Solving optimal stopping problems via randomization and empirical dual optimization

Denis Belomestny [*], Christian Bender [†] and John Schoenmakers [‡]

October 18, 2021

Abstract

In this paper we consider optimal stopping problems in their dual form. In this way we reformulate the optimal stopping problem as a problem of stochastic average approximation (SAA) which can be solved via linear programming. By randomizing the initial value of the underlying process, we enforce solutions with zero variance while preserving the linear programming structure of the problem. A careful analysis of the randomized SAA algorithm shows that it enjoys favorable properties such as faster convergence rates and reduced complexity as compared to the non randomized procedure. We illustrate the performance of our algorithm on several benchmark examples.

Keywords: Optimal stopping, duality, stochastic average approximation, randomization
MSC: 91G60, 65C05, 60G40

1 Introduction

Since the emergence of complexly structured callable products in the financial industry, also known as American style derivatives, the last decades have seen a huge development of numerical methods for solving optimal stopping problems. Indeed, the evaluation of virtually all callable derivatives that encounter in the financial markets may be mathematically translated to the solution of an optimal stopping problem,

$$y^* := \sup_{\tau \in \mathcal{T}} \mathsf{E}[Z_\tau]. \tag{1.1}$$

In (1.1) $Z = (Z_j)_{j=0,\ldots,J}$, $J \in \mathbb{N}_+$, is a nonnegative square-integrable stochastic process on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_j)_{j=0,\ldots,J}, \mathsf{P})$, adapted to $\mathbb{F}$, and $\mathcal{T}$ is the set of all $\mathbb{F}$-stopping times $\tau \in \{0, \ldots, J\}$. For notational convenience, it is assumed that $\mathcal{F}_0$ is trivial, $Z_0 = 0$, and that $\mathsf{P}(\{Z_i > 0\}) > 0$ for some $i = 1, \ldots, J$, hence $y^* > 0$.

At the cutting edge around the beginning of this century various simulation based methods that aimed at construction of the optimal exercise boundary or exercise strategy were developed. Let us mention, among other popular works, Longstaff and Schwartz [2001], Tsitsiklis and Van Roy [2001], Andersen [1999], and Broadie and Glasserman [2004]. These methods are usually called primal since they provide lower approximations to (1.1) due to their very nature.

[*] Duisburg-Essen University, Essen, denis.belomestny@uni-due.de
[†] Saarland University, Saarbrücken, bender@math.uni-sb.de
[‡] WIAS, Berlin, schoenma@wias-berlin.de

In this paper we focus on the dual representation of the optimal stopping problem (1.1), i.e. the stochastic minimization problem

$$y^* = \inf_{M \in \mathcal{M}} \mathsf{E}[\max_{i=0,\ldots,J}(Z_i - M_i)], \tag{1.2}$$

where $\mathcal{M}$ is the set of all $(\mathcal{F}_i)$-martingales starting in 0 at $i = 0$, see Rogers [2002], Haugh and Kogan [2004] and the short recap in Section 2. We also refer to Brown et al. [2010] and Rogers [2007] for different approaches to the general theory of information relaxation duality in stochastic control.

Several Monte-Carlo algorithms for constructing upper biased estimators for y^* based on the minimization problem (1.2) have been suggested in the literature. They typically consist of two steps:

I Apply some numerical method to pick a martingale $\hat{M}$ (typically depending on some training sample $\mathcal{D}_n$ of size n) which is close to optimality:

$$\mathsf{E}[\max_{i=0,\ldots,J}(Z_i - \hat{M}_i)] \approx y^* \tag{1.3}$$

II Estimate $\mathsf{E}[\max_{i=0,\ldots,J}(Z_i - \hat{M}_i)]$ by the sample mean, applying a new independent sample (testing sample) of size N.

All the existing dual Monte Carlo algorithms can be divided into two broad categories depending on how the martingale $\hat{M}$ is constructed. In the first class of algorithms, see for example Andersen and Broadie [2004], Belomestny et al. [2009], and Glasserman [2003], Belomestny and Schoenmakers [2018] for further references, the choice of the martingale $\hat{M}$ is based on approximating the so-called Doob martingale:

$$M_j^* = \sum_{i=1}^{j} Y_i^* - \mathsf{E}[Y_i^* | \mathcal{F}_{i-1}], \quad j = 0, \ldots J, \tag{1.4}$$

where

$$Y_J^* = Z_J, \quad Y_j^* = \max\{Z_j, \mathsf{E}[Y_{j+1}^* | \mathcal{F}_j]\}, \quad j = J-1, \ldots, 0, \tag{1.5}$$

is the *Snell envelope* process. A particular feature of the Doob martingale is that it solves (1.2) and moreover satisfies (see Section 2)

$$y^* = \max_{i=0,\ldots,J}(Z_i - M_i^*) \quad \text{almost surely.} \tag{1.6}$$

Because of (1.6) we say that the Doob martingale is *surely* or *strongly* optimal.

In the second class of algorithms, one tries to solve the dual optimization problem (1.2) directly using methods of stochastic approximation and some parametric subclasses of $\mathcal{M}$. See for example Joshi and Theis [2002], who use a family of discounted swap rates for dual pricing of Bermudan swaptions, Desai et al. [2012] discussed below, and more recently Lelong [2018] where a family of martingales is constructed by using chaos expansions. Let us briefly discuss further the existing literature around the methods in this second class. One of the first extensive studies is carried out by Desai et al. [2012], where the authors essentially apply the *Stochastic Average Approximation* (SAA) method in the spirit of Shapiro [2003], next use nested Monte Carlo to construct a suitable finite dimensional linear space of martingales, and then cast the resulting minimization problem into a *linear program.* However, as demonstrated somewhat later in Schoenmakers et al. [2013] in a toy example, the approach in Desai et al. [2012] may end up with martingales $\hat{M}$ that are close to optimality in the weak sense (1.3) but with the variance of the random variable $\max_{i=0,\ldots,J}(Z_i - \hat{M}_i)$ being relatively high. In contrast, due to (1.6), for a martingale that is close to the Doob martingale M^* (in the L_2-sense for instance) this variance will be close to zero. As a consequence, for such a martingale the estimation in step II can be done more efficiently. Moreover, in Schoenmakers et al. [2013] it is shown that if $\mathsf{Var} \max_{i=0,\ldots,J}(Z_i - \hat{M}_i) = 0$, then (1.6) holds for $\hat{M}$ and if $\mathsf{Var} \max_{i=0,\ldots,J}(Z_i - \hat{M}_i) \approx 0$, then (1.3) applies (under some mild conditions). Thus, it may be considered desirable to find martingales that are "close" to the Doob martingale, or at least "close" to a surely optimal martingale. Belomestny [2013] proposes a modification based on variance penalization while keeping the resulting minimization problem convex. An extension of the latter approach is considered in Belomestny et al. [2019] where the authors study a family of different penalization methods and their performance. However, due to penalization, the original dual empirical objective function is modified to a computationally more expensive object, and minimization by the standard linear programming may become unfeasible. In a recent study Belomestny and Schoenmakers [2021] suggest to minimize over a randomized class of martingales instead of penalizing the cost criterion. Based on a complete characterization of the sets of all weakly and surely optimal martingales for the original dual problem, they identify an optimal randomization that allows for sorting out the Doob martingale as the only minimizer in the randomized martingale class. However, the identification of the optimal randomization in this approach requires knowledge of the Snell envelope. In practice, this randomization may generally only be approximated, and the theoretical quantification of the resulting martingale based on such an approximation has not yet been studied.

The goal of the present paper is a numerically efficient way to deal with the potentially high variance of the martingales found by minimizing (1.2) directly via stochastic averaging. The key idea is to randomize the initial value Z_0 of the cashflow process. The rationale behind this approach can be easily explained. If we replace the constant value $Z_0 = 0$ by the (unknown) constant y^*, then immediate exercise at time zero becomes optimal and any optimal martingale M^* for the new dual problem satisfies

$$\mathsf{E}[\max\{y^*, \max_{i=0,\ldots,J}(Z_i - M_i^*)\}] = y^*$$

Hence, $\max_{i=0,\ldots,J}(Z_i - M_i^*) \leq y^*$ almost surely, but

$$\mathsf{E}[\max_{i=0,\ldots,J}(Z_i - M_i^*)] \geq y^*$$

by duality for the original problem with $Z_0 = 0$, which implies that $\max_{i=0,\ldots,J}(Z_i - M_i^*) = y^*$ almost surely. Thus, any optimal martingale for the problem with Z_0 replaced by y^* is surely optimal for the original problem with $Z_0 = 0$. Since y^* is the quantity which we wish to compute numerically, it must be considered as unknown and, hence, cannot be used as the new initial value for Z. Instead, we sample the initial value of Z randomly from a distribution whose support is large enough to contain y^*. As demonstrated in Section 4 below, such an initial randomization is still sufficient to ensure that the optimal martingales of the dual problem with randomized Z_0 are exactly the surely optimal martingales of the original dual formulation. These observations motivate to apply the stochastic average approximation to the randomized dual problem in order to benefit from the zero variance property of surely optimal martingales without changing the structure of the original problem.

Finally, let us stress that direct empirical minimization for solving numerically the information relaxation dual has recently been applied beyond the optimal stopping problem, see e.g. Yang et al. [2019] for applications in merchant energy production or Chandramouli [2019] for the pricing of options with several exercise rights. We consider our randomization approach for the dual problem (1.2) as prototypically and expect that it can also be applied to more general stochastic minimization problems. The key ingredient is to design the randomization in such a way that some kind of Bernstein condition (linking the original problem and the randomized problem) is enforced as in Theorem 4.2 below. We also refer to Bartlett and Mendelson [2006] for a discussion on how to improve convergence rates for empirical minimization problems in the presence of the Bernstein condition.

The paper is organized as follows: In Section 2, we briefly review the dual minimization problem, explain why (weakly) optimal martingales with arbitrarily large variance typically exist, and relate our randomization approach to variance penalization. In Section 3, we present a class of stylized examples, in which the Doob martingale is the only surely optimal martingale. However, the stochastic average method applied to the original dual formulation provably fails to converge to the Doob martingale in this example class, while it does converge after initial randomization. Section 4 is devoted to a detailed study of the initial randomization technique. In particular, we derive bounds for the bias and the variance of the original problem in terms of the bias of the randomized problem. These bounds are also crucial for studying the convergence behavior of the empirical randomized dual minimization problem (i.e., after replacing the expectation by the sample mean), see Section 5. In Theorem 5.1, we show that empirical randomized dual problem may converge to the theoretical randomized dual problem at a faster rate than the Monte-Carlo rate of $n^{-1/2}$ in the training sample, even if the martingales are parameterized by a compact set in an infinite-dimensional metric space. The benefits of the initial randomization are demonstrated numerically in Section 6. We run the algorithm into convergence in a toy example, but also illustrate how to exploit the favorable variance properties of our algorithm in a standard multi-asset Bermudan option pricing framework. All proofs are presented in Section 7.

2 Recap of duality, lurking troubles by solving (1.2), and key idea

The representation (1.2) and the almost sure property (1.6) of the Doob martingale (1.4) are by now classical and follow directly from the following two observations:
(i) Let τ^* be an optimal stopping time in (1.1). Then for any martingale M with $M_0 = 0$ we have

$$y^* = \mathsf{E}[Z_{\tau^*}] = \mathsf{E}[Z_{\tau^*} - M_{\tau^*}] \leq \mathsf{E}[\max_{i=0,\dots,J}(Z_i - M_i)] \qquad \text{and}$$

(ii) for the Doob martingale (1.4) it holds that

$$\begin{aligned}
\max_{i=0,\dots,J}(Z_i - M_j^*) &= \max_{i=0,\dots,J}\Big(Z_i - \sum_{j=1}^{i} Y_j^* + \sum_{j=1}^{i} \mathsf{E}[Y_j^*|\mathcal{F}_{j-1}]\Big) \\
&= Y_0^* + \max_{i=0,\dots,J}\Big(Z_i - Y_i^* + \sum_{j=0}^{i-1}(\mathsf{E}[Y_{j+1}^*|\mathcal{F}_j] - Y_j^*)\Big) \\
&\leq Y_0^* = y^*,
\end{aligned}$$

by (1.5). We now expose the troubles that one may encounter when solving (1.2) by straightforward minimizing the corresponding Monte Carlo estimate over a class of martingales M. Denote by

$$\mathcal{M}^\circ := \{M \in \mathcal{M} : \ \mathsf{E}[\max_{i=0,\dots,J}(Z_i - M_i)] = y^*\}$$

the set of all (weakly) optimal martingales. It turns out that the set $\mathcal{M}^\circ$ contains martingales that can be found with positive probability and which have arbitrary high variance. More formally, we prove the following statement in Section 7.2.

Proposition 2.1 *If there is a sequence of events $(B_k)_{k\in\mathbb{N}}$ in $\mathcal{F}_1$ such that $\mathsf{P}(B_k) \in (0,1)$ for every $k \in \mathbb{N}$ and $\lim_{k\to\infty} \mathsf{P}(B_k) = 1$, then,*

$$\sup_{M\in\mathcal{M}^\circ} \mathsf{Var}\left(\max_{i=0,\ldots,J}(Z_i - M_i)\right) = +\infty,$$

that is, the variances of weakly optimal martingales can be arbitrarily large.

For example, the assumption of the above proposition is already satisfied if there is an $\mathcal{F}_1$-measurable standard Gaussian random variable ξ (in this case we set $B_k := \{\xi \leq k\}$). The above considerations suggest favoring numerical methods which try to approximate the Doob martingale of Y^* directly. On the other hand, one of the most popular algorithms of this kind, Andersen and Broadie [2004], requires nested Monte Carlo and a preliminary (usually regression based) estimate of the Snell envelope. The aim of our paper is to modify the plain SAA method in such a way that

- it can provably benefit from the zero variance property of the Doob martingale of Y^*;
- and it does so, without changing the structure of the problem. In particular, the Sample Average Version of the modified minimization problem can still be solved via linear programming on finite dimensional subspaces on $\mathcal{M}$ as was highlighted by Desai et al. [2012] for the SAA version of (1.2).

The key idea is to randomize the initial value Z_0 of the cashflow, that is, to replace Z by

$$Z_i^{(A)} := \begin{cases} A, & i = 0, \\ Z_i, & i > 0, \end{cases}$$

where A is a nonnegative random variable independent of $\mathcal{F}_J$. To illustrate that this randomization behaves similarly to a variance penalty, let us consider the randomized dual problem

$$\inf_{M\in\mathcal{M}} \mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(A)} - M_i)]$$

in the case where A is uniformly distributed on $[0, K]$ for some constant $K > 0$. If we restrict the minimization to martingales that satisfy $\max_{j=0,\ldots,J}(Z_j - M_j) \leq K$ a.s. (cp. the discussion in Remark 4.1 below), then

$$\mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(A)} - M_i)] = \frac{y^*}{K}\left(\mathsf{E}[\max_{i=0,\ldots,J}(Z_i - M_i)] + \frac{1}{2y^*}\mathsf{E}[(\max_{i=0,\ldots,J}(Z_i - M_i) - y^*)^2]\right) + const. \tag{2.1}$$

Hence, the randomized dual problem is essentially the same as minimizing

$$M \mapsto \mathsf{E}[\max_{i=0,\ldots,J}(Z_i - M_i)] + \frac{1}{2y^*}\mathsf{E}[(\max_{i=0,\ldots,J}(Z_i - M_i) - y^*)^2].$$

In particular, if M additionally is weakly optimal, then

$$\mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(A)} - M_i)] = \frac{1}{2K}\mathsf{Var}(\max_{i=0,\ldots,J}(Z_i - M_i)) + const. \tag{2.2}$$

Note that Eq. (2.1) is a consequence of the straightforward identity

$$\mathsf{E}[\max\{A, c\}] = \frac{K}{2} + \frac{c^2}{2K}, \quad c \in [0, K],$$

since, by independence of A and $\mathcal{F}_J$,

$$\mathsf{E}[\max_{i=0,\dots,J}(Z_i^{(A)} - M_i)] = \mathsf{E}[\mathsf{E}[\max\{A, \max_{i=0,\dots,J}(Z_i - M_i)\}|\mathcal{F}_J]] = \frac{K}{2} + \frac{1}{2K}\mathsf{E}[(\max_{i=0,\dots,J}(Z_i - M_i))^2]$$
$$= \frac{y^*}{K}\left(\mathsf{E}[\max_{i=0,\dots,J}(Z_i - M_i)] + \frac{1}{2y^*}\mathsf{E}[(\max_{i=0,\dots,J}(Z_i - M_i) - y^*)^2]\right) + const.$$

3 A stylized yet illustrative example

In this section we consider an example showing the drawbacks of applying straightforwardly the Sample Average Approximation method (see Shapiro [2003]) to the problem

$$\inf_{\gamma} \mathsf{E}\left[\max_{0 \le j \le J}(Z_j - M_j(\gamma))\right],$$

where $(M(\gamma))_\gamma$ is a family of martingales from $\mathcal{M}$. In particular, this example demonstrates that one may end up with convergence to weakly but not surely optimal martingales, even if the Doob martingale is contained in the family $(M(\gamma))_\gamma$. It also illustrates how randomization of Z_0 can serve as a remedy in such a situation.

For complete transparency, we consider a stylized Bermudan option with only two exercise dates. That is, we take $J = 2$ with $Z_0 = 0$, $Z_1 > 0$, $Z_2 > 0$ a.s., so that the Snell Y^* envelope reads

$$Y_2^* = Z_2, \quad Y_1^* = \max\left(Z_1, \mathsf{E}_{\mathcal{F}_1}[Z_2]\right), \quad Y_0^* = y^* = \mathsf{E}\left[\max\left(Z_1, \mathsf{E}_{\mathcal{F}_1}[Z_2]\right)\right].$$

For the rest we keep the example fairly general and realistic by merely assuming that Y_1^* has a density which is strictly positive in an interval $(y_-, y_+) \subset (0, \infty)$ and zero in $[0, \infty) \diagdown (y_-, y_+)$. Hence $y^* \in (y_-, y_+)$. For the Doob martingale we thus have

$$M_1^* = Y_1^* - y^*, \quad M_2^* - M_1^* = Z_2 - \mathsf{E}_{\mathcal{F}_1}[Z_2]. \tag{3.1}$$

As a martingale family we consider

$$M_1(\gamma) = \gamma M_1^*, \quad M_2(\gamma) - M_1(\gamma) = M_2^* - M_1^*, \quad \gamma \in \mathbb{R}. \tag{3.2}$$

One may argue that this family of martingales is rather artificial and unrealistic, as it has only one parameter and contains the Doob martingale (for $\gamma^* = 1$). However, it is perfectly well suited for our illustration purposes. As we will show below, if the number of training paths for the sample average grows to infinity, the martingales that plain SAA delivers, converge to weakly optimal martingales but not to surely optimal ones. This in spite of the fact that the Doob martingale is a member of the family $(M(\gamma))$ we optimize over! In more specific terms we will prove the following.

Proposition 3.1 *Let us set*

$$\mathcal{Z}(\gamma) = \max(0, Z_1 - M_1(\gamma), Z_2 - M_2(\gamma)),$$

and consider

$$\mathcal{Z}_n(\gamma) := \frac{1}{n}\sum_{i=1}^{n} \mathcal{Z}^{(i)}(\gamma)$$

for a sample $\mathcal{Z}^{(i)}(\gamma) := \max(0, Z_1^{(i)} - M_1^{(i)}(\gamma), Z_2^{(i)} - M_2^{(i)}(\gamma))$, $i = 1, ..., n$. *Further define*

$$\gamma_+ := \frac{y_+}{y_+ - y^*} = 1 + \frac{y^*}{y_+ - y^*} > 1 \quad \textit{and} \quad \gamma_- := \frac{y_-}{y_- - y^*} < 0. \tag{3.3}$$

(i) One then has that

$$\mathsf{E}\left[\mathcal{Z}(\gamma)\right] = y^* \ \textit{and} \ \mathsf{Var}\left[\mathcal{Z}(\gamma)\right] = (1-\gamma)^2 \mathsf{Var}\left[Y_1^*\right] \quad \textit{for} \quad \gamma_- \leq \gamma \leq \gamma_+.$$

Hence the only surely optimal martingale is obtained for $\gamma = 1$, *that is, for the Doob martingale.*
(ii) With probability one there is a unique

$$\gamma_n^{\textit{inf}} = \arg\min_{\gamma \in \mathbb{R}} \mathcal{Z}_n(\gamma),$$

and one has that $\gamma_n^{\textit{inf}} \notin [\gamma_-, \gamma_+]$. *Furthermore, one has that*

$$\mathsf{E}\left[\mathcal{Z}(\gamma_n^{\textit{inf}})\right] \overset{n\to\infty}{\longrightarrow} y^* \quad \textit{and} \quad \gamma_n^{\textit{inf}} \overset{n\to\infty}{\longrightarrow} \{\gamma_-, \gamma_+\} \quad \textit{almost surely} \tag{3.4}$$

(meaning that the Euclidean distance between $\gamma_n^{\textit{inf}}$ *and the set* $\{\gamma_-, \gamma_+\}$ *converges to zero), while*

$$\mathsf{Var}\left[\mathcal{Z}(\gamma_\pm)\right] = \left(\frac{y^*}{y^* - y_\pm}\right)^2 \mathsf{Var}\left[Y_1^*\right] > 0. \tag{3.5}$$

The proof of Proposition 3.1 is given in Section 7.3.

We now illustrate the effect of randomization in this explicit example. Let us choose a random variable A, which is uniformly distributed on $[0, K]$ for some $K > y^*$ and independent of $\mathcal{F}_2$, and define

$$\mathcal{Z}_A(\gamma) = \max(A, Z_1 - M_1(\gamma), Z_2 - M_2(\gamma)). \tag{3.6}$$

Thanks to (7.4) below, there is an $\epsilon > 0$ such that $\mathcal{Z}(\gamma) \leq K$ *a.s.* for every $\gamma \in (1-\epsilon, 1+\epsilon) \subset [\gamma_-, \gamma_+]$. By (2.2),

$$\mathsf{E}[\mathcal{Z}_A(\gamma)] = \frac{1}{2K}\mathsf{Var}[\max_{i=0,\ldots,J}(Z_i - M_i(\gamma))] + \textit{const.}, \quad \gamma \in (1-\epsilon, 1+\epsilon),$$

since the corresponding martingales $M(\gamma)$ are weakly optimal by Proposition 3.1, (i). The variance estimates in Proposition 3.1, (i), now imply that $\gamma \mapsto \mathsf{E}[\mathcal{Z}_A(\gamma)]$ has a unique local minimum in $(1-\epsilon, 1+\epsilon)$ at $\gamma = 1$. Hence, by convexity, the expectation of (3.6) has a strict global minimum at $\gamma = 1$. If we now apply the SAA approach to the problem

$$\inf_\gamma \mathsf{E}\left[\mathcal{Z}_A(\gamma)\right],$$

then Thm. 4 in Shapiro [2003] straightforwardly implies that the approximate sequence of martingales converges to the Doob martingale of the original (non-randomized) problem as the size of the training sample tends to infinity.

4 Cash-flow randomization

In this section, we discuss in more detail the cashflow randomization motivated at the end of Section 2. Suppose A is a nonnegative square-integrable random variable independent of $\mathcal{F}_J$

and define $Z_0^{(A)} = A$ and $Z_j^{(A)} = Z_j$ for $j = 1, \dots, J$. In this section, we study the minimization problem

$$y_A^* := \inf_{M \in \mathcal{M}} \mathsf{E}[\max_{i=0,\dots,J} (Z_i^{(A)} - M_i)] \tag{4.1}$$

and its connections to the original problem (1.2). Note that the only difference between both problems is the value of the cash-flow process at time 0.

We first show that the minimizers to the auxiliary problem (4.1) are exactly those minimizers to the original problem (1.2), which share the zero variance property of the Doob martingale of Y^*, that is, which belong to

$$\mathcal{M}^{\circ\circ} := \{M \in \mathcal{M} : \max_{i=0,\dots,J} (Z_i - M_i) = y^*, \ a.s.\} \tag{4.2}$$

i.e. $M^* \in \mathcal{M}^{\circ\circ}$. The following theorem additionally entails that every close-to-optimal martingale for the auxiliary problem (4.1) is also close-to-optimal for the original problem (1.2).

Theorem 4.1 *Suppose A is a nonnegative, integrable random variable independent of $\mathcal{F}_J$ such that $\mathsf{P}(\{A \in (y^* - \epsilon, y^*]\}) > 0$ for every $\epsilon > 0$. Then:*

(i) Optimality: *The value of the auxiliary minimization problem (4.1) is given by*

$$\inf_{M \in \mathcal{M}} \mathsf{E}[\max_{i=0,\dots,J} (Z_i^{(A)} - M_i)] = \mathsf{E}[\max\{y^*, A\}],$$

and M is a minimizing martingale for (4.1) *if and only if $M \in \mathcal{M}^{\circ\circ}$.*

(ii) Bias: *For every $M \in \mathcal{M}$ it holds that*

$$\begin{aligned} &\mathsf{E}[\max_{j=0,\dots,J} (Z_j - M_j)] - y^* \\ &\leq \frac{1}{\mathsf{P}(\{A \leq y^*\})} \left(\mathsf{E}[\max_{j=0,\dots,J} (Z_j^{(A)} - M_j)] - \mathsf{E}[\max\{A, y^*\}] \right). \end{aligned}$$

The next theorem is a key to our error analysis. It states that, under appropriate assumptions on the distribution of A, every close-to-optimal martingale for the auxiliary problem has automatically small variance for the original problem. For the exact statement, we introduce the set $\mathcal{M}_{\leq K}$ of all $(\mathcal{F}_i)$-martingales M such that $M_0 = 0$ and

$$\max_{i=0,\dots,J} (Z_i - M_i) \leq K, \quad a.s. \tag{4.3}$$

Theorem 4.2 *Suppose A is $(0, \infty)$-valued, integrable random variable independent of $\mathbb{F}$ with strictly positive, continuous density f_A. Then, for every $K > y^*$ and $M \in \mathcal{M}_{\leq K}$*

$$\begin{aligned} &\mathsf{Var}(\max_{j=0,\dots,J} (Z_j - M_j)) \leq \mathsf{E}[|\max_{j=0,\dots,J} (Z_j - M_j) - y^*|^2] \\ &\leq \frac{8}{3} \left(\min_{x \in [y^*/3, K]} f_A(x) \right)^{-1} \left(\mathsf{E}[\max_{j=0,\dots,J} (Z_j^{(A)} - M_j)] - \mathsf{E}[\max\{A, y^*\}] \right) \end{aligned}$$

Remark 4.1 *Suppose that Z is bounded, i.e. there is a constant $K_Z > 0$ such that*

$$\max_{i=0,\dots,J} Z_i \leq K_Z.$$

If a martingale $M \in \mathcal{M}$ is bounded from below, that is, there is a constant $K_M > 0$ such that

$$\min_{i=0,\dots,J} M_i \geq -K_M,$$

then, $M \in \mathcal{M}_{\leq K}$ for every $K \geq K_M + K_Z$. Note that, for every $M^{\circ\circ} \in \mathcal{M}^{\circ\circ}$,

$$\min_{i=0,\dots,J} M_i^{\circ\circ} \geq -y^* \geq -K_Z$$

So we can restrict the optimization to martingales which are bounded from below by $-K_Z$, and these are included in $\mathcal{M}_{\leq 2K_Z}$

Discussion: Suppose that Z is bounded by some constant K_Z and $K \geq 2K_Z$. We assume that some algorithm is fixed which takes a training sample $\mathcal{D}_n$ of size n as input and computes a martingale in $\mathcal{M}_{\leq K}$ as output. Hence, we may think of this algorithm as a family of maps

$$\hat{M}^n : \mathcal{D}_n \to \mathcal{M}_{\leq K}, \quad n \in \mathbb{N}.$$

Conditionally on the realization of the training sample $\mathcal{D}_n$, we generate N independent copies $Z^l - \hat{M}^{n,l}$, $l = 1, \dots N$, of $Z - \hat{M}$ (testing sample) and study the upper biased estimator

$$y_{n,N}^{(up)} = \frac{1}{N} \sum_{l=1}^{N} \max_{j=0,\dots J} (Z_j^l - \hat{M}_j^{n,l})$$

of y^*. For a complexity analysis based on Theorems 4.1 and 4.2, we assume that we have some control on the bias of the randomized problem, i.e. there are constants c, $\gamma > 0$ such that for every $n \in \mathbb{N}$

$$\mathsf{E}[\max_{j=0,\dots,J} (Z_j^{(A)} - \hat{M}_j^n) | \mathcal{D}_n] - \mathsf{E}[\max\{A, y^*\}] \leq cn^{-\gamma} \tag{4.4}$$

with probability larger or equal to $1 - n^{-2\gamma}$. Then, by Theorems 4.1 and 4.2 and thanks to (4.4),

$$\mathsf{E}[|y_{n,N}^{(up)} - y^*|^2] \leq \frac{c^2}{\mathsf{P}(\{A \leq y^*\})^2} n^{-2\gamma} + \frac{8c}{3} \left(\min_{x \in [y^*/3, K]} f_A(x) \right) \frac{n^{-\gamma}}{N} + (y^* + K)^2 n^{-2\gamma} \tag{4.5}$$

In order to match the error terms we may choose N proportionally to n^γ, and obtain a root mean-square error (RMSE) of the order

$$\sqrt{\mathsf{E}[|y_{n,N}^{(up)} - y^*|^2]} \lessapprox \frac{1}{N}$$

in the size of the testing sample. As this rate beats the standard Monte Carlo rate of 1/2, we observe that, by building approximate martingales based on the bias of the randomized problem as error criterion, we can provably benefit from the zero-variance property of surely optimal martingales.

Let us now assume that

- the cost of the algorithm $\hat{M}^n$ for choosing the martingale grows as n^ρ in the size of the training sample;

- the cost for evaluating a single trajectory of $(\hat{M}_j^n)_{j=0,\ldots,J}$ grows as n^θ.

Then, in the typical situation $\gamma \le 1$ and $\rho \ge \theta + 1$ (i.e., the algorithm for choosing the martingale based on a sample of size n is at least as expensive as evaluating n trajectories of this martingale), the choice $N \approx n^\gamma$ is easily seen to be optimal. Therefore the complexity for achieving a RMSE error of size $\epsilon > 0$ is of the order $\epsilon^{-\rho/\gamma}$. We shall demonstrate by a theoretical example (Remark 5.4) and by a numerical test case (Section 6.1) below, that the complexity for solving the dual problem (1.2) may (under suitable assumptions) even grow at an order ρ/γ smaller than 2, when solving the randomized dual problem via direct empirical minimization.

5 Randomized empirical dual optimization

In this section, we study the bias estimate (4.4), when the martingales are constructed by direct empirical minimization of the randomized dual problem. Our analysis crucially depends on the variance estimate in Theorem 4.2.

We start by fixing a metric space Ψ and a family $(M_j(\psi))_{j=,0,\ldots,J}$ of martingales $M(\psi) \in \mathcal{M}$ parameterized by $\psi \in \Psi$. That is, $M(\psi)$ is adapted to $\mathbb{F}$ and satisfies $M_0(\psi) = 0$ for all $\psi \in \Psi$. Now let A be a nonnegative random variable independent of $\mathbb{F}$ and define

$$\mathcal{Z}_A(\psi) := \max_{j=0,\ldots,J} \left(Z_j^A - M_j(\psi) \right),$$

where $Z_0^A = A$ and $Z_j^A = Z_j$ for $j = 1, \ldots, J$. Let Ψ_0 be a bounded subset of Ψ. In view of Theorem 4.1 we consider the optimization problem

$$\arg\inf_{\psi \in \Psi_0} \mathcal{Q}_A(\psi) \text{ with } \mathcal{Q}_A(\psi) := \mathsf{E}[\mathcal{Z}_A(\psi)] \text{ for any } \psi \in \Psi \tag{5.1}$$

together with the original (non-randomized) dual problem

$$\arg\inf_{\psi \in \Psi_0} \mathcal{Q}(\psi) \text{ with } \mathcal{Q}(\psi) := \mathsf{E}[\mathcal{Z}(\psi)] \text{ for any } \psi \in \Psi, \tag{5.2}$$

where

$$\mathcal{Z}(\psi) = \max_{j=0,\ldots,J} \left(Z_j - M_j(\psi) \right).$$

Since the expectation in (5.1) ((5.2)) can not be computed in closed form, we shall replace (5.1) ((5.2)) via the corresponding empirical dual optimization problem (EDO) on a set of trajectories. To this end, we define the product space $(\Omega^{\mathbb{N}}, \widetilde{\mathbb{F}}^{\mathbb{N}}, \mathsf{P}^{\mathbb{N}})$, where (with slight abuse of notation) $\widetilde{\mathbb{F}} = \mathbb{F} \vee \sigma(A)$, and its natural projections

$$\Pi_i(\underline{\omega}) = \omega_i, \quad \underline{\omega} = (\omega_n)_{n \in \mathbb{N}} \in \Omega^{\mathbb{N}}$$

as well as the processes $\mathcal{Z}_A^{(i)}$, $i = 1, 2, \ldots$ on $(\Omega^{\mathbb{N}}, \widetilde{\mathbb{F}}^{\mathbb{N}}, \mathsf{P}^{\mathbb{N}}) \times \Psi$ via

$$\mathcal{Z}_A^{(i)}(\underline{\omega}, \psi) := \max_{j=0,\ldots,J} \left(Z_j^{A(\Pi_i(\underline{\omega}))}(\Pi_i(\underline{\omega})) - M_j(\Pi_i(\underline{\omega}); \psi) \right).$$

Let ψ_n^A denote one of the random solutions of the EDO problem

$$\inf_{\psi \in \Psi_0} \mathcal{Q}_{A,n}(\psi) := \inf_{\psi \in \Psi_0} \left\{ \frac{1}{n} \sum_{i=1}^{n} \mathcal{Z}_A^{(i)}(\underline{\omega}, \psi) \right\}, \tag{5.3}$$

If $n \to \infty$ then this optimization problem becomes $\mathsf{P}^{\mathbb{N}}$-a.s. close to the optimization problem (5.1), and we denote by $\psi^{\circ\circ}$ one of its (deterministic) surely optimal solutions. Let us now analyze the properties of the measurable selector ψ_n^A. Set for any $n \in \mathbb{N}$,

$$\mathcal{D}_n : \Omega^{\mathbb{N}} \to (\mathbb{R}^n)^{\Psi}, \quad \underline{\omega} \mapsto \left(\mathcal{Z}_A^{(1)}(\underline{\omega}, \psi), \ldots, \mathcal{Z}_A^{(n)}(\underline{\omega}, \psi) \right)_{\psi \in \Psi}.$$

The mapping $\mathcal{D}_n$ can be interpreted as a set of Monte Carlo paths of the process $\mathcal{Z}_A(\psi)$ used to construct ψ_n^A. We assume that there is $\psi^{\circ\circ} \in \Psi$ such that $M(\psi^{\circ\circ}) \in \mathcal{M}^{\circ\circ}$ and hence

$$\psi^{\circ\circ} \in \arg\min_{\psi \in \Psi} \mathcal{Q}_A(\psi),$$

that is, $\mathcal{Q}_A(\psi^{\circ\circ}) = \mathsf{E}\left[\max(A, y^*)\right]$. Let us introduce the (semi-)metric

$$d(\psi, \psi') := \mathsf{E}[|\mathcal{Z}(\psi) - \mathcal{Z}(\psi')|^2]^{1/2}. \tag{5.4}$$

We further need the following quantitative uniqueness assumption. There exists $\lambda_{\min}$ such that

$$\mathcal{Q}_A(\psi) - \mathcal{Q}_A(\psi^{\circ\circ}) \geq \lambda_{\min}\, d^2\left(\psi, \psi^{\circ\circ}\right). \tag{5.5}$$

In view of Theorem 4.2, the inequality (5.5) holds with $\lambda_{\min} = \frac{3}{8}\left(\min_{x \in [Y^*/3, K]} f_A(x)\right)$, provided that (4.3) holds with $K > y^*$. Now we are prepared to analyze the properties of the minimizer ψ_n^A.

For a bounded subset Ψ_0 in a metric space Ψ endowed with a metric ρ, the covering number $\mathcal{N}(\delta, \Psi_0, \rho)$ is defined as the smallest number of balls of radius δ in the ρ-metric needed to cover Ψ_0, that is, the smallest value of N such that there exist $\psi_1, \ldots, \psi_N \in \Psi_0$, satisfying $\min_{j=1,\ldots,N} \rho(\psi, \psi_j) \leq \delta$ for any $\psi \in \Psi_0$. The following result concerns the optimization problem (5.3) and shows how close is its solution ψ_n^A to $\psi^{\circ\circ}$.

Theorem 5.1 *Assume that* (5.5) *holds with* $\lambda_{\min} \leq 1$, $|\mathcal{Z}| \leq H_{\max}$ *a.s. for a constant* $H_{\max} \geq 1$ *and*

$$\sup_{\substack{\psi, \psi' \in \Psi_0 \\ d(\psi, \psi') \neq 0}} \mathsf{E}\left[\exp\left(\frac{|\mathcal{Z}_A(\psi) - \mathcal{Z}_A(\psi')|^2}{B d^2(\psi, \psi')}\right)\right] \leq L_{\max} \tag{5.6}$$

for some constant $B > 0$. *Suppose also that, for all* $0 < u \leq 2H_{\max}$,

$$\log[1 + \mathcal{N}(\Psi_0, d, u)] \leq \varkappa \left(\frac{1}{u}\right)^{\alpha} \tag{5.7}$$

for some constants $\varkappa = \varkappa(\Psi_0) \geq 1$ *and* $\alpha \in (0, 2)$. *Set* $\overline{\psi} \in \arg\inf_{\psi \in \Psi_0} \mathcal{Q}_A(\psi)$, *then for all* $n \geq n_0$ *and all* $t > 0$, *with probability at least* $1 - 4e^{-t}$,

$$0 \leq \mathcal{Q}_A(\psi_n^A) - \mathcal{Q}_A(\psi^{\circ\circ}) \lesssim \max\left\{ \lambda_{\min}^{-(2-\alpha)/(2+\alpha)} \left(\frac{\varkappa L_{\max}^2}{n}\right)^{2/(2+\alpha)}, \frac{t H_{\max}^2}{n\, \lambda_{\min}} \right\}, \tag{5.8}$$

provided that $\mathcal{Q}_A(\overline{\psi}) - \mathcal{Q}_A(\psi^{\circ\circ}) \lesssim n^{-2/(2+\alpha)}$, *where* $\lesssim$ *stands for inequality up to a constant not depending on* n, t.

Remark 5.2 *Let us note that the set* Ψ_0 *may grow with* n *to make* $\overline{\psi}$ *satisfy* $\mathcal{Q}_A(\overline{\psi}) - \mathcal{Q}_A(\psi^{\circ\circ}) \lesssim n^{-2/(2+\alpha)}$, *then the constant* $\varkappa$ *can also increase in dependence on* n.

Upon getting ψ_n^A, we generate a new set of N trajectories of the process $\mathcal{Z}(\psi_n^A)$ independent of $\mathcal{D}_n$ and consider the Monte Carlo estimate of Y^*

$$Y_{N,n}^A := \frac{1}{N}\sum_{l=1}^{N} \mathcal{Z}^{(l)}(\psi_n^A).$$

Then, in view of Theorem 5.1 and the discussion at the end of Section 4, we obtain

$$\mathsf{E}[|y^* - Y_{N,n}^A|^2] \lesssim \left(\frac{\kappa}{n}\right)^{4/(2+\alpha)} + \left(\frac{\kappa}{n}\right)^{2/(2+\alpha)} \frac{1}{N}, \tag{5.9}$$

and recall that the choice $N \approx (n/\kappa)^{2/(2+\alpha)}$ is optimal in typical settings.

Example 5.3 *Consider a linear parametric family of martingales*

$$M_j(\boldsymbol{\alpha}) := \alpha_1 M_j^1 + \ldots + \alpha_K M_j^K, \quad j = 0, \ldots, J, \quad \boldsymbol{\alpha} \in \mathbb{R}^K, \tag{5.10}$$

where $M^1, \ldots, M^K$ are given $\mathbb{F}$-martingales. In this case the optimization problem (5.3) *boils down to*

$$\operatorname*{arg\,min}_{\boldsymbol{\alpha}\in\mathbb{R}^K} \left\{ \frac{1}{n}\sum_{i=1}^{n} \max_{0\le j\le J} \left(Z_j^{A,(i)} - M_j^{(i)}(\boldsymbol{\alpha}) \right) \right\} \tag{5.11}$$

which in turn can be written as a linear program:

$$\operatorname*{arg\,min}_{\boldsymbol{\alpha}\in\mathbb{R}^K,\, z\in\mathbb{R}^n} \frac{1}{n}\sum_{i=1}^{n} z_i \tag{5.12}$$

under the constraints

$$z_i \ge Z_j^{A,(i)} - M_j^{(i)}(\boldsymbol{\alpha}), \quad i = 1, \ldots, n, \quad j = 0, \ldots, J.$$

Indeed, this is an optimization problem of the type $\arg\min_x f^\top x$ *under* $\mathcal{A}x \le b$ *with* $x = (z_1, \ldots, z_n, \alpha_1, \ldots, \alpha_K) \in \mathbb{R}^{n+K}$, $f = (1/n, \ldots, 1/n, 0, \ldots, 0)$, *a very sparse matrix* $\mathcal{A} \in \mathbb{R}^{n\cdot(J+1)} \times \mathbb{R}^{n+K}$ *with* $n \cdot J \cdot K + n \cdot (J+1)$ *nonzero entries and a vector* $b \in \mathbb{R}^{n\cdot(J+1)}$. *This linear programming formulation has already been exploited by Desai et al. [2012] for the non-randomized empirical dual minimization (i.e., $A = 0$). They also argue that due to the sparse structure of the matrix $\mathcal{A}$, the cost for solving this linear program is of the order n for a fixed linear parametric martingale family.*

In the setting of this example, the entropy bound (5.7) *changes to*

$$\log[1 + \mathcal{N}(\Psi_0, d, u)] \le K \log\left(\frac{1}{u}\right) \tag{5.13}$$

where Ψ_0 is a bounded subset in $\mathbb{R}^K$ and the rate in (5.8) *becomes of order $O(K/n)$ with a hidden constant depending on $\lambda_{\min}$ and $L_{\max}$. In order to fulfill the condition $\mathcal{Q}_A(\overline{\psi}) - \mathcal{Q}_A(\psi^{\circ\circ}) \lesssim n^{-2/(2+\alpha)}$, we need to take an increasing sequence $K = K(n)$. The growth of K in n depends on approximation properties of the family $M(\boldsymbol{\alpha})$, $\boldsymbol{\alpha} \in \mathbb{R}^K$ w.r.t. to the set $\mathcal{M}^{\circ\circ}$. As a result the bound* (5.9) *becomes*

$$\mathsf{E}[|y^* - Y_{N,n}^A|^2] \lesssim \left(\frac{K}{n}\right)^2 + \frac{K}{n}\frac{1}{N}.$$

Remark 5.4 *Let us revisit the stylized example in Section 3. Restrict the martingale family to* $\gamma \in [-\gamma_{\max}, \gamma_{\max}]$ *with* $\gamma_{\max}$ *large enough such that* $\gamma_{\max} > \max(|\gamma_-|, \gamma_+)$*, and consider an independent random variable* A *with strictly positive density* f_A *on* $(0, \infty)$ *as in Theorem 4.1. It is easily seen from (7.4) that with* $K = y_+(1 + \gamma_{\max})$*,* $M(\gamma) \in \mathcal{M}_{\leq K}$ *for* $\gamma \in [-\gamma_{\max}, \gamma_{\max}]$*. Thus, (5.9) applies and our randomized algorithm, that is solving the minimization problem*

$$\gamma_{A,n}^{\inf} := \operatorname*{arg\,inf}_{\gamma \in [-\gamma_{\max}, \gamma_{\max}]} \frac{1}{n} \sum_{i=1}^{n} \max(A^{(i)}, \mathcal{Z}^{(i)}(\gamma)),$$

and subsequently computing the upper bound estimate

$$y_{n,N}^{(A)} := \frac{1}{N} \sum_{i=1}^{N} \mathcal{Z}^{(i)}(\gamma_{A,n}^{\inf}) \tag{5.14}$$

based on an independent trajectory sample of size $N \approx n^{1/(2+\alpha)}$*, yields a root mean-square error*

$$RMSE \lesssim n^{-2/(2+\alpha)}, \qquad \alpha \in (0, 2). \tag{5.15}$$

Since the stylized martingale family (3.2) is spanned by a fixed finite number of martingales and contains the Doob martingale, we can take $\rho = 1$ *and* $\theta = 0$ *in the complexity analysis at the end of Section 4 and can choose* $\alpha = 0$*, see the previous example. Hence our algorithm achieves a complexity of* ϵ^{-1} *for a RMSE of size* ϵ *in this example.*

In contrast, without randomization we always end up with a test estimator (i.e. (5.14) for $A = 0$*) with positive variance (3.5). As a consequence, no matter how fast the convergence speed of the SAA method for the non-randomized minimization actually is, the resulting complexity will always be bounded from below by the standard Monte Carlo rate of* ϵ^{-2}*.*

Example 5.5 (Martingales based on stochastic integrals) *Let* $(S_t)_{t\geq 0}$ *denote a d-dimensional diffusion process solving the following system of SDE's:*

$$dS_t = a(t, S_t)dt + \sigma(t, S_t)\, dW_t, \quad S_0 = x_0, \tag{5.16}$$

where $a : [0, T] \times \mathbb{R}^D \to \mathbb{R}^D$ *and* $\sigma : [0, T] \times \mathbb{R}^D \to \mathbb{R}^{D \times m}$ *are Lipschitz-continuous in space and 1/2 Hölder-continuous in time, with* m *denoting the dimension of the Brownian motion* $W = (W_1, \ldots, W_m)^T$*. Then the martingale representation theorem implies that any square integrable martingale* $(M_t)_{t\geq 0}$ *with respect to the filtration* $(\mathcal{F}_t)_{t\geq 0}$ *generated by* $(W_t)_{t\geq 0}$ *such that* $M_0 = 0$ *can be represented as*

$$M_t = \int_0^t G_s\, dW_s, \quad t \in [0, T], \tag{5.17}$$

where (G_s) *is an adapted to* $(\mathcal{F}_t)_{t\geq 0}$ *square integrable on* $[0, T]$ *process. Under some conditions it can be shown using the Itô formula, that the Doob martingale* (M_t^*) *of the Snell process* $V_t^* = \operatorname{ess\,sup}_{\tau \in \mathcal{T}, \tau \geq t} \mathsf{E}\left[f(S_\tau) \,|\, \mathcal{F}_t\right]$ *for a function* $f : \mathbb{R}^d \to \mathbb{R}$ *has representation* (5.17) *on* $[0, T]$ *with* $G_s = G(s, X_s)$ *and some measurable function* $G : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ *such that* (G_s) *is square integrable on* $[0, T]$*, see Ye and Zhou [2015]. Hence it seems reasonable to parameterize square integrable martingales adapted to* $(\mathcal{F}_t)_{t\in[0,T]}$ *by a class* Ψ *of functions* $\psi(t, x) = (\psi_1(t, x), \ldots, \psi_m(t, x))$ *satisfying*

$$\int_0^T \mathsf{E}\left[|\psi(t, S_t)|^2\right] dt < \infty$$

in the following way

$$M_t^c = M_t^c(\psi) = \int_0^t \psi(u, S_u)\, dW_u.$$

Note that such type of representations was already used to solve optimal stopping/control problem in dual formulation, see e.g. Wang and Caflisch [2010] and Ye and Zhou [2015]. Also they were used to construct control variates in variance reduction methods, see e.g. Vidales et al. [2021]. By discretizing the time we can consider a family of martingales

$$M_j(\psi) = M_{t_j}^c(\psi), \quad j = 0, \ldots, J, \tag{5.18}$$

where $0 = t_0 < t_1 < \ldots < t_J = T$. *It is clear that* $(M_j(\psi))$ *is adapted to the filtration* $(\mathcal{F}_{t_j})$.

Denote by $\mathcal{H}_p^s([0,T]\times\mathbb{R}^D)$ *the Sobolev space of functions defined on* $[0,T]\times\mathbb{R}^D$, *that is, the set of functions* $f \in L_p([0,T]\times\mathbb{R}^D)$, *such that for every multi-index* r *with* $|r| \le s$ *the mixed partial derivative* $D^r f$ *exists and is in* $L_p([0,T]\times\mathbb{R}^D)$. *Further let* $\beta \in \mathbb{R}$ *and* $\langle x\rangle^\beta = (1+|x|^2)^{\beta/2}$ *for all* $x \in \mathbb{R}^D$. *For* $s - D/p > 0$ *we define the weighted Sobolev-space*

$$\mathcal{H}_p^{s,\beta}([0,T]\times\mathbb{R}^D) = \left\{ f : [0,T]\times\mathbb{R}^D \to \mathbb{R}^m \mid f\cdot\langle x\rangle^\beta \in \mathcal{H}_p^s([0,T]\times\mathbb{R}^D)\right\}.$$

Let π_t *denote the density function of* S_t. *We have for the distance* d *from* (5.4),

$$\begin{aligned} d(\psi,\psi') &\lesssim \left(\int_0^T \mathsf{E}\left|\psi-\psi'\right|^2(t,S_t)dt\right)^{1/2} \\ &= \left(\int_0^T \int_{\mathbb{R}^d} \left|\psi-\psi'\right|^2(t,x)\pi_t(x)\,dx\,dt\right)^{1/2} = \|\psi-\psi'\|_{L_2([0,T]\times\mathbb{R}^D,\mu)}, \end{aligned}$$

where μ *is a finite measure on* $[0,T]\times\mathbb{R}^D$ *with density* $\pi_t(x)$. *Let* $\widetilde{\Psi}$ *be a bounded subset of* $\mathcal{H}_2^{s,\beta}([0,T]\times\mathbb{R}^D)$ *for some* $\beta \le 0$, *then due to [Nickl and Pötscher, 2007, Corollary 4], we have*

$$\log(\mathcal{N}(\widetilde{\Psi}, d, \varepsilon)) \lesssim \varepsilon^{-(D+1)/s}, \quad \varepsilon \to 0, \tag{5.19}$$

provided that

$$\sqrt{\int_{[0,T]}\int_{\mathbb{R}^D} \langle x\rangle^{2(\varrho-\beta)}\pi_t(x)\,dx\,dt} < \infty \tag{5.20}$$

with $\varrho > s - (D+1)/2$. *It is well known (see Theorem 8 in Friedman [2008]) that if the diffusion coefficient* σ *is uniformly elliptic, the coefficients* a *and* σ *are infinitely differentiable in* $[0,T]\times\mathbb{R}^D$ *with bounded derivatives of any order, then* $\partial_t^s\partial_x^r\pi_t(x)$ *exists for all natural* r *and* s. *Moreover, it holds for all* $x \in \mathbb{R}^D$ *and* $t > 0$,

$$|\partial_t^s\partial_x^r\pi_t(x)| \lesssim \frac{1}{t^{(D+|r|)/2+s}}\exp\left(-c\,\frac{|x-x_0|^2}{t}\right), \quad \textit{for some } c > 0.$$

where $\lesssim$ *means that the above inequality holds up to a constant, only depending on* s *and* r. *Hence* (5.20) *is satisfied for an arbitrary large* $\varrho \ge \beta$. *As a result, the condition* (5.7) *holds for* $\widetilde{\Psi} \subset \mathcal{H}_2^s([0,T]\times\mathbb{R}^D) \cap L_\infty([0,T]\times\mathbb{R}^D)$ *with* $\alpha = (D+1)/s$. *Moreover,* $\alpha \in (0,2)$ *if* $D \le 2s-1$, *that is, we the functions* ψ *are sufficiently smooth. As to the condition* (5.6), *it can be directly checked in many situations using the inequality (see Section 4.13 in Liptser and Shiryayev [1989])*

$$\sup_{\substack{\psi,\psi'\in\Psi_0\\ d(\psi,\psi')\neq 0}} \mathsf{E}\left[\exp\left(\frac{|M_\tau(\psi)-M_\tau(\psi')|^2}{\mathsf{E}|M_\tau(\psi)-M_\tau(\psi')|^2}\right)\right] \le L'_{\max}$$

which holds for any stopping time $\tau \in \{0,\ldots,J\}$ *and a finite constant* $L'_{\max} > 0$.

6 Numerical examples

In this section, we illustrate our theoretical results by two numerical examples. The first one is a toy example with two time steps and with a single uniformly distributed random variable as the only source of randomness. In this toy example, the excellent convergence properties of the randomized empirical dual minimization can be illustrated, when the martingale dictionary exhausts a complete orthonormal system. The second numerical example is a typical pricing problem of a Bermudan knock-out max-call option on several stocks. This problem has been previously treated by Desai et al. [2012] and we show that, by adding our randomization, one can significantly improve the numerical results in this example.

6.1 A toy example

In this toy example originally due to Schoenmakers et al. [2013], the cashflow is given by

$$Z_0 = 0, \quad Z_1 = 2U, \quad Z_2 = 1,$$

where U is uniformly distributed on the unit interval. One easily computes the Snell envelope Y^* and its Doob martingale M^* as

$$y^* = Y_0^* = 1.25, \quad Y_1^* = 1 + (2U-1)_+, \quad Y_2^* = 1$$

and

$$M_0^* = 0, \quad M_1^* = M_2^* = (2U-1)_+ - 0.25.$$

Obviously, this example satisfies the assumptions of Proposition 2.1, and so there are weakly optimal martingales leading to a pathwise maximum with arbitrarily large variance for the original (non-randomized) problem.

Since the only source of randomness in this example is a uniform random variable, we may construct martingale dictionaries in terms of the complete orthonormal system of Legendre polynomials for $L^2(\sigma(2U-1), \mathsf{P})$. Let

$$\mathcal{L}_k(x) = \sum_{\kappa=1}^{\lfloor k/2 \rfloor} (-1)^\kappa \frac{(2k-2\kappa)!}{(k-\kappa)!(k-2\kappa)!\kappa! 2^k} x^{k-2\kappa}$$

denote the Legendre polynomial of degree k (normalized to have value 1 at $x = 1$). Then, $(\sqrt{2k+1}\mathcal{L}_k(2U-1))_{k\in\mathbb{N}_0}$ is an orthonormal basis for $L^2(\sigma(2U-1), \mathsf{P})$. For a fixed maximal polynomial degree K, we consider the martingale families

$$\mathcal{M}^{(K)} = \left\{ M^{(K)}(\psi) := \sum_{k=1}^{K} \psi_k \tilde{M}^k, \quad (\psi_1, \ldots, \psi_K) \in \mathbb{R}^K \right\},$$

where $\tilde{M}_0^k = 0$ and $\tilde{M}_1^k = \tilde{M}_2^k = \sqrt{2k+1}\mathcal{L}_k(2U-1)$.

We first perform a direct empirical dual minimization without randomization as originally suggested by Desai et al. [2012]. For $K = 1, 3, 5, 7, 9, 11$, we simulate $n = n(K) = 2000K^2$ independent copies U_ν of U (which we refer to as training sample or optimization sample) and consider the empirical minimization problem

$$\inf_{\psi\in\mathbb{R}^K} \frac{1}{n} \sum_{\nu=1}^{n} \left(\max\{2U_\nu, 1\} - \sum_{k=1}^{K} \psi_k \sqrt{2k+1}\mathcal{L}_k(2U_\nu - 1) \right)_+ ,$$

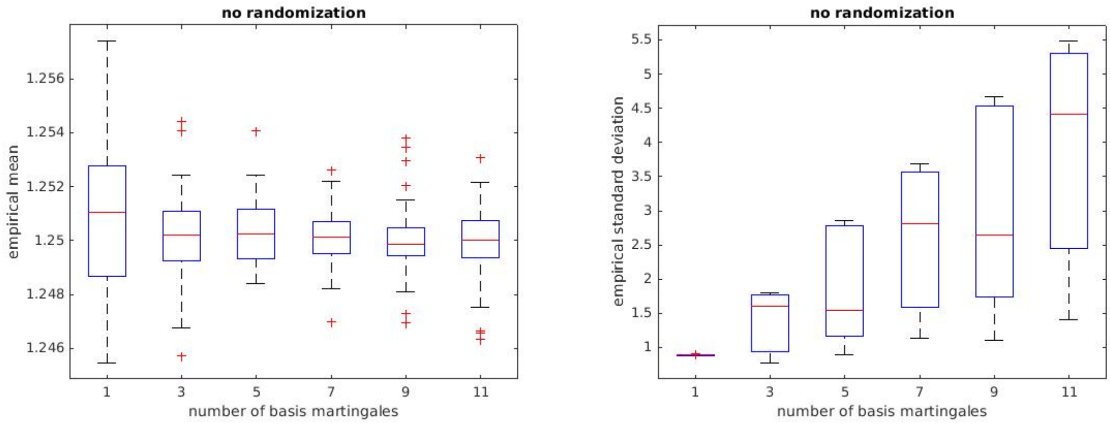

Figure 1: Boxplots of the empirical mean and the empirical standard deviation as a function of the number of basis martingales (without randomization)

noting that $\max_{i=0,1,2}(Z_i - M_i^{(K)}(\psi)) = (\max\{2U, 1\} - M_1^{(K)}(\psi))_+$. We solve the empirical minimization problem by applying MATLAB®'s interior point algorithm to the corresponding linear program (cp. Example 5.3) and denote the resulting coefficient vector by ψ_n. Then, we generate a new independent sample of $N = N(K) = 100000K^2$ uniform random variables in order to estimate the mean and the standard deviations of $\max_{i=0,1,2}(Z_i - M_i^{(K)}(\psi_n))$. The whole algorithm is then repeated 60 times, and the boxplots in Figure 1 exhibit the range of the 60 empirical means and empirical standard deviations computed this way.

The display on the left-hand side shows that the empirical minimization is successful in the sense that the empirical means are close to the correct value $y^* = 1.25$, and that the approximation quality tends to improve as the number of basis martingales increases. However, the display on the right-hand side illustrates that the range and the median of the empirical standard deviations increase with the number of basis martingales. One explanation is the following: For a larger polynomial degree K, there is an increasing number of martingales in $\mathcal{M}^{(K)}$, which are 'close' to some weakly optimal martingale, and the interior point algorithm appears to have a tendency to come up with a variety of different 'close-to-optimal' martingales favoring those with a large variance. This example, thus, demonstrates the need for some variance regularization.

We therefore apply the randomized empirical minimization, which was introduced and studied in the previous sections, and choose $A = |\xi|$, where ξ is Gaussian with mean 1 and variance 1 as the initial distribution for $Z^{(A)}$. In view of Theorem 5.1, the optimal parameter choice for the number $n(K)$ of samples for the empirical minimization depends on the behavior of the bias of the randomized problem

$$\inf_{\psi \in \mathbb{R}^K} \mathsf{E}[\max_{i=0,1,2}(Z_i^{(A)} - M_i^{(K)}(\psi))] - \mathsf{E}[\max\{A, y^*\}] =: B(K)$$

as function in the polynomial degree K. A direct computation based on the Rodrigues formula and integration by parts shows that

$$M_1^* = \sum_{k=1}^{\infty} c_k^* \sqrt{2k+1} \mathcal{L}_k(2U_\nu - 1),$$

where

$$c_1^* = \frac{1}{2\sqrt{3}}, \quad c_{2k}^* = \frac{\sqrt{4k+1}(2k-2)!(-1)^{k-1}}{2^{2k+1}(k+1)!(k-1)!}, \quad c_{2k+1}^* = 0, \quad k \geq 1.$$

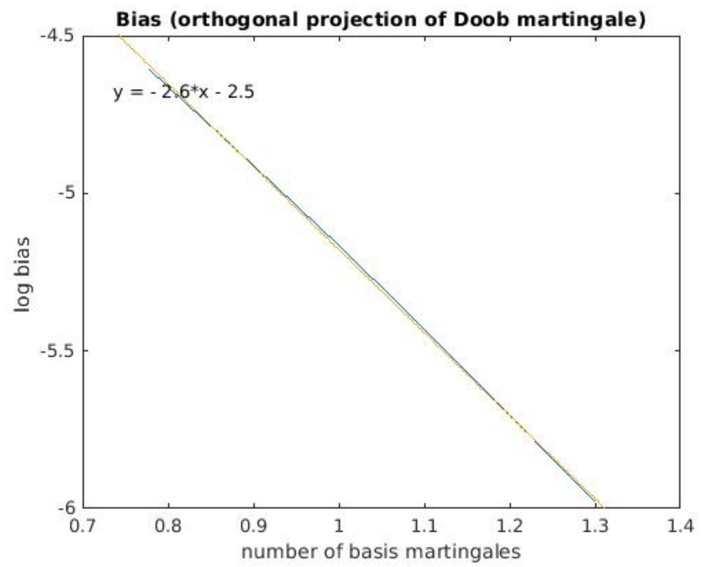

Figure 2: log-log plot of the bias for the orthogonal projection $M^{(K)}(c_1^*,\dots,c_K^*)$ of the Doob martingale as function of the number of basis martingales (with randomization).

Figure 2 shows a log-log plot of the bias of the pathwise maximum corresponding to the orthogonal projection of M^* on $\mathcal{M}^{(K)}$, i.e. the mapping

$$K \mapsto \mathsf{E}[\max_{i=0,1,2}(Z_i^{(A)} - M_i^{(K)}(c_1^*,\dots,c_K^*))] - \mathsf{E}[\max\{A, y^*\}]$$

It suggests that $B(K) = O(K^{-2.6})$. Tailoring the number of training samples $n = n(K)$ for the optimization to the limiting case $\alpha = 0$ in Theorem 5.1 (cp. Example 5.3), we let $n(K) = 500 \cdot \lceil K^{2.6} \rceil$. The constraint matrix of the corresponding linear program is sparse with the number of non-zero entries growing proportionally to $Kn = n^{18/13}$. The log-log-plot of the average run time of 60 repetitions of the empirical minimization (Figure 3) roughly confirms that the run time is of the order $18/13 \approx 1.38$ in the number of training samples.

We again simulate a second independent sample of N trajectories and apply it to estimate the mean of $\max_{i=0,1,2}(Z_i - M_i^{(K)}(\psi_n^A))$, where ψ_n^A is the (random) coefficient vector computed by MATLAB®'s interior point algorithm applied to the empirical randomized problem based on the training sample. The discussion following Theorem 5.1 (with $\alpha = 0$) suggests that the root mean-square error between the sample mean and the true value y^* is of the order $1/n + \sqrt{1/n \cdot 1/N}$. We, thus, choose $N = 2000 \cdot \lceil K^{2.6} \rceil$ proportionally to n to calibrate the algorithm to an error of the order $1/n$. Note that the cost for the evaluation of a single path $M^{(K)}(\psi_n^A)$ grows linearly in K in practice for polynomials of low degree and so the cost for this evaluation step of the algorithm is of the same order as for the optimization step.

A log-log plot of the empirical root mean-square error between sample mean and the true value y^* (over 60 runs of the algorithm) as a function of the number n of optimization samples is shown in Figure 4. The empirical rate of 0.88 of the root mean-square error in the number of samples is slightly below the optimal limiting rate of 1 derived from Theorem 5.1, but, nonetheless we obtain a rate of about 0.68 for the root mean-square error in the run time for this toy example, which even beats the Monte-Carlo rate of 1/2 for computing a single expectation. To summarize, this toy example shows that the randomization of the initial value Z_0 of the cashflow in the direct empirical dual minimization does not only stabilize the algorithm, but significantly improves its convergence behavior.

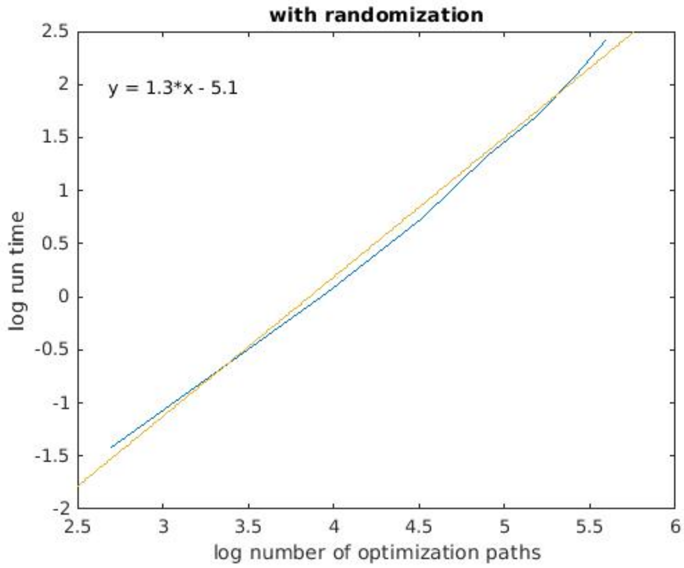

Figure 3: log-log plot of the average run time (60 repetitions) as a function of the number of samples for the empirical minimization (with randomization).

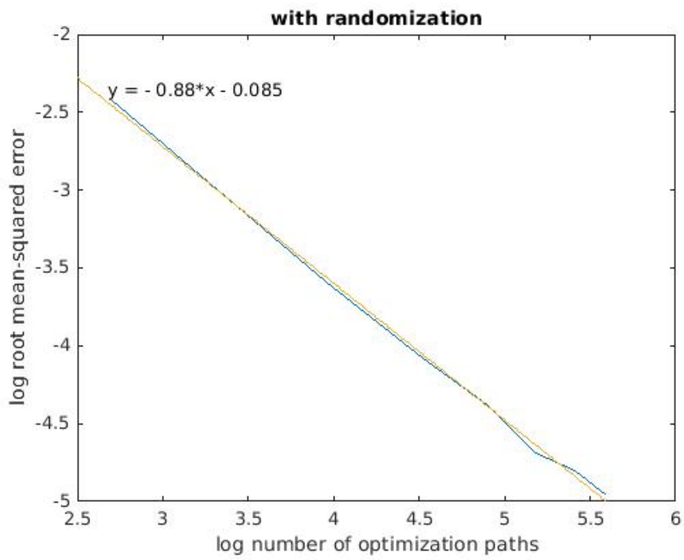

Figure 4: log-log plot of the empirical root mean-square error (over 60 runs) as a function of the number of samples for the empirical minimization (with randomization).

6.2 A Bermudan max-call option

The second numerical example is a typical test case of a pricing problem for a multi-asset Bermudan option and is taken from Desai et al. [2012].

The discounted cashflow of a knock-out max-call option on D stocks $S^1, \dots S^D$ is given by

$$Z_i = e^{-rt_i}\left(\max_{d=1,\dots,D} S^d_{t_i} - K_1\right)_+ \cdot \prod_{j=1}^{i} \mathbf{1}_{\{\max_{d=1,\dots,D} S^d_{t_j} \le K_2\}}, \quad i = 1, \dots J.$$

The option can be exercised at times $t_i = iT/J$, $i = 1, \dots, J$ for some maturity $T > 0$, r denotes the riskless interest rate, K_1 is the strike price of the option, and K_2 is the knock-out level. Note that the option expires worthless, if one of the stock prices exceeds the knock out level K_2, before the option is exercised. We assume that the stock prices follow independent identically distributed Black-Scholes models (under the risk-neutral pricing measure), i.e.

$$S^d_t = s_0 e^{\sigma W^d_t + (r-\sigma^2/2)t}, \quad 0 \le t \le T,$$

where W^d, $d = 1, \dots, D$, are independent standard Brownian motions, s_0 is the initial stock price and σ is the stock volatility. Following one of the specifications in Desai et al. [2012], we choose:

$$D = 4, \quad J = 54, \quad T = 3, \quad K_1 = s_0 = 100, \quad K_2 = 170, \quad r = 0.05, \quad \sigma = 0.2.$$

Desai et al. [2012] state a price estimate with a negative bias of 41.541 and one with a positive bias of 43.853 for this problem, where the latter one was computed via direct empirical dual minimization and serves as a benchmark for our experiments.

Following a standard procedure in financial engineering, we do not attempt to choose generic basis martingales and to run the algorithm into convergence, but rather fix a problem-specific martingale basis. We aim at showing that the variance benefits of the randomized dual minimization also lead to significant improvements in this pre-limit situation. Specifically, we construct martingale families as follows. Denote by

$$y_i = \prod_{j=1}^{i} \mathbf{1}_{\{\max_{d=1,\dots,D} S^d_{t_j} \le K_2\}}$$

the indicator function that the option has not been knocked-out up to time t_i. Let

$$\Delta M^1_i = y_i \left(\max_{d=1,\dots,D} S^d_{t_i} - K_1\right)_+ - E\left[y_i \left(\max_{d=1,\dots,D} S^d_{t_i} - K_1\right)_+ \middle| \mathcal{F}_{t_{i-1}}\right]$$

and

$$\Delta M^2_i = y_i \max_{d=1,\dots,D} S^d_{t_i} - E[y_i \max_{d=1,\dots,D} S^d_{t_i} | \mathcal{F}_{t_{i-1}}],$$

where the conditional expectations are estimated by one layer of nested Monte-Carlo with 500 inner simulations in the practical implementation. We consider the two martingale families

$$\tilde{M}_i(\psi) = \psi_1 \sum_{j=1}^{i} \Delta M^1_j + \psi_2 \sum_{j=1}^{i} \Delta M^2_j, \quad \psi \in \mathbb{R}^2$$

and

$$M_i(\psi) = \sum_{j=1}^{i} \psi_j \Delta M^1_j + \sum_{j=1}^{i} \psi_{J+j} \Delta M^2_j, \quad \psi \in \mathbb{R}^{2J}$$

In the first family, we build two martingales based on the martingale increments ΔM^1, resp. ΔM^2 and fit two parameters, one for each martingale. We refer to this family as the global case or time-independent case below. In the second family, we fit one parameter for each increment ΔM_i^ι of each martingale, and refer to this situation as the local case or time-dependent case. In this time-dependent case, we have to fit a total of 108 parameters (since $J = 54$).

We compare below the cases without randomization of Z_0 and with randomization. In the case of initial randomization we let $Z_0^{(A)} := A := |\xi|$, where ξ is Gaussian with mean a and variance b^2. In order to study the influence of the randomization distribution we consider the three cases

$$(a, b^2) \in \{(30, 40), (30, 4), (40, 40)\}.$$

In the figures below, we denote the various settings by:

- *global:* the two-parametric martingale family $\tilde{M}(\psi)$, $\psi \in \mathbb{R}^2$ and no initial randomization.
- (a, b^2): the time-dependent martingale family $M(\psi)$, $\psi \in \mathbb{R}^{108}$ and the initial randomization based on the Gaussian ξ with mean a and variance b^2.

The general implementation is analogous to the first numerical example. In the training step, we generate n independent copies $(A^\nu, Z_i^\nu, \Delta M_i^{1,\nu}, \Delta M_i^{2,\nu},\ i = 1, \ldots, J)$, $\nu = 1, \ldots, n$, of $(A, Z_i, \Delta M_i^1, \Delta M_i^2,\ i = 1, \ldots, J)$. In the randomized, time-dependent case, we consider the randomized empirical dual minimization problem

$$\operatorname*{arg\,inf}_{\psi \in \mathbb{R}^{108}} \left(\frac{1}{n} \sum_{\nu=1}^{n} \max_{i=0,\ldots,J} (Z_i^{(A),\nu} - M_i^\nu(\psi)) \right)$$

and find a minimizer $\psi^{*,n}$ by applying MATLAB®'s interior point algorithm to the linear programming formulation of this problem. We proceed in the same way in the non-randomized case, replacing $Z^{(A)}$ by Z, and in the global basis case, replacing $M^\nu(\psi)$ by the two-parametric family $\tilde{M}^\nu(\psi)$.

In the testing step, we simulate (with a slight abuse of notation) a new independent sample $(Z_i^\mu, \Delta M_i^{1,\mu}, \Delta M_i^{2,\mu},\ i = 1, \ldots, J)$, $\mu = 1, \ldots, N$, of $(Z_i, \Delta M_i^1, \Delta M_i^2,\ i = 1, \ldots, J)$. This sample is then applied to compute the empirical mean

$$Y_{N,n} = \frac{1}{N} \sum_{\mu=1}^{N} \max_{i=0,\ldots,J} (Z_i^\mu - M_i^\mu(\psi^{*,n})),$$

the empirical standard deviation

$$\Sigma_{N,n} = \sqrt{\frac{1}{N-1} \sum_{\mu=1}^{N} \left(\max_{i=0,\ldots,J} (Z_i^\mu - M_i^\mu(\psi^{*,n})) - Y_{N,n} \right)^2}$$

of the pathwise maximum for the original non-randomized problem, and the 97.5% upper confidence bound $Y_{N,n} + 1.96 \cdot \Sigma_{N,n}$ for the option price (with M replaced by $\tilde{M}$ in the global basis case).

In our simulation study, we let $N = 10n$ and vary the number of training paths n. We repeat 20 runs of each algorithm specification. The results are summarized in Table 1 and Figures 5 and 6, which show boxplots of the 20 empirical means and empirical standard deviations of the different algorithm specifications as the number of training (or, optimization) paths varies from $n = 2000$ to $n = 6000$. The 'global' specification (two basis martingales, no randomization) achieves (upper biased) approximations for the option price Y_0^*, which slightly improve

Algorithm \ n	2000	4000	6000
global	43.89	43.84	43.82
no randomization	*44.07*	*43.97*	*43.93*
local	44.67	44.03	43.92
no randomization	*45.87*	*44.63*	*44.37*
local	43.88	43.74	43.68
randomization (30,40)	*44.08*	*43.84*	*43.75*
local	43.85	43.76	43.71
randomization (40,40)	*43.96*	*43.84*	*43.77*
local	43.90	43.75	43.69
randomization (30,4)	*44.09*	*43.90*	*43.78*

Table 1: Upper biased estimate (empirical mean; solid) and 97.5 % upper confidence bound (italic) for the option price in dependence of the number of training samples n for various algorithm specifications. The numbers represent the average over the 20 runs of each algorithms.

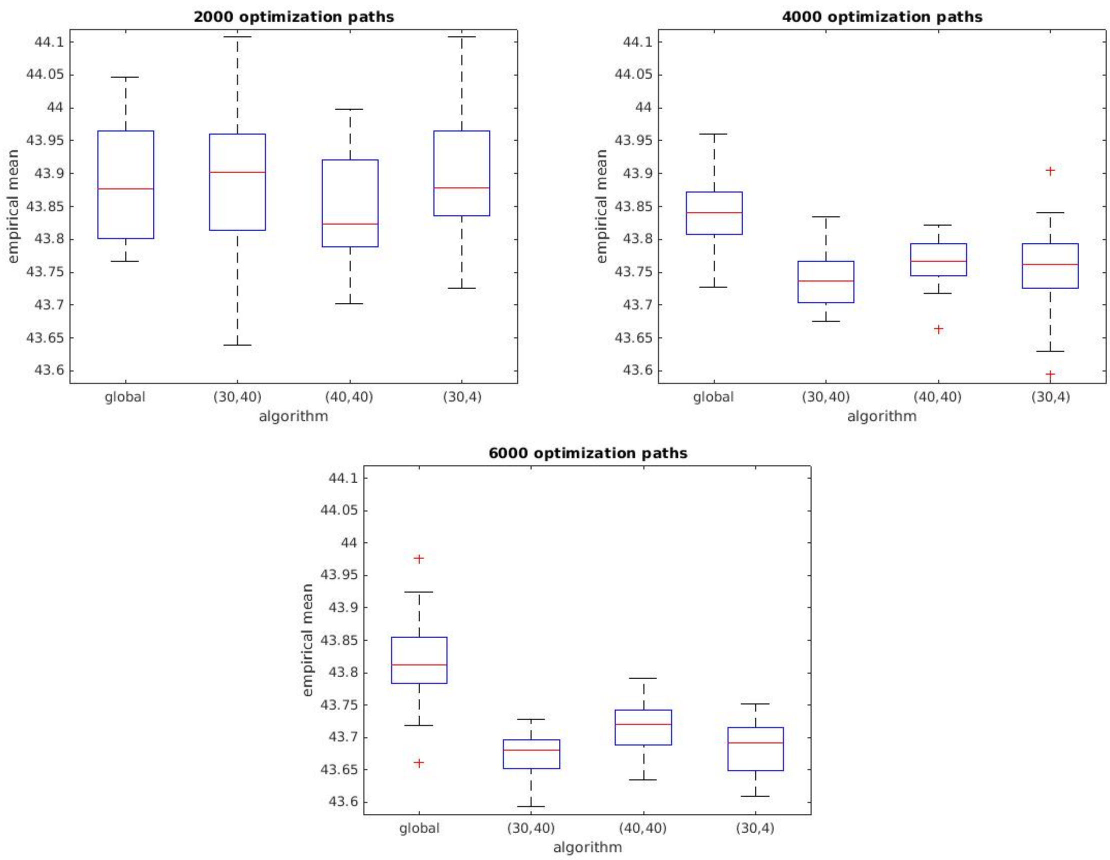

Figure 5: Boxplot of the 20 empirical means for different algorithm specifications for $n = 2000, 4000, 6000$.

as n increases and vary around 43.83. As expected, these estimates without randomization are comparable to the bounds obtained by Desai et al. [2012] with similar, but different basis martingales. In the presence of the randomization, the same moderate number of optimization paths is sufficient to stably fit the time-dependent martingale family with 108 parameters. This larger family leads to an improved upper biased estimate of around 4.7 for the three different randomizations (when $n = 6000$). The variability of the empirical means for each algorithm specification, observed in Figure 5 can be easily explained by the central location of the empirical standard deviation (shown in Figure 6) in conjunction with the moderate number of evaluation paths N. Figure 6 also shows the variance reduction effect of the randomization: The median of the empirical standard deviations under each of the three radomizations is significantly below the one for the nonrandomized global specification. In line with the theoretical result in Theorem 4.2, the variance reduction effect but also a higher stability of the optimization step (for which we may take a low variability of the empirical standard deviations as an indicator) are more pronounced when the randomization distribution has a relatively flat density ($b^2 = 40$ compared to $b^2 = 4$). We also remark that, in each of the algorithm specifications, the run time increases (by and large) linearly in the number of training paths which is proportional to the number of non-zero entries in the sparse constraint matrix of the linear program. Quite surprisingly, there are no significant differences in the run time for the global case (2 parameters, no randomization) and the three time-dependent specifications (108 parameters) with randomization in our implementation. At $n = 6000$ optimization paths, the largest average run-time (over the 20 runs) has even been observed for the time-independent case without randomization.

For comparison, Figure 7 exhibits boxplots for the empirical means and standard deviations for the time-dependent martingale basis run without randomization. The moderate number of optimization paths is clearly not sufficient to obtain a stable fit of the 108 parameters without randomization (note the different scales on the y-axis compared to Figures 5 and 6).

To summarize, in the typical situation in financial engineering, where a small number of basis martingales has been fixed, the initial randomization brings in the opportunity to improve the upper bound estimates by stably fitting time-dependent parameters (one for each martingale increment), which is out of reach for the direct dual minimization algorithm of Desai et al. [2012] with a moderate size of the optimization sample.

7 Proofs

7.1 Preliminaries

For a generic initial time $i = 0, ..., J$ the optimal stopping problem is defined as

$$Y_i^* = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_i} \mathsf{E}[Z_\tau | \mathcal{F}_i],$$

where $\mathcal{T}_i$ denotes the set of $\{i, \ldots, J\}$-valued stopping times. By arguments analogue to (i) and (ii) in Section 2, one obviously has that for every $M \in \mathcal{M}$ and $i = 0, \ldots, J$,

$$\mathsf{E}[\max_{j=i,\ldots,J}(Z_j - (M_j - M_i))|\mathcal{F}_i] \geq \sup_{\tau \in \mathcal{T}_i} \mathsf{E}[\max_{j=i,\ldots,J}(Z_\tau - (M_\tau - M_i))|\mathcal{F}_i] = Y_i^*, \tag{7.1}$$

and moreover for $i = 0, \ldots, J$ one has for the Doob martingale (1.4)

$$\max_{j=i,\ldots,J}(Z_j - (M_j^* - M_i^*)) = Y_i^*, \quad a.s.,. \tag{7.2}$$

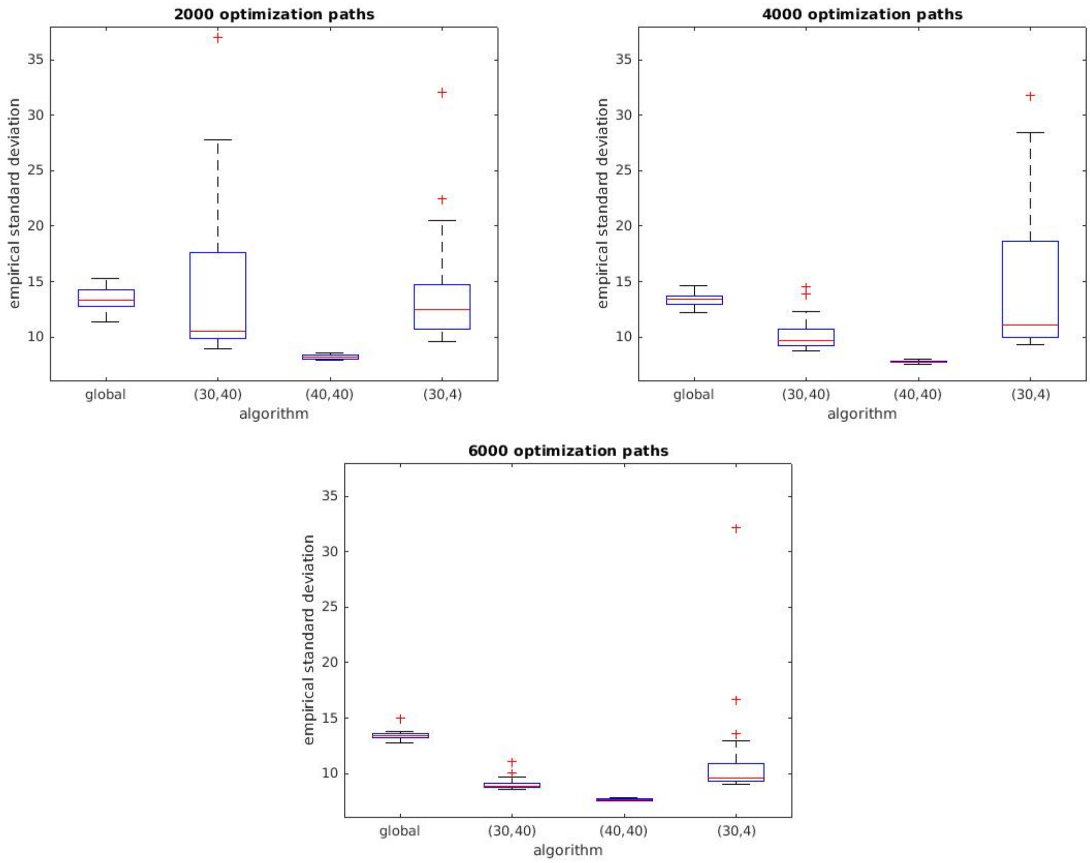

Figure 6: Boxplot of the 20 empirical standard deviations for different algorithm specifications for $n = 2000, 4000, 6000$.

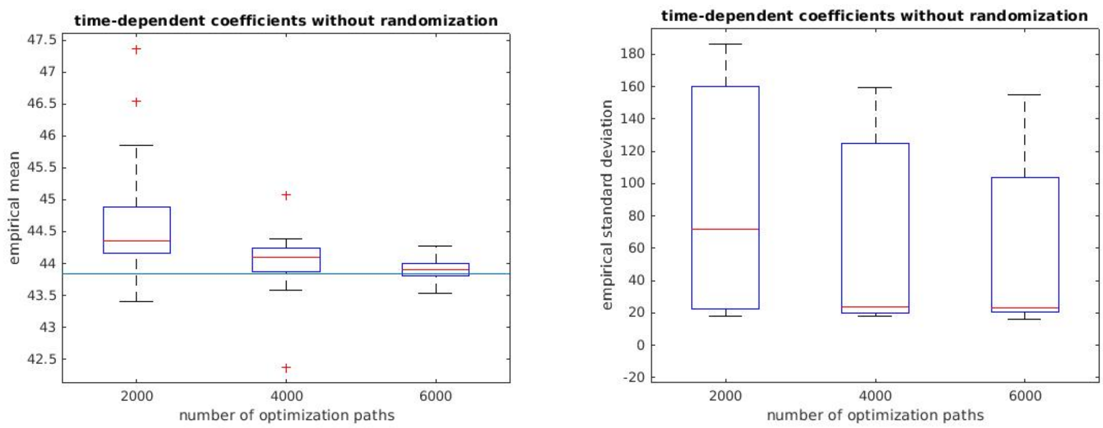

Figure 7: Boxplot of the 20 empirical means (left) and standard deviations (right) for $n = 2000, 4000, 6000$ for the time dependent martingale family $M(\psi)$ without randomization. The solid line (left) shows the average empirical mean obtained from the initial randomization $(30, 40)$ with $n = 6000$ for comparison.

We now provide a characterization of the set $\mathcal{M}^{\circ\circ}$ of surely optimal martingales, which is crucial for the proof of Theorem 4.1. To this end, we introduce the notation

$$\mathcal{M}^{\circ\circ}(a) := \{M \in \mathcal{M} : \max_{i=0,\ldots,J}(Z_i^{(a)} - M_i) = \max\{a, y^*\},\ a.s.\}$$
$$\mathcal{M}^{\circ}(a) := \{M \in \mathcal{M} : E[\max_{i=0,\ldots,J}(Z_i^{(a)} - M_i)] = \max\{a, y^*\}\}$$

for sets of (surely and weakly) optimal martingales, after changing the initial value of the cashflow to the real number $a \geq 0$.

Lemma 7.1 *(i)* $\mathcal{M}^{\circ\circ} = \mathcal{M}^{\circ}(y^*) = \bigcap_{a\geq 0} \mathcal{M}^{\circ\circ}(a)$
(ii) If $M \in \mathcal{M} \setminus \mathcal{M}^{\circ\circ}$*, then there is an* $\epsilon > 0$ *such that* $M \in \mathcal{M} \setminus \mathcal{M}^{\circ}(a)$ *for every* $a \in (y^* - \epsilon, y^* + \epsilon)$.

Proof. (i) Denote

$$\tilde{\mathcal{M}} := \{M \in \mathcal{M} : \max_{1=1,\ldots,J}(Z_i - M_i) = y^*,\ a.s.\}.$$

Since the inclusions $\bigcap_{a\geq 0} \mathcal{M}^{\circ\circ}(a) \subset \mathcal{M}^{\circ\circ}$ and $\bigcap_{a\geq 0} \mathcal{M}^{\circ\circ}(a) \subset \mathcal{M}^{\circ}(y^*)$ are obvious, it suffices to show the inclusions

$$\mathcal{M}^{\circ}(y^*) \subset \tilde{\mathcal{M}}, \quad \mathcal{M}^{\circ\circ} \subset \tilde{\mathcal{M}}, \quad \tilde{\mathcal{M}} \subset \bigcap_{a\geq 0} \mathcal{M}^{\circ\circ}(a).$$

$\mathcal{M}^{\circ}(y^*) \subset \tilde{\mathcal{M}}$: Note first that, for every $M \in \mathcal{M}$,

$$\mathsf{E}[\max_{j=1,\ldots,n}(Z_j - M_j)] \geq y^*. \tag{7.3}$$

Indeed, taking expectation in (7.1) for $i = 1$ yields

$$\mathsf{E}[\max_{j=1,\ldots,n}(Z_j - M_j)] \geq \mathsf{E}[Y_1^*].$$

However $y^* = \max\{\mathsf{E}[Y_1^*], Z_0\} = \mathsf{E}[Y_1^*]$, since $Z_0 = 0$ and Y^* inherits the nonnegativity from Z. We now fix some $M \in \mathcal{M}^{\circ}(y^*)$. Then,

$$\mathsf{E}[\max\{y^*, \max_{i=1,\ldots,n}(Z_i - M_i)\}] = \mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(y^*)} - M_i)] = y^*,$$

which yields $\max_{i=1,\ldots,n}(Z_i - M_i)\} \leq y^*$ a.s. Now (7.3) implies

$$\max_{i=1,\ldots,n}(Z_i - M_i) = y^*, \quad a.s.,$$

i.e. $M \in \tilde{\mathcal{M}}$.
$\mathcal{M}^{\circ\circ} \subset \tilde{\mathcal{M}}$: If $M \in \mathcal{M}^{\circ\circ}$, then

$$\max_{i=0,\ldots,n}(Z_i - M_i) = y^*, \quad a.s.$$

As $y^* > 0$ and $(Z_0 - M_0) = 0$, we obtain

$$\max_{i=1,\ldots,n}(Z_i - M_i) = y^*, \quad a.s.,$$

i.e. $M \in \tilde{\mathcal{M}}$.

$\tilde{\mathcal{M}} \subset \bigcap_{a\geq 0} \mathcal{M}^{\circ\circ}(a)$: Let $M \in \tilde{\mathcal{M}}$. Then, for every $a \geq 0$,

$$\max_{i=0,\dots,J}(Z_i^{(a)} - M_i) = \max\{a, \max_{i=1,\dots,n}(Z_i - M_i)\} = \max\{a, y^*\}, \quad a.s.,$$

i.e. $M \in \mathcal{M}^{\circ\circ}(a)$.

(ii) By (i), if $M \notin \mathcal{M}^{\circ\circ}$, then it is not optimal for $Z^{(y^*)}$, i.e. the map

$$f : \mathbb{R} \to \mathbb{R}, \quad a \mapsto \mathsf{E}[\max_{i=0,\dots,J}(Z_i^{(a)} - M_i)] - \max\{a, y^*\}$$

satisfies: $f(y^*) > 0$. Since f is continuous, there is an $\epsilon > 0$ such that $f(a) > 0$ for every $a \in (y^* - \epsilon, y^* + \epsilon)$, or, equivalently, $M \notin \mathcal{M}^\circ(a)$ for every $a \in (y^* - \epsilon, y^* + \epsilon)$. ■

7.2 Proof of Proposition 2.1

We recall that M^* denotes the Doob martingale of the Snell envelope Y^* of Z. Let $b_k := \mathsf{P}(B_k)$. We define a sequence of martingales by $M_0^k = 0$ and

$$M_i^k = M_i^* + y^*\mathbf{1}_{B_k} - \frac{y^* b_k}{1-b_k}\mathbf{1}_{B_k^c}, \quad i = 1, \dots, J.$$

As M^* is a martingale starting at 0, $B_k \in \mathcal{F}_1$ and $\mathcal{F}_0$ is trivial, the martingale property of M^k is a simple consequence of

$$\mathsf{E}\left[y^*\mathbf{1}_{B_k} - \frac{y^* b_k}{1-b_k}\mathbf{1}_{B_k^c}\right] = 0.$$

Since $Z_0 = 0$ and $y^* > 0$, we observe that $Y_1^* - M_1^* = y^*$. Then, by (7.1),

$$\begin{aligned}\max_{i=0,\dots,J}(Z_i - M_i^k) &= \left(\max_{i=1,\dots,J}(Z_i - M_i^*) - y^*\mathbf{1}_{B_k} + \frac{y^* b_k}{1-b_k}\mathbf{1}_{B_k^c}\right)_+ \\ &= \left(Y_1^* - M_1^* - y^*\mathbf{1}_{B_k} + \frac{y^* b_k}{1-b_k}\mathbf{1}_{B_k^c}\right)_+ = y^* - y^*\mathbf{1}_{B_k} + \frac{y^* b_k}{1-b_k}\mathbf{1}_{B_k^c}.\end{aligned}$$

Hence,

$$\mathsf{E}\left[\max_{i=0,\dots,J}(Z_i - M_i^k)\right] = y^*,$$

i.e. $M^k \in \mathcal{M}^\circ$ for every $k \in \mathbb{N}$. Moreover,

$$\mathsf{Var}\left(\max_{i=0,\dots,J}(Z_i - M_i^k)\right) = y^* b_k \left(1 + \frac{b_k}{1-b_k}\right) \to \infty, \quad (k \to \infty).$$

7.3 Proof of Proposition 3.1

Let us observe that

$$\begin{aligned}\max(Z_1 - M_1(\gamma), Z_2 - M_2(\gamma)) &= -M_1(\gamma) + \max(Z_1, Z_2 + M_1(\gamma) - M_2(\gamma)) \\ &= -\gamma M_1^* + \max(Z_1, Z_2 + \mathsf{E}_{\mathcal{F}_1}[Z_2] - Z_2) \\ &= -\gamma Y_1^* + \gamma y^* + \max(Z_1, \mathsf{E}_{\mathcal{F}_1}[Z_2]) \\ &= (1-\gamma)Y_1^* + \gamma y^* \\ &= Y_1^* + \gamma(y^* - Y_1^*).\end{aligned}$$

Since $0 \le y_- < Y_1^* < y_+$ almost surely, it is not difficult to see that

$$Y_1^* + \gamma(y^* - Y_1^*) = (1-\gamma)Y_1^* + \gamma y^* \ge 0 \quad \text{almost surely if} \quad \gamma_- \le \gamma \le \gamma_+.$$

One thus has $\mathcal{Z}(\gamma) = Y_1^* + \gamma\,(y^* - Y_1^*)$ for $\gamma_- \le \gamma \le \gamma_+$ and so, for $\gamma_- \le \gamma \le \gamma_+$, $\mathsf{E}\left[\mathcal{Z}(\gamma)\right] = y^*$ and $\mathsf{Var}\left[\mathcal{Z}(\gamma)\right] = (1-\gamma)^2\mathsf{Var}\left[Y_1^*\right]$, i.e. (i) is proved.

For arbitrary $\gamma \in \mathbb{R}$ we have

$$\mathcal{Z}(\gamma) = \left(Y_1^* + \gamma\,(y^* - Y_1^*)\right)^+, \tag{7.4}$$

which is piecewise linear in γ, with a kink at $Y_1^*/(Y_1^* - y^*)$ (which is almost surely finite). We so consider the estimator

$$\mathcal{Z}_n(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_1^{*(i)} + \gamma(y^* - Y_1^{*(i)})\right)^+.$$

Obviously, the function $\gamma \to \mathcal{Z}_n(\gamma)$ is nonnegative and convex. Moreover, since Y_1^* has a density, we have almost surely a set of mutually different kink points

$$\gamma^{(i)} = \gamma^{(i)}_{\text{kink}} = \frac{Y_1^{*(i)}}{Y_1^{*(i)} - y^*} = 1 + \frac{y^*}{Y_1^{*(i)} - y^*}, \quad i = 1, ..., n.$$

Thus, we see that

$$\gamma^{(i)} > \gamma_+ \quad \text{if} \quad Y_1^{*(i)} > y^* \quad \text{and} \quad \gamma^{(i)} < \gamma_- \quad \text{if} \quad Y_1^{*(i)} < y^*,$$

hence $\gamma^{(i)}_{\text{kink}} \notin [\gamma_-, \gamma_+]$, $i = 1, ..., n$ with probability one. Suppose that $\mathcal{Z}_n(\gamma)$ is constant on an interval $[\gamma^{(i_0)}, \gamma^{(i_0')}]$ for two neighbored kink points $\gamma^{(i_0)} < \gamma^{(i_0')}$ (hence no further kinks in between). Then, it is easy to see by convexity that

$$\mathcal{Z}_n(\gamma) = \min_{\gamma' \in \mathbb{R}} \mathcal{Z}_n(\gamma') \quad \text{for} \quad \gamma^{(i_0)} \le \gamma \le \gamma^{(i_0')}.$$

Moreover, since $\mathcal{Z}_n(\gamma)$ is differentiable on $(\gamma^{(i_0)}, \gamma^{(i_0')})$ we must have for $\gamma^{(i_0)} < \gamma < \gamma^{(i_0')}$,

$$\begin{aligned} \mathcal{Z}_n'(\gamma) = 0 &= \frac{1}{n}\sum_{i=1}^{n}(y^* - Y_1^{*(i)})H\left(Y_1^{*(i)} + \gamma(y^* - Y_1^{*(i)})\right) \\ &= \frac{1}{n}\sum_{\gamma^{(i)} \le \gamma^{(i_0)}}(y^* - Y_1^{*(i)})H\left(Y_1^{*(i)} + \gamma(y^* - Y_1^{*(i)})\right) \\ &+ \frac{1}{n}\sum_{\gamma^{(i)} > \gamma^{(i_0)}}(y^* - Y_1^{*(i)})H\left(Y_1^{*(i)} + \gamma(y^* - Y_1^{*(i)})\right), \end{aligned} \tag{7.5}$$

where $H = 1_{[0,\infty)}$ denotes the Heaviside function. It is clear that,

$$\begin{gathered} Y_1^{*(i)} + \gamma(y^* - Y_1^{*(i)}) > 0 \quad \text{if and only if} \\ \left(y^* - Y_1^{*(i)} > 0 \quad \text{and} \quad \gamma > \gamma^{(i)}\right) \quad \text{or} \quad \left(y^* - Y_1^{*(i)} < 0 \quad \text{and} \quad \gamma < \gamma^{(i)}\right). \end{gathered} \tag{7.6}$$

Thus (7.5) implies

$$\sum_{\gamma^{(i)} \le \gamma^{(i_0)}}(y^* - Y_1^{*(i)})1_{\left\{y^* - Y_1^{*(i)} > 0\right\}} = \sum_{\gamma^{(i)} > \gamma^{(i_0)}}(Y_1^{*(i)} - y^*)1_{\left\{y^* - Y_1^{*(i)} < 0\right\}},$$

or,

$$\sum_{i=1}^{n}(y^* - Y_1^{*(i)})1_{\left\{y^*-Y_1^{*(i)}\neq 0\right\}} = 0$$

which can only be satisfied on an event of probability zero. We thus conclude that with probability one there is a unique $i_n^{\inf}$ such that

$$\gamma_n^{\inf} := \gamma^{(i_n^{\inf})} \in \left\{\gamma^{(i)} : i = 1, ...n\right\} \quad \text{with} \quad \gamma_n^{\inf} = \arg\min_{\gamma\in\mathbb{R}} \mathcal{Z}_n(\gamma),$$

and in particular $\gamma_n^{\inf} \notin [\gamma_-, \gamma_+]$.

Let us next show that

$$\arg\min_{\gamma\in\mathbb{R}} \mathsf{E}\left[\mathcal{Z}(\gamma)\right] = [\gamma_-, \gamma_+].$$

By (i), it is enough to show that for any $\gamma > \gamma_+$ or $\gamma < \gamma_-$ one has that $\mathsf{E}\left[\mathcal{Z}(\gamma)\right] > y^*$. Indeed, suppose that for some $\gamma > \gamma_+$ one has

$$\mathsf{E}\left[\mathcal{Z}(\gamma)\right] = \mathsf{E}\left[(Y_1^* + \gamma\left(y^* - Y_1^*\right))^+\right] = y^*.$$

It then follows by $y^* = \mathsf{E}\left[Y_1^*\right]$ that

$$\mathsf{E}\left[(Y_1^* + \gamma\left(y^* - Y_1^*\right))^+ - (Y_1^* + \gamma\left(y^* - Y_1^*\right))\right] = 0. \tag{7.7}$$

Then, since for any real x, $x^+ - x \geq 0$, (7.7) implies $(Y_1^* + \gamma\left(y^* - Y_1^*\right))^+ = Y_1^* + \gamma\left(y^* - Y_1^*\right)$ almost surely. However, for $\gamma > \gamma_+$ one has $Y_1^* + \gamma\left(y^* - Y_1^*\right) < 0$ for $Y_1^* > y^* + y^*/(\gamma - 1)$, hence for an event with positive probability, since $y^* + y^*/(\gamma - 1) < y_+$ for $\gamma > \gamma_+$. The case $\gamma < \gamma_-$ goes analogue.

After the above preparations we may now straightforwardly apply Thm. 4 in Shapiro [2003] and conclude that

$$\mathsf{E}\left[\mathcal{Z}(\gamma_n^{i^{\inf}})\right] \overset{n\to\infty}{\to} \inf_{\gamma\in\mathbb{R}} \mathsf{E}\left[\mathcal{Z}(\gamma)\right] = y^* \quad \text{and,}$$
$$[\gamma_-, \gamma_+] \not\ni \gamma_n^{i^{\inf}} \to \{\gamma_-, \gamma_+\} \quad \text{for} \quad n \to \infty \quad a.s. \tag{7.8}$$

The expression in (3.5) follows directly from (i).

7.4 Proof of Theorem 4.1

(i) Note that, by Fubini's theorem,

$$\Phi(M) := \mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(A)} - M_i)] = \int \mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(a)} - M_i)]P_A(da), \quad M \in \mathcal{M}, \tag{7.9}$$

where P_A denotes the distribution of A. If $M \notin \mathcal{M}^{\circ\circ}$, then, by Lemma 7.1(ii), there is an $\epsilon > 0$ such that

$$\mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(a)} - M_i)] > \max\{a, y^*\}, \quad a \in (y^* - \epsilon, y^* + \epsilon).$$

By assumption $P_A((y^* - \epsilon, y^* + \epsilon)) > 0$. Moreover, by the duality (1.2) with $Z^{(a)}$ in place of Z,

$$\mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(a)} - M_i)] \geq \max\{a, y^*\}, \quad a \in \mathbb{R}.$$

Hence,

$$\int \mathsf{E}[\max_{i=0,\ldots,J}(Z_i^{(a)} - M_i)]P_A(da) > \mathsf{E}[\max\{A, y^*\}],$$

which, in view of (7.9), implies $\Phi(M) > \mathsf{E}[\max\{A, y^*\}]$.

If $M \in \mathcal{M}^{\circ\circ}$, then, by Lemma 7.1-(i), $M \in \mathcal{M}^{\circ\circ}(a)$ for every $a \geq 0$, i.e.

$$\mathsf{E}[\max_{i=0,\dots,J}(Z_i^{(a)} - M_i)] = \max\{a, y^*\}, \quad a \geq 0.$$

Integrating the previous identity with respect to P_A yields, thanks to (7.9): $\Phi(M) = \mathsf{E}[\max\{A, y^*\}]$.

2. Consider the function

$$\varphi(a) = \mathsf{E}[\max\{A, y^* + a\} - \max\{A, y^*\}], \quad a \geq 0.$$

Note that

$$\varphi(a) \geq \mathsf{E}[(\max\{A, y^* + a\} - \max\{A, y^*\})\mathbf{1}_{\{A \leq y^*\}}] = a\mathsf{P}(\{A \leq y^*\}), \quad a \geq 0. \tag{7.10}$$

Let $a := \mathsf{E}[\max_{j=0,\dots,J}(Z_j - M_j)] - y^*$ for some fixed $M \in \mathcal{M}$, and note that $\mathsf{P}(\{A \leq y^*\}) > 0$ by assumption. Then we obtain by (7.10) and by Jensen's inequality, exploiting the independence of A and $\mathcal{F}_J$,

$$\begin{aligned}
&\mathsf{E}[\max_{j=0,\dots,J}(Z_j - M_j)] - y^* \\
&\leq \frac{1}{\mathsf{P}(\{A \leq y^*\})}\mathsf{E}[\max\{A, \mathsf{E}[\max_{j=0,\dots,J}(Z_j - M_j)|A]\} - \max\{A, y^*\}] \\
&\leq \frac{1}{\mathsf{P}(\{A \leq y^*\})}\mathsf{E}[\max\{A, \max_{j=0,\dots,J}(Z_j - M_j)\} - \max\{A, y^*\}] \\
&= \frac{1}{\mathsf{P}(\{A \leq y^*\})}\left(\mathsf{E}[\max_{j=0,\dots,J}(Z_j^{(A)} - M_j)] - \mathsf{E}[\max\{A, y^*\}]\right),
\end{aligned}$$

which concludes the proof of the bias estimate.

7.5 Proof of Theorem 4.2

Consider the function

$$\varphi(a) = \mathsf{E}[\max\{A, y^* + a\} - \max\{A, y^*\}], \quad a \in \mathbb{R}.$$

We note that

$$\varphi(a) = \int_{y^*}^{y^*+a} (y^* + a - u)f(u)du + a\mathsf{P}(\{A \leq y^*\}), \quad a \in \mathbb{R}, \tag{7.11}$$

which follows for $a \geq 0$ by a direct computation. The identity also holds for $a < 0$, since

$$\begin{aligned}
\varphi(a) &= \int_{y^*+a}^{y^*} (u - y^*)f(u)du + a\mathsf{P}((\{A \leq y^* + a\}) \\
&= \int_{y^*}^{y^*+a} (y^* + a - u)f(u)du + a\mathsf{P}(\{A \leq y^*\}).
\end{aligned}$$

We now fix $K > y^*$ and $M \in \mathcal{M}_{\leq K}$, and define

$$\Delta := \max_{j=0,\dots,J}(Z_j - M_j) - y^*.$$

As $\mathsf{E}[\Delta] \geq 0$, Eq. (7.11) yields

$$\mathsf{E}[\varphi(\Delta)] \geq \mathsf{E}\left[\int_{y^*}^{y^*+\Delta} (y^* + \Delta - u) f(u) du\right].$$

Let $f^* := \min_{x \in [y^*/3, K]} f(x)$ and note that $f^* > 0$. Then, on the one hand,

$$\begin{aligned}
&\mathsf{E}\left[\int_{y^*}^{y^*+\Delta} (y^* + \Delta - u) f(u) du \mathbf{1}_{\{\Delta \geq -2y^*/3\}}\right] \\
&\geq f^* \mathsf{E}\left[\int_{y^*}^{y^*+\Delta} (y^* + \Delta - u) du \mathbf{1}_{\{\Delta \geq -2y^*/3\}}\right] = \frac{f^*}{2} \mathsf{E}[\Delta^2 \mathbf{1}_{\{\Delta \geq -2y^*/3\}}].
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\mathsf{E}\left[\int_{y^*}^{y^*+\Delta} (y^* + \Delta - u) f(u) du \mathbf{1}_{\{\Delta < -2y^*/3\}}\right] \\
&\geq \mathsf{E}\left[\int_{y^*/3}^{y^*} -(y^* + \Delta - u) f(u) du \mathbf{1}_{\{\Delta < -2y^*/3\}}\right] \\
&\geq \frac{f^*}{2} \mathsf{E}[\Delta^2 \mathbf{1}_{\{\Delta < -2y^*/3\}}] - \frac{f^*}{2} \mathsf{E}[(\Delta + 2y^*/3)^2 \mathbf{1}_{\{\Delta < -2y^*/3\}}].
\end{aligned}$$

Gathering terms, we obtain

$$\mathsf{E}[\Delta^2] \leq \frac{2}{f^*} \mathsf{E}[\varphi(\Delta)] + \mathsf{E}[(\Delta + 2y^*/3)^2 \mathbf{1}_{\{\Delta < -2y^*/3\}}]. \tag{7.12}$$

As $\Delta + y^* \geq 0$, we observe that

$$-y^*/3 \leq \Delta + 2y^*/3 < 0 \quad \text{on} \quad \{\Delta < -2y^*/3\}.$$

Thus, by Markov's inequality,

$$\mathsf{E}[(\Delta + 2y^*/3)^2 \mathbf{1}_{\{\Delta < -2y^*/3\}}] \leq \frac{(y^*)^2}{9} \mathsf{P}(\{|\Delta| > 2y^*/3\}) \leq \frac{1}{4} \mathsf{E}[\Delta^2].$$

Hence, in view of (7.12) and the independence of A and Δ,

$$\begin{aligned}
\mathsf{E}[\Delta^2] &\leq \frac{8}{3f^*}\left(\mathsf{E}[\max\{A, \max_{j=0,\ldots,J}(Z_j - M_j)\} - \max\{A, y^*\}]\right) \\
&= \frac{8}{3f^*}\left(\mathsf{E}[\max_{j=0,\ldots,J}(Z_j^{(A)} - M_j)] - \mathsf{E}[\max\{A, y^*\}]\right).
\end{aligned}$$

7.6 Proof of Theorem 5.1

Let

$$\delta_n := \mathcal{Q}_A(\psi_n^A) - \mathcal{Q}_A(\psi^{\circ\circ}) \leq H_{\max}.$$

Notice that by the definition of ψ_n^A we get

$$\begin{aligned}
\delta_n &\leq \left(\mathcal{Q}_A(\psi_n^A) - \mathcal{Q}_A(\psi^{\circ\circ})\right) - \left(\mathcal{Q}_{A,n}(\psi_n^A) - \mathcal{Q}_{A,n}(\overline{\psi})\right) \\
&= T_n(\psi_n^A) - T_n(\overline{\psi}) + \mathcal{Q}_A(\overline{\psi}) - \mathcal{Q}_A(\psi^{\circ\circ})
\end{aligned}$$

with $T_n(\psi) := (\mathsf{E} - \mathsf{E}_n)(\mathcal{Z}_A(\psi) - \mathcal{Z}_A(\psi^{\circ\circ}))$ and

$$\mathsf{E}_n \mathcal{Z}_A(\psi) := \frac{1}{n} \sum_{i=1}^{n} \mathcal{Z}_A^{(i)}(\cdot, \psi).$$

Set $\varepsilon_n = n^{-2/(2+\alpha)}$. Since $0 \leq \mathcal{Q}_A(\overline{\psi}) - \mathcal{Q}_A(\psi^{\circ\circ}) \lesssim n^{-2/(2+\alpha)}$, we have

$$\delta_n \lesssim \phi_n(\delta_n),$$

where, for any $\delta > 0$,

$$\phi_n(\delta) := \sup_{\psi \in \Psi(\delta)} T_n(\psi) + |T_n(\overline{\psi})| + \varepsilon_n, \quad \Psi(\delta) := \big\{\psi \in \Psi : \ \mathcal{Q}_A(\psi) - \mathcal{Q}_A(\psi^{\circ\circ}) \leq \delta\big\}. \tag{7.13}$$

The main technical result we will invoke, in order to bound $\bar{\delta}_n$, is the following lemma. This result follows from the combination of arguments presented in Theorem 4.1, Corollary 4.1, and Theorem 4.3 in [Koltchinskii, 2011].

Lemma 7.2 *Let $\{\phi(\delta) : \delta \geq 0\}$ be non-negative random variables (indexed by all deterministic $\delta \geq 0$) such that, almost surely, $\phi(\delta) \leq \phi(\delta')$ if $\delta \leq \delta'$. Let $\{\beta(\delta, t) : \delta \geq 0, t \geq 0\}$, be (deterministic) real numbers such that*

$$\Pr\big(\phi(\delta) \geq \beta(\delta, t)\big) \leq e^{-t}. \tag{7.14}$$

Finally, let $\hat{\delta}$ be a nonnegative random variable, a priori upper bounded by a constant $\bar{\delta} > 0$, and such that, almost surely,

$$\hat{\delta} \leq \phi(\hat{\delta}).$$

Then defining, for all $t \geq 0$,

$$\beta(t) := \inf \left\{ \tau > 0 : \sup_{\delta \geq \tau} \frac{\beta\left(\delta, \frac{t\delta}{\tau}\right)}{\delta} \leq \frac{1}{2} \right\}, \tag{7.15}$$

we obtain, for all $t \geq 0$,

$$\Pr\big(\hat{\delta} \geq \beta(t)\big) \leq 2e^{-t}.$$

According to Theorem 7.2, it remains to bound $\phi_n(\delta)$ with high probability for any fixed $\delta > 0$. The next lemma reduces the problem to that of bounding the expected suprema of empirical processes.

Lemma 7.3 *Suppose that Assumption (5.5) holds and $|\mathcal{Z}| \leq H_{\max}$ a.s. Then, for every $\delta > 0$ and any $t > 0$,*

$$\Pr\big(\phi_n(\delta) \geq \beta_n(\delta, t)\big) \leq 2e^{-t},$$

where,

$$\beta_n(\delta, t) := B \left(\mathsf{E} \sup_{\psi \in \Psi(\delta)} (\mathsf{E} - \mathsf{E}_n)\Delta(\psi) + \sqrt{\frac{H_{\max} t(\delta + \varepsilon_n)}{\lambda_{\min} n}} + \frac{H_{\max}^2 t}{n} + \frac{1}{n^{2/(2+\alpha)}} \right),$$

$\Delta(\psi) := \mathcal{Z}_A(\psi) - \mathcal{Z}_A(\psi^{\circ\circ})$ and $B > 0$ is a universal constant.

Proof. We have $|\Delta(\psi)| \le 2H_{\max}$ for all $\psi \in \Psi$. It then follows from the Bernstein inequality and a version of Talagrand's inequality due to Bousquet (see Theorem A.1) that, with probability at least $1 - 2e^{-t}$,

$$\phi_n(\delta) \lesssim \underbrace{\mathsf{E}\phi_n(\delta) + \sqrt{\frac{t}{n}\big(\sigma^2(\delta) + H_{\max}\,\mathsf{E}\phi_n(\delta)\big)}}_{\text{Bound for } T_n(\psi_n^A)} + \underbrace{\frac{H_{\max}t}{n} + \sqrt{\frac{t}{n}\mathsf{E}[\Delta^2(\overline{\psi})]}}_{\text{Bound for } T_n(\overline{\psi})} + \varepsilon_n, \tag{7.16}$$

where

$$\sigma^2(\delta) = \sup_{\psi\in\Psi(\delta)} \mathsf{E}[\Delta(\psi)]^2.$$

Using basic inequalities $\sqrt{u+v} \le \sqrt{u} + \sqrt{v}$ and $2\sqrt{uv} \le u + v$ for positive numbers u and v, we further deduce that, with probability at least $1 - e^{-t}$,

$$\phi_n(\delta) \lesssim \mathsf{E}\phi_n(\delta) + \left(\sigma(\delta) + \sqrt{\mathsf{E}[\Delta^2(\overline{\psi})]}\right)\sqrt{\frac{t}{n}} + \frac{H_{\max}t}{n}. \tag{7.17}$$

Now let us upper bound the terms $\sigma^2(\delta)$ and $\mathsf{E}[\Delta^2(\overline{\psi})]$. Due to (5.5) we have for every $\psi \in \Psi$,

$$\mathsf{E}[\Delta(\psi)]^2 = \mathsf{E}[|\mathcal{Z}_A(\psi) - \mathcal{Z}_A(\psi^{\circ\circ})|^2] \le d^2(\psi, \psi^{\circ\circ}) \le \lambda_{\min}^{-1}[\mathcal{Q}_A(\psi) - \mathcal{Q}_A(\psi^{\circ\circ})],$$

so that

$$\sigma^2(\delta) \le \lambda_{\min}^{-1}\delta. \tag{7.18}$$

Combining (7.17) with (7.18) and using the fact that $\mathcal{Q}_A(\overline{\psi}) - \mathcal{Q}_A(\psi^{\circ\circ}) \lesssim \varepsilon_n$, we deduce that

$$\phi_n(\delta) \lesssim \mathsf{E}\sup_{\psi\in\Psi(\delta)} (\mathsf{E} - \mathsf{E}_n)\Delta(\psi) + \sqrt{\frac{H_{\max}t(\delta+\varepsilon_n)}{\lambda_{\min}n}} + \frac{H_{\max}^2 t}{n}. \tag{7.19}$$

■

Let us now bound the expected suprema of empirical processes given in $\beta_n(\delta, t)$ under the assumptions on the entropy which follows from Proposition 5.2 in Viens and Vizcarra [2007].

Lemma 7.4 *It holds under assumptions of Theorem 5.1,*

$$\mathsf{E}\sup_{\psi\in\Psi(\delta)} (\mathsf{E} - \mathsf{E}_n)\Delta(\psi) \le \frac{L_{\max}}{\sqrt{n}}\int_0^{\sqrt{\delta/\lambda_{\min}}} \sqrt{\log\left[1 + \mathcal{N}(\Psi_0, d; \varepsilon)\right]}\, d\varepsilon$$

for some constant $C > 0$.

Using the assumption (5.7), we get

$$\frac{\beta_n(\delta, t\delta/\tau)}{\delta} \lesssim \frac{L_{\max}\sqrt{\varkappa(\Psi_0)}}{\sqrt{n\delta^{1+\alpha/2}\lambda_{\min}^{1-\alpha/2}}} + \sqrt{\frac{tH_{\max}}{n\lambda_{\min}\tau}} + \sqrt{\frac{\varepsilon_n tH_{\max}}{n\lambda_{\min}\delta\tau}} + H_{\max}^2\frac{t}{n\tau} + \frac{1}{n^{2/(2+\alpha)}\delta}$$

yielding for

$$\beta_n(t) := \inf\left\{\tau > 0 : \sup_{\delta\ge\tau}\frac{\beta_n\left(\delta, \frac{t\delta}{\tau}\right)}{\delta} \le \frac{1}{2}\right\}$$

the inequality

$$\beta_n(t) \lesssim \max\left\{\lambda_{\min}^{-(2-\alpha)/(2+\alpha)} L_{\max}^{4/(2+\alpha)}\left(\frac{\varkappa(\Psi_0)}{n}\right)^{2/(2+\alpha)}, \frac{tH_{\max}}{n\,(1\wedge\lambda_{\min})}\right\}$$

for n large enough.

A Supplementary Material

In this section we have compiled some standard facts on empirical processes. We start with the well-known concentration result, known as Bousquet's form of Talagrand's inequality for empirical processes. It involves a notion of variance of the empirical process

$$\sigma_{\mathcal{H}} := \sup_{h \in \mathcal{H}} \sqrt{\mathsf{E} h^2},$$

which plays a crucial role in many modern proof techniques involving the local behavior of the supremum of empirical process. The proof of the following lemma can be found in [Bousquet, 2002] or [Giné and Nickl, 2016].

Lemma A.1 *Suppose that all functions in $\mathcal{H}$ are $[a, b]$-valued, for some $a < b$. Then, for all $n \geq 1$ and all $t > 0$,*

$$\sup_{h \in \mathcal{H}} (\mathsf{E} - \mathsf{E}_n) h \leq \mathsf{E} \sup_{h \in \mathcal{H}} (\mathsf{E} - \mathsf{E}_n) h + \sqrt{\frac{2t}{n} \Big(\sigma_{\mathcal{H}}^2 + 2(b-a)\, \mathsf{E} \sup_{h \in \mathcal{H}} (\mathsf{E} - \mathsf{E}_n) h \Big)} + \frac{(b-a)t}{3n},$$

with probability larger than $1 - e^{-t}$.

Lemma A.2 *Let (X_t), $t \geq 0$ be an $\mathbb{R}$-valued, continuous square integrable martingale with $X_0 = 0$, and τ a bounded stopping time. Then, it holds for every $\varepsilon > 0$, $\Gamma > 0$,*

$$\mathsf{P} \left(\sup_{t \in [0, \tau]} |X_t| > \varepsilon, [X]_\tau \leq \Gamma \right) \leq 2 \exp \left(-\frac{\varepsilon^2}{2\Gamma} \right), \tag{A.1}$$

where $([X]_t)_{t \geq 0}$ is quadratic variation process of X.

References

Leif Andersen. A simple approach to the pricing of Bermudan swaptions in the multi-factor Libor market model. *Journal of Computational Finance*, 3:5–32, 1999.

Leif Andersen and Mark Broadie. A Primal-Dual Simulation Algorithm for Pricing Multi-Dimensional American Options. *Management Science*, 50(9):1222–1234, 2004.

Peter L. Bartlett and Shahar Mendelson. Empirical minimization. *Probab. Theory Related Fields*, 135(3):311–334, 2006. ISSN 0178-8051.

Denis Belomestny. Solving optimal stopping problems via empirical dual optimization. *The Annals of Applied Probability*, 23(5):1988–2019, 2013.

Denis Belomestny and John Schoenmakers. *Advanced simulation-based methods for optimal stopping and control.* Palgrave Macmillan, London, 2018.

Denis Belomestny and John Schoenmakers. From optimal martingales to randomized dual optimal stopping, 2021. arXiv: 2102.01533.

Denis Belomestny, Christian Bender, and John Schoenmakers. True upper bounds for Bermudan products via non-nested Monte Carlo. *Math. Finance*, 19(1):53–71, 2009.

Denis Belomestny, Roland Hildebrand, and John Schoenmakers. Optimal stopping via pathwise dual empirical maximisation. *Appl. Math. Optim.*, 79(3):715–741, 2019.

Olivier Bousquet. A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Math. Acad. Sci. Paris*, 334(6):495–500, 2002.

M. Broadie and P. Glasserman. A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance*, 7(4):35–72, 2004.

David B. Brown, James E. Smith, and Peng Sun. Information relaxations and duality in stochastic dynamic programs. *Oper. Res.*, 58(4, part 1):785–801, 2010.

Shyam S. Chandramouli. A convex optimization approach to multiple stopping: Pricing chooser caps and swing options. *SSRN Preprint*, 2019.

V.V. Desai, V.F. Farias, and C.C. Moallemi. Pathwise optimization for optimal stopping problems. *Management Science*, 58(12):2292–2308, 2012.

Avner Friedman. *Partial differential equations of parabolic type*. Courier Dover Publications, 2008.

Evarist Giné and Richard Nickl. *Mathematical Foundations of Infinite-Dimensional Statistical Models*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, New York, 2016.

Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53. Springer Science & Business Media, 2003.

Martin Haugh and Leonid Kogan. Pricing American options: A duality approach. *Oper. Res.*, 52(2):258–270, 2004.

Mark Joshi and Jochen Theis. Bounding bermudan swaptions in a swap-rate marketmodel. *Quantitative Finance*, 2(5):371–361, 2002.

Vladimir Koltchinskii. *Oracle inequalities in empirical risk minimization and sparse recovery problems*, volume 2033 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 38th Probability Summer School held in Saint-Flour, 2008.

Jérôme Lelong. Dual pricing of American options by Wiener chaos expansion. *SIAM J. Financ. Math.*, 9(2):493–519, 2018.

R Liptser and AN Shiryayev. *Theory of martingales*. Mathematics and its Applications. Kluwer, Dordrecht, 1989.

Francis A. Longstaff and Eduardo S. Schwartz. Valuing American options by simulation: a simple least-squares approach. *Review of Financial Studies*, 14(1):113–147, 2001.

Richard Nickl and Benedikt Pötscher. Bracketing metric entropy rates and empirical central limit theorems for function classes of besov-and sobolev-type. *Journal of Theoretical Probability*, 20(3):177–199, 2007.

L. C. G. Rogers. Pathwise stochastic optimal control. *SIAM Journal on Control and Optimization*, 46(3):1116–1132, 2007.

Leonard C. G. Rogers. Monte Carlo valuation of American options. *Mathematical Finance*, 12 (3):271–286, 2002.

John Schoenmakers, Jianing Zhang, and Junbo Huang. Optimal dual martingales, their analysis, and application to new algorithms for Bermudan products. *SIAM J. Financial Math.*, 4(1):86–116, 2013.

Alexander Shapiro. Monte carlo sampling methods. In A. Ruszczyński and A. Shapiro, editors, *Handbooks in OR & MS*, chapter 6, pages 353–425. Elsevier Science B.V., 2003.

J. Tsitsiklis and B. Van Roy. Regression methods for pricing complex American style options. *IEEE Trans. Neural. Net.*, 12(14):694–703, 2001.

Marc Sabate Vidales, David Siska, and Lukasz Szpruch. Unbiased deep solvers for parametric pdes, 2021. arXiv 1810.05094.

Frederi G Viens and Andrew B Vizcarra. Supremum concentration inequality and modulus of continuity for sub-nth chaos processes. *Journal of Functional Analysis*, 248(1):1–26, 2007.

Yang Wang and Russel Caflisch. Fast computation of upper bounds for American-style options without nested simulation. *Available at SSRN 1612336*, 2010.

Bo Yang, Selvaprabu Nadarajah, and Nicola Secomandi. Pathwise optimization for merchant energy production. *SSRN Preprint*, 2019.

Fan Ye and Enlu Zhou. Information relaxation and dual formulation of controlled Markov diffusions. *IEEE Transactions on Automatic Control*, 60(10):2676–2691, 2015.