

# Adaptive importance sampling with forward-backward stochastic differential equations

Omar Kebiri<sup>1</sup>, Lara Neureither<sup>1</sup>, and Carsten Hartmann<sup>1</sup>

Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus, Germany,

**Abstract.** We describe an adaptive importance sampling algorithm for rare events that is based on a dual stochastic control formulation of a path sampling problem. Specifically, we focus on path functionals that have the form of cumulate generating functions, which appear relevant in the context of, e.g. molecular dynamics, and we discuss the construction of an optimal (i.e. minimum variance) change of measure by solving a stochastic control problem. We show that the associated semi-linear dynamic programming equations admit an equivalent formulation as a system of uncoupled forward-backward stochastic differential equations that can be solved efficiently by a least squares Monte Carlo algorithm. We illustrate the approach with a suitable numerical example and discuss the extension of the algorithm to high-dimensional systems.

**Keywords:** Importance sampling, rare events, path sampling, forward-backward SDE, least squares Monte Carlo, model reduction

## 1 Introduction

Importance sampling of rare events is one of the challenges in computational statistical mechanics, including molecular dynamics [16], material science [11] or climate modelling [28]. Concrete examples include the study of critical phase transitions in many-particle systems or the estimation of small transition probabilities in protein folding. Estimating small probabilities by Monte Carlo is tricky, because the standard deviation of the corresponding statistical estimator is typically larger than the quantity to be estimated. One technique to improve the efficiency of estimators for small probabilities is importance sampling. Here the idea is to sample from another distribution under which the rare event is no longer rare and then correct (i.e. reweight) the estimator with the appropriate likelihood ratio. Designing such a change of measure so that the variance of the reweighted estimator stays bounded is not at all a trivial task, and several methods have been developed to cope with this issue; for an overview, we refer to the standard textbooks [1, 24] and the references therein.

Here we consider adaptive importance sampling strategies where the change of measure is mediated by an exponential tilting of the reference probability measure. For stochastic differential equations, this exponential tilting can be

interpreted as a control that changes the drift of the stochastic dynamics. Adaptive importance sampling has been predominantly studied in the context of small noise diffusions, for which the optimal control can be computed from the zero viscosity limit of the corresponding dynamic programming equation [8, 9, 27]. In this case the value function of the zero viscosity (deterministic) control problem is equal to the large deviations rate function that describes the exponential tails of the rare events under consideration, and as a consequence, the change of measure captures the rare events statistics and results in estimators that, under certain assumptions, have uniformly bounded relative error.

Here we follow a different route, in that we do not resort to large deviations asymptotics, but rather try to compute the exact change of measure from a suitable approximation of the dynamic programming equation that is underlying the stochastic control problem; cf. [16, 29]. In contrast to our previous works [16, 29], in which the change of measure has been obtained by solving the corresponding variational problem directly, we here focus on the reformulation of the underlying dynamic programming equation as a system of forward-backward stochastic differential equations (see, e.g. [21, 25]) that is solved by a least squares Monte Carlo algorithm [4, 13]. Our approach is partly inspired by related duality techniques in financial mathematics [17, 23], but exploits the specific duality structure of the change of measure problem; see also [22] for a survey of related approaches in financial mathematics.

The paper is organised as follows: In Section 2 we introduce our stochastic dynamics, the corresponding path space free energy and its dual variational characterisation. Section 3 deals with the formulation of the free energy sampling problem and the (dual) optimal control problem as a forward-backward stochastic differential equation (FBSDE, in short). The numerical solution of the FBSDE that can be used to either directly compute the free energy or to approximate the optimal control that generates the minimum variance importance sampling scheme is the topic of Section 4, with a simple numerical illustration presented in Section 5. The article concludes in Section 6 with a short summary and a discussion of open problems and future work.

## 2 Importance sampling in path space

Let  $X = (X_s)_{s \geq 0}$  be the solution of

$$dX_s = b(X_s, s) ds + \sigma(X_s) dB_s, \quad X_0 = x, \quad (1)$$

where  $X_s \in \mathbb{R}^d$ ,  $b$  and  $\sigma$  are smooth drift and noise coefficients, and  $B$  is an  $m$ -dimensional standard Brownian motion where in general  $m \leq d$ . Our standard example will be a non-degenerate diffusion in an energy landscape,

$$dX_s = -\nabla U(X_s) ds + \sigma dB_s, \quad X_0 = x, \quad (2)$$

with smooth potential energy function  $U$  and  $\sigma > 0$  constant. We assume throughout this paper that the functions  $b, \sigma, U$  are such that Equations (1)

or (2) have unique strong solutions for all  $s \geq 0$ . Now let  $W$  be a continuous functional

$$W_\tau(X) = \int_0^\tau f(X_s, s) ds + g(X_\tau), \quad (3)$$

of  $X$  up to some bounded stopping time  $\tau$  where  $f, g$  are bounded and sufficiently smooth, real valued functions.

**Definition 1 (Path space free energy).** *Let  $X$  be the solution of Equation (1) and let  $W_\tau = W_\tau(X)$  be defined by Equation (3). The quantity*

$$\gamma = -\log \mathbf{E}[\exp(-W_\tau)] \quad (4)$$

*is called the free energy of  $W_\tau$  where the expectation is understood with respect to the realizations of (1) for given initial condition  $X_0 = x$ .*

### 2.1 Donsker-Varadhan variational formula for the free energy

The adaptive importance sampling strategy described below is based on a variational characterization of (4) in terms of a change of measure. To make it precise, 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  $B$  in (1). We denote the expectation with respect to  $P$  by  $\mathbf{E}[\cdot]$ . In abstract form, the Donsker-Varadhan variational principle [10] states

$$\gamma = \inf_{Q \ll P} \{\mathbf{E}_Q[W_\tau] + D(Q|P)\}, \quad (5)$$

where  $Q \ll P$  stands for absolute continuity of  $Q$  with respect to  $P$ , and

$$D(Q|P) = \begin{cases} \int_\Omega \log \frac{dQ}{dP}(\omega) dQ(\omega) & \text{if } Q \ll P \\ +\infty & \text{else.} \end{cases} \quad (6)$$

denotes the *relative entropy* or *Kullback-Leibler divergence* between  $Q$  and  $P$ . Note that  $D(Q|P) = \infty$  when  $Q$  is not absolutely continuous with respect to  $P$ , therefore it is sufficient to take the infimum in (5) over all path space measures  $Q \ll P$ . If  $W_\tau \geq 0$ , it is a simple convexity argument (see, e.g., [7]), which shows that the minimum in Equation (5) is attained at  $Q^*$  given by

$$\left. \frac{dQ^*}{dP} \right|_{\mathcal{F}_\tau} = \exp(\gamma - W_\tau), \quad (7)$$

where  $\varphi|_{\mathcal{F}_\tau}$  denotes the restriction of the path space density  $\varphi = dQ^*/dP$  to the  $\sigma$ -algebra  $\mathcal{F}_\tau \subset \mathcal{E}$  that is generated by the Brownian motion  $B$  up to time  $\tau$ .<sup>1</sup>

<sup>1</sup> More precisely,  $\varphi|_{\mathcal{F}_\tau}$  is understood as the restriction of the measure  $Q^*$  defined by  $dQ^* = \varphi dP$  to the  $\sigma$ -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  $\sigma$ -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  $t$ .

By the strict convexity of the exponential function, it holds that  $Q^*$ -a.s. [15]

$$\mathbf{E} [\exp(-W_\tau)] = \exp(-W_\tau) \left( \frac{dQ^*}{dP} \Big|_{\mathcal{F}_\tau} \right)^{-1} \quad (8)$$

or, equivalently,

$$\gamma = W_\tau + \log \left( \frac{dQ^*}{dP} \Big|_{\mathcal{F}_\tau} \right). \quad (9)$$

That is,  $Q^*$  defines a zero-variance change of measure. (Note that the inverse of the Radon-Nikodym derivative in (8) exists since  $W_\tau$  is bounded.)

## 2.2 Related stochastic control problem

The only admissible change of measure from  $P$  to  $Q$  such that  $D(Q|P) < \infty$  results in a change of the drift in Equation (1). Specifically, let  $u$  be a process with values in  $\mathbb{R}^m$  that is adapted to  $B$  and that satisfies

$$\mathbf{E} \left[ \exp \left( \frac{1}{2} \int_0^\tau |u_s|^2 ds \right) \right] < \infty. \quad (10)$$

Further define the auxiliary process

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

so that (1) can be expressed as

$$dX_s = (b(X_s, s) + \sigma(X_s)u_s) ds + \sigma(X_s)dB_s^u, \quad X_0 = x. \quad (11)$$

By construction,  $B^u$  is not a Brownian motion under  $P$ , but by Girsanov's theorem (see, e.g., [19], Theorem 8.6.4) there exists a measure  $Q$  defined by

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_\tau} = \exp \left( \int_0^t u_s \cdot dB_s^u + \frac{1}{2} \int_0^t |u_s|^2 ds \right) \quad (12)$$

so that  $B^u$  is a standard Brownian motion under  $Q$ . (The Novikov condition (10) guarantees that  $Q$  is a probability measure.) Inserting (12) into (5), using that  $B^u$  is a Brownian motion with respect to  $Q$ , it follows that (cf. [6, 7]):

$$\gamma = \inf_u \mathbf{E}_Q \left[ \int_0^\tau f(X_s, s) + \frac{1}{2} |u_s|^2 ds + g(X_\tau) \right], \quad (13)$$

with  $X$  being the solution of Equation (11). 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

$$\gamma = \inf_u \mathbf{E} \left[ \int_0^\tau f(X_s^u, s) + \frac{1}{2} |u_s|^2 ds + g(X_\tau^u) \right]. \quad (14)$$

where  $X^u$  is the solution of the controlled SDE

$$dX_s^u = (b(X_s^u, s) + \sigma(X_s^u)u_s) ds + \sigma(X_s^u)dB_s, \quad X_0^u = x, \quad (15)$$

with  $B$  being the standard  $m$ -dimensional Brownian motion under  $P$ . The Donsker-Varadhan variational principle (5) and zero-variance property (8) of the probability measure  $Q^*$ , for which equality in (5) is attained, have the following stochastic control analogue (see [15, Thm. 3.1]):

**Theorem 1.** *Let  $T > 0$  and  $\tau_O = \inf\{s > 0: X_s^u \notin O\}$  for an open and bounded set  $O \subset \mathbb{R}^n$  with smooth boundary  $\partial O$ . Further define  $\tau = \tau_O \wedge T$  and*

$$\Psi(x, t) = \mathbf{E} \left[ \exp \left( - \int_t^\tau f(X_s, s) ds - g(X_\tau) \right) \middle| X_t = x \right] \quad (16)$$

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

$$u_s^* = \sigma(X_s^{u^*})^T \nabla_x \log \Psi(X_s^{u^*}, s) \quad (17)$$

and (15) yields a zero variance estimator, i.e.,

$$\Psi(x, 0) = \exp \left( - \int_0^\tau f(X_s^{u^*}, s) ds - g(X_\tau^{u^*}) \right) \left( \frac{dQ^*}{dP} \middle|_{\mathcal{F}_\tau} \right)^{-1} Q^* \text{-a.s.} \quad (18)$$

### 3 From dynamic programming to forward-backward SDE

Following the route taken by [6], it can be shown that the control  $u^*$  in (17) is the unique minimiser of the following stochastic control problem: minimize

$$J(u) = \mathbf{E} \left[ \int_0^\tau f(X_s^u, s) + \frac{1}{2} |u_s|^2 ds + g(X_\tau^u) \right] \quad (19)$$

over all measurable and square integrable Markovian controls  $u$ , such that the controlled SDE (15) has a unique strong solution. Now let

$$V(x, t) = \min_u \mathbf{E} \left[ \int_t^\tau f(X_s^u, s) + \frac{1}{2} |u_s|^2 ds + g(X_\tau^u) \middle| X_t^u = x \right] \quad (20)$$

be the associated value function (or: optimal cost-to-go). Further define  $E = O \times [0, T)$  and let  $\partial E^+ = (\partial O \times [0, T)) \cup (O \times \{T\})$  be the terminal set of the augmented process  $(X_s^u, s)_{s \geq 0}$ , such that  $\tau = \tau_O \wedge T$  can be recast as

$$\tau = \inf\{s > 0: (X_s^u, s) \notin E\} \quad (21)$$

Assuming sufficient regularity of the coefficients  $b, \sigma, f, g$  and  $\partial O$ , a necessary and sufficient condition for  $u = u^*$  being optimal is that (see [12, Sec. VI.5])

$$u_s^* = -\sigma(X_s^{u^*})^T \nabla_x V(X_s^{u^*}, s) \quad (22)$$

where  $V \in C^{2,1}(E) \cap C(\partial E^+)$  solves the dynamic programming equation

$$\begin{aligned} \partial_t V + LV + h(s, x, V, \sigma^T \nabla_x V) &= 0 \quad \text{in } E \\ V &= g \quad \text{on } \partial E^+, \end{aligned} \quad (23)$$

with nonlinearity

$$h(s, x, y, z) = -\frac{1}{2}|z|^2 + f(x, s) \quad (24)$$

and the infinitesimal generator of the control-free process  $X_t$ ,

$$L = \frac{1}{2} \sigma \sigma^T : \nabla_x^2 + b \cdot \nabla_x. \quad (25)$$

For the derivation of (22)–(23) from the Feynman-Kac representation formula for the free energy (4), we refer to [14, Sec. 6].

### 3.1 FBSDE representation of the dynamic programming equation

We will now recast the semi-linear, parabolic boundary value problem for  $V \in C^{2,1}(E) \cap C(\partial E^+)$ . To this end, define the processes

$$Y_s = V(X_s, s), \quad Z_s = -\sigma(X_s)^T \nabla_x V(X_s, s) \quad (26)$$

with  $X$  denoting the solution of the uncontrolled SDE (1) with infinitesimal generator (25). Applying Ito's formula to  $Y$ , using that  $V$  is a classical solution to (23), we obtain the following backward SDE (BSDE)

$$dY_s = -h(s, X_s, Y_s, Z_s) ds + Z_s \cdot dB_s, \quad Y_\tau = g(X_\tau) \quad (27)$$

for the pair  $(Y, Z)$ . Note that, by definition,  $Y$  is continuous and adapted to  $X$ , and  $Z$  is predictable and a.s. square integrable, i.e.,

$$\int_0^\tau |Z_s|^2 ds < \infty, \quad (28)$$

in accordance with the interpretation of  $Z_s$  as a control variable. Further note that (27) must be understood as a *backward* SDE rather than a *time-reversed* SDE, since  $Y_s$  at time  $s < \tau$  is measurable with respect to the filtration generated by the Brownian motion  $(B_r)_{0 \leq r \leq s}$ . However  $Y_s$  depends on  $B_\tau$  via the terminal condition  $Y_\tau = g(X_\tau)$ , which requires a larger filtration.

Since  $g$  is bounded, [5, Prop. 3.1] and the results in [18] entail existence and uniqueness of (27); see also [2, 3]. and by exploiting the specific form of the nonlinearity (24) that appears as the driver  $h$  in the backward SDE (27) 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 (23):

$$\begin{aligned} dX_s &= b(X_s, s) ds + \sigma(X_s) dB_s, \quad X_t = x \\ dY_s &= -f(X_s, s) ds + \frac{1}{2} |Z_s|^2 ds + Z_s \cdot dB_s, \quad Y_\tau = g(X_\tau). \end{aligned} \quad (29)$$

The solution to (29) now is a triplet  $(X, Y, Z)$ , and since  $Y$  is adapted, it follows that  $Y_t$  is a deterministic function of the initial data  $(x, t)$  only. As a consequence (see e.g. [20]),

$$Y_t = V(x, t) \quad (\text{a.s.}) \quad (30)$$

equals the value function of our control problem. Recalling Theorem 1, a straight consequence of Equations (14) and (20) therefore is:

**Proposition 1.** *The free energy (4) is equal to*

$$\gamma = \mathbf{E}[Y_0], \quad (31)$$

where the expectation is over the initial conditions  $X_0$  in  $Y_0 = V(X_0, 0)$ .

*Remark 1.* A remark on the role of the control variable  $Z_s$  in the BSDE is in order. In (27), let  $h = 0$  and consider a random variable  $\xi$  that is square-integrable and  $\mathcal{F}_\tau$ -measurable where  $\mathcal{F}_s$  is the  $\sigma$ -Algebra generated by  $(B_r)_{0 \leq r \leq s}$ . Ignoring the measurability for a second, a pair of processes  $(Y, Z)$  satisfying

$$dY_s = Z_s \cdot dB_s, \quad Y_\tau = \xi. \quad (32)$$

is  $(Y, Z) \equiv (\xi, 0)$ , but then  $Y$  is not adapted unless the terminal condition  $\xi$  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

$$Y_t = \mathbf{E}[\xi] + \int_0^t \tilde{Z}_s \cdot dB_s, \quad (33)$$

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

$$Y_t = \xi - \int_t^\tau \tilde{Z}_s \cdot dB_s, \quad (34)$$

or, equivalently,

$$dY_t = \tilde{Z}_s \cdot dB_s, \quad Y_\tau = \xi. \quad (35)$$

Hence  $Z_s = \tilde{Z}_s$  in (32) is indeed a control variable that makes  $Y$  adapted.

*Remark 2.* The forward-backward SDE (or: FBSDE) (29) is called *uncoupled* since the forward SDE does not depend on the solution to the associated BSDE—a property that will be exploited in various ways later on.

### 3.2 Importance sampling in path space, cont'd.

The role of the process  $Z$  in the FBSDE representation of the dynamic programming equation is not only to guarantee that  $Y$  in (29) is adapted, so that

$Y_t = V(x, t)$  is the value function, but it can be literally interpreted as a control since  $Z_t = -\sigma(X_t)^T \nabla_x V(X_t, t)$ . We could compute the optimal control for the zero-variance importance sampling estimator (18) by solving (29) with initial condition  $X_t = X_t^u$  on-the-fly, in which case one has to compute the solution of (29) in parallel to the solution of (15). Depending on the nature of the system (in particular the state space dimension) this on-the-fly-computation, though computationally demanding, may be nonetheless a sensible alternative to numerical schemes that seek to approximate the value function by globally supported basis functions, which may be an ill-conditioned problem, e.g. if the majority of the trajectories are known to reside inside a small set.

As an alternative that we discuss in detail later on, we suggest to define a feedback control for the controlled SDE (15) by

$$u_t = -\sigma(X_t^u)^T \nabla_x V_K(X_t^u, t), \quad (36)$$

where

$$V_K(x, t) = \sum_{k=1}^K \alpha_k(t) \phi_k(x) \quad (37)$$

with  $\alpha_k \in \mathbb{R}$  and continuously differentiable (e.g. radial) basis functions  $\phi_k$  is an approximation ansatz for the value function. Then, by Girsanov's Theorem,

$$\mathbf{E}[\exp(-W_\tau)] = \mathbf{E}_Q[\exp(-L_\tau^u - W_\tau^u)] \quad (38)$$

where  $L_\tau^u = \log(dQ/dP)$  is the log likelihood of the change of measure from  $P$  to  $Q$  on  $\mathcal{F}_\tau$ , as given by (12). By continuity of the functional (38), we expect that any unbiased estimator of the right hand side of (38) will have a considerably smaller variance than the plain vanilla estimator (based on independent draws from  $P$ ), provided that  $V_K \approx V$  approximates the value function.

## 4 Least-squares Monte Carlo

In this section we discuss the numerical discretization of the uncoupled FBSDE (29), following an approach that was first suggested by Gobet et al. [13] and later on refined by several authors; here we suggest a semi-parametric approach with radial basis functions based on the work by Bender and Steiner [4].

### 4.1 Time stepping scheme

The fact that the FBSDE (29) is decoupled implies that it can be discretised by an explicit time-stepping algorithm. Here we utilise a variant of the least-squares Monte Carlo algorithm proposed in [13]. The convergence of the numerical schemes for an FBSDE with quadratic nonlinearities in the driver has been analysed in [26]. The least-squares Monte Carlo scheme is based on the Euler discretization of (29), specifically,

$$\begin{aligned} \hat{X}_{n+1} &= \hat{X}_n + \Delta t b(\hat{X}_n, t_n) + \sqrt{\Delta t} \sigma(\hat{X}_n) \xi_{n+1} \\ \hat{Y}_{n+1} &= \hat{Y}_n - \Delta t h(\hat{X}_n, \hat{Y}_n, \hat{Z}_n) + \sqrt{\Delta t} \hat{Z}_n \cdot \xi_{n+1} \end{aligned} \quad (39)$$



where  $(\hat{X}_n, \hat{Y}_n, \hat{Z}_n)$  denotes the numerical discretization of the joint process  $(X_s, Y_s, Z_s)$ , where we set  $X_s \equiv X_{\tau_O}$  for  $s \in (\tau_O, T]$  when  $\tau_O < T$ , and  $(\xi_i)_{i \geq 1}$  is an i.i.d. sequence of normalized Gaussian random variables. Now let

$$\mathcal{F}_n = \sigma(\{\hat{B}_k : 0 \leq k \leq n\})$$

be the  $\sigma$ -algebra generated by the discrete Brownian motion  $\hat{B}_n := \sqrt{\Delta t} \sum_{i \leq n} \xi_i$ . By definition, the continuous-time process  $(X_s, Y_s, Z_s)$  is adapted to the filtration generated by  $(B_r)_{0 \leq r \leq s}$ . For the discretised process, this implies

$$\hat{Y}_n = \mathbf{E}[\hat{Y}_n | \mathcal{F}_n] = \mathbf{E}[\hat{Y}_{n+1} + \Delta t h(\hat{X}_n, \hat{Y}_n, \hat{Z}_n) | \mathcal{F}_n], \quad (40)$$

using that  $\hat{Z}_n$  is independent of  $\xi_{n+1}$ . 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

$$\hat{Y}_n = \mathbf{E}[\hat{Y}_{n+1} + \Delta t h(\hat{X}_n, \hat{Y}_{n+1}, \hat{Z}_{n+1}) | \mathcal{F}_n]. \quad (41)$$

Note that we can use the identification of  $Z$  with the optimal control (36) and replace  $\hat{Z}_{n+1}$  in the last equation by

$$\hat{Z}_{n+1} = -\sigma(\hat{X}_{n+1})^T \nabla V_K(\hat{X}_{n+1}, t_{n+1}), \quad (42)$$

where  $V_K$  is given by the parametric ansatz (37).

*Remark 3.* If an explicit representation of  $\hat{Z}_n$  such as (42) is not available, 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 (39) 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

$$0 = \mathbf{E}\left[\xi_{n+1}(Y_{n+1} - \sqrt{\Delta t} \hat{Z}_n \cdot \xi_{n+1}) | \mathcal{F}_n\right] \quad (43)$$

or, equivalently,

$$\hat{Z}_n = \frac{1}{\sqrt{\Delta t}} \mathbf{E}[\xi_{n+1} Y_{n+1} | \mathcal{F}_n]. \quad (44)$$

Together with (41) or, alternatively, with

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

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

## 4.2 Conditional expectation

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

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

(Hence the name *least-squares Monte Carlo*.) Here measurability with respect to  $\mathcal{F}_n$  means that  $(\hat{Y}_n, \hat{Z}_n)$  can be expressed as functions of  $\hat{X}_n$ . In view of the ansatz (37) and equation (41), this suggests the approximation scheme

$$\hat{Y}_n \approx \operatorname{argmin}_{Y=Y(\hat{X}_n)} \frac{1}{M} \sum_{m=1}^M \left| Y - \hat{Y}_{n+1}^{(m)} - \Delta t h(\hat{X}_n^{(m)}, \hat{Y}_{n+1}^{(m)}, \hat{Z}_{n+1}^{(m)}) \right|^2, \quad (46)$$

where the data at time  $t_{n+1}$  is given in form of  $M$  independent realizations of the forward process,  $\hat{X}_n^{(m)}$ ,  $m = 1, \dots, M$ , the resulting values for  $\hat{Y}_{n+1}$ ,

$$\hat{Y}_{n+1}^{(m)} = \sum_{k=1}^K \alpha_k(t_{n+1}) \phi_k(\hat{X}_{n+1}^{(m)}), \quad (47)$$

and

$$\hat{Z}_{n+1}^{(m)} = -\sigma(\hat{X}_{n+1}^{(m)})^T \sum_{k=1}^K \alpha_k(t_{n+1}) \nabla \phi_k(\hat{X}_{n+1}^{(m)}). \quad (48)$$

At time  $T := N\Delta t$ , the data are determined by the terminal cost:

$$\hat{Y}_N^{(m)} = g(X_N^{(m)}), \quad \hat{Z}_N^{(m)} = -\sigma(\hat{X}_N^{(m)})^T \nabla g(X_N^{(m)}) \quad (49)$$

Note that we have defined the forward process so that all trajectories have length  $T$ , but the realizations may be constant between  $\tau_O$  and the terminal time  $T$ .

The unknowns that have to be computed in every iteration step are the coefficients  $\alpha_k$ , which makes them functions of time, i.e.  $\alpha_k = \alpha_k(t_{n+1})$ . We call  $\hat{\alpha} = (\alpha_1, \dots, \alpha_K)$  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

$$\hat{\alpha}(t_n) = \operatorname{argmin}_{\alpha \in \mathbb{R}^K} \|A_n \alpha - b_n\|^2, \quad (50)$$

with coefficients

$$A_n = \left( \phi_k(\hat{X}_n^{(m)}) \right)_{m=1, \dots, M; k=1, \dots, K} \quad (51)$$

and data

$$b_n = \left( \hat{Y}_{n+1}^{(m)} + \Delta t h(\hat{X}_n^{(m)}, \hat{Y}_{n+1}^{(m)}, \hat{Z}_{n+1}^{(m)}) \right)_{m=1, \dots, M}. \quad (52)$$

Assuming that the coefficient matrix  $A_n \in \mathbb{R}^{M \times K}$ ,  $K \leq M$  defined by (51) has maximum rank  $K$ , then the solution to (50) is given by

$$\hat{\alpha}(t_n) = (A_n^T A_n)^{-1} A_n^T b_n. \quad (53)$$

The thus defined scheme that is summarised in Algorithm 1 is strongly convergent of order 1/2 as  $\Delta t \rightarrow 0$  and  $M, K \rightarrow \infty$ ; see [13]. Controlling the approximation quality for finite values  $\Delta t, M, K$ , however, requires a careful adjustment of the simulation parameters and appropriate 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 the next section.

---

**Algorithm 1** Least-squares Monte Carlo
 

---

Define  $K, M, N$  and  $\Delta t = T/M$ .

Set initial condition  $x \in \mathbb{R}^d$ .

Choose radial basis functions  $\{\phi_k \in C^1(\mathbb{R}^d, \mathbb{R}) : k = 1, \dots, K\}$ .

Generate  $M$  independent realisations  $\hat{X}^{(1)}, \dots, \hat{X}^{(M)}$  of length  $N$  from

$$\hat{X}_{n+1} = \hat{X}_n + \Delta t b(\hat{X}_n, t_n) + \sqrt{\Delta t} \sigma(\hat{X}_n) \xi_{n+1}, \quad \hat{X}_0 = x.$$

Initialise BSDE by

$$\hat{Y}_N^{(m)} = g(\hat{X}_N^{(m)}), \quad \hat{Z}_N^{(m)} = -\sigma(\hat{X}_N^{(m)})^T \nabla g(\hat{X}_N^{(m)}).$$

**for**  $n = N - 1$ : **1 do**

Assemble linear system  $A_n \hat{\alpha}(t_n) = b_n$  according to (50)–(52).

Evaluate  $\hat{Y}_n^{(m)}$  and  $\hat{Z}_n^{(m)}$  according to

$$\hat{Y}_n^{(m)} = \sum_{k=1}^K \alpha_k(t_n) \phi_k(\hat{X}_n^{(m)}), \quad \hat{Z}_n^{(m)} = -\sigma(\hat{X}_n^{(m)})^T \sum_{k=1}^K \alpha_k(t_n) \nabla \phi_k(\hat{X}_n^{(k)}).$$

If necessary, adapt basis functions  $\phi_k$ .

**end for**

---

*Remark 4.* In some cases, generating forward trajectories that sample the terminal condition  $g(\cdot)$  may be difficult because of metastability or the like. In this case, it may be possible to change the drift of the forward SDE from  $b$  to, say,  $b_0$  where  $b_0$  is chosen such that generating forward trajectories is easy. Assuming that the noise coefficient  $\sigma$  is square and invertible, it is easy to see that the dynamic programming PDE (23) can be recast as

$$\begin{aligned} \partial_t V + \tilde{L}V + \tilde{h}(s, x, V, \sigma^T \nabla_x V) &= 0 \quad \text{in } E \\ V &= g \quad \text{on } \partial E^+, \end{aligned}$$

where

$$\tilde{L} = L - (b - b_0) \cdot \nabla$$

is the generator of a forward SDE with drift  $b_0$ , and

$$\tilde{h}(x, y, z) = h(x, y, z) + \sigma(x)^{-1} (b(x) - b_0(x)) \cdot z.$$

is the driver of the corresponding backward SDE. Changing the drift may be moreover advantageous in connection with the martingale basis approach of Bender and Steiner in [4] the authors have suggested to use basis functions that are defined as conditional expectations of certain linearly independent candidate functions over the forward process, which makes the basis functions martingales. Computing the martingale basis, however, comes with a large computational

overhead, which is why the authors consider only cases in which the conditional expectations can be computed analytically. Changing the drift of the forward SDE may thus be used to simplify the forward dynamics so that its distribution becomes analytically tractable.

## 5 Numerical illustration

We shall illustrate the previous considerations with a standard example. To this end, we consider a one dimensional diffusion in the double-well potential  $U(x) = (x^2 - 1)^2$  that is governed by the equation

$$dX_s = -\nabla U(X_s)ds + \sigma dB_s, \quad X_0 = x, \quad (54)$$

and want to compute the probability of exiting from the left well  $O = \{x < 0\}$  before time  $T < \infty$ . More specifically, we set  $f \equiv 0$  and  $g(x) = -\log(\mathbf{1}_{\partial O}(x))$  in equation (3) and define the bounded stopping time  $\tau = \tau_O \wedge T$  to be the minimum of the first exit time  $\tau_O$  of the set  $O$  and the terminal time  $T$ . Note that  $\tau_O$  is a.s. finite since the potential  $U$  is growing sufficiently fast at infinity, so that  $(X_s)_{s \geq 0}$  is positive recurrent.

For the equivalent stochastic control problem with the cost

$$J(u) = \mathbf{E} \left[ \frac{1}{2} \int_0^\tau |u_s|^2 ds - \log(\mathbf{1}_{\partial O}(X_\tau)) \right] \quad (55)$$

and the controlled process

$$dX_s^u = (\sigma u_s - \nabla U(X_s^u)) ds + \sigma dB_s, \quad X_0^u = x, \quad (56)$$

this means that the control  $u$  seeks to push the process towards the set boundary  $\partial O$  when  $s \approx T$  and the process has not yet left the set  $O$ , for otherwise there will an infinite cost to pay.

Since such an infinite terminal cost is numerically difficult to handle, we consider a regularised control problem and replace  $g$  by  $g^\varepsilon = -\log(\mathbf{1}_{\partial O}(x) + \varepsilon)$ ; for the numerical calculations, we choose  $\varepsilon = 0.01$ . The duality relation (5) between the control value  $\gamma^\varepsilon = \min_u J(u)$  for fixed initial data  $X_0 = x$  and the transition probability  $P(\tau_O < T)$  then reads

$$P(\tau_O < T | X_0 = x) = \exp(-\gamma^\varepsilon) - \varepsilon \quad (57)$$

We will compare the results from the FBSDE solution for  $\gamma^\varepsilon$  with a reference solution that is obtained from numerically solving the linear PDE

$$\left( \frac{\partial}{\partial t} - L \right) u(x, t) = 0, \quad (x, t) \in O \times [0, T] \quad (58)$$

together with the boundary conditions<sup>2</sup>

$$\begin{aligned} u(0, t) &= 1, & t &\in [0, T) \\ u(x, 0) &= 0, & x &\in O. \end{aligned} \quad (59)$$

<sup>2</sup> For the numerical computation, we add reflecting boundary conditions at  $x = -L$  for some  $L > 0$ , the precise value of which does not affect the results (assuming that it is sufficiently large, say,  $L > 3$ ) since the potential has a 4-th order growth.

Then

$$u(x, T) = P(\tau_O < T | X_0 = x). \quad (60)$$

Table 1 below shows the reference value  $V_{ref}^\varepsilon(0, x) := -\log(u(T, x) + \varepsilon)$ , together with the corresponding FBSDE solution. The procedure to obtain the FBSDE solution is described in Algorithm 1, and the table displays the results for different values of  $K, M, N = \lceil T/\Delta t \rceil$ . As basis functions we choose

$$\phi_{k,n}^{\mu_k, \delta}(x) = \exp\left(-\frac{(\mu_k - x)^2}{2\delta}\right), \quad (61)$$

where  $\delta = 0.1$  is fixed but  $\mu_k = \mu_k(n)$  varies with time such that the forward process can be well covered by the basis functions. More precisely, the centres of the basis functions are chosen by simulating  $K$  additional independent forward trajectories  $X^{(k)}, k = 1, \dots, K$  and letting  $\mu_k(n) = X_n^{(k)}$ . We let the whole algorithm run 20 times and compute empirical mean and variance of  $V^\varepsilon$ , denoted by  $\bar{V}^\varepsilon$  and  $S^2(V^\varepsilon)$ . The results are shown in the table.

Table 1: Numerical results for the FBSDE scheme described in Algorithm 1.

	$V_{ref}^\varepsilon(0, x)$	$\bar{V}^\varepsilon(0, x)$	$S^2(V^\varepsilon(0, x))$
$K = 8, M = 300, T = 5, \Delta t = 10^{-3},$ $x = -1, \sigma = 1$	0.3949	0.3748	$10^{-3}$
$K = 5, M = 300, T = 1, \Delta t = 10^{-3},$ $x = -1, \sigma = 1$	1.7450	1.6446	0.0248
$K = 5, M = 400, T = 1, \Delta t = 10^{-4},$ $x = -1, \sigma = 0.6$	4.3030	4.5779	$10^{-3}$
$K = 6, M = 450, T = 1, \Delta t = 10^{-4},$ $x = -1, \sigma = 0.5$	4.5793	4.6044	$5 \cdot 10^{-4}$

Overall we find that the FBSDE scheme results in a fairly good approximation of the value function and, as a consequence of the smoothness of the basis functions, of the optimal control. Moreover, due to the adaptive choice of the basis functions  $\{\phi^{\mu_k, \delta}\}$ , the results do not seem to be very sensitive to the noise intensity  $\sigma$  or the time horizon  $T$ . Speaking of which, we stress that increasing the number of basis functions  $K$  is not always advisable, since the matrix  $A$  in (51) can easily become rank deficient, especially if  $\sigma$  is small and the trajectories stay close together. Therefore it is crucial to check the rank of  $A$  in the simulation and to set  $K$  to the value of the maximally observed rank.

### 5.1 Computational issues

Let us also discuss the fact that we set  $X_s \equiv X_{\tau_O}$  for  $s \in (\tau_O, T]$  when  $\tau_O < T$  again in more detail. Setting the forward trajectories constant from the exit time on, allows to include the terminal condition  $g(X_{\tau \wedge T})$  into the least squares problem at time  $T$ , i.e. into the initialisation step  $b_N$ , for all backward trajectories. It seems that this stabilizes the solution of the backward trajectory  $\hat{Y}$ .

Another approach, following the equations more closely, would be to start each backward trajectories individually from either  $\tau$  or  $T$  depending on whether the corresponding forward trajectory  $\hat{X}$  has made an exit or not. This approach induces numerical problems though, because the data  $b_n$  in the least squares problem as given by (52), which in the case that all backward trajectories were starting from  $T$  was dominated by the positive term  $\hat{Y}_{n+1}$ , is now perturbed at the different exit times by the negative value  $-\log(\varepsilon)$ . This renders the solution  $\hat{\alpha}_n$  of the linear equation (53) rougher, which in turn leads to fluctuations in the solution of  $\hat{Y}_n$  and  $\hat{Z}_n$  which can build up and eventually lead to an explosion of the solutions.

Let us further make suggestions how to efficiently treat the case when  $T$  is large. We will resort to the ideas of Remark 4 here, which suggests to modify the drift  $b$  to  $b_0$  such that under the new drift the event which determines the stopping time  $\tau$  is not rare anymore. Assume now, that for all trajectories  $\hat{X}^{(m)}$ ,  $m = 1, \dots, M$  the family of stopping times

$$\tau_O^m = \left\{ s > 0 : X_s^{(m)} \notin O \right\} \quad (62)$$

is dominated by  $T$  in the sense that

$$\tilde{T} := \max\{\tau_O^m : m = 1, \dots, M\} \ll T. \quad (63)$$

Then the terminal condition  $g$  is completely determined at time  $\tilde{T}$  and the same is true for the backward dynamics. Hence, we suggest in case that  $T$  is large to modify the drift such that  $\tilde{T}$  will be small and run the algorithm only up to time  $\tilde{T}$ . In this case we propose to start each backward trajectory individually from the corresponding exit time on. The matrix  $A_n$  is then of size  $K \times M_n$  where

$$M_n = \left| \left\{ m : \hat{X}_{n-1}^{(m)} \in O \right\} \right| \quad (64)$$

is the number of trajectories which have not left the set  $O$  up to timestep  $n$ . This ensures that  $A$  is not rank deficient at these times which would be the case if we set all trajectory constant after the exit, due to the definition of  $A_n$  with

$$(A_n)_{k,m} = \phi_{k,n}^{\mu_k, \delta}(\hat{X}_n^{(m)}) \quad (65)$$

because the basis functions are evaluated at the same constant value for all these trajectories. To the best of our knowledge, the approximation error of the least squares Monte Carlo algorithms with random stopping times has not been analysed so far, and we leave this topic for future work.

We want to add that in contrast to the complexity of numerically solving the HJB equation, which grows exponentially in the dimension  $d$ , the complexity of solving the FBSDE is determined by solving the SDE and linear equations, i.e. is at most cubic in  $d$ .

## 6 Conclusion and outlook

We have presented a numerical method to compute the free energy of path space functionals of a diffusion process where the functionals may depend on

paths having a random length. Free energies of path space functionals appear in connection with rare event simulation and, as a guiding example for this article, we have considered exit probabilities that are relevant in the context of molecular dynamics or risk analysis.

The approach for efficiently computing path space free energies is based on a variational characterisation of the free energy as the value function of an optimal control problem or, equivalently, as an adaptive importance sampling strategy that is based on the optimal control of the aforementioned stochastic control problem; as we have argued, the importance sampling estimator for the free energy enjoys a minimum variance property under the optimal control. Our numerical strategy for solving the underlying stochastic control problem is based on the reformulation of the corresponding semi-linear dynamic programming equation as a forward-backward stochastic differential equation, which can be solved quite efficiently using a least squares Monte Carlo method. For our guiding example, the reformulation of the adaptive importance sampling algorithm as a forward-backward SDE showed promising results.

We have discussed several options that can help to improve the convergence of the least squares algorithm. For example, we have discussed the option of changing the drift of the forward SDE by modifying the cost functional of the corresponding control problem; while this does not change the dynamic programming equation of the underlying control problem, the corresponding forward-backward stochastic differential equations are different, and it is possible to control the speed of convergence of the numerical method in this way, by controlling the random length of the forward trajectories.

Another aspect that we have only briefly touched upon is the choice of the basis functions for the least squares algorithm. A convenient choice are martingale basis functions that, by definition, are non-parametric and adaptive. Evaluating the martingale basis requires to compute on-the-fly conditional expectations and it is possible to change the drift of the forward SDE so as to avoid numerically expensive computations of the conditional expectations. In this article we used a semi-parametric approach, future research should address the non-parametric one. Another interesting topic concerns sampling problems on an infinite time horizon, which can be represented by a stopping time for hitting an *impossible set*, a set which the dynamics can never reach.

We believe that forward-backward SDE are an interesting numerical and analytical tool for applications in computational statistical mechanics that connects such diverse topics as control, filtering and estimation. A specific feature of the proposed method is that the corresponding forward-backward SDE are decoupled, which leaves room for combining the aforementioned tasks with coarse-graining and model reduction techniques. We leave all this for future work.

**Acknowledgement** This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through the grant CRC 1114 "Scaling Cascades in Complex Systems", Project A05 "Probing scales in equilibrated systems by

optimal nonequilibrium forcing”. O.K. received funding from the EU-METALIC II Programme.

## References

1. A. Asmussen and P. Glynn *Stochastic Simulation: Algorithms and Analysis* Springer, New York, (2007).
2. K. Bahlali , B. Gherbal , B. Mezerdi, *Existence of optimal controls for systems driven by FBSDEs*, Sys. Control Letters 60 (2011) pp.344-349. Fields 103 273-283 (1995).
3. K. Bahlali, O. Kebiri, A. Mtiraoui: *Existence of an optimal Control for a system driven by a degenerate coupled Forward-Backward Stochastic Differential Equations*, C. R. Acad. Sci. Paris, Ser. I (2016).
4. C. Bender, J. Steiner: *Least-Squares Monte Carlo for BSDEs*. In: Carmona et al. (Eds.), *Numerical Methods in Finance*, Springer, 257-289 (2012).
5. A. Bensoussan, L. Boccardo, F. Murat, *Homogenization of elliptic equations with principal part not in divergence form and hamiltonian with quadratic growth*, Commun. Pure Appl. Math. 39 769-805 (1986).
6. M. Boué and P. Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.*, 26(4):1641–1659, (1998).
7. P. Dai Pra., L. Meneghini, and W. J. Runggaldier, *Connections between stochastic control and dynamic games*, Math. Control Signals Systems, 9, pp. 303-326 (1996).
8. Dupuis, P.; Wang, H. Importance sampling, large deviations, and differential games. *Stoch. Int. J. Probab. Stoch. Proc.* 76, 481–508 (2004).
9. Dupuis, P.; Wang, H. Subsolutions of an Isaacs equation and efficient schemes for importance sampling. *Math. Oper. Res.* 32, 723–757 (2007).
10. R. S. Ellis. Entropy, *Large Deviations and Statistical Mechanics*, volume 271 of Grundlehren der mathematischen Wissenschaften. Springer, New York, (1985).
11. S. Engelund and R. Rackwitz. A benchmark study on importance sampling techniques in structural reliability. *Structural Safety* 12: 255–276, (1993).
12. W. H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*. Applications of mathematics. Springer, New York, 2nd edition, (2006).
13. E. Gobet, J.-P. Lemor, and X. Warin, Xavier, *A regression-based Monte Carlo method to solve backward stochastic differential equations*, Ann. Appl. Probab. 15, 2172–2202, (2005).
14. C. Hartmann, R. Banisch, M. Sarich, Th. Badowski, Ch. Schtte. *Characterization of rare events in molecular dynamics* Entropy 16, 350–376 (2014).
15. C. Hartmann, L. Richter, Ch. Schtte, W. Zhang. *Variational characterization of free energy: theory and algorithms* Entropy 19, 626–653, (2017).
16. C. Hartmann and C. Schütte, *Efficient rare event simulation by optimal nonequilibrium forcing*, J. Stat. Mech. Theor. Exp., 2012, p. P11004 (2012).
17. M.B. Haugh and L. Kogan. Pricing American Options: A Duality Approach. *Oper. Res.* 52, 258–270 (2004).
18. Kobylanski, M. Backward stochastic differential equations and partial differential equations with quadratic growth *Annals of Probability*, 28(2):558–602 (2000).
19. B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer, 6th ed., (2010).



20. E. Pardoux, S. Peng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, in: B.L. Rozovskii, R.B. Sowers (Eds.), *Stochastic Partial Differential Equations and their Applications*, Lecture Notes in Control and Information Sciences 176, Springer, Berlin, (1992).
21. S. Peng, *Backward Stochastic Differential Equations and Applications to Optimal Control*, *Appl Math Optim* 27, pp.125-144 (1993).
22. H. Pham, *Continuous-time stochastic control and optimization with financial applications*, *Stochastic modelling and applied probability*, Springer, Berlin, Heidelberg, (2009).
23. L.C.G. Rogers. Monte Carlo valuation of American options. *Mathematical Finance* **12**, 271–286 (2002).
24. R.Y. Rubinstein, Dirk P. Kroese. *Simulation and the Monte Carlo Method*, John Wiley & Sons, Hoboken (2008).
25. N. Touzi *Optimal stochastic control, stochastic target problem, and backward differential equation*, (2010).
26. P. Turkedjiev: *Numerical methods for backward stochastic differential equations of quadratic and locally Lipschitz type*, Dissertation, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II (2013).
27. Vanden-Eijnden, E.; Weare, J. Rare Event Simulation of Small Noise Diffusions. *Commun. Pure Appl. Math.* 65, 1770–1803 (2012).
28. J. Wouters and F. Bouchet. Rare event computation in deterministic chaotic systems using genealogical particle analysis. *J. Phys. A* **49**:374002, (2016).
29. W. Zhang, H. Wang, C. Hartmann, M. Weber and Ch. Schütte, *Applications of the cross-entropy method to importance sampling and optimal control of diffusions*, *SIAM J. Sci. Comput.* 36, A2654-A2672 (2014).