Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Optimal dual martingales and their stability; fast evaluation of Bermudan products via dual backward regression

John G. M. Schoenmakers, Junbo Huang

submitted: December 9, 2010

Weierstrass Institute
for Applied Analysis and Stochastics
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: John.Schoenmakers@wias-berlin.de
Junbo.Huang@wias-berlin.de

Preprint No. 1574
Berlin 2010

2000 *Mathematics Subject Classification.* 62L15, 65C05, 91B28.

Key words and phrases. Bermudan options, duality, Monte Carlo simulation, linear regression.

Supported by DFG Research Center MATHEON "Mathematics for Key Technologies" in Berlin.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In this paper we introduce and study the concept of optimal and surely optimal dual martingales in the context of dual valuation of Bermudan options. We provide a theorem which give conditions for a martingale to be surely optimal, and a stability theorem concerning martingales which are near to be surely optimal in a sense. Guided by these theorems we develop a regression based backward construction of such a martingale in a Wiener environment. In turn this martingale may be utilized for computing upper bounds by non-nested Monte Carlo. As a by-product, the algorithm also provides approximations to continuation values of the product, which in turn determine a stopping policy. Hence, we obtain lower bounds at the same time. The proposed algorithm is pure dual in the sense that it doesn't require an (input) approximation to the Snell envelope, is quite easy to implement, and in a numerical study we show that, regarding the computed upper bounds, it is comparable with the method of Belomestny, et. al. (2009).

1 Introduction

It is well-known that evaluation of Bermudan callable derivatives comes down to solving an optimal stopping problem. For many callable exotic products, for example interest products, the underlying state space is high-dimensional however. As such these products are usually very hard to solve with deterministic (PDE) methods and therefore simulation based (Monte Carlo) methods are called for. The first developments in this respect concentrated on the construction of a 'good' exercise policy. We mention, among others, regression based methods by Carriere (1996), Longstaff and Schwartz (2001) and Tsistsiklis and Van Roy (1999), the stochastic mesh method of Broadie and Glasserman (2004), and quantization algorithms by Bally and Pages (2003). Especially for very high dimension, Kolodko and Schoenmakers (2004) developed a policy improvement approach which can be effectively combined with Longstaff and Schwartz (2001) for example (see Bender et al. (2008) and Bender et al. (2006)).

As a common feature, the aforementioned simulation methods provide lower biased estimates for the Bermudan product under consideration. As a new breakthrough, Rogers (2001) and Haugh and Kogan (2004) introduced a dual approach, which comes down to minimization over a set of martingales rather than maximization over a family of stopping times. By its very nature the dual approach gives upper biased estimates for the Bermudan product and after its appearance several numerical algorithms for computing dual upper bounds are proposed. Probably the most popular one is the method of Andersen and Broadie (2004), although this method requires nested Monte Carlo simulation (see also Kolodko and Schoenmakers (2004) and Schoenmakers (2005)). In a Wiener environment, Belomestny, et. al. (2009) provide a fast generic method for computing dual upper bounds by non-nested simulation.

The algorithms for computing dual upper bounds so far have in common that they start with some given 'good enough' approximation of the Snell envelope and then construct the Doob-martingale due to this approximation. In a recent paper Rogers (2010), points out how to construct a particular 'good' martingale via a sequence of martingales which are constant on an ever bigger time interval. In this construction no input approximation to the Snell envelope is used. The approach proposed in this paper has some flavor of the method of Rogers (2010), in the sense that no approximation to the Snell envelope is used either.

Moreover, a detailed analysis of the almost sure property of the dual representation is provided and a new (backward) regression algorithm for constructing a dual martingale is proposed.

We introduce the concept of a surely optimal martingale, loosely speaking, a martingale that minimizes the dual representation with a particular almost sure property (Section 3). In this respect we will point out that a martingale which minimizes the dual representation is not necessarily surely optimal, and on the other hand, a surely optimal martingale is generally not unique. As one of the main contributions in this paper we provide a characterizing criterion that guarantees that a martingale is surely optimal (Theorem 5). In the application of the Andersen and Broadie (2004) algorithm one generally observes that, the better the constructed dual martingale, the lower the variance of the upper bound estimator. Actually this observation was not well studied from a mathematical point of view by now. In Section 4 we study this phenomenon and, as a next contribution, give an explanation of it by a convergence or stability Theorem 8. Finally, based on Theorem 5 and Theorem 8 we develop an algorithm in Section 5 for constructing a martingale which is in a sense near to be surely optimal by a regression procedure in a Wiener environment. As a by-product of this regression procedure we obtain estimations of continuation values as well. As a result we end up with a procedure which computes upper bound as well as lower bounds at the same time by non-nested simulation. As such this new procedure is quite easy to implement and may be considered as an interesting alternative to the non-nested method of Belomestny, et. al. (2009), where a dual martingale is obtained by constructing a discretized Clark-Ocone derivative of some (input) approximation to the Snell envelope via regression. In a numerical study (Section 6) we demonstrate at two multi-dimensional examples that our new method is of the same quality as the one in Belomestny, et. al. (2009) regarding the upper bounds, and it moreover produces very fast good lower bounds. In Section 2 we start with a short resume of well-known facts on Bermudan derivatives.

2 Bermudan derivatives and optimal stopping

Let $(Z_i : i = 0, 1, \ldots, T)$[1] be a non-negative stochastic process in discrete time on a filtered probability space $(\Omega, \mathcal{F}, P)$, adapted to a filtration $\mathbb{F} := (\mathcal{F}_i \,:\, 0 \leq i \leq T)$ which satisfies $E|Z_i| < \infty$, for $0 \leq i \leq T$. The measure P may be considered as a pricing measure and the process Z may be seen as a (discounted) cash-flow, which an investor may exercise once in the time set $\{0, ..., T\}$. As such he is faced with a Bermudan derivative. As a well known fact, a fair price of such a derivative is the value of the Snell envelope process

$$Y_i^* = \sup_{\tau \in \{i, \ldots, T\},} EZ_\tau, \quad 0 \leq i \leq T, \tag{2.1}$$

at time $i = 0$. In (2.1), τ denotes a stopping time, $E_i := E_{\mathcal{F}_i}$ denotes conditional expectation with respect to the σ-algebra $\mathcal{F}_i$, and $\sup\ (\inf)$ is to be understood as *essential supremum* (*infimum*) if it ranges over an uncountable family of random variables. Let us recall some well-known facts (e.g. see Neveu (1975)).

1 The Snell envelope Y^* of Z is the smallest super-martingale that dominates Z.

2 A family of optimal stopping times is given by

$$\tau_i^* = \inf\{j : j \geq i, \quad Z_j \geq Y_j^*\}, \quad 0 \leq i \leq T.$$

[1] For notational convenience we have chosen for this stylized time set. The reader may reformulate all statements and results in this paper for a general discrete time set $\{T_0, T_1, \ldots, T_J\}$ in a trivial way.

In particular,

$$Y_i^* = EZ_{\tau_i^*}, \quad 0 \leq i \leq T,$$

and the above family is the family of first optimal stopping times if several optimal stopping families exist.

The optimal stopping problem (2.1) has a natural interpretation in the point of view of the option holder: He seeks for an optimal exercise strategy which optimizes his expected pay-off. On the other hand, the seller of the option rather seeks for the minimal cash amount (smallest super-martingale) he has to have at hand in any case the holder of the option exercises.

3 Duality and surely optimal martingales

We briefly recall the dual approach proposed by Rogers (2001) and, independently, Haugh and Kogan (2004). The dual approach is based on the following observation. For any martingale (M_j) with $M_0 = 0$ we have,

$$Y_0^* = \sup_{\tau \in \{0,\dots,T\}} E_0 Z_\tau \leq \sup_{\tau \in \{0,\dots,T\}} E_0 \left(Z_\tau - M_\tau \right) \leq E_0 \max_{0 \leq j \leq T} \left(Z_j - M_j \right) \tag{3.2}$$

hence the right-hand side provides a (dual) upper bound for Y_0^*.

Rogers (2001) and independently Haugh and Kogan (2004) showed, that the equality in (3.2) is attained at the martingale part of the Doob decomposition of Y^*, i.e. $Y_j^* = Y_0^* + M_j^* - A_j^*$ where M^* is a martingale with $M_0^* = 0$, and A^* is predictable with $A_0^* = 0$. I.e.,

$$M_j^* = \sum_{l=1}^{j} \left(Y_l^* - E_{l-1} Y_l^* \right), \quad A_j^* = \sum_{l=1}^{j} \left(Y_{l-1}^* - E_{l-1} Y_l^* \right), \tag{3.3}$$

where A^* is non-decreasing due to the Bellman principle. In addition, they showed that

$$Y_0^* = \max_{0 \leq j \leq T} \left(Z_j - M_j^* \right) \quad \text{a.s.} \tag{3.4}$$

The next lemma by Kolodko and Schoenmakers (2006) provides a somewhat more general class of super-martingales, which turn (3.2) into an equality such that moreover (3.4) holds.

Lemma 1. *Let $\mathcal{S}$ be the set of super-martingales S with $S_0 = 0$. Let $S \in \mathcal{S}$ be such that $Z_j - Y_0^* \leq S_j$, $1 \leq j \leq T$. Then,*

$$Y_0^* = \max_{0 \leq j \leq T} (Z_j - S_j) \quad a.s. \tag{3.5}$$

For the proof see Kolodko and Schoenmakers (2006).

Examples 2. *Obviously, by taking for S the Doob martingale (3.3), Lemma 1 applies. But, this is not the only one. For example, in the case $Z > 0$ a.s. we may also take*

$$S_j = (N_j^* - 1) Y_0^*$$

where N^* is the multiplicative Doob part of the Snel envelope, i.e. $Y_j^* = Y_0^* N_j^* B_j^*$ for a martingale N^* with $N_0^* = 1$ and predictable B^* with $B_0^* = 1$. Hence

$$N_j^* = \prod_{l=1}^{j} \frac{Y_l^*}{E_{l-1} Y_l^*}, \quad B_j^* = \prod_{l=1}^{j} \frac{E_{l-1} Y_l^*}{Y_{l-1}^*}. \tag{3.6}$$

Indeed, since B^* is non-increasing due to the Bellman principle, we have

$$S_j = Y_0^* \left(\frac{Y_j^*}{Y_0^* B_j^*} - 1 \right) \geq Y_0^* \left(\frac{Y_j^*}{Y_0^*} - 1 \right) = Y_j^* - Y_0^* \geq Z_j - Y_0^*$$

and so Lemma 1 applies again.

The multiplicative Doob decomposition in (3.6) is used by Jamshidian (2007) for constructing a multiplicative dual representation. In a comparative study, Chen and Glasserman (2007) pointed out however, that from a numerical point of view additive dual algorithms perform better due to the nice almost sure property (3.4).

Remark 3. *It is **not** true that for any martingale M which turns (3.2) into equality the almost sure statement (3.4) holds. As a simple counterexample consider $T = 1$, $Z_0 = 0$, $Z_1 = 2$, $M_0 = 0$, and $M_1 = \pm 1$ each with probability $1/2$. Indeed, we see that $Y_0^* = 2 = E_0(2 - M_1) = E_0 \max(0, 2 - M_1)$, but, $Y_0^* \neq \max(0, 2 - M_1)$ a.s.*

In order to have a unified dual representation for the Snell envelope Y_i^* at any i, it is convenient to drop the assumption that martingales start at zero. We then may restate the dual theorem as

$$Y_i^* = \inf_{M \in \mathcal{M}} E_i \max_{i \leq j \leq T} (Z_j - M_j + M_i) \tag{3.7}$$

$$= \max_{i \leq j \leq T} \left(Z_j - M_j^* + M_i^* \right) \text{ a.s.,} \tag{3.8}$$

for all i, $0 \leq i \leq T$, where $\mathcal{M}$ is the set of all martingales and M^* is the Doob martingale part of Y^*.

In view of Remark 3 and Examples 2, a martingale for which the infimum (3.7) is attained must not necessarily satisfy the almost sure property (3.8), and, martingales which do satisfy (3.8) are generally not unique.

Definition 4. *We say that a martingale M is **surely optimal** for the Snell envelope Y^* at a time i, $0 \leq i \leq T$, if (3.8) holds.*

Obviously, the Doob martingale of Y^* is surely optimal at each i, $0 \leq i \leq T$. The next theorem provides a characterization of surely optimal martingales.

Theorem 5. *Let Y^* be the Snell envelope of the cash-flow Z and let M be any martingale. Then, for every $i \in \{0, ..., T\}$ the following statement holds:*

For any set $\mathcal{A}_i \in \mathcal{F}_i$ we have

$$1_{\mathcal{A}_i} \max_{i \leq j \leq T} (Z_j - M_j + M_i) \in \mathcal{F}_i \quad \Longrightarrow \quad 1_{\mathcal{A}_i} \max_{i \leq j \leq T} (Z_j - M_j + M_i) = 1_{\mathcal{A}_i} Y_i^*.$$

Proof. We use backward induction on the number i. If $i = T$ the statement reads, for any $\mathcal{A}_T \in \mathcal{F}_T$ we have $1_{\mathcal{A}_T} Z_T \in \mathcal{F}_T \quad \Longrightarrow \quad 1_{\mathcal{A}_T} Z_T = 1_{\mathcal{A}_T} Y_T^*$ which is trivially true. Suppose the statement holds for some $i + 1$ with $0 \leq i < T$ and assume for an arbitrary but fixed set $\mathcal{A}_i \in \mathcal{F}_i$ that

$$1_{\mathcal{A}_i} \vartheta_i := 1_{\mathcal{A}_i} \max_{i \leq j \leq T} (Z_j - M_j + M_i) \in \mathcal{F}_i.$$

We then have

$$\begin{aligned}1_{\mathcal{A}_i}\vartheta_i &= 1_{\mathcal{A}_i}\max(Z_i,\ M_i - M_{i+1} + \max_{i+1\le j\le T}(Z_j - M_j + M_{i+1}))\\ &= 1_{\mathcal{A}_i}\max(Z_i,\ M_i - M_{i+1} + \vartheta_{i+1}) \in \mathcal{F}_i.\end{aligned}$$

We may then consider the following $\mathcal{F}_i$ measurable events:

$$\begin{aligned}&\mathcal{A}_i \cap A^{\mathrm{i)}} \text{ with } A^{\mathrm{i)}} := \{\vartheta_i = Z_i\} \cap \{M_i - M_{i+1} + \vartheta_{i+1} \le Z_i\}, \quad \text{and}\\ &\mathcal{A}_i \cap A^{\mathrm{ii)}} \text{ with } A^{\mathrm{ii)}} = \Omega \backslash A^{\mathrm{i)}}\\ &\qquad := \{\vartheta_i = M_i - M_{i+1} + \vartheta_{i+1}\} \cap \{Z_i < M_i - M_{i+1} + \vartheta_{i+1}\}.\end{aligned}$$

By taking $\mathcal{F}_i$-conditional expectations it follows that

$$1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} Z_i = E_i\, 1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} Z_i \ge E_i\,(M_i - M_{i+1} + \vartheta_{i+1})\, 1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} = 1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} E_i\, \vartheta_{i+1}. \tag{3.9}$$

Since M is a martingale, $E_{i+1}\,\vartheta_{i+1}$ is an upper bound for Y^*_{i+1} by (3.7), and we thus have,

$$E_i\,\vartheta_{i+1} = E_i E_{i+1}\,\vartheta_{i+1} \ge E_i Y^*_{i+1},$$

which yields combined with (3.9),

$$1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} E_i Y^*_{i+1} \le 1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} Z_i$$

and so

$$1_{\mathcal{A}_i\cap A^{\mathrm{i)}}}\vartheta_i = 1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} Z_i = 1_{\mathcal{A}_i\cap A^{\mathrm{i)}}} Y^*_i. \tag{3.10}$$

On the other hand, we notice that $1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}}\vartheta_{i+1} \in \mathcal{F}_{i+1}$ so we have by induction that $1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}}\vartheta_{i+1} = 1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}} Y^*_{i+1}$. It then follows that

$$\begin{aligned}1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}}\vartheta_i &= E_i\, 1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}}\vartheta_i = E_i\, 1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}}\,(M_i - M_{i+1} + \vartheta_{i+1}) = 1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}} E_i \vartheta_{i+1}\\ &= E_i\, 1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}}\vartheta_{i+1} = E_i\, 1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}} Y^*_{i+1} = 1_{\mathcal{A}_i\cap A^{\mathrm{ii)}}} E_i Y^*_{i+1}.\end{aligned} \tag{3.11}$$

Combining (3.10) and (3.11) finally yields $1_{\mathcal{A}_i}\vartheta_i = 1_{\mathcal{A}_i} Y^*_i$. ■

As an immediate consequence of Theorem 5 we obtain:

Corollary 6. *Let Y^*, Z, and martingale M be as in Theorem 5. For $i \in \{0, ..., T\}$ it holds,*

$$\max_{i\le j\le T}(Z_j - M_j + M_i) \in \mathcal{F}_i \quad \Longrightarrow \quad M \textit{ is surely optimal at } i.$$

Remark 7. *Any martingale M is trivially surely optimal at $i = T$. However, it is **not** true that sure optimality for some i with $i < T$ implies sure optimality at $i+1$. As a counterexample let us consider $T = 2$, and $Z_0 = 4$, $Z_1 = 0$, $Z_2 = 2$. Take as martingale $M_0 = 0$, $M_1 = \pm 1$, each with probability $1/2$, and $M_2 = M_1 \pm 1$, each with probability $1/2$ conditional M_1. Then $\max_{0\le j\le 2}(Z_j - M_j + M_0) = 4$ a.s., hence by Theorem 6 we have $Y^*_0 = 4$ and thus M is surely optimal at $i = 0$. But, $\max_{1\le j\le 2}(Z_j - M_j + M_1) = 2 - M_2 + M_1 \notin \mathcal{F}_1$, so M is not surely optimal for Y^* at $i = 1$.*

4 Stability of surely optimal martingales

In equivalent terms, Corollary 6 states that, if a martingale M is such that the conditional variance of

$$\vartheta_i^{(M)} := \max_{i\leq j\leq T}\left(Z_j - M_j + M_i\right), \quad i = 0, ..., T,$$

is zero, i.e.

$$Var_i\left(\vartheta_i^{(M)}\right) := E_i\left(\vartheta_i - E_i\vartheta_i\right)^2 = 0, \text{ a.s., } i = 0, ..., T,$$

then $\vartheta_i^{(M)} = Y_i^*$, for $i = 0, ..., T$. Hence the martingale M is surely optimal for $i = 0, ..., T$. In this section we present a stability result for martingales M which are, loosely speaking, near to be surely optimal in the sense that each i, $\mathrm{Var}_i\left(\vartheta_i^{(M)}\right)$ is small. More specifically, for a sequence of martingales $(M^{(n)})_{n\geq 1}$ we provide mild conditions which guarantee that the corresponding upper bounds converge to the Snell envelope (although the sequence of martingales $(M^{(n)})$ does not need to converge itself). We have the following theorem.

Theorem 8. *Suppose that for each* i, $0 \leq i \leq T$, *Var*$_i(\vartheta_i^{(n)}) \xrightarrow{P} 0$, *if* $n \to \infty$, *where* $\vartheta_i^{(n)} := \max_{i\leq j\leq T}\left(Z_j - M_j^{(n)} + M_i^{(n)}\right)$. *In addition, suppose that for each* i, *the sequence of martingales* $\left(M_i^{(n)}\right)_{n\geq 1}$ *is uniformly integrable for* $n \geq 1$. *It then holds* $\left(\vartheta_i^{(n)}\right)_{n\geq 1}$ *is uniformly integrable for each* i, *and*

$$E_i\,\vartheta_i^{(n)} \xrightarrow{L_1} Y_i^*, \quad i = 0, ..., T.$$

Proof. We will prove the theorem by backward induction on i. For $i = T$ there is nothing to prove. Suppose the theorem is proved for $i+1$, $i < T$. Let us consider

$$\begin{aligned}
\vartheta_i^{(n)} - Z_i &:= \max_{i\leq j\leq T}\left(Z_j - M_j^{(n)} + M_i^{(n)}\right) - Z_i \\
&= \left(M_i^{(n)} - M_{i+1}^{(n)} + \max_{i+1\leq j\leq T}\left(Z_j - Z_i - M_j^{(n)} + M_{i+1}^{(n)}\right)\right)_+ \\
&= \left(\vartheta_{i+1}^{(n)} - Z_i + M_i^{(n)} - M_{i+1}^{(n)}\right)_+, \qquad (4.12)
\end{aligned}$$

and define

$$\psi_i^{(n)} := \vartheta_{i+1}^{(n)} - Z_i + M_i^{(n)} - M_{i+1}^{(n)}. \qquad (4.13)$$

Due to the induction hypothesis $E_{i+1}\vartheta_{i+1}^{(n)} \xrightarrow{L_1} Y_{i+1}^*$ with $\vartheta_{i+1}^{(n)}$ being uniformly integrable. Thus, $\psi_i^{(n)}$ and $\vartheta_i^{(n)}$ are uniformly integrability due to (4.12), (4.13), and the uniform integrability of the martingales. So it holds,

$$E_i\,\psi_i^{(n)} = E_i\,E_{i+1}\vartheta_{i+1}^{(n)} - Z_i \xrightarrow{L_1} E_iY_{i+1}^* - Z_i. \qquad (4.14)$$

We will then show that

$$E_i\left(\psi_i^{(n)}\right)_+ \xrightarrow{L_1} \left(E_iY_{i+1}^* - Z_i\right)_+, \qquad (4.15)$$

which in turn implies by (4.12),

$$E_i\,\vartheta_i^{(n)} \xrightarrow{L_1} Z_i + \left(E_iY_{i+1}^* - Z_i\right)_+ = \max\left(Z_i, E_iY_{i+1}^*\right) = Y_i^*.$$

Since the family $E_i \left(\psi_i^{(n)}\right)_+$ is uniformly integrable too, it is enough to show that

$$E_i \left(\psi_i^{(n)}\right)_+ \xrightarrow{P} \left(E_i Y_{i+1}^* - Z_i\right)_+ . \tag{4.16}$$

From (4.14) it follows

$$\left(E_i \, \psi_i^{(n)}\right)_+ \xrightarrow{P} \left(E_i Y_{i+1}^* - Z_i\right)_+ ,$$

so it is sufficient to prove that

$$E_i \left(\psi_i^{(n)}\right)_+ - \left(E_i \, \psi_i^{(n)}\right)_+ \xrightarrow{P} 0.$$

On the one side, by Jensen's inequality, we have

$$E_i \left(\psi_i^{(n)}\right)_+ \geq \left(E_i \psi_i^{(n)}\right)_+ \quad \text{a.s.}$$

Now take an arbitrary $\epsilon > 0$ and consider the inequality

$$\begin{aligned} &1_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} 1_{\left\{\left(\psi_i^{(n)}\right)_+ \leq \left(E_i \, \psi_i^{(n)}\right)_+\right\}} \\ &\leq 1_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} 1_{\left\{\left(\psi_i^{(n)}\right)_+ \leq E_i\left(\psi_i^{(n)}\right)_+ - \epsilon\right\}}. \end{aligned} \tag{4.17}$$

A conditional version of Chebyschev's inequality implies that

$$\begin{aligned} I_n := 1&_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} E_i \, 1_{\left\{\left(\psi_i^{(n)}\right)_+ \leq E_i\left(\psi_i^{(n)}\right)_+ - \epsilon\right\}} \\ &\leq 1_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} \frac{\mathsf{Var}_i \left(\left(\psi_i^{(n)}\right)_+\right)}{\epsilon^2} \\ &= 1_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} \frac{\mathsf{Var}_i \left(\vartheta_i^{(n)}\right)}{\epsilon^2} \xrightarrow{P} 0, \end{aligned}$$

by the assumptions of the Theorem. Since obviously $0 \leq I_n \leq 1$, this implies that $I_n \xrightarrow{L_1} 0$. Then note that (see (4.17))

$$\begin{aligned} &1_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} 1_{\left\{\left(\psi_i^{(n)}\right)_- > 0\right\}} \\ &\leq 1_{\left\{E_i \, \psi_i^{(n)} > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} 1_{\left\{\left(\psi_i^{(n)}\right)_+ \leq \left(E_i \, \psi_i^{(n)}\right)_+\right\}}. \end{aligned}$$

As a consequence it follows that

$$0 \leq E_i 1_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} 1_{\left\{\left(\psi_i^{(n)}\right)_- > 0\right\}} \leq I_n \xrightarrow{L_1} 0. \tag{4.18}$$

Now since the family $\left(\psi_i^{(n)}\right)_-$ is uniformly integrable also, it is not difficult to see that (4.18) implies

$$1_{\left\{E_i\left(\psi_i^{(n)}\right)_+ > \left(E_i \, \psi_i^{(n)}\right)_+ + \epsilon\right\}} E_i \left(\psi_i^{(n)}\right)_- \xrightarrow{L_1} 0. \tag{4.19}$$

Next we consider

$$
\begin{aligned}
&1_{\left\{E_i\left(\psi_i^{(n)}\right)_+>\left(E_i\,\psi_i^{(n)}\right)_++\epsilon\right\}}E_i\left(\psi_i^{(n)}\right)_- \\
&=1_{\left\{E_i\left(\psi_i^{(n)}\right)_+>\left(E_i\,\psi_i^{(n)}\right)_++\epsilon\right\}}\left[E_i\left(\psi_i^{(n)}\right)_+-E_i\,\psi_i^{(n)}\right] \\
&\geq 1_{\left\{E_i\left(\psi_i^{(n)}\right)_+>\left(E_i\,\psi_i^{(n)}\right)_++\epsilon\right\}}\left[\left(E_i\,\psi_i^{(n)}\right)_+-E_i\,\psi_i^{(n)}+\epsilon\right] \\
&=1_{\left\{E_i\left(\psi_i^{(n)}\right)_+>\left(E_i\,\psi_i^{(n)}\right)_++\epsilon\right\}}\left[\left(E_i\,\psi_i^{(n)}\right)_-+\epsilon\right] \\
&\geq \epsilon 1_{\left\{E_i\left(\psi_i^{(n)}\right)_+>\left(E_i\,\psi_i^{(n)}\right)_++\epsilon\right\}}\geq 0,
\end{aligned}
$$

and so by (4.19),

$$P\left(E_i\left(\psi_i^{(n)}\right)_+>\left(E_i\,\psi_i^{(n)}\right)_++\epsilon\right)\to 0.$$

■

The following simple example illustrates that Theorem 8 would not be true when the uniform integrability condition is dropped.

Example 9. *Take* $T=1$, $Z_0=Z_1=0$, $M_1^{(n)}=:-\xi_n$ *with* $E_0\xi_n=0$, $n=1,2,\dots$ *Then we have*

$$\vartheta_0^{(n)}=\max(Z_0-M_0^{(n)},Z_1-M_1^{(n)})=\max(0,\xi^{(n)})=\xi_+^{(n)}.$$

Now take

$$\xi^{(n)}=\begin{cases}1 \text{ with Prob. } \frac{n-1}{n}\\ 1-n \text{ with Prob. } \frac{1}{n}\end{cases}$$

(hence $E_0\xi^{(n)}=0$*). Then,* $v_0^{(n)}=E_0(\xi_+^{(n)})^2-\left(E_0\xi_+^{(n)}\right)^2=\frac{n-1}{n}-\left(\frac{n-1}{n}\right)^2=\frac{n-1}{n^2}\to 0,$ *whereas* $E_0\vartheta_0^{(n)}=E_0\xi_n^+=\frac{n-1}{n}\to 1.$ *Clearly, for each* $K>1$, $E_0\left|M_1^{(n)}\right|1_{\left\{\left|M_1^{(n)}\right|>K\right\}}\geq\frac{n-1}{n}1_{\{n-1>K\}}\to 1$ *as* $n\to\infty$, *hence the* $\left(M_1^{(n)}\right)$ *are not uniformly integrable.*

Remark 10. *Theorem 8 is important in practical situations, for instance, if there exists some underlying (multi-dimensional) Markovian structure with respect to some (multi-dimensional) Wiener filtration. In this environment we may consider the following class of uniformly integrable martingales.*

Let X *be a* D*-dimensional Markov process adapted to a filtration generated by an* m*-dimensional Brownian motion* W *and let the function* $c(\cdot,\cdot):\mathbb{R}_{\geq 0}\times\mathbb{R}^D\to\mathbb{R}_{\geq 0}$ *be such that* $E_0\int_0^T c^2(s,X_s)ds<\infty$. *Then the class* $\mathcal{M}^{UI}$ *of martingales defined by*

$$M\in\mathcal{M}^{UI}:\Longleftrightarrow M_t=\int_0^t b^{\mathsf{T}}(s,X_s)dW_s,\quad 0\leq t\leq T,\ \text{ for } b \text{ with }\ |b|\leq c$$

is uniformly integrable due to the criterion of de la Vallée Poussin, since

$$\sup_{M\in\mathcal{M}^{UI}}E\left|M_t\right|^2\leq\int_0^T E_0c^2(s,X_s)ds<\infty,\quad 0\leq t\leq T.$$

Remark 11. *Consider any class of uniformly integrable martingales $\mathcal{M}^{UI}$. For example the one considered in Remark 10. As the topology of convergence in probability is metrizable by the Ky Fan Metric (e.g., see Dudley (2002))*

$$d_P(X,Y) := \inf\{\epsilon > 0 :\ P(|X-Y| > \epsilon) \le \epsilon\},$$

one may restate Theorem 8 as follows. For any $\epsilon > 0$ there exist a $\delta > 0$ such that

$$\left[M \in \mathcal{M}^{UI} \ \wedge \ \max_{0\le i\le T} d_P(\mathit{Var}_i(\vartheta_i^{(M)}), 0) < \delta \right] \Longrightarrow \max_{0\le i\le T} \left\| \vartheta_i^{(M)} - Y_i^* \right\|_{L_1} < \epsilon.$$

In view of Remark 11, Theorem 8 may be considered as a stability theorem related to the statement of Corollary 6.

5 Dual backward construction of a nearly sure optimal martingale and pricing of a Bermudan derivative

Let us suppose we are in an environment as outlined in Remark 10 and assume in addition that the cash-flow process is of the form

$$Z_i = Z_i(X_i).$$

We then propose a dual backward algorithm for constructing a martingale $\left(\widehat{M}_i\right)$ such that the corresponding conditional variances $\mathrm{Var}_i\left(\widehat{\vartheta}_i\right)$ (see Theorem 8) are low. This martingale may then be used for computing a price upper bound by linear Monte Carlo. As a by-product, the algorithm also provides a set of continuation functions which may next be used to construct an exercise policy and to simulate a lower bound in the usual way.

On a pseudo algorithmic level, we construct $\left(\widehat{M}_i\right)_{i=0,\dots,T}$ by backward induction as follows. At $i = T$ we simply set $\widehat{\vartheta}_T = \widehat{M}_T = Z(T, X_T)$, hence $\mathrm{Var}_T\left(\widehat{\vartheta}_T\right) = 0$. Suppose we have constructed for $i < T$ the martingale $\left(\widehat{M}_j\right)_{i+1\le j\le T}$ such that $\mathrm{Var}_{i+1}\left(\widehat{\vartheta}_{i+1}\right)$ is low. Then for any extension $\widehat{M}_i$ such that $\left(\widehat{M}_j\right)_{i\le j\le T}$ is a martingale we consider

$$\begin{aligned}
\widehat{\vartheta}_i &:= \max_{i\le j\le T}\left(Z_j(X_j) - \widehat{M}_j + \widehat{M}_i\right) \\
&= \max\left(Z_i(X_i), \widehat{M}_i - \widehat{M}_{i+1} + \max_{i+1\le j\le T}\left(Z_j(X_j) - \widehat{M}_j + \widehat{M}_{i+1}\right)\right) \\
&= \max\left(Z_i(X_i), \widehat{M}_i - \widehat{M}_{i+1} + \widehat{\vartheta}_{i+1}\right)
\end{aligned}$$

following (4.12). In the spirit of Belomestny, et. al. (2009) we choose a set of (possibly time dependent) basis functions

$$(\varphi_k(t,x))_{1\le k\le K} := \left(\varphi_k^{(d)}(t,x)\right)_{d=1,\dots,D,\,1\le k\le K}$$

(where one might take for each d the same set of functions of course), and set

$$\widehat{M}_{i+1} - \widehat{M}_i = \sum_{k=1}^{K} \beta_k \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u)\, dW_u.$$

We then require that

$$\max\left(Z_i(X_i), \widehat{\vartheta}_{i+1} - \sum_{k=1}^{K} \beta_k \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u) dW_u\right) \tag{5.20}$$

is "approximately" $\mathcal{F}_i$-measurable. Moreover, since the Snell envelope Y_i^* is X_i measurable, we will even require that (5.20) is "approximately" X_i measurable. For this we consider in addition an auxiliary set of basis functions $\left(\varphi_k^{(0)}(t,x)\right)_{1\leq k\leq K}$ and, in fact, estimate approximately the continuation pay-off by the regression procedure

$$(\widehat{\gamma}_i, \widehat{\beta}_i) := \underset{\gamma,\beta}{\arg\min} E\left[\widehat{\vartheta}_{i+1} - \sum_{k=1}^{K} \gamma_k \varphi_k^{(0)}(i, X_i) - \sum_{k=1}^{K} \beta_k \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u) dW_u\right]^2.$$

Next we set

$$\widehat{M}_i = \widehat{M}_{i+1} - \sum_{k=1}^{K} \widehat{\beta}_{i,k} \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u) dW_u, \quad \text{and then set}$$

$$\widehat{\vartheta}_i = \max\left(Z_i(X_i), \widehat{\vartheta}_{i+1} - \sum_{k=1}^{K} \widehat{\beta}_k \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u) dW_u\right).$$

In the context of a real Monte Carlo algorithm, the procedure is as follows: Given a Monte Carlo sample of trajectories $(X_r^{(m)}, 0 \leq r \leq T, m = 1, ..., M)$ we estimate

$$\begin{aligned}(\widetilde{\gamma}_i, \widetilde{\beta}_i) := \underset{\gamma,\beta}{\arg\min} \sum_{m=1}^{M} \Bigg[&\widehat{\vartheta}_{i+1}^{(m)} - \sum_{k=1}^{K} \gamma_k \varphi_k^{(0)}(i, X_i^{(m)}) \\ &- \sum_{k=1}^{K} \beta_k \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u^{(m)}) dW_u^{(m)}\Bigg]^2,\end{aligned} \tag{5.21}$$

where the $\widehat{\vartheta}_{i+1}^{(m)}$, $m = 1, ..., M$, are assumed to be already constructed. For a detailed description of Monte Carlo based regression procedures we refer to Glasserman (2003). In (5.21) the Wiener integrals may be approximated as usual by

$$\int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u^{(m)}) dW_u^{(m)} \approx \sum_{l=0}^{\delta^{-1}-1} \varphi_k^{\mathsf{T}}(i + l\delta, X_{i+l\delta}^{(m)}) \left(W_{i+(l+1)\delta}^{(m)} - W_{i+l\delta}^{(m)}\right), \tag{5.22}$$

for a small enough $\delta > 0$. Finally we set

$$\widehat{M}_i = \widehat{M}_{i+1} - \sum_{k=1}^{K} \widetilde{\beta}_{i,k} \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u) dW_u, \text{ and then set}$$

$$\widehat{\vartheta}_i^{(m)} = \max\left(Z_i(X_i^{(m)}), \widehat{\vartheta}_{i+1}^{(m)} - \sum_{k=1}^{K} \widehat{\beta}_k \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u^{(m)}) dW_u^{(m)}\right), \ \ m = 1, ..., M.$$

By working backward from $i = T$ to $i = 0$, the above sequential regression procedure yields a (nearly surely) martingale

$$\widehat{M}_r = \sum_{i=1}^{r-1} \sum_{k=1}^{K} \widetilde{\beta}_{i,k} \int_i^{i+1} \varphi_k^{\mathsf{T}}(u, X_u) dW_u, \tag{5.23}$$

and, as a by-product, an additional system of approximations to the continuation value functions,

$$\mathcal{C}_i(x) \approx \sum_{k=1}^{K} \widetilde{\gamma}_{i,k} \varphi_k^{(0)}(i,x). \tag{5.24}$$

The martingale (5.23) may be used to compute a tight dual upper bound at $i = 0$, by starting a new simulation of trajectories $\widetilde{X}_i^{(m)}$, $\widetilde{m} = 1, ..., \widetilde{M}$, and computing

$$Y_0^{up} \approx \frac{1}{\widetilde{M}} \sum_{\widetilde{m}=1}^{\widetilde{M}} \max_{0 \le j \le T} \left(Z_j(\widetilde{X}_j^{(\widetilde{m})}) - \sum_{r=1}^{j-1} \sum_{k=1}^{K} \widetilde{\beta}_{r,k} \int_r^{r+1} \varphi_k^{\mathsf{T}}(u, \widetilde{X}_u^{(\widetilde{m})}) d\widetilde{W}_u^{(\widetilde{m})} \right)$$

(see (5.22) for approximation of the Wiener integrals).

On the other side, based on the approximate continuation functions (5.24), we may define an exercise strategy

$$\tau_0 := \inf\{i : 0 \le i \le T, \;\; Z_i(X_i) \ge \mathcal{C}_i(X_i)\},$$

and simulate a lower biased price estimate,

$$Y_0^{low} \approx \frac{1}{\widetilde{M}} \sum_{\widetilde{m}=1}^{\widetilde{M}} Z_{\tau_0^{(\widetilde{m})}}(\widetilde{X}_{\tau_0^{(\widetilde{m})}}^{(\widetilde{m})}), \quad \text{where}$$

$$\tau_0^{(\widetilde{m})} = \inf\{i : 0 \le i \le T, \;\; Z_i(\widetilde{X}_i^{(\widetilde{m})}) \ge \mathcal{C}_i(\widetilde{X}_i^{(\widetilde{m})})\}, \;\; \widetilde{m} = 1, ..., \widetilde{M}.$$

Remark 12. *We have no doubt that the convergence of the regression algorithm presented in this section can be proved in the spirit of Belomestny, et. al. (2010) or Clement et al. (2002).*

6 Numerical examples

In this section we test our algorithm at two benchmark examples: Bermudan basket-put on 5 assets and Bermudan max-call on 2 assets, which are also considered in Bender et al. (2006a) and Belomestny, et. al. (2009), respectively. In both examples, the risk-neutral dynamic of each asset is governed by

$$dX_t^d = (r - \delta) X_t^d dt + \sigma X_t^d dW_t^d, \quad d = 1, ..., D,$$

where D is the number of assets, W_t^d, $d = 1, ..., D$, are independent one-dimensional Brownian motions and r, δ and σ are constants. Exercise opportunities are equally spaced at times $T_j = \frac{jT}{J}$, $j = 0, ..., J$. The discounted payoff from exercise at time t is given by

$$Z_t(X_t) = e^{-rt}(K - \frac{X_t^1 + \ldots + X_t^D}{D})^+ \quad \text{for the Bermudan basket-put}$$

and

$$Z_t(X_t) = e^{-rt}(\max(X_t^1, \ldots, X_t^D) - K)^+ \quad \text{for the Bermudan max-call,}$$

where $X_t = (X_t^1, \ldots, X_t^D)$.

For each numerical implementation, the time interval $[T_j, T_{j+1}]$, $j = 0, \ldots, J-1$, is supposed to be partitioned into L equal subintervals of width $\Delta t = \frac{T}{N}$ with $N = J \times L$. The numerical procedure can be described briefly as follows. We first simulate M independent samples of Brownian increments

$\{(\Delta W_i^{1,(m)}, \ldots, \Delta W_i^{D,(m)}),\ i = 1, \ldots, N\}$, $m = 1, \ldots, M$. Then the trajectories of $X_i^{(m)} = (X_i^{1,(m)}, \ldots, X_i^{D,(m)})$, $i = 1, \ldots, N$, $m = 1, \ldots, M$ are given by

$$X_i^{d,(m)} = X_{i-1}^{d,(m)} \exp\{(r - \delta - \frac{1}{2}\sigma^2)\Delta t + \sigma \Delta W_i^{d,(m)}\}$$

for $d = 1, \ldots, D$ and initial data X_0. We run the regression procedure as explained in Section 5. We notice that the Wiener integrals should be approximated by using the same Brownian increments as above. Once we obtain the estimation of coefficients $(\widetilde{\gamma}_i, \widetilde{\beta}_i)$, we simulate a new set of $\widetilde{M}$ independent Brownian increments and trajectories $X_i^{(m)}$ and compute the upper and lower bounds Y_0^{up}, Y_0^{low} as shown in Section 5.

As one can expect, the choice of basis functions is crucial to obtain tight upper and lower bounds. In this respect, special information on the pricing problem may help us finding suitable basis functions. Suppose $E_t\left(Z_T(X_T)\right) = f(t, X_t)$ for $0 \le t \le T = T_J$. Then, by Itô's formula and the fact that $E_t(Z_T)$ is a martingale we have

$$Z_T(X_T) - E_{T_{J-1}}\left(Z_T(X_T)\right) = \sum_{d=1}^{D} \sigma \int_{T_{J-1}}^{T} f_{x^d}(t, X_t) X_t^d dW_t^d.$$

Recall that $\widehat{\vartheta}_T = Z_T$ and $E_{T_{J-1}}\left(Z_T(X_T)\right)$ can be expressed in the following form

$$E_{T_{J-1}}\left(Z_T(X_T)\right) = e^{-rT_{J-1}} EP(T_{J-1}, X_{T_{J-1}}; T),$$

where $EP(t, x; T)$ is the price of the corresponding European option with maturity T at time t. Thus, it is natural to choose from time T to time T_{J-1} European option values for the basis $\left(\varphi_k^{(0)}(t, x)\right)_{1 \le k \le K}$, and the corresponding European deltas multiplied by the value of the underlying asset for the basis $(\varphi_k(t, x))_{1 \le k \le K}$. Although for the following steps ($t < T_{J-1}$) there is no easy way to predict optimal choices of $(\varphi_k^{(0)})$ and (φ_k), the above analysis suggests us to always include the still-alive European options into the basis $(\varphi_k^{(0)})$ and include the information on the European deltas into the basis (φ_k). In fact, these choices of basis functions were already proposed in Belomestny, et. al. (2009).

6.1 Bermudan basket-put

In this example, we take the following parameter values,

$$r = 0.05, \quad \delta = 0, \quad \sigma = 0.2, \quad D = 5, \quad T = 3,$$

and

$$X_0^1 = \ldots = X_0^D = x_0, \quad K = 100.$$

For $T_j \le t < T_{j+1}$, $j = 0, \ldots, J-1$, we choose $\{1, Pol_3(X_t), Pol_3(EP(t, X_t; T_{j+1})), Pol_3(EP(t, X_t; T_J))\}$ as the basis functions $\varphi_k^{(0)}$, where $Pol_3(y)$ denotes the set of polynomials of degree up to n in the components of a vector y and $EP(t, X; T)$ denotes the (approximated) value of a European basket-put with maturity T at time t. We use $\{1, X_t^d \frac{\partial EP(t, X_t; T_{j+1})}{\partial X_t^d}, X_t^d \frac{\partial EP(t, X_t; T_J)}{\partial X_t^d}, d = 1, \ldots, D\}$ as basis functions φ_k.

Since there is no closed-form formula for the still-alive European basket-put, we use the moment-matching method to approximate their values (e.g., see Brigo et al. (2004), and Lord (2005)). Let $S_t = \dfrac{X_t^1 + \ldots + X_t^D}{D}$, and consider another asset G_t whose risk-neutral dynamic follows

$$dG_t = rG_t dt + \tilde{\sigma} G_t,$$

where $\tilde{\sigma}$ is a constant. The value of the European put on this asset can be easily computed by the Black-Scholes formula, that is,

$$E[e^{-rT}(K - G_T)^+] = BS(G_0, r, \tilde{\sigma}, K, T). \tag{6.25}$$

If S_T and G_T have the same moments up to two, then the Black-Scholes price in (6.25) can be regarded as a good approximation for the value of the European basket-put $E\left(e^{-rT}(K - S_T)^+\right)$. Since

$$E(S_T) = \frac{1}{D}\sum_{d=1}^{D} X_0^d e^{rT},$$

$$E(S_T^2) = \frac{1}{D^2} e^{2rT}\left(\sum_{i,j=1}^{D} X_0^i X_0^j \exp(1_{i=j}\sigma^2 T)\right)$$

and

$$E(G_T) = G_0 e^{rT}, \qquad E(G_T^2) = G_0^2 e^{2rT + \tilde{\sigma}^2 T},$$

we can simply set

$$G_0 = \frac{1}{D}\sum_{d=1}^{D} X_0^d$$

and

$$\tilde{\sigma}^2 = \frac{1}{T}\ln\left(\frac{1}{(\sum_{d=1}^{D} X_0^d)^2}\sum_{i,j=1}^{D} X_0^i X_0^j \exp(1_{i=j}\sigma^2 T)\right).$$

The European deltas can be approximated by

$$\frac{\partial BS}{\partial G_0}\frac{\partial G_0}{\partial X_0^d} = -\mathcal{N}(-d_1)\frac{1}{D}, \quad d = 1, \ldots, D,$$

where $d_1 = \dfrac{\ln(\frac{G_0}{K}) + (r + \frac{\sigma^2}{2})T}{\tilde{\sigma}\sqrt{T}}$ and $\mathcal{N}$ denotes the cumulative standard normal distribution function. The numerical results are shown in Table 1. 100000 simulations are used for the regression procedure and 100000 simulations for computing the upper and lower bounds. Values in parentheses are the standard errors. The last column in the table shows the lower and upper bounds obtained in Bender et al. (2006a).

6.2 Bermudan max-call

We use the same parameter values as in Section 6.1 except $\delta = 0.1$ and $D = 2$. As in the previous example we use European (call) options as basis functions in (φ^0) and the corresponding deltas in

Table 1: Lower and upper bounds for Bermudan basket-put on 5 assets with parameters $r = 0.05$, $\delta = 0$, $\sigma = 0.2$, $K = 100$, $T = 3$ and different J and x_0

J	x_0	Lower Bound (SD)	Upper Bound (SD)	BKS Price Interval
3	90	10.0000 (0.0000)	10.0000 (0.0000)	[10.000,10.004]
	100	2.1649 (0.0119)	2.1817 (0.0015)	[2.154,2.164]
	110	0.5357 (0.0060)	0.5584 (0.0008)	[0.535,0.540]
6	90	10.0000 (0.0000)	10.0007 (0.0001)	[10.000,10.000]
	100	2.3986 (0.0112)	2.4336 (0.0013)	[2.359,2.412]
	110	0.5870 (0.0061)	0.5994 (0.0007)	[0.569,0.580]
9	90	10.0000 (0.0000)	10.0024 (0.0001)	[10.000,10.005]
	100	2.4862 (0.0109)	2.5197 (0.0012)	[2.385,2.502]
	110	0.6006 (0.0060)	0.6164 (0.0006)	[0.577,0.600]

the basis (φ_k). The value of European max-call options is computed by the following formula (Johnson (1987)).

$$\sum_{l=1}^{D} X_0^l \frac{e^{-\delta T}}{\sqrt{2\pi}} \int_{(-\infty, d_+^l]} \exp[-\frac{1}{2}z^2] \prod_{\substack{l'=1 \\ l' \neq l}}^{D} \mathcal{N}\left(\frac{\ln \frac{X_0^l}{X_0^{l'}}}{\sigma\sqrt{T}} - z + \sigma\sqrt{T} \right) dz$$
$$- Ke^{-rT} + Ke^{-rT} \prod_{l=1}^{D} \left(1 - \mathcal{N}\left(d_-^l\right)\right),$$

where

$$d_-^l := \frac{\ln \frac{X_0^l}{K} + (r - \delta - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, \quad d_+^l = d_-^l + \sigma\sqrt{T}.$$

We use the central difference quote $\dfrac{f(x+\frac{h}{2}) - f(x-\frac{h}{2})}{h}$ (or the forward difference quote $\dfrac{f(x+h) - f(x)}{h}$) to approximate the European deltas. Note that the rounding errors resulting from too small values of h may give a totally different basis function. So special attentions need to be paid for choosing a suitable h. The numerical results are shown in Table 2. They are based on 1000 simulations for the regression procedure, 1000 simulations for computing the upper bound, 100000 simulations for computing the lower bound. The price intervals in the last column are quoted from Andersen and Broadie (2004).

Table 2: Lower and upper bounds for Bermudan max-call on 2 assets with parameters $r = 0.05$, $\delta = 0.1$, $\sigma = 0.2$, $K = 100$, $T = 3$ and different x_0

x_0	Lower Bound (SD)	Upper Bound (SD)	A&B Price Interval
90	8.1027 (0.0380)	8.1369 (0.0379)	[8.053, 8.082]
100	13.8049 (0.0475)	14.0501 (0.0467)	[13.892, 13.934]
110	21.3208 (0.0556)	21.4860 (0.0531)	[21.316, 21.359]

6.3 Conclusion

The numerical results presented in Tables 1, 2 due to our new algorithm may be considered satisfactory, given the required low computation times, which are in the order of minutes (in a C++ compiled implementation). In this respect it should be noted that the upper bounds in Bender et al. (2006a) (Table 1) and Andersen and Broadie (2004) (Table 2) are computed with nested Monte Carlo simulation requiring computation times in the order of hours. Moreover, the algorithm delivers very fast and surprisingly good lower bounds while the upper bounds are comparable with the ones obtained with the algorithm in Belomestny, et. al. (2009). Needless to say that, as for the method of Belomestny, et. al. (2009), the performance of the here presented algorithm will highly depend on the choice of the basis functions. An in depth treatment of this issue is considered beyond scope however.

Acknowledgements

J.S. is grateful to Denis Belomestny for interesting discussions on this topic, and to Christian Bender for an inspiring remark, which led to counterexample Remark 9.

References

L. Andersen, M. Broadie (2004). A primal-dual simulation algorithm for pricing multidimensional American options. Management Sciences, **50**, No. 9, 1222-1234.

V. Bally, G. Pages (2003). A quantization algorithm for solving multidimensional discrete optimal stopping problem. Bernoulli, **9**(6), 1003-1049.

D. Belomestny, A. Kolodko, and J. Schoenmakers (2010). Regression Methods for Stochastic Control Problems and their Convergence Analysis. SIAM J. Control Optim., Vol. 48, No. 5, pp. 3562–3588.

D. Belomestny, C. Bender, and J. Schoenmakers (2009). True upper bounds for Bermudam products via non-nested Monte Carlo. Mathematical Finance, **19**, No. 1, 53–71.

C. Bender, A. Kolodko, J. Schoenmakers (2008). Enhanced policy iteration for American options via scenario selection. Quant. Finance **8**, No. 2, 135–146.

C. Bender, A. Kolodko, J. Schoenmakers (2006). Iterating cancellable snowballs and related exotics. RISK, September 2006 pp. 126–130.

C. Bender, A. Kolodko, J. Schoenmakers (2006a). Policy iteration for american options: overview. Monte Carlo Methods and Appl., **12**, No. 5-6, pp. 347–362.

D. Brigo, F. Mercurio, F. Rapisarda, and R. Scotti (2004). Approximated moment-matching dynamics for basket-options pricing. Quantitative Finance, 4, 1–16.

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

J. Carriere (1996). Valuation of early-exercise price of options using simulations and nonparametric regression. Insurance: Mathematics and Economics, **19**, 19-30.

N. Chen and P. Glasserman (2007). Additive and Multiplicative Duals for American Option Pricing. Finance and Stochastics, **11**, 153—179.

E. Clément, D. Lamberton, P. Protter (2002). An analysis of a least squares regression algorithm for American option pricing. Finance and Stochastics, **6**, 449-471.

R.M. Dudley (2002). Real analysis and probability. Cambridge University Press.

P. Glasserman (2003). Monte Carlo Methods in Financial Engineering. Springer.

M. Haugh, L. Kogan (2004). Pricing American options: a duality approach. Operations Research, **52**, No. 2, 258–270.

F. Jamshidian (2007). The duality of optimal exercise and domineering claims: A Doob-Meyer decomposition approach to the Snell envelope. Stochastics, **79**, No. 1–2, 27–60.

H. Johnson (1987). Options on the maximum or the minimum of several assets. Journal of Financial and Quantitative Analysis 22, 227–83.

A. Kolodko, J. Schoenmakers (2004). Upper bounds for Bermudan style derivatives. Monte Carlo Methods and Appl., **10**, No. 3–4, 331-343.

A. Kolodko, J. Schoenmakers (2006). Iterative construction of the optimal Bermudan stopping time. Finance and Stochastics, **10**, 27–49.

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

R. Lord (2005). A motivation for conditional moment matching. In M. Vanmaele, A. De Schepper, J. Dhaene, H. Reynaerts, W. Schoutens, and P. Van Goethem, editors, Proceedings of the 3rd Actuarial and Financial Mathematics Day, Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, 69–94.

J. Neveu (1975). Discrete Parameter Martingales. North-Holland, Amsterdam.

L.C.G. Rogers (2001). Monte Carlo valuation of American options. Mathematical Finance, **12**, 271–286.

L.C.G. Rogers (2010). Dual Valuation and Hedging of Bermudan Options. SIAM J. Financial Math., **1**, 604–608.

J. Schoenmakers (2005). Robust Libor Modelling and Pricing of Derivative Products. Chapman & Hall/CRC.

J. Tsitsiklis, B. Van Roy (1999). Regression methods for pricing complex American style options. IEEE Trans. Neural. Net., **12**, 694–703.