Available online at www.sciencedirect.com

ScienceDirect

European Journal of Operational Research 00 (2016) 1–30

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

Dynamic Convexification within Nested Benders Decomposition using Lagrangian Relaxation: An Application to the Strategic Bidding Problem

Gregory Steeger, Steffen Rebennack[a]

Colorado School of Mines, Division of Economics and Business, Golden, CO 80401

[a]*srebenna@mines.edu*

Abstract

Many decomposition algorithms like Benders decomposition and stochastic dual dynamic programming are limited to convex optimization problems. In this paper, we utilize a dynamic convexification method that makes use of Lagrangian relaxation to overcome this limitation and enables the modeling of non-convex multi-stage problems using decomposition algorithms. Though the algorithm is confined by the duality gap of the problem being studied, the computed upper bounds (for maximization problems) are at least as good as those found via an LP relaxation approach. We apply the method to the strategic bidding problem for a hydroelectric producer, in which we ask: What is the revenue-maximizing production schedule for a single price-maker hydroelectric producer in a deregulated, bid-based market? Because the price-maker's future revenue function has a sawtooth shape, we model it using mixed-integer linear programming. To remedy the non-concavity issues associated with modeling the future revenue function as a mixed-integer linear program, we model the price-maker's bidding decision utilizing both Benders decomposition and Lagrangian relaxation. We demonstrate the utility of our algorithm through an illustrative example and through three case studies in which we model electricity markets in El Salvador, Honduras, and Nicaragua.

Keywords: Lagrangian relaxation, mixed-integer linear programming, Benders decomposition, hydrothermal scheduling, hydroelectric, bidding problem, strategic bidding problem, single price-maker, Stochastic Dual Dynamic Programming (SDDP)

1. Introduction

Power generation companies are no different than other firms in that they seek to maximize their profit. What is different about power generation companies is the market through which they sell their goods. Many hydroelectric companies sell their energy through a specific type of procurement auction, commonly referred to as the day-ahead electricity market. Though day-ahead electricity markets vary (Banal-Estañol and Micola 2009), we assume market-clearing prices and quantities are derived in the following manner: (i) producers submit price and/or quantity bids for the following day; (ii) the market operator adjusts the market-clearing price, while minimizing costs, until the total energy dispatched is equal to the total demand; (iii) energy producers are paid for the amount of energy they dispatch, according to the market-clearing price.

The problem in which producers submit bids to the day-ahead electricity market is termed the *bidding problem*. Few study this problem for price-maker hydroelectric producers over the medium term, because of the difficulty associated with modeling non-convex dynamic problems that possess uncertainty. More specifically, what is missing in the literature is a method that determines a coherent hydroelectric operations policy, over horizons lasting longer than one year, in a strategic bidding framework. In this vein, we ask: What is the revenue-maximizing bidding schedule for a single price-maker hydroelectric producer in a deregulated market?

Hydroelectric producers are different than all other energy producers because of their cost-free ability to store energy (water), ignoring evaporation. Because of this, hydroelectric producers must consider the impact of selling water (energy) today versus selling it in the future. The impact, also referred to as opportunity cost, associated with the use of water today creates a time-coupling between the present and the future. Consequently, one must consider the impacts of today's water use on future revenue over a medium term (one to five years). Over the medium term, hydroelectric producers experience large variations in reservoir inflows, due mainly to varying amounts of rainfall. Since reservoir inflows impact the price-maker hydroelectric producer's bidding decisions, the examination of different possible inflow scenarios is crucial.

Dynamic Programming (DP) is attractive in hydroelectric modeling because each stage in the time-coupled problem can be separated and solved using a multi-stage decomposition scheme. For a given stage, DP algorithms recursively compute the maximum possible revenue in all future stages via so-called Future Revenue Functions (FRF) and predefined values for the state variables. Alternatively, one can use a decomposition algorithm to dynamically approximate FRFs, *e.g.*, Benders decomposition (Benders 1962).

Benders decomposition, however, is limited in that it requires that the FRFs be concave, because it utilizes linear optimality cuts. In our problem, FRFs are piecewise linear and non-concave. Consequently, modeling FRFs requires Integer Programming (IP) formulations. Practitioners typically resort to either Linear Programming (LP) relaxations or concave overestimators. Both approaches concavify the objective function through making each stages' revenue function concave. These *static overestimations* typically yield larger cumulative errors than a *dynamic overestimation*.

The present paper proposes a novel dynamic overestimation that overcomes the aforementioned convexity limitation and enables the use of Benders decomposition for non-convex multi-stage problems, in general, and for the bidding problem, in particular. In the bidding problem, we are able to dynamically overestimate the revenue by using Lagrangian relaxation to dualize the constraints that link stages together in the given time horizon. Instead of overestimating the revenue function in each stage, we use a representation of the immediate revenue function and overestimate only the resulting FRF. Furthermore, for any given stage, the resulting Lagrangian objective function can be approximated using Benders optimality cuts. Paired together, we utilize Lagrangian relaxation and Benders decomposition to model and solve the bidding problem for hydroelectric producers over the medium term. Though combining Lagrangian relaxation and Benders decomposition is not new to hydrothermal scheduling problems (Cerisola et al. 2012, Thomé 2013), its application to the bidding problem is. Our tailored Benders decomposition scheme solves the bidding problem by dualizing the constraints (water-balance equations) that link stages together in the time horizon.

Ultimately, our method is to be used in conjunction with a decomposition algorithm that incorporates inflow uncertainty and is capable of modeling large systems over the medium term (one to five years), *i.e.*, Stochastic Dual Dynamic Programming (SDDP) (Pereira and Pinto 1991). However, in this paper we assume inflows are known so that we can compute an optimal solution to the problem. Only then can we demonstrate the value of our approach by comparing the bounds our approach yields with both an optimal solution and the bounds that are found from existing methods.

Our method advances decomposition algorithms by alleviating their greatest limitation through dynamic convexification. We

- present a dynamic convexification using Lagrangian relaxation for multi-stage non-convex problems (Algorithm 1).

Specifically, our work strengthens the body of literature on the bidding problem by

- applying the method introduced in Cerisola et al. (2012) to solve the bidding problem, over the medium term, for a single price-maker hydroelectric producer and by proving the method's correctness (Section 5 and Theorem 1);
- providing an in-depth analysis of the proposed algorithm and documenting the cases in which our approach is useful, *i.e.,* computes better bounds and solutions, through theoretical results and case studies (Sections 5.3-6, and Theorem 3); and
- proving that the approaches that utilize LP relaxation or tight concave overestimators for the revenue function yield identical upper bounds (Theorem 2).

The remainder of the paper is organized as follows: Section 2 surveys the relevant literature. Next, Section 3 introduces our methodology, in a general framework. Following this, Sections 4 and 5 present our specific application. Section 6 demonstrates the usefulness of our approach through an illustrative example and three case studies. Section 7 concludes the paper. Following the body of the paper, Appendix B through Appendix D present proofs of our main theorems, and Appendix E lists the data required to replicate our case studies.

We represent fixed values (or parameters) with uppercase letters, decision variables with lowercase letters, and vectors of parameters or decision variables with boldface. Additionally, we use an apostrophe $'$ to denote the transpose of a vector to avoid confusion between the transpose operator and the time index. Lastly, energy producers that have the ability to impact prices with their bids are termed *price makers* and energy producers that have no influence on prices are termed *price takers*.

2. Literature Review

This paper focuses on the bidding problem for a market in which only one producer is a price maker and the remainder are price takers. Before reviewing the literature that is most relevant to this problem (the single price-maker case), we emphasize the differences between it and the case in which multiple producers have some degree of market power. In the single price-maker case, it is optimal for the price takers to bid at their marginal cost and, as a result, the price maker's revenue is determined solely on their own bids and the optimal solution may be found via multiple different optimization techniques (Li et al. 2011).

When the market consists of multiple price-maker producers, the revenue each price maker receives depends not only on their own bids but also on the bids from all other price makers. In this paradigm, all price makers bid simultaneously and the resulting "game" is typically modeled via non-cooperative economic game theory. The goal is to find a Nash equilibrium (Nash Jr 1951); a set of bids from the price makers such that, given all other price makers' bids, no price maker can increase their revenue by changing their bid. There may be zero, one, multiple, or even an infinite number of Nash equilibria (Yang et al. 2012, Steeger and Rebennack 2015). This type of economic game is drastically different than the bidding problem for a single price maker. It is not necessarily even thought of as an optimization problem (Mallozzi 2013). The remainder of the literature review focuses on the single price maker case.

The most involved variant of the bidding problem is for price-maker hydro producers, in which the problem is multi-stage, non-convex, and stochastic (Steeger et al. 2014). Stochastic Dynamic Programming (SDP) is well suited for multi-stage, stochastic problems with nonlinearities. Unfortunately, SDP suffers from "curses-of-dimensionality" (Yakowitz 1982). In other words, as the number of reservoirs that we are modeling and/or the number of stages over which we solve the problem increase, the number of required state variable discretizations grow exponentially and the problem becomes computationally intractable. For this reason, SDP is typically limited in its applications to systems with ten or fewer hydro plants, *i.e.*, water reservoirs. And because many real-world systems contain more than ten hydro plants, often significantly more, SDP is not used to model large systems in practice.

2.1. Stochastic Dual Dynamic Programming

In hydro-scheduling problems the state space is comprised of the water reservoir levels at the beginning of each stage, assuming stage-wise independent inflows. When modeling hydro-scheduling problems, one needs to consider various combinations of reservoir levels for each reservoir, and this leads to an exponential number of combinations of reservoir levels – a curse-of-dimensionality in the state-space. Benders decomposition counters the curse-of-dimensionality by automatically and iteratively computing reservoir levels at values which are candidates for an optimal solution.

The stochastic nature of the inflows also creates curse-of-dimensionality concerns. Modeling the uncertainty via a tree may lead to an exponential number of possible realizations, if the tree is not "thinned" out. To overcome the curse-of-dimensionality inherent to stochastic dynamic programming approaches, Pereira and Pinto (1991) proposed their SDDP algorithm. SDDP is a Nested Benders Decomposition (NBD) algorithm which utilizes sampling approaches to counter the aforementioned curse-of-dimensionality (Birge 1980, Rebennack 2016).

SDDP is an established method in hydrothermal scheduling and allows for the detailed modeling of the reservoir system, though it is limited to convex problems. This limitation poses challenges in modeling many important real-world aspects, such as nonlinear head effects. Many papers have been published, specifically, in the last decade, extending the SDDP methodology to include (linearized) transmission constraints (Granville et al. 2003), natural gas supply, demand and transportation networks (Bezerra et al. 2006), alternate sampling strategies (Homem-de-Mello et al. 2011), emission quotas (Rebennack et al. 2012), risk constraints (Shapiro et al. 2013), or expansion planning (Rebennack 2014). SDDP-type algorithms have been dominating practical applications, specifically in Central America, for more than a decade (Maceira et al. 2008).

Several important algorithms have been developed which are related to the SDDP methodology, including general decomposition methods (Pereira and Pinto 1985, Morton 1996), abridged nested decomposition (Donohue 1996, Donohue and Birge 2006), cutting-plane and partial-sampling (Chen and Powell 1999), generalized dual dynamic programming (Velásquez 2002), constructive dual dynamic programming (Read 1989), and approximate dual dynamic programming (Löhndorf et al. 2013).

2.2. Lagrangian Relaxation

To combat the challenges associated with the non-convexities inherent in our problem, we turn to Lagrangian relaxation. Geoffrion (1974) provides a fundamental review of Lagrangian relaxation and Fisher (2004) reviews more recent advances in this area. The seminal work resulted from studying non-convex problems and the Traveling Salesman Problem in the 1960's and 1970's (Held and Karp 1970). In most recent applications, Lagrangian relaxation is used for solving large-scale problems, where complicating constraints are dualized so that the resulting problem is easier to solve (Van Roy 1983). In many cases, the original problem is too large to solve efficiently and Lagrangian relaxation is used to solve the problem by decomposing it into a number of solvable independent subproblems (Muckstadt and Koenig (1977), Nowak and Römisch (2000), Takriti and Birge (2000), and Cerisola et al. (2009)).

Lagrangian relaxation is not new to hydroelectric producer modeling. The approach is often used to decompose problems into manageable subproblems, especially in the presence of integer variables or nonlinearities. In these instances, the problem is decomposed by relaxing the water balance (or coupling, or hydro conservation) constraints (Nilsson and Sjelvgren 1996, Li et al. 1997, Borghetti et al. 2003, Belloni et al. 2003, Cerisola et al. 2012, Thomé 2013). Nilsson and Sjelvgren (1996) and Li et al. (1997) solve mixed-integer problems, due to the incorporation of unit startup constraints. Borghetti et al. (2003) and Belloni et al. (2003) study a similar problem but focus on updating the Lagrangian multipliers via disaggregated bundle methods.

Nonlinearities arise when modeling reservoir head effects and are often important to include, especially when modeling reservoirs with small capacities (Diniz and Maceira 2008, Catalão et al. 2009). Cerisola et al. (2012) and Thomé (2013) incorporate nonlinear reservoir head effects into their hydrothermal scheduling models. Both minimize cost while satisfying demand. The head effects in these models create Future Cost Functions (FCFs) that are nonlinear and, thus, not necessarily convex. In both the present work and the work by Cerisola and Thomé, the water balance constraints are dualized in order to make the FRFs concave (for a maximization problem) or FCFs convex (for a minimization problem). Note that the proposed approach does not convexify the subproblems solved – the subproblems remain non-convex (MILPs in Cerisola, Thomé, and in our model). The key idea is to "concavify" the FRFs or to convexify the FCFs in the state variables, *i.e.*, water reservoir levels.

Cerisola et al. propose three methods for dealing with the non-convexities. In methods one and two, they solve a linear relaxation for each subproblem to obtain dual information. By solving the corresponding Lagrangian subproblem for these dual values, they aim to strengthen the Benders optimality cuts through parallel shifting. Cerisola et al.'s third method optimizes the dual values, but the manner in which this is done is not explicitly stated and computational results are not reported. Thomé (2013) analyzes and compares two different convexification approaches: FCF convexification and component convexification. Thomé also focuses on ways to find better Lagrangian multipliers and thus tighter bounds on the associated Benders' cuts (or FCFs).

Like in "cross decomposition," we couple Lagrangian relaxation and Benders decomposition, but in a different fashion (Van Roy 1983). Cross decomposition is a modified version of Benders decomposition that aims to decrease computational time through simultaneously exploiting both primal and dual information. The method makes use of integer variables in the first stage and Lagrangian relaxation in subsequent stages. Lagrangian relaxation is used to relax a subset of complicating constraints. Cross decomposition solves Mixed-Integer Linear Programming (MILP) problems exactly in a finite number of iterations.

Though several authors have used Lagrangian relaxation in hydrothermal scheduling, most use it to solve short-term unit commitment problems. We are the first to apply this approach to profit maximization in a deregulated, bid-based market (*i.e.,* the bidding problem). Our problem is nonlinear due to a nonlinear revenue function and the revenue function is nonlinear because energy is sold via a bidding scheme. In addition, our work distinguishes itself from the papers above in that it provides a rigorous mathematical and computational analysis of our general decomposition algorithm and its connection to other relaxations as well as the original problem, through quantifying optimality and duality gaps. It is important to note that quantification of optimality gaps is only possible when the original problem can be solved to optimality. For this reason, we model systems that are small enough to solve to optimality and do not include stochastic inflows.

3. Methodology: Dynamic Convexification within Benders Decomposition using Lagrangian Relaxation

Suppose we desire to solve the T-stage problem

$$(M) \qquad \mathcal{M}(\mathbf{X}_1) := \max_{\mathbf{x}_{t+1},\mathbf{y}_t \geq 0} \sum_{t=1}^{T} \mathbf{f}_t(\mathbf{x}_{t+1}, \mathbf{y}_t)$$

$$\text{s.t.} \quad \mathbf{h}_t(\mathbf{x}_t) - \mathbf{g}_{1t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0} \qquad (\lambda_t \geq 0) \qquad \forall\, t \tag{1}$$

$$\mathbf{g}_{2t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0} \qquad \forall\, t, \tag{2}$$

in which functions $\mathbf{f}_t : \mathbb{R}^{n^1_{t+1}+n^2_t} \to \mathbb{R}$, $\mathbf{h}_t : \mathbb{R}^{n^1_t} \to \mathbb{R}^{m^1_t}$, $\mathbf{g}_{1t} : \mathbb{R}^{n^1_{t+1}+n^2_t} \to \mathbb{R}^{m^1_t}$, and $\mathbf{g}_{2t} : \mathbb{R}^{n^1_{t+1}+n^2_t} \to \mathbb{R}^{m^2_t}$; and $n_t = n^1_{t+1} + n^2_t$ and $m_t = m^1_t + m^2_t$ denote the number of decision variables and the number of constraints for stage t. Additionally, note that $\mathbf{x}_1$ is given as the parameter $\mathbf{X}_1$, and λ_t is an m^1_t-dimensional vector of dual variables associated with constraints (1). In (M), we assume that there exists a solution x^*_{t+1} and y^*_t that satisfies (1)-(2), for any value x_t. This is known as complete recourse (actually, relative complete recourse is sufficient for our purposes). Thus, with this assumption, we conclude that the feasible region of (M) is non-empty. Further, we assume that the feasible region as well as $\mathcal{M}(\mathbf{X}_1)$ is bounded for any values of x_t. The general framework shown in (M) fits LP's, MILP's, non-linear programs, and mixed-integer non-linear programs, and has a block diagonal structure. We exploit this structure to develop a decomposition technique, based on NBD.

Decomposition algorithms make use of recursive formulations. Writing (M) in a recursive form yields

$$(M_t) \qquad \eta_t(\mathbf{X}_t) := \max_{\mathbf{x}_{t+1},\mathbf{y}_t \geq \mathbf{0}} \mathbf{f}_t(\mathbf{x}_{t+1}, \mathbf{y}_t) + \eta_{t+1}(\mathbf{x}_{t+1})$$

$$\text{s.t.} \quad \mathbf{h}_t(\mathbf{X}_t) - \mathbf{g}_{1t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0} \qquad (\lambda_t \geq 0) \tag{3}$$

$$\mathbf{g}_{2t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0}, \tag{4}$$

for $t = 1, \ldots, T$ and $\eta_{T+1}(\cdot) \equiv 0$. It is important to recognize that $\mathbf{x}_{t+1}$ in (3) is a vector of decision variables in stage t, but a vector of parameters in stage $t + 1$.

The limitation, however, with Benders-decomposition-type algorithms is that they require the objective function to be concave (for a maximization problem). Decomposition algorithms are limited in this way because they use Benders optimality cuts, which are linear, to estimate the future value function, *i.e.,* $\eta_{t+1}(\cdot)$. We overcome this limitation by using Lagrangian relaxation in a unique way.

3.1. Lagrangian Relaxation

For fixed $\lambda_1, \ldots, \lambda_T$, the Lagrangian relaxation of (M) found by dualizing (1) reads

$$(\mathscr{L}_\lambda) \qquad \mathscr{L}_1(\lambda_1, \ldots, \lambda_T, \mathbf{X}_1) := \sup_{\mathbf{x}_{t+1},\mathbf{y}_t \geq \mathbf{0}} \sum_{t=1}^{T} \Big(\mathbf{f}_t(\mathbf{x}_{t+1}, \mathbf{y}_t) - \lambda'_t\Big(\mathbf{h}_t(\mathbf{X}_t) - \mathbf{g}_{1t}(\mathbf{x}_{t+1}, \mathbf{y}_t)\Big)\Big)$$

$$\text{s.t.} \quad \mathbf{g}_{2t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0} \qquad \forall\, t.$$

Recall that we denote the transpose of λ_t as λ'_t. For any fixed $\lambda_1, \ldots, \lambda_T \geq \mathbf{0}$, ($\mathscr{L}_\lambda$) is a relaxation of ($M$), and, therefore $\mathscr{L}_1(\lambda_1, \ldots, \lambda_T, \mathbf{X}_1)$ provides an upper bound on $\mathcal{M}(\mathbf{X}_1)$. To find the tightest possible upper bound obtainable through this Lagrangian relaxation, we seek the set of multipliers that solve

$$(\mathscr{L}) \qquad \mathscr{L}_1(\mathbf{X}_1) := \min_{\lambda_1,\ldots,\lambda_T \geq \mathbf{0}} \mathscr{L}_1(\lambda_1, \ldots, \lambda_T, \mathbf{X}_1).$$

Rather than explicitly solving ($\mathscr{L}$) to find the optimal set of Lagrangian multipliers, we may use a subgradient method (Polyak 1969, Fisher 1985), a surrogate subgradient method (Zhao et al. 1999), or a bundle method (Lemaréchal and Sagastizábal 1997, Kaskavelis and Caramanis 1998). The practitioner typically tailors the method to the application being studied, *cf.* Section 5.1.

In order to ensure the concavity of function $\eta_t(\mathbf{X}_t)$ in $\mathbf{X}_t$, one typically uses a convex relaxation of the feasible region and a concave overestimator for $f_t(\cdot)$. As such, these relaxations are static. Rather than taking this approach, we resolve the non-concavity issues with respect to $\mathbf{X}_t$ by employing Lagrangian relaxation dynamically for $t = 1, \dots, T$,

$$(\mathscr{L}_{\lambda t}) \quad \mathscr{L}_t(\lambda_t, \dots, \lambda_T, \mathbf{X}_t) := \sup_{\mathbf{x}_{t+1}, \mathbf{y}_t \geq 0} \mathbf{f}_t(\mathbf{x}_{t+1}, \mathbf{y}_t) - \lambda_t'\Big(\mathbf{h}_t(\mathbf{X}_t) - \mathbf{g}_{1t}(\mathbf{x}_{t+1}, \mathbf{y}_t)\Big) + \mathscr{L}_{t+1}(\lambda_{t+1}, \dots, \lambda_T, \mathbf{x}_{t+1}) \quad \text{s.t. } \mathbf{g}_{2t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0}.$$

This relaxed formulation overestimates $\eta_t(\cdot)$ for any $\lambda_1, \dots, \lambda_T \geq \mathbf{0}$. To find the tightest bound using this recursive formulation we must find the set of optimal $\lambda_t, \dots, \lambda_T$ that solve

$$(\mathscr{L}_t) \qquad \mathscr{L}_t(\mathbf{X}_t) := \min_{\lambda_t, \dots, \lambda_T \geq \mathbf{0}} \mathscr{L}_t(\lambda_t, \dots, \lambda_T, \mathbf{X}_t).$$

This is done recursively so that in stage t we only seek the optimal values for λ_t, because the optimal values for $\lambda_{t+1}, \dots, \lambda_T$ have already been determined and are fixed in this stage of the recursive solve.

Interestingly, in the special case in which $\mathbf{h}_t(\mathbf{X}_t)$ is convex in $\mathbf{X}_t$, for fixed $\lambda_t, \dots, \lambda_T \geq \mathbf{0}$, function $\mathscr{L}_t(\lambda_t, \dots, \lambda_T, \mathbf{X}_t)$ is concave in $\mathbf{X}_t$. Consequently, $\mathscr{L}_t(\mathbf{X}_t)$ is the minimum of a collection of functions that are all concave in $\mathbf{X}_t$, which implies $\mathscr{L}_t(\mathbf{X}_t)$ is concave. This is summarized in

Remark 1. *If $\mathbf{h}_t(\mathbf{X}_t)$ is a convex function for any $t = 1, \dots, T$, then the function $\mathscr{L}_t(\mathbf{X}_t)$ is concave in $\mathbf{X}_t$ for any $t = 1, \dots, T$.*

Note, replacing the future value function $\mathscr{L}_{t+1}(\cdot)$ in ($\mathscr{L}_t$) by any overestimator does not affect the shape of $\mathscr{L}_t(\cdot)$.

3.2. Benders Optimality Cuts

Consider the relaxed $t+1$-stage problem, for any $\lambda_{t+1} \geq \mathbf{0}$

$$(\hat{\mathscr{L}}_{\lambda, t+1}) \qquad \hat{\mathscr{L}}_{t+1}(\lambda_{t+1}, \mathbf{X}_{t+1}) := \sup_{\mathbf{x}_{t+2}, \mathbf{y}_{t+1} \geq 0} \mathbf{f}_{t+1}(\mathbf{x}_{t+2}, \mathbf{y}_{t+1}) - \lambda_{t+1}'\Big(\mathbf{h}_{t+1}(\mathbf{X}_{t+1}) - \mathbf{g}_{1,t+1}(\mathbf{x}_{t+2}, \mathbf{y}_{t+1})\Big) + \hat{\mathscr{L}}_{t+2}(\mathbf{x}_{t+2}) \quad \text{s.t. } \mathbf{g}_{2,t+1}(\mathbf{x}_{t+2}, \mathbf{y}_{t+1}) \leq \mathbf{0},$$

in which $\hat{\mathscr{L}}_{t+2}(\mathbf{x}_{t+2})$ overestimates $\mathscr{L}_{t+2}(\mathbf{x}_{t+2})$. We shift the focus to the $t+1$-stage problem so that we can develop the methodology used to compute the Benders optimality cuts that approximate the FRF.

We approximate (overestimate) $\hat{\mathscr{L}}_{t+1}(\lambda_{t+1}, \mathbf{X}_{t+1})$ using the Benders optimality cut

$$-\lambda_{t+1}' \mathbf{h}_{t+1}(\chi_{t+1}) + \gamma_{t+1}^{\text{const}},$$

in which $\gamma_{t+1}^{\text{const}}$ is calculated via

$$\gamma_{t+1}^{\text{const}} := \hat{\mathscr{L}}_{t+1}(\lambda_{t+1}, \mathbf{X}_{t+1}) + \lambda_{t+1}' \mathbf{h}_{t+1}(\mathbf{X}_{t+1}). \tag{5}$$

In this construct, in accordance with our notation convention, $\mathbf{X}_{t+1}$ is a parameter and χ_{t+1} is a decision variable. The validity of this optimality cut, which is not necessarily linear, is established in

Theorem 1. *The function or "cut"*

$$-\lambda_{t+1}' \mathbf{h}_{t+1}(\chi_{t+1}) + \gamma_{t+1}^{\text{const}} \tag{6}$$

overestimates $\eta_{t+1}(\chi_{t+1})$, for any $t = 1, \dots, T$.

The proof for Theorem 1 is given in Appendix B. It follows from the proof that, for any $\lambda_{t+1} \geq \mathbf{0}$, the Benders optimality cut (6) is valid as long as $(\hat{\mathscr{L}}_{\lambda,t+1})$ has been solved to global optimality or when an upper bound on $\hat{\mathscr{L}}_{t+1}(\lambda_{t+1}, \mathbf{X}_{t+1})$ is used in the cut constant calculation (5). We summarize this in

Remark 2. *The Benders optimality cuts are valid for any $\lambda_{t+1} \geq \mathbf{0}$.*

Given $L-1$ Benders optimality cuts, we incorporate them into (M_t) and write

$$(\hat{M}_t) \qquad \hat{\eta}_t(\mathbf{X}_t) := \max_{\mathbf{x}_{t+1}, \mathbf{y}_t \geq 0} \; \mathbf{f}_t(\mathbf{x}_{t+1}, \mathbf{y}_t) + \hat{\eta}_{t+1}$$

$$\text{s.t.} \quad (3), (4)$$

$$\hat{\eta}_{t+1} \leq -\lambda'_{l,t+1} \mathbf{h}_{t+1}(\mathbf{x}_{t+1}) + \gamma^{\text{const}}_{l,t+1} \qquad l = 1, \ldots, L-1. \tag{7}$$

Similarly, we rewrite formulation $(\mathscr{L}_t)$ as

$$(\hat{\mathscr{L}}_t) \qquad \hat{\mathscr{L}}_t(\mathbf{X}_t) := \min_{\lambda_t \geq \mathbf{0}} \; \sup_{\mathbf{x}_{t+1}, \mathbf{y}_t \geq 0} \; \mathbf{f}_t(\mathbf{x}_{t+1}, \mathbf{y}_t) - \lambda'_t\Big(\mathbf{h}_t(\mathbf{X}_t) - \mathbf{g}_{1t}(\mathbf{x}_{t+1}, \mathbf{y}_t)\Big) + \hat{\eta}_{t+1}$$

$$\text{s.t.} \quad (4), (7).$$

3.3. Benders Decomposition

We now explain how the Benders optimality cuts given in (7) are used in conjunction with NBD (see Algorithm 1). Each iteration of the algorithm is divided into a *forward pass* and a *backward pass*. The $\mathbf{X}_{t+1}$ values link the subproblems in the forward and backward passes.

In the forward pass, the goal is to compute a feasible solution to the original problem (M), yielding a lower bound on $\mathcal{M}(\mathbf{X}_1)$. In the L^{th} iteration, we solve the global optimization problem given in $(\hat{M}_t)$ for each stage t. In $(\hat{M}_t)$ the $L-1$ cuts in (7) are computed in the previous $L-1$ backward pass iterations. The forward pass also yields trial values for the state variables, $\mathbf{X}_{t+1}$, which are required for the backward pass.

In the backward pass, we use the trial $\mathbf{X}_{t+1}$ values and compute (near) optimal Lagrangian multipliers to generate valid Benders optimality cuts, which overestimate function $\mathscr{L}_{t+1}(\mathbf{X}_{t+1})$ and, thus, $\eta_{t+1}(\mathbf{X}_{t+1})$. Specifically, the cut for stage t during the L^{th} iteration, is generated by solving $(\hat{\mathscr{L}}_{t+1})$. Each time we solve $(\hat{\mathscr{L}}_{t+1})$, we generate exactly one Benders optimality cut for the future value function. An upper bound on $\mathcal{M}(\mathbf{X}_1)$ and on $\mathscr{L}_1(\mathbf{X}_1)$ is found by solving $(\hat{M}_1)$ for the given initial value $\mathbf{X}_1$.

Depending on the shape of the $\mathbf{f}_t(\cdot)$, $\mathbf{g}_{1t}(\cdot)$, $\mathbf{g}_{2t}(\cdot)$, and $\mathbf{h}_t(\cdot)$ functions, there may be a non-zero duality gap defined by

$$\mathscr{L}_1(\mathbf{X}_1) - \mathcal{M}(\mathbf{X}_1) \geq 0. \tag{8}$$

Thus, in general, the lower bound $\underline{z}$ and the upper bound $\overline{z}$ do not converge and the stopping criteria in Algorithm 1 has to be chosen carefully (*e.g.,* $\overline{z} - \underline{z} <$ a specified value).

Algorithm 1 computes a feasible solution to (M) (*i.e.,* a lower bound) and an upper bound on $\mathcal{M}(\mathbf{X}_1)$. The bounds are obtained by computing optimality cuts on a relaxed problem and by incorporating these cuts into the original recursive problem. In doing so, we overestimate the future value function in order to generate a feasible solution to the original problem (M). If the problem possesses a non-zero duality gap, as defined in (8), then Algorithm 1 may neither solve (M) nor $(\mathscr{L})$ to optimality, regardless of whether or not optimal Lagrangian multipliers have been obtained. Assuming optimal Lagrangian multipliers are obtained, Algorithm 1 can be adjusted to solve problem $(\mathscr{L})$ to any given accuracy in finitely many forward-backward passes (under some mild conditions, for instance, that $\mathscr{L}_t(\mathbf{X}_t)$ is a continuous function in $\mathbf{X}_t$ on an appropriate compactum) through solving $(\hat{\mathscr{L}}_t)$ in the forward pass instead of $(\hat{M}_t)$ and by defining appropriate stopping criteria.

For stage t, if optimal Lagrangian multipliers are used to generate cut (7) for some value $\mathbf{X}_t$, then the cut is tight for $\hat{\mathscr{L}}_t(\mathbf{X}_t)$. Thus, for stage T, the computed cut is also tight for $\mathscr{L}_T(\mathbf{X}_T)$. Eventually, after multiple backward-forward passes, and given that the stopping criteria has not been satisfied, the computed cuts in Algorithm 1 are tight for all other stages at specific $\mathbf{X}_t$ values.

Algorithm 1 Benders Decomposition with Lagrangian Relaxation

Step (0). Initialize: $\hat{\eta}_{t+1} = 0,\ \forall\ t$ for first forward pass only, $\underline{z} = 0$, $\overline{z} = +\infty$, and $\iota = 1$.
while Stopping criteria not satisfied **do**
  Forward Pass
  for $t = 1, \ldots, T$ **do**
    Step (1). Solve the t-stage (global) optimization problem $(\hat{M}_t)$. Let $\mathbf{X}_{t+1,\iota} := \mathbf{x}^*_{t+1}$ and $\mathbf{y}_{t\iota} = \mathbf{y}^*_t$.
  end for
  Step (2). Update lower bound $\underline{z}$ on $\mathcal{M}(\mathbf{X}_1)$:

$$\underline{z} \leftarrow \max \left\{ \underline{z}, \sum_{t=1}^{T} \mathbf{f}_t(\mathbf{X}_{t+1,\iota}, \mathbf{y}_{t\iota}) \right\}.$$

  Backward Pass
  for $t = T, T-1, \ldots, 2$ **do**
    Step (3). With the stored $\mathbf{X}_{t\iota}$ values from the forward pass, solve the t-stage Lagrangian problem $(\hat{\mathscr{L}}_t)$ for (near) optimal Lagrangian multipliers.
    Step (4). Calculate a new Benders optimality cut (7) for stage t using the Lagrange multipliers and objective function value obtained from *Step (3)*.
  end for
  Step (5). Solve the first-stage problem $(\hat{M}_1)$.
  Step (6). Calculate upper bound $\overline{z}$ on $\mathcal{M}(\mathbf{X}_1)$:

$$\overline{z} = \hat{\eta}_1(\mathbf{X}_1).$$

  Step (7). Increment the iteration count $\iota \leftarrow \iota + 1$.
end while
Step (8). **Exit** with $\underline{z}, \overline{z}, \mathbf{X}_{t+1,\iota-1}$, and $\mathbf{y}_{t,\iota-1}$.

3.4. Linear and Mixed-Integer Linear Programming

There are two special cases that are worth discussing in greater detail, namely when the original problem (M) is an LP or a MILP. In each case, if we assume that $\mathbf{h}_t(\cdot)$ and $\mathbf{g}_{1t}(\cdot)$ are linear functions, and $\mathbf{g}_{2t}(\cdot)$ models the integrality restrictions (when (M) is a MILP), then $(\mathscr{L}_\lambda)$ has a linear objective function and optimizes over a polytope with finitely many (and possibly integer) extreme points. Thus, the Lagrangian function is piecewise-linear. This is summarized in

Proposition 1. *If (M) is an LP or a MILP then the function $\mathscr{L}_t(\mathbf{X}_t)$ is piecewise-linear and concave in $\mathbf{X}_t$, for any $t = 1, \ldots, T$.*

In these cases, replacing the future value function in $(\mathscr{L}_t)$ by a piecewise-linear and concave overestimator does not affect the shape of $\mathscr{L}_t(\cdot)$. It follows that a finite number of Benders optimality cuts are sufficient to represent $\mathscr{L}_t(\cdot)$ exactly.

If (M) is an LP, then the duality gap (8) is zero (strong duality holds). In this case, if the optimal Lagrangian multipliers are obtained and an appropriate stopping criterion is chosen, Algorithm 1 solves problem (M) to optimality. More specifically, for LP's, the presented methodology is a (computationally inefficient) generalization of classical NBD, in which dual variable values are obtained via Lagrangian relaxation instead of using duality theory.

In general, MILP problems do not possess a zero duality gap. Thus, Algorithm 1 is not guaranteed to compute an optimal solution for MILP problems. However, our Lagrangian approach dominates an LP relaxation-based method. Dualizing (1), in the LP relaxation, yields a corresponding Lagrangian which has the same objective function value as the LP relaxation (when using optimal Lagrangian multipliers) because LPs possess a zero duality gap; strong duality holds. Moreover, our Lagrangian is more restrictive, because integrality is enforced, and thus leads to an upper bound which is not worse than the LP relaxation. Since, in our application (M) is a MILP, we highlight

Corollary 1. *Let (M) be a MILP. Solving the Lagrangian relaxation $(\mathscr{L})$ yields an upper bound for (M) which is at least as tight as the upper bound obtained from solving its LP relaxation $(\overline{M})$.*

The proof of Corollary 1 is given in Appendix A. Note that $(\overline{M})$ is formed by replacing the constraints in the MILP (M) that require specified variables to be integral by weaker constraints allowing these same variables to be continuous.

3.5. Stochastic Dual Dynamic Programming

If we desire to model parameters that are uncertain, then Algorithm 1 can be modified accordingly. With this change, uncertainty is considered via a scenario tree and the algorithm becomes an extension of SDDP. SDDP, as originally proposed, models the uncertainty as stagewise independent. In doing so, SDDP is appealing because it exploits this special structure by (i) sampling a small subset of scenarios from the scenario tree in the forward pass, and (ii) sharing cuts amongst scenarios in the backward pass. When applying this methodology, the expectation operator, which is based on scenario realizations and their associated probabilities, is used and scenarios are considered in both the forward pass and when computing the optimality cuts in the backward pass. If the uncertainty is assumed to be stagewise independent, these are essentially the only changes that need to be made to Algorithm 1.

However, if the uncertainty is assumed to be stagewise dependent (Infanger and Morton 1996, Queiroz and Morton 2013, Lohmann et al. 2016), then past inflows become additional state variables for the FRFs or the FCFs. With stagewise-dependent uncertainty, the dimensionality of the FRFs or the FCFs are increased and the FRFs or the FCFs might be non-concave or non-convex in the additional state variables. To concavify or convexify the FRFs or the FCFs, in these additional arguments, all constraints that contain past inflow state variables must be dualized as well. These constraints are the Bender cuts.

4. Application: The Bidding Problem

We are now ready to present the formulation for the problem of interest. We assume the electricity market in which the single price-maker hydro producer sells energy, consists of J price-taker thermal producers. The hydro producer is assumed to be a price maker because it controls a large portion of the energy offered in the market. This assumption fits several electricity markets in Central America. Alternatively, with minor alterations, we could model a market that consists of multiple price makers and use the method described below to solve each price-maker's best response problem, assuming the actions from all rivals are known. Doing so would extend the work in Bushnell (2003) to a bid-based market and is an area for future research. To begin, we examine the problem over a one-stage time horizon. We make use of the following notation:

Indices and Sets:

$j = 1, \ldots, J$ index for price-taker thermal producers

Parameters:

$\overline{D}$ demand in market [GWh]

$\overline{G}_j$ quantity bid (capacity) for producer j, $j = 1, \ldots, J$ [GWh]

C_j price bid (variable operating cost) for producer j, $j = 1, \ldots, J$ [\$/GWh]

Decision Variables:

e hydro production (quantity bid) for price-maker [GWh]

g_j quantity of energy produced (sold in market) for producer j [GWh]

π^d market-clearing price (dual of demand constraint) [\$/GWh]

4.1. Market Clearing and Immediate Revenue Function

The price maker seeks to maximize revenue, or mathematically,

$$\max \; R(e) = \pi^d \cdot e.$$

The difficulty is that the market-clearing price, π^d, depends on how much the price maker bids, e, and vice versa. To remedy this, we seek to determine how the revenue function changes with different values of e, which can be determined by studying the market-clearing formulation. In the day-ahead market, the operator satisfies demand while minimizing cost and the market-clearing price, π^d, is determined based on the value of e. The market-clearing formulation, for a fixed bid quantity from the price maker, E_k, is given as

$$\begin{aligned} \min \quad & \sum_{j=1}^{J} C_j g_j \\ \text{s.t.} \quad & \sum_{j=1}^{J} g_j = \overline{D} - E_k && (\pi_k^d) \\ & 0 \le g_j \le \overline{G}_j \qquad \forall j. \end{aligned}$$

The set of g_j, $j = 1, \ldots, J$, such that $g_j > 0$, in a solution to the market-clearing formulation, is referred to as the market-clearing dispatch. The variable π^d is given as the dual (marginal value) of the demand constraint, *i.e.,* π^d is the highest price bid from the set of producers that are part of the market-clearing dispatch. Regardless of what the producers bid, all producers that are part of the market-clearing dispatch are paid the market-clearing price π^d.

We assume the price-maker producer submits their bid at zero price, so that all of the quantity they offer (given adequate demand) is dispatched. Additionally, we assume that all of the price-taker producers offer quantity bids equal to their capacity at a price equal to their variable cost, *i.e.*, they bid according to their optimal bidding strategy (Gross and Finlay 2000). Knowledge of each price-taker j's bid allows us to populate parameters C_j and $\overline{G}_j$ in the market-clearing formulation and, consequently, determines the shape of the price maker's revenue function. This is a large assumption. In practice, competitors' variable costs and, thus, their bids are not known.

In Figure 1, we plot the revenue curve $R(e)$ (solid gray line), *i.e.*, the revenue as a function of the price maker's quantity bids. The function has a sawtooth shape, since as the price maker's bid increases, the most costly price-taker producer that is part of the market-clearing dispatch "drops out." Because the market-clearing price is equivalent to the most costly price bid in the market-clearing dispatch, when price-taker producers are forced out of the dispatch by the price maker, the market-clearing price (slope of the revenue curve) decreases.

Because we know the optimal behavior of the price maker, we can provide a representation of the revenue curve by eliminating "downward" steps, as depicted with the dashed black lines in Figure 1. For our intents and purposes, this is an exact representation because the agent's optimal generation decisions will not include inefficient amounts (spillage is not restricted). For example, any quantity of hydro production between break point $\dot{E}_{1t}$ and break point $\dot{E}_{2t}$ will not be chosen because the same revenue can be achieved by producing the lesser amount $\dot{E}_{1t}$. The regions shaded in gray denote efficient quantities (*i.e.,* generation decisions that may be optimal).

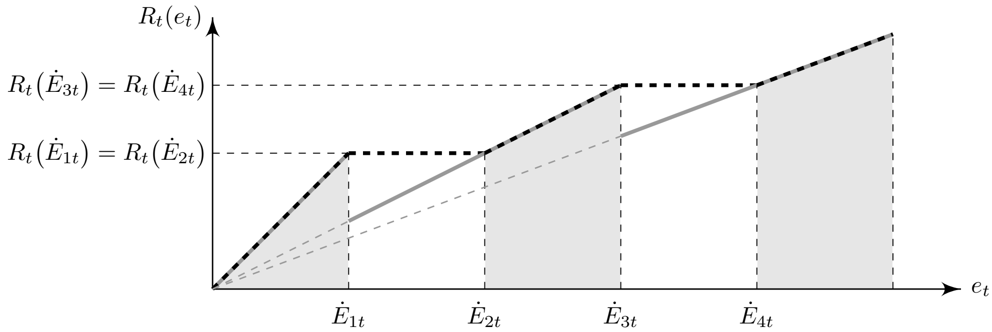

Figure 1. Representation of the revenue function

4.2. Formulation

With the above knowledge of the revenue function, we present the bidding problem formulation for a single price-maker producer. For a time horizon lasting longer than one stage, we distinguish between the immediate revenue function and the FRF. For a given stage, the immediate revenue function yields the revenue gained in that stage, is a function of production in that stage, and is found in the exact manner discussed in Section 4.1. The FRF yields the maximum achievable revenue in all future stages, based on the amount of water available in the future stages, which is a function of current production. The notation we use in this formulation and throughout the remainder of this paper is defined below.

Indices and Sets:

$i = 1, \ldots, I$ index for reservoirs operated by the price-maker

$k = 0, \ldots, K$ index for the breakpoints of the revenue function $R_t(e_t)$

$t = 1, \ldots, T$ index for stages

$\mathbb{U}_i$ set of reservoirs that are immediately upstream of reservoir i

Parameters:

$\dot{E}_{kt}$ k^{th} breakpoint of the revenue function, in stage t [GWh]

$\dot{R}_{kt}$ value of the revenue function at the k^{th} breakpoint, in stage t [\$/stage]

$\underline{V}_i$ minimum storage storage level for reservoir i [hm^3]

$\overline{V}_i$ maximum storage level for reservoir i [hm^3]

V_{i1} initial storage level for reservoir i [hm^3]

$\overline{U}_i$ water turbine outflow capacity for reservoir i [hm^3]

A_{it} inflow for reservoir i, in stage t [hm^3]

ρ_i production coefficient (assumed constant) for reservoir i [GWh/hm^3]

Functions:

$R_t(e_t)$ immediate revenue function, based on total production $e_t = \sum_{i=1}^{I} e_{it}$, for stage t [\$/stage]

Special Ordered Set Variables of type 2 (SOS2):

μ_{kt} indicator to determine which segment k of the revenue function e_t is on [-]

Decision Variables:

e_t total price-maker hydro production (bid) in stage t [GWh]

e_{it} price-maker hydro production (bid) from reservoir i in stage t [GWh]

s_{it} spillage from reservoir i in stage t [hm^3]

v_{it} volume of reservoir i at the beginning of stage t [hm^3]

λ_{it} dual variable associated with reservoir i in stage t [\$/hm^3]

The problem we wish to solve, the bidding problem, is given as

$$(B^R) \qquad \beta_1(\mathbf{V}_1) := \max \sum_{t=1}^{T} R_t(e_t)$$

$$\text{s.t.} \quad v_{i,t+1} = v_{it} + A_{it} - \frac{e_{it}}{\rho_i} - s_{it} + \sum_{m \in \mathbb{U}_i} \left(\frac{e_{mt}}{\rho_m} + s_{mt} \right) \qquad \forall i, t \qquad (9)$$

$$\underline{V}_i \le v_{i,t+1} \le \overline{V}_i \qquad \forall i, t \qquad (10)$$

$$0 \leq \frac{e_{it}}{\rho_i} \leq \overline{U}_i \qquad \forall i, t \qquad (11)$$

$$e_t = \sum_{i=1}^{I} e_{it} \qquad \forall t \qquad (12)$$

$$s_{it} \geq 0 \qquad \forall i, t, \qquad (13)$$

where $v_{i1} \equiv V_{i1}$. Constraint (9) is the water balance constraint that couples our decisions in stage t with all past and future stages. We assume that water used for production or spilled from reservoir $m \in \mathbb{U}_i$, immediately upstream from reservoir i, is available to reservoir i in the same stage; *cf.* Section 5.3.2. Constraints (10) and (11) guarantee that we do not exceed reservoir capacity and that the water outflow does not exceed the capacity limitations of the water turbine, for any of the I reservoirs. Constraint (12) sums production from each of the price-maker's reservoirs yielding total production. We now discuss the specific integer restrictions required to model the non-concave revenue $R_t(\cdot)$.

The bidding problem (B^R) can be equivalently written as

$$(B) \qquad \beta_1(\mathbf{V}_1) := \max \sum_{t=1}^{T} \sum_{k=1}^{K} \mu_{kt} \dot{R}_{kt}$$

$$\text{s.t.} \quad (9) - (13)$$

$$e_t = \sum_{k=1}^{K} \mu_{kt} \dot{E}_{kt} \qquad \forall t \qquad (14)$$

$$\sum_{k=0}^{K} \mu_{kt} = 1 \qquad \forall t \qquad (15)$$

$$0 \leq \mu_{kt} \leq 1 \qquad \forall k, t \qquad (16)$$

$$\{\mu_{0t}, \ldots, \mu_{Kt}\} \in \text{SOS2} \qquad \forall t. \qquad (17)$$

When we refer to solving the monolith formulation, we are referring to solving (B) directly, *i.e.,* without any tailored (decomposition) algorithm, using off-the-shelf software. In (B), the revenue for each stage is found via the methodology discussed in Section 4.1 and the $\dot{E}_{kt}$, and $\dot{R}_{kt}$ values are parameters that are found a priori, based on our knowledge of the price takers' bids. Constraints (14) through (17) ensure that we model the piecewise-linear segments of the revenue function correctly. SOS2 variables are an ordered set of non-negative variables, of which at most two can be non-zero (Beale 1963). And, if two are non-zero they must be consecutive in their ordering.

Constraints (17) can be modeled with binary variables

$$\phi_{kt} = \begin{cases} 1, & \text{if } e_t \text{ is in segment } k, \\ 0, & \text{otherwise} \end{cases} \qquad \forall k = 1, \ldots, K, t,$$

and the following constraints to ensure $\phi_k = 1$ if and only if e_t is on line segment k:

$$\sum_{k=1}^{K} \phi_{kt} = 1 \qquad \forall t \qquad (18)$$

$$\mu_{0t} \leq \phi_{1t} \qquad \forall t \qquad (19)$$

$$\mu_{kt} \leq \phi_{kt} + \phi_{k+1,t} \qquad k = 1, \ldots, K-1, \ \forall t \qquad (20)$$

$$\mu_{Kt} \leq \phi_{Kt} \qquad \forall t \qquad (21)$$

$$\phi_{kt} \in \{0, 1\} \qquad \forall k = 1, \ldots, K, t. \qquad (22)$$

Many off-the-shelf MILP solvers (such as XPRESS-MP, CPLEX and GUROBI) allow users to declare decision variables as SOS2 without having to specify an explicit model like (18)-(22). In other words, the binary variables ϕ_{kt} do not need to be added to the model. The off-the-shelf solver then deals with the SOS2 variables in its own way, *e.g.*, by using a model like (18)-(22) or by exploiting the special structure via tailored branching rules. Other mathematical programming formulations for the immediate revenue function, $R_t(e_t)$, become attractive when one realizes that the function is continuous and piecewise-linear in the hydro production. These include incremental (Padberg 2000), disaggregated (Sherali 2001), and logarithmic approaches (Vielma and Nemhauser 2011).

Note that the MILP presented in (B) is a special case of (M) in Section 3. In (B), $\sum_{k,t} \mu_{kt} \dot{R}_{kt}$ is analogous to $\mathbf{f}_t(\cdot)$ and constraints (10)-(17) are analogous to $\mathbf{g}_{2t}(\cdot)$. The water balance constraint given in (9) is a special case of (1) in which $h_t(\cdot)$ is an identity for the first I constraints and a negative identity for the remaining I constraints, in order to convert equation (9) into inequality constraints (1). Thus, the associated Lagrangian multiplier used in this application is unrestricted in sign.

4.3. LP Relaxation and Concave Overestimation

Current state-of-the-art MILP solvers are able to solve (B) with an acceptable computational effort, even for large reservoir systems and many stages. However, when the reservoir inflows A_{it} are unknown, stochastic programming techniques are the methods of choice for solving (B). Treating inflows via stochastic programming techniques leads to scenario-tree-type approaches. Unfortunately, the resulting MILP problems are several orders of magnitude too large for current state-of-the-art MILP solvers. Consequently, we rely on a decomposition algorithm, namely Benders decomposition.

To overcome issues with non-concavity, in a different context, Newham (2008) solves the LP relaxation of (B),

$$(\overline{B}) \qquad z(\overline{B}) := \max \sum_{t=1}^{T} \sum_{k=1}^{K} \mu_{kt} \dot{R}_{kt}$$
$$\text{s.t. } (9)-(16), (18)-(21)$$
$$0 \le \phi_{kt} \le 1 \qquad \forall k = 1, \ldots, K, t\,.$$

The LP relaxation creates an outer approximation of the feasible region of (B), which makes it a convex set. The dual values from (9) are then used to generate Benders optimality cuts.

Flach et al. (2010) suggest another approach to convexify (B). Flach et al. employ a concave, piecewise-linear overestimator to approximate the immediate revenue function $R_t(\cdot)$, which leads to the LP formulation (C):

$$(C) \qquad z(C) := \max \sum_{t=1}^{T} \sum_{n=1}^{N} \alpha_{nt} \tilde{R}_{nt}$$
$$\text{s.t. } (9)-(13)$$
$$e_t = \sum_{n=1}^{N} \alpha_{nt} \tilde{E}_{nt} \qquad \forall t$$
$$\sum_{n=0}^{N} \alpha_{nt} = 1 \qquad \forall t$$
$$0 \le \alpha_{nt} \le 1 \qquad \forall\, n, t\,.$$

The breakpoints $\tilde{E}_{nt}$ and $\tilde{R}_{nt}$ are chosen so that the resulting segments of the piecewise-linear overestimated revenue curve have decreasing slopes. By selecting $\tilde{E}_{nt}$ and $\tilde{R}_{nt}$ in this manner, Flach et al. ensure the overestimation is concave. It is important to note that the values for the parameters $\tilde{E}_{nt}$ and $\tilde{R}_{nt}$ may differ from the values of the parameters $\dot{E}_{kt}$ and $\dot{R}_{kt}$ found in (B). Though the formulations in ($\overline{B}$) and (C) appear to be different, they result in the same optimal objective function value, which we formalize in

Theorem 2. *The optimal objective function values of the LP relaxation formulation ($\overline{B}$) and the concave overestimator formulation (C) are the same, i.e., $z(\overline{B}) = z(C)$.*

Theorem 2 holds because the tightest concave overestimator of a (continuous,) piecewise, affine function can be formed by using the convex combinations of the function values at the breakpoints. This is formalized in the proof for Theorem 2, presented in Appendix C. In addition to the upper bound derived from ($\overline{B}$) and (C) on the bidding problem (B), NBD applied to ($\overline{B}$) and (C) can be adapted to yield several lower bounds, *cf.* Section 5.2.

5. Application Methodology: Solving the Bidding Problem using Benders Decomposition and Lagrangian Relaxation

We desire a decomposition method that can solve the bidding problem (B) for hydro producers with significant storage (ten or more reservoirs), over time horizons of one year or longer with weekly or monthly fidelity. Both of the approaches devised by Newham and Flach et al. aim to do this, but can result in large optimality gaps because the revenue function they utilize is an approximation. In our approach, we model the revenue exactly and then employ Lagrangian relaxation through dualizing the water balance constraints that connect the different stages with each other to quell non-concavity issues, *i.e.*, we apply Algorithm 1 to the bidding problem (B). We write (B) in the following recursive form, corresponding to (M_t), for $t = 1, \dots, T$:

$$(B_t) \qquad \beta_t(\mathbf{V}_t) := \max \sum_{k=1}^{K} \mu_{kt}\dot{R}_{kt} + \beta_{t+1}(\mathbf{v}_{t+1})$$

$$\text{s.t.} \quad \sum_{i=1}^{I} e_{it} = \sum_{k=1}^{K} \mu_{kt}\dot{E}_{kt} \tag{23}$$

$$\sum_{k=0}^{K} \mu_{kt} = 1 \tag{24}$$

$$v_{i,t+1} = V_{it} + A_{it} - \frac{e_{it}}{\rho_i} - s_{it} + \sum_{m\in\mathbb{U}_i}\left(\frac{e_{mt}}{\rho_m} + s_{mt}\right) \qquad (\lambda_{it}) \ \forall i \tag{25}$$

$$\underline{V}_i \le v_{i,t+1} \le \overline{V}_i \qquad \forall i \tag{26}$$

$$0 \le \frac{e_{it}}{\rho_i} \le \overline{U}_i \qquad \forall i \tag{27}$$

$$0 \le \mu_{kt} \le 1 \qquad \forall k \tag{28}$$

$$s_{it} \ge 0 \qquad \forall i \tag{29}$$

$$\{\mu_{0t}, \dots, \mu_{Kt}\} \in \text{SOS2}, \tag{30}$$

in which $\beta_{T+1}(\cdot) \equiv 0$. Since the water balance equations (25) in problem (B_t) link the decisions between stages, we represent this time-dependency via the state variables $\mathbf{V}_t$. Note, problem (B_t) is feasible for any vector $\mathbf{V}_t$, with $\underline{\mathbf{V}} \le \mathbf{V}_t \le \overline{\mathbf{V}}$, because $A_{it} \ge 0$ and $\dot{E}_{0t} = 0$.

The Benders optimality cuts for stage t are computed by solving the following $t+1$-stage problem, for a given $\boldsymbol{\lambda}_{t+1}$ and $\mathbf{V}_{t+1}$:

$$(\hat{\mathscr{L}}_{\lambda,t+1}) \quad \hat{\mathscr{L}}_{t+1}(\boldsymbol{\lambda}_{t+1}, \mathbf{V}_{t+1}) = \max_{e_{t+1},\mathbf{v}_{t+2},\mathbf{s}_{t+1}} \Bigg[\sum_{k=1}^{K} \mu_{k,t+1}\dot{R}_{k,t+1} + \sum_{i=1}^{I} \lambda_{i,t+1}\left(V_{i,t+1} + A_{i,t+1} - \frac{e_{i,t+1}}{\rho_i} - v_{i,t+2} - s_{i,t+1} + \sum_{m\in\mathbb{U}_i}\left(\frac{e_{m,t+1}}{\rho_m} + s_{m,t+1}\right)\right) + \hat{\beta}_{t+2}(\mathbf{v}_{t+2})\Bigg]$$

$$\text{s.t.} \quad (23), (24), (26) - (30).$$

Function $\hat{\beta}_{t+2}(\mathbf{v}_{t+2})$ is the approximate FRF within stage $t+1$ that approximates the cumulative revenues of stages $t+2, \dots, T$, *i.e.,* the approximation of $\beta_{t+2}(\mathbf{v}_{t+2})$. Function $\hat{\beta}_{t+1}(\cdot)$ is then approximated, for each fixed value of $\boldsymbol{\lambda}_{t+1}$ and $\mathbf{V}_{t+1}$, via the linear cut

$$\sum_{i=1}^{I} \gamma_{it}^{\text{slope}} V_{i,t+1} + \gamma_t^{\text{const}}, \tag{31}$$

with slope $\gamma_{it}^{\text{slope}}$ and constant term γ_t^{const}. These parameters are calculated as

$$\gamma_{it}^{\text{slope}} = \lambda_{i,t+1} \qquad \text{and} \qquad \gamma_t^{\text{const}} = \hat{\mathscr{L}}_{t+1}(\boldsymbol{\lambda}_{t+1}, \mathbf{V}_{t+1}) - \sum_{i=1}^{I} \lambda_{i,t+1} V_{i,t+1}.$$

Note that cut (31) is valid for any water reservoir level (*cf.* Theorem 1), not only within its bounds $\underline{\mathbf{V}}$ and $\overline{\mathbf{V}}$. This has important implications for the tightness of the computed cuts found through Lagrangian relaxation, *cf.* Section 5.3.

5.1. Near-Optimal Lagrangian Multipliers

For given (initial) water reservoir levels in stage t, $\mathbf{V}_t$, we compute near-optimal Lagrangian multipliers in an effort to generate tight Benders cuts. To obtain these multipliers, we use an iterative method based on the subgradient method (see Algorithm 2). It is important to note that Corollary 1 is only guaranteed to hold if we can obtain and use the optimal multipliers.

We begin by solving the LP relaxation ($\overline{B}_t$) of (B_t) and use the associated dual variable values as the first trial values for the Lagrangian multipliers, for stage t and given value $\mathbf{V}_t$. By doing so, we ensure that our upper bounds are always as good as the upper bounds obtained by solving the LP relaxation instead. Solving the Lagrangian MILP problem for these trial values may lead to a parallel shift of the Benders optimality cut. In *Steps (2)-(5)*, we iteratively compute different trial values for the Lagrangian multipliers in order to compute a tighter cut, with a different slope than the one given by the LP relaxation in *Step (1)*. In *Step (4)*, $\boldsymbol{\varepsilon}_{t\kappa}$ is the difference between the left and right-hand side value of the water balance equation, for iteration κ. As the stopping criterion, we propose to terminate Algorithm 2 after a fixed number of iterations, or when the change in each multiplier is "sufficiently" small – whichever is satisfied first.

Algorithm 2 Update for Lagrangian multipliers (for stage t and water reservoir levels $\mathbf{V}_t$)

Step (0). Initialize: $LBest = +\infty$, $\kappa = 1$, $\theta_1 = \frac{1}{2}$.
Step (1). Solve t-stage LP ($\overline{B}_t$) with the computed cuts approximating function $\beta_{t+1}(\cdot)$ and obtain duals $\overline{\boldsymbol{\lambda}}_t$ associated with (25); assign $\boldsymbol{\lambda}_{t\kappa} = \overline{\boldsymbol{\lambda}}_t$.
while Stopping criteria not satisfied **do**
  Step (2). Solve the MILP Lagrangian problem ($\hat{\mathscr{L}}_{\lambda t}$) for stage t using $\boldsymbol{\lambda}_{t\kappa}$.
  Step (3). Determine if $\hat{\mathscr{L}}_t(\boldsymbol{\lambda}_{t\kappa}, \mathbf{V}_t)$ yields a bound tighter than the bound obtained thus far:
    if $\hat{\mathscr{L}}_t(\boldsymbol{\lambda}_{t\kappa}, \mathbf{V}_t) < LBest$, **then** $LBest = \hat{\mathscr{L}}_t(\boldsymbol{\lambda}_{t\kappa}, \mathbf{V}_t)$ and $\boldsymbol{\lambda}_t^* = \boldsymbol{\lambda}_{t\kappa}$.
  Step (4). Update multipliers and step size:
    $\boldsymbol{\lambda}_{t,\kappa+1} = \max\{\mathbf{0}, \boldsymbol{\lambda}_{t\kappa} - \theta_\kappa \boldsymbol{\varepsilon}_{t\kappa}\}$ and $\theta_{\kappa+1} = \frac{1}{\kappa+2}$.
  Step (5). Increment the iteration count $\kappa \leftarrow \kappa + 1$.
end while
Step (6). **Exit** with $LBest$ and $\boldsymbol{\lambda}_t^*$.

5.2. Bounds

We now focus on comparing the lower and upper bounds obtained by solving either the LP relaxation, using a concave overestimator, or our Lagrangian relaxation approach. Generally speaking, lower bounds are computed using the best feasible solution to the original problem found via each respective approach and upper bounds are computed through solving each respective approaches relaxed formulation of the problem. In this context, of special interest is formulation ($\hat{B}_t$) which is (B_t) in which the FRF $\beta_{t+1}(\cdot)$ is approximated via Benders optimality cuts obtained from solving an associated relaxed problem. Whenever we refer to ($\hat{B}_t$), we explicitly state which relaxed problem is being used to generate the optimality cuts. Table 1 describes each of the resulting bounds.

From Corollary 1 and Theorem 2, we know that

$$\overline{z} \leq z(\overline{B}) = z(C).$$

Because alternate (primal) optima may exist for ($\overline{B}$) and (C), there is no relationship between $\check{\underline{z}}_{\overline{B}}$ and $\check{\underline{z}}_{C}$. Similarly, when using NBD, due to the possibility of alternate (dual) optima, the resulting Benders optimality cuts generated when solving $(\overline{B})$ and (C) may differ. Consequently, pairs $\underline{z}_{\overline{B}}$ and $\underline{z}_C$ as well as $\overline{z}_{\overline{B}}$ and $\overline{z}_C$ do not have a predefined relationship.

We know that any Benders cut (6) computed with Algorithm 1 is at least as tight as a cut computed using the LP relaxation and note that the same logic holds for the cuts computed using the concave overestimation approach (see Section 5.1). However, this in only true when using the same state variable value. For any other state variable value, we do not know which of the two cuts is tighter. In addition, both cuts (6) and the cuts obtained from an LP relaxation overestimate the true future value functions. Consequently, we cannot say which of the lower bounds $\underline{z}$ or $\underline{z}_{\overline{B}}$ is better. We expect that, in most cases, the tighter cuts (6) produce a better feasible solution and, hence, more restrictive lower bounds.

Approach	**Symbol**	**Bound**	**Method**	**Description**
LP	$\check{\underline{z}}_{\overline{B}}$	Lower	LP	Solve $(\overline{B})$; evaluate exact revenue
Relaxation $(\overline{B})$	$\underline{z}_{\overline{B}}$	Lower	NBD, MILPs	Solve $(\overline{B})$ using NBD to optimality; perform one forward pass to solve MILP $(\hat{B}_t)$ with computed cuts
	$\overline{z}_{\overline{B}}$	Upper	NBD, MILP	Solve $(\overline{B})$ using NBD to optimality; solve MILP $(\hat{B}_1)$ with computed cuts
	$z(\overline{B})$	Upper	LP	Optimal objective function value of $(\overline{B})$, see Section 4.2
Concave	$\check{\underline{z}}_C$	Lower	LP	Solve (C); evaluate exact revenue
Overestimator (C)	$\underline{z}_C$	Lower	NBD, MILPs	Solve (C) using NBD to optimality; perform one forward pass to solve MILP $(\hat{B}_t)$ with computed cuts
	$\overline{z}_C$	Upper	NBD, MILP	Solve $(\overline{B})$ using NBD to optimality; solve MILP $(\hat{B}_1)$ with computed cuts
	$z(C)$	Upper	LP	Optimal objective function value of (C), see Section 4.2
Lagrangian	$\underline{z}$	Lower	Algorithm 1	(Best) Step (2) of Algorithm 1
Relaxation	$\overline{z}$	Upper	Algorithm 1	(Final) Step (6) of Algorithm 1

Table 1. Bounds on (B) obtained from different solution techniques. NBD refers to nested Benders decomposition for LPs.

5.3. Discussion: When is this Approach Useful for the Bidding Problem?

At this point you may ask: If LP relaxation, in some cases, can provide an upper bound as good as Lagrangian relaxation for the bidding problem, then when does Lagrangian relaxation provide a tighter upper bound and a better solution, *i.e.,* when is this approach useful? When there is only one reservoir, the answer is never. Whereas, in most cases, with more than one reservoir, the Lagrangian will provide tighter upper bounds.

5.3.1. One Reservoir

If the hydro producer only operates one reservoir, then the optimal objective function value using the tightest concave overestimator for the immediate revenue function and the optimal objective function value of the Lagrangian relaxation are the same. It follows, from Theorem 2, that they also equal the optimal objective function value of the LP relaxation.

To continue this discussion, we assume that the FRF is given and analyze the problem for the first stage only. The general case, for T stages, is then derived by induction. In this framework, problem (B) simplifies to

$$(G) \qquad \max_{e,v,s} \; R(e) + \beta(v)$$

$$\text{s.t.} \quad v = V_0 + A - \frac{e}{\rho} - s \tag{32}$$

$$\underline{V} \leq v \leq \overline{V} \tag{33}$$

$$0 \leq \frac{e}{\rho} \leq \overline{U} \tag{34}$$

$$s \geq 0 \,. \tag{35}$$

Parameter V_0, is the initial volume of the reservoir and A is the water inflow during the stage. Let $(\breve{G})$ denote formulation (G), where the tightest concave overestimator approximates $R(e) : [0, \rho\overline{U}] \to \mathbb{R}$, and let $z(\breve{G})$ denote $(\breve{G})$'s optimal objective function value.

Dualizing the water balance equations (32) yields the Lagrangian relaxation

$$(\mathscr{L}_G) \qquad z(\mathscr{L}_G) := \min_{\lambda} \left[\max_{e,v,s} \; R(e) + \lambda \left(V_0 + A - \frac{e}{\rho} - v - s \right) + \beta(v) \right]$$
$$\text{s.t.} \;\; (33) - (35) \; .$$

We are now ready to state

Theorem 3. *With one reservoir, the optimal objective function value of* (G) *using the tightest concave overestimator for* $R(e) : [0, \rho\overline{U}] \to \mathbb{R}$*, equals the optimal objective function value of the Lagrangian relaxation formulation* $(\mathscr{L}_G)$*, i.e.,* $z(\breve{G}) = z(\mathscr{L}_G)$ *for all* V_0.

The proof of Theorem 3 is given in Appendix D. Tightness of the constructed concave overestimator for $R(e)$ is crucial. If we construct a tight concave overestimator for $R(e)$ on an interval $\mathcal{I} \supsetneq [0, \rho\overline{U}]$, then Lagrangian relaxation might yield a smaller (*i.e.*, better) objective function value. This holds true, even if the exact revenue function $R(e)$ is used on interval $\mathcal{I}$ instead of the interval $[0, \rho\overline{U}]$. In this case, the Lagrangian is able to automatically detect the precise bounds for the immediate revenue function, which potentially results in a better solution. Consequently, this may result in a tighter cut when used in conjunction with Benders decomposition. The importance of this observation is made clear in the next subsection and this is the very reason why the Lagrangian approach outperforms static, a priori approximation of $R(e)$, with multiple reservoirs.

Similarly, in the LP relaxation $(\overline{B})$, the representation of $R(e)$ needs to be exact on $[0, \rho\overline{U}]$, *e.g.*, there needs to be a breakpoint at $V_0 = 0$ and at $V_0 = \rho\overline{U}$. As was the case with one reservoir, the LP and Lagrangian relaxation objective function values are the same, and $\lambda^* = \overline{\lambda}$.

Theorem 3 can be generalized to include lower bounds $\underline{U}$ on the hydro plants' generation. However, if this is done, care has to be taken regarding the problem's feasibility. Also, other non-concave revenue functions can be used, but, in these cases, the LP relaxation and the concave overestimation approach may not yield the same results.

5.3.2. Multiple Reservoirs

If the hydro producer operates more than one reservoir, then our approach tends to yield better solutions and bounds for three main reasons.

First, our approach may yield tighter bounds and better solutions, because, rather than overestimating the immediate revenue function, we represent it exactly. By representing the immediate revenue functions exactly we create more accurate FRF approximations. This dynamic overestimation approximates or overestimates the resulting FRFs as opposed to approximating or overestimating the immediate revenue functions in each stage, *i.e.*, static overestimation. Since dynamic overestimation may yield tighter bounds, it also may generate tighter optimality cuts.

Second, the Lagrangian approach takes into account the cumulative output capacity of the hydro plants. By doing so, it may detect that a certain portion of the immediate revenue function can never be reached for specific inflow scenarios. This is consistent with the above discussion on the single reservoir case – the immediate revenue function contains this information implicitly in the upper bound $\rho\overline{U}$. In contrast to the cumulative output capacity of the hydro plants, the Lagrangian cannot detect cases in which a certain portion of the immediate revenue function can never be reached, resulting from bounds on the reservoir volumes. This is because the computed cuts are valid for any reservoir levels, *cf.* Theorem 1. These facts, together, allow for the computation of tighter cuts compared to the LP relaxation.

Third, for cases in which the producer can increase revenue in the current stage *and* simultaneously increase revenue in future stages, by focusing on the concavity of the cumulative revenue and not the concavity of the revenue in each stage, our approach creates a "smoothing" effect that tends to yield better bounds and solutions. This can occur, for example, if we remove the assumption that upstream water, used for production or spilled, is available in the same stage. Typically, practitioners (including us, in this paper) assume that when a hydro producer uses water to produce energy in the current stage, that water is immediately available to the downstream reservoirs. This makes it so that, in any particular stage, the hydro producer must choose to increase revenue by either producing now *or* by saving the water to be used in future stages. If, instead, the practitioner assumes that it takes some amount of time (say one stage) for the water being used for production in the current stage to flow to the downstream reservoirs, then the hydro producer's decision changes slightly. In this construct, if the hydro producer produces now it will increase revenue in the current stage *and* increase revenue in the future stage as well, since production in the current stage makes more water available in the next stage. Through adding delays in the resulting inflows from upstream activity, there is a good chance that our approach will yields tighter upper bounds and better solutions than LP relaxation due to the smoothing effect. Figure 2 helps to explain this smoothing effect. In this simple example, by making the cumulative function concave instead of making the function concave in each stage and summing the results, we are able to approximate the function without error.

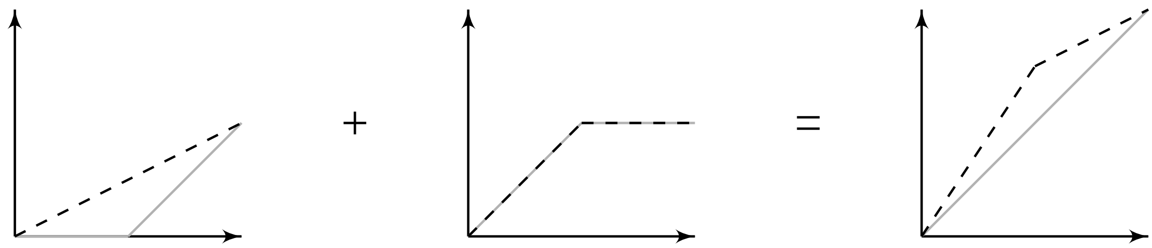

Figure 2. Non-concave function smoothing. Solid lines depict the exact function while dashed lines depict the function's concave overestimator.

6. Computational Results

Our computational tests are performed on a standard desktop computer with an Intel(R) Core(TM) i3 CPU @ 2.10 GHz processor, with 4GB of RAM, running on Windows 7. We implemented the monolith formulation and the NBD algorithms (LP relaxation, concave overestimation, and Lagrangian relaxation) in GAMS 24.1.3 and solved the resulting LPs and MILPs using CPLEX version 12.5.1.0. We terminate the NBD algorithms after a fixed number of iterations.

6.1. Illustrative Example

A small example illustrates some of the finer points of our methodology. In the example, we compare the optimal solution with the solution we obtain from our approach and the solution found from solving the LP relaxation or via using a concave overestimator. Assume there is one price-maker hydro producer that controls two independent reservoirs, and that the time horizon of interest is only two stages long. Additionally, assume there is one price-taker thermal producer in the first stage, and two price-taker thermal producers in the second (see Table 2). These parameters result in the immediate revenue functions depicted in Figure 3.

Thermal Parameters		$\overline{D}$	j	$\overline{G}_j$	C_j	g_j	
	Period 1	400	1	400	95	400	
	Period 2	400	1	300	150	300	
			2	110	300	100	
Hydro Parameters		V_{0i}	$\underline{V}_i$	$\overline{V}_i$	ρ_i	$\overline{U}_i$	A_{it}
Reservoir 1	Period 1	0	0	500	1	175	225
	Period 2						25
Reservoir 2	Period 1	0	0	500	1	225	0
	Period 2						0

Table 2. Example parameters

Solving the problem via the monolith, via the LP relaxation, via the use of a concave overestimator, or via NBD with Lagrangian relaxation results in the variable and function values given in Table 3. Observe that solving the LP relaxation or using a concave overestimator, in this case, results in the same solution and bounds (see Sections 4.3 and 5.2). Solving the problem using NBD coupled with Lagrangian relaxation results in a tighter upper bound and

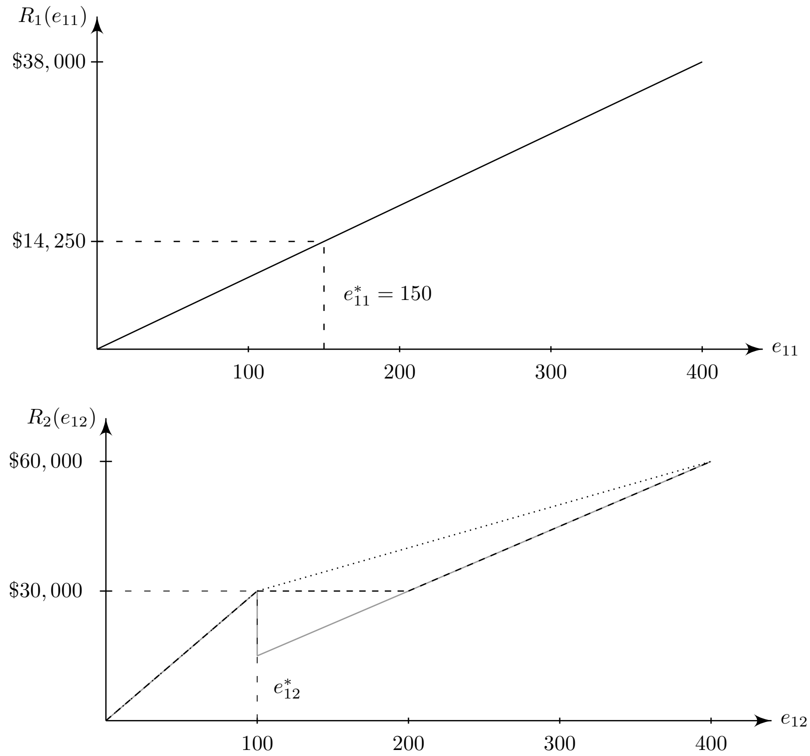

Figure 3. Example immediate revenue function (solid line), representation (dashed line) of the immediate revenue function, and concave overestimator (dotted line). Stage one is on the top and stage two is on the bottom. Since the immediate revenue function is linear for stage one, the function, its representation, and the concave overestimator are all the same.

Monolith		**LP**		**Overestimator**		**Lagrangian**	
Variable	**Level**	**Variable**	**Level**	**Variable**	**Level**	**Variable**	**Level**
e_{11}^*	150	e_{11}	75	e_{11}	75	e_{11}	150
e_{12}^*	100	e_{12}	175	e_{12}	175	e_{12}	100
R_1^*	$14,250	R_1	$7,125	R_1	$7,125	R_1	$14,250
R_2^*	$30,000	R_2	$37,500	R_2	$37,500	R_2	$30,000
v_{12}^*	75	v_{12}	150	v_{12}	150	v_{12}	75
v_{13}^*	0	v_{13}	0	v_{13}	0	v_{13}	0
β_1^*	$44,250	$\check{\underline{z}}_{\overline{B}}$	$37,125	$\check{\underline{z}}_C$	$37,125		-
	-	$\underline{z}_{\overline{B}}$	$37,125	$\underline{z}_C$	$37,125	$\underline{z}$	$44,250
	-	$\overline{z}_{\overline{B}}$	$44,625	$\overline{z}_C$	$44,625	$\overline{z}$	$44,250
	-	$z(\overline{B})$	$44,625	$z(C)$	$44,625		-
	-	Gap	20.20 %	Gap	20.20 %	Gap	0 %

Table 3. Variable and function values found when solving the MILP (left) vs. the LP relaxation (left middle) vs. solving using a concave overestimator (right middle) vs. Benders decomposition and Lagrangian relaxation (right). Since $e_{2t}^* = 0, v_{2t}^* = 0 \quad \forall\ t$, we do not report these values.

the Lagrangian is exact, *i.e.*, solves the bidding problem (B) to proven optimality. This is neither the case for the LP relaxation nor the concave overestimator.

The optimality cuts generated by the LP relaxation and the Lagrangian relaxation are given in Figure 4. We construct this example so that reservoir two has an initial water volume of zero and no inflow so that we can plot the Benders optimality cuts in two dimensions. Consequently, Figure 4 is a projection onto reservoir one's space. From this plot, we see that the cuts generated by the Lagrangian are always at least as tight as the cuts generated by the LP relaxation (or through using a concave overestimator). The cuts generated by the LP relaxation and the concave overestimator are identical. The horizontal cut generated by the Lagrangian is exact, and this is why we have a zero optimality gap. The cut is exact because, by representing the immediate revenue function exactly, we create an exact approximation of the FRF and consequently generate exact Benders optimality cuts. Through solving the LP relaxation (or the concave overestimator formulation), we overestimate the immediate and the future revenue. In Figure 3, we can see how the overestimation results in selecting inefficient production quantities ($e_{11} = 75$ and $e_{12} = 175$ vs. $e^*_{11} = 150$ and $e^*_{12} = 100$). The overestimation also leads to Benders optimality cuts that are not as tight as those found via Lagrangian relaxation.

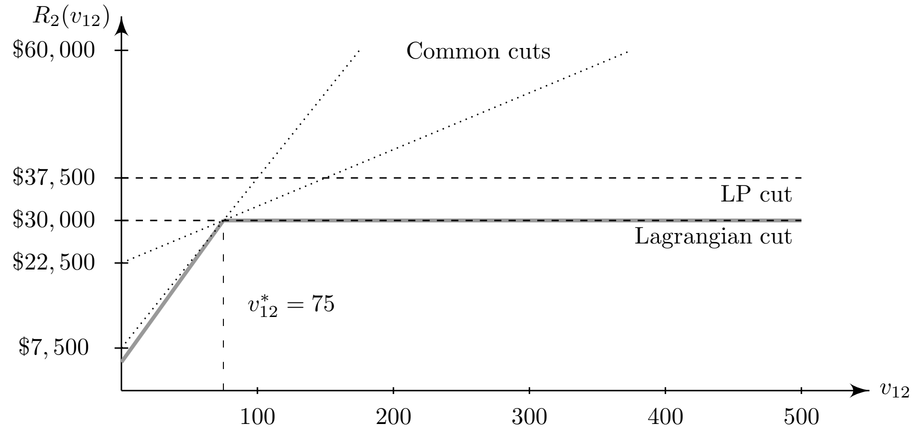

Figure 4. LP relaxation and Lagrangian relaxation optimality cuts. The FRF is denoted by the solid gray line. "Common cuts" are the cuts that are generated from both formulations. The "LP cut" is the unique cut generated from the LP relaxation solve while the "Lagrangian cut" is the unique cut generated from Lagrangian relaxation solve.

6.2. Case Studies

To demonstrate the value of our approach we model the electricity markets for three Central American countries: El Salvador, Honduras, and Nicaragua. In each solve, we compare the solution found from our approach and the solution found from solving the LP relaxation with the optimal solution (*i.e.,* the solution found from solving the monolith). Since we have already demonstrated the relationship between the LP relaxation and the concave overestimator, we now center our focus on the differences between the LP relaxation and the Lagrangian.

We assume a time horizon of two years and solve the problem in monthly steps for "low," "medium-low," "medium," "medium-high," and "high" inflow scenarios. As such, each stage represents a one-month period. Additionally, in each market, we assume that there is only one price-maker hydro producer that owns and operates a series of reservoirs and run-of-the-river plants. Run-of-the-river plants are hydro plants with little or no storage. They are modeled as reservoirs with no storage, *i.e.,* $\underline{V} = \overline{V} = 0$. The thermal and hydro parameters used to model each of these countries electricity markets are given in Appendix E. Initial reservoir levels are determined based on historical data. Last, if the price-taker thermal producers alone cannot satisfy demand, we assume the price-maker hydro producer must at least produce enough energy, in each period, to meet the residual demand. In this case and in our model, the price-maker hydro producer is penalized for not meeting the residual demand and only earns revenue for production levels above the residual demand. Of the three countries that we model, this only occurs for Honduras.

The resulting bounds found from solving the LP relaxation, the Lagrangian relaxation, and the monolith (optimal solution) for each of the three countries and the five different inflow scenarios are given in Table 4. Table 5 compares the optimality gaps and differences between bounds for each of these solves individually and for each country. The optimality gaps and the differences between the bounds for the LP relaxation solves are computed as follows:

$$\text{Gap} = \frac{\overline{z}_{\overline{B}} - \underline{z}_{\overline{B}}}{\underline{z}_{\overline{B}}}, \qquad \tilde{\text{Gap}} = \frac{z(\overline{B}) - \check{\underline{z}}_{\overline{B}}}{\check{\underline{z}}_{\overline{B}}},$$

$$\Delta\text{LB} = \frac{\text{Optimal} - \underline{z}_{\overline{B}}}{\underline{z}_{\overline{B}}}, \text{ and} \qquad \Delta\text{UB} = \frac{\overline{z}_{\overline{B}} - \text{Optimal}}{\text{Optimal}}.$$

Similarly, we compute the optimality gaps and the differences between the bounds for the Lagrangian relaxation solves as

$$\text{Gap} = \frac{\overline{z} - \underline{z}}{\underline{z}},$$

$$\Delta\text{LB} = \frac{\text{Optimal} - \underline{z}}{\underline{z}}, \text{ and} \qquad \Delta\text{UB} = \frac{\overline{z} - \text{Optimal}}{\text{Optimal}}.$$

From the tables we observe that the Lagrangian function produces significantly smaller optimality gaps. Also, in all but one case, the lower bound found from solving the Lagrangian is much closer to the actual solution than the lower bound found from solving the LP relaxation. This is an important distinction because the lower bounds represent the strategy that would be implemented. The smaller gaps and better solutions however come at the cost of longer computation times. Note that in line 4 of Table 4, the upper bound from the Lagrangian is smaller than the solution found from solving the monolith. Also note that the Lagrangian lower and upper bounds match. This may occur when we use near-optimal Lagrangian multipliers. Computation times from solving the monolith (B), directly, are not reported because they are approximately the same as the times reported for solving the LP relaxation. We expect the differences between the gaps to grow as we lengthen the time horizon over which we study the problem and as we model larger systems.

		LP relaxation (Overestimator)				**Lagrangian**		**Monolith**
		$\check{\underline{z}}_{\overline{B}}$	$\underline{z}_{\overline{B}}$	$\overline{z}_{\overline{B}}$	$z(\overline{B})$	$\underline{z}$	$\overline{z}$	$z(B)$
El Salvador	Low inflow	\$921.50	\$990.87	\$1,024.82	\$1,024.82	\$998.06	\$1,020.07	\$1,001.33
	Medium-low inflow	\$922.65	\$995.45	\$1,035.21	\$1,035.21	\$1,000.72	\$1,034.15	\$1,005.12
	Medium inflow	\$914.15	\$996.42	\$1,034.88	\$1,034.88	\$1,001.23	\$1,030.79	\$1,007.29
	Medium-high inflow	\$955.44	\$1,035.57	\$1,063.09	\$1,063.09	\$1,038.32	\$1,063.09	\$1,040.62
	High inflow	\$976.58	\$1,069.72	\$1,098.74	\$1,098.74	\$1,068.36	\$1,098.74	\$1,074.21
Honduras	Low inflow	\$265.25	\$299.59	\$340.66	\$341.36	\$304.57	\$330.43	\$312.03
	Medium-low inflow	\$317.06	\$339.31	\$371.89	\$372.90	\$346.26	\$367.45	\$347.09
	Medium inflow	\$326.78	\$343.86	\$378.23	\$379.15	\$355.25	\$372.65	\$355.39
	Medium-high inflow	\$379.03	\$393.72	\$413.75	\$414.64	\$394.08	\$411.75	\$394.90
	High inflow	\$384.90	\$395.99	\$416.11	\$417.21	\$396.50	\$415.41	\$396.77
Nicaragua	Low inflow	\$188.10	\$188.90	\$191.46	\$191.46	\$189.21	\$190.47	\$189.25
	Medium-low inflow	\$193.08	\$193.56	\$197.05	\$197.05	\$194.91	\$196.44	\$194.93
	Medium inflow	\$194.64	\$196.26	\$198.73	\$198.73	\$196.62	\$198.08	\$197.10
	Medium-high inflow	\$199.10	\$200.45	\$203.43	\$203.43	\$201.01	\$202.77	\$201.81
	High inflow	\$225.22	\$226.64	\$229.19	\$229.19	\$227.55	\$229.19	\$228.08

Table 4. Lower and upper bounds found from solving the LP relaxation, the Lagrangian relaxation and the monolith [\$M].

Figure 5 shows the evolution of the lower and upper bounds for the Lagrangian approach, for the computations provided in Tables 4 and 5. Observe that the largest improvements are achieved in the first four iterations, while almost no improvement is made after the fifth iteration. This relatively fast convergence is promising.

7. Conclusion

We present a methodology that combines Lagrangian relaxation and nested Benders decomposition to solve multi-stage non-convex problems. Through utilizing Lagrangian relaxation in this manner we propose a way to overcome the convexity limitations associated with Benders decomposition algorithms. We apply our methodology to the problem in which a single price-maker hydro producer seeks to maximize revenue in the day-ahead electricity market.

Our dynamic convexification method yields efficient solutions to the bidding problem for a single price-maker hydro producer, over the medium term. In many cases, our methodology yields tighter optimality cuts and consequently better upper bounds and better solutions than existing methods.

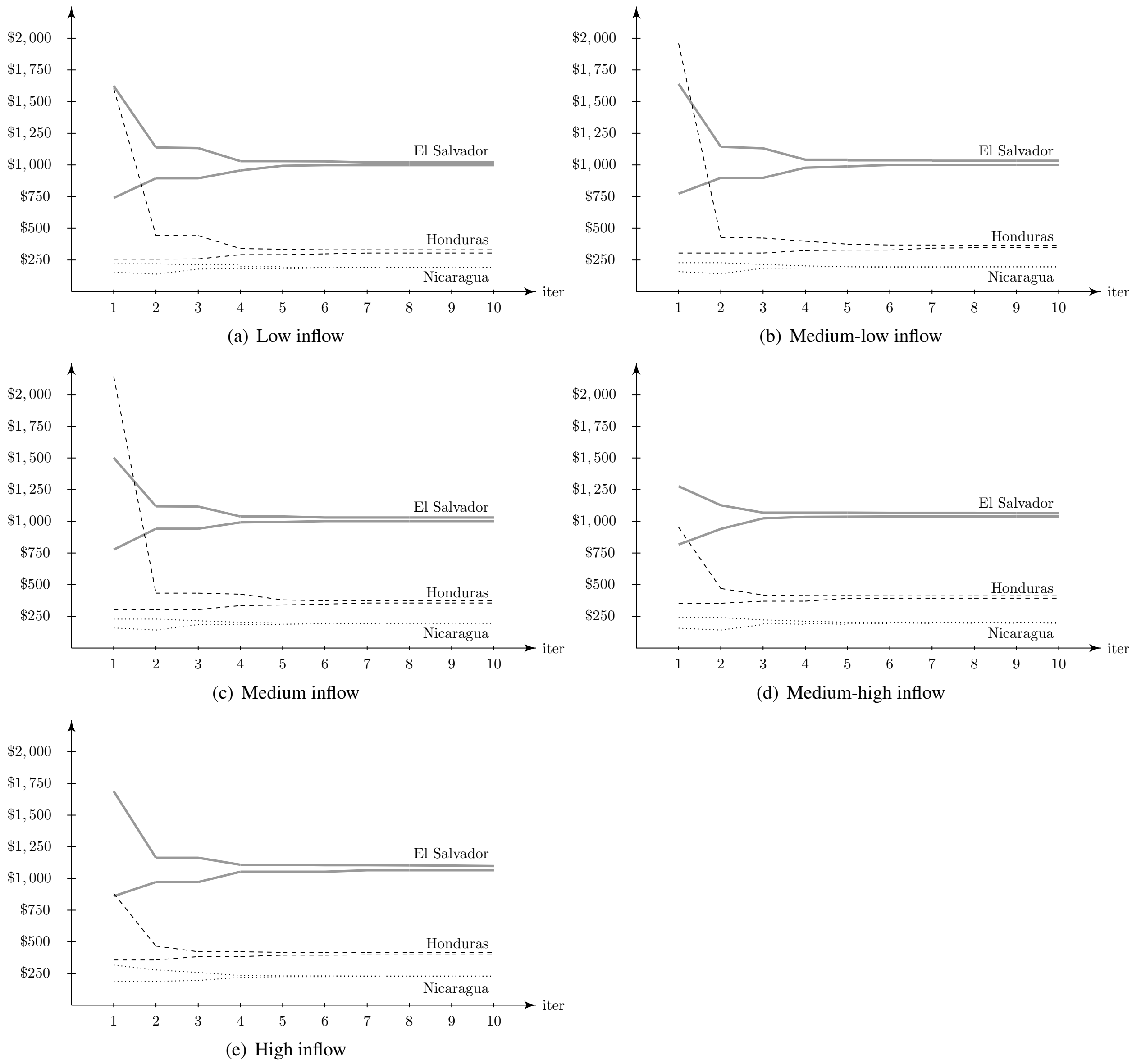

Figure 5. Evolution of the lower and upper bounds found from solving the Lagrangian relaxation for the five different inflow scenarios and the three countries of interest.

To emphasize the fundamentals of our approach, and to have an optimal solution as a benchmark with which to compare our results, we chose to implement deterministic inflows. Future work will incorporate uncertainty in reservoir inflows. Ultimately, this approach is to be paired with SDDP so that large hydro systems can be modeled and solved for price-maker producers. Future efforts in this area could (i) study the problem and adapt our approach for cases with multiple price makers, (ii) extend the bidding problem to incorporate other non-convex model characteristics, *e.g.*, power flow constraints (Frank and Rebennack 2015), (iii) seek tailored methods for updating the Lagrangian multipliers, thus reducing computation times, and/or (iv) apply our method to different market structures (Aïd et al. 2011).

		LP relaxation (Overestimator)					Lagrangian			
		Gap	Gãp	Δ LB	Δ UB	Time	Gap	Δ LB	Δ UB	Time
El Salvador	Low inflow	3.99%	11.21%	0.97%	2.99%	00:08	3.34%	0.44%	2.89%	12:51
	Medium-low inflow	3.43%	12.20%	1.06%	2.35%	00:08	2.21%	0.33%	1.87%	13:13
	Medium inflow	3.86%	13.21%	1.09%	2.74%	00:08	2.95%	0.61%	2.33%	12:52
	Medium-high inflow	2.66%	11.27%	0.49%	2.16%	00:07	2.39%	0.22%	2.16%	12:40
	High inflow	2.71%	12.51%	0.42%	2.28%	00:07	2.84%	0.55%	2.28%	11:37
	Average	3.33%	12.08%	0.81%	2.50%	00:07	2.75%	0.43%	2.31%	12:39
Honduras	Low inflow	13.71%	28.70%	4.15%	9.18%	00:06	8.49%	2.45%	5.90%	22:00
	Medium-low inflow	9.60%	17.61%	2.29%	7.14%	00:05	6.12%	0.24%	5.87%	22:50
	Medium inflow	10.00%	16.02%	3.35%	6.43%	00:10	4.90%	0.04%	4.86%	22:54
	Medium-high inflow	5.09%	9.40%	0.30%	4.78%	00:08	4.49%	0.21%	4.27%	22:12
	High inflow	5.08%	8.39%	0.20%	4.87%	00:09	4.77%	0.07%	4.70%	21:48
	Average	8.70%	16.02%	2.06%	6.48%	00:08	5.75%	0.60%	5.12%	22:21
Nicaragua	Low inflow	1.35%	1.78%	0.18%	1.17%	00:04	0.66%	0.02%	0.64%	08:02
	Medium-low inflow	1.80%	2.06%	0.71%	1.09%	00:07	0.78%	0.01%	0.77%	08:32
	Medium inflow	1.26%	2.10%	0.43%	0.83%	00:07	0.59%	0.09%	0.50%	07:56
	Medium-high inflow	1.49%	2.17%	0.68%	0.80%	00:07	0.88%	0.40%	0.48%	05:58
	High inflow	1.12%	1.76%	0.64%	0.49%	00:02	0.72%	0.24%	0.49%	04:27
	Average	1.40%	1.97%	0.53%	0.88%	00:05	0.73%	0.15%	0.58%	06:59

Table 5. Optimality gaps and computation times

Acknowledgements

The authors would like to thank two anonymous reviewers for their valuable comments and suggestions.

Appendix A. Proof of Corollary 1

Corollary 1. *Let* (M) *be a MILP. Solving the Lagrangian relaxation* $(\mathscr{L})$ *yields an upper bound for* (M) *which is at least as tight as the upper bound obtained from solving its LP relaxation* $(\overline{M})$*.*

Proof. We can write (M) as

$$(M) \qquad \mathcal{M}(\mathbf{X}_1) := \max_{\mathbf{x}_{t+1},\mathbf{y}_t \geq 0} \sum_{t=1}^{T} \mathbf{f}_t(\mathbf{x}_{t+1}, \mathbf{y}_t)$$

$$\text{s.t. } \mathbf{h}_t(\mathbf{x}_t) - \mathbf{g}_{1t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0} \qquad \forall\, t \tag{A.1}$$

$$\mathbf{g}_{2t}(\mathbf{x}_{t+1}, \mathbf{y}_t) \leq \mathbf{0} \qquad \forall\, t \tag{A.2}$$

$$\mathbf{x}_{t+1} \in \mathbb{R}^{n_{t+1}^{1,\text{cont}}} \times \mathbb{Z}^{n_{t+1}^{1,\text{int}}} \qquad \forall\, t \tag{A.3}$$

$$\mathbf{y}_t \in \mathbb{R}^{n_t^{2,\text{cont}}} \times \mathbb{Z}^{n_t^{2,\text{int}}} \qquad \forall\, t, \tag{A.4}$$

with *linear* functions $\mathbf{f}_t$, $\mathbf{h}_t$, $\mathbf{g}_{1t}$, and $\mathbf{g}_{2t}$ as well as $n_{t+1}^1 = n_{t+1}^{1,\text{cont}} + n_{t+1}^{1,\text{int}}$ and $n_t^2 = n_t^{2,\text{cont}} + n_t^{2,\text{int}}$.

The LP relaxation $(\overline{M})$ is obtained by using a continuous domain relaxation in (M), *i.e.*, (A.3) and (A.4) get replaced by

$$\mathbf{x}_{t+1} \in \mathbb{R}^{n_{t+1}^{1,\text{cont}}+n_{t+1}^{1,\text{int}}} \qquad \text{and} \qquad \mathbf{y}_t \in \mathbb{R}^{n_t^{2,\text{cont}}+n_t^{2,\text{int}}} \qquad \forall\, t.$$

Let the optimal objective function value of $(\overline{M})$ be denoted by $\overline{\mathcal{M}}(\mathbf{X}_1)$.

Next, we obtain the Lagrangian relaxation of $(\overline{M})$ by dualizing (A.1). We denote this Lagrangian by $(\overline{\mathscr{L}})$. Let its optimal objective function be $\overline{\mathscr{L}}_1(\mathbf{X}_1)$. Since $(\overline{M})$ is an LP, strong duality holds, *i.e.*,

$$\overline{\mathscr{L}}_1(\mathbf{X}_1) = \overline{\mathcal{M}}(\mathbf{X}_1).$$

By adding the integrality restriction (A.3) and (A.4) to $(\overline{\mathscr{L}})$, we obtain the Lagrangian relaxation $(\mathscr{L})$. This yields

$$\mathscr{L}_1(\mathbf{X}_1) \leq \overline{\mathcal{M}}(\mathbf{X}_1).$$

□

Appendix B. Proof of Theorem 1

Theorem 1. *The function or "cut"*

$$-\lambda'_{t+1}\mathbf{h}_{t+1}(\chi_{t+1}) + \gamma_{t+1}^{\text{const}}$$

overestimates $\eta_{t+1}(\chi_{t+1})$*, for any* $t = 1, \dots, T$.

Proof. We start by examining the last stage T with initial value $\mathbf{X}_T^o$. Denote the optimal decision variable values to $\mathscr{L}_T(\lambda_T, \mathbf{X}_T^o)$ by $\mathbf{x}_{T+1}^*$ and $\mathbf{y}_T^*$. Since $\mathscr{L}_T(\lambda_T, \cdot)$ is a relaxation of the function $\eta_T(\cdot)$, we know that (for a maximization problem)

$$\eta_T(\chi_T) \leq \mathscr{L}_T(\lambda_T, \chi_T) \quad \forall \ \lambda_T \geq \mathbf{0}, \quad \text{and} \ \forall \ \chi_T \ .$$

By adding "0" to the Lagrangian, we obtain

$$\begin{aligned}
\mathscr{L}_T(\lambda_T, \chi_T) &= \mathbf{f}_T(\mathbf{x}_{T+1}^*, \mathbf{y}_T^*) - \lambda'_T\Big(\mathbf{h}_T(\chi_T) - \mathbf{g}_{1T}(\mathbf{x}_{T+1}^*, \mathbf{y}_T^*)\Big) + \lambda'_T\mathbf{h}_T(\mathbf{X}_T^o) - \lambda'_T\mathbf{h}_T(\mathbf{X}_T^o) \qquad \text{(B.1)}\\
&= \mathscr{L}_T(\lambda_T, \mathbf{X}_T^o) + \lambda'_T\mathbf{h}_T(\mathbf{X}_T^o) - \lambda'_T\mathbf{h}_T(\chi_T)\\
&= -\lambda'_T\mathbf{h}_T(\chi_T) + \gamma_T^{\text{const}} \ .
\end{aligned}$$

Note that in (B.1), changing the initial parameter from $\mathbf{X}_T^o$ to χ_T does not change the optimal decision variable values $\mathbf{x}_{T+1}^*$ and $\mathbf{y}_T^*$ because the $\lambda'_T\mathbf{h}_T(\chi_T)$ term is a constant in the objective function. Thus, we conclude that

$$\eta_T(\chi_T) \leq -\lambda'_T\mathbf{h}_T(\chi_T) + \gamma_T^{\text{const}} \quad \forall \ \chi_T \ ,$$

and all cuts we generate for the last stage T are valid.

It follows that the generated cuts are valid for any stage t. This is made evident through backwards induction and by observing that the computed cuts overestimate function $\eta_{t+1}(\cdot)$. □

Appendix C. Proof of Theorem 2

Theorem 2. *The optimal objective function values of the LP relaxation formulation* $(\overline{B})$ *and the concave overestimator formulation* (C) *are the same, i.e.,* $z(\overline{B}) = z(C)$*.*

Proof. We begin this proof by making the following observations about the set of breakpoints and revenue function values in each of the two formulations:

$$\begin{aligned}
&\{\tilde{E}_{nt}\} \subset \{\dot{E}_{kt}\} \text{ and } \tilde{E}_{0t} = \dot{E}_{0t} = 0, \tilde{E}_{Nt} = \dot{E}_{Kt} \qquad \forall \ t;\\
&\{\tilde{R}_{nt}\} \subset \{\dot{R}_{kt}\} \text{ and } \tilde{R}_{0t} = \dot{R}_{0t} = 0, \tilde{R}_{Nt} = \dot{R}_{Kt} \qquad \forall \ t \ .
\end{aligned}$$

Given any solution to $(\overline{B})$, we can construct a solution to (C) such that $z(C) = z(\overline{B})$ and vice versa. To see why this is the case suppose, without loss of generality, $T = 1$ and we have an optimal solution to $(\overline{B})$, $z(\overline{B})$, where e_1^* lies somewhere between breakpoints p and r (in which the breakpoints are not necessarily subsequent). Since we are maximizing, the optimal e_1^* could not lie between breakpoints that are not in $\{\tilde{E}_{nt}\}$; thus, breakpoints p and r are guaranteed to be in the set of breakpoints in (C). Therefore, to construct our solution in (C) where $z(C) = z(\overline{B})$ we set $\alpha_{p1} = \mu_{p1}$ and $\alpha_{r1} = \mu_{r1}$.

In other words, the convex combination of breakpoints that make up our value for e_t is the same regardless of whether we are solving using $(\overline{B})$ or (C) and the immediate revenue function representation, $R_t(e_t)$, is also the same. □

Appendix D. Proof of Theorem 3

Theorem 3. *With one reservoir, the optimal objective function value of* (G) *using the tightest concave overestimator for* $R(e) : [0, \rho\overline{U}] \rightarrow \mathbb{R}$*, equals the optimal objective function value of the Lagrangian relaxation formulation* $(\mathscr{L}_G)$*, i.e.,* $z(\breve{G}) = z(\mathscr{L}_G)$ *for all* V_0*.*

Proof. The theorem holds, if and only if there exits no V_0 such that the objective function of the Lagrangian is strictly less than the objective of the tight concave overestimated problem, *i.e.,* $\nexists\ V_0$ such that $z(\mathscr{L}_G) < z(\breve{G})$.

We prove this theorem by analyzing the following three cases:

(i.) The optimal solution of the Lagrangian only involves $R(e)$, *i.e.*, $v = \underline{V}$. In words, the hydro producer does not save any water for future stages.
From Theorem 2 we know that

$$\{\tilde{E}_{nT}\} \subset \{\dot{E}_{kT}\} \ .$$

We also know that both objective functions in $(\mathscr{L}_G)$ and $(\breve{G})$ are concave. Now, because $(\breve{G})$ is a convex optimization problem (maximizing a concave function over a polyhedron), strong duality holds for $(\breve{G})$, (Boyd and Vandenberghe 2004). Consequently, the associated Lagrangian dual problem yields the same optimal objective function value as $(\breve{G})$. Because $(\mathscr{L}_G)$ only dualizes constrains (32) and $R(e)$ is modeled exactly (*i.e.*, not overestimated), we have that

$$z(\mathscr{L}_G) \leq z(\breve{G}).$$

But it is not possible for $z(\mathscr{L}_G) < z(\breve{G})$ if $\{\tilde{E}_{nT}\} \subset \{\dot{E}_{kT}\}$; thus,

$$z(\mathscr{L}_G) = z(\breve{G}) \ .$$

In other words, the concave overestimator is, by construction, the tightest concave approximation of the immediate revenue function. So the Lagrangian relaxation cannot be better but, by strong duality, we know it cannot be worse. Therefore, it must be the same.

(ii.) The optimal solution of the Lagrangian only involves $\beta(v)$, *i.e.*, $\frac{e}{\rho} = 0$. Specifically stated, the hydro producer saves as much water as possible for future stages.
Assume there exists V_0 such that $z(\mathscr{L}_G) < z(\breve{G})$. This contradicts the fact that the tightest concave overestimator is used to approximate $R(e)$ in $(\breve{G})$ because the same $\beta(\cdot)$ function is used in both formulations.

(iii.) The optimal solution of the Lagrangian involves both $R(e)$ and $\beta(v)$. In words, the hydro producer produces something in the current stage and saves some water for future stages.
Again, assume there exists V_0 such that $z(\mathscr{L}_G) < z(G)$. Let the optimal objective function value of the Lagrangian be $z(\mathscr{L}_G)$ and its solution be e^*, v^* and ε^*. Now, fix the decision variable $v = \underline{v}$ by increasing the spillage. Increasing the spillage has no impact on the optimal values for e^* and ε^*. The Lagrangian then yields an objective function value of $z(\mathscr{L}_G) - \beta(v^*) + \beta(\underline{V})$. This defines an optimal solution for $\tilde{V}_0 = V_0 - v^* - \underline{V}$. Because both the Lagrangian and the LP relaxation have the same function $\beta(\cdot)$ and the Lagrangian yields a valid cut, the cut computed by the Lagrangian is a strictly better, concave (because it is linear) overestimator for $R(e)$ for $\tilde{V}_0$. Again, this contradicts the fact that the tightest concave overestimator is used to approximate $R(e)$ in $(\breve{G})$.

□

Appendix E. Data: Thermal and Hydro Parameters

The thermal and hydro parameters used to model each of these countries' electricity markets are given in Tables E.6-E.8 and Figures E.6-E.8. In the figures, reservoirs are depicted as triangles, while run-of-the-river plants are depicted as squares only. Demand and reservoir inflow parameters used to model and solve each of our case studies can be found in the online appendix.

j	**Name**	$\overline{G}_j$	C_j
1	ACAJ-M	104.90	151,179.51
2	ACAJ-U	114.68	322,220.99
3	SOYA	10.76	175,482.49
4	TALN	72.96	142,497.25
5	NEPO	106.87	168,020.36
6	TEXT	30.88	154,235.77
7	BERL	80.35	2,728.25
8	AHUA	72.17	3,256.33
9	CESSA	24.11	171,616.60
10	GCSA	8.33	190,096.44
11	BORE	9.46	200,884.75
12	HILC	4.76	180,208.72
13	CASSA	33.48	419.00
14	LANG	30.13	414.00
15	LCAB	7.44	832.00
16	CHAP	7.44	574.00
	Total	718.72	
	Min	4.76	414.00
	Max	114.68	322,220.99

Table E.6. El Salvador's thermal parameters. Capacities shown are for a 30-day month.

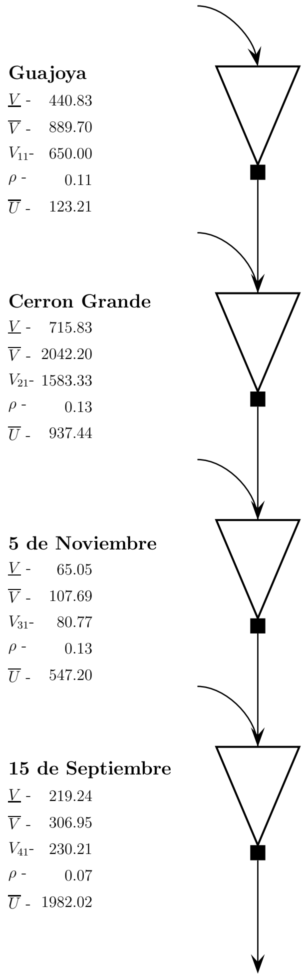

Figure E.6. El Salvador's hydro profile and parameters. Turbining capacities shown are for a 30-day month.

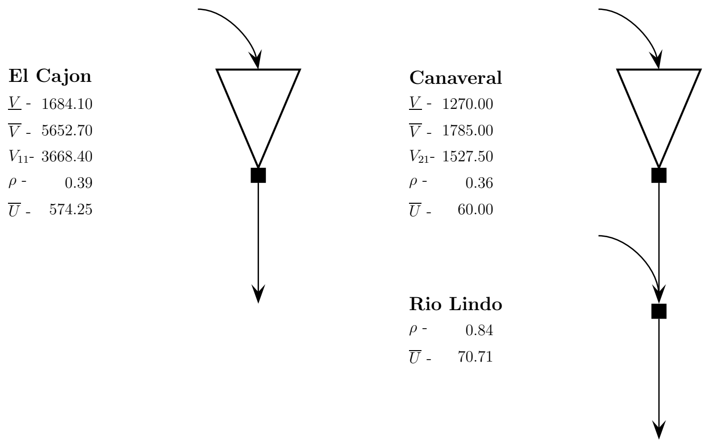

Figure E.7. Honduras' hydro profile and parameters. Run-of-the-river plants not included in the profile are not connected to reservoirs. Water turbine outflow capacities shown are based on a 30-day month.

i	Name	ρ_i	$\overline{U}_i$
4	Nispero	0.37	45.26
5	Nacaome	0.06	189.10
6	Esperanza	1.27	7.50
7	Cuyamapa	0.83	9.37
8	Cuyamel	0.31	18.75
9	Peq-hidro	0.52	29.30

j	Name	$\overline{G}_j$	C_j
1	Lufussa1	29.39	162,954.50
2	Lufussa2	57.29	82,375.00
3	Emce2	40.92	80,125.00
4	Enersa	159.96	81,225.00
5	Ceiba	17.86	91,725.00
6	Lufu3-210	156.24	75,225.00
7	Elcosa	59.52	85,275.00
8	Puert ENE	11.90	253,536.80
9	Puert MEX	7.44	299,155.00
	Total	540.52	
	Min	7.44	75,225.00
	Max	159.96	299,155.00

Table E.7. Honduras' run-of-the-river hydro parameters and thermal parameters. Since water turbine outflow and thermal producer capacities change based on the number of days in the month, we only show these parameters for a month with 30 days.

j	Name	$\overline{G}_j$	C_j
1	PNI-U1	37.20	100,500.00
2	PNI-U2	37.20	103,650.00
3	PMANAG-U3	31.99	117,825.00
4	PMANAG-U4	3.72	87,375.00
5	PMANAG-U5	3.72	86,850.00
6	PBRISA-U1	17.86	229,767.00
7	PBRISA-U2	28.27	179,695.20
8	PMOMOTOM	23.06	56,860.00
9	PAMFELS	42.41	83,850.00
10	PCENT1	37.20	79,287.50
11	PCENT2	13.76	84,225.00
12	PTIPITAPA	37.87	76,700.00
13	MTR	26.04	27,000.00
14	HChavez	44.64	162,109.20
15	CG Canal	4.46	85,625.00
16	Albanisa	44.64	86,400.00
	Total	434.05	
	Min	3.72	27,000.00
	Max	44.64	229,767.20

Table E.8. Nicaragua's thermal parameters. Capacities shown are for a 30-day month.

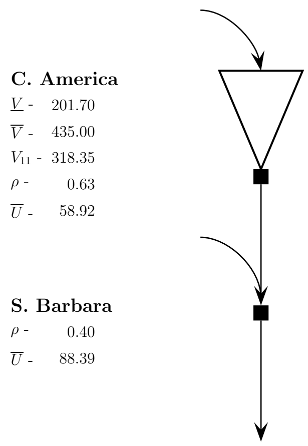

Figure E.8. Nicaragua's hydro profile and parameters. Turbining capacities shown are for a 30-day month.

References

R. Aïd, G. Chemla, A. Porchet, N. Touzi, Hedging and vertical integration in electricity markets, Management Science 57 (2011) 1438–1452.

A. Banal-Estañol, A.R. Micola, Composition of electricity generation portfolios, pivotal dynamics, and market prices, Management Science 55 (2009) 1813–1831.

E. Beale, Two transportation problems, in: Proceedings of the Third International Conference on Operational Research 1963, Dunod, Paris and English Universities Press, London, 1963, pp. 780–788.

A. Belloni, A.D.S. Lima, M.P. Maceira, C.A. Sagastizábal, Bundle relaxation and primal recovery in unit commitment problems. the brazilian case, Annals of Operations Research 120 (2003) 21–44.

J.F. Benders, Partitioning procedures for solving mixed-variables programming problems, Numerische Mathematik 4 (1962) 238–252.

B. Bezerra, R. Kelman, L. Barroso, B. Flach, M. Latorre, N. Campodonico, M. Pereira, Integrated electricity-gas operations planning in hydrothermal systems, in: Symposium Proceedings for Specialists in Electric Operational and Expansion Planning (SEPOPE), Brazil.

J.R. Birge, Solution methods for stochastic dynamic linear programs., Ph.D. thesis, Stanford University, 1980.

A. Borghetti, A. Frangioni, F. Lacalandra, C.A. Nucci, Lagrangian heuristics based on disaggregated bundle methods for hydrothermal unit commitment, IEEE Transactions on Power Systems 18 (2003) 313–323.

S.P. Boyd, L. Vandenberghe, Convex optimization, Cambridge university press, 2004.

J. Bushnell, A mixed complementarity model of hydrothermal electricity competition in the western United States, Operations Research 51 (2003) 80–93.

J. Catalão, S. Mariano, V. Mendes, L. Ferreira, Scheduling of head-sensitive cascaded hydro systems: a nonlinear approach, IEEE Transactions on Power Systems 24 (2009) 337–346.

S. Cerisola, A. Baíllo, J. Fernández-López, A. Ramos, R. Gollmer, Stochastic power generation unit commitment in electricity markets: A novel formulation and a comparison of solution methods, Operations Research 57 (2009) 32–46.

S. Cerisola, J.M. Latorre, A. Ramos, Stochastic dual dynamic programming applied to nonconvex hydrothermal models, European Journal of Operational Research 218 (2012) 687–697.

Z.L. Chen, W. Powell, A convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse, Journal of Optimization Theory and Applications 103 (1999) 497–524.

A. Diniz, M. Maceira, A four-dimensional model of hydro generation for the short-term hydrothermal dispatch problem considering head and spillage effects, IEEE Transactions on Power Systems 23 (2008) 1298–1308.

C. Donohue, Stochastic network programming and the dynamic vehicle allocation problem, Ph.D. thesis, University of Michigan, 1996.

C. Donohue, J. Birge, The abridged nested decomposition method for multistage stochastic linear programs with relatively complete recourse, Algorithmic Operations Research 1 (2006) 20–30.

M.L. Fisher, An applications oriented guide to Lagrangian relaxation, Interfaces 15 (1985) 10–21.

M.L. Fisher, The Lagrangian relaxation method for solving integer programming problems, Management Science 50 (2004) 1861–1871.

B.C. Flach, L.A. Barroso, M.V.F. Pereira, Long-term optimal allocation of hydro generation for a price-maker company in a competitive market: Latest developments and a stochastic dual dynamic programming approach, IET Generation, Transmission & Distribution 4 (2010) 1–16.

S.M. Frank, S. Rebennack, Optimal design of mixed AC-DC distribution systems for commercial buildings: A nonconvex generalized Benders decomposition approach, European Journal of Operational Research 242 (2015) 710–729.

A. Geoffrion, Lagrangean relaxation for integer programming, in: M. Balinski (Ed.), Approaches to Integer Programming, volume 2 of *Mathematical Programming Studies*, Springer Berlin Heidelberg, 1974, pp. 82–114.

S. Granville, G. Oliveira, L. Thomé, N. Campodónico, M. Latorre, M.V. Pereira, L. Barroso, Stochastic optimization of transmission constrained and large scale hydrothermal systems in a competitive framework, in: IEEE Power Engineering Society General Meeting.

G. Gross, D. Finlay, Generation supply bidding in perfectly competitive electricity markets, Computational and Mathematical Organization Theory 6 (2000) 83–98.

M. Held, R.M. Karp, The Traveling Salesman Problem and minimum spanning trees, Operations Research 18 (1970) 1138–1162.

T. Homem-de-Mello, V.L. de Matos, E.C. Finardi, Sampling strategies and stopping criteria for stochastic dual dynamic programming: A case study in long-term hydrothermal scheduling, Energy Systems 2 (2011) 1–31.

G. Infanger, D. Morton, Cut sharing for multistage stochastic linear programs with interstage dependency, Mathematical Programming 75 (1996) 241–256.

C.A. Kaskavelis, M.C. Caramanis, Efficient Lagrangian relaxation algorithms for industry size job-shop scheduling problems, IIE Transactions 30 (1998) 1085–1097.

C. Lemaréchal, C. Sagastizábal, Variable metric bundle methods: from conceptual to implementable forms, Mathematical Programming 76 (1997) 393–410.

C.A. Li, A.J. Svoboda, C.L. Tseng, R.B. Johnson, E. Hsu, Hydro unit commitment in hydro-thermal optimization, IEEE Transactions on Power Systems 12 (1997) 764–769.

G. Li, J. Shi, X. Qu, Modeling methods for genco bidding strategy optimization in the liberalized electricity spot market-a state-of-the-art review, Energy 36 (2011) 4686–4700.

T. Lohmann, A.S. Hering, S. Rebennack, Spatio-temporal hydro-inflow forecasting of multireservoir inflows for hydro-thermal scheduling, European Journal of Operational Research 255 (2016) 243–258.

N. Löhndorf, D. Wozabal, S. Minner, Optimizing trading decisions for hydro storage systems using approximate dual Dynamic Programming, Operations Research 61 (2013) 810–823.

M. Maceira, V. Duarte, D. Penna, L. Moraes, A. Melo, Ten years of application of Stochastic Dual Dynamic Programming in official and agent studies in Brazil - description of the NEWAVE program, in: 16th Power Systems Computation Conference - PSCC.

L. Mallozzi, An application of optimization theory to the study of equilibria for games: a survey, Central European Journal of Operations Research 21 (2013) 523–539.

D. Morton, An enhanced decomposition algorithm for multistage stochastic hydroelectric scheduling, Annals of Operations Research 64 (1996) 211–235.

J.A. Muckstadt, S.A. Koenig, An appplication of Lagrangian relaxation to scheduling in power-generation systems, Operations Research 25 (1977) 387–403.

J.F. Nash Jr, Non-cooperative games, The Annals of Mathematics 54 (1951) 286–295.

N. Newham, Power System Investment Planning using Stochastic Dual Dynamic Programming, Ph.D. thesis, University of Canterbury, Christchurch, New Zealand, 2008.

O. Nilsson, D. Sjelvgren, Mixed-integer programming applied to short-term planning of a hydro-thermal system, IEEE Transactions on Power Systems 11 (1996) 281–286.

M.P. Nowak, W. Römisch, Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty, Annals of Operations Research 100 (2000) 251–272.

M. Padberg, Approximating separable nonlinear functions via mixed zero-one programs, Operations Research Letters 27 (2000) 1–5.

M.V.F. Pereira, L.M.V.G. Pinto, Stochastic optimization of a multireservoir hydroelectric system: A decomposition approach, Water Resources Research 21 (1985) 779–792.

M.V.F. Pereira, L.M.V.G. Pinto, Multi-stage stochastic optimization applied to energy planning, Mathematical Programming 52 (1991) 359–375.

B.T. Polyak, Minimization of unsmooth functionals, USSR Computational Mathematics and Mathematical Physics 9 (1969) 14–29.

A. Queiroz, D. Morton, Sharing cuts under aggregated forecast when decomposing multi-stage stochastic programs, Operations Research Letters 41 (2013) 311–316.

E. Read, A dual approach to Stochastic Dynamic Programming for reservoir release scheduling, in: A. Esogbue (Ed.), Dynamic programming for optimal water resources system management, Prentice Hall, NY, 1989, pp. 361–372.

S. Rebennack, Generation expansion planning under uncertainty with emissions quotas, Electric Power Systems Research 114 (2014) 78–85.

S. Rebennack, Combining sampling-based and scenario-based nested Benders decomposition methods: application to Stochastic Dual Dynamic Programming, Mathematical Programming 156 (2016) 343–389.

S. Rebennack, B. Flach, M. Pereira, P. Pardalos, Hydro-thermal scheduling under CO2 emission constraints, IEEE Transactions on Power Systems 27 (2012) 58–68.

A. Shapiro, W. Tekaya, J. da Costa, M. Soares, Risk neutral and risk averse stochastic dual dynamic programming method, European Journal of Operational Research 224 (2013) 375–391.

H. Sherali, On mixed-integer zero-one representations for separable lower-semicontinuous piecewise-linear functions, Operations Research Letters 28 (2001) 155–160.

G. Steeger, L.A. Barroso, S. Rebennack, Optimal bidding strategies for hydro-electric producers: A literature survey, IEEE Transactions on Power Systems 29 (2014) 1758–1766.

G. Steeger, S. Rebennack, Strategic bidding for multiple price-maker hydroelectric producers, IIE Transactions 47 (2015) 1013–1031.

S. Takriti, J.R. Birge, Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic problems, Operations Research 48 (2000) 91–98.

F.S. Thomé, Represenação de Não-convexidades no planejamento da operação hidrotérmica utilizando pdde, Ph.D. thesis, Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa Engenharia, 2013.

T.J. Van Roy, Cross decomposition for mixed integer programming, Mathematical Programming 25 (1983) 46–63.

J. Velásquez, GDDP: Generalized Dual Dynamic Programming theory, Annals of Operations Research 117 (2002) 21–31.

J.P. Vielma, G.L. Nemhauser, Modeling disjunctive constraints with a logarithmic number of binary variables and constraints, Mathematical Programming 128 (2011) 49–72.

S. Yakowitz, Dynamic Programming applications in water resources, Water Resources Research 18 (1982) 673–696.

Y. Yang, Y. Zhang, F. Li, H. Chen, Computing all Nash equilibria of multiplayer games in electricity markets by solving polynomial equations, IEEE Transactions on Power Systems 27 (2012) 81–91.

X. Zhao, P.B. Luh, J. Wang, Surrogate gradient algorithm for Lagrangian relaxation, Journal of Optimization Theory and Applications 100 (1999) 699–712.