Monte Carlo methods via a dual approach for some discrete time stochastic control problems

L. G. Gyurkó[1], B. M. Hambly[2], J. H. Witte[3]

Abstract

We consider a class of discrete time stochastic control problems motivated by some financial applications. We use a pathwise stochastic control approach to provide a dual formulation of the problem. This enables us to develop a numerical technique for obtaining an estimate of the value function which improves on purely regression based methods. We demonstrate the competitiveness of the method on the example of a gas storage valuation problem.

1 Introduction

The numerical pricing of options with early exercise features, such as American options, is a challenging problem, especially when the dimension of the underlying asset increases. There is a large body of literature which discusses this problem from different points of view, beginning with techniques aimed at solving the dynamic programming problem using trees or the associated Hamilton-Jacobi-Bellman equation. Over the past decade, there has been a lot of activity in developing Monte Carlo techniques for such optimal stopping problems. The most popular have been basis function regression methods initially proposed in [13] and [21]. If these methods are used to provide an approximate optimal exercise strategy, they naturally provide lower bounds for prices. Thus they were soon followed by dual methods [16, 12] designed to find upper bounds. An account of these methods can be found in [11].

Following on from the development of dual methods for American options, there has been a strand of research extending these ideas to multiple optimal stopping problems, which correspond to options with multiple exercise features [15]. The dual method proceeds via the idea of pathwise optimization, which originated in [10]. This pathwise optimization method was developed in a general setting in [17] where it was applied to more general stochastic control problems.

In this paper, our aim is to consider a subclass of such stochastic control problems for which we can develop a relatively simple dual approach and which leads to numerical algorithms for the efficient computation of the value function.

We were originally motivated by option pricing problems in the electricity market. In that setting contracts such as swing options give the holder certain rights to exercise variable amounts through the lifetime of the contract. The dual approach, initiated in [15], used a simplistic swing contract in which a single exercise was allowed on each day, with the total number of exercise rights over the lifetime of the contract constrained.

[1]Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK.
E-mail: gyurko@maths.ox.ac.uk

[2]Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK.
E-mail: hambly@maths.ox.ac.uk

[3]Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK.
E-mail: witte@maths.ox.ac.uk

In [1, 5], this was extended to multiple discrete exercises on a given day. Other recent developments have seen a move to continuous time [6] and a 'pure martingale' dual formulation of the problem [19].

Our first aim in this paper is to provide a more general formulation of the dual problem in discrete time which allows exercise of continuous amounts and contains the 'pure martingale' approach.

Our second aim is to provide a useful numerical approach to this type of problem. Having moved beyond the multiple optimal stopping problem to a more general stochastic control formulation, the space of controls is now potentially of dimension greater than one, and consequently more difficult to handle. Instead of a purely binary decision (or at most a finite set of decisions) at each time point, we have the possibility of choosing from a Euclidean space (in the electricity context, this is exercising a real amount corresponding to a volume of power). Our dual formulation of the problem leads naturally to an upper bound on the value function. We develop a technique based on being given an a priori estimate for the value function, say typically an estimate obtained via basis function regression, and converting this to an improved estimate via the dual.

In order to produce the a priori estimate, the method uses least squares regression and Monte Carlo techniques, an extension of the approach due to Longstaff-Schwarz [13] and Tsitsiklis and van Roy [21]; we use test functions that depend on both the underlying factor and the control value. We note that this idea has been considered by Boogert and Jong [7]; however, Boogert and Jong did not develop the extended regression based method in detail, but worked with regression depending only on the underlying factor for several discrete values of the control. Belomestny et al. [4] have also developed a family of least squares regression and Monte-Carlo based numerical algorithms. The algorithm in [4] can be applied to more general discrete time control problems than the ones we consider in this paper. However, as in [7], Belomestny et al. regress the conditional expectation arising in the dynamic programming principle using test functions depending on the underlying factor only. When applied to the same control problem, with the right choice of test functions and grid in the space of underlying factor and control, we found that our extended regression based method performs better than the method in [7] or the method in [4], especially when the control is high dimensional.

The a priori estimate is used as an input to the dual formulation based upper bound. The implementation of the dual estimate requires the numerical solution of several independent deterministic optimal control problems. We note that these control problems can be solved simultaneously, and, hence, it is well suited for a parallel implementation.

As an application, we will focus on one example in this paper, namely natural gas storage valuation. The owner of a natural gas storage facility is faced with an optimal control problem in order to maximize the return from running the facility. The demand for natural gas is seasonal with high demand and prices in the winter, and low demand in the summer. The operator of a facility will want to buy and store gas when it is cheaper over the summer, and then sell gas into the market when the price is higher in the winter. The operation of the facility is thus a control problem where, on a given day, the operator has the decision to inject or produce a volume of gas, given the current price of gas. Thus, we have the set up of a stochastic control problem of the type we consider here. We chose the particular gas storage problem as a numerical example in order to compare the results of our probabilistic approach to the results of the partial differential equation based methods (cf. [8, 20]). In general, we expect the probabilistic approach

to perform better than the PDE methods when the dimension of the underlying factor and/or the dimension of the control is high.

Our numerical example demonstrates that the dual formulation based upper bound is sharper than the one we get from the a priori estimate at comparable computational expense. This empirical observation justifies the potential benefit of computing the dual formulation based estimate in practice.

The outline of the paper is as follows. We will begin with the setup for the problem in Section 2 and follow this with the dual formulation in Section 3. We obtain our main representation in Theorem 3.1, and then derive a version which can be used for the Monte Carlo based numerical technique in Lemma 3.3. We follow this with a discussion of the numerical technique itself in Section 4. Finally, we apply the approach to the gas storage problem in the last section.

2 Discrete time decoupled stochastic control problems

We consider an economy in discrete time defined up to a finite time horizon T. We assume a financial market described by the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathcal{T}}, \mathbb{P})$, where $\mathcal{T} = \{0, 1, \ldots, T\}$. We take $(X_t)_{t\in\mathcal{T}}$ to be an $\mathbb{R}^d$-valued discrete time Markov chain representing the price of the underlying assets and any other variables that affect the dynamics of the underlyings. We assume that the filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$ is generated by X. Moreover, we assume that $\mathbb{P}$ is a risk neutral pricing measure, and write $\mathbb{E}_t(X) = \mathbb{E}(X|\mathcal{F}_t)$ for any random variable X on our probability space. Throughout the paper, we will assume that interest rates are 0.

We phrase our problem in the language of options, even though it is a standard stochastic control problem of maximizing a reward obtained from a randomly evolving system. The payoff of the option (or the reward for the position) $H_t(h, X_t)$ at time $t \in \mathcal{T}$ is a function of the underlying $(X_t)_{t\in\mathcal{T}}$ and the exercise amount $h \in \mathbb{R}^l$, which is chosen by the holder of the option subject to certain constraints. The problems that we consider are decoupled in that the decision at time t regarding h_t has no impact on the evolution of the underlying state of the economy X_s for $s > t$.

The set of *admissible exercise decisions* available at a given time is defined by a (set-valued) function K on $\mathcal{T} \times \mathbb{R}^l \times \mathbb{R}^d$ that takes values in the set of subsets of $\mathbb{R}^l$. The control process $(y_t)_{t\in\mathcal{T}} \subset \mathbb{R}^l$ is defined by $t \geq 0$, $y_{t+1} := y_t - h_t$, for the exercise amount h_t. We will write $K_t(y_t, X_t(\omega))$, or, if needed, $K(t, y_t, X_t(\omega))$, for the given set of admissible exercise decisions depending on t, the state of the underlyings X_t, and the value of the control process y_t. The initial value y_0 and the constraints $K_t(\cdot, \cdot)$ for $t \in \mathcal{T}$ are determined by the option contract.

Definition 2.1 (K-admissible exercise policy)**.** *A policy, or exercise strategy, $\pi = (h_t, \ldots, h_T)$ is a* K-admissible exercise policy *on $\{t, \ldots, T\}$ started at y if it satisfies all of the following properties.*

(i) h_s is $\mathcal{F}_s$-measurable for $s = t, \ldots, T$,

(ii) and $h_s(\omega) \in K_s(y_s, X_s(\omega))$ for all $s = t, \ldots, T$ and for all ω in a set of probability one,

where y_s is defined recursively by $y_t = y$ and $y_{s+1} = y_s - h_s$ for $s = t, \ldots, T-1$.

The set of such policies is denoted by $\mathcal{P}_{K,y,t}$.

Thus, a K-admissible exercise policy π on the time-set $\{t, \dots, T\}$ is defined by the $\mathcal{F}$-adapted process $(h_s)_{s=t,\dots,T}$ describing the exercise decisions at times $t, \dots, T$, and the value of such an exercise policy π at time t is given by

$$V_t^{K,\pi}(y_t, x) = \mathbb{E}\left[\sum_{s=t}^{T} H_s(h_s, X_s)|X_t = x\right]. \tag{2.1}$$

In the particular examples considered in this paper, the set K_t will be a line segment in $\mathbb{R}$ or a quadrant of $\mathbb{R}^2$.

We are now in a position to define the value function $V_t^{K,*}(y_t, X_t)$ at time t of the option satisfying the constraints K.

Definition 2.2. *We define the value function to be*

$$V_t^{K,*}(y, x) = \sup_{\pi \in \mathcal{P}_{K,y,t}} V_t^{K,\pi}(y, x) = \sup_{\pi \in \mathcal{P}_{K,y,t}} \mathbb{E}\left[\sum_{s=t}^{T} H_s(h_s, X_s)|X_t = x\right], \ (y, x) \in \mathbb{R}^l \times \mathbb{R}^d.$$

For simplicity, we make the following assumption.

Assumption 2.3. *There exists a set Y_0^K of initial control values and a bound C such that*

$$\mathbb{E}[|H_s(h, X_s)|] < C \ \ \forall s \in \mathcal{T}, h \in K_s(y, X_s),$$

for all (y, X_s) reachable at time s from $Y_0^K \times \{X_0(\omega)|\omega \in \Omega\}$ by a K-admissible policy.

This is enough to ensure the existence of the value function and the dynamic programming principle. Weaker assumptions which guarantee existence would be possible, but are not the focus of this paper.

In order to indicate the type of problems that fit into this framework, we give three examples. In the final section, we will focus on the first.

Gas storage valuation:

Natural gas storage valuation and optimal operation can be formulated as an option contract as described above. In particular, let $X_t \in \mathbb{R}$ denote the spot price of natural gas at time t, and let $y_t \in \mathbb{R}$ denote the amount of gas stored in the facility at time t. In that case, with h_t denoting the change of level between t and $t+1$, the payoff is defined by

$$H_t(h_t, X_t) = h_t X_t.$$

More accurate models may also take into account the loss of gas occurring at injection.

At time t, the set $K_t(y_t, X_t)$ is determined by the maximum and minimum capacity of the gas storage facility, and by the injection/production rate depending on the stored amount X_t. A continuous time description of this problem was given in [14] and in [20]. In section 5.1, we present a time-discretized version.

Swing option pricing:

In the electricity market, a swing option enables the holder to protect themselves against the risk of price spikes if they are exposed to the spot price of electricity X. The simplest versions give their holder the right, for a specified period of time, to purchase each day

(on- or off- peak time) electricity at a fixed price K (strike price). In this case, the payoff is that of a call $(X_t - K)^+$. When exercising a swing option at a time t, the amount purchased may vary (or swing) between a minimum volume, m_t, and a maximum volume, M_t, while the total quantity purchased for the period must remain within minimum $\bar{m}$ and maximum $\bar{M}$ volume levels. Thus, we have

$$H_t(h_t, X_t) = h_t \max(X_t - K, 0),$$

with $m_t \leq h_t \leq M_t$ and $\bar{m} \leq \sum_{t=0}^{T} h_t \leq \bar{M}$. The set K_t is the line segment determined by these constraints.

Optimal liquidation:

A similar approach can be pursued to model the optimal liquidation of a portfolio of dependent assets. Let $X_t = (X_t^1, X_t^2) \in \mathbb{R}^2$ denote the value of two bonds by the same issuer with different issue dates. We assume that, whenever X^1 is traded, there is a temporary price impact on both, X^1 and X^2, referred to as a *multi-asset price impact*. For example, if $h_t = (h_t^1, h_t^2) \in \mathbb{R}^+ \times \mathbb{R}^+$ denotes the quantities sold of the bonds X^1 and X^2, respectively, the payoff is

$$H_t(h_t, X_t) = h_t^1 X_t^1 + h_t^2 X_t^2 - (h_t^T \Lambda h_t)^\beta$$

for some $\beta > 1/2$ and some matrix $\Lambda \in \mathbb{R}^{2\times 2}$. The reader is referred to [18] for further details. The one-dimensional case is considered in [2].

In this setting, the set K_t is determined by the total volume of each bond that we hold and wish to liquidate, and is hence a subset of $\mathbb{R}^2$.

Many models (see for example [3]) consider modelling the permanent price impact of trades on top of incorporating the temporary impact. In the presence of permanent impact, the optimal liquidation problem is typically formulated as a coupled stochastic control problem. The approach based on coupled control problems falls beyond the scope of this paper. For certain models of permanent price impact (e.g. non-resilient impact), the a priori estimate presented in Section 4.2 can be easily adapted. Deriving a dual formulation, however, is less straightforward.

3 Dual formulation

Definition 2.2 represents the value of the option as the supremum over the set of admissible exercise policies. We now develop a dual for this problem that represents the option value as an infimum over a space of martingale-valued functions. Let $\mathbf{M}$ denote the space of functions defined on $\mathbb{R}^l$ and taking values in the space $\mathcal{M}_0$ of martingales which are adapted to the filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$ and null at time 0. For $M \in \mathbf{M}$, $y \in \mathbb{R}^l$, $t \in \mathcal{T}$, M_t^y denotes the time-t value of $M(y) \in \mathcal{M}_0$.

Theorem 3.1. *Let K be a function defined on $\mathcal{T} \times \mathbb{R}^l \times \mathbb{R}^d$ and taking values in the set of subsets of $\mathbb{R}^l$. Then, for all $y \in Y_0^K$, the value $V_0^{K,*}(y, X_0)$ of the option at time 0 almost surely satisfies the following.*

$$V_0^{K,*}(y,x) = \inf_{M\in\mathbf{M}} \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \sum_{t=0}^{T-1} (H_t(h_t, X_t) - M_{t+1}^{y_{t+1}} + M_t^{y_{t+1}}) + H_T(h_T, X_T) \middle| X_0 = x\right]. \tag{3.1}$$

Moreover,

$$V_0^{K,*}(y,x) = \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \sum_{t=0}^{T-1}(H_t(h_t,X_t) - M_{t+1}^{K,y_{t+1}} + M_t^{K,y_{t+1}}) + H_T(h_T,X_T)\middle| X_0 = x\right], \tag{3.2}$$

where $M^{K,y} \in \mathcal{M}_0$ *such that*

$$M_{t+1}^{K,y} := M_t^{K,y} + V_{t+1}^{K,*}(y,X_{t+1}) - \mathbb{E}_t\left[V_{t+1}^{K,*}(y,X_{t+1})\right].$$

Proof. We follow a similar approach to that of Rogers [17]. We have

$$\begin{aligned}
V_0^{K,*}(y,x) &= \sup_{\pi\in\mathcal{P}_{K,y,0}} \mathbb{E}\left[\sum_{s=0}^{T} H_s(h_s,X_s)\middle| X_0 = x\right] \\
&= \sup_{\pi\in\mathcal{P}_{K,y,0}} \mathbb{E}\left[\sum_{s=0}^{T-1}\left(H_s(h_s,X_s) - M_{s+1}^{y_{s+1}} + M_s^{y_{s+1}}\right) + H_T(h_T,X_T)\middle| X_0 = x\right] \\
&\leq \mathbb{E}\left[\left\{\sup_{\pi\in\mathcal{P}_{K,y,0}} \sum_{s=0}^{T-1}\left(H_s(h_s,X_s) - M_{s+1}^{y_{s+1}} + M_s^{y_{s+1}}\right) + H_T(h_T,X_T)\right\}\middle| X_0 = x\right].
\end{aligned}$$

As this holds for all martingales M^w, we then have

$$V_0^{K,*}(y,x) \leq \inf_{M\in\mathbf{M}} \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \sum_{s=0}^{T-1}\left(H_s(h_s,X_s) - M_{s+1}^{y_{s+1}} + M_s^{y_{s+1}}\right) + H_T(h_T,X_T)\middle| X_0 = x\right].$$

To see that the inequality holds the other way around, we consider a particular family of martingales. The one that we take is $\{M_t^{K,y_t}, t \in \mathcal{T}\setminus T\}$ from the Doob decomposition of the value function. Thus, its increments are given by

$$\Delta M_t^{K,y_{t+1}} = M_{t+1}^{K,y_{t+1}} - M_t^{K,y_{t+1}} = V_{t+1}^{K,*}(y_{t+1},X_{t+1}) - \mathbb{E}_t\left[V_{t+1}^{K,*}(y_{t+1},X_{t+1})\right].$$

Using this martingale, we have

$$\begin{aligned}
&\inf_{M\in\mathbf{M}} \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \left\{\sum_{s=0}^{T-1}(H_s(h_s,X_s) - M_{s+1}^{y_{s+1}} + M_s^{y_{s+1}}) + H_T(h_T,X_T)\right\}\middle| X_0 = x\right] \\
&\quad\leq \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \left\{\sum_{s=0}^{T-1}(H_s(h_s,X_s) - \Delta M_s^{K,y_{s+1}}) + H_T(h_T,X_T)\right\}\middle| X_0 = x\right] \\
&\quad= \mathbb{E}\Bigg[\sup_{\pi\in\mathcal{P}_{K,y,0}} \Bigg\{\sum_{s=0}^{T-1}\left(H_s(h_s,X_s) - V_{s+1}^{K,*}(y_{s+1},X_{s+1}) + \mathbb{E}_s\left[V_{s+1}^{K,*}(y_{s+1},X_{s+1})\right]\right) \\
&\qquad\qquad + H_T(h_T,X_t)\Bigg\}\Bigg| X_0 = x\Bigg].
\end{aligned}$$

By the definition of the value function $V_t^{K,*}(\cdot,\cdot)$, for any $(y,x) \in \mathbb{R}^l \times \mathbb{R}^d$, $t \in \mathcal{T}$, and $h \in K_t(y,x)$, we have

$$V_t^{K,*}(y,x) \geq H_t(h,x) + \sup_{\pi\in\mathcal{P}_{K,y-h,t+1}} \mathbb{E}\left[\sum_{s=t+1}^{T} H_s(h_s,X_s)\middle| X_t = x\right],$$

and, therefore,

$$V_t^{K,*}(y,x) \geq H_t(h,x) + \mathbb{E}\left[V_{t+1}^{K,*}(y-h, X_{t+1})|X_t = x\right].$$

Hence,

$$\begin{aligned}
&\inf_{M\in\mathbf{M}} \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \left\{\sum_{s=0}^{T-1}(H_s(h_s,X_s) - M_{s+1}^{y_{s+1}} + M_s^{y_{s+1}}) + H_T(h_T,X_T)\right\}\middle| X_0 = x\right] \\
&\quad\leq \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \left\{\sum_{s=0}^{T-1}\left(V_s^{K,*}(y_s,X_s) - V_{s+1}^{K,*}(y_{s+1},X_{s+1})\right) + H_T(h_T,X_T)\right\}\middle| X_0 = x\right] \\
&\quad= V_0^{K,*}(y,x) + \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \left\{H_T(h_T,X_T) - V_T^{K,*}(h_T,X_T)\right\}\right].
\end{aligned}$$

Now, using the fact that at T we must have $V_T^{K,*}(y,x) = H_T(y,x)$, we have

$$\inf_{M\in\mathbf{M}} \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,0}} \left\{\sum_{s=0}^{T-1}(H_s(h_s,X_s) - M_{s+1}^{y_{s+1}} + M_s^{y_{s+1}}) + H_T(h_T,X_T)\right\}\middle| X_0 = x\right] \leq V_0^{K,*}(y,x)$$

as required. □

Remark 3.2. Consider the specification of the multiple stopping problem in [19]. The payoff function is $H_t(h,x) = hx$ for $t \in \mathcal{T}$, the control satisfies $0 < y_0 \leq T+1$ and takes non-negative integer values, and the constraint sets are defined by

$$K_t(y,x) = K_t(y) = \begin{cases} \{1\} & \text{if } y \geq T-t+1, \\ \{0,1\} & \text{if } T-t+1 > y > 0, \\ \{0\} & \text{if } y = 0. \end{cases}$$

In this special case, the payoff value is either 0 or X_t. Hence, (3.2) simplifies to the following.

$$V_0^{K,*}(y_0,x) = \inf_{M^1,\ldots,M^k\in\mathcal{M}_0} \mathbb{E}\left[\max_{0\leq t_1<\cdots<t_{y_0}\leq T} \sum_{k=1}^{y_0}(X_{t_k} - M_{t_{k+1}}^{y_0-k} + M_{t_k}^{y_0-k})\middle| X_0 = x\right].$$

The dual formulation in this form coincides with the result obtained in [19].

In general, we need to solve the deterministic control problem along the path in order to use this approach. If we have a good approximation to the value function, then we can use the martingale arising from its Doob decomposition, as this will be an approximation to the optimal martingale.

We note that, if we are given a set of approximations to the value function, we can bound the error made in the upper bound arising from the dual formulation in terms of what are essentially the errors in the dynamic programming equations.

More specifically, let $V_t(y,x)$, $t = 0,\ldots,T$ be a set of approximations to the value function and denote by $V_t^{\uparrow}(y,x)$ the associated upper bound on the value function.

Lemma 3.3. *The difference between the a priori estimate for the value function and the associated estimate arising from the dual formulation can be expressed as*

$$V_t^{\uparrow}(y,x) - V_t(y,x) = \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,t}} \sum_{s=t}^{T-1} H_s(h_s,X_s) + \mathbb{E}_s\left[V_{s+1}(y_{s+1},X_{s+1})\right] - V_s(y_s,X_s)\middle| X_t = x\right].$$

Proof. We choose the martingale in the upper bound to be the one generated by $V_t(y,x)$, and, thus,

$$M_{t+1}^y - M_t^y = V_{t+1}(y,X_{t+1}) - \mathbb{E}_t\left[V_{t+1}(y,X_{t+1})\right].$$

Substituting this into the dual formulation given in Theorem 3.1, we have

$$\begin{aligned}
V_t^{\uparrow}(y,x) &= \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,t}} \sum_{s=t}^{T-1} \left\{H_s(h_s,X_s) - M_{s+1}^{y_{s+1}} + M_s^{y_{s+1}}\right\} + H_T(h_T,X_T)\middle| X_t = x\right] \\
&= \mathbb{E}\Bigg[\sup_{\pi\in\mathcal{P}_{K,y,t}} \sum_{s=t}^{T-1} \left\{H_s(h_s,X_s) - V_{s+1}(y_{s+1},X_{s+1}) + \mathbb{E}_s\left[V_{s+1}(y_{s+1},X_{s+1})\right]\right\} \\
&\qquad + H_T(h_T,X_T)\Bigg| X_t = x\Bigg] \\
&= \mathbb{E}\Bigg[\sup_{\pi\in\mathcal{P}_{K,y,t}} \sum_{s=t}^{T-1} \Big\{H_s(h_s,X_s) - V_{s+1}(y_{s+1},X_{s+1}) + V_s(y_s,X_s) \\
&\qquad + \mathbb{E}_s\left[V_{s+1}(y_{s+1},X_{s+1})\right] - V_s(y_s,X_s)\Big\} + H_T(h_T,X_T)\Bigg| X_t = x\Bigg] \\
&= \mathbb{E}\Bigg[\sup_{\pi\in\mathcal{P}_{K,y,t}} \sum_{s=t}^{T-1} \left\{H_s(h_s,X_s) + \mathbb{E}_s\left[V_{s+1}(y_{s+1},X_{s+1})\right] - V_s(y_s,X_s)\right\} \\
&\qquad + V_t(y_t,X_t) - V_T(y_T,X_T) + H_T(h_T,X_T)\Bigg| X_t = x\Bigg] \\
&= V_t(y,x) + \mathbb{E}\left[\sup_{\pi\in\mathcal{P}_{K,y,t}} \sum_{s=t}^{T-1} \left\{H_s(h_s,X_s) + \mathbb{E}_s\left[V_{s+1}(y_{s+1},X_{s+1})\right] - V_s(y_s,X_s)\right\}\middle| X_t = x\right]
\end{aligned}$$

as $V_T(y_T,X_T) = H_T(h_T,X_T)$, giving the required result. □

4 The numerical approach

We now present a numerical implementation of the dual upper bound derived in Lemma 3.3. The lemma gives a representation of the difference between an upper bound $V_0^{\uparrow}(y,x)$ and another (a priori) approximation $V_t(y,x)$ of $V_t^{K,*}(y,x)$. In Section 4.1, we present a numerical method that approximates $V_0^{\uparrow}(y,x)$ given that approximations of the functions $V_t(y,x)$ and

$$(x,y) \mapsto \mathbb{E}\left[V_{t+1}(y,X_{t+1})|X_t = x\right] \tag{4.1}$$

are available.

In Section 4.2, we introduce an approach to generate an a priori estimate $V_t(y,x)$ and an approximation of the conditional expectation (4.1).

4.1 Estimating the dual upper bound

In this section, we assume that a set of a priori approximations $V_t(y,x)$ is available, i.e., for $t=0,\dots,T-1$, the function $V_t(y,x)$ represents an approximation of $V_t^{K,*}(y,x)$. Furthermore, we assume that, for $t=0,\dots,T-1$, the function $V_t(x,y)$ and (an estimate of)

$$(x,y)\mapsto \mathbb{E}\left[V_{t+1}(y,X_{t+1})|X_t=x\right]$$

can be computed for any time-t reachable pair (x,y).

Under such assumptions, we introduce a numerical method that implements the upper estimate $V_0^{\uparrow}(y,x)$ derived in Lemma 3.3. Lemma 3.3 requires the estimation of a path-wise optimum. Hence, given a trajectory $x_\cdot=\{x_0,\dots,x_T\}$, we aim to approximate the function

$$F_t(y,x_\cdot):=\sup_{\pi\in\mathcal{P}_{K,y,t}}\sum_{s=t}^{T-1}\left\{H_s(h_s,x_s)+\mathbb{E}\left[V_{s+1}(y_{s+1},X_{s+1})|X_s=x_s\right]-V_s(y_s,x_s)\right\}$$

recursively for $t=T,T-1,\dots,0$. The optimization algorithm is based on the following path-wise dynamic programming principle.

$$\begin{aligned}
&F_T(y,x_\cdot)=0,\\
&F_t(y,x_\cdot)=\sup_{\pi\in\mathcal{P}_{K,y,t}}\left\{\sum_{s=t}^{T-1}\left\{H_s(h_s,x_s)+\mathbb{E}\left[V_{s+1}(y_{s+1},X_{s+1})|X_s=x_s\right]-V_s(y_s,x_s)\right\}\right\}\\
&\qquad=\sup_{h\in K_t(y,x_t)}\Big\{H_t(h,x_t)+\mathbb{E}\left[V_{t+1}(y-h,X_{t+1})|X_t=x_t\right]\\
&\qquad\qquad-V_t(y,x_t)+F_{t+1}(y-h,x_\cdot)\Big\},
\end{aligned}\tag{4.2}$$

and

$$V_0^{\uparrow}(y,x)=V_0(y,x)+\mathbb{E}\left[F_0(y,X_\cdot)|X_0=x\right].\tag{4.3}$$

Based on (4.2) and (4.3), we are now in a position to formulate the following algorithm.

Algorithm 4.1. *Generate N independent trajectories $x_\cdot^i$, $i=1,\dots,N$ of the process X started at a fixed X_0. For $i=1,\dots,N$*

1. *Set $t=T$, and define $y\mapsto\widehat{F}_T(y,x_\cdot^i)=0$.*
2. *Set $t-1\to t$.*
3. *Define a finite gird $\mathcal{G}_t^y\subseteq Dom(\widehat{F}_t(\cdot,x_\cdot^i))\subseteq\mathbb{R}^l$ (see Remark 4.2), and for each $y\in\mathcal{G}_t^y$ solve the optimization problem*

$$\begin{aligned}
\overline{F}_t(y,x_\cdot^i)=\sup_{\substack{y-h\in Dom(\widehat{F}_{t+1}(\cdot,x_\cdot^i))\\ h\in K_t(y,x_t^i)}}&\Big\{H_t(h,x_t^i)+\mathbb{E}\left[V_{t+1}(y-h,X_{t+1})|X_t=x_t^i\right]\\
&-V_t(y,x_t^i)+\widehat{F}_{t+1}(y-h,x_\cdot^i)\Big\}
\end{aligned}$$

4. *Given the set $\{(y,\overline{F}_t(y,x_\cdot^i))|y\in\mathcal{G}_t^y\}$, define (interpolate) $\widehat{F}_t(\cdot,x_\cdot^i)$ on the whole domain $Dom(\widehat{F}_t(\cdot,x_\cdot^i))$ (see Remark 4.2).*

5. If $t \geq 1$, continue with 2, otherwise finish.

Once $y \mapsto \widehat{F}_0(y, x^i_\cdot)$ is defined for all $i = 1, \ldots, N$, we approximate $V_0^{\uparrow}(y, X_0)$ by the Monte-Carlo average

$$V_0(y, X_0) + \tfrac{1}{N} \sum_{i=1}^{N} \widehat{F}_0(y, x^i_\cdot).$$

Clearly, the main challenge in the implementation of Algorithm 4.1 is the solution of the optimization problem in step 3.

Remark 4.2. The particular implementations of the above algorithm differ in

(i) the specification of the domain $\mathrm{Dom}(\widehat{F}_t(\cdot, x^i_\cdot))$,

(ii) the definition of $\mathcal{G}_t^y$,

(iii) the approximation of the solution to the optimization problem in point 3 of the algorithm,

(iv) and the method applied in point 4 of the algorithm.

Two possible versions of Algorithm 4.1 are presented in Sections 4.1.1 and 4.1.2.

4.1.1 Implementation I: Discretization of the control

One possible approach is to discretized the problem in the control. We define $\mathcal{G}_0^y$ as an (equidistant) grid contained in the set of initial control values of interest. Then, recursively for $t = 0, \ldots, T-1$, we define $\mathcal{G}_{t+1}^y$ to be an (equidistant) grid contained in the set

$$\{y + h | y \in \mathcal{G}_t^y,\ h \in K_t(y, x_t)\}.$$

Furthermore, $\mathrm{Dom}(\widehat{F}_t(\cdot, x^i_\cdot))$ is defined to be the same as $\mathcal{G}_t^y$; this specification implies that the optimization problem in step 3 of Algorithm 4.1 is an optimization over a finite set; moreover, $\overline{F}_t(\cdot, \cdot) = \widehat{F}_t(\cdot, \cdot)$ for $t = 0, \ldots, T$.

Remark 4.3. The choice of $\mathcal{G}_t^y$ depends on the constraints of the problem. For instance, in the case of the gas storage problem, there is a well defined lower and upper limit of y; $\mathcal{G}_t^y$ can be an equidistant grid in this region.

4.1.2 Implementation II: Parametric curve fitting

Here, we define $\mathcal{G}_t^y$ similarly to the previous version. However, we assume $\widehat{F}_t(\cdot, \cdot)$ to be a parametric surface of the following form.

$$\widehat{F}_t(y, x^i_\cdot) = \sum_{r=1}^{R} \lambda^i_{t,r} \phi_r(y)$$

for some vector of parameters $\Lambda^i_t = (\lambda^i_{t,1}, \ldots, \lambda^i_{t,R})$ depending on the trajectory $x^i_\cdot$ and for some set of test functions $(\phi_1, \ldots, \phi_R)$ with domains in $\mathbb{R}^l$, implying

$$\mathrm{Dom}(\widehat{F}_t(\cdot, x^i_\cdot)) = \bigcap_{r=1}^{R} \mathrm{Dom}(\phi_r).$$

The accuracy of the algorithm is sensitive to the choice of test functions; more specifically, different settings may have different optimal sets of test functions, and the (numerical) solution of the optimization problem in step 3 of Algorithm 4.1 should be adapted to the particular choice of test functions.

Point 4 of Algorithm 4.1 is implemented via a least squares regression, i.e., we define Λ_t^i to minimize the expression

$$\sum_{y \in \mathcal{G}_t^y} \left[\overline{F}_t(y, x_\cdot^i) - \sum_{r=1}^{R} \lambda_{t,r}^i \phi_r(y) \right]^2 .$$

Remark 4.4. As we increase the number R of independent test functions and the number N of simulated trajectories, we anticipate that $\widehat{F}_t(\cdot,\cdot)$ converges to $F_t(\cdot,\cdots)$ for $t = 0,\ldots,T$. However, $\widehat{F}_t(\cdot,\cdot)$ is likely to estimate $F_t(\cdot,\cdots)$ from below, and, therefore, our method is likely to result in a low-biased estimate of the dual formulation based upper bound.

4.2 An a priori estimate

As stated at the beginning of Section 4.1, the solution of (4.2) requires computable functions $V_t(y,x)$ and

$$G_t(y,x) := \mathbb{E}\left[V_{t+1}(y, X_{t+1}) | X_t = x\right]$$

approximating $V_t^{K,*}(y,x)$ and $\mathbb{E}\left[V_{t+1}^{K,*}(y, X_{t+1}) | X_t = x\right]$, respectively, for $t = 0,\ldots,T-1$. We suggest the following method, which is based on the dynamic programming formulation.

Definition 4.5. *(Dynamic Programming Formulation)*

$$V_t^{K,*}(y,x) := \begin{cases} \sup_{h \in K_T(y,x)} H_T(h,x) & \text{if } t = T, \\ \sup_{h \in K_t(y,x)} \left\{ H_t(h,x) + \mathbb{E}\left[V_{t+1}^{K,*}(y-h, X_{t+1}) | X_t = x\right] \right\} & \text{if } 0 \le t \le T-1. \end{cases} \tag{4.4}$$

For the computation of the conditional expectation in the above formulation, we introduce a slightly extended version of the standard least squares regression based Monte Carlo method [13, 21, 9]. Our construction yields an a priori estimate $V_0(\cdot,\cdot)$ that approximates $V_0^{K,*}(\cdot,\cdot)$ for a bounded set S of initial X_0 values, where S is contained in the support of the law of X_0.

Algorithm 4.6. *Define a set $\mathcal{G}_0^x$ of distinct initial values of X in S (see Section 4.2.1) and generate independent trajectories $x_\cdot^i$, $i = 1,\ldots,N$, of the process X, with $x_0^i = x$ for each $x \in \mathcal{G}_0^x$. Define $\mathcal{G}_t^x = \{x_t^i \mid 1 \le i \le N\}$ for $t = 1,\ldots,T$, where $N = |\mathcal{G}_0^x|$. Furthermore, for $t = 1,\ldots,T$, define a finite set*

$$\mathcal{G}_t^{yx} \subseteq Dom(V_t(\cdot,\cdot)) \subseteq \mathbb{R}^l \times \mathbb{R}^d$$

such that, for all $(y,x) \in \mathcal{G}_t^{yx}$, we have $x \in \mathcal{G}_t^x$ (see Remark 4.7). Then, proceed as follows.

1. *Set $t = T$ and define*
$$V_T(y,x) = \sup_{h \in K_T(y,x)} H_T(h,x).$$

2. *Set $t-1 \to t$.*

3. *Given the set*
$$\left\{ \left(y, x_t, x_{t+1}, V_{t+1}(y, x_{t+1})\right) \mid (y, x_{t+1}) \in \mathcal{G}_{t+1}^{yx} \right\},$$
define a function $\widehat{G}_t(y,x)$ approximating $G_t(y,x)$ on $Dom(\widehat{G}_t(\cdot,\cdot))$ (see Remark 4.7).

4. *For each $(y,x) \in \mathcal{G}_t^{yx}$, solve the optimization problem*
$$\overline{V}_t(y,x) = \sup_{\substack{y-\hat{y} \in K_t(y,x) \\ (\hat{y},x) \in Dom(\widehat{G}_t(\cdot,\cdot))}} \left\{ H_t(y-\hat{y},x) + \widehat{G}_t(\hat{y},x) \right\}. \tag{4.5}$$

5. *$\overline{V}_t(\cdot,\cdot)$ is only defined on $\mathcal{G}_t^{yx}$. Given the set*
$$\left\{ \left(y, x, \overline{V}_t(y,x)\right) \mid (y,x) \in \mathcal{G}_t^{yx} \right\},$$
define the function $V_t(\cdot,\cdot)$ on its domain (see Remark 4.7).

6. *If $t \geq 1$, then continue with step 2; else, $V_0(y,x)$ results in an a priori approximation.*

The above outline of Algorithm 4.6 leaves some choice as to how certain things are done in detail; in particular, this includes the following points.

Remark 4.7. The particular implementations of Algorithm 4.6 differ in

(i) the construction of the set $\mathcal{G}_t^{yx}$,

(ii) the construction of the function $\widehat{G}_t(\cdot,\cdot)$ in step 4,

(iii) and the construction of the function $V_t(\cdot,\cdot)$ in step 5.

In Section 4.2.1, we implement a particular version of Algorithm 4.6 for the a priori estimate.

4.2.1 Choosing an implementation

Since the three items described in Remark 4.7 are closely connected, we discuss them together.

In a similar way to the classic least squares regression based approach (cf. [13, 21, 9]), we approximate the function $G_t(\cdot,\cdot)$ by an orthogonal projection onto a function space spanned by a set of test functions $\{\psi_1, \ldots, \psi_Q\}$, where, for $q = 1, \ldots, Q$, ψ_q is defined on $\mathrm{Dom}(G_t(\cdot,\cdot)) \subseteq \mathbb{R}^l \times \mathbb{R}^d$, and

$$\widehat{G}_t(y,x) = \sum_{q=1}^{Q} \gamma_{t,q} \psi_q(y,x) \approx G_t(y,x) = \mathbb{E}\left[V_{t+1}(y, X_{t+1}) | X_t = x\right]. \tag{4.6}$$

In contrast to [13, 21, 9], where the orthogonal projection at time t is determined by the distribution of X_t, we have to deal with the control variable as well. We define the projection to minimize

$$\mathbb{E}_{Y,Z}\left[\left(\mathbb{E}[V_{t+1}(Y,X_{t+1})|X_t=Z]-\sum_{q=1}^{Q}\gamma_{t,q}\psi_q(Y,Z)\right)^2\right], \tag{4.7}$$

where Z and Y are independent random variables. In most applications, the set reachable by y_{t+1} is bounded, and, therefore, in order to try to obtain uniform accuracy across the reachable set, we will take Y to be uniformly distributed on this bounded set. The distribution of Z can be defined to coincide with the distribution of X_t. However, in many applications, such as the numerical example in Section 5, only the distribution of X_t conditioned on particular values of X_0 is specified; here, we assume that the law of X_0 is uniform on a certain set S.

Formula (4.7) suggests that, by increasing the number of appropriately chosen test functions $\psi_1, \psi_2, \ldots$, $\widehat{G}_t(\cdot,\cdot)$ approximates the conditional expectation

$$(y,x)\mapsto \mathbb{E}[V_{t+1}(y,X_{t+1})|X_t=x]$$

in the mean square sense with respect to the joint measure of Y and Z; for our particular choice of test functions, see Section 5.2.

In order to determine the regression coefficients $\gamma_{t,q}$ for $q=1,\ldots,Q$, we observe that, when (4.7) is minimized, we have

$$\begin{aligned}
&\frac{\partial}{\partial\gamma_{t,r}}\mathbb{E}_{Y,Z}\left[\left(\mathbb{E}[V_{t+1}(Y,X_{t+1})|X_t=Z]-\sum_{q=1}^{Q}\gamma_{t,q}\psi_q(Y,Z)\right)^2\right]\\
&\quad=2\mathbb{E}_{Y,Z}\big[\mathbb{E}[V_{t+1}(Y,X_{t+1})|X_t=Z]\psi_r(Y,Z)\big]-2\mathbb{E}_{Y,Z}\left[\sum_{q=1}^{Q}\gamma_{t,q}\psi_q(Y,Z)\psi_r(Y,Z)\right]\\
&\quad=2\mathbb{E}_{Y,Z}\big[V_{t+1}(Y,X_{t+1})\psi_r(Y,Z)\big]-2\sum_{q=1}^{Q}\gamma_{t,q}\mathbb{E}_{Y,Z}\left[\psi_q(Y,Z)\psi_r(Y,Z)\right]=0.
\end{aligned}$$

Hence, $\gamma_t=(\gamma_{t,1},\ldots,\gamma_{t,Q})^T$ satisfies the linear equation

$$B_{V,\psi}=B_\psi\gamma_t, \tag{4.8}$$

where

$$B_{V,\psi}=\mathbb{E}_{Y,Z}\big[V_{t+1}(Y,X_{t+1})\psi(Y,Z)\big]$$

and

$$B_\psi=\mathbb{E}_{Y,Z}\big[\psi(Y,Z)\psi(Y,Z)^T\big]$$

for $\psi(x,y)=(\psi_1(x,y),\ldots,\psi_Q(x,y))^T$.

When estimating the regression coefficients, we replace $B_{V,\psi}$ and B_ψ in (4.8) with their Monte-Carlo estimates

$$\widehat{B}_{V,\psi}:=\tfrac{1}{|\mathcal{G}_{t+1}^{yx}|}\sum_{(y,x_{t+1})\in\mathcal{G}_{t+1}^{yx}}V_{t+1}(y,x_{t+1})\psi(y,x_t) \tag{4.9}$$

$$\text{and}\quad \widehat{B}_{\psi}:=\tfrac{1}{|\mathcal{G}_{t+1}^{yx}|}\sum_{(y,x_{t+1})\in\mathcal{G}_{t+1}^{yx}}\psi(y,x_t)\psi(y,x_t)^T. \tag{4.10}$$

The choice of $\mathcal{G}_0^x$ and $\mathcal{G}_t^{yx}$, $t = 0, \ldots, T$, determines how accurately $\widehat{B}_{V,\psi}$ and $\widehat{B}_\psi$ approximate $B_{V,\psi}$ and B_ψ, respectively. When implementing the method, we consider

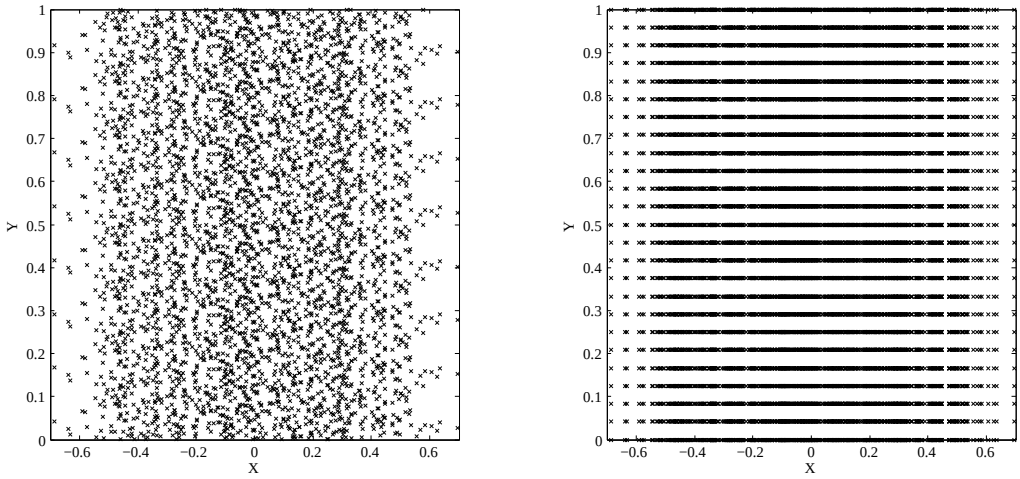
i) $\mathcal{G}_0^x$ to be randomly sampled from the law of X_0, or $\mathcal{G}_0^x$ to be a low discrepancy sequence in S,

ii) x_t to be randomly sampled from the conditional distribution $X_t|X_0$,

iii) and y to be independent of x_t and randomly sampled from the uniform distribution on the support of y_t, or to be a low discrepancy sequence[4] in the support of y_t; we generated a small number (1 to 10) of y items for each x.

Remark 4.8. Initially, we looked at defining $\mathcal{G}_t^{yx} = \mathcal{G}_t^y \times \mathcal{G}_t^x$, for some set $\mathcal{G}_t^y$. However, the numerical results showed that to achieve a given accuracy, a large enough $\mathcal{G}_t^y$ is required, resulting in a set $\mathcal{G}_t^y \times \mathcal{G}_t^x$ significantly larger than the size of $\mathcal{G}_t^{yx}$ constructed in the version described prior to this remark (calibrated to yield the same accuracy).

Figure 1 demonstrates the difference between $\mathcal{G}_t^y \times \mathcal{G}_t^x$ and the set $\mathcal{G}_t^{yx}$ described before this remark. We observe that $\mathcal{G}_t^{yx}$ yields a better coverage with fewer grid points.



(a) $\mathcal{G}_t^{yx}$ grid, y points generated rank-1 lattice rule (4000 points in total, 8 y-items per each x item)

(b) $\mathcal{G}_t^{yx} = \mathcal{G}_t^y \times \mathcal{G}_t^x$ (12500 points in total, 25 y-items per each x item)

Figure 1: Different constructions of $\mathcal{G}_t^{yx}$ based on the same $\mathcal{G}_t^x$, assuming equidistant $\mathcal{G}_0^x \subset [-0.5, 0.5]$ grid and Gaussian conditional distribution $(X_t|X_0)$.

What remains is to be specified are the particulars of step 5 of Algorithm 4.6, i.e., to define $V_t(\cdot, \cdot)$ given

$$\left\{ \left(y, x, \overline{V}_t(y, x)\right) \mid (y, x) \in \mathcal{G}_t^{yx} \right\}.$$

To do this, one can use interpolation, or one can fit a parametric surface to the graph of $\overline{V}_t(\cdot, \cdot)$; we consider the parametric representation

$$V_t(y, x) = \sum_{q=1}^{Q} \beta_{t,q} \psi_q(y, x),$$

[4]We tested rank-1 lattices, see Section 5.

choosing $\beta_{t,q}$, $q = 1, \dots, Q$, to minimize the mean square error

$$\sum_{(y,x)\in\mathcal{G}_t^{yx}} \left(\overline{V}_t(y,x) - \sum_{q=1}^{Q} \beta_{t,q}\psi_q(y,x) \right)^2,$$

i.e., we define $V_t(\cdot,\cdot)$ by another least squares regression.

Remark 4.9. For the numerical computation of the dual formulation based approach, to ensure that we have an upper bound, it is essential that the estimate of

$$V_t(y,x) - \mathbb{E}[V_{t+1}(y, X_{t+1})|X_t = x] \tag{4.11}$$

is a martingale increment, that is it has zero expectation.

As introduced in this section, the function

$$V_t(y,x) - \widehat{G}_t(y,x) \tag{4.12}$$

is a biased estimate of (4.11). When (4.12) is used for computing the dual upper bound - in particular, when $\widehat{G}_t(\cdot,\cdot)$ is a poor estimate of $G_t(\cdot,\cdot)$ - the error due to this bias might be larger than the statistical error.

In such cases, one can replace $\widehat{G}_t(\cdot,\cdot)$ with the following estimate

$$\widetilde{G}_t(y,x) = \tfrac{1}{L} \sum_{i=1}^{L} V_{t+1}(y, X_{t+1}^{(i)})$$

where $\{X_{t+1}^{(i)}, i = 1, \dots, L\}$ for some L is an i.i.d. sample from the conditional distribution of $(X_{t+1}|x)$.

4.2.2 Other choices in the implementation

Multivariate regression similar to (4.6) has been mentioned in [7]. However, [7] does not pursue the same route as presented above, but rather restricts attention to a least squares regression that uses test functions depending only on the underlying factor X, and, for each value of y in a finite set $\mathcal{G}^y \subset \mathbb{R}^l$, a separate simpler regression is computed. The extension of $\overline{V}_t(\cdot,\cdot)$ to the whole domain of $V_t(\cdot,\cdot)$ is not considered (step 5 in Algorithm 4.6). [7] restricts the optimization problem (4.5) in step 5 to $\mathcal{G}^y$.

Computing regressions which are based on test functions depending only on X is less expensive than a regression with high number of (y, X)-dependent test functions (see Section 5.2 for the implementation of the a priori method). However, for accurate estimates, a fine grid $\mathcal{G}^y$ is required, and, hence, a high number of simple regressions needs to be computed. With carefully chosen $\mathcal{G}_t^{yx}$-grid (see Remark 4.8) and a suitable set of (y, X)-dependent test functions, our version attains the same accuracy at significantly lower cost.

4.2.3 A note on low biased methods

The a priori estimates $V_t(y,x)$ and $\mathbb{E}[V_{t+1}(y, X_{t+1})|X_t = x]$, described above, typically result in a high biased estimate of the value function. However, the outcome of the a

priori method can be applied to generate a low biased estimate. By definition, for any $y_0 \in Y_0^K$ and for any policy $\pi \in \mathcal{P}_{K,y_0,0}$, $V_0^{K,\pi}(y_0,\cdot)$ is a low biased estimate of $V_0^{K,*}(y_0,\cdot)$.

The a priori estimate generates a policy as follows. For $t \in \{0,\dots,T-1\}$, a reachable pair (y,x), and $\epsilon > 0$, there exists at least one value $\hat{h}_t \in K_t(y,x)$ that satisfies

$$H_t(\hat{h}_t,x) + \mathbb{E}[V_{t+1}(y-\hat{h}_t, X_{t+1})|X_t = x] \\ \sup_{h \in K_t(y,x)} \{H_t(h,x) + \mathbb{E}[V_{t+1}(y-h,X_{t+1})|X_t = x]\} - \epsilon. \tag{4.13}$$

Given a starting value $y_0 \in Y_0^K$, and assuming that the supremum exists almost surely for reachable pairs, (4.13) determines a exercise policy $\hat{\pi} = (\hat{h}_0,\dots,\hat{h}_T) \in \mathcal{P}_{K,y_0,0}$. As ϵ approaches 0, $V_0^{K,\hat{\pi}}(\cdot,\cdot)$ converges to $V_0^{K,*}(\cdot,\cdot)$. This convergence result motivates the following low-biased algorithm.

Algorithm 4.10. *Fix $y_0 \in Y_0^K$ and $\epsilon > 0$. Generate N independent trajectories $x_\cdot^i$, $i = 1,\dots,N$, of the process X started at a fixed X_0. For $i = 1,\dots,N$,*

1. *set $t = 0$, and $V^i = 0$,*
2. *for $x = x_t^i$ and $y = y_t$, find a $\hat{h}_t$ that satisfies* (4.13)*,*
3. *set $V^i + H_t(\hat{h}_t, x_t^i) \to V^i$.*
4. *and, if $t = T$, then stop; else, set $t+1 \to t$, and continue with step 2.*

Once this routine has been executed for all $i = 1,\dots,N$, the Monte-Carlo average

$$V_0^{\downarrow}(y_0,X_0) := \tfrac{1}{N}\sum_{i=1}^{N} V^i$$

approximates (up to statistical error due to sampling variance) a low biased estimate at time 0 for initial control value y_0 and initial factor value x.

5 Numerical results

In this section, we discuss the gas storage example following [20], and we compare the numerical performance of the implementation of both, the a priori method and the method based on the dual formulation.

The gas storage problem as well as related probabilistic numerical methods have also been discussed in [14] and in [7].

5.1 The gas storage problem

The natural gas storage problem addresses the optimal utilization of certain types of storage facilities. We assume relatively high deliverability and high injection rates. In particular, given the price X_t of gas and the amount y_t of *working gas* in the inventory at time t, we aim to optimize the production (injection) amount for the given day, for each day over a year.

We introduce the following notation.

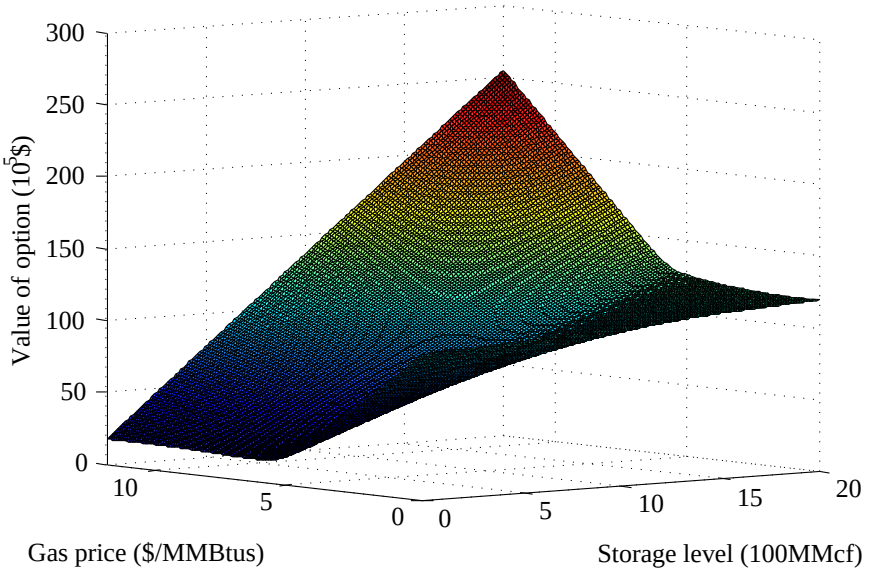


Figure 2: A priori estimate of the value of option at time 0.

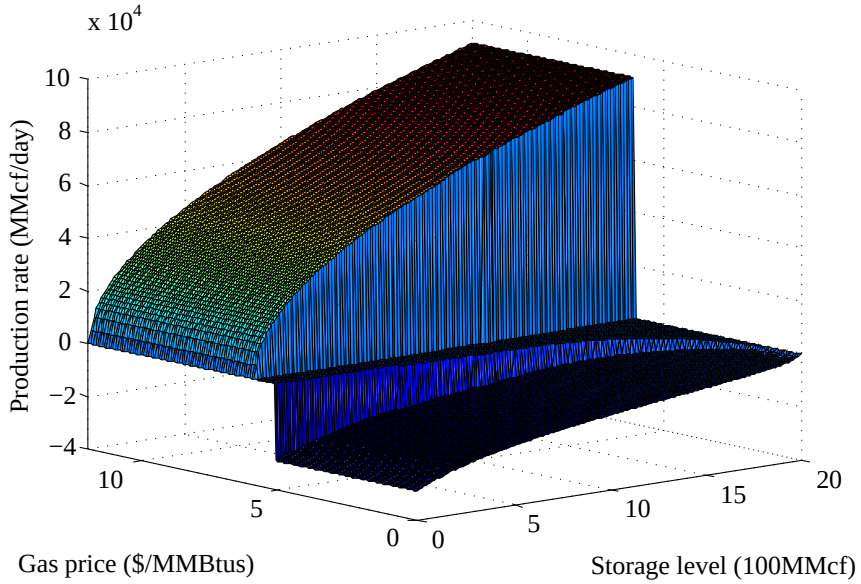


Figure 3: A priori estimate of the optimal rate of production at time 0.

- c, the rate of production if $c > 0$, or the rate of injection if $c < 0$. The rate is measured in million cubic feet per day (MMcf/day).
- y_{base}, base gas requirement (built into the facility and cannot be removed).
- $y_{\max}$, the maximum storage capacity of the facility on top of the base gas level.
- $c_{\max}(y)$, the maximum production rate at storage level y.
- $c_{\min}(y)$, the maximum injection rate at storage level y.
- $a(y, c)$, the rate of gas that is lost given production at rate $c > 0$ or injection at rate $c < 0$.
- r, the discount rate.

As in [20], we consider a facility with working gas capacity of $y_{\max} = 2000$MMcf and with base gas requirement $y_{\text{base}} = 500$MMcf. The maximum production rate (attainable at maximum capacity) is known to be $c_{\max}(y_{\max}) = 250$MMcf/day, whereas the maximum injection rate (attainable at minimum capacity) is $c_{\min}(0) = -80$MMcf/day. The facility is available for one year, and a decision on gas production/injection is made daily, i.e., $\mathcal{T} = \{0, 1, \ldots, 365\}$.

We assume that the loss rate satisfies

$$a(y, c) = a(c) = \begin{cases} 0 & \text{if } c \geq 0, \\ 1.7 & \text{if } c < 0. \end{cases}$$

In the discrete-time formulation[5], we approximate the daily delivered/injected amount by

$$h_t = y_t - y_{t+1} \approx c\Delta t, \tag{5.1}$$

i.e., the unit of time is assumed to be a day (including weekend days), which means $\Delta t = 1$.

The daily constraints on gas production and injection are derived from the ideal gas law and Bernuolli's law (the reader is referred to Section 3 in [20] for details[6]). In particular,

$$c_{\max}(y) = C_0\sqrt{y}, \tag{5.2}$$

where $C_0 = c_{\max}(y_{\max})/\sqrt{y_{\max}}$. Moreover,

$$c_{\min}(y) = -C_1\sqrt{\frac{1}{y + y_{\text{base}}} + C_2}, \tag{5.3}$$

where $C_2 = -1/(y_{\max} + y_{\text{base}})$ and

$$C_1 = c_{\min}(0)/\sqrt{\frac{1}{y_{\text{base}}} + C_2}.$$

[5] In [20], the continuous time production/injection is described by an ordinary differential equation. The discrete-time formulation is an approximation of the solution to that ODE.

[6] Note that, in this paper, the time unit is daily, whereas in [20] the time is measured in years.

Combining (5.2) and (5.3) with (5.1), we get the constraint set for the amount of gas that can be produced/injected during a day:

$$K_t(y,x) = K(y) = [-\min\{c_{\min}(y)\Delta t, y_{\max} - y\}, \min\{c_{\max}(y)\Delta T, y\}]. \tag{5.4}$$

The payoff function is defined by $H_t(\cdot,\cdot) = 0$ for $t = T$, and

$$H_t(h_t, X_t) = \begin{cases} e^{-rt} h_t X_t & \text{if } h \geq 0, \\ e^{-rt}(h_t - a(h_t)\Delta t) X_t & \text{if } h_t < 0, \end{cases} \tag{5.5}$$

for $t = 0, \ldots, T-1$, incorporating the value of the loss of gas at injection.

The discount rate is assumed to be 10%.

In practice, gas prices are quoted in "dollars per million British thermal units" ($/MMBtus). We note that 1000 MMBtus are roughly equivalent to 1 MMcf.

The calculations in [20] are based on the gas price model

$$\mathrm{d}X_t = \alpha(\beta - X_t)\mathrm{d}t + \gamma X_t \mathrm{d}B_t + (J_t - X_t)\mathrm{d}q_t, \tag{5.6}$$

where $t \mapsto q_t$ is a Poisson process with intensity rate λ and independent of the Brownian motion B_t. Moreover, J_t is normally distributed with mean μ and variance σ^2 independent of B_t and q_t. In our implementation, we rescaled the parameters of [20] to daily time-scale: $\alpha = 0.25/365$, $\beta = 2.5$, $\gamma = 0.2/\sqrt{365}$, $\lambda = 2/365$, $\mu = 64$, and $\sigma^2 = 4$.

Remark 5.1. Since the payoff function is piece-wise linear in h and the constraints sets are bounded (uniformly in t) for any K-admissible policy π, the following bounds are satisfied for all $t \in \mathcal{T}$, $x \in \mathbb{R}^+$, $y \in [0, y_{\max}]$.

$$-\infty < (-c_{\min}(0) - a(-1))\Delta t \sum_{s=t}^{T} \mathbb{E}[X_s | X_t = x] \leq V_t^\pi(y,x) \leq c_{\max}(y_{\max})\Delta t \sum_{s=t}^{T} \mathbb{E}[X_s | X_t = x] < \infty.$$

These inequalities imply that the value function is well defined, and the dynamic programming principle holds for this particular formulation of the gas storage problem.

5.2 The a priori estimate

We computed the a priori estimate as follows.

First, we ran the method using an equidistant initial grid $\mathcal{G}_0^x$ in the price region $[0, 12]$ of interest ([20] presents results in this price interval). However, we found that the absolute value of the second derivative of $V_0(\cdot, \dot{)}$ with respect to gas price was large in the price interval $[5, 7]$, and close to zero otherwise; therefore, we decided to refine the grid in the middle region. In particular, we chose an initial grid $\mathcal{G}_0^x$ that had 2500 equidistant points in the interval $[0, 5]$, 5000 equidistant points in $[5, 7]$, and 2500 equidistant points on $[7, 12]$.

The gas price trajectories $x_\cdot^i$ for $i = 1, \ldots, 10000$ were simulated using the Euler time-discretisation

$$x_{t+1}^i = x_t^i + \alpha(\beta - x_t^i)\Delta t + \gamma x_t^i \Delta B_t^i + (J_t^i - x_t^i)\Delta q_t^i,$$

X_0	Range of $V_0^{\uparrow}(y_0, X_0) - V_0^{\downarrow}(y_0, X_0)$	Range of $\text{stdev}(V_0^{\uparrow}(y_0, X_0))$	Range of $\text{stdev}(V_0^{\downarrow}(y_0, X_0))$
3	$[1.224, 3.781]$	$[0.115, 0.121]$	$[0.188, 0.208]$
6	$[1.758, 3.677]$	$[0.116, 0.128]$	$[0.121, 0.133]$
9	$[2.174, 4.276]$	$[0.060, 0.076]$	$[0.115, 0.118]$

Table 1: Comparison of high-biased and low-biased estimates: ranges of differences and ranges of estimated standard deviation over the domain $y_0 \in [0, 20]$ measured in \$/MMBtus.

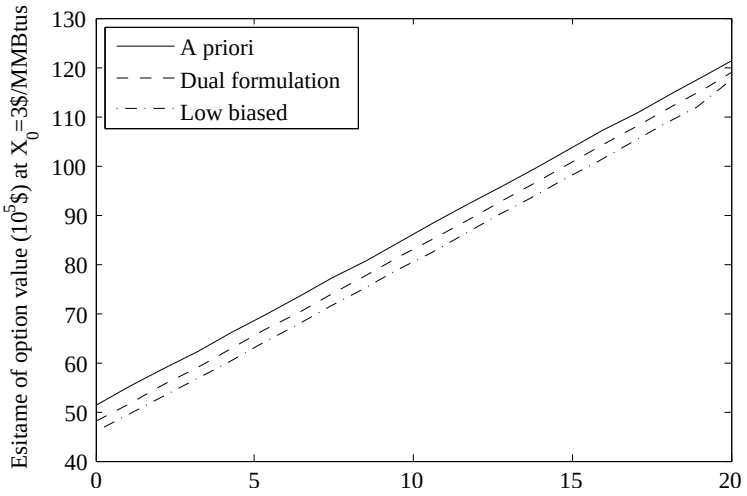


Figure 4: Comparison of three methods, $X_0 = 3$\$/MMBtus.

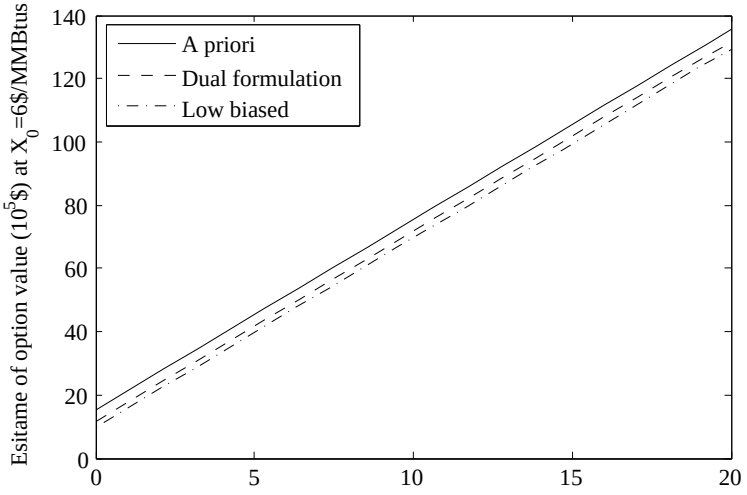


Figure 5: Comparison of three methods, $X_0 = 6$\$/MMBtus.

where ΔB_t^i are independent Brownian increments on a unit time step ($\Delta t = 1$), J_t^i are

drawn from the distribution of J_t, and Δq_t^i drawn from the distribution

$$\Delta q_t^i = \begin{cases} 0 & \text{with probability } 1 - \lambda \Delta t, \\ 1 & \text{with probability } \lambda \Delta t. \end{cases}$$

In order to generate the grid $\mathcal{G}_t^{yx}$, at each time step, we generated a low discrepancy sequence (using a rank 1 lattice rule with random offset, see [11]) of length $|\mathcal{G}_t^x| \times N_y$, and assigned N_y y-points to each of the elements in $\mathcal{G}_t^x$. We tested the method with $N_y = 3, 6, 21$.

Initially, we considered using polynomial test functions for the regression. However, we found that these test functions did not capture well neither the conditional expectation function nor the value function. Therefore, we decided to use test functions that are polynomial on patches and constant outside the patches. We partitioned the (y, x) domain $[0, 20] \times [0, 12]$ into smaller rectangles

$$\begin{array}{cc} [0, 10] \times [0, 5] & [10, 20] \times [0, 5] \\ [0, 10] \times [5, 7] & [10, 20] \times [5, 7] \\ [0, 10] \times [7, 14] & [10, 20] \times [7, 14]. \end{array}$$

On each rectangle, we used the following polynomials: 1, x, y, x^2, y^2, xy, x^2y, y^2x, and x^2y^2. In addition to these polynomials, on the patches in the second row, we also used x^3, x^3y, and x^3y^2. Although defining functions locally on small rectangles leads to a relatively high number of test functions, the matrix is sparse, and the evaluation is tractable.

In step 4 of Algorithm 4.6, we simply compared the outcome of three scenarios: $h = 0$, $h = \min K_t(y, x)$, and $h = \max K_t(y, x)$; i.e., we assumed *bang-bang controls*. We also tested replacing the supremum with the maximum over finer grids in $K_t(y, x)$; however, these tests did not result in significantly different option values.

The numerical results corresponding to $t = 0$, $N_y = 6$, and bang-bang controls are shown in Figures 2 and 3. Comparing these figures to the plots on page 235 in [20], we find that our a priori method slightly overestimates the option value. Given that, in order to estimate the values at time 0, the a priori method uses information from later times, it is likely to be a *high biased* method (see comments on the least squares regression based methods in [11]).

5.3 The dual upper bound

We implemented the version of the method based on the dual formulation as specified in Section 4.1.2 for three different initial gas prices (3$/MMBtus, 6$/MMBtus and 9$/MMBtus). In each case, we generated $N = 10000$ gas price trajectories. For $\mathcal{G}_t^y$, we used a fixed equidistant grid in $[0, 20]$ with $N_y = 320$ points.

For the parametric curve fitting component, we partitioned the control interval $[0, 20]$ into three shorter intervals ($[0, 7]$, $[7, 14]$, and $[14, 20]$), and on each small interval we used the following polynomials as test functions: 1, y, y^2, and y^3.

In order to compute the optimization in step 3 of Algorithm 4.1, we approximated the supremum with the maximum on a finite grid in $K_t(y, x)$. This grid can be chosen to be finer than $\mathcal{G}_t^y$. With other optimization techniques, even more accurate estimates can be computed.

In order to estimate the accuracy of the method, we ran the algorithm using finer $\mathcal{G}_t^y$ grids but the same set of gas price trajectories, more test functions defined locally on finer partitions, and more accurate optimization. Since the refined specifications resulted in absolute differences that were around $10\% - 15\%$ of the standard deviation of the results, we consider the refined estimates numerically equivalent to our reference results.

We also computed low biased estimates following the method described in Section 4.2.3 using the a priori value functions and a sample of 50000 gas price trajectories.

The results are given in Table 1 and plotted in Figures 4, 5, and 6. For each case ($X_0 = 3, 6, 9$), Table 1 describes the range of differences of the high-biased and low-biased estimates over the range of control $y \in [0, 20]$. We also provide the range of estimated standard deviations to indicate the order of magnitude of the statistical error. We note that a conservative upper and lower bound can be computed by adding three times its standard deviation to the upper estimate and subtracting three times its standard deviation from the low-biased estimate.

Figures 4, 5, and 6 suggest that, in some cases, the dual formulation method based estimate results in a sharper upper bound compared to the estimates of the a priori method. The upper and lower estimates are consistent with the numerical results of [20].

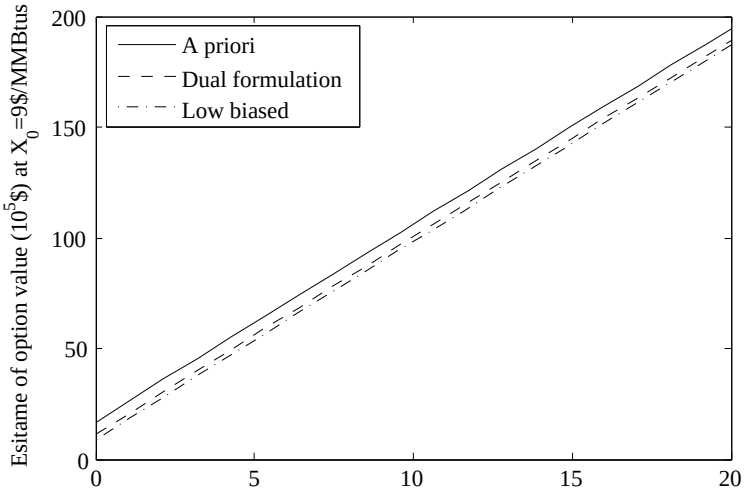


Figure 6: Comparison of three methods, $X_0 = 9\$/\text{MMBtus}$.

References

[1] Aleksandrov, N., and Hambly, B.M. *A dual approach to multiple exercise options under constraints.* Math. Methods Oper. Res. **71** (2010), 503–533.

[2] Aleksandrov, N., and Hambly, B.M. *Liquidity modelling and optimal liquidation in bond markets.* preprint, (2010).

[3] Almgren, R. & Chriss, N., *Optimal execution of portfolio transactions*, Journal of Risk, **3** (2000/2001), 5–39.

[4] Belomestny, D., Kolodko, A. and Schoenmakers, J., *Regression Methods for Stochastic Control Problems and Their Convergence Analysis*, SIAM J. Control Optim. **48** (2009/10), 3562–3588.

[5] Bender, C. *Dual pricing of multi-exercise options under volume constraints*, Finance Stoch. **15** (2011), 1–26.

[6] Bender, C. *Primal and Dual Pricing of Multiple Exercise Options in Continuous Time*, SIAM J. Finan. Math., **2** (2011), 562–586.

[7] Boogert A. and Jong C., *Gas Storage Valuation Using a Monte Carlo Method*, J. Derivatives, **15** (2008), 81–98.

[8] Chen, Z. and Forsyth, P.A., *A semi-Lagrangian approach for natural gas storage valuation and optimal operation*, SIAM J. Sci. Comput., **30** (2007), 339–368.

[9] Clement, E., Lamberton, D. and Protter, P.*An Analysis of a Least Squares regression Method for American Option Pricing*, Finance Stoch. **6** (2002), 449–471.

[10] Davis, M. H. A. and Karatzas, I. *A deterministic approach to optimal stopping* In: Probability, statistics and optimisation, 455–466, Wiley Ser. Probab. Math. Statist., Wiley, Chichester, (1994).

[11] Glasserman, P., *Monte Carlo Methods in Financial Engineering*, Springer, (2003)

[12] Haugh, M. B. and Kogan, L. *Pricing American options: a duality approach.* Oper. Res. 52 (2004), 258–270.

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

[14] Ludovski, M. & Carmona, R., *Valuation of Energy Storage: An Optimal Switching Approach* Quantitative Finance, **10** (2010), 359–374.

[15] Meinshausen, N. & Hambly, B.M., *Monte Carlo methods for the valuation of multiple exercise options*, Math. Finance, **14** (2004), 557–583.

[16] Rogers L.C.G. *Monte Carlo valuation of american options*, Math. Finance, **12** (2002), 271–286.

[17] Rogers, L. C. G., *Pathwise stochastic optimal control*, SIAM J. Control Optim., **46** (2007), 1116–1132.

[18] Schöneborn, T., *Adaptive basket liquidation*, preprint 2011, `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1343985`

[19] Schonmakers, J., *A pure martingale dual for multiple stopping*, Finance Stoch., Published online: 30 November 2010, DOI: 10.1007/s00780-010-0149-1

[20] Thompson, M., Davison, M. and Rasmussen, H., *Natural Gas Storage Valuation and Optimization: A Real Options application*, Naval Research Logistics, **56** (2009), 226–238.

[21] Tsitsiklis J. N. and Van Roy B., *Regression Methods for Pricing Complex American-Style Options*, IEEE Trans. on Neural Networks, **12** (2001), 694–703.