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

Preprint ISSN 2198-5855

Regression based duality approach to optimal control with application to hydro electricity storage

Roland Hildebrand[1], John Schoenmakers[2], Jianing Zhang[3], Fabian Dickmann[2]

submitted: November 8, 2016

[1] LJK, CNRS / Université Grenoble-Alpes
700 avenue Centrale
38041 Domaine Universitaire de Saint Martin d'Hères
France
E-Mail: roland.hildebrand@imag.fr

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: John.Schoenmakers@wias-berlin.de
cfdickmann@googlemail.com

[3] Technische Universität München
Boltzmannstr. 3
85748 Garching
Germany
E-Mail: j.zhang@tum.de

No. 2330
Berlin 2016

2010 *Mathematics Subject Classification.* 49J21, 65C05.

Key words and phrases. Stochastic optimal control, dual martingale method, hydro electricity storage.

R. H. and J. S. acknowledge support by Research Center MATHEON "Mathematics for Key Technologies" in Berlin.

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

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In this paper we consider the problem of optimal control of stochastic processes. We employ the dual martingale method brought forward in [Brown, Smith, and Sun, 2010]. The martingale constituting the solution of the dual problem is determined by linear regression within a Monte-Carlo approach. We apply the solution algorithm to a model of a hydro electricity storage and production system coupled with a model of the electricity wholesale market.

1 Introduction

The dual martingale approach was developed in [8] and [7], see also [4] which in fact contained it in germ. Originally this approach was designed for the optimal stopping problem which arises in the pricing of American options. Desai et al. [5] introduced a *pathwise optimization* method for solving this kind of optimal stopping problems. The dual martingale approach was generalized to general problems of optimal control of stochastic processes in [2].

In this approach, a dual problem is formulated over the space of martingales. A feasible solution of the dual problem yields an upper bound for the solution of the original primal problem. In practice, the optimization is performed over a finite-dimensional subspace of martingales. A sample of paths of the underlying stochastic process is produced by a Monte-Carlo simulation, and the expectation is replaced by the empirical mean.

In this contribution we study the convergence properties of the dual martingale method for general problems of optimal control. We establish conditions for convergence and bounds on the error in the objective value of the dual problem.

We use the obtained convergence results to design a Monte-Carlo based algorithm for computing an approximation of the optimal the dual martingale. This dual martingale is computed recursively backwards in time. In the spirit of previous work [1] the martingale increment is sought as a linear combination of a known and fixed set of basis martingale increments and the coefficients are determined by a linear regression. Note, however, that these coefficients are functions of the adopted control policy and can in general be quite complicated.

We apply the designed algorithm to the control of a model of a hydro electricity storage and production system which is coupled with a model of the electricity market. The system consists of a chain of linearly arranged water reservoirs. Here the topology of the reservoirs is chosen for simplicity and is not a limitation required by the proposed method, which can handle also more complicated networks. Electricity can be produced by releasing water from an upper reservoir to a lower one, and stored by pumping water from a lower reservoir to an upper one. The inflow in the reservoirs as well as the electricity price are coupled stochastic processes which are driven by some Brownian motion. The amount of sold or bought electricity is each day determined by the realization of the price and depends on a piece-wise linear bid curve which is submitted by the producers a day before. The control variables which have to be decided on daily are the amounts of water which is to be released or pumped from or to the reservoirs and the bid curve which is to be submitted for the next day. This model has been considered in [9], which determined an approximation of the optimal strategy based on a large database of past realizations of the inflow and price curves.

In this contribution, we present a Monte-Carlo solution method which implements the dual martingale approach proposed in [2]. We exploit a specific property of the model, namely that both the future expected pay-off and the feasible set at some fixed time t depend only on the current water levels and the submitted bid curve at this time instant. This allows us to limit the complexity of the dependence of the regression coefficients on the implemented past control decisions. All constraints and the pay-off function are linear in the decision variables. This implies that the feasible set of the arising optimization problems is a (random) polyhedron. We show that the expected future payoff is a concave function of the past control decisions and approximate it by a linear support function. This allows us to write all optimization subproblems arising in the recursions as linear programs. The remainder of the paper is organized as follows. In the next section we review the dual martingale method in application to general optimal control problems, as proposed in [2]. In Section 3 we prove our convergence results. In Section 4 we design a recursive Monte-Carlo algorithm based on linear regression. In Section 5 we describe the models of the hydro-electricity storage and production system and the model of the electricity market and the proposed solution algorithm. In the Appendix we collect technical details of the linear programs and provide some background material on sensitivity analysis of the solution of linear programs.

2 Duality for optimal control

In this section we review the dual martingale methods introduced in [2] for computing upper bounds to the considered control problem. First we formalize the considered class of problems.

Consider a probability space $(\Omega, (\mathcal{F}_t)_{t=0,\dots,T}, \mathbb{P})$. Let $(A_i, \mathcal{A}_i)$, $i = 0, ..., T$, be a sequence of measurable control spaces, $\mathbf{A}$ be the canonically measurable product space, i.e., $\mathbf{A} := A_0 \times \cdot \cdot \times A_T$, and $U : \mathbf{A} \times \Omega \to \mathbb{R}$ a reward map satisfying

$$U(\mathbf{a}, \cdot) \text{ is } \mathcal{F}\text{-measurable and} \quad \mathbb{E}_0 \sup_{\mathbf{a}\in\mathbf{A}} |U(\mathbf{a}, \cdot)| < \infty.$$

(with $\mathcal{F} := \mathcal{F}_T$, $\sup :=$ ess.sup, $\mathbb{E}_t := \mathbb{E}_{\mathcal{F}_t}$, $t = 0, ..., T$). For each $t = 0, ..., T$,

$$\mathfrak{p}_t := \{\text{random variables } \alpha : \Omega \longrightarrow A_t,\ \alpha \text{ is } \mathcal{F}_t\text{-measurable}\}.$$

The mapping

$$\boldsymbol{\alpha} : \Omega \longrightarrow (\alpha_0, ..., \alpha_T) \in \mathbf{A}$$

is said to be a policy whenever for each $t = 0, ..., T$, $\alpha_t \in \mathfrak{p}_t$. The set of all policies is denoted by $\mathfrak{P}$, i.e.,

$$\mathfrak{P} = \mathfrak{p}_0 \times \cdots \times \mathfrak{p}_T$$

For convenience we also define $\mathfrak{P}^{(t)} = \mathfrak{p}_t \times \cdots \times \mathfrak{p}_T$. The optimal control problem may now be formulated as

$$V_0^* := \sup_{\boldsymbol{\alpha} \in \mathfrak{P}} \mathbb{E}_0 \, U(\boldsymbol{\alpha}).$$

For any $t = 0, ..., T$, and $a_j \in A_j$, $j = 0, ..., t-1$, we also consider the control problem,

$$V_t^*(\mathbf{a}_{t-1}) := \sup_{\boldsymbol{\alpha}^{(t)} \in \mathfrak{P}^{(t)}} \mathbb{E}_t \, U(\mathbf{a}_{t-1}, \boldsymbol{\alpha}^{(t)})$$

with

$$\boldsymbol{\alpha}^{(t)} := (\alpha_t, ..., \alpha_T) \quad \text{and} \quad \mathbf{a}_t := (a_0, ..., a_t) \in A_0 \times \cdots \times A_t =: \mathbf{A}_t.$$

The Bellman principle thus becomes

$$\begin{aligned} V_T^*(\mathbf{a}_{T-1}) &= \sup_{\boldsymbol{\alpha}^{(T)} \in \mathfrak{P}^{(T)}} U(\mathbf{a}_{T-1}, \alpha_T) = \sup_{a \in A_T} U(\mathbf{a}_{T-1}, a), \\ V_t^*(\mathbf{a}_{t-1}) &= \sup_{\alpha_t \in \mathfrak{p}_t} \sup_{\boldsymbol{\alpha}^{(t+1)} \in \mathfrak{P}^{(t+1)}} \mathbb{E}_t \, \mathbb{E}_{t+1} U(\mathbf{a}_{t-1}, \alpha_t, \boldsymbol{\alpha}_{t+1}) \\ &= \sup_{a_t \in A_t} \mathbb{E}_t \, V_{t+1}^*(\mathbf{a}_{t-1}, a_t) \quad \text{for } t = 0, ..., T-1. \end{aligned}$$

Theorem 1 (Duality) *Let us define for $t = 1, ..., T$, and $\mathbf{a} \in \mathbf{A}$,*

$$\xi_t^*(\mathbf{a}) := \xi_t^*(\mathbf{a}_{t-1}) := V_t^*(\mathbf{a}_{t-1}) - \mathbb{E}_{t-1} \, V_t^*(\mathbf{a}_{t-1}).$$

Then we have

(I) $V_0^* = \sup_{\mathbf{a} \in \mathbf{A}} \left(U(\mathbf{a}) - \sum_{t=1}^T \xi_t^*(\mathbf{a}) \right)$ *a.s.*

Moreover, suppose for any $\mathbf{a} \in \mathbf{A}$ and $t = 1, ..., T$, there is given an $\mathcal{F}_t$-measurable random variable of the form

$$\xi_t(\mathbf{a}) = \xi_t(\mathbf{a}_{t-1}), \quad \textit{such that} \quad \mathbb{E}_{t-1}\xi_t(\mathbf{a}) = \mathbb{E}_{t-1}\xi_t(\mathbf{a}_{t-1}) = 0.$$

Then we have

(II) $V_0^* \leq \mathbb{E}_0 \, \sup_{\mathbf{a} \in \mathbf{A}} \left(U(\mathbf{a}) - \sum_{t=1}^T \xi_t(\mathbf{a}) \right).$

Proof. Note that for any $\boldsymbol{\alpha} \in \mathfrak{P}$, and $t = 1, ..., T$,

$$\mathbb{E}_0 \, \xi_t(\boldsymbol{\alpha}) = \mathbb{E}_0 \, \mathbb{E}_{t-1} \, \xi_t(\alpha_0, ..., \alpha_{t-1}) = 0,$$

so we have

$$\begin{aligned} V_0^* := \sup_{\boldsymbol{\alpha} \in \mathfrak{P}} \mathbb{E}_0 \, U(\boldsymbol{\alpha}) &= \sup_{\boldsymbol{\alpha} \in \mathfrak{P}} \mathbb{E}_0 \left(U(\boldsymbol{\alpha}) - \sum_{t=1}^T \xi_t(\boldsymbol{\alpha}) \right) \\ &\leq \mathbb{E}_0 \sup_{\mathbf{a} \in \mathbf{A}} \left(U(\mathbf{a}) - \sum_{t=1}^T \xi_t(\mathbf{a}) \right), \end{aligned}$$

hence (II) follows. To prove (I), consider

$$
\begin{aligned}
&\sup_{\mathbf{a}\in\mathbf{A}}\left(U(\mathbf{a})-\sum_{t=1}^{T}\xi_t^*(\mathbf{a})\right)\\
&=\sup_{\mathbf{a}\in\mathbf{A}}\left(U(\mathbf{a})-\sum_{t=1}^{T}\left(V_t^*(\mathbf{a}_{t-1})-\mathbb{E}_{t-1}V_t^*(\mathbf{a}_{t-1})\right)\right)\\
&=V_0^*+\sup_{\mathbf{a}\in\mathbf{A}}\left(\underbrace{U(\mathbf{a}_T)-V_T^*(\mathbf{a}_{T-1})}_{\leq 0\text{ due to Bellman Principle}}+\sum_{t=0}^{T-1}\underbrace{\left(\mathbb{E}_t V_{t+1}^*(\mathbf{a}_t)-V_t^*(\mathbf{a}_{t-1})\right)}_{\leq 0\text{ due to Bellman Principle}}\right)\\
&\leq V_0^*.
\end{aligned}
$$

Hence by (II), (I) follows. ∎

3 Characterization and convergence of optimal penalties

For an arbitrary j, $1\leq j\leq T$, and arbitrary $\mathbf{a}\in\mathbf{A}$ consider a system of random variables

$$\xi_j(\mathbf{a}):=\xi_j(\mathbf{a}_{j-1})\in\mathcal{F}_j,\quad\text{such that}\quad\mathbb{E}_{j-1}\xi_j(\mathbf{a})=\mathbb{E}_{j-1}\xi_j(\mathbf{a}_{j-1})=0.$$

Let us further introduce $\mathbf{A}^{(t)}:=A_t\times\cdots\times A_T$, and thus denote

$$\mathbf{a}=(\mathbf{a}_{t-1},\mathbf{a}^{(t)})=(\mathbf{a}_{t-1},a_t,\mathbf{a}^{(t+1)}).$$

Let us define for $t\leq T$

$$\theta_t(\mathbf{a}):=\theta_t(\mathbf{a}_{t-1}):=\sup_{\mathbf{a}^{(t)}\in\mathbf{A}^{(t)}}\left(U(\mathbf{a})-\sum_{j=t+1}^{T}\xi_j(\mathbf{a})\right).\tag{1}$$

So, for $t<T$ we have

$$
\begin{aligned}
\theta_t(\mathbf{a}_{t-1})&=\sup_{a_t\in A_t}\left\{\sup_{\mathbf{a}^{(t+1)}\in\mathbf{A}^{(t+1)}}\left(U(\mathbf{a})-\sum_{j=t+2}^{T}\xi_j(\mathbf{a})\right)-\xi_{t+1}(\mathbf{a})\right\}\\
&=\sup_{a_t\in A_t}\left(\theta_{t+1}(\mathbf{a}_{t-1},a_t)-\xi_{t+1}(\mathbf{a}_{t-1},a_t)\right).
\end{aligned}\tag{2}
$$

Proposition 2 *If for t, $0\leq t\leq T$, the system $\left(\xi_{j+1}(\mathbf{a})\right)_{t\leq j<T}$ is such that*

$$\theta_{j+1}(\mathbf{a}_j)-\xi_{j+1}(\mathbf{a}_j)\ \textit{ is }\ \mathcal{F}_j\textit{-measurable for } t\leq j<T,\tag{3}$$

then we have

$$
\begin{aligned}
&\theta_j(\mathbf{a}_{j-1})=V_j^*(\mathbf{a}_{j-1})\quad\textit{for}\quad t\leq j\leq T,\quad\textit{and}\\
&\xi_{j+1}(\mathbf{a})=V_{j+1}^*(\mathbf{a}_j)-\mathbb{E}_jV_{j+1}^*(\mathbf{a}_j)\quad\textit{for}\quad t\leq j<T.
\end{aligned}\tag{4}
$$

Proof. The assertion follows by induction from the Bellman principle. For $t = T$ the statement (4) is obvious. Suppose the statement holds for $0 < t \leq T$. Assume the system $(\xi_{j+1}(\mathbf{a}))_{t-1 \leq j < T}$ is such that (3) holds for time $t-1$. Then in particular $(\xi_{j+1}(\mathbf{a}))_{t \leq j < T}$ satisfies (3) for time t, and so by the induction hypothesis we have (4). We thus obtain by (2) and (4),

$$\theta_{t-1}(\mathbf{a}_{t-2}) = \sup_{a_{t-1} \in A_{t-1}} \left(V_t^*(\mathbf{a}_{t-2}, a_{t-1}) - \xi_t(\mathbf{a}_{t-2}, a_{t-1}) \right). \tag{5}$$

Due to the assumption for $j = t-1$, $\xi_t(\mathbf{a}) = \theta_t(\mathbf{a}_{t-1}) - \xi_t(\mathbf{a}_{t-1})$ is $\mathcal{F}_{t-1}$ measurable, and due to (4), $\theta_t(\mathbf{a}_{t-1}) = V_t^*(\mathbf{a}_{t-1})$. We so have that $V_t^*(\mathbf{a}_{t-1}) - \xi_t(\mathbf{a}_{t-1})$ is $\mathcal{F}_{t-1}$ measurable, i.e.,

$$\begin{aligned} V_t^*(\mathbf{a}_{t-1}) - \xi_t(\mathbf{a}_{t-1}) &= \mathbb{E}_{t-1}\left[V_t^*(\mathbf{a}_{t-1}) - \xi_t(\mathbf{a}_{t-1})\right] \\ &= \mathbb{E}_{t-1}\left[V_t^*(\mathbf{a}_{t-1})\right], \end{aligned}$$

whence

$$\xi_t(\mathbf{a}) = V_t^*(\mathbf{a}_{t-1}) - \mathbb{E}_{t-1} V_t^*(\mathbf{a}_{t-1}),$$

and from (5),

$$\theta_{t-1}(\mathbf{a}_{t-2}) = \sup_{a_{t-1} \in A_{t-1}} \mathbb{E}_{t-1}\left[V_t^*(\mathbf{a}_{t-2}, a_{t-1})\right] = V_{t-1}^*(\mathbf{a}_{t-2}),$$

so (4) holds for time $t-1$ also. ■

Theorem 3 *Let there be given a system* $\left(\xi^n(\mathbf{a})\right)_{n \in \mathbb{N}} := \left(\xi_t^n(\mathbf{a})\right)_{0 \leq t \leq T}$ *with* $\mathbf{a} \in \mathbf{A}$ *and* $\xi_t^n(\mathbf{a}) = \xi_t^n(\mathbf{a}_{t-1})$, *and with the property*

$$\mathbb{E}_t\, \xi_{t+1}^n(\mathbf{a}) = 0 \quad \textit{for all } 0 \leq t < T,$$

where the sequence $(\xi_t^n(\mathbf{a}))_{n \in \mathbb{N}}$ *is uniformly integrable. Further introduce*

$$\theta_t^n(\mathbf{a}) := \theta_t^n(\mathbf{a}_{t-1}) := \sup_{\mathbf{a}^{(t)} \in \mathbf{A}^{(t)}} \left(U(\mathbf{a}) - \sum_{j=t+1}^{T} \xi_j^n(\mathbf{a}) \right),$$

cf. (1). Assume that for t, $0 \leq t < T$, *the system* $\left(\xi_{j+1}^n(\mathbf{a})\right)_{t \leq j < T}$ *satisfies for each* $\mathbf{a} \in \mathbf{A}$,

$$\textit{Var}_j \left(\theta_{j+1}^n(\mathbf{a}) - \xi_{j+1}^n(\mathbf{a}) \right) \xrightarrow{P} 0 \quad \textit{for all } t \leq j < T \quad \textit{with}$$

We then have for each $\mathbf{a} \in \mathbf{A}$,

$$\theta_j^n(\mathbf{a}) \xrightarrow{L_1} V_j^*(\mathbf{a}_{j-1}) \quad \textit{for} \quad t \leq j \leq T \tag{6}$$

$$\xi_{j+1}^n(\mathbf{a}) \xrightarrow{L_1} V_{j+1}^*(\mathbf{a}_j) - \mathbb{E}_j V_{j+1}^*(\mathbf{a}_j) \quad \textit{for} \quad t \leq j < T, \tag{7}$$

where we assume that ($\mathbf{A}$ is such that) the sequence

$$\theta_j^n(\mathbf{a}) = \sup_{\mathbf{a}^{(j)} \in \mathbf{A}^{(j)}} \Big(U(\mathbf{a}) - \sum_{l=j+1}^{T} \xi_l^n(\mathbf{a}) \Big), \quad n \in \mathbb{N},$$

is uniformly integrable too.

Proof. For $t = T$ it holds

$$\theta_T(\mathbf{a}_{T-1}) = \sup_{a_t \in A_T} U(\mathbf{a}_{T-1}, a_T) = V_T^*(\mathbf{a}_{T-1}),$$

hence (6), and (7) is trivially true. Suppose the theorem is proved for $0 < t \leq T$ and that for each $\mathbf{a} \in \mathbf{A}$,

$$\text{Var}_j \left(\theta_{j+1}^n(\mathbf{a}) - \xi_{j+1}^n(\mathbf{a})\right) \xrightarrow{P} 0 \quad \text{for all } t-1 \leq j < T.$$

The as induction hypothesis (6) and (7) holds for time t. Take an arbitrary $\mathbf{a} \in \mathbf{A}$. Consider for $\varepsilon > 0$ the set

$$A^n(t) := \left\{ |\theta_t^n(\mathbf{a}) - \xi_t^n(\mathbf{a}) - \mathbb{E}_{t-1}\, \theta_t^n(\mathbf{a})| > \varepsilon \right\},$$

By the conditional version of Chebyshev's inequality, we have

$$\mathbb{E}_{t-1}\, 1_{A^n(t)} = P_{t-1}\Big[A^n(t)\Big] \leq \frac{\text{Var}_{t-1}\Big(\theta_t^n(\mathbf{a}) - \xi_t^n(\mathbf{a})\Big)}{\varepsilon^2} \xrightarrow{P} 0,$$

Since $\mathbb{E}_{t-1}\, 1_{A^n(t)}$ $n = 1, 2, ...$ is uniformly bounded, hence uniformly integrable, it follows that

$$1_{A^n(t)} \xrightarrow{L_1} 0, \quad \text{i.e. } P(A^n(t)) \longrightarrow 0$$

that is, $\theta_t^n(\mathbf{a}) - \xi_t^n(\mathbf{a}) - \mathbb{E}_{t-1}\, \theta_t^n(\mathbf{a}) \to 0$ in probability. By induction,

$$\theta_t^n(\mathbf{a}) \xrightarrow{L_1} V_t^*(a_0, \ldots, a_{t-1})$$

This implies

$$\xi_t^n(\mathbf{a}) \xrightarrow{L_1} V_t^*(\mathbf{a}_{t-1}) - \mathbb{E}_{t-1}\, V_t^*(\mathbf{a}_{t-1})$$

since in addition the sequence $(\xi_t^n(\mathbf{a}))_{n=1,2,..}$ is UI by assumption. So (7) is proved for time $t-1$. So we have

$$\theta_t^n(\mathbf{a}_{t-2}, a_{t-1}) - \xi_t^n(\mathbf{a}_{t-2}, a_{t-1}) \xrightarrow{L_1} \mathbb{E}_{t-1}\, V_t^*(\mathbf{a}_{t-1}).$$

Hence it follows that,

$$\begin{aligned} \theta_{t-1}^n(\mathbf{a}_{t-2}) &:= \sup_{a_{t-1} \in A_{t-1}} \left(\theta_t^n(\mathbf{a}_{t-2}, a_{t-1}) - \xi_t^n(\mathbf{a}_{t-2}, a_{t-1})\ \right) \\ &\xrightarrow{L_1} \sup_{a_{t-1} \in A_{t-1}} \mathbb{E}_{t-1}\, V_t^*(\mathbf{a}_{t-2}, a_{t-1}) = V_{t-1}^*(\mathbf{a}_{t-2}), \end{aligned}$$

that is (6). Finally note that $\mathbf{a} \in \mathbf{A}$ was arbitrary. ∎

Theorem 4 *Let a system* $\big(\xi(\mathbf{a})\big)_{\mathbf{a} \in \mathbf{A}}$ *(cf. Proposition 2) satisfy for some* $p > 1$,

$$\sup_{\mathbf{a} \in \mathbf{A}} \left\|\xi_{j+1}(\mathbf{a})\right\|_p < C \text{ for all } 0 \leq j < T. \tag{8}$$

We also assume that ($\mathbf{A}$ is such that) the

$$\sup_{\mathbf{a} \in \mathbf{A}} \left\|\theta_j(\mathbf{a})\right\|_p < C \text{ for all } 0 \leq j \leq T. \tag{9}$$

Assume that for t, $0 \le t < T$, *we have that*

$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\,\textit{Var}_j\left(\theta_{j+1}(\mathbf{a}) - \xi_{j+1}(\mathbf{a})\right) \le \epsilon^2 \ \textit{ for all } \ t \le j < T. \tag{10}$$

We then have that

$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\left|\theta_j(\mathbf{a}_{j-1}) - V_j^*(\mathbf{a}_{j-1})\right| \le K_t\,\epsilon^{\frac{2}{3-2/p}}, \ \textit{ for } \ t \le j \le T, \tag{11}$$

$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\left|\xi_{j+1}^n(\mathbf{a}_j) - V_{j+1}^*(\mathbf{a}_j)\right. \tag{12}$$

$$\left. + \mathbb{E}_j V_{j+1}^*(\mathbf{a}_j)\right| \le F_t\,\epsilon^{1-\frac{1}{3-2/p}}, \ \textit{ for } \ t \le j < T,$$

for some constants $K_t, F_t > 0$ *depending on* t *only.*

Proof. For $t = T$ the statement is obvious (cf. the proof of Theorem 4). Suppose the theorem is proved for $0 < t \le T$ and that

$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\mathrm{Var}_j\left(\theta_{j+1}(\mathbf{a}) - \xi_{j+1}(\mathbf{a})\right) \le \epsilon^2 \ \text{ for all } \ t-1 \le j < T.$$

so (11) and (12) holds for time t. We now prove (6) and (7) for time $t-1$. Take an arbitrary $\mathbf{a} \in \mathbf{A}$. Consider the set

$$A(t) := \left\{ \left|\theta_t(\mathbf{a}) - \xi_t(\mathbf{a}) - \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right| > \epsilon^{1-\frac{1}{3-2/p}} \right\}.$$

It holds that

$$\begin{aligned} \mathbb{E}_{t-1}\left|\theta_t(\mathbf{a}) - \xi_t(\mathbf{a}) - \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right| &= \mathbb{E}_{t-1}\left|\theta_t(\mathbf{a}) - \xi_t(\mathbf{a}) - \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right| 1_{A(t)} \\ &\quad + \mathbb{E}_{t-1}\left|\theta_t(\mathbf{a}) - \xi_t(\mathbf{a}) - \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right| 1_{\Omega\setminus A(t)} \\ &\le \mathbb{E}_{t-1}\left|\theta_t(\mathbf{a}) - \xi_t(\mathbf{a}) - \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right| 1_{A(t)} + \epsilon^{1-\frac{1}{3-2/p}}. \end{aligned}$$

Thus, by Hölder's inequality with $q = (1-1/p)^{-1}$,

$$\begin{aligned} &\mathbb{E}\left|\theta_t(\mathbf{a}) - \xi_t(\mathbf{a}) - \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right| \\ &\le \sqrt[p]{\mathbb{E}\left|\theta_t(\mathbf{a}) - \xi_t(\mathbf{a}) - \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right|^p}\,\sqrt[q]{\mathbb{E}1_{A(t)}} + \epsilon^{1-\frac{1}{3-2/p}} \\ &\le 3C\left(\frac{\mathbb{E}\mathrm{Var}_{t-1}\Big(\theta_t(\mathbf{a}) - \xi_t(\mathbf{a})\Big)}{\epsilon^{2-\frac{2}{3-2/p}}}\right)^{1/q} + \epsilon^{1-\frac{1}{3-2/p}} \\ &\le (3C+1)\,\epsilon^{1-\frac{1}{3-2/p}} \end{aligned} \tag{13}$$

The induction hypothesis implies,

$$\mathbb{E}\left|\theta_t(\mathbf{a}) - V_t^*(a_0,\ldots,a_{t-1})\right| \le K_t\,\epsilon^{1-\frac{1}{3-2/p}},$$

hence, by (13),

$$\begin{aligned} &\mathbb{E}\left|\xi_t(\mathbf{a}_{t-1}) - V_t^*(\mathbf{a}_{t-1}) + \mathbb{E}_{t-1}V_t^*(\mathbf{a}_{t-1})\right| \\ &\le \mathbb{E}\left|\theta_t(\mathbf{a}) - V_t^*(\mathbf{a}_{t-1})\right| + \mathbb{E}\left|\xi_t(\mathbf{a}) - \theta_t(\mathbf{a}) + \mathbb{E}_{t-1}\,\theta_t(\mathbf{a})\right| \\ &+ \mathbb{E}\left|\mathbb{E}_{t-1}\left(V_t^*(\mathbf{a}_{t-1}) - \ \theta_t(\mathbf{a}_{t-1})\right)\right| \le (2K_t + 3C + 1)\ \epsilon^{1-\frac{1}{3-2/p}}. \end{aligned}$$

So since

$$\theta_t(\mathbf{a}_{t-1}) - \xi_t(\mathbf{a}_{t-1}) - \mathbb{E}_{t-1}V_t^*(\mathbf{a}_{t-1})$$
$$= \theta_t(\mathbf{a}_{t-1}) - V_t^*(\mathbf{a}_{t-1}) - (\xi_t(\mathbf{a}_{t-1}) - V_t^*(\mathbf{a}_{t-1}) + \mathbb{E}_{t-1}V_t^*(\mathbf{a}_{t-1}))$$

we have

$$\mathbb{E}\left|\theta_t(\mathbf{a}_{t-1}) - \xi_t(\mathbf{a}_{t-1}) - \mathbb{E}_{t-1}V_t^*(\mathbf{a}_{t-1})\right|$$
$$\leq (3K_t + 3C + 1)\,\epsilon^{1-\frac{1}{3-2/p}} =: F_{t-1}\epsilon^{1-\frac{1}{3-2/p}},$$

i.e. since $\mathbf{a} \in \mathbf{A}$ was arbitrary, we get (12) for time $t-1$. Next we have

$$\mathbb{E}\left|\theta_{t-1}(\mathbf{a}) - V_{t-1}^*(a_0,\ldots,a_{t-2})\right| = \mathbb{E}\left|\theta_{t-1}(\mathbf{a}) - \sup_{a_{t-1}\in A_{t-1}} \mathbb{E}_{t-1}V_t^*(a_0,\ldots,a_{t-1})\right|$$
$$= \mathbb{E}\left|\sup_{a_{t-1}\in A_{t-1}} \left(\theta_t(a_0,\ldots,a_{t-2},a_{t-1}) - \xi_t(a_0,\ldots,a_{t-2},a_{t-1})\,\right)\right.$$
$$\left. - \sup_{a_{t-1}\in A_{t-1}} \mathbb{E}_{t-1}V_t^*(a_0,\ldots,a_{t-1})\right| \leq C_{A_{t-1}}F_{t-1}\epsilon^{1-\frac{1}{3-2/p}} =: K_{t-1}\epsilon^{1-\frac{1}{3-2/p}},$$

where $C_{A_{t-1}}$ is roughly the cardinality of A_{t-1} (because A is nice (finite is enough)). We so have, again since $\mathbf{a} \in \mathbf{A}$ was arbitrary,

$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\left|\theta_{t-1}(\mathbf{a}_{t-2}) - V_{t-1}^*(\mathbf{a}_{t-2})\right| \leq K_{t-1}\epsilon^{1-\frac{1}{3-2/p}},$$

i.e. (11). ■

Theorem 4 can be formulated also as follows.

Corollary 5 *Suppose that the system $\left(\xi^n(\mathbf{a})\right)_{n\in\mathbb{N}}$ is such that (8) and (9) are uniformly satisfied for each $n \in \mathbb{N}$, and that*

$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\,\mathit{Var}_j\left(\theta_{j+1}^n(\mathbf{a}) - \xi_{j+1}^n(\mathbf{a})\right) \leq \frac{D^2}{n^2} \text{ for all } t \leq j < T,\ \ n = 1,2,3,...,$$

where $D > 0$ is constant. We then have for $n = 1,2,3,...$,

$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\left|\theta_j^n(\mathbf{a}_{j-1}) - V_j^*(\mathbf{a}_{j-1})\right| \leq K_t\left(\frac{D}{n}\right)^{1-\frac{1}{3-2/p}}, \text{ for } t \leq j \leq T,$$
$$\sup_{\mathbf{a}\in\mathbf{A}} \mathbb{E}\left|\xi_{j+1}^n(\mathbf{a}_j) - V_{j+1}^*(\mathbf{a}_j) + \mathbb{E}_j V_{j+1}^*(\mathbf{a}_j)\right| \leq F_t\left(\frac{D}{n}\right)^{1-\frac{1}{3-2/p}}, \text{ for } t \leq j < T,$$

for some constants $K_t, F_t > 0$ depending on t only.

Theorem 3, Theorem 4, and Corollary 5 are important in practical situations, for instance, for (possibly high dimensional) underlyings of jump-diffusion type in a Lévy-Itô setup. In this environment we may construct a class of uniformly integrable martingale increments. The exercise dates $0 < s_1 < \cdots < s_T$ are for notational convenience identified with the index numbers $0 < 1 < \cdots < T$.

Let W be an m-dimensional Brownian motion and let N denote a Poisson random measure in $\mathbb{R}^q$, independent of W, with (deterministic) compensator measure $\nu(s, du)ds$ such that

$$\int_0^s \int_{\mathbb{R}^q} (|u|^2 \wedge |u|)\nu(s, du)ds < \infty, \quad s \geq 0.$$

Let $(\mathcal{F}_t)_{0\leq t\leq T}$ be the filtration generated by W and N, augmented by null sets. Now let X be a D-dimensional Markov process, adapted to $(\mathcal{F}_t)$, and consider the mappings $c : [0, T] \times \mathbb{R}^D \to \mathbb{R}_{\geq 0}$ and $d : [0, T] \times \mathbb{R}^D \times \mathbb{R}^q \to \mathbb{R}_{\geq 0}$ satisfying

$$\mathbb{E} \int_0^T |c(s, X_s)|^2 ds + E \int_0^T \int_{\mathbb{R}^q} |d(s, X_s, u)|^2 \nu(s, du)ds < \infty.$$

We define a class of uniformly integrable elementary martingale increments, $\mathcal{M}^{UI}$, as the set of all martingale increments

$$\mathfrak{M} := \left(\mathfrak{m}_{j+1}\left(\varphi^c, \varphi^d\right) : 0 \leq j < T\right)$$

defined by

$$\mathfrak{m}_{j+1}\left(\varphi^c, \varphi^d\right) = \int_{s_j}^{s_{j+1}} \varphi^c(s, X_s) \cdot dW_s + \int_{s_j}^{s_{j+1}} \int_{\mathbb{R}^q} \varphi^d(s, X_s, u) \cdot \tilde{N}(ds, du), \quad \text{where} \tag{14}$$

$$\varphi^c : [0, T] \times \mathbb{R}^D \to \mathbb{R}^m \quad \text{and} \quad \varphi^d : [0, T] \times \mathbb{R}^D \times \mathbb{R}^q \to \mathbb{R}^q \quad \text{with}$$
$$|\varphi^c| \leq c \quad \text{and} \quad \left|\varphi^d\right| \leq d,$$

with "$\cdot$" denoting scalar product, and where $\tilde{N} = N - \nu$ is the compensated Poisson measure. Note that we have indeed

$$\mathbb{E}_j \mathfrak{m}_{j+1}\left(\varphi^c, \varphi^d\right) = 0.$$

The quadratic variation of $\mathfrak{m}_{j+1}\left(\varphi^c, \varphi^d\right)$ satisfies with the help of the BDG inequalities (see, e.g., [10, Theorem 48])

$$\begin{aligned}
\mathbb{E}|\mathfrak{m}_{j+1}\left(\varphi^c, \varphi^d\right)|^2 &\leq \mathbb{E} \sup_{s_j\leq t\leq s_{j+1}} \Big| \int_{s_j}^t \varphi^c(s, X_s) \cdot dW_s + \int_{s_j}^t \int_{\mathbb{R}^q} \varphi^d(s, X_s, u) \cdot \tilde{N}(ds, du)\Big|^2 \\
&\leq C\Big(\mathbb{E} \int_{t_j}^{t_{j+1}} |\varphi^c(s, X_s)|^2 ds + \mathbb{E} \int_{t_j}^{t_{j+1}} \int_{\mathbb{R}^q} |\varphi^d(s, X_s, u)|^2 \nu(s, du)ds\Big) \\
&\leq C\Big(\mathbb{E} \int_{t_j}^{t_{j+1}} |c(s, X_s)|^2 ds + \mathbb{E} \int_{t_j}^{t_{j+1}} \int_{\mathbb{R}^q} |d(s, X_s, u)|^2 \nu(s, du)ds\Big) < C,
\end{aligned}$$

where $C > 0$ denotes a generic constant which may vary from line to line. We then have for every $0 \leq j < T$

$$\sup_{\mathfrak{m} \in \mathfrak{M}} \mathbb{E}|\mathfrak{m}|^2 < C < \infty,$$

Finally, an application of the de la Vallée Poussin criterion yields that $\mathfrak{M}$ is indeed a family of uniformly integrable martingales.

4 Monte Carlo based regression algorithm for an additive structure

Suppose that we have a reward functional of the form

$$U(\mathbf{a}) = \sum_{t=0}^{T} C_t(X_t, \mathbf{a}_t, \omega),$$

where for each x and $\mathbf{a}_t$ the random variable $C_t(x, \mathbf{a}_t, \cdot)$ is assumed to be $\mathcal{F}_t$-measurable and $C_t(\cdot, \cdot, \cdot)$ is assumed to be bounded from above. We assume that the random argument in C_t stems from certain random constraints such that $C_t(X_t, \mathbf{a}_t, \omega) = -\infty$ for $\mathbf{a}_t$ such that the constraint is violated. More specifically, we assume that

$$C_t(X_t, \mathbf{a}_t, \omega) = C^{(0)}(X_t, \mathbf{a}_t) 1_{\mathbf{a}_t \in \mathcal{C}_t} - \infty \cdot 1_{\mathbf{a}_t \notin \mathcal{C}_t}$$

for $\mathcal{F}_t$-measurable random sets $\mathcal{C}_t$. Then we have

$$\begin{aligned} V_T(\mathbf{a}_{T-1}) &= \sup_{a_T \in A_T} U(\mathbf{a}_{T-1}, a_T) \\ &= \sum_{t=0}^{T-1} C_t(X_t, \mathbf{a}_t) + \mathcal{V}_T(\mathbf{a}_{T-1}) \end{aligned}$$

with

$$\mathcal{V}_T(\mathbf{a}_{T-1}) := \sup_{a_T \in A_T} C_T(X_T, (\mathbf{a}_{T-1}, a_T)).$$

For $0 \leq j < T$ we have generically,

$$\begin{aligned} V_j(\mathbf{a}_{j-1}) &= \sum_{t=0}^{j-1} C_t(X_t, \mathbf{a}_t) + \mathcal{V}_j(\mathbf{a}_{j-1}), \quad \text{where} \\ \mathcal{V}_j(\mathbf{a}_{j-1}) &= \sup_{a_j \in A_j} \left[C_j(X_j, (\mathbf{a}_{j-1}, a_j)) + \mathbb{E}_j \mathcal{V}_{j+1}(\mathbf{a}_j) \right]. \end{aligned} \tag{15}$$

The relation (15) may be considered as the Bellman principle in terms of the $\mathcal{V}_j$. Now let us construct in a Wiener environment generated by an m-dimensional Wiener process a backward regression algorithm for $t = T, ..., 0$. The extension to a jump-diffusion setting will be straightforward and therefore omitted.

Algorithm:

Step 1: We simulate a Monte Carlo ensemble of trajectories $(X_t^{(n)}, 0 \leq t \leq T, n = 1, ..., N)$ ($X_t^{(n)} := X_{s_t}^{(n)}$, $\mathcal{F}_t := \mathcal{F}_{s_t}$ etc.) and initialize for $n = 1, ..., N$, at time (index) T,

$$\vartheta_T^{(n)}(\mathbf{a}_{T-1}) := \mathcal{V}_T^{(n)}(\mathbf{a}_{T-1}) = \sup_{a_T \in A_T} C_T(X_T^{(n)}, \mathbf{a}_{T-1}, a_T), \quad \text{that corresponds to}$$

$$\theta_T^{(n)}(\mathbf{a}_{T-1}) := \vartheta_T^{(n)}(\mathbf{a}_{T-1}) + \sum_{r=0}^{T-1} C_r(X_r^{(n)}, \mathbf{a}_r).$$

In view of (3), at $t = T$ all objects are trivially $\mathcal{F}_T$-measurable so there is nothing to do further. Suppose that for $t,\ 0 < t \leq T$, estimations $\widehat{\vartheta}_j^{(n)}(\mathbf{a}_{j-1}),\ t \leq j \leq T$, and “martingale increments” $\widehat{\xi}_{j+1}^{(n)}(\mathbf{a}),\ t \leq j < T,\ n = 1, ..., N$, are constructed, such that for $t \leq j < T$, (3), the expected $\mathcal{F}_j$-conditional variance of

$$\widehat{\theta}_{j+1}^{(n)}(\mathbf{a}_j) - \widehat{\xi}_{j+1}^{(n)}(\mathbf{a}_j) = \widehat{\vartheta}_{j+1}^{(n)}(\mathbf{a}_j) + \sum_{r=0}^{j} C_r(X_r^{(n)}, \mathbf{a}_r) - \widehat{\xi}_{j+1}^{(n)}(\mathbf{a}_j)$$

is “small”. Since the $\mathcal{F}_j$-conditional variance of the sample $\sum_{r=0}^{j} C_r(X_r^{(n)}, \mathbf{a}_r)$ is trivially zero, this is equivalent with saying that the $\mathcal{F}_j$-conditional variance of

$$\widehat{\vartheta}_{j+1}^{(n)}(\mathbf{a}_j) - \widehat{\xi}_{j+1}^{(n)}(\mathbf{a}_j)$$

is “small”. We then construct a next martingale increment $\widehat{\xi}_t(\mathbf{a})$, such that the expected $\mathcal{F}_{t-1}$-conditional variance of

$$\widehat{\theta}_t^{(n)}(\mathbf{a}_{t-1}) - \widehat{\xi}_t^{(n)}(\mathbf{a}_{t-1}) \text{ is “small”, and } \mathbb{E}_{t-1}\widehat{\xi}_t(\mathbf{a}) = 0,$$

in the following way. For this it is enough to make sure that $\mathcal{F}_{t-1}$-conditional variance of

$$\widehat{\vartheta}_t^{(n)}(\mathbf{a}_{t-1}) - \widehat{\xi}_t^{(n)}(\mathbf{a}_{t-1}) \text{ is “small”..}$$

We consider two systems of basis functions $(\psi_k(t,x))_{1 \leq k \leq K}$, $(\varphi_k(t,x))_{1 \leq k \leq K}$ (that may coincide in principle) and carry out the linear regression procedure

$$\begin{aligned}
&(\widehat{\gamma}_{t-1}(\mathbf{a}_{t-1}), \widehat{\beta}_{t-1}(\mathbf{a}_{t-1})) := \\
&\arg\min_{\gamma \in \mathbb{R}^K, \beta \in \mathbb{R}^{K \times m}} \sum_{n=1}^{N} \left[\widehat{\vartheta}_t^{(n)}(\mathbf{a}_{t-1}) - \sum_{k=1}^{K} \gamma^k \psi_k(s_{t-1}, X_{t-1}^{(n)}) - \sum_{k=1}^{K} \sum_{l=1}^{m} \beta^{k,l} \mathfrak{m}_t^{l,k,(n)} \right]^2 = \\
&\arg\min_{\gamma \in \mathbb{R}^K, \beta \in \mathbb{R}^{K \times m}} \sum_{n=1}^{N} \left[\widehat{\vartheta}_t^{(n)}(\mathbf{a}_{t-1}) - \gamma \cdot \psi(s_{t-1}, X_{t-1}^{(n)}) - \mathrm{Tr}\left[\beta \mathfrak{m}_t^{(n)}\right] \right]^2,
\end{aligned}$$

where the martingale increments $\mathfrak{m}_t := \left[\mathfrak{m}_t^{l,k}\right]_{l=1,...,m,\, k=1,...,K}$ are defined by

$$\mathfrak{m}_t^{l,k} := \int_{s_{t-1}}^{s_t} \varphi_k(s, X_s) dW_s^l, \quad l = 1, ..., m,\ k = 1, ..., K.$$

We then set on each trajectory $n = 1, ..., N,$

$$\widehat{\xi}_t^{(n)}(\mathbf{a}_{t-1}) := \widehat{\xi}_t^{(n)}(\mathbf{a}_{t-2}, a_{t-1}) := \mathsf{Tr}\left[\widehat{\beta}_{t-1}(\mathbf{a}_{t-2}, a_{t-1})\mathfrak{m}_t^{(n)}\right]$$

and then set

$$\widehat{\vartheta}_{t-1}^{(n)}(\mathbf{a}_{t-2}) := \sup_{a_{t-1}\in A_{t-1}} \left(\widehat{\vartheta}_t^{(n)}(\mathbf{a}_{t-2}, a_{t-1}) + C_{t-1}(X_{t-1}^{(n)}, (\mathbf{a}_{t-2}, a_{t-1})) - \widehat{\xi}_t^{(n)}(\mathbf{a}_{t-2}, a_{t-1})\right)$$

that corresponds to

$$\widehat{\theta}_{t-1}^{(n)}(\mathbf{a}_{t-2}) := \widehat{\vartheta}_{t-1}^{(n)}(\mathbf{a}_{t-2}) + \sum_{r=0}^{t-2} C(X_r^{(n)}, \mathbf{a}_r)$$

which then has “nearly” $\mathcal{F}_{t-1}$-conditional variance zero. We so proceed all the way down to $t = 0$ and obtain a system satisfying Proposition 2 in an approximate sense. In particular we end up with a system of coefficients

$$(\widehat{\gamma}_{t-1}(\mathbf{a}_{t-1}), \widehat{\beta}_{t-1}(\mathbf{a}_{t-1})) \in \mathbb{R}^K \times \mathbb{R}^{K\times m}, \quad 1 \leq t \leq T.$$

Step 2: We simulate a new Monte Carlo ensemble of trajectories $(\widetilde{X}_t^{(\widetilde{n})}, 0 \leq t \leq T, \widetilde{n} = 1, ..., \widetilde{N})$ and construct the martingale increment samples

$$\widetilde{\xi}_t^{(\widetilde{n})}(\mathbf{a}_{t-1}) = \mathsf{Tr}\left[\widehat{\beta}_{t-1}(\mathbf{a}_{t-1})\widetilde{\mathfrak{m}}_t^{(\widetilde{n})}\right], \quad \mathbf{a}_{t-1} \in \mathbf{A}_{t-1}, \quad 1 \leq t \leq T, \quad \widetilde{n} = 1, ..., \widetilde{N},$$

with

$$\widetilde{\mathfrak{m}}_t^{l,k(\widetilde{n})} := \int_{s_{t-1}}^{s_t} \varphi_k(s, X_s^{(\widetilde{n})}) dW_s^{l,(\widetilde{n})}.$$

As an upper biased estimate we then construct

$$\widehat{V}_0^{up} := \frac{1}{\widetilde{N}} \sum_{\widetilde{n}=1}^{\widetilde{N}} \max_{\mathbf{a}\in\mathbf{A}} \left(U^{(\widetilde{n})}(\mathbf{a}) - \sum_{t=1}^{T} \widetilde{\xi}_t^{(\widetilde{n})}(\mathbf{a})\right).$$

Step 3: Note that for $0 \leq t < T$ we have the approximation

$$\begin{aligned}\mathbb{E}_t V_{t+1}(\mathbf{a}_{t-1}, a_t) &= \sum_{r=0}^{t-1} C_r(X_r, \mathbf{a}_r) + \mathbb{E}_t \mathcal{V}_{t+1}(\mathbf{a}_{t-1}, a_t) + C_t(X_t, \mathbf{a}_{t-1}, a_t) \\ &\approx \sum_{r=0}^{t-1} C_r(X_r, \mathbf{a}_r) + \widehat{\gamma}_t(\mathbf{a}_{t-1}, a_t) \cdot \psi(s_t, X_t) + C_t(X_t, \mathbf{a}_{t-1}, a_t).\end{aligned}$$

So, construct recursively on each trajectory

$$\alpha_t^{(\widetilde{n})} := \arg\max_{a_t\in A_t} \begin{pmatrix} \widehat{\gamma}_t(\alpha_0^{(\widetilde{n})}, ..., \alpha_{t-1}^{(\widetilde{n})}, a_t) \cdot \psi(s_t, X_t^{(\widetilde{n})}) \\ +C_t(X_t, \alpha_0^{(\widetilde{n})}, ..., \alpha_{t-1}^{(\widetilde{n})}, a_t) \end{pmatrix}, \quad t = 0, ..., T,$$

and then with $\boldsymbol{\alpha}^{(\widetilde{n})} := \left(\alpha_0^{(\widetilde{n})}, ..., \alpha_T^{(\widetilde{n})}\right)$, the lower biased estimate

$$\widehat{V}_0^{low} := \frac{1}{\widetilde{N}} \sum_{\widetilde{n}=1}^{\widetilde{N}} U^{(\widetilde{n})}(\boldsymbol{\alpha}^{(\widetilde{n})}).$$

5 Model of hydro-electricity storage system

Our numerical experiments are carried out on an example from optimizing operations for hydro storage systems as described in [9]. We consider a power generating company which runs a network of hydro storage plants and participates in the trade on a wholesale electricity market. The aim of the power generating company is to maximize its expected profits from trading electricity by efficiently operating its hydro storage system.

We consider the following stylized model. Let $T \in \mathbb{N}$ denote the maturity day of the planning horizon and let $t \in \{0, 1, \dots, T\}$ denote a day within the planning period. Let $(X_t)_{t=0,\dots,T}$ be a process which models the dynamics of environmental factors and let $(P_t)_{t=0,\dots,T}$ be a process which models the evolution of electricity prices. We assume P_t to be adapted to the σ-field $\mathcal{F}_t = \sigma(X_u : u \le t)$ for $t = 0, \dots, T$. At any stage $0 \le t < T$, the power generating company submits a bidding curve $b_t(\cdot)$. That is, the company commits itself at time t to deliver a volume $b_t(P_{t+1})$ at time $t+1$. A positive volume means that the company is going to sell electricity, and a negative value means that the company is ready to buy it. Since any continuous function can be approximated by a piece-wise linear function with an arbitrary accuracy, the bidding curve will be taken to be piece-wise constant and parametrized in the following way. For a fixed grid of linearly ordered prices ρ_l that are assumed to be positive, i.e.

$$0 < \rho_l < \rho_{l+1}, \quad l = 1, \dots, L-1,$$

the piece-wise linear curve b_t is determined by

$$b_t(P_{t+1}) = \begin{cases} b_t^1, & \text{if } P_{t+1} < \rho_1 \\ b_t^{l-1} + \frac{b_t^l - b_t^{l-1}}{\rho_l - \rho_{l-1}}(P_{t+1} - \rho_{l-1}), & \text{if } \rho_{l-1} \le P_{t+1} \le \rho_l, \\ b_t^L, & \text{if } P_{t+1} \ge \rho_L \end{cases} \tag{16}$$

where b_t^l denotes the volume to deliver at the following day $t+1$, associated to the grid price ρ_l. We further impose the monotonicity constraint

$$b_t^l \le b_t^{l+1}, \quad l = 1, \dots, L-1. \tag{17}$$

To guarantee generation of the volume x_t at time t, $1 \le t \le T$, the company can operate its hydro storage system. Let there be $J \in \mathbb{N}$ linearly arranged reservoirs and let $c_{t,j}$ denote the charge and $d_{t,j}$ the discharge decision into and from reservoir $j \in \{1, \dots, J\}$. The reservoirs are assumed to be linearly upwards ordered, that is, reservoir j can only discharge to reservoir $j-1$ and charge to reservoir $j+1$. Let the total natural inflow (rain for example) on day t be given by $\zeta_j I_t$ in reservoir j, where $\zeta_j > 0$ are fixed coefficients. It is assumed that reservoirs have a maximum water capacity U_j^R. The storage state of reservoir $j \in \{1, \dots, J\}$ after the operations of day t is thus given by

$$R_t^j = R_{t-1}^j - d_{t,j} + d_{t,j+1} + c_{t,j} - c_{t,j+1} + \zeta_j I_t + o_{t,j+1} - o_{t,j}, \tag{18}$$
$$\text{with} \quad o_{t,J+1} = d_{t,J+1} = c_{t,J+1} = c_{t,1} = 0, \;\; c_{t,j}, d_{t,j}, o_{t,j}, R_t^j \ge 0$$

for $t \in \{1, \ldots, T\}$, where $o_{t,j}$ is the overspill of reservoir j. The net trading volume at time t has to satisfy

$$b_{t-1}(P_t) = \sum_{j=1}^{J} \left(\eta_j^d d_{t,j} - \eta_j^c c_{t,j}\right), \tag{19}$$

where η_j^d, η_j^c are constant conversion factors for the turbine and the pump attached to the j-th reservoir and $\xi_j > \eta_j$. In order to guarantee that this volume can be realized at time t for sure, regardless the inflow I_t, the bidding curve submitted at time $t-1$ need to be such that

$$\left[b_{t-1}^1, b_{t-1}^L\right] \subset \text{Range}\left\{\sum_{j=1}^{J} \eta_j^d d_j - \eta_j^c c_j : \begin{array}{l} R^j = R_{t-1}^j - d_j + d_{j+1} + c_j - c_{j+1} + o_{j+1} - o_j, \\ o_{J+1} = d_{J+1} = c_{J+1} = c_1 = 0, \;\; c_j, d_j, o_j, R^j \geq 0 \end{array}\right\}. \tag{20}$$

For the range (20) of the amount of electricity that can be sold or bought at time t we hence assume the worst case $I_t = 0$. Denote further by U_j^c and U_j^d the maximum pump and turbine capacities, respectively. We thus have to satisfy the additional constraints

$$\begin{aligned} R_t^j &\leq U_j^R, \\ d_{t,j} &\leq U_j^d, \\ c_{t,j} &\leq U_j^c. \end{aligned} \tag{21}$$

The reservoir levels R_0^j at $t = 0$ are assumed to be known. The goal of the power generating company is to maximize the expected profits from trading electricity under constraints (16),(17),(18),(19),(20),(21). This optimization problem is equivalent with a recursive Bellman principle of the form (15). Let us denote the controls by

$$\boldsymbol{\pi}_t := (\pi_0, \ldots, \pi_t) \quad \text{with} \quad \pi_t := (c_t, d_t, o_t, b_t), \quad t = 1, \ldots, T-1, \; \pi_0 = (b_0), \; \pi_T = (c_T, d_T, o_T).$$

Note that at time T the bidding curve b_T is void since in the present setup there is no electricity to deliver at time $T+1$. Further, $\kappa_t := (\kappa_{t,1}, \ldots, \kappa_{t,J})$ for $\kappa \in \{c, d, o\}$ and $b_t = (b_t^1, \ldots, b_t^L)$, and let us denote the constraints by

$$\boldsymbol{\Pi}_t := \left\{\boldsymbol{\pi}_t : \text{for } 0 \leq s \leq t, \; \pi_s \text{ satisfies (16),(17),(18),(19),(20),(21) with } t \text{ replaced by } s\right\}.$$

By construction we have for a generic state t that

$$\boldsymbol{\pi}_{t-1} \in \boldsymbol{\Pi}_{t-1} \Longrightarrow \mathcal{V}_t\left(\boldsymbol{\pi}_{t-1}\right) > -\infty, \quad 1 \leq t \leq T.$$

For a clear exposition we will carry out the initialization and the first backward step in detail. At time T we have in view of (16), (19) (with slight abuse of notation),

$$\mathcal{V}_T = \mathcal{V}_T\left(\boldsymbol{\pi}_{T-1}\right) = \max_{(\boldsymbol{\pi}_{T-1}, \pi_T) \in \boldsymbol{\Pi}_T} C_T(X_T, \boldsymbol{\pi}_{T-1}, \pi_T) = P_T b_{T-1}(P_T), \quad \boldsymbol{\pi}_{T-1} \in \boldsymbol{\Pi}_{T-1}.$$

Next, at time $T-1$ (15) reads,

$$\begin{aligned} \mathcal{V}_{T-1}(\boldsymbol{\pi}_{T-2}) &= \sup_{(\boldsymbol{\pi}_{T-2}, \pi_{T-1}) \in \boldsymbol{\Pi}_{T-1}} \left[C_{T-1}(X_{T-1}, \boldsymbol{\pi}_{T-2}, \pi_{T-1}) + \mathbb{E}_{T-1}\mathcal{V}_T(\boldsymbol{\pi}_{T-1})\right] \\ &= \sup_{(\boldsymbol{\pi}_{T-2}, \pi_{T-1}) \in \boldsymbol{\Pi}_{T-1}} \left[P_{T-1} b_{T-2}(P_{T-1}) + \mathbb{E}_{T-1}\left[P_T b_{T-1}(P_T)\right]\right] \\ &= P_{T-1} b_{T-2}(P_{T-1}) + \sup_{(\boldsymbol{\pi}_{T-2}, (c_{T-1}, d_{T-1}, o_{T-1}, b_{T-1})) \in \boldsymbol{\Pi}_{T-1}} \mathbb{E}_{T-1}\left[P_T b_{T-1}(P_T)\right]. \end{aligned}$$

For a generic $1 \leq t < T$ we get,

$$\mathcal{V}_t(\boldsymbol{\pi}_{t-1}) = P_t b_{t-1}(P_t) + \sup_{(\boldsymbol{\pi}_{t-1},\pi_t)\in\boldsymbol{\Pi}_t} \mathbb{E}_t \mathcal{V}_{t+1}(\boldsymbol{\pi}_{t-1},\pi_t).$$

The time $t = 0$ needs separate consideration,

$$\mathcal{V}_0 = \sup_{\pi_0\in\Pi_0} \mathbb{E}_0 \mathcal{V}_1(\pi_0).$$

Let us now describe the model of the stochastic process. We define $X_t := (G_t, T_t, I_t)$ where G_t denotes the gas price, I_t denotes the inflow and T_t the temperature. In detail, we assume that

$$\begin{aligned}
\log G_t \quad &= \log G_0 + \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t^{(1)}, \\
T_t \quad &= T_0 + \int_0^t \varsigma_0 \left(\varsigma_1 + \varsigma_2\sqrt{1+\frac{\omega^2}{\varsigma_0^2}}\sin(\omega s + \varsigma_3 + \arctan\frac{\omega}{\varsigma_0}) - T_s\right) ds + \varsigma_4 W_t^{(2)}, \\
I_t \quad &= \left(I_0 + \int_0^t \upsilon_0 \left(\upsilon_1 + \upsilon_2\sqrt{1+\frac{\omega^2}{\upsilon_0^2}}\sin(\omega s + \upsilon_3 + \arctan\frac{\omega}{\upsilon_0}) - I_s\right) ds + \upsilon_4 W_t^{(3)}\right)^+,
\end{aligned}$$

where $\omega = \frac{2\pi}{365}$ is the frequency of the seasonality, $W_t = (W_t^{(1)}, W_t^{(2)}, W_t^{(3)})$ is a three dimensional standard Brownian motion, and the electricity price is modeled as

$$P_t = p_G G_t + p_T T_t + p_I I_t.$$

Here μ, σ, ς_i, υ_i, p_i are known fixed coefficients, $\varsigma_0, \upsilon_0 > 0$. The models for the temperature and the inflow are chosen such that these quantities track the sinusoidal reference curves

$$\begin{aligned}
T_t^{ref} &= \varsigma_1 + \varsigma_2 \sin(\omega t + \varsigma_3), \\
I_t^{ref} &= \upsilon_1 + \upsilon_2 \sin(\omega t + \upsilon_3)
\end{aligned}$$

reflecting the influence of annual seasonality.

5.1 Solution algorithm

In this subsection we present a problem-specific regression algorithm for the hydro storage problem in the spirit of Section 4. Note that at time t, all information generated by the preceding decisions which is relevant for future decisions is encoded by the water levels R_{t-1}^j and the bid curve b_{t-1} which was submitted the previous day.

Feasibility set. When at the beginning of day t the electricity price P_t and the inflow I_t become known, the producer has to decide how to choose the design variables $d_{t,j}, c_{t,j}, o_{t,j}$ in order to produce the needed amount of electricity $b_{t-1}(P_t)$ and to submit the bidding curve b_t, while respecting constraints (17),(18),(19),(20),(21). These constraints are linear in the decision variables, i.e. of the form $B_1\pi \le q_1$, $B_2\pi = q_2$ with fixed coefficient matrices B_1 and B_2. However, the right-hand sides of these constraints depend linearly on the random variables P_t, I_t. This kind of random linear programs with right-hand side depending linearly on the random parameter is well studied, and we give some background information on this type of problems in Appendix A. The maximal amount of electricity that can be produced is limited by the available water in the reservoirs, encoded by the water levels R_t^j, and by the capacities U_j^d of the turbines. In the considered model we do not allow the producer to cover any deficit by buying on the spot market, and hence we are dealing with hard constraints. Let us determine the upper bound of the range (20). Emptying the reservoirs completely, i.e. $R^j \equiv 0$, and assuming discharge of water only, i.e $c_j \equiv 0$, yields by virtue of the balance equation

$$d_j - d_{j+1} - o_{j+1} + o_j = R_{t-1}^j, \quad o_{J+1} = d_{J+1} = c_{J+1} = c_1 = 0.$$

So $\sum_{j=1}^J \eta_j^d d_j$ attains its maximum for

$$d_j = \min\left(\sum_{k=j}^J R_{t-1}^k, U_j^d\right), \qquad o_j = \left(\sum_{k=j}^J R_{t-1}^k - U_j^d\right)^+,$$

which yields the upper bound

$$\sum_{j=1}^J \eta_j^d \min\left(\sum_{k=j}^J R_{t-1}^k, U_j^d\right).$$

It is not difficult to see that the range (20) is unbounded from below (in principle). Indeed, When buying and storing electricity we do not encounter a hard constraint, since any excessive water charged into an upstream reservoir will return downstream by virtue of the overspill. However, it does not make sense, to charge more water than can be stored and discharged later, leading to a finite lower bound on (20) in practice.

We are now going to describe the algorithm.

Step 1: First simulate trajectories $X_t^{(n)}, 0 \le t \le T, n = 1, ..., N$. We then proceed recursively, constructing the dual martingale penalties from $t = T$ backwards in time down to $t = 0$.

For $t = T$ and $n = 1, ..., N$, initialize

$$\widehat{\vartheta}_T^{(n)}(b_{T-1}, R_{T-1}) = P_T^{(n)} \cdot b_{T-1}(P_T^{(n)}). \tag{22}$$

Note that (22) is linear in (b_{T-1}, R_{T-1}).

For general $0 < t \le T$ we assume that $\widehat{\vartheta}_t^{(n)}(b_{t-1}, R_{t-1})$, $n = 1, ..., N$, has already been constructed as a *linear* function. We then compute functions

$$(\widehat{\gamma}_{t-1}(b_{t-1}, R_{t-1}), \widehat{\beta}_{t-1}(b_{t-1}, R_{t-1})) :=$$
$$\underset{\gamma\in\mathbb{R}^K, \beta\in\mathbb{R}^{K\times m}}{\arg\min} \sum_{n=1}^N \left[\widehat{\vartheta}_t^{(n)}(b_{t-1}, R_{t-1}) - \sum_{k=1}^K \gamma^k \psi_k(X_{t-1}^{(n)}) - \sum_{k=1}^K \sum_{l=1}^m \beta^{k,l} \mathfrak{m}_t^{l,k,(n)}\right]^2$$

by linear regression, where $\psi_k = \varphi_k$, $m = 3$, and

$$\mathfrak{m}_t^{l,k,(n)} := \int_{t-1}^{t} \varphi_k(X_u^{(n)}) dW_u^{l,(n)}$$

are the realizations of the basis martingales. Further, set for $n = 1, ..., N$,

$$\widehat{\xi}_t^{(n)}(b_{t-1}, R_{t-1}) = \sum_{k=1}^{K} \sum_{l=1}^{m} \widehat{\beta}_{t-1}^{k,l}(b_{t-1}, R_{t-1}) \mathfrak{m}_t^{l,k,(n)}.$$

Notice that $\widehat{\gamma}_{t-1}$ and $\widehat{\beta}_{t-1}$, and hence also $\widehat{\xi}_t^{(n)}$ are linear in (b_{t-1}, R_{t-1}). Next we compute for $n = 1, ..., N$,

$$\begin{aligned} \widetilde{\vartheta}_{t-1}^{(n)}(b_{t-2}, R_{t-2}) &:= P_{t-1}^{(n)} \cdot b_{t-2}(P_{t-1}^{(n)}) + \sup_{b_{t-1}, R_{t-1} \in \Pi_{t-1}} \left(\widehat{\vartheta}_t^{(n)}(b_{t-1}, R_{t-1}) - \widehat{\xi}_t^{(n)}(b_{t-1}, R_{t-1}) \right) \qquad (23) \\ &= P_{t-1}^{(n)} \cdot b_{t-2}(P_{t-1}^{(n)}) + \\ &\sup_{b_{t-1}, R_{t-1} \in \Pi_{t-1}} \left(\widehat{\vartheta}_t^{(n)}(b_{t-1}, R_{t-1}) - \sum_{k=1}^{K} \sum_{l=1}^{m} \widehat{\beta}_{t-1}^{k,l}(b_{t-1}, R_{t-1}) \mathfrak{m}_t^{l,k,(n)} \right). \end{aligned}$$

Here $\widetilde{\vartheta}_0^{(n)}$ is just a number, as it lacks any arguments, and $\widetilde{\vartheta}_1^{(n)}$ effectively depends only on b_0, because R_0 is fixed and not a design variable. Note also that $\widetilde{\vartheta}_{t-1}^{(n)}(b_{t-2}, R_{t-2})$ depends only on those components of b_{t-2} which correspond to the end-points of the price interval containing the price $P_{t-1}^{(n)}$ on the price grid.

Note that the maximization problem in (23) is a parameterized linear program, because all constraints as well as the cost function are linear in the decision variables π_{t-1}, R_{t-1}. The parameters of this program are the water levels R_{t-2} and the bid curve b_{t-2} from the previous day. These parameters enter only in the right-hand sides of the equality constraints, and their dependence is linear. From general properties of the solution of linear programs it then follows that the objective value and therefore also the function $\widetilde{\vartheta}_{t-1}^{(n)}(b_{t-2}, R_{t-2})$ is concave in (b_{t-2}, R_{t-2}). We may then construct a *linear* approximation $\widehat{\vartheta}_{t-1}^{(n)}(b_{t-2}, R_{t-2})$ of $\widetilde{\vartheta}_{t-1}^{(n)}(b_{t-2}, R_{t-2})$ by computing a supporting hyperplane to its hypograph. It then follows that $\widehat{\vartheta}_{t-1}^{(n)}$ is an upper bound for $\widetilde{\vartheta}_{t-1}^{(n)}$. For details we refer to the Appendix at the end of this paper.

In this manner we proceed in backward time up to $t = 0$ and obtain a system of coefficients

$$(\widehat{\gamma}_{t-1}(b_{t-1}, R_{t-1}), \widehat{\beta}_{t-1}(b_{t-1}, R_{t-1})) \in \mathbb{R}^K \times \mathbb{R}^{K \times m}, \quad 1 \leq t \leq T.$$

Step 2: We simulate a new Monte Carlo ensemble of trajectories $\{\widetilde{X}_t^{(\widetilde{n})}\}_{0 \leq t \leq T, \widetilde{n}=1,...,\widetilde{N}}$ and construct the martingale increment samples

$$\widetilde{\xi}_t^{(\widetilde{n})}(b_{t-1}, R_{t-1}) = \sum_{k=1}^{K} \sum_{l=1}^{m} \widehat{\beta}_{t-1}^{k,l}(b_{t-1}, R_{t-1}) \widetilde{\mathfrak{m}}_t^{l,k,(\widetilde{n})}, \quad 1 \leq t \leq T, \quad \widetilde{n} = 1, ..., \widetilde{N},$$

with

$$\widetilde{\mathfrak{m}}_t^{l,k,(\widetilde{n})} := \int_{s_{t-1}}^{s_t} \varphi_k(\widetilde{X}_s^{(\widetilde{n})}) dW_s^{l,(\widetilde{n})}.$$

As an upper biased estimate we then construct

$$\widehat{V}_0^{up} := \frac{1}{\widetilde{N}} \sum_{\widetilde{n}=1}^{\widetilde{N}} \max_{b_0,(b_1,R_1),\dots,(b_{T-1},R_{T-1})} \left(\sum_{t=1}^{T} P_t^{(\widetilde{n})} b_{t-1}(P_t^{(\widetilde{n})}) - \sum_{t=1}^{T} \widetilde{\xi}_t^{(\widetilde{n})}(b_{t-1}, R_{t-1}) \right).$$

Step 3: Construct recursively from $t = 0$ to $t = T - 1$ on each trajectory

$$(b_t^{(\widetilde{n})}, R_t^{(\widetilde{n})}) := \arg\max_{(b_t,R_t)} \sum_{k=1}^{K} \widehat{\gamma}_t^k(b_t, R_t)\psi_k(X_t^{(\widetilde{n})}),$$

and compute the lower biased estimate

$$\widehat{V}_0^{low} := \frac{1}{\widetilde{N}} \sum_{\widetilde{n}=1}^{\widetilde{N}} \sum_{t=1}^{T} P_t^{(\widetilde{n})} b_{t-1}^{(\widetilde{n})}(P_t^{(\widetilde{n})}).$$

Note that for $0 \leq t < T$ we have the approximation

$$\mathbb{E}_t V_{t+1}^* \approx \sum_{k=1}^{K} \widehat{\gamma}_t^k(b_t, R_t)\psi_k(X_t).$$

6 Numerical tests

We test our algorithm in a setting with four time steps $t = 0, \dots, 3$, $J = 2, \dots, 4$ reservoirs, rising energy prices and large capacities. In this setting, the expected revenue heavily depends on the employed control strategy. Intuitively, the best strategy is saving water at the first time steps and releasing it later. The basis functions used in the regression procedure in Step 1 above are chosen to be polynomials of order 2 in the components of the driving process. The Monte-Carlo simulations comprised 1000 paths for each value of J and the initial conditions were such that each reservoir is half full at the beginning. Pumping water up and down in a cycle corresponds to an efficiency of 50%, i.e., the consumption of the pumps is twice as high as the efficiency of the turbines.

The parameters of the simulation are summarized in Table 1.

The gas price is a geometric Brownian motion with drift $\mu = 2.5$ and variance $\sigma = 0.005$. Therefore the gas price rises sharply over the considered time horizon, and it becomes much more profitable to produce and sell later within the horizon. The parameters of the temperature and inflow are chosen such that the mean value of these processes matches the observations in Vaduz, Liechtenstein.

L	ρ_l	η^d	η^c	R_0	U^R	U^d	U^c
4	$30 : 500 : 2030$	1	2	2500	5000	2000	2000

Table 1: Parameter values.

For each value of the number of reservoirs $J = 2, 3, 4$, an upper bound on the value of the optimal control problem has been computed as described in Step 2, and a lower bound has been computed as in Step 3. These bounds have then been refined in a local search over the coefficients β and γ. In addition, the performance of a simple strategy has been recorded which consists in selling on each day as much electricity as possible while respecting the constraints.

	$J = 2$	$J = 3$	$J = 4$
Upper Bound	8165493	12360195	16554896
Lower Bound	6748607	9669602	12551734
Optimized Upper Bound	8125700	12309000	16499000
Optimized Lower Bound	7126300	10345000	13580000
Simple Strategy	1793736	5993296	10213987

Table 2: Results of the simulation with and without final local search.

From the results summarized in Table 2 it can be seen that the bounds computed by our algorithm cannot be significantly improved by a local search, especially the upper bound. This means that the algorithm finds a martingale that gives close to optimal performance. The control strategy computed from the dual martingale outperforms the simple strategy and is in fact nearly optimal, as the close-ness of the upper and lower bounds demonstrates.

References

[1] Denis Belomestny, Roland Hildebrand, and John Schoenmakers. Optimal stopping via pathwise dual empirical maximisation. Preprint 2043, WIAS, 2014.

[2] David B. Brown, J. Smith, and P. Sun. Information relaxations and duality in stochastic dynamic programs. *Oper. Res.*, 58(4):785–801, 2010.

[3] George B. Dantzig and Mukund N. Thapa. *Linear Programming 2: Theory and Extensions*. Springer series in Operations Research. Springer, New York, Berlin, Heidelberg, 2003.

[4] M.H.A. Davis and I. Karatzas. *A deterministic approach to optimal stopping*, chapter 33, pages 455–466. Wiley and Sons, Chichester, 1994.

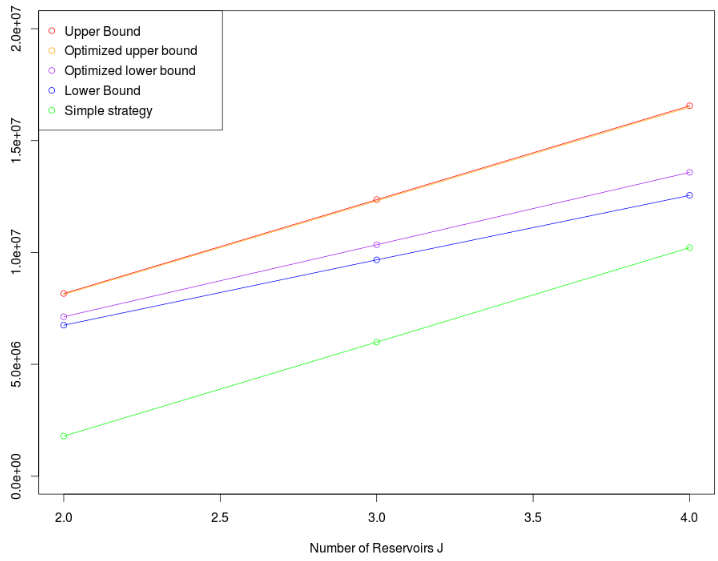

Figure 1: Graphical representation of the results in Table 2.

[5] V.V. Desai, V.F. Farias, and C.C. Moallemi. Pathwise optimization for optimal stopping problems. *Manage. Sci.*, 58(12):2292–2308, 2012.

[6] Tomas Gal. *Postoptimal Analyses, Parametric Programming and Related Topics*. McGraw-Hill, New York, 1979.

[7] M. Haugh and L. Kogan. Pricing American options: A duality approach. *Oper. Res.*, 52:258–270, 2004.

[8] Rogers L.C.G. Monte Carlo valuation of American options. *Math. Financ.*, 12(3):271–286, 2002.

[9] N. Löhndorf, D. Wozabal, and S. Minner. Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming. *Oper. Res.*, 61(4):810–823, 2013.

[10] Philipp Protter. *Stochastic Integration and Differential Equations*. Springer, Berlin, Heidelberg, New York, 2004.

A Parameterized linear programs

Consider a linear program

$$\max\langle c, x\rangle : \quad A_{eq}x = b_{eq},\ A_{ineq}x \le b_{ineq},$$

where the right-hand sides of the equality constraints depend on a parameter y, $b_{eq} = b_{eq}(y)$. Then the optimal solution and the optimal value of the linear program also depend on this parameter. The dependence of the solution of an optimization problem on its data is the subject of sensitivity analysis. For linear programs, sensitivity analysis has been considered in [6], see also [3] for the case when the dependence on the parameter is linear.

We are interested in the case when b_{eq} is a linear function, $b_{eq}(y) = By + b_{eq0}$. Pass to the dual program

$$\min\langle\lambda_{eq}, b_{eq}(y)\rangle + \langle\lambda_{ineq}, b_{ineq}\rangle : \quad A_{eq}^T\lambda_{eq} + A_{ineq}^T\lambda_{ineq} = c,\ \lambda_{ineq} \geq 0.$$

The feasible set of this dual program does no more depend on the parameter y, but now the cost function is concave in y. Note also that if the dual program is feasible, then it has the same optimal value as the primal program, because Linear Programming exhibits a zero duality gap.

Let $F(y)$ be the optimal value of the primal and dual programs above as a function of the parameter. Fix $y = \hat{y}$ and consider the solution $\lambda_{eq}^*, \lambda_{ineq}^*$ of the dual program with this parameter value. This solution is a feasible point of the dual program for any parameter value. We then get for arbitrary y that

$$\begin{aligned} F(y) &\leq \langle\lambda_{eq}^*, b_{eq}(y)\rangle + \langle\lambda_{ineq}^*, b_{ineq}\rangle = \langle\lambda_{eq}^*, b_{eq}(\hat{y})\rangle + \langle\lambda_{ineq}^*, b_{ineq}\rangle + \left\langle\lambda_{eq}^*, B(y-\hat{y})\right\rangle \\ &= F(\hat{y}) + \left\langle B^T\lambda_{eq}^*, y-\hat{y}\right\rangle. \end{aligned}$$

It follows that the vector $B^T\lambda_{eq}^*$ is a super-gradient of the concave function $F(y)$ at $y = \hat{y}$.