tau}^{u^*} = Z_\tau^{u^*}$ and we have

$$d\zeta_s^{u^*} = -f(X_s^{u^*},s)ds - dZ_{s,\tau}^{u^*}\,.$$

As a consequence, using the continuity of the process as $s\downarrow 0$,

$$\zeta_\tau^{u^*} = \zeta_0^{u^*} - Z_\tau^{u^*} - \int_0^\tau f(X_s^{u^*},s)ds\,. \tag{A8}$$

By definition of ζ_s^u, the initial value $\zeta_0^{u^*} = -\log\Psi(X_0^{u^*},0) = -\log\Psi(x,0)$ is deterministic. Moreover $\zeta_\tau^{u^*} = -\log\Psi(X_\tau^{u^*},\tau) = g(X_\tau^{u^*})$, which in combination with (A8) yields

$$-\log\Psi(x,0) = g(X_\tau^{u^*}) + \int_0^\tau f(X_s^{u^*},s)ds - Z_\tau^{u^*}\,. \tag{A9}$$

Rearranging the terms in the last equation, we find

$$\Psi(x,0) = \exp\left(-Z_\tau^{u^*} - \int_0^\tau f(X_s^{u^*},s)ds - g(X_\tau^{u^*})\right)\,, \tag{A10}$$

with probability one, which yields the assertion of Theorem 3.

References

1. R. Banisch and C. Hartmann. A sparse Markov chain approximation of LQ-type stochastic control problems. *Math. Control Relat. Fields*, 6: 363–389, 2016.
2. C. Bender, J. Steiner. *Least-Squares Monte Carlo for BSDEs*. In: Carmona et al. (Eds.), Numerical Methods in Finance, Springer (2012), pp. 257-289, 2012.
3. D.P. Bertsekas. Approximate policy iteration: a survey and some new methods. *J. Control Theory Appl.*, 9(3):310–355, 2011.

4. M. Boué and P. Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.*, 26(4):1641–1659, 1998.
5. G. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60:2721–2726, 1999.
6. P. Dai Pra., L. Meneghini, and W.J. Runggaldier. Connections between stochastic control and dynamic games. *Math. Control Signals Systems*, 9:303–326, 1996.
7. J.D. Deuschel and D.W. Stroock. *Large Deviations.* Academic Press, New York, 1989.
8. E. Darve, D. Rodriguez-Gomez, and A. Pohorille. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.*, 128:144120, 2008.
9. P. Van Carey. *Statistical Thermodynamics and Microscale Thermophysics.* Cambridge University Press, Cambridge, 1999.
10. L. Delle Site, G. Ciccotti, and C. Hartmann. Partitioning a macroscopic system into independent subsystems. *J. Stat. Mech. Theor. Exp.*, 2017:083201,2017.
11. A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, Springer-Verlag, New York, 2000.
12. P. Dupuis and R.S. Ellis, *A Weak Conergence Approach to the Theory of Large Deviations*, John Wiley & Sons, New York, 1997.
13. P. Dupuis and K. Spiliopoulos. Large deviations for multiscale diffusion via weak convergence methods. *Stoch. Proc. Appl.*, 122:1947–1987, 2012.
14. P. Dupuis and H. Wang. Importance sampling, large deviations, and differential games. *Stoch. and Stoch. Reports.*, 76:481–508, 2004.
15. P. Dupuis and H. Wang. Subsolutions of an Isaacs equation and efficient schemes for importance sampling. *Math. Oper. Res.*, 32:1–35, 2007.
16. E. Gobet and P. Turkedjiev, *Adaptive importance sampling in least-squares Monte Carlo algorithms for backward stochastic differential equations*, Stoch. Proc. Appl. 127, 1171–1203, 2005
17. N. El Karoui, S. Hamadène, and A. Matoussi. Backward stochastic differential equations and applications. *Appl. Math. Optim.*, 27:267–320, 2008.
18. W. H. Fleming and H. M.Soner. *Controlled Markov processes and viscosity solutions*. Springer, New York, 2006.
19. P. Glasserman. *Monte Carlo Methods in Financial Engineering.* Springer, New York, 2004.
20. J.M. Hammersely, and K.W. Morton. Poor Man's Monte Carlo. *Journal of the Royal Statistical Society*, Series B (16) 23-38, 1954.
21. C. Hartmann, R. Banisch, M. Sarich, T. Badowski, and Ch. Schütte. Characterization of rare events in molecular dynamics. *Entropy*, 16:350–376, 2014.
22. C. Hartmann and C. Schütte. Efficient rare event simulation by optimal nonequilibrium forcing. *J. Stat. Mech. Theor. Exp.*, 2012:P11004, 2012.
23. C. Hartmann, J.C. Latorre, G.A. Pavliotis, and W. Zhang. Optimal control of multiscale systems using reduced-order models. *J. Comput. Dyn.*, 1(2):279–306, 2014.
24. C. Hartmann, C. Schütte, and W. Zhang. Model reduction algorithms for optimal control and importance sampling of diffusions. *Nonlinearity*, 29(8):2298-2326, 2016.
25. C. Hartmann, C. Schütte, M. Weber, and W. Zhang. Importance sampling in path space for diffusion processes with slow-fast variables. *Probab. Theory Related Fields*, DOI: https://doi.org/10.1007/s00440-017-0755-3, 2017.
26. C. Hartmann, O. Kebiri, and L. Neureither. Importance sampling of rare events using least squares Monte Carlo. *Under preparation*, 2018.
27. C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997.
28. R. Jordan and D. Kinderlehrer. An extended variational principle. In: *Partial Differential Equations and Applications*, P. Marcellini, G.G. Talenti, and E. Vesintini (Eds.), Lecture Notes in Pure and Applied Mathematics 177, Marcel Dekker, New York, NY, (1996), pp. 187–200, 1996.
29. B. Lapeyre, E. Pardoux, and R. Sentis. *Méthodes de Monte Carlo pour les équations de transport et de diffusion.* Springer, Berlin, 1998.
30. T. Lelièvre, M. Rousset, G. Stoltz. Computation of free energy profiles with parallel adaptive dynamics. *J. Chem. Phys.*, 126:134111, 2007.
31. T. Lelièvre, M. Rousset, G. Stoltz. Long-time convergence of an adaptive biasing force methods. *Nonlinearity*, 21:1155–1181, 2008.

32. T. Lelièvre and G. Stolz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numerica*, 25:1–160, 2016.
33. H. C. Lie. *On a strongly convex approximation of a stochastic optimal control problem for importance sampling of metastable diffusions.* PhD Thesis, Department of Mathematics & Computer Science, Freie Universität Berlin, 2016.
34. J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, New York, 1999.
35. H. Oberhofer, C. Dellago. Optimum bias for fast-switching free energy calculations. *Comput. Phys. Commun.*, 179:41–45, 2008.
36. B. Øksendal. *Stochastic Differential Equations: An Introduction With Applications.* Springer, Berlin, 2003.
37. M.N. Rosenbluth and A.W. Rosenbluth. Monte Carlo Calculations of the Average Extension of Molecular Chains. *Journal of Chemical Physics*, (23) 356-359, 1955.
38. O. Papaspiliopoulos and G.O. Roberts. Importance sampling techniques for estimation of diffusions models. *University of Warwick, Centre for Research in Statistical Methodology, Working papers*, No.28, 2009.
39. L. Richter. *Efficient statistical estimation using stochastic control and optimization.* MSc Thesis, Department of Mathematics & Computer Science, Freie Universität Berlin, 2016.
40. G.O. Roberts and R.L. Tweedie. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 1996.
41. C. Schütte, S. Winkelmann, and C. Hartmann. Optimal control of molecular dynamics using Markov state models. *Math. Program. Ser. B*, 134:259–282, 2012.
42. E. Vanden-Eijnden and J. Weare. Rare Event Simulation of Small Noise Diffusions. *Commun. Pure Appl. Math.*, 65(12):1770–1803, 2012.
43. W. Zhang, C. Hartmann, and C. Schütte. Effective dynamics along given reaction coordinates, and reaction rate theory. *Faraday Discuss.*, 195:365–394, 2016.
44. W. Zhang, H. Wang, C. Hartmann, M. Weber and C. Schütte. Applications of the cross-entropy method to importance sampling and optimal control of diffusions. *SIAM J. Sci. Comput.*, 36(6):A2654-A2672, 2014.