Iterative Construction of the Optimal Bermudan Stopping Time[1]

Anastasia Kolodko[*,**], **John Schoenmakers**[*]

[*]Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin. (e-mail: {kolodko,schoenma}@wias-berlin.de).
[**]Institute of Computational Mathematics and Mathematical Geophysics, Prosp. Lavrentjeva 6, 630090 Novosibirsk, Russia.

Abstract. We present an iterative procedure for computing the optimal Bermudan stopping time, hence the Bermudan Snell envelope. The method produces an increasing sequence of approximations of the Snell envelope from below, which coincide with the Snell envelope after finitely many steps. Then, by duality, the method induces a convergent sequence of upper bounds as well. In a Markovian setting the presented iterative procedure allows to calculate approximative solutions with only a few nestings of conditionals expectations and is therefore tailor-made for a plain Monte-Carlo implementation. The method presented may be considered generic for all discrete optimal stopping problems. The power of the procedure is demonstrated at Bermudan swaptions in a full factor LIBOR market model.

Key words: Bermudan options, optimal stopping, Monte Carlo simulation, LIBOR market model

JEL Classification: G13

Mathematics Subject Classification (2000): 62L15, 65C05, 91B28

1 Introduction

Pricing of American style derivatives on a high dimensional system of underlyings is considered a perennial problem for the last decades. As a matter of fact, such high dimensional options are difficult, if not impossible, to compute by standard PDE methods. For high dimensional European options an almost canonical alternative to PDEs is Monte Carlo simulation. However, for American options,

[1]Supported by the DFG Research Center "Mathematics for key technologies" (FZT 86) in Berlin.

Monte Carlo simulation is less trivial since the (optimal) exercise boundary is usually unknown.

In the past literature, various Monte Carlo algorithms for pricing American options are developed. Many of these approaches are related to so called backward dynamic programming which comes down to a recursive representation of the Snell envelope. Among these methods we mention the random tree method and the stochastic mesh method of Broadie & Glasserman (2000, 2004), the cross-sectional least squares algorithm by Longstaff & Schwarz (2001), and the quantization algorithm by Bally & Pages (2003).

As an alternative to backward dynamic programming, one may search for a suitable parametric family of exercise boundaries and then maximize the solutions of the corresponding family of boundary value problems over the parameters. In Andersen (1999) this concept has been applied in the context of Bermudan swaptions in a LIBOR market model. Bermudan options are in fact American options with a discrete set of exercise dates.

In 2001 Rogers proposed a method for American options which is based on a dual representation of the Snell envelope. Independently, Haugh & Kogan (2001) proposed a similar dual procedure for Bermudan derivatives. The dual approach was a new impulse for the research on Monte Carlo methods for American options in fact, since it enables Monte Carlo simulation of tight upper bounds due to given tight lower bounds. More recently, another (multiplicative) dual procedure based on change of numeraires is developed by Jamshidian (2003, 2004).

Further recent papers on (Monte Carlo) methods for high-dimensional American and Bermudan options include Joshi & Theis (2002), Milstein, Reiß & Schoenmakers (2003), Kolodko & Schoenmakers (2003, 2004a), Belomestny & Milstein (2004), Berridge & Schumacher (2004), and for a more detailed and general overview we refer to Glasserman (2003) and the references therein.

The central result in this paper is an iterative construction of the Bermudan Snell envelope via a sequence of stopping times which increases to the (first) optimal stopping time. In each iteration step we improve a whole family of stopping times (τ_i), where i runs through the set of exercise dates, and τ_i is the stopping time for the Bermudan which is not exercised before date i. In fact, the proposed improvement is inspired by a canonical exercise policy for Bermudan swaptions which is already not far from optimal usually; *exercise as soon as the cash-flow dominates all the Europeans ahead.* The thus obtained sequence of stopping families naturally induces an increasing sequence of lower approximations of the Snell envelope. This sequence even coincides with the Snell envelope after finitely many steps. However, the main issue is that after each iteration step we obtain an improved approximation of the Snell envelope which ranges over *all* exercise dates. This in contrast to the backward dynamic program which requires, for obtaining a value for the Snell envelope at the initial date, a number of steps equal to the total number of exercise dates. As a consequence (explained in detail in Remark 2.1), for a larger number of exercise dates, for example ten or twenty, a plain Monte Carlo implementation of the latter method, where conditional expectations are computed by Monte Carlo, would require astronomical computation times. Further, by the dual approach

we deduce from the constructed sequence of lower bounds a sequence of upper bounds which converges to the Snell envelope as well.

The whole analysis is based on a finite set of exercise dates, but since a continuous time American option may be approximated by a Bermudan option with a fine grid of exercise dates, one may in principle apply the method to American options also. The same remark applies to an infinite (perpetual) discrete optimal stopping problem which may be approximated by a discrete problem for a finite set of exercise dates (Shiryayev (1978)). We further underline that the presented method is quite generic and can in principle be applied to any discrete (not necessarily financial) optimal stopping problem.

The presented iterative method may be considered as a pseudo-algorithm since for an implementation conditional expectations need to be computed along trajectories. In a Markovian setting these conditional expectations may be computed (approximately) by Monte Carlo and the thus resulting algorithm is spelled out. In a subsequent paper, Bender & Schoenmakers (2004), the stability of this algorithm is shown and, moreover, the present iteration method is extended for multiple callable structures. We implement the algorithm for Bermudan swaptions in a multi-factor LIBOR market model. It turns out that with a one-degree nested Monte Carlo simulation (two iterations) very accurate Bermudan prices (within 1%) may be obtained in reasonable times.

It should be noted that the in this paper proposed method is a kind of "policy iteration" which is rather different from policy iteration as in Howard (1960) and Puterman (1994). In particular, due to its special nature, our method gives usually good results with only a very few iterations. Moreover, the procedure does not rely on explicit knowledge of underlying transition kernels and is essentially dimension independent. These are important features since in various applications, for example, in the case of an underlying LIBOR process the dimension is usually high and transition kernels are not explicitly available.

The paper is organized as follows. In Section 2 we formulate the discrete optimal stopping problem and consider its implementation for Bermudan style derivatives. In Section 3 we prove the following corner stone result: Based on a given family of stopping times which satisfies some natural conditions, one can construct a next stopping family (by a procedure inspired by a canonical exercise policy for Bermudan swaptions), which is closer to the optimal family as the one we have started with. By using this result, we construct in Section 4 a sequence of stopping families and lower bounds for the Snell envelope and show that by this sequence the Snell envelope is attained after a finite number of steps. Then, in Section 5, we recall the dual upper bound representation for the Snell envelope by Rogers, Haugh & Kogan and include a small extension. Based on the dual approach, the convergent lower bounds constructed in Section 4, and an approximation theorem we deduce a sequence of upper bounds which converges to the Snell envelope from above (also in finitely many steps). In Section 6 we describe a Monte Carlo implementation of the iterative procedure in Section 4, for the usual situation where the underlying process is Markovian. Finally, in Section 7, we apply our method to Bermudan swaptions in a LIBOR market model. We also give a numerical comparison with Andersen's (1999)

lower bound method and its dual considered by Andersen and Broadie (2001).

2 Discrete optimal stopping and its application to Bermudan style derivatives

Let us consider a discrete non-negative random reward process $(Z_i)_{0\le i\le k}$ with state space $\mathbb{R}$, adapted to some discrete filtration $\mathbb{F} = (\mathcal{F}_i)_{0\le i\le k}$. For a fixed i, $0 \le i \le k$, an $\mathbb{F}$-stopping time τ_i^* is called an *optimal stopping time* in the discrete set of *exercise dates* $\{i, \dots, k\}$, if

$$Y_i^* := E^i Z_{\tau_i^*} = \sup_{\tau\in\{i,\dots,k\}} E^i Z_\tau. \tag{1}$$

In (1), E^i denotes conditional expectation with respect to $\mathcal{F}_i$ and the supremum is taken over all $\mathbb{F}$-stopping times τ with values in the set $\{i, ..., k\}$. The sequence of optimal stopping times $(\tau_i^*)_{0\le i\le k}$ is called an *optimal stopping family.* For technical reasons we assume that Z_i has finite expectation for each i, $0 \le i \le k$. The process Y^*, called the *Snell envelope process*, is a supermartingale. This can be seen as follows. Let $i < j$, then it holds

$$E^i Y_j^* = E^i E^j Z_{\tau_j^*} = E^i Z_{\tau_j^*} \le \sup_{\tau\in\{i,\dots,k\}} E^i Z_\tau = Y_i^*.$$

The Snell envelope Y^* can be constructed by the following algorithm, called *backward dynamic programming* (see e.g. Shiryayev (1978), Elliot & Kopp (1999)). At the last exercise date we have trivially $Y_k^* = Z_k$, and for $0 \le j < k$,

$$Y_j^* = \max\left(Z_j, E^j Y_{j+1}^*\right). \tag{2}$$

An optimal stopping family is represented by

$$\tau_i^* = \inf\left\{j, i \le j \le k : Y_j^* \le Z_j\right\}. \tag{3}$$

In fact, (3) denotes the *first optimal* stop after time index i.

As a canonical application of discrete optimal stopping in finance we consider the Bermudan pricing problem. Let $L(t)$ be an $\mathbb{R}^D$-valued random process on a finite time interval $[0, T_\infty]$, $T_\infty < \infty$, adapted to a filtration $(\mathcal{F}(t))_{0\le t\le T_\infty}$ which satisfies the usual conditions. For example, L can be a system of asset prices, but also a not explicitly tradable object such as the term structure of interest rates, or a system of LIBOR rates. Consider a set of dates $\mathbb{T} := \{\mathcal{T}_0, \dots, \mathcal{T}_n\}$ with $0 \le \mathcal{T}_0 < \cdots < \mathcal{T}_n \le T_\infty$. An option issued at time $t = 0$ to exercise a cash-flow $C_{\mathcal{T}_\tau} := C(\mathcal{T}_\tau, L(\mathcal{T}_\tau))$ at a date $\mathcal{T}_\tau \in \mathbb{T}$, to be decided by the option holder, is called a *Bermudan style derivative*. With respect to a pricing measure P connected with some pricing numeraire B, the value of the Bermudan derivative at a future time point t, when the option is not exercised before t, is given by

$$V(t) := B(t) \sup_{\tau\in\{\kappa(t),\dots,n\}} E^{\mathcal{F}(t)} \frac{C_{\mathcal{T}_\tau}}{B(\mathcal{T}_\tau)} \tag{4}$$

with $\kappa(t) := \min\{m : \mathcal{T}_m \geq t\}$. Note that $V(t)$ can also be seen as the price of a Bermudan option newly issued at time t, with exercise opportunities $\mathcal{T}_{\kappa(t)}, \ldots, \mathcal{T}_n$. The fact that (4) can be considered as the fair price for the Bermudan derivative is due to general no-arbitrage principles, e.g. see Duffie (2001). The discounted Bermudan price can be expressed as the solution of a discrete optimal stopping problem (1) as follows. If $t = \mathcal{T}_i$ for some $0 \leq i \leq n$, define $k := n - i$, and

$$\begin{aligned} Z_0 := \frac{C_{\mathcal{T}_i}}{B(\mathcal{T}_i)}, \quad \ldots, \quad Z_k := \frac{C_{\mathcal{T}_n}}{B(\mathcal{T}_n)}, \\ \mathcal{F}_0 := \mathcal{F}(\mathcal{T}_i), \quad \ldots, \quad \mathcal{F}_k := \mathcal{F}(\mathcal{T}_n). \end{aligned}$$

If $\mathcal{T}_i < t < \mathcal{T}_{i+1}$ for some $0 \leq i < n$, define $k := n - i$, and

$$\begin{aligned} Z_0 := 0,\ Z_1 := \frac{C_{\mathcal{T}_{i+1}}}{B(\mathcal{T}_{i+1})}, \quad \ldots, \quad Z_k := \frac{C_{\mathcal{T}_n}}{B(\mathcal{T}_n)}, \\ \mathcal{F}_0 := \mathcal{F}(t),\ \mathcal{F}_1 := \mathcal{F}(\mathcal{T}_{i+1}), \quad \ldots, \quad \mathcal{F}_k := \mathcal{F}(\mathcal{T}_n). \end{aligned}$$

Remark 2.1 For a Markovian sequence $(L(\mathcal{T}_i), B(\mathcal{T}_i))_{0 \leq i \leq k}$ (e.g. a LIBOR model with B being the spot LIBOR rolling-over account) it would be possible, in principle, to construct the Snell envelope $Y^* = V/B$ by plain Monte Carlo simulation of the backward dynamic program (2) in the following way. Step 1: for each state $(L(\mathcal{T}_{k-1}), B(\mathcal{T}_{k-1}))$, Y^*_{k-1} can be estimated via (2) by simulation of Y^*_k $(= Z_k)$ under the conditional measure $P^{(L(\mathcal{T}_{k-1}), B(\mathcal{T}_{k-1}))}$. Next, Step 2: for each state $(L(\mathcal{T}_{k-2}), B(\mathcal{T}_{k-2}))$, Y^*_{k-2} can be estimated via (2) by simulation of Y^*_{k-1} under the conditional measure $P^{(L(\mathcal{T}_{k-2}), B(\mathcal{T}_{k-2}))}$. However, note that each simulation in Step 2 requires an estimation of Y^*_{k-1}, hence a simulation according to Step 1. As a result, Y^*_{k-2} requires a (one-fold) nested Monte Carlo simulation. Thus proceeding we would obtain Y^*_0 via a $(k-1)$-fold nested Monte Carlo simulation. But, the complexity of this procedure would be tremendous. For example, when using N simulations in each backward step we would need N^k samples for the value of Y^*_0. As an example $N = 10\ 000$ and $k = 10$ would lead to more than 10^{40} samples!

3 A one step improvement upon a given family of stopping times

We now study the optimal stopping problem (1). With respect to the discrete filtration $(\mathcal{F}_i)_{0 \leq i \leq k}$ we consider a family of integer valued stopping indexes (τ_i), with the following properties,

$$\begin{aligned} i \leq \tau_i \leq k,\ \tau_k = k, \\ \tau_i > i \Rightarrow \tau_i = \tau_{i+1}, \qquad 0 \leq i < k, \end{aligned} \tag{5}$$

and the process

$$Y_i := E^{\mathcal{F}_i} Z_{\tau_i}. \tag{6}$$

For the Bermudan problem (4) one could take for example

$$\tau_i := \inf\{j \geq i : (\mathcal{T}_j, L(\mathcal{T}_j)) \in R\} \wedge k,$$

where R is a certain region in $\mathbb{R}^{D+1}$, or, as a more trivial example, the family $\tau_i \equiv i$. Generally, the process (Y_i) is a lower approximation of the Snell envelope process (Y_i^*) due to the family of (sub-optimal) stopping times (τ_i). Based on the family (τ_i) we are going to construct a new family $(\widehat{\tau}_i)$ satisfying (5), which induces a new approximation of the Snell envelope.

We fix a window parameter κ; an integer with $1 \leq \kappa \leq k$, and then introduce an intermediate process

$$\widetilde{Y}_i := \max_{p:\, i \leq p \leq \min(i+\kappa,k)} E^i Z_{\tau_p}. \tag{7}$$

Using $\widetilde{Y}_i$ as an exercise criterion we define a new family of stopping indexes

$$\begin{aligned} \widehat{\tau}_i : &= \inf\{j : i \leq j \leq k,\ \widetilde{Y}_j \leq Z_j\} \\ &= \inf\{j : i \leq j \leq k,\ \max_{p:\, j \leq p \leq \min(j+\kappa,k)} E^j Z_{\tau_p} \leq Z_j\}, \quad 0 \leq i \leq k, \end{aligned} \tag{8}$$

and consider the process

$$\widehat{Y}_i := E^i Z_{\widehat{\tau}_i} \tag{9}$$

as a next approximation of the Snell envelope. Clearly, the family $(\widehat{\tau}_i)$ satisfies the properties (5) as well. Let us take $\kappa = k$. Then for Bermudan derivatives, as an example, the trivial family $\tau_i \equiv i$ gives for $\widetilde{Y}$ the maximum of still alive Europeans and for $\widehat{Y}$ the second "canonical example" in Kolodko & Schoenmakers (2003, 2004a). As another example, $\tau_i \equiv k$ gives for $\widetilde{Y}$ the European option process due to the last exercise date k and

$$\widehat{\tau}_i := \inf\{j : i \leq j \leq k,\ E^j Z_k \leq Z_j\}, \quad 0 \leq i \leq k.$$

By the next theorem, $(\widehat{Y}_i)$ is generally an improvement of (Y_i).

Theorem 3.1 *Let (τ_i) be a family of stopping times with the property* (5) *and (Y_i) be given by* (6)*. Let the processes $(\widetilde{Y}_i)$ and $(\widehat{Y}_i)$ be defined by* (7) and (9)*, respectively. Then, it holds*

$$Y_i \leq \widetilde{Y}_i \leq \widehat{Y}_i \leq Y_i^*, \quad 0 \leq i \leq k.$$

Proof.
The inequalities $Y_i \leq \widetilde{Y}_i$ and $\widehat{Y}_i \leq Y_i^*$ are trivial. We only need to show the middle inequality. We use induction in i. Due to the definition of $\widetilde{Y}$ and $\widehat{Y}$, we have $\widetilde{Y}_k = \widehat{Y}_k = Z_k$. Suppose that $\widetilde{Y}_i \leq \widehat{Y}_i$ for some i with $0 < i \leq k$. We will then show that $\widetilde{Y}_{i-1} \leq \widehat{Y}_{i-1}$. Let us write

$$\begin{aligned} \widehat{Y}_{i-1} &= E^{i-1} Z_{\widehat{\tau}_{i-1}} = 1_{\widehat{\tau}_{i-1}=i-1} Z_{i-1} + 1_{\widehat{\tau}_{i-1}>i-1} E^{i-1} E^i Z_{\widehat{\tau}_i} \\ &= 1_{\widehat{\tau}_{i-1}=i-1} Z_{i-1} + 1_{\widehat{\tau}_{i-1}>i-1} E^{i-1} \widehat{Y}_i. \end{aligned}$$

Then, by induction and Jensen's inequality,

$$\begin{aligned}\widehat{Y}_{i-1} &\geq 1_{\widehat{\tau}_{i-1}=i-1}Z_{i-1} + 1_{\widehat{\tau}_{i-1}>i-1}E^{i-1}\widetilde{Y}_i \\ &= 1_{\widehat{\tau}_{i-1}=i-1}Z_{i-1} + 1_{\widehat{\tau}_{i-1}>i-1}E^{i-1}\max_{p:\, i\leq p\leq \min(i+\kappa,k)} E^i Z_{\tau_p} \\ &\geq 1_{\widehat{\tau}_{i-1}=i-1}\widetilde{Y}_{i-1} + 1_{\widehat{\tau}_{i-1}>i-1}\max_{p:\, i\leq p\leq \min(i+\kappa,k)} E^{i-1} Z_{\tau_p}, \qquad (10)\end{aligned}$$

since for $\widehat{\tau}_{i-1} = i-1$ we have $i-1 = \inf\{j : i-1 \leq j \leq k,\ \widetilde{Y}_j \leq Z_j\}$, and so $\widetilde{Y}_{i-1} \leq Z_{i-1}$. We may write (10) as

$$\widehat{Y}_{i-1} \geq \widetilde{Y}_{i-1} + 1_{\widehat{\tau}_{i-1}>i-1}\Big(\max_{i\leq p\leq \min(i+\kappa,k)} E^{i-1}Z_{\tau_p} - \max_{i-1\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p}\Big).$$

Thus, after showing that $\widehat{\tau}_{i-1} > i-1$ implies

$$E^{i-1}Z_{\tau_{i-1}} \leq \max_{p:i\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p},$$

it follows that $\widehat{Y}_{i-1} \geq \widetilde{Y}_{i-1}$. It holds,

$$\begin{aligned}E^{i-1}Z_{\tau_{i-1}} &= 1_{\tau_{i-1}=i-1}Z_{i-1} + 1_{\tau_{i-1}>i-1}E^{i-1}Z_{\tau_i} \\ &\leq 1_{\tau_{i-1}=i-1}Z_{i-1} + 1_{\tau_{i-1}>i-1}\max_{p:i\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p}. \qquad (11)\end{aligned}$$

Then, on the set $\{\widehat{\tau}_{i-1} > i-1\}$ we have

$$Z_{i-1} < \max_{p:\, i-1\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p},$$

so, if $(\widehat{\tau}_{i-1} > i-1) \wedge (\ \tau_{i-1} = i-1)$, it follows that

$$Z_{i-1} < \max(Z_{i-1}, \max_{p:\, i\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p}).$$

Hence, if $(\widehat{\tau}_{i-1} > i-1) \wedge (\tau_{i-1} = i-1)$,

$$Z_{i-1} < \max_{p:\, i\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p}.$$

Thus, from (11) we have on the set $\{\widehat{\tau}_{i-1} > i-1\}$,

$$\begin{aligned}E^{i-1}Z_{\tau_{i-1}} &\leq 1_{\tau_{i-1}=i-1}\max_{p:\, i\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p} \\ &+ 1_{\tau_{i-1}>i-1}\max_{p:\, i\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p} \\ &= \max_{p:\, i\leq p\leq \min(i-1+\kappa,k)} E^{i-1}Z_{\tau_p}.\end{aligned}$$

■

The following corollary is a simple consequence of Theorem 3.1.

Corollary 3.2 *It holds* $Z_i \leq \widehat{Y}_i$, *for* $0 \leq i \leq k$.

Proof. Suppose $Z_i > \widehat{Y}_i$, then by Theorem 3.1, $Z_i > \widetilde{Y}_i$. So by the construction (8) we must have $\widehat{\tau}_i = i$ and thus $Z_i = E^i Z_{\widehat{\tau}_i} = \widehat{Y}_i$, hence a contradiction. ■

In the above analysis the window parameter κ was fixed. When κ is increased, more information of the input exercise policy is used in the improvement step (8). Therefore, it is intuitively clear that in general the quality of the corresponding improvement step will advance when κ increases. We have the following proposition.

Proposition 3.3 *Consider κ, κ' with $0 \leq \kappa < \kappa' \leq k$. Let $\hat{\tau}_i$ and $\hat{Y}_i$ be as in (8)-(9), and let $\hat{\tau}'_i$ and $\hat{Y}'_i$ be constructed according to (8)-(9) for the window parameter κ'. Then we have $\hat{Y}_i \leq \hat{Y}'_i$, $0 \leq i \leq k$.*

Proof. Obviously, we have $\hat{\tau}'_i \geq \hat{\tau}_i$ by construction. Hence, by (5) and Corollary 3.2 it follows that $\hat{Y}'_i = E^i Z_{\hat{\tau}'_i} = \sum_{p=i}^k E^i 1_{\hat{\tau}_i=p} Z_{\hat{\tau}'_p} = \sum_{p=i}^k E^i 1_{\hat{\tau}_i=p} E^p Z_{\hat{\tau}'_p}$ $\geq \sum_{p=i}^k E^i 1_{\hat{\tau}_i=p} Z_p = E^i Z_{\hat{\tau}_i} = \hat{Y}_i$. ■

As a result, in general the best policy improvement (8)-(9) is obtained by making in (8) a forward sweep over all exercise dates; hence by choosing $\kappa = k$. Indeed, when the input policy exercises much too early and κ is too small, it may happen that the improved policy (8) exercises much too early again, for instance, in situations where the reward process has low expectations in the near future but higher expectations on longer terms. In contrast, when κ is large enough, e.g. $\kappa = k$, such high future expectations can be detected by (8), and then the improved policy will exercise later. In Section 5 we will show this by an explicit example (Example 5.4), in the context of an iterative procedure based on (8)-(9).

Finally, we note that by choosing the smallest κ, hence $\kappa = 1$, (8) boils down to an improvement which basically goes back to Howard (1960). See also Irle (1980) and Putterman (1994).

4 Iterative construction of the optimal stopping time and the Snell envelope process

Naturally, we may construct by induction via the procedure (8)-(9) a sequence of pairs

$$\Big((\tau_i^{(m)})_{0 \leq i \leq k}, (Y_i^{(m)})_{0 \leq i \leq k} \Big)_{m=0,1,2,\ldots}$$

in the following way: Start with some family of stopping times $(\tau_i^{(0)})_{0 \leq i \leq k}$, which satisfies (5). A canonical starting family is obtained, for example, by taking $\tau_i^{(0)} \equiv i$. Suppose that for $m \geq 0$ the pair $\Big((\tau_i^{(m)}), (Y_i^{(m)}) \Big)$ is constructed, where

$$Y_i^{(m)} := E^i Z_{\tau_i^{(m)}}, \quad 0 \leq i \leq k,$$

and the stopping time family $(\tau_i^{(m)})$ satisfies (5). Then define

$$\begin{aligned} \tau_i^{(m+1)} &:= \inf\{j : i \leq j \leq k, \max_{p:\, j \leq p \leq \min(j+\kappa,k)} E^j Z_{\tau_p^{(m)}} \leq Z_j\} \\ &= \inf\{j : i \leq j \leq k,\ \widetilde{Y}_j^{(m+1)} \leq Z_j\}, \quad 0 \leq i \leq k, \end{aligned} \tag{12}$$

with

$$\widetilde{Y}_i^{(m+1)} := \max_{p:\, i\le p\le \min(i+\kappa,k)} E^i Z_{\tau_p^{(m)}}$$

being an intermediate dummy process and

$$Y_i^{(m+1)} := E^i Z_{\tau_i^{(m+1)}}, \quad 0 \le i \le k.$$

Clearly, the family $(\tau_i^{(m+1)})_{0\le i\le k}$ satisfies (5) as well.

Due to Theorem 3.1 the above constructed sequence satisfies

$$Y_i^{(0)} \le Y_i^{(m)} \le \widetilde{Y}_i^{(m+1)} \le Y_i^{(m+1)} \le Y_i^*, \quad m \ge 0, \quad 0 \le i \le k, \tag{13}$$

and, due to Corollary 3.2,

$$Z_i \le Y_i^{(m)}, \quad m \ge 1, \quad 0 \le i \le k. \tag{14}$$

By the following proposition, for each fixed i the sequence $(\tau_i^{(m)})_{m\ge1}$ constructed above is also nondecreasing in m and bounded from above by any optimal stopping time τ_i^*. It is possible, however, that $\tau_i^* < \tau_i^{(0)}$ (e.g. when $\tau_i^{(0)} \equiv k$).

Proposition 4.1 *Let (τ_i^*) be an optimal stopping family. For each $m \ge 1$, and i, $0 \le i \le k$, we have*

$$\tau_i^{(m)} \le \tau_i^{(m+1)} \le \tau_i^*.$$

Proof. Suppose that $\tau_i^* < \tau_i^{(m)}$ for some $m \ge 1$ and some i with $0 \le i \le k$. Then, by (13) and the definition of $\tau_i^{(m)}$,

$$Y^*_{\tau_i^*} \ge \widetilde{Y}^m_{\tau_i^*} > Z_{\tau_i^*},$$

so τ_i^* is not optimal, hence a contradiction. Thus, the right inequality is proved. Next suppose $\tau_i^{(m+1)} < \tau_i^{(m)}$ for some $m \ge 1$ and some i with $0 \le i \le k$. Then, by the definition of $\tau_i^{(m)}$ we have

$$\widetilde{Y}^{(m)}_{\tau_i^{(m+1)}} > Z_{\tau_i^{(m+1)}}.$$

On the other hand, according the definition of $\tau_i^{(m+1)}$, we have

$$\widetilde{Y}^{(m+1)}_{\tau_i^{(m+1)}} \le Z_{\tau_i^{(m+1)}}.$$

So, we get $\widetilde{Y}^{(m)}_{\tau_i^{(m+1)}} > \widetilde{Y}^{(m+1)}_{\tau_i^{(m+1)}}$, which contradicts (13). ∎

The next lemma formulates a "quasi-supermartingale property" for $Y^{(m)}$, and $\widetilde{Y}^{(m)}$.

Lemma 4.2 *For $m \geq 0$ and $0 \leq i < k$ it holds*

$$Y_i^{(m+1)} \geq \widetilde{Y}_i^{(m+1)} \geq E^i Y_{i+1}^{(m)} \geq E^i \widetilde{Y}_{i+1}^{(m)},$$

where $\tilde{Y}^{(0)}$ has to be interpreted as 0.

Proof. Using (13) we have
$Y_i^{(m+1)} \geq \widetilde{Y}_i^{(m+1)} = \max_{p:\, i \leq p \leq \min(i+\kappa,k)} E^i Z_{\tau_p^{(m)}} \geq E^i Z_{\tau_{i+1}^{(m)}} = E^i E^{i+1} Z_{\tau_{i+1}^{(m)}} = E^i Y_{i+1}^{(m)} \geq E^i \widetilde{Y}_{i+1}^{(m)}$. ∎

Due to the inequality chain (13) and Proposition 4.1 there exist a limit lower bound process Y^∞ and a limit family of stopping times (τ_i^∞),

$$Y_i^\infty := (\text{a.s.}) \lim_{m \uparrow \infty} \uparrow Y_i^{(m)} \quad \text{and} \quad \tau_i^\infty := (\text{a.s.}) \lim_{m \uparrow \infty} \uparrow \tau_i^{(m)}, \quad 0 \leq i \leq k, \quad (15)$$

where the uparrows indicate that the respective sequences are non-decreasing. Further it is clear that the family (τ_i^∞) satisfies (5). Since $\tau_i^{(m)}$ is an integer valued random variable in the set $\{i, \ldots, k\}$ for each m, we have almost surely, $\tau_i^{(m)}(\omega) = \tau_i^\infty(\omega)$ for $m > N(\omega)$. Therefore, $Z_{\tau_i^{(m)}} \to Z_{\tau_i^\infty}$ a.s., and so by dominated convergence we have,

$$Y_i^\infty = (\text{a.s.}) \lim_{m \uparrow \infty} \uparrow E^i Z_{\tau_i^{(m)}} = E^i Z_{\tau_i^\infty}, \quad 0 \leq i \leq k. \quad (16)$$

Funnily, it turns out that after at most k iterations the limit Y^∞ is attained and coincides with the Snell envelope. Hence, the limiting procedure (16) terminates.[2] We will show this by backward induction starting from the last exercise date using the quasi-supermartingale property Lemma 4.2. So, for the sequence of stopping families we have an analogue property; Proposition 4.4. As an alternative, we derive this directly from the iterative construction (12) by backward induction.

Proposition 4.3 *For i, $0 \leq i \leq k$, the following identity holds,*

$$Y_i^{(m)} = Y_i^\infty = Y_i^*, \quad \textit{for } m \geq k - i.$$

Proof. We use backward induction over i. For $i = k$ we have trivially $Y_k^{(m)} = Z_k = Y_k^*$ for $m \geq 0$. Suppose the assertion is already proved for some i with $0 < i \leq k$. Now let $m \geq k-(i-1)$, hence $m-1 \geq k-i \geq 0$. Then, by Lemma 4.2 and the induction hypothesis it follows that $Y_{i-1}^{(m)} \geq E^{i-1} Y_i^{(m-1)} = E^{i-1} Y_i^*$. On the other hand, by Corollary 3.2 we have $Y_{i-1}^{(m)} \geq Z_{i-1}$ since $m \geq 1$. Thus, $Y_{i-1}^{(m)} = Y_{i-1}^*$. ∎

[2] This issue was pointed out by a perceptive referee, and also by Christian Bender practically right after the submission of the first version.

Proposition 4.4 *For i, $0 \le i \le k$, it holds*

$$\tau_i^{(m)} = \tau_i^{\infty} = \tau_i^*, \quad \textit{for } m \ge k - i,$$

where τ_i^* is the first optimal stopping time after i, introduced in (3).

Proof. For $i = k$ we have trivially $\tau_k^{(m)} = \tau_k^*$ for $m \ge 0$. Suppose the assertion is already proved for all j, $i \le j \le k$, where $0 < i \le k$. Now let $m \ge k - (i-1)$, hence $m - 1 \ge k - i$. Then, by defining the event $A := \{Z_{i-1} \ge \max_{p:\, i-1 \le p \le \min(i-1+\kappa, k)} E^{i-1} Z_{\tau_p^{(m-1)}}\}$ and using the induction hypothesis, we may write

$$\begin{aligned}
\tau_{i-1}^{(m)} &= \inf\{j : i-1 \le j \le k, Z_j \ge \max_{p:\, j \le p \le \min(j+\kappa, k)} E^j Z_{\tau_p^{(m-1)}}\} \\
&= 1_A (i-1) + 1_{A^c} \inf\{j : i \le j \le k, Z_j \ge \max_{p:\, j \le p \le \min(j+\kappa, k)} E^j Z_{\tau_p^*}\} \\
&= 1_A \tau_{i-1}^* + 1_{A^c} \tau_i^* = \tau_{i-1}^*,
\end{aligned}$$

since $A \subset \{Z_{i-1} \ge E^{i-1} Z_{\tau_i^*}\}$ by the induction hypothesis again, and $A^c \subset \{\tau_{i-1}^* > i-1\}$. We thus have $\tau_{i-1}^* = i-1$ on A and $\tau_{i-1}^* = \tau_i^*$ on A^c. ∎

The following lemma gives a representation for the distance between two iterations.

Lemma 4.5 *For m, n with $1 \le m \le n$, and i, $0 \le i \le k$, we have*

$$0 \le Y_i^{(n)} - Y_i^{(m)} = E^i \sum_{p=\tau_i^{(m)}}^{\tau_i^{(n)}-1} 1_{\tau_p^{(m)}=p} (E^p Y_{p+1}^{(m)} - Y_p^{(m)}).$$

Proof. Due to Proposition 4.1 we have for $1 \le m \le n$, $\tau_i^{(n)} \le \tau_{\tau_i^{(n)}}^{(m)} \le \tau_{\tau_i^{(n)}}^{(n)} = \tau_i^{(n)}$. Hence, it holds $\tau_{\tau_i^{(n)}}^{(m)} = \tau_i^{(n)}$, and so $Y_i^{(n)} = E^i Z_{\tau_i^{(n)}} = E^i Z_{\tau_{\tau_i^{(n)}}^{(m)}} = \sum_{l=i}^k E^i 1_{\tau_i^{(n)}=l} Z_{\tau_l^{(m)}} = \sum_{l=i}^k E^i 1_{\tau_i^{(n)}=l} E^l Z_{\tau_l^{(m)}} = \sum_{l=i}^k E^i 1_{\tau_i^{(n)}=l} Y_l^{(m)} = E^i Y_{\tau_i^{(n)}}^{(m)}$. We thus may write (with empty sums being zero),

$$\begin{aligned}
0 \le Y_i^{(n)} - Y_i^{(m)} &= E^i Y_{\tau_i^{(n)}}^{(m)} - E^i Y_{\tau_i^{(m)}}^{(m)} \\
&= E^i \sum_{p=\tau_i^{(m)}}^{\tau_i^{(n)}-1} (Y_{p+1}^{(m)} - Y_p^{(m)}) \\
&= \sum_{\alpha \le \beta} \sum_{p=\alpha}^{\beta-1} E^i 1_{\tau_i^{(m)}=\alpha} 1_{\tau_i^{(n)}=\beta} (E^p Y_{p+1}^{(m)} - Y_p^{(m)}) \\
&= E^i \sum_{p=\tau_i^{(m)}}^{\tau_i^{(n)}-1} 1_{\tau_p^{(m)}=p} (E^p Y_{p+1}^{(m)} - Y_p^{(m)}),
\end{aligned}$$

where we have used in the final step that $\tau_p^{(m)} > p$ implies $E^p Y_{p+1}^{(m)} = Y_p^{(m)}$. ∎

By letting $n \uparrow \infty$ and using Proposition 4.1 and Proposition 4.3, Lemma 4.5 yields a representation for the distance between the m-th iteration and the Snell envelope. In fact, Lemma 4.5 states that the difference $Y_i^{(n)} - Y_i^{(m)}$, or $Y_i^* - Y_i^{(m)}$ respectively, can be seen as an aggregation of errors made due to exercise decisions according to policy $\tau^{(m)}$, before the generally better exercise time $\tau_i^{(n)}$, or τ_i^* respectively.

Iterative procedure (12)-(13) versus backward dynamic program

Note that on the one hand, in view of Proposition 4.3, the iterative procedure requires k iterations for constructing the whole Snell envelope. On the other hand, the backward dynamic program requires k recursion steps backward for constructing the whole Snell envelope. In this respect, if all occurring conditional expectations would be easily tractable (or, let us say if they would be given by God), the backward dynamic program should be preferred of course, as it doesn't require a forward sweep over all exercise dates at each step. However, particularly in high dimensional problems, the conditional expectations occurring in the backward dynamic program are usually difficult to evaluate, from a technical as well as from a computational point of view. For instance, in the Longstaff-Schwartz regression method one needs to choose an appropriate system of basis functions suiting to the problem under consideration. This system may be rather complicated for high dimensional problems (such as Bermudans due to a 40-factor LIBOR model). The alternative, plain Monte Carlo simulation of these conditional expectations in each backward recursion step, leads to an explosion of computation time since the degree of nesting increases with each step, see Remark 2.1. Yet, the very power of the iteration procedure based on (12)-(13), started up with $\tau_i^{(0)} \equiv i$ for example, lays in the following facts:

- For any m (in particular for low m, e.g. $1, 2, 3$), iteration $Y^{(m)}$ gives an approximation of the *whole* Snell envelope $Y_0^*, \ldots, Y_k^*$, with $Y_{k-m}^{(m)} = Y_{k-m}^*$, $\ldots$, $Y_k^{(m)} = Y_k^*$, while, in contrast, m backward steps of the backward dynamic program give the values $Y_{k-m}^*, \ldots, Y_k^*$, *and not more.*

- For fixed i the approximation $Y_i^{(m)}$ increases, and thus *improves*, in each iteration step m.

- By taking $\kappa = k$ we usually get very close to the Snell envelope with only a few iterations (e.g. $m = 1, 2, 3$), due to the particular nature of the iteration procedure (12)-(13), which is inspired by a canonical stopping strategy for Bermudan swaptions.

In fact, the iteration procedure (12)-(13) may be regarded as an extension of the backward dynamic program, see Figure 1. Thus, as we will see in detail in Section 6, in a Markovian setting the iteration procedure (12)-(13) allows for approximating the whole Snell envelope by plain Monte Carlo simulation using

a few iterations (e.g. $m = 1, 2, 3$). Moreover, as we will see in Section 7, we may get surprisingly good results!

		—	*Exercise*	*date*	→		
		0	1	$\cdots$	$k-2$	$k-1$	k
	0	$\mathbf{Y_0^{(0)}}$	$\mathbf{Y_1^{(0)}}$	$\cdots$	$\mathbf{Y_{k-2}^{(0)}}$	$\mathbf{Y_{k-1}^{(0)}}$	Y_k^*
	1	$\mathbf{Y_0^{(1)}}$	$\mathbf{Y_1^{(1)}}$	$\cdots$	$\mathbf{Y_{k-2}^{(1)}}$	Y_{k-1}^*	Y_k^*
|	2	$\mathbf{Y_0^{(2)}}$	$\mathbf{Y_1^{(2)}}$		Y_{k-2}^*	Y_{k-1}^*	Y_k^*
Iteration	·	·	·		·	·	·
level	·	·	·		·	·	·
↓	·	·	·		·	·	·
	$k-1$	$\mathbf{Y_0^{(k-1)}}$	Y_1^*	$\cdots$	Y_{k-2}^*	Y_{k-1}^*	Y_k^*
	k	Y_0^*	Y_1^*	$\cdots$	Y_{k-2}^*	Y_{k-1}^*	Y_k^*

Figure 1. The iterative procedure as an extension (bold terms) of the Backward Dynamic Program

5 Convergent upper bounds by the dual approach

We now deduce by a dual approach a sequence of upper bound processes from the sequence of lower approximations of the Snell envelope, $Y^{(m)}$, $m = 0, 1, \ldots$, constructed in Section 4. This dual approach, developed in the works of Davis & Karatzas (1994), Rogers (2001), Haugh & Kogan (2001), is based on the following observation. For any supermartingale $(S_j)_{0 \le j \le k}$ with $S_0 = 0$ we have,

$$\begin{aligned} Y_0^* &= \sup_{\tau \in \{0,\ldots,k\}} E^0 Z_\tau \le \sup_{\tau \in \{0,\ldots,k\}} E^0 (Z_\tau - S_\tau) \\ &\le E^0 \max_{0 \le j \le k} (Z_j - S_j), \end{aligned} \tag{17}$$

hence the right-hand side provides a (dual) upper bound for Y_0^*. Rogers (2001) and independently Haugh & Kogan (2001) show, that the equality in (17) is attained at the martingale part of the Doob-Meyer decomposition of Y^*,

$$M_0^* := 0; \quad M_j^* := \sum_{l=1}^{j} (Y_l^* - E^{l-1} Y_l^*), \quad 1 \le j \le k, \tag{18}$$

and also at the shifted Snell envelope process

$$S_j^* := Y_j^* - Y_0^*, \quad 0 \le j \le k. \tag{19}$$

The next lemma provides a somewhat more general class of supermartingales, which turn (17) into an equality. Moreover, we show that the equality holds almost sure.

Lemma 5.1 *Let $\mathcal{S}$ be the set of supermartingales S with $S_0 = 0$. Let $S \in \mathcal{S}$ be such that $Z_j - Y_0^* \leq S_j$ for each j: $1 \leq j \leq k$. Then,*

$$Y_0^* = \max_{0 \leq j \leq k} (Z_j - S_j) \quad a.s. \tag{20}$$

Proof. By the assumptions we have $Z_j \leq Y_0^* + S_j$ for each j: $0 \leq j \leq k$, and so by (17),

$$0 \leq E^0 \max_{0 \leq j \leq k} (Z_j - S_j) - Y_0^* \leq 0.$$

Hence, we have $E^0 \max_{0 \leq j \leq k} (Z_j - S_j - Y_0^*) = 0$ and $Z_j - S_j - Y_0^* \leq 0$ for each j: $0 \leq j \leq k$, which yields (20). ■

Note that both (18) and (19) satisfy the conditions of Lemma 5.1, for more examples see Kolodko & Schoenmakers (2004).

The duality representation provides a simple way to estimate the Snell envelope from above, using a lower approximation process denoted by $\overline{Y}$, hence $\overline{Y} \leq Y^*$. Let $\overline{M}$ be the martingale part of the Doob-Meyer decomposition of $\overline{Y}$, satisfying

$$\begin{aligned} \overline{M}_0 &= 0; \\ \overline{M}_j &= \overline{M}_{j-1} + \overline{Y}_j - E^{j-1}\overline{Y}_j \\ &= \sum_{l=1}^{j} \overline{Y}_l - \sum_{l=1}^{j} E^{l-1}\overline{Y}_l, \quad 1 \leq j \leq k. \end{aligned}$$

Then, according to (17),

$$Y_0^* \leq E^0 \max_{0 \leq j \leq k} (Z_j - \overline{M}_j) =: \overline{Y}_0^{up}. \tag{21}$$

Let us now consider the sequence of lower bound processes $Y^{(m)}$, $m = 0, 1, \ldots$ from Section 4. Analogue to (21) we may deduce a sequence of upper bound processes,

$$\begin{aligned} Y_i^{(m),up} &:= E^i \max_{i \leq j \leq k} (Z_j - \sum_{l=i+1}^{j} Y_l^{(m)} + \sum_{l=i+1}^{j} E^{l-1} Y_l^{(m)}) \\ &= Y_i^{(m)} + E^i \max_{i \leq j \leq k} (Z_j - Y_j^{(m)} + \sum_{l=i}^{j-1} (E^l Y_{l+1}^{(m)} - Y_l^{(m)})) \qquad (22) \\ &=: Y_i^{(m)} + \Delta_i^{(m)}, \qquad 0 \leq i \leq k. \end{aligned}$$

Let us also define $\widehat{\Delta}_i^{(m)} := Y_i^{(m),up} - Y_i^{(m+1)}$. By the next proposition $\Delta_i^{(m)}$ and $\widehat{\Delta}_i^{(m)}$ depend in a sense on how far the lower bound process $Y^{(m)}$ is away from being a supermartingale.

Proposition 5.2 *For $m \geq 1$, $0 \leq i \leq k$, it holds,*

$$0 \quad \leq \quad \Delta_i^{(m)} \leq E^i \sum_{j=i}^{k-1} 1_{\tau_j^{(m)}=j} \max\left(E^j Y_{j+1}^{(m)} - Y_j^{(m)},\ 0\right), \tag{23}$$

$$0 \quad \leq \quad \widehat{\Delta}_i^{(m)} \leq E^i \sum_{j=i+1}^{k-1} 1_{\tau_j^{(m)}=j} \max\left(E^j Y_{j+1}^{(m)} - Y_j^{(m)},\ 0\right) - 1_{\tau_i^{(m)}>i}(Y_i^{(m+1)} - Y_i^{(m)}). \tag{24}$$

Proof. By property (5) it holds $Y_j^{(m)} = E^j Y_{j+1}^{(m)}$ on the set $\{\tau_j^{(m)} > j\}$. So, the first inequality follows from (22) and (14). From Lemma 4.2 and (13) we have

$$1_{\tau_i^{(m)}=i} \max\left(E^i Y_{i+1}^{(m)} - Y_i^{(m)},\ 0\right) - 1_{\tau_i^{(m)}=i}(Y_i^{(m+1)} - Y_i^{(m)}) \leq 0,$$

and then the second inequality follows from $\widehat{\Delta}_i^{(m)} = \Delta_i^{(m)} - (Y_i^{(m+1)} - Y_i^{(m)})$ and (23). ■

We note that estimation (23) is basically the same as in Kolodko & Schoenmakers (2003, 2004a). As an immediate consequence of Proposition 5.2 and Proposition 4.3, we obtain

Proposition 5.3 *For $0 \leq i \leq k$, we have the following identity,*

$$Y_i^{(m),up} = Y_i^*, \qquad for \quad m \geq \max(k-i-1, 0).$$

Proof. For $i = k$ and $i = k-1$ this follows trivially from (22) and the convention of empty sums being zero. Due to (24) and Lemma 4.2 it holds

$$\widehat{\Delta}_i^{(m)} \leq E^i \sum_{j=i+1}^{k-1} (Y_j^{(m+1)} - Y_j^{(m)}).$$

Then we can use straightforward backward induction starting from $i = k-1$. ■

The choice of window parameter κ

It should be noted that all results in Sections 3-5 hold for any window parameter κ between 1 and k. However, due to Proposition 3.3 and Example 5.4 below, it is clear that $\kappa = k$, hence a forward sweep over all exercise dates, is in general the optimal choice.

Example 5.4 Consider a reward process Z with the following properties: $k \geq 2$, $Z_0 = 1$, Z_j is a martingale for $0 \leq j \leq k-2$ with $0 \leq Z_j \leq 2$, Z_{k-1} is independent of Z_j, $0 \leq j \leq k-2$ with $EZ_{k-1} = 3$, and $Z_k = 1$. So, obviously, $\tau_i^* = k-1$ for $0 \leq i < k-1$. Let us now start with the initial policy $\tau_i^{(0)} = i$. Then, it is easily seen that for an iteration (12) based on $\kappa = 1$ we have $\tau_0^{(m)} = 0$ for $0 \leq m \leq k-2$, and $\tau_0^{(k-1)} = k-1$. Hence, $k-1$ iterations are required for obtaining the optimal stopping time at $j = 0$. In contrast, iteration (12) based on $\kappa = k$ yields the optimal stopping time already in one iteration step.

For Example 5.4 in a Markovian setting, a plain Monte Carlo implementation of the iterative procedure in the spirit of Section 6 below, based on $\kappa = 1$, would give rise to about k-fold nested simulations, hence computation times exploding with k. In practice, nevertheless, due to special features of a problem under consideration it may happen that a suitable choice of κ smaller than k gives computationally more efficient results. In the case of Bermudan swaptions, for example, the underlying Europeans far away from the monitor date have usually relatively low value. So a κ smaller than k may be computationally more efficient in this case.

6 Monte Carlo Algorithm for the Iterative Method

We now describe how to construct an approximation of the exercise policy $(\tau_i^{(m)})_{0\le i\le k}$, given by (12), by the Monte Carlo method. For this we assume, that Z is a function of an $\mathbb{R}^D$-valued Markov random sequence $(L_i)_{0\le i\le k}$ adapted to the filtration $\mathbb{F} = (\mathcal{F}_i)_{0\le i\le k}$, i.e.

$$Z_i = h_i(L_i), \quad 0 \le i \le k.$$

This assumption is fulfilled in practically all applications, see also Remark 2.1.

For $m \ge 0$ and i, $0 \le i \le k$, we introduce the *random* set of dates

$$\begin{aligned}\Theta_i^{(m)} \quad := \quad & \{\eta_j : \quad \eta_1 = \tau_i^{(m)}; \quad \text{for } j \ge 1, \text{ if } \eta_j < k \text{ then } \eta_{j+1} = \tau_{\eta_j+1}^{(m)}\} \\ & \subset \{i, \dots, k\},\end{aligned}$$

at which the exercise policy $\tau^{(m)}$ says "exercise". Obviously, for $m \ge 1$, the set $\Theta_i^{(m)}$ can be represented as

$$\Theta_i^{(m)} = \{j \ge i : Z_j \ge \widetilde{Y}_j^{(m)}\}, \quad 0 \le i \le k,$$

and, due to the nondecreasingness of the sequence $(\widetilde{Y}_i^{(m)})_{m\ge1}$, we have

$$\Theta_i^{(m+1)} \subset \Theta_i^{(m)} \subset \Theta_i^{(1)}, \quad m \ge 1, \quad 0 \le i \le k. \tag{25}$$

For each $m \ge 0$ and i, $0 \le i \le k$, the stopping times $\tau_p^{(m)}$, $i \le p \le k$, may be retrieved from $\Theta_i^{(m)}$ via

$$\tau_p^{(m)} = \inf \Theta_i^{(m)} \cap \{p, \dots, k\}.$$

For a generic trajectory $L^{(\omega)}$ of the process L starting afresh at date i in $L_i^{(\omega)}$, and $m \ge 1$, we construct the set $\Theta_i^{(m)}$ by a recursive procedure $R(m)$. For this we fix a sequence of numbers N_m, $m \ge 1$.

Initial procedure $R(1)$**:**
We construct the set $\Theta_i^{(1)}$ on a trajectory $L^{(\omega)}$, starting in $L_i^{(\omega)}$ at date i, by checking the condition

$$Z_j \geq \widetilde{Y}_j^{(1)} = \max_{j \leq p \leq \min(j+\kappa,k)} E^j Z_{\tau_p^{(0)}}$$

on $L^{(\omega)}$ at each exercise date j, $j \geq i$, for the given initial stopping family $\tau^{(0)}$. In general the conditional expectations on the righthand side need to be estimated by Monte Carlo. However, in many cases explicit formulas or good explicit approximations are available. For instance, in the case of Bermudan swaptions and trivial initial stopping family $\tau_i^{(0)} \equiv i$, we need the values of still-alive Europeans which can be approximated within 0.3% relative by explicit formulas (e.g., see Schoenmakers (2005)).

Recursion step $R(m+1)$**:**
We here construct recursively the set $\Theta_i^{(m+1)}$ on a trajectory $L^{(\omega)}$ which starts in $L_i^{(\omega)}$ at date i.

Construct $\Theta_i^{(m)}$ on $L^{(\omega)}$ by $R(m)$;

for $j \in \Theta_i^{(m)}$ **do**

begin

Simulate N_m trajectories $\xi^{(\alpha)} = (\xi_l^{(\alpha)})_{l=j}^k$, $1 \leq \alpha \leq N_m$, under the conditional measure $P^{L_j^{(\omega)}}$ (hence $\xi_j^{(\alpha)} = L_j^{(\omega)}$ for each α);

for $\alpha = 1, \ldots, N_m$ **do**

begin

Construct $\Theta_j^{(m,\alpha)}$ on $\xi^{(\alpha)}$ by $R(m)$;

for $p = j, \ldots, \min(j+\kappa, k)$ **do**

$Estimate\_tau[p,\alpha] := \inf \Theta_j^{(m,\alpha)} \cap \{p, \ldots, k\}$

end;

for $p = j, \ldots, \min(j+\kappa, k)$ **do**

$Estimate\_Y[p] := \frac{1}{N_m} \sum_{\alpha=1}^{N_m} h_{Estimate\_tau[p,\alpha]} \left(\xi^{(\alpha)}_{Estimate\_tau[p,\alpha]} \right)$;

$Estimate\_Ytilde[j] := \max_{j \leq p \leq \min(j+\kappa,k)} Estimate\_Y[p]$;

if $h_j(L_j^{(\omega)}) \geq Estimate\_Ytilde[j]$ **then** $j \in \Theta_i^{(m+1)}$

end.

Note, that $\Theta_i^{(m)}$ on $L^{(\omega)}$ and $\Theta_j^{(m,\alpha)}$ on $\xi^{(\alpha)}$ in the procedure $R(m+1)$ are constructed by calling $R(m)$ recursively.

By computing $\Theta_0^{(m)}$ on a set of trajectories of the process L using the above described procedure, we can estimate the lower bound $Y_0^{(m)}$, for $m \geq 1$, as

follows by Monte Carlo. Simulate for $\alpha = 1, \ldots, N$ a trajectory $(L_j^{(\alpha)})_{0 \le j \le k}$ under the measure P^{L_0}. Then, for each trajectory $L^{(\alpha)}$ determine the index $\eta^{(m,\alpha)} := \inf \Theta_0^{(m,\alpha)}$, which is an estimate of $\tau_0^{(m)}$ on this trajectory. Next, compute

$$Y_0^{(m)} \approx \frac{1}{N} \sum_{\alpha=1}^{N} h_{\eta^{(m,\alpha)}} \left(L_{\eta^{(m,\alpha)}}^{(\alpha)} \right).$$

Worst case complexity analysis of the algorithm

Let us fix an iteration level $m \ge 1$ and take $\kappa = k$. For simplicity we take $N_p = N$ for $1 \le p \le m$. Let $C_{p,i}$, $1 \le p \le m$, be the cost of the recursive procedure $R(p)$ starting at exercise date i, and c be the cost, on average, of simulating the process L from one exercise date to the next. Then, for $p = 1$ we have roughly, when always $\#\Theta_i^{(1)} \approx k - i$ and k is not too small,

$$C_{1,i} \approx \frac{(k-i)^{1+\xi} N^{\xi}}{(1+\xi)!} c, \tag{26}$$

where $\xi = 0$ if (quasi-)analytical formulas for $E^i Z_{\tau_j^{(0)}}$, $i \le j \le k$, are available, and $\xi = 1$ otherwise. From the recursive pseudo-code it follows that in the worst case, when always $\#\Theta_i^{(p)} \approx k - i$, we have

$$C_{p+1,i} \approx \sum_{j=i}^{k} N C_{p,j}, \quad 1 \le p < m \tag{27}$$

(we assume $N \gg 1$). Then, it is easy to show by induction from (26), (27), and the asymptotic identity $\sum_{j=1}^{l} j^p = \frac{l^{p+1}}{p+1}(1 + O(l^{-1}))$ for $l \to \infty$ and fixed p, $p \ge 1$, that

$$C_{m,i} \approx \frac{(k-i)^{m+\xi} N^{m-1+\xi}}{(m+\xi)!}\, c$$

for larger k. So, the total cost of estimating $Y_0^{(m)}$ is in worst cases approximately

$$Cost(Y_0^{(m)}) \approx \frac{k^{m+\xi} N^{m+\xi}}{(m+\xi)!}\, c. \tag{28}$$

We emphasize that (28) holds for fixed m and k large enough, but not for fixed k and $m \to \infty$! For the important case $m = 2$ and $\xi = 0$, which applies for example in Section 7, (28) yields $Cost(Y_0^{(2)}) \approx N^2 k^2 c/2$.

Remark 6.1 The cost estimate (28) is really "worst case". In usual applications the inclusion sequence (25) thins out rapidly from the right to the left and, as a consequence, the algorithm performs usually much faster as one would expect on the basis of (28). We note further that, in practice, it turns out more efficient to take N_{p+1} much smaller than N_p rather than equal (see for instance Section 7). However, a detailed analysis of this issue may depend on the particular problem under consideration and is considered beyond the scope of this article.

Remark 6.2 *(variance reduced Monte Carlo simulation of $Y^{(m)}$)* We can reduce the number of Monte Carlo simulations for $Y^{(m)}$ by using the following variance reduced representation,

$$Y_i^{(m)} = E^i Z_{\tau_i^{(m)}} = E^i Z_{\tau_i^{(m-1)}} + E^i(Z_{\tau_i^{(m)}} - Z_{\tau_i^{(m-1)}}). \tag{29}$$

One can expect that $Z_{\tau_i^{(m-1)}}$ and $Z_{\tau_i^{(m)}}$ are strongly correlated and thus the variance of $(Z_{\tau_i^{(m)}} - Z_{\tau_i^{(m-1)}})$ will be less than the variance of $Z_{\tau_i^{(m)}}$. So, the computation of $E^i(Z_{\tau_i^{(m)}} - Z_{\tau_i^{(m-1)}})$ for a given accuracy, can usually be done with less Monte Carlo simulations than needed for direct simulation of $E^i Z_{\tau_i^{(m)}}$.

Remark 6.3 *(stability of the algorithm)* An important issue remaining is the stability of the presented Monte Carlo algorithm, which is derived from the pseudo algorithm in Section 4 by replacing mathematical expectations with Monte Carlo approximations in fact. The stability question is highly non-trivial and has been studied in the subsequent work Bender & Schoenmakers (2004). From the results in this paper it follows that the here presented algorithm is numerically stable indeed.

7 A numerical example: Bermudan swaptions in the LIBOR market model

Let us first recall the LIBOR Market Model with respect to a tenor structure $0 < T_1 < \ldots < T_n$ in the spot LIBOR measure P^*. The dynamics of the forward LIBORs $L_i(t)$, defined in the interval $[0, T_i]$ for $1 \le i < n$, are governed by the following system of SDE's (e.g., see Jamshidian (1997)),

$$dL_i = \sum_{j=\kappa(t)}^{i} \frac{\delta_j L_i L_j \, \gamma_i \cdot \gamma_j}{1 + \delta_j L_j} \, dt + L_i \, \gamma_i \cdot dW^*, \tag{30}$$

where $\delta_i = T_{i+1} - T_i$ are day count fractions, $t \to \gamma_i(t) = (\gamma_{i,1}(t), \ldots, \gamma_{i,d}(t))$ are deterministic volatility vector functions defined in $[0, T_i]$ (called factor loadings), and $\kappa(t) := \min\{m : T_m \ge t\}$ denotes the next reset date at time t. In (30), $(W^*(t) \mid 0 \le t \le T_{n-1})$ is a standard d-dimensional Wiener process under the measure P^* with d, $1 \le d < n$, being the number of driving factors. The spot LIBOR measure is induced by the numeraire

$$B_*(t) := \frac{B_{\kappa(t)}(t)}{B_1(0)} \prod_{i=1}^{\kappa(t)-1} (1 + \delta_i L_i(T_i)) \,,$$

where $B_i(t)$ is the value of a zero coupon bond with face value \$1 at time $t \le T_i$.

A (payer) swaption contract with maturity T_i and strike θ with principal \$1 gives the right to contract at T_i for paying a fixed coupon θ and receiving floating

LIBOR at the settlement dates $T_{i+1},\ldots,T_n$. So the contract is equivalent with a pay-off at maturity T_i given by

$$S_{T_i,T_n}(T_i) := \left(\sum_{j=i}^{n-1} B_{j+1}(T_i)\delta_j \left(L_j(T_i)-\theta\right)\right)^+ .$$

A Bermudan swaption issued at $t=0$ gives the right to exercise a cash-flow

$$B_*(\mathcal{T}_\tau)Z_\tau := S_{\mathcal{T}_\tau,T_n}(\mathcal{T}_\tau)$$

at an exercise date $\mathcal{T}_\tau \in \{\mathcal{T}_1,\ldots,\mathcal{T}_k\} \subset \{T_1,\ldots,T_n\}$, to be decided by the option holder.

Since the system $(L(T_i), B_*(T_i))$ is Markovian we may compute Bermudan swaption prices in the LIBOR model (30) by the algorithm described in Section 6. For our simulation experiments we use the following LIBOR volatility structure,

$$\gamma_i(t) = cg(T_i-t)e_i, \quad \text{where} \quad g(s) = g_\infty + (1-g_\infty+as)e^{-bs}$$

is a parametric volatility function proposed by Rebonato (1999), and e_i are d-dimensional unit vectors, decomposing some input correlation matrix of rank d. For generating LIBOR models with different numbers of factors d, we take as a basis a correlation structure of the form

$$\rho_{ij} = \exp(-\varphi|i-j|); \qquad i,j = 1,\ldots,n-1 \tag{31}$$

which has full-rank for $\varphi > 0$, and then for a particular choice of d we deduce from ρ a rank-d correlation matrix ρ^d with decomposition $\rho^d_{ij} = e_i \cdot e_j$, $1 \le i,j < n$, by principal component analysis. We note that instead of (31) it is possible to use more general and economically more realistic correlation structures. For instance the parametric structures of Schoenmakers & Coffey (2003).

We take the following model parameters: A flat 10% initial LIBOR curve over a 40 period quarterly tenor structure, and the parameters

$$n = 41,\ \delta_i = 0.25,\ c = 0.2,\ a = 1.5,\ b = 3.5,\ g_\infty = 0.5,\ \varphi = 0.0413. \tag{32}$$

We consider Bermudan swaptions with yearly exercise opportunities, hence $\mathcal{T}_i = T_{4i}$, $i = 1,\ldots,10$. For a "practically exact" numerical integration of the SDE (30), we used the log-Euler scheme with $\Delta t = \delta/5$ (e.g., see also Kurbanmuradov, Sabelfeld and Schoenmakers (2002)).

By the algorithm in Section 6, based on $\kappa = k$, we now compute, starting from $\tau_i^{(0)} \equiv i$, two successive lower bounds $Y_0^{(m)}$: $Y_0^{(1)}$ and $Y_0^{(2)}$, and the dual upper bound $Y_0^{(1),up}$ for different number of factors d and for different types of options: in-the-money (ITM) with $\theta = 0.08$, at-the-money (ATM) with $\theta = 0.1$ and out-of-the-money (OTM) with $\theta = 0.12$.

Although closed form expressions for European swaptions do not exist in a LIBOR market model, there do exist very accurate (typically better, than 0.3%

relative) approximate formulas (see e.g. Schoenmakers (2005)), which we use for the computation of $\tau^{(1)}$. Thus, $Y_0^{(1)}$ can be computed by an ordinary Monte Carlo simulation and for $Y_0^{(2)}$ and $Y_0^{(1),up}$ we can use Monte Carlo simulation with only one degree of nesting. For $Y_0^{(1)}$ we construct 10^6 trajectories. Next, we construct $Y_0^{(2)}$ using variance reduction technique (29), where we used for estimating the expectation in the second term a number of (outer) Monte Carlo simulations varying from 2000 to 15 000 (2000 for ITM, 10 000 for ATM, 15 000 for ITM) in order to keep standard deviations within 0.5%. For each outer trajectory, we use 100 nested (inner) simulations. Further, $Y_0^{(1),up} - Y_0^{(1)}$ is simulated with 100 inner and 1000 outer simulations to get comparable standard deviations. We compare the results with the lower bound $Y_{A,0} = EZ_{\tau_{A,0}}$ and the dual upper bound $Y_{A,0}^{up}$ due to the stopping family, constructed by strategy I of the Andersen method (see Andersen (1999)):

$$\tau_{A,i} = \inf\{j \geq i : B_*(\mathcal{T}_j)Z_j \geq H_j\}.$$

The sequence of constants H_j is pre-computed using 10^5 simulations. We use 10^6 Monte Carlo trajectories for $Y_{A,0}$ and 5000 Monte Carlo trajectories (with 1000 inner simulations) for $Y_{A,0}^{up} - Y_{A,0}$, in order to get comparable standard deviations. The results are presented in Table 1.

We conclude, that for ITM swaptions the gap between the first iteration of the lower bound $Y_0^{(1)}$ and its dual $Y_0^{(1),up}$ is less than 0.5% relative, while for ATM and OTM swaptions it is about 2% relative. However, the second iteration, the lower bound $Y_0^{(2)}$, gives the Bermudan price with an accuracy better, than 1% relative in all cases. Further we note that for multi-factor models even the first iteration $Y_0^{(1)}$ is higher than $Y_{A,0}$. The dual upper bounds $Y_0^{(1),up}$ and $Y_{A,0}^{up}$ are almost the same, however. The dual upper bounds are computed via a standard upper bound estimator as in Andersen & Broadie (2001). For more efficient dual estimators, see Kolodko & Schoenmakers (2003, 2004a). The computing times for generating the values in Table 1 are reported in Table 2.

Table 1. (all the values are in basis points)

θ	d	$Y_0^{(1)}$ (SD)	$Y_0^{(2)}$ (SD)	$Y_0^{(1),up}$ (SD)	$Y_{A,0}$ (SD)	$Y_{A,0}^{up}$ (SD)
0.08 (ITM)	1	1104.6(0.5)	1108.9(2.4)	1109.4(0.7)	1107.7(0.5)	1109.0(0.5)
	2	1098.6(0.4)	1100.5(2.4)	1103.7(0.7)	1097.5(0.4)	1104.1(0.6)
	10	1094.4(0.4)	1096.9(2.1)	1098.1(0.6)	1093.0(0.4)	1099.5(0.6)
	40	1093.6(0.4)	1096.1(2.0)	1096.6(0.6)	1092.9(0.4)	1098.2(0.5)
0.10 (ATM)	1	374.3(0.4)	381.2(1.6)	382.9(0.8)	381.2(0.4)	383.1(0.4)
	2	357.9(0.3)	364.4(1.5)	366.4(0.8)	354.6(0.4)	367.4(0.6)
	10	337.8(0.3)	343.5(1.3)	345.6(0.7)	331.9(0.3)	348.2(0.6)
	40	332.6(0.3)	338.7(1.2)	341.2(0.8)	327.0(0.3)	342.7(0.6)
0.12 (OTM)	1	119.0(0.2)	121.0(0.6)	121.3(0.4)	120.5(0.2)	121.1(0.1)
	2	112.7(0.2)	113.8(0.5)	114.9(0.4)	110.0(0.2)	114.4(0.3)
	10	100.2(0.2)	100.7(0.4)	101.5(0.3)	95.7(0.2)	102.1(0.3)
	40	96.5(0.2)	96.9(0.4)	97.7(0.3)	92.2(0.2)	98.1(0.3)

Table 2. Computing times (minutes) for the values in Table 1 (Pentium III).

θ	d	$Y_0^{(1)}$	$Y_0^{(2)} - Y_0^{(1)}$	$Y_0^{(1),up} - Y_0^{(1)}$	$Y_{A,0}$	$Y_{A,0}^{up} - Y_{A,0}$
0.08 (ITM)	1	7.6	10.6	3.0	12.4	14.9
	2	7.2	11.0	3.1	11.7	14.7
	10	7.7	12.8	3.6	12.7	17.7
	40	10.4	27.2	6.1	17.8	29.9
0.10 (ATM)	1	21.4	27.1	5.2	26.6	26.2
	2	21.7	38.6	5.4	26.6	26.3
	10	24.7	47.6	6.6	30.25	33.0
	40	34.5	79.4	11.4	45.7	57.7
0.12 (OTM)	1	30.5	9.9	6.9	33.9	33.3
	2	31.4	8.5	7.2	34.2	34.2
	10	36.6	9.8	11.5	39.4	43.4
	40	63.4	17.3	21.6	58.7	78.1

Computational time

We note that, in order to exclude systematic errors due to Euler simulation of the LIBOR SDE, we have taken relatively small time steps ($\Delta t = \delta/5$). For an acceptable accuracy in practice one may take larger time steps, for example $\Delta t = \delta$ (e.g. see Kurbanmuradov, Sabelfeld & Schoenmakers (2002)). Therefore, the computation times in Table 2 may be substantially reduced by optimizing Δt and, further, by optimizing the trade of between the effort used for the computation of $Y^{(1)}$ and $Y^{(2)} - Y^{(1)}$, by optimizing the choice of the window parameter κ, and by other variance reduction techniques (e.g., antithetic variables). However, we consider these numerical optimizations beyond the scope of this article and in any case we expect that, when the producers of microprocessor chips keep "riding the exponential", computation of higher order iterations, hence practically exact Bermudan prices, will become feasible in the near future.

Acknowledgements

The present paper strongly benefited from Christian Bender who gave many suggestions which have been led to a substantially improved presentation, and showed the termination of the iterative procedure from Lemma 4.5. We thank the anonymous referees for their useful remarks and pointing out the termination issue also. J.S. is grateful to Martin Schweizer for helpful discussions and suggestions. This work is supported by the DFG Research Center MATHEON in Berlin.

References

[1] Andersen, L.: A simple approach to the pricing of Bermudan swaptions in the multifactor LIBOR market model. Journal of Computational Finance, **3**(2), 5–32 (1999)

[2] Andersen, L., Broadie, M.: A primal-dual simulation algorithm for pricing multidimensional American options. Management Science, **50**(9), 1222–1234 (2004)

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

[4] Belomestny, D., Milstein, G.N.: Monte Carlo evaluation of American options using consumption processes. To appear in Int. Journal of Theoretical and Applied Finance, WIAS-Preprint No. 930 (2004)

[5] Bender, C., Schoenmakers, J.: An iterative algorithm for multiple stopping: Convergence and stability. WIAS-Preprint No. 991 (2004)

[6] Berridge, S.J., Schumacher, J.M.: An irregular grid approach for pricing high-dimensional American options. Working paper (2004)

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

[8] Broadie, M., Glasserman, P., Ha, Z.: Pricing American options by simulations using a stochastic mesh with optimized weights. In: Probabilistic Constrained Optimization (ed. S. Uryasev): Methodology and Applications. Norwell: Kluwer Publishers 2000, pp. 32–50

[9] Davis, M.H.A., Karatzas, I.: A Deterministic Approach to Optimal Stopping. In: Probability, Statistics and Optimisation (ed. F.P. Kelly). NewYork Chichester: John Wiley & Sons Ltd. 1994, pp. 455–466

[10] Duffie, D.: Dynamic Asset Pricing Theory. Princeton: Princeton University Press 2001

[11] Elliot, R.J., Kopp, P.E.: Mathematics of Financial Markets. Berlin Heidelberg NewYork: Springer 1999

[12] Glasserman, P.: Monte Carlo Methods in Financial Engineering. Berlin Heidelberg NewYork: Springer 2003

[13] Haugh, M.B., Kogan, L.: Pricing American options: a duality approach. Operations Research, **52**(2), 258–270 (2004)

[14] Howard, R.: Dynamic programming and Markov processes. Cambridge, Massachusetts: MIT Press 1960

[15] Irle, A.: On the Best Choice Problem with Random Population Size. Zeitschrift für Operations Research, **24**, 177–190 (1980)

[16] Jamshidian, F.: LIBOR and swap market models and measures. Finance and Stochastics, **1**, 293–330 (1997)

[17] Jamshidian, F.: Minimax optimality of Bermudan claims: and arbitrage argument approach without probability theory. Working Paper (2003)

[18] Jamshidian, F.: Numeraire-invariant option pricing & american, bermudan, and trigger stream rollover. Working paper (2004)

[19] Joshi, M.S., Theis, J.: Bounding Bermudan swaptions in a swap-rate market model. Quantitative Finance, **2**(5), 371-361 (2002)

[20] Kolodko, A., Schoenmakers, J.: An Efficient Dual Monte Carlo Upper Bound for Bermudan Style Derivative. WIAS-Preprint No. 877 (2003)

[21] Kolodko, A., Schoenmakers, J.: Iterative construction of the optimal Bermudan stopping time. WIAS-Preprint No. 926 (2004)

[22] Kolodko, A., Schoenmakers, J.: Upper Bound for Bermudan Style Derivatives. Monte Carlo Methods and Applications, **10**(3-4), 331–343 (2004a)

[23] Kurbanmuradov, O., Sabelfeld, K., Schoenmakers, J.: Lognormal approximations to Libor market models. Journal of Computational Finance, **6**(1), 69–100 (2002)

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

[25] Milstein, G.N., Reiß, O., Schoenmakers, J.: A new Monte Carlo method for American options. Int. Journal of Theoretical and Applied Finance, **7**(5), 591–614 (2004)

[26] Puterman, M.: Markov decision processes. NewYork: Wiley 1994

[27] Rebonato, R.: Volatility and Correlation. NewYork Chichester: John Wiley & Sons Ltd. 1999

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

[29] Schoenmakers, J.: Robust Libor Modelling and Pricing of Derivative Products. BocaRaton London NewYork Singapore: Chapman & Hall – CRC Press 2005

[30] Schoenmakers, J., Coffey, B.: Systematic Generation of Parametric Correlation Structures for the LIBOR Market Model. International Journal of Theoretical and Applied Finance, **6**(5), 507-519 (2003)

[31] Shiryayev, A.N.: Optimal Stopping Rules. Berlin Heidelberg NewYork: Springer-Verlag 1978