

# A Dual Approximate Dynamic Programming Approach to Multi-stage Stochastic Unit Commitment

Jagdish Ramakrishnan<sup>1</sup> and James Luedtke<sup>2</sup>

June 22, 2018

## Abstract

We study the multi-stage stochastic unit commitment problem in which commitment and generation decisions can be made and adjusted in each time period. We formulate this problem as a Markov decision process, which is “weakly-coupled” in the sense that if the demand constraint is relaxed, the problem decomposes into a separate, low-dimensional, Markov decision process for each generator. We demonstrate how the dual approximate dynamic programming method of Barty, Carpentier, and Girardeau (*RAIRO Operations Research*, 44:167-183, 2010) can be adapted to obtain bounds and a policy for this problem. Previous approaches have let the Lagrange multipliers depend only on time; this can result in weak lower bounds. Other approaches have let the multipliers depend on the entire history of past random observations; although this provides a strong lower bound, its ability to handle a large number of sample paths or scenarios is limited. We demonstrate how to bridge these approaches for the stochastic unit commitment problem by letting the multipliers depend on the current observed demand. This allows a good tradeoff between strong lower bounds and good scalability with the number of scenarios. We illustrate this approach numerically on a 168-stage stochastic unit commitment problem, including minimum uptime, downtime, and ramping constraints.

## 1 Introduction

The unit commitment problem is an important problem in operation of power systems and has been studied extensively. Due to the presence of both integer and continuous variables, it remains a challenging problem to solve. The basic problem is to determine the on/off status and generation amounts of a collection of interconnected generators so that demands are met while minimizing the total generation cost. An important feature of the unit commitment problem is the generator constraints, which include constraints on the minimum and maximum generation amount, minimum and maximum number of consecutive periods the

---

<sup>1</sup>Walmart Labs, San Bruno, CA, jramakrishnan@walmartlabs.com. This work was completed while the author was with the Wisconsin Institute for Discovery, University of Wisconsin-Madison.

<sup>2</sup>Department of Industrial and Systems Engineering, University of Wisconsin-Madison, jim.luedtke@wisc.edu.

generator can be on or off (so called “min up/down constraints”), and bounds on the change in generation level from one period to the next (ramping constraints). In this paper, we consider a version of the problem in which there is a single aggregate amount of demand to be met in each time period. More complicated models also consider the transmission network and its associated constraints, to ensure the generated electricity can be feasibly distributed to the demand locations in the grid.

In the deterministic unit commitment problem, the future demands are modeled as known quantities. A significant amount of literature has focused on this problem, see e.g., [11, 13, 15, 2, 33]. Stochastic formulations model the future demands as a sequence of random variables. A sequence of possible demands over time is known as a demand scenario. As the number of demand scenarios grow, the optimization model becomes very challenging. While our discussion is limited to handling demand uncertainties, there are a number of other uncertainties that can be modeled in the unit commitment problem. For example, there are models that take into account generator failures [30], weather variations [29], price spikes in the spot market [16], and availability of renewable energy [3]. There is a vast amount of literature on the stochastic unit commitment problem, see e.g., [24, 19, 21, 31, 26]. A popular approach is to use a two-stage stochastic programming model [10], where the first stage typically consists of generator on/off decisions, while the second stage consists of power dispatch decisions (and perhaps also, on/off decision for quick-start generators) [10, 32]. These models are appropriate when commitment decisions must be fixed for the entire planning horizon.

Multi-stage models can accurately model a longer time horizon and dependencies between time periods; this modeling approach can be useful when generator commitment decisions may be adjusted frequently. However, with the increased complexity, large instances of the problem (e.g., having many generators or many time periods) are very challenging to solve. This limited scalability is due to the exponential increase in the demand scenarios with the number of stages. Note that we can view the two-stage model to be a restriction on the multi-stage model where the generator on/off decisions are restricted to be decided in advance.

We begin with a Markov Decision Process formulation [22] of the multi-stage stochastic unit commitment problem. Direct solution of this model is impractical for even modest-size instances, since the size of the state-space grows exponentially with the number of generators in the system. We therefore investigate an approximation approach that can yield a policy, along with a bound on how far it is from the optimal policy. In particular, we apply the Dual Approximate Dynamic Programming (DADP) approach proposed in [5],

which leads to an approach that decomposes the problem into a separate MDP problem for each generator in the system by relaxing the constraints that demands must be met in each time period. The key to this approach is to allow the Lagrangian multipliers to depend on a “summary” of the history of observed demands up to that time period, allowing a trade-off (by choosing the summary) between the complexity of solving the relaxed problem and the quality of the lower bound achieved. This approach is related to the relaxation approach in [14, 1], but in their work the Lagrange multipliers only depend on the time period, which can result in weak lower bounds. On the other hand, in [27] the Lagrange multipliers depend on the time period and the scenario of demand outcomes up to that time period. This approach can yield strong bounds, but is not practical for instances with many stages because the number of sample paths to a time period may grow exponentially with the number of stages. The DADP approach has been applied on a small energy problem with hydraulic plants and thermal units for illustrative purposes in [4]. However, it does not capture many of the complexities in the stochastic unit commitment problem, such as min up/down and ramping constraints. We present a numerical illustration on a large-scale 168-stage stochastic unit commitment problem. For bound comparisons, we generate a feasible policy and obtain upper bounds by using the value function from the DADP approach as an approximate future value function for a one-step lookahead policy. We show that this approach provides good lower and upper bounds for the stochastic unit commitment problem and provides good scalability with the number of generators.

The remainder of this paper is organized as follows. The problem formulation is given in section 2. The application of DADP to this problem is derived in section 3, and our numerical illustration is presented in section 4.

## 2 Formulation of the Stochastic Unit Commitment Problem

We assume there are  $n$  generators and  $T$  time periods indexed by  $t = 1, \dots, T$ . We consider a model that ensures that total power generation is sufficient to meet total demand in each period, but does not consider the transmission network. We define the random parameter  $D_t$  as the electric load or demand at time  $t$ , which is an element of the space  $\mathcal{D}_t = \{\delta_t(r), r = 1, 2, \dots, R\}$  where  $\delta_t(r)$  is the  $r$ th possible demand realization in time period  $t$ . Demand is modeled as a Markovian process, i.e., the distribution of the random demand  $D_{t+1}$  depends on  $D_t$ . We define the Markovian demand distribution  $P_t(w|d)$  as the probability that  $D_{t+1} = w$ , given  $D_t = d$ , for  $w \in \mathcal{D}_{t+1}$  and  $d \in \mathcal{D}_t$ , for  $t = 1, \dots, T$ . In our formulation, we assume that there

is a single load to be satisfied, so the demand takes on a scalar value.

At time  $t$ , the state of the system is given by the vector  $\mathbf{x}_t = (\mathbf{y}_t, \mathbf{q}_t, D_t)$ , where  $\mathbf{y}_t = (y_t^1, y_t^2, \dots, y_t^n)$  is a vector of generator statuses,  $\mathbf{q}_t = (q_t^1, q_t^2, \dots, q_t^n)$  is a vector of generator production levels in the previous period, and  $D_t$  is the current aggregate demand. The current demand  $D_t$  is assumed to be observed at the end of the previous stage, and so is included as part of the state vector, so that decisions in stage  $t$  may depend on the observed value of  $D_t$ . The vector  $\mathbf{q}_t$  of previous production levels is used to enforce ramp up and down constraints for each generator. The minimum and maximum generation levels from each generator  $i = 1, \dots, n$  are denoted  $b_{\min}^i$  and  $b_{\max}^i$ , respectively, and hence  $q_t^i \in [b_{\min}^i, b_{\max}^i]$ . The vector  $\mathbf{y}_t$  keeps track of how long each generator has been on or off and is needed to enforce minimum up and down time constraints. We also view the state as a vector of three tuples of the form  $\mathbf{x}_t = (x_t^1, x_t^2, \dots, x_t^n)$ , where  $x_t^i = (y_t^i, q_t^i, D_t)$ . Note that a generator can also represent external trading on the spot market, whether it is buying or selling electric load for a price. In this case, the cost of producing power in such a unit would then represent the cost of buying (positive cost) or the profit from selling (negative cost).

We denote the minimum up and down time for generator  $i$  to be  $\bar{l}_i$  and  $l_i$ , respectively. Let  $y_t^i = (\alpha_t^i, \beta_t^i)$ , where  $\alpha_t^i \in [0, \dots, \bar{l}_i]$  represents the number of periods the generator has been on, and  $\beta_t^i \in [0, \dots, l_i]$  represents the number of periods the generator has been off. Either  $\alpha_t^i$  or  $\beta_t^i$  must be zero at any point in time, but they cannot be zero simultaneously. If the generator has been on for more than  $\bar{l}_i$  time periods, then  $(\alpha_t^i, \beta_t^i) = (\bar{l}_i, 0)$ , meaning the generator can be turned off. Similarly, if the generator has been off for more than  $l_i$  time periods, then  $(\alpha_t^i, \beta_t^i) = (0, l_i)$ , meaning the generator can be turned on. If the generator is on and must remain on for some more time,  $\alpha_t^i$  will be a positive integer but strictly less than  $\bar{l}_i$  and  $\beta_t^i$  will be zero, and vice versa if the generator is off. For initialization purposes, we could set  $y_1^i = (0, l_i)$ , which would mean the generator has remained off for long enough that it can be turned on.

In summary, the state space in period  $t$  for each generator  $i = 1, \dots, n$  is defined as

$$\mathcal{X}_t^i = \{(y_t^i, q_t^i, d) : y_t^i = (\alpha_t^i, \beta_t^i), \alpha_t^i \in [0, \dots, \bar{l}_i], \beta_t^i \in [0, \dots, l_i], q_t^i \in [b_{\min}^i, b_{\max}^i], d \in \mathcal{D}_t\},$$

for  $t = 2, \dots, T$ , and  $\mathcal{X}_1^i = \{(y_1^i, q_1^i, d) : y_1^i = (0, l_i), q_1^i = 0, d = 0\}$ . The overall state  $\mathbf{x}_t$  is a member of the state space  $\mathcal{X}_t$ , which is the Cartesian product of the individual state spaces  $\mathcal{X}_t^i$ , i.e.,  $\mathbf{x}_t \in \mathcal{X}_t = \times_{i=1}^n \mathcal{X}_t^i$ .

The actions at time  $t$  are denoted by the vector  $\mathbf{a}_t = (\mathbf{z}_t, \mathbf{u}_t)$ , where  $\mathbf{z}_t = (z_t^1, z_t^2, \dots, z_t^n)$  is a vector of generator production levels, and  $\mathbf{u}_t = (u_t^1, u_t^2, \dots, u_t^n)$  is a vector of binary generator on/off decisions for

the next stage. Here,  $u_t^i = 0$  means generator  $i$  is off, and  $u_t^i = 1$  means the generator is on. We assume that the on/off decisions are made for the next period, whereas the generation decisions are made for the current period, after observing the current demand  $D_t$ . Thus, at time  $t$ , the generation decisions  $\mathbf{z}_t$  are for the current period  $t$ , and the commitment decisions  $\mathbf{u}_t$  are for period  $t + 1$ .

For each generator  $i$ , we enforce the following: minimum and maximum production level bounds, minimum up and down time constraints, and ramp up and down production level constraints. The minimum and maximum production levels for generator  $i$  are enforced with the following constraint:

$$\underline{u}(y_t^i)b_{\min}^i \leq z_t^i \leq \underline{u}(y_t^i)b_{\max}^i, \quad (1)$$

where  $\underline{u}(y_t^i)$  is the applied commitment decision  $u_{t-1}^i$  which equals 1 when  $y_t^i = (j, 0)$  for  $j = 1, \dots, \bar{l}_i$  and 0 otherwise. In the above constraint, if the previous commitment decision was  $\underline{u}(y_t^i) = 0$ , the production level is set to 0; otherwise, the production level remains between its minimum and maximum levels. We enforce minimum up and down constraints by requiring:

$$\underline{I}(y_t^i) \leq u_t^i \leq \bar{I}(y_t^i), \quad (2)$$

where  $\underline{I}(y_t^i)$  equals 1 if  $y_t^i = (j, 0)$  for  $j = 1, \dots, \bar{l}_i - 1$  and 0 otherwise, and  $\bar{I}(y_t^i)$  equals 0 if  $y_t^i = (0, j)$  for  $j = 1, \dots, \underline{l}_i - 1$  and 1 otherwise. For ramp up and down constraints, we enforce:

$$q_t^i - r_d^i - (1 - \underline{u}(y_t^i))b_{\min}^i \leq z_t^i \leq q_t^i + r_u^i + \bar{w}(y_t^i)b_{\min}^i, \quad (3)$$

where  $\underline{u}(y_t^i)$  is the applied commitment decision  $u_{t-1}^i$  as before,  $\bar{w}(y_t^i)$  is a ‘‘turn on’’ indicator that is 1 if  $y_t^i = (1, 0)$  and 0 otherwise, and  $r_u^i$  and  $r_d^i$  are ramp up and down amounts for generator  $i$ , respectively.

At each time period, we enforce a linking constraint that ensures the sum of the production levels from each generator satisfies the demands observed:

$$\sum_{i=1}^n z_t^i = D_t. \quad (4)$$

Since we assume  $D_1 = 0$  and all the generators are initially off in the first stage, the above constraint would mean  $z_1^i = 0$  for all  $i$ . This model assumes total generation should *exactly* meet the load. An alternate

constraint could ensure total generation to *at least* meet the load; we discuss minor changes in the solution approach if this were modeled in a later section. This model may be extended to allow  $z_t^i$ , and  $D_t$  to be vectors, e.g., if we have multiple electric loads, although we focus on the scalar case.

In summary, the control space for each generator  $i = 1, \dots, n$  is defined as

$$\mathcal{A}_t^i(x_t^i) = \{(z_t^i, u_t^i) : (1), (2), (3), u_t^i \in \{0, 1\}\}.$$

Then, the overall feasible action space in stage  $t$  is defined as

$$\bar{\mathcal{A}}_t(\mathbf{x}_t) = \{\mathbf{a}_t = (a_t^1, a_t^2, \dots, a_t^n) \in \times_{i=1}^n \mathcal{A}_t^i(x_t^i) : \sum_{i=1}^n z_t^i = D_t\}.$$

To initialize the model, in the first period,  $t = 1$ , we assume we only make commitment decisions. Thus, we assume  $D_1 = 0$  and hence  $z_1^i = 0$  for all  $i$ .

We now define the state update equations. For the state  $y_t^i$ , we have:

$$y_{t+1}^i = \begin{cases} (\alpha_t^i + 1, 0), & \text{if } 0 < \alpha_t^i < \bar{l}_i \\ (0, \beta_t^i + 1), & \text{if } 0 < \beta_t^i < \underline{l}_i \\ (\bar{l}_i, 0), & \text{if } \alpha_t^i = \bar{l}_i, u_t^i = 1 \\ (0, 1), & \text{if } \alpha_t^i = \bar{l}_i, u_t^i = 0 \\ (0, \underline{l}_i), & \text{if } \beta_t^i = \underline{l}_i, u_t^i = 0 \\ (1, 0), & \text{if } \beta_t^i = \underline{l}_i, u_t^i = 1. \end{cases}$$

The state update equations for  $q_t^i$  representing previous production levels is  $q_{t+1}^i = z_t^i$ . The overall update equation  $\mathbf{x}_{t+1} = \mathbf{f}_t(\mathbf{x}_t, \mathbf{a}_t) = (f_t^1(x_t^1, a_t^1), \dots, f_t^n(x_t^n, a_t^n))$  represents all of the above update equations taken together.

At time  $t$  for generator  $i$ , the cost  $g_t^i(x_t^i, a_t^i)$  is the total expected generation cost. We define  $\bar{c}_i$  to be the no load cost (fixed cost for generator being on),  $h_i$  to be a fixed cost for turning on generator  $i$  when it is off, and  $F_t^i(z)$  to be the generation cost of producing  $z$ . We model  $F_t^i(z)$  as a piecewise linear function of  $z$ . For generator  $i$ , we evaluate price-quantity bids over a grid and denote these points  $(b_k^i, c_k^i)$ , for  $k = 0, \dots, K_i$ , where  $b_k^i$  is the  $k$ th generation level and  $c_k^i$  is the cost associated with it. These are the breakpoints of the

piecewise linear generation cost function. Note that based on the previous notation, we have  $b_{\min}^i = b_0^i$  and  $b_{\max}^i = b_{K_i}^i$ , for  $i = 1, \dots, n$ . The cost for time period  $t$  is incurred after implementing the controls  $\mathbf{a}_t$ . The startup cost  $H_t^i(y_t^i, u_t^i)$  is defined as

$$H_t^i(y_t^i, u_t^i) = \begin{cases} h_i, & \text{if } y_t^i = (0, \underline{l}_i), u_t^i = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Now we define the total cost incurred per time period as the sum of the start up cost, no load cost, and the generation cost:

$$g_t^i(x_t^i, a_t^i) = H_t^i(y_t^i, u_t^i) + \bar{c}_i u_t^i + F_t^i(z_t^i),$$

for  $t = 1, \dots, T-1$ . For the last period  $T$ , the on/off decisions are irrelevant since commitment decisions are determined for the next stage; thus, only the cost associated with production level decisions are incurred:

$$g_T^i(x_T^i, a_T^i) = F_T^i(z_T^i).$$

The overall cost incurred at time  $t$  is the sum of the individual costs, i.e.,  $g_t(\mathbf{x}_t, \mathbf{a}_t) = \sum_{i=1}^n g_t^i(x_t^i, a_t^i)$ .

Thus, we formulate the stochastic unit commitment problem as

$$\min_{\pi} \mathbb{E} \left[ \sum_{t=1}^T g_t(\mathbf{x}_t, \mathbf{a}_t) \right],$$

where  $\pi = \{(\zeta_0, \boldsymbol{\mu}_0), \dots, (\zeta_{T-1}, \boldsymbol{\mu}_{T-1})\}$  represents an admissible policy, where  $(\zeta_t, \boldsymbol{\mu}_t)$  maps the state  $\mathbf{x}_t$  into actions  $(\mathbf{z}_t, \mathbf{u}_t) = (\zeta_t(\mathbf{x}_t), \boldsymbol{\mu}_t(\mathbf{x}_t))$  such that  $(\zeta_t(\mathbf{x}_t), \boldsymbol{\mu}_t(\mathbf{x}_t)) \in \bar{\mathcal{A}}_t(\mathbf{x}_t)$  for all  $\mathbf{x}_t \in \mathcal{X}_t$ . Note that we have not included a terminal cost associated with being in a potential undesirable state after applying the sequence of decisions; this would be a straightforward addition to the cost, e.g.,  $\mathbb{E}[g_{T+1}(\mathbf{x}_{T+1})]$ . If we define  $J_t(\mathbf{x}_t)$  to be the minimum expected cost-to-go when the system is in state  $\mathbf{x}_t \in \mathcal{X}_t$ , then  $J_t(\mathbf{x}_t)$  satisfies the dynamic programming (DP) recursion

$$J_t(\mathbf{x}_t) = \min_{\mathbf{a}_t \in \bar{\mathcal{A}}_t(\mathbf{x}_t)} \left\{ \mathbb{E} [g_t(\mathbf{x}_t, \mathbf{a}_t) + J_{t+1}(\mathbf{f}_t(\mathbf{x}_t, \mathbf{a}_t)) | \mathbf{x}_t, \mathbf{a}_t] \right\}, \quad (5)$$

for  $t = 1, \dots, T$ , where  $J_{T+1}(\mathbf{x}_{T+1}) = 0$ , and the expectation is taken with respect to the probability distribution  $P_t(D_{t+1} | D_t)$ .

The notation described in this section is summarized in Table 1.

Table 1: Notation

Constant	Description
$y_t^i$	state variable indicating status of $i$ th generator at time $t$
$q_t^i$	state variable indicating generator $i$ 's previous production level at time $t$
$D_t$	observed demand at time $t$
$z_t^i$	production level decision of $i$ th generator at time $t$
$u_t^i$	binary on/off decision of $i$ th generator at time $t$
$r_u^i$	maximum ramp up amount for generator $i$
$r_d^i$	maximum ramp down amount for generator $i$
$\bar{l}_i$	minimum up time for generator $i$ (minimum time generator must stay on after being turned on)
$\underline{l}_i$	minimum down time for generator $i$ (minimum time generator must stay off after being turned off)
$\bar{c}_i$	no load cost for generator $i$ (fixed cost for generator being on)
$h_i$	turn on cost for generator $i$ (additional cost for turning on generator when it is off)

### 3 Dual Approximate Dynamic Programming Approach

We now describe how we adapt the dual approximate dynamic programming approach [5, 12, 17] to obtain a policy and optimality bound for the stochastic unit commitment problem. While the DADP approach has been applied previously to a hydraulic valley example and simple small-scale power management problem, the problem did not have any integer variables and did not capture the complexities including min / up down and ramping constraints. We show for the first time its effectiveness on a large-scale stochastic unit commitment problem.

In time period  $t$ , an exact approach using the original DP recursion (5) would result in a total number of states of  $|\mathcal{X}_t| = \prod_{i=1}^n |\mathcal{X}_t^i|$ . Even with a relatively small number of states for each subproblem, this solution approach would quickly become computational intractable because of the number of states growing exponentially with the number of generators. In the DADP approach, a Lagrangian relaxation approach is used and the resulting subproblems are solved independently. This approach instead solves a problem with  $|\mathcal{X}_t^i|$  for each generator  $i = 1, \dots, n$ . Having solved the relaxed problem, the approach provides a lower bound on the original optimal objective. We can obtain a primal policy (and hence an upper bound) by using a one step lookahead policy by using an approximate value function derived from the Lagrangian relaxation solution.



The main idea of the DADP approach is to introduce an additional state:

$$v_t = \tilde{f}_t(v_{t-1}, D_t), \quad (6)$$

which summarizes the exogenous information process  $D_1, D_2, \dots, D_t$ , for  $t = 2, \dots, T$ , with an initial state  $v_1 = D_1 = 0$ . For each  $t = 2, \dots, T$ , we assume  $v_t$  lies within a finite set of values, denoted by the set  $\mathcal{V}_t$ . For the DADP approach to be computationally practical, the size of  $\mathcal{V}_t$  must not be too large. We let the Lagrange multipliers depend on  $v_t$ , and hence define  $\lambda_t : \mathcal{V}_t \rightarrow \mathbb{R}$ , for  $t = 1, \dots, T$ . Here, if there are multiple linking constraints,  $\lambda_t$  would be vector-valued, with each element representing multipliers for each linking constraint. We assume that knowing  $v_t$  is sufficient to know the distribution of  $D_r$  for any  $r > t$ . This is trivially satisfied if the random demands are stage-wise independent. More generally, if knowing  $v_t$  implies we know  $D_t$  (e.g.,  $v_t$  may be a vector containing  $D_t$  as one component), then this is implied by the Markovian assumption.

We let  $\boldsymbol{\lambda} = [\lambda_t]_{t=1}^T$ , be the collection of all Lagrangian multipliers. For a fixed  $\boldsymbol{\lambda}$ , the Lagrangian problem is:

$$\min_{\pi} \mathbb{E} \left\{ \sum_{t=1}^T \left[ g_t(\mathbf{x}_t, \mathbf{a}_t) + \lambda_t(v_t)^\top \left( \sum_{i=1}^n z_t^i - D_t \right) \right] \right\}, \quad (7)$$

where  $\pi$  represents the class of admissible policies over the feasible control space  $\mathcal{A}_t(\mathbf{x}_t)$ . Now, the Lagrangian recursion is:

$$L_{T+1}(\mathbf{x}_{T+1}, v_{T+1}; \boldsymbol{\lambda}) = 0,$$

$$L_t(\mathbf{x}_t, v_t; \boldsymbol{\lambda}) = \min_{\mathbf{a}_t \in \mathcal{A}_t(\mathbf{x}_t)} \mathbb{E} \left[ g_t(\mathbf{x}_t, \mathbf{a}_t) + \lambda_t(v_t)^\top \left( \sum_{i=1}^n z_t^i - D_t \right) + L_{t+1}(\mathbf{x}_{t+1}, v_{t+1}; \boldsymbol{\lambda}) \mid \mathbf{x}_t, v_t, \mathbf{a}_t \right] \quad (8)$$

for  $t = 1, \dots, T$ .

**Remark 1.** *The representation (6) is very general; note that by letting  $v_t = [v_{t-1}, D_t]$ , we have a multiplier for every sequence in the exogenous demand process.*

**Remark 2.** *Since (4) is an equality constraint, we let  $\lambda_t$  to be a free variable. However, we could have allowed inequality linking constraints as well, in which case we would have ensured  $\lambda_t \geq 0$ .*

### 3.1 Decomposition and Structural Properties

We present the key results that are needed for applying the DADP approach to this problem. See [12, 17] for further results.

The following result shows that under this representation, the Lagrangian problem decomposes into  $n$  individual subproblems. In the theorem, the notation  $\mathbb{E}[\lambda_r(v_r)D_r \mid v_t]$  for  $1 \leq t \leq r \leq T$  represents the expected value of  $\lambda_r(v_r)D_r$  given that the state of the demand process in stage  $t$  is  $v_t$ , where the expectation is taken with respect to the random outcomes  $D_{t+1}, \dots, D_r$ . Note that for  $r = t$ , this term is simply  $\lambda_t(v_t)D_t$ .

**Theorem 3.** *The Lagrangian recursion decouples as follows:*

$$L_t(\mathbf{x}_t, v_t; \boldsymbol{\lambda}) = \sum_{i=1}^n L_t^i(x_t^i, v_t; \boldsymbol{\lambda}) - \sum_{r=t}^T \mathbb{E}[\lambda_r(v_r)D_r \mid v_t],$$

for  $t = 1, \dots, T$ , where

$$L_{T+1}^i(x_{T+1}^i, v_{T+1}; \boldsymbol{\lambda}) = 0,$$

and for  $t = T, \dots, 1$

$$L_t^i(x_t^i, v_t; \boldsymbol{\lambda}) = \min_{a_t^i \in \mathcal{A}_t^i(x_t^i)} \mathbb{E}[g_t^i(x_t^i, a_t^i) + \lambda_t(v_t)z_t^i + L_{t+1}^i(x_{t+1}^i, v_{t+1}; \boldsymbol{\lambda}) \mid x_t^i, v_t, a_t^i]. \quad (9)$$

*Proof.* We proceed by induction. For the base case, we have by definition

$$L_{T+1}(\mathbf{x}_{T+1}, v_{T+1}; \boldsymbol{\lambda}) = 0 = \sum_{i=1}^n L_{T+1}^i(x_{T+1}^i, v_{T+1}; \boldsymbol{\lambda}).$$

Now, assume the statement in the theorem holds for time  $t + 1$ . Then, we have

$$\begin{aligned}
L_t(\mathbf{x}_t, v_t; \boldsymbol{\lambda}) &= \min_{\mathbf{a}_t \in \mathcal{A}_t(\mathbf{x}_t)} \mathbb{E} \left[ g_t(\mathbf{x}_t, \mathbf{a}_t) + \lambda_t(v_t) \left( \sum_{i=1}^n z_t^i - D_t \right) + L_{t+1}(\mathbf{x}_{t+1}, v_{t+1}; \boldsymbol{\lambda}) \mid \mathbf{x}_t, v_t, \mathbf{a}_t \right] \\
&= \min_{\mathbf{a}_t \in \mathcal{A}_t(\mathbf{x}_t)} \mathbb{E} \left[ \sum_{i=1}^n g_t^i(x_t^i, a_t^i) + \lambda_t(v_t) \left( \sum_{i=1}^n z_t^i - D_t \right) + \sum_{i=1}^n L_{t+1}^i(x_{t+1}^i, v_{t+1}; \boldsymbol{\lambda}) \right. \\
&\quad \left. - \sum_{r=t+1}^T \mathbb{E}[\lambda_r(v_r) D_r \mid v_{t+1}] \mid \mathbf{x}_t, v_t, \mathbf{a}_t \right] \\
&= \min_{\mathbf{a}_t \in \mathcal{A}_t(\mathbf{x}_t)} \sum_{i=1}^n \mathbb{E} \left[ g_t^i(x_t^i, a_t^i) + \lambda_t(v_t) z_t^i + L_{t+1}^i(x_{t+1}^i, v_{t+1}; \boldsymbol{\lambda}) \mid x_t^i, v_t, a_t^i \right] \\
&\quad - \sum_{r=t}^T \mathbb{E}[\lambda_r(v_r) D_r \mid v_t] \\
&= \sum_{i=1}^n L_t^i(x_t^i, v_t; \boldsymbol{\lambda}) - \sum_{r=t}^T \mathbb{E}[\lambda_r(v_r) D_r \mid v_t],
\end{aligned}$$

where  $L_t^i(x_t^i, v_t; \boldsymbol{\lambda})$  satisfies (9), as desired.  $\square$

In particular, Theorem 3 implies that

$$L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda}) = \sum_{i=1}^n L_1^i(x_1^i, v_1; \boldsymbol{\lambda}) - \sum_{t=1}^T \mathbb{E}[\lambda_t(v_t) D_t \mid v_1]. \quad (10)$$

The importance of Theorem 3 is that, for fixed  $\boldsymbol{\lambda}$ ,  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$  can be evaluated by solving  $n$  independent Markov decision problems, each with a relatively small state space. The term  $\sum_{t=1}^T \mathbb{E}[\lambda_t(v_t) D_t \mid v_1]$  is independent of the decision process, and so can be estimated via simulation.

We next present two important structural properties of  $\mathcal{L}$ .

**Theorem 4.** *We have that*

1.  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda}) \leq J_1(\mathbf{x}_1)$  for all  $\boldsymbol{\lambda}$ .
2.  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$  is concave function of  $\boldsymbol{\lambda}$ .

*Proof.* These results follow from standard Lagrangian theory [8, 9]. For any feasible policy, we have

$\sum_{i=1}^n z_t^i = D_t$  for all  $D_t, t = 1, \dots, T$ . Thus, by definition

$$\begin{aligned} L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda}) &= \min_{\pi} \mathbb{E} \left[ \sum_{t=1}^T g_t(\mathbf{x}_t, \mathbf{a}_t) + \lambda_t(v_t) \left( \sum_{i=1}^n z_t^i - D_t \right) \right], \\ &\leq \mathbb{E} \left[ \sum_{t=1}^T g_t(\mathbf{x}_t, (\boldsymbol{\zeta}_t(\mathbf{x}_t, v_t), \boldsymbol{\mu}_t(\mathbf{x}_t, v_t))) \right], \end{aligned}$$

for a feasible policy  $(\boldsymbol{\zeta}_t, \boldsymbol{\mu}_t)$  where  $\boldsymbol{\zeta}_t$  is the policy associated with production levels and  $\boldsymbol{\mu}_t$  is associated with the commitment decisions. Since the above is true for any feasible policy, it also holds for an optimal policy  $(\boldsymbol{\zeta}_t^*(\mathbf{x}_t, v_t), \boldsymbol{\mu}_t^*(\mathbf{x}_t, v_t))$ . We now have

$$L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda}) \leq \mathbb{E} \left[ \sum_{t=1}^T g_t(\mathbf{x}_t, (\boldsymbol{\zeta}_t^*(\mathbf{x}_t, v_t), \boldsymbol{\mu}_t^*(\mathbf{x}_t, v_t))) \right] = J_1(\mathbf{x}_1).$$

For the second claim, we proceed by induction, and use the recursive definition of  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$  given in (8). For the base case, we see that  $L_{T+1}(\mathbf{x}_{T+1}, v_{T+1}; \boldsymbol{\lambda})$  is clearly concave in  $\boldsymbol{\lambda}$ . Now, suppose  $L_{t+1}(\mathbf{x}_{t+1}, v_{t+1}; \boldsymbol{\lambda})$  is concave in  $\boldsymbol{\lambda}$ . Then, the expected value term in (8) is a concave function of  $\boldsymbol{\lambda}$ .  $L_t(\mathbf{x}_t, v_t; \boldsymbol{\lambda})$  is concave because it is a minimum of concave functions of  $\boldsymbol{\lambda}$ .  $\square$

We find the best lower bound and thus, maximize  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$  over  $\boldsymbol{\lambda}$ . We define

$$\mathcal{L}_1 = \max_{\boldsymbol{\lambda}} L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda}).$$

From Theorem 4, it follows that  $\mathcal{L}_1 \leq J_1(\mathbf{x}_1)$ .

Because  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$  is concave in  $\boldsymbol{\lambda}$ , but not necessarily smooth, it can be maximized using supergradient-based methods. In general, it is difficult to determine the supergradient exactly. However, we can obtain an unbiased stochastic estimator of a supergradient using sampling. The following theorem shows how to compute an unbiased estimator of a supergradient of  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$ .

**Theorem 5.** *Suppose  $\pi = \{(\boldsymbol{\zeta}_1, \boldsymbol{\mu}_1), \dots, (\boldsymbol{\zeta}_T, \boldsymbol{\mu}_T)\}$  is the optimal policy for the the Lagrangian relaxation problem  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$ , where  $\boldsymbol{\zeta}_t$  is associated with generator production levels and  $\boldsymbol{\mu}_t$  is associated with commitment decisions. Here, each subproblem  $i$  has its policy  $\pi^i = \{(\zeta_1^i, \mu_1^i), \dots, (\zeta_T^i, \mu_T^i)\}$ . An unbiased estimator of a supergradient of  $\mathcal{L}$  at  $\boldsymbol{\lambda}$  is*

$$\left[ \sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t \right]_{t=1}^T.$$

*Proof.* For any  $\hat{\lambda}$ , we have

$$\begin{aligned}
L_1(\mathbf{x}_1, v_1; \hat{\lambda}) &\leq \mathbb{E} \sum_{t=1}^T \left[ g_t(\mathbf{x}_t, (\zeta_t(\mathbf{x}_t, v_t), \boldsymbol{\mu}_t(\mathbf{x}_t, v_t))) + \hat{\lambda}_t^\top \left( \sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t \right) \right] \\
&= \mathbb{E} \sum_{t=1}^T \left[ g_t(\mathbf{x}_t, (\zeta_t(\mathbf{x}_t, v_t), \boldsymbol{\mu}_t(\mathbf{x}_t, v_t))) + \lambda_t^\top \left( \sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t \right) \right. \\
&\quad \left. + (\hat{\lambda}_t - \lambda_t)^\top \mathbb{E} \left( \sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t \right) \right] \\
&= L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda}) + \sum_{t=1}^T (\hat{\lambda}_t - \lambda_t)^\top \mathbb{E} \left( \sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t \right),
\end{aligned}$$

where the first inequality follows because  $\pi$  is a feasible, but not necessarily optimal, policy for  $\hat{\lambda}$ . It follows that  $\left[ \mathbb{E} \left( \sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t \right) \right]_{t=1}^T$  is a supergradient at  $\boldsymbol{\lambda}$ .  $\square$

The importance of Theorem 3 is that a *stochastic* supergradient method [18, 7, 8] can then be applied to maximize  $L_1(\mathbf{x}_1, v_1; \boldsymbol{\lambda})$ . For a fixed  $\boldsymbol{\lambda}$ , we can solve the Lagrangian recursion via a decoupled approach given in Theorem 3. Then, for any simulated sample path of the random variables  $D_t$ , for  $t = 1, \dots, T$ , the vector of demand violations  $\left[ \sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t \right]_{t=1}^T$ , is an unbiased estimate of a supergradient of  $L_1$  at  $\boldsymbol{\lambda}$ . To get a better (reduced variance) estimate of the supergradient, we can use batch gradient averages, i.e., simulate many sample paths and average the demand violations to obtain a supergradient estimate.

### 3.2 Implementation of the DADP Approach

We solve the decoupled MDP subproblems using the recursion equation given in (9) using standard dynamic programming. The only detail we need to deal with is that, as stated, the state space is not finite, so we cannot directly enumerate all states when calculating the value function. To address this, for generator  $i$ , and for each possible value of  $y_t^i$  and  $D_t \in \mathcal{D}_t$ , we discretize  $q_t^i$  at the same points as for the cost function  $g_t^i(x_t^i, a_t^i)$ , i.e.,  $b_k^i$ , for  $k = 1, \dots, K_i$ , for every  $t$ . We denote the discretized version of the set  $\mathcal{X}_t$  by  $\tilde{\mathcal{X}}_t$ .

Because both the piecewise-linear cost functions have the same set of break points in each time period, these break points are inherited by the value function, and hence the minimization in (9) always has a solution at one of those break points or at the bounds. To avoid having a solution at the bounds, we relax the bound constraints to allow one discretization point above and below the ramping bounds. Thus, although we

restrict the production levels (and state variables) to lie at one of the break points, the value we obtain from the algorithm will still be a lower bound on the optimal value. Suppose  $\tilde{\mathcal{A}}_t^i(x_t^i)$  is the discretized version of the set  $\mathcal{A}_t^i(x_t^i)$ , where  $z_t^i$  only takes on the values  $b_k^i$ , for  $k = 1, \dots, K_i$ , or  $b_0^i = 0$  if the generator is off, and the set is modified so that the ramping constraint (3) is relaxed to include one discretization point less than or equal to the lower bound and one point greater than or equal to the upper bound. Thus, we solve the discretized version of (9) by simply taking a pointwise minimum over the objective evaluated at their break points. The overall DADP algorithm is given in pseudocode in the Algorithm 1 box below. The “parfor” loops indicate loops that can be parallelized.

---

**Algorithm 1** DADP Algorithm

---

**Parameters:**  $\rho, \eta, \text{maxIters}, \text{batchSize}$

Initialize  $\lambda_t^1(v_1)$ , for  $t = 1, \dots, T$

**for** ( $k = 1; k \leftarrow k + 1; k \leq \text{maxIters}$ ) **do**

**parfor** ( $i = 1; i \leftarrow i + 1; i \leq n$ ) **do**

**for** ( $t = T; t \leftarrow t - 1; t > 0$ ) **do**

**for**  $((x, v) = ((l, b_k^i, d), v) \in (\tilde{\mathcal{X}}_t, \mathcal{V}_t))$  **do**

$$L_t^i(x, v; \lambda) = \min_{a_t^i \in \tilde{\mathcal{A}}_t^i(x)} [g_t^i(x, a_t^i) + \lambda_t^k(v)z_t^i + \sum_{r \in \mathcal{D}_{t+1}} P_t(r|d)L_{t+1}^i(f_t^i(x, a_t^i), \tilde{f}_t(v, r); \lambda)]$$

            and let  $(\zeta_t^i(x, v), \mu_t^i(x, v))$  be an optimal action.

**end for**

**end for**

**end parfor**

$g_t(v_t) = 0$

**parfor** ( $j = 1; j \leftarrow j + 1; j \leq \text{batchSize}$ ) **do**

**for** ( $t = 1; t \leftarrow t + 1; t \leq T$ ) **do**

            Simulate  $D_t$

            Update states  $\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}, (\zeta_{t-1}^i(\mathbf{x}_{t-1}, v_{t-1}), \mu_{t-1}^i(\mathbf{x}_{t-1}, v_{t-1})))$ ,  $v_t = \tilde{f}_t(v_{t-1}, D_t)$

$g_t(v_t) = g_t(v_t) + (\sum_{i=1}^n \zeta_t^i(x_t^i, v_t) - D_t) / \text{batchSize}$

**end for**

**end parfor**

$\lambda_t^k(v_t) \leftarrow \lambda_t^{k-1}(v_t) + \rho\eta^k g_t(v_t)$ , where  $\rho$  and  $0 < \eta < 1$  are step size parameters

**end for**

$\mathcal{L}_1 = \sum_{i=1}^n L_1^i(x_1^i, v_1; \lambda) - \sum_{t=1}^T \mathbb{E}[\lambda_t(v_t)D_t | v_1]$ , where initial states  $x_1^i = ((0, l_i), 0, 0)$  and  $v_1 = 0$ .

**return**  $\mathcal{L}_1, (\zeta_t(\mathbf{x}_t, v_t), \mu_t(\mathbf{x}_t, v_t))$ , for  $t = 1, \dots, T$

---

### 3.3 One-step Lookahead Policy and an Upper Bound

Solving the Lagrangian relaxation (7) gives a lower bound to the optimal value. However, the policy obtained from solving this problem via a DADP approach is not guaranteed to be feasible. One way to obtain a feasible policy is to approximate the future value function with the relaxed value functions, i.e., with

$L_t(\mathbf{x}_t, v_t; \boldsymbol{\lambda})$ , for  $t = 1, \dots, T$ . We use a one-step lookahead policy [6]. We first use the DADP approach to optimize for  $\boldsymbol{\lambda}$  and solve  $n$  independent MDP and their associated relaxed value functions  $L_t(\mathbf{x}_t, v_t; \boldsymbol{\lambda})$ , for  $t = 1, \dots, T$ . Then, we simulate a sample path  $D_1, \dots, D_T$ . For time  $t = 1, \dots, T$ , we solve the following:

$$\hat{J}_t(\mathbf{x}_t) = \min_{\mathbf{a}_t \in \overline{\mathcal{A}}_t(\mathbf{x}_t)} \left\{ \mathbb{E} [g_t(\mathbf{x}_t, \mathbf{a}_t) + L_{t+1}(\mathbf{x}_{t+1}, v_{t+1}; \boldsymbol{\lambda}) | \mathbf{x}_t, \mathbf{a}_t] \right\}, \quad (11)$$

for the state  $\mathbf{x}_t$  which determines the generation decisions  $z_t^i$  for the current period and the commitment decisions  $u_t^i$  that will determine the available generators in the next period. Problem (11) can be formulated as a deterministic mixed-integer program with decision variables  $(z_t^i, u_t^i)$  for  $i = 1, \dots, n$  and with constraints (1)-(4). Note that in this policy we enforce the correct lower and upper bounds implied by ramping, and hence the policy is feasible. The piecewise linear cost function and relaxed value functions  $L_{t+1}(\mathbf{x}_{t+1}, v_{t+1}; \boldsymbol{\lambda})$  are modeled with turn on variables and non-negative variables, similar to what is done for the perfect information MIP in section 4.3. Since this lookahead policy is a feasible policy, we can obtain a stochastic upper bound by generating a large number of samples and averaging the resulting overall costs from applying the above policy.

The Lagrangian dual problem may be solved offline once to generate lower bounds, and the obtained value function approximation may be used for the one-step lookahead policy through the whole time horizon. However, in cases where the demand model does not exactly follow the assumed distribution, it may be advantageous to re-optimize the Lagrangian dual every so often with updated demand information.

### 3.4 Discussion

The DADP approach is an extension of the approach by Adelman and Mersereau [1], which describes using a state independent multiplier  $\lambda_t$  for Lagrangian decomposition. Their paper shows that using a state dependent multiplier, e.g.,  $\lambda_t(\mathbf{x}_t)$ , does not result in decomposition. However, in the DADP approach, dependence on only the exogenous demand history or some function of it, e.g.,  $v_t$ , results in the desirable decomposition property.

Our model formulation for the stochastic unit commitment problem closely follows the one in Takriti et al [28]. Their approach also uses Lagrangian decomposition, but the multipliers depend on the scenario. This dependence results in an exponential increase in the number of scenarios with the decision stages, and the solution approach in [28] quickly becomes intractable, or requires an overly coarse representation of the

stochastic process. For example, if we assume there are 10 demand scenarios for every hour of the week (168 periods), we would have a total of  $10^{168}$  multipliers, each associated with a scenario. Thus, this limits the approach to a relatively small number of decision stages.

The scenario-based approach is also special case of the DADP approach when the state  $v_{t+1}$  consists of the full demand history:

$$v_t = [D_1, \dots, D_{t-1}, D_t].$$

As mentioned earlier, this representation results in a huge state space and is therefore limited to a small number of stages. The key in the DADP approach is the selection of a good “summary function”  $f_t(v_{t-1}, D_t)$  for summarizing the demand process up to stage  $t$ . This requires finding a tradeoff between letting  $v_t$  represent the full demand history and ignoring the history completely (which is the state independent multiplier case in Adelman and Mersereau [1]).

## 4 Numerical Illustration

All implementations and problem instances can be obtained at <https://github.com/jramak/dual-adp-suc>.

### 4.1 Problem Data

We use generator data from the FERC eLibrary Docket Number AD10-12, ACCNNUM 20120222-4012. This included min up/down times, ramp up/down amounts, no load costs, turn on costs, and up to 10 pairs of price-quantity bids. Of the 1011 generators, we randomly selected generators for our 15, 30, and 50 generator test cases. We obtained 2013 hourly demand data from the PJM Interconnection ISO, which is the regional transmission organization for the eastern electricity market. In order to model realistic demand fluctuations, we averaged out demand for each of the 168 hours of the week and normalized it by the maximum demand (see Figure 1). Note that we have  $T = 169$  because in the first period  $D_1 = 0$  and  $z_1^i = 0$  for all  $i$ , and only the on/off decisions  $\mathbf{u}_1$  for the next (second) stage are determined. To create a reasonable problem for each test case, we scaled this normalized demand by a percentage of the maximum combined generation level of all generators,  $TotCap$ , in the test set. We denote  $\mu$  to be the percentage of the maximum generation level, e.g., the scaling factor was  $\mu * TotCap$ . We assume demand is independent between time periods and sample 10 possible demand scenarios for each period. To generate demand



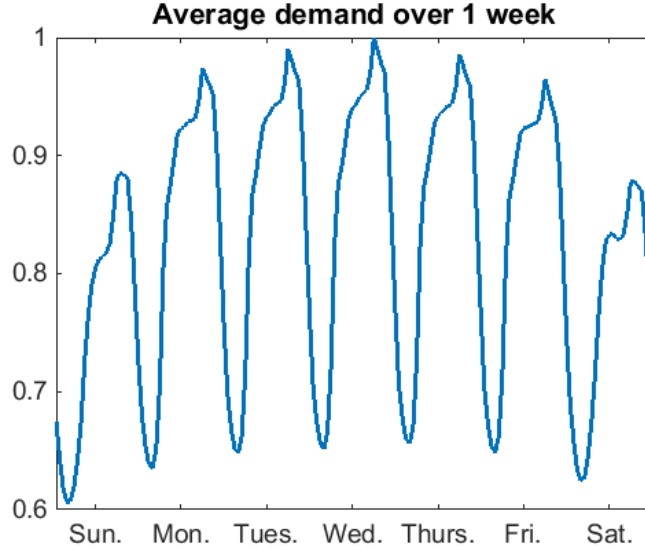


Figure 1: Average hourly demand from the 2013 PJM Interconnection ISO normalized by the maximum average demand.

scenarios for each time  $t$ , we evaluated 10 points using the Legendre-Gauss Quadrature in the interval  $[d_{scaled} - 4\sigma d_{scaled}, d_{scaled} + 4\sigma d_{scaled}]$ , where  $d_{scaled}$  is the scaled mean demand and  $\sigma$  represents the percentage of the variation in the mean demand. With the choices  $\mu = 0.4, 0.6, 0.8$  and  $\sigma = 0.15, 0.20, 0.25$ , we had a total of 9 instances for each test case. For modeling the generation cost, we evaluated the price-quantity bids over a uniform grid between the minimum and maximum generation levels (i.e.,  $b_{\min}^i$  and  $b_{\max}^i$ ) using 50 points.

## 4.2 DADP-based Bounds

For the DADP approach, we use  $v_t = D_t$ , for  $t = 1, \dots, T$ , which with 10 possible demand values per period increases the state space by 10 compared to the state independent case. Due to having  $v_t = D_t$  and the stagewise independence assumption for our numerical example, we calculate for a fixed  $\lambda$  the expected value term exactly in the Lagrangian function (10):

$$\sum_{t=1}^T \mathbb{E}[\lambda_t(v_t)D_t | v_1] = \sum_{t=1}^T \mathbb{E}[\lambda_t(D_t)D_t] = \sum_{t=1}^T \sum_{\delta \in \mathcal{D}_t} [P_t(\delta) \lambda_t(\delta) \delta]$$

Because of the exact calculation of the expected value term, the lower bounds produced by the DADP approach are deterministic.

For the stochastic supergradient method, we used a batch of 1000 sample paths to estimate the supergradient  $g_t$  at each step. For the supergradient estimation, we used a serial implementation, but this could be easily parallelized. The  $k$ th iteration is defined as

$$\lambda_t^k(D_t) = \lambda_t^{k-1}(D_t) + \rho\eta^k g_t,$$

for  $t = 1, \dots, T$ , where  $\rho$  and  $0 < \eta < 1$  are step size parameters. For large enough  $\rho$  and  $\eta$ , the supergradient method is guaranteed to converge although larger values can result in slower convergence. We used  $\rho = 50/(\mu * TotCap)$ ,  $\eta = 0.99$ , and set 250 as the maximum number of iterations. The parameter choices we chose resulted in apparent convergence and provided good enough solution quality and times, and we did not optimize the choice of the parameters further. Similar to the approach in [28], we initialize the multipliers  $\lambda_t(D_t)$  by approximating the cost in period  $t$  for generator  $i$  by a linear function and solving the approximate problem. The slope of the linearized cost function at time  $t$  for generator  $i$  is

$$(h_i + \bar{c}_i + F_t^i(b_{\max}^i))/b_{\max}^i.$$

The above slopes are sorted in increasing order, and the demand  $D_t$  is fulfilled by generators in this order. We initialize  $\lambda_t(D_t)$  to be the slope of the last unit used.

We implemented the state independent Lagrangian approach from [1] and the DADP approach in MATLAB 64-bit R2014b. For each instance, we used the HTCCondor framework to schedule a job on a machine with at least 4 CPUs, 4 GB RAM, and 12 GB disk space. Both the Lagrangian approach and the DADP approach ran on the same machine one after the other for fair time comparisons. Within each instance, for the generator subproblems, we used MATLAB's parpool with 4 workers. We modeled the one-step lookahead DADP MIP described in Subsection 3.3 in GAMS and used the solver CPLEX 12.6. For each of the 500 sample paths, we solved a sequence of MIPs, one for each time period, by using system calls to GAMS from MATLAB.

### 4.3 Perfect Information Bound

For comparison to the DADP approach, we also implemented the perfect information approach that provides a lower bound. In this approach, we assume perfect knowledge of a sample path (i.e., demand realizations). For a given sample path, the stochastic unit commitment problem becomes a MIP. We simulated 100 sample paths and averaged the resulting costs to obtain a stochastic lower bound, and a 95% confidence interval around this stochastic bound. Note that this approach only provides a stochastic lower bound and does not generate an implementable policy (and therefore, an associated upper bound).

We provide the perfect information MIP formulation below. Here again  $z_t^i$  refers to the generation level in time period  $t$ ,  $u_t^i$  refers to the commitment decision for time period  $t + 1$ , and  $D_t$  refers to the demand observed immediately before determining the generation level  $z_t^i$ . To model start up and shut down costs, we introduce additional turn on variables  $w_t^i$ , for  $i = 1, \dots, n$  and  $t = 1, \dots, T$ , that have shown to result in stronger relaxations (see [23]). We model piecewise linear functions with the locally ideal MIP formulation suggested in [25].

$$\begin{aligned}
& \underset{u, z, \gamma, e, w}{\text{minimize}} && \sum_{i=1}^n \sum_{t=1}^T [e_t^i + \bar{c}_i u_t^i + h_i w_t^i] \\
& \text{subject to} && \sum_{i=1}^n z_t^i = D_t, \quad t = 1, \dots, T, && \text{(demand satisfaction)} \\
& && w_0^i = u_0^i, \quad i = 1, \dots, n, \\
& && w_t^i \geq u_t^i - u_{t-1}^i, \quad i = 1, \dots, n, \quad t = 1, \dots, T, && \text{(turn on variables)} \\
& && \sum_{r=(t'-l_i+1)^+}^{t'} w_r^i \leq u_t^i, \quad i = 1, \dots, n, \quad t' = 0, \dots, T-1, && \text{(turn on inequalities)} \\
& && \sum_{r=(t'-l_i+1)^+}^{t'} w_r^i \leq 1 - u_{t-l_i}^i, \quad i = 1, \dots, n, \quad t' = l_i, \dots, T-1, && \text{(turn off inequalities)} \\
& && z_{t-1}^i - r_d^i - (1 - u_{t-1}^i) b_{\min}^i \leq z_t^i, \quad i = 1, \dots, n, \quad t = 1, \dots, T, && \text{(ramp down constraints)} \\
& && z_t^i \leq z_{t-1}^i + r_u^i + w_{t-1}^i b_{\min}^i, \quad i = 1, \dots, n, \quad t = 1, \dots, T, && \text{(ramp up constraints)} \\
& && \sum_{k=0}^{K_i} \gamma_{t,k}^i = u_{t-1}^i, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \\
& && z_t^i = \sum_{k=0}^{K_i} b_k^i \gamma_{t,k}^i, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \\
& && e_t^i = \sum_{k=0}^{K_i} c_k^i \gamma_{t,k}^i, \quad i = 1, \dots, n, \quad t = 1, \dots, T, && \text{(PWL cost)} \\
& && u_t^i \in \{0, 1\}, \quad i = 1, \dots, n, \quad t = 0, \dots, T, \\
& && w_t^i \in \{0, 1\}, \quad i = 1, \dots, n, \quad t = 0, \dots, T, \\
& && \gamma_{t,k}^i \geq 0, \quad k = 0, \dots, K_i, \quad i = 1, \dots, n, \quad t = 1, \dots, T,
\end{aligned}$$

where we define  $(x)^+ = \max(0, x)$ .

For fair comparison, the perfect information MIP ran after DADP lower bound on the same machine scheduled by HTCondor. The MIP was modeled in GAMS, solved using CPLEX 12.6, and was called through system calls in MATLAB. The MIPs for each sample path ran in parallel through MATLAB's 4 parfor workers. We limited the MIP solver to use a single CPU.

## 4.4 Results

We report the bounds obtained from each of the approaches for the 27 test instances in Table 2. For the lower bounds,  $\lambda_t$  refers to the state independent Lagrangian bound [1] and  $\lambda_t(v_t)$  refers to the DADP approach. Since the perfect bound is stochastic, we report the average under the column  $\text{mean}_{PI_{info}}$  and the half-width from a 95% confidence interval under the column  $\text{HW}_{PI_{info}}$ . It's defined as  $\text{HW}_{PI_{info}} = 1.96 \sigma_{PI_{info}}/\sqrt{N}$ , where  $\sigma_{PI_{info}}$  is the sample standard deviation and  $N = 100$  is the number of sample paths. For the upper bounds, we used the one step lookahead approach with the value function obtained from the DADP approach. The half-width is again defined similarly except we used  $N = 500$  sample paths. We see that the DADP approach provides improved lower bounds over the perfect information and state-independent approach. The upper bounds show that we are not too far from closing the optimality gap.

In Table 3, we report the solve times for the lower bounds. The results in each row were obtained on the same machine, so comparing solve times comparisons between approaches within each row are meaningful. However, different instances may have been run on different machines, so we should not compare solve times in different rows to each other. For most instances, the state-independent approach is faster than the DADP and perfect information approaches. The DADP approach is slower than the perfect information bound for the 15 generator case, but faster for the 30 generator instances. For the 50 generator instances, it was surprising that in comparison to the DADP approach the perfect information approach had a comparable speed for most instances and was even faster for a few of them. After looking into this further, we found that the 30 generator instances had a different mix of generators than the 50 generator instances, which made them more difficult to solve. In particular, a few of the generators were long term generators that had large minimum and maximum generation levels and once turned on had to remain on for the remainder of the time horizon. Overall, the DADP approach provides better bounds, and scales similarly and possibly better than the perfect information bounds. For the one-step lookahead upper bound, we solved a MIP for each time period for each of the 500 sample paths. The average solve time for each MIP in the first 50 generator instance was about 1.1 seconds, indicating that this may be practical for implementation.

## 5 Conclusion

In our numerical results, we included a single coupling demand constraint. If there are multiple loads, we could in principle have a coupling constraint for each one and relax each a different sets of multipliers. However, this would create further variables to optimize the Lagrangian function, and the obtained bounds may become weaker. Future work could address handling the case of multiple coupling constraints efficiently.

In this paper, we assumed that the demands were independent from one time period to another. A simple extension, such as having weather states that indicate the demand distribution, is possible. However, more complex modeling such as a two-level Markov model with hidden states [20] would require further investigation.

## 6 Acknowledgements

Thanks to Dr. Yanchao Liu for providing the FERC generator data. This work is supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under contract number DE-AC02-06CH11357.

## References

- [1] D. Adelman and A. J. Mersereau. Relaxations of weakly coupled stochastic dynamic programs. *Operations Research*, 56:712–727, 2008.
- [2] B. F. W. Allen J. Wood and G. B. Shebl. *Power Generation, Operation, and Control*. Wiley, 3 edition, 2013.
- [3] C. Archer, H. Simo, W. Kempton, W. Powell, and M. Dvorak. The challenge of integrating offshore wind power in the u.s. electric grid. part i: Wind forecast error. *Renewable Energy*, 103:346 – 360, 2017.
- [4] K. Barty, P. Carpentier, G. Cohen, and P. Girardeau. Price decomposition in large-scale stochastic optimal control. 2010.
- [5] K. Barty, P. Carpentier, and P. Girardeau. Decomposition of large-scale stochastic optimal control problems. *RAIRO Operations Research*, 44:167–183, 2010.

- [6] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2005.
- [7] D. P. Bertsekas. *Convex Optimization Algorithms*. Athena Scientific, 2015.
- [8] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, third edition, 2016.
- [9] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [10] C. C. Care and R. Schultz. A two-stage stochastic program for unit commitment under uncertainty in a hydro-thermal power system. In *PREPRINT SC 98-11, KONRAD-ZUSE-ZENTRUM FUR INFORMATIONSTECHNIK*, pages 98–13, 1998.
- [11] A. Frangioni, C. Gentile, and F. Lacalandra. Tighter approximated milp formulations for unit commitment problems. *IEEE Transactions on Power Systems*, 24(1):105–113, Feb 2009.
- [12] P. Girardeau. *Résolution de grands problèmes en optimisation stochastique dynamique et synthèse de lois de commande*. PhD thesis, Université Paris-Est, 2010.
- [13] S. Guo, X. Guan, and Q. Zhai. The necessary and sufficient conditions for determining feasible solutions to unit commitment problems with ramping constraints. In *IEEE Power Engineering Society General Meeting, 2005*, pages 344–349 Vol. 1, June 2005.
- [14] J. T. Hawkins. *A Lagrangian Decomposition Approach to Weakly Coupled Dynamic Optimization Problems and its Applications*. PhD thesis, Massachusetts Institute of Technology, 2003.
- [15] Y. Huang, P. M. Pardalos, and Q. P. Zheng. Electrical power unit commitment. 2017.
- [16] J. H. Kim and W. B. Powell. An hour-ahead prediction model for heavy-tailed spot prices. *Energy Economics*, 33(6):1252 – 1266, 2011.
- [17] V. Leclère. *Contributions to Decomposition Methods in Stochastic Optimization*. PhD thesis, Université Paris-Est, 2014.
- [18] S. S. Methods. Notes for ee364b, stanford university, winter 2006-07, April 2014.
- [19] J. M. Morales, A. J. Conejo, and J. Perez-Ruiz. Economic valuation of reserves in power systems with high penetration of wind power. *IEEE Transactions on Power Systems*, 24(2):900–910, May 2009.

- [20] J. Nascimento, D. Jiang, J. Durante, and W. Powell. A modified SDDP method for a risk-averse high-dimensional multistage stochastic energy storage optimization problem. 2017 INFORMS Computing Society Conference, 2017.
- [21] A. Papavasiliou, S. S. Oren, and R. P. O’Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Transactions on Power Systems*, 26(4):2197–2206, Nov 2011.
- [22] M. L. Puterman. *Markov Decision Processes*. John Wiley & Sons, Inc., 2005.
- [23] D. Rajan and S. Takriti. Minimum up/down polytopes of the unit commitment problem with start-up costs. *Research Report RC23628, IBM*, 2005.
- [24] T. Shiina and J. R. Birge. Stochastic unit commitment problem. *International Transactions in Operational Research*, 11(1):19–32, 2004.
- [25] S. Sridhar, J. Linderoth, and J. Luedtke. Locally ideal formulations for piecewise linear functions with indicator variables. *Operations Research Letters*, 41:627–632, 2013.
- [26] M. Tahanan, W. Ackooij, A. Frangioni, and F. Lacalandra. Large-scale unit commitment under uncertainty. *4OR*, 13(2):115–171, 2015.
- [27] S. Takriti and J. R. Birge. Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic programs. *Operations Research*, 48:091–098, 2000.
- [28] S. Takriti, B. Krasenbrink, and L. X.-Y. Wu. Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Research*, 48:268–280, 2000.
- [29] C. Ukun, A. Botterud, and J. R. Birge. An improved stochastic unit commitment formulation to accommodate wind uncertainty. *IEEE Transactions on Power Systems*, 31(4):2507–2517, July 2016.
- [30] L. Wang and C. Singh. Unit commitment considering generator outages through a mixed-integer particle swarm optimization algorithm. In *2006 IEEE Region 5 Conference*, pages 261–266, April 2006.
- [31] Q. P. Zheng, J. Wang, and A. L. Liu. Stochastic optimization for unit commitment—a review. *IEEE Transactions on Power Systems*, 30:1913–1924, 2015.



- [32] Q. P. Zheng, J. Wang, P. M. Pardalos, and Y. Guan. A decomposition approach to the two-stage stochastic unit commitment problem. *Annals of Operations Research*, 210(1):387–410, 2012.
- [33] J. Zhu. *Optimization of Power System Operation*. Wiley, 2 edition, 2015.

Table 2: Comparison of lower bounds from the state-independent Lagrangian approach, the DADP approach, and the perfect information bound. The upper bounds are generated from the one-step lookahead policy using the value function generated from the DADP approach.

$\mu$	$\sigma$	# Gen	Lower Bounds [\$ millions]				Upper Bound [\$ millions]	
			$\lambda_t$	$\lambda_t(v_t)$	mean $PI_{Info}$	HW $PI_{Info}$	mean	HW
0.4	0.15	15	9.17	9.65	9.48	0.03	9.94	0.02
0.4	0.20	15	9.17	10.17	9.73	0.04	10.53	0.02
0.4	0.25	15	9.17	10.85	10.11	0.04	11.33	0.03
0.6	0.15	15	14.96	17.75	16.89	0.08	18.08	0.05
0.6	0.20	15	14.97	19.84	18.29	0.13	20.19	0.07
0.6	0.25	15	14.96	22.52	19.79	0.19	22.70	0.10
0.8	0.15	15	26.99	35.04	33.76	0.31	35.80	0.14
0.8	0.20	15	26.98	39.93	37.98	0.41	41.52	0.19
0.8	0.25	15	26.99	45.15	43.07	0.54	46.95	0.26
0.4	0.15	30	7.15	8.99	8.49	0.04	9.31	0.03
0.4	0.20	30	7.15	10.07	9.17	0.05	10.73	0.05
0.4	0.25	30	7.15	11.17	9.91	0.07	12.53	0.08
0.6	0.15	30	11.54	15.09	13.85	0.06	15.90	0.06
0.6	0.20	30	11.53	16.95	15.16	0.10	18.24	0.09
0.6	0.25	30	11.54	18.90	16.63	0.16	20.76	0.12
0.8	0.15	30	16.75	23.87	22.27	0.26	25.59	0.13
0.8	0.20	30	16.75	28.87	26.52	0.33	31.64	0.23
0.8	0.25	30	16.75	34.12	31.48	0.46	37.50	0.22
0.4	0.15	50	11.59	12.93	12.29	0.05	13.03	0.03
0.4	0.20	50	11.59	13.93	12.85	0.06	13.96	0.05
0.4	0.25	50	11.59	15.12	13.47	0.08	15.18	0.06
0.6	0.15	50	20.40	23.65	22.17	0.12	24.50	0.08
0.6	0.20	50	20.40	26.45	23.79	0.19	28.15	0.15
0.6	0.25	50	20.40	30.05	26.21	0.26	31.03	0.14
0.8	0.15	50	31.86	47.02	43.87	0.53	47.98	0.27
0.8	0.20	50	31.85	56.42	51.97	0.82	57.81	0.43
0.8	0.25	50	31.84	67.13	62.13	1.06	68.42	0.55

Table 3: Solve times for the state-independent Lagrangian approach, the DADP approach, and the perfect information bound reported in minutes.

$\mu$	$\sigma$	# Gen	Solve Time [min]		
			$\lambda_t$	$\lambda_t(v_t)$	PInfo
0.4	0.15	15	3.2	40.8	3.7
0.4	0.20	15	7.6	151.5	7.3
0.4	0.25	15	7.6	150.9	12.8
0.6	0.15	15	7.9	150.5	4.8
0.6	0.20	15	3.2	40.4	2.4
0.6	0.25	15	12.8	99.4	7.8
0.8	0.15	15	5.3	88.4	3.9
0.8	0.20	15	7.8	137.1	3.7
0.8	0.25	15	8.5	136.4	3.8
0.4	0.15	30	20.3	137.9	986.2
0.4	0.20	30	51.6	341.1	1264.1
0.4	0.25	30	66.1	316.3	1120.5
0.6	0.15	30	66.5	315.1	480.5
0.6	0.20	30	67.1	316.2	1171.7
0.6	0.25	30	20.0	158.4	1379.1
0.8	0.15	30	20.2	160.3	151.7
0.8	0.20	30	48.1	347.1	435.7
0.8	0.25	30	47.9	347.9	515.9
0.4	0.15	50	34.8	271.7	225.9
0.4	0.20	50	21.8	249.4	263.5
0.4	0.25	50	21.6	248.9	281.6
0.6	0.15	50	22.2	250.0	174.2
0.6	0.20	50	18.9	256.5	101.1
0.6	0.25	50	23.9	281.7	58.7
0.8	0.15	50	10.3	134.6	8.8
0.8	0.20	50	22.6	250.7	21.1
0.8	0.25	50	22.3	253.7	20.9