Variational approach to rare event simulation using least-squares regression

Carsten Hartmann,[1] Omar Kebiri,[1] Lara Neureither,[1] and Lorenz Richter[1,2]

[1] *Institute of Mathematics, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus, Germany*

[2] *Institute of Mathematics, Freie Universität Berlin, Berlin, Germany*

(Dated: 16 April 2019)

We propose an adaptive importance sampling scheme for the simulation of rare events when the underlying dynamics is given by a diffusion. The scheme is based on a Gibbs variational principle that is used to determine the optimal (i.e. zero-variance) change of measure and exploits the fact that the latter can be rephrased as a stochastic optimal control problem. The control problem can be solved by a stochastic approximation algorithm, using the Feynman-Kac representation of the associated dynamic programming equations, and we discuss numerical aspects for high-dimensional problems along with simple toy examples.

When computing small probabilities associated with rare events by Monte Carlo it so happens that the variance of the estimator is of the same order as the quantity of interest. Importance sampling is a means to reduce the variance of the Monte Carlo estimator by sampling from an alternative probability distribution under which the rare event is no longer rare. The estimator must then be corrected by an appropriate reweighting that depends on the likelihood ratio between the two distributions and, depending on this change of measure, the variance of the estimator may easily increase rather than decrease. e.g. when the two probability distributions are (almost) non-overlapping. The Gibbs variational principle links the cumulant generating function (or: free energy) of a random variable with an entropy minimisation principle, and it characterises a probability measure that leads to importance sampling estimators with minimum variance. When the underlying probability measure is the law of a diffusion process, the variational principle can be rephrased as a stochastic optimal control problem, with the optimal control inducing the change of measure that minimises the variance. In this article, we discuss the properties of the control problem and propose a numerical method to solve it. The numerical method is based on a nonlinear Feynman-Kac representation of the underlying dynamic programming equation in terms of a pair of forward-backward stochastic differential equations that can be solved by least-squares regression. At first glance solving a stochastic control problem may be more difficult than the original sampling problem, however it turns out that the reformulation of the sampling problem opens a completely new toolbox of numerical methods and approximation algorithms that can be combined with Monte Carlo sampling in a iterative fashion and thus leads to efficient algorithms.

I. INTRODUCTION

The estimation of small probabilities associated with rare events is among the most difficult problems in computational statistics. Typical examples of rare event probabilities, the precise estimation of which is important, involve protein folding, phase transitions in materials or large-scale atmospheric events, such as extreme heat waves or hurricanes. The hallmark of these rare events is that the average waiting time between the events is orders of magnitude longer than the characteristic timescale of the system—especially the timescale

of the switching event itself—which renders the direct numerical simulation of rare events often infeasible.

We can distinguish between two major classes of sampling techniques: splitting methods such as RESTART[44] or Adaptive Multilevel Splitting[8] that decompose state space, but are still essentially based on the underlying probability distribution, and biasing methods, such as importance sampling[35] or the adaptive biasing force method[7] that enhance the rare events under consideration by perturbing the underlying probability distribution and thus altering the rare events statistics; see Juneja and Shahabuddin[28] for an overview. We should also mention sequential Monte-Carlo[6] that combines both worlds and that can be embedded into a splitting-like framework.

In this article, we focus on the second class of methods, namely importance sampling. Specifically, we consider diffusion processes and quantities that have the form of a cumulant generating function (or: thermodynamic free energy) and which are characterised by a Gibbs variational principle on a suitable subspace of the space of probability measures. The Gibbs variational principle expresses a fundamental duality between cumulant generating functions and relative entropy known as the Donsker-Varadhan principle in large deviations theory.[14] In our case, the variational principle is a constrained entropy minimisation problem, the minimiser of which defines an optimal change of measure that leads to minimum (i.e. zero) variance estimators of the quantity of interest.[24] The connection between the zero variance estimator and the Gibbs variational principle is essentially of theoretical interest, because the normalisation constant of the optimal change of measure depends on the quantity of interest.

In order to turn the Gibbs principle into a workable numerical method, we interpret the variational principle as a stochastic optimal control problem, with the unique optimal control force (or: bias) generating the zero-variance probability measure. Specifically, we propose a reformulation of the semilinear dynamic programming equations of the optimal control problem as a pair of uncoupled forward-backward stochastic differential equations (FBSDE) that can be solved by Monte Carlo.[32] The advantage of the FBSDE approach is that it offers good control of the variance of the resulting estimators at low additional numerical cost. One of the key results of this paper is that the control that is obtained from the solution to the FBSDE acts as a control variate that, when augmented by an additional bias, produces a whole family of zero-variance estimators. We discuss several variants of the FBSDE method, based on a parametric formulation of the least-squares Monte Carlo algorithms by Gobet et al.[22] and a deep learning based algorithm that is due to E et al.[13].

Related work

The idea of exploiting the variational formulation of cumulant generating functions to devise feedback control based importance sampling strategies for rare events goes back to Dupuis and Wang[11] who suggested to approximate the optimal change of measure by vanishing viscosity solutions or subsolutions of the associated dynamic programming equations. The thus obtained change of measure can be shown to converge to the optimal exponential change of measure as the probability of the rare event goes to zero, which is implied by the fact that the zero-viscosity solution to the dynamic programming equation is the associated large deviations rate function of the rare event. As a consequence, the resulting estimators are either asymptotically efficient[43], when the change of measure is based on the exact viscosity solution, or log asymptotically efficient[12], when the viscosity solution is approximated by a subsolution. The development of state-dependent importance sampling schemes, that in the context of diffusion processes can be considered as the small noise limit of the control approach considered in this article, was triggered by the observation that an exponential change of measure based on an exponential tilting with a constant tilting parameter may perform worse than standard Monte Carlo.[20]

The relation between large deviations principles and control has been pointed out quite early in the work by Fleming and co-workers[16–18] and later on in the context of risk-sensitive

control,[9,27,45,46] and we should note that the underlying duality relation has also been exploited to recast certain stochastic control problems as linear elliptic or parabolic boundary value problems[31,38,40,41] or to solve data assimilation problems.[29,30,39] Using FBSDE numerics to solve the dynamic programming equations associated with certain stochastic control problems, similar to the ones considered in this paper, has been recently suggested in Exarchos and Theodorou[15], Huré et al.[25] and Bachouch et al.[1].

Outline of the article

The article is structured a follows: In Section II we explain the basic importance sampling problem for stochastic processes and, in case of a diffusion process, characterise the optimal change of measure in terms of a solution to an optimal control problem. Section III is devoted to the reformulation of the optimal control problem, or more precisely to the reformulation of the associated dynamic programming equation in form of an FBSDE pair; the main result of this section is that we show that there is a family of equivalent FBSDE pairs that lead to zero-variance importance sampling estimators. The numerical discretisation of the FBSDE is discussed in Section IV and illustrated with a few numerical examples in Section V. Conclusions are given in Section VI. The article contains an appendix in which the relation between the optimal change of measure and Doob's h-transform is briefly explained.

II. RARE EVENT SIMULATION

Let $(\Omega, \mathcal{E}, P)$ be a probability space, on which we consider an $\mathbb{R}^d$-valued stochastic process $X = (X_s)_{s \geq 0}$. Suppose that we want to compute a small probability, such as the probability of hitting a set $C \subset \mathbb{R}^d$,

$$\theta = P(X_\tau \in C), \tag{1}$$

where τ is some a.s. finite stopping time $\tau < \infty$. For example, τ may the first hitting time $\tau_{A \cup C}$ of either of the two disjoint sets $A, C \subset \mathbb{R}^d$, in which case θ is the probability to reach C before A, in other words: the committor probability; if τ is the minimum of the first hitting time τ_C of C and a finite time $T \in (0, \infty)$, then θ is the probability that $\tau_C < T$. (We assume throughout that all subsets are measurable.)

We assume that $\theta \ll 1$, and without digging into the details of large deviations theory, we call $X_\tau \in C$ a *rare event*, simply because of this assumption that implies that θ is difficult to compute numerically. To understand why this is the case, consider the Monte Carlo approximation of the parameter θ: given N independent realisations $X(\omega_1), \ldots, X(\omega_N)$ of X,

$$\theta_N = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_C(X_\tau(\omega_i)) \tag{2}$$

is an unbiased estimator of θ that converges a.s. to θ by the law of large numbers. Moreover the variance of the estimator decreases with rate $1/N$ since

$$\mathrm{Var}(\theta_N) = \frac{1}{N}\theta(1 - \theta) \leq \frac{1}{4N}.$$

The last equation reflects the typical Monte Carlo rate of convergence. Nevertheless the relative error (or: relative standard deviation) is unbounded as a function of θ:

$$\delta := \frac{\sqrt{\mathrm{Var}(\theta_N)}}{\mathbf{E}[\theta_N]} \sim \frac{1}{N\sqrt{\theta}} \quad \text{as} \quad \theta \to 0.$$

Here $\mathbf{E}[\cdot]$ denotes the expectation with respect to the probability P. The bottom line is that computing small probabilities such as (1) is difficult, since the number of Monte Carlo samples that is required to obtain an accurate estimate grows with $1/\sqrt{\theta}$.

A. Importance sampling

The idea of importance sampling is to reduce the variance of (2) by drawing the samples from another probability measure, say, Q under which the event is no longer rare. Let Q be absolutely continuous with respect to P, so that the likelihood ratio $\varphi(\omega) = (dQ/dP)(\omega)$ exists. We further assume that $\varphi > 0$ on the set $\{\omega \in \Omega \colon X_\tau(\omega) \in C\}$. Then, letting $\mathbf{E}_Q[\cdot]$ denote the expectation with respect to Q, it holds that

$$P(X_\tau \in C) = \mathbf{E}[\mathbf{1}_C(X_\tau)] = \mathbf{E}_Q[\mathbf{1}_C(X_\tau)\varphi^{-1}]. \tag{3}$$

The last equality gives rise to the importance sampling estimator

$$\hat{\theta}_N = \frac{1}{N}\sum_{i=1}^N \mathbf{1}_C(X_\tau(\hat{\omega}_i))\varphi^{-1}(\hat{\omega}_i), \tag{4}$$

where the realisations $X(\hat{\omega}_i)$ or $\hat{\omega}_i$, respectively, are independent draws from the new probability measure Q. It is easy to see that the estimator (4) is unbiased under Q, i.e. $\mathbf{E}_Q[\hat{\theta}_N] = \theta$. Moreover choosing Q, such that

$$\frac{dQ}{dP} = \frac{\mathbf{1}_C(X_\tau)}{\theta},$$

the resulting importance sampling estimator $\hat{\theta}_N$ has zero variance under Q, i.e., $\mathrm{Var}_Q(\hat{\theta}_N) = 0$. We call the change of measure $Q = Q^*$ that reduces the variance to zero – and gives the correct answer already for $N = 1$ – the *optimal change of measure*.

Note, however, that the optimal change of measure depends on the sought quantity θ, which is not surprising as it completely removes the randomness from the estimator, but which renders the result somewhat useless. Further notice that $Q^*(\cdot) = P(\cdot|X_\tau \in C)$, in other words, the optimal change of measure is given by conditioning the original measure P on the rare event $X_\tau \in C$.

We will later on discuss the question how to devise approximations to the optimal change of measure.

B. Importance sampling in path space

Throughout the rest of this paper we assume that X is governed by a stochastic differential equation (SDE)

$$dX_s = b(X_s)ds + \sigma(X_s)dB_s, \quad X_0 = x, \tag{5}$$

where the coefficients b and σ are such that (5) has a unique strong solution. For simplicity we further assume that $(\sigma\sigma^T)(\cdot)\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}$ has a uniformly bounded inverse. Our standard example will be a non-degenerate diffusion in an energy landscape,

$$dX_s = -\nabla U(X_s)ds + \sigma dB_s, \quad X_0 = x, \tag{6}$$

with smooth potential energy U and $\sigma > 0$ constant.

We will now generalise the previous considerations to more general properties and functionals of (5). To this end, let $O \subset \mathbb{R}^d$ denote an open and bounded set with smooth

boundary ∂O such that $C \subset \partial O$; we define

$$\tau = \inf\{t > 0 \colon X_t \notin O\} \tag{7}$$

to be the first exit time of the set O and call W the continuous functional

$$W_\tau(X) = \int_0^\tau f(X_s)\, ds + g(X_\tau)\,, \tag{8}$$

of X where f, g are bounded and sufficiently smooth, real valued functions. Our aim is to estimate the free energy

$$F(x) = -\log \mathbf{E}_x \left[\exp(-W_\tau)\right] \tag{9}$$

that can be considered a scaled version of the cumulant generating function of W_τ, where the expectation is understood with respect to the realisations of the Brownian motion $B = (B_s)_{s \geq 0}$ in the uncontrolled SDE (5) for a given initial condition $X_0 = x$. Now, let X^u be the solution of the controlled SDE

$$dX_s^u = (b(X_s^u) + \sigma(X_s^u)u_s)\, ds + \sigma(X_s^u)dB_s\,, \tag{10}$$

with initial data $X_0^u = x$. By Jensen's inequality,[23]

$$F(x) \leq \mathbf{E}[W_\tau^u] + H(Q|P)\,, \tag{11}$$

where $H(Q|P)$ denotes the relative entropy or Kullback-Leibler divergence between the probability measures Q and P, restricted to the history $\mathcal{F}_\tau$ of the stopped process, and we have introduced the shorthands $Q = Q^u$ and $W_\tau^u = W_\tau(X^u)$ to denote quantities generated by the controlled SDE (10).

The inequality (11) is the basis for the famous Gibbs variational principle—also known as the Donsker-Varadhan principle in its dual form—that relates the free energy with the (relative) entropy. It can be shown that equality in (11) is attained if and only if Q belongs to the exponential family, with $dQ \propto \exp(-W_\tau)dP$.

Now, informally, Girsanov's Theorem states that

$$\mathbf{E}_x[\exp(-W_\tau)] = \mathbf{E}_x[\exp(L_\tau^u - W_\tau^u)] \tag{12}$$

where the expectation on the right hand side is taken over all realisations of the controlled process, and

$$L_\tau^u = -\int_0^\tau u_s \cdot dB_s - \frac{1}{2}\int_0^\tau |u_s|^2\, ds\,, \tag{13}$$

denotes the log likelihood ratio between the realisations of the controlled SDE (10) and the uncontrolled SDE (5); see Ikeda and Watanabe [26], Ch. IV.4 or Appendix C for an informal derivation of the relation (12).

The following variational characterisation of the free energy is a straightforward generalisation of the previous considerations and characterises the optimal change of measure for the free energy from P to Q in terms of the solution to an optimal control problem.[5,24]

Theorem II.1 (Hartmann et al. [24]) *Assuming sufficient regularity of the coefficients f, g, b, σ and the boundary of the set $O \subset \mathbb{R}^d$, the free energy is the value function of the following optimal control problem: minimise*

$$J(u) = \mathbf{E}\left[\int_0^\tau \left(f(X_s^u) + \frac{1}{2}|u_s|^2\right) ds + g(X_\tau^u)\right] \tag{14}$$

where X^u is the solution of the controlled SDE (10) with $X_0^u = x$. That is, $F(x) = V(x)$

where $V = \min_u J(u)$. The minimiser u^* is unique and given by the feedback law

$$u_s^* = -\sigma(X_s^u)^T \nabla V(X_s^u).$$

Moreover, with probability one,

$$\mathbf{E}_x[\exp(-W_\tau)] = \exp\left(L_\tau^{u^*} - W_\tau^{u^*}\right). \tag{15}$$

In other words, the optimal control u^* generates a path space measure $Q = Q^*$ that yields a zero-variance importance sampling estimator via the identity (12). We refer to Appendix B for a formal derivation of the underlying stochastic optimal control problem.

Importance sampling estimators

In practice, one will not have access to the optimal control and an exact simulation of the process X^u, but rather use a numerical approximation. In this case, the variance of the importance sampling estimator will be small, but not zero. Given N statistically independent numerical approximations $\hat{X}^{u,1}, \ldots, \hat{X}^{u,N}$ of X^u, all starting at $\hat{X}_0^{u,i} = x$, an estimator for the free energy (9) that replaces (15) is

$$\hat{F}_N(x) = -\log \hat{\Psi}_N(x), \tag{16}$$

with

$$\hat{\Psi}_N(x) = \frac{1}{N} \sum_{i=1}^{N} \exp\left(\hat{L}_\tau^{u,i} - \hat{W}_\tau^{u,i}\right) \tag{17}$$

being an unbiased estimator of the moment generating function $\Psi = \mathbf{E}[\exp(-W_\tau)]$. Here $\hat{L}_\tau^{u,i}$ and $\hat{W}_\tau^{u,i}$ denote the numerical approximations of the log likelihood L_τ^u and the path functional W_τ^u. Note that, even though the estimator (17) is unbiased, the estimator for F is not as it follows by Jensen's inequality that

$$\mathbf{E}\left[\hat{F}_N(x)\right] \geq -\log \mathbf{E}\left[\hat{\Psi}_N(x)\right] = F(x). \tag{18}$$

Another biased estimator of F is

$$\tilde{F}_N(x) = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{W}_\tau^{u,i} - \hat{L}_\tau^{u,i}\right), \tag{19}$$

where the bias depends on how close u is to the optimal control u^*. If u is a good approximation of u^*, this estimator may turn out to be advantageous in terms of variance. Note that, by the central limit theorem, both (17) and (19) are asymptotically normal.

III. NONLINEAR FEYNMAN-KAC FORMULA

The aim of this section is to give an alternative characterisation of the dual optimal control problem in Theorem II.1 that (a) leads to a practical stochastic approximation algorithm for computing the optimal control, even for high-dimensional problems, and (b) that gives rise to an interpretation of the optimal change of measure in the context of control variates that may have implications for the numerical implementation of adaptive importance sampling schemes.

Applying the dynamic programming principle (see e.g. Fleming and Soner [19], Sec. IV.5) to the stochastic control problem (14), it follows that the value function $V = \min_u J(u)$ or,

equivalently, the free energy F solves the stationary HJB equation

$$LV + h(x, \sigma^T \nabla V) = 0, \quad x \in O$$
$$V = g, \quad x \in \partial O,$$

(20)

with the generator

$$L = \frac{1}{2}\sigma\sigma^T : \nabla^2 + b \cdot \nabla$$

(21)

and the nonlinearity

$$h(x, z) = -\frac{1}{2}|z|^2 + f(x).$$

(22)

The boundary value problem (20)–(22) is the straighforward generalisation of the dynamic programming equation (B5) to path functionals of the form (8), and under suitable regularity assumptions, it can be shown[19] that it has a classical solution $V \in C^2(O) \cap C(\partial O)$.

We will now reformulate the HJB equation as an equivalent system of forward-backward stochastic differential equations (FBSDE) that will be the basis of the numerical approximation of the optimal change of measure. To this end, we define the processes

$$Y_s = V(X_s), \quad Z_s = \sigma(X_s)^T \nabla V(X_s)$$

(23)

Now, by Ito's formula, the value function $V(X_s)$ satisfies

$$dV(X_s) = (LV)(X_s) + (\sigma^T \nabla V)(X_s) \cdot dB_s,$$

(24)

which upon inserting (20) and (23) yields the following backward stochastic differential equation (BSDE)

$$dY_s = -h(X_s, Z_s)ds + Z_s \cdot dB_s,$$

(25)

for the pair $(Y, Z) = (Y_s, Z_s)_{s \geq 0}$. By construction, the equation comes with the terminal condition

$$Y_\tau = g(X_\tau),$$

(26)

where X denotes the solution to the uncontrolled forward SDE (5). Note that, by definition, Y is continuous and adapted to X, and Z is predictable and square integrable, in accordance with the interpretation of Z_s as a control variable. Further note that (25) must be understood as a *backward* SDE rather than a *time-reversed* SDE, since, by definition, Y_s at time $s < \tau$ is measurable with respect to the filtration generated by the Brownian motion $(B_r)_{0 \leqslant r \leqslant s}$, whereas a time-reversed version of Y_s would depend on B_τ via the terminal condition $Y_\tau = g(X_\tau)$, which would require a larger filtration.

By exploiting the specific form of the nonlinearity (22) that appears as the driver h in the backward SDE (25) and the fact that the forward process X is independent of (Y, Z), we obtain the following representation of the solution to the dynamic programming equation (20)–(22):

$$dX_s = b(X_s)ds + \sigma(X_s)\, dB_s$$
$$dY_s = \left(\frac{1}{2}|Z_s|^2 - f(X_s)\right) ds + Z_s \cdot dB_s,$$

(27)

with boundary data

$$X_0 = x, \quad Y_\tau = g(X_\tau).$$

(28)

The solution to (27)–(28) now is a triplet (X, Y, Z), and since Y is adapted, it follows that Y_0 is a deterministic function of the initial data $X_0 = x$ only. Since g is bounded, the results in Kobylanski [34] entail existence and uniqueness of the FBSDE (27); see also Delbaen et al. [10] for the case of unbounded terminal cost. As a consequence, $Y_0 = V(x)$ equals the value function of our control problem.

Remark 1 *A remark on the role of the control variable Z_s in the BSDE is in order. In (25), let $h = 0$ and consider a random variable ξ that is square-integrable and $\mathcal{F}_\tau$-measurable where $\mathcal{F}_s$ is the σ-Algebra generated by $(B_r)_{0 \leqslant r \leqslant s}$. Ignoring the measurability for a second, a pair of processes (Y, Z) satisfying*

$$dY_s = Z_s \cdot dB_s, \quad Y_\tau = \xi. \tag{29}$$

is $(Y, Z) \equiv (\xi, 0)$, but then Y is not adapted unless the terminal condition ξ is a.s. constant, because Y_t for any $t < \tau$ is not measurable with respect to $\mathcal{F}_s \subset \mathcal{F}_\tau$. An adapted version of Y can be obtained by replacing $Y_t = \xi$ by its best approximation in L^2, i.e. by the projection $Y_t = \mathbf{E}[\xi | \mathcal{F}_t]$. Since the thus defined process Y is a martingale with respect to our filtration, the martingale representation theorem asserts that Y_t must be of the form

$$Y_t = \mathbf{E}[\xi] + \int_0^t \tilde{Z}_s \cdot dB_s, \tag{30}$$

for some unique, predictable process $\tilde{Z}$. Subtracting the last equation from $Y_\tau = \xi$ yields

$$Y_t = \xi - \int_t^\tau \tilde{Z}_s \cdot dB_s, \tag{31}$$

or, equivalently,

$$dY_t = \tilde{Z}_t \cdot dB_t, \quad Y_\tau = \xi. \tag{32}$$

Hence $Z_s = \tilde{Z}_s$ in (29) is indeed a control variable that makes Y adapted.

Remark 2 *The above setting includes cases such as exit probabilities $P(\tau < T)$, in which case the free energy becomes explicitly time-dependent via the initial conditions $X_t = x$. We only need to replace $O \subset \mathbb{R}^d$ by $O \times [0, T) \subset \mathbb{R}^d \times [0, \infty)$ and X by an augmented process $\tilde{X}$ with $\tilde{X}_t = (X_t, t)$ that includes time as an extra state variable. Accordingly, the elliptic operator L must be replaced by the parabolic operator $\tilde{L} = \partial / \partial t + L$.*

A. From importance sampling to control variates

The role of the process Z in the FBSDE representation of the dynamic programming equation is not only to guarantee that Y in (27) is adapted, so that $Y_0 = V(x, 0)$ is the value function, but it can be literally interpreted as a control since $Z_s = \sigma(X_s)^T \nabla V(X_s)$, even though it is evaluated along the uncontrolled process X rather than the controlled process X^u.

We will now show that the control Z_s plays the role of a control variate that produces a zero-variance estimator.

Proposition III.1 *Consider the solution (X, Y, Z) of the FBSDE (27)–(28). Further let*

$$L_\tau^Z = -\int_0^\tau Z_s \cdot dB_s - \frac{1}{2} \int_0^\tau |Z_s|^2 \, ds$$

Then, with probability one,

$$\mathbf{E}_x \left[\exp\left(-W_\tau\right) \right] = \exp\left(-L_\tau^Z - W_\tau\right)$$

or, equivalently,

$$F(x) = L_\tau^Z + W_\tau \,,$$

where $F(x) = Y_0$ is the free energy (9).

Proof: Using (27)–(28), L_τ^Z can be recast as

$$L_\tau^Z = -\int_0^\tau Z_s \cdot dB_s - \frac{1}{2}\int_0^\tau |Z_s|^2 \, ds$$

$$= Y_0 - \int_0^\tau f(X_s) \, ds - g(X_\tau)$$

where we have used that $Y_\tau = g(X_\tau)$. Therefore, using the identification of Y_0 with the free energy $F(x)$, we have

$$\exp\!\big(-L_\tau^Z - W_\tau\big) = \exp(-F(x)) \,,$$

which holds with probability one. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

B. Importance sampling within control variates

Even though the importance sampling and the control variate based estimators look very similar, there is an important difference, in that $Z_s = \sigma(X_s)^T \nabla V(X_s)$ is a function of the uncontrolled rather than the controlled process. Thus the second approach does not involve a change of measure which may be advantageous when the existence of the Radon-Nikodym derivative is not guaranteed, which, for example, may be the case when the stopping time τ is either unbounded or can become very large with a non-negligible probability. (Note that the controlled process need not be simulated at all.)

Yet, in some situations it may be difficult to sample the terminal condition $g(X_\tau)$ by forward trajectories, in which case it may be advantageous to use importance sampling, either instead of or within the control variate scheme. To better understand the relation between the two approaches we do a change of drift in the FBSDE, so that the associated HJB equation remains the same. Specifically, consider a change of drift of the form

$$b \mapsto b + \sigma v \tag{33}$$

for some adapted process $v = (v_s)_{s\geq 0}$ that may or may not depend on the state of the process $X^v = (X_s^v)_{s\geq 0}$ with the new drift. Under this change of drift, using the identification $Z_s^v = \sigma(X_s^v)^T \nabla V(X_s^v)$, the original FBSDE (27) turns into

$$\begin{aligned} dX_s^v &= (b(X_s^v) + \sigma(X_s^v)v_s)\, ds + \sigma(X_s^v)\, dB_s \\ dY_s^v &= -h^v(X_s^v, Y_s^v, Z_s^v)\, ds + Z_s^v \cdot dB_s \,, \end{aligned} \tag{34}$$

with the driver

$$h^v(x, y, z) = -\frac{1}{2}|z|^2 - z \cdot v + f(x) \tag{35}$$

and boundary data (28), with X replaced X^v. It can be easily checked that (34)–(35) and (27) represent the same HJB equation (20).

The change of drift furnishes an exponential change of measure in the free energy functional. We will now show that, for any reasonable choice of an adapted control v, say, bounded and continuous, *every* estimator of the form

$$\mathbf{E}_x[\exp\left(-W_\tau\right)] = \mathbf{E}_x\big[\exp\big(\langle v, Z^v\rangle - L_\tau^H - W_\tau^v\big)\big] \,, \tag{36}$$

with $H_s = Z_s^v$ has zero variance where the expectation on the right hand side is taken over the realisations of the FBSDE (34) with initial conditions $X_0^v = x$, and

$$\langle v, Z^v \rangle = \int_0^\tau v_s \cdot Z_s^v \, ds \,. \tag{37}$$

Proposition III.2 *Let v be adapted and such that the FBSDE (34) with driver (35) has a unique strong solution. Then, with probability one,*

$$\mathbf{E}_x \left[\exp \left(-W_\tau \right) \right] = \exp \left(\langle v, Z^v \rangle - L_\tau^H - W_\tau^v \right) \,. \tag{38}$$

Proof: The argument is essentially the same as in the proof of Proposition III.1. Substituting the expressions for $L_\tau^H - \langle v, Z^v \rangle$ in the backward part of the FBSDE (34), we conclude that

$$L_\tau^H - \langle v, Z^v \rangle = Y_0 - \int_0^\tau f(X_s^v) \, ds - g(X_\tau^v)$$
$$= Y_0^v - W_\tau^v \,.$$

Thus, almost surely,

$$\exp \left(\langle v, Z^v \rangle - L_\tau^H - W_\tau^v \right) = \exp(-Y_0^v) \,,$$

where $Y_0^v = Y_0 = F(x)$, since (34)–(35) is a Feynman-Kac representation of the HJB equation (20). $\qquad\square$

Hence we can change the drift of the forward SDE by modifying the control, without affecting the variance of the free energy estimator. Having a zero-variance estimator is of course only useful under the assumption that it is possible to solve the BSDE associated with (27) or (34), and changing the drift is also a means to reduce the variance of the numerical scheme for the BSDE. Similar ideas along these lines have been suggested by Bender and Moseler[3] who use a change of the drift, with the aim of reducing the variance of the BSDE simulation.

More importance sampling estimators

Along the lines of the considerations in Section II B, we define the standard (biased) estimator for $G := Y_0^v$ as

$$\hat{G}_N(x) = -\log \Phi_N(x) \,, \tag{39}$$

with

$$\hat{\Phi}_N(x) = \frac{1}{N} \sum_{i=1}^N \exp \left(\langle v, \hat{Z}^{v,i} \rangle_N - \hat{L}_\tau^{H,i} - \hat{W}_\tau^{v,i} \right) \,. \tag{40}$$

Here $\hat{L}_\tau^{H,i}$ and $\hat{W}_\tau^{v,i}$ stand for the discretisations of L_τ^H and W_τ^v, and the bilinear term $\langle v, \hat{Z}^{v,i} \rangle_N$ denotes the numerical approximations of the scalar product (37) by a suitable quadrature rule. If an accurate approximation of the control Z is available, another biased estimator of G that may have a smaller variance than (39) is

$$\tilde{G}_N(x) = \frac{1}{N} \sum_{i=1}^N \left(\hat{W}_\tau^{v,i} + \hat{L}_\tau^{H,i} - \langle v, \hat{Z}^{v,i} \rangle_N \right) \,. \tag{41}$$

Note that due to the occurence of the bilinear term, none of the estimators will in general be unbiased for fixed N.

IV. LEAST-SQUARES REGRESSION

We now discuss the numerical discretisation of (27) and (34). The fact that both FBSDE are decoupled implies that they can be discretised by an explicit time-stepping scheme. Specifically, we discuss two different approaches: a Monte Carlo approach that is based on a backward iteration that involves the numerical computation of conditional expectations using least-squares and that was first suggested by Gobet et al.[22] and later on refined by several authors[2,4,21], and a deep learning method that seeks to approximate the BSDE solution (Y, Z) by a neural network with a quadratic loss function, as suggested by E et al. [13]. The convergence of the numerical schemes for an FBSDE with quadratic nonlinearities in the driver has been analysed by Turkedjiev [42].

For the ease of notation, we confine our discussion to the FBSDE (27) and then comment on the difference to (34) whenever necessary. Thus consider the Euler-Maruyama discretisation

$$
\begin{aligned}
\hat{X}_{n+1} &= \hat{X}_n + \Delta t\, b(\hat{X}_n) + \sqrt{\Delta t}\, \sigma(\hat{X}_n)\xi_{n+1} \\
\hat{Y}_{n+1} &= \hat{Y}_n - \Delta t\, h(\hat{X}_n, \hat{Y}_n, \hat{Z}_n) + \sqrt{\Delta t}\, \hat{Z}_n \cdot \xi_{n+1}\,,
\end{aligned}
\tag{42}
$$

of (27) where $(\xi_i)_{i \geqslant 1}$ is an i.i.d. sequence of normalised Gaussian random variables and $(\hat{X}_n, \hat{Y}_n, \hat{Z}_n)$ denotes the numerical discretisation of $(X_{t_n}, Y_{t_n}, Z_{t_n})$.

To fix notation, we denote by $\eta = \inf\{n > 0 \colon \hat{X}_n \notin O\}$ the discrete-time approximation to τ, such that $\tau \approx \eta \Delta t$. Further let $n_{\max}$ be the maximum iteration number, with $T_{\max}/\Delta t$ where $T_{\max}$ is the maximum simulation time that should be chosen sufficiently large, so that either $Q(\eta < n_{\max})$ or $P(\eta < n_{\max})$ are close to one (say, between 0.9 and 1), depending on whether the controlled or uncontrolled forward process is simulated.

A. Parametric least-squares Monte Carlo

The least-squares Monte Carlo (LSMC) scheme is based on a parametric representation

$$
V_K(x) = \sum_{k=1}^{K} \alpha_k \phi_k(x)\,, \quad \alpha_k \in \mathbb{R}\,,
\tag{43}
$$

of the value function V (or the free energy F) as a linear combination of finitely many basis functions $\phi_1, \ldots, \phi_K \colon \mathbb{R}^n \to \mathbb{R}$. We assume that the ϕ_k are continuously differentiable, so that we can express the control by the gradient of V_K.

Now let us introduce the shorthand

$$
\mathbf{E}[\cdot | \hat{X}_n] = \mathbf{E}[\cdot | \hat{\mathcal{F}}_n]
$$

for the conditional expectation with respect to the σ-algebra $\hat{\mathcal{F}}_n = \sigma(\{\hat{B}_k \colon 0 \leqslant k \leqslant n\})$ that is generated by the discrete Brownian motion $\hat{B}_n := \sqrt{\Delta t} \sum_{i \leqslant n} \xi_i$. By definition, the continuous-time process (X_s, Y_s, Z_s) is adapted to the filtration generated by $(B_r)_{0 \leqslant r \leqslant s}$. For the discretised process, this implies (cf. Remark 1)

$$
\hat{Y}_n = \mathbf{E}\big[\hat{Y}_n | \hat{X}_n\big]
\tag{44}
$$

so that, with (42),

$$\hat{Y}_n = \mathbf{E}\big[\hat{Y}_{n+1} + \Delta t\, h(\hat{X}_n, \hat{Y}_n, \hat{Z}_n)|\hat{X}_n\big] \tag{45}$$

using that $\hat{Z}_n$ is independent of ξ_{n+1}. In order to compute $\hat{Y}_n$ from $\hat{Y}_{n+1}$, it is convenient to replace $(\hat{Y}_n, \hat{Z}_n)$ on the right hand side by $(\hat{Y}_{n+1}, \hat{Z}_{n+1})$, so that we end up with the fully explicit time stepping scheme

$$\hat{Y}_n := \mathbf{E}\big[\hat{Y}_{n+1} + \Delta t\, h(\hat{X}_n, \hat{Y}_{n+1}, \hat{Z}_{n+1})|\hat{X}_n\big], \tag{46}$$

which is equivalent to (45) up to terms of order $(\Delta t)^2$.

Note that we can use the identification of $-Z$ with the control and replace $\hat{Z}_{n+1}$ in the last equation by

$$\hat{Z}_{n+1} = \sigma(\hat{X}_{n+1})^T \nabla V_K(\hat{X}_{n+1}), \tag{47}$$

where V_K is given by the parametric ansatz (43).

1. Conditional expectation

We next address the question how to compute the conditional expectations with respect to $\hat{\mathcal{F}}_n$. To this end, we recall that the conditional expectation can be characterised as a best approximation in L^2:

$$\mathbf{E}\big[S|\hat{X}_n\big] = \underset{Y \in L^2,\, \hat{\mathcal{F}}_n\text{-measurable}}{\operatorname{argmin}} \mathbf{E}[|Y - S|^2].$$

(Hence the name *least-squares Monte Carlo*.) Here measurability with respect to $\hat{\mathcal{F}}_n$ means that $(\hat{Y}_n, \hat{Z}_n)$ can be expressed as functions of $\hat{X}_0, \hat{X}_1, \ldots, \hat{X}_n$. In view of (23), this suggests the approximation scheme

$$\hat{Y}_n \approx \underset{Y=Y(\hat{X}_n)}{\operatorname{argmin}} \frac{1}{M} \sum_{m=1}^{M} \big|Y - b_n^{(m)}\big|^2, \tag{48}$$

with the shorthand

$$b_n^{(m)} = \hat{Y}_{n+1}^{(m)} + \Delta t\, h\big(\hat{X}_n^{(m)}, \hat{Y}_{n+1}^{(m)}, \hat{Z}_{n+1}^{(m)}\big). \tag{49}$$

Here the superscript in parentheses is used to label the M independent realisations of the forward process, $\hat{X}$, the resulting values for the backward process,

$$\hat{Y}_{n+1}^{(m)} = \sum_{k=1}^{K} \alpha_k(t_{n+1}) \phi_k\big(\hat{X}_{n+1}^{(m)}\big), \tag{50}$$

and the control,

$$\hat{Z}_{n+1}^{(m)} = \sigma\big(\hat{X}_{n+1}^{(m)}\big)^T \sum_{k=1}^{K} \alpha_k(t_{n+1}) \nabla \phi_k\big(\hat{X}_{n+1}^{(m)}\big). \tag{51}$$

At the terminal time $n_{\max}$, the data are determined by

$$\hat{Y}_{n_{\max}}^{(m)} = g\big(\hat{X}_{n_{\max}}^{(m)}\big) \tag{52}$$

and

$$\hat{Z}_{n_{\max}}^{(m)} = \sigma\big(\hat{X}_{n_{\max}}^{(m)}\big)^T \nabla g\big(\hat{X}_{n_{\max}}^{(m)}\big), \tag{53}$$

where only those realisations are taken into account that have *not* yet reached the boundary, i.e. ceased to exist.

2. LSMC algorithm

The unknown coefficients α_k have to be computed in every iteration step which makes them functions of time (i.e. $\alpha_k = \alpha_{k,n}$), even though the value function is not explicitly time dependent. We call $\hat{\alpha}_n = (\alpha_{1,n}, \dots, \alpha_{K,n})$ the vector of the unknowns, so that the least-squares problem that has to be solved in the n-th step of the backward iteration is of the form

$$\hat{\alpha}_n = \underset{\alpha \in \mathbb{R}^K}{\operatorname{argmin}} |A_n \alpha - b_n|^2 \,, \tag{54}$$

with coefficients

$$A_n = \Big(\phi_k\big(\hat{X}_n^{(m)}\big)\Big)_{m=1,\dots,M;k=1,\dots,K} \tag{55}$$

and data

$$b_n = \Big(b_n^{(1)}, \dots, b_n^{(M)}\Big). \tag{56}$$

Assuming that the coefficient matrix $A_n \in \mathbb{R}^{M \times K}$, $K \leqslant M$ defined by (55) has maximum rank K, then the solution to (54) is given by

$$\hat{\alpha}_n = \big(A_n^T A_n\big)^{-1} A_n^T b_n \,. \tag{57}$$

As has been shown by Gobet et al.[22], the thus defined scheme is strongly convergent of order $1/2$ as $\Delta t \to 0$ and $M, K \to \infty$. Controlling the approximation quality for finite values $\Delta t, M, K$, however, requires a careful adjustment of the simulation parameters and basis functions, especially with regard to the condition number of the matrix A_n, and we will discuss suitable strategies to determine a good basis in Section V.

Remark 3 *If an explicit representation of $\hat{Z}_n$ such as (47) is not available, which, for example, is the case when the noise coefficient $\sigma = \sigma(x)$ is controlled too, it is possible to derive a time stepping scheme for $(\hat{Y}_n, \hat{Z}_n)$ in the following way: multiplying the second equation in (42) by $\xi_{n+1} \in \mathbb{R}^m$ from the left, taking expectations and using the fact that $\hat{Y}_n$ is adapted, it follows that*

$$0 = \mathbf{E}\Big[\xi_{n+1}\big(\hat{Y}_{n+1} - \sqrt{\Delta t}\hat{Z}_n \cdot \xi_{n+1}\big)\big|\hat{X}_n\Big] \tag{58}$$

or, equivalently,

$$\hat{Z}_n = \frac{1}{\sqrt{\Delta t}}\mathbf{E}\Big[\xi_{n+1}\hat{Y}_{n+1}\big|\hat{X}_n\Big] \,. \tag{59}$$

Together with (46) or, alternatively, with

$$\hat{Y}_n = \mathbf{E}\big[\hat{Y}_{n+1} + \Delta t\, h(\hat{X}_n, \hat{Y}_{n+1}, \hat{Z}_n)|\hat{X}_n\big] \,, \tag{60}$$

we have a fully explicit scheme for $(\hat{Y}_n, \hat{Z}_n)$.

B. Deep learning based shooting method

As an alternative we discuss a modification of the deep learning based approach that has been proposed by E et al. [13] and that is basically a clever implementation of a shooting method for two-point boundary value problems.

The idea is to approximate $Y_{n\Delta t}$ for every $n = 0, \ldots, n_{\max} - 1$ by a random variable $\mathcal{Y}_n = \mathcal{Y}_n^\vartheta(x)$ that depends on parameters $\vartheta = (\vartheta_Y, \vartheta_Z) \in \mathbb{R} \times \mathbb{R}^p$ and the initial condition $\hat{X}_0 = x$ and that satisfies the *forward* iteration

$$\mathcal{Y}_{n+1} = \mathcal{Y}_n - \Delta t\, h(\hat{X}_n, \mathcal{Y}_n, \mathcal{Z}_n) + \sqrt{\Delta t}\, \mathcal{Z}_n \cdot \xi_{n+1}\,. \tag{61}$$

Here we model $\mathcal{Y}_0 = \vartheta_Y$ with a single parameter and $\mathcal{Z}_n = \mathcal{Z}^{\vartheta_Z}(\hat{X}_n)$ as a neural net approximation of $Z_{n\Delta t}$, where ϑ is chosen so as to minimise the quadratic loss function

$$\ell(\vartheta) = \mathbf{E}\big[|\mathcal{Y}_\eta - g(\hat{X}_\eta)|^2\big]\,. \tag{62}$$

The choice of the loss function (62) is motivated by the fact that the exact FBSDE solution satisfies

$$\mathbf{E}\big[|Y_\tau - g(X_\tau)|^2\big] = 0\,. \tag{63}$$

Therefore, by construction, the approximants will be adapted, with the property

$$\mathcal{Y}_n \approx V(\hat{X}_n)\,, \quad \mathcal{Z}_n \approx (\sigma^T \nabla V)(\hat{X}_n)\,, \tag{64}$$

assuming that Δt is sufficiently small, that sufficiently many training samples of $\hat{X}_\eta$ are available to approximate the expectation in (62), and that the trained neural network is sufficiently rich (i.e. that p is sufficiently large). Understanding the approximation (64) in more detail will be a subject of future research.

1. Stochastic gradient descent

We define the central objects of the method and explain how to compute the optimal parameters. To this end, let $(\Omega, \mathcal{E}, P)$ be our generic probability space on which the family of random variables $\xi_n \colon \Omega \to \mathbb{R}^d$ that appear in the BSDE (61) is defined. Further let $\mathcal{Y}_n^\vartheta \colon \Omega \to \mathbb{R}$ and $\mathcal{Z}^\vartheta \colon \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ be random fields parametrised by $\vartheta = (\vartheta_Y, \vartheta_Z) \in \mathbb{R} \times \mathbb{R}^p$ that satisfy (61).

Letting $\hat{X}_n^{(1)}, \hat{X}_n^{(2)}, \ldots, \hat{X}_n^{(M)}$ denote independent and identically distributed realisations of the forward dynamics, an unbiased estimator of (62) is given by

$$\hat{\ell}(\vartheta) = \frac{1}{M} \sum_{m=1}^M |\mathcal{Y}_\eta^{(m)} - g(\hat{X}_\eta^{(m)})|^2\,. \tag{65}$$

We suppose that the random function $\hat{\ell}$ is differentiable in ϑ, which allows us to minimise the loss by doing stochastic gradient descent

$$\vartheta^{(i+1)} = \vartheta^{(i)} - \gamma^{(i)} \nabla \hat{\ell}(\vartheta^{(i)})\,, \tag{66}$$

where the step size or learning rate $\gamma^{(i)} \to 0$ is decreasing and satisfies the usual divergence condition

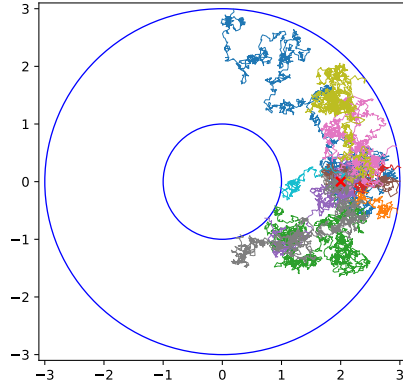$$\sum_{m=1}^\infty \gamma^{(i)} = \infty\,. \tag{67}$$

FIG. 1. Typical realisations of the 2-dimensional committor problem for $a = 1$ and $c = 3$.

By construction $\nabla \hat{\ell}(\vartheta^{(i)})$ is an unbiased estimator of the exact gradient $\nabla \ell(\vartheta^{(i)})$, when conditioned on the current iterate $\vartheta^{(i)}$. The thus described algorithm is the most basic one, but it can be augmented in various ways, e.g. by using adaptive moment estimation.[33]

V. ILLUSTRATIVE EXAMPLES

We consider three different toy examples, one of which involves pure drift-less Brownian motion and a random stopping time with a non-trivial terminal condition, one an Ornstein-Uhlenbeck process on a finite (deterministic) time horizon and another one a metastable overdamped Langevin dynamics. (The code can be found online at `https://github.com/lorenzrichter/BSDE`.)

A. Committor equation

Let $X_t = x + B_t$, $t \geq 0$ be a Brownian motion in $\mathbb{R}^d$, and consider the open and bounded set

$$O = \{x \in \mathbb{R}^d \colon a < |x| < c\} \subset \mathbb{R}^d.$$

We define the sets $A = \{x \in \mathbb{R}^d \colon |x| \leq a\}$ and $C = \{x \in \mathbb{R}^d \colon |x| \geq c\}$ for $a < c$ and denote by

$$\tau = \inf\{t > 0 \colon X_t \in A \cup C\}$$

the first hitting time of $A \cup C$ (see Figure 1). Letting the stopping times τ_A, τ_C be the first hitting times of A, C, the committor probability from A to C as a function of the initial condition x is given by the function

$$h(x) = P_x(\tau_C < \tau_A) = \mathbf{E}_x[\mathbf{1}_C(X_\tau)].$$

The committor function solves the elliptic boundary value problem on $\overline{O} = O \cup \partial A \cup \partial C$:

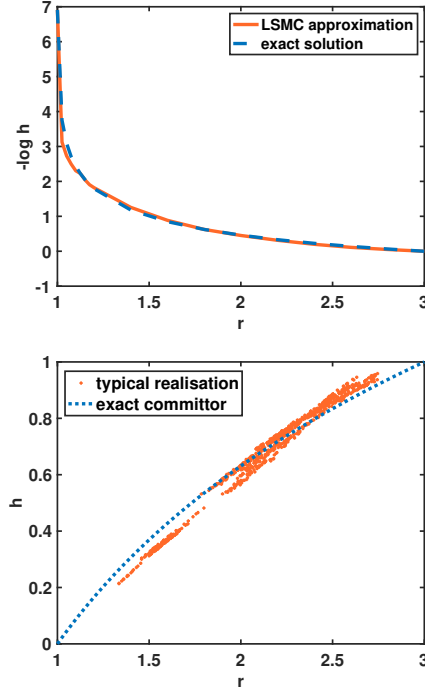$$Lh = 0, \quad h|_{\partial A} = 0, \; h|_{\partial C} = 1, \tag{68}$$

FIG. 2. LSMC approximation of the value function (upper panel) and the committor (lower panel). The lower panel also shows $\exp(-(\hat{Y}_n)_{n \geq 0})$ for a single realisation $(\hat{Y}_n^{(m)})_{n \geq 0}$, as a function of the absolute value process $(|\hat{X}_n^{(m)}|)_{n \geq 0}$.

with $L = \Delta/2$ being infinitesimal generator of X. By the spherical symmetry of the problem, the committor is a function of $r = |x|$ only. Using that the Laplacian of a function $h = h(r)$ can be recast as

$$\Delta h(r) = h''(r) + \frac{d-1}{r} h'(r) , \tag{69}$$

the committor equation can be integrated twice to yield the explicit solution

$$h_d(r) = \frac{a^2 - r^{2-d} a^d}{a^2 - c^{2-d} a^d} , \ d \in \mathbf{N} \tag{70}$$

For $d = 1$, the function is linear. For $d = 2$, both enumerator and denominator are zero, but u has a well-defined limit (computable by l'Hopital's rule), namely,

$$h_2(r) = \frac{\log(a) - \log(r)}{\log(a) - \log(c)} . \tag{71}$$

For $d \to \infty$, the solution h_d converges to the constant 1 on $(a, c]$ and zero for $r = a$.

The associated FBSDE has vanishing running cost, $f = 0$, and non-smooth terminal cost, $g = -\log \mathbf{1}_C$. The numerical solution can be computed after an appropriate regularisation of the logarithm. We choose $g^\epsilon = -\log(\mathbf{1}_C + \epsilon)$ for an arbitrary $\epsilon > 0$. The solution to the thus regularised BSDE is related to (70) by

$$Y_x^\epsilon = -\log(h_d + \epsilon) . \tag{72}$$

We apply the LSMC algorithm described in Section IV A where we run M independent

forward realisations of the discrete Brownian motion

$$\hat{X}_{n+1} = \hat{X}_n + \sqrt{\Delta t}\, \xi_{n+1} \,, \quad \hat{X}_0 = x \,, \tag{73}$$

with $\Delta t = 0.005$. The maximum length of the trajectories, $T_{\max} = \Delta t n_{\max}$, is set equal to 0.5 times the mean first exit time from a d-dimensional hypersphere of radius c that is given by $(c^2 - |x|^2)/d$, so that most, but not all trajectories have exited from O by time $T_{\max}$. In order to solve the corresponding BSDE

$$\hat{Y}^\epsilon_{n+1} = \hat{Y}^\epsilon_n + \frac{\Delta t}{2}|\hat{Z}^\epsilon_n|^2 + \sqrt{\Delta t}\,\hat{Z}^\epsilon_n \cdot \xi_{n+1} \,, \tag{74}$$

with terminal condition

$$\hat{Y}^\epsilon_{n_{\max}} = g^\epsilon(\hat{X}_{n_{\max}}) \,, \tag{75}$$

an adaptive basis of smooth ansatz functions ϕ_k, $k = 1, \ldots, K$ is constructed in the following way: For every $n \in \{1, \ldots, n_{\max}\}$ we compute the empirical mean $\bar{X}_n$ over the M active realisations of the forward process $\hat{X}_n$, and we define Gaussian ansatz functions

$$\phi_k(x) = \mathcal{N}(m_k, v^2) \tag{76}$$

with constant variance $v^2 = 1$ and mean

$$m_k(n) = \bar{X}_n - \delta + \frac{2\delta}{K-1}(k-1) \,, \tag{77}$$

where $\delta = v$ is kept fixed throughout the simulation. (In some cases, it may pay off to set δ equal to the empirical standard deviation of $\hat{X}_n$ for every n.) The last equation admits a straightforward generalisation to the multidimensional case if it is interpreted component-wise. Another strategy in the multidimensional case that has proven useful is to place the basis functions so that, for each component, their means or centre points are equidistributed between the minimum and maximum values of the forward trajectories.

The upper panel of Figure 2 shows the value function (free energy) $-\log \hat{Y}_0 \approx V(r)$ as a function of the initial radius $r = |x|$ for the parameters $a = 1$ and $c = 3$; the simulation parameters were set to $K = 5$ (number of basis functions), $M = 1000$ (number of realisations), $\delta = 1$ (spreading of basis functions), and $v = 2$ (variance of Gaussian basis function). Note that even though we use globally supported radial basis functions to represent the solution of the backward SDE, the approximation of $V(|x|)$ is meaningful only in a small neighbourhood of the initial value $\hat{X}_0 = x$. Nevertheless it is possible to obtain a coarse representation of the value function or the committor function along single realisations, using that $\hat{Y}_n \approx V(|\hat{X}_n|)$, by definition of the backward process (see lower panel of Figure 2).

We tested the LSMC algorithm for a 10-dimensional example, with a single initial value $x \in \mathbb{R}^{10}$ with $|x| = 1.5$, $a = 1$, $c = 2$. Figure 3 illustrates the bias coming from the fact that all forward realisations have finite length $T_{\max} = n_{\max}\Delta t$. The bias can be reduced by increasing $n_{\max}$, at the expense of increasing the computational overhead and the variance of the estimator as the variance of the LSMC coefficients (57) increases when the number of alive (i.e. non-stopped) realisations decreases. Note that the relative error in this case is below 1%.

The deep learning based algorithm did not produce any reproducible results on the committor example.
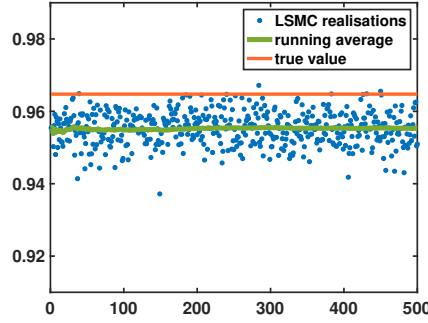
FIG. 3. LSMC approximation of the 10-dimensional value function and the resulting committor probability.

B. Ornstein-Uhlenbeck process

An example for which the deep learning based shooting method is applicable is when the stopping time is deterministic (cf. Remark 2). Specifically, we consider the computation of the conditional expectation

$$\mathbf{E}[\exp(-\alpha X_T)|X_0 = x] \tag{78}$$

under the one-dimensional linear dynamics

$$dX_t = (\mu - X_t)dt + \sigma dB_t, \tag{79}$$

where we assume $\alpha, \mu, \sigma \in \mathbb{R}$ to be time-independent. Since the transition probability density of the process $X_T^{x,t} := (X_T|X_t = x)$ is explicitly known for all times, namely

$$X_T^{x,t} \sim \mathcal{N}\left((x - \mu)e^{t-T} + \mu, \frac{\sigma^2}{2}\left(1 - e^{2(t-T)}\right)\right), \tag{80}$$

it is straightforward to compute the corresponding (now explicitly time-dependent) value function

$$V(x,t) = \alpha((x-\mu)e^{t-T} + \mu) - \frac{\alpha^2 \sigma^2}{4}\left(1 - e^{2(t-T)}\right) \tag{81}$$

and therefore the optimal control

$$u^*(x,t) = -\sigma \alpha e^{t-T}, \tag{82}$$

which remarkably does not depend on x.

We apply the shooting method with values $x = 0, \alpha = 1, \mu = 0, \sigma = \sqrt{2}, T = 5$ and $\Delta t = 0.05$ by identifying the terminal costs $g(x) = \alpha x$. Contrary to the explanation above, which aims at a hitting time example, here the control is explicitly time-dependent and we therefore need time-dependent approximations $\mathcal{Z}_n = \mathcal{Z}_n^{\vartheta_{Z_n}}(\hat{X}_n)$. For those we choose multiple fully connected neural networks $\mathcal{Z}_n^{\vartheta_{Z_n}} : \mathbb{R} \to \mathbb{R}$, each with one hidden layer, batch normalisation and $p = 105$ parameters, that are supposed to approximate $\hat{Z}_n$ for $n = 1, \ldots, N$, as well as the single parameter $\vartheta_Y \in \mathbb{R}$ that shall approximate $\hat{Y}_0$.

Additionally to considering only uncontrolled forward trajectories, we add the control $v_s = -Z_s$ as described in (34)-(35). More precisely, we use the approximation of the
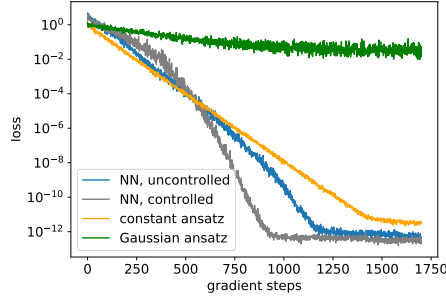
FIG. 4. Loss functions in the shooting method for the OU example for different choices of function classes. For the neural network approximation we compare the controlled and the uncontrolled forward trajectories.

optimal control from a previous iteration step

$$\mathcal{Z}_n^{\vartheta_{Z_n}^{(i)}} = \mathcal{Z}_n^{\vartheta_{Z_n}^{(i)}}(\hat{X}_n) \qquad (83)$$

when simulating the forward trajectories for the $(i+1)$-th gradient step, i.e. we simulate the two processes

$$\hat{X}_{n+1} = \hat{X}_n + \Delta t \left(-\hat{X}_n - \sigma \mathcal{Z}_n^{\vartheta_{Z_n}^{(i)}}(\hat{X}_n) \right) + \sigma \sqrt{\Delta t}\, \xi_{n+1}$$

$$\hat{Y}_{n+1} = \hat{Y}_n - \frac{\Delta t}{2} \left(\mathcal{Z}_n^{\vartheta_{Z_n}^{(i)}}(\hat{X}_n) \right)^2 + \sqrt{\Delta t}\, \mathcal{Z}_n^{\vartheta_{Z_n}^{(i)}}(\hat{X}_n)\xi_{n+1}$$

with $\hat{X}_0 = x, \hat{Y}_0 = \vartheta_Z^{(i)}$. In our simulations we observe that the added control in the forward trajectories can accelerate the convergence of the loss function (65) as shown in Figure 4. We are able to drive the loss to zero with the Adam optimiser[33] and a batch size $M = 50$. In the plots of Figure 5 we see a good agreement of the true optimal control function (82) and its neural network approximation.

In the shooting method, one can of course also use ansatz functions for the approximation of $\hat{Z}_n$, namely

$$\mathcal{Z}_n^{\vartheta_{Z_n}} = \sum_{k=1}^{K} \vartheta_{Z_n}\phi_k(\cdot), \qquad (84)$$

where now $\vartheta_{Z_n} \in \mathbb{R}$. For the Ornstein-Uhlenbeck example, we compare the previous deep learning attempt with choosing two different sets of ansatz functions, once equidistant Gaussians as in (76) and once the "correct" basis function $\phi(x) = 1$, which we identify due to the knowledge of the exact optimal control (82). In both cases the gradient of the loss (65) with respect to ϑ can be computed analytically and stochastic gradient descent can be performed as described above. In Figures 4 and 5 we see that both attempts yield reasonable results, however, when using Gaussians we are not able to drive the loss very close to zero. For a comparison, we additionally approximate this toy example with the LSMC attempt, choosing the same parameters as in the shooting method, and realise that this method is less robust with respect to the choice of ansatz functions and the time discretisation of the stochastic process (cf. Figure 5).

In practice and in particular in higher dimensions it is of course much more difficult to choose ansatz functions appropriately and a priori bounds for the approximation are not available. The application of neural networks to higher dimensional processes on the other
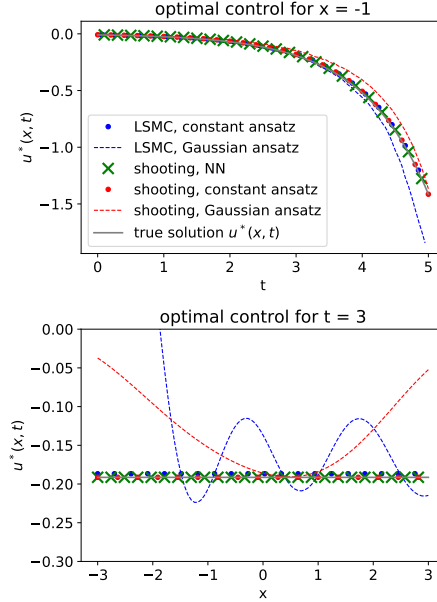
FIG. 5. Approximations of the optimal control for the OU example with shooting and the LSMC methods using different ansatz functions (upper panel: fixed x; lower panel: fixed t).

hand is straightforward, however, the optimisation can become more difficult especially if the dimensions strongly interact. Particularly interesting will be the study of the shooting method in the context of metastable processes.

C. Double-well potential

As an example of a rare event we consider computing the probability of leaving a metastable set before time T,

$$\psi(x,t) = P(\tau_O < T|X_t = x),$$

where the dynamics is given by the Langevin equation

$$dX_t = -\nabla U(X_t)dt + \sigma dW_t \tag{85}$$

with a potential $U(x) = (x^2 - 1)^2$ and a random stopping time $\tau_O = \inf\{t > 0 : X_t \notin O\}, O = (\infty, 0)$. We recall that leaving a metastable set scales exponentially with the energy barrier ΔU and the inverse of the diffusion coefficient σ by Kramers law, namely

$$\lim_{\sigma \to 0} \sigma^2 \log \mathbf{E}[\tau_O] = 2\Delta U. \tag{86}$$

The overall stopping time is defined by $\tau = \min\{\tau_O, T\}$. Referring to the notation in (8) this corresponds to choosing $f(x) = 0$ and $g(x) = -\log(\mathbb{1}_{\partial O}(x))$ and since the latter expression is difficult to handle numerically we consider the reguralized problem by taking $g^\epsilon(x) = -\log(\mathbb{1}_{\partial O}(x) + \epsilon)$ for a small $\epsilon > 0$ and note that $\psi(x,t) = \psi^\epsilon(x,t) - \epsilon$ and $V(x,t) = -\log(\exp(-V^\epsilon(x,t)) - \epsilon)$. We also note that the choice of ϵ can have a significant effect on the corresponding optimal control as illustrated in Figure 6 for the choice of $\sigma = 0.2$.

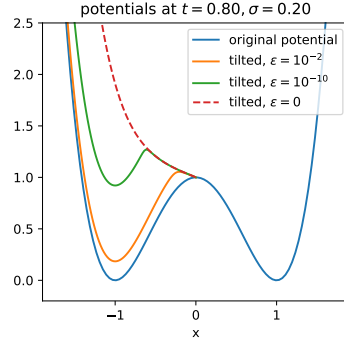By the Feynman-Kac theorem (see e.g. Øksendal[36], Thm. 8.2.1), the function $\psi(x,t)$

FIG. 6. The original double-well potential for fixed t and its optimally tilted versions for different regularization values ϵ.
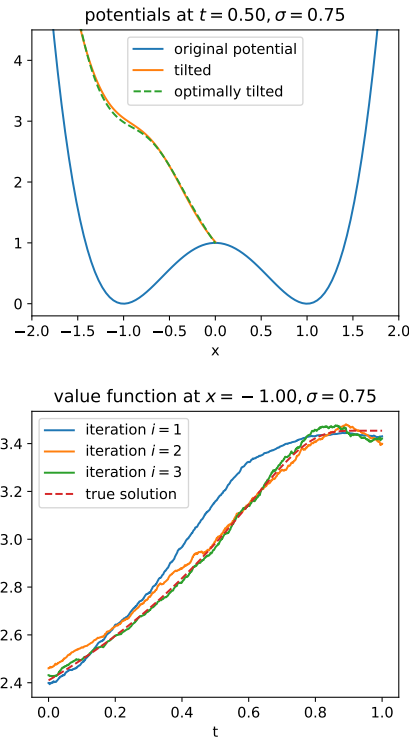

FIG. 7. Top: The original double-well potential and its two tilted versions for foxed t (exact optimal potential and its numerical approximation). Bottom: The approximations of the value functions for fixed x with the iterated LSMC algorithm.

fullfills the linear parabolic evolution equation

$$\left(\frac{\partial}{\partial t} + L\right)\psi(x,t) = 0, \qquad (x,t) \in O \times [0,T] \tag{87}$$

with the boundary conditions

$$\begin{aligned} \psi(0,t) &= 1, \quad t \in [0,T], \\ \psi(x,T) &= 0, \quad x \in O. \end{aligned} \tag{88}$$

	estimate	relative error	trajectories hit
MC	2.42×10^{-4}	61.08	0.02 %
IS	2.54×10^{-4}	2.76	68.15 %

TABLE I. Hitting probabilities: Comparison of brute-force Monte-Carlo (MC) and importance sampling (IS) using a rough FBSDE approximation of the optimal control.

We numerically approach this problem by using the LSMC algorithm explained in section IV A 2, which we additionally iterate by using a previously found approximation as a control variate as explained in (34)-(35). More precisely, after the first iteration, LSMC provides approximations for $\hat{Y}_n$, $\hat{Z}_n$ for $0 \leq n \leq N$, and we can use $-\hat{Z}_n$, corresponding to the optimal control, as an additional drift in the forward process to run LSMC once again and repeat until convergence. As a small modification to the above described algorithm we choose random initial points $\hat{X}_0 \sim \mathrm{Unif}(-1.5, 0)$, which make the algorithm more stable since in particular the matrix inversion in (57) is easier if trajectories are more spread out.

In our simulation, we choose $K = 5$ equidistant Gaussian functions $\phi_k(x)$ as in (76) and let $\epsilon = 0.01, T = 1, \Delta t = 0.001, \sigma = 0.75, K = 1000$. A reference solution is computed by a numerical discretization of (87). In the bottom panel of Figure 7 we see that after the second iteration we get quite close to the true value function, however, we have no guarantee for such a behavior and depending on σ we have observed that the stability of the algorithm crucially depends on the clever choice of ansatz functions and a good initial guess of a drift in the forward process. Convergence analysis of the iteration procedure is a question for further research.

As an alternative strategy for computing the rare event probabilities that we are after, which is also suitable in the case where the value function approximation does not seem to converge, one can resort to importance sampling as an additional step: The LSMC algorithm provides an approximation of the control as in (51) and we can use this—even if potentially suboptimal—in a Girsanov reweighting such as in (13). We illustrate this for $\sigma = 0.5$, for which the value function approximation itself did not yield satisfactory results. In Table I we compare the importance sampling approach to naive Monte Carlo, where one does not add any drift to the forward trajectories. Here the true value is

$$\psi(-1, 0) = \mathbb{P}(\tau < T | X_0 = -1) = 2.62 \times 10^{-4} \,,$$

and we realize that the importance sampling approach brings a significant reduction of the relative error by, roughly, a factor of 20, as a consequence of which the amount of samples needed in order to reach a given accuracy is reduced by a factor of 400.

VI. CONCLUSIONS

We have given a proof of concept that it is possible compute the optimal change of measure for rare event simulation problems with deterministic or random stopping time by solving an associated stochastic optimal control problem. The latter can be recast as a forward-backward stochastic differential equation (FBSDE) that has a nice interpretation in terms of control variates. The FBSDE can be solved by least-squares regression, and we have tested two numerical schemes: a least-squares Monte Carlo algorithm that uses predefined basis functions to represent the solution of the optimal control problem and that can be applied to—potentially high-dimensional—problems with random stopping time and non-smooth terminal cost, and a deep learning based shooting method that can be applied to systems with deterministic finite time horizon. Let us stress that both algorithms can be combined with each other, but despite of their obvious appeal, none of the methods presented should be considered as a black box algorithm that works without any a priori knowledge about the system; a careful choice of the basis functions or the hyperparameters is crucial for the variational problems to converge. Therefore future research ought to address

these questions as well as the generalisation of the deep learning algorithm to problems with random stopping time.

Appendix A: Conditioning and Doob's h-transform

The optimal change of measure that minimises the variance of an importance sampling estimator can be interpreted as a conditional probability under rather general assumptions. Specifically, let X be a Markov process in $\mathbb{R}^d$ with infinitesimal generator L. Introducing the shorthand $P_x(\cdot) = P(\cdot|X_0 = x)$, we define the function

$$h(x) = P_x(X_\tau \in C) \,, \tag{A1}$$

which is only a slight variation of the formula (1), in that all paths start at $X_0 = x$. Then, for any sufficiently small $s > 0$, it follows by the Markov property of X that

$$
\begin{aligned}
h(x) &= \int P_x(X_\tau \in C|X_s = y)\, dP_x(X_s = y) \\
&= \int P_y(X_\tau \in C)\, dP_x(X_s = y) \\
&= \mathbf{E}_x[h(X_s)] \,,
\end{aligned}
$$

where $\mathbf{E}_x[\cdot] = \mathbf{E}[\cdot|X_0 = x]$ denotes the expectation over all paths of X starting at $X_0 = x$. As a consequence,

$$(Lh)(x) = \lim_{s \searrow 0} \frac{1}{s} \left(\mathbf{E}_x[h(X_s)] - h(x) \right) = 0 \,, \tag{A2}$$

which implies that h is harmonic.

For simplicity, we suppose that the transition kernel of X has a smooth and strictly positive transition probability $p_s(x, \cdot)$ on $\mathbb{R}^d$ for any $s > 0$, with

$$P_x(X_s \in A) = \int_A p_s(x, y)\, dy \,, \quad A \subset \mathbb{R}^d \,,$$

and we define

$$p_s^h(x, y) = p_s(x, y) \frac{h(y)}{h(x)} \,, \quad x, y \in \mathbb{R}^d \,.$$

Then, since h is harmonic,

$$p_s^h \geq 0 \quad \text{and} \quad \int p_s^h(x, y)\, dy = \frac{\mathbf{E}_x[h(X_s)]}{h(x)} = 1 \,,$$

that is, $p_s^h(x, \cdot)$ is a transition probability density for every $s \geq 0$; we denote by $P_x^h(X_s \in \cdot)$ the corresponding transition probability and by $X^h = (X_s^h)_{s \geq 0}$ the corresponding process. The transformation

$$p_s(x, y) \mapsto p_s(x, y) h(y)/h(x)$$

is called an h-transform, and the transformed process X^h enjoys the familiar zero-variance property:

$$h(x) = \mathbf{1}_C(X_\tau^h)\varphi^{-1} \quad \text{a.s.}$$

Here $\varphi = dP^h/dP$ denotes the likelihood ratio between the path measures P^h and P.

As before this optimal importance sampling change of measure amounts to a conditional probability.

Lemma A.1 *It holds that*

$$P_x^h(X_\tau \in C) = 1 \quad \forall x \in \mathbb{R}^d. \tag{A3}$$

Moreover $P_x^h(\cdot) = P_x(\cdot|X_\tau \in C)$, i.e. the law of X^h is the law of X conditioned on $\{X_\tau \in C\}$.

Proof: Let

$$\mathbf{E}_x^h[g(X_s)] = \frac{1}{h(x)}\mathbf{E}_x[g(X_s)h(X_s)],$$

the expectation with respect to P_x^h of any bounded and measurable function $g\colon \mathbb{R}^d \to \mathbb{R}$. Setting $g(x) = \mathbf{1}_C(x)$, it suffices to show that

$$\mathbf{E}_x^h[\mathbf{1}_C(X_\tau)] = 1 \quad \forall x \in \mathbb{R}^d.$$

Then, by the optional stopping theorem,

$$\begin{aligned}
\mathbf{E}_x^h[\mathbf{1}_C(X_\tau)] &= \frac{1}{h(x)}\mathbf{E}_x[\mathbf{1}_C(X_\tau)h(X_\tau)] \\
&= \frac{1}{h(x)}\mathbf{E}_x[\mathbf{1}_C(X_\tau)],
\end{aligned}$$

and the last expression is equal to one by definition of h. The rest of the proof is omitted for brevity. $\qquad\square$

Appendix B: Conditioning of diffusions

We will now characterise the h-transform in concrete situations, specifically, when X is a diffusion. To this end, we will show that the h-transform can be realised by a change of drift in the SDE (5). By definition, the function h is harmonic, and so it can be characterised as the solution to an elliptic boundary value problem, with the second-order differential operator

$$L\phi = \frac{1}{2}\sigma\sigma^T : \nabla^2\phi + b \cdot \nabla\phi, \quad \phi \in \mathcal{D}(L). \tag{B1}$$

We let $O \subset \mathbb{R}^d$ denote an open and bounded set with $C \subset \partial O$, and we define

$$\tau = \inf\{t > 0\colon X_t \notin O\} \tag{B2}$$

to be the first exit time of the set O. Then h solves the boundary value problem

$$\begin{aligned}
Lh &= 0, & x \in O \\
h &= \mathbf{1}_C, & x \in \partial O.
\end{aligned} \tag{B3}$$

For reasons that will become clear in a moment, we need h to be strictly positive. Therefore we define a regularised indicator function $\mathbf{1}_C^\epsilon = \mathbf{1}_C + \epsilon$. Further assuming that $\sigma\sigma^T$ is invertible with uniformly bounded inverse the operator L is uniformly elliptic and thus the regularised boundary value problem

$$
\begin{aligned}
Lh^\epsilon &= 0\,, & x \in O \\
h^\epsilon &= \mathbf{1}_C^\epsilon\,, & x \in \partial O\,.
\end{aligned}
\tag{B4}
$$

Then, $h^\epsilon = h + \epsilon$ which, by the strong maximum principle, is a strictly positive function on the closure $\overline{O}$. Now define the function $V^\epsilon = -\log h^\epsilon$ that solves the nonlinear elliptic boundary value problem

$$
\begin{aligned}
LV^\epsilon - \frac{1}{2}|\nabla V^\epsilon|_a^2 &= 0\,, & x \in O \\
V^\epsilon &= -\log \mathbf{1}_C^\epsilon\,, & x \in \partial O\,,
\end{aligned}
\tag{B5}
$$

where we have introduced the shorthands $a = \sigma\sigma^T$ and $|v|_a^2 = |a^{1/2}v|^2$. Noting that

$$
-\frac{1}{2}|\nabla V^\epsilon|_a^2 = \min_{u \in \mathbb{R}^d} \left\{ (\sigma u) \cdot \nabla V^\epsilon + \frac{1}{2}|u|^2 \right\}
\tag{B6}
$$

we realise that (B5) is the dynamic programming equation or Hamilton-Jacobi-Bellman (HJB) equation of the following optimal control problem: minimise the cost

$$
J^\epsilon(u) = \mathbf{E}\left[\frac{1}{2}\int_0^\tau |u_s|^2 dt - \log \mathbf{1}_C^\epsilon(X_\tau^u) \right]\,,
\tag{B7}
$$

subject to

$$
dX_s^u = \left(b(X_s^u) + \sigma(X_s^u)u_s \right) ds + \sigma(X_s^u)dB_s\,,
\tag{B8}
$$

with initial data $X_0^u = x$. The optimal control $u^* = (u_s^*)_{s\geq 0}$ is given by the minimiser in (B6):

$$
u_s^* = -\sigma(X_s^u)^T \nabla V^\epsilon(X_s^u)\,.
\tag{B9}
$$

Letting $\epsilon \to 0$ in (B5), the function $V^\epsilon = \min_u J^\epsilon(u)$, considered as function of the initial conditions, converges to the viscosity solution of dynamic programming equation (B5), with $\mathbf{1}_C^\epsilon$ replaced by $\mathbf{1}_C$. Bearing in mind that $V^\epsilon = -\log h^\epsilon$, it follows that, as $\epsilon \to 0$, the optimal control (B9) realises the h-transform, in that the (weak) solution of the controlled SDE

$$
dZ_s = \left(b + (\sigma\sigma^T)\nabla \log h \right)(Z_s)ds + \sigma(Z_s)dB_s\,,
\tag{B10}
$$

with initial condition $Z_0 = x$ has the same law as X^h.

Appendix C: Finite-dimensional Girsanov formula

We will explain the basic idea behind Girsanov's Theorem and the change of measure formulae (12)–(13) for finite-dimensional Gaussian measures, partly following an idea in Papaspiliopoulos and Roberts[37].

Let μ 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 $\mu_B = \mu \circ B^{-1}$. Given a (deterministic) vector $b \in \mathbb{R}^d$ and a matrix $\sigma \in \mathbb{R}^{d\times m}$,

we define a new random variable $X \colon \Omega \to \mathbb{R}^d$ by

$$X(\omega) = b + \sigma B(\omega). \tag{C1}$$

Since B is Gaussian, so is X, with mean b and covariance $C = \sigma \sigma^T$. Now let $u \in \mathbb{R}^d$ and define the shifted Gaussian random variable

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

and consider the alternative representation

$$X(\omega) = b^u + \sigma B^u(\omega) \tag{C2}$$

of X that is equivalent to (C1) if and only if

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

has a solution (that may not be unique though). The idea of Girsanov's Theorem is to seek a probability measure $\nu \ll \mu$ such that B^u is standard Gaussian under ν, and we claim that such a ν should have the property

$$\frac{d\nu}{d\mu}(\omega) = \exp\left(u \cdot B(\omega) - \frac{1}{2}|u|^2 \right) \tag{C3}$$

or, equivalently,

$$\frac{d\nu}{d\mu}(\omega) = \exp\left(u \cdot B^u(\omega) + \frac{1}{2}|u|^2 \right). \tag{C4}$$

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

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

Indeed, since B is standard Gaussian under P, it follows that the probability $\nu(B^u \in A)$ is equal to

$$\int_{\{\omega \,:\, B^u(\omega) \in A\}} \exp\left(u \cdot B(\omega) - \frac{1}{2}|u|^2 \right) d\mu(\omega)$$

$$= \int_{\{\omega \,:\, B(\omega) - u \in A\}} \exp\left(u \cdot B(\omega) - \frac{1}{2}|u|^2 \right) d\mu(\omega)$$

$$= \frac{1}{(2\pi)^{m/2}} \int_{\{x \,:\, x - u \in A\}} \exp\left(u \cdot x - \frac{1}{2}|u|^2 - \frac{1}{2}|x|^2 \right) dx$$

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

$$= \frac{1}{(2\pi)^{m/2}} \int_A \exp\left(-\frac{|y|^2}{2} \right) dy,$$

showing that B^u has a standard Gaussian distribution under ν. Hence, by the definition of

ν, it holds that

$$\mathbf{E}[f(X)] = \mathbf{E}_\nu\left[f(X)\exp\left(-u\cdot B^u(\omega) - \frac{1}{2}|u|^2\right)\right] \qquad (C5)$$

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

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

Since the distribution of the pair (X^u, B) under μ is the same as the distribution of the pair (X, B^u) with $X = b^u + \sigma B^u$ under ν, the identity (C5) entails that

$$\mathbf{E}[f(X)] = \mathbf{E}\left[f(X^u)\exp\left(-u\cdot B(\omega) - \frac{1}{2}|u|^2\right)\right], \qquad (C6)$$

which is the finite dimensional analogue of (12).

REFERENCES

[1] Achref Bachouch, Côme Huré, Nicolas Langrené, and Huyen Pham. Deep neural networks algorithms for stochastic control problems on finite horizon, Part 2: numerical applications. *arXiv e-prints*, art. arXiv:1812.05916, Dec 2018.

[2] Christian Bender and Robert Denk. A forward scheme for backward SDEs. *Stoch. Proc. Appl.*, 117(12): 1793 – 1812, 2007.

[3] Christian Bender and Thilo Moseler. Importance sampling for backward SDEs. *Stoch. Anal. Appl.*, 28 (2):226–253, 2010.

[4] Christian Bender and Jessica Steiner. Least-squares monte carlo for backward sdes. In René A. Carmona, Pierre Del Moral, Peng Hu, and Nadia Oudjane, editors, *Numerical Methods in Finance*, pages 257–289. Springer Berlin Heidelberg, 2012.

[5] Michelle Boué and Paul Dupuis. A variational representation for certain functionals of brownian motion. *Ann. Probab.*, 26(4):1641–1659, 1998.

[6] F. Cérou, P. Del Moral, T. Furon, and A. Guyader. Sequential monte carlo for rare event estimation. *Stat. Comput.*, 22(3):795–808, 2012.

[7] Jeffrey Comer, James C. Gumbart, Jérôme Hénin, Tony Lelièvre, Andrew Pohorille, and Christophe Chipot. The adaptive biasing force method: Everything you always wanted to know but were afraid to ask. *The Journal of Physical Chemistry B*, 119(3):1129–1151, 2015.

[8] Frédéric Crou and Arnaud Guyader. Adaptive multilevel splitting for rare event analysis. *Stoch. Anal. Appl.*, 25(2):417–443, 2007.

[9] P. Dai Pra, L. Meneghini, and W. Runggaldier. Connections between stochastic control and dynamic games. *Math. Control Signals Systems*, 9:303–326, 1996.

[10] Freddy Delbaen, Ying Hu, and Adrien Richou. On the uniqueness of solutions to quadratic bsdes with convex generators and unbounded terminal conditions. *Ann. Inst. H. Poincar Probab. Statist.*, 47(2): 559–574, 2011.

[11] Paul Dupuis and Hui Wang. Importance sampling, large deviations, and differential games. *Stochastics*, 76(6):481–508, 2004.

[12] Paul Dupuis and Hui Wang. Subsolutions of an isaacs equation and efficient schemes for importance sampling. *Math. Oper. Res.*, 32(3):723–757, 2007.

[13] Weinan E, Jiequn Han, and Arnulf Jentzen. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communication in Mathematics and Statistics*, 5(4):349–380, 2017.

[14] Richard S. Ellis. *Entropy, Large Deviations and Statistical Mechanics.* Springer, Berlin, 1985.

[15] Ioannis Exarchos and Evangelos A. Theodorou. Stochastic optimal control via forward and backward stochastic differential equations and importance sampling. *Automatica*, 87:159 – 165, 2018.

[16] W.H. Fleming. Exit probabilities and optimal stochastic control. *Appl. Math. Optim.*, 4:329–346, 1977.

[17] W.H. Fleming and W.M. McEneaney. Risk-sensitive control on an infinite time horizon. *SIAM J. Control Optim.*, 33:1881–1915, 1995.

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

[19] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions.* Springer, 2006.

[20] P. Glasserman and Y. Wang. Counterexamples in importance sampling for large deviations probabilities. *Ann. Appl. Probab.*, 7:pp. 731–746, 1997.

[21] Emmanuel Gobet and Plamen Turkedjiev. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Math. Comput.*, 85(299):1359–1391, 2016.

[22] Emmanuel Gobet, Jean-Philippe Lemor, and Xavier Warin. A regression-based monte carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15(3):2172–2202, 2005.

[23] Carsten Hartmann and Christof Schütte. Efficient rare event simulation by optimal nonequilibrium forcing. *J. Stat. Mech. Theor. Exp.*, 2012:P11004, 2012.

[24] Carsten Hartmann, Lorenz Richter, Christof Schütte, and Wei Zhang. Variational characterization of free energy: Theory and algorithms. *Entropy*, 19(11), 2017.

[25] Côme Huré, Huyên Pham, Achref Bachouch, and Nicolas Langrené. Deep neural networks algorithms for stochastic control problems on finite horizon, part I: convergence analysis. *arXiv e-prints*, art. arXiv:1812.04300, Dec 2018.

[26] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland/KodanshaElsevier, 1989.

[27] Matthew James. Asymptotic analysis of nonlinear stochastic risk-sensitive control and differential games. *Math. Control Signals Systems*, 5:401–417, 1992.

[28] S. Juneja and P. Shahabuddin. Rare-event simulation techniques: An introduction and recent advances. In Shane G. Henderson and Barry L. Nelson, editors, *Simulation*, volume 13 of *Handbooks in Operations Research and Management Science*, pages 291 – 350. Elsevier, 2006.

[29] H. J. Kappen and H. C. Ruiz. Adaptive importance sampling for control and inference. *J. Stat. Phys.*, 162(5):1244–1266, 2016.

[30] Hilbert J. Kappen, Vicenç Gómez, and Manfred Opper. Optimal control as a graphical model inference problem. *Machine Learning*, 87(2):159–182, May 2012.

[31] H.J. Kappen. Path integrals and symmetry breaking for optimal control theory. *J. Stat. Mech. Theor. Exp.*, 2005(11):P11011, 2005.

[32] O. Kebiri, L. Neureither, and C Hartmann. Adaptive importance sampling with forward backward stochastic differential equations. *Proceedings of the IHP Trimester "Stochastic Dynamics Out of Equilibrium", Institute Henri Poincaré (accepted)*, 2018.

[33] Diederick P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

[34] Magdalena Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.*, 28(2):558–602, 2000.

[35] Pierre L'Ecuyer, Michel Mandjes, and Bruno Tuffin. *Importance Sampling in Rare Event Simulation*, chapter 2, pages 17–38. John Wiley & Sons, Ltd, 2009.

[36] B. Øksendal. *Stochastic differential equations : an introduction with applications*. Springer, Berlin, 2003.

[37] Omiros Papaspiliopoulos and Gareth Roberts. Importance sampling techniques for estimation of diffusion models. In *Statistical Methods for Stochastic Differential Equations*, pages 329–357. Chapman and Hall/CRC, 2012.

[38] K. Rawlik, M. Toussaint, and S. Vijayakumar. On stochastic optimal control and reinforcement learning by approximate inference. In *Proc. Robotics: Science and Systems Conference (R:SS '12)*, 2012.

[39] Sebastian Reich. Data assimilation-the schrödinger perspective. *arXiv Preprint arXiv:1807.08351*, 2018.

[40] C. Schütte, S. Winkelmann, and C. Hartmann. Optimal control of molecular dynamics using Markov state models. *Math. Program. Ser. B*, 134:259–282, 2012.

[41] E. Todorov. Efficient computation of optimal actions. *Proc. Natl. Acad. Sci. USA*, 106(28):11478–11483, 2009.

[42] Plamen Turkedjiev. *Numerical methods for backward stochastic differential equations of quadratic and locally Lipschitz type*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2013.

[43] Eric Vanden-Eijnden and Jonathan Weare. Rare event simulation of small noise diffusions. *Communications Pure Appl. Math.*, 65(12):1770–1803, 2012.

[44] Manuel Villén-Altamirano and José Villén-Altamirano. Restart: A straightforward method for fast simulation of rare events. In *Proceedings of the 26th Conference on Winter Simulation*, WSC '94, pages 282–289, San Diego, CA, USA, 1994. Society for Computer Simulation International.

[45] P. Whittle. Risk-sensitivity, large deviations and stochastic control. *Eur. J. Oper. Res.*, 73:295–303, 1994.

[46] P. Whittle. Risk-sensitivity, a strangely pervasive concept. *Macroecon. Dyn.*, 6:5–18, 2002.