Monte Carlo methods for the valuation of multiple exercise options

Author(s):
Meinshausen, Nicolai; Hambly, B.M.

Publication Date:
2002

Permanent Link:
https://doi.org/10.3929/ethz-a-004448704

Monte Carlo methods for the valuation of multiple exercise options

by

N. Meinshausen[1] and B.M. Hambly[2]

Research Report No. 111
December 2002

Seminar für Statistik

Eidgenössische Technische Hochschule (ETH)

CH-8092 Zürich

Switzerland

[1]ETH Zürich, Seminar für Statistik, 8092 Zürich, Switzerland
[2]Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK

Monte Carlo methods for the valuation of multiple exercise options

N. Meinshausen[‡]and B.M. Hambly[§]

Seminar für Statistik
ETH Zentrum
CH-8092 Zürich, Switzerland

December 2002

Abstract

We discuss Monte-Carlo methods for pricing 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, showing that this improves on earlier work in the single exercise case. 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, cannot be applied when the dimension of the problem gets large. The alternative approach via Monte-Carlo methods works 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 [11] and this has been followed by a number of papers, [1], [3], [4], [12], [8], 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, [6], [9], has been to consider the dual problem and, using this, try to improve the Monte-Carlo approach and produce a high biased estimate for the price.

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 high biased bound on the price. The algorithm is then analysed to show that we can improve on the estimate of the convergence given in [6]. 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 [8], we can construct a reasonable approximation to the optimal martingale.

[‡]ETH Zürich, Seminar für Statistik, 8092 Zürich, Switzerland

[§]Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK

We will consider two examples where our results can be applied. 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 prices can then be bounded, with 90% confidence and are seen to be within 1-2 % of each other using around 1000 sample paths.

Our second example is a very similar product from the energy derivatives market; a swing type option. 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 show how this product falls into our framework and this allows bounds for prices with high and low biases to be obtained. Here we assume there are 1000 possible exercise times and up to 100 exercise possibilities and we still see only a small gap between the upper and lower estimates.

We use simple models to illustrate our ideas but the extension to models with higher dimensionality or more complicated dynamics is straightforward. We note that as our techniques are probabilistic, it is possible that our high biased bound is actually lower than the true value, but the mean of the distribution from which it comes is higher than the true value.

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 $n \in \mathbb{N}$. 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. We will also assume that at $t = 0$ we have $n < T+1$, so that initially there are more possible exercise dates than exercise opportunities. Thus 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, $Z : [0, \ldots, T] \times S \to \mathbb{R}$, the reward earned by using an exercise possibility when the underlying process is at a particular point in space and time. We will assume that Z is bounded and independent of n.

Our aim is to maximize the expected payoff from the multiple exercise opportunities and we begin by defining the optimization problem we will solve.

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

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

We note that $V_t^{\pi,n} : E \to \mathbb{R}$ 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 2.1 *The* value function $V_t^{\star,n}$, *with n remaining exercise possibilities is the expected payoff under an optimal policy,*

$$V_t^{\star,n}(x) = \sup_{\pi} \mathbb{E}_t\big[\sum_{m=1}^{n} Z_{\tau_m} | X_t = x\big]. \tag{2}$$

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

Definition 2.2 (Marginal Value) *The marginal value $\Delta V_t^{\pi,n}$ is defined for every policy π and for $n \geq 1$ as:*

$$\Delta V_t^{\pi,n} \quad = \quad V_t^n - V_t^{n-1}.$$

For $n = 1$ this amounts to $\Delta V_t^1 = V_t^1$. Under an optimal policy $\pi^\star$ 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.

3 The algorithm and error analysis

Our aim is to find a high biased estimate of the value function using duality ideas. We tackle the problem sequentially and find a high 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 3.1 *The marginal value $\Delta V_0^{\star,n}$ is equal to:*

$$\Delta V_0^{\star,n} = \inf_{\pi} \inf_{M \in H_0} \mathbb{E}_t\big[\max_{u \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_u - M_u\big)\big],$$

where $\mathcal{T} = \{0, \ldots, T\}$ is the set of possible exercise dates, $0 \leq \tau_{n-1} < \ldots < \tau_1$ 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 $\pi^\star$ and the martingale $M^\star$ whose increment at time t is the martingale part of the marginal value function:

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

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

Remark 3.2 This is an extension of the single exercise case (n=1) where the dual problem, found in [9, 6], is given by

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

3.1 The algorithm

According to Theorem 3.1, sampling from the quantity

$$\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M_t\big) \tag{3}$$

with $\mathcal{T} = \{0, \ldots, T\}$ yields, for any policy and martingale, a high-biased estimate of the marginal value $\Delta V_0^{\star,n}$. The bias vanishes as soon as we use the optimal policy $\pi^\star$ and the martingale $M^\star$ as specified in Theorem 3.1.

Approximations to both can be obtained from an approximation ΔV_t^n to the marginal value function $\Delta V_t^{\star,n}$. We assume such an approximation to be available for the moment.

An approximation π to the optimal policy $\pi^\star$ can then be obtained using the approximated value function. The option is exercised if the estimated payoff under exercise is, according to the approximated value function, larger than the expected payoff under continuation.

Definition 3.3 *The policy π that is derived from an approximation to the marginal value function is defined by the stopping times $\tau_m, m = 1, \ldots, n$ as*

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

This policy can also be used to obtain a low-biased estimate of the value function by averaging the return obtained by applying it to simulated sample paths of the underlying Markov chain.

An approximation to the martingale which achieves the infimum can also be obtained from an approximation to the value function. Theorem 3.1 states that the infimum is attained for the martingale $M^\star$, with $M_0^\star = 0$ and increments

$$M_{t+1}^\star - M_t^\star = \Delta V_{t+1}^{\star,m} - \mathbb{E}_t[\Delta V_{t+1}^{\star,m}], \tag{4}$$

where m is the largest number such that $t > \tau_m$. We consider the martingale with increments given by replacing the marginal value function by its approximation and replacing the expectation by a sample mean over independent paths.

Definition 3.4 *Let $X_{t+1}^{(i)}$, $i = 1, \ldots, N$ be N independent samples of the underlying Markov chain conditional upon $X_t = x_t$ and $\Delta V_{t+1}^{m\,(i)}$ be the corresponding value under the approximated marginal-value function. Then the martingale M is defined by $M_0 = 0$ and for $1 \le t < T$ and $x_{t+1} = X_{t+1}^{(1)}$ as*

$$M'_{t+1} - M'_t \quad = \quad \Delta V_{t+1}^{m\,(1)} - \frac{1}{N}\sum_{i=1}^{N} \Delta V_{t+1}^{m\,(i)}, \tag{5}$$

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

We still have to specify how to obtain an approximation to the marginal value function. Methods based on value-function regression [8, 13] seem to be suitable but other methods, such as dynamic programming, are possible in a problem of reduced dimensionality. As we are going to employ value-function regression in the numerical examples, in particular the algorithm proposed by Longstaff and Schwartz [8], we briefly describe the main idea behind this approach. The algorithm uses some fixed set of sample paths and works backwards in time, starting at terminal time $t = T$, where the value function is simply the payoff from immediate exercise. A regression with a finite set of basis functions is computed. Then the algorithm proceeds iteratively backward in time. At time $t < T$ approximations for the value function at times larger than t are thus available and can be used, in the sense of definition 3.3, to approximate the optimal policy and hence the payoff for the sample paths under this policy. These payoffs are then used again to determine the regression at time t. Convergence of the precise algorithm in [8] has recently been proven in [5].

3.2 Error Analysis

We examine what bias is incurred by using these approximations instead of the optimal policy $\pi^\star$ and martingale $M^\star$. Using Theorem 3.1 we can define the bias in the marginal value under approximations to the optimal policy and martingale.

Definition 3.5 (Bias) *For any policy $\pi = \{\tau_{n-1}, \dots, \tau_1\}$ and martingale M_t the bias B_t^n of the estimated value is denoted by*

$$B_t^n(\pi, M) = \mathbb{E}_t\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \dots, \tau_1\})} (Z_t - M_t)\big] - \Delta V_t^{\star,n},$$

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

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

Definition 3.6 *We define the distance D_V between $\Delta V^\star$ and the given approximation ΔV as*

$$D_V = \sup_{\substack{m=1,\dots,n,\\ x \in E,\, 0 \le t < T}} \left|\Delta V_t^{\star,m}(x) - \Delta V_t^m(x)\right| \tag{6}$$

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

$$D_\pi^n = V_0^{\star,n} - V_0^{\pi,n} \tag{7}$$

We denote by σ_N the expected standard deviation of the Monte-Carlo estimate in the approximation (5) of the martingale M':

$$\sigma_N^2 = \sup_{\substack{m=1,\dots,n,\\ x \in E,\, 0 \le t < T}} \mathbb{E}_t\Big[\Big(\mathbb{E}_t[\Delta V_{t+1}^m] - \frac{1}{N}\sum_{i=1}^{N} \Delta V_{t+1}^{m\,(i)}\Big)^2\Big]$$

Theorem 3.7 *The bias of the marginal value estimate is bounded by:*

$$B_0^n(\pi, M') \quad \le \quad D_\pi^{n-1} + 2\sqrt{(4D_V^2 + \sigma_N^2)\, T} \tag{8}$$

The proof of this result can be found in Section 6.

Remark 3.8 For the single-exercise case $D_\pi^0 = 0$ as no stopping time has to be chosen. The scaling is an improvement on [6] who obtained a linear scaling in T.

4 Numerical Results

We now apply the proposed method to two products in the financial markets.

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. While the flexible cap exercises each caplet automatically if the interest rate is above K, at the payment date, the chooser flexible cap provides the holder with the decision at the 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 was calibrated to market data. We use a two-factor additive Gaussian model from [2], a variant of the Longstaff-Schwartz model [7]. 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. 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 $t = 0$.

The value function approximation is obtained by using the Longstaff-Schwartz algorithm [8] with 1000 sample paths. The basis functions for the algorithm are chosen to be:

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

The low-biased estimate is the average, discounted, payoff over 1000 new sample paths, using the stopping times of the Longstaff-Schwartz algorithm. The same stopping times are then used to compute the high-biased estimate. We use $N = 50$ paths in the approximation of the value function at each timestep to construct the martingale M'. The quantity

$$\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1}, \ldots, \tau_1\})} \left(Z_t - M_t' \right) \tag{9}$$

with $\mathcal{T} = \{1, 2, \ldots, 40\}$ is calculated for 20 sample paths. They are then averaged to form the estimate of the marginal value. The marginal values are summed to obtain a high-biased estimate for the total value of the option. The 90% confidence interval for both the low- and high-biased estimator is the interval centered at the mean over all sample paths and with a width that is obtained as the square root of the summed variance of all estimates of the necessary marginal values, divided by the square root of the number of samples (1000 for the high- and 20 for the low-biased estimate) and multiplied by the factor 1.645. The 90% confidence interval for the true value stretches from the lower boundary of the confidence interval of the low-biased estimate to the upper boundary of the confidence interval of the high-biased estimate. The constants are set to $a = 5, b = 2, \sigma = 0.05, \eta = 0.02, \rho = 0.2$, with $\phi_t = 0.05$ for all time and the strike given by $K = 0.05$.

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 [2] Theorem 4.2.1 and the interest rate dynamics in the T-forward measure given in [2] Lemma 4.2.2.

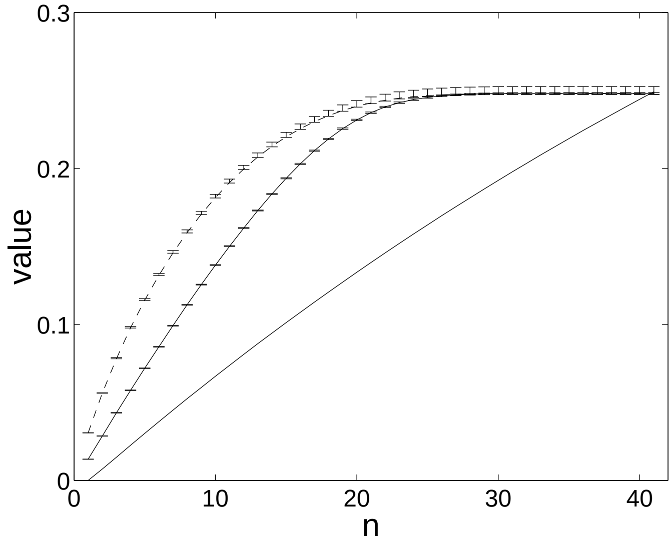

Figure 1: The confidence interval and low-biased estimate for the value of a caplet as a function of the number n of caplets for the chooser flexible cap (topmost line). Also shown are the mean and confidence intervals for the value of an option with automatic exercise on the first n exercise possibilities (flexible cap), while the lowest line shows the exact value of a cap with n caplets.

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 high- and low-biased estimates obtained through our algorithm.

exercise possibilities	low-biased estimate	high-biased estimate	relative difference	90% confidence interval
1	0.03047	0.03049	0.0007	[0.03040 , 0.03053]
2	0.05598	0.05612	0.0025	[0.05587 , 0.05618]
3	0.07811	0.07858	0.0059	[0.07797 , 0.07865]
4	0.09783	0.09854	0.0073	[0.09766 , 0.09862]
5	0.11556	0.11641	0.0074	[0.11536 , 0.11650]
6	0.13146	0.13257	0.0084	[0.13124 , 0.13267]
7	0.14591	0.14725	0.0092	[0.14566 , 0.14735]
8	0.15896	0.16053	0.0099	[0.15869 , 0.16064]
9	0.17072	0.17248	0.010	[0.17043 , 0.17259]
10	0.18140	0.18332	0.010	[0.18109 , 0.18344]
15	0.22029	0.22301	0.012	[0.21989 , 0.22314]
20	0.23985	0.24341	0.014	[0.23937 , 0.24355]
25	0.24678	0.25086	0.016	[0.24626 , 0.25101]
30	0.24801	0.25230	0.017	[0.24747 , 0.25245]
35	0.24807	0.25243	0.017	[0.24754 , 0.25257]
40	0.24807	0.25244	0.017	[0.24754 , 0.25258]

The table above shows the numerical results for n, the number of caplets that can be exercised, ranging from just one to the maximal possible number $n = 40$. From left to right we give the low- and high-biased estimates of the value at time $t = 0$, their relative difference and the 90% confidence interval for the true price of the option. We note that the size of the confidence interval of the value of the chooser flexible cap is mainly determined by the difference between low- and

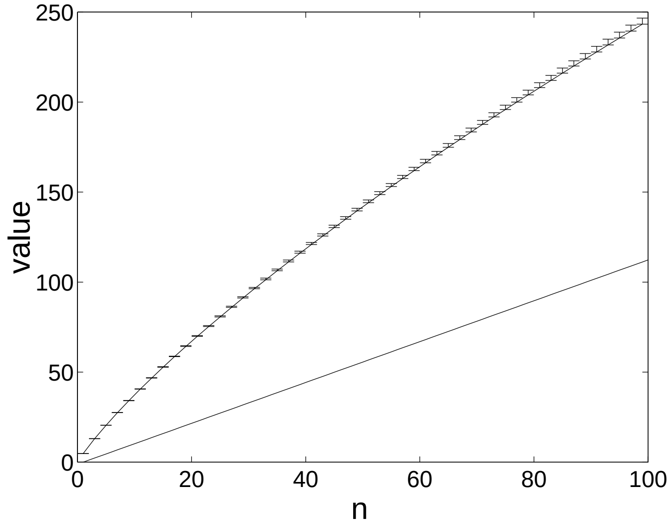

Figure 2: The confidence interval and low-biased point estimate for the value as a function of the number n of total exercise rights in the case of the Swing option. The lower line shows the exact value of an option with automatic exercise on the first n exercise possibilities.

high- biased estimate of the marginal value and much less so by the variance of these estimates. The obtained values for the chooser flexible cap are compared in Figure 1 to the corresponding values of a flexible cap and a cap.

It can be seen that the difference between low- and high-biased estimate of the value of the chooser flexible cap is below 1.7% 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 for the sake of discussion is the so-called full swing option that entitles its holder 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 maxima l 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 possibilities.

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:

$$\log S_{t+1} = (1-k)(\log S_{t-1} - \mu) + \mu + \sigma W_t, \tag{10}$$

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 are chosen to be:

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

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

exercise possibilities	low-biased estimate	high-biased estimate	relative difference	90% confidence interval
1	4.777	4.790	0.002	[4.775 , 4.793]
2	9.024	9.085	0.006	[9.019 , 9.089]
3	12.970	13.094	0.009	[12.963 , 13.098]
4	16.786	16.899	0.006	[16.778 , 16.904]
5	20.455	20.573	0.005	[20.445 , 20.578]
10	37.334	37.531	0.005	[37.316 , 37.537]
15	52.713	52.999	0.005	[52.686 , 53.006]
20	67.105	67.513	0.006	[67.070 , 67.521]
30	93.742	94.507	0.008	[93.691 , 94.515]
40	118.457	119.611	0.009	[118.391 , 119.620]
50	141.832	143.345	0.010	[141.750 , 143.355]
60	164.112	166.020	0.011	[164.015 , 166.031]
70	185.511	187.711	0.011	[185.399 , 187.723]
80	206.045	208.682	0.012	[205.917 , 208.695]
90	225.900	228.965	0.013	[225.757 , 228.978]
100	245.157	248.630	0.013	[244.999 , 248.644]

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 low- and high-biased estimate remains below 1.3%. 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 3.7.

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 The proof of Theorem 3.1

We establish the duality result through a sequence of lemmas.

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

Definition 5.1 (Continuation Value) *The* continuation value $Q_t^{\star,n}$ *is the expectation of the value function one timestep later,*

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

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

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

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

With this definition and (2) we can write the value function $V_t^{\star,n}$ with n remaining exercise possibilities as

$$V_t^{\star,n}(x) = \sup_{t \leq \tau_n < T} \mathbb{E}\big[Z_{\tau_n} + Q_{\tau_n}^{\star,n-1}|X_t = x\big], \tag{11}$$

$$\text{or as} \quad V_t^{\star,n}(x) = \max\big\{Z_t(x) + Q_t^{\star,n-1}(x), Q_t^{\star,n}(x)\big\}. \tag{12}$$

The second expression is easily seen as $Z_t + Q_t^{\star,n-1}$ 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 $Q_t^{\star,n}$ on the other hand is the expected total payoff at time t under continuation without exercise at time t. The optimal policy $\pi^\star$ takes, by definition, the value-maximizing decision, hence the value at time t is the maximum of the two quantities.

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

$$V_t^{\star,n} = V_0^{\star,n} + M_t^{\star,n} - A_t^{\star,n}, \tag{13}$$

where $M_t^{\star,n}$ is a martingale and $A_t^{\star,n}$ a previsible increasing process, both zero at $t = 0$. We denote for $n \geq 1$ by $\Delta M_t^{\star,n}$ the difference $M_t^{\star,n} - M_t^{\star,n-1}$ and likewise $\Delta A_t^{\star,n} = A_t^{\star,n} - A_t^{\star,n-1}$.

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

Proposition 5.2 *The marginal value $\Delta V_t^{\star,n}$ is equal to*

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

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

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 5.2: Using (11), the martingale property of $M_t^{\star,n}$ and the optional stopping theorem

$$\begin{aligned}
\Delta V_t^{\star,n} &= V_t^{\star,n} - V_t^{\star,n-1} \\
&\overset{(11)}{=} \sup_{t\leq\tau\leq T} \mathbb{E}_t\big[Z_\tau + \mathbb{E}_\tau[V_{\tau+1}^{\star,n-1}]\big] - V_t^{\star,n-1} \\
&= \sup_{t\leq\tau\leq T} \mathbb{E}_t\big[Z_\tau + V_\tau^{\star,n-1} + (\mathbb{E}_\tau[V_{\tau+1}^{\star,n-1}] - V_\tau^{\star,n-1})\big] - V_t^{\star,n-1}.
\end{aligned}$$

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

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

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

$$\begin{aligned}
\Delta V_t^{\star,n} &= \sup_{t\leq\tau\leq T} \mathbb{E}_t\big[Z_\tau + V_t^{\star,n-1} - (A_{\tau+1}^{\star,n-1} - A_t^{\star,n-1})\big] - V_t^{\star,n-1} \\
&= \sup_{t\leq\tau\leq T} \mathbb{E}_t\big[Z_\tau - A_{\tau+1}^{\star,n-1}\big] + A_t^{\star,n-1},
\end{aligned} \tag{14}$$

which completes the proof. □

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

Lemma 5.3 *The increments of the process $A_t^{\star,n}$ can be expressed as*

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

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

Proof. The Doob decomposition allows us to write

$$\begin{aligned} A^{\star,n}_{t+1} - A^{\star,n}_{t} &= V^{\star,n}_t - Q^{\star,n}_t \\ &\overset{(12)}{=} \max\left\{Z_t + Q^{\star,n-1}_t, Q^{\star,n}_t\right\} - Q^{\star,n}_t \\ &= \left[Z_t - Q^{\star,n}_t + Q^{\star,n-1}_t\right]_+ \\ &= \left[Z_t - \Delta Q^{\star,n}_t\right]_+. \end{aligned}$$

□

The preceding Lemma is now used to prove the intuitive fact that marginal values decrease as the number of exercise opportunities remaining increases.

Proposition 5.4 *The marginal value is a decreasing function of the number of exercise opportunities remaining, in that for $n \geq 2$,*

$$\begin{aligned} \Delta V^{\star,n}_t &\leq \Delta V^{\star,n-1}_t, \\ \textit{and also}\quad \Delta A^{\star,n}_{t+1} &\geq \Delta A^{\star,n}_t. \end{aligned} \tag{15}$$

Proof. The proof is by induction. We begin by showing $\Delta V^{\star,2}_t \leq \Delta V^{\star,1}_t = V^{\star,1}_t$. Using Proposition 5.2, the marginal value can be written as

$$\Delta V^{\star,2}_t = \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau - A^{\star,1}_{\tau+1}] + A^{\star,1}_t.$$

As $A^{\star,1}_t$ is an increasing process, $A^{\star,1}_{\tau+1} - A^{\star,1}_t \geq 0,\ \forall\, \tau \geq t$ and therefore

$$\begin{aligned} \Delta V^{\star,2}_t &= \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau - A^{\star,1}_{\tau+1}] + A^{\star,1}_t \\ &\leq \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau - A^{\star,1}_t] + A^{\star,1}_t \\ &= \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau] = \Delta V^{\star,1}_t, \end{aligned}$$

as required.

As $\Delta A^{\star,0}_t \equiv 0$ and $\Delta A^{\star,1}_t$ is an increasing process we have the first step in the induction for our second result. We now assume, by the inductive hypothesis, that for $n \geq 2$, $\Delta Q^{\star,n-1}_t \leq \Delta Q^{\star,n-2}_t$ and we show that then $\Delta V^{\star,n}_t \leq \Delta V^{\star,n-1}_t$ and $\Delta A^{\star,n}_{t+1} \geq \Delta A^{\star,n}_t$. By Lemma 5.3 the increments can be written as

$$A^{\star,n-1}_{t+1} - A^{\star,n-1}_t = \left[Z_t - \Delta Q^{\star,n-1}_t\right]_+,$$

and by the inductive hypothesis it follows that

$$\begin{aligned} A^{\star,n-1}_{t+1} - A^{\star,n-1}_t &= \left[Z_t - \Delta Q^{\star,n-1}_t\right]_+ \\ &\geq \left[Z_t - \Delta Q^{\star,n-2}_t\right]_+ = A^{\star,n-2}_{t+1} - A^{\star,n-2}_t. \end{aligned} \tag{16}$$

By rearranging we have the induction for our second claim. We now apply (16) to the representation for the marginal value given in Proposition 5.2

$$\begin{aligned} \Delta V^{\star,n}_t &= \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau - A^{\star,n-1}_{\tau+1}] + A^{\star,n-1}_t \\ &\leq \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau - A^{\star,n-2}_{\tau+1}] + A^{\star,n-2}_t \\ &= \Delta V^{\star,n-1}_t, \end{aligned} \tag{17}$$

completing our induction for $\Delta V^{\star,n}_t$. □

Lemma 5.5 *The marginal value process is a supermartingale. That is for all $n \geq 1$*

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

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

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

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

$$\Delta V_t^{\star,n} - \Delta Q_t^{\star,n} = \Delta A_{t+1}^{\star,n} - \Delta A_t^{\star,n}, \tag{18}$$

which is positive by (15), completing the proof. □

Lemma 5.6 *It holds that*

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

Proof. This uses Proposition 5.4 and Lemma 5.3,

$$\begin{aligned}
\Delta V_t^{\star,n} &= V_t^{\star,n} - V_t^{\star,n-1} \\
&\overset{\text{Propos. 5.2}}{=} \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{\star,n-1} + A_t^{\star,n-1}] \\
&= \max\{Z_t - (A_{t+1}^{\star,n-1} - A_t^{\star,n-1}), \\
&\qquad \mathbb{E}_t[\sup_{t+1\leq\tau\leq T} \mathbb{E}_{t+1}[Z_\tau - A_{\tau+1}^{\star,n-1} + A_t^{\star,n-1}]]\} \\
&\overset{\text{Propos. 5.2}}{=} \max\{Z_t - (A_{t+1}^{\star,n-1} - A_t^{\star,n-1}), \Delta Q_t^{\star,n} - (A_{t+1}^{\star,n-1} - A_t^{\star,n-1})\} \\
&\overset{\text{Lemma 5.3}}{=} \max\{Z_t - [Z_t - \Delta Q_t^{\star,n-1}]_+, \Delta Q_t^{\star,n} - (A_{t+1}^{\star,n-1} - A_t^{\star,n-1})\} \\
&\leq \max\{\Delta Q_t^{\star,n-1}, \Delta Q_t^{\star,n}\} \\
&\overset{\text{Prop. 5.4}}{=} \Delta Q_t^{\star,n-1},
\end{aligned} \tag{19}$$

which completes the proof. □

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

Proposition 5.7 *The marginal value can be expressed as*

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

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

We prove the proposition in three parts.

Lemma 5.8 *It holds for all stopping times τ bounded by T and all martingales $M_t \in H_0$ that*

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

Proof. Let ϑ and τ denote stopping times. By using Proposition 5.2, and the fact that $A_t^{\star,n-1}$ is increasing and zero at time $t=0$,

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

By Lemma 5.6, $\Delta V_\vartheta^{\star,n} \leq \mathbb{E}_\vartheta[\Delta V_{\vartheta+1}^{\star,n-1}]$. Introducing a martingale M zero at $t=0$,

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

□

To complete the proof of the first part of Proposition 5.7 it has to be shown that the following inequality holds as well.

Lemma 5.9 *The marginal value can be bounded from below by*

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

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

Proof. For any stopping time τ and martingale M,

$$\begin{aligned}
&\inf_{0\leq\tau\leq T}\inf_{M\in H_0^1}\mathbb{E}_0\big[\max_{0\leq t\leq\tau}\big(Z_t\mathbf{1}_{\{t<\tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]\mathbf{1}_{\{t=\tau\}} - M_t^n\big)\big]\\
&\quad\leq \mathbb{E}_0\big[\max_{0\leq t\leq\tau^\star}\big(Z_t\mathbf{1}_{\{t<\tau^\star\}} + \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]\mathbf{1}_{\{t=\tau^\star\}} - \Delta M_t^{\star,n}\big)\big].
\end{aligned}\tag{20}$$

The stopping time $\tau^\star$ is chosen as $\min\{t : A_{t+1}^{\star,n-1} > 0\}$. Hence it holds that $A_t^{\star,n-1} = 0$ for $t \leq \tau^\star$ and $A_{\tau^\star+1}^{\star,n-1} > 0$. It follows from Lemma 5.3 that

$$\begin{aligned}
\mathbb{E}_{\tau^\star}[\Delta V_{\tau^\star+1}^{\star,n-1}] &= Z_{\tau^\star} - A_{\tau^\star+1}^{\star,n-1} + A_{\tau^\star}^{\star,n-1}\\
&= Z_{\tau^\star} - A_{\tau^\star+1}^{\star,n-1}.
\end{aligned}\tag{21}$$

Using this in (20) it follows that

$$\begin{aligned}
&\inf_{0\leq\tau\leq T}\inf_{M\in H_0^1}\mathbb{E}_0\big[\max_{0\leq t\leq\tau}\big(Z_t\mathbf{1}_{\{t<\tau\}} + \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]\mathbf{1}_{\{t=\tau\}} - M_t^n\big)\big]\\
&\quad\leq \mathbb{E}_0\big[\max_{0\leq t\leq\tau^\star}\big(Z_t - A_{t+1}^{\star,n-1} - \Delta M_t^{\star,n}\big)\big].
\end{aligned}\tag{22}$$

As $A_t^{\star,n-1}$ is positive, by Proposition 5.2, $\Delta V_t^{\star,n} \geq \sup_{t\leq\tau\leq T} \mathbb{E}_t[Z_\tau - A_{\tau+1}^{\star,n-1}]$. Thus, by the Doob decomposition,

$$\begin{aligned} Z_t - A_{t+1}^{\star,n-1} &\leq \Delta V_t^{\star,n} \\ &= \Delta V_0^{\star,n} + \Delta M_t^{\star,n} - \Delta A_t^{\star,n}. \end{aligned} \tag{23}$$

Hence

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

and, replacing this in (22),

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

From Proposition 5.4 it is clear that $\Delta A_t^{\star,n}$ is a positive, increasing process. Hence

$$\mathbb{E}_0\big[\max_{0\leq t\leq\tau^\star} \big(\Delta V_0^{\star,n} - \Delta A_{t+1}^{\star,n}\big)\big] \leq \Delta V_0^{\star,n},$$

which completes the proof. □

Finally we consider the optimal stopping time.

Lemma 5.10 *The stopping time $\tau^\star = \min\{t : A_{t+1}^{\star,n-1} > 0\}$ is identical to the stopping time $\tau_{n-1}^\star$ under an optimal policy $\pi^\star$.*

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 possibilities is $Q_t^{\star,n-1}$. The value under exercise is $Z_t + Q_t^{\star,n-2}$. The option is therefore exercised under an optimal policy if $Z_t + Q_t^{\star,n-2} > Q_t^{\star,n-1}$, that is if $Z_t > \Delta Q_t^{\star,n-1}$. Thus the stopping time $\tau_{n-1}^\star$ is equal to

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

On the other hand the increments of the previsible process $A_t^{\star,n-1}$ are, by Lemma 5.3, equal to

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

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

This completes the proof of Proposition 5.7.

Now we have to show that Theorem 3.1 follows from Proposition 5.7. We need the following definition:

Definition 5.11 (Single-Sample Estimates) *The single-sample estimates $\widehat{\Delta V}_t^n$ of the marginal value $\Delta V_t^{\star,n}$ are defined for any stopping time τ_{n-1} and martingale M zero at time t as*

$$\widehat{\Delta V}_t^n = \max_{u\in(\mathcal{T}\setminus\{\tau_{n-1},\ldots,\tau_1\})} \big(Z_u - M_u\big),$$

where $\mathcal{T} = \{t,\ldots,T\}$. Under the optimal policy $\pi^\star$ and martingale $M^\star$, as specified in Theorem 3.1, we denote the resulting single-sample estimate by

$$\widehat{\Delta V}_t^{\star,n} = \max_{u\in(\mathcal{T}\setminus\{\tau_{n-1}^\star,\ldots,\tau_1^\star\})} \big(Z_u - M_u^\star\big),$$

Lemma 5.12 *The single-sample estimate is greater than or equal to the true marginal value for any path, that is* $\widehat{\Delta V}_t^n \geq \Delta V_t^{\star,n}$ *along any path.*

Proof. The proof is by induction. For $n = 1$ the claim is clearly true. For $n > 1$ we rewrite the single-sample estimate, in a form similar to the r.h.s. of Proposition 5.7, as

$$\begin{aligned}
\widehat{\Delta V}_t^n &= \max_{u \in (\{t,\ldots,T\} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_u - M_u\big) \\
&= \max\Big\{ \max_{t \leq u < \tau_{n-1}} \big(Z_u - M_u\big), \max_{v \in (\{\tau_{n-1}+1,\ldots,T\} \setminus \{\tau_{n-2},\ldots,\tau_1\})} \big(Z_v - M_v\big)\Big\} \\
&= \max\Big\{ \max_{0 \leq u < \tau_{n-1}} \big(Z_u - M_u\big), \widehat{\Delta V}_{\tau_{n-1}+1}^{n-1} - M_{\tau_{n-1}+1}\Big\} \\
&= \max_{t \leq u \leq \tau_{n-1}} \big(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + (\widehat{\Delta V}_{u+1}^{n-1} - M_{u+1} + M_u)\mathbf{1}_{\{u = \tau_{n-1}\}} - M_u\big) \qquad (24)
\end{aligned}$$

We have for some other martingale $\widetilde{M}_u$ that

$$\begin{aligned}
& \max_{t \leq u \leq \tau_{n-1}} \big(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + (\widehat{\Delta V}_{u+1}^{n-1} - M_{u+1} + M_u)\mathbf{1}_{\{u = \tau_{n-1}\}} - M_u\big) \\
= & \max_{t \leq u \leq \tau_{n-1}} \big(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + \mathbb{E}_u[\widehat{\Delta V}_{u+1}^{n-1}]\mathbf{1}_{\{u = \tau_{n-1}\}} - \widetilde{M}_u\big)
\end{aligned}$$

By induction hypothesis $\mathbb{E}_u[\widehat{\Delta V}_{u+1}^{n-1}] \geq \mathbb{E}_u[\Delta V_{u+1}^{\star,n-1}]$. Hence

$$\begin{aligned}
& \max_{t \leq u \leq \tau_{n-1}} \big(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + \mathbb{E}_u[\widehat{\Delta V}_{u+1}^{n-1}]\mathbf{1}_{\{u = \tau_{n-1}\}} - \widetilde{M}_u\big) \\
\geq & \max_{t \leq u \leq \tau_{n-1}} \big(Z_u \mathbf{1}_{\{u < \tau_{n-1}\}} + \mathbb{E}_u[\Delta V_{u+1}^{\star,n-1}]\mathbf{1}_{\{u = \tau_{n-1}\}} - \widetilde{M}_u\big) \\
\geq & \inf_{0 \leq \tau \leq T} \inf_{M \in H_0} \max_{t \leq u \leq \tau} \big(Z_u \mathbf{1}_{\{u < \tau\}} + \mathbb{E}_u[\Delta V_{u+1}^{\star,n-1}]\mathbf{1}_{\{u = \tau\}} - M_u\big) \\
= & \Delta V_t^{\star,n},
\end{aligned}$$

which completes the proof. □

Lemma 5.13 *Under an optimal policy* $\pi^\star$ *and martingale* $\Delta M_t^{\star,n}$*, the single-sample path estimate is constant for any path. In particular,*

$$\widehat{\Delta V}_0^{\star,n} = \Delta V_0^{\star,n},$$

for any path.

Proof. The proof is by induction. For $n = 1$, we have that $M_t^\star = \Delta M_t^{\star,1}$. By the Doob decomposition $V_t^{\star,1} = V_0^{\star,1} + M_t^{\star,1} - A_t^{\star,1}$ of the value function $V_t^{\star,1} = \sup_{t \leq \tau \leq T} \mathbb{E}_t[Z_\tau]$, it is clear that $Z_t - M_t^{\star,1} \leq V_0^{\star,1} - A_t^{\star,1}$ for all $1 \leq t \leq T$. As $A_t^{\star,1}$ is a positive, increasing process, $Z_t - M_t^{\star,1} \leq V_0^{\star,1}$ for $0 \leq t \leq T$. Hence $\max_{0 \leq t \leq T} \big(Z_t - M_t^{\star,1}\big) \leq V_0^{\star,1}$ for any path. On the other hand it was shown that

$$\mathbb{E}_0[\max_{0 \leq t \leq T} \big(Z_t - M_t^{\star,1}\big)] = V_0^{\star,1}.$$

Hence

$$\max_{0\le t\le T}\left(Z_t - M_t^{\star,1}\right) = V_0^{\star,1}$$

for any path and the claim follows for $n=1$. For $n>1$ we have from (24) that

$$\widehat{\Delta V}_0^{\star,n} \quad = \quad \max_{0\le t\le \tau_{n-1}^\star}\left(Z_t \mathbf{1}_{\{t<\tau_{n-1}^\star\}} + (\widehat{\Delta V}_{t+1}^{n-1} - M_{t+1}^\star + M_t^\star)\mathbf{1}_{\{t=\tau_{n-1}\}} - M_t^\star\right) \tag{25}$$

We assume, by the inductive hypothesis, that the claim holds for $n-1$ in that we have $\widehat{\Delta V}_{t+1}^{n-1} = \Delta V_{t+1}^{\star,n-1}$. Hence

$$\widehat{\Delta V}_0^{\star,n} \quad = \quad \max_{0\le t\le \tau_{n-1}^\star}\left(Z_t \mathbf{1}_{\{t<\tau_{n-1}^\star\}} + (\Delta V_{t+1}^{\star,n-1} - M_{t+1}^\star + M_t^\star)\mathbf{1}_{\{t=\tau_{n-1}\}} - M_t^\star\right).$$

By definition of $M^\star$, when $t=\tau_{n-1}$, the increment $M_{t+1}^\star + M_t^\star$ is identical to $\Delta M_{t+1}^{\star,n-1} - \Delta M_t^{\star,n-1}$. Thus $\Delta V_{t+1}^{\star,n-1} - M_{t+1}^\star + M_t^\star = \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]$ for $t=\tau_{n-1}$ and

$$\widehat{\Delta V}_0^{\star,n} \quad = \quad \max_{0\le t\le \tau_{n-1}^\star}\left(Z_t \mathbf{1}_{\{t<\tau_{n-1}^\star\}} + \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]\mathbf{1}_{\{t=\tau_{n-1}\}} - \Delta M_t^{\star,n}\right).$$

The proof then follows along the same lines as for $n=1$. Using (21), the positivity of $A_t^{\star,n}$ and the inequality (23) from the Doob decomposition, we can deduce that

$$\begin{aligned}\widehat{\Delta V}_0^{\star,n} &\le \max_{0\le t\le \tau_{n-1}^\star}\left(Z_t - A_{t+1}^{\star,n-1} - \Delta M_t^{\star,n}\right)\\ &\le \Delta V_0^{\star,n}.\end{aligned}$$

Together with Lemma 5.12 it follows that $\widehat{\Delta V}_t^{\star,n} = \Delta V_t^{\star,n}$, which completes the proof. □

6 Proof of Theorem 3.7

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

Proposition 6.1 *The effect of an error in the martingale approximation can be bounded, for any policy* $\pi = \{\tau_{n-1},\dots,\tau_1\}$, *by*

$$\begin{aligned}\mathbb{E}_0\Big[\max_{t\in(\mathcal{T}\setminus\{\tau_{n-1},\dots,\tau_1\})}\left(Z_t - M_t'\right)\Big] - \mathbb{E}_0\Big[\max_{t\in(\mathcal{T}\setminus\{\tau_{n-1},\dots,\tau_1\})}\left(Z_t - M_t^\star\right)\Big]&\\ \le 2\sqrt{(4D_V^2+\sigma_N^2)\,T},&\end{aligned} \tag{26}$$

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

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

Proposition 6.2 *The bias in the marginal value of using a possibly non-optimal policy is bounded by*

$$\begin{aligned}&\mathbb{E}_0\Big[\max_{t\in(\mathcal{T}\setminus\{\tau_{n-1},\dots,\tau_1\})}\left(Z_t - M_t^\star\right)\Big]\\ &-\mathbb{E}_0\Big[\max_{t\in(\mathcal{T}\setminus\{\tau_{n-1}^\star,\dots,\tau_1^\star\})}\left(Z_t - M_t^\star\right)\Big] \le D_\pi^{n-1},\end{aligned} \tag{27}$$

By definition of the bias Propositions 6.1 and 6.2 together prove Theorem 3.7. To prove the first Proposition, the difference between the approximating and correct martingale is defined as:

Definition 6.3 *The process R_t is defined as the difference $R_t = M'_t - M^\star_t$ between the optimal martingale $M^\star_t$ and the approximation M'_t.*

Lemma 6.4 *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*

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

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

Proof. As both M'_t and $M^\star_t$ 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:

$$\begin{aligned} R_{t+1} - R_t &= \big(M'_{t+1} - M'_t\big) - \big(M^\star_{t+1} - M^\star_t\big) \\ &= \big(\Delta V^m_{t+1} - \frac{1}{p}[\sum_{i=1}^{p} \Delta V^{m\,(i)}_{t+1}]\big) - \big(\Delta V^{\star,m}_{t+1} - \mathbb{E}_t[\Delta V^{\star,m}_{t+1}]\big) \\ &= \big(\Delta V^m_{t+1} - \Delta V^{\star,m}_{t+1}\big) + \big(\mathbb{E}_t[\Delta V^{\star,m}_{t+1}] - \mathbb{E}_t[\Delta V^m_{t+1}]\big) \\ &\quad + \big(\mathbb{E}_t[\Delta V^m_{t+1}] - \frac{1}{p}[\sum_{i=1}^{p} \Delta V^{m\,(i)}_{t+1}]\big) \end{aligned}$$

The first two terms in brackets are each bounded in absolute value by D_V. The last term has mean 0 and second moment bounded by σ_N^2. Thus the second moment of the increment is bounded by $\mathbb{E}_t[(R_{t+1} - R_t)^2] \leq 4D_V^2 + \sigma_N{}^2$. The last claim follows by orthogonality of the martingale increments. □

Now our first Proposition can be proven.
Proof of Proposition 6.1: Using the definition of R_t it follows that

$$\begin{aligned} & \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M'_t\big)\big] \\ = \; & \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M^\star_t - R_t\big)\big] \\ \leq \; & \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M^\star_t\big)\big] + \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} |R_t|\big] \\ \leq \; & \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M^\star_t\big)\big] + \mathbb{E}_0\big[\max_{t \in \mathcal{T}} |R_t|\big] \\ \leq \; & \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M^\star_t\big)\big] + \big(\mathbb{E}_0\big[\max_{0 \leq t \leq T} (R_t)^2\big]\big)^{\frac{1}{2}}. \end{aligned}$$

As R_t is a martingale, $(R_t)^2$ is a nonnegative submartingale. Both process are bounded and Doob's submartingale inequality (e.g. [10] p.464) can be applied to obtain

$$\mathbb{E}_0[\max_{0 \leq t \leq T} (R^n_t)^2] \leq \, 4\, \mathbb{E}_0[(R^n_T)^2].$$

As $M^\star$ is the martingale at which the infimum of the marginal value is attained it follows from Lemma 6.4 that

$$\begin{aligned} & \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M'_t\big)\big] - \mathbb{E}_0\big[\max_{t \in (\mathcal{T} \setminus \{\tau_{n-1},\ldots,\tau_1\})} \big(Z_t - M^\star_t\big)\big] \\ & \qquad \leq 2\, \big(\mathbb{E}_0\big[(R^n_T)^2\big]\big)^{\frac{1}{2}} \\ & \qquad \leq 2\, \sqrt{(4D_V^2 + \sigma_N{}^2)\, T}, \end{aligned}$$

which completes the proof. ◻

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

Definition 6.5 (Loss function) *The loss function $L : \mathcal{S} \to \mathbb{R}$ is defined at time t for policy $\pi = \{\tau_n, \tau_{n-1}, \dots, \tau_1\}$ as*

$$L_t^m = \begin{cases} \left[Z_t - \mathbb{E}_t[\Delta V_{t+1}^{\star,m}]\right]_+ & t < \tau_m \\ \left[\mathbb{E}_t[\Delta V_{t+1}^{\star,m}] - Z_t\right]_+ & t = \tau_m \end{cases} \tag{28}$$

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

Lemma 6.6 *The difference in the value of n remaining exercise possibilities under an optimal policy $\pi^\star$ and policy $\pi = \{\tau_n, \dots, \tau_1\}$ is equal to*

$$V_t^{\star,n} - V_t^{\pi,n} \quad = \quad \mathbb{E}_t\big[\sum_{u=t}^{T} L_u^{m(t)}\big],$$

where $m(t)$ is the number of remaining exercise possibilities 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 $t = T$. Assume that there are m remaining exercise possibilities at time t. Either $t = \tau_m$, if the m-th remaining exercise possibility is used by policy π at time t, or $t < \tau_m$, if the continuation decision is made. In this latter case the value under policy π is $V_t^{\pi,m} = \mathbb{E}_t[V_{t+1}^{\pi,m}]$. The value under the optimal, value-maximizing policy $\pi^\star$ is $V_t^{\star,m} = \max\{Z_t + \mathbb{E}_t[V_{t+1}^{\star,m-1}], \mathbb{E}_t[V_{t+1}^{\star,m}]\}$. Hence

$$\begin{aligned} V_t^{\star,m} - V_t^{\pi,m} &= \mathbb{E}_t[V_{t+1}^{\star,m} - V_{t+1}^{\pi,m}] + [Z_t - \mathbb{E}_t[\Delta V_{t+1}^{\star,m}]]_+ \\ &= \mathbb{E}_t[D_{t+1}^{\pi,m}] + L_t^{m(t)}. \end{aligned} \tag{29}$$

As the exercise opportunity is not used at time t, $m(t+1) = m(t)$ and, by the induction hypothesis, it follows that $V_t^{\star,m} - V_t^{\pi,m} = \mathbb{E}_t\big[\sum_{u=t}^{T} L_u^{m(t)}\big]$. If on the other hand $t = \tau_m$, then the exercise possibility is used under policy π and $V_t^{\pi,m} = Z_t + \mathbb{E}_t[V_{t+1}^{\pi,m}]$. The value under the optimal policy is again $V_t^{\star,m} = \max\{Z_t + \mathbb{E}_t[V_{t+1}^{\star,m-1}, \mathbb{E}_t[V_{t+1}^{\star,m}]]\}$. Hence for $t = \tau_m$

$$\begin{aligned} V_t^{\star,m} - V_t^{\pi,m} &= \mathbb{E}_t[V_{t+1}^{\star,m-1} - V_{t+1}^{\pi,m-1}] + [\mathbb{E}_t[\Delta V_{t+1}^{\star,m}] - Z_t]_+ \\ &= \mathbb{E}_t[D_{t+1}^{\pi,m-1}] + L_t^{m(t)}. \end{aligned} \tag{30}$$

As the exercise opportunity was used under policy π at time t, $m(t+1) = m(t) - 1$. Hence it follows, again by the inductive hypothesis, that $V_t^{\star,m} - V_t^{\pi,m} = \mathbb{E}_t\big[\sum_{u=t}^{T} L_u^{m(t)}\big]$. This completes the proof. ◻

We recall that the single-sample estimates $\widehat{\Delta V}_t^n$ of the marginal value $\Delta Q_t^{\star,n}$ were defined for any policy π and martingale M_t as

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

where $\mathcal{T} = \{t, \dots, T\}$. The following lemma completes the proof of Proposition 6.2.

Lemma 6.7 *If the martingale $M_t^\star$ is used in the computation of the single-sample estimate,*

$$\widehat{\Delta V}_0^n \leq \Delta V_0^{\star,n} + \sum_{t=0}^{T} L_t^{m(t)},$$

along any path. $m(t)$ with $m(0) = n-1$ is the number of remaining exercise rights under policy $\pi = \{\tau_{n-2}, \ldots, \tau_1\}$.

Proof. For $n = 1$ there is nothing to prove as L^0 vanishes identically and, according to Lemma 5.13,

$$\widehat{\Delta V}_0^1 = \Delta V_0^{\star,1}.$$

For $n > 1$ it was previously shown, see equation (25), that the single-sample estimate can be written as

$$\widehat{\Delta V}_0^n = \max_{0\leq t\leq \tau_{n-1}} \left(Z_t \mathbf{1}_{\{t<\tau_{n-1}\}} + (\widehat{\Delta V}_{t+1}^{n-1} - \Delta M_{t+1}^{\star,n} + \Delta M_t^{\star,n})\mathbf{1}_{\{t=\tau_{n-1}\}} - \Delta M_t^{\star,n}\right).$$

By the inductive hypothesis we assume that $\widehat{\Delta V}_{t+1}^{n-1} \leq \Delta V_{t+1}^{\star,n-1} + \sum_{t=\tau_{n-1}+1}^{T} L_t^{n-1}$. As furthermore $\Delta V_{t+1}^{\star,n-1} - \Delta M_{t+1}^{\star,n} + \Delta M_t^{\star,n} = \mathbb{E}_t[\Delta V_{t+1}^{\star,n}]$, it suffices to show that

$$\begin{aligned} &\max_{0\leq t\leq \tau_{n-1}} \left(Z_t \mathbf{1}_{\{t<\tau_{n-1}\}} + \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]\mathbf{1}_{\{t=\tau_{n-1}\}} - \Delta M_t^{\star,n}\right) \\ \leq\ & \Delta V_0^{\star,n} + \sum_{t=0}^{\tau_{n-1}} L_t^{n-1}. \end{aligned} \tag{31}$$

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

$$\begin{aligned} &\max_{0\leq t\leq \tau_{n-1}} \left(Z_t \mathbf{1}_{\{t<\tau_{n-1}\}} + \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]\mathbf{1}_{\{t=\tau_{n-1}\}} - \Delta M_t^{\star,n}\right) \\ &\leq \max_{0\leq t\leq \tau_{n-1}} \left((Z_t - A_{t+1}^{\star,n-1})\mathbf{1}_{\{t<\tau_{n-1}\}} + \right. \\ &\qquad \left.(\mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}] - A_t^{\star,n-1})\mathbf{1}_{\{u=\tau_{n-1}\}} - \Delta M_t^{\star,n}\right) + A_{\tau_{n-1}}^{\star,n-1}. \end{aligned} \tag{32}$$

From Lemma 5.3, $A_{t+1}^{\star,n-1} - A_t^{\star,n-1} = [Z_t - \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]]_+$. As $A_0^{\star,n-1} = 0$, we have

$$A_{\tau_{n-1}}^{\star,n-1} = \sum_{t=0}^{(\tau_{n-1})-1} L_t^{n-1}.$$

On the other hand we can deduce that

$$\begin{aligned} \mathbb{E}_{\tau_{n-1}}[\Delta V_{(\tau_{n-1})+1}^{\star,n-1}] - A_{\tau_{n-1}}^{\star,n-1} &= Z_{\tau_{n-1}} - A_{\tau_{n-1}+1}^{\star,n-1} + [\mathbb{E}_{\tau_{n-1}}[\Delta V_{\tau_{n-1}+1}^{\star,n-1}] - Z_{\tau_{n-1}}]_+ \\ &= Z_{\tau_{n-1}} - A_{\tau_{n-1}+1}^{\star,n-1} + L_{\tau_{n-1}}^{n-1}. \end{aligned} \tag{33}$$

Putting the last two observations back into equation (32), it follows that

$$\begin{aligned} &\max_{0\leq t\leq \tau_{n-1}} \left((Z_t - A_{t+1}^{\star,n-1})\mathbf{1}_{\{t<\tau_{n-1}\}} + (\mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}] - A_t^{\star,n-1})\mathbf{1}_{\{u=\tau_{n-1}\}} - \Delta M_t^{\star,n}\right) + A_{\tau_{n-1}}^{\star,n-1} \\ &\qquad \leq \max_{0\leq t\leq \tau_{n-1}} \left(Z_t - A_{t+1}^{\star,n-1} - \Delta M_t^{\star,n}\right) + \sum_{t=0}^{\tau_{n-1}} L_t^{n-1}. \end{aligned} \tag{34}$$

In a similar way to (23), it follows from Proposition 5.2 that

$$\max_{0\leq t\leq \tau_{n-1}} \left(Z_t - A_{t+1}^{\star,n-1} - \Delta M_t^{\star,n}\right) \leq \Delta V_0^{\star,n},$$

and hence, combining (34) and (32), it follows that

$$\max_{0\le t\le \tau_{n-1}} \big(Z_t \mathbf{1}_{\{t<\tau_{n-1}\}} + \mathbb{E}_t[\Delta V_{t+1}^{\star,n-1}]\mathbf{1}_{\{t=\tau_{n-1}\}} - \Delta M_t^{\star,n}\big) \le \Delta V_0^{\star,n} + \sum_{t=0}^{\tau_{n-1}} L_t^{n-1}.$$

This shows that (31) holds and hence completes the proof of Lemma 6.7. □

Proposition 6.2 follows then from the two preceding lemmas. Taking expectations in Lemma 6.7 we have that

$$\mathbb{E}_0[\widehat{\Delta V}_0^n] \le \Delta V_0^{\star,n} + \mathbb{E}_0[\sum_{t=0}^{T} L_t^{m(t)}].$$

The last term is, according to Lemma 6.6, equal to $D_\pi^n = V_0^{\star,n} - V_0^{\pi,n}$. Hence Proposition 6.2 is proven. This completes the proof of Theorem 3.7.

References

[1] J. Barraquand and D. Martineau (1995): Numerical Valuation of High Dimensional Multivariate American Securities, *J. Fin. and Quant. Anal.*, 30, 383–405.

[2] D. Brigo and F. Mercurio (2001): *Interest rate models - Theory and practice*, Springer-Verlag.

[3] M. Broadie and P. Glasserman (1997): Pricing American-style securities using simulation, *J. Econ. Dyn. Cont.*, 21, 1323–1352.

[4] M. Broadie and P. Glasserman (1997): A Stochastic Mesh Method for Pricing High-Dimensional American Options, *Preprint*.

[5] 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.

[6] M. B. Haugh and L. Kogan (2001): Pricing American Options: A Duality Approach, *Preprint*.

[7] F.A. Longstaff and E.S. Schwartz (1992): Interest Rate Volatility and the Term Stucture: A Two-Factor General Equilibrium Model, *J. Finance*, 47, 1259–1282.

[8] F. A. Longstaff and E. S. Schwartz (2001): Valuing American Options by Simulation: A Simple Least-Squares Approach, *Rev. Fin. Studies*, 14, 113–147.

[9] L.C.G. Rogers (2002): Monte Carlo valuation of American options, *Mathematical Finance*, 12, 271–286.

[10] A.N. Shiryaev (1984): *Probability*, Springer-Verlag.

[11] J.A. Tilley (1993): Valuing American Options In A Path Simulation Model, *Trans. Soc. Act.*, 45, 83–104.

[12] 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.

[13] J. Tsitsiklis (2001): Regression Methods for Pricing Complex American-Style Options, *IEEE Trans. Neural Net.*, 12, 694–703.