Exploiting Identical Generators in Unit Commitment

Ben Knueven and Jim Ostrowski

Department of Industrial and Systems Engineering
University of Tennessee, Knoxville, TN 37996
bknueven@vols.utk.edu jostrows@utk.edu

Jean-Paul Watson

Discrete Math and Optimization
Sandia National Laboratories, Albuquerque, NM 87185
jwatson@sandia.gov

June 30, 2017

Abstract

We present sufficient conditions under which thermal generators can be aggregated in mixed-integer linear programming (MILP) formulations of the unit commitment (UC) problem, while maintaining feasibility and optimality for the original disaggregated problem. Aggregating thermal generators with identical characteristics (e.g., minimum/maximum power output, minimum up/down-time, and cost curves) into a single unit reduces redundancy in the search space caused by both exact symmetry (permutations of generator schedules) and certain classes of mutually non-dominated solutions. We study the impact of aggregation on two large-scale UC instances, one from the academic literature and another based on real-world data. Our computational tests demonstrate that when present, identical generators can negatively affect the performance of modern MILP solvers on UC formulations. Further, we show that our reformation of the UC MILP through aggregation is an effective method for mitigating this source of difficulty.

Nomenclature

Indices and Sets

$g \in \mathcal{G}$	Thermal generators.
$t \in \mathcal{T}$	Hourly time steps: $1, \ldots, \mathbf{T}$.
$[t, t') \in \mathcal{Y}^g$	Feasible intervals of operation for generator g with respect to its minimum uptime, that is, $[t, t') \in \mathcal{T} \times \mathcal{T}$ such that $t' \geq t + \mathbf{UT}^g$.

Parameters

$\mathbf{D}_t$	Load (demand) at time t (MW).
$\mathbf{DT}^g$	Minimum down time for generator g (h).
$\overline{\mathbf{P}}^g$	Maximum power output for generator g (MW).
$\underline{\mathbf{P}}^g$	Minimum power output for generator g (MW).
$\mathbf{RD}^g$	Ramp-down rate for generator g (MW/h).
$\mathbf{RU}^g$	Ramp-up rate for generator g (MW/h).
$\mathbf{SD}^g$	Shutdown ramp rate for generator g (MW/h).
$\mathbf{SU}^g$	Startup ramp rate for generator g (MW/h).
$\mathbf{TC}^g$	Time down after which generator g goes cold (h).
$\mathbf{UT}^g$	Minimum up time for generator g (h).

Variables

p_t^g	Power above minimum for generator g at time t (MW).
u_t^g	Commitment status of generator g at time t, $\in \{0, 1\}$.
v_t^g	Startup status of generator g at time t, $\in \{0, 1\}$.
w_t^g	Shutdown status of generator g at time t, $\in \{0, 1\}$.
$y_{[t,t')}^g$	Indicator arc for startup at time t, shutdown at time t', committed for $i \in [t, t')$, for generator g, $\in \{0, 1\}$, $[t, t') \in \mathcal{Y}^g$.

1 Introduction

Unit commitment (UC) is a core optimization problem in power systems operations, in which the objective is to determine an on/off schedule for thermal generating units that minimizes production costs while satisfying constraints related to generator performance characteristics and power flow physics [1]. UC is now widely modeled as a mixed-integer linear program (MILP) and solved using commercial branch-and-cut technologies, e.g., such as those available in the Gurobi [2] and CPLEX [3] software packages. Due to its criticality, significant research has been dedicated over the past 15 years toward improving the quality of MILP formulations of UC, specifically focusing on the strength of the associated LP relaxation – as this is strongly correlated with computational difficulty. Recent examples of the progress in state-of-the-art of UC formulations are reported in [4–8].

Most of the UC research to date has focused on the analysis of generator ramping and startup cost polytopes. An alternative and orthogonal approach, however, considers the impact of symmetry in the UC MILP model induced by the presence of multiple generators with identical physical characteristics and other properties. Here, we consider two generators to be identical if they have identical performance parameters and cost coefficients. We will show that initial status can be ignored for our purposes. As we discuss subsequently, such identical generators are found in both academic UC instances and those based on real-world data. Further, developing an understanding of exact symmetry is a necessary first step toward developing formulations that consider partial symmetry, which is more pervasive in practice.

Aggregating thermal generators with identical characteristics into a single unit reduces redundancy in the search space caused by both exact symmetry (permutations of generator schedules) and certain classes of mutually non-dominated solutions – both of which can cause performance issues when commercial MILP solvers are employed to solve UC [9, 10]. Further, degeneracy in the solution space induced by symmetry is known to cause convergence difficulties for decomposition-based solvers for stochastic UC, e.g., see [11].

In the context of exact MILP-based solution methods for UC, one approach to addressing identical generator symmetry is through the introduction of advanced branching strategies when exploring the branch-and-bound search tree. For example, [9] introduces "modified orbital branching," which strengthens orbital branching for cases when the problem's symmetry group contains additional structure, as is the case with UC [12]. While such an approach has promise, its implementation is non-trivial, and when using CPLEX callbacks many of the solver's advanced features are disabled, which can result in slower solve-times overall. Further, the approach is not possible to implement in Gurobi as the callback interface does not allow the user to access or decide on a branching variable [2].

Another approach to addressing the presence of symmetry in UC MILP formulations involves the introduction of symmetry-breaking inequalities. For example, [10] adds static symmetry-breaking inequalities to a UC MILP formulation. Their approach eliminates some, but not all of, of the redundancy in the branch-and-bound tree induced by symmetry, leading to faster computational times overall for highly symmetric instances. However, symmetry-breaking inequalities have the disadvantage that they increase the size of the LP relaxation. We compare their inequalities to the method we propose in this paper in the appendix.

In this paper we propose addressing identical generators in UC MILP formulations through a novel aggregation approach. While this idea is certainly not new (e.g., see [14–17]), we introduce conditions under which such an aggregation can be done exactly – that is, after simple post-processing we have a provably optimal solution to the original disaggregated problem. Our results are a consequence of recent advances in convex hull formulations for commitment of a single generator [6,8,18–20]. We provide conditions under which "orbital shrinking" [21] can be done exactly for UC. Finally, we show that our approach can significantly reduce the solve times required for large-scale UC instances, considering a state-of-the-art UC MILP formulation.

The remainder of this paper is organized as follows. In Section 2 we review the UC problem, and demonstrate how naive aggregation can result in an infeasible solution to the original disaggregated UC problem. Section 3 explores the sufficient conditions for feasible and optimal disaggregation and introduces simple algorithms for disaggregation. These base results are expanded for more cases in the appendix. Section 4 considers how the presence of multiple identical generators can result in symmetric and non-symmetric solutions with the same objective function value, and how aggregation models can express these solutions concurrently. In Section 5 we present computational results for two sets of large-scale UC instances, one from the academic literature and another based on real-world data. We then conclude in Section 6 with a summary of our contributions.

2 Unit Commitment Formulation

UC MILP formulations reported in the literature are typically expressed in the general form

$$\min \sum_{g \in G} \mathbf{c}^g(p^g) \tag{1a}$$

subject to

$$\sum_{g \in \mathcal{G}} p_t^g = \mathbf{D}_t \qquad \forall t \in \mathcal{T} \tag{1b}$$

$$p^g \in \Pi^g \qquad \forall g \in \mathcal{G} \tag{1c}$$

where $\mathbf{c}^g(p^g)$ denotes the costs associated with thermal generator g producing a (vector) output of p^g over the scheduling horizon and Π^g denotes the set of feasible schedules for generator g. When it is clear from context that we are referencing a single generator, we drop the superscript g on parameters and variables. A significant portion of the UC literature focuses on how to model the feasible sets Π^g, where there is a trade-off between the model size versus the tightness of the formulation.

For the purposes of exposition we will focus on the simplified version of UC above, but in the appendix we go into more detail, including reserves, piecewise linear operating costs, and time-dependent startup costs.

The ability to aggregate identical generators is dependent on which UC MILP formulation is considered. For instance, consider the basic "3-bin" formulation for a generator [4, 22]:

$$
\begin{aligned}
&\underline{\mathbf{P}} u_t \le p_t \le \overline{\mathbf{P}} u_t, && \forall t \in \mathcal{T} && \text{(2a)}\\
&p_t - p_{t-1} \le \mathbf{RU} u_{t-1} + \mathbf{SU} v_t, && \forall t \in \mathcal{T} && \text{(2b)}\\
&p_{t-1} - p_t \le \mathbf{RD} u_t + \mathbf{SD} w_t, && \forall t \in \mathcal{T} && \text{(2c)}\\
&u_t - u_{t-1} = v_t - w_t, && \forall t \in \mathcal{T} && \text{(2d)}\\
&\sum_{i=t-\mathbf{UT}+1}^{t} v_i \le u_t, && \forall t \in [\mathbf{UT}, \mathbf{T}] && \text{(2e)}\\
&\sum_{i=t-\mathbf{DT}+1}^{t} w_i \le 1 - u_t, && \forall t \in [\mathbf{DT}, \mathbf{T}] && \text{(2f)}\\
&p_t \in \mathbb{R}_+,\ u_t,\ v_t,\ w_t \in \{0,1\} && \forall t \in \mathcal{T} && \text{(2g)}
\end{aligned}
$$

where Constraints (2a) enforce minimum / maximum generator output, Constraints (2b, 2c) enforce ramping limits, Constraints (2d) enforce logical constraints on u, v, and w, and Constraints (2e, 2f) enforce minimum up / down times.

Consider the case of two generators with identical performance and cost parameters. We can always model this situation by treating each generator individually. However, it would be desirable if we could aggregate the generators to exploit the symmetry structure. Unfortunately, the above model does not present a straightforward way to accomplish this. Consider a 5 time period case where two generators each have $\underline{\mathbf{P}} = \mathbf{SU} = \mathbf{SD} = 100$, $\overline{\mathbf{P}} = 200$, and $\mathbf{RU} = \mathbf{RD} = 50$. Ideally, we would prefer to let the u, v, and w variables represent how many of the generators remain on, are turned on, and are turned off at a given time. Suppose, then, that we use formulation (2) but allow u_t, v_t, and w_t variables to take values in $\{0, 1, 2\}$. Then, consider the following feasible solution to this simple aggregated UC model:

$$
U = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \; P = \begin{pmatrix} 200 \\ 300 \\ 400 \\ 300 \\ 200 \end{pmatrix}. \tag{3}
$$

The solution to the aggregated model is clearly not feasible in the disaggregated model, as both generators must produce at full capacity in time period 3 and the generator that started up at time 2 cannot ramp to full capacity by time period 3. The problem with this naive aggregated formulation is that when one generator is operating at full capacity, the other is able to “steal” its ramping capability. Unfortunately, tighter descriptions in the 3-bin space do not overcome this problem.

Some UC MILP formulations do allow for variable aggregation without violating the structure of the base problem. For example, consider the formulation

$$
\begin{aligned}
& \mathbf{A}^{[a,b)} p^{[a,b)} \leq \mathbf{b}^{[a,b)} y_{[a,b)} && \forall [a,b) \in \mathcal{Y} && \text{(4a)} \\
& \sum_{[a,b) \in \mathcal{Y}} p_t^{[a,b)} = p_t && \forall t \in \mathcal{T} && \text{(4b)} \\
& \sum_{\{[a,b) \in \mathcal{Y} \mid t \in [a, b+\mathbf{DT})\}} y_{[a,b)} \leq 1 && \forall t \in \mathcal{T} && \text{(4c)} \\
& y_{[a,b)} \in \{0,1\} && \forall [a,b) \in \mathcal{Y}, && \text{(4d)}
\end{aligned}
$$

where the polytope $\{p^{[a,b)} \in \mathbb{R}_+^T \mid \mathbf{A}^{[a,b)} p^{[a,b)} \leq \mathbf{b}^{[a,b)}\}$ describes the feasible production of the generator if it is turned on at time a, turned off at time b, and consistently on during the interval $[a, b)$. Constraints (4a) enforce the appropriate ramping and minimum/maximum power output given that the generator is on during the $[a, b)$ time interval. Constraints (4c) ensure that minimum up and downtime constraints are met. While the above formulation is large (it contains $O(|\mathcal{T}|^3)$ many variables and constraints), it is provably tight [8].

As observed in [8], Constraints (4c) correspond to clique inequalities for an interval graph. Because interval graphs are totally unimodular [23], it follows that the matrix given by Constraints (4c) is totally unimodular, and thus has the integer decomposition property [24]. Therefore, we can model k-many identical generators by letting the Y variables be general integers in the range $[0, k]$ and rewriting Constraints (4c) as

$$
\sum_{\{[a,b) \in \mathcal{Y} \mid t \in [a, b+\mathbf{DT})\}} Y_{[a,b)} \leq k \qquad \forall t \in \mathcal{T}. \tag{5}
$$

In this context, $Y_{[a,b)}$ represents *how many* of the generators are on during the interval $[a, b)$.[1] Because there are separate power variables for each on-interval $[a, b)$, this formulation overcomes the problems seen in (3).

Note that the size of the above formulation is heavily dependent on generator minimum up and down times. For generators with small minimum up and down times, the above formulation is very large, so much so that the benefits of a tight model are outweighed by model size. However, for generators with moderate minimum up and downtimes (say 8 hours), the above model is quite tractable.

Given that we see many generators with small minimum up and downtimes (say 2 hours each), it is worthwhile to ask *when* the 3-bin model is decomposable. As the example in (3) suggests, the problem is with the ramping. Consider the traditional 3-bin formulation for generators with redundant ramping

[1] We use capital letters to represent aggregated variables.

constraints, when $\mathbf{UT} \geq 2$ [5]:

$$\underline{\mathbf{P}} u_t \leq p_t, \qquad \forall t \in \mathcal{T} \tag{6a}$$

$$p_t \leq \overline{\mathbf{P}} u_t + (\mathbf{SU} - \overline{\mathbf{P}}) v_t + (\mathbf{SD} - \overline{\mathbf{P}}) w_{t+1} \qquad \forall t \in \mathcal{T} \tag{6b}$$

$$u_t - u_{t-1} = v_t - w_t \qquad \forall t \in \mathcal{T} \tag{6c}$$

$$\sum_{i=t-\mathbf{UT}+1}^{t} v_i \leq u_t \qquad \forall t \in [\mathbf{UT}, \mathbf{T}] \tag{6d}$$

$$\sum_{i=t-\mathbf{DT}+1}^{t} w_i \leq 1 - u_t \qquad \forall t \in [\mathbf{DT}, \mathbf{T}] \tag{6e}$$

$$p_t \in \mathbb{R}_+,\ u_t,\ v_t,\ w_t \in \{0, 1\} \qquad \forall t \in \mathcal{T}. \tag{6f}$$

This formulation has the property that the constraint matrix defined by (6c, 6d, 6e) is totally unimodular [18], and so it too has the integer decomposition property [24]. We discuss the case when $\mathbf{UT} = 1$ in the appendix for clarity of exposition here.

In any case, the total unimodularity of Constraints (4c) and (6c, 6d, 6e) only ensure that the on-off schedules can be decomposed, and tell us nothing about whether (and if so, how) the aggregated power output can be disaggregated into a feasible production schedule for each generator in the aggregation. In the following section we explore these issues.

3 Disaggregating Solutions

While the results from [24] give conditions for when a UC schedule can be decomposed, it does not suggest a constructive approach to performing the decomposition. We now outline how to decompose solutions to the aggregate UC formulation into individual generator schedules. We first provide theorems regarding the relationship between schedules and the power output of identical generators.

Theorem 1. *Consider identical generators $g_1, g_2 \in \mathcal{G}$ on at time t, and assume their production costs are increasing and convex. If possible, $p_t^{g_1} = p_t^{g_2}$. If not, then least one of the nominal or startup/shutdown ramping constraints is binding for generator g_1 or g_2 at time t in an optimal solution.*

Proof. By contradiction, consider an optimal schedule in which $p_t^{g_1} > p_t^{g_2}$ and with no binding nominal or startup/shutdown ramping constraints at time t. Then, we may decrease $p_t^{g_1}$ by epsilon and increase $p_t^{g_2}$ by epsilon without affecting impacting feasibility. Furthermore, because production costs are increasing and convex, this new solution is no worse than the original. □

From Theorem 1, if two identical generators start up at time a and shut down at time b, then their power outputs in the interval $[a, b)$ are identical. Theorem 1 also applies to fast-ramping generators, i.e., generators that are not ramp-limited, as the lack of ramping constraints ensures if two identical generators are on in a given time period, then they must have the same power output. This result suggests that allowing u to be a general integer is also sufficient. The only exception to this rule is when there is a binding startup/shutdown rate.

Theorem 2. *Suppose generator g_1 is turned off at time t. If identical generator g_2 can also be turned off at time t, there exists an optimal solution where the generator that has been on for the least amount of time is turned off.*

Proof. Suppose identical $g_1, g_2 \in \mathcal{G}$ have been on at time t for at least $\mathbf{UT}(= \mathbf{UT}^{g_1} = \mathbf{UT}^{g_2})$ time periods, starting at time t_0, and that generator g_1 has been on longer. If generator g_1 is turned off at time t, then there exists a t' with $t_0 \leq t' \leq t$ such that $p^{g_2}_{t'-1} \leq p^{g_1}_{t'-1}$ and $p^{g_1}_{t'} < p^{g_2}_{t'}$. Notice that permuting g_1 and g_2 for all $t \geq t'$ does not affect the objective value, and does not change the power output. Finally, the permuted solution satisfies the ramping constraints since $p^{g_1}_{t'} - p^{g_2}_{t'-1} < p^{g_2}_{t'} - p^{g_2}_{t'-1}$ and $p^{g_2}_{t'} - p^{g_1}_{t'-1} \leq p^{g_2}_{t'} - p^{g_2}_{t'-1}$, satisfying ramp up, and $p^{g_1}_{t'-1} - p^{g_2}_{t'} < p^{g_1}_{t'-1} - p^{g_1}_{t'}$ and $p^{g_2}_{t'-1} - p^{g_1}_{t'} \leq p^{g_1}_{t'-1} - p^{g_1}_{t'}$, satisfying ramp down. □

Theorem 3. *Suppose generator g_1 is turned on at time t. If an identical generator g_2 can also be turned on at time t, and there are no time-dependent startup costs for g_1 and g_2, then there exists an optimal solution where g_2 is turned on at t.*

Proof. Suppose identical $g_1, g_2 \in \mathcal{G}$ at time t have been off for at least $\mathbf{DT}(= \mathbf{DT}^{g_1} = \mathbf{DT}^{g_2})$ time periods. Then for all $t' \in \mathcal{T}$ such that $t' \geq t$, we can permute the schedules for g_1 and g_2 without changing the objective value or the power output, in which case g_2 is turned on at time t. □

Theorems 2 and 3 illustrate how we can interchange parts of a UC schedule that involve identical generators, and not lose optimality. We explore this issue further in Section 4. The implications of the above theorems also provide some strong direction for the characteristics a disaggregation method of aggregated UC schedules that maintains both feasibility and optimality. Before formalizing this procedure, we first introduce some additional notation. Suppose $\mathcal{K} \subset \mathcal{G}$ such that all generators in $\mathcal{K}$ have identical properties, except for initial status. We again use capital letters to represent aggregated variables, and the superscript $\mathcal{K}$ to represent the parameters shared among the generators (e.g., $\mathbf{DT}^{\mathcal{K}} = \mathbf{DT}^g$, $\mathbf{SD}^{\mathcal{K}} = \mathbf{SD}^g$, etc., for $g \in \mathcal{K}$). In this context, we now illustrate how to decompose schedules for both the extended formulation and the 3-bin formulation for fast-ramping generators.

3.1 Extended Formulation (EF)

Let $Y = \sum_{g\in\mathcal{K}} y^g$, $P = \sum_{g\in\mathcal{K}} p^g$, and $P^{[a,b)} = \sum_{g\in\mathcal{K}} p^{g,[a,b)}$ $\forall [a,b) \in \mathcal{Y}^{\mathcal{K}}$. Consider the aggregated extended formulation for the generators in $\mathcal{K}$:

$$\mathbf{A}^{[a,b)} P^{[a,b)} \leq \mathbf{b}^{[a,b)} Y_{[a,b)} \qquad \forall [a,b) \in \mathcal{Y}^{\mathcal{K}} \tag{7a}$$

$$\sum_{[a,b)\in\mathcal{Y}^{\mathcal{K}}} P_t^{[a,b)} = P_t \qquad \forall t \in \mathcal{T} \tag{7b}$$

$$\sum_{\{[a,b)\in\mathcal{Y}^{\mathcal{K}} \mid t\in[a,b+\mathbf{DT})\}} Y_{[a,b)} \leq |\mathcal{K}| \qquad \forall t \in \mathcal{T} \tag{7c}$$

$$Y_{[a,b)} \in \{0,\ldots,|\mathcal{K}|\} \qquad \forall [a,b) \in \mathcal{Y}^{\mathcal{K}}. \tag{7d}$$

Algorithm 1 demonstrates how to construct feasible schedules given a solution (Y^*, P^*) to (7), by "peeling-off" a feasible solution (y^g, p^g) for generator g and leaving behind a feasible solution $(\hat{Y}, \hat{P})$ to (7) for $\mathcal{K}\setminus\{g\}$. The essential logic of the method is to always take feasible startups when available (line 10), thus ensuring the remaining aggregated solution is feasible. Theorem 3 ensures optimality of this approach. Additionally, Theorem 1 allows us to assign to each generator on in an interval $[a,b)$ the average of the power output across all generators on during interval $[a,b)$ (line 9) while maintaining optimality and feasibility.

Algorithm 1 (Peel Off EF) Constructs feasible generator schedules from a solution of (7).

Initialize all y^g, $p^{g,[a,b)}$ to 0.
Initialize all $\hat{Y}$ to Y^*, $\hat{P}^{[a,b)}$ to $P^{*[a,b)}$ respectively.
$t \leftarrow \min\{i \mid Y^*_{[i,j)} \geq 1\}$
while $t \leq \mathbf{T}$ **do**
5: $\quad t' \leftarrow \min\{j \mid Y^*_{[t,j)} \geq 1\}$
$\quad y^g_{[t,t')} \leftarrow 1$; $\hat{Y}_{[t,t')} \leftarrow Y^*_{[t,t')} - 1$;
$\quad$ **for** $i \in [t,t') \cap \mathcal{T}$ **do**
$\quad\quad p_i^{g,[t,t')} \leftarrow P_i^{*[t,t')} / Y^*_{[t,t')}$
$\quad\quad \hat{P}_i^{[t,t')} \leftarrow P_i^{*[t,t')} - p_i^{g,[t,t')}$
10: $\quad t \leftarrow \min\{i \geq t' + \mathbf{DT} \mid Y^*_{[i,j)} \geq 1\}$
$\hat{P}_t \leftarrow \sum_{[a,b)\in\mathcal{Y}} \hat{P}_t^{[a,b)}$, $\forall t \in \mathcal{T}$
$p_t^g \leftarrow \sum_{[a,b)\in\mathcal{Y}} p_t^{g,[a,b)}$, $\forall t \in \mathcal{T}$

Note that having different initial conditions is not an issue as long as we extend $\mathcal{Y}^{\mathcal{K}}$ back to include intervals when the generators in $\mathcal{K}$ last started or last ran (adding duplicate intervals if necessary for two generators that started at the same time period but have different outputs at $t = 0$).

3.2 3-Bin for Fast-Ramping

Let $U = \sum_{g\in\mathcal{K}} u^g$, $V = \sum_{g\in\mathcal{K}} v^g$, and $W = \sum_{g\in\mathcal{K}} w^g$. We will first consider the aggregated 3-bin model for commitment status:

$$U_t - U_{t-1} = V_t - W_t \qquad \forall t \in \mathcal{T} \tag{8a}$$

$$\sum_{i=t-\mathbf{UT}+1}^{t} V_i \leq U_t \qquad \forall t \in [\mathbf{UT}^{\mathcal{K}}, \mathbf{T}] \tag{8b}$$

$$\sum_{i=t-\mathbf{DT}+1}^{t} W_i \leq |\mathcal{K}| - U_t \qquad \forall t \in [\mathbf{DT}^{\mathcal{K}}, \mathbf{T}] \tag{8c}$$

$$U_t, V_t, W_t \in \{0, \dots, |\mathcal{K}|\} \qquad \forall t \in \mathcal{T} \tag{8d}$$

Algorithm 2 demonstrates how to disaggregate a solution (U^*, V^*, W^*) to (8) by constructing a feasible 3-bin schedule for a generator g and leaving a feasible solution $(\hat{U}, \hat{V}, \hat{W})$ to (8) after g is removed from $\mathcal{K}$. Similar to Algorithm 1, it essentially takes shutdowns whenever possible when on (lines 7 – 11) and startups whenever possible when off (lines 12 – 14). Thus, the schedule is clearly feasible for g, and taking startups/shutdowns whenever possible ensures (8) remains feasible for $\mathcal{K} \setminus \{g\}$ (and so all the integer bounds decrease by 1), and as before, Theorems 2 and 3 establish optimality.

Algorithm 2 (Peel Off 3-bin) Constructs feasible generator schedules from a solution of (8).

1: **Initialize** all u^g, v^g, w^g to 0.
2: **Initialize** all $\hat{U}$, $\hat{V}$, $\hat{W}$ to U^*, V^*, W^* respectively.
3: **if** $U_1^* \geq 1$ **then**
4:     $u_1^g \leftarrow 1$; $\hat{U}_1 \leftarrow U_1^* - 1$;
5: $t \leftarrow 2$
6: **while** $t \leq \mathbf{T}$ **do**
7:     **if** $u_{t-1}^g = 1$ **then**
8:         **if** $\sum_{i=t-\mathbf{UT}^{\mathcal{K}}+1}^{t-1} v_i = 0$ **and** $W_t^* \geq 1$ **then**
9:             $w_t^g \leftarrow 1$; $\hat{W}_t \leftarrow W_t^* - 1$;
10:         **else**
11:             $u_t \leftarrow 1$; $\hat{U}_t \leftarrow U_t^* - 1$;
12:     **else** ▷ $u_{t-1} = 0$
13:         **if** $\sum_{i=t-\mathbf{DT}^{\mathcal{K}}+1}^{t-1} w_i = 0$ **and** $V_t^* \geq 1$ **then**
14:             $u_t, v_t \leftarrow 1$; $\hat{U}_t \leftarrow U_t^* - 1$; $\hat{V}_t \leftarrow V_t^* - 1$;
15:     $t \leftarrow t + 1$;

Note that historical data can be leveraged after the first time period. If the generator is on at $t = 1$, then we can arbitrarily assign it a historical startup v_t. Similarly, if the generator is off at $t = 1$, we can assign a historical shutdown w_t. Because ramping constraints are not active, initial conditions can be arbitrarily assigned based on whether the generator is on or not. We expand this result and provide a proof of correctness – including the case with time-dependent startup costs – in the appendix.

Consider the power output for an aggregated set of identical fast-ramping generators. Along with (8), we have the aggregated power $P = \sum_{g \in \mathcal{K}} p^g$ with the constraints for $\mathbf{UT}^{\mathcal{K}} \geq 2$:

$$\underline{\mathbf{P}}^{\mathcal{K}} U_t \leq P_t \qquad \forall t \in \mathcal{T} \tag{9a}$$

$$P_t \leq \overline{\mathbf{P}}^{\mathcal{K}} U_t + (\mathbf{SU}^{\mathcal{K}} - \overline{\mathbf{P}}^{\mathcal{K}}) V_t + (\mathbf{SD}^{\mathcal{K}} - \overline{\mathbf{P}}^{\mathcal{K}}) W_{t+1}, \qquad \forall t \in \mathcal{T}. \tag{9b}$$

Because (9) is a sum of constraints, it is clearly valid. Further, using the result from Algorithm 2 along with Theorem 1, we can construct a feasible and optimal disaggregation for power output. For simplicity suppose $\mathbf{SU}^{\mathcal{K}} = \mathbf{SD}^{\mathcal{K}}$; we handle the remaining two cases in the appendix. If $u_t^g = 1$, $v_t^g = 0$, and $w_{t+1}^g = 0$ then

$$p_t^g = \frac{P_t^* - \min\{\mathbf{SU}^{\mathcal{K}}, P_t^*/U_t^*\} \cdot V_t^* - \min\{\mathbf{SD}^{\mathcal{K}}, P_t^*/U_t^*\} \cdot W_{t+1}^*}{U_t^* - V_t^* - W_{t+1}^*}. \tag{10}$$

If the generator is just starting up such that $v_t^g = 1$, then $p_t^g = \min\{\mathbf{SU}^{\mathcal{K}}, P_t^*/U_t^*\}$, and similarly if shutting down ($w_{t+1}^g = 1$), then $p_t^g = \min\{\mathbf{SD}^{\mathcal{K}}, P_t^*/U_t^*\}$. If $u_t^g = 0$ then $p_t^g = 0$. Reserves and piecewise linear costs can be disaggregated in a similar fashion, as can generators with $\mathbf{UT}^{\mathcal{K}} = 1$; see the appendix for details.

4 The Potential Impact of Aggregation

Aggregating identical generators provides a way of efficiently exploiting the presence of symmetry in UC, leading to several computational advantages. First, UC instances with large numbers of identical generators may have alternative optimal solutions. Consider the case where two identical generators in a UC instance have different schedules. Permuting these schedules will lead to different solutions in a disaggregated UC model. If there are k generators of the same type, then there may be as many as $k!$ different optimal solutions. While a variety of methods are used to combat the effects of symmetry in MILP models, these generally rely on symmetry-breaking cuts or clever branching, and experience has shown that explicitly aggregating symmetry away is the most successful way of exploiting symmetry – when possible.

We observe that aggregate UC solutions may encode more than just symmetric solutions found by permuting generator schedules. For example, consider the U variables defined in formulation (3), and assume generators are not ramp-constrained. Assuming $\mathbf{UT}^{\mathcal{K}} \leq 3$, there are two ways to feasibly disaggregate the on/off solution. The first is

$$u^{g_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, u^{g_2} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \tag{11}$$

and the other is:

$$u^{g_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, u^{g_2} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \tag{12}$$

These solutions are not symmetric, but they do have identical objective function values. As is the case with symmetric solutions, such mutually non-dominating solutions may lead to more tree exploration in the branch-and-cut process. Aggregation allows us to consider these solutions simultaneously.

Now, we further illustrate the disaggregation process. When $\mathbf{UT}^{\mathcal{K}} \leq 3$, Algorithm 2 decomposes the aggregate solution to formulation (3) – again assuming redundant ramping limits – to (11). Then, by (10) the power output associated with solution (11) is given as

$$p^{g_1} = \begin{pmatrix} 200 \\ 200 \\ 200 \\ 200 \\ 200 \end{pmatrix}, p^{g_2} = \begin{pmatrix} 0 \\ 100 \\ 200 \\ 100 \\ 0 \end{pmatrix}. \tag{13}$$

On the other hand, if $\mathbf{UT}^{\mathcal{K}} > 3$, then Algorithm 2 with equation (10) yields

$$p^{g_1} = \begin{pmatrix} 200 \\ 200 \\ 200 \\ 100 \\ 0 \end{pmatrix}, p^{g_2} = \begin{pmatrix} 0 \\ 100 \\ 200 \\ 200 \\ 200 \end{pmatrix}. \tag{14}$$

Table 1: Ostrowski Instances: Generator Performance Data

Gen	$\overline{\mathbf{P}}$ (MW)	$\underline{\mathbf{P}}$ (MW)	**UT**(**DT**) (h)	**SU**(**SD**) (MW/h)	**RU**(**RD**) (MW/h)	**TC** (h)
1	455	150	8	150	225	14
2	455	150	8	150	225	14
3	130	20	5	20	50	10
4	130	20	5	20	50	10
5	162	25	6	25	60	11
6	80	20	3	20	60	8
7	85	25	3	25	60	6
8	55	10	1	20	135	2

5 Computational Experiments

To test the effectiveness of our aggregation approach, we selected two unit commitment test sets from the literature. The first, described further in [25], is based on real-world data gathered from the California Independent System Operator (CAISO). Each of the twenty instances share a set of 610 thermal generators. We consider five 48-hour demand scenarios, crossed with static reserve requirements that varied between 0%, 1%, 3%, and 5% of system demand. The scenarios labeled with dates correspond to real-world load profiles from the corresponding calendar date. In contrast, "Scenario400" is a hypothetical scenario where wind supply is on average 40% of demand (we use a "net-load" formulation for wind supply). We approximated the quadratic cost curves provided with two piecewise segments. Each thermal generator has two startup categories. We refer to these instances as "CAISO" instances.

The other UC test set is taken from [4]. The instances in this test case – which we refer to as the "Ostrowski" instances – are constructed by replicating the thermal generators in UC instances originally introduced in [26] and [22]. These generators have been used as a baseline to create large UC instances through replication in much of the UC literature [5, 27–36]. The parameters and cost curves used for the eight base generators are provided in Tables 1 and 2. In Table 2, a, b, and c denote the coefficients of the quadratic cost function such that $\mathbf{c}^p(p_t) = a^2 p_t + b p_t + c$; $\mathbf{c}^H$ and $\mathbf{c}^C$ respectively denote the hot- and cold-startup costs. The number of copies of each generator type in each of the twenty instances is specified in Table 3. The number of thermal generators in the Ostrowski instances ranges from 28 to 187. The demand curve for each instance is given as a percentage of total system capacity, as reported in Table 4, and for each instance the reserve level was fixed to 3% of system demand. We use a two segment piecewise approximation for production costs. In contrast to the CAISO instances, the Ostrowski instances consider only a 24 hour scheduling horizon.

We consider the base UC MILP formulation described in [25], which represents the state-of-the-art. The performance of our aggregation approach is analyzed relative to this baseline.

Table 2: Ostrowski Instances: Generator Cost Data

Gen	c (\$)	b (\$/MW)	a (\$/MW^2)	$\mathbf{c}^H$ (\$)	$\mathbf{c}^C$ (\$)
1	1000	16.19	0.00048	4500	9000
2	970	17.26	0.00031	5000	10000
3	700	16.60	0.00200	550	1100
4	680	16.50	0.00211	560	1120
5	450	19.70	0.00398	900	1800
6	370	22.26	0.00712	170	340
7	480	27.74	0.00079	260	520
8	660	25.92	0.00413	30	60

Table 3: Ostrowski Instances: Number of Generators of Each Type

	Generator								Total
Problem	1	2	3	4	5	6	7	8	Gens
1	12	11	0	0	1	4	0	0	28
2	13	15	2	0	4	0	0	1	35
3	15	13	2	6	3	1	1	3	44
4	15	11	0	1	4	5	6	3	45
5	15	13	3	7	5	3	2	1	49
6	10	10	2	5	7	5	6	5	50
7	17	16	1	3	1	7	2	4	51
8	17	10	6	5	2	1	3	7	51
9	12	17	4	7	5	2	0	5	52
10	13	12	5	7	2	5	4	6	54
11	46	45	8	0	5	0	12	16	132
12	40	54	14	8	3	15	9	13	156
13	50	41	19	11	4	4	12	15	156
14	51	58	17	19	16	1	2	1	165
15	43	46	17	15	13	15	6	12	167
16	50	59	8	15	1	18	4	17	172
17	53	50	17	15	16	5	14	12	182
18	45	57	19	7	19	19	5	11	182
19	58	50	15	7	16	18	7	12	183
20	55	48	18	5	18	17	15	11	187

Table 4: Ostrowski Instances: Demand (% of Total Capacity)

Time	1	2	3	4	5	6	7	8	9	10	11	12
Demand	71%	65%	62%	60%	58%	58%	60%	64%	73%	80%	82%	83%
Time	13	14	15	16	17	18	19	20	21	22	23	24
Demand	82%	80%	79%	79%	83%	91%	90%	88%	85%	84%	79%	74%

All computational experiments were conducted on a Dell PowerEdge T620 server with two Intel Xeon E5-2670 processors, for a total of 16 cores, 32 threads, and 256GB of RAM, running the Ubuntu 14.04.5 Linux operating system. The Gurobi 7.0.1 MILP solver [2] was used in all experiments, and the solver was allowed to use all 32 threads in each experimental trial.

5.1 CAISO Instances

Of the total 610 thermal generators in the CAISO instances, there is no symmetry among the 36 slow-ramping generators, but the 574 fast-ramping generators do have some non-trivial symmetry. Aggregation allows us to reduce these 574 fast-ramping generators to 429 aggregated generators, with the largest aggregation representing 8 physical generators. Overall, we obtained 36 aggregated generators. In Table 5 we report on a subset of the aggregated generators, specifically excluding those for which we can only aggregate two generators. The CAISO test set is dominated by flexible, fast-ramping generators, with $\mathbf{UT} = \mathbf{DT} = 1$, $\mathbf{TC} = 2$, $\mathbf{SU}, \mathbf{SD} \geq \overline{\mathbf{P}}$, and $\mathbf{RU}, \mathbf{RD} \geq (\overline{\mathbf{P}} - \underline{\mathbf{P}})$. Consequently, we omit this information for purposes of brevity. Similar to Table 2, a, b, and c denote coefficients for the quadratic cost function, while $\mathbf{c}^H$ and $\mathbf{c}^C$ respectively denote the hot- and cold-startup costs. Because we are not using the EF UC formulation, our aggregation approach allows us to reduce the size relative to the standard 3-bin formulation by 24%.

In Table 6 we report the wall-clock time and number of branch-and-cut nodes explored before termination for the respective formulations, where "3-bin" denotes the UC formulation proposed in [25] and "3-bin+A" denotes the aggregation formulation for fast-ramping generators introduced in this paper. We left all Gurobi parameter settings at their default value, such that the solver terminated when the optimality gap was less than or equal to 0.01%. We did not impose a time limit for these experiments. Despite their size, we observe that these instances are not difficult given the current state-of-the-art UC formulation and a modern commercial MILP solver. Yet, we do observe a geometric mean improvement of 37% in wall clock time with aggregation, across the twenty CAISO instances. Aggregation is only slower in only two of the twenty instances, and in both cases the difference is minimal.

Next, considering the number of nodes explored during the branch-and-cut search process, we see that neither formulation has an advantage. We conjecture this is likely due to role of Gurobi's incumbent-finding heuristics. Specifically, because the typical CAISO instance has a root gap of <0.01% [25], Gurobi only needs to find a high-quality solution before terminating, and does not need to expend significant effort proving that a solution is optimal within optimality gap tolerance.

5.2 Ostrowski Instances

For the Ostrowski instances, generators 1 – 5 have non-redundant ramping constraints, so we use the extended formulation in our aggregation, and for generators 6 – 8 we use the 3-bin fast-ramping aggregation. In all

Table 5: CAISO: Selected Generator Performance Characteristics

Number Identical	$\overline{\mathbf{P}}$ (MW)	$\underline{\mathbf{P}}$ (MW)	c $*10^3$	b ($/MW)$*10^3$	a ($/MW^2)$*10^3$	$\mathbf{c}^C$ $*10^3$	$\mathbf{c}^H$ $*10^3$
8	106.3	47.835	2.19187	0.02498	0.0000581	3.189	4.252
7	3.1	0.93	0.00662	0.03977	0	0.093	0.124
6	1.4	0.35	0.00662	0.04452	0	0.07	0.098
5	100	45	2.06698	0.02498	0.0000593	3.0	4.0
5	3.5	1.05	0.00415	0.04206	0	0.0105	0.14
4	180	81	3.59086	0.02402	0.0000297	5.4	7.2
4	75	74.25	0.13711	0.02726	0	2.25	3.00
4	70	53.9	0.21868	0.02493	0.0000148	2.1	2.8
4	49.9	22.455	1.09041	0.02568	0.0000636	1.497	1.996
4	49.5	22.275	1.05231	0.02497	0.0000611	1.485	1.98
4	12.2	3.05	0.04126	0.04326	0	0.61	0.854
4	11.2	5.04	0.24371	0.0295	0.0000496	0.336	0.448
4	11	4.95	0.23929	0.02495	0.0000533	0.33	0.44
4	10.75	4.8375	0.23392	0.02495	0.0000517	0.3225	0.43
3	100	45	1.7298	0.02381	0.0000672	3.0	4.0
3	78	77.22	0.06097	0.02831	0	2.34	3.12
3	21.69	5.4225	0.07335	0.04326	0	1.0845	1.5183
3	21.6	21.384	-0.76024	0.06329	0	0.648	0.864
3	16.27	4.0675	0.05502	0.04326	0	0.8135	1.1389
3	3.4	1.02	0.00381	0.03977	0	0.102	0.136
3	1.3	0.6	-0.00728	0.04197	0	0.039	0.052

cases there are no more than 8 generators in the aggregated model. Due to the difference in difficulty relative to the CAISO instances, we impose a time limit of 900 seconds for these experiments.

We report the results of experiments on the Ostrowski instances in Table 7, recording the terminating optimality gap in parentheses for cases when the time limit was reached. Even with a state-of-the-art UC MILP formulation (i.e., that of [25]) and 900 seconds of wall-clock time, Gurobi fails to establish optimality within tolerance for over half of the 20 instances. Given that the Ostrowski instances have half the number of time periods and far fewer thermal generators than the CAISO instances, one might expect these instances to be easier. However, as demonstrated in [9,10] and Table 7, even a modern MILP solver with sophisticated, general symmetry detection routines cannot handle UC instances with large numbers of identical generators. In comparison, our aggregation approach significantly reduces the difficulty of these instances, to the point where they can be solved in at most two minutes of wall clock time. Further, our aggregation approach requires far fewer nodes during the branch-and-cut process, often by an order of magnitude or more. As reported in the appendix, aggregation also outperforms the static symmetry breaking inequalities proposed in [10] for the 3-bin formulation.

Table 6: Computational Results for CAISO UC Instances

	Time (s)		Nodes	
Instance	**3-bin**	**3-bin+A**	**3-bin**	**3-bin+A**
2014-09-01 0%	31.35	**14.25**	0	0
2014-12-01 0%	25.77	**12.38**	0	0
2015-03-01 0%	24.08	**14.27**	0	0
2015-06-01 0%	13.11	**8.50**	0	0
Scenario400 0%	27.29	**23.63**	0	0
2014-09-01 1%	20.52	**16.44**	0	0
2014-12-01 1%	38.48	**24.69**	95	**0**
2015-03-01 1%	21.75	**19.11**	0	0
2015-06-01 1%	39.87	**15.59**	47	**0**
Scenario400 1%	47.54	**44.63**	**0**	1438
2014-09-01 3%	81.47	**38.27**	**7**	122
2014-12-01 3%	65.01	**36.53**	1292	**125**
2015-03-01 3%	50.79	**25.04**	0	0
2015-06-01 3%	87.25	**41.23**	**0**	115
Scenario400 3%	131.28	**69.45**	2055	**880**
2014-09-01 5%	47.07	**30.95**	95	**7**
2014-12-01 5%	83.87	**66.90**	**1203**	3978
2015-03-01 5%	80.57	**21.65**	923	**0**
2015-06-01 5%	**26.99**	43.79	**0**	402
Scenario400 5%	**115.53**	118.51	**3867**	4225
Geometric Mean:	43.85	**27.55**		

6 Conclusion

We have shown that symmetry due to the presence of identical generators is present in both real-world and academic UC instances, and we posited an aggregation method that can mitigate computational issues induced by this symmetry. While modern MILP solvers possess sophisticated symmetry-detection technology, they are unable to address this form of UC symmetry. Our aggregation approach requires a fairly straightforward reformulation of the UC MILP, and disaggregation is straightforward. Thus, our approach is viable in practice for addressing symmetry in UC.

Acknowledgments

B. Knueven and J. Ostrowski were supported by NSF award number 1332662. B. Knueven was additionally supported the U.S. Department of Energy (DOE), Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education (ORISE) for the DOE. ORISE is managed by ORAU under contract number DE-AC05-06OR23100. The work of J.-P. Watson was

Table 7: Computational Results for Ostrowski UC Instances

	Time (s)		Nodes	
Instance	**3-bin**	**EF/3-bin+A**	**3-bin**	**EF/3-bin+A**
1	**8.44**	14.02	1509	**68**
2	154.75	**21.07**	48129	**157**
3	703.94	**100.33**	316704	**4464**
4	**14.84**	17.28	8532	**60**
5	143.18	**57.22**	131320	**4350**
6	95.41	**28.00**	62394	**72**
7	(0.0238%)	**119.22**	535361*	**4854**
8	(0.0107%)	**71.00**	1378310*	**9267**
9	(0.0169%)	**125.63**	819798*	**12217**
10	(0.0327%)	**82.89**	751319*	**11549**
11	(0.0186%)	**18.76**	73976*	**1155**
12	(0.0240%)	**22.91**	42729*	**460**
13	(0.0266%)	**74.43**	41325*	**6464**
14	(0.0144%)	**19.75**	41469*	**15**
15	780.76	**39.63**	120599	**3091**
16	(0.0162%)	**90.31**	42102*	**2597**
17	154.37	**27.88**	2114	**1059**
18	(0.0121%)	**22.36**	60651*	**151**
19	(0.0195%)	**21.30**	42683*	**2436**
20	106.44	**18.46**	527	**0**
Geometric Mean:	>349.77	**38.03**		

supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under contract number KJ0401000 through the project "Multifaceted Mathematics for Complex Energy Systems", and the Grid Modernization Initiative of the U.S. DOE, under project 1.4.26, as part of the Grid Modernization Laboratory Consortium, a strategic partnership between DOE and the national laboratories to bring together leading experts, technologies, and resources to collaborate on the goal of modernizing the nation's grid. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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.

References

[1] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power Generation, Operation and Control.* John Wiley & Sons, 2013.

[2] Gurobi Optimization, Inc., "Gurobi optimizer reference manual," 2016. [Online]. Available: http://www.gurobi.com

[3] International Business Machines Corporation, "IBM CPLEX Optimizer," 2017. [Online]. Available: https://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/

[4] J. Ostrowski, M. F. Anjos, and A. Vannelli, "Tight mixed integer linear programming formulations for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, p. 39, 2012.

[5] G. Morales-España, J. M. Latorre, and A. Ramos, "Tight and compact MILP formulation for the thermal unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4897–4908, 2013.

[6] C. Gentile, G. Morales-Espana, and A. Ramos, "A tight MIP formulation of the unit commitment problem with start-up and shut-down constraints," *EURO Journal on Computational Optimization*, vol. 5, no. 1–2, pp. 177–201, 2017.

[7] G. Morales-España, C. Gentile, and A. Ramos, "Tight MIP formulations of the power-based unit commitment problem," *OR Spectrum*, pp. 1–22, 2015.

[8] B. Knueven, J. Ostrowski, and J. Wang, "Generating cuts from the ramping polytope for the unit commitment problem," 2016.

[9] J. Ostrowski, M. F. Anjos, and A. Vannelli, "Modified orbital branching for structured symmetry with an application to unit commitment," *Mathematical Programming*, vol. 150, no. 1, pp. 99–129, 2015.

[10] R. M. Lima and A. Q. Novais, "Symmetry breaking in MILP formulations for unit commitment problems," *Computers & Chemical Engineering*, vol. 85, pp. 162–176, 2016.

[11] K. Cheung, D. Gade, C. Silva-Monroy, S. Ryan, J. Watson, R. Wets, and D. Woodruff, "Toward scalable stochastic unit commitment - part 2: Solver configuration and performance assessment," *Energy Systems*, vol. 6, no. 3, pp. 417–438, 2015.

[12] J. Ostrowski, J. Linderoth, F. Rossi, and S. Smriglio, "Orbital branching," *Mathematical Programming*, vol. 126, no. 1, pp. 147–178, 2011.

[13] B. Knueven, J. Ostrowski, and J.-P. Watson, "Online companion for exploiting identical generators in unit commitment," 2017.

[14] S. Sen and D. Kothari, "An equivalencing technique for solving the large-scale thermal unit commitment problem," in *The Next Generation of Electric Power Unit Commitment Models.* Springer, 2002, pp. 211–225.

[15] J. Garcia-Gonzalez, R. M. R. de la Muela, L. M. Santos, and A. M. Gonzalez, "Stochastic joint optimization of wind generation and pumped-storage units in an electricity market," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 460–468, 2008.

[16] A. Shortt and M. O'Malley, "Impact of variable generation in generation resource planning models," in *Power and Energy Society General Meeting, 2010 IEEE*. IEEE, 2010, pp. 1–6.

[17] B. Palmintier and M. Webster, "Impact of unit commitment constraints on generation expansion planning with renewables," in *Power and Energy Society General Meeting, 2011 IEEE*. IEEE, 2011, pp. 1–7.

[18] P. Malkin, "Minimum runtime and stoptime polyhedra," *Report, CORE, Université catholique de Louvain*, 2003.

[19] J. Lee, J. Leung, and F. Margot, "Min-up/min-down polytopes," *Discret. Optim.*, vol. 1, no. 1, pp. 77–85, Jun. 2004.

[20] D. Rajan and S. Takriti, "Minimum up/down polytopes of the unit commitment problem with start-up costs," *IBM Res. Rep*, 2005.

[21] M. Fischetti and L. Liberti, "Orbital shrinking," in *International Symposium on Combinatorial Optimization*. Springer, 2012, pp. 48–58.

[22] M. Carrion and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1371–1378, Aug 2006.

[23] M. C. Golumbic, *Algorithmic graph theory and perfect graphs*. Elsevier, 2004, vol. 57.

[24] S. Baum and L. E. Trotter Jr, "Integer rounding and polyhedral decomposition for totally unimodular systems," in *Optimization and Operations Research*. Springer, 1978, pp. 15–23.

[25] B. Knueven, J. Ostrowski, and J.-P. Watson, "A novel matching formulation for startup costs in unit commitment," 2017.

[26] S. A. Kazarlis, A. Bakirtzis, and V. Petridis, "A genetic algorithm solution to the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 83–92, 1996.

[27] A. Borghetti, A. Frangioni, F. Lacalandra, A. Lodi, S. Martello, C. Nucci, and A. Trebbi, "Lagrangian relaxation and tabu search approaches for the unit commitment problem," in *Power Tech Proceedings, 2001 IEEE Porto*, vol. 3. IEEE, 2001, pp. 7–pp.

[28] A. Borghetti, A. Frangioni, F. Lacalandra, and C. A. Nucci, "Lagrangian heuristics based on disaggregated bundle methods for hydrothermal unit commitment," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 313–323, 2003.

[29] A. Frangioni and C. Gentile, "Perspective cuts for a class of convex 0–1 mixed integer programs," *Mathematical Programming*, vol. 106, no. 2, pp. 225–236, 2006.

[30] A. Frangioni, C. Gentile, and F. Lacalandra, "Solving unit commitment problems with general ramp constraints," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 5, pp. 316–326, 2008.

[31] A. Frangioni and C. Gentile, "A computational comparison of reformulations of the perspective relaxation: SOCP vs. cutting planes," *Operations Research Letters*, vol. 37, no. 3, pp. 206–210, 2009.

[32] A. Frangioni, C. Gentile, and F. Lacalandra, "Tighter approximated MILP formulations for unit commitment problems," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 105–113, 2009.

[33] R. Jabr, "Tight polyhedral approximation for mixed-integer linear programming unit commitment formulations," *IET Generation, Transmission & Distribution*, vol. 6, no. 11, pp. 1104–1111, 2012.

[34] L. Yang, J. Jian, Y. Zhu, and Z. Dong, "Tight relaxation method for unit commitment problem using reformulation and lift-and-project," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 13–23, 2015.

[35] L. Yang, J. Jian, Y. Wang, and Z. Dong, "Projected mixed integer programming formulations for unit commitment problem," *International Journal of Electrical Power & Energy Systems*, vol. 68, pp. 195–202, 2015.

[36] M. J. Feizollahi, M. Costley, S. Ahmed, and S. Grijalva, "Large-scale decentralized unit commitment," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 97–106, 2015.

[37] Y. Pochet and L. A. Wolsey, *Production planning by mixed integer programming.* Springer Science & Business Media, 2006.

Appendix

In this appendix we detail how reserves, piecewise linear operating costs, and time-dependent startup costs can be exactly aggregated. Additionally we specify how to disaggregate solutions for fast-ramping generators with different startup and shutdown ramp rates. For ease of presentation this appendix (mostly) is self-contained.

Nomenclature

Indices and Sets

$g \in \mathcal{G}$	Thermal generators
$l \in \mathcal{L}^g$	Piecewise production cost intervals for generator g: $1, \dots, \mathbf{L}_g$.
$s \in \mathcal{S}^g$	Startup categories for generator g, from hottest (1) to coldest ($\mathbf{S}_g$).
$t \in \mathcal{T}$	Hourly time steps: $1, \dots, \mathbf{T}$.
$[t, t') \in \mathcal{X}^g$	Feasible intervals of operation for generator g with respect to its minimum downtime, that is, $[t, t') \in \mathcal{T} \times \mathcal{T}$ such that $t' \geq t + \mathbf{DT}^g$, including times (as necessary) before and after the planning period $\mathcal{T}$.
$[t, t') \in \mathcal{Y}^g$	Feasible intervals of operation for generator g with respect to its minimum uptime, that is, $[t, t') \in \mathcal{T} \times \mathcal{T}$ such that $t' \geq t + \mathbf{UT}^g$, including times (as necessary) before and after the planning period $\mathcal{T}$.

Parameters

$\mathbf{c}^{l,g}$	Cost coefficient for piecewise segment l for generator g (\$/MWh).
$\mathbf{c}^{R,g}$	Cost of generator g running and operating at minimum production $\underline{\mathbf{P}}_g$ (\$/h).
$\mathbf{c}^{s,g}$	Startup cost of category s for generator g (\$).
$\mathbf{D}_t$	Load (demand) at time t (MW).
$\mathbf{DT}^g$	Minimum down time for generator g (h).
$\overline{\mathbf{P}}^g$	Maximum power output for generator g (MW).
$\overline{\mathbf{P}}^{l,g}$	Maximum power available for piecewise segment l for generator g (MW) ($\overline{\mathbf{P}}^{0,g} = \underline{\mathbf{P}}^g$).
$\underline{\mathbf{P}}^g$	Minimum power output for generator g (MW).
$\mathbf{R}_t$	Spinning reserve at time t (MW).
$\mathbf{RD}^g$	Ramp-down rate for generator g (MW/h).
$\mathbf{RU}^g$	Ramp-up rate for generator g (MW/h).
$\mathbf{SD}^g$	Shutdown rate for generator g (MW/h).
$\mathbf{SU}^g$	Startup rate for generator g (MW/h).

$\mathbf{TC}^g$	Time down after which generator g goes cold, i.e., enters state S^g.
$\underline{\mathbf{T}}^{s,g}$	Time offline after which the startup category s is available ($\underline{\mathbf{T}}^{1,g} = \mathbf{DT}^g$, $\underline{\mathbf{T}}^{S^g,g} = \mathbf{TC}^g$)
$\overline{\mathbf{T}}^{s,g}$	Time offline after which the startup category s is no longer available ($= \underline{\mathbf{T}}^{s+1,g}$, $\overline{\mathbf{T}}^{S_g,g} = +\infty$)
$\mathbf{UT}^g$	Minimum run time for generator g (h).

Variables

p_t^g	Power output for generator g at time t (MW).
$p_t^{l,g}$	Power from piecewise interval l for generator g at time t (MW).
r_t^g	Spinning reserves provided by generator g at time t (MW), ≥ 0.
u_t^g	Commitment status of generator g at time t, $\in \{0,1\}$.
v_t^g	Startup status of generator g at time t, $\in \{0,1\}$.
w_t^g	Shutdown status of generator g at time t, $\in \{0,1\}$.
$c_t^{SU,g}$	Startup cost for generator g at time t (\$), ≥ 0.
$x_{[t,t')}^g$	Indicator arc for shutdown at time t, startup at time t', uncommitted for $i \in [t,t')$, for generator g, $\in \{0,1\}$, $[t,t') \in \mathcal{X}^g$.
$y_{[t,t')}^g$	Indicator arc for startup at time t, shutdown at time t', committed for $i \in [t,t')$, for generator g, $\in \{0,1\}$, $[t,t') \in \mathcal{Y}^g$.

A Unit Commitment Formulation

This section lays out the basic unit commitment formulation we consider for the computational tests above.

$$\min \sum_{g \in G} \mathbf{c}^g(p^g) \tag{15a}$$

subject to:

$$\sum_{g \in \mathcal{G}} p_t^g = \mathbf{D}_t \qquad \forall t \in \mathcal{T} \tag{15b}$$

$$\sum_{g \in \mathcal{G}} r_t^g \geq \mathbf{R}_t \qquad \forall t \in \mathcal{T} \tag{15c}$$

$$(p^g, r^g) \in \Pi^g \qquad \forall g \in \mathcal{G}, \tag{15d}$$

where $\mathbf{c}^g(p^g)$ is the cost function of generator g producing an output of p^g over the time horizon $\mathcal{T}$, and Π^g represents the set of feasible schedules for generator g. Here $\mathbf{c}^g$ includes possibly piecewise linear convex production costs, as well as time-dependent startup costs. We will see how we can modify the formulations in the main body to allow for these additional modeling features while still maintaining the property that we can disaggregate solutions to aggregated generators.

For completeness we restate the theorems from the main text.

Theorem 1. *Consider identical generators $g_1, g_2 \in \mathcal{G}$ on at time t, and assume their production costs are increasing and convex. If possible, $p_t^{g_1} = p_t^{g_2}$. If not, then least one of the nominal or startup/shutdown ramping constraints is binding for generator g_1 or g_2 at time t in an optimal solution.*

Theorem 2. *Suppose generator g_1 is turned off at time t. If identical generator g_2 can also be turned off at time t, there exists an optimal solution where the generator that has been on for the least amount of time is turned off.*

Theorem 3. *Suppose generator g_1 is turned on at time t. If an identical generator g_2 can also be turned on at time t, and there are no time-dependent startup costs for g_1 and g_2, then there exists an optimal solution where g_2 is turned on at t.*

While Theorem 3 is useful for when start-up costs are not time-dependent, this is often not a realistic assumption. Therefore, we have the following two theorems which will be of use when start-up costs are time-dependent.

Theorem 4. *Suppose generator g_1 is turned on at time t, and has been off for at least $\mathbf{TC}(= \mathbf{TC}^{g_1} = \mathbf{TC}^{g_2})$ time periods. If identical generator g_2 can also be turned on at time t and has been off for at least $\mathbf{TC}$ time periods, there exists an optimal solution where where g_2 is turned on at time t.*

Proof. Similar to the proof of Theorem 3. □

Theorem 5. *If identical generators $g_1, g_2 \in \mathcal{G}$ both shut down at time t and g_1 starts up at time $t_1 \geq t + \mathbf{DT}$ $(\mathbf{DT} = \mathbf{DT}^{g_1} = \mathbf{DT}^{g_2})$ and g_2 starts up at time $t_2 \geq t + \mathbf{DT}$, then an equally good solution exists where g_1 starts up at time t_2 and g_2 starts up at time t_1.*

Proof. Like in the proof of Theorem 3, we may permute the remainder of each generator's schedule without affecting feasibility or the objective value. □

B Disaggregating the Extended Formulation

First, we consider the extended formulation, which is more straightforward than disaggregating the 3-bin formulation. We can add reserves and piecewise linear production costs by adding new variables $r_t^{[a,b),g}$ $\forall [a,b) \in \mathcal{Y}^g, \ \forall t \in \mathcal{T}$ and $p^{[a,b),l,g}$ $\forall [a,b) \in \mathcal{Y}^g, \ \forall l \in \mathcal{L}^g, \ \forall t \in \mathcal{T}$. We add time-dependent startup costs by replacing the packing polytope we used above with the shortest path polytope [37]. The resulting formulation is

$$\mathbf{A}^{[a,b)}p^{[a,b)} + \mathbf{A}'^{[a,b)}r^{[a,b)} + \sum_{l\in\mathcal{L}} \mathbf{A}^{[a,b),l}p^{[a,b),l} \leq \mathbf{b}^{[a,b)}y_{[a,b)} \qquad \forall [a,b) \in \mathcal{Y} \tag{16a}$$

$$\sum_{[a,b)\in\mathcal{Y}} p_t^{[a,b)} = p_t \qquad \forall t \in \mathcal{T} \tag{16b}$$

$$\sum_{[a,b)\in\mathcal{Y}} r_t^{[a,b)} = r_t \qquad \forall t \in \mathcal{T} \tag{16c}$$

$$\sum_{\{[c,d)\in\mathcal{X} \mid t=d\}} x_{[c,d)} = \sum_{\{[a,b)\in\mathcal{Y} \mid t=a\}} y_{[a,b)} \qquad \forall t \in \mathcal{T} \tag{16d}$$

$$\sum_{\{[a,b)\in\mathcal{Y} \mid t=b\}} y_{[a,b)} = \sum_{\{[c,d)\in\mathcal{X} \mid t=c\}} x_{[c,d)} \qquad \forall t \in \mathcal{T} \tag{16e}$$

$$\sum_{\{[a,b)\in\mathcal{Y} \mid a\leq 0\}} y_{[a,b)} + \sum_{\{[c,d)\in\mathcal{X} \mid c\leq 0\}} x_{[c,d)} = 1 \tag{16f}$$

$$\sum_{\{[a,b)\in\mathcal{Y} \mid b>\mathbf{T}\}} y_{[a,b)} + \sum_{\{[c,d)\in\mathcal{X} \mid d>\mathbf{T}\}} x_{[c,d)} = 1, \tag{16g}$$

where the polytope

$$\{p^{[a,b)}, r^{[a,b)}, p^{[a,b),1}, \ldots, p^{[a,b),\mathbf{L}} \in \mathbb{R}_+ \mid \mathbf{A}^{[a,b)}p^{[a,b)} + \mathbf{A}'^{[a,b)}r^{[a,b)} + \sum_{l\in\mathcal{L}} \mathbf{A}^{[a,b),l}p^{[a,b),l} \leq \mathbf{b}^{[a,b)}\} \tag{17}$$

represents feasible production given that the generator is turned on at time a and turned off at time b. Piecewise production costs can then be handled by placing the appropriate objective coefficient on the $p^{[a,b),l}$ variables and time-dependent startup costs are accounted for by placing the appropriate objective coefficient on the $x_{[c,d)}$ variables.

Similar to the EF presented in the main text, we see that the underlying shortest path polytope (16e, 16e, 16f, 16g) with nonnegativity, has a totally unimodular constraint matrix, and thus has the integer decomposition property [24]. Hence if we have k generators have identical parameters, we can replace (16f) and (16g) with

$$\sum_{\{[a,b)\in\mathcal{Y} \mid a\leq 0\}} Y_{[a,b)} + \sum_{\{[c,d)\in\mathcal{X} \mid c\leq 0\}} X_{[c,d)} = k \tag{18a}$$

$$\sum_{\{[a,b)\in\mathcal{Y} \mid b>\mathbf{T}\}} Y_{[a,b)} + \sum_{\{[c,d)\in\mathcal{X} \mid d>\mathbf{T}\}} X_{[c,d)} = k, \tag{18b}$$

thus pushing k units of flow through the graph. Changing these right-hand-sides doesn't affect the integrality of the full polytope (16a – 16e), (18) [8].

As before, allowing capital variables to represent aggregated variables for identical generators, now $Y_{[a,b)}$ represents *how many* of the generators are on during the interval $[a, b)$ and $X_{[a,b)}$ represents how many generators are off during the interval $[a, b)$. Since there are separate power variables for each on interval $[a, b)$, like before with the EF in the main text, Theorem 1 enables us to disaggregate power easily once the status variables are disaggregated.

Letting $\mathcal{K} \subset \mathcal{G}$ be some set of identical generators, consider the extended formulation for these aggregated generators

$$\mathbf{A}^{[a,b)}P^{[a,b)} + \mathbf{A'}^{[a,b)}R^{[a,b)} + \sum_{l\in\mathcal{L}} \mathbf{A}^{[a,b),l}P^{[a,b),l} \leq \mathbf{b}^{[a,b)}Y_{[a,b)} \qquad \forall [a,b) \in \mathcal{Y} \tag{19a}$$

$$\sum_{[a,b)\in\mathcal{Y}} P_t^{[a,b)} = P_t \qquad \forall t \in \mathcal{T} \tag{19b}$$

$$\sum_{[a,b)\in\mathcal{Y}} R_t^{[a,b)} = R_t \qquad \forall t \in \mathcal{T} \tag{19c}$$

$$\sum_{\{[c,d)\in\mathcal{X} \ | \ t=d\}} X_{[c,d)} = \sum_{\{[a,b)\in\mathcal{Y} \ | \ t=a\}} Y_{[a,b)} \qquad \forall t \in \mathcal{T} \tag{19d}$$

$$\sum_{\{[a,b)\in\mathcal{Y} \ | \ t=b\}} Y_{[a,b)} = \sum_{\{[c,d)\in\mathcal{X} \ | \ t=c\}} X_{[c,d)} \qquad \forall t \in \mathcal{T} \tag{19e}$$

$$\sum_{\{[a,b)\in\mathcal{Y} \ | \ a\leq 0\}} Y_{[a,b)} + \sum_{\{[c,d)\in\mathcal{X} \ | \ c\leq 0\}} X_{[c,d)} = |\mathcal{K}| \tag{19f}$$

$$\sum_{\{[a,b)\in\mathcal{Y} \ | \ b>\mathbf{T}\}} Y_{[a,b)} + \sum_{\{[c,d)\in\mathcal{X} \ | \ d>\mathbf{T}\}} X_{[c,d)} = |\mathcal{K}|. \tag{19g}$$

We can then write down an easy algorithm to decompose solutions to (19).

After running Algorithm 3 $|\mathcal{K}| - 1$ times we are left with $|\mathcal{K}|$ feasible (and by Theorem 1 optimal) schedules, one for each generator in $\mathcal{K}$. We formalize this in Theorem 6.

First we need a simple lemma regarding the decomposability of polytopes.

Lemma 1. *Let P a polytope such that $P := \{x \in \mathbb{R}^n_+ \mid Ax \leq b\}$ and for $k \geq 1$ define $kP := \{x \mid \frac{1}{k}x \in P\} = \{x \in \mathbb{R}^n_+ \mid Ax \leq kb\}$. If $y \in kP$, then $\frac{1}{k}y \in P$ and $\frac{(k-1)}{k}y \in (k-1)P$.*

Proof. It suffices to notice that $\frac{1}{k}y$ is feasible for the system $\{x \in \mathbb{R}^n_+ \mid Ay - (k-1)b \leq Ax \leq b\}$. □

Now we turn to the main result.

Theorem 6. *Algorithm 3 returns a feasible solution for* (16) *and a feasible solution for* (19) *for the remaining $\mathcal{K} \setminus \{g\}$ generators. That is, after applying Algorithm 3 $|\mathcal{K}| - 1$ times we have a feasible and optimal solution for every $g \in \mathcal{K}$.*

Proof. Notice the feasible s, t path leaves the equalities (16d – 16g) and (19d – 19g) feasible for x^g, y^g and $\hat{X}$, $\hat{Y}$ respectively. Further, the solutions constructed for the $p^{g,[a,b)}$, $r^{g,[a,b)}$, and $p^{g,[a,b),l}$ variables (lines 6

Algorithm 3 (Peel Off EF) Constructs feasible generator schedules from a solution of (19).

Initialize all $\hat{P}^{[a,b)}$ to $P^{*[a,b)}$ and all $p^{g,[a,b)}$ to 0.
Find a feasible s, t path based on (19d – 19g) and store in x^g, y^g.
$\hat{X} \leftarrow X^* - x^g$, $\hat{Y} \leftarrow Y^* - y^g$
for $[a,b) \in \mathcal{Y}$ with $y^g_{[a,b)} = 1$ **do**
5: **for** $t \in [a,b) \cap \mathcal{T}$ **do**
$p_t^{g,[a,b)} \leftarrow P_t^{*[a,b)}/Y^*_{[a,b)}$; $\hat{P}_t^{[a,b)} \leftarrow P_t^{*[a,b)} - p_t^{g,[a,b)}$
$r_t^{g,[a,b)} \leftarrow R_t^{*[a,b)}/Y^*_{[a,b)}$; $\hat{R}_t^{[a,b)} \leftarrow R_t^{*[a,b)} - r_t^{g,[a,b)}$
for $l \in \mathcal{L}$ **do**
$p_t^{g,[a,b),l} \leftarrow P_t^{*[a,b),l}/Y^*_{[a,b)}$; $\hat{P}_t^{[a,b),l} \leftarrow P_t^{*[a,b),l} - p_t^{g,[a,b),l}$
10: **for** $t \in \mathcal{T}$ **do**
$\hat{P}_t \leftarrow \sum_{[a,b)\in\mathcal{Y}} \hat{P}_t^{[a,b)}$
$\hat{R}_t \leftarrow \sum_{[a,b)\in\mathcal{Y}} \hat{R}_t^{[a,b)}$
$p_t^g \leftarrow \sum_{[a,b)\in\mathcal{Y}} p_t^{g,[a,b)}$
$r_t^g \leftarrow \sum_{[a,b)\in\mathcal{Y}} r_t^{g,[a,b)}$

– 9) are exactly of the type prescribed by Lemma 1, and so both these and the $\hat{P}^{[a,b)}$, $\hat{R}^{[a,b)}$, and $\hat{P}^{[a,b),l}$ variables are feasible for (16a) and (19a) respectively. Theorem 1 ensures this assignment is optimal as well. The equalities (16b, 16c) and (19b, 19c) follow from lines 11 – 14.

The last statement follows from inducting on the size of $\mathcal{K}$. □

C Disaggregating the 3-bin polytope

Recalling the traditional 3-bin formulation for fast ramping generators (when $\mathbf{UT}^g \geq 2$) [5, 6]:

$$\underline{\mathbf{P}}^g u_t^g \leq p_t^g, \qquad \forall t \in \mathcal{T}, \tag{20a}$$

$$p_t^g + r_t^g \leq \overline{\mathbf{P}}^g u_t^g + (\mathbf{SU}^g - \overline{\mathbf{P}}^g) v_t^g + (\mathbf{SD}^g - \overline{\mathbf{P}}^g) w_{t+1}^g, \qquad \forall t \in \mathcal{T}, \tag{20b}$$

$$u_t^g - u_{t-1}^g = v_t^g - w_t^g, \qquad \forall t \in \mathcal{T}, \tag{20c}$$

$$\sum_{i=t-\mathbf{UT}+1}^{t} v_i^g \leq u_t^g, \qquad \forall t \in [\mathbf{UT}^g, \mathbf{T}], \tag{20d}$$

$$\sum_{i=t-\mathbf{DT}+1}^{t} w_i^g \leq 1 - u_t^g, \qquad \forall t \in [\mathbf{DT}^g, \mathbf{T}], \tag{20e}$$

$$p_t^g,\ r_t^g \in \mathbb{R}_+, \qquad \forall t \in \mathcal{T}, \tag{20f}$$

$$u_t^g,\ v_t^g,\ w_t^g \in \{0, 1\}, \qquad \forall t \in \mathcal{T}. \tag{20g}$$

It has the property that the constraint matrix defined by (20c, 20d, 20e) is totally unimodular [18], and so it too has the integer decomposition property [24]. When $\mathbf{UT}^g = 1$, (20b) is replaced with the following [5,6]:

$$p_t^g + r_t^g \leq \overline{\mathbf{P}}^g u_t^g + (\mathbf{SU}^g - \overline{\mathbf{P}}^g) v_t^g, \qquad \forall t \in \mathcal{T}, \tag{21a}$$

$$p_t^g + r_t^g \leq \overline{\mathbf{P}}^g u_t^g + (\mathbf{SD}^g - \overline{\mathbf{P}}^g) w_{t+1}^g, \qquad \forall t \in \mathcal{T}. \tag{21b}$$

C.1 Generator Production Cost Function

The convex production costs are typically approximated by piecewise linear costs. This is done by partitioning the interval $[\underline{\mathbf{P}}, \overline{\mathbf{P}}]$ into L subintervals with breakpoints $\overline{\mathbf{P}}^l$, with $\overline{\mathbf{P}}^0 = \underline{\mathbf{P}}$, $\overline{\mathbf{P}}^L = \overline{\mathbf{P}}$, and $\overline{\mathbf{P}}^l < \overline{\mathbf{P}}^{l+1}$. Let $\mathbf{c}^l$ be the marginal cost for the segment $[\overline{\mathbf{P}}^{l-1}, \overline{\mathbf{P}}^l]$. The variable c_t, then, represents the production cost for time t given the following constraints:

$$c_t = \sum_{l=1}^{k} \mathbf{c}^l p_t^l \qquad \forall t \in \mathcal{T} \tag{22a}$$

$$p_t = \underline{\mathbf{P}} u_t + \sum_{l=1}^{k} p_t^l \qquad \forall t \in \mathcal{T} \tag{22b}$$

$$0 \leq p_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1}) u_t \qquad \forall l \in \mathcal{L},\ \forall t \in \mathcal{T} \tag{22c}$$

The cost functions of the generators need to be modified to account for the aggregation. Fortunately, using the intuition behind Theorem 1, the lack of ramping constraints ensure if two identical generators are on in a given time period, they must have the same production, meaning that allowing u to be a general integer is also sufficient. The only exception to this rule is when there is a startup/shutdown rate. Without loss of generality assume $\mathbf{SU} = \mathbf{SD} = \overline{\mathbf{P}}^{l'}$. We substitute (22c) by

$$p_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1}) u_t \qquad \forall l \in [l'],\ \forall t \in \mathcal{T} \tag{23a}$$

$$p_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(u_t - v_t - w_{t+1}) \qquad \forall l > l',\ \forall t \in \mathcal{T} \tag{23b}$$

$$p_t^l \geq 0 \qquad \forall l \in \mathcal{L},\ \forall t \in \mathcal{T} \tag{23c}$$

when $\mathbf{UT} \geq 2$ and

$$p_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1}) u_t \qquad \forall l \in [l'],\ \forall t \in \mathcal{T} \tag{24a}$$

$$p_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(u_t - v_t) \qquad \forall l > l',\ \forall t \in \mathcal{T} \tag{24b}$$

$$p_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(u_t - w_{t+1}) \qquad \forall l > l',\ \forall t \in \mathcal{T} \tag{24c}$$

$$p^l_t \geq 0 \qquad \forall l \in \mathcal{L},\ \forall t \in \mathcal{T} \tag{24d}$$

when $\mathbf{UT} = 1$. When the generator has just turned on (about to turn off), then its power output cannot be above $\mathbf{SU}$ ($\mathbf{SD}$). Hence constraints of the form (23b) and (24b, 24c) cut off these solutions in the p^l variables just as (20b) and (21a, 21b) do for the p variables.

C.2 Generator Startup Costs

There are many different proposed formulations for time-dependent startup costs. One of which, [37], provides a perfect (and totally unimodular) formulation. It is:

$$\sum_{\{t' | [t,t') \in \mathcal{Y}\}} y_{[t,t')} = v_t \qquad \forall t \in \mathcal{T} \tag{25a}$$

$$\sum_{\{t' | [t',t) \in \mathcal{Y}\}} y_{[t',t)} = w_t \qquad \forall t \in \mathcal{T} \tag{25b}$$

$$\sum_{\{t' | [t',t) \in \mathcal{X}\}} x_{[t',t)} = v_t \qquad \forall t \in \mathcal{T} \tag{25c}$$

$$\sum_{\{t' | [t,t') \in \mathcal{X}\}} x_{[t,t')} = w_t \qquad \forall t \in \mathcal{T} \tag{25d}$$

$$\sum_{\{[\tau,\tau') \in \mathcal{Y} | t \in [\tau,\tau')\}} y_{[\tau,\tau')} = u_t \qquad \forall t \in \mathcal{T}. \tag{25e}$$

Like the extended formulation in the generator model, this formulation can be quite large, containing $O(T^2)$ many variables. However, after adding these constraints to 3-bin model for fast-ramping generators, the resulting formulation still satisfies the integer decomposition property, and we can show this aggregation is valid for the generator schedule.

[25] suggest a more compact formulation of startup costs:

$$\sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} x_{[t',t)} \leq v_t \qquad \forall t \in \mathcal{T}, \tag{26a}$$

$$\sum_{t'=t+\mathbf{DT}}^{t+\mathbf{TC}-1} x_{[t,t')} \leq w_t \qquad \forall t \in \mathcal{T}, \tag{26b}$$

(where the sums are understood to be taken over valid t') and the objective function is

$$c^{SU}_t = \mathbf{c}^S v_t + \sum_{s=1}^{S-1} (\mathbf{c}^s - \mathbf{c}^S) \left(\sum_{t'=t-\overline{\mathbf{T}}^s+1}^{t-\underline{\mathbf{T}}^s} x_{[t',t)} \right) \qquad \forall t \in \mathcal{T}, \tag{26c}$$

and dropping the remaining $x_{[t,t')}$ for which $t' \geq t + \mathbf{TC}$. The resulting computational experiments suggest that this formulation dominates (25) computationally, in addition to the other formulations examined. The reason is that while (26) is not an integer polytope, it is integer "in the right direction," that is, any fractional vertex for 3-bin models using (26) will be dominated by an integer solution for all reasonable objective coefficients (the fractionally enters the formulation by allowing cooler startups to be assigned even when the generator is still hot).

C.3 Disaggregating schedules

In this section we will show how to decompose solutions to the aggregated formulation with various common features, such as time-dependent startup costs, reserves, and piecewise linear production costs. Suppose $\mathcal{K} \subset \mathcal{G}$ such that all generators in $\mathcal{K}$ have identical properties (save initial status). Let $U = \sum_{g \in \mathcal{K}} u^g$, $V = \sum_{g \in \mathcal{K}} v^g$, $W = \sum_{g \in \mathcal{K}} w^g$, and $X = \sum_{g \in \mathcal{K}} x^g$. Let $\mathbf{UT} = \mathbf{UT}^g$, $\mathbf{DT} = \mathbf{DT}^g$, and $\mathbf{TC} = \mathbf{TC}^g$ for some (every) $g \in \mathcal{K}$.

First consider the aggregated 3-bin model for commitment status with startup costs:

$$U_t - U_{t-1} = V_t - W_t \qquad \forall t \in \mathcal{T} \tag{27a}$$

$$\sum_{i=t-\mathbf{UT}+1}^{t} V_i \leq U_t \qquad \forall t \in [\mathbf{UT}, \mathbf{T}] \tag{27b}$$

$$\sum_{i=t-\mathbf{DT}+1}^{t} W_i \leq |\mathcal{K}| - U_t \qquad \forall t \in [\mathbf{DT}, \mathbf{T}] \tag{27c}$$

$$\sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} X_{[t',t)} \leq V_t \qquad \forall t \in \mathcal{T}, \forall g \in \mathcal{G} \tag{27d}$$

$$\sum_{t'=t+\mathbf{DT}}^{t+\mathbf{TC}-1} X_{[t,t')} \leq W_t \qquad \forall t \in \mathcal{T}, \forall g \in \mathcal{G} \tag{27e}$$

$$U_t, V_t, W_t \in \{0, \ldots, |\mathcal{K}|\} \qquad \forall t \in \mathcal{T} \tag{27f}$$

$$X_{[t,t')} \in \{0, \ldots, |\mathcal{K}|\} \qquad \forall\, t, t' \in \mathcal{T}^2 \text{ with } \mathbf{DT} \leq t' - t < \mathbf{TC}. \tag{27g}$$

Algorithm 4 demonstrates how to disaggregate a solution to (27). We establish the correctness of Algorithm 4 with the next theorem.

Algorithm 4 (PeelOff) Constructs feasible generator schedules from a solution of (27).

Initialize all u^g, v^g, w^g, x^g to 0.
Initialize all $\hat{U}$, $\hat{V}$, $\hat{W}$, $\hat{X}$ to U^*, V^*, W^*, X^*, respectively.
if $U_1^* \geq 1$ **then**
  $u_1^g \leftarrow 1$; $\hat{U}_1 \leftarrow U_1^* - 1$;
  Assign historical startup v_t, for $-\mathbf{UT} < t \leq 0$.
else
  Assign historical shutdown w_t, for $-\mathbf{DT} < t \leq 0$.
$t \leftarrow 2$
while $t \leq \mathbf{T}$ **do**
  if $u_{t-1}^g = 1$ **then** ▷ If on in the previous period
    if $\sum_{i=t-\mathbf{UT}+1}^{t-1} v_i = 0$ **and** $W_t^* \geq 1$ **then** ▷ If we can turn off and a turn off is available
      $w_t^g \leftarrow 1$; $\hat{W}_t \leftarrow W_t^* - 1$; ▷ Turn off
      if $\exists t'$ s.t. $t + TD \leq t' < t + \mathbf{TC}$ **and** $X_{[t,t')}^* \geq 1$ **then** ▷ If there is a hot-start available
        $u_{t'}^g, v_{t'}^g \leftarrow 1$; $x_{[t,t')}^g \leftarrow 1$; ▷ Take it
        $\hat{U}_{t'} \leftarrow U_{t'}^* - 1$; $\hat{V}_{t'} \leftarrow V_{t'}^* - 1$;
        $\hat{X}_{[t,t')} \leftarrow X_{[t,t')}^* - 1$;
        $t \leftarrow t' + 1$;
      else if $\exists t' \geq t + \mathbf{TC}$ s.t. $\sum_{t'=t+\mathbf{DT}}^{t+\mathbf{TC}-1} X_{[t,t')}^* < V_{t'}^*$ **then** ▷ Else take some cold start, if possible
        $u_{t'}^g, v_{t'}^g \leftarrow 1$; $\hat{U}_{t'} \leftarrow U_{t'}^* - 1$; $\hat{V}_{t'} \leftarrow V_{t'}^* - 1$;
        $t \leftarrow t' + 1$;
      else ▷ If not, stay off for the rest of the time horizon
        $t \leftarrow \mathbf{T} + 1$;
    else ▷ If we could not turn off or a turn off was not available
      $u_t \leftarrow 1$; $\hat{U}_t \leftarrow U_t^* - 1$; ▷ Stay on
      $t \leftarrow t + 1$;
  else ▷ $u_{t-1} = 0$, i.e., off previously
    if $\sum_{i=t-\mathbf{DT}+1}^{t-1} w_i = 0$ **and** $V_t^* \geq 1$ **then** ▷ If we can turn on and a turn on is available
      $u_t, v_t \leftarrow 1$; $\hat{U}_t \leftarrow U_t^* - 1$; $\hat{V}_t \leftarrow V_t^* - 1$; ▷ Turn on
      if $\exists t' \in (t - \mathbf{TC}, t - \mathbf{DT}] \cap \mathbb{Z}$ s.t. $X_{[t',t)}^* \geq 1$ **then** ▷ If there is a historical hot-start
        $x_{[t',t)}^g \leftarrow 1$; $\hat{X}_{[t',t)} \leftarrow X_{[t',t)}^* - 1$; ▷ Take it
        $w_{t'}^g \leftarrow 1$; $\hat{W}_{t'} \leftarrow W_{t'}^* - 1$; ▷ Assign historical data for $t \leq -\mathbf{DT}$
      $t \leftarrow t + 1$;
    else ▷ If no turn on feasible or available
      $t \leftarrow t + 1$; ▷ Stay off

Theorem 7. *Suppose (U^*, V^*, W^*, X^*) is a feasible solution for (27). Then for every $g \in \mathcal{K}$ there exist $(u^{g*}, v^{g*}, w^{g*}, x^{g*})$ feasible for the minimum up-time/down-time system for $\mathcal{K}$:*

$$
\begin{aligned}
& u_t^g - u_{t-1}^g = v_t^g - w_t^g && \forall t \in \mathcal{T},\ \forall g \in \mathcal{K} \\
& \sum_{i=t-\mathbf{UT}+1}^{t} v_i^g \leq u_t^g && \forall t \in [\mathbf{UT}, \mathbf{T}],\ \forall g \in \mathcal{K} \\
& \sum_{i=t-\mathbf{DT}+1}^{t} w_i^g \leq 1 - u_t^g && \forall t \in [\mathbf{DT}, \mathbf{T}],\ \forall g \in \mathcal{K} \\
& \sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} x_{[t',t)} \leq v_t^g && \forall t \in \mathcal{T},\ \forall g \in \mathcal{K} \\
& \sum_{t'=t+\mathbf{DT}}^{t+\mathbf{TC}-1} x_{[t,t')} \leq w_t^g && \forall t \in \mathcal{T},\ \forall g \in \mathcal{K} \\
& u_t^g, v_t^g, w_t^g \in \{0,1\} && \forall t \in \mathcal{T},\ \forall g \in \mathcal{K} \\
& x_{[t,t')} \in \{0, \ldots, |\mathcal{K}|\} && \forall\ t, t' \in \mathcal{T}^2 \text{ with } \mathbf{DT} \leq t' - t < \mathbf{TC}.
\end{aligned}
$$

Proof. Clearly this is true when $|\mathcal{K}| = 1$. We will proceed by induction on the size of $\mathcal{K}$, "peeling off" feasible binary vectors and leaving behind a still feasible crushed system. Suppose (U^*, V^*, W^*, X^*) is feasible (27), and $|\mathcal{K}| = I$. We wish to find a feasible solution to the following system:

$$
\begin{aligned}
& U_t^* = \hat{U}_t + u_t^g,\ V_t^* = \hat{V}_t + v_t^g,\ W_t^* = \hat{W}_t + w_t^g && \forall t \in \mathcal{T} && \text{(28a)} \\
& X_{[t,t')}^* = \hat{X}_{[t,t')} + x_{[t,t')}, && \forall\ t, t' \in \mathcal{T}^2 \text{ with } \mathbf{DT} \leq t' - t < \mathbf{TC} && \text{(28b)}
\end{aligned}
$$

$$
\begin{aligned}
& u_t^g - u_{t-1}^g = v_t^g - w_t^g && \forall t \in \mathcal{T} && \text{(29a)} \\
& \sum_{i=t-\mathbf{UT}+1}^{t} v_i^g \leq u_t^g && \forall t \in [\mathbf{UT}, \mathbf{T}] && \text{(29b)} \\
& \sum_{i=t-\mathbf{DT}+1}^{t} w_i^g \leq 1 - u_t^g && \forall t \in [\mathbf{DT}, \mathbf{T}] && \text{(29c)} \\
& \sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} x_{[t',t)} \leq v_t^g && \forall t \in \mathcal{T} && \text{(29d)} \\
& \sum_{t'=t+\mathbf{DT}}^{t+\mathbf{TC}-1} x_{[t,t')} \leq w_t^g && \forall t \in \mathcal{T} && \text{(29e)} \\
& u_t^g, v_t^g, w_t^g \in \{0,1\}\ \forall t \in \mathcal{T}, \quad x_{[t,t')}^g \in \{0,1\},\ \forall\ t, t' \in \mathcal{T}^2 \text{ with } \mathbf{DT} \leq t' - t < \mathbf{TC} && && \text{(29f)}
\end{aligned}
$$

$$
\begin{aligned}
& \hat{U}_t - \hat{U}_{t-1} = \hat{V}_t - \hat{W}_t && \forall t \in \mathcal{T} && \text{(30a)}\\
& \sum_{i=t-\mathbf{UT}+1}^{t} \hat{V}_i \le \hat{U}_t && \forall t \in [\mathbf{UT}, \mathbf{T}] && \text{(30b)}\\
& \sum_{i=t-\mathbf{DT}+1}^{t} \hat{W}_i \le (I-1) - \hat{U}_t && \forall t \in [\mathbf{DT}, \mathbf{T}] && \text{(30c)}\\
& \sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} \hat{X}_{[t',t)} \le \hat{V}_t && \forall t \in \mathcal{T} && \text{(30d)}\\
& \sum_{t'=t+\mathbf{DT}}^{t+\mathbf{TC}-1} \hat{X}_{[t,t')} \le \hat{W}_t && \forall t \in \mathcal{T} && \text{(30e)}\\
& \hat{U}_t, \hat{V}_t, \hat{W}_t \in \{0, \dots, I-1\}\ \forall t \in \mathcal{T}, \quad \hat{X}_{[t,t')} \in \{0, \dots, I-1\} \forall\ t, t' \in \mathcal{T}^2 \text{ with } \mathbf{DT} \le t' - t < \mathbf{TC}. && && \text{(30f)}
\end{aligned}
$$

Algorithm 4 constructs a feasible solution to (28, 29, 30) from a solution of (27). To see this, first notice that the solution returned by Algorithm 4 always has (28). Similarly, Algorithm 4 constructs a feasible solution for (u^g, v^g, w^g, x^g), so (29) holds. Further, (27a) and (29a) together imply (30a). Notice that given the bounds on $\hat{U}$ and (30b - 30e), we get the bounds on $\hat{V}$, $\hat{W}$, and $\hat{X}$, and the proof is finished. Therefore we check the bounds on $\hat{U}$ and (30b–30e), proceeding by contraction each time.

$\hat{U}_t \le I-1$: Let t be the first time period such that $\hat{U}_t > I-1$ (notice $t > 1$ by line 8 in Algorithm 4). Then $\hat{U}_t = I$ and further, $u_t = 0$, $U_t^* = I$. Now by (27c), $\sum_{i=t-\mathbf{DT}+1}^{t} W_i^* \le I - I = 0$. Therefore $W_i^* = 0,\ \forall i \in [t-\mathbf{DT}+1, t]$, yielding $w_i^g = 0,\ \forall i \in [t-\mathbf{DT}+1, t]$. Since $U_t^* = I$ and $W_t^* = 0$, (27a) gives $I = U_{t-1}^* + V_t^*$. If $V_t^* > 0$, since g is eligible for a turn-on, this contracts line 27 in Algorithm 4. If $V_t^* = 0$, then $U_{t-1}^* > I-1$, contracting the minimality of t.

(30b): Suppose there is t with $\sum_{i=t-\mathbf{UT}+1} \hat{V}_i > \hat{U}_i$. We have the following relation:

$$
U_t^* \overset{v_i^g \ge 0}{\ge} U_t^* - \sum_{i=t-\mathbf{UT}+1}^{t} v_i^g \overset{(27b)}{\ge} \sum_{i=t-\mathbf{UT}+1}^{t} (V_i^* - v_i^g) \overset{(28a)}{=} \sum_{i=t-\mathbf{UT}+1}^{t} \hat{V}_t > \hat{U}_t \overset{(28a)}{=} U_t^* - u_t^g \overset{u_t^g \le 1}{\ge} U_t^* - 1 \tag{31}
$$

Since the far left and far right differ by only 1, and all quantities are integer, in order for the strict inequality to hold, all weak inequalities in (31) must be equalities. Therefore we have (a) $u_t^g = 1$, (b) $\sum_{i=t-\mathbf{UT}+1}^{t} v_i = 0$, and (c) $\sum_{i=t-\mathbf{UT}+1}^{t} V_i^* = U_t^*$. Together (a) and (b) imply $u_i^g = 1$ and $U_i^* \ge 1$ for every $i \in \{t-\mathbf{UT}, \dots, t\}$. Therefore generator g started-up most recently at some time $\hat{i}$ such that $\hat{i} \le t - \mathbf{UT}$; i.e $v_{\hat{i}} = 1$. Line 11 of Algorithm 4 imply then that $W_i^* = 0$ for every $i \in \{\hat{i} + \mathbf{UT}, \dots, t\}$, and hence (d) $U_i^* - U_{i-1}^* = V_i^*$ for every $i \in \{\hat{i} + \mathbf{UT}, \dots, t\}$ by equation (27a). There are but two cases then.

Case 1. Suppose $\hat{i} + \mathbf{UT} \leq t - \mathbf{UT} + 1$. Then $\sum_{i=t-\mathbf{UT}+1}^{t} V_i^* = U_t^* - U_{t-\mathbf{UT}}^*$ by (d). By (c) then $U_{t-\mathbf{UT}}^* = 0$ but (a) and (b) give $U_{t-\mathbf{UT}}^* \geq 1$.

Case 2. Suppose $t - \mathbf{UT} + 1 < \hat{i} + \mathbf{UT}$. We have:

$$U_t^* \overset{(c)}{=} \sum_{i=t-\mathbf{UT}+1}^{t} V_i^* = \sum_{i=t-\mathbf{UT}+1}^{\hat{i}+\mathbf{UT}-1} V_i^* + \sum_{i=\hat{i}+\mathbf{UT}}^{t} V_i^* \overset{(d)}{=} \sum_{i=t-\mathbf{UT}+1}^{\hat{i}+\mathbf{UT}-1} V_i^* + U_t^* - U_{\hat{i}+\mathbf{UT}-1}^* \tag{32}$$

Subtracting U_t^* from both sides we get the relation $0 = \sum_{i=t-\mathbf{UT}+1}^{\hat{i}+\mathbf{UT}-1} V_i^* - U_{\hat{i}+\mathbf{UT}-1}^*$. Finally then, we see

$$A0 = \sum_{i=t-\mathbf{UT}+1}^{\hat{i}+\mathbf{UT}-1} V_i^* - U_{\hat{i}+\mathbf{UT}-1}^* \overset{(27b)}{\leq} \sum_{i=t-\mathbf{UT}+1}^{\hat{i}+\mathbf{UT}-1} V_i^* - \sum_{i=\hat{i}}^{\hat{i}+\mathbf{UT}-1} V_i^* = - \sum_{i=\hat{i}}^{t-\mathbf{UT}} V_i^* \overset{v_{\hat{i}}=1}{\leq} -1, \tag{33}$$

yielding the contraction desired.

(30c): Very similar to the proof for (30b).

(30d): Supposing there is t with $\sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} \hat{X}_{[t',t)} > \hat{V}_t$ we can use the same technique from the proof of (30b) to get the following string of inequalities:

$$\begin{aligned} V_t^* \geq V_t^* - \sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} x_{[t',t)}^g &\overset{(27d)}{\geq} \sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} X_{[t',t)}^* - \sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} x_{[t',t)}^g \\ \overset{(28b)}{=} \sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} \hat{X}_{[t',t)} > \hat{V}_t \overset{(28a)}{=} V_t^* - v_t^g \geq V_t^* - 1 \end{aligned} \tag{34}$$

Hence we may conclude (a) $v_t^g = 1$, (b) $\sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} x_{[t',t)}^g = 0$, and (c) $\sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} X_{[t',t)}^* = V_t^*$. We have (a) implies that $V_t^* \geq 1$, so $\sum_{t'=t-\mathbf{TC}+1}^{t-\mathbf{DT}} X_{[t',t)}^* \geq 1$, so $\exists t' \in \{t - \mathbf{DT} + 1, \ldots, t - \mathbf{TC}\}$ such that $X_{[t',t)}^* \geq 1$. But line 14 in Algorithm 4 then sets $x_{[t',t)}^g = 1$, a contraction.

(30e): This is the same as (30d).

Hence Algorithm 4 constructs a feasible solution for (28, 29, 30). We can then proceed in this manner until $I = 1$, proving the theorem. □

Notice that Algorithm 4 constructs solutions exactly of the type in Theorems 2 and 4. Theorem 5 justifies the arbitrary choice in lines 13 – 20 for which shutdown/startup path the generator takes. Hence we do not loose anything in optimality or feasibility by aggregating a fast-ramping generator's status variables, even in the presence of time-dependent startup costs. Next we will see how to disaggregate the power and reserve variables.

C.4 Disaggregating Power and Reserves

C.4.1 Disaggregating Power when $\mathbf{UT} \geq 2$

Consider the power output for an aggregated set of identical fast-ramping generators. First we will consider the case when $\mathbf{UT} \geq 2$. Along with (27), we the aggregated power $P = \sum_{g \in \mathcal{K}} p^g$ and reserves $R = \sum_{g \in \mathcal{K}} r^g$ with the constraints:

$$\underline{\mathbf{P}} U_t \leq P_t \qquad \forall t \in \mathcal{T}, \tag{35a}$$

$$P_t + R_t \leq \overline{\mathbf{P}} U_t + (\mathbf{SU} - \overline{\mathbf{P}}) V_t + (\mathbf{SD} - \overline{\mathbf{P}}) W_{t+1}, \qquad \forall t \in \mathcal{T}. \tag{35b}$$

Since (35) is just a sum of constraints, it is clearly valid.

Algorithm 5a (Peel Off Power) Constructs feasible generator schedule from a solution of (35) when $\mathbf{SU} \geq \mathbf{SD}$.

for $g \in \mathcal{K},\ t \in \mathcal{T}$ **do**
  if $P_t^*/U_t^* \leq \mathbf{SD}$ **then** ▷ If the average power is less than $\mathbf{SD}$
    if $u_t^g = 1$ **then**
      $p_t^g \leftarrow P_t^*/U_t^*$ ▷ Give all generators on average power
    else
      $p_t^g \leftarrow 0$
  else if $(P_t^* - \mathbf{SD} \cdot W_{t+1}^*)/U_t^* \leq \mathbf{SU}$ **then** ▷ If not, check if remaining average power $\leq \mathbf{SU}$
    if $w_{t+1}^g = 1$ **then**
      $p_t^g \leftarrow \mathbf{SD}$ ▷ Give generators shutting down $\mathbf{SD}$
    else if $u_t^g = 1$ **then**
      $p_t^g \leftarrow (P_t^* - \mathbf{SD} \cdot W_{t+1}^*)/U_t^*$ ▷ Give all others on remaining average power
    else
      $p_t^g \leftarrow 0$
  else ▷ $(P_t^* - \mathbf{SD} \cdot W_{t+1}^*)/U_t^* > \mathbf{SU}$, so we need separate out generators starting
    if $w_{t+1}^g = 1$ **then**
      $p_t^g \leftarrow \mathbf{SD}$ ▷ Give generators shutting down $\mathbf{SD}$
    else if $v_t^g = 1$ **then**
      $p_t^g \leftarrow \mathbf{SU}$ ▷ Give (remaining) generators starting up $\mathbf{SU}$
    else if $u_t^g = 1$ **then**
      $p_t^g \leftarrow (P_t^* - \mathbf{SU} \cdot V_t^* - \mathbf{SD} \cdot W_{t+1}^*)/U_t^*$ ▷ all others on get remaining average power
    else
      $p_t^g \leftarrow 0$

If $\mathbf{SU} \geq \mathbf{SD}$, then Algorithm 5a demonstrates how to disaggregate power. On the other hand, when $\mathbf{SD} \geq \mathbf{SU}$ disaggregation can be done in an analogous fashion, as shown in Algorithm 5b. The essential logic of Algorithms 5a and 5b is that of Theorem 1. That is, either the power outputs of all generators are equal, or either the startup- and/or shutdown-ramping constraints are active. When $\mathbf{SU} = \mathbf{SD}$ then Algorithms 5a and 5b give the same result.

Algorithm 5b (Peel Off Power) Constructs feasible generator schedule from a solution of (35) when $\mathbf{SD} \geq \mathbf{SU}$.

for $g \in \mathcal{K},\ t \in \mathcal{T}$ **do**
  if $P_t^*/U_t^* \leq \mathbf{SU}$ **then** ▷ If the average power is less than **SU**
    if $u_t^g = 1$ **then**
      $p_t^g \leftarrow P_t^*/U_t^*$ ▷ Give all generators on average power
    else
      $p_t^g \leftarrow 0$
  else if $(P_t^* - \mathbf{SU} \cdot V_t^*)/U_t^* \leq \mathbf{SD}$ **then** ▷ If not, check if remaining average power $\leq$ **SD**
    if $v_t^g = 1$ **then**
      $p_t^g \leftarrow \mathbf{SU}$ ▷ Give generators starting up **SU**
    else if $u_t^g = 1$ **then**
      $p_t^g \leftarrow (P_t^* - \mathbf{SU} \cdot V_t^*)/U_t^*$ ▷ Give all others on remaining average power
    else
      $p_t^g \leftarrow 0$
  else ▷ $(P_t^* - \mathbf{SU} \cdot V_t^*)/U_t^* > \mathbf{SD}$, so we need separate out generators stopping
    if $v_t^g = 1$ **then**
      $p_t^g \leftarrow \mathbf{SU}$ ▷ Give generators starting up **SU**
    else if $w_{t+1}^g = 1$ **then**
      $p_t^g \leftarrow \mathbf{SD}$ ▷ Give (remaining) generators shutting down **SD**
    else if $u_t^g = 1$ **then**
      $p_t^g \leftarrow (P_t^* - \mathbf{SU} \cdot V_t^* - \mathbf{SD} \cdot W_{t+1}^*)/U_t^*$ ▷ all others on get remaining average power
    else
      $p_t^g \leftarrow 0$

C.4.2 Disaggregating Power when $\mathbf{UT} = 1$

When $\mathbf{UT} = 1$, we need consider a modified version of the aggregated generator's production constraint, as is the case for a single generator [5, 6]. Again using the aggregated variables from before, consider the aggregated production constraints

$$\underline{\mathbf{P}}U_t \leq P_t \qquad \forall t \in \mathcal{T} \tag{36a}$$

$$P_t + R_t \leq \overline{\mathbf{P}}U_t + (\mathbf{SU} - \overline{\mathbf{P}})V_t, \qquad \forall t \in \mathcal{T} \tag{36b}$$

$$P_t + R_t \leq \overline{\mathbf{P}}U_t + (\mathbf{SD} - \overline{\mathbf{P}})W_{t+1}, \qquad \forall t \in \mathcal{T}. \tag{36c}$$

When $\mathbf{SU} \geq \mathbf{SD}$, we can again use Algorithm 5a, except we need modify line 20 to

$$p_t^g \leftarrow (P_t^* - \mathbf{SU} \cdot \min\{V_t^* - W_{t+1}^*,\ 0\} - \mathbf{SD} \cdot W_{t+1}^*)/U_t^*. \tag{37}$$

The correctness of Algorithm 5a the modification above to line 20 follows from Theorem 1 and Algorithm 4. That is, $\min\{V_t^*,\ W_{t+1}^*\}$ generators turn on at time t and turn off at time $t+1$, and hence $\min\{V_t^* - W_{t+1}^*,\ 0\}$ will get $\mathbf{SU}$ power at line 18.

Similarly when $\mathbf{SD} \geq \mathbf{SU}$, we can use Algorithm 5b, modifying line 20 to

$$p_t^g \leftarrow (P_t^* - \mathbf{SU} \cdot V_t^* - \mathbf{SD} \cdot \min\{W_{t+1}^* - V_t^*,\ 0\})/U_t^*. \tag{38}$$

The correctness of Algorithm 5b with the modification to line 20 is exactly analogous to that in the $\mathbf{SU} \geq \mathbf{SD}$ case above.

C.4.3 Disaggregating Reserves

Having disaggregated power, reserves r^g can be disaggregated by considering $(p_t^g + r_t^g)$ and $(P_t^* + R_t^*)$ in Algorithm 5a or 5b (or their modified analogs when $\mathbf{UT} = 1$) above in place of p_t^g and P_t^* respectively.

C.5 Disaggregating Piecewise Linear Production Costs

C.5.1 $\mathbf{UT} \geq 2$

In the case of piecewise production costs we can modify (23) by considering the aggregated piecewise production variables $P^l = \sum_{g \in \mathcal{K}} p^{l,g}$. We consider the sum of constraints of the form (23), recalling l' is such that $\mathbf{SU} = \mathbf{SD} = \overline{\mathbf{P}}^{l'}$

$$P_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})U_t \qquad \forall l \in [l'],\ \forall t \in \mathcal{T} \tag{39a}$$

$$P_t^l \le (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(U_t - V_t - W_{t+1}) \qquad \forall l > l',\ \forall t \in \mathcal{T} \tag{39b}$$

$$P_t^l \ge 0 \qquad \forall l \in \mathcal{L},\ \forall t \in \mathcal{T} \tag{39c}$$

If the generator is on and did not just turn on nor is about to turn off ($u_t^g = 1,\ v_t^g = 0,\ w_{t+1}^g = 0$), then $p_t^{l,g} = P_t^{*l}/U_t^*$. If the generator just turned on ($u_t^g = 1,\ v_t^g = 1,\ w_{t+1}^g = 0$), then

$$p_t^{l,g} = P_t^{*l}/U_t^* \qquad \forall l \in [l'],\ \forall t \in \mathcal{T} \tag{40a}$$

$$p_t^{l,g} = 0 \qquad \forall l > l',\ \forall t \in \mathcal{T}, \tag{40b}$$

and similarly if the generator is just about to turn off ($u_t^g = 1,\ v_t^g = 0,\ w_{t+1}^g = 1$) Notice that with the aggregated linking constraints for piecewise production

$$P_t = \underline{\mathbf{P}}U_t + \sum_{l\in\mathcal{L}} P_t^l \qquad \forall t \in \mathcal{T},$$

the assignment given by (40) is compatible with Algorithms 5a and 5b. $p_t^{l,g}$ is 0 for all l, t when the generator is off ($u_t^g = 0$).

Consider the case when $\mathbf{SU} > \mathbf{SD}$. Letting $\overline{\mathbf{P}}^{l^{\mathbf{SU}}} = \mathbf{SU}$ and $\overline{\mathbf{P}}^{l^{\mathbf{SD}}} = \mathbf{SD}$, ($l^{\mathbf{SU}} > l^{\mathbf{SD}}$) we may use a modified version of (39)

$$P_t^l \le (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})U_t \qquad \forall l \in [l^{\mathbf{SD}}],\ \forall t \in \mathcal{T} \tag{41a}$$

$$P_t^l \le (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(U_t - W_{t+1}) \qquad \forall l \in (l^{\mathbf{SD}}, l^{\mathbf{SU}}],\ \forall t \in \mathcal{T} \tag{41b}$$

$$P_t^l \le (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(U_t - V_t - W_{t+1}) \qquad \forall l > l^{\mathbf{SU}},\ \forall t \in \mathcal{T} \tag{41c}$$

$$P_t^l \ge 0 \qquad \forall l \in \mathcal{L},\ \forall t \in \mathcal{T}. \tag{41d}$$

Like before, if the generator did not just turn on nor is about to turn off ($u_t^g = 1,\ v_t^g = 0,\ w_{t+1}^g = 0$), then $p^{l,g} = P_t^{*l}/U_t^*$. If the generator just turned on ($u_t^g = 1,\ v_t^g = 1,\ w_{t+1}^g = 0$), then

$$p_t^{l,g} = P_t^{*l}/U_t^* \qquad \forall l \in [l^{\mathbf{SU}}],\ \forall t \in \mathcal{T} \tag{42a}$$

$$p_t^{l,g} = 0 \qquad \forall l > l^{\mathbf{SU}},\ \forall t \in \mathcal{T}. \tag{42b}$$

If the generator is just about to turn off ($u_t^g = 1,\ v_t^g = 0,\ w_{t+1}^g = 1$), then

$$p_t^{l,g} = P_t^{*l}/U_t^* \qquad \forall l \in [l^{\mathbf{SD}}],\ \forall t \in \mathcal{T} \tag{43a}$$

$$p_t^{l,g} = 0 \qquad \forall l > l^{\mathbf{SD}},\ \forall t \in \mathcal{T}. \tag{43b}$$

Finally, $p_t^{l,g} = 0$ for all l, t when the generator g is off ($u_t^g = 0$).

The case when $\mathbf{SD} > \mathbf{SU}$ can be handled similarly.

C.5.2 $\mathbf{UT} = 1$

When $\mathbf{UT} = 1$, the aggregated piecewise power constraints need to be modified as well. Under the assumption $\overline{\mathbf{P}}^{l'} = \mathbf{SU} = \mathbf{SD}$, consider the aggregated version of (24)

$$P_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})U_t \qquad \forall l \in [l'],\ \forall t \in \mathcal{T} \tag{44a}$$
$$P_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(U_t - V_t) \qquad \forall l > l',\ \forall t \in \mathcal{T} \tag{44b}$$
$$P_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(U_t - W_{t+1}) \qquad \forall l > l',\ \forall t \in \mathcal{T} \tag{44c}$$
$$P_t^l \geq 0 \qquad \forall l \in \mathcal{L},\ \forall t \in \mathcal{T}. \tag{44d}$$

Without loss of generality, we can assign the piecewise power making the same modifications that were necessary for the power production. In particular, we have that $p_t^{l,g} = P_t^{*l}/U_t^*$ if $u_t^g = 1,\ v_t^g = w_{t+1}^g = 0$. If is a startup or shutdown (or both), ($u_t^g = 1,\ v_t^g$ and/or $w_{t+1}^g = 1$), then we can use (40).

Assuming $\mathbf{SU} \neq \mathbf{SD}$, we can introduce $l^{\mathbf{SU}}$ and $l^{\mathbf{SD}}$ as before for (41)

$$P_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})U_t \qquad \forall l \in [\max\{l^{\mathbf{SU}}, l^{\mathbf{SD}}\}],\ \forall t \in \mathcal{T} \tag{45a}$$
$$P_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(U_t - V_t) \qquad \forall l > l^{\mathbf{SU}},\ \forall t \in \mathcal{T} \tag{45b}$$
$$P_t^l \leq (\overline{\mathbf{P}}^l - \overline{\mathbf{P}}^{l-1})(U_t - W_{t+1}) \qquad \forall l > l^{\mathbf{SD}},\ \forall t \in \mathcal{T} \tag{45c}$$
$$P_t^l \geq 0 \qquad \forall l \in \mathcal{L},\ \forall t \in \mathcal{T}. \tag{45d}$$

If we have $u_t^g = 1,\ v_t^g = 0,\ w_{t+1}^g = 0$, then $p_t^{l,g} = P_t^{*l}/U_t^{*l}$. Then there are three other cases to consider. Suppose $\mathbf{SU} > \mathbf{SD}$. If the generator is just starting and does not shutdown ($u_t^g = 1, v_t^g = 1, w_{t+1}^g = 0$), then (42) applies, and if the generator is shutting down ($u_t^g = 1,\ v_t^g = 0$ or $1,\ w_{t+1}^g = 1$), then (43) applies.

The case when $\mathbf{SD} > \mathbf{SU}$ is handled analogously. That is, if the generator is shutting down and did just startup ($u_t^g = 1,\ v_t^g = 0,\ w_{t+1}^g = 1$), then (43) is used, and if the generator is starting up ($u_t^g = 1,\ v_t^g = 1, w_t^g = 0$ or 1) then (42) applies.

D Additional Computational Tests

In this section we present some additional computational results to complement those in the main text.

D.1 Symmetry Breaking Inequalities

In addition to the formulations considered before, we consider the addition of the "S3" variables and inequalities from [10] to the base 3-bin UC formulation. Lima and Novais [10] propose introducing new variables yon^g which indicate if generator g ever turned on during the time horizon, along with inequalities to enforce this

$$\sum_{t \in \mathcal{T}} u_t^g \geq yon^g \qquad \forall g \in \mathcal{G} \tag{46}$$

$$u_t^g \leq yon^g \qquad \forall t \in \mathcal{T},\ \forall g \in \mathcal{G}. \tag{47}$$

They then propose the following two symmetry-breaking inequalities for each set of identical generators $\mathcal{K}$

$$yon^g \geq yon^{g+1} \qquad \forall g, g{+}1 \in \mathcal{K} \tag{48}$$

$$\sum_{t \in \mathcal{T}} u_t^g \geq \sum_{t \in \mathcal{T}} u_t^{g+1} \qquad \forall g, g{+}1 \in \mathcal{K}, \tag{49}$$

where $g, g{+}1 \in \mathcal{K}$ is understood to be two consecutive generators in $\mathcal{K}$. (48) enforces that if generator $g{+}1$ ever turns on then generator g does as well, and (49) enforces that generator g is scheduled in at least as many time periods as generator $g{+}1$. As pointed out in [10], this eliminates many, but not every, source of symmetry in UC.

D.2 Computational Results

Here we present computational results for the instances tested in the main text with the addition of the "S3" variables and inequalities from [10], which we label as "3-bin+SBC". The computational platform is as described in the main text.

D.2.1 CAISO Instances

In Table 8 we report the computational results for the CAISO instances described above. As we can see, in almost all cases the symmetry-breaking constraints are unhelpful. This is to be expected since these instances have a relatively tight root optimality gap, and the extra constraints serve to slow effective cut generation and heuristic search at the root node. Overall they serve to slow the solver down over 3-bin, though in one instance the 3-bin+SBC variant finds a high-quality solution fastest and with fewest nodes.

Table 8: Computational Results for CAISO UC Instances

	Time (s)			Nodes		
Instance	**3-bin**	**3-bin+SBC**	**3-bin+A**	**3-bin**	**3-bin+SBC**	**3-bin+A**
2014-09-01 0%	31.35	34.07	**14.25**	0	0	0
2014-12-01 0%	25.77	29.36	**12.38**	0	0	0
2015-03-01 0%	24.08	34.47	**14.27**	0	0	0
2015-06-01 0%	13.11	14.25	**8.50**	0	0	0
Scenario400 0%	27.29	35.99	**23.63**	0	0	0
2014-09-01 1%	20.52	29.38	**16.44**	0	0	0
2014-12-01 1%	38.48	86.41	**24.69**	95	566	**0**
2015-03-01 1%	21.75	35.76	**19.11**	0	0	0
2015-06-01 1%	39.87	30.42	**15.59**	47	**0**	**0**
Scenario400 1%	47.54	57.85	**44.63**	**0**	154	1438
2014-09-01 3%	81.47	92.60	**38.27**	**7**	696	122
2014-12-01 3%	65.01	120.03	**36.53**	1292	**95**	125
2015-03-01 3%	50.79	77.96	**25.04**	**0**	3	**0**
2015-06-01 3%	87.25	147.05	**41.23**	**0**	79	115
Scenario400 3%	131.28	147.61	**69.45**	2055	**140**	880
2014-09-01 5%	47.07	69.51	**30.95**	95	176	**7**
2014-12-01 5%	83.87	132.97	**66.90**	1203	**79**	3978
2015-03-01 5%	80.57	72.70	**21.65**	923	**0**	**0**
2015-06-01 5%	**26.99**	96.56	43.79	**0**	162	402
Scenario400 5%	115.53	**95.50**	118.51	3867	**31**	4225
Geometric Mean:	43.85	59.69	**27.55**			

D.2.2 Ostrowski Instances

As in the main text, we set a time limit of 900 seconds for the Ostrowski instances and report the terminating MIP gap in parentheses when the solver terminates at the time limit. In Table 9 we report the computational results for the Ostrowski instances from the main text. As reported in [10], the symmetry-breaking constraints are helpful overall for the smaller Ostrowski instances (1–10), but they perform worse than 3-bin on the larger instances (11–20). In comparison, EF/3-bin+A has a relatively flat performance profile across all 20 instances, suggesting that, when handled properly, identical generators can be leveraged to significantly reduce the computational burden of UC.

Table 9: Computational Results for Ostrowski UC Instances

	Time (s)			**Nodes**		
Instance	**3-bin**	**3-bin+SBC**	**EF/3-bin+A**	**3-bin**	**3-bin+SBC**	**EF/3-bin+A**
1	**8.44**	54.82	14.02	1509	13700	**68**
2	154.75	44.73	**21.07**	48129	10370	**157**
3	703.94	127.23	**100.33**	316704	25802	**4464**
4	**14.84**	109.15	17.28	8532	26297	**60**
5	143.18	101.73	**57.22**	131320	17784	**4350**
6	95.41	45.03	**28.00**	62394	7160	**72**
7	(0.0238%)	270.34	**119.22**	535361*	75097	**4854**
8	(0.0107%)	**57.36**	71.00	1378310*	29362	**9267**
9	(0.0169%)	167.56	**125.63**	819798*	49318	**12217**
10	(0.0327%)	287.18	**82.89**	751319*	133909	**11549**
11	(0.0186%)	(0.0220%)	**18.76**	73976*	20838*	**1155**
12	(0.0240%)	(0.0265%)	**22.91**	42729*	7505*	**460**
13	(0.0266%)	(0.0264%)	**74.43**	41325*	13420*	**6464**
14	(0.0144%)	(0.0203%)	**19.75**	41469*	4127*	**15**
15	780.76	(0.0105%)	**39.63**	120599	35111*	**3091**
16	(0.0162%)	(0.0223%)	**90.31**	42102*	8770*	**2597**
17	154.37	237.90	**27.88**	2114	2185	**1059**
18	(0.0121%)	(0.0214%)	**22.36**	60651*	3972*	**151**
19	(0.0195%)	(0.0250%)	**21.30**	42683*	6025*	**2436**
20	106.44	628.20	**18.46**	527	4288	**0**
Geometric Mean:	>349.77	>277.82	**38.03**			