

A Computationally Efficient Algorithm for Computing Convex Hull Prices

Bernard Knueven^{a,1,5}, James Ostrowski^{b,2}, Anya Castillo^{c,3,5},
Jean-Paul Watson^{d,4,5}

^a*National Renewable Energy Laboratory, Golden, CO 80401 USA*

^b*University of Tennessee, Knoxville, TN 37996 USA*

^c*Sandia National Laboratories, Albuquerque, NM 87185 USA*

^d*Lawrence Livermore National Laboratory, Livermore, CA 94550 USA*

Email addresses: Bernard.Knueven@nrel.gov (Bernard Knueven),
jostrows@utk.edu (James Ostrowski), anya.castillo@gmail.com (Anya Castillo),
watson61@llnl.gov (Jean-Paul Watson)

¹This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308.

²James Ostrowski is supported by DOE grant DE-SC0018175

³Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

⁴This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes. This work was performed in part under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

⁵B. Knueven, A. Castillo, and J.-P. Watson were supported by DOE's Office of Electricity Advanced Grid Modeling (AGM) program. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Abstract

Electricity markets worldwide allow participants to bid non-convex production offers. While non-convex offers can more accurately reflect a resource’s capabilities, they create challenges for market clearing processes. For example, system operators may be required to execute side payments to participants whose costs are not covered through energy sales as determined via traditional locational marginal pricing schemes. Convex hull pricing minimizes this and other types of side payments while providing uniform (i.e., locationally and temporally consistent) prices. Computing convex hull prices involves solving either a large-scale linear program or the Lagrangian dual of the corresponding non-convex scheduling problem. Further, the former approach requires explicit descriptions of market participants’ convex hulls.

While linear programs for computing convex hull prices are large, their structure is naturally decomposable by generators. Here, we propose and empirically analyze a Benders decomposition approach to computing convex hull prices that leverages recent advances in convex hull formulations for thermal generating units. We demonstrate across a large set of test instances that our decomposition approach only requires modest computational effort, obtaining solutions at least an order of magnitude faster than the equivalent large-scale linear programming approach. Overall, we provide a computationally feasible method for computing convex hull prices for industrial scale market clearing problems, enabling the possibility of practical adoption of this advanced pricing mechanism.

Keywords: Convex hull pricing, electricity markets, unit commitment, Benders decomposition, linear programming.

1. Introduction

In worldwide electricity markets, including independent system operators (ISOs) in the United States and organized power exchanges in Europe, price and commitment/dispatch signals ideally uphold the following properties:

- (1) Commitment and dispatch signals are given by a globally optimal solution resulting from day-ahead and real-time (balancing) market clearing.

- (2) Price signals incentivize profit-maximizing generating units to follow their commitment and dispatch signals.
- (3) Price signals are *uniform*, i.e., any two generating units at the same location and time receive identical price signals.

Property (1) ensures that overall cost is minimized while more expensive resources are not scheduled over less expensive resources. Property (2) ensures that market participants follow the market operator’s commitment/dispatch signal because it aligns with their own incentives. Property (3) ensures that prices are non-discriminatory and transparent: the former enforces a basic market fairness principle (resources are paid identically for identical products), while the latter ensures no potential revenue is hidden from those making long-term investment decisions. Although for illustrative purposes this work focuses on the unit commitment (UC) and economic dispatch (ED) problems present in the United States market clearing contexts, a similar approach could be used for the power pool structure used in throughout the world and for the “block bid” structure currently used in European power exchanges [1, 2].

If an optimization problem to determine commitment/dispatch signals is convex, then satisfying these properties is in practice computationally tractable: global solutions to convex optimization problems can typically found in polynomial time, satisfying Property (1), while price signals satisfying Property (2) and Property (3) are given by dual optimal values associated with system constraints. Most convex optimization algorithms determine optimal dual values as part of any computed optimality certificate, so no additional computational effort is generally required once optimal commitment/dispatch signals are determined.

If the optimization problem to determine commitment/dispatch signals is non-convex, the situation is much more complex. Non-convexities in power systems market models arise from the following two sources: (1) non-convexities due to thermal generating unit operating characteristics and (2) non-convexities due to the transmission system model (e.g., AC power flow). In practice, market operators linearize the second set of non-convexities, but not the first. Hence, typical long-term (day-ahead) market clearing involves solving a mixed-integer linear program (MILP). Some short-term (real-time) market models may also allow non-convexities to address units with fast-start capabilities. With a modern commercial MILP solver and advanced UC formulation, market operators can usually obtain high-quality solutions (e.g.,

$\leq 0.1\%$ optimality gap) to the corresponding MILP optimization models. However, due to the primal degeneracy present in most UC and ED models [3, 4], locating and proving the global optimality of a solution is often much more difficult. Hence, in practice, we often cannot satisfy property (1) above. Determining price signals satisfying (2) and (3) in the presence of non-convexities is even more challenging – and in general, such price signals might not mathematically exist [5].

Addressing these challenges has made for an active area of recent research [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. However, these problems are not yet solved. In practice, system operators usually decide on prices that are approximately uniform and incentivize most market participants to follow their assigned commitment/dispatch signals. Side payments and/or penalty mechanisms are then employed on an individual participant basis for those generators not properly incentivized. In most markets, payments are introduced for those participants incentivized to under-produce and penalties are introduced for participants incentivized to over-produce.

Obtaining globally optimal solutions to the unit commitment problem as required by Property (1) is the subject of its own line of research so we will not consider it further here. In the context of Property (2) and Property (3), convex hull pricing (CHP) is a conceptually attractive approach. Convex hull prices use the tightest convex relaxation for each generator’s feasible operating region and its cost function to determine the price of energy and other grid services. CHP minimizes specific side payments, which increases market transparency; the latter is consistent with Property (3). Consequently, lost opportunity costs are usually lower under CHP, leading to price signals that incentivize more generators to follow their commitment/dispatch signals. However, there are known issues with CHP, e.g., commitment/dispatch solutions are not used in their calculation. This leads to some counter-intuitive results, such as off-line generators being marginal in pricing problems, and non-binding system constraints from the commitment/dispatch problem having non-zero shadow prices [10]. Further, solving the convex hull pricing problem (CHPP) has proven difficult in practice; [13] presents an exact method that formulates a large-scale linear program (LP). Both [13] and [14] propose decomposition methods which are heuristics when ramping constraints are present.

While in this work we will not explicitly compare CHP and other pricing mechanisms, we do point out a potential benefit of CHP. We note that any pricing scheme that uses sub-optimal commitments/dispatches as a basis is

as capricious as the selection of that non-optimal commitment/dispatch [6]. Given that most commitment/dispatch solutions are determined using MILP heuristics, it is difficult to justify a price signal based on the corresponding globally sub-optimal solution [3]. That said, this approach could be adapted to other pricing schemes which depend on explicit knowledge of the convex hull for market participants [15].

In this paper, we will set aside the question of whether convex hull prices are “the” price signal that should be used in practice (see [12] for a recent small-scale comparison against other pricing schemes), and instead turn to the practical matter of their implementation: Can they be efficiently computed for an industrial-scale (e.g., an ISO in the US) market clearing problem? We will ultimately answer this question in the affirmative, with a few caveats, by introducing and empirically analyzing the computational performance of a Benders decomposition approach for solving the CHPP [16]. We also demonstrate that in practice, the price signal obtained using “approximated” convex hull prices (aCHP) is very similar to the CHP signal, and that the difference between aCHP and CHP is empirically related to the strength of the convex hull relaxation.

The remainder of this paper is organized as follows. Section 2 details the sets, parameters, and variables that will be used throughout this manuscript. We then begin in Section 3 with a description of CHP and the closely related aCHP modeling and computation. In Section 4, we consider convex hull descriptions for thermal generators. We introduce our Benders decomposition algorithm for the CHPP in Section 5. In Section 6, we conduct an empirical analysis of the efficiency of our Benders decomposition algorithm on a range of large-scale instances. In Section 7 we demonstrate the difference between various aCHP methods and CHP. Finally, we conclude with a summary of our results in Section 8.

2. Nomenclature

2.1. Indices and Sets

$g \in \mathcal{G}$	Thermal generators.
$l \in \mathcal{L}_g$	Piecewise production cost intervals for generator g : $1, \dots, L_g$.
$t \in \mathcal{T}$	Hourly time steps: $1, \dots, T$.
$[a, b] \in \mathcal{Y}^g$	Feasible intervals of operation for generator g with respect to its minimum up-time, i.e., $[a, b] \in \mathcal{T} \times \mathcal{T}$ such that $b \geq a + UT^g - 1$.

$[c, d] \in \mathcal{Z}^g$ Feasible intervals of operation for generator g with respect to its minimum down-time, i.e., $[c, d] \in \mathcal{T} \times \mathcal{T}$ such that $d \geq c + DT^g - 1$.

2.2. Parameters

C_g^l Cost coefficient for piecewise segment l of generator g (\$/MWh).
 $C_g^{s,[c,d]}$ Start-up cost for generator g given non-operation in the interval $[c, d] \in \mathcal{Z}^g$ (\$).
 D Vector of loads (MW).
 DT^g Minimum down-time for generator g (h).
 \overline{P}^g Maximum power output for generator g (MW).
 \underline{P}^g Minimum power output for generator g (MW).
 \overline{P}_g^l Maximum power for piecewise segment l of generator g (MW).
 RD^g Ramp-down rate for generator g (MW/h).
 RU^g Ramp-up rate for generator g (MW/h).
 SD^g Shut-down ramp rate for generator g (MW/h).
 SU^g Start-up ramp rate for generator g (MW/h).
 UT^g Minimum up-time for generator g (h).

2.3. Variables

c_t^g Total cost of operation for generator g at time t (\$).
 $c_t^{g,p}$ Power production cost for generator g at time t (\$).
 $c_t^{g,s}$ Start-up cost for generator g at time t (\$).
 p_t^g Power output for generator g at time t (MW).
 $p_t^{g,l}$ Power from piecewise interval l for generator g at time t (MW), ≥ 0 .
 w_t^g Commitment status of generator g at time t , $\in [0, 1]$.
 v_t^g Start-up status of generator g at time t , $\in [0, 1]$.
 w_t^g Shut-down status of generator g at time t , $\in [0, 1]$.
 $y_{[a,b]}^g$ Indicator arc for start-up at time a , stopping at time b (shut-down at time $b + 1$), committed for $i \in [a, b]$, for generator g , $\in [0, 1]$, $[a, b] \in \mathcal{Y}^g$.
 $z_{[c,d]}^g$ Indicator arc for shut-down at time c , ending at time d (start-up at time $d + 1$), uncommitted for $i \in [c, d]$, for generator g , $\in [0, 1]$, $[c, d] \in \mathcal{Z}^g$.

3. A Primal Pricing Problem

Consider the following generic UC MILP formulation

$$z^{\text{UC}} = \min \sum_{g \in \mathcal{G}} c^g \quad (\text{UC.1})$$

subject to

$$\sum_{g \in \mathcal{G}} (A^g p^g + B^g u^g) = D \quad (\text{UC.2})$$

$$(u^g, p^g, c^g) \in \Pi^g, \forall g \in \mathcal{G} \quad (\text{UC.3})$$

where p^g denotes the (continuous) dispatch vector for generator g , u^g denotes the (binary) commitment vector for generator g , and c^g denotes the cost vector of the schedule associated with p^g and u^g . Let Π^g denote the set of feasible schedules and costs for generator g , per Constraint (UC.3). Constraint (UC.2) abstractly represents the set of system balance constraints with any required auxiliary variables s and right-hand side D . Finally, Objective (UC.1) minimizes system operations costs. We will assume the relationships (UC.2) are linear, although the most general form of (UC) is non-linear and defines nodal power injections for an AC transmission network. We will also assume throughout that problem (UC) is feasible.

Convex hull prices are derived from the Lagrangian relaxation of (UC) with respect to the system constraints:

$$L(\pi) = \min \sum_{g \in \mathcal{G}} c^g + \pi^T (D - A^g p^g - B^g u^g) \quad (\text{LR.1})$$

subject to

$$(u^g, p^g, c^g) \in \Pi^g, \forall g \in \mathcal{G}, \quad (\text{LR.2})$$

with associated Lagrangian dual:

$$z^{\text{LD}} = \max_{\pi} L(\pi). \quad (\text{LD})$$

Convex hull prices are defined to be the optimal solutions π to (LD) [5, 6, 7, 10].

Let $\text{conv}(\cdot)$ denote the convex hull of the given set. As shown in [12], the CHPP can be solved via the following linear program (LP):

$$z^{\text{CH}} = \min \sum_{g \in \mathcal{G}} c^g \quad (\text{CH.1})$$

subject to

$$\sum_{g \in \mathcal{G}} (A^g p^g + B^g u^g) = D \quad (\pi^{\text{CH}}) \quad (\text{CH.2})$$

$$(u^g, p^g, c^g) \in \text{conv}(\Pi^g), \quad \forall g \in \mathcal{G}. \quad (\text{CH.3})$$

The optimal dual vector π^{CH} associated with the system constraints (CH.2) are also CHPs. This fact follows from an observation of [17], which states that (CH) is the LP dual of (LD), such that (when both problems are bounded) an optimal solution π^* of (LD) is an optimal dual vector for the constraints (CH.2), and an optimal dual vector $\pi^{\text{CH},*}$ is an optimal solution for (LD).

Problem (CH) is a LP because (CH.2) is a linear set of constraints and the convex hull of the mixed-integer set Π^g is polyhedral, for every g . A traditional problem with this “primal” approach for CHPP is that one needs explicit knowledge of $\text{conv}(\Pi^g)$ for each participating generator in the market clearing problem. Recent work has made this requirement less of an issue than in the past. In particular, it is now known that there exists a polynomial-sized convex hull formulation for a general thermal generator, i.e., with minimum/maximum power levels, minimum/maximum up-time/down-time constraints, and ramping constraints [18]. Specifically, as long as the dispatch for a turn-on at time t_1 and a turn-off at time t_2 can be represented using a compact polyhedron, expressing $\text{conv}(\Pi^g)$ is relatively straightforward. We will discuss below how this assumption can be relaxed.

In practice, the convex hull characterization for thermal generators established in [18] is large: for a generator with minimum up-time and irredundant ramping constraints, $\mathcal{O}(T^3)$ variables and constraints are required. Hence solving Problem (CH) as a monolithic large-scale LP becomes impractical, especially for market clearing problems with realistic-sized systems and longer time-horizons, e.g., as is the case for day-ahead.

To address this issue, one may compute an “approximated” convex hull price (aCHP). The word approximated is used loosely in this context, because

to the best of the knowledge of the authors, there is no formal approximation theorem or guarantee underlying this approach. This method has also been referred to as the “dispatchable model” [5].

To demonstrate, let $\mathcal{R}(\Pi^g)$ denote a MILP relaxation for the set Π^g , i.e., $\Pi^g = \{(u^g, p^g, c^g) \in \mathcal{R}(\Pi^g) \mid u^g \in \mathbb{Z}\}$ and $\mathcal{R}(\Pi^g)$ is a polyhedron. If $\mathcal{R}(\Pi^g) \neq \text{conv}(\Pi^g)$ for some $g \in \mathcal{G}$, we can compute aCHPs by solving the linear program

$$z^{\text{LPR}} = \min \sum_{g \in \mathcal{G}} c^g \quad (\text{LPR.1})$$

subject to

$$\sum_{g \in \mathcal{G}} (A^g p^g + B^g u^g) = D \quad (\pi^{\text{aCHP}(\mathcal{R})}) \quad (\text{LPR.2})$$

$$(u^g, p^g, c^g) \in \mathcal{R}(\Pi^g), \forall g \in \mathcal{G}. \quad (\text{LPR.3})$$

Again, the prices are defined as the optimal dual vector $\pi^{\text{aCHP}(\mathcal{R})}$. We note that the following relationships between the optimal objectives of the aforementioned problems (when all are feasible): $z^{\text{aCHP}(\mathcal{R})} \leq z^{\text{CH}} = z^{\text{LD}} \leq z^{\text{UC}}$; the first follows from the fact that (LPR) is a relaxation of (CH), the second from the strong duality between (CH) and (LD), and the last by the weak duality between (LD) and (UC) [17].

As demonstrated in [4, 12], aCHPs provide adequate incentives and minimize wealth transfers when compared to other pricing methods. One disadvantage of using aCHPs is that for any two relaxations $\mathcal{R}^1(\Pi^g)$ and $\mathcal{R}^2(\Pi^g)$, the prices given by (LPR) will, in general, be different [19, 20]. This is particularly relevant in practice as it is common for market clearing UC and ED models to change their descriptions of $\mathcal{R}(\Pi^g)$ as better formulations are devised – e.g., see [21] for a recent comprehensive overview of UC formulations. Thus, market operators employing aCHP – or a variant thereof – should be very careful about any adjustments to \mathcal{R} . Even if they are computationally beneficial for solving problem (UC), they may have unanticipated side-effects for the pricing problem (LPR), warranting further study.

Finally, we note that while tighter UC formulations will generally increase the objective function value of problem (LPR), it is difficult to know a priori if this will result in an increase in prices. For example, in an over-generation scenario tightening the generator formulation may result in a

decrease in prices by forcing even more generation on-line (when it is not needed) in problem (LPR).

4. Constructing Thermal Generator Convex Hull Formulations

As indicated previously, one potential drawback to using Problem (CH) to solve the CHPP is that it requires explicit knowledge of the convex hull for each generator. Because any real generator has a bounded feasible region, it is always possible to construct an explicit convex hull representation using disjunctive programming [22, 10]. Doing so, however, requires complete enumeration of the feasible discrete points in Π^g , which may be exponential in length of the scheduling horizon – leading to unreasonably sized linear programming problems.

However, one can express $\text{conv}(\Pi^g)$ with a polynomial-sized extended formulation for large classes of generators by using some theorems from integer programming. In certain untypical cases (e.g., inter-temporal maximum energy output or inter-temporal fuel constraints), no polynomial-sized convex hull may be known. For these cases, one can always resort to the aforementioned disjunctive programming result [22, 10]. We now review these results and some “tools of the trade” that can be used to express $\text{conv}(\Pi^g)$ for a single generator g , and will drop the index g for convenience.

The “simplest” type of generator has minimum and maximum generation levels \underline{P} and \overline{P} and linear production costs C . A convex hull description for such a generator is as follows:

$$c = \sum_{t \in \mathcal{T}} Cp_t \tag{1a}$$

$$\underline{P}u_t \leq p_t \leq \overline{P}u_t \quad \forall t \in \mathcal{T} \tag{1b}$$

$$u_t \in [0, 1] \quad \forall t \in \mathcal{T}. \tag{1c}$$

Note that if u_t is restricted to be binary, then this is a valid generator formulation. This result is of course not new or interesting. However, we now demonstrate how to prove this is a convex hull description for this generator by applying a simple lemma twice, which is given below.

Lemma 1. [23, Lemma 4] *Suppose $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ is an integer polyhedron. We introduce new variables $y \in \mathbb{R}^k$ such that $Q = \{(x, y) \in \mathbb{R}^{n+k} \mid l^i x \leq y_i \leq u^i x, i = 1, \dots, k\}$, with $l^i x \leq u^i x$ for every $x \in P$,*

$i = 1, \dots, k$. (That is, each new variable y_i has at most one upper bound $u^i x$ and one lower bound $l^i x$.) Then $P \cap Q$ is an integer polyhedron in x . If $l^i x, u^i x \in \mathbb{Z}$ for every $x \in P \cap \mathbb{Z}$ for some $i \in \{1, \dots, k\}$ then $P \cap Q$ is also integer in y_i .

We apply Lemma 1 as follows. First, consider Constraint (1c), which is the convex hull of $\{u \mid u \in \{0, 1\}^T\}$. (This is because the only vertices of the T -dimensional hypercube described by Constraint (1c) have integer values, and exactly correspond to on/off sequences for this generator.) Now we can arrive at (1) being integer in u by applying Lemma 1 twice: the first time with $P = \{u \in \mathbb{R}^T \mid (1c)\}$, $Q = \{(u, p) \in \mathbb{R}^{2T} \mid (1b)\}$ and the second time with $P = \{(u, p) \mid (1b), (1c)\}$, $Q = \{(u, p, c) \in \mathbb{R}^{2T+1} \mid (1a)\}$.

Lemma 1 effectively shows the following: if the inter-temporal constraints on generator g can be expressed in the integer variables alone, and there is a known convex hull description for the inter-temporal constraints in the integer variables, then basic dispatch and linear costs can be added without affecting the integrality of the integer variables. Gentile et. al [23] exploit Lemma 1 to obtain a convex hull representation for a generator with minimum up/down-times in addition to start-up and shut-down ramp rates. This result is extended using the same lemma to a generator with piecewise production costs in [21]. However, if the inter-temporal constraints cannot be expressed in the integer variables alone, we can use the next Lemma.

Lemma 2. [24, 25, 22, 18, 26] Let $P^i := \{x^i \in \mathbb{R}^n \mid A^i x^i \leq b^i\}$, $i = 1, \dots, m$ be a collection of bounded polyhedra and let $\Gamma \subset \mathbb{R}_+^m$ be a non-empty polyhedron in the non-negative orthant. Consider the set $P := \cup_{\gamma \in \Gamma} \{\sum_{i=1}^m \gamma_i P^i\}$ and define the polyhedron

$$Y := \begin{cases} A^i x^i \leq \gamma_i b^i, & i \in [m] \\ \sum_{i=1}^m x^i = x \\ (\gamma_1, \dots, \gamma_m) = \gamma \in \Gamma. \end{cases} \quad (2)$$

Then the following are true:

- (1) $\text{proj}_x(Y) = P$ and P is a polyhedron
- (2) If Γ is an integer polytope, then all the vertices of Y have integer γ
- (3) If Γ and all P^i , $i = 1, \dots, m$ are integer polytopes with integer coefficient constraint matrices and right-hand sides, then Y is an integer polytope.

In the context of market clearing, it is natural to see Γ as the polytope defining the relationships between the discrete decision variables, with the polytopes P^i representing the feasible dispatches given a particular discrete decision γ_i . Nearly any generator's convex hull can be represented in the form of equation (2) by simply enumerating the discrete decision space and associating each point in it with a γ_i , which is exactly the disjunctive programming result [22, 10]. Sometimes the discrete and continuous decisions are additive in the “right” way such that a full enumeration of the discrete decision space is not required – this is the basic insight from [18], which provides a polynomial-sized extended formulation for a generator with ramping constraints, minimum up-time/down-time constraints, piecewise production costs, and time-varying start-up costs. We now state this result in a general form.

Define the economic dispatch polytope $ED^{[a,b]}$ for a generator starting at time a and stopping at time b (off at $t = b + 1$) as follows:

$$\begin{aligned}
p_t^{[a,b]} &\leq \bar{P} && t \in [a + 1, b - 1] \cap \mathcal{T} \\
p_a^{[a,b]} &\leq SU \\
p_b^{[a,b]} &\leq SD \\
p_t^{[a,b]} &\leq 0 && t \in \mathcal{T}, t \notin [a, b] \\
-p_t^{[a,b]} &\leq -\underline{P} && t \in [a, b] \cap \mathcal{T} \\
-p_t^{[a,b]} &\leq 0 && t \in \mathcal{T}, t \notin [a, b] \\
-p_t^{[a,b]} &\leq 0 && t \in \mathcal{T} \\
p_t^{[a,b]} - p_{t-1}^{[a,b]} &\leq RU && t \in [a + 1, b] \cap \mathcal{T} \\
p_{t-1}^{[a,b]} - p_t^{[a,b]} &\leq RD && t \in [a + 1, b] \cap \mathcal{T} \\
p_t^{l,[a,b]} &\leq \bar{P}^l - \bar{P}^{l-1} && l \in \mathcal{L}, t \in [a, b] \cap \mathcal{T} \\
p_t^{l,[a,b]} &\leq 0 && l \in \mathcal{L}, t \in \mathcal{T}, t \notin [a, b] \\
-p_t^{l,[a,b]} &\leq 0 && l \in \mathcal{L}, t \in \mathcal{T} \\
\sum_{l \in \mathcal{L}} p_t^{l,[a,b]} &\leq p_t^{[a,b]} && t \in [a, b] \cap \mathcal{T} \\
-\sum_{l \in \mathcal{L}} p_t^{l,[a,b]} &\leq -p_t^{[a,b]} && t \in [a, b] \cap \mathcal{T}
\end{aligned}$$

$$\begin{aligned}
\sum_{l \in \mathcal{L}} C^l p_t^{l,[a,b]} - c_t^{p,[a,b]} &\leq 0 & t \in [a,b] \cap \mathcal{T} \\
-\sum_{l \in \mathcal{L}} C^l p_t^{l,[a,b]} + c_t^{p,[a,b]} &\leq 0 & t \in [a,b] \cap \mathcal{T} \\
-c_t^{p,[a,b]} &\leq 0 & t \in \mathcal{T}, t \notin [a,b] \\
c_t^{p,[a,b]} &\leq 0 & t \in \mathcal{T}, t \notin [a,b].
\end{aligned}$$

We summarize the above system as $ED^{[a,b]} = \{p^{[a,b]}, p^{\mathcal{L},[a,b]}, c^{[a,b]} \mid D^{[a,b]}(p^{[a,b]}, p^{\mathcal{L},[a,b]}, c^{p,[a,b]}) \leq f^{[a,b]}\}$. The additional superscript $[a,b]$ indicates that the variable denotes the specified quantity *given* the generator starts at time a and stops at time b .

We then define the generator polytope G as:

$$D^{[a,b]}(p^{[a,b]}, p^{\mathcal{L},[a,b]}, c^{p,[a,b]}) \leq f^{[a,b]} y^{[a,b]} \quad [a,b] \in \mathcal{Y} \quad (\text{GP.1})$$

$$\sum_{[a,b] \in \mathcal{Y}} p_t^{[a,b]} = p_t \quad t \in \mathcal{T} \quad (\text{GP.2})$$

$$\sum_{[a,b] \in \mathcal{Y}} p_t^{l,[a,b]} = p_t^l \quad l \in \mathcal{L}, t \in \mathcal{T} \quad (\text{GP.3})$$

$$\sum_{[a,b] \in \mathcal{Y}} c_t^{p,[a,b]} = c_t^p \quad t \in \mathcal{T} \quad (\text{GP.4})$$

$$\sum_{\{[c,d] \in \mathcal{Z} \mid t-1=d\}} C^{s,[c,d]} z^{[c,d]} = c_t^s \quad t \in \mathcal{T} \quad (\text{GP.5})$$

$$\sum_{\{[a,b] \in \mathcal{Y} \mid t=a\}} y^{[a,b]} - \sum_{\{[c,d] \in \mathcal{Z} \mid t-1=d\}} z^{[c,d]} = 0 \quad t \in \mathcal{T} \quad (\text{GP.6})$$

$$\sum_{\{[c,d] \in \mathcal{Z} \mid t=c\}} z^{[c,d]} - \sum_{\{[a,b] \in \mathcal{Y} \mid t-1=b\}} y^{[a,b]} = 0 \quad t \in \mathcal{T} \quad (\text{GP.7})$$

$$\sum_{\{[a,b] \in \mathcal{Y} \mid a < 1\}} y^{[a,b]} + \sum_{\{[c,d] \in \mathcal{Z} \mid c < 1\}} z^{[c,d]} = 1 \quad (\text{GP.8})$$

$$\sum_{\{[a,b] \in \mathcal{Y} \mid b > T\}} y^{[a,b]} + \sum_{\{[c,d] \in \mathcal{Z} \mid d > T\}} z^{[c,d]} = 1 \quad (\text{GP.9})$$

$$y^{[a,b]} \geq 0 \quad [a,b] \in \mathcal{Y}; \quad z^{[c,d]} \geq 0 \quad [c,d] \in \mathcal{Z}. \quad (\text{GP.10})$$

By applying Lemma 2, we can show the polytope G defined in Equation (GP) has integer vertices, and hence is a convex hull description.

Theorem 1. [27, 28, 18, 29] *The generator polytope G is integer in y and z .*

Proof. Equations (GP.6)–(GP.10) are a shortest path polytope on a diagraph given initial arcs defined by the previous generator state. Equation (GP.5) simply adds an additional set of variables to (GP.6)–(GP.10) in the form of Lemma 1. So (GP.5)–(GP.10) is a bounded integer polytope with integer y and z vertices, and all $ED^{[a,b]}$ are bounded polytopes. Hence Lemma 2 ensures G is a polytope with integer y and z vertices. \square

Remark 1. *We can add various ancillary service products (and prices for providing them) to G , provided $ED^{[a,b]}$ remains a bounded polytope.*

As mentioned before, certain types of inter-temporal constraints (maximum energy, fuel) may still require enumerating all possible on-off sequences. However, even in this case we may not need to construct such formulations if the generator happens to have integer commitment variables when using some relaxation of its convex hull, as explained in the next section.

5. Benders Decomposition for CHPP

Applying Lemmas 1 and 2, one can explicitly characterize the convex hull for any generator with a bounded feasible region, and compute π^{CH} by solving Problem (CH). However, for ramping-constrained generators, the smallest representation for each $g \in \mathcal{G}$ may be $\mathcal{O}(T^3)$ in size. For a power system with several hundred ramping-constrained generators, even if each has a known polynomial-sized convex hull, problem (CH) is a large-scale LP with tens of millions of non-zeros in the constraint matrix. Further, most of the variables and constraints in a convex hull description of a ramping-constrained generator, like (GP), will be non-active or non-binding in an optimal solution – suggesting alternative approaches to direct solution of the large-scale LP could be effective.

Modern UC MILP formulations provide relatively tight approximations of $\text{conv}(\Pi^g)$. This observation motivates our proposed Benders decomposition approach to solve (CH), which is outlined as follows. For the Benders master

problem, we select a relatively tight (but more compact than (GP)) MILP formulation – which we will denote \mathcal{R}^* – and construct following problem:

$$z^{\text{BD}} = \min \sum_{g \in \mathcal{G}} c^g \quad (\text{BD.1})$$

subject to

$$\sum_{g \in \mathcal{G}} (A^g p^g + B^g u^g) = D \quad (\pi^{\text{BD}}) \quad (\text{BD.2})$$

$$(u^g, p^g, c^g) \in \mathcal{R}^*(\Pi^g), \quad \forall g \in \mathcal{G} \quad (\text{BD.3})$$

$$\alpha_{g,i}^T u^g + \beta_{g,i}^T p^g + \delta_{g,i}^T c^g \leq \eta_{g,i}, \quad \forall i \in F_g, \forall g \in \mathcal{G}, \quad (\text{BD.4})$$

where F_g is the set of Benders cuts for generator g produced thus far. Initially, $F_g = \emptyset, \forall g \in \mathcal{G}$.

Relaxing our notation slightly by allowing u^g to be the vector of binary decision variables in \mathcal{R}^* , p^g be the vector of dispatch variables in \mathcal{R}^* , and c^g be the vector of all cost variables in \mathcal{R}^* , we can certify a proposed master solution $(\hat{u}^g, \hat{p}^g, \hat{c}^g) \in \text{conv}(\Pi^g)$ one of three ways.

- (1) For generator g , we may know that $\mathcal{R}^*(\Pi^g) = \text{conv}(\Pi^g)$. In this case, there is no need for a Benders subproblem.
- (2) For generator g , we may have \hat{u}^g is binary in the master problem solution. Because $\mathcal{R}^*(\Pi^g)$ and $\text{conv}(\Pi^g)$ must agree when \hat{u}^g is binary, this certifies $(\hat{u}^g, \hat{p}^g, \hat{c}^g) \in \text{conv}(\Pi^g)$, and again a Benders subproblem is not needed.
- (3) If (1) and (2) are both false for generator g , then we can test $(\hat{u}^g, \hat{p}^g, \hat{c}^g) \in \text{conv}(\Pi^g)$ through a Benders feasibility subproblem.

For the Benders subproblems, we adopt the LP from [18, Eq. (22)] for the feasibility cut-generation subproblems. We can the generator polytope G to either certify $(\hat{u}^g, \hat{p}^g, \hat{c}^g) \in \text{conv}(\Pi^g)$, or generate a valid cut of the form

$$\alpha_{g,i}^T u^g + \beta_{g,i}^T p^g + \delta_{g,i}^T c^g \leq \eta_{g,i}, \quad (3)$$

to be added to the set F_g (and hence the master problem), which eliminates $(\hat{u}^g, \hat{p}^g, \hat{c}^g)$ from the feasible region as $\alpha_{g,i}^T \hat{u}^g + \beta_{g,i}^T \hat{p}^g + \delta_{g,i}^T \hat{c}^g > \eta_{g,i}$. We then update the master with the new cuts for every g with a violation, and

Algorithm 1 (BENDERS CHP) Solves the CHPP using Benders decomposition.

```

 $\mathcal{G}^R \leftarrow \{g \in \mathcal{G} \mid g \text{ has irredundant ramping constraints}\}$ 
 $F_g \leftarrow \emptyset$ 
 $cuts \leftarrow True$ 
while  $cuts$  do
     $cuts \leftarrow False$ 
    Solve master problem (BD)
    for  $g \in \mathcal{G}^R$  do
        Solve feasibility subproblem (GP) for  $g$  with
             $(\hat{u}^g, \hat{p}^g, \hat{c}^g)$  fixed from solution of master
        if feasibility subproblem is infeasible then
            add cut  $\alpha_{g,i}^T u^g + \beta_{g,i}^T p^g + \delta_{g,i}^T c^g \leq \eta_{g,i}$  to  $F_g$ 
             $cuts \leftarrow True$ 
    return Primal values  $(u^*, p^*, c^*) := \{(\hat{u}^g, \hat{p}^g, \hat{c}^g) \in \text{conv}(\Pi^g)\}_{g \in \mathcal{G}}$  and
        dual values  $\pi^{BD}$  of Constraint (LPR.2)

```

continue until all the subproblems certify that each generator's (fractional) commitment/dispatch is in its convex hull.

This process is summarized in Algorithm 1. The following subsections present some theoretical properties and implementation details.

5.1. Theoretical Properties

Theorem 2 (Benders Procedure[16]). *If $\text{conv}(\Pi^g)$ is a polytope for every $g \in \mathcal{G}$, and the CGLP solver returns vertex solutions, then Algorithm 1 converges in a finite number of iterations. Further, when converged, the returned primal values (u^*, p^*, c^*) are an optimal solution to (CH).*

Theorem 3. *The prices π^{BD} returned by Algorithm 1 are convex hull prices.*

Proof. At every iteration, the prices π^{BD} are feasible for (LD). At the termination of Algorithm 1, $z^{BD} = z^{CH}$ by Theorem 2 and $z^{CH} = z^{LD}$ by the strong duality between (CH) and (LD), so $z^{BD} = z^{LD}$. Because (BD) is a relaxation of (CH), by [17, Theorem 1(b)], $z^{BD} \leq L(\pi^{BD})$. But as $L(\pi^{BD}) \leq z^{LD}$ and $z^{BD} = z^{LD}$, it follows that $L(\pi^{BD}) = z^{LD}$. Therefore π^{BD} is an optimal solution to (LD).

As convex hull prices are optimal solutions to (LD), π^{BD} are convex hull prices at the termination of Algorithm 1. \square

Remark 2. *The proof of Theorem 3 shows that for any iterative algorithm which (i) maintains a relaxation of (CH) at every iteration and (ii) converges in objective to z^{CH} , computes convex hull prices as part of its dual optimality certificate at the final iteration.*

In our computational experiments described below, the assumptions of Theorems 2 and 3 are satisfied, so both methods solve the CHPP. In particular, the generators in our test set have the following standard features: minimum/maximum power, spinning reserve, ramping constraints, minimum up/down times, time-dependent start-up costs, and piecewise production costs. This is in contrast to the iterative algorithms in [13] and [14], which in general are heuristics when ramping constraints are present; however, Remark 2 shows that variants of the IAC1 and IAC2 algorithms from [13] (with less stringent stopping rules) will converge to convex hull prices.

5.2. Formulating \mathcal{R}^*

For \mathcal{R}^* , we adopt the “Tight” formulation from [21] with a slight modification described below. Hence, the Benders master problem is just (LPR). If a generator is not ramping-constrained, this formulation is known to always have optimal LP solutions which are in $\text{conv}(\Pi^g)$ for generators with minimum up/down-times, piecewise production costs, start-up/shut-down ramping limits, and non-decreasing time-dependent start-up costs [30, 31]. Lemma 1 can be applied sequentially (over $t \in \mathcal{T}$) to add a basic spinning up reserve to this formulation while maintaining integrality in the optimal solutions. Hence, for any generator with irredundant (i.e., non-trivial, such that they may be active in an optimal solution) ramping constraints, in any optimal solution, the fractional $(\hat{u}^g, \hat{p}^g, \hat{c}^g) \in \mathcal{R}^*(\Pi^g)$ certifies $(\hat{u}^g, \hat{p}^g, \hat{c}^g) \in \text{conv}(\Pi^g)$.

5.3. Enhancements to \mathcal{R}^* for the Benders Master Problem

To tighten the piecewise production cost curves for a ramping-constrained generator, we take the suggestion from [21, Sect. 3.6] and use the start-up and shut-down ramping trajectories to tighten the bounds on the p_t^l variables. Let $T^{RU} = \left\lfloor \frac{\bar{P}-SU}{RU} \right\rfloor$ and $T^{RD} = \left\lfloor \frac{\bar{P}-SU}{RD} \right\rfloor$, and suppose $UT \geq T^{RU} + T^{RD} + 2$.

The following is such a strengthening of the piecewise cost-curve:

$$p_t^l \leq (\bar{P}^l - \bar{P}^{l-1})u_t - \sum_{i=0}^{TRU} C^v(l, i)v_{t-i} - \sum_{i=0}^{TRD} C^w(l, i)w_{t+1+i} \quad (4)$$

where

$$C^v(l, i) := \begin{cases} 0 & \bar{P}^l \leq SU + iRU \\ \bar{P}^l - (SU + iRU) & \bar{P}^{l-1} < SU + iRU < \bar{P}^l \\ \bar{P}^l - \bar{P}^{l-1} & \bar{P}^{l-1} \geq SU + iRU \end{cases}$$

and

$$C^w(l, i) := \begin{cases} 0 & \bar{P}^l \leq SD + iRD \\ \bar{P}^l - (SD + iRD) & \bar{P}^{l-1} < SD + iRD < \bar{P}^l \\ \bar{P}^l - \bar{P}^{l-1} & \bar{P}^{l-1} \geq SD + iRD \end{cases}.$$

When $UT < T^{RU} + T^{RD} + 2$, we can drop some of the terms of (4) so long as the number of consecutive v and w terms (in time) is no more than UT .

The impetus behind adding constraints like (4) is to lower the total number of Benders cuts that must be added to the master problem, while still keeping the master problem relatively compact.

5.4. Formulation of the Benders Subproblems

We formulate the Benders subproblems using the generator formulation (GP) as a basis. For each ramping-constrained generator g , we take the values $(\hat{u}^g, \hat{p}^g, \hat{c}^g)$ from the master problem, fix the associated variables in (GP), and minimize the distance to feasibility. If the minimum distance to feasibility is 0, then the associated point $(\hat{u}^g, \hat{p}^g, \hat{c}^g)$ is the convex hull of generator g . If not, we can use the dual information from the subproblem to generate a linear cut which eliminates the point $(\hat{u}^g, \hat{p}^g, \hat{c}^g)$ from the feasible region, which is then added to the master problem. Details of the subproblem formulation can be found in [18, Eq. (22)].

6. Computational Experiments

We now assess the empirical performance of our Benders decomposition approach to the alternative of directly solving the following large-scale LP:

for all ramping-constrained generator $g \in G$ (such that we cannot be sure a solution is in $\text{conv}(\Pi^g)$), we formulate its polytope and add the associated constraints to Problem (LPR). This modified version of Problem (LPR) will have the same optimal objective function value as that of Problem (CH). Therefore, for any optimal solution, $(u^g, p^g, c^g) \in \text{conv}(\Pi^g)$ for every $g \in \mathcal{G}$. We refer to this formulation as the “extensive form” (EF) for computing CHPs. We report summary results for computing exact CHPs using both the proposed Benders decomposition approach and the EF approach described above.

6.1. Instances

We use the same UC test instances as reported in [21], but consider variants with a 24 hour scheduling horizon in addition to the standard 48 hour scheduling horizon. The smallest system, which we refer to as the modified RTS-GMLC [32, 33], has 73 thermal units – of which 26 have irredundant ramping constraints – at an hourly time horizon. The system is modified from [33] by dividing all ramping rates by 3 to allow for some generators with irredundant ramping constraints. The medium sized system, which we refer to as CAISO – derived from public data obtained from the California Independent System Operator in the US, have 610 thermal units, of which only 20 have irredundant ramping constraints at an hourly time horizon. Finally, the largest system, which we refer to as FERC – these instances were released by the US Federal Energy Regulatory Commission, consists of two sets of generators. The first set (Winter) has 934 thermal units (of which 472 have irredundant hourly ramping constraints) and the second (Summer) has 978 thermal unit (of which 504 have irredundant hourly ramping constraints). From a computational perspective the FERC instances are the most challenging on paper, as they require the most ramping polytopes. For more details, the reader is referred to [21]. All three sets of instances are publicly available as part of the IEEE PES Power Grid Lib - Unit Commitment benchmark library (<https://github.com/power-grid-lib/pglib-uc>).

6.2. Computational Enhancements

Due to consistently superior performance, all LPs were initially solved using Gurobi’s [34] barrier optimizer with crossover. All master problem updates associated with Benders decomposition were solved with dual simplex. Each Benders subproblem was initially solved using the barrier method, with the additional parameters `presolve=2` (the maximum intensity) and

`crossover=4`. The later setting directs Gurobi to execute crossover by conducting a primal push first, followed by a dual push, and then uses dual simplex for cleanup. For our Benders subproblems, this parameter change had a significant impact on crossover performance. Finally, if the Benders subproblem for generator g yields a cut, we discard the existing simplex basis and execute the next solve for generator g using barrier. Conversely, if a Benders subproblem for generator g does not yield a cut, we keep the existing simplex basis and execute the subsequent subproblem solve for generator g using simplex. The reasoning behind this last enhancement is as follows: a cut generated on the last pass moves the master problem solution for the next pass “far enough” away that the basis is no longer useful. On the other hand, if no cut was generated on the last pass, we would expect the next solution from the master problem to be “close” to the prior solution, such that the prior basis is still useful.

6.3. 24-Hour Instances

6.3.1. Computational Platform

We conducted all experiments with 24 hour scheduling horizon instances using a standard laptop: a MacBook Pro with a 2.9 GHz Intel Core i7 processor and 16GB of RAM, running macOS 10.12.6. Gurobi 8.1.1 was used to solve all LPs on this platform, and no other computationally intensive programs were running while the experiments were being conducted. All experiments with the 48-hour scheduling horizon instances were conducted on a 64-core Linux workstation. The shift in machine was necessary, as there was insufficient memory on the laptop to execute the EF method. As all the models and algorithms were implemented in Python, for both the EF and our proposed Benders decomposition approach we report the total wall clock time expended solving LPs by Gurobi.

6.3.2. Computational Results

We now present and analyze summary computational results across our three sets of test instances. In Tables 1–3 under the “Benders” section, “Cut Time (s)” refers to the total time (in seconds) expended solving cut-generation subproblems, “Cuts Added” refers to the total number of cuts added to the master problem, “Iterations” refers to the total number of Benders decomposition iterations, and “Total Time (s)” refers to the cumulative Gurobi LP solve time (in seconds) for both the subproblems and the initial master problem solve and updates. The label “Time (s)” under the “EF”

Table 1: Summary computational results for computing CHPs on the (modified) RTS-GMLC set of generators, 24 hour instances.

	Benders				EF
	Cut Time (s)	Cuts Added	Iterations	Total Time (s)	Time (s)
min	0.01	0	1	0.37	17.13
mean	0.10	2.79	2.62	0.74	22.28
max	0.30	12	7	1.26	39.90

Table 2: Summary computational results for computing CHPs on the CAISO set of generators, 24 hour instances.

	Benders				EF
	Cut Time (s)	Cuts Added	Iterations	Total Time (s)	Time (s)
min	0.00	0	1	1.75	17.58
mean	0.05	0.15	1.15	2.70	24.23
max	0.19	1	2	3.87	35.08

Table 3: Summary computational results for computing CHPs on both FERC sets of generators, 24 hour instances.

	Benders				EF
	Cut Time (s)	Cuts Added	Iterations	Total Time (s)	Time (s)
min	0.74	20	6	13.55	733.4
mean	3.11	115.79	14.20	20.55	1553.8
max	13.27	577	47	40.62	2932.2

section refers to the total Gurobi LP solve time (in seconds) for the EF. For each of these statistics, we report the minimum (min), arithmetic mean (mean), and maximum (max) across the instances in each test set for each category.

Table 1 reports summary results for the RTS-GMLC system. The results indicate that for this set of instances the Benders decomposition approach is, on average and in the worst case, ~ 30 times faster than the EF and only requires a modest number of cuts. Table 2 reports the summary results for the CAISO system. Here, the performance difference is more modest, with Benders decomposition being only ~ 10 times faster than the EF. Because of

Table 4: Summary computational results for computing CHPs on the (modified) RTS-GMLC set of generators, 48 hour instances.

	Benders			EF	
	Cut Time (s)	Cuts Added	Iterations	Total Time (s)	Time (s)
min	0.80	0	1	1.85	369.9
mean	4.27	5.29	4.0	6.59	473.7
max	18.49	24	16	21.34	645.5

the small percentage ($\sim 5\%$) of ramping-constrained generators, the Benders decomposition approach typically does not generate any cuts, certifying that those ramping-constrained generators are already in their respective convex hulls.

The summary results for the FERC system, in Table 3, show the most extreme differences. Recall that for these instances ramping-constrained generators comprise approximately half of the fleet. Here, our proposed Benders decomposition approach is ~ 70 times faster in both average and worst cases. For the EF approach, 7 of the 24 instances required more than a half hour to solve the LP, which is not a feasible time-scale for ISOs to implement CHP. Conversely, even with this many ramping-constrained generators, the worst case performance for the proposed Benders decomposition approach is less than a minute, and the process could be terminated with some solution (i.e., prices) after any iteration. In contrast, it is much more difficult to construct a pricing solution from a partial LP solve result.

6.4. 48-Hour Instances

6.4.1. Computational Platform

To test the scalability of our Benders decomposition approach, we next analyze results for computing CHPs in the context of 48-hour scheduling horizons. As indicated earlier, these problems could not be solved on our laptop in the case of the EF due to large memory requirements. Consequently, we shifted our experimental platform to a 64-core Linux workstation with Intel Xeon 2.3GHz processors and 512GB of RAM. Gurobi 8.0.0 was used to solve all LPs on this platform. As with the 24-hour instances, we report the cumulative time spent on solving the LPs by Gurobi.

Table 5: Summary computational results for computing CHPs on the CAISO set of generators, 48 hour instances.

	Benders				EF
	Cut Time (s)	Cuts Added	Iterations	Total Time (s)	Time (s)
min	0.00	0	1	5.91	317.3
mean	1.31	0.0	1.0	11.55	537.5
max	4.72	0	1	17.42	745.5

Table 6: Summary computational results for computing CHPs on both FERC sets of generators, 48 hour instances.

	Benders				EF
	Cut Time (s)	Cuts Added	Iterations	Total Time (s)	Time (s)
min	13.5	42	5	61.1	*
mean	93.4	184.6	20.2	165.0	*
max	480.5	1084	86	658.1	*

6.4.2. Computational Results

The computational results reported in Tables 4–6 are similar to those in Tables 1–3. For the RTS-GMLC instances (Table 4), our Benders decomposition approach is, on average, ~ 70 times faster than the EF approach, and is still ~ 30 times faster in the worst case. For the CAISO instances (Table 5), our Benders decomposition approach is ~ 40 times faster in the average and worst cases. Interestingly, in all 20 of these cases, the formulation \mathcal{R}^* was tight enough such that no Benders cuts needed to be added to the master problem. In a few of the cases, all of the ramping-constrained generators had integer solutions in the master problem, resulting in a minimum cut time of 0.0.

Results for the 48-hour FERC instances are reported in Table 6. With half of the fleet comprising ramping-constrained generators, both approaches find these instances to be the most difficult. For our Benders decomposition approach, the worst case requires just under 11 minutes, with a majority of time expended solving cut generation subproblems. As both the EF and individual cut generation subproblems are highly degenerate (hence the importance of using an interior point method for both), a possible research path is mitigating this degeneracy to improve performance. In contrast to

our Benders decomposition approach, the EF formulation could not solve any instance within an hour – as indicated by the “*” in every field. The EFs for these instances have over 70 million variables and constraints, and over 275 million non-zeros in the constraint matrices. The pre-solved models still have over 150 million non-zeros, and typically barrier does not even initiate within the hour limit; most or all of the time is spent on pre-solve. Hence, while the EF method is straightforward to implement, for a problem with significant ramping-constrained units, it still may be larger than a modern commercial LP code can compute in a reasonable amount of time.

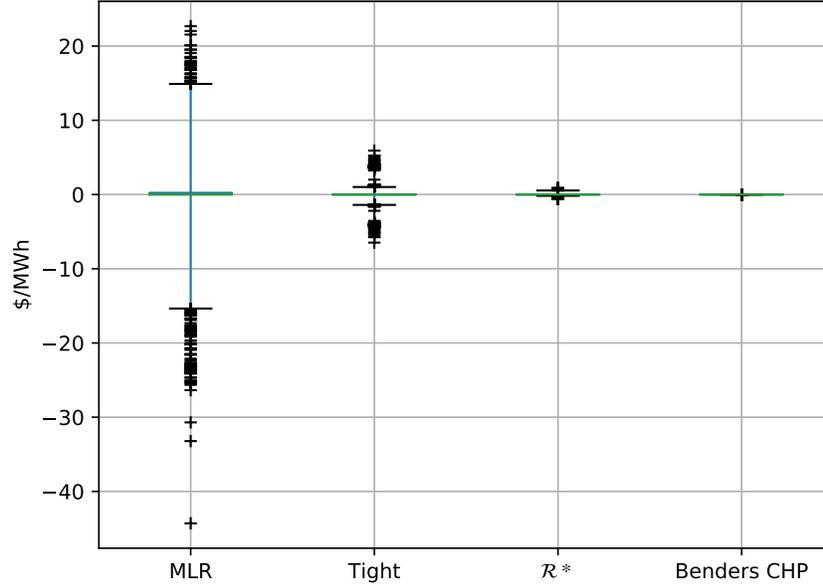
7. Analysis of Approximate Convex Hull Prices

In the section, we present some analysis of different “approximate” convex hull pricing schemes based on (LPR). The first, denoted “MLR”, sets \mathcal{R} to the unit commitment formulation presented in [35]. The second, denoted “Tight”, uses the Tight formulation from [21]. We compare these approximations to the Benders master problem (with no cuts), denoted as \mathcal{R}^* , and the formulation resulting from the Benders procedure applied to \mathcal{R}^* in Section 5. In comparing performance, we measure the difference in shadow price between these formulations, respectively, and the EF computation of convex hull prices. Note that even under Theorem 3, the proposed Benders approach for CHPP and the EF approach for CHPP method may yield different prices in the presence of dual degeneracy, though we did not observe this in any of the test instances. In every case we just consider the 24-hour variant of each test instance.

7.1. RTS-GMLC

In Figure 1 we detail the energy prices against those found by the EF approach to CHPP for the RTS-GMLC instances. First, we notice that the proposed Benders approach produces the exact same prices as the EF approach, giving some empirical evidence for the validity of Theorem 3. We also observe that tighter approximations of $\text{conv}(\Pi^g)$, in general, result in prices that are closer to CHPs. However, none of the formulations examined are biased in a particular way—most values are tightly clustered around CHPs. That said, the “MLR” formulation has differences which could be considered significant within 99% of the observations, often being \$10/MWh different than convex hull prices. The Tight and \mathcal{R}^* formulations are better approximators: on these instances \mathcal{R}^* is never more \$1/MWh different than the

Figure 1: Difference in energy prices in every hour against the EF approach across all 24-hour RTS-GMLC instances. Whiskers are drawn to cover 99% of observations.



computed CHPs, and Tight is never \$7/MWh different than the computed CHPs.

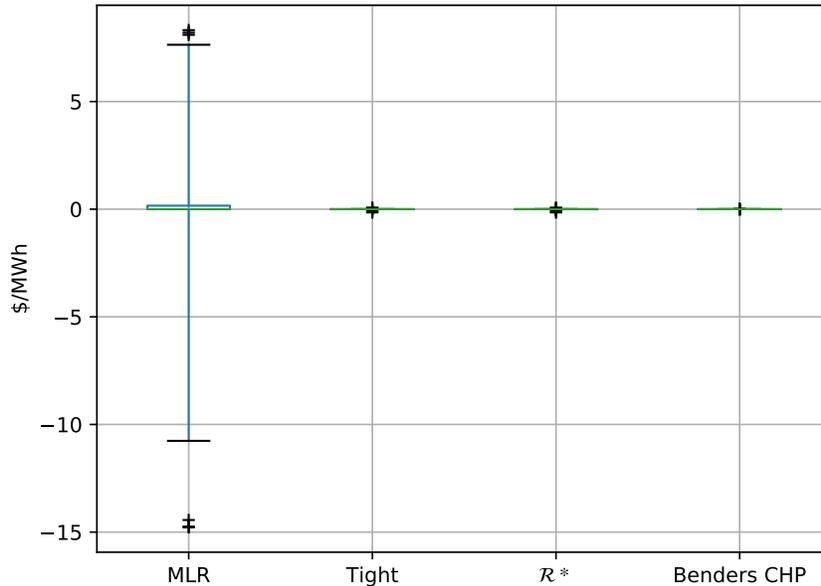
7.2. CAISO

Figure 2 is similar to Figure 1, and details the differences in energy prices against those found by the EF for the CAISO instances. We see a similar pattern to from the RTS-GMLC instances, where again the MLR formulation does a poor job overall approximating CHPs, but the Tight and \mathcal{R}^* formulations are very good approximators. This can be explained by the tight startup cost formulation shared by Tight and \mathcal{R}^* , as compared to that in MLR (see [30] for more details), and because the CAISO instances have very few generators with irredundant hourly ramping constraints. Hence, for most generators, the Tight formulation already ensures that most generators are within their convex hulls at LP optimum.

7.3. FERC

Again, Figure 3 is analogous to Figures 1 and 2, only examining pricing variation across the FERC instances. Unlike the RTS-GMLC and CAISO

Figure 2: Difference in energy prices in every hour against the EF approach across all 24-hour CAISO instances. Whiskers are drawn to cover 99% of observations.

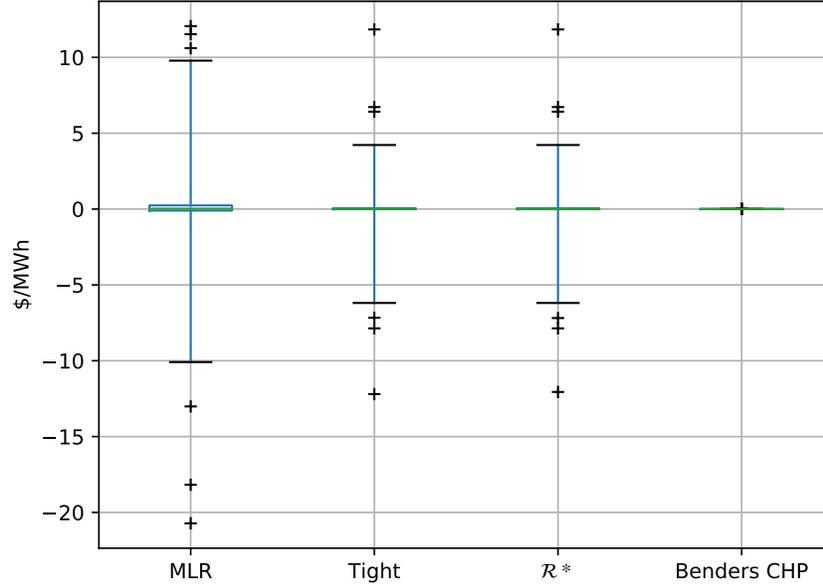


instances, Tight and \mathcal{R}^* do not provide significantly more improvement as CHP approximators over MLR. Approximately half of the generators in the FERC system have irredundant hourly ramping constraints, and further, the system is less flexible than RTS-GMLC or CAISO, which negates the importance of the startup cost formulation, especially at a 24-hour time horizon. While Tight and \mathcal{R}^* perform better as approximators than MLR, both have observed worst-case differences of \$12/MWh.

7.4. Summary Results

Overall then, we can conclude that while aCHPs based on \mathcal{R}^* may be “good” approximators for CHPs, especially for systems with mostly flexible generation, having a significant amount of ramping-constrained generation can weaken their approximation value. Further, one would expect similar differences for any market item (e.g., reserves) for which the convex hull of the generators is not well-understood or is not well-approximated by the typically compact unit commitment models. As such, being able to sufficiently formulate and compute the convex hull, as we demonstrate in this work,

Figure 3: Difference in energy prices in every hour against the EF approach across all 24-hour FERC instances. Whiskers are drawn to cover 99% of observations.



is essential for improved pricing schemes. The proposed Benders CHP is guaranteed to outperform the aCHP approaches, and as a result could be executed for exact convex hull pricing schemes or even improved approximated convex hull pricing schemes in which the problem is not solved to completion due to runtime limitations.

8. Conclusion

In summary, we have explored the theory and practice needed to efficiently compute convex hull prices for large-scale market clearing problems. To maintain scope for this particular study, we have not explored the complete implications of implementing a convex hull pricing scheme, though we have demonstrated that such an implementation is presently possible within time limits required by system operators. Additionally, we demonstrated empirically that “approximated” convex hull prices can be both good and bad approximators for CHPs, depending on the generator parameters and the unit commitment relaxed they are based on. Further, our implementation

could be refined: the Benders subproblems could be computed in parallel, and additional research should be conducted to reduced the subproblem solve times for longer time horizons.

Overall, our results conclusively indicate that computational intractability is no longer a barrier for the implementation of convex hull prices in real-world market clearing contexts.

References

- [1] P. N. Biskas, D. I. Chatzigiannis, A. G. Bakirtzis, European electricity market integration with mixed market designs—part i: Formulation, *IEEE Transactions on Power Systems* 29 (1) (2013) 458–465.
- [2] N. E. Koltsaklis, A. S. Dagoumas, Incorporating unit commitment aspects to the european electricity markets algorithm: An optimization model for the joint clearing of energy and reserve markets, *Applied energy* 231 (2018) 235–258.
- [3] R. Sioshansi, R. O’Neill, S. S. Oren, Economic consequences of alternative solution methods for centralized unit commitment in day-ahead electricity markets, *IEEE Trans. Power Syst.* 23 (2) (2008) 344–352.
- [4] B. Eldridge, R. O’Neill, B. F. Hobbs, Near-optimal scheduling in day-ahead markets: Pricing models and payment redistribution bounds, *IEEE Trans. Power Syst.* (2019).
- [5] P. R. Gribik, W. W. Hogan, S. L. Pope, Market-clearing electricity prices and energy uplift (2007).
URL sites.hks.harvard.edu/fs/whogan/Gribik_Hogan_Pope_Price_Uplift_123107.pdf
- [6] R. B. Johnson, S. S. Oren, A. J. Svoboda, et al., Equity and efficiency of unit commitment in competitive electricity markets, *Utilities Policy* 6 (1997) 9–20.
- [7] W. W. Hogan, B. J. Ring, On minimum-uplift pricing for electricity markets, *Electricity Policy Group* (2003).
- [8] R. P. O’Neill, P. M. Sotkiewicz, B. F. Hobbs, M. H. Rothkopf, W. R. Stewart Jr, Efficient market-clearing prices in markets with nonconvexities, *European journal of operational research* 164 (1) (2005) 269–285.

- [9] M. Cadwalader, P. Gribik, W. Hogan, S. Pope, Extended LMP and financial transmission rights, Harvard Univ., Cambridge, MA, USA, working paper (2010).
- [10] D. A. Schiro, T. Zheng, F. Zhao, E. Litvinov, Convex hull pricing in electricity markets: Formulation, analysis, and implementation challenges, *IEEE Trans. Power Syst.* 31 (5) (2016) 4068–4075.
- [11] R. B. Hytowitz, R. P. O’Neill, B. Eldridge, A. Castillo, Multi-period dual pricing algorithm in non-convex ISO markets, Tech. rep., Working Paper, Johns Hopkins University, Baltimore, MD (2017).
- [12] B. Hua, R. Baldick, A convex primal formulation for convex hull pricing, *IEEE Trans. Power Syst.* 32 (5) (2017) 3814–3823.
- [13] Y. Yu, Y. Guan, Y. Chen, An extended integral unit commitment formulation and an iterative algorithm for convex hull pricing, *IEEE Trans. Power Syst.* (2020).
- [14] C. Alvarez, F. Mancilla-David, P. Escalona, A. Angulo, A Bienstock-Zuckerberg-based algorithm for solving a network-flow formulation of the convex hull pricing problem, *IEEE Trans. Power Syst.* (2019).
- [15] Z. Yang, T. Zheng, J. Yu, K. Xie, A unified approach to pricing under nonconvexity, *IEEE Transactions on Power Systems* 34 (5) (2019) 3417–3427.
- [16] J. F. Benders, Partitioning procedures for solving mixed-variables programming problems, *Numerische mathematik* 4 (1) (1962) 238–252.
- [17] A. Geoffrion, Lagrangean relaxation for integer programming, *Math. Program. Stud.* 2 (1974) 82–114.
- [18] B. Knueven, J. Ostrowski, J. Wang, The ramping polytope and cut generation for the unit commitment problem, *INFORMS Journal on Computing* 30 (4) (2018) 739–749.
- [19] T. Zheng, F. Zhao, D. Schiro, E. Litvinov, The hidden properties of fast start pricing, *Increasing Market and Planning Efficiency and Enhancing Resilience through Improved Software* (June 2018).

- [20] J. Mays, D. Morton, R. P. O’Neill, Investment effects of pricing schemes for non-convex markets (2018).
- [21] B. Knueven, J. Ostrowski, J.-P. Watson, On mixed integer programming formulations for the unit commitment problem, *INFORMS Journal on Computing* (2019) to appear.
- [22] E. Balas, Disjunctive programming: Properties of the convex hull of feasible points, *Discrete Appl Math.* 89 (1) (1998) 3–44.
- [23] C. Gentile, G. Morales-Espana, A. Ramos, A tight MIP formulation of the unit commitment problem with start-up and shut-down constraints, *EURO Journal on Computational Optimization* 5 (1–2) (2017) 177–201.
- [24] Y. Pochet, L. A. Wolsey, Lot-sizing with constant batches: Formulation and valid inequalities, *Mathematics of Operations Research* 18 (4) (1993) 767–785.
- [25] T. L. Magnanti, L. A. Wolsey, Optimal trees, *Handbooks in Ops. Res. and Mgmt. Sci.* 7 (1995) 503–615.
- [26] Y. Faenza, G. Oriolo, G. Stauffer, The hidden matching structure of the composition of strips: a polyhedral perspective, in: *14th Aussois Workshop on Combinatorial Optimization*, 2010.
- [27] Y. Pochet, L. A. Wolsey, *Production planning by mixed integer programming*, Springer Science & Business Media, 2006.
- [28] A. Frangioni, C. Gentile, New MIP formulations for the single-unit commitment problems with ramping constraints, *IASI Research Report* 15-06 (2015).
- [29] Y. Guan, K. Pan, K. Zhou, Polynomial time algorithms and extended formulations for unit commitment problems, *IIE Transactions* 50 (8) (2018) 735–751.
- [30] B. Knueven, J. Ostrowski, J.-P. Watson, A novel matching formulation for startup costs in unit commitment, *Mathematical Programming Computation* (2020) 1–24.
- [31] B. Knueven, J. Ostrowski, J.-P. Watson, Online supplement for a novel matching formulation for startup costs in unit commitment (2018).

- [32] C. Barrows, A. Bloom, A. Ehlen, J. Ikaheimo, J. Jorgenson, D. Krishnamurthy, J. Lau, B. McBennett, M. O’Connell, E. Preston, et al., The ieeereliability test system: A proposed 2019 update, *IEEE Trans. Power Syst.* (2019).
- [33] RTS-GMLC: Reliability test system - grid modernization lab consortium (August 2018).
URL github.com/GridMod/RTS-GMLC
- [34] Gurobi Optimization, Inc., Gurobi optimizer reference manual (2019).
URL www.gurobi.com
- [35] G. Morales-España, J. M. Latorre, A. Ramos, Tight and compact MILP formulation for the thermal unit commitment problem, *IEEE Trans. Power Syst.* 28 (4) (2013) 4897–4908.