Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Relaxations of Approximate Linear Programs for the Real Option Management of Commodity Storage

Selvaprabu Nadarajah, François Margot, Nicola Secomandi

MANAGEMENT SCIENCE
Vol. 61, No. 12, December 2015, pp. 3054–3076
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.2136

Relaxations of Approximate Linear Programs for the Real Option Management of Commodity Storage

Selvaprabu Nadarajah
College of Business Administration, University of Illinois at Chicago, Chicago, Illinois 60607,
selvan@uic.edu

François Margot, Nicola Secomandi
Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
{fmargot@andrew.cmu.edu, ns7@andrew.cmu.edu}

The real option management of commodity conversion assets gives rise to intractable Markov decision processes (MDPs), in part because of the use of high-dimensional models of commodity forward curve evolution, as commonly done in practice. Focusing on commodity storage, we identify a deficiency of approximate linear programming (ALP), which we address by developing a novel approach to derive relaxations of approximate linear programs. We apply our approach to obtain a class of tractable ALP relaxations, also subsuming an existing method. We provide theoretical support for the use of these ALP relaxations rather than their associated approximate linear programs. Applied to existing natural gas storage instances, our ALP relaxations significantly outperform their corresponding approximate linear programs. Our best ALP relaxation is both near optimal and competitive with, albeit slower than, state-of-the-art methods for computing heuristic policies and lower bounds on the value of commodity storage, but is more directly applicable for dual (upper) bound estimation than these methods. Our approach is potentially relevant for the approximate solution of MDPs that arise in the real option management of other commodity conversion assets, as well as the valuation of real and financial options that depend on forward curve dynamics.

Keywords: programming; linear; applications; dynamic programming; Markov; finance; asset pricing; industries; petroleum-natural gas
History: Received June 20, 2012; accepted November 22, 2014, by Dimitris Bertsimas, optimization. Published online in *Articles in Advance* July 9, 2015.

1. Introduction

Real options are models of projects that exhibit managerial flexibility (Dixit and Pindyck 1994). In commodity settings, this flexibility arises from the ability to adapt the operating policy of commodity conversion assets to the uncertain evolution of commodity prices. For example, consider a merchant that manages a natural gas storage asset (Maragos 2002). This merchant can purchase natural gas from the wholesale market at a given price, and store it for future resale into this market at a higher price. Other examples of commodity conversion assets include assets that produce, transport, ship, and procure energy sources, agricultural products, and metals.

Managing commodity conversion assets as real options (Smith and McCardle 1999, Geman 2005) gives rise to, generally, intractable Markov decision processes (MDPs). In a given stage, the state of such an MDP includes both endogenous and exogenous information. The endogenous information describes the current operating conditions of the conversion asset, and the exogenous information represents current market conditions. Changes in the endogenous information are caused by managerial decisions. The exogenous information evolves as a result of market dynamics. The MDP intractability is in part due to the common use in practice of high-dimensional models of the evolution of the exogenous information (Eydeland and Wolyniec 2003). To illustrate, consider the MDP for the real option management of a commodity storage asset formulated by Lai, Margot, and Secomandi (2010; hereinafter LMS) using a multimaturity version of the Black (1976) model of futures price evolution. This is the problem we focus on in this paper. The endogenous information is the asset available inventory at a given date, a one-dimensional variable; the exogenous information is the commodity forward curve at a given time, an object with much higher dimensionality than inventory. Approximations are thus typically needed to heuristically solve such MDPs.

Approximate linear programing (ALP; Schweitzer and Seidmann 1985, de Farias and Van Roy 2003) is an approach that approximates the primal linear program associated with an MDP (Manne 1960, Puterman

1994) by applying a lower-dimensional representation of its variables. Solving an approximate linear program (also abbreviated to ALP for expositional convenience) provides a value function approximation that can be used to obtain a heuristic control policy and estimate lower and upper bounds on the value of an optimal policy (see Bertsekas 2007, Brown et al. 2010, Powell 2011, and references therein). Applications of this approach include Trick and Zin (1997) in economics; Adelman (2004) and Adelman and Klabjan (2012) in inventory routing and control; Adelman (2007), Farias and Van Roy (2007), Zhang and Adelman (2009), and Adelman and Mersereau (2013) in revenue management; Morrison and Kumar (1999), de Farias and Van Roy (2001, 2003), Moallemi et al. (2008), and Veatch (2015) in queuing control; and Restrepo (2008, Chap. 4) in ambulance redeployment. To the best of our knowledge, ALP has not yet been applied to approximately solve MDPs that arise in the real option management of commodity conversion assets.

We focus on the use of ALP for the real option management of commodity storage starting from the MDP of LMS. We analyze the optimal solution sets of both the ALP dual and the dual of the linear program associated with a discretized version of this MDP, which we refer to as the exact dual. The optimal solutions of the exact dual are in one-to-one correspondence with the state-action probability distributions induced by optimal policies (Puterman 1994). In contrast, we find that all the optimal solutions of the ALP dual may correspond to state-action probabilities induced by ill-defined policies with undefined actions at some reachable states. ALP can thus yield low quality value function approximations that lead to poor control policies and bounds. Motivated by this insight, we develop a novel approximate dynamic programming approach that (i) addresses this deficiency of the ALP dual by adding constraints to this model to approximate a key property of the exact dual, and (ii) obtains a value function approximation by solving the ALP relaxation corresponding to the primal linear program associated with this restricted ALP dual.

We apply our approach using two look-up table value function approximations that, in the spirit of LMS, neglect part of the forward curve at each stage and state—one retains only the spot price and the other also the prompt-month futures price. We propose a class of ALP relaxations, which we label *constraint-based* ALP relaxations. We derive three constraint-based ALP relaxations and equivalently reformulate them as recursive optimization models, which we refer to as approximate dynamic programs (ADPs). Two of these ADPs are new. Our third constraint-based ALP relaxation yields the LMS ADP, which we label as storage ADP (SADP). We bound the difference between the exact value function and the value functions of each of these ADPs when future prices evolve according to the multimaturity Black (1976) model. This analysis provides theoretical support for the use of these ADPs rather than their respective ALPs, as well as the superiority of our ADP based on the spot and prompt-month futures prices compared to the other ADPs.

We numerically evaluate our approach on the LMS natural gas instances. Our results are encouraging. Our ADPs significantly outperform their corresponding ALPs, in terms of both the estimated lower and upper bounds. Compared to the other ADPs, our ADP that uses both the spot and prompt-month futures prices yields better upper bounds and substantially better lower bounds, most of which are near optimal. In addition, this ADP relies less on periodic reoptimizations than the other ADPs to obtain near optimal lower bounds, and is thus a better approximation of the commodity storage MDP than these other models. This ADP is also competitive with two state-of-the-art techniques for computing a heuristic operating policy and a lower bound on the value of commodity storage: the rolling (reoptimized) intrinsic (RI) policy (see, e.g., LMS, Wu et al. 2012, and references therein) and the least squares Monte Carlo (LSMC) approach (Longstaff and Schwartz 2001; see Boogert and de Jong 2008, 2011/2012, and Cortazar et al. 2008 for commodity storage and real option applications). Although typically slower than both of these methods for lower bound estimation, our approach is more directly applicable for dual (upper) bound estimation (see Brown et al. 2010 and references therein) because it gives explicit value function approximations, whereas these methods do not. Our research thus adds to the literature on commodity storage real option valuation and management (see, e.g., Chen and Forsyth 2007, Boogert and de Jong 2008, Thompson et al. 2009, Carmona and Ludkovski 2010, LMS, Secomandi 2010, Birge 2011, Boogert and de Jong 2011/2012, Felix and Weber 2012, Wu et al. 2012, Secomandi et al. 2015).

The use of relaxations in ALP is relatively new and the literature is scant: Petrik and Zilberstein (2009) propose an ALP relaxation method with an objective function that penalizes constraint violation; Desai et al. (2012) relax an ALP by imposing a budget on the amount of constraint violation. In contrast, we introduce a general approach for deriving ALP relaxations from ALP dual restrictions and our constraint-based ALP relaxation class does not penalize constraint violations and is not based on the idea of budgeted constraint violations.

De Farias and Van Roy (2003) and Desai et al. (2012) also derive error bounds for ALP and ALP relaxations, respectively. In contrast to de Farias and

Van Roy (2003) our error bounds are for ALP relaxations. Different from the error bounds of Desai et al. (2012), ours rely on the recursive structure of versions of the ADPs that correspond to our constraint-based ALP relaxations. Although related to the error bounds of Tsitsiklis and Van Roy (2001), Munos (2007), and Petrik (2012) for other approximate dynamic programming methods, our bounds differ from theirs in the details of the bounding term.

Even if our focus is on commodity storage, our proposed methodology is potentially relevant for the approximate solution of intractable MDPs that arise in the real option management of other commodity conversion assets, as well as the valuation of real and financial options that depend on forward curve dynamics; that is, MDPs whose states include both endogenous and exogenous information. Examples are commodity processing assets, energy swing options, put-call Bermudan options, and mortgages and interest rate caps and floors (see, e.g., Longstaff and Schwartz 2001, Jaillet et al. 2004, Cortazar et al. 2008, Devalkar et al. 2011).

We provide background material in §2. We discuss the ALP associated with the storage MDP in §3 and analyze it in §4. We describe our ALP relaxation approach and apply it using look-up table value function approximations in §5. We discuss our performance bounds in §6. We present our numerical results in §7. We conclude in §8. Appendix A analyzes the computational complexity of solving our constraint-based ALP relaxations as ADPs and estimating lower and upper bounds using look-up table value function approximations. Appendix B includes proofs.

2. Background Material

In §§2.1 and 2.2 we present the commodity storage MDP and the numerical bounding approach that we use. These subsections are in part based on §§2 and 4.4 in LMS.

2.1. Commodity Storage MDP

A commodity storage asset provides a merchant with the option to purchase and inject, store (do nothing), and withdraw and sell a commodity during a predetermined finite time horizon, while respecting injection and withdrawal capacity limits, as well as inventory constraints. For instance, in the United States, natural gas storage is rented via contracts with finite terms (see, e.g., Maragos 2002). The merchant's goal is to maximize the market value of the storage asset. We model this valuation problem as an MDP. Purchases and injections and withdrawals and sales give rise to cash flows. The storage asset has N possible dates with cash flows. The ith cash flow occurs at time T_i, $i \in \mathcal{I} := \{0, \ldots, N-1\}$. Each such time is also the maturity of a futures contract. Since the trading times in our model coincide with monthly futures maturity dates, we discretize time into monthly intervals. We denote by $F_{i,j}$ the price at time T_i of a futures contract maturing at time T_j, $j \geq i$. The forward curve is the collection of futures prices $F_i := \{F_{i,i}, F_{i,i+1}, \ldots, F_{i,N-1}\}$. We define $F_N := 0$.

Set $\mathcal{I}$ is the stage set. The inventory level at stage zero is the given singleton x_0. The set of inventory levels at every other stage $i \in \mathcal{I}\backslash\{0\}$ is $\mathcal{X}' := [0, \bar{x}]$, where 0 and $\bar{x} \in \mathbb{R}_+$ represent the minimum and maximum inventory levels, respectively. The (absolute value of) the injection capacity C^I (< 0) and the withdrawal capacity C^W (> 0) represent the maximum amounts that can be injected and withdrawn in between two successive stages, respectively. An action a corresponds to an inventory change during this time period. A positive action represents a withdraw-and-sell decision, a negative action a purchase-and-inject decision, and the zero action is the do-nothing decision. Define $\cdot \wedge \cdot := \min\{\cdot, \cdot\}$ and $\cdot \vee \cdot := \max\{\cdot, \cdot\}$. The set of feasible injections, withdrawals, and overall actions are $\mathcal{A}^I(x) := [C^I \vee (x - \bar{x}), 0]$, $\mathcal{A}^W(x) := [0, x \wedge C^W]$, and $\mathcal{A}'(x) := \mathcal{A}^I(x) \cup \mathcal{A}^W(x)$, respectively.

The immediate reward from taking action a at time T_i is the function $r(a, s_i)$, where $s_i \equiv F_{i,i}$ is the spot price at this time. The coefficients $\alpha^W \in (0, 1]$ and $\alpha^I \geq 1$ model commodity losses associated with withdrawals and injections, respectively. The coefficients c^W and c^I represent withdrawal and injection marginal costs, respectively. Given an action a and a spot price s, the immediate reward $r(a, s)$ is $(\alpha^I s + c^I)a$ if $a < 0$, 0 if $a = 0$, and $(\alpha^W s - c^W)a$ if $a > 0$.

Let Π denote the set of all the feasible storage policies. Given the initial state (x_0, F_0), valuing a storage asset entails finding a feasible policy that achieves the maximum time T_0 $(:= 0)$ market value of this asset in this state, $V_0(x_0, F_0)$. Thus, we are interested in solving the following problem:

$$V_0(x_0, F_0) := \max_{\pi \in \Pi} \sum_{i \in \mathcal{I}} \delta^i \, \mathbb{E}\big[r(A_i^\pi(x_i^\pi, F_i), s_i) \,\big|\, x_0, F_0\big], \quad (1)$$

where δ is the risk-free discount factor from time T_i back to time T_{i-1}, $\forall\, i \in \mathcal{I}\backslash\{0\}$; $\mathbb{E}$ is the expectation under the risk-neutral measure for the forward curve evolution (this measure is unique when the commodity market is complete, see, e.g., Björk 2004, p. 122, which we assume to be the case in this paper); x_i^π is the inventory level at stage i when using policy π; and $A_i^\pi(\cdot, \cdot)$ is the decision rule of policy π for stage i.

In our MDP formulation, committing on date T_i to perform a physical trade on date $T_j > T_i$ does not add any value, because the payoff from purchasing and injecting or withdrawing and selling the commodity is linear in the transacted price, given the size of a trade, and we use risk-neutral valuation. For example, the time T_i market value of committing on date T_i to withdraw one unit of commodity on date T_j, $j > i$, and selling this unit forward on date T_i at price $F_{i,j}$ is $\delta^{j-i}(\alpha^W F_{i,j} - c^W)$, and is also the time T_i market value of the cash flow $\alpha^W s_j - c^W$ from withdrawing and selling one unit of commodity at the time T_j spot price s_j, because $\mathbb{E}[s_j \mid F_{i,j}] = F_{i,j}$ (Shreve 2004, p. 244).

When C^I, C^W, and $\bar{x}$ are integer multiples of a maximal number $Q \in \mathbb{R}_+$, Lemma 1 in Secomandi et al. (2015) establishes that we can optimally discretize the continuous inventory set $\mathscr{X}'$ into the finite set $\mathscr{X} := \{0, Q, 2Q, \ldots, \bar{x}\}$, and the continuous feasible action set $\mathscr{A}'(x)$ for inventory level $x \in \mathscr{X}$ into the finite set $\mathscr{A}(x) := \{[C^I \vee (x-\bar{x})], [C^I \vee (x-\bar{x})] + Q, \ldots, [x \wedge C^W]\}$. This assumption holds in the rest of this paper. We can thus replace $\mathscr{X}'$ and $\mathscr{A}'(x)$ in (1) by $\mathscr{X}$ and $\mathscr{A}(x)$, respectively, without sacrificing optimality. Moreover, an optimal policy for problem (1) can be obtained by solving a stochastic dynamic program that uses the sets $\mathscr{X}$ and $\mathscr{A}(\cdot)$. Letting $V_i(x_i, F_i)$ be the optimal value function in stage i and state (x_i, F_i), this stochastic dynamic program, which we refer to as the exact dynamic program (EDP), is

$$V_i(x_i,F_i) = \max_{a_i \in \mathscr{A}(x_i)} \{ r(a_i,s_i) + \delta \mathbb{E}[V_{i+1}(x_i - a_i, F_{i+1}) \mid F_i] \}, \quad (2)$$

$\forall (i, x_i, F_i) \in \mathscr{I} \times \mathscr{X} \times \mathbb{R}_+^{N-i}$, with boundary conditions $V_N(x_N, F_N) := 0$, $\forall x_N \in \mathscr{X}$.

Consistent with the practice-based literature (Eydeland and Wolyniec 2003, Chap. 5; Gray and Khandelwal 2004; and the discussion in LMS), we assume that EDP is formulated using a full-dimensional model of the risk-neutral evolution of the forward curve. An example is the multimaturity version of the Black (1976) model of futures price evolution used by LMS, which we also use for our computational experiments. In this continuous time model, the time t futures price with maturity at time T_i is denoted by $F(t, T_i)$ ($F(T_{i'}, T_i) \equiv F_{i',i}$ for $i', i \in \mathscr{I}$, $i' \leq i$). This price evolves during the interval $[0, T_i]$ as a driftless geometric Brownian motion with maturity specific and constant volatility $\sigma_i > 0$ and standard Brownian motion increment $dZ_i(t)$. The instantaneous correlation between the standard Brownian motion increments $dZ_i(t)$ and $dZ_j(t)$ corresponding to the futures prices with maturities T_i and T_j, $i \neq j$, is $\rho_{ij} \in (-1, 1)$ ($\rho_{ii} = 1$). This model is

$$\frac{dF(t, T_i)}{F(t, T_i)} = \sigma_i \, dZ_i(t), \quad \forall i \in \mathscr{I}, \quad (3)$$

$$dZ_i(t)\, dZ_j(t) = \rho_{ij}\, dt, \quad \forall i, j \in \mathscr{I}, i \neq j. \quad (4)$$

Property 2.1, which is easy to verify, states that in price model (3)–(4) *each* futures price in the forward curve evolves in a Markovian fashion.

Property 2.1. *In model (3)–(4), at a given stage $i \in \mathscr{I}$ and for a given maturity T_j with $j \in \{i+1, \ldots, N-1\}$, the futures price $F_{i,j}$ is sufficient to obtain the probability distribution of the random futures price $F_{i+1,j}$.*

We use a version of Property 2.1 in §§2.2 and 5 to simplify the computation of expectations. This property also holds for other common futures price evolution models used in real option applications (see, e.g., Cortazar and Schwartz 1994). Model (3)–(4) can be extended by making time dependent the constant volatilities and instantaneous correlations without affecting Property 2.1.

2.2. Numerical Bounding Approach

In general, computing an optimal policy for EDP under a price model such as (3)–(4) is intractable. We now describe Monte Carlo simulation procedures for estimating lower and upper bounds on the EDP optimal value function in the initial stage and state given an approximation of the EDP value function. We use this bounding approach in §7 to numerically assess the quality of the policies associated with the models presented in §§5.2 and 5.3. The lower bound estimation relies on the Monte Carlo simulation of a heuristic policy that is *greedy* with respect to this value function approximation (see Bertsekas 2007, §6.1.1; Powell 2011). The upper bound estimation applies the information relaxation and duality approach (see Brown et al. 2010 and references therein) based on this value function approximation. We illustrate these procedures using the value function approximation $\hat{V}_i(x_i, s_i)$, which we assume is available. This function only uses the spot price s_i from the forward curve F_i. Nevertheless, the same approach extends in a straightforward manner to value function approximations that depend on a larger subset of prices in this forward curve.

Consider the lower bound estimation. Given an inventory level x_i and a forward curve F_i in stage i, we use $\hat{V}_i(x_i, s_i)$ to compute a greedy action by solving the greedy optimization problem

$$\max_{a_i \in \mathscr{A}(x_i)} \{ r(a_i, s_i) + \delta \mathbb{E}[\hat{V}_{i+1}(x_i - a_i, s_{i+1}) \mid F_{i,i+1}] \}, \quad (5)$$

where $F_{i,i+1}$ is sufficient for computing the expectation by Property 2.1. We obtain (5) from (2) by replacing $V_{i+1}(\cdot,\cdot)$ with $\hat{V}_{i+1}(\cdot,\cdot)$ and F_i with $F_{i,i+1}$. In computations, we numerically approximate the expectation in (5) using Rubinstein (1994) lattices, as discussed in Appendix A. We apply the action $a_i(x_i, s_i)$ computed in (5) (breaking ties by picking $a_i(x_i, s_i)$ such that the inventory change $|a_i(x_i, s_i)|$ is minimized) and sample the forward curve F_{i+1} to obtain the new state $(x_i - a_i(x_i, s_i), F_{i+1})$ in stage $i+1$. Starting from the initial stage and state, we continue in this fashion until time T_{N-1} (included). We then discount back to time zero the cash flows generated by this process, and add them up. We repeat this process for multiple forward curve samples and average the sample discounted total cash flows to estimate the value of the greedy policy, that is, the policy defined by the greedy action in each stage and state. This average provides us with an unbiased estimate of a greedy (lower) bound on the value of storage, $V_0(x_0, F_0)$.

When a value function approximation is computed by an approximate dynamic programming model it is sometimes possible to generate an improved greedy bound estimate by sequentially reoptimizing this model to update its value function approximations within the Monte Carlo simulation used for lower bound estimation (see, e.g., LMS). Specifically, solving such a model at time T_i yields value function approximations for stages i through $N-1$. However, we only implement the greedy action induced by the stage i value function approximation. At time T_{i+1}, we reoptimize the "residual" model, that is, the one defined over the remaining stages $i+1$ through $N-1$, given the inventory level resulting from performing this action and the newly available forward curve. We repeat this procedure until time T_{N-1}. Repeating this process over multiple forward curve sample paths allows us to estimate a reoptimized greedy bound.

For upper bound estimation, we sample a sequence of spot price and prompt-month futures price pairs $P_0 := ((s_i, F_{i,i+1}))_{i=0}^{N-1}$ starting from the forward curve F_0 at time zero. We use our value function approximation $\hat{V}_{i+1}(x_{i+1}, s_{i+1})$ to define the following dual penalty for executing the feasible action a_i in stage i and state (x_i, F_i) given knowledge of the prompt-month futures price $F_{i,i+1}$ and the stage $i+1$ spot price s_{i+1}: $\mathsf{P}_i(x_i, a_i, s_{i+1}, F_{i,i+1}) := \hat{V}_{i+1}(x_i - a_i, s_{i+1}) - \mathbb{E}[\hat{V}_{i+1}(x_i - a_i, s_{i+1}) \mid F_{i,i+1}]$, where $F_{i,i+1}$ is sufficient for computing the expectation by Property 2.1. For computational purposes, we numerically approximate the expectation in the dual penalty definition using Rubinstein (1994) lattices (see Appendix A). This penalty approximates the value of knowing the next stage spot price when performing this action. Then, we solve the following deterministic dynamic program given the sequence P_0 (see Brown et al. 2010, LMS):

$$U_i(x_i; P_0) = \max_{a_i \in \mathcal{A}(x)} \{ r(a_i, s_i) - \mathsf{P}_i(x_i, a_i, s_{i+1}, F_{i,i+1}) + \delta U_{i+1}(x_i - a_i; P_0) \}, \quad (6)$$

$\forall (i, x_i) \in \mathcal{I} \times \mathcal{X}$, with boundary conditions $U_N(x_N; P_0) := 0$, $\forall x_N \in \mathcal{X}$. In (6), the immediate reward $r(a_i, s_i)$ is modified by the penalty $\mathsf{P}_i(x_i, a_i, s_{i+1}, F_{i,i+1})$ for using the future information available in P_0. We solve a collection of deterministic dynamic programs (6), each one corresponding to a sampled sequence P_0. We estimate a dual (upper) bound on the value of storage as the average of the value functions of these deterministic dynamic programs in the initial stage and state; that is, we estimate $\mathbb{E}[U_0(x_0; P_0) \mid F_0]$, where the expectation is taken with respect to the risk-neutral distribution of the random sequence P_0 conditional on F_0.

3. ALP

In this section we apply the ALP approach for heuristically solving MDPs with finite state and action spaces (Schweitzer and Seidmann 1985, de Farias and Van Roy 2003). EDP has a finite action space but its state space is in part continuous. To be able to apply the ALP approach, we discretize the forward curve part of the EDP state to obtain a discretized version of EDP (DDP). We let $\mathcal{F}_i \subset \mathbb{R}_+^{N-i}$ represent the finite set of forward curves at time T_i, and $\mathcal{F}_{i,j} \subset \mathbb{R}_+$ the corresponding finite set for the futures price $F_{i,j}$. We denote by $\{\Pr(F_{i+1} \mid F_i),\ F_{i+1} \in \mathcal{F}_{i+1}\}$ the probability mass function of the random vector F_{i+1} on the set $\mathcal{F}_{i+1}$ conditional on the forward curve $F_i \in \mathcal{F}_i$. We make Assumption 3.1 to ensure that all the forward curves in our discretized sets have positive probability and their evolution satisfies a version of Property (2.1), which holds in our numerical study discussed in §7 (see Appendix A for details).

Assumption 3.1. *It holds that* $\Pr(F_{i+1} \mid F_i) > 0\ \forall (F_i, F_{i+1}) \in \mathcal{F}_i \times \mathcal{F}_{i+1}$. *Moreover, at a given stage* $i \in \mathcal{I}$ *and for a given maturity* T_j *with* $j \in \{i+1, \ldots, N-1\}$, *the futures price* $F_{i,j}$ *is sufficient to obtain the probability mass function of the random futures price* $F_{i+1,j}$.

To simplify the notational burden, for the most part in the rest of this paper we omit the sets on which a tuple is defined. For example, we write (i, x_i, F_i, a_i) in lieu of $(i, x_i, F_i, a_i) \in \mathcal{I} \times \mathcal{X} \times \mathcal{F}_i \times \mathcal{A}(x_i)$. We write $(\cdot)_{-(i)}$ to indicate that i is excluded from $\mathcal{I}$ in the tuple ground set. Replacing the continuous forward curve sets that define EDP with their discretized versions yields DDP. Letting $V_i^D(x_i, F_i)$ be the DDP optimal value function in stage i and state (x_i, F_i), DDP is

$$V_i^D(x_i, F_i) = \max_{a_i} \{ r(a_i, s_i) + \delta \mathbb{E}[V_{i+1}^D(x_i - a_i, F_{i+1}) \mid F_i] \}, \quad (7)$$

$\forall (i, x_i, F_i)$, with boundary conditions $V_N^D(x_N, F_N) := 0$, $\forall x_N$. The expectation in (7) is expressed with respect to the probability mass function $\{\Pr(F_{i+1} \mid F_i),\ F_{i+1}\}$, even though our notation does not make it explicit.

It is well known (Manne 1960, Puterman 1994, §6.9) that DDP can be reformulated as a linear program, which we refer to as the exact primal linear program (PLP). PLP has one variable for every stage and state and one constraint for every stage, state, and action. We refer to the PLP dual (Puterman 1994, p. 223) as DLP.

Solving PLP or DLP is typically intractable because of the exponential number of variables and constraints as a function of the number of futures prices in the forward curve. Computational tractability dictates approximating these models.

Following the ADP literature (Schweitzer and Seidmann 1985, de Farias and Van Roy 2003), PLP can be approximated by replacing its variables with lower-dimensional approximations defined as linear combinations of a manageable number of basis functions. Let $\psi_{i,x_i,b}$: $\mathbb{R}^{N-i} \to \mathbb{R}$ be the bth basis function corresponding to the pair (i, x_i) that takes F_i as argument. There are B_i basis functions for each stage i, that is, $b \in \{1, \ldots, B_i\}$. The weight associated with the bth basis function for each pair (i, x_i) is $\beta_{i,x_i,b} \in \mathbb{R}$. The value function approximation is $\sum_b \psi_{i,x_i,b}(F_i)\beta_{i,x_i,b}$. Since the relevant state at stage zero is (x_0, F_0), we choose $B_0 = 1$ and $\psi_{0,x_0,1}(F_0) = 1$ without loss of generality. The value function approximation weights solve the following ALP:

$$\min_{\beta} \quad \beta_{0,x_0,1} \tag{8}$$

$$\text{s.t.} \quad \sum_b \psi_{N-1,x_{N-1},b}(F_{N-1})\beta_{N-1,x_{N-1},b} \ge r(a_{N-1}, s_{N-1}), \quad \forall (x_{N-1}, F_{N-1}, a_{N-1}), \tag{9}$$

$$\sum_b \psi_{i,x_i,b}(F_i)\beta_{i,x_i,b} \ge r(a_i, s_i) + \delta \mathbb{E}\left[\sum_b \psi_{i+1,x_i-a_i,b}(F_{i+1})\beta_{i+1,x_i-a_i,b} \,\middle|\, F_i\right], \quad \forall (i, x_i, F_i, a_i)_{-(N-1)}. \tag{10}$$

The objective function (8) minimizes the approximate value function corresponding to the initial stage and state. Constraints (9)–(10) can be obtained from DDP as follows: for each triple (i, x_i, F_i) express the maximization over the set $\mathcal{A}(x_i)$ in (7) as $|\mathcal{A}(x_i)|$ inequalities, where $|\cdot|$ denotes set cardinality, one for each a_i, and then replace $V_i(x_i, F_i)$ by $\sum_b \psi_{i,x_i,b}(F_i)\beta_{i,x_i,b}$. These constraints ensure that the ALP value function is an upper bound on the DDP value function approximation at every stage and state (de Farias and Van Roy 2003).

Let $\mathbb{1}(\cdot)$ represent the indicator function that evaluates to one when its argument is true and zero otherwise. Denoting by $w_i(x_i, F_i, a_i)$ the dual variable of the ALP constraint corresponding to (i, x_i, F_i, a_i), the dual of this ALP (DALP) is

$$\max_{w} \quad \sum_{(i,x_i,F_i,a_i)} r(a_i, s_i) w_i(x_i, F_i, a_i) \tag{11}$$

$$\text{s.t.} \quad \sum_{a_0} w_0(x_0, F_0, a_0) = 1, \tag{12}$$

$$\sum_{F_i} \psi_{i,x_i,b}(F_i)\left[\sum_{a_i} w_i(x_i, F_i, a_i)\right] = \sum_{F_i} \psi_{i,x_i,b}(F_i)\left[\delta \sum_{F_{i-1}} \Pr(F_i \mid F_{i-1}) \cdot \sum_{(x_{i-1},a_{i-1})} \mathbb{1}(x_{i-1} - a_{i-1} = x_i) w_{i-1}(x_{i-1}, F_{i-1}, a_{i-1})\right], \quad \forall (i, x_i, b)_{-(0)}, \tag{13}$$

$$w_i(x_i, F_i, a_i) \ge 0, \quad \forall (i, x_i, F_i, a_i). \tag{14}$$

It can be verified that the DALP objective function (11) and the constraints (12) and (14) are identical to the corresponding DLP objective function and constraints (the DLP formulation is not given here for brevity). In contrast, the flow conservation constraints (13) differ from the DLP flow conservation constraints. Specifically, for each pair (i, x_i) DALP has one constraint (13) for each basis function, and the DALP constraint corresponding to the triple (i, x_i, b) is a linear combination of the DLP flow conservation constraints corresponding to the triples in the set $\{(i, x_i, F_i), F_i\}$ taken using the coefficients $\psi_{i,x_i,b}(F_i)$. Therefore, DALP is a relaxation of DLP.

4. ALP Analysis

In this section we analyze the ALP and DALP models introduced in §3. We begin this analysis by discussing the relationship between feasible DLP solutions and feasible DDP policies. Following a DDP feasible policy starting from the initial stage and state induces a collection of probability mass functions defined over the feasible state and action spaces in each stage. Given a feasible DDP policy π and such a probability mass function, we denote by $\Pr^{\pi}(x_i, F_i, a_i)$ the time zero probability of visiting state (x_i, F_i) in stage i and taking action a_i under policy π (this probability depends on the initial stage and state but we suppress this dependence from our notation for expositional convenience). Therefore, a feasible DDP policy π can be equivalently specified by the set of probabilities $\{\Pr^{\pi}(x_i, F_i, a_i), (i, x_i, F_i, a_i)\}$. It follows from Theorem 6.9.1 in Puterman (1994, p. 224) that the set of feasible DLP solutions encodes the set of feasible DDP policies: there is a one-to-one correspondence between feasible DDP policies and feasible DLP solutions. In particular, for every feasible DDP policy π there exists a feasible DLP solution u such that

$$u_i(x_i, F_i, a_i) = \delta^i \Pr^{\pi}(x_i, F_i, a_i), \quad \forall (i, x_i, F_i, a_i). \tag{15}$$

It follows from the equalities (15) that every optimal DDP policy is related to an optimal DLP solution in this manner.

Let $\Pr^*(x_i, F_i, a_i)$ denote the probability of visiting state (x_i, F_i) in stage i and taking action a_i under an optimal DDP policy. For this optimal policy, we now investigate whether there exists an optimal DALP solution w^* that satisfies a condition analogous to (15), that is,

$$w_i^*(x_i, F_i, a_i) = \delta^i \Pr{}^*(x_i, F_i, a_i), \quad \forall (i, x_i, F_i, a_i). \tag{16}$$

We make Assumption 4.1 to ensure feasibility of ALP (de Farias and Van Roy 2003):

Assumption 4.1. *It holds that $\psi_{i,x_i,1} = 1$, $\forall (i, x_i)$.*

We denote by $\mathcal{F}_i^=(\beta^*)$ the set of stage i forward curves for which at least one ALP constraint corresponding to stage i holds as an equality at an ALP optimal solution β^*. Proposition 4.2 is useful to identify possible violations of (16) by the set of optimal DALP solutions.

Proposition 4.2. *Suppose Assumption 4.1 holds. For every feasible DALP solution w it holds that*

$$\sum_{(x_i, F_i, a_i)} w_i(x_i, F_i, a_i) = \delta^i, \quad \forall i. \tag{17}$$

Moreover, for every optimal DALP solution w^* it holds that

$$\sum_{(x_i, a_i)} w_i^*(x_i, F_i, a_i) = 0, \quad \forall (i, F_i) \in \mathcal{I} \times \{\mathcal{F}_i \backslash \mathcal{F}_i^=(\beta^*)\}. \tag{18}$$

Condition (17) states that a feasible DALP solution specifies a collection of discounted probability mass functions defined over the DDP state and action spaces in each stage. Suppressing its dependence on F_0 for notational convenience, let $\Pr(F_i)$ denote the probability of observing the forward curve F_i. Condition (18) implies that the collection of probability mass functions corresponding to the set of optimal DALP solutions violates (16) when the set $\mathcal{F}_i^=(\beta^*)$ is a proper subset of $\mathcal{F}_i$, because the conditions

$$\sum_{(x_i, a_i)} w_i^*(x_i, F_i, a_i) = \delta^i \Pr(F_i), \quad \forall (i, F_i), \tag{19}$$

obtained by summing both sides of (16) over (x_i, a_i), are necessary for the validity of (16) and $\Pr(F_i) > 0$ by Assumption 3.1. Optimal DALP solutions can thus correspond to ill-defined policies that have undefined actions at some reachable states. In particular, comparing (18) and (19) reveals a distortion between the collections of the discounted probability mass functions associated with every optimal DALP solution and with an optimal DDP policy. Specifically, at stage i the probability mass function w_i^* assigns zero probability to the forward curves in set $\{\mathcal{F}_i \backslash \mathcal{F}_i^=(\beta^*)\}$, hence inflating the probability it attributes to the forward curves in set $\mathcal{F}_i^=(\beta^*)$, because $w_i^*(\cdot)$ satisfies (17). Hence, this distortion can lead to poor value function approximations, as further discussed in §5.2. In contrast, Proposition 4.3 states that when at least one optimal DDP policy and one optimal DALP solution satisfy (16) every optimal ALP solution (β^*) leads to an optimal greedy policy. (The equality $\mathcal{F}_i^=(\beta^*) = \mathcal{F}_i$ is necessary for (16) to hold, which follows from the proof of Proposition 4.3 and complementary slackness.) We denote by $\Pi^g(\beta^*)$ the set of greedy policies induced by the value function approximation specified by β_i^*.

Proposition 4.3. *If an optimal DDP policy and an optimal DALP solution satisfy (16), then for every ALP optimal solution β^* there exists a deterministic optimal DDP policy π^* that is greedy with respect to the value function approximation defined by β^*; that is, $\pi^* \in \Pi^g(\beta^*)$.*

Proposition 4.3 provides a sufficient condition for the optimality of an ALP-based greedy policy. This condition is satisfied even if the ALP value function approximation differs from the DDP value function but their slopes with respect to inventory coincide. Thus, Proposition 4.3 suggests that, if possible, it may be useful to require an optimal DALP solution to be consistent, in the sense of (16), with a deterministic optimal DDP policy.

5. ALP Relaxations

In this section we present our approach to derive ALP relaxations (§5.1), formulate and analyze two ALPs based on look-up table value function approximations (§5.2), and apply our relaxation approach to these ALPs (§5.3). Table 1 summarizes the correspondence between the ALPs and the ADPs analyzed in §§5 and 7.

5.1. Approach for Deriving ALP Relaxations

Before delving into the details of how we derive our ALP relaxations, we briefly summarize the end result of this process: Given a value function approximation architecture (set of basis functions), (i) we introduce variables that allow deviations between the left- and the right-hand sides of the ALP constraints in addition to what is allowed by the inequality sign of these constraints, and (ii) ensure that averages of such deviations equal zero when taken according to particular probability distributions, which are the main modeling lever in our approach.

Motivated by our analysis in §4, we would like to add constraints to DALP requiring that its feasible solutions match the discounted probability mass function induced by an optimal policy for DDP. The

Table 1 **Correspondence Between the ALPs and the ADPs Analyzed in §§5 and 7**

ALP	ALP relaxation	Number of prices in the value function approximation
ADP0	SADP, ADP1	1
ADP0′	ADP2	2

specific constraints that we would like to add to DALP are

$$w_i(x_i, F_i, a_i) = \delta^i \Pr^*(x_i, F_i, a_i), \quad \forall\,(i, x_i, F_i, a_i). \tag{20}$$

Although the probability on the right-hand side of (20) is unknown in applications, we proceed temporarily ignoring this important fact.

Let $d_i(x_i, F_i, a_i)$ be the dual variable associated with the constraint in (20) corresponding to (i, x_i, F_i, a_i). The dual of the DALP restricted by constraints (20) is the ALP relaxation

$$\min_{\beta, d} \; \beta_{0, x_0, 1} + \sum_{(i, x_i, F_i, a_i)} \delta^i \Pr^*(x_i, F_i, a_i) d_i(x_i, F_i, a_i) \tag{21}$$

$$\text{s.t.} \;\; \beta_N = 0, \tag{22}$$

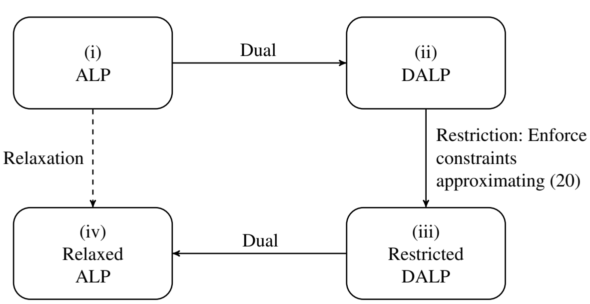
$$\sum_b \psi_{i, x_i, b}(F_i)\beta_{i, x_i, b} + d_i(x_i, F_i, a_i) \geq r(a_i, s_i) + \delta\,\mathbb{E}\left[\sum_b \psi_{i+1, x_i - a_i, b}(F_{i+1})\beta_{i+1, x_i - a_i, b} \,\middle|\, F_i\right], \quad \forall\,(i, x_i, F_i, a_i). \tag{23}$$

Compared to ALP, that is, (8)–(10), the linear program (21)–(23) includes the variables $d_i(x_i, F_i, a_i)$ (i) on the left-hand side of its constraints (23) and (ii) in a term in its objective function that penalizes relaxations of constraints (23) when $d_i(x_i, F_i, a_i)$ is strictly positive and rewards the tightening of these constraints when $d_i(x_i, F_i, a_i)$ is strictly negative.

The ALP relaxation (21)–(23) is impractical because it depends on the unknown terms $\Pr^*(\cdot, \cdot, \cdot)$, in addition to having exponentially many variables $d_i(x_i, F_i, a_i)$ and constraints (23). We thus focus on deriving practical ALP relaxations by adding constraints to DALP that approximate (20) and avoid this exponential growth in the number of variables and constraints in the resulting ALP relaxation. Our approach is summarized in Figure 1. The solid arrows in this figure show the process of constructing an ALP relaxation: (i) starting from ALP, (ii) we formulate DALP, (iii) restrict it in the stated manner, and (iv) take the dual of this restriction to obtain an ALP relaxation.

Our approach relaxes the ALP constraints (10), which ensure that the ALP value function is an upper bound on the DDP value function at each stage and state (de Farias and Van Roy 2003). Therefore, unlike ALP, the value function approximation obtained by solving an ALP relaxation may not provide an upper bound on the DDP value function at every stage and state. Nevertheless, if the constraints used to restrict DALP are implied by (20), it can be verified that (i) the optimal objective function value of the ALP relaxation provides an upper bound on the DDP optimal value function at the initial stage and state, $V_0^D(x_0, F_0)$, and (ii) this upper bound is no worse than the corresponding ALP upper bound. An ALP relaxation, derived from applying our approach, with this upper bounding property can be found in Nadarajah (2014, Appendix A.1).

Figure 1 Schematic Illustration of the ALP Relaxation Framework



5.2. ALPs Based on Look-Up Table Value Function Approximations

In the rest of this paper we focus on using low-dimensional look-up table value function approximations: *discrete grids* that in each stage depend on the inventory level and at most the first two futures prices in the forward curve. In light of Assumption 3.1, look-up table value function approximations are appealing because they result in a dimensionality reduction that makes tractable (i) computing the expectation in (23) and (ii) solving the resulting linear program.

Our starting point is an ALP formulated using the look-up table value function approximation $\phi_i(x_i, s_i)$, which in stage i depends on the inventory x_i and the spot price s_i, as in LMS. This value function approximation corresponds to choosing for each pair (i, x_i) as many indicator basis functions as there are elements in the spot price set $\mathcal{F}_{i,i}$ ($B_i = |\mathcal{F}_{i,i}|$), such that the indicator basis function associated with a specific spot price $s_i \in \mathcal{F}_{i,i}$ returns one when its argument equals s_i and zero otherwise. The value of $\phi_i(x_i, s_i)$ is the weight of each such indicator basis function. By Assumption 3.1, the expectation $\mathbb{E}[\phi_{i+1}(\cdot, s_{i+1}) \mid F_i]$ can be simplified to $\mathbb{E}[\phi_{i+1}(\cdot, s_{i+1}) \mid F_{i,i+1}]$. The corresponding ALP, which has a much smaller number of constraints than the ALP (8)–(10), is

$$\min_{\phi} \; \phi_0(x_0, s_0) \tag{24}$$

$$\text{s.t.} \;\; \phi_{N-1}(x_{N-1}, s_{N-1}) \geq r(a_{N-1}, s_{N-1}), \quad \forall\,(x_{N-1}, s_{N-1}, a_{N-1}), \tag{25}$$

$$\phi_i(x_i, s_i) \geq r(a_i, s_i) + \delta\,\mathbb{E}[\phi_{i+1}(x_i - a_i, s_{i+1}) \mid F_{i,i+1}], \quad \forall\,(i, x_i, s_i, F_{i,i+1}, a_i)_{-(N-1)}. \tag{26}$$

Proposition 5.1 states that an optimal solution to the ALP (24)–(26) can be obtained by solving a recursive optimization model based on a modification of the Bellman operator associated with DDP (model (7)). We refer to this recursive optimization model and other related such models as ADPs. We label by ADP0 the following ADP associated with the ALP (24)–(26):

$$\phi_i^{\mathrm{ADP0}}(x_i, s_i) = \max_{F_{i,i+1}} \Big\{ \max_{a_i} r(a_i, s_i) + \delta \mathbb{E}\big[\phi_{i+1}^{\mathrm{ADP0}}(x_i - a_i, s_{i+1}) \,|\, F_{i,i+1}\big] \Big\}, \tag{27}$$

$\forall (i, x_i, s_i)$, with $\phi_N^{\mathrm{ADP0}}(x_N, s_N) := 0$, $\forall x_N$.

Proposition 5.1. *The terms* $\phi_i^{\mathrm{ADP0}}(x_i, s_i)$ *optimally solve* (24)–(26).

Proposition 5.1 sheds some light on how the ALP (24)–(26) modifies the DDP Bellman operator when determining a value function approximation. Specifically, ADP0 has two maximizations: the first over the price $F_{i,i+1}$ and the second over the action a_i. The second maximization is analogous to the maximization in DDP (see (7)). By Proposition 5.1, the first maximization implies that the ALP (24)–(26) treats the exogenous futures price $F_{i,i+1}$ as a choice, which is unrealistic. Moreover, given a pair (x_i, s_i), we have verified numerically on the instances discussed in §7 that the maximizer in this optimization is typically the largest price in the set $\mathcal{F}_{i,i+1}$, which has a low probability of occurring given s_i. In other words, on our instances the ALP (24)–(26) yields value function approximations that are determined by unlikely prompt-month futures prices. Therefore, the ALP (24)–(26) seems a particularly poor model. Our ALP relaxation approach addresses this issue.

We also formulate the following ALP using the value function approximation $\phi_i(x_i, s_i, F_{i,i+1})$, which compared to $\phi_i(x_i, s_i)$ is a look-up table that in every stage also depends on the prompt-month futures price:

$$\min_{\phi} \ \phi_0(x_0, s_0, F_{0,1}) \tag{28}$$

$$\text{s.t.} \ \phi_{N-1}(x_{N-1}, s_{N-1}) \ge r(a_{N-1}, s_{N-1}), \quad \forall (x_{N-1}, s_{N-1}, a_{N-1}), \tag{29}$$

$$\phi_{N-2}(x_{N-2}, s_{N-2}, F_{N-2,N-1}) \ge r(a_{N-2}, s_{N-2}) + \delta \mathbb{E}\big[\phi_{N-1}(x_{N-2} - a_{N-2}, s_{N-1}) \,\big|\, F_{N-2,N-1}\big], \quad \forall (x_{N-2}, s_{N-2}, F_{N-2,N-1}, a_{N-2}), \tag{30}$$

$$\phi_i(x_i, s_i, F_{i,i+1}) \ge r(a_i, s_i) + \delta \mathbb{E}\big[\phi_{i+1}(x_i - a_i, s_{i+1}, F_{i+1,i+2}) \,\big|\, F_{i,i+1}, F_{i,i+2}\big], \quad \forall i \in \mathcal{I} \setminus \{N-1, N-2\}, (x_i, s_i, F_{i,i+1}, F_{i,i+2}, a_i). \tag{31}$$

An optimal solution to this ALP can be computed by solving ADP0′, which is analogous to ADP0 and is formulated as

$$\phi_{N-1}^{\mathrm{ADP0}'}(x_{N-1}, s_{N-1}) = \max_{a_{N-1}} r(a_{N-1}, s_{N-1}), \quad \forall (x_{N-1}, s_{N-1}),$$

$$\phi_{N-2}^{\mathrm{ADP0}'}(x_{N-2}, s_{N-2}, F_{N-2,N-1}) = \max_{a_{N-2}} \Big\{ r(a_{N-2}, s_{N-2}) + \delta \mathbb{E}\big[\phi_{N-1}^{\mathrm{ADP0}'}(x_{N-2} - a_{N-2}, s_{N-1}) \,\big|\, F_{N-2,N-1}\big] \Big\}, \quad \forall (x_{N-2}, s_{N-2}, F_{N-2,N-1}),$$

$$\phi_i^{\mathrm{ADP0}'}(x_i, s_i, F_{i,i+1}) = \max_{F_{i,i+2}} \Big\{ \max_{a_i} r(a_i, s_i) + \delta \mathbb{E}\big[\phi_{i+1}^{\mathrm{ADP0}'}(x_i - a_i, s_{i+1}, F_{i+1,i+2}) \,\big|\, F_{i,i+1}, F_{i,i+2}\big] \Big\}, \quad \forall i \in \mathcal{I} \setminus \{N-1, N-2\}, (x_i, s_i, F_{i,i+1}),$$

with $\phi_N^{\mathrm{ADP0}'}(x_N, s_N) := 0$, $\forall x_N$. ADP0′ suffers from a drawback similar to the one discussed for ADP0. Analogous to the ALP (24)–(26), the ALP (28)–(31) is thus a poor model. Our ALP relaxation approach tackles this deficiency.

5.3. Constraint-Based ALP Relaxations

By applying our relaxation framework outlined in Figure 1 to the duals of the ALPs (24)–(26) and (28)–(31), we now derive a class of ALP relaxations that include additional constraints that control the extent to which the original constraints of these ALPs are relaxed. We thus label this class as *constraint-based* ALP relaxations. These ALP relaxations have equivalent ADP reformulations that (i) allow us to interpret how these relaxations overcome the shortcomings of ADP0 and ADP0′ pointed in §5.2; (ii) have optimal policies that share the double base-stock target structure of a DDP optimal policy (see Secomandi et al. 2015, Lemma 1, for the DDP optimal policy structure; our claim is easy to verify; see also LMS, Theorem 1(a)); and (iii) are easier to solve using backward recursion than their corresponding linear programming formulations, because of the low dimensionality of the endogenous state and action spaces—this solution approach is thus suitable for problems with this feature.

We denote $\{F_{i,i+2}, \ldots, F_{i,N-1}\}$ as $F_i \setminus \{s_i, F_{i,i+1}\}$ and the sums $\sum_{F_i \setminus \{s_i, F_{i,i+1}\}} \Pr^*(x_i, F_i, a_i)$ and $\sum_{(F_i \setminus \{s_i, F_{i,i+1}\}, a_i)} \Pr^*(x_i, F_i, a_i)$ as $\Pr^*(x_i, s_i, F_{i,i+1}, a_i)$ and $\Pr^*(x_i, s_i, F_{i,i+1})$, respectively. The dual of the ALP (24)–(26) can be obtained from DALP (model (11)–(14)) by applying Assumption 3.1 and the look-up table value function approximation discussed in §5.2. Specifically, each DALP constraint (13) becomes

$$\sum_{(F_{i,i+1}, a_i)} \sum_{F_i \setminus \{s_i, F_{i,i+1}\}} w_i(x_i, F_i, a_i) = \delta \sum_{(s_{i-1}, F_{i-1,i})} \Pr(s_i \,|\, F_{i-1,i}) \sum_{(x_{i-1}, a_{i-1})} \mathbb{1}(x_{i-1} - a_{i-1} = x_i) \cdot \sum_{F_{i-1} \setminus \{s_{i-1}, F_{i-1,i}\}} w_{i-1}(x_{i-1}, F_{i-1}, a_{i-1}).$$

The term $\sum_{F_i \backslash \{s_i, F_{i,i+1}\}} w_i(x_i, F_i, a_i)$ factors out of each such constraint, as well as out of the DALP objective function (11) once expressed accordingly. Replacing this term by a single variable $w_i(x_i, s_i, F_{i,i+1}, a_i)$ makes the resulting DALP (11)–(14) identical to the dual of the ALP (24)–(26). Applying a similar factorization and substitution after summing both sides of constraints (20) over $F_i \backslash \{s_i, F_{i,i+1}\}$ gives

$$w_i(x_i, s_i, F_{i,i+1}, a_i) = \delta^i \Pr^*(x_i, s_i, F_{i,i+1}, a_i), \quad \forall (i, x_i, s_i, F_{i,i+1}, a_i). \quad (32)$$

We obtain the constraints that we add to the dual of the ALP (24)–(26) by approximating constraints (32) in three steps. In the first step we sum both sides of constraints (32) over the feasible actions:

$$\sum_{a_i} w_i(x_i, s_i, F_{i,i+1}, a_i) = \delta^i \Pr^*(x_i, s_i, F_{i,i+1}), \quad \forall (i, x_i, s_i, F_{i,i+1}). \quad (33)$$

In the second step we express the discounted probabilities $\delta^i \Pr^*(x_i, s_i, F_{i,i+1})$ on the right-hand side of (33) as $\delta^i \Pr^*(F_{i,i+1} \mid x_i, s_i) \cdot \Pr^*(x_i, s_i)$ and replace the term $\delta^i \Pr^*(x_i, s_i)$ by the new decision variable $\theta_i(x_i, s_i)$, hence obtaining the constraints

$$\sum_{a_i} w_i(x_i, s_i, F_{i,i+1}, a_i) = \Pr^*(F_{i,i+1} \mid x_i, s_i)\theta_i(x_i, s_i), \quad \forall (i, x_i, s_i, F_{i,i+1}). \quad (34)$$

Constraints (34) relax constraints (32), because the decision variable $\theta_i(x_i, s_i)$ can take values other than $\delta^i \Pr^*(x_i, s_i)$. In the third step we approximate the unknown probability $\Pr^*(F_{i,i+1} \mid x_i, s_i)$ on the right-hand side of (34) by a known probability $p(F_{i,i+1} \mid s_i, F_0)$, which we discuss below. The specific constraints that we add to the dual of the ALP (24)–(26) are thus

$$\sum_{a_i} w_i(x_i, s_i, F_{i,i+1}, a_i) = p(F_{i,i+1} \mid s_i, F_0)\theta_i(x_i, s_i), \quad \forall (i, x_i, s_i, F_{i,i+1}). \quad (35)$$

Constraints (34) are implied by (32) (because (32) implies (33), which implies (34)) and, hence, are satisfied by at least one optimal solution of the exact dual, DLP. In contrast, constraints (35) may not be implied by (32), because the probability $p(F_{i,i+1} \mid s_i, F_0)$ does not depend on the stage i inventory level obtained by an optimal policy. More specifically, $\Pr^*(F_{i,i+1} \mid x_i, s_i)$, which is shorthand for $\Pr^*(F_{i,i+1} \mid x_i^* = x_i, s_i)$, may differ from $\Pr(F_{i,i+1} \mid s_i)$ in general because the (random) inventory level x_i^* reached in stage i by following an optimal policy may be correlated with the prompt-month futures price $F_{i,i+1}$. As a consequence, the optimal objective function of the resulting ALP relaxation may not be an upper bound on the corresponding DDP optimal value function at the initial stage and state, $V_0^D(x_0, F_0)$.

The ALP relaxation obtained by adding constraints (35) to the dual of the ALP (24)–(26) is

$$\min_{\beta, d} \ \phi_0(x_0, s_0) \quad (36)$$

$$\text{s.t.} \ \phi_N(x_N, s_N) = 0, \quad \forall x_N, \quad (37)$$

$$\phi_i(x_i, s_i) + d_i(x_i, s_i, F_{i,i+1}) \geq r(a_i, s_i) + \delta \mathbb{E}\big[\phi_{i+1}(x_i - a_i, s_{i+1}) \,\big|\, F_{i,i+1}\big], \quad \forall (i, x_i, s_i, F_{i,i+1}, a_i), \quad (38)$$

$$\sum_{F_{i,i+1}} p(F_{i,i+1} \mid s_i, F_0)\, d_i(x_i, s_i, F_{i,i+1}) = 0, \quad \forall (i, x_i, s_i). \quad (39)$$

This relaxed ALP differs from the ALP relaxation (21)–(23) in two aspects. First, compared to the decision variables $d_i(x_i, F_i, a_i)$, the variables $d_i(x_i, s_i, F_{i,i+1})$ do not depend on the action a_i and the set of futures prices $F_i \backslash \{s_i, F_{i,i+1}\}$. Second, the amount of relaxation in (36)–(39) is controlled by the constraints (39), whereas in (21)–(23) it is regulated by the second term in the objective function (21). Specifically, the constraints (39) set to zero the weighted average of the variables in set $\{d_i(x_i, s_i, F_{i,i+1}), \ F_{i,i+1}\}$ where the weights are the probabilities in set $\{p(F_{i,i+1} \mid s_i, F_0), F_{i,i+1}\}$ for each pair (i, x_i, s_i). Hence, large relaxations of a subset of constraints (39) corresponding to forward curves that occur with low probability can be balanced by small restrictions of a subset of constraints (39) corresponding to forward curves that happen with high probability.

The ALP relaxation of Desai et al. (2012) also includes variables that relax the ALP constraints, but these variables are nonnegative and hence only capture violations of ALP constraints. In contrast to model (36)–(39), the model of these authors uses a constraint to impose a budget on an average of the constraint violations.

Proposition 5.2 states that an optimal solution to the constraint-based ALP relaxation (36)–(39) can be computed by solving the following ADP, which depends on the conditional probability mass function $\{p(F_{i,i+1} \mid s_i, F_0), F_{i,i+1}\}$:

$$\phi_i^p(x_i, s_i) = \sum_{F_{i,i+1}} p(F_{i,i+1} \mid s_i, F_0) \cdot \max_{a_i} \big\{ r(a_i, s_i) + \delta \mathbb{E}\big[\phi_{i+1}^p(x_i - a_i, s_{i+1}) \big| F_{i,i+1}\big]\big\}, \quad (40)$$

$\forall (i, x_i, s_i)$, with $\phi_N^p(x_N, s_N) := 0$, $\forall x_N$. We define $d_i^p(x_i, s_i, F_{i,i+1})$ as

$$\max_{a_i} \{ r(a_i, s_i) + \delta \mathbb{E}[\phi_{i+1}^p(x_i - a_i, s_{i+1}) \mid F_{i,i+1}]\} - \phi_i^p(x_i, s_i).$$

Proposition 5.2. *The terms $\phi_i^p(x_i, s_i)$ and $d_i^p(x_i, s_i, F_{i,i+1})$ optimally solve* (36)–(39).

In light of Proposition 5.2, comparing (27) and (40) reveals that the constraint-based ALP relaxation (36)–(39) effectively replaces the "odd" maximization over the set $\mathcal{F}_{i,i+1}$ in (27) with an expectation taken with respect to the probability mass function $\{p(F_{i,i+1} \mid s_i, F_0), F_{i,i+1}\}$.

Different constraint-based ALP relaxations can be obtained from (40) by varying the choice of the conditional probability mass function $\{p(F_{i,i+1} \mid s_i, F_0), F_{i,i+1}\}$. We consider the following choices for $p(F_{i,i+1} \mid s_i, F_0)$:

$$\Pr(F_{i,i+1} \mid s_i, F_{0,i+1}), \tag{41}$$

$$\mathbb{1}(F_{i,i+1} = \mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}]). \tag{42}$$

The term $\Pr(F_{i,i+1} \mid s_i, F_{0,i+1})$ is the conditional probability of $F_{i,i+1}$ given s_i and $F_{0,i+1}$ under the forward curve probability mass function discussed at the beginning of §3; $\mathbb{1}(F_{i,i+1} = \mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}])$ is a degenerate conditional probability mass function on the set $\mathcal{F}_{i,i+1}$ that places all its mass on the value $\mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}]$. (This expectation is also under the forward curve probability mass function discussed in §3 and is assumed to be in set $\mathcal{F}_{i,i+1}$.)

Using (41), that is, letting $p(F_{i,i+1} \mid s_i, F_0) = \Pr(F_{i,i+1} \mid s_i, F_{0,i+1})$ in (40), gives the following constraint-based ALP relaxation:

$$\phi_i^{\mathrm{SADP}}(x_i, s_i) = \mathbb{E}\Big[\max_{a_i}\big\{r(a_i, s_i) + \delta\,\mathbb{E}\big[\phi_{i+1}^{\mathrm{SADP}}(x_i - a_i, s_{i+1}) \mid F_{i,i+1}\big]\big\} \mid s_i, F_{0,i+1}\Big], \tag{43}$$

$\forall (i, x_i, s_i)$, with $\phi_N^{\mathrm{SADP}}(x_N, s_N) := 0$, $\forall x_N$. This is the ADP of LMS, that is, SADP (hence the superscript on ϕ_i and ϕ_{i+1} in (43); LMS do not show that SADP is an ALP relaxation).

Using (42), that is, letting $p(F_{i,i+1} \mid s_i, F_0) = \mathbb{1}(F_{i,i+1} = \mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}])$ in (40), gives ADP1:

$$\phi_i^{\mathrm{ADP1}}(x_i, s_i) = \max_{a_i}\big\{r(a_i, s_i) + \delta\,\mathbb{E}\big[\phi_{i+1}^{\mathrm{ADP1}}(x_i - a_i, s_{i+1}) \mid \mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}]\big]\big\}, \tag{44}$$

$\forall (i, x_i, s_i)$, with $\phi_N^{\mathrm{ADP1}}(x_N, s_N) := 0$, $\forall x_N$. ADP1 is a new model.

Our two choices of $p(F_{i,i+1} \mid s_i, F_{0,i+1})$ can be interpreted as restricting the amount of information revealed by the stochastic process (3)–(4) that an ALP relaxation uses to obtain a value function approximation. In both cases, at time T_i the joint probability mass function of the price pair $(s_i, F_{i,i+1})$ conditional on $(F_{0,i}, F_{0,i+1})$ is replaced by the marginal probability mass function of the spot price s_i given $F_{0,i}$ and a conditional probability mass function of the futures price $F_{i,i+1}$ given $(s_i, F_{0,i+1})$: the one based on (41) for SADP and (42) for ADP1.

We now extend ADP1 to ADP2, starting from the ALP (28)–(31). Specifically, we add to the dual of this ALP the following constraints that are related to the constraints (35) used in the derivation of ADP1:

$$\sum_{a_i} w_i(x_i, s_i, F_{i,i+1}, F_{i,i+2}, a_i) = \mathbb{1}\big(F_{i,i+2} = \mathbb{E}\big[F_{i,i+2} \mid s_i, F_{i,i+1}, F_0\big]\big)\theta_i(x_i, s_i, F_{i,i+1}), \quad \forall (i, x_i, s_i, F_{i,i+1}, F_{i,i+2}).$$

ADP2, which can be shown to be equivalent to the ALP relaxation corresponding to this ALP dual restriction, is

$$\phi_{N-1}^{\mathrm{ADP2}}(x_{N-1}, s_{N-1}) = \max_{a_{N-1}} r(a_{N-1}, s_{N-1}), \quad \forall (x_{N-1}, s_{N-1}), \tag{45}$$

$$\phi_{N-2}^{\mathrm{ADP2}}(x_{N-2}, s_{N-2}, F_{N-2,N-1}) = \max_{a_{N-2}}\big\{r(a_{N-2}, s_{N-2}) + \delta\,\mathbb{E}\big[\phi_{N-1}^{\mathrm{ADP2}}(x_{N-2} - a_{N-2}, s_{N-1}) \,\big|\, F_{N-2,N-1}\big]\big\}, \quad \forall (x_{N-2}, s_{N-2}, F_{N-2,N-1}), \tag{46}$$

$$\phi_i^{\mathrm{ADP2}}(x_i, s_i, F_{i,i+1}) = \max_{a_i}\big\{r(a_i, s_i) + \delta\,\mathbb{E}\big[\phi_{i+1}^{\mathrm{ADP2}}(x_i - a_i, s_{i+1}, F_{i+1,i+2}) \,\big|\, F_{i,i+1}, \mathbb{E}\big[F_{i,i+2} \,\big|\, s_i, F_{i,i+1}, F_{0,i+2}\big]\big]\big\}, \quad \forall i \in \mathcal{I} \setminus \{N-2, N-1\}, (x_i, s_i, F_{i,i+1}), \tag{47}$$

with $\phi_N^{\mathrm{ADP2}}(x_N, s_N) := 0$, $\forall x_N$.

6. Error Bound Analysis

In this section we analyze the value function approximations obtained by versions of the ADPs discussed in §§5.2 and 5.3: ADP0, SADP, ADP1, ADP0′, and ADP2. Our analysis provides insights into the relative performance of these ADPs, in particular the benefit of the constraint-based ALP relaxations proposed in §5.3 relative to their corresponding ALPs, and sheds light on when SADP, ADP1, and ADP2 can be expected to perform well.

Let l represent an ADP in the set $\mathcal{L} := \{$SADP, ADP1, ADP2$\}$. Consistent with how EDP is formulated, we analyze versions of ADP0, ADP0′, and the ADPs in set $\mathcal{L}$ reformulated assuming that the forward curve F_i at each stage i belongs to $\mathbb{R}_+^{N-i}$ instead of the finite set $\mathcal{F}_i$ (and hence the first "max" in (27) is assumed to be replaced with "sup"; an analogous substitution is assumed for ADP0′). For simplicity, we continue to use the same labels for these reformulated ADPs. Under a mild assumption satisfied by price model (3)–(4), Proposition 6.1 compares the value function approximations of ADP0 and ADP0′ against the value function of EDP and the value function approximations of the ADPs in set $\mathcal{L}$. The mild assumption in Proposition 6.1 is that the distributions of the random variables $s_{i+1} \mid F_{i,\,i+1}$ and $s_{i+2} \mid F_{i,\,i+2}$ are stochastically increasing in $F_{i,\,i+1}$ and $F_{i,\,i+2}$, respectively (see, e.g., Topkis 1998, Lemma 3.9.1(b)).

Proposition 6.1. (i) *If the distribution of $s_{i+1} \mid F_{i,\,i+1}$ is stochastically increasing in $F_{i,\,i+1} \in \mathbb{R}_+$, for each $i \in \mathcal{I}_{-(N-1)}$, then the ADP0 value function is unbounded in every state in stages zero through $N-2$.* (ii) *If the distribution of $s_{i+2} \mid F_{i,\,i+2}$ is stochastically increasing in $F_{i,\,i+2} \in \mathbb{R}_+$, for each $i \in \mathcal{I}\backslash\{N-1,\ N-2\}$, then the ADP0′ value function is unbounded in every state in stages zero through $N-3$.* (iii) *The value functions of EDP and the ADPs in set $\mathcal{L}$ are bounded in every stage and state.*

Parts (i) and (ii) of Proposition 6.1 are consistent with the discussion given after Proposition 5.1. Together with part (iii) they suggest that there is potential benefit in using constraint-based ALP relaxations rather than an ALP. The value functions of ADP0 and ADP0′ are unbounded because of the assumption made in this section that prices are continuous and unbounded—if prices were bounded these value functions would also be bounded but might still be of low quality, as we find in our numerical experiments discussed in §§7.2 and 7.3, where we solve discretized versions of these models. We thus focus on providing approximation guarantees for the value functions of the ADPs in set $\mathcal{L}$. Our approximation guarantees are based on the norm $\|g\|_{\mathbb{E},\,\infty}$, which we define as $\max_x \mathbb{E}[g(x,F_i)\mid F_0]$, where $g(x,F_i)$ is a generic function with support on the stage i EDP state space and the expectation is with respect to the distribution of F_i conditional on F_0. Specifically, we analyze the following errors between the stage i EDP value function V_i and each ADP l value function ϕ_i^l:

$$\|V_i-\phi_i^l\|_{\mathbb{E},\,\infty} := \begin{cases} \max\limits_{x_i} \mathbb{E}\big[|V_i(x_i,F_i)-\phi_i^l(x_i,s_i)|\,\big|\,F_0\big], \\ \quad l\in\{\text{ADP1},\text{SADP}\}, \\ \max\limits_{x_i} \mathbb{E}\big[|V_i(x_i,F_i)-\phi_i^{\text{ADP2}}(x_i,s_i,F_{i,\,i+1})|\,\big|\,F_0\big], \\ \quad l=\text{ADP2}. \end{cases}$$

We refer to $\|V_i-\phi_i^l\|_{\mathbb{E},\,\infty}$ as the l-error at stage i.

Our analysis is based on the concept of *ideal* value function approximation $\phi_i^{l,\,V}$ for each ADP $l\in\mathcal{L}$. Each such function is defined by replacing with V_{i+1} the function ϕ_{i+1}^l on the right-hand side of each ADP l recursion, that is, (43), (44), and (45)–(47) reformulated as discussed above, and modifying the conditional expectations accordingly:

$$\phi_i^{\text{SADP},\,V}(x_i,s_i) := \mathbb{E}\Big[\max_{a_i}\{r(a_i,s_i) + \delta\,\mathbb{E}\big[V_{i+1}(x_i-a_i,F_{i+1})\,\big|\,F_i\big]\}\,|\,s_i,F_0\Big],$$

$$\phi_i^{\text{ADP1},\,V}(x_i,s_i) := \max_{a_i}\{r(a_i,s_i) + \delta\,\mathbb{E}\big[V_{i+1}(x_i-a_i,F_{i+1})\,\big|\,\bar{F}_i'(s_i,F_0)\big]\},$$

$$\phi_i^{\text{ADP2},\,V}(x_i,s_i,F_{i,\,i+1}) := \max_{a_i}\{r(a_i,s_i)+\delta\,\mathbb{E}\big[V_{i+1}(x_i-a_i,F_{i+1})\,\big|\,F_{i,\,i+1},\bar{F}_i''(s_i,F_{i,\,i+1},F_0)\big]\},$$

where defining $F_i' := \{F_{i,i+1},F_{i,i+2},\ldots,F_{i,N-1}\}$ we denote as $\bar{F}_i'(s_i,F_0)$ the quantity $\mathbb{E}[F_i'|s_i,F_0]$, and defining $F_i'' := \{F_{i,i+2},F_{i,i+3},\ldots,F_{i,N-1}\}$ we label as $\bar{F}_i''(s_i,F_{i,i+1},F_0)$ the term $\mathbb{E}[F_i''|s_i,F_{i,i+1},F_0]$.

We now bound the various l-errors using recursive functions that depend on the absolute value of the differences between the EDP value function and the l ideal value function approximations. These recursive functions for SADP and ADP1 are defined, $\forall\,(i,x_i,F_i)_{-(N-1)}$, as

$$\gamma_i^{\text{SADP}}(x_i,F_i) := \big|V_i(x_i,F_i)-\phi_i^{\text{SADP},\,V}(x_i,s_i)\big| + \delta\,\mathbb{E}\Big[\max_{x_{i+1}}\mathbb{E}\big[\gamma_{i+1}^{\text{SADP}}(x_{i+1},F_{i+1})|F_i'\big]|s_i,F_0\Big],$$

$$\gamma_i^{\text{ADP1}}(x_i,F_i) := \big|V_i(x_i,F_i)-\phi_i^{\text{ADP1},\,V}(x_i,s_i)\big| + \delta\max_{x_{i+1}}\mathbb{E}\big[\gamma_{i+1}^{\text{ADP1}}(x_{i+1},F_{i+1})|\bar{F}_i'(s_i,F_0)\big],$$

with boundary conditions $\gamma_{N-1}^l(\cdot) := 0$, $l\in\{\text{SADP},\text{ADP1}\}$. For ADP2, this recursive function is defined, $\forall\,i\in\mathcal{I}\backslash\{N-1,N-2\}$, (x_i,F_i), as

$$\gamma_i^{\text{ADP2}}(x_i,F_i) := \big|V_i(x_i,F_i)-\phi_i^{\text{ADP2},\,V}(x_i,s_i,F_{i,\,i+1})\big| + \delta\max_{x_{i+1}}\mathbb{E}\big[\gamma_{i+1}^{\text{ADP2}}(x_{i+1},F_{i+1})\,\big|\,F_{i,\,i+1},\bar{F}_i''(s_i,F_{i,\,i+1},F_0)\big],$$

with boundary conditions $\gamma_i^{\text{ADP2}}(\cdot)\equiv 0$, $\forall\,i\in\{N-2,N-1\}$. As shown in Appendix B, our bound on the stage i l-error is

$$\|V_i-\phi_i^l\|_{\mathbb{E},\infty} \le \|\gamma_i^l\|_{\mathbb{E},\,\infty}, \tag{48}$$

$\forall\,(i,l)\in\mathcal{I}\times\mathcal{L}$. The bound (48) formalizes the intuition that the ADPs in set $\mathcal{L}$ incur an error when the exact value function differs from its corresponding l ideal value function approximation at a given stage and state. This bound is finite (by Proposition 6.1, part (iii)).

The bound (48) is zero in the last stage $(N-1)$ for all the ADPs in set $\mathcal{L}$ and also in the penultimate stage $(N-2)$ for ADP2. Proposition 6.2 identifies limiting regimes under price model (3)–(4) for which the bound (48) tends to zero in all other stages. We denote by ρ the matrix of the instantaneous correlations between the standard Brownian motion increments of price model (3)–(4). We let $\bar{\rho}$ be a rank-two matrix ρ such that each of its elements $\bar{\rho}_{i,i+1}$ satisfies $|\bar{\rho}_{i,i+1}| < 1$. We also denote by 1 a matrix of ones that is compatible with ρ.

Proposition 6.2. *Under price model* (3)–(4) *it holds that* (i) $\lim_{\rho\to 1}\|\gamma_i^l\|_{\mathbb{E},\infty}=0$ *for all* $l\in\mathcal{L}$ *and* $i\in\mathcal{I}\setminus\{N-1\}$, *and* (ii) $\lim_{\rho\to\bar{\rho}}\|\gamma_i^{\mathrm{ADP2}}\|_{\mathbb{E},\infty}=0$ *for all* $i\in\mathcal{I}\setminus\{N-1,N-2\}$.

Part (i) of Proposition 6.2 suggests that the ADPs in $\mathcal{L}$ should perform near optimally in the practically relevant limiting regime when the instantaneous correlations in price model (3)–(4) are sufficiently large and positive (close to one)—for instance, these correlations are high when estimated on natural gas futures prices data (see, e.g., LMS). The intuition for this conclusion is that at the stated limit there is a single source of uncertainty in price model (3)–(4), and hence the current spot price is a sufficient statistic for the future evolution of the entire forward curve. Part (ii) of this proposition suggests that ADP2 is also near optimal even if not all the instantaneous correlation coefficients in price model (3)–(4) are large and positive. The intuition here is that at the asserted limit there are two sources of uncertainty in this price model and the spot price and the prompt-month futures price, which are defining elements of the ADP2 value function, are not sufficient statistics for each other. Because the ADP1 and SADP value functions depend on the spot price but not the prompt-month futures price, this result thus provides theoretical support for ADP2 outperforming both SADP and ADP1.

Our use of an ideal value function approximation to bound the l-error is similar in spirit to the approach taken in de Farias and Van Roy (2003) and Desai et al. (2012) to bound the error incurred by the value function approximation determined by their models. However, the error bounds of these authors do not apply to the ADPs in set $\mathcal{L}$ because (i) de Farias and Van Roy (2003) provide bounds for an ALP whereas we analyze ADPs corresponding to ALP relaxations, and (ii) Desai et al. (2012) study an ALP relaxation that is different from the ones that we consider, as discussed in §5.3. Likewise, our bound (48) is specific to the ADPs considered here.

Our error bounds are also related to the analysis of Tsitsiklis and Van Roy (2001) and Munos (2007) of the approximate value iteration algorithm and the investigation of Petrik (2012) of a method that approximates a value function by optimizing a metric related to the quality of its corresponding greedy policy. The bounding terms of Tsitsiklis and Van Roy (2001) are similar to ours but are based on a weighted two-norm and a Bellman operator specific to the American option valuation context they consider. The error bounds of Munos (2007) and Petrik (2012) rely on differences between an approximate value function and the function obtained by applying the Bellman operator to this approximate value function. In contrast, our bounding terms are defined by differences between the exact value function and the function returned from a single application of a modified Bellman operator to the exact value function.

7. Numerical Results

In this section we discuss our computational results. We present our test instances in §7.1. In §§7.2 and 7.3, respectively, we investigate the upper and lower bounding performance of the models presented in §§5.2 and 5.3. In §7.3 we also consider the RI policy and the LSMC approach. We analyze the computational requirement of all these models in §7.4.

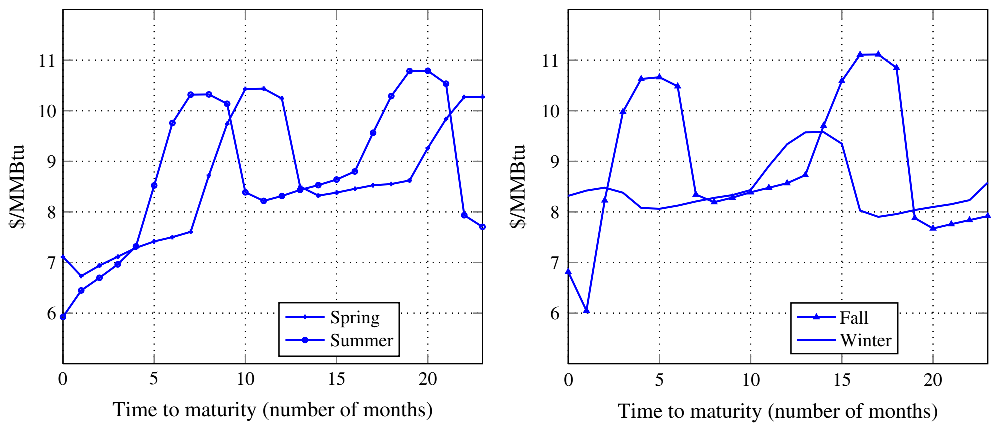
7.1. Instances

We consider the 24-stage LMS instances, which are based on natural gas data from the New York Mercantile Exchange (NYMEX) and the energy trading literature. Each instance is identified by a season (spring, summer, fall, or winter) and one of the three injection and withdrawal capacity pairs $(-0.15, 0.30)$, $(-0.30, 0.60)$, and $(-0.45, 0.90)$, labeled 1, 2, and 3, respectively. The injection and withdrawal costs are \$0.02 and \$0.01 per MMBtu, respectively, and the injection and withdrawal loss coefficients are 1.01 and 0.99, respectively. The initial inventory level is zero. These instances are based on the futures price model (3)–(4). Figure 2 displays the NYMEX natural gas forward curves on March 1, 2006 (spring), June 1, 2006 (summer), August 31, 2006 (fall), and December 1, 2006 (winter); the estimated volatilities and instantaneous correlations for this model can be found in LMS. The risk-free discount rates are 4.74%, 5.05%, 5.01%, and 4.87% in spring, summer, fall, and winter, respectively.

7.2. Upper Bounds

Like LMS, we use 10,000 forward curve sample paths to obtain our dual bound estimates on the value of storage in the initial stage and state. Across all the considered instances, the ADP0-based dual bound estimates are between 30% and 690% larger than the worst dual bound estimates obtained with ADP1 and SADP, and the ADP0′-based dual bound estimates are between 21% and 600% larger than the ADP2-based dual bound estimates. Thus, on these instances, the value function approximations of the considered ALP relaxations lead to substantially tighter dual bound estimates than the value function approximations of their respective ALPs. These findings are consistent with our error bound analysis carried out in §6.

Figure 2 (Color online) NYMEX Natural Gas Forward Curves on March 1, 2006 (Spring), June 1, 2006 (Summer), August 31, 2006 (Fall), and December 1, 2006 (Winter)



Source. LMS.

We denote by UBS, UB1, and UB2 the dual bound estimates associated with SADP, ADP1, and ADP2, respectively. Figure 3 displays UBS and UB1 on all the considered instances as percentages of UB2, which is tighter than all the other estimated upper bounds and, as will become evident in §7.3, is near optimal. The error bars in this figure indicate the standard errors, also reported as percentages of UB2. UBS and UB1 are indistinguishable on all the instances after accounting for sampling variability. UB2 is better than both UBS and UB1 by an average of 2.82% on the winter instances, whereas this average is smaller on the other instances. We are thus able to obtain substantially improved upper bound estimates compared to LMS on the winter instances. The observed quality of UB2 relative to the ones of UBS and UB1 is consistent with our error bound analysis performed in §6.

7.3. Lower Bounds

We also use 10,000 sample paths to obtain our lower bound estimates on the value of storage in the initial stage and state. Across all the considered instances, the ADP0-based lower bound estimates are between 25% and 100% smaller than the worst lower bound

Figure 3 (Color online) Estimated Upper Bounds and Their Standard Errors (Error Bars)

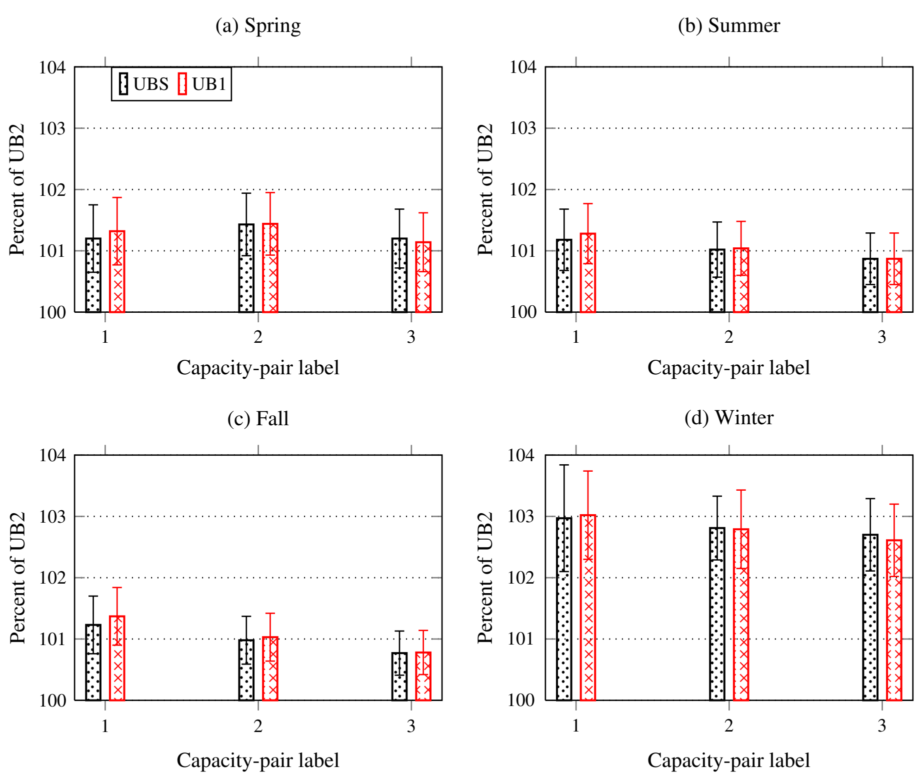
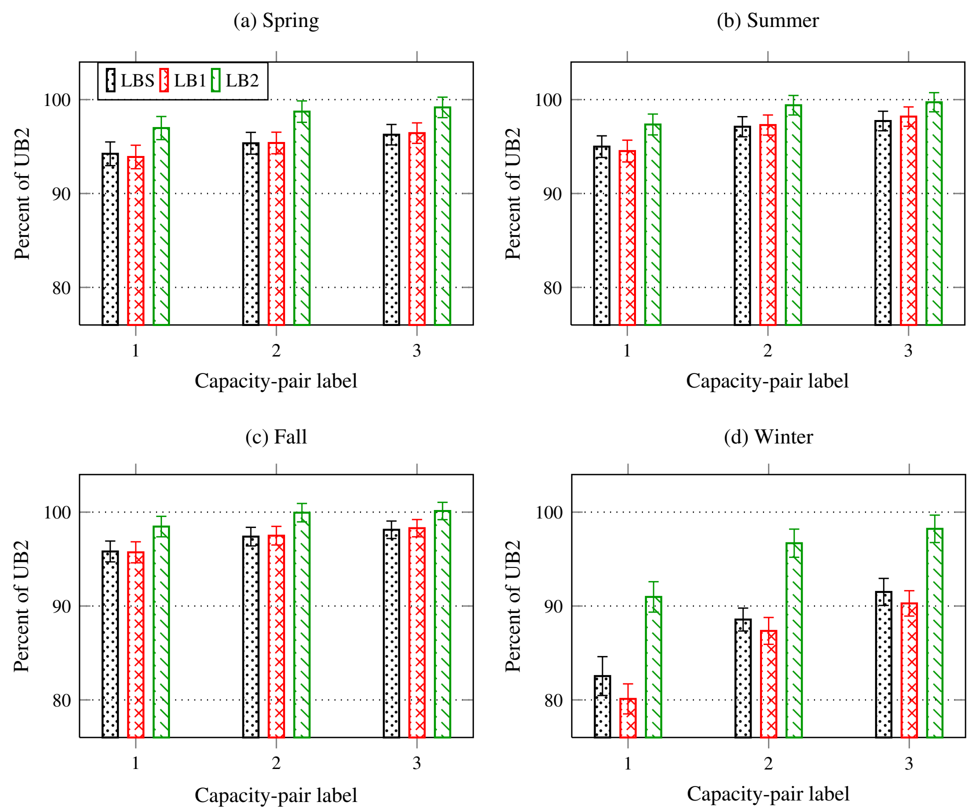


Figure 4 (Color online) Estimated Lower Bounds and Their Standard Errors (Error Bars) Without Reoptimization

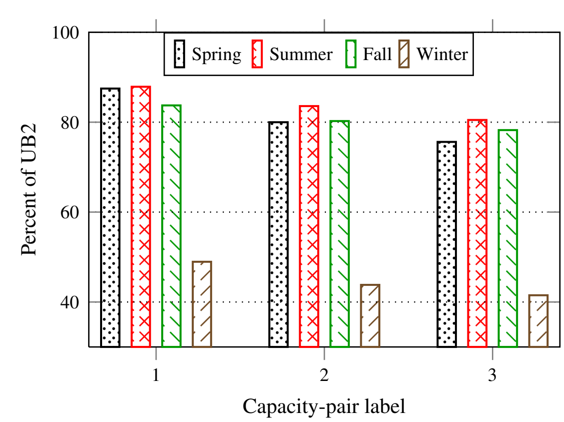


estimates obtained with ADP1 and SADP, and the ADP0′-based lower bound estimates are between 5% and 89% smaller than the ADP2-based lower bound estimates. The control policies obtained from the ALP relaxations are thus substantially better than the control policies based on their respective ALPs on these instances. These results are in line with our error bound analysis performed in §6.

We denote by LBS, LB1, and LB2 the lower bound estimates obtained using SADP, ADP1, and ADP2, respectively. Figure 4 displays these estimates as percentages of UB2. The error bars in this figure indicate the standard errors of these estimates as percentages of UB2. The difference between LBS and LB1 is less than one standard error (expressed as a ratio of UB2) on the spring, summer, and fall instances, whereas LB1 is weaker than LBS by no more than 2.44% of UB2 on the winter instances. LB2 outperforms both LBS and LB1 on all the considered instances: the improvement of LB2 on LBS is 2.00 to 3.36% across the spring, summer, and fall instances, and 6.72 to 8.43% on the winter instances. The improvements of LB2 on LB1 are similar on the spring, summer, and fall instances, but are larger on the winter instances. These results suggest that ADP2 is a better model than both SADP and ADP1, with maximum suboptimality gaps of 3.03% of UB2 on the spring, summer, and fall instances, and 9.03% of UB2 on the winter instances. In contrast, these suboptimalities are 5.77% and 17.46% for SADP, and 6.11% and 19.89% for ADP1.

The relative performance of ADP2 against ADP1 and SADP is consistent with part (ii) of Proposition 6.2. To shed some more light on the difference between the ADP2-based and SADP/ADP1-based estimated lower bounds on the winter instances relative to the other instances, Figure 5 reports for each instance the intrinsic value of storage, that is, the value of storage due to seasonality (deterministic price variability). This value is obtained by solving a

Figure 5 (Color online) Intrinsic Values

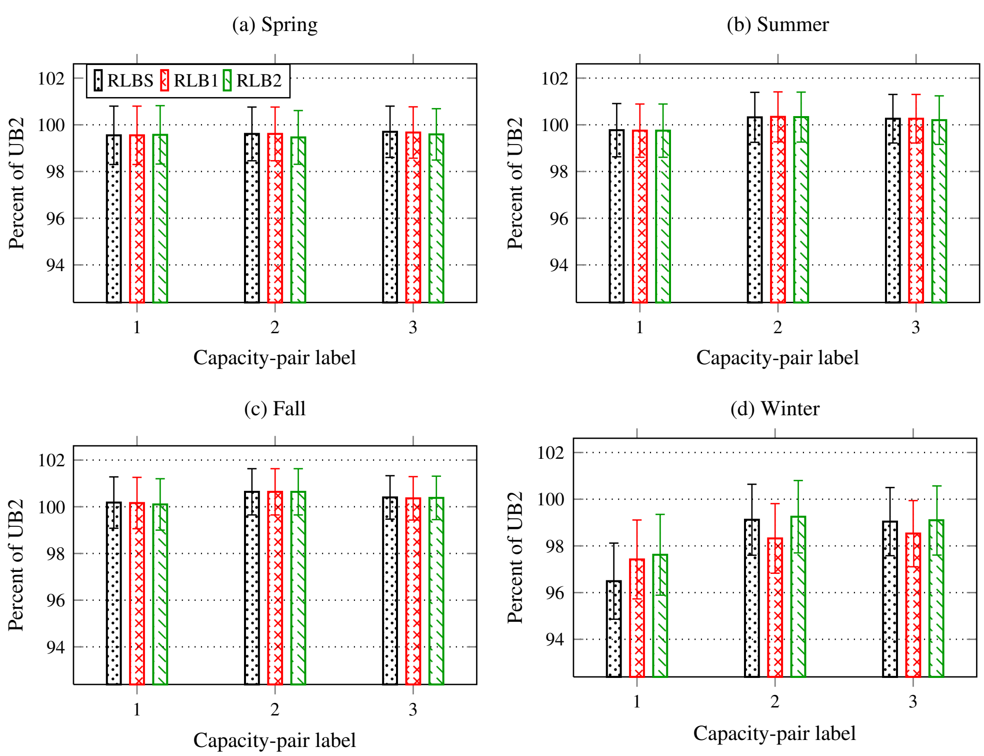


deterministic version of EDP, model (2), in which the spot price for each stage is fixed to its corresponding futures price in the initial (time zero) forward curve. Because futures prices are risk-neutral expected spot prices (Shreve 2004, p. 244), the value function of this deterministic model at the initial stage and state is the exact value of applying its optimal policy, which is known as the intrinsic policy, when prices evolve stochastically. The computed intrinsic values are less than 50% of UB2 on the winter instances, whereas they are at least 75% of UB2 on the remaining instances. The lower winter intrinsic value relative to the other seasons is due to the less pronounced seasonality (humps) in the winter forward curve compared to the forward curves for the other seasons (see Figure 2). Thus, a substantially larger portion of the storage value is attributable to price uncertainty for the winter instances than for the other instances. In other words, capturing the evolution of the forward curve appears to be more important on the winter instances than on the other instances. Because the ADP2 value function approximation depends on both the spot and prompt futures prices, whereas the ones of SADP and ADP1 depend only on the spot price, ADP2 is better able to capture the evolution of the forward curve.

We denote by RLBS, RLB1, and RLB2 the estimates of the reoptimization versions of LBS, LB1, and LB2, respectively. Figure 6 displays these reoptimization-based lower bound estimates and their standard errors as percentages of UB2 (some of the reported lower bound estimates exceed UB2 because of Monte Carlo sampling error). RLBS, RLB1, and RLB2 are almost tight on the spring, summer, and fall instances, which brings to light the near optimality of UB2. RLB2 is slightly better than both RLBS and RLB1 on the winter instances, with a maximum optimality gap of 2.38% of UB2 compared to 3.51% for RLBS and 2.58% for RLB1. Further, LB2 is worse than RLB2 by 0.20 to 6.65% of UB2 on all the instances, and LBS and LB1, respectively, fall below RLBS and RLB1 by 2.29 to 13.94% and 1.28 to 14.51% of UB2 on all the instances. Thus, whereas reoptimization can be useful even for ADP2, it appears to be less critical for ADP2 than it is for SADP and ADP1 to obtain near optimal lower bound estimates and policies.

We now compare the ADP2-based lower bounds against the ones estimated using the RI policy and the LSMC approach (see §1 for references). The LSMC method approximates the continuation function, that is, $\delta\mathbb{E}[V_{i+1}(x_{i+1}, F_{i+1}) \mid F_i]$ for each (i, x_{i+1}, F_i), using basis functions that for every stage and inventory level include polynomials of order one and two in each futures price. Across all the considered instances, the averages of the estimated lower bounds (as percentages of UB2) corresponding to the RI policy and the LSMC method, respectively, are 99.14% and 98.83% (the standard errors of the individual lower bound estimates vary between 0.77% and 1.76% of UB2). The analogous averages for LB2 and RLB2

Figure 6 (Color online) Estimated Lower Bounds and Their Standard Errors (Error Bars) with Reoptimization



are 97.98% and 99.59%, respectively. The ADP2-based lower bounds are thus competitive with the ones obtained by these state-of-the-art techniques.

7.4. Computational Times

The models that we solve numerically are formulated on discretized state and action spaces (but we estimate our bounds via Monte Carlo simulation from the continuous price model (3)–(4)). As LMS, we optimally discretize the feasible inventory set into 21 equally spaced points, however further reducing the considered inventory levels by eliminating at each stage the ones that cannot be feasibly reached from the initial stage and state, and obtain discretized price sets using binomial lattices (see Appendix A for details), but also applying lattice restrictions (Levy 2004) to shorten the time required to solve ADP2. Our computational setup is a 64 bits PowerEdge R515 with 12 AMD Opteron 4176 2.4 GHz processors, of which we used only one, with 64GB of memory, the Linux Fedora 15 operating system, and the g++ 4.6.1 20110908 (Red Hat 4.6.1-9) compiler. The reported SADP results are based on running the LMS code within our computational setup. Given their poor bounding performance, we exclude ADP0 and ADP0′ from the ensuing discussion.

Table 2 summarizes the computational effort for SADP, ADP1, ADP2, the RI policy, and the LSMC method. Without reoptimization, ADP1 and the LSMC method approximate a value/continuation function in a comparable amount of time, which is at least 100 times and 30 times smaller, respectively, than the time taken by SADP and ADP2. ADP1 estimates an upper bound nine and six times faster than SADP and ADP2, respectively. SADP, ADP1, and the LSMC method exhibit similar effort when estimating a lower bound, and this effort is at least three times smaller than the requirement of ADP2. The ADP2 lattice restrictions yield a speed up equal to one order of magnitude without appreciably affecting the estimated bounds. With reoptimization, the estimation of a lower bound using the RI policy is at least 3, 12, and 34 times faster than estimating a lower bound using SADP, ADP1, and ADP2, respectively. In contrast to the no-reoptimization case, ADP2 is slower than SADP when computing the reoptimized lower bounds, because the ADP2 lattice restrictions are comparably less effective for the coarse price grids employed when reoptimizing these models (see LMS).

Table 2 Central Processing Unit Seconds (Ranges) for SADP, ADP1, ADP2, the RI Policy, and the LSMC Method

	No reoptimization			
Model	Continuation value function approximation	Upper bound	Lower bound	Reoptimized lower bound
SADP	120–122	153–194	3–5	425–498
ADP1	1	9–16	1	90–93
ADP2	36–53	102–158	18–21	1,187–1,201
LSMC	1	—	4–9	—
RI	—	—	—	16–34

8. Conclusions

Real option management of commodity storage assets is an important practical problem that, in general, gives rise to an intractable MDP when using high-dimensional models of commodity forward curve evolution. We develop a novel approximate dynamic programming approach to derive ALP relaxations. Our approach relies on approximately enforcing on the ALP dual a property of the exact dual. We derive tractable ALP relaxations by applying our approach using low-dimensional look-up table value function approximations, subsuming an existing approximate dynamic programming model. We derive error bounds that provide theoretical support for using our ALP relaxations rather than their respective ALPs. Our numerical results on existing natural gas storage instances are promising, showing that our ALP relaxations substantially outperform their associated ALPs, with our best ALP relaxation matching or improving on the best lower and upper bounds available in the literature for these instances, and are also competitive with, but typically slower than, state-of-the-art methods for obtaining heuristic policies and estimating lower bounds on the value of commodity storage. Our research has potential relevance beyond this specific application.

Acknowledgments

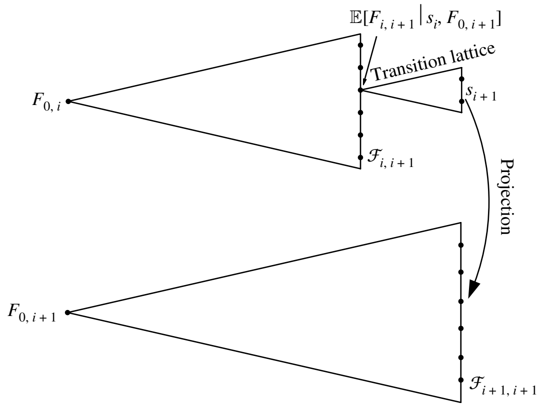
The authors thank the editors and reviewers for their constructive feedback that led to a substantially improved version of this paper. This research was supported by the National Science Foundation [Grant CMMI 1129163].

Appendix A. Computational Complexity Analysis

In this section we discuss the computational complexity of solving the ALP relaxations presented in §5.3 and estimating greedy and dual bounds. This complexity depends on the specific technique used for discretizing the relevant price sets. Our computational study in §7 is based on discretizing the multimaturity Black (1976) price model (3)–(4) via Rubinstein (1994) binomial lattices. We thus focus on this discretization approach. Our analysis uses the (easy to establish) property that the policies associated with SADP, ADP1, and ADP2 share the base-stock target structure of an optimal DDP policy (for details, see Proposition 4 and Lemma 2 in Secomandi et al. 2015).

Consider ADP1. Figure A.1 illustrates our discretization approach. We construct the set $\mathcal{F}_{i,i}$ (discretize $\mathbb{R}_+$) by evolving the time zero futures price $F_{0,i}$ using a two-dimensional

Figure A.1 Illustration of Our Discretization Approach for ADP1



Rubinstein binomial tree based on the volatility σ_i (see the top part of Figure A.1). Let m_i be the number of time steps used to discretize the time interval $[0, T_i]$. Building this lattice results in a set $\mathscr{F}_{i,i}$ with $m_i + 1$ values and requires $O(m_i)$ operations.

At each stage i, solving ADP1 entails executing the following steps:

Step 1. Determine a probability mass function with support $\mathscr{F}_{i+1,i+1}$ for the random variable s_{i+1} given $\mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}]$ for all s_i.

Step 2. Compute the optimal ADP1 base-stock targets for all s_i.

Step 3. Evaluate $\phi_i^{\text{ADP1}}(x_i, s_i)$ for all (x_i, s_i).

In Step 1, we evolve a two-dimensional Rubinstein lattice, starting from each price $\mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}]$, referred to as the transition lattice, by using m time steps to discretize the interval $[T_i, T_{i+1}]$ (see the top part of Figure A.1). We compute each price $\mathbb{E}[F_{i,i+1} \mid s_i, F_{0,i+1}]$ in closed form in $O(1)$ operations under the price model (3)–(4). Each transition lattice yields a discretization of s_{i+1} with $m+1$ values. Building all the $m_i + 1$ transition lattices thus takes $O(m_i \cdot m)$ operations. To obtain the distribution of s_{i+1} given $\mathbb{E}[F_{i,i+1}|s_i, F_{0,i+1}]$ with support on $\mathscr{F}_{i+1,i+1}$, we project each price s_{i+1} in each transition lattice onto the set $\mathscr{F}_{i+1,i+1}$ by rounding each such price to the closest spot price in $\mathscr{F}_{i+1,i+1}$ (see Figure A.1). The set $\mathscr{F}_{i+1,i+1}$ is constructed in a manner analogous to how we generate the set $\mathscr{F}_{i,i}$, but using the parameters m_{i+1}, T_{i+1}, $F_{0,i+1}$, and σ_{i+1} (see the bottom part of Figure A.1). Since the s_{i+1} values in each transition lattice and the set $\mathscr{F}_{i+1,i+1}$ are sorted, this projection takes a total of $O(m_{i+1} \cdot m)$ operations. Therefore, the time complexity for Step 1 is $O(m_i \cdot m + m_{i+1} \cdot m)$.

Executing Step 2 requires performing the maximization in (44) at inventory levels 0 and $\bar{x}$ with the injection and withdrawal capacities relaxed to $-\bar{x}$ and $\bar{x}$, respectively, which requires $O(m_i \cdot |\mathscr{X}| \cdot m)$ operations. Executing Step 3 also takes $O(m_i \cdot |\mathscr{X}| \cdot m)$ operations. Therefore, computing $\phi_i^{\text{ADP1}}(x_i, s_i)$ for all (x_i, s_i) involves $O(m \cdot (m_i + m_{i+1} + 2 \cdot m_i \cdot |\mathscr{X}|))$ operations. Using $m' := \max_{i \in \mathscr{I}} m_i$, this number of operations simplifies to $O(m' \cdot |\mathscr{X}| \cdot m)$, since $|\mathscr{X}| \geq 2$. Thus, for an N-stage problem, solving ADP1 entails $O(N \cdot m' \cdot |\mathscr{X}| \cdot m)$ operations.

For SADP and ADP2, each set $\mathscr{F}_{i,i} \times \mathscr{F}_{i,i+1}$ is determined using a three-dimensional Rubinstein lattice. For SADP, two-dimensional binomial lattices and projections yield the probability mass function of s_{i+1} conditional on each of the $(m_i + 1)^2$ values of $F_{i,i+1}$. For ADP2 we use three-dimensional lattices and projections to obtain the joint probability mass function of each pair $(s_{i+1}, F_{i+1,i+2})$ on the support $\mathscr{F}_{i+1,i+1} \times \mathscr{F}_{i+1,i+2}$ conditional on each pair $(F_{i,i+1}, \mathbb{E}[F_{i,i+2} \mid s_i, F_{i,i+1}, F_{0,i+2}])$. An analysis similar to the one performed for ADP1 shows that SADP and ADP2 can be solved in $O(N \cdot (m')^2 \cdot |\mathscr{X}|^2 \cdot m)$ and $O(N \cdot (m')^2 \cdot |\mathscr{X}|^2 \cdot m^2)$ operations, respectively. Thus, the ordering of SADP, ADP1, and ADP2 in terms of increasing computational complexity is ADP1, SADP, and ADP2.

The operations count for estimating upper and lower bounds depends on the number of prices included in a look-up table value function approximation. Let n^{S} denote the number of price sample paths used in a Monte Carlo simulation used to estimate a greedy bound and a dual bound (see §2.2). Different from how we obtain each discretization $\mathscr{F}_{i,i}$, this simulation is based on evolving the entire forward curve. A simple analysis shows that estimating lower and upper bounds, respectively, when using the look-up table value function approximation $\phi_i(x_i, s_i)$ requires $O(n^{\text{S}} \cdot N \cdot \log m' + n^{\text{S}} \cdot N \cdot |\mathscr{X}| \cdot m)$ and $O(n^{\text{S}} \cdot N \cdot |\mathscr{X}| \cdot \log m' + n^{\text{S}} \cdot N \cdot |\mathscr{X}|^2 \cdot m)$ operations ($O(\log m')$ operations are needed by binary search when projecting a transition lattice); doing this estimation when using the look-up table value function approximation $\phi_i(x_i, s_i, F_{i,i+1})$ involves $O(n^{\text{S}} \cdot N \cdot \log m' \cdot m + n^{\text{S}} \cdot N \cdot |\mathscr{X}| \cdot m^2)$ and $O(n^{\text{S}} \cdot N \cdot |\mathscr{X}| \cdot \log m' \cdot m + n^{\text{S}} \cdot N \cdot |\mathscr{X}|^2 \cdot m^2)$ operations, respectively. Therefore, estimating dual bounds is more costly than estimating greedy bounds, due to the computation of the dual value function used in (6) at each inventory level in the set $\mathscr{X}$ and for all the stages in set $\mathscr{I}$ given a price sample path P_0. Reasonable values of the parameters n^{S}, $|\mathscr{X}|$, and m' satisfy $n^{\text{S}} \cdot |\mathscr{X}| \geq m'$. Hence, estimating dual bounds is also more costly than solving each of SADP, ADP1, and ADP2.

Appendix B. Proofs

Proof of Proposition 4.2. Suppose Assumption 4.1 is true. We proceed by induction to prove (17). The result is clearly true at stage zero. Suppose the result is also true for all stages $1, \ldots, i-1$. At stage i, for a given (i, x_i) the DALP constraint (13) corresponding to the first basis function, that is, $b = 1$, is

$$\sum_{(F_i, a_i)} w_i(x_i, F_i, a_i) = \delta \underbrace{\sum_{F_{i-1}} \sum_{F_i} \Pr(F_i | F_{i-1})}_{=1} \sum_{(x_{i-1}, a_{i-1})} \mathbb{1}(x_{i-1} - a_{i-1} = x_i) \cdot w_{i-1}(x_{i-1}, F_{i-1}, a_{i-1}).$$

Summing over x_i on both sides of this constraint and simplifying gives

$$\begin{aligned} \sum_{(x_i, F_i, a_i)} w_i(x_i, F_i, a_i) &= \delta \sum_{(x_{i-1}, F_{i-1}, a_{i-1})} \sum_{x_i} \mathbb{1}(x_{i-1} - a_{i-1} = x_i) \cdot w_{i-1}(x_{i-1}, F_{i-1}, a_{i-1}) \\ &= \delta \sum_{(x_{i-1}, F_{i-1}, a_{i-1})} w_{i-1}(x_{i-1}, F_{i-1}, a_{i-1}) = \delta^i, \end{aligned}$$

where the last equality follows from the induction hypothesis. Condition (17) is thus true for stage i. This condition holds for all the stages by the principle of mathematical induction.

We proceed by contradiction to prove (18). Suppose there exists an optimal solution w^* to DALP and a forward curve $\check{F}_i \notin \mathcal{F}_i^{=}(\beta^*)$ such that $\sum_{(x_i,a_i)} w_i^*(x_i, \check{F}_i, a_i) > 0$. This implies that there exists at least one pair $(\check{x}_i, \check{a}_i)$ such that $w_i^*(\check{x}_i, \check{F}_i, \check{a}_i) > 0$. Since (14) and (17) imply that the feasible set of DALP is bounded, we can write $w^* = \sum_{j\in\mathcal{J}} \lambda_j w^j$, where $\sum_{j\in\mathcal{J}} \lambda_j = 1$, $\lambda_j \geq 0$, $\mathcal{J}$ is the index set for the set of basic feasible solutions, and w^j is the jth basic feasible solution. The optimality of w^* implies that every w^j such that $\lambda_j > 0$ must also be a basic optimal solution. Further, the inequality $w_i^*(\check{x}_i, \check{F}_i, \check{a}_i) > 0$ implies that at least one of these basic optimal solutions must satisfy $w_i^j(\check{x}_i, \check{F}_i, \check{a}_i) > 0$. It follows from complementary slackness that the primal constraint corresponding to $(i, \check{x}_i, \check{F}_i, \check{a}_i)$ holds as an equality, which contradicts $\check{F}_i \notin \mathcal{F}_i^{=}(\beta^*)$. □

Proof of Proposition 4.3. Suppose there exists a DDP optimal policy and a DALP optimal solution that satisfy (16). This assumption implies that the DLP and DALP optimal objective function values match. As discussed in §3, DALP is a relaxation of DLP. Hence, every optimal solution of DLP is also optimal for DALP. Let w^* be the basic DALP optimal solution that corresponds to the deterministic DDP optimal policy π^*. Since π^* is deterministic it can be equivalently represented by the set of stage-state-action tuples $\mathcal{K} := \{(i, x_i, F_i, a_i)\colon w_i^*(x_i, F_i, a_i) > 0\}$. By complementary slackness, the ALP constraints corresponding to tuples in $\mathcal{K}$ hold as equalities. Hence, for each stage-state-action tuple (i, x_i, F_i, a_i) in set $\mathcal{K}$ the action a_i is greedy optimal at stage i and state (x_i, F_i) with respect to the value function approximation corresponding to β^*. Therefore, starting from state (x_0, F_0) in stage zero, at each visited state in each stage we can choose an action that is greedy optimal relative to the value function approximation given by β^* such that the encountered stage-state-action tuples belong to $\mathcal{K}$. Hence, it holds that $\pi^* \in \Pi^g(\beta^*)$. □

Proof of Proposition 5.1. The constraints (25) and (26), respectively, can be equivalently rewritten as

$$\phi_{N-1}(x_{N-1}, s_{N-1}) \geq \max_{a_{N-1}} r(a_{N-1}, s_{N-1}), \quad \forall (x_{N-1}, s_{N-1}), \quad \text{(B1)}$$

$$\phi_i(x_i, s_i) \geq \max_{F_{i,i+1}} \Big\{ \max_{a_i} \big\{ r(a_i, s_i) + \delta \mathbb{E}\big[\phi_{i+1}(x_i - a_i, s_{i+1}) \,\big|\, F_{i,i+1}\big]\big\}\Big\}, \quad \forall (i, x_i, s_i)_{-(N-1)}. \quad \text{(B2)}$$

These inequalities hold as equalities when evaluated using ϕ^{ADP0}. Moreover, (i) the variable $\phi_{N-1}(x_{N-1}, s_{N-1})$ is implicitly multiplied by one, a positive coefficient, on the left-hand side of inequalities (B1); and (ii) the variable $\phi_i(x_i, s_i)$ is multiplied by positive coefficients on both the left-hand sides of the inequalities (B2) corresponding to (i, x_i, s_i) and the right-hand sides of the inequalities (B2) corresponding to stage $i-1$. Therefore, a feasible solution of (24)–(26) for which the constraints (B1)–(B2) do not hold as equalities has an objective function value greater than or equal to $\phi_0^{\text{ADP0}}(x_0, s_0)$. Hence, ϕ^{ADP0} is an optimal solution of (24)–(26). □

Proof of Proposition 5.2. The constraints (38) provide lower bounds on the values of the $d_i(x_i, s_i, F_{i,i+1})$ variables. Using such lower bound in (39) yields the following inequalities:

$$\phi_i(x_i, s_i) \geq \sum_{F_{i,i+1}} p(F_{i,i+1} \,|\, s_i, F_0) \cdot \max_{a_i} \big\{ r(a_i, s_i) + \delta \mathbb{E}\big[\phi_{i+1}(x_i - a_i, s_{i+1}) \,\big|\, F_{i,i+1}\big]\big\}, \quad \forall (i, x_i, s_i).$$

The solution (ϕ^p, d^p) is feasible to (36)–(39) and makes these inequalities hold as equalities. The rest of the proof is analogous to the proof of Proposition 5.1. □

Proof of Proposition 6.1. (i) Define $(\cdot)^+ := \max(0, \cdot)$. It holds that $\phi_{N-1}^{\text{ADP0}}(x_{N-1}, s_{N-1}) = (\alpha^W s_{N-1} - c^W)^+ x_{N-1}$ for all $x_{N-1} \in \mathcal{X}$ and $s_{N-1} \in \mathbb{R}_+$, since $\phi_N^{\text{ADP0}}(x_N, s_N) \equiv 0$ for all $x_N \in \mathcal{X}$. At stage $N-2$ for $x_{N-2} \in \mathcal{X}\backslash\{0\}$ we have

$$\begin{aligned}
&\phi_{N-2}^{\text{ADP0}}(x_{N-2}, s_{N-2}) \\
&= \sup_{F_{N-2,N-1} \in \mathbb{R}_+} \Big\{ \max_a \big\{ r(a, s_{N-2}) + \delta \mathbb{E}\big[\phi_{N-1}^{\text{ADP0}}(x_{N-2} - a, s_{N-1}) \,\big|\, F_{N-2,N-1}\big]\big\}\Big\} \\
&= \sup_{F_{N-2,N-1} \in \mathbb{R}_+} \left\{ \max_a \left\{ r(a, s_{N-2}) + \alpha^W \delta (x_{N-2} - a) \mathbb{E}\left[\left(s_{N-1} - \frac{c^W}{\alpha^W}\right)^+ \,\Big|\, F_{N-2,N-1}\right]\right\}\right\} \\
&\geq \alpha^W \delta x_{N-2} \sup_{F_{N-2,N-1} \in \mathbb{R}_+} \mathbb{E}\left[\left(s_{N-1} - \frac{c^W}{\alpha^W}\right)^+ \,\Big|\, F_{N-2,N-1}\right],
\end{aligned}$$

where we obtain the inequality by noting that the do-nothing decision is feasible and from $r(0, s_{N-2}) = 0$. The term $\mathbb{E}[(s_{N-1} - c^W/\alpha^W)^+ \,|\, F_{N-2,N-1}]$ is an increasing function of $F_{N-2,N-1}$ under the assumption that the conditional distribution of s_{N-1} given $F_{N-2,N-1}$ is stochastically increasing in $F_{N-2,N-1}$ (Topkis 1998, Corollary 3.9.1(a)). It follows that $\phi_{N-2}^{\text{ADP0}}(x_{N-2}, s_{N-2}) = \infty$, for all $x_{N-2} \in \mathcal{X}\backslash\{0\}$ and $s_{N-2} \in \mathbb{R}_+$. To show that $\phi_{N-2}^{\text{ADP0}}(0, s_{N-2}) = \infty$ we use a similar argument with the feasible action equal to C^I instead of 0. It follows that $\phi_i^{\text{ADP0}}(x_i, s_i)$ is also equal to infinity for all $i \in \mathcal{I}$, $x_i \in \mathcal{X}$, and $s_i \in \mathbb{R}_+$.

(ii) The proof for this part is similar to the proof of part (i).

(iii) We have $0 \leq V_i(x_i, F_i) \leq \bar{x}\mathbb{E}[(\sum_{j=i}^{N-1} \delta^{j-i} s_j) \,|\, F_i] = \bar{x}(\sum_{j=i}^{N-1} \delta^{j-i} F_{ij})$, where the first inequality holds because doing nothing at every stage and state is a feasible policy with zero value; the second inequality is true because the stage i value of selling $\bar{x}$ at every stage without incurring the withdrawal loss and marginal cost provides a trivial upper bound on the stage i optimal value function; and the equality follows from the property that $\mathbb{E}[s_j \,|\, F_i]$ is equal to $F_{i,j}$ (Shreve 2004, p. 244). It can be shown in an analogous manner that the value functions of the ADPs in set $\mathcal{L}$ are bounded. We omit these derivations for brevity. □

We use Lemma B.1 to derive inequality (48).

Lemma B.1. *Let the functions f and g be defined on a finite set $\mathcal{Z}$. It holds that* $|\max_{z\in\mathcal{Z}} f(z) - \max_{z\in\mathcal{Z}} g(z)| \leq \max_{z\in\mathcal{Z}} |f(z) - g(z)|$.

Proof. Let $z_1 \in \arg\max_{z\in\mathscr{Z}} f(z)$ and $z_2 \in \arg\max_{z\in\mathscr{Z}} g(z)$. It holds that

$$f(z_1)-g(z_2) \le f(z_1)-g(z_1) \le \max_{z\in\mathscr{Z}}\{f(z)-g(z)\} \le \max_{z\in\mathscr{Z}}|f(z)-g(z)|.$$

Following the same steps starting from $g(z_2)-f(z_1)$ yields $g(z_2)-f(z_1)\le \max_{z\in\mathscr{Z}}|f(z)-g(z)|$. □

Derivation of Inequality (48). Consider ADP1. To prove the claimed bound, it suffices to prove that

$$\left|V_i(x_i,F_i)-\phi_i^{\text{ADP1}}(x_i,s_i)\right| \le \gamma_i^{\text{ADP1}}(x_i,F_i) \tag{B3}$$

holds for all (i,x_i,F_i), because $\gamma_i^{\text{ADP1}}(x_i,F_i)\ge 0$. We establish (B3) by induction on the number of stages. Inequality (B3) holds as an equality at stage $N-1$ with both sides equal to zero. Make the induction hypothesis that this inequality also holds for stages $i+1,\ldots,N-2$. At stage i we have

$$\begin{aligned}&\left|V_i(x_i,F_i)-\phi_i^{\text{ADP1}}(x_i,s_i)\right|\\ &\quad\le \left|V_i(x_i,F_i)-\phi_i^{\text{ADP1},V}(x_i,s_i)\right|\\ &\qquad+\left|\phi_i^{\text{ADP1},V}(x_i,s_i)-\phi_i^{\text{ADP1}}(x_i,s_i)\right|.\end{aligned} \tag{B4}$$

We bound $|\phi_i^{\text{ADP1},V}(x_i,s_i)-\phi_i^{\text{ADP1}}(x_i,s_i)|$ as

$$\begin{aligned}&\left|\phi_i^{\text{ADP1},V}(x_i,s_i)-\phi_i^{\text{ADP1}}(x_i,s_i)\right|\\ &=\Big|\max_{a_i}\{r(a_i,s_i)+\delta\mathbb{E}[V_{i+1}(x_i-a_i,F_{i+1})\,|\,\bar{F}_i'(s_i,F_0)]\}\\ &\qquad-\max_{a_i}\{r(a_i,s_i)+\delta\mathbb{E}[\phi_{i+1}^{\text{ADP1}}(x_i-a_i,s_{i+1})\,|\,\bar{F}_i'(s_i,F_0)]\}\Big|\\ &\le\delta\max_{a_i}\left|\mathbb{E}[V_{i+1}(x_i-a_i,F_{i+1})-\phi_{i+1}^{\text{ADP1}}(x_i-a_i,s_{i+1})\,|\,\bar{F}_i'(s_i,F_0)]\right|\\ &\le\delta\max_{x_{i+1}}\mathbb{E}[|V_{i+1}(x_{i+1},F_{i+1})-\phi_{i+1}^{\text{ADP1}}(x_{i+1},s_{i+1})|\,|\,\bar{F}_i'(s_i,F_0)]\\ &\le\delta\max_{x_{i+1}}\mathbb{E}[\gamma_{i+1}^{\text{ADP1}}(x_{i+1},F_{i+1})\,|\,\bar{F}_i'(s_i,F_0)],\end{aligned} \tag{B5}$$

where the first inequality follows from Lemma B.1, the second from the modulus inequality (Resnick 1999, p. 128) and $x_i-a_i\in\mathscr{X}$, and the third from the induction hypothesis. Using (B4), (B5), and the definition of $\gamma_i^{\text{ADP1}}(x_i,F_i)$ yields $|V_i(x_i,F_i)-\phi_i^{\text{ADP1}}(x_i,s_i)|\le\gamma_i^{\text{ADP1}}(x_i,F_i)$. The validity of inequality (B3) in all other stages follows from the principle of mathematical induction. The proofs of the bounds for the remaining ADPs are similar and are omitted for brevity.

We use Lemmas B.2–B.4 in the proof of Proposition 6.2. We define $\rho_i := \{\rho_{i,j}, j>i\}$ and use the vector expressions $\rho_i = 1$ and $\rho_i \to 1$, where 1 is a compatible vector of ones, instead of $\rho_{i,j}=1, \forall j>i$ and $\rho_{i,j}\to 1, \forall j>i$, respectively. Extending our notation $\bar{F}_i'(s_i,F_0)$, we denote by $\bar{F}_i'(s_i,F_k)$, with $i>k$, the expectation $\mathbb{E}[F_i'\,|\,s_i,F_k]$, where the stage i futures price vector F_i' is random and the conditioning information is the stage i spot price s_i and the stage k forward curve F_k. Similarly, we denote the expectation $\mathbb{E}[F_i'\,|\,s_i,\bar{F}_{i-1}'(s_{i-1},F_0)]$ by $\bar{F}_i'(s_i,s_{i-1},F_0)$. Based on these notations, we use $\bar{F}_i'^{(\rho_i=1)}(s_i,F_k)$ and $\bar{F}_i'^{(\rho_i=1)}(s_i,s_{i-1},F_0)$ to denote the limits $\lim_{\rho_i\to 1}\mathbb{E}[F_i'\,|\,s_i,F_k]$ and $\lim_{\rho_i\to1}\mathbb{E}[F_i'\,|\,s_i,\bar{F}_{i-1}'(s_{i-1},F_0)]$, respectively.

Lemma B.2. *Under price model (3)–(4): (i) The function $V_i(x_i,\cdot)$ is continuous. Given i and x_i, the random variable $V_i(x_i,F_i)$ is uniformly integrable. (ii) Given s_i and F_k, with $i>k$, the random vector F_i' converges in distribution to the deterministic vector $\bar{F}_i'^{(\rho_i=1)}(s_i,F_k)$ when $\rho_i\to1$. (iii) Given s_i and $\bar{F}_{i-1}'(s_{i-1},F_0)$, the random vector F_i' converges in distribution to the deterministic vector $\bar{F}_i'^{(\rho_i=1)}(s_i,s_{i-1},F_0)$ when $\rho_i\to1$. (iv) Given s_i and $\bar{F}_{i-1}'(s_{i-1},F_0)$, the vector $\bar{F}_i'^{(\rho_i=1)}(s_i,s_{i-1},F_0)$ tends to the vector $\bar{F}_i'^{(\rho_i=1)}(s_i,F_0)$ as $\rho_{i-1}\to1$. (v) Given s_i and $\bar{F}_{i-1}'(s_{i-1},F_0)$, the vector $\bar{F}_i'(s_i,F_0)$ tends to the vector $\bar{F}_i'^{(\rho_i=1)}(s_i,F_0)$ as $\rho_i\to1$.*

Proof. (i) The continuity of $V_i(x_i,\cdot)$ follows from Proposition 5 in Secomandi et al. (2015). The uniform integrability of $V_i(x_i,F_i)$ holds by part (iii) of Proposition 6.1 and the fact that futures prices are uniformly integrable under price model (3)–(4) (see the dominated families criterion in Resnick 1999, p. 183).

(ii)–(iii) In each case, pick an element of the relevant random vector. This element is a lognormal random variable with mean and variance that are functions of the volatilities and instantaneous correlations of price model (3)–(4). In the limit, it can be easily verified that this variance tends to zero and this mean tends to its respective claimed constant. Thus, we have the stated convergence in distribution (see Resnick 1999, p. 249).

Parts (iv) and (v) follow because the elements of the stated vectors are continuous functions of the stated instantaneous correlation coefficients. □

Lemma B.3. *Let $f(\cdot)$ be a real valued and continuous function and X_n a sequence of uniformly integrable random vectors. Suppose that the random variable $f(X_n)$ is uniformly integrable and $\lim_{n\to\infty}X_n$ converges in distribution to the deterministic vector $\bar{X}$. Then, $\lim_{n\to\infty}\mathbb{E}[|f(X_n)-f(\bar{X})|]=0$.*

Proof. Because convergence in distribution to a constant implies convergence in probability (see Resnick 1999, Proposition 8.5.2), we have that $\lim_{n\to\infty}X_n$ converges in probability to the constant vector $\bar{X}$. Using this result and the continuity of f, it follows from part (ii) of Corollary 6.3.1 in Resnick (1999) that $f(X_n)$ converges in probability to the constant $f(\bar{X})$. This result and the uniform integrability of $f(X_n)$ allows us to use Theorem 6.6.1 in Resnick (1999) to obtain the claimed result. □

Lemma B.4 holds by the properties of (multivariate) lognormal random variables. We assume that all the time intervals $[T_i, T_{i+1}]$ have the same length ΔT.

Lemma B.4. *Under price model (3)–(4), it holds for $j>i>k$ that*

$$\bar{F}_{i,j}(s_i,F_k)=F_{k,j}\left(\frac{s_i}{F_{k,i}}\right)^{\rho_{i,j}\sigma_j/\sigma_i}\cdot\exp\left(\frac{\rho_{i,j}\sigma_j(i-k)\Delta T}{2}(\sigma_i-\rho_{i,j}\sigma_j)\right), \tag{B6}$$

$$\mathbb{E}[s_i^{\sigma_j/\sigma_i}\,|\,F_{i-1,i}]=F_{i-1,i}^{\sigma_j/\sigma_i}\exp\left(\frac{\sigma_i\Delta T}{2}(\sigma_j-\sigma_i)\right), \tag{B7}$$

$$\begin{aligned}&\mathbb{E}[F_{i-1,i}^{\sigma_j/\sigma_i}\,|\,s_{i-1},F_0]\\ &=F_{0,i}^{\sigma_j/\sigma_i}\left(\frac{s_{i-1}}{F_{0,i-1}}\right)^{\sigma_j\rho_{i-1,i}/\sigma_{i-1}}\\ &\quad\cdot\exp\left(\frac{\sigma_j(i-1)\Delta T}{2}(\rho_{i-1,i}\sigma_{i-1}-\sigma_i+\sigma_j(1-\rho_{i-1,i}^2))\right).\end{aligned} \tag{B8}$$

Proof of Proposition 6.2. (i) Consider ADP1. Note that the equivalence $\phi_i^{\mathrm{ADP1},V}(x_i,s_i)\equiv V_i(x_i,s_i,\bar F_i'(s_i,F_0))$ follows from the definitions of $\phi_i^{\mathrm{ADP1},V}$ in §6 and V_i in §2.1. At a stage i, using this equivalence and the definition of γ_i^{ADP1} we have

$$\|\gamma_i^{\mathrm{ADP1}}\|_{\mathbb{E},\infty}\equiv\max_{x_i}\mathbb{E}\big[|V_i(x_i,F_i)-V_i(x_i,s_i,\bar F_i'(s_i,F_0))|\,|F_0\big] +\delta\mathbb{E}\Big[\max_{x_{i+1}}\mathbb{E}\big[\gamma_i^{\mathrm{ADP1}}(x_{i+1},F_{i+1})\,|\,\bar F_i'(s_i,F_0)\big]\,|F_0\Big]. \quad \text{(B9)}$$

We now show that the right-hand side of (B9) tends to zero as $\rho\to1$. For the first term on the right-hand side of (B9), we have

$$\begin{aligned}
&\lim_{\rho\to1}\max_{x_i}\mathbb{E}\big[|V_i(x_i,F_i)-V_i(x_i,s_i,\bar F_i'(s_i,F_0))|\,|F_0\big]\\
&\le\max_{x_i}\lim_{\rho\to1}\mathbb{E}\big[\mathbb{E}\big[|V_i(x_i,F_i)-V_i(x_i,s_i,\bar F_i'^{(\rho_i=1)}(s_i,F_0))|\,|s_i,F_0\big]\,|F_0\big]\\
&\quad+\max_{x_i}\lim_{\rho\to1}\mathbb{E}\big[\mathbb{E}\big[|V_i(x_i,s_i,\bar F_i'^{(\rho_i=1)}(s_i,F_0))\\
&\qquad\qquad-V_i(x_i,s_i,\bar F_i'(s_i,F_0))|\,|s_i,F_0\big]\,|F_0\big]\\
&=\max_{x_i}\lim_{\rho\to1}\mathbb{E}\big[\mathbb{E}\big[|V_i(x_i,F_i)-V_i(x_i,s_i,\bar F_i'^{(\rho_i=1)}(s_i,F_0))|\,|s_i,F_0\big]\,|F_0\big]\\
&=0,
\end{aligned}$$

where the inequality follows from swapping the limit and maximization, which is allowed because the maximization is over the finite set $\mathscr{X}$, iterated expectations based on conditioning on s_i, and applying the triangle inequality; the first equality holds by the continuity of $V_i(x_i,\cdot)$ established in part (i) of Lemma B.2 and the fact that, given s_i and F_0, $\bar F_i'(s_i,F_0)$ is a continuous function of ρ_i that tends to $\bar F_i'^{(\rho_i=1)}(s_i,F_0)$ as $\rho_i\to1$ (see (B6) in Lemma B.4); and the final equality follows by applying parts (i) and (ii) of Lemma B.2 and Lemma B.3. Thus, the limit of the first term on the right-hand side of (B9) must be zero.

The second term on the right-hand side of (B9) can be rewritten as an expression that depends only on the exact value function V_i by using the recursive definition of γ_i^{ADP1}. This expression is a sum of terms, one for each $j\in\{i+1,\dots,N-2\}$, where the jth term is a sequence of iterated expectations and the innermost expectation is

$$\mathbb{E}\big[|V_j(x_j,s_j,F_j')-V_j(x_j,s_j,\bar F_j'(s_j,F_0))|\,|\,\bar F_{j-1}'(s_{j-1},F_0)\big]. \quad \text{(B10)}$$

We show that the limit of (B10) as $\rho_{j-1}\to1$ and $\rho_j\to1$ is zero for any fixed x_j as follows (we write the limit explicitly once and then suppress its argument in the remaining expressions):

$$\begin{aligned}
&\lim_{\substack{\rho_{j-1}\to1\\ \rho_j\to1}}\mathbb{E}\big[|V_j(x_j,s_j,F_j')-V_j(x_j,s_j,\bar F_j'(s_j,F_0))|\,|\,\bar F_{j-1}'(s_{j-1},F_0)\big]\\
&=\lim\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,F_j')-V_j(x_j,s_j,\bar F_j'(s_j,F_0))|\,|\,s_j,\\
&\qquad\qquad\bar F_{j-1}'(s_{j-1},F_0)\big]\,|\,\bar F_{j-1}'(s_{j-1},F_0)\big]\\
&\le\lim\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,F_j')-V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,s_{j-1},F_0))|\,|\,s_j,\\
&\qquad\qquad\bar F_{j-1}'(s_{j-1},F_0)\big]\,|\,\bar F_{j-1}'(s_{j-1},F_0)\big]\\
&\quad+\lim\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,s_{j-1},F_0))\\
&\qquad\qquad-V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_0))|\,|\,s_j,\\
&\qquad\qquad\bar F_{j-1}'(s_{j-1},F_0)\big]\,|\,\bar F_{j-1}'(s_{j-1},F_0)\big]\\
&\quad+\lim\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_0))-V_j(x_j,s_j,\bar F_j'(s_j,F_0))|\,|s_j,\\
&\qquad\qquad\bar F_{j-1}'(s_{j-1},F_0)\big]\,|\,\bar F_{j-1}'(s_{j-1},F_0)\big]\\
&=0,
\end{aligned}$$

where the first equality follows from iterated expectations based on conditioning on s_j; the inequality from applying the triangle inequality twice; and the second equality from Lemma B.3, part (i) of Lemma B.2, and parts (iii)–(v) of Lemma B.2. Thus, the limit of the second term on the right-hand side of (B9) must be zero, which completes the proof for ADP1. The proof for ADP2 is similar and is thus omitted for brevity.

Consider SADP. At stage i, by definition of γ_i^{SADP} we have

$$\begin{aligned}
&\|\gamma_i^{\mathrm{SADP}}\|_{\mathbb{E},\infty}\\
&\quad=\max_{x_i}\mathbb{E}\big[|V_i(x_i,F_i)-V_i^{\mathrm{SADP}}(x_i,s_i)|\,|\,F_0\big]\\
&\qquad+\delta\,\mathbb{E}\Big[\mathbb{E}\Big[\max_{x_{i+1}}\mathbb{E}\big[\gamma_{i+1}^{\mathrm{SADP}}(x_{i+1},F_{i+1})\,|\,F_i'\big]\,|\,s_i,F_0\Big]\,|\,F_0\Big]. \quad \text{(B11)}
\end{aligned}$$

We show that each of the terms on the right-hand side of (B11) tends to zero as $\rho\to1$. The proof of this claim for the first term is similar to the proof given above for ADP1, and is thus omitted. The second term can be rewritten as an expression involving only the exact value function V_i using the recursive definition of γ_i^{SADP}. This expression is a sum of terms, one for each $j\in\{i+1,\dots,N-2\}$, where the jth term is a sequence of iterated expectations and the innermost expectation is

$$\mathbb{E}\Big[\max_{x_j}\mathbb{E}\big[|V_j(x_j,s_j,F_j')-\phi_j^{\mathrm{SADP},V}(x_j,s_j)|\,|F_{j-1}'\big]\Big|s_{j-1},F_0\Big]. \quad \text{(B12)}$$

We now show that the limit of (B12) as $\rho_{j-1}\to1$ and $\rho_j\to1$ is zero for any fixed x_j. We have

$$\begin{aligned}
&\lim_{\substack{\rho_{j-1}\to1\\ \rho_j\to1}}\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,F_j')-\phi_j^{\mathrm{SADP},V}(x_j,s_j)|\,|F_{j-1}'\big]\,|s_{j-1},F_0\big]\\
&=\lim_{\substack{\rho_{j-1}\to1\\ \rho_j\to1}}\mathbb{E}\big[\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,F_j')\\
&\qquad\qquad-\phi_j^{\mathrm{SADP},V}(x_j,s_j)|\,|s_j,F_{j-1}'\big]\,|F_{j-1}'\big]\,|s_{j-1},F_0\big]\\
&\le\lim_{\substack{\rho_{j-1}\to1\\ \rho_j\to1}}\mathbb{E}\big[\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,F_j')\\
&\qquad\qquad-V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_{j-1}'))|\,|s_j,F_{j-1}'\big]\,|F_{j-1}'\big]\,|s_{j-1},F_0\big]\\
&\quad+\lim_{\substack{\rho_{j-1}\to1\\ \rho_j\to1}}\mathbb{E}\big[\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_{j-1}'))\\
&\qquad\qquad-V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_0))|\,|s_j,F_{j-1}'\big]\,|F_{j-1}'\big]\,|s_{j-1},F_0\big]\\
&\quad+\lim_{\substack{\rho_{j-1}\to1\\ \rho_j\to1}}\mathbb{E}\big[\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_0))\\
&\qquad\qquad-\phi_j^{\mathrm{SADP},V}(x_j,s_j)|\,|s_j,F_{j-1}'\big]\,|F_{j-1}'\big]\,|s_{j-1},F_0\big] \quad \text{(B13)}\\
&=\lim_{\rho_{j-1}\to1}\mathbb{E}\big[\mathbb{E}\big[|V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_{j-1}'))\\
&\qquad\qquad-V_j(x_j,s_j,\bar F_j'^{(\rho_j=1)}(s_j,F_0))|\,|F_{j-1}'\big]\,|s_{j-1},F_0\big]\\
&\le C'\sum_{k=j+1}^{N-2}\lim_{\rho_{j-1}\to1}\mathbb{E}\big[\mathbb{E}\big[|\bar F_{j,k}'^{(\rho_{j,k}=1)}(s_j,F_{j-1}')\\
&\qquad\qquad-\bar F_{j,k}^{(\rho_{j,k}=1)}(s_j,F_0)|\,|F_{j-1,j},F_{j-1,k}\big]\,|s_{j-1},F_{0,j},F_{0,k}\big]\\
&:=\mathrm{RHS}, \quad \text{(B14)}
\end{aligned}$$

where the first equality follows from iterated expectations and conditioning on s_j; the first inequality from applying the triangle inequality twice; the second equality from

(i) the first term in (B13) being zero by parts (i) and (ii) of Lemma B.2 and Lemma B.3, (ii) unconditioning on s_j in the second term of (B13), and (iii) the third term in (B13) being zero by the equivalence $\phi_i^{\text{SADP},V}(x_i,s_i) \equiv \mathbb{E}[V_i(x_i,s_i,F_i')\,|\,s_i,F_0]$, parts (i) and (ii) of Lemma B.2, and Lemma B.3; the last inequality follows from the Lipschitz continuity of $V_j(x_i,\cdot)$ with Lipschitz constant $C' := \alpha^I \max\{|C^I|, C^W\}$, which is a straightforward modification of Proposition 5 in Secomandi et al. (2015).

We now proceed to show that the term labeled RHS in (B14) is zero. We use the notation $\mathbb{E}^{(\rho_{j-1,j}=1)}[\cdot]$ to indicate the limit of $\mathbb{E}[\cdot]$ as $\rho_{j-1,j}$ tends to 1. We have the following:

$$
\begin{aligned}
&C'\sum_{k=j+1}^{N-2}\lim_{\rho_{j-1}\to 1}\mathbb{E}\Bigg[\Bigg|\frac{F_{j-1,k}}{F_{j-1,j}^{\sigma_k/\sigma_j}}\exp\bigg(\frac{\sigma_k\Delta T}{2}(\sigma_j-\sigma_k)\bigg)-\frac{F_{0,k}}{F_{0,j}^{\sigma_k/\sigma_j}}\\
&\qquad\cdot\exp\bigg(\frac{\sigma_k j\Delta T}{2}(\sigma_j-\sigma_k)\bigg)\Bigg|\mathbb{E}\big[s_j^{\sigma_k/\sigma_j}\,|F_{j-1,j}\big]\Big|s_{j-1},F_{0,j},F_{0,k}\Bigg]\\
&=C'\sum_{k=j+1}^{N-2}\lim_{\rho_{j-1}\to 1}\mathbb{E}\Bigg[\Bigg|F_{j-1,k}-\frac{F_{0,k}}{F_{0,j}^{\sigma_k/\sigma_j}}F_{j-1,j}^{\sigma_k/\sigma_j}\\
&\qquad\qquad\cdot\exp\bigg(\frac{\sigma_k(j-1)\Delta T}{2}(\sigma_j-\sigma_k)\bigg)d\Bigg|\,\Big|s_{j-1},F_{0,j},F_{0,k}\Bigg]\\
&=C'\sum_{k=j+1}^{N-2}\Bigg(\lim_{\rho_{j-1}\to 1}\mathbb{E}\Bigg[\Bigg(\bigg(F_{j-1,k}-\frac{F_{0,k}}{F_{0,j}^{\sigma_k/\sigma_j}}F_{j-1,j}^{\sigma_k/\sigma_j}\\
&\qquad\qquad\cdot\exp\bigg(\frac{\sigma_k(j-1)\Delta T}{2}(\sigma_j-\sigma_k)\bigg)\bigg)\Bigg)^+\Bigg|s_{j-1},F_{0,j},F_{0,k}\Bigg]\\
&\quad+\lim_{\rho_{j-1}\to 1}\mathbb{E}\Bigg[\Bigg(\frac{F_{0,k}}{F_{0,j}^{\sigma_k/\sigma_j}}F_{j-1,j}^{\sigma_k/\sigma_j}\exp\bigg(\frac{\sigma_k(j-1)\Delta T}{2}(\sigma_j-\sigma_k)\bigg)\\
&\qquad\qquad\qquad-F_{j-1,k}\Bigg)^+\Bigg|s_{j-1},F_{0,j},F_{0,k}\Bigg]\Bigg)\\
&=C'\sum_{k=j+1}^{N-2}\Bigg(\Bigg(\bar{F}_{j-1,k}^{(\rho_{j-1,k}=1)}(s_{j-1},F_0)\\
&\qquad\qquad-\frac{F_{0,k}}{F_{0,j}^{\sigma_k/\sigma_j}}\exp\bigg(\frac{\sigma_k(j-1)\Delta T}{2}(\sigma_j-\sigma_k)\bigg)\\
&\qquad\qquad\qquad\cdot\mathbb{E}^{(\rho_{j-1,j}=1)}\big[F_{j-1,j}^{\sigma_k/\sigma_j}\,|s_{j-1},F_0\big]\Bigg)^+\\
&\quad+\Bigg(\frac{F_{0,k}}{F_{0,j}^{\sigma_k/\sigma_j}}\exp\bigg(\frac{\sigma_k(j-1)\Delta T}{2}(\sigma_j-\sigma_k)\bigg)\\
&\qquad\cdot\mathbb{E}^{(\rho_{j-1,j}=1)}\big[F_{j-1,j}^{\sigma_k/\sigma_j}\,|s_{j-1},F_0\big]-\bar{F}_{j-1,k}^{(\rho_{j-1,k}=1)}(s_{j-1},F_0)\Bigg)^+\Bigg)\\
&=C'\sum_{k=j+1}^{N-2}\Bigg|\bar{F}_{j-1,k}^{(\rho_{j-1,k}=1)}(s_{j-1},F_0)-\frac{F_{0,k}}{F_{0,j}^{\sigma_k/\sigma_j}}\\
&\qquad\cdot\exp\bigg(\frac{\sigma_k(j-1)\Delta T}{2}(\sigma_j-\sigma_k)\bigg)\mathbb{E}^{(\rho_{j-1,j}=1)}\big[F_{j-1,j}^{\sigma_k/\sigma_j}\,|s_{j-1},F_0\big]\Bigg|\\
&=0,
\end{aligned}
$$

where the first equality follows from using (B6) and factoring the term within the absolute value out of the inner expectation given that this term is deterministic given F_{j-1}; the second follows from using (B7) and simplifying; the third by splitting the absolute value into a sum of two positive parts; the fourth by evaluating the resulting expectations using the exchange option formula of Margrabe (1978) and evaluating their respective limit as $\rho_{j-1}\to 1$; the fifth by expressing the sum of the two positive parts as an absolute value; and the sixth by using (B8) and simplifying. Hence, expression (B12) tends to zero when $\rho\to 1$.

(ii) Since the limiting matrix is rank two and $|\bar{\rho}_{i,i+1}|<1$, we have convergence of $F_i''|s_i,F_{i,i+1}$ in distribution to $\bar{F}_i''(s_i,F_{i,i+1})$ as $\rho\to\bar{\rho}$ (recall that $F_i''\equiv\{F_{i,i+1},\ldots,F_{N-1}\}$). The rest of the proof of this part is analogous to the proof of part (i), and is thus omitted for brevity. □

References

Adelman D (2004) A price-directed approach to stochastic inventory/routing. *Oper. Res.* 52(4):499–514.

Adelman D (2007) Dynamic bid prices in revenue management. *Oper. Res.* 55(4):647–661.

Adelman D, Klabjan D (2012) Computing near optimal policies in generalized joint replenishment. *INFORMS J. Comput.* 24(1):148–164.

Adelman D, Mersereau A (2013) Dynamic capacity allocation to customers who remember past service. *Management Sci.* 59(3):592–612.

Bertsekas DP (2007) *Dynamic Programming and Optimal Control*, 3rd ed., Vol. 2 (Athena Scientific, Belmont, MA).

Birge J (2011) Quasi-convex stochastic dynamic programming. Working paper, University of Chicago, Chicago.

Björk T (2004) *Arbitrage Theory in Continuous Time* (Oxford University Press, Oxford, Oxfordshire, UK).

Black F (1976) The pricing of commodity contracts. *J. Financial Econom.* 3(1–2):167–179.

Boogert A, de Jong C (2008) Gas storage valuation using a Monte Carlo method. *J. Derivatives* 15(3):81–98.

Boogert A, de Jong C (2011/2012) Gas storage valuation using a multifactor price process. *J. Energy Markets* 4(4):29–52.

Brown DB, Smith JE, Sun P (2010) Information relaxations and duality in stochastic dynamic programs. *Oper. Res.* 58(4):785–801.

Carmona R, Ludkovski M (2010) Valuation of energy storage: An optimal switching approach. *Quant. Finance* 10(4):359–374.

Chen Z, Forsyth PA (2007) A semi-Lagrangian approach for natural gas storage valuation and optimal operation. *SIAM J. Scientific Comput.* 30(1):339–368.

Cortazar G, Schwartz ES (1994) The valuation of commodity contingent claims. *J. Derivatives* 1(4):27–39.

Cortazar G, Gravet M, Urzua J (2008) The valuation of multidimensional American real options using the LSM simulation method. *Computers Oper. Res.* 35(1):113–129.

de Farias DP, Van Roy B (2001) On constraint sampling for the linear programming approach to approximate dynamic programming. *Math. Oper. Res.* 29(3):462–478.

de Farias DP, Van Roy B (2003) The linear programming approach to approximate dynamic programming. *Oper. Res.* 51(6):850–865.

Desai VV, Farias VF, Moallemi CC (2012) Approximate dynamic programming via a smoothed approximate linear program. *Oper. Res.* 60(3):655–674.

Devalkar SK, Anupindi R, Sinha A (2011) Integrated optimization of procurement, processing, and trade of commodities. *Oper. Res.* 59(6):1369–1381.

Dixit AK, Pindyck RS (1994) *Investment Under Uncertainty* (Princeton University Press, Princeton, NJ).

Eydeland A, Wolyniec K (2003) *Energy and Power Risk Management* (John Wiley & Sons, Hoboken, NJ).

Farias VF, Van Roy B (2007) An approximate dynamic programming approach to network revenue management. Working paper, Stanford University, Stanford, CA.

Felix B, Weber C (2012) Gas storage valuation applying numerically constructed recombining trees. *Eur. J. Oper. Res.* 216(1):178–187.

Geman H (2005) *Commodities and Commodity Derivatives: Modelling and Pricing for Agriculturals, Metals, and Energy* (Wiley, Chichester, UK).

Gray J, Khandelwal P (2004) Towards a realistic gas storage model. *Commodities Now* (June):1–4.

Jaillet P, Ronn EI, Tompaidis S (2004) Valuation of commodity-based swing options. *Management Sci.* 50(7):909–921.

Lai G, Margot F, Secomandi N (2010) An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation. *Oper. Res.* 58(3):564–582.

Levy G (2004) *Computational Finance: Numerical Methods for Pricing Financial Instruments* (Butterworth-Heinemann, Oxford, UK).

Longstaff FA, Schwartz ES (2001) Valuing American options by simulation: A simple least-squares approach. *Rev. Financial Stud.* 14(1):113–147.

Manne AS (1960) Linear programming and sequential decisions. *Management Sci.* 60(3):259–267.

Maragos S (2002) Valuation of the operational flexibility of natural gas storage reservoirs. Ronn E, ed. *Real Options and Energy Management Using Options Methodology to Enhance Capital Budgeting Decisions* (Risk Publications, London, UK), 431–456.

Margrabe W (1978) The value of an option to exchange one asset for another. *J. Finance* 33(1):177–186.

Moallemi CC, Kumar S, Van Roy B (2008) Approximate and data-driven dynamic programming for queuing networks. Working paper, Stanford University, Stanford, CA.

Morrison JR, Kumar PR (1999) New linear program performance bounds for queuing networks. *J. Optim. Theory Appl.* 100(3): 575–597.

Munos R (2007) Performance bounds in L_p-norm for approximate value iteration. *SIAM J. Control Optim.* 46(2):541–561.

Nadarajah S (2014) Approximate dynamic programming for commodity and energy merchant operations. Ph.D. thesis, Carnegie Mellon University, Pittsburgh.

Petrik M (2012) Distributionally robust approach to approximate dynamic programming. Langford J, Pineau J, eds. *Proc. 29th Internat. Conf. Machine Learning, Edinburgh, Scotland*, 497–504.

Petrik M, Zilberstein S (2009) Constraint relaxation in approximate linear programs. Bottou L, Littman M, eds. *Proc. 26th Internat. Conf. Machine Learning, Montréal, Québec*, 809–816.

Powell WB (2011) *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, 2nd ed. (John Wiley & Sons, Hoboken, NJ).

Puterman M (1994) *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, New York).

Resnick S (1999) *A Probability Path* (Birkhäuser, Boston).

Restrepo M (2008) Computational methods for static allocation and real-time redeployment of ambulances. Ph.D. thesis, Cornell University, Ithaca, NY.

Rubinstein M (1994) Return to OZ. *Risk Magazine*.

Schweitzer PJ, Seidmann A (1985) Generalized polynomial approximations in Markovian decision processes. *J. Math. Anal. Appl.* 110(2):568–582.

Secomandi N (2010) Optimal commodity trading with a capacitated storage asset. *Management Sci.* 56(3):1090–1049.

Secomandi N, Lai G, Margot F, Scheller-Wolf A, Seppi DJ (2015) Merchant commodity storage and term-structure model error. *Manufacturing Service Oper. Management* 17(3):302–320.

Shreve S (2004) *Stochastic Calculus for Finance II: Continuous-Time Models* (Springer, New York).

Smith J, McCardle K (1999) Options in the real world: Lessons learned in evaluating oil and gas investments. *Oper. Res.* 47(1): 1–15.

Thompson M, Davison M, Rasmussen H (2009) Natural gas storage valuation and optimization: A real options application. *Naval Res. Logist.* 56(3):226–238.

Topkis D (1998) *Supermodularity and Complementarity* (Princeton University Press, Princeton, NJ).

Trick MA, Zin SE (1997) Spline approximations to value functions. *Macroeconomic Dynamics* 1(1):255–277.

Tsitsiklis J, Van Roy B (2001) Regression methods for pricing complex American-style options. *IEEE Trans. Neural Networks* 12(4):694–703.

Veatch MH (2015) Approximate linear programming for networks: Average cost bounds. *Computers Oper. Res.* 63(November): 32–45.

Wu O, Wang D, Qin Z (2012) Seasonal energy storage operations with limited flexibility: The price adjusted rolling intrinsic policy. *Manufacturing Service Oper. Management* 14(3):455–471.

Zhang D, Adelman D (2009) An approximate dynamic programming approach to network revenue management with customer choice. *Transportation Sci.* 43(3):381–394.