

N. Meinshausen¹ and B.M. Hambly²**Abstract**

We discuss Monte Carlo methods for valuing options with multiple exercise features in discrete time. By extending the recently developed duality ideas for American option pricing we show how to obtain estimates on the prices of such options using Monte Carlo techniques. We prove convergence of our approach and estimate the error. The methods are applied to options in the energy and interest rate derivative markets.

1 Introduction

A difficult problem in derivatives pricing has been to price high dimensional American style options quickly and accurately. The methodology for pricing European contracts numerically, such as trees and finite difference methods for partial differential equations, leads to methods that are prohibitively time consuming when the dimension of the problem gets large. The alternative approach via Monte Carlo methods works more effectively in high dimensions but the technique is difficult to apply to options with early exercise features.

There is now an extensive literature on the high dimensional American option pricing problem. The first work in this direction was Tilley (1993) and this has been followed by a number of papers (Barraquand and Martineau 1995; Broadie and Glasserman 1997a; Broadie and Glasserman 1997b; Tsitsiklis and Van Roy 1999; Longstaff and Schwartz 2001), using techniques for approximation of the exercise boundary. As the price is the supremum over the return from all possible exercise strategies, these techniques determine a non-optimal exercise strategy and therefore naturally give a lower bound on the price. A recent development (Andersen and Broadie 2001; Haugh and Kogan 2001; Rogers 2002), has been to consider the dual problem and, using this, try to improve the Monte Carlo approach and produce a positive-biased estimate for the price.

¹Seminar für Statistik, Department of Mathematics, ETH Zürich, 8092 Zürich, Switzerland

²Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK

At first sight this seems impractical as the dual problem is infinite dimensional, involving minimising over a space of martingales. Fortunately it appears to be possible to choose martingales which are close to optimal to get accurate approximations, which enable this technique to be potentially useful.

We extend this theory to the multiple exercise case. We are able to obtain an expression for the value of the option as the infimum over a choice of stopping times and martingales. We then develop an algorithm using this which gives a positive-biased bound on the price. The key idea that we use is to express the martingale in terms of the value function and, as we have accurate approximations to the value function using recent techniques, such as those in Longstaff and Schwartz (2001), we can construct a reasonable approximation to the optimal martingale.

We will consider two examples where our results can be applied. We use simple models to illustrate our ideas but the extension to models with higher dimensionality or more complicated dynamics is straightforward. The first is a chooser flexible cap, a product from the interest rate market which gives the holder the right to exercise a certain number of caplets over the life of the contract. In this setting we may consider having any number of exercise opportunities up to the total number of exercise dates, which is 40 in our example. The positive-biased and negative-biased estimates are seen to be within 1-2 % of each other using around 1000 sample paths. A tight 99% confidence interval for the price can thus be formed.

Our second example is a very similar product from the energy derivatives market; a swing type option. We consider a stylized version of this product; for our purposes a swing option gives the holder a certain number of exercise opportunities at which the holder has the right to purchase power at a given price. We treat the problem as determining the expected value of this option under a suitable pricing measure. We assume there are 1000 possible exercise times and up to 100 exercise possibilities and obtain confidence intervals which show that there is still only a small gap between the upper and lower estimates. Naturally significant differences remain to be analyzed in any application to contracts actually traded in the energy markets.

The structure of the paper is as follows. We begin by setting up the discrete

time formulation of the problem as a Markov decision process. Our next step is to give the main results concerning the approximation of the value function and the methods for the implementation of our approach. Once we have this we can analyse the algorithm, proving convergence and obtaining an estimate of the error. We follow with the numerical results and conclude with the proofs of the results given in section 3.

2 Preliminaries

In this paper we will work with a discrete time Markov decision process which we now specify. Let E denote the space in which our underlying process lives, this will always be a subset of \mathbb{R}^d . The underlying process $\{X_t\}$ is a discrete time Markov chain, taking values in E , on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t=0}^T, \mathbb{P})$. The problems that we consider are always defined over finite lifetimes and hence time is always bounded by the finite maturity date T . We will write \mathbb{E} for expectation with respect to \mathbb{P} and $\mathbb{E}_t = \mathbb{E}(\cdot | \mathcal{F}_t)$ for the conditional expectation at time t .

As we are interested in multiple exercise problems we will need to keep track of the number of exercise opportunities available which we label by $n \in \mathbb{N}$. We will assume that it is possible to use an exercise opportunity at time 0, and thus up to maturity T , there are $T + 1$ possible exercise times. The natural state space for our Markov decision process is $S = E \times \mathbb{N}$. As time evolves, and we apply our actions, the decision process will move on S according to the dynamics inherited from the probability law for X .

We also have \mathcal{A} , a set of actions which can be taken. In this setting we take $\mathcal{A} = \{0, 1\}$, where we interpret 0 as exercise and 1 as continuation. Note that if $n = 0$, then the only possible action is to continue. The action space can alternatively be viewed as determining stopping times τ at which to use an exercise opportunity. We also have a payoff function, $Z : [0, \dots, T] \times S \rightarrow \mathbb{R}$, the reward earned by using an exercise opportunity when the underlying process is at a particular point in space and time. We will assume that $\mathbb{E}(\max_{t=0, \dots, T} |Z_t|) < \infty$ and that Z is independent of n .

Our aim is to maximize the expected payoff from the multiple exercise op-

portunities and we begin by defining the optimization problem we will solve.

We call a set of stopping times $\{\tau_n, \dots, \tau_1\}$ with $\tau_n < \dots < \tau_1$ a policy π . τ_m determines the time where the m -th remaining exercise opportunity is used under policy π . The expected payoff under policy π is,

$$V_t^{\pi,n}(x) = \mathbb{E}_t \left[\sum_{m=1}^n Z_{\tau_m} | X_t = x \right].$$

We note that $V_t^{\pi,n} : E \rightarrow \mathbb{R}$ is a function of $x \in E$, the state of the underlying Markov chain at time t .

The option pricing problem of finding the value function is then an optimization problem, to find the exercise policy π for which the expected payoff is maximized.

Definition 1 *The value function $V_t^{\star,n}$, with n remaining exercise opportunities is the expected payoff under an optimal policy,*

$$(1) \quad V_t^{\star,n}(x) = \sup_{\pi} \mathbb{E}_t \left[\sum_{m=1}^n Z_{\tau_m} | X_t = x \right].$$

If the supremum is attained, we denote the corresponding optimal policy by $\pi^ = \{\tau_n^*, \dots, \tau_1^*\}$.*

Definition 2 *The marginal value $\Delta V_t^{\pi,n}$ is defined for every policy π and for $n \geq 1$ as:*

$$(2) \quad \Delta V_t^{\pi,n}(x) = V_t^{\pi,n}(x) - V_t^{\pi,n-1}(x).$$

For $n = 1$ this amounts to $\Delta V_t^{\pi,1} = V_t^{\pi,1}$. Under an optimal policy π^ we denote the marginal value by $\Delta V_t^{\star,n}$.*

The marginal value thus specifies the additional payoff that can be expected from having one more exercise right (or alternatively, the loss that is incurred from having one less exercise right).

3 Methods

In general no unbiased estimate of the value of the multiple exercise option is available. We thus try to bound the value of the option from above and from

below as well as possible. Assume that two random variables $V_0^{\uparrow,n}$ and $V_0^{\downarrow,n}$ are given with positive and negative bias respectively, in that

$$(3) \quad \begin{aligned} \mathbb{E}[V_0^{\downarrow,n}] &\leq V_0^{*,n}, \\ \mathbb{E}[V_0^{\uparrow,n}] &\geq V_0^{*,n}. \end{aligned}$$

Let μ^\uparrow be the mean and σ^\uparrow be the standard deviation of $V_0^{\uparrow,n}$ over N independent samples. We define μ^\downarrow and σ^\downarrow similarly for $V_0^{\downarrow,n}$. An (asymptotically correct) $1 - \alpha$ confidence interval for the true value $V_0^{*,n}$ is then given by

$$(4) \quad \left[\mu^\downarrow - \beta \frac{\sigma^\downarrow}{\sqrt{N}}, \mu^\uparrow + \beta \frac{\sigma^\uparrow}{\sqrt{N}} \right],$$

where $\beta = \Phi^{-1}(1 - \alpha/2)$ is the $1 - \alpha/2$ -quantile of the standard normal distribution.

Note that the width of the confidence interval is determined by two contributions, the gap between the expected value of the positive- and negative-biased estimate on the one hand and the width of the distribution of the positive- or negative-biased estimate. The latter contribution becomes negligible as the number of samples N used to determine the bound increases. Hence we are interested in finding estimates with a small or vanishing bias.

A negative-biased estimate is readily obtained as the payoff along a random path under any given policy $\pi = \{\tau_n, \dots, \tau_1\}$. Indeed it follows from Definition 1 that the random variable

$$(5) \quad V_0^{\downarrow,n} = \sum_{m=1}^n Z_{\tau_m}$$

has a negative bias. The bias vanishes as soon as the optimal policy in the sense of Definition 1 is used. We will return later to the question of how to obtain an approximation to the optimal policy.

3.1 Positive-Biased Estimate

Our aim in the following is to find a positive-biased estimate of the value function using duality ideas. We tackle the problem sequentially and find a positive-biased estimate for the marginal value. We begin with a key theorem that gives a dual formulation for the marginal value. The proof can be found in Section 5.

Theorem 1 *The marginal value $\Delta V_0^{*,n}$ is equal to:*

$$\Delta V_0^{*,n} = \inf_{\pi} \inf_{M \in H_0} \mathbb{E}_0 \left[\max_{u \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_u - M_u) \right],$$

where $\mathcal{T} = \{0, \dots, T\}$ is the set of possible exercise dates, $0 \leq \tau_{n-1} < \dots < \tau_1$ are stopping times, $\{M_t\}$ is a martingale and H_0 is the set of all martingales that are zero at time $t = 0$. Moreover, the infimum is attained for the optimal policy π^* and the martingale M^* with $M_0^* = 0$, whose increment at time t is the martingale part of the marginal value function:

$$(6) \quad M_t^* - M_{t-1}^* = \Delta V_t^{*,m} - \mathbb{E}_{t-1}[\Delta V_t^{*,m}],$$

where m is the largest natural number such that $t \leq \tau_m$.

Remark 1 This is an extension of the single exercise case ($n=1$) where the dual problem, found in Haugh and Kogan (2001) and Rogers (2002), is given by

$$V_0^{*,1} = \inf_{M \in H_0} \mathbb{E}_t \left[\max_{0 \leq t < T} (Z_t - M_t) \right].$$

A positive-biased estimate of the value of an option with multiple exercise possibilities can easily be found with help of Theorem 1.

Definition 3 *We define for any policy π and martingale M with $M_t = 0$ the random variable $\Delta V_t^{\uparrow,n}$ by*

$$\Delta V_t^{\uparrow,n} = \max_{u \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_u - M_u),$$

where $\mathcal{T} = \{t, \dots, T\}$

According to Theorem 1, it holds that for any policy and martingale

$$\mathbb{E}[\Delta V_0^{\uparrow,n}] \geq \Delta V_0^{*,n}.$$

Summing up the marginal values, we obtain the random variable $V_0^{\uparrow,n}$,

$$(7) \quad V_0^{\uparrow,n} = \sum_{m=1}^n \Delta V_0^{\uparrow,m}.$$

This random variable is positive-biased in the sense of equation (3). The magnitude of the bias depends on the choice of the policy π and martingale M in (7) and vanishes for the optimal policy π^* and the martingale M^* as specified in Theorem 1.

3.2 Policy and Martingale Construction

The magnitude of both the negative bias of $V_0^{\downarrow,n}$ and the positive bias of $V_0^{\uparrow,n}$ depend on the quality of the optimal policy-approximation and, in the case of $V_0^{\uparrow,n}$, additionally on the quality of the optimal martingale-approximation.

We will show in the following how such approximations can be obtained from an approximation ΔV_t^n to the marginal value function $\Delta V_t^{*,n}$. The resulting policy and martingale are optimal if the approximation ΔV_t^n coincides with the true value function $\Delta V_t^{*,n}$. Hence the bias of both the negative-biased and positive-biased estimate of the value function vanish for an optimal approximation to the value function.

An approximation to the value function can be obtained in several ways. Methods based on value-function regression (Longstaff and Schwartz 2001; Tsitsiklis and Van Roy 2001) seem to be suitable but other methods, such as dynamic programming, are possible in a problem of reduced dimensionality. We are going to employ value-function regression in the numerical examples, in particular the algorithm proposed by Longstaff and Schwartz (2001). Convergence of the algorithm in Longstaff and Schwartz (2001) has recently been proven in Clement, Lamberton, and Protter (2002).

We assume an approximation to the value function to be available for the following.

Optimal Policy Approximation. Assuming that the true value function is known, the optimal policy is given by exercising the option if and only if the payoff from immediate exercise is larger than the expected marginal value under continuation.

An approximation to the optimal policy can be obtained by replacing the marginal value $\Delta V_t^{*,n}$ by an approximation ΔV_t^n .

Definition 4 *The policy $\pi = \{\tau_n, \dots, \tau_1\}$ is defined by the stopping time*

$$\tau_n = \min\{t : Z_t > \mathbb{E}_t[\Delta V_{t+1}^n]\},$$

and, for $m = 1, \dots, n-1$, by

$$\tau_m = \min\{t > \tau_{m+1} : Z_t > \mathbb{E}_t[\Delta V_{t+1}^m]\},$$

In the case that the approximation to the value function is the true value function, we clearly have that $\pi = \pi^*$ and the optimal policy is obtained.

Optimal Martingale Approximation. An approximation to the martingale of equation (6) can also be obtained from an approximation to the value function. Theorem 1 states that the infimum is attained for the martingale M^* , with $M_0^* = 0$ and increments

$$M_t^* - M_{t-1}^* = \Delta V_t^{*,m} - \mathbb{E}_{t-1}[\Delta V_t^{*,m}],$$

where m is the largest number such that $t \leq \tau_m$.

We consider as an approximation to M^* the martingale M with increments given by replacing the marginal value function by its approximation and replacing the expectation by a Monte Carlo approximation.

Definition 5 *The martingale M is defined by $M_0 = 0$ and for $1 \leq t < T$ by the increments*

$$M_t - M_{t-1} = \Delta V_t^m - E_{t-1}^m,$$

where m is the largest natural number such that $t \leq \tau_m$ and E_{t-1}^m is a Monte Carlo approximation to $\mathbb{E}_{t-1}[\Delta V_t^m]$. In particular, let $X_t^i, i = 1, \dots, k$ be independent samples from the distribution of X_t conditional on X_{t-1} . Then

$$(8) \quad E_{t-1}^m = \frac{1}{k} \sum_{i=1}^k \Delta V_t^m(X_t^i).$$

If the approximation to the value function is identical to the true value function, we obtain the optimal martingale M^* in the limit as $k \rightarrow \infty$, where k is the number of samples in the Monte Carlo approximation (8).

3.3 Computational Complexity

Once an approximation to the value function is found, the effort required for computing a single realization of the negative-biased $V_0^{\downarrow, n}$ consists essentially of computing a realisation of the Markov chain X_t and the corresponding payoff Z_t . Thus the computational cost for the negative-biased estimate increases linearly with the number of timesteps T .

The same holds true for the positive-biased estimate, as the computation of the positive-biased estimate only incurs the additional cost of computing the martingale M , and again this increases only linearly with T .

The cost for computing a single realization of the martingale M is additionally roughly linear in k , where k is the number of independent samples used in the Monte Carlo approximation (8). This seemingly makes accurate approximations to the optimal martingale prohibitive. Note, however, that the more accurate the approximation to the optimal martingale the smaller the variance of $V_0^{\uparrow,n}$. In fact, by Lemma 7 and 8, the variance (and the bias) of $V_0^{\uparrow,n}$ vanishes in the limit of an optimal approximation. Thus fewer samples of $V_0^{\uparrow,n}$ are needed in the limit of large k , offsetting the computational burden incurred by the Monte Carlo approximation (8) in the construction of the martingale.

3.4 Error Analysis

We examine the magnitude of the positive bias of $\Delta V_0^{\uparrow,n}$ incurred by using these approximations instead of the optimal policy π^* and martingale M^* . Using Theorem 1 we can define the bias in the marginal value under approximations to the optimal policy and martingale.

Definition 6 For any policy $\pi = \{\tau_{n-1}, \dots, \tau_1\}$ and martingale M , the bias at time t with n exercise opportunities remaining $B_t^n(\pi, M)$ of the estimated marginal value is denoted by

$$B_t^n(\pi, M) = \mathbb{E}_t[\Delta V_t^{\uparrow,n}] - \Delta V_t^{*,n}.$$

In order to estimate the bias we define two distance functions.

Definition 7 We define the distance D_V between ΔV^* and the given approximation ΔV as

$$D_V = \sup_{\substack{m=1, \dots, n, \\ x \in E, 0 \leq t < T}} |\Delta V_t^{*,m}(x) - \Delta V_t^m(x)|$$

The distance D_π^n between the policy $\pi = \{\tau_n, \dots, \tau_1\}$ and the optimal policy $\pi^* = \{\tau_n^*, \dots, \tau_1^*\}$ is

$$D_\pi^n = V_0^{*,n} - V_0^{\pi,n}$$

We denote by σ_k a bound on the standard deviation of the Monte Carlo estimate in the approximation (8) of the martingale M :

$$\sigma_k^2 = \sup_{\substack{m=1, \dots, n, \\ x \in E, 1 \leq t < T}} \mathbb{E}_{t-1} [(E_{t-1}^m - \mathbb{E}_{t-1}[\Delta V_t^m])^2]$$

Theorem 2 *The bias of the marginal value estimate is bounded above, in that*

$$B_0^n(\pi, M) \leq D_\pi^{n-1} + 2\sqrt{(4D_V^2 + \sigma_k^2) T}$$

The proof of this result can be found in Section 6. We note that the bound will only be non-trivial when D_V and σ_k exist. These exist in the case of bounded payoff, but also under weaker moment conditions. For instance σ_k will exist provided that the payoff process is L^2 -bounded. They will exist for call and put payoffs under the standard Black-Scholes model and for the options and models we discuss in this paper.

For the single-exercise case, $D_\pi^{n-1} = 0$ and we are left with the second term only. This is consistent with the fact that no stopping time has to be chosen for the evaluation of $\Delta V_0^{\uparrow,1}$ and hence the bias of the positive-biased estimate is independent of the quality of the optimal-policy approximation.

If the quality of the value function approximation is not affected by the number T of timesteps, it is clear by Theorem 2 that the bias of the estimated marginal values and hence of $V_0^{\uparrow,n}$ scales maximally with the square root of the number T of timesteps.

4 Numerical Results

We now apply the proposed method to two products in the financial markets. First we give a complete description of the algorithm employed.

The valuation involves three main steps.

Step 1: Fitting the Value Function. We employ value-function regression, in particular the method by Longstaff and Schwartz (2001) to obtain an approximation to the marginal value function.

A fixed number of independent paths (here 1000) of the underlying Markov Chain $X_t, t = 0, \dots, T$ are computed and stored with common starting value $X_0 = x_0$, where x_0 is the point at which the value of the option is to be evaluated.

Instead of directly fitting the marginal value function, we compute an approximation to the continuation value Q_t^m , see Definition 8. The continuation value is smoother than the value function. Furthermore the continuation value is of direct use in the optimal policy approximation, see Definition 4. An approximation to the value function is then readily obtained by equation (11) and an approximation to the marginal values by equation (2).

The algorithm works backwards in time, starting at terminal time $t = T$, where the continuation value vanishes identically. At time $t < T$ an approximation to the continuation value is already available for times larger than t . Hence, using Definition 4, an approximation to the optimal policy is available. For each value for the number of exercise opportunities $m = 1, \dots, n$, we do the following:

- (1) The payoff under the approximated optimal policy (and the additional constraint $\tau_m > t$ as we are interested in the continuation value) is calculated for each path.
- (2) A linear regression model is then fitted for the continuation value. Thus for each timestep we have n different linear regression models. The set of basis function $\{\psi_1, \dots, \psi_l\}$ for the regression model is the same for all values of t and m but chosen differently for different option contracts.

Step 2: Computing the negative-biased estimate. A new set of independent paths (here 1000) with $X_0 = x_0$ is computed. For each path the accumulated payoff $V_0^{\downarrow, n}$, equation (5), is computed under policy π of Definition 4, using the approximation to the continuation value obtained from Step 1. The average μ^\downarrow and the standard deviation σ^\downarrow for these $N = 1000$ realizations of $V_0^{\downarrow, n}$ are calculated.

Step 3: Computing the positive-biased estimate. Another new set of independent paths (here 20) with $X_0 = x_0$ is computed. For each path the following is done. For every initial number $m = 1, \dots, n$ of exercise opportunities, the martingale M_t is computed according to Definition 5 with a constant $k = 50$ in the Monte Carlo approximation (8). The value of $\Delta V_0^{\uparrow, m}$ is then computed for every value of $m = 1, \dots, n$ according to equation (7) and $V_0^{\uparrow, n}$ is obtained by summing up the marginal values, see equation (7). Hence we obtain in to-

tal $N = 20$ independent realizations of the random variable $V_0^{\uparrow,n}$. Again the average μ^\uparrow and standard deviation σ^\uparrow over these realizations are computed.

The confidence interval is then formed according to equation (4).

Note that the number of paths ($N = 20$) used to evaluate the positive-biased estimate is small compared to the number of paths ($N = 1000$) used to evaluate the negative-biased estimate. On the one hand this ensures that the effort for the computation of μ^\uparrow and μ^\downarrow is comparable, as the martingale approximation (with $k = 50$) in the case of the positive-biased estimate makes the computation of a single realization of $V_0^{\uparrow,n}$ about k times as costly as the computation of a single realization of $V_0^{\downarrow,n}$. On the other hand, the variance of $V_0^{\uparrow,n}$ is magnitudes smaller than the variance of $V_0^{\downarrow,n}$. In fact, the variance of $V_0^{\uparrow,n}$ vanishes if the approximation to the policy π^* and martingale M^* is optimal (as shown in Lemma 8). Hence many fewer paths are needed in the case of the positive-biased estimate to achieve the same degree of accuracy as for the negative-biased estimate.

4.1 Chooser Flexible Cap

The chooser flexible cap is a product in the interest rate market which enables the holder to protect themselves against adverse movements in the interest rate. An interest rate cap is a sequence of caplets at, for instance, quarterly intervals over the lifetime of the option. For our purposes the i -th caplet provides the holder at time T_i with a payment of the notional multiplied by the difference in the current interest rate R and the fixed strike K , if it is positive. A flexible cap (sometimes called an autocap or a limit cap) is similar to the cap with the additional feature that at most n caplets will be exercised over the lifetime of the option, where $n < T$ with T being the total number of possible caplets in the lifetime of the option. These caplets are exercised automatically if the interest rate is above K at the payment date. The chooser flexible cap allows greater flexibility in that it provides the holder with the decision at each payment date of whether to exercise the caplet or to spare the caplet for use at a later time when the interest rate might be substantially above K and hence the associated payoff larger. The caplets will expire worthless, however, if not used before the end of the option contract period.

The chooser flexible cap is thus a product suited to the pricing approach we have developed here and we illustrate our methods with a particular example. We take the lifetime of the option to be $T = 40$, and here the total number of possible exercise dates also to be $T = 40$. This corresponds to an option with lifetime 10 years and quarterly exercise rights.

To price the option, a particular model for the interest rate R_t has to be specified. We will assume that we are directly modelling the market in the martingale measure as would be the case if the model were calibrated to market data. By way of a simple two-dimensional example, we use a two-factor additive Gaussian model from Brigo and Mercurio (2001), a variant of the Longstaff-Schwartz model (Longstaff and Schwartz 1992). The dynamics of the interest rate are given by $R_t = \phi_t + S_t + U_t$ with ϕ_t a deterministic time varying rate and $S_0 = U_0 = 0$ with

$$\begin{aligned} dS_t &= -aS_t dt + \sigma dW_t^s \\ dU_t &= -bU_t dt + \eta dW_t^u, \end{aligned}$$

where W^s and W^u are Brownian motions with correlation $dW_t^s dW_t^u = \rho dt$. The payoff of the chooser flexible cap under an exercise decision is $\max\{R_t - K, 0\}$, where R_t is the interest rate at the beginning of the t -th quarter. The constants are set to $a = 5, b = 2, \sigma = 0.05, \eta = 0.02, \rho = 0.2$ and $\phi_t = 0.05$ for all time. The strike is given by $K = 0.05$.

As the model is Gaussian we can compute the distribution of the rate at each quarter and thus set up our discrete pricing framework with 40 time steps. (If the model is non-Gaussian, we could just discretise the continuous model). We compute in the following the value of the chooser flexible cap at $t = 0$.

We note that for the chosen constants and for $n = 40$ the product is a standard cap and the price can be computed explicitly for this Gaussian model. This is done by using a change of numeraire technique with the bond prices of Brigo and Mercurio (2001) Theorem 4.2.1 and the interest rate dynamics in the T -forward measure given in Brigo and Mercurio (2001) Lemma 4.2.2. In our case the resulting pricing formula gives for $n = T = 40$ the value at $t = 0$ of 0.2492, or 2492 basis points, which lies between the positive- and negative-biased estimates obtained through our algorithm.

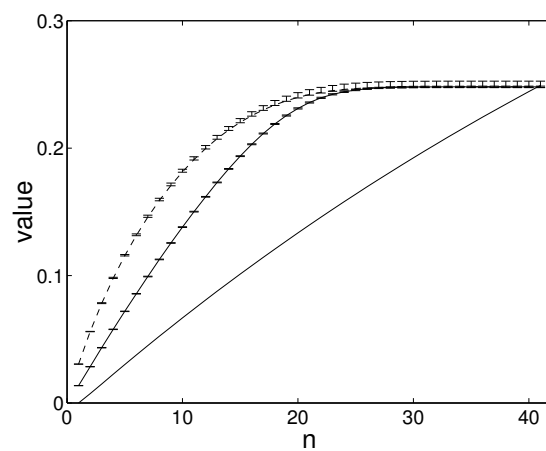


Figure 1: The negative-biased estimate for the value of a caplet as a function of the number n of caplets for the chooser flexible cap (topmost line). The 99% confidence interval is plotted for each value of n as a vertical bar. Also shown are the mean and confidence intervals for the value of an option with automatic exercise on the first n exercise opportunities with positive payoff (flexible cap), while the lowest line shows the exact value of a cap with n caplets on the first n exercise dates.

The basis functions for the linear regression model are chosen as:

$$\begin{aligned}\Psi_1(u_t, s_t) &= 1 \\ \Psi_2(u_t, s_t) &= u_t \\ \Psi_3(u_t, s_t) &= s_t \\ \Psi_4(u_t, s_t) &= (u_t + s_t)^2.\end{aligned}$$

The numerical valuation then follows the three steps described above.

exercise possibilities	negative-biased estimate	positive-biased estimate	relative difference	99% confidence interval
1	0.03047	0.03049	0.0007	[0.03036 , 0.03055]
2	0.05598	0.05612	0.0025	[0.05581 , 0.05621]
3	0.07811	0.07858	0.0059	[0.07789 , 0.07869]
4	0.09783	0.09854	0.0073	[0.09756 , 0.09867]
5	0.11556	0.11641	0.0074	[0.11525 , 0.11655]
6	0.13146	0.13257	0.0084	[0.13112 , 0.13273]
7	0.14591	0.14725	0.0092	[0.14552 , 0.14741]
8	0.15896	0.16053	0.0099	[0.15854 , 0.16070]
9	0.17072	0.17248	0.010	[0.17027 , 0.17265]
10	0.18140	0.18332	0.010	[0.18092 , 0.18351]
15	0.22029	0.22301	0.012	[0.21966 , 0.22321]
20	0.23985	0.24341	0.014	[0.23910 , 0.24363]
25	0.24678	0.25086	0.016	[0.24597 , 0.25109]
30	0.24801	0.25230	0.017	[0.24717 , 0.25253]
35	0.24807	0.25243	0.017	[0.24724 , 0.25265]
40	0.24807	0.25244	0.017	[0.24724 , 0.25266]

The table above shows the numerical results for n , the number of caplets that can be exercised, ranging from $n = 1$ to the maximal possible number $n = 40$. From left to right we give the negative and positive-biased estimates of the value at time $t = 0$, their relative difference and the 99% confidence interval for the true price of the option.

The values we obtain for the chooser flexible cap are compared in Figure 1

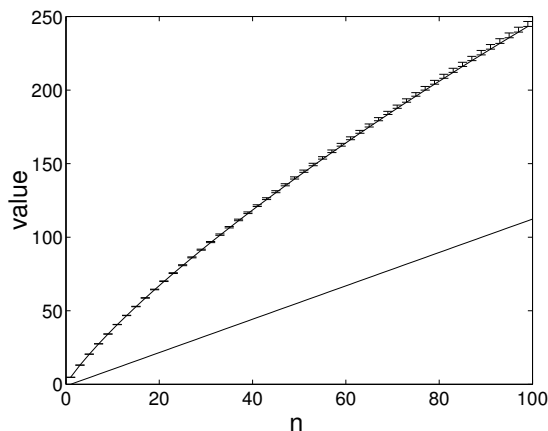


Figure 2: The confidence interval and negative-biased point estimate for the value of the Swing option as a function of the number n of total exercise rights. The lower line shows the exact value of an option with automatic exercise on the first n exercise possibilities.

to the corresponding values of a flexible cap and a cap. It can be seen that the difference between the negative- and positive-biased estimate of the value of the chooser flexible cap is below 2% for all n despite the fact that the approximation to the value function was obtained with a very simple regression architecture.

4.2 Swing option

The second example is a Swing option, a product in the energy market. There are several variants of swing options. The one we focus on here is a stylized version of swing options actually traded in the energy market. Its holder is entitled to buy, over a specified period of T days, on each day a certain quantity of energy/electricity, for a fixed price K . There is a constraint on the maximal amount of energy that can be purchased over the lifetime of the Swing option, in that energy can only be purchased on n days. Here we consider a period of $T = 1000$ days and up to $n = 100$ exercise opportunities. We note that in the electricity market there are no natural hedging instruments and we will therefore compute the value of the swing contract as the expectation under a given pricing measure.

There are a number of models that have been proposed for energy prices and

any model could be used here. For simplicity we take a type of AR(1) model for the logarithm of the energy price in which the price S_t on day t is taken to be of the form:

$$(9) \quad \log S_{t+1} = (1 - k)(\log S_{t-1} - \mu) + \mu + \sigma W_t,$$

where, W_t is a normally distributed random variable with unit variance. The constants in the model are set to $\sigma = 0.5$, $k = 0.9$ and $\mu = 0$, $K = 0$ and $S_1 = 1$.

The basis functions for the linear regression model are chosen as:

$$\begin{aligned} \Psi_1(s_t) &= 1 \\ \Psi_2(s_t) &= s_t. \end{aligned}$$

The negative- and positive-biased estimates are otherwise obtained in the same way as for the chooser flexible cap.

exercise possibilities	negative-biased estimate	positive-biased estimate	relative difference	99% confidence interval
1	4.777	4.790	0.002	[4.773 , 4.794]
2	9.024	9.085	0.006	[9.016 , 9.091]
3	12.970	13.094	0.009	[12.959 , 13.100]
4	16.786	16.899	0.006	[16.773 , 16.906]
5	20.455	20.573	0.005	[20.439 , 20.580]
10	37.334	37.531	0.005	[37.305 , 37.540]
15	52.713	52.999	0.005	[52.670 , 53.009]
20	67.105	67.513	0.006	[67.050 , 67.525]
30	93.742	94.507	0.008	[93.662 , 94.519]
40	118.457	119.611	0.009	[118.353 , 119.625]
50	141.832	143.345	0.010	[141.703 , 143.360]
60	164.112	166.020	0.011	[163.960 , 166.037]
70	185.511	187.711	0.011	[185.335 , 187.729]
80	206.045	208.682	0.012	[205.844 , 208.702]
90	225.900	228.965	0.013	[225.676 , 228.985]
100	245.157	248.630	0.013	[244.910 , 248.651]

The numerical results are shown in the table above and are displayed in Figure 2. It can be seen that, despite the large number of timesteps, the relative

difference between the negative- and positive-biased estimate remains below 1.5%. This illustrates that confidence intervals remain tight even when the number of possible exercise dates is increased substantially, in agreement with our theoretical finding in Theorem 2.

We remark that the models we have used in our examples are very simple and the problems could be solved by other means. However the advantage of our approach is that it allows prices to be obtained when the models are extended to higher dimensions.

5 Proof of Theorem 1

We establish the duality result through a sequence of lemmas.

We introduce first the so-called continuation, or Q-value

Definition 8 *The continuation value $Q_t^{*,n}$ is the expectation of the value function one timestep later,*

$$(10) \quad Q_t^{*,n}(x) = \begin{cases} \mathbb{E}[V_{t+1}^{*,n}|X_t = x] & t < T, \\ 0 & t = T. \end{cases}$$

The marginal continuation value $\Delta Q_t^{,n}$ is defined for $n \geq 1$ as:*

$$\Delta Q_t^{*,n} = Q_t^{*,n} - Q_t^{*,n-1}.$$

For $n = 1$ this amounts to $\Delta Q_t^{,1} = Q_t^{*,1}$.*

With this definition and (1) we can write the value function $V_t^{*,n}$ with n remaining exercise opportunities as

$$(11) \quad \begin{aligned} V_t^{*,n}(x) &= \sup_{t \leq \tau_n < T} \mathbb{E}[Z_{\tau_n} + Q_{\tau_n}^{*,n-1}|X_t = x], \\ \text{or as } V_t^{*,n}(x) &= \max \{Z_t(x) + Q_t^{*,n-1}(x), Q_t^{*,n}(x)\}. \end{aligned}$$

The second expression is easily seen as $Z_t + Q_t^{*,n-1}$ is the payoff Z_t under exercise of the n -th exercise opportunity at time t plus the expected future payoff with the remaining $n-1$ exercise possibilities. The quantity $Q_t^{*,n}$ on the other hand is the expected total payoff at time t under continuation without exercise at time t . The optimal policy π^* takes, by definition, the value-maximizing decision, hence the value at time t is the maximum of the two quantities.

The value process $V_t^{*,n}$ is the Snell envelope of the payoff function $Z_t + Q_t^{*,n-1}$ and is thus a supermartingale. The value process can then be written, using the Doob decomposition as

$$V_t^{*,n} = V_0^{*,n} + M_t^{*,n} - A_t^{*,n},$$

where $M_t^{*,n}$ is a martingale and $A_t^{*,n}$ a previsible increasing process, both zero at $t = 0$. For $n \geq 1$ we write $\Delta M_t^{*,n} = M_t^{*,n} - M_t^{*,n-1}$ and likewise $\Delta A_t^{*,n} = A_t^{*,n} - A_t^{*,n-1}$.

It is useful to restate the marginal value as an optimal stopping problem that involves the Doob decomposition of the value process $V_t^{*,n-1}$.

Proposition 1 *The marginal value $\Delta V_t^{*,n}$ is equal to*

$$\Delta V_t^{*,n} = \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{*,n-1}] + A_t^{*,n-1},$$

where $A_t^{*,n-1}$ is the previsible increasing process in the Doob decomposition of $V_t^{*,n-1}$.

We note that here and throughout the paper we can apply the optional stopping theorem as the stopping times we use are always bounded.

Proof of Proposition 1: Using (10), the martingale property of $M_t^{*,n}$ and the optional stopping theorem

$$\begin{aligned} \Delta V_t^{*,n} &= V_t^{*,n} - V_t^{*,n-1} \\ &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau + \mathbb{E}_\tau[V_{\tau+1}^{*,n-1}]] - V_t^{*,n-1} \\ &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau + V_\tau^{*,n-1} + (\mathbb{E}_\tau[V_{\tau+1}^{*,n-1}] - V_\tau^{*,n-1})] - V_t^{*,n-1}. \end{aligned}$$

As $\mathbb{E}_\tau[V_{\tau+1}^{*,n-1}] - V_\tau^{*,n-1}$ is just the previsible part of the value process,

$$\begin{aligned} \Delta V_t^{*,n} &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau + V_\tau^{*,n-1} - (A_{\tau+1}^{*,n-1} - A_\tau^{*,n-1})] - V_t^{*,n-1} \\ &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau + V_t^{*,n-1} + (M_\tau^{*,n-1} - M_t^{*,n-1}) - \\ &\quad (A_\tau^{*,n-1} - A_t^{*,n-1}) - (A_{\tau+1}^{*,n-1} - A_\tau^{*,n-1})] - V_t^{*,n-1} \\ &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau + V_t^{*,n-1} + (M_\tau^{*,n-1} - M_t^{*,n-1}) - \\ &\quad (A_{\tau+1}^{*,n-1} - A_t^{*,n-1})] - V_t^{*,n-1}. \end{aligned}$$

Using the optional stopping theorem, $\mathbb{E}_t[M_\tau^{*,n-1} - M_t^{*,n-1}] = 0$ and hence

$$\begin{aligned}\Delta V_t^{*,n} &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau + V_t^{*,n-1} - (A_{\tau+1}^{*,n-1} - A_t^{*,n-1})] - V_t^{*,n-1} \\ &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{*,n-1}] + A_t^{*,n-1},\end{aligned}$$

which completes the proof. \square

The $(n+1)$ -th marginal value thus depends on the previsible part $A_t^{*,n}$ of the value function with n exercise possibilities remaining. We can capture the fact that the process $A_t^{*,n}$ only increases when continuation is the sub-optimal decision, in the following Lemma.

Lemma 1 *The increments of the process $A_t^{*,n}$ can be expressed as*

$$A_{t+1}^{*,n} - A_t^{*,n} = [Z_t - \Delta Q_t^{*,n}]_+,$$

with the convention that $[\cdot]_+ := \max\{\cdot, 0\}$.

Proof. Using (11), the Doob decomposition allows us to write

$$\begin{aligned}A_{t+1}^{*,n} - A_t^{*,n} &= V_t^{*,n} - Q_t^{*,n} \\ &= \max\{Z_t + Q_t^{*,n-1}, Q_t^{*,n}\} - Q_t^{*,n} \\ &= [Z_t - Q_t^{*,n} + Q_t^{*,n-1}]_+ \\ &= [Z_t - \Delta Q_t^{*,n}]_+.\end{aligned}$$

\square

Using Proposition 1, we now show the intuitive fact that marginal values decrease as the number of remaining exercise opportunities increases.

Proposition 2 *The marginal value is a decreasing function of the number of exercise opportunities remaining, in that for $n \geq 2$,*

$$\Delta V_t^{*,n} \leq \Delta V_t^{*,n-1},$$

Proof. The proof is by induction. We begin by proving the claim for $n = 2$.

Using Proposition 1, the marginal value can be written as

$$\Delta V_t^{*,2} = \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{*,1}] + A_t^{*,1}.$$

As $A_t^{*,1}$ is an increasing process, $A_{\tau+1}^{*,1} - A_t^{*,1} \geq 0$, $\forall \tau \geq t$ and therefore

$$\begin{aligned} \Delta V_t^{*,2} &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{*,1}] + A_t^{*,1} \\ &\leq \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_t^{*,1}] + A_t^{*,1} \\ &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau] = V_t^{*,1} = \Delta V_t^{*,1}, \end{aligned}$$

as required.

For $n > 2$ we can assume by the inductive hypothesis that $\Delta V_{t+1}^{*,n-1} \leq \Delta V_{t+1}^{*,n-2}$. By Definition 8 this implies

$$\Delta Q_t^{*,n-1} \leq \Delta Q_t^{*,n-2}.$$

Hence, by Lemma 1,

$$\begin{aligned} A_{t+1}^{*,n-1} - A_t^{*,n-1} &= [Z_t - \Delta Q_t^{*,n-1}]_+ \\ &\geq [Z_t - \Delta Q_t^{*,n-2}]_+ \\ (12) \qquad \qquad \qquad &= A_{t+1}^{*,n-2} - A_t^{*,n-2}. \end{aligned}$$

Using Proposition 1,

$$\begin{aligned} \Delta V_t^{*,n} &= \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{*,n-1}] + A_t^{*,n-1} \\ &\leq \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{*,n-2}] + A_t^{*,n-2} \\ &= \Delta V_t^{*,n-1}, \end{aligned}$$

completing the proof. □

Lemma 2 *The marginal value process is a supermartingale. That is for all $n \geq 1$*

$$\Delta V_t^{*,n} \geq \Delta Q_t^{*,n}.$$

Proof. This is a simple application of the Doob decomposition. As the difference of two martingales is still a martingale we have that

$$\Delta M_t^{*,n} = \Delta V_t^{*,n} - \Delta V_0^{*,n} + \Delta A_t^{*,n},$$

is a martingale. Taking conditional expectations of $\Delta M_{t+1}^{*,n}$ and equating we have

$$\Delta V_t^{*,n} - \Delta Q_t^{*,n} = \Delta A_{t+1}^{*,n} - \Delta A_t^{*,n}.$$

Using the same argument as in (12), we have that $\Delta Q_{t+1}^{*,n} \geq \Delta Q_t^{*,n}$, completing the proof. \square

Lemma 3 *It holds that*

$$\Delta V_t^{*,n} \leq \Delta Q_t^{*,n-1}.$$

Proof. This uses Proposition 1, 2, and Lemma 1,

$$\begin{aligned}
\Delta V_t^{*,n} &= V_t^{*,n} - V_t^{*,n-1} \\
&= \sup_{t \leq \tau \leq T} \mathbb{E}_t [Z_\tau - A_{\tau+1}^{*,n-1} + A_t^{*,n-1}] \\
&= \max \{ Z_t - (A_{t+1}^{*,n-1} - A_t^{*,n-1}), \\
&\quad \mathbb{E}_t [\sup_{t+1 \leq \tau \leq T} \mathbb{E}_{t+1} [Z_\tau - A_{\tau+1}^{*,n-1} + A_t^{*,n-1}]] \} \\
&= \max \{ Z_t - (A_{t+1}^{*,n-1} - A_t^{*,n-1}), \Delta Q_t^{*,n} - (A_{t+1}^{*,n-1} - A_t^{*,n-1}) \} \\
&= \max \{ Z_t - [Z_t - \Delta Q_t^{*,n-1}]_+, \Delta Q_t^{*,n} - (A_{t+1}^{*,n-1} - A_t^{*,n-1}) \} \\
&\leq \max \{ \Delta Q_t^{*,n-1}, \Delta Q_t^{*,n} \} \\
(13) \quad &= \Delta Q_t^{*,n-1},
\end{aligned}$$

which completes the proof. \square

The following proposition is the basis for the computation of an upper bound for the marginal values $\Delta V_t^{*,n}$. We write $\mathbf{1}_A$ for the indicator of the event A .

Proposition 3 *The marginal value can be expressed as*

$$\Delta V_0^{*,n} = \inf_{0 \leq \tau \leq T} \inf_{M \in H_0} \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau} (Z_t \mathbf{1}_{\{t < \tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau\}} - M_t) \right],$$

where τ is a stopping time and M_t a martingale with $M_0 = 0$. Moreover, the infimum is attained for the martingale $\Delta M_t^{*,n} = M_t^{*,n} - M_t^{*,n-1}$ and stopping time $\tau = \min\{t : A_{t+1}^{*,n-1} > 0\}$.

We prove the proposition in two parts.

Lemma 4 *It holds for all stopping times τ bounded by T and all martingales $M_t \in H_0$ that*

$$\Delta V_0^{*,n} \leq \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau} (Z_t \mathbf{1}_{\{t < \tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau\}} - M_t) \right].$$

Proof. Let ϑ, ϑ' and τ denote stopping times. By using Proposition 1, splitting at an intermediate stopping time and the fact that $A_t^{*,n-1}$ is increasing and zero at time $t = 0$,

$$\begin{aligned} \Delta V_0^{*,n} &= \sup_{0 \leq \tau \leq T} \mathbb{E}_0[Z_\tau - A_{\tau+1}^{*,n-1}] \\ &= \sup_{0 \leq \vartheta \leq \tau} \mathbb{E}_0[(Z_\vartheta - A_{\vartheta+1}^{*,n-1}) \mathbf{1}_{\{\vartheta < \tau\}} + \sup_{\tau \leq \vartheta' \leq T} \mathbb{E}_\tau[Z_{\vartheta'} - A_{\vartheta'+1}^{*,n-1}] \mathbf{1}_{\{\vartheta = \tau\}}] \\ &\leq \sup_{0 \leq \vartheta \leq \tau} \mathbb{E}_0[Z_\vartheta \mathbf{1}_{\{\vartheta < \tau\}} + (\sup_{\tau \leq \vartheta' \leq T} \mathbb{E}_\tau[Z_{\vartheta'} - A_{\vartheta'+1}^{*,n-1}] + A_\tau^{*,n-1}) \mathbf{1}_{\{\vartheta = \tau\}}] \\ &= \sup_{0 \leq \vartheta \leq \tau} \mathbb{E}_0[Z_\vartheta \mathbf{1}_{\{\vartheta < \tau\}} + \Delta V_\vartheta^{*,n} \mathbf{1}_{\{\vartheta = \tau\}}]. \end{aligned}$$

By Lemma 3, $\Delta V_\vartheta^{*,n} \leq \mathbb{E}_\vartheta[\Delta V_{\vartheta+1}^{*,n-1}]$. Introducing a martingale M , which is zero at $t = 0$,

$$\begin{aligned} \Delta V_0^{*,n} &\leq \sup_{0 \leq \vartheta \leq \tau} \mathbb{E}_0[Z_\vartheta \mathbf{1}_{\{\vartheta < \tau\}} + \mathbb{E}_\vartheta[\Delta V_{\vartheta+1}^{*,n-1}] \mathbf{1}_{\{\vartheta = \tau\}}] \\ &= \sup_{0 \leq \vartheta \leq \tau} \mathbb{E}_0[Z_\vartheta \mathbf{1}_{\{\vartheta < \tau\}} + \mathbb{E}_\vartheta[\Delta V_{\vartheta+1}^{*,n-1}] \mathbf{1}_{\{\vartheta = \tau\}} - M_\vartheta] \\ &\leq \mathbb{E}_0 \max_{0 \leq t \leq \tau} (Z_t \mathbf{1}_{\{t < \tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau\}} - M_t). \end{aligned}$$

□

To complete the proof of the first part of Proposition 3 it has to be shown that the following inequality holds as well.

Lemma 5 *The marginal value can be bounded from below by*

$$\Delta V_0^{*,n} \geq \inf_{0 \leq \tau \leq T} \inf_{M \in H_0^1} \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau} (Z_t \mathbf{1}_{\{t < \tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau\}} - M_t) \right].$$

Moreover, the infimum is attained for $\tau^* = \min\{t : A_{t+1}^{*,n-1} > 0\}$ and $\Delta M_t^{*,n}$.

Proof. For any stopping time τ and martingale M ,

$$(14) \quad \begin{aligned} & \inf_{0 \leq \tau \leq T} \inf_{M \in H_0^1} \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau} (Z_t \mathbf{1}_{\{t < \tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau\}} - M_t^n) \right] \\ & \leq \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau^*} (Z_t \mathbf{1}_{\{t < \tau^*\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau^*\}} - \Delta M_t^{*,n}) \right]. \end{aligned}$$

We define now a stopping time τ^* by $\tau^* = \min\{t : A_{t+1}^{*,n-1} > 0\}$. It will be shown later in Lemma 6 that this stopping time is in fact equivalent to the stopping time τ_{n-1}^* under an optimal policy. It clearly holds for τ^* that $A_{\tau^*}^{*,n-1} = 0$. It follows then from Lemma 1 that

$$\begin{aligned} \mathbb{E}_{\tau^*}[\Delta V_{\tau^*+1}^{*,n-1}] &= Z_{\tau^*} - A_{\tau^*+1}^{*,n-1} + A_{\tau^*}^{*,n-1} \\ &= Z_{\tau^*} - A_{\tau^*+1}^{*,n-1}. \end{aligned}$$

Using this in (14),

$$(15) \quad \begin{aligned} & \inf_{0 \leq \tau \leq T} \inf_{M \in H_0^1} \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau} (Z_t \mathbf{1}_{\{t < \tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau\}} - M_t^n) \right] \\ & \leq \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau^*} (Z_t - A_{t+1}^{*,n-1} - \Delta M_t^{*,n}) \right]. \end{aligned}$$

As $A_t^{*,n-1}$ is positive, by Proposition 1, $\Delta V_t^{*,n} \geq \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{*,n-1}]$.

Thus, by the Doob decomposition,

$$\begin{aligned} Z_t - A_{t+1}^{*,n-1} &\leq \Delta V_t^{*,n} \\ &= \Delta V_0^{*,n} + \Delta M_t^{*,n} - \Delta A_t^{*,n}. \end{aligned}$$

Hence

$$Z_t - A_{t+1}^{*,n-1} - \Delta M_t^{*,n} \leq \Delta V_0^{*,n} - \Delta A_t^{*,n},$$

and, putting this into (15),

$$\begin{aligned} & \inf_{0 \leq \tau \leq T} \inf_{M \in H_0^1} \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau} (Z_t \mathbf{1}_{\{t < \tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{*,n-1}] \mathbf{1}_{\{t=\tau\}} - M_t^n) \right] \\ & \leq \mathbb{E}_0 \left[\max_{0 \leq t \leq \tau^*} (\Delta V_0^{*,n} - \Delta A_t^{*,n}) \right]. \end{aligned}$$

From Proposition 2 it is clear that $\Delta A_t^{*,n}$ is a positive, increasing process. Hence

$$\mathbb{E}_0 \left[\max_{0 \leq t \leq \tau^*} (\Delta V_0^{*,n} - \Delta A_t^{*,n}) \right] \leq \Delta V_0^{*,n},$$

which completes the proof. \square

Finally we consider the optimal stopping time.

Lemma 6 *The stopping time $\tau^* = \min\{t : A_{t+1}^{*,n-1} > 0\}$ is identical to the stopping time τ_{n-1}^* under an optimal policy π^* .*

Proof. The optimal, value-maximizing policy exercises the option if the value under exercise is greater than the continuation value. The continuation value under $n - 1$ exercise opportunities is $Q_t^{*,n-1}$. The value under exercise is $Z_t + Q_t^{*,n-2}$. The option is therefore exercised under an optimal policy if $Z_t + Q_t^{*,n-2} > Q_t^{*,n-1}$, that is if $Z_t > \Delta Q_t^{*,n-1}$. Thus the stopping time τ_{n-1}^* is equal to

$$\tau_{n-1}^* = \min\{t : Z_t > \Delta Q_t^{*,n-1}\}.$$

On the other hand the increments of the previsible process $A_t^{*,n-1}$ are, by Lemma 1, equal to

$$A_{t+1}^{*,n-1} - A_t^{*,n-1} = [Z_t - \Delta Q_{t+1}^{*,n-1}]_+.$$

Hence we have $\tau^* = \tau_{n-1}^*$ as desired. \square

This completes the proof of Proposition 3.

It remains to show that Theorem 1 follows from Proposition 3. We recall Definition 3,

$$\Delta V_t^{\uparrow,n} = \max_{u \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_u - M_u),$$

where $\mathcal{T} = \{t, \dots, T\}$. Using this notation it suffices to show that upper and lower bounds are the same. We will show firstly in Lemma 7, that

$$\mathbb{E}[\Delta V_0^{\uparrow,n}] \geq \Delta V_0^{*,n}.$$

Secondly, we show that if the optimal policy π^* and martingale M^* are used in the computation of $\Delta V_t^{\uparrow, n}$,

$$\mathbb{E}_0[\Delta V_0^{\uparrow, n}] \leq \Delta V_0^{*, n}.$$

In fact we will be able to show with Lemma 7 and 8 that $\Delta V_0^{\uparrow, n} = \Delta V_0^{*, n}$ holds almost surely in this latter case.

Lemma 7 *For any martingale M with $M_0 = 0$ and any policy $\pi = \{\tau_n, \dots, \tau_1\}$,*

$$\mathbb{E}_0[\Delta V_0^{\uparrow, n}] \geq \Delta V_0^{*, n}$$

Proof. The proof is by induction. For $n = 1$ the claim is clearly true. For $n > 1$ we have

$$\begin{aligned} \mathbb{E}_0[\Delta V_0^{\uparrow, n}] &= \mathbb{E}_0\left[\max_{t \in (\{0, \dots, T\} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t)\right] \\ &= \mathbb{E}_0\left[\max\left\{\max_{0 \leq t < \tau_{n-1}} (Z_t - M_t), \max_{u \in (\{\tau_{n-1}+1, \dots, T\} \setminus \{\tau_{n-2}, \dots, \tau_1\})} (Z_u - M_u)\right\}\right] \\ &= \mathbb{E}_0\left[\max\left\{\max_{0 \leq t < \tau_{n-1}} (Z_t - M_t), \right. \right. \\ (16) \quad &\left. \left. \max_{u \in (\{\tau_{n-1}+1, \dots, T\} \setminus \{\tau_{n-2}, \dots, \tau_1\})} (Z_u - M_u + M_{\tau_{n-1}+1}) - M_{\tau_{n-1}+1}\right\}\right]. \end{aligned}$$

The martingale $M_u - M_{\tau_{n-1}+1}$ is clearly zero at time $u = \tau_{n-1} + 1$. By the inductive hypothesis we can thus conclude that

$$\mathbb{E}_{\tau_{n-1}+1}\left[\max_{u \in (\{\tau_{n-1}+1, \dots, T\} \setminus \{\tau_{n-2}, \dots, \tau_1\})} (Z_u - M_u + M_{\tau_{n-1}+1})\right] \geq \Delta V_{\tau_{n-1}+1}^{*, n}.$$

Using the tower property,

$$\mathbb{E}_{\tau_{n-1}}\left[\max_{u \in (\{\tau_{n-1}+1, \dots, T\} \setminus \{\tau_{n-2}, \dots, \tau_1\})} (Z_u - M_u + M_{\tau_{n-1}+1})\right] \geq \mathbb{E}_{\tau_{n-1}}[\Delta V_{\tau_{n-1}+1}^{*, n}].$$

Hence there exists some \mathcal{F}_T -measurable random variable U with $\mathbb{E}_{\tau_{n-1}}[U] = 0$ such that

$$\max_{u \in (\{\tau_{n-1}+1, \dots, T\} \setminus \{\tau_{n-2}, \dots, \tau_1\})} (Z_u - M_u + M_{\tau_{n-1}+1}) \geq \mathbb{E}_{\tau_{n-1}}[\Delta V_{\tau_{n-1}+1}^{*, n-1}] + U.$$

Using this in (16), we have

$$\begin{aligned} \mathbb{E}_0[\Delta V_0^{\uparrow, n}] &\geq \mathbb{E}_0\left[\max\left\{\max_{0 \leq t < \tau_{n-1}} (Z_t - M_t), \mathbb{E}_{\tau_{n-1}}[\Delta V_{\tau_{n-1}+1}^{*, n-1}] + U - M_{\tau_{n-1}+1}\right\}\right] \\ &= \mathbb{E}_0\left[\max\left\{\max_{0 \leq t < \tau_{n-1}} (Z_t - M_t), \mathbb{E}_{\tau_{n-1}}[\Delta V_{\tau_{n-1}+1}^{*, n-1}] + U + M_{\tau_{n-1}} - M_{\tau_{n-1}+1} - M_{\tau_{n-1}}\right\}\right]. \end{aligned}$$

By the properties of U and M , there exists a process Y_t with $Y_t = 0$ for $t < \tau_{n-1}$ and $\mathbb{E}_{t-1}[Y_t] = 0$ for $t = \tau_{n-1}$ such that

$$\begin{aligned}\mathbb{E}_0[\Delta V_0^{\uparrow, n}] &\geq \mathbb{E}_0\left[\max\left\{\max_{0 \leq t < \tau_{n-1}} (Z_t + Y_t - M_t), \mathbb{E}_{\tau_{n-1}}[\Delta V_{\tau_{n-1}+1}^{*, n-1}] + Y_{\tau_{n-1}} - M_{\tau_{n-1}}\right\}\right] \\ &= \mathbb{E}_0\left[\max_{0 \leq t \leq \tau_{n-1}} (Z_t \mathbf{1}_{\{t < \tau_{n-1}\}} + \mathbb{E}_t[\Delta V_{t+1}^{*, n-1}] \mathbf{1}_{\{t = \tau_{n-1}\}} - (M_t - Y_t))\right].\end{aligned}$$

Setting $M'_t = M_t - Y_t$ for $t \leq \tau_{n-1}$, we see that M' is a martingale such that

$$\mathbb{E}_0[\Delta V_0^{\uparrow, n}] \geq \mathbb{E}_0\left[\max_{0 \leq t \leq \tau_{n-1}} (Z_t \mathbf{1}_{\{t < \tau_{n-1}\}} + \mathbb{E}_t[\Delta V_{t+1}^{*, n-1}] \mathbf{1}_{\{t = \tau_{n-1}\}} - M'_t)\right].$$

As the choice of policy π and martingale M was arbitrary we have

$$\begin{aligned}\mathbb{E}_0[\Delta V_t^{\uparrow, n}] &\geq \inf_{0 \leq \tau \leq T} \inf_{M \in H_0} \mathbb{E}_0\left[\max_{0 \leq t \leq \tau_{n-1}} (Z_t \mathbf{1}_{\{t < \tau_{n-1}\}} + \mathbb{E}_t[\Delta V_{t+1}^{*, n-1}] \mathbf{1}_{\{t = \tau_{n-1}\}} - M_t)\right] \\ &= \Delta V_0^{*, n},\end{aligned}$$

where the last step follows from Proposition 3. \square

Lemma 8 *Under an optimal policy and martingale, the random variable $\Delta V_t^{\uparrow, n}$ is smaller than the true marginal value,*

$$\Delta V_0^{\uparrow, n} \leq \Delta V_0^{*, n}.$$

Together with Lemma 7 this establishes that the positive-biased estimate is almost surely equal to the true marginal value under an optimal policy and martingale.

Proof. The proof is by induction. For $n = 1$ and under an optimal martingale M^* , the random variable $\Delta V_0^{\uparrow, n}$ can be written as

$$\Delta V_0^{\uparrow, n} = \max_{0 \leq t \leq T} (Z_t - M_t^*),$$

where $M_t^* = \Delta M_t^{*, 1} = M_t^{*, 1}$. By the Doob decomposition of the value function $\Delta V_0^{*, 1} = V_0^{*, 1} = \sup_{0 \leq \tau \leq T} \mathbb{E}_0[Z_\tau]$,

$$V_t^{*, 1} = V_0^{*, 1} + M_t^{*, 1} - A_t^{*, 1},$$

it is clear that $Z_t - M_t^{*, 1} \leq V_0^{*, 1} - A_t^{*, 1}$ for all $0 \leq t \leq T$. As $A_t^{*, 1}$ is a positive, increasing process, $Z_t - M_t^* \leq V_0^{*, 1}$. Hence, for any path,

$$\Delta V_0^{\uparrow, 1} = \max_{0 \leq t \leq T} (Z_t - M_t^{*, 1}) \leq V_0^{*, 1},$$

and the claim is proven for $n = 1$.

For $n > 1$ we have

$$\begin{aligned}
\Delta V_0^{\uparrow, n} &= \max_{t \in (\{0, \dots, T\} \setminus \{\tau_{n-1}^*, \dots, \tau_1^*\})} (Z_t - M_t^*) \\
&= \max \left\{ \max_{0 \leq t < \tau_{n-1}^*} (Z_t - M_t^*), \max_{u \in (\{\tau_{n-1}^*+1, \dots, T\} \setminus \{\tau_{n-2}^*, \dots, \tau_1^*\})} (Z_u - M_u^*) \right\}. \\
&= \max \left\{ \max_{0 \leq t < \tau_{n-1}^*} (Z_t - M_t^*), \right. \\
(17) \quad &\left. \max_{u \in (\{\tau_{n-1}^*+1, \dots, T\} \setminus \{\tau_{n-2}^*, \dots, \tau_1^*\})} (Z_u - M_u^* + M_{\tau_{n-1}^*+1}^*) - M_{\tau_{n-1}^*+1}^* \right\}.
\end{aligned}$$

By the induction hypothesis,

$$\Delta V_{\tau_{n-1}^*+1}^{\uparrow, n-1} = \max_{u \in (\{\tau_{n-1}^*+1, \dots, T\} \setminus \{\tau_{n-2}^*, \dots, \tau_1^*\})} (Z_u - M_u^* + M_{\tau_{n-1}^*+1}^*) \leq \Delta V_{\tau_{n-1}^*+1}^{*, n-1},$$

and hence

$$\Delta V_0^{\uparrow, n} \leq \max_{0 \leq t \leq \tau_{n-1}^*} (Z_t \mathbf{1}_{\{t < \tau_{n-1}^*\}} + (\Delta V_{t+1}^{*, n-1} + M_t^* - M_{t+1}^*) \mathbf{1}_{\{t = \tau_{n-1}^*\}} - M_t^*).$$

By definition of M^* , when $t = \tau_{n-1}^*$, the increment $M_{t+1}^* - M_t^*$ is identical to

$$\Delta M_{t+1}^{*, n-1} - \Delta M_t^{*, n-1} = \Delta V_{t+1}^{*, n-1} - \mathbb{E}_t[\Delta V_{t+1}^{*, n-1}].$$

Thus $\Delta V_{t+1}^{*, n-1} - M_{t+1}^* + M_t^* = \mathbb{E}_t[\Delta V_{t+1}^{*, n-1}]$ for $t = \tau_{n-1}^*$ and

$$\Delta V_0^{\uparrow, n} \leq \max_{0 \leq t \leq \tau_{n-1}^*} (Z_t \mathbf{1}_{\{t < \tau_{n-1}^*\}} + \mathbb{E}_t[\Delta V_{t+1}^{*, n-1}] \mathbf{1}_{\{t = \tau_{n-1}^*\}} - \Delta M_t^{*, n}).$$

The proof then follows along the same lines as for $n = 1$. Using (15), the positivity of $A_t^{*, n}$ and the inequality (16) from the Doob decomposition, we can deduce that

$$\begin{aligned}
\Delta V_0^{\uparrow, n} &\leq \max_{0 \leq t \leq \tau_{n-1}^*} (Z_t - A_{t+1}^{*, n-1} - \Delta M_t^{*, n}) \\
&\leq \Delta V_0^{*, n},
\end{aligned}$$

which completes the proof. Together with Lemma 7 it follows that $\Delta V_0^{\uparrow, n} = \Delta V_0^{*, n}$ almost surely for the optimal policy π^* and martingale M^* . \square

6 Proof of Theorem 2

The proof of the error bound consists of two Propositions. The first examines the effect of an error in the martingale approximation.

Proposition 4 *The effect of an error in the martingale approximation can be bounded, for any policy $\pi = \{\tau_{n-1}, \dots, \tau_1\}$, by*

$$\begin{aligned} \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t) \right] - \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t^*) \right] \\ \leq 2\sqrt{(4D_V^2 + \sigma_k^2) T}, \end{aligned}$$

where $\mathcal{T} = \{0, \dots, T\}$.

The second Proposition bounds the remaining bias due to an error in the optimal-policy approximation.

Proposition 5 *The bias in the marginal value of using a possibly non-optimal policy is bounded by*

$$\mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t^*) \right] - \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}^*, \dots, \tau_1^*\})} (Z_t - M_t^*) \right] \leq D_\pi^{n-1},$$

By definition of the bias Propositions 4 and 5 together prove Theorem 2. To prove the first Proposition, the difference between the approximating and correct martingale is defined as:

Definition 9 *The process R_t is defined as the difference $R_t = M_t - M_t^*$ between the optimal martingale M_t^* and the approximation M_t .*

To avoid the trivial bound we assume that the quantities D_V and σ_k exist.

Lemma 9 *For all $n \in \mathbb{N}$, R_t is a martingale and $R_0 = 0$. Moreover the second moments of the increments $R_{t+1} - R_t$ are bounded by*

$$\mathbb{E}_t[(R_{t+1} - R_t)^2] \leq 4D_V^2 + \sigma_k^2$$

and hence $\mathbb{E}_0[(R_t)^2] \leq (4D_V^2 + \sigma_k^2)t$.

Proof. The Monte Carlo approximation (8) to $\mathbb{E}_t[\Delta V_{t+1}^m]$ is given by

$$E_t^m = \frac{1}{k} \sum_1^k \Delta V^m(X_{t+1}^i),$$

where $X_{t+1}^i, i = 1, \dots, k$ are independent samples from the distribution of X_{t+1} conditional on X_t .

As both M_t and M_t^* are martingales that are null at 0, R_t is a martingale which is null at 0. The increment of the process can be written for $\tau_m \geq t > \tau_{m-1}$ as:

$$\begin{aligned}
R_{t+1} - R_t &= (M_{t+1} - M_t) - (M_{t+1}^* - M_t^*) \\
&= (\Delta V_{t+1}^m - E_t^m) - (\Delta V_{t+1}^{*,m} - \mathbb{E}_t[\Delta V_{t+1}^{*,m}]) \\
&= (\Delta V_{t+1}^m - \Delta V_{t+1}^{*,m}) + (\mathbb{E}_t[\Delta V_{t+1}^{*,m}] - \mathbb{E}_t[\Delta V_{t+1}^m]) \\
&\quad + (\mathbb{E}_t[\Delta V_{t+1}^m] - E_t^m)
\end{aligned}$$

The first two terms in brackets are each bounded in absolute value by D_V . Note that the last term $\mathbb{E}_t[\Delta V_{t+1}^m] - E_t^m$ is the error incurred by the Monte Carlo approximation E_t^m to $\mathbb{E}_t[\Delta V_{t+1}^m]$. This error has mean 0, a second moment bounded by σ_k^2 and is independent of $(\Delta V_{t+1}^m - \Delta V_{t+1}^{*,m})$. Thus the second moment of the increment is bounded by $\mathbb{E}_t[(R_{t+1} - R_t)^2] \leq 4D_V^2 + \sigma_k^2$ and the claim follows. \square

Now our first Proposition can be proven.

Proof of Proposition 4: Using the definition of R_t it follows that

$$\begin{aligned}
&\mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t) \right] \\
&= \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t^* - R_t) \right] \\
&\leq \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t^*) \right] + \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} |R_t| \right] \\
&\leq \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t^*) \right] + \mathbb{E}_0 \left[\max_{t \in \mathcal{T}} |R_t| \right] \\
(18) \quad &\leq \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t^*) \right] + \left(\mathbb{E}_0 \left[\max_{0 \leq t \leq T} (R_t)^2 \right] \right)^{\frac{1}{2}}.
\end{aligned}$$

As R_t is a martingale, $(R_t)^2$ is a nonnegative submartingale (by the existence of D_V and σ_k). Thus Doob's submartingale inequality (e.g., Shiryaev 1984 p.464) can be applied to obtain

$$\mathbb{E}_0 \left[\max_{0 \leq t \leq T} (R_t)^2 \right] \leq 4 \mathbb{E}_0 \left[(R_T^n)^2 \right].$$

As M^* is the martingale at which the infimum of the marginal value is attained it follows that by combining (18) and (19), and applying Lemma 9, we have

$$\begin{aligned}
&\mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t) \right] - \mathbb{E}_0 \left[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t^*) \right] \\
&\leq 2 \left(\mathbb{E}_0 \left[(R_T^n)^2 \right] \right)^{\frac{1}{2}} \\
&\leq 2 \sqrt{(4D_V^2 + \sigma_k^2) T},
\end{aligned}$$

which completes the proof. \square

We now prove the second Proposition. First we define a loss function

Definition 10 *The loss function $L : \mathcal{S} \rightarrow \mathbb{R}$ is defined at time t for policy $\pi = \{\tau_n, \tau_{n-1}, \dots, \tau_1\}$ as*

$$L_t^m = \begin{cases} [Z_t - \mathbb{E}_t[\Delta V_{t+1}^{*,m}]]_+ & t < \tau_m \\ [\mathbb{E}_t[\Delta V_{t+1}^{*,m}] - Z_t]_+ & t = \tau_m \\ 0 & t > \tau_m \end{cases}$$

If $m = 0$ we define L_t^0 to vanish identically.

Lemma 10 *The difference in the value of $m(t)$ remaining exercise opportunities under an optimal policy π^* and policy $\pi = \{\tau_n, \dots, \tau_1\}$ is equal to*

$$V_t^{*,m(t)} - V_t^{\pi,m(t)} = \mathbb{E}_t \left[\sum_{u=t}^T L_u^{m(u)} \right],$$

where $m(t)$ is the number of remaining exercise opportunities at time t . Hence $m(t)$ is the smallest natural number such that $t \leq \tau_m$ under policy $\pi = \{\tau_n, \dots, \tau_1\}$.

Proof. The proof is by induction. The claim clearly holds for $t = T$. We assume in the following that the claim holds for $t + 1$. There are $m(t)$ remaining exercise opportunities at time t . Either $t = \tau_{m(t)}$, if the m -th remaining exercise opportunity is used by policy π at time t , or $t < \tau_{m(t)}$, if the continuation decision is made. In this latter case the value under policy π is $V_t^{\pi,m(t)} = \mathbb{E}_t[V_{t+1}^{\pi,m(t)}]$. The value under the optimal, value-maximizing policy π^* is on the other hand equal to

$$V_t^{*,m(t)} = \max\{Z_t + \mathbb{E}_t[V_{t+1}^{*,m(t)-1}], \mathbb{E}_t[V_{t+1}^{*,m(t)}]\}.$$

Hence

$$\begin{aligned} V_t^{*,m(t)} - V_t^{\pi,m(t)} &= \mathbb{E}_t[V_{t+1}^{*,m(t)} - V_{t+1}^{\pi,m(t)}] + [Z_t - \mathbb{E}_t[\Delta V_{t+1}^{*,m(t)}]]_+ \\ &= \mathbb{E}_t[V_{t+1}^{*,m(t)} - V_{t+1}^{\pi,m(t)}] + L_t^{m(t)}. \end{aligned}$$

As the exercise opportunity is not used at time t , $m(t+1) = m(t)$ and it follows by the induction hypothesis that

$$\begin{aligned} V_t^{*,m(t)} - V_t^{\pi,m(t)} &= \mathbb{E}_t[V_{t+1}^{*,m(t+1)} - V_{t+1}^{\pi,m(t+1)}] + L_t^{m(t)} \\ &= \mathbb{E}_t[\mathbb{E}_{t+1}[\sum_{u=t+1}^T L_u^{m(u)}]] + L_t^{m(t)} \end{aligned}$$

By the tower property the r.h.s. is equal to $\mathbb{E}_t[\sum_{u=t}^T L_u^{m(u)}]$ and the claim follows.

If on the other hand $t = \tau_{m(t)}$, then the exercise opportunity is used under policy π and $V_t^{\pi,m(t)} = Z_t + \mathbb{E}_t[V_{t+1}^{\pi,m(t)-1}]$. The value under the optimal policy is again

$$V_t^{*,m(t)} = \max\{Z_t + \mathbb{E}_t[V_{t+1}^{*,m(t)-1}], \mathbb{E}_t[V_{t+1}^{*,m(t)}]\}.$$

Hence for $t = \tau_{m(t)}$

$$\begin{aligned} V_t^{*,m(t)} - V_t^{\pi,m(t)} &= \mathbb{E}_t[V_{t+1}^{*,m(t)-1} - V_{t+1}^{\pi,m(t)-1}] + [\mathbb{E}_t[\Delta V_{t+1}^{*,m(t)}] - Z_t]_+ \\ &= \mathbb{E}_t[V_{t+1}^{*,m(t)-1} - V_{t+1}^{\pi,m(t)-1}] + L_t^{m(t)}. \end{aligned}$$

As the exercise opportunity was used under policy π at time t , $m(t+1) = m(t) - 1$. The claim follows now as in the case $t < \tau_{m(t)}$. \square

The following lemma completes the proof of Proposition 5.

Lemma 11 *If the martingale M_t^* is used in the computation of $\Delta V_0^{\uparrow,n}$,*

$$\Delta V_t^{\uparrow,m(t)} \leq \Delta V_t^{*,m(t)} + \sum_{u=t}^T L_u^{m(u)-1},$$

along any path. By $m(t)$ with $m(t) = n - 1$ we denote again the number of remaining exercise rights under policy $\pi = \{\tau_{n-2}, \dots, \tau_1\}$.

Proof. For $n = 1$ there is nothing to prove as L^0 vanishes identically and, according to Lemma 8,

$$\Delta V_t^{\uparrow,1} \leq \Delta V_t^{*,1}.$$

For $n > 1$, using (17) and

$$\Delta V_{\tau_{n-1}+1}^{\uparrow,n-1} = \max_{u \in (\{\tau_{n-1}^*+1, \dots, T\} \setminus \{\tau_{n-2}^*, \dots, \tau_1^*\})} (Z_u - M_u^* + M_{\tau_{n-1}+1}^*),$$

it follows that

$$\begin{aligned}\Delta V_t^{\uparrow, n} &= \max_{t \leq u \leq \tau_{n-1}} \left((Z_u - M_u^*) \mathbf{1}_{\{u < \tau_{n-1}\}} + (\Delta V_{u+1}^{\uparrow, n-1} - M_{u+1}^*) \mathbf{1}_{\{u = \tau_{n-1}\}} \right). \\ &= \max_{t \leq u \leq \tau_{n-1}} \left(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + (\Delta V_{u+1}^{\uparrow, n-1} - M_{u+1}^* + M_u^*) \mathbf{1}_{\{u = \tau_{n-1}\}} - M_u^* \right).\end{aligned}$$

By the definition of $m(t)$ it is clear that $m(u) = n$ for $u \leq \tau_{n-1}$ and $m(u) = n - 1$ for $u = \tau_{n-1} + 1$. By the inductive hypothesis we have that $\Delta V_{\tau_{n-1}+1}^{\uparrow, n-1} \leq \Delta V_{\tau_{n-1}+1}^{*, n-1} + \sum_{u=\tau_{n-1}+1}^T L_u^{m(u)-1}$. As furthermore $\Delta V_{u+1}^{*, n-1} - M_{u+1}^* + M_u^* = \mathbb{E}_t[\Delta V_{u+1}^{*, n}]$ and $M_u^* = \Delta M_u^{*, n}$ for $t \leq u \leq \tau_{n-1}$, it suffices to show that

$$\max_{t \leq u \leq \tau_{n-1}} \left(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + \mathbb{E}_u[\Delta V_{u+1}^{*, n-1}] \mathbf{1}_{\{u = \tau_{n-1}\}} - \Delta M_u^{*, n} \right) \leq \Delta V_t^{*, n} + \sum_{u=t}^{\tau_{n-1}} L_u^{n-1}.$$

As $A_t^{*, n-1}$ is a nonnegative increasing process, we have

$$\begin{aligned}(20) \quad & \max_{t \leq u \leq \tau_{n-1}} \left(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + \mathbb{E}_u[\Delta V_{u+1}^{*, n-1}] \mathbf{1}_{\{u = \tau_{n-1}\}} - \Delta M_u^{*, n} \right) \\ & \leq \max_{t \leq u \leq \tau_{n-1}} \left((Z_u - A_{u+1}^{*, n-1}) \mathbf{1}_{\{u < \tau_{n-1}\}} + \right. \\ & \quad \left. (\mathbb{E}_u[\Delta V_{u+1}^{*, n-1}] - A_u^{*, n-1}) \mathbf{1}_{\{u = \tau_{n-1}\}} - \Delta M_u^{*, n} \right) + A_{\tau_{n-1}}^{*, n-1}.\end{aligned}$$

From Lemma 1, $A_{t+1}^{*, n-1} - A_t^{*, n-1} = [Z_t - \mathbb{E}_t[\Delta V_{t+1}^{*, n-1}]]_+$. Hence

$$A_{\tau_{n-1}}^{*, n-1} - A_t^{*, n-1} = \sum_{u=t}^{(\tau_{n-1})-1} L_u^{n-1}.$$

On the other hand we can deduce that

$$\begin{aligned}\mathbb{E}_{\tau_{n-1}}[\Delta V_{(\tau_{n-1})+1}^{*, n-1}] - A_{\tau_{n-1}}^{*, n-1} &= Z_{\tau_{n-1}} - A_{\tau_{n-1}+1}^{*, n-1} + [\mathbb{E}_{\tau_{n-1}}[\Delta V_{\tau_{n-1}+1}^{*, n-1}] - Z_{\tau_{n-1}}]_+ \\ &= Z_{\tau_{n-1}} - A_{\tau_{n-1}+1}^{*, n-1} + L_{\tau_{n-1}}^{n-1}.\end{aligned}$$

Putting the last two observations back into equation (20), it follows that

$$\begin{aligned}(21) \quad & \max_{t \leq u \leq \tau_{n-1}} \left((Z_u - A_{u+1}^{*, n-1}) \mathbf{1}_{\{u < \tau_{n-1}\}} + (\mathbb{E}_u[\Delta V_{u+1}^{*, n-1}] - A_u^{*, n-1}) \mathbf{1}_{\{u = \tau_{n-1}\}} - \Delta M_u^{*, n} \right) + A_{\tau_{n-1}}^{*, n-1} \\ & \leq \max_{t \leq u \leq \tau_{n-1}} \left(Z_u - A_{u+1}^{*, n-1} - \Delta M_u^{*, n} \right) + \sum_{u=t}^{\tau_{n-1}} L_u^{n-1}.\end{aligned}$$

In a similar way to (16), it follows from Proposition 1 that

$$\max_{t \leq u \leq \tau_{n-1}} \left(Z_u - A_{u+1}^{*, n-1} - \Delta M_u^{*, n} \right) \leq \Delta V_t^{*, n},$$

and hence, combining (21) and (20), it follows that

$$\max_{t \leq u \leq \tau_{n-1}} \left(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + \mathbb{E}_u[\Delta V_{u+1}^{*, n-1}] \mathbf{1}_{\{u = \tau_{n-1}\}} - \Delta M_u^{*, n} \right) \leq \Delta V_t^{*, n} + \sum_{u=t}^{\tau_{n-1}} L_u^{n-1}.$$

This shows that (19) holds and hence completes the proof of Lemma 11. \square

Proposition 5 then follows from the two preceding lemmas. Taking expectations in Lemma 11 we have that

$$\mathbb{E}_t[\Delta V_t^{\uparrow, n}] \leq \Delta V_t^{*, n} + \mathbb{E}_t\left[\sum_{u=t}^T L_u^{m(u)-1}\right].$$

The last term is, according to Lemma 10, for $t = 0$ equal to $D_{\pi}^{n-1} = V_0^{*, n-1} - V_0^{\pi, n-1}$. Hence Proposition 5 is proven. This completes the proof of Theorem 2.

References

- [1] L. Andersen and M. Broadie (2001): “A Primal-Dual Simulation Algorithm for Pricing Multi-Dimensional American Options”, *Working paper*, Columbia University.
- [2] J. Barraquand and D. Martineau (1995): “Numerical Valuation of High Dimensional Multivariate American Securities”, *J. Fin. and Quant. Anal.*, 30, 383–405.
- [3] D. Brigo and F. Mercurio (2001): *Interest rate models - Theory and practice*, Springer-Verlag.
- [4] M. Broadie and P. Glasserman (1997a): “Pricing American-style securities using simulation”, *J. Econ. Dyn. Cont.*, 21, 1323–1352.
- [5] M. Broadie and P. Glasserman (1997b): “A Stochastic Mesh Method for Pricing High-Dimensional American Options”, *Preprint*.
- [6] E. Clement and D. Lamberton and P. Protter (2002): “An Analysis of a least squares regression method for American option pricing”, *Fin. Stoch.*, 6, 449–471.
- [7] M.B. Haugh and L. Kogan (2001): “Pricing American Options: A Duality Approach”, forthcoming in *Operations Research*.
- [8] F.A. Longstaff and E.S. Schwartz (1992): “Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model”, *J. Finance*, 47, 1259–1282.

- [9] F.A. Longstaff and E.S. Schwartz (2001): “Valuing American Options by Simulation: A Simple Least-Squares Approach”, *Rev. Fin. Studies*, 14, 113–147.
- [10] L.C.G. Rogers (2002): “Monte Carlo valuation of American options”, *Mathematical Finance*, 12, 271–286.
- [11] A.N. Shiryaev (1984): *Probability*, Springer-Verlag.
- [12] J.A. Tilley (1993): “Valuing American Options In A Path Simulation Model”, *Trans. Soc. Act.*, 45, 83–104.
- [13] J. Tsitsiklis and B. Van Roy (1999): “Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives”, *IEEE Trans. Auto. Control*, 44, 1840–1851.
- [14] J. Tsitsiklis and B. Van Roy (2001): “Regression Methods for Pricing Complex American-Style Options”, *IEEE Trans. Neural Net.*, 12, 694–703.