https://dx.doi.org/10.1287/ijoc.2022.1235

A Polyhedral Study on Fuel-Constrained Unit Commitment

Kai Pan
Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, kai.pan@polyu.edu.hk

Ming Zhao
Department of Business Administration, Alfred Lerner College of Business and Economics, University of Delaware, Newark, DE 19716, USA, mzhao@udel.edu

Chung-Lun Li
Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, chung-lun.li@polyu.edu.hk

Feng Qiu
Energy Systems Division, Argonne National Laboratory, Lemont, IL 60439, USA, fqiu@anl.gov

The electricity production of a thermal generator is often constrained by the available fuel supply. These fuel constraints impose a maximum bound on the energy output over multiple time periods. Fuel constraints are increasingly important in electricity markets, due to two main reasons. First, as more natural gas-fired generators join the deregulated market, there is often competition for natural gas supply from other sectors (e.g., residential and manufacturing heating). Second, as more environmental and emission regulations are being placed on fossil fuel-fired generators, fuel supply is becoming more limited. However, there are few studies that consider the fuel constraints in the unit commitment problem from the perspective of computational analysis. To address the challenge faced by an independent power producer with a limited fuel supply, we study a fuel-constrained self-scheduling unit commitment (FSUC) problem where the production decisions are coupled across multiple time periods. We provide a complexity analysis of the FSUC problem and conduct a comprehensive polyhedral study by deriving strong valid inequalities. We demonstrate the effectiveness of our proposed inequalities as cutting planes in solving various multistage stochastic FSUC problems.

Key words: Unit commitment; fuel supply constraint; cutting planes; convex hull

1. Introduction

With 42% of the total electricity generation capacity, natural gas-fired generators provided 34% of total electricity generation in 2016, surpassing coal to become the largest generation source (US EIA 2017). As reported in US EIA (2021a), both the generation capacity and output of natural gas-fired generators continue to grow. Due to the environmental friendliness and continued cost-competitiveness of natural gas compared to coal, an increasing portion of independent power producers (IPPs) in the energy market are replacing coal-fired generators with natural gas-fired ones. Natural gas-fired generators are flexible and efficient in electricity generation because their fast ramping capabilities allow IPPs to mitigate the effect of uncertainties in today's complex power systems. A natural gas-fired generator can largely mitigate the effect of the electricity generation uncertainty caused by the increasing penetration of renewable energy. More important, a natural gas-fired generator can be efficiently self-scheduled by IPPs in order to maximize the total profit obtained from selling electricity in times of volatile electricity prices.

Although IPPs may enjoy the benefits of natural gas-fired generators, the profits of power generation largely depend on the availability of natural gas supply, and circumstances in which a natural gas-fired generator encounters a limited fuel supply are not uncommon in practice. IPPs with natural gas-fired generators are often considered lower priority than residential, commercial, and industrial users (US EIA 2021b) that are experiencing the same fuel shortage. The situation is more severe in cold weather when natural gas is mainly used for heating. For example, significant amounts of planned electricity generation from natural gas-fired generators became unavailable due to curtailed natural gas supply during the 2014 "polar vortex" weather conditions (NERC 2014). As more IPPs with natural gas-fired generators join power systems, the limited natural gas supply imposes an increasing challenge for an IPP's power generation scheduling, which needs to solve a self-scheduling unit commitment (UC) problem for the optimal generation schedule.

The UC problem, often formulated as a mixed-integer program (MIP) (e.g., Bixby 2010, Carlson et al. 2012, Li et al. 2021, Xavier et al. 2021), has received ample attention because it is widely used for power generation scheduling, electricity market clearing, operational reliability assessment, expansion planning, and other activities. As a fundamental problem in power system operations, the UC problem determines the optimal commitment (online/offline status) and production levels of generators while respecting the characteristics of the generators and physics of the power system (e.g., generation upper/lower bounds, minimum-up/-down time limits, ramping constraints). Because of the increasing importance of natural gas-fired generators, fuel supply constraints are receiving increased attention in industry practice. Fuel supply constraints, imposing a maximum bound on the energy output of a generator over multiple time periods, were first introduced in the late 1980s to reflect the daily or hourly limits on maximum fuel availability due to contractual obligations or technological constraints (Cohen & Wan 1987, Aoki et al. 1987, 1989). These constraints complicate the UC problem, but because the production decisions of the generator are now coupled across different time periods of the operation horizon, they offer relevant analyses for many applications. For example, Shahidehpour et al. (2005) show the short-term impact of natural gas prices on power generation scheduling by running security-constrained UC problems. Chen et al. (2019) incorporate gas flow into UC problems to optimize generator scheduling, given the supply constraints imposed by natural gas pipeline flows, and use linearized gas flow equations to derive local marginal prices. Zhao et al. (2017) study the impact of uncertain gas supply availability induced by gas transmission congestion and extremely high gas prices during peak demand.

Various solution approaches have been proposed to solve deterministic fuel-constrained UC problems. A majority of these studies adopt the Lagrangian relaxation (LR) method by relaxing several complicated constraints (including the fuel constraints) in the problem. An early attempt can be found in Cohen & Wan (1987), in which the load balance, reserve, and fuel constraints are relaxed and penalized in the objective function using Lagrange multipliers. There are some follow-up studies using the LR method, including Aoki et al. (1987, 1989), Ružić & Rajaković (1991), Kuloor et al. (1992), Baldick (1995), Shaw (1995), and Lu & Shahidehpour (2005), among others. A related method often used is the augmented LR method, which additionally adds quadratic penalty terms to the objective function of the relaxed problem. For instance, Ma & Shahidehpour (1999) add quadratic terms to penalize the load balance violation, and Fu et al. (2005) add quadratic terms to penalize the generation difference violation, while both studies penalize the fuel constraints using Lagrange multipliers. Another method commonly used for solving the UC problem with no fuel constraints is dynamic programming (Padhy 2004, Saravanan et al. 2013), and such method is also used in Al-Kalaani et al. (1996) for solving the fuel-constrained UC problem. Specifically, Al-Kalaani et al. (1996) derive an approach to transfer the fuel constraints into unit capacity limits and then approximate the original fuel-constrained UC problem by a UC problem with no fuel constraints, which is further solved by dynamic programming. Lee (1991) proposes a sequential commitment approach for solving the fuel-constrained UC problem, in which a tentative commitment schedule for each candidate unit is first determined, and then the most advantageous unit is sequentially committed by evaluating the economic benefit of each unit. Li et al. (1997) propose a unit decommitment approach for solving the fuel-constrained UC problem, where the units are turned off one at a time to reduce the total cost until no further cost reduction is possible. Vemuri & Lemonidis (1992) solve a fuel-constrained UC problem by decomposing it into two subproblems—a linear fuel dispatch subproblem and a unit commitment subproblem—and solve them iteratively until the algorithm converges to a near-optimal solution. Heuristic search methods are often used for solving fuel-constrained UC problems. For example, Amjady & Nasiri-Rad (2011) develop a solution method by integrating particle swarm optimization and genetic algorithm, and Bai & Shahidehpour (1997) develop a solution method by integrating tabu search with the augmented LR method. Stochastic fuel-constrained UC problems have also been studied. Takriti et al. (2000) and Wu et al. (2007, 2008) consider various types of uncertainties and propose different stochastic fuel-constrained UC models, and they apply LR methods to solve the problems. Saneifard et al. (1997) propose a fuzzy logic approach to tackle uncertainties in fuel-constrained UC problems. Our paper differs from these fuel-constrained UC research in that we apply polyhedral theory to a fuel-constrained UC problem and derive strong valid inequalities that can improve the effectiveness of a branch-and-cut solution process.

In this paper, we consider an IPP that faces a fuel supply constraint when self-scheduling its generator to maximize its total profits from selling electricity. Specifically, we study a fuel-constrained self-scheduling unit commitment (FSUC) model, which represents a core structure of power system operations with limited fuel supply. It is worth noting that, because our model considers an upper bound of the total power generation over the operation horizon, the setting can be extended to any generators (e.g., coal-fired ones) with limits on fuel supply or carbon emissions. For example, many environmental regulations require that the total generation of an IPP in one day cannot exceed a certain upper limit (El-Keib et al. 1994, Pulgar-Painemal 2005, Kockar et al. 2009). It is also worth noting that our model is applicable to hydropower production scheduling, where the fuel supply constraint is used for limiting the amount of water that can be used from the reservoir during the planning horizon.

Because cutting plane is an efficient approach to tightening an MIP formulation and speeding up the corresponding branch-and-cut algorithm (Nemhauser & Wolsey 1988), we conduct a comprehensive polyhedral study of the FSUC by deriving strong valid inequalities as cutting planes and convex hull descriptions to improve the computational performance of the problem with the FSUC embedded. There have been some studies on the polyhedral structure of the traditional UC model without fuel constraints. In particular, Lee et al. (2004) and Rajan & Takriti (2005) provide convex hull descriptions for the polytopes with only minimum-up/-down time constraints, while a more general result incorporating time-dependent and bounded up/down time limits is provided by Queyranne & Wolsey (2017). Morales-España et al. (2013) and Gentile et al. (2017) focus on tightening the generation upper bound constraints. Ostrowski et al. (2012), Damcı-Kurt et al. (2016), Pan et al. (2016), Pan & Guan (2016a), and Huang et al. (2021a) derive strong valid inequalities to strengthen various ramping constraints. Knueven et al. (2020) perform a comprehensive review of the various MIP models and evaluate their computational advantages. However, there is no work that studies the polyhedral structures of the UC problem with fuel constraints, which complicate the analysis. To fill this gap and help efficiently solve the practical problems, we investigate the FSUC problem by considering both the generator's physical characteristics and its fuel supply constraint, and we derive several families of strong valid inequalities to strengthen the original formulation by focusing on the impact of fuel constraint.

Our main contributions can be summarized as follows:

1. We provide a complexity analysis to show that the FSUC problem is NP-hard. To our knowledge, we are the first to show that a self-scheduling UC considering a single generator is NP-hard.

2. We derive a family of strong valid inequalities (semi-continuous inequalities) that are strong enough to describe the convex hull of the polytope by including only the generation upper/lower bounds and fuel constraint.
3. We derive another two families of strong valid inequalities for the FSUC problem, namely look forward inequalities and look backward inequalities, by considering the minimum-up/-down time limits, generation lower/upper bounds, ramping constraints, and fuel constraint. The look forward (resp. backward) inequalities bound the generation amount in a time period by considering the generator's online/offline status after (resp. before) that period.
4. We conduct extensive computational experiments to demonstrate the effectiveness of our proposed inequalities in improving branch-and-cut algorithms for solving various multistage stochastic FSUC problems.

The remainder of this paper is organized as follows. Section 2 introduces the MIP formulation of the FSUC problem and the corresponding complexity analysis results. Section 3 develops several families of strong valid inequalities for the FSUC. Section 4 demonstrates the effectiveness of the proposed strong valid inequalities through extensive computational experiments. Section 5 summarizes our results. All the proofs are presented in the Online Appendix.

2. The Model

The FSUC problem in its MIP formulation is a core of many planning and operational problems supporting power system operations. In this section, we model this core part as an MIP and summarize the difficulty of solving this model. We will show the corresponding applications of this model in Section 4 when we test our results derived from this core model.

The model involves binary decisions on a generator's online/offline status and the generation output in each time period when the generator is online. The physical characteristics of our focal generator are defined as follows. Let $L > 0$ and $\ell > 0$ be the minimum-up and minimum-down time limits, respectively, i.e., once the generator is online, it must stay online for at least L time periods. Similarly, once the generator is offline, it must stay offline for at least ℓ time periods. Let $\overline{C}$ and $\underline{C}$ be the upper and lower limits, respectively, on the amount of electricity that can be generated when the generator is online, where $\overline{C} > \underline{C} > 0$. Let $V > 0$ be the ramp-up/down rate limit in the stable generation region; that is, the generation amounts in two consecutive time periods must not differ from each other by more than V if the generator is online in these two periods. Let $\overline{V}$ be the start-up/shut-down ramp rate limit; that is, immediately after the generator starts up and immediately before the generator shuts down, the generation amount must not exceed $\overline{V}$. We assume that $V \leq \overline{V}$ and $\underline{C} < \overline{V} < \underline{C} + V$, both of which hold in most industrial settings. Let $\overline{U} \geq 0$ and $\underline{U} \geq 0$ be the start-up and shut-down costs, respectively, of the generator.

The operation horizon has T time periods, and we define $\mathcal{T} = \{1,\ldots,T\}$. For each $t \in \mathcal{T}$, we define the following quantities:

- c_t: the fixed cost incurred if the generator is online in period t ($c_t > 0$);
- ξ_t: the per-unit electricity price in period t ($\xi_t \geq 0$);
- $f_t(x_t)$: the non-decreasing convex piecewise linear generation cost in period t, where $f_t(0) = 0$;
- y_t: a binary decision variable such that $y_t = 1$ if the generator is online in period t, and $y_t = 0$ otherwise;
- u_t: a binary decision variable such that $u_t = 1$ if the generator starts up in period t, and $u_t = 0$ otherwise;
- x_t: a continuous decision variable indicating the generation amount in period t.

Note that the generation cost is a non-decreasing convex piecewise linear function of x_t. The piecewise linear cost is commonly used in practice to approximate the quadratic cost function $c_t'' x_t^2 + c_t' x_t$ (Carrión & Arroyo 2006).

Let Q be the maximum fuel supply (in terms of power output) over the entire operation horizon of T periods. In addition, we let $\bar{L} = \max\{L, \ell\}$ and assume that the values of $u_{-\bar{L}+2}, u_{-\bar{L}+3}, \ldots, u_{-1}, u_0, y_{-\ell+1}, y_{-\ell+2}, \ldots, y_{-1}, y_0, x_0$ are given as initial conditions. The FSUC problem can be formulated as follows:

$$\text{Problem (1):}\quad \min \quad \sum_{t=1}^{T} \Big[f_t(x_t) + c_t y_t + \overline{U} u_t + \underline{U}(y_{t-1} - y_t + u_t) - \xi_t x_t \Big] \tag{1a}$$

$$\text{s.t.} \quad \sum_{i=t-L+1}^{t} u_i \leq y_t, \ \forall t \in \mathcal{T}, \tag{1b}$$

$$\sum_{i=t-\ell+1}^{t} u_i \leq 1 - y_{t-\ell}, \ \forall t \in \mathcal{T}, \tag{1c}$$

$$y_t - y_{t-1} - u_t \leq 0, \ \forall t \in \mathcal{T}, \tag{1d}$$

$$-x_t + \underline{C} y_t \leq 0, \ \forall t \in \mathcal{T}, \tag{1e}$$

$$x_t - \overline{C} y_t \leq 0, \ \forall t \in \mathcal{T}, \tag{1f}$$

$$x_t - x_{t-1} \leq V y_{t-1} + \overline{V}(1 - y_{t-1}), \ \forall t \in \mathcal{T}, \tag{1g}$$

$$x_{t-1} - x_t \leq V y_t + \overline{V}(1 - y_t), \ \forall t \in \mathcal{T}, \tag{1h}$$

$$\sum_{t=1}^{T} x_t \leq Q, \tag{1i}$$

$$y_t \in \{0,1\}, \ u_t \in \{0,1\}, \ \forall t \in \mathcal{T}. \tag{1j}$$

Objective function (1a) minimizes the total cost, including the fixed and variable generation costs and the start-up and shut-down costs, minus the total revenue obtained from selling the electricity. Here, the term "$\underline{U}(y_{t-1} - y_t + u_t)$" is the shut-down cost in period t. Note that constraints (1b) and (1d), together with the minimization objective, ensure that the shut-down cost $\underline{U}$ is incurred in period t (i.e., $y_{t-1} - y_t + u_t = 1$) if and only if $y_{t-1} = 1$ and $y_t = u_t = 0$. Constraints (1b) and (1c) model the requirements of the minimum-up and minimum-down time limits, respectively. The minimum-up time limit requires that if the generator starts up in period $t-L+1$ (i.e., $y_{t-L} = 0$, $y_{t-L+1} = 1$), then it needs to stay online in the following L consecutive periods (including $t-L+1$) until period t; and the minimum-down time limit requires that if the generator shuts down in period $t-\ell+1$ (i.e., $y_{t-\ell} = 1$ and $y_{t-\ell+1} = 0$), then it needs to stay offline in the following ℓ consecutive periods (including $t-\ell+1$) until period t. Constraints (1d) describe the relations between online status and start-up action. It requires $u_t = 1$ when $y_t = 1$ and $y_{t-1} = 0$. Constraints (1e) and (1f) model the generation lower and upper bounds if the generator is online. Constraints (1g) and (1h) model the ramp-up and ramp-down rate limits, respectively, between two consecutive time periods. The ramp-up rate limit requires that if the generator is online in period $t-1$ (i.e., $y_{t-1} = 1$), then the generation increment from $t-1$ to t should be no more than V; otherwise, it should be no more than $\overline{V}$. The ramp-down rate limit requires that if the generator is online in period t (i.e., $y_t = 1$), then the generation decrement from $t-1$ to t should be no more than V; otherwise, it should be no more than $\overline{V}$. Constraint (1i) models the fuel supply limit.

We assume that all the given parameters in Problem (1) are rational numbers. To avoid trivial cases, we assume that $L, \ell \leq T-1$ and $(L+1)\underline{C} < Q < T\overline{C}$. Note that the objective function of Problem (1) is piecewise linear. Following the existing literature (Carrión & Arroyo 2006, Frangioni et al. 2008), Problem (1) can be converted into an MIP. Note also that constraint (1i) follows the literature (e.g., Lee 1991, Kuloor et al. 1992, Vemuri & Lemonidis 1992, Zhao et al. 2017) to impose an upper bound on the total power generation over the operation horizon, and accordingly helps maintain the whole model as an MIP. To focus on the effects of limited gas supply, here we only impose the gas supply upper limit and omit the lower limit that is also considered by Lee (1991) and Kuloor et al. (1992).

The following proposition states the computational complexity of the problem.

Proposition 1. *Problem (1) is NP-hard.*

Remark 1. Existing studies such as Frangioni & Gentile (2006), Damcı-Kurt et al. (2016), and Guan et al. (2018) have shown that when there is no fuel supply constraint, similar variants of Problem (1) with different setups of initial conditions can be solved polynomially in $\mathcal{O}(T^3)$ time. When constraint (1i) is removed, we can follow their approaches to solving Problem (1) in polynomial time, because the problem can be reduced to a shortest-path problem on an acyclic network with $\mathcal{O}(T^2)$ nodes and $\mathcal{O}(T^2)$ arcs, where each node, denoted by a time-index pair (h,k) such that $k \geq h+L-1$, represents a time period $\{h, h+1, \ldots, k\}$ in which the generator is online, and each arc represents a transition between two nodes that has to satisfy the minimum-down time requirement. By Proposition 1, Problem (1) is NP-hard. This implies that the fuel supply constraint significantly increases the problem complexity. In addition, the fuel supply constraint may increase the computational burden substantially when Problem (1) is embedded in other large-scale problems.

Remark 2. From the proof of Proposition 1, it is not difficult to see that the NP-hardness proof remains valid when constraints (1b), (1c), (1d), (1g), and (1h) are removed from Problem (1) and when $\overline{U} = \underline{U} = 0$. Hence, the major factors that contribute to the high computational complexity of the problem are (i) the fuel supply limit constraint (1i), (ii) the non-linearity of the generation cost function $f_t(x_t)$, and (iii) the time-dependency of the cost parameters.

Remark 3. If $Q \geq T\overline{C}$, then constraint (1i) becomes redundant, and by Remark 1, Problem (1) becomes polynomial-time solvable. If $Q \leq (L+1)\underline{C}$, then there are at most two stable generation regions in the operation horizon. In this case, there are $\mathcal{O}(T^4)$ possible combinations of start-up and shut-down periods. Once the start-up and shut-down periods are known, the generation amounts in different periods can be determined in polynomial time by a linear program. In this case, Problem (1) is also polynomial-time solvable. Hence, in this paper, we focus on the case where $(L+1)\underline{C} < Q < T\overline{C}$.

3. Strong Valid Inequalities

In this section, we focus on developing and analyzing various families of strong valid inequalities that can be applied to help efficiently solve Problem (1), as well as other extensions with Problem (1) embedded (e.g., multistage stochastic FSUC). For notational convenience, we let $\mathbb{R}$ and $\mathbb{Z}$ denote the set of real numbers and the set of integers, respectively. We define $[n_1, n_2]_{\mathbb{Z}}$ as the set of integers between integers n_1 and n_2. That is, $[n_1, n_2]_{\mathbb{Z}} = \{n_1, n_1+1, \cdots, n_2\}$ if $n_1 \leq n_2$, and $[n_1, n_2]_{\mathbb{Z}} = \emptyset$ otherwise. We also define $[n]^+ = n$ if $n \geq 0$, and $[n]^+ = 0$ otherwise.

Let $\mathcal{D}_0 = \{(x, y, u) \in \mathbb{R}^T \times \{0,1\}^T \times \{0,1\}^T : \text{(1b)} - \text{(1i)}\}$. We consider the following linear relaxation of set $\mathcal{D}_0$:

$$\mathcal{D} = \Big\{ \ (x, y, u) \in \mathbb{R}^T \times \{0,1\}^T \times \{0,1\}^{T-1} :$$

$$\sum_{i=t-L+1}^{t} u_i \leq y_t, \ \forall t \in [L+1, T]_{\mathbb{Z}}, \tag{2a}$$

$$\sum_{i=t-\ell+1}^{t} u_i \leq 1 - y_{t-\ell}, \ \forall t \in [\ell+1, T]_{\mathbb{Z}}, \tag{2b}$$

$$y_t - y_{t-1} - u_t \leq 0, \ \forall t \in \mathcal{T} \setminus \{1\}, \tag{2c}$$

$$-x_t + \underline{C} y_t \leq 0,\ \forall t \in \mathcal{T}, \tag{2d}$$

$$x_t - \overline{C} y_t \leq 0,\ \forall t \in \mathcal{T}, \tag{2e}$$

$$x_t - x_{t-1} \leq V y_{t-1} + \overline{V}(1 - y_{t-1}),\ \forall t \in \mathcal{T} \setminus \{1\}, \tag{2f}$$

$$x_{t-1} - x_t \leq V y_t + \overline{V}(1 - y_t),\ \forall t \in \mathcal{T} \setminus \{1\}, \tag{2g}$$

$$\sum_{t=1}^{T} x_t \leq Q \Big\}. \tag{2h}$$

Note that variable u_1 is not included in $\mathcal{D}$ and that $\mathcal{D}$ is $(3T-1)$-dimensional. Let $\mathrm{conv}(\mathcal{D})$ be the convex hull of $\mathcal{D}$. Let $\mathcal{D}' = \{(x, y, u) \in \mathbb{R}^T \times [0,1]^T \times [0,1]^{T-1} : (2a) - (2h)\}$, which is a linear relaxation of $\mathcal{D}$.

Proposition 2. *The polytope* $\mathrm{conv}(\mathcal{D})$ *is full dimensional.*

In the following, we develop several families of strong valid inequalities for $\mathrm{conv}(\mathcal{D})$. Note that the set $\mathcal{D}$ contains a subset of the variables and constraints in $\mathcal{D}_0$, and that $\mathcal{D}$ is independent of $u_{-\bar{L}+2}, u_{-\bar{L}+3}, \ldots, u_{-1}, u_0, y_{-\ell+1}, y_{-\ell+2}, \ldots, y_{-1}, y_0, x_0$. It follows that any strong valid inequalities for $\mathrm{conv}(\mathcal{D})$ are also valid for $\mathrm{conv}(\mathcal{D}_0)$ and can be used to help solve Problem (1), regardless of the initial conditions. Note also that Rajan & Takriti (2005) prove that constraints (2a) – (2c) provide the convex hull description of all feasible solutions in the space of y and u variables when considering only the minimum-up/-down time limits with start-up costs. In Section 3.1, we derive valid inequalities by considering a relaxation of set $\mathcal{D}$. In Section 3.2, we derive two families of valid inequalities for $\mathrm{conv}(\mathcal{D})$.

3.1. Semi-Continuous Inequalities

The generation outputs in the operation horizon, $x_t \in \{0\} \cup [\underline{C}, \overline{C}], \forall t \in \mathcal{T}$, are represented by a set of semi-continuous variables (Beale 1985) and are linked by the fuel supply constraint (2h). This gives rise to a semi-continuous knapsack set defined as $\mathcal{D}_{sc} = \{x \in \mathbb{R}^T : (2h);\ x_t \in \{0\} \cup [\underline{C}, \overline{C}]\ \forall t \in \mathcal{T}\}$, which plays a central role in our model. De Farias & Zhao (2013) have conducted a polyhedral study of a general semi-continuous knapsack problem. However, simply applying their inequalities to the set $\mathcal{D}_{sc}$ only gives us valid inequalities. In this subsection, we exploit the unique structure of the set $\mathcal{D}_{sc}$, characterize all of its extreme points, and then derive its convex hull description.

We refer to an inequality as a *semi-continuous inequality* if it is facet-defining for $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$. Denote $\mathcal{D}'_{sc} = \{(x, y, u) \in \mathbb{R}^T \times \mathbb{R}^T \times \mathbb{R}^{T-1} : x \in \mathcal{D}_{sc}\}$. Clearly, $\mathcal{D} \subseteq \mathcal{D}'_{sc}$. Thus, any valid inequality for $\mathrm{conv}(\mathcal{D}'_{sc})$ is also valid for $\mathrm{conv}(\mathcal{D})$. Because $\mathcal{D}'_{sc}$ simply lifts $\mathcal{D}_{sc}$ by introducing $2T-1$ dimensions of y and u with no constraints enforced in these dimensions, any valid inequality for $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ is valid for $\mathrm{conv}(\mathcal{D}'_{sc})$. Hence, those inequalities that are valid for $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ are also valid for $\mathrm{conv}(\mathcal{D})$.

Let $\lambda = \lfloor Q/\overline{C} \rfloor$, $\lambda' = \lceil (\underline{C} + \lambda\overline{C} - Q)/(\overline{C} - \underline{C}) \rceil$, and $\theta^* = Q - \lambda\overline{C} + \lambda'(\overline{C} - \underline{C})$. The following lemma characterizes all the extreme points of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$.

Lemma 1. *Each extreme point of* $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ *must belong to one of the following two categories:*

1. *No more than* λ *components of the extreme point are equal to* $\overline{C}$*, and the other components are 0.*
2. *One component of the extreme point is equal to* θ^**. Among the remaining* $T-1$ *components,* $\lambda-\lambda'$ *components are equal to* $\overline{C}$*, and* λ' *components are equal to* $\underline{C}$*, while the other components are 0.*

Based on this characterization of the extreme points, the following proposition offers a family of inequalities that are valid for $\mathrm{conv}(\mathcal{D})$ and strengthen (2h).

Proposition 3. *If* $\lambda \geq \lambda'$*, then for any* $\mathcal{T}_1 \subseteq \mathcal{T}$ *such that* $\lambda - \lambda' + 1 \leq |\mathcal{T}_1| \leq \lambda$*, the inequality*

$$\sum_{t\in\mathcal{T}_1} x_t + \rho \sum_{t\in\mathcal{T}\setminus\mathcal{T}_1} x_t \leq \zeta \tag{3}$$

is a semi-continuous inequality and is valid for $\mathrm{conv}(\mathcal{D})$*, where*

$$\rho = 1 - \frac{Q-\lambda\overline{C}}{\underline{C}-(\lambda-|\mathcal{T}_1|)(\overline{C}-\underline{C})} \text{ and } \zeta = \big[|\mathcal{T}_1| + \rho(\lambda - |\mathcal{T}_1|)\big]\overline{C}.$$

Example 1. Let $T=6$, $Q=19$, $\underline{C}=4$, and $\overline{C}=6$. Then, $\lambda=3$ and $\lambda'=2$. Thus, inequality (3) holds for any $\mathcal{T}_1$ with $2 \leq |\mathcal{T}_1| \leq 3$. We have $(\rho,\zeta)=(1/2,15)$ when $|\mathcal{T}_1|=2$, and $(\rho,\zeta)=(3/4,18)$ when $|\mathcal{T}_1|=3$. For example, we obtain the following two valid inequalities if we set $\mathcal{T}_1=\{1,2\}$ and $\mathcal{T}_1=\{1,3,5\}$ respectively:

$$x_1 + x_2 + 0.50x_3 + 0.50x_4 + 0.50x_5 + 0.50x_6 \leq 15;$$
$$x_1 + 0.75x_2 + x_3 + 0.75x_4 + x_5 + 0.75x_6 \leq 18.$$

The family of inequalities (3), along with (2h) and the condition $0 \leq x_t \leq \overline{C}$, is sufficient to provide the linear description of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$, as stated in the following proposition.

Proposition 4. *The polytope* $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ *is equal to* $\{x\in\mathbb{R}^T : \text{(2h)};\ \text{(3)}\ \forall \mathcal{T}_1 \subseteq \mathcal{T} \text{ s.t. } \lambda-\lambda'+1 \leq |\mathcal{T}_1| \leq \lambda;\ 0 \leq x_t \leq \overline{C}\ \forall t\in\mathcal{T}\}$*.*

Note that the size of the inequality family (3) is exponential in T. The following proposition states that the separation can be done in polynomial time.

Proposition 5. *Given an optimal solution of the linear programming (LP) relaxation of Problem (1), there exists an* $\mathcal{O}(T\log T)$ *time algorithm to find the most violated inequality (3), if any.*

Inequality (3) is usually not facet-defining for $\mathrm{conv}(\mathcal{D})$ because the ramping constraints are not considered in $\mathcal{D}_{sc}$. In Section 3.2, we introduce two families of strong valid inequalities for $\mathrm{conv}(\mathcal{D})$ that take into account the supply limit, minimum-up/-down time, generation lower/upper bounds, and ramping constraints.

3.2. Look Forward and Look Backward Inequalities

Here we define some functions and variables that will be used for describing our strong valid inequalities.

Definition 1. For $n_1, n_2 \in \mathbb{Z}$, define

$$
\begin{aligned}
\Delta(n_1,n_2) &= \frac{n_1(n_1+1)}{2} + \frac{n_2(n_2+1)}{2};\\
\omega_1(n_1,n_2) &= \underline{C}\,(n_1+n_2+1) + \left[\frac{n_1(n_1-1)+n_2(n_2-1)}{2} + \max\{n_1,n_2\}\right] V;\\
\omega_2(n_1,n_2) &= \underline{C}\,(n_1+n_2+1) + \left[\frac{|n_1(n_1+1)-n_2(n_2+1)|}{2} + n_1 n_2\right] V;\\
\omega_3(n_1,n_2) &= 2\underline{C} + \overline{V}\,(n_1+n_2+1) + \left[\frac{n_1(n_1-1)+n_2(n_2-1)}{2} + \max\{n_1,n_2\}\right] V;\\
\omega_4(n_1,n_2) &= \underline{C} + \overline{V}\,(n_1+n_2+1) + \left[\frac{|n_1(n_1+1)-n_2(n_2+1)|}{2} + n_1 n_2\right] V.
\end{aligned}
$$

Definition 2. For any $t \in \mathcal{T}$, define

$$
\alpha_t = \begin{cases} \max\{n\in\mathbb{Z} : \omega_1(n,n)\le Q\}, & \text{if } Q < Q_{B1};\\ \max\{n\in\mathbb{Z} : \omega_2(n,T-t)\le Q\}, & \text{if } Q_{B1}\le Q < Q_{B2} \text{ and } t-2 \ge T-t;\\ t-1, & \text{otherwise;}\end{cases}
$$

$$
\beta_t = \begin{cases} \max\{n\in\mathbb{Z} : \omega_1(n,n)\le Q\}, & \text{if } Q < Q_{B1};\\ \max\{n\in\mathbb{Z} : \omega_2(t-1,n)\le Q\}, & \text{if } Q_{B1}\le Q < Q_{B2} \text{ and } t \le T-t;\\ T-t, & \text{otherwise;}\end{cases}
$$

$$
\tau_{t1} = \begin{cases} 1, & \text{if } (\alpha_t \ne t-1,\ \beta_t \ne T-t, \text{ and } \omega_3(\alpha_t,\beta_t)\le Q)\\ & \text{or } (\alpha_t \ne t-1,\ \beta_t = T-t, \text{ and } \omega_4(\alpha_t,\beta_t)\le Q);\\ 0, & \text{otherwise;}\end{cases}
$$

$$
\tau_{t2} = \begin{cases} 1, & \text{if } (\alpha_t \ne t-1,\ \beta_t \ne T-t, \text{ and } \omega_3(\alpha_t,\beta_t)\le Q)\\ & \text{or } (\alpha_t = t-1,\ \beta_t \ne T-t, \text{ and } \omega_4(\alpha_t,\beta_t)\le Q);\\ 0, & \text{otherwise}\end{cases}
$$

where $Q_{B1} = \omega_1(\min\{t-1,T-t\},\min\{t-1,T-t\})$ and $Q_{B2} = \omega_2(t-1,T-t)$.

Proposition 6. *Given any $t \in \mathcal{T}$, (i) the values of α_t, β_t, τ_{t1}, and τ_{t2} can be obtained in $\mathcal{O}(1)$ time, provided that the floor function and square root function can be evaluated in constant time, and (ii) $\alpha_t, \beta_t \ge 0$.*

The quantities α_t, β_t, τ_{t1}, and τ_{t2} are used for describing the ramping pattern of the generator if it is online for the maximum number of periods allowed by the supply limit. For instance, if we would like to maximize the generation amount in period t with a given limited fuel supply limit Q, then α_t and τ_{t1} (resp. β_t and τ_{t2}) are used to describe the maximum number of online time periods before (resp. after) period t. Specifically, the generation amount will increase continuously at the maximum ramp-up rate V from period $t-\alpha_t$ to period t, and then decrease continuously at the maximum ramp-down rate V from period t to period $t+\beta_t$. If the generator is also able to be online in period $t-\alpha_t-1$ (resp. $t+\beta_t+1$) under the supply limit, then $\tau_{t1}=1$ (resp. $\tau_{t2}=1$). Note that in such a case, the generation difference between periods $t-\alpha_t-1$ and $t-\alpha_t$ (resp. between periods $t+\beta_t$ and $t+\beta_t+1$) is strictly smaller than V. The following proposition provides some properties of these quantities.

PROPOSITION 7. *For any $t\in\mathcal{T}$, (i) if $Q<Q_{B1}$, then $\alpha_t=\beta_t$ and $\tau_{t1}=\tau_{t2}$; (ii) if $Q_{B1}\le Q<Q_{B2}$, then $Q\ge\omega_2(\alpha_t,\beta_t)$, and $\tau_{t1}=0$ or $\tau_{t2}=0$; (iii) $\alpha_t+\tau_{t1}\le t-1$ and $\beta_t+\tau_{t2}\le T-t$.*

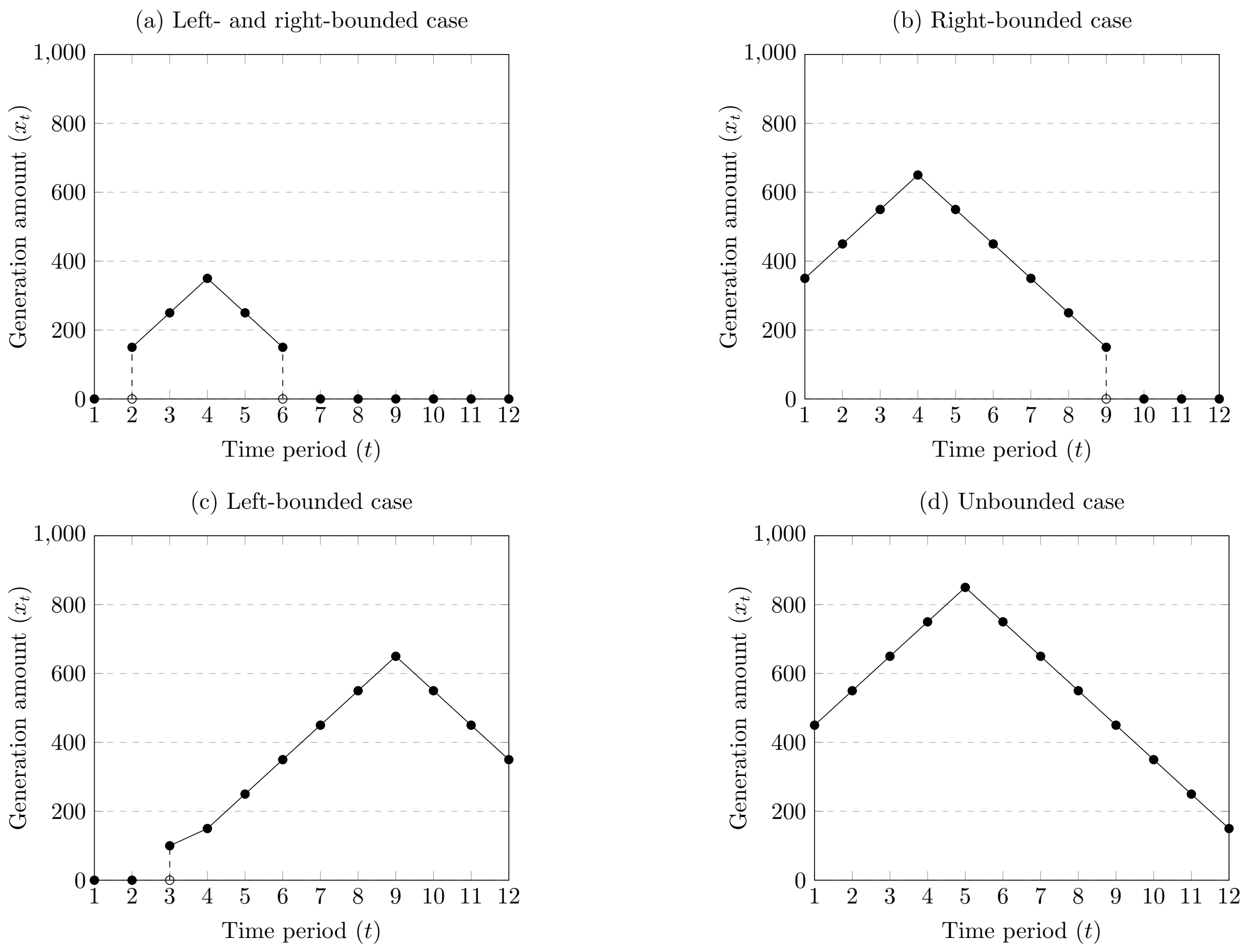

Figure 1 Ramping Patterns

Figure 1 depicts four exemplary cases, where we consider a generator with $\overline{C}=1000$, $\underline{C}=100$, $\overline{V}=150$, $V=100$, and $L=\ell=1$, and an operation horizon of $T=12$.

- First, consider a low fuel supply limit $Q=1155$, and suppose that we want to maximize the generation amount in period $t=4$. Note that $Q_{B1}=\omega_1(3,3)=1600$, and thus $Q<Q_{B1}$. It is easy to verify that $\alpha_t=\beta_t=2$ and $\tau_{t1}=\tau_{t2}=0$. As shown in Figure 1(a), the generator should

start up in period 2 (i.e., the period $t-\alpha_t-\tau_{t1}$) with the generation amount 150, ramp up at the maximum ramp-up rate until period 4, ramp down at the maximum ramp-down rate until period 6 (i.e., the period $t+\beta_t+\tau_{t2}$), and shut down in period 7. We refer to such a case as the "left- and right-bounded case."

- Second, consider a medium fuel supply limit $Q=3755$, and suppose that we want to maximize the generation amount in period $t=4$. Note that $Q_{B1}=\omega_1(3,3)=1600$ and $Q_{B2}=\omega_2(3,8)=6600$, and thus $Q_{B1}\leq Q<Q_{B2}$. It is easy to verify that $\alpha_t=3$, $\beta_t=5$, and $\tau_{t1}=\tau_{t2}=0$. As shown in Figure 1(b), the generator should stay online starting in period 1 (i.e., the period $t-\alpha_t-\tau_{t1}$), ramp up at the maximum ramp-up rate until period 4, ramp down at the maximum ramp-down rate until the generation amount reaches 150 in period 9 (i.e., the period $t+\beta_t+\tau_{t2}$), and shut down in period 10. We refer to such a case as the "right-bounded case."
- Third, consider a slightly larger fuel supply limit $Q=3850$, and suppose that we want to maximize the generation amount in period $t=9$. Note that $Q_{B1}=\omega_1(3,3)=1600$, $Q_{B2}=\omega_2(8,3)=6600$, and $\omega_4(5,3)=3850$, and thus $Q_{B1}\leq Q<Q_{B2}$ and $\omega_4(5,3)\leq Q$. It is easy to verify that $\alpha_t=5$, $\beta_t=3$, $\tau_{t1}=1$, and $\tau_{t2}=0$. As shown in Figure 1(c), the generator should start up in period 3 (i.e., the period $t-\alpha_t-\tau_{t1}$) with the generation amount 100, ramp up to the generation amount 150 in period 4, ramp up at the maximum ramp-up rate from period 4 until period 9, and ramp down at the maximum ramp-down rate until period 12 (i.e., the period $t+\beta_t+\tau_{t2}$). We refer to such a case as the "left-bounded case."
- Fourth, consider a large fuel supply limit $Q=5805$, and suppose that we want to maximize the generation amount in period $t=5$. Note that $Q_{B2}=\omega_2(4,7)=5800$, and thus $Q\geq Q_{B2}$. It is easy to verify that $\alpha_t=4$, $\beta_t=7$, and $\tau_{t1}=\tau_{t2}=0$. As shown in Figure 1(d), the generator should stay online throughout the entire operation horizon. It should ramp up at the maximum ramp-up rate from period 1 (i.e., period $t-\alpha_t-\tau_{t1}$) until period 5, and then ramp down at the maximum ramp-down rate until period 12 (i.e., the period $t+\beta_t+\tau_{t2}$). We refer to such a case as the "unbounded case."

We now derive strong valid inequalities to bound the generation amount x_t in period t. Note that if we attempt to reach the largest possible generation amount x_t in period t, then the generator should be online immediately before and after period t, and the ramp-up (resp. ramp-down) rate from periods $t-i$ to t (resp. from periods t to $t+j$) should be set equal to its maximum possible value V for some $i\in[0,t-1]_{\mathbb{Z}}$ (resp. $j\in[0,T-t]_{\mathbb{Z}}$). When there is no supply restriction, the values of i and j can be as large as $t-1$ and $T-t$, respectively. However, with the supply limit constraint (2h), we can restrict the values of i and j to tighter ranges. We can shrink the range for i from $[0,t-1]_{\mathbb{Z}}$ to $[0,\alpha_t]_{\mathbb{Z}}$ and shrink the range for j from $[0,T-t]_{\mathbb{Z}}$ to $[0,\beta_t]_{\mathbb{Z}}$. With such a reduced number of online time periods before and after period t, we are able to derive tighter upper bounds on x_t. The following propositions present two families of strong valid inequalities developed based on this idea. Proposition 8 presents inequalities that bound the generation amount x_t in period t by considering the generator's online/offline status after period t, while Proposition 9 presents inequalities that bound the generation amount x_t in period t by considering the generator's online/offline status before period t. We refer to the families of inequalities in Propositions 8 and 9 as *look forward inequalities* and *look backward inequalities*, respectively.

Proposition 8. *Consider any* $L \geq 2$, $t \in [L+1, T]_{\mathbb{Z}}$, $k \in [2, \beta_t]_{\mathbb{Z}}$, $S' \subseteq [t', t+k-1]_{\mathbb{Z}}$, *and* $S = [t+1, t'-1]_{\mathbb{Z}} \cup S'$, *where* $t' = \min\{t+L, t+k\}$. *The inequality*

$$x_t \leq V\left(\sum_{j=[L-k]^+}^{L-1} \min\{L-1-j, j\} u_{t-j} + \sum_{j=1}^{L-k-1} j u_{t-j}\right) + \sum_{i \in S} (d_i - i) V \left(y_i - \sum_{j=0}^{L-1} u_{i-j}\right)$$
$$+ \left(\frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1} - (k-1)V - \overline{V}\right)\left(y_{t+k} - \sum_{j=0}^{L-1} u_{t+k-j}\right) + \overline{V} y_t \tag{4}$$

is valid for $\operatorname{conv}(\mathcal{D})$, *where* $d_i = \min\{a \in S \cup \{t+k\} : a > i\}$ *for any* $i \in S$.

Remark 4. Under certain conditions, the valid inequality stated in Proposition 8 is facet-defining. Specifically, if $Q \geq Q_{B2}$, $\overline{C} \geq [Q + \Delta(t-1, T-t)V]/T$, $t+k = T$, $k \geq (L+1)/2$, and $S' = [t', t+k-1]_{\mathbb{Z}}$, then inequality (4) is facet-defining for $\operatorname{conv}(\mathcal{D})$. The proof is provided in Appendix A.10.

For notational convenience, we define an auxiliary variable $w_t = y_{t-1} - y_t + u_t$ to represent the shut-down action in period t. Thus, $w_t = 1$ when the generator is online in period $t-1$ and offline in period t (i.e., $y_{t-1} = 1$ and $y_t = 0$), and $w_t = 0$ otherwise.

Proposition 9. *Consider any* $L \geq 2$, $t \in [2, T-L]_{\mathbb{Z}}$, $k \in [2, \alpha_t]_{\mathbb{Z}}$, $S' \subseteq [t-k+1, t']_{\mathbb{Z}}$, *and* $S = S' \cup [t'+1, t-1]_{\mathbb{Z}}$, *where* $t' = \max\{t-L, t-k\}$. *The inequality*

$$x_t \leq V\left(\sum_{j=[L-k]^+ + 1}^{L} \min\{L-j, j-1\} w_{t+j} + \sum_{j=1}^{L-k} (j-1) w_{t+j}\right) + \sum_{i \in S} (i - d_i) V \left(y_i - \sum_{j=1}^{L} w_{i+j}\right)$$
$$+ \left(\frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1} - (k-1)V - \overline{V}\right)\left(y_{t-k} - \sum_{j=1}^{L} w_{t-k+j}\right) + \overline{V} y_t \tag{5}$$

is valid for $\operatorname{conv}(\mathcal{D})$, *where* $d_i = \max\{a \in \{t-k\} \cup S : a < i\}$ *for any* $i \in S$.

Example 2. Consider an FSUC problem with $T = 8$, a generator with $\underline{C} = 13$, $\overline{C} = 111$, $V = 11$, $\overline{V} = 18$, $L = \ell = 2$, and a fuel supply limit $Q = 296$. When $t = 5$, it is easy to verify that $\alpha_5 = 4$, $\beta_5 =$

3, $\tau_{51}=0$, and $\tau_{52}=0$. In Proposition 8, when $k=3$, we have $t'=7$, $S'\subseteq\{7\}$, and $S=\{6\}\cup S'$. We obtain the following two look forward inequalities below if we set $S'=\emptyset$ and $S'=\{7\}$, respectively:

$$x_5 - 18y_5 - 22y_6 - 19y_8 + 22u_5 + 22u_6 + 19u_7 + 19u_8 \leq 0;$$
$$x_5 - 18y_5 - 11y_6 - 11y_7 - 19y_8 + 11u_5 + 22u_6 + 30u_7 + 19u_8 \leq 0.$$

Note that according to Remark 4, the second inequality is facet-defining for $\text{conv}(\mathcal{D})$. When $t=6$, it is easy to verify that $\alpha_6=4$, $\beta_6=2$, $\tau_{61}=0$, and $\tau_{62}=0$. In Proposition 9, when $k=3$, we have $t'=4$, $S'\subseteq\{4\}$, and $S=S'\cup\{5\}$. We obtain the look backward inequalities below if we set $S'=\emptyset$ and $S'=\{4\}$, respectively:

$$x_6 - \frac{159}{7}y_5 - 18y_6 - 22y_7 + \frac{159}{7}u_4 + \frac{159}{7}u_5 + 22u_6 + 22u_7 \leq 0;$$
$$x_6 - \frac{159}{7}y_5 - 29y_6 - 11y_7 + \frac{159}{7}u_4 + \frac{236}{7}u_5 + 22u_6 + 11u_7 \leq 0.$$

The sizes of the inequality families (4) and (5) are exponential in T. The following propositions state that the separation can be done in polynomial time.

Proposition 10. *Given an optimal solution of the LP relaxation of Problem* (1)*, there exists an* $\mathcal{O}(T^4)$ *time algorithm to find the most violated inequality* (4)*, if any.*

Proposition 11. *Given an optimal solution of the LP relaxation of Problem* (1)*, there exists an* $\mathcal{O}(T^4)$ *time algorithm to find the most violated inequality* (5)*, if any.*

The strong valid inequalities (3)–(5) can be added to commercial solvers as user cuts. In our computational experiments, as described in the next section, we select a subset of these inequalities and add them to the user cut pool of the CPLEX optimizer, from which CPLEX can efficiently select the effective user cuts to help solve the problem in the branch-and-bound process, leading to a branch-and-cut algorithm.

4. Computational Experiments

In this section, we conduct computational experiments to demonstrate the effectiveness of our proposed strong valid inequalities in solving application problems, by focusing on solving the multistage stochastic FSUC problems faced by IPPs. We provide a description of the multistage stochastic FSUC problem in Section 4.1, describe the computational setting in Section 4.2, and present the computational results in Sections 4.3 and 4.4. All the instance data used in this section are publicly available in Pan et al. (2022).

4.1. Problem Description

We consider an IPP that owns a natural gas-fired generator. The IPP is facing volatile electricity prices when it devises an optimal natural gas procurement strategy for the next planning horizon. More importantly, it has to evaluate the gas procurement strategy by solving a stochastic FSUC problem, in which the uncertainty comes from the electricity price volatility. Since the planning horizon comprises multiple days and the power generation schedule is arranged on a day-by-day basis, the IPP faces a multistage stochastic FSUC problem.

In this computational study, we consider such a multistage stochastic FSUC problem. We let R denote the number of stages in the planning horizon, where each stage represents one day. For each stage $r = 1, \ldots, R$, there are T time periods (e.g., $T = 24$ hours). Given a gas supply Q_r (i.e., the gas procurement strategy) in each stage r, we use a multistage stochastic scenario tree as shown in Figure 2 to represent electricity price uncertainty. The corresponding daily generation schedule is obtained at each node of the scenario tree. All the constraints in set $\mathcal{D}$ should be respected in each stage of the problem. It follows that our strong valid inequalities derived in Section 3 are also valid for the whole problem and are therefore able to help solve this multistage problem. Note that such a scenario tree has been applied in various practical settings to solve power system problems under uncertainty; see, for example, Takriti et al. (1996), Pan & Guan (2016b), Zou et al. (2019a), and Huang et al. (2021b).

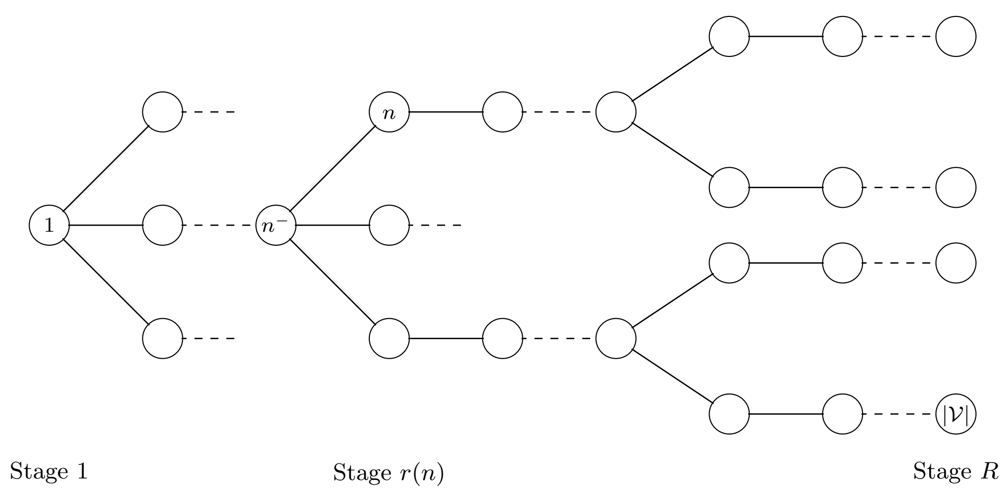

Figure 2 Multistage Stochastic Scenario Tree

Let $\mathcal{A} = (\mathcal{V}, \mathcal{E})$ denote the scenario tree with R stages, where $\mathcal{V}$ is the collection of all scenario nodes. Each node $n \in \mathcal{V}$ in stage r of the tree provides the state of the system that can be distinguished by the information available up to stage r, with the root node denoted by 1. The root node, which is the only node in stage 1, represents the current state of the system. For each node $n \in \mathcal{V}$, we let $r(n)$ denote its stage, n^- denote its unique parent node, and $p_n \in [0,1]$ denote the probability of occurrence of the state corresponding to node n. We let 1^- represent a dummy parent node of root node 1. Note that $r(1)=1$, and $\sum_{n\in\mathcal{V}:r(n)=r'} p_n = 1$ for any stage r'. The decisions corresponding to each node $n \in \mathcal{V}$ are made after observing the realization of the problem parameters along the path from the root node to the current node n, but are nonanticipative with respect to future realizations.

To formulate the multistage stochastic FSUC problem, we use all the notations defined in Section 2 and add a superscript n to each decision variable to indicate the scenario node $n \in \mathcal{V}$. In addition, we assume that the initial conditions are given as $u^{1^-}_{T-\bar{L}+2}, u^{1^-}_{T-\bar{L}+3}, \ldots, u^{1^-}_{T-1}, u^{1^-}_{T}, y^{1^-}_{T-\ell+1}, y^{1^-}_{T-\ell+2}, \ldots, y^{1^-}_{T-1}, y^{1^-}_{T}, x^{1^-}_{T}$. The mathematical formulation for the multistage stochastic FSUC problem is given as follows:

$$
\begin{aligned}
\text{Problem (6):}\quad \min \quad & \sum_{n\in\mathcal{V}} p_n \Bigg[\left[f_1(x^n_1) + c_1 y^n_1 + \overline{U} u^n_1 + \underline{U}(y^{n^-}_T - y^n_1 + u^n_1) - \xi^n_1 x^n_1 \right] \\
& \qquad + \sum_{t=2}^{T} \left[f_t(x^n_t) + c_t y^n_t + \overline{U} u^n_t + \underline{U}(y^n_{t-1} - y^n_t + u^n_t) - \xi^n_t x^n_t \right] \Bigg] && \text{(6a)} \\
\text{s.t.} \quad & \sum_{i=t-L+1}^{t} u^n_i \le y^n_t, \ \forall n \in \mathcal{V}, \ t \in [L+1, T]_{\mathbb{Z}}, && \text{(6b)} \\
& \sum_{i=t-\ell+1}^{t} u^n_i \le 1 - y^n_{t-\ell}, \ \forall n \in \mathcal{V}, \ t \in [\ell+1, T]_{\mathbb{Z}}, && \text{(6c)} \\
& y^n_t - y^n_{t-1} - u^n_t \le 0, \ \forall n \in \mathcal{V}, \ t \in \mathcal{T} \setminus \{1\}, && \text{(6d)} \\
& -x^n_t + \underline{C} y^n_t \le 0, \ \forall n \in \mathcal{V}, \ t \in \mathcal{T}, && \text{(6e)} \\
& x^n_t - \overline{C} y^n_t \le 0, \ \forall n \in \mathcal{V}, \ t \in \mathcal{T}, && \text{(6f)} \\
& x^n_t - x^n_{t-1} \le V y^n_{t-1} + \overline{V}(1 - y^n_{t-1}), \ \forall n \in \mathcal{V}, \ t \in \mathcal{T} \setminus \{1\}, && \text{(6g)} \\
& x^n_{t-1} - x^n_t \le V y^n_t + \overline{V}(1 - y^n_t), \ \forall n \in \mathcal{V}, \ t \in \mathcal{T} \setminus \{1\}, && \text{(6h)} \\
& \sum_{t=1}^{T} x^n_t \le Q_{r(n)}, \ \forall n \in \mathcal{V}, && \text{(6i)} \\
& \sum_{i=T-(L-t)+1}^{T} u^{n^-}_i + \sum_{i=1}^{t} u^n_i \le y^n_t, \ \forall n \in \mathcal{V}, \ t \in [1, L]_{\mathbb{Z}}, && \text{(6j)} \\
& \sum_{i=T-(\ell-t)+1}^{T} u^{n^-}_i + \sum_{i=1}^{t} u^n_i \le 1 - y^{n^-}_{T-(\ell-t)}, \ \forall n \in \mathcal{V}, \ t \in [1, \ell]_{\mathbb{Z}}, && \text{(6k)} \\
& y^n_1 - y^{n^-}_T - u^n_1 \le 0, \ \forall n \in \mathcal{V}, && \text{(6l)} \\
& x^n_1 - x^{n^-}_T \le V y^{n^-}_T + \overline{V}(1 - y^{n^-}_T), \ \forall n \in \mathcal{V}, && \text{(6m)} \\
& x^{n^-}_T - x^n_1 \le V y^n_1 + \overline{V}(1 - y^n_1), \ \forall n \in \mathcal{V}, && \text{(6n)} \\
& y^n_t \in \{0,1\}, \ u^n_t \in \{0,1\}, \ \forall n \in \mathcal{V}, \ t \in \mathcal{T}. && \text{(6o)}
\end{aligned}
$$

In objective function (6a), the per-unit electricity price ξ_t^n is dependent on n, while other objective coefficients are scenario-independent. For each $n \in \mathcal{V}$, constraints (6b)–(6i) repeat those constraints in set $\mathcal{D}$ (i.e., (2a)–(2h)). Constraints (6j)–(6n) are the minimum-up/-down time and ramping rate constraints between a given scenario node $n \in \mathcal{V}$ and its parent node n^-.

Problem (6) is a generalization of the deterministic multistage FSUC problem in which no uncertainty is considered. Note that there are other ways (e.g., stochastic dynamic programming) to formulate the multistage stochastic FSUC problem, and Problem (6) is a deterministic equivalent formulation that can be solved as a deterministic integer program via commercial solvers.

The valid inequalities derived in Section 3 are also valid for Problem (6) and can be used to help solve the problem. However, adding too many inequalities may increase the computational time, because the resulting problem will become very large. Since the size of each of the inequality families (3)–(5) is an exponential function of T, in our computational experiments we use an offline selection process to select a subset of these valid inequalities and add them to the user cut pool of CPLEX. Specifically, for the family of inequality (3), we let $(\pi(1), \dots, \pi(T))$ be a permutation of $(1, \dots, T)$ such that $\xi^*_{\pi(1)} \geq \dots \geq \xi^*_{\pi(T)}$, and we select those inequalities that satisfy $\mathcal{T}_1 = \{\pi(1), \dots, \pi(s)\}$ with $\lambda - \lambda' + 1 \leq s \leq \lambda$. This selection is used because when solving our problem, a high electricity price in a period t will potentially lead to a large generation amount, and the corresponding LP relaxation solution will more likely violate inequality (3). Hence, those inequalities associated with set $\mathcal{T}_1$ that include time periods with high electricity prices should be added to cut off the LP relaxation solutions. For the family of inequality (4), for every $t \in [L+1, T]_{\mathbb{Z}}$ and $k \in [2, \beta_t]_{\mathbb{Z}}$, we select those inequalities that satisfy $S' = [t', t+k-1]_{\mathbb{Z}}$, where $t' = \min\{t+L, t+k\}$, because according to the conditions presented in Remark 4, they are more likely to be facet-defining for $\text{conv}(\mathcal{D})$ and are more efficient. Similarly, for the family of inequality (5), for every $t \in [2, T-L]_{\mathbb{Z}}$ and $k \in [2, \alpha_t]_{\mathbb{Z}}$, we select those inequalities that satisfy $S' = [t-k+1, t']_{\mathbb{Z}}$, where $t' = \max\{t-L, t-k\}$.

4.2. Data Generation

In this subsection, we create data instances based on the modified IEEE 118-bus system available at motor.ece.iit.edu/data/SCUC_118. Four different natural gas-fired generators are selected for the experiments, and we use Gi $(i = 1, \dots, 4)$ to denote them. The physical characteristics of these generators are provided in Table 1. In all test instances, we set $T = 24$. In each period $t \in \mathcal{T}$ of each node $n \in \mathcal{V}$, the convex non-increasing piecewise linear function $f_t(\cdot)$ in (6a) is obtained by approximating the quadratic cost function $c''_t(x_t^n)^2 + c'_t x_t^n$, where the c''_t and c'_t values are shown in Table 1. We apply a method developed by Carrión & Arroyo (2006) to perform the piecewise linear approximation, using eight line segments with the x-coordinates of the breakpoints spread

Table 1 Generator Data

Gen.	L (h)	ℓ (h)	$\underline{C}$ (MW)	$\overline{C}$ (MW)	V (MW/h)	$\overline{V}$ (MW/h)	$\underline{U}$ (\$/h)	$\overline{U}$ (\$/h)	c''_t (\$/MW^{2}h)	c'_t (\$/MWh)	c_t (\$/h)
G1	4	4	10	65	15	17.5	0	600	0.0398	19.7	75
G2	3	3	25	150	32	41	0	800	0.0211	16.5	120
G3	5	5	50	310	70	85	0	1600	0.0031	17.26	192.5
G4	4	4	59	440	51	76	0	300	0.02	22	100

evenly between the lower bound $\underline{C}$ and the upper bound $\overline{C}$ (see formulation (6)–(11) in Carrión & Arroyo 2006).

Generators G1–G3 are used in Section 4.3 to solve relatively small instances of Problem (6), while generator G4 is used in Section 4.4 to solve larger instances. In both subsections, different scenarios of the uncertain electricity price are created for the scenario tree in Figure 2. For simplicity, we assume that the electricity prices are uniformly distributed, since the uniform distribution is often used in the literature for modeling electricity prices and for generating electricity prices in computational studies; see, for example, Ren & Galiana (2004a,b), Pan & Guan (2016b), Melamed et al. (2018), and De Souza et al. (2021). Specifically, we let the electricity price ξ_t^n be uniformly distributed on $[0, 40]$ in each period t of each scenario node n. The fuel supply limit Q_r is uniformly distributed on $[0, T\overline{C}/3]$ in each stage r. In addition, to test the variations of the proposed instances, we consider different numbers of stages and scenarios. For each parameter setting, we create a scenario tree, generate five test instances with different electricity prices and fuel supply limits, and report the average result. All of the computational experiments are conducted on a computer node with Intel(R) Xeon(R) CPU E5-2637 v3 at 3.50GHz and sixteen cores. IBM ILOG CPLEX 12.9 with a single thread is used as the MIP solver and the addressable memory is 8GB.

Next, we describe how the scenario trees are created in our computational study. For ease of exposition, we let $\mathcal{V}(r)$ denote the set of nodes in stage r. Those nodes in stage R do not have children and are called leaf nodes, while those nodes in the other stages have children and are called non-leaf nodes. The path from a leaf node to the root node forms a complete scenario. Thus, $|\mathcal{V}(R)|$ is the total number of scenarios. For each small test instance discussed in Section 4.3, the scenario tree is generated as follows. Given the number of stages R and the number of scenarios $|\mathcal{V}(R)|$, we perform the following steps:

(i) Create a deterministic scenario with one child per non-leaf node.

(ii) Among those nodes with one child, select one node randomly with equal probability, and let r be the stage of the selected node.

(iii) Add a child node to the selected node, add a child node to the newly added child node, and so on, until the newly added child node is in stage R. This leads to a new scenario in the tree.

(iv) Repeat steps (ii) and (iii) until the number of scenarios equals $|\mathcal{V}(R)|$.

(v) If a scenario node has two children, then we assign a conditional probability of 0.5 to each of them. If a scenario node has one child, then we assign a conditional probability of 1 to it. We then determine the probability of occurrence p_n of each node n using these conditional probabilities.

For the larger instances discussed in Section 4.4, we create larger scenario trees by having k children per non-leaf node, where $k = 2, 3, 4$, and assigning a conditional probability of $1/k$ to each non-leaf node. Thus, there are k^{R-1} scenarios in the scenario tree.

4.3. Small Instances

In this subsection, we present the results of the first part of our computational study, in which we consider test instances with relatively small scenario trees and short planning horizons. We consider generators G1–G3, and set $R = 6, 8, 10$ and $|\mathcal{V}(R)| = 16, 27, 32$. Thus, there are 27 parameter settings and $27 \times 5 = 135$ test instances in total. We use CPLEX to solve formulation (6) of these instances, and we compare the computational performance of two approaches. The first approach is "Default CPLEX," where Problem (6) is solved by CPLEX without any of our valid inequalities added. The second approach is "Branch-and-Cut," where Problem (6) is solved by CPLEX with all of our selected valid inequalities added as user cuts at each scenario node $n \in \mathcal{V}$. We set the time limit to one hour per run and use CPLEX's default optimality gap criterion (i.e., 0.01%).

We first report the extent to which our valid inequalities can tighten the relaxation of the feasible region of Problem (6). In Table 2, the "LP (%)" columns report the average LP relaxation gap of the original problem obtained by solving the five test instances using the "Default CPLEX" approach. The "Cut (%)" columns report the average LP relaxation gap of the five test instances

Table 2 Tightness of LP Relaxation

R	$\|\mathcal{V}(R)\|$	G1			G2			G3		
		LP (%)	Cut (%)	Pct (%)	LP (%)	Cut (%)	Pct (%)	LP (%)	Cut (%)	Pct (%)
6	16	16.0	10.3	35.1	6.2	3.6	43.9	3.5	2.3	32.6
	24	15.1	8.2	43.6	9.7	5.8	41.1	8.7	4.4	50.7
	32	9.8	5.7	42.1	5.7	3.0	47.8	4.3	2.3	46.8
8	16	14.2	9.3	32.8	5.4	3.3	39.4	8.8	5.8	33.2
	24	17.3	10.1	39.8	6.7	3.9	39.9	6.2	3.8	40.3
	32	12.4	7.7	37.7	4.2	2.3	46.2	5.0	2.5	49.6
10	16	14.1	7.8	43.7	4.3	2.7	37.0	7.6	4.1	47.0
	24	14.6	8.2	44.2	9.7	5.4	45.6	9.8	5.2	46.8
	32	10.8	5.9	44.8	7.4	3.9	47.5	9.4	4.9	46.6

Note: Each row represents the average result of five instances.

after adding our strong valid inequalities (i.e., using the “Branch-and-Cut” approach). Here, the LP relaxation gap of the original problem is defined as $(Z_{\text{LP}} - Z_{\text{MILP}})/Z_{\text{MILP}} \times 100\%$, and the LP relaxation gap using the “Branch-and-Cut” approach is defined as $(Z_{\text{LP}}^{\text{Cut}} - Z_{\text{MILP}})/Z_{\text{MILP}} \times 100\%$, where Z_{LP} is the objective value of the original LP relaxation without adding our valid inequalities, $Z_{\text{LP}}^{\text{Cut}}$ is the objective value of the LP relaxation after adding our valid inequalities, and Z_{MILP} is the objective value of Problem (6) with the best integer solution obtained by CPLEX using either approach. The average percentage reduction is reported in the “Pct (%)” columns, where the percentage reduction is given by

$$\frac{(\text{the “LP (\%)” value}) - (\text{the “Cut (\%)” value})}{\text{the “LP (\%)” value}} \times 100\%.$$

From this table, we observe that our valid inequalities can help reduce the average LP relaxation gap by around 40% in all 27 parameter settings.

Next, we report the extent to which our branch-and-cut approach can speed up the solution process. In Table 3, the “TGap (%)” columns report the average terminating gap of the five test instances when CPLEX reaches the terminating criterion. The “# Nodes” columns report the average number of branch-and-bound nodes explored by CPLEX among the five test instances, and the “Time” columns report the average computational time in seconds. The “# Cuts” column reports the average number of user cuts used by CPLEX. The numbers in square brackets in the “TGap (%)” columns indicate the number of instances (out of five) that are not solved to optimality within the one hour time limit. From this table, we observe that our “Branch-and-Cut” approach reduces the computational time significantly and solves many more instances to optimality within the time limit than the “Default CPLEX” does. In addition, the reduced terminating gaps and the reduced number of branch-and-bound nodes indicate that our valid inequalities can help tighten the LP relaxation solved at each branch-and-bound node and thus help reduce the searching space.

4.4. Large Instances

In this subsection, we present the results of the second part of our computational study, in which we consider test instances with relatively large scenario trees and long planning horizons. We use generator G4, and set $R = 21, 26, 31$. As mentioned in Section 4.2, in this part we make use of larger scenario trees with k children per non-leaf node, where $k = 2, 3, 4$. The deterministic equivalent formulation for these large instances is extremely difficult to solve, so we adopt stochastic dynamic programming approaches. Specifically, we assume stage-wise independence within the multistage stochastic FSUC problem and adopt the stochastic dual dynamic integer programming (SDDiP) approach developed by Zou et al. (2019b) accordingly.

Table 3 Results of the Branch-and-Cut Scheme (Small Instances)

R	Gen.	$\lvert\mathcal{V}(R)\rvert$	Default CPLEX			Branch-and-Cut			
			TGap (%)	# Nodes	Time	TGap (%)	# Nodes	Time	# Cuts
6	G1	16	0.10 [1]	100295.8	1006.8	0.06 [1]	82306.4	918.9	140.9
		24	0.15 [3]	54986.3	2748.8	0.08 [1]	38598.6	1286.9	162.8
		32	0.01 [1]	100105.1	1537.1	0.01	49803.7	725.5	151.3
	G2	16	0.01	162143.2	1183.3	0.01	12423.2	82.0	102.5
		24	0.05 [2]	45023.7	2608.1	0.01	14261.2	960.1	150.1
		32	0.01	52097.4	335.2	0.01	13352.0	86.4	157.3
	G3	16	0.01	112577.0	1386.2	0.01	28269.6	302.6	166.3
		24	0.21 [4]	53994.2	1808.5	0.03 [1]	25421.8	976.1	263.6
		32	0.04 [3]	166577.0	2180.1	0.01	8644.9	91.2	311.2
8	G1	16	0.02 [1]	87035.4	1084.0	0.01	64898.6	717.5	147.6
		24	0.16 [4]	143292.9	3001.8	0.05 [1]	19731.4	2301.6	279.4
		32	0.08 [2]	94422.5	1814.5	0.02 [1]	90849.5	912.4	200.8
	G2	16	0.01 [1]	86558.8	1046.1	0.01	2253.8	37.2	132.9
		24	0.12 [3]	61974.1	2376.5	0.04 [1]	56473.7	1574.7	282.4
		32	0.01	29515.0	578.4	0.01	10579.9	192.4	164.8
	G3	16	0.04 [2]	140908.5	1827.9	0.01	68435.8	812.9	213.5
		24	0.27 [5]	65381.4	2217.5	0.03 [1]	69368.1	1163.2	262.7
		32	0.06 [2]	95113.1	2447.1	0.02 [1]	86772.9	1609.8	479.2
10	G1	16	0.02 [4]	171001.8	2899.6	0.01	47802.6	717.1	198.5
		24	0.31 [5]	182754.3	3600.0	0.14 [4]	149000.1	3119.3	197.8
		32	0.13 [5]	98715.3	3600.0	0.08 [3]	92987.7	2259.2	291.3
	G2	16	1.29 [3]	117048.8	2295.6	0.01	93606.5	764.5	258.4
		24	1.65 [5]	78232.2	3600.0	0.95 [3]	82954.7	2547.3	480.7
		32	1.46 [5]	65809.3	3600.0	0.95 [5]	66960.7	3600.0	1525.1
	G3	16	0.40 [5]	182251.7	3600.0	0.05 [1]	36316.7	729.1	296.2
		24	1.15 [5]	75155.5	3600.0	0.35 [3]	57176.7	2163.6	254.2
		32	0.87 [5]	76780.6	3600.0	0.42 [5]	64414.7	3600.0	474.1

Note: Each row represents the average result of five instances. The number in square brackets indicates the number of instances that are not solved to optimality.

Given a multistage stochastic integer program formulated over a scenario tree, the SDDiP algorithm solves it by iteratively solving subproblems until a certain convergence criterion is met. In each iteration, there are three steps: a sampling step, a forward step, and a backward step. In the sampling step, a subset of all the scenarios is randomly sampled based on a certain probability distribution (uniform distribution is used in our experiments). Here, a scenario represents a path from the root node to a leaf node in the scenario tree. Thus, after this step, we have a number of sampled paths (i.e., sampled scenarios). In the forward step, for each sampled scenario the algorithm proceeds stage-wise from stage 1 to stage R by solving a dynamic programming recursion with an approximated expected cost-to-go function at each node of the sampled scenario. Such a dynamic program recursion is solved as an MIP, where the approximated expected cost-to-go function (i.e., function $\psi_n^i(\cdot)$ in problem (3.1a)–(3.1d) in Zou et al. (2019b)) is represented by a decision variable (i.e., variable θ_n in (3.2a)–(3.2b) in Zou et al. (2019b)) together with a set of linear constraints (i.e., constraints (3.2a)–(3.2b) in Zou et al. (2019b)). After completing the forward step for all the sampled scenarios, we have a statistical upper bound of the overall problem. In the backward step, for each sampled scenario the algorithm proceeds stage-wise from stage R to stage 1. For each scenario node at stage r ($r = 1, \ldots, R$) of a sampled scenario, the algorithm first solves a relaxation of the MIP (using LP relaxation and Lagrangian relaxation) at this node as well as its sibling nodes, which share a parent node with this node. It then derives three types of cuts, namely strengthened Benders' cuts, integer optimality cuts, and Lagrangian cuts, and then obtains an updated approximated expected cost-to-go function (with the derived cuts added as linear constraints) for the shared parent node. Note that the forward problem solved at the root node of the scenario tree provides a lower bound of the entire problem. The algorithm terminates when this lower bound is sufficiently close to the statistical upper bound.

Note that our valid inequalities presented in Section 3 are valid for the subproblem solved at each stage using the SDDiP approach. In this part of our computational study, we solve Problem (6) using two approaches and compare their computational performance. The first approach is "Default SDDiP," where Problem (6) is solved by the SDDiP algorithm without any of our valid inequalities added. The second approach is "Branch-and-Cut in SDDiP," where Problem (6) is solved by the SDDiP algorithm with our strong valid inequalities added as user cuts at each scenario node. When solving each test instance with these two approaches, we try a different number of sampled paths in the forward step. For example, if the test instance has 21 days and 3 children at each non-leaf node, then the total number of sampled paths is 3^{20}, and the sampling step of the SDDiP algorithm will select a number of sampled paths among them. In our experiments, we set the numbers of sampled paths to 5, 10, and 15.

When implementing the SDDiP algorithm, we use all the cuts derived in Zou et al. (2019b) (i.e., strengthened Benders' cuts, integer optimality cuts, and Lagrangian cuts) in the backward step. We set the minimum and maximum numbers of iterations to 50 and 100, respectively. The MIP is solved by CPLEX with the optimality gap set to 0.5%, and the Lagrangian relaxation problem is solved by a basic subgradient algorithm with an optimality tolerance of 0.5%. The time limit for solving each single MIP is set to 1800 seconds, while the time limit for each complete run is set to 18000 seconds. In addition, we follow the SDDiP enhancements designed by Zou et al. (2019a) to use backward parallelization to improve computational performance.

Table 4 Results of the Branch-and-Cut Scheme (Large Instances)

R	# Ch	# SP	Default SDDiP LB	Default SDDiP Stat UB	Default SDDiP Gap (%)	Default SDDiP Time	Default SDDiP # Ite	Branch-and-Cut in SDDiP LB	Branch-and-Cut in SDDiP Stat UB	Branch-and-Cut in SDDiP Gap (%)	Branch-and-Cut in SDDiP Time	Branch-and-Cut in SDDiP # Ite
21	2	5	-295899	-284801	3.96	746.8	73.0	-290205	-280415	3.57	274.9	78.3
		10	-312873	-305559	2.44	871.4	74.7	-307357	-304262	0.99	288.1	72.3
		15	-311051	-306480	1.53	575.6	71.7	-307009	-303003	1.38	429.3	74.0
	3	5	-297723	-291163	2.20	1540.8	82.3	-294800	-290668	1.43	476.3	80.0
		10	-301063	-294263	2.41	1674.5	80.7	-298264	-294127	1.45	415.3	70.3
		15	-312922	-304690	2.76	1823.3	71.7	-309218	-302689	2.19	404.3	71.7
	4	5	-307659	-299390	2.79	1379.8	83.0	-302267	-291793	3.67	390.2	87.0
		10	-306937	-299284	2.50	2653.0	96.0	-301390	-294396	2.38	653.4	77.3
		15	-308048	-303431	1.48	2935.6	85.7	-304402	-299573	1.55	924.0	90.3
26	2	5	-455197	-442475	2.92	1568.5	72.7	-447031	-438758	1.90	333.8	76.7
		10	-448403	-443363	1.19	1297.2	71.7	-440003	-436162	0.92	332.5	72.7
		15	-434416	-428614	1.39	1077.6	65.7	-426238	-421751	1.07	441.3	69.0
	3	5	-454520	-438995	3.44	1840.0	70.3	-445641	-433224	2.96	600.3	80.0
		10	-445226	-435651	2.23	2349.5	81.0	-438191	-427613	2.45	686.0	70.7
		15	-444461	-437280	1.59	2446.4	80.0	-436806	-430234	1.60	872.8	78.3
	4	5	-460480	-448386	2.56	3677.1	80.3	-453204	-446734	1.46	766.1	79.0
		10	-458493	-450062	1.81	6472.4	87.7	-450240	-437432	2.80	681.1	76.3
		15	-427557	-422748	1.15	2866.6	77.0	-420202	-414637	1.35	816.8	76.7
31	2	5	-557917	-549861	1.49	2523.0	83.0	-547894	-541660	1.13	848.4	73.7
		10	-607778	-602622	0.83	1170.9	73.0	-597588	-592653	0.83	614.8	78.7
		15	-569424	-562669	1.26	1631.1	72.0	-558540	-552801	1.06	627.6	71.7
	3	5	-558473	-548619	1.73	3014.0	86.7	-549132	-542066	1.34	662.6	75.7
		10	-577292	-569453	1.41	1862.0	83.0	-565894	-556384	1.64	956.1	78.7
		15	-586400	-576273	1.68	6071.6	87.7	-576629	-568623	1.39	1900.2	69.3
	4	5	-570628	-555320	2.85	5625.8	86.0	-563840	-552450	2.08	1217.2	83.7
		10	-617597	-607393	1.75	4474.7	83.3	-607451	-597057	1.69	1909.7	87.3
		15	-565101	-558802	1.16	6555.2	96.0	-556050	-550471	1.01	2906.8	77.3

Note: Each row represents the average result of five instances.

The computational results are presented in Table 4, with each row summarizing the results of five test instances. The "# Ch" column indicates the number of children of each non-leaf node in the scenario tree. The "# SP" column indicates the number of different sampled paths selected by the sampling step. The "LB" columns report the average value of the lower bounds obtained for the five test instances, where each test instance's lower bound is obtained by selecting the best lower bound value generated by different iterations of the SDDiP algorithm. As mentioned by Zou et al. (2019b), when a forward step of the SDDiP algorithm is completed, the objective values corresponding to all sampling scenarios are obtained. Based on these objective values, a 95%-confidence range for the objective values is constructed by assuming that they are normally distributed. The right-end point of this range corresponding to the forward step in the last iteration is the statistical upper bound obtained for the test instance. The “Stat UB” columns report the average value of the statistical upper bounds obtained for the five test instances. The “Gap (%)” columns report the average SDDiP gap. The SDDiP gap is the relative gap between “LB” and “Stat UB,” and is given by

$$\frac{\text{(the “Stat UB” value)} - \text{(the “LB” value)}}{|\text{the “Stat UB” value}|} \times 100\%.$$

The “Time” columns report the average computational time of the SDDiP algorithm per instance in seconds, and the “# Ite” columns report the average number of iterations per instance. From this table, we observe that the computational time is significantly reduced by adding our valid inequalities to the SDDiP algorithm. Our branch-and-cut approach also helps reduce the SDDiP gap in most of the cases.

5. Conclusions

Natural gas-fired generators are increasingly entering electricity markets due to their environmental friendliness, high flexibility, and affordable running costs. However, natural gas is mainly used in residential, commercial, and industrial sectors rather than for electricity generation, thus limiting supplies for IPPs. The limited fuel supply becomes an increasing challenge for IPPs that own gas-fired generators, because production decisions in different time periods in the operation horizon are linked. More important, such a linking constraint occurs frequently for other types of generators in different ways in terms of coal supply limit, carbon emission limit, and pollutant limit. In this paper, the challenge of the fuel constraint is addressed by a comprehensive polyhedral study in which several families of strong valid inequalities are derived. Extensive computational experiments have been performed to demonstrate the effectiveness of our proposed inequalities in solving practical problems (i.e., multistage stochastic FSUC problems) in various settings. Meanwhile, our model provides a scenario analysis tool to efficiently explore a good fuel procurement strategy. In our model, the fuel supply limit Q is an input parameter, and the IPP can vary the value of Q to perform sensitivity analyses. For each value of Q, the IPP can apply our derived inequalities to efficiently solve Problem (1) and obtain the corresponding optimal profit. It can then devise a good fuel procurement strategy by choosing an appropriate value of Q.

This research can be extended in various directions. First, our semi-continuous inequalities are constructed by including constraint (2h) and the upper- and lower-limit requirements of the x

variables in the semi-continuous knapsack set $\mathcal{D}_{sc}$. It would be interesting to consider other constraints in the set $\mathcal{D}$ and include some of them in $\mathcal{D}_{sc}$. This leads to a new convex hull $\text{conv}(\mathcal{D}_{\text{sc}})$, which may provide us with stronger valid inequalities. Second, our look forward and look backward inequalities involve only one x variable. It would be interesting to derive strong valid inequalities that contain more than one x variable. Third, with significant gas supply and electricity price uncertainties, it would be an interesting problem to consider fuel procurement and unit commitment together. In our computational experiments, we have assumed that the fuel procurement strategy is given, and then solved the multistage stochastic FSUC problems to efficiently evaluate this strategy. Integrating fuel procurement and unit commitment decisions will lead to an even more challenging problem. Fourth, it would be appealing to consider the fuel supply limit enforced over multiple generators. When multiple gas-fired generators share the same source of fuel supply or share a single carbon emission limit, the fuel supply limit may jointly constrain the schedule of all the generators.

Acknowledgments

The authors thank the Area Editor, Associate Editor, and two anonymous referees for their valuable comments and suggestions. Kai Pan was supported in part by the Research Grants Council of Hong Kong [Grant 15501920]; Feng Qiu was supported in part by the U.S. Department of Energy Advanced Grid Modeling Program under Grant DE-OE0000875.

References

Al-Kalaani, Y., Villaseca, F. E., & Renovich, F. (1996). Storage and delivery constrained unit commitment. *IEEE Transactions on Power Systems*, 11(2), 1059–1066.

Amjady, N. & Nasiri-Rad, H. (2011). Security constrained unit commitment by a new adaptive hybrid stochastic search technique. *Energy Conversion and Management*, 52(2), 1097–1106.

Aoki, K., Itoh, M., Satoh, T., Nara, K., & Kanezashi, M. (1989). Optimal long-term unit commitment in large scale systems including fuel constrained thermal and pumped-storage hydro. *IEEE Transactions on Power Systems*, 4(3), 1065–1073.

Aoki, K., Satoh, T., Itoh, M., Ichimori, T., & Masegi, K. (1987). Unit commitment in a large-scale power system including fuel constrained thermal and pumped-storage hydro. *IEEE Transactions on Power Systems*, 2(4), 1077–1084.

Bai, X. & Shahidehpour, S. M. (1997). Extended neighborhood search algorithm for constrained unit commitment. *International Journal of Electrical Power & Energy Systems*, 19(5), 349–356.

Baldick, R. (1995). The generalized unit commitment problem. *IEEE Transactions on Power Systems*, 10(1), 465–475.

Beale, E. M. L. (1985). The evolution of mathematical programming systems. *Journal of the Operational Research Society*, 36(5), 357–366.

Bixby, R. E. (2010). *Mixed-integer programming: It works better than you may think.* Accessed June 23, 2022, https://cms.ferc.gov/sites/default/files/2020-05/20100609110044-Bixby%2C%2520Gurobi%2520Optimization_0.pdf.

Carlson, B., Chen, Y., Hong, M., Jones, R., Larson, K., Ma, X., Nieuwesteeg, P., Song, H., Sperry, K., Tackett, M., Taylor, D., Wan, J., & Zak, E. (2012). MISO unlocks billions in savings through the application of operations research for energy and ancillary services markets. *Interfaces*, 42(1), 58–73.

Carrión, M. & Arroyo, J. M. (2006). A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. *IEEE Transactions on Power Systems*, 21(3), 1371–1378.

Chen, S., Conejo, A. J., Sioshansi, R., & Wei, Z. (2019). Unit commitment with an enhanced natural gas-flow model. *IEEE Transactions on Power Systems*, 34(5), 3729–3738.

Cohen, A. I. & Wan, S. H. (1987). A method for solving the fuel constrained unit commitment problem. *IEEE Transactions on Power Systems*, 2(3), 608–614.

Damcı-Kurt, P., Küçükyavuz, S., Rajan, D., & Atamtürk, A. (2016). A polyhedral study of production ramping. *Mathematical Programming*, 158(1–2), 175–205.

De Farias, I. R. & Zhao, M. (2013). A polyhedral study of the semi-continuous knapsack problem. *Mathematical Programming*, 142(1–2), 169–203.

De Souza, M. F. Z., Cañizares, C. A., & Bhattacharya, K. (2021). Self-scheduling models of a CAES facility under uncertainties. *IEEE Transactions on Power Systems*, 36(4), 3607–3617.

El-Keib, A. A., Ma, H., & Hart, J. L. (1994). Economic dispatch in view of the Clean Air Act of 1990. *IEEE Transactions on Power Systems*, 9(2), 972–978.

Frangioni, A. & Gentile, C. (2006). Solving nonlinear single-unit commitment problems with ramping constraints. *Operations Research*, 54(4), 767–775.

Frangioni, A., Gentile, C., & Lacalandra, F. (2008). Tighter approximated MILP formulations for unit commitment problems. *IEEE Transactions on Power Systems*, 24(1), 105–113.

Fu, Y., Shahidehpour, M., & Li, Z. (2005). Security-constrained unit commitment with AC constraints. *IEEE Transactions on Power Systems*, 20(2), 1001–1013.

Garey, M. R. & Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness.* WH Freeman & Co., New York.

Gentile, C., Morales-España, G., & Ramos, A. (2017). A tight MIP formulation of the unit commitment problem with start-up and shut-down constraints. *EURO Journal on Computational Optimization*, 5(1–2), 177–201.

Guan, Y., Pan, K., & Zhou, K. (2018). Polynomial time algorithms and extended formulations for unit commitment problems. *IISE Transactions*, 50(8), 735–751.

Huang, J., Pan, K., & Guan, Y. (2021a). Cutting planes for security-constrained unit commitment with regulation reserve. *IISE Transactions*, 53(4), 437–452.

Huang, J., Pan, K., & Guan, Y. (2021b). Multistage stochastic power generation scheduling co-optimizing energy and ancillary services. *INFORMS Journal on Computing*, 33(1), 352–369.

Knueven, B., Ostrowski, J., & Watson, J.-P. (2020). On mixed-integer programming formulations for the unit commitment problem. *INFORMS Journal on Computing*, 32(4), 857–876.

Kockar, I., Conejo, A. J., & McDonald, J. R. (2009). Influence of the emissions trading scheme on generation scheduling. *International Journal of Electrical Power & Energy Systems*, 31(9), 465–473.

Kuloor, S., Hope, G. S., & Malik, O. P. (1992). Environmentally constrained unit commitment. *IEE Proceedings C (Generation, Transmission and Distribution)*, 139(2), 122–128.

Lee, F. N. (1991). The coordination of multiple constrained fuels. *IEEE Transactions on Power Systems*, 6(2), 699–707.

Lee, J., Leung, J., & Margot, F. (2004). Min-up/min-down polytopes. *Discrete Optimization*, 1(1), 77–85.

Li, C., Zhang, M., & Hedman, K. (2021). Extreme ray feasibility cuts for unit commitment with uncertainty. *INFORMS Journal on Computing*, 33(3), 1037–1055.

Li, C.-A., Johnson, R. B., & Svoboda, A. J. (1997). A new unit commitment method. *IEEE Transactions on Power Systems*, 12(1), 113–119.

Lu, B. & Shahidehpour, M. (2005). Unit commitment with flexible generating units. *IEEE Transactions on Power Systems*, 20(2), 1022–1034.

Ma, H. & Shahidehpour, S. M. (1999). Unit commitment with transmission security and voltage constraints. *IEEE Transactions on Power Systems*, 14(2), 757–764.

Melamed, M., Ben-Tal, A., & Golany, B. (2018). A multi-period unit commitment problem under a new hybrid uncertainty set for a renewable energy source. *Renewable Energy*, 118, 909–917.

Morales-España, G., Latorre, J. M., & Ramos, A. (2013). Tight and compact MILP formulation for the thermal unit commitment problem. *IEEE Transactions on Power Systems*, 28(4), 4897–4908.

Nemhauser, G. L. & Wolsey, L. A. (1988). *Integer and Combinatorial Optimization*. John Wiley & Sons, New York.

NERC (2014). *Polar vortex review*. Accessed June 23, 2022, https://www.nerc.com/pa/rrm/January%202014%20Polar%20Vortex%20Review/Polar_Vortex_Review_29_Sept_2014_Final.pdf.

Ostrowski, J., Anjos, M. F., & Vannelli, A. (2012). Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Transactions on Power Systems*, 27(1), 39–46.

Padhy, N. P. (2004). Unit commitment—A bibliographical survey. *IEEE Transactions on Power Systems*, 19(2), 1196–1205.

Pan, K. & Guan, Y. (2016a). A polyhedral study of the integrated minimum-up/-down time and ramping polytope. *arXiv preprint arXiv:1604.02184*.

Pan, K. & Guan, Y. (2016b). Strong formulations for multistage stochastic self-scheduling unit commitment. *Operations Research*, 64(6), 1482–1498.

Pan, K., Guan, Y., Watson, J.-P., & Wang, J. (2016). Strengthened MILP formulation for certain gas turbine unit commitment problems. *IEEE Transactions on Power Systems*, 31(2), 1440–1448.

Pan, K., Zhao, M., Li, C.-L., & Qiu, F. (2022). A polyhedral study on fuel-constrained unit commitment. *INFORMS Journal on Computing*. https://doi.org/10.5281/zenodo.6896199. available for download at https://github.com/INFORMSJoC/2021.0219

Pulgar-Painemal, H. A. (2005). Short-term generation scheduling under a SO_2 emissions allowances market. *Electric Power Systems Research*, 74(2), 257–265.

Queyranne, M. & Wolsey, L. A. (2017). Tight MIP formulations for bounded up/down times and interval-dependent start-ups. *Mathematical Programming*, 164(1–2), 129–155.

Rajan, D. & Takriti, S. (2005). *Minimum up/down polytopes of the unit commitment problem with start-up costs*. Technical Report, *IBM*.

Ren, Y. & Galiana, F. D. (2004a). Pay-as-bid versus marginal pricing—Part I: Strategic generator offers. *IEEE Transactions on Power Systems*, 19(4), 1771–1776.

Ren, Y. & Galiana, F. D. (2004b). Pay-as-bid versus marginal pricing—Part II: Market behavior under strategic generator offers. *IEEE Transactions on Power Systems*, 19(4), 1777–1783.

Ružić, S. & Rajaković, N. (1991). A new approach for solving extended unit commitment problem. *IEEE Transactions on Power Systems*, 6(1), 269–277.

Saneifard, S., Prasad, N. R., & Smolleck, H. A. (1997). A fuzzy logic approach to unit commitment. *IEEE Transactions on Power Systems*, 12(2), 988–995.

Saravanan, B., Das, S., Sikri, S., & Kothari, D. P. (2013). A solution to the unit commitment problem—A review. *Frontiers in Energy*, 7(2), 223–236.

Shahidehpour, M., Fu, Y., & Wiedman, T. (2005). Impact of natural gas infrastructure on electric power systems. *Proceedings of the IEEE*, 93(5), 1042–1056.

Shaw, J. J. (1995). A direct method for security-constrained unit commitment. *IEEE Transactions on Power Systems*, 10(3), 1329–1342.

Takriti, S., Birge, J. R., & Long, E. (1996). A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(3), 1497–1508.

Takriti, S., Krasenbrink, B., & Wu, L. S.-Y. (2000). Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Research*, 48(2), 268–280.

US EIA (2017). *Natural gas generators make up the largest share of overall U.S. generation capacity (December 18)*. Accessed June 23, 2022, https://www.eia.gov/todayinenergy/detail.php?id=34172.

US EIA (2021a). *Electricity*. Accessed June 23, 2022, https://www.eia.gov/electricity/data.php#gencapacity.

US EIA (2021b). *Natural gas consumption by end use*. Accessed June 23, 2022, https://www.eia.gov/dnav/ng/ng_cons_sum_dcu_nus_a.htm.

Vemuri, S. & Lemonidis, L. (1992). Fuel constrained unit commitment. *IEEE Transactions on Power Systems*, 7(1), 410–415.

Wu, L., Shahidehpour, M., & Li, T. (2007). Stochastic security-constrained unit commitment. *IEEE Transactions on Power Systems*, 22(2), 800–811.

Wu, L., Shahidehpour, M., & Li, T. (2008). Cost of reliability analysis based on stochastic unit commitment. *IEEE Transactions on Power Systems*, 23(3), 1364–1374.

Xavier, Á. S., Qiu, F., & Ahmed, S. (2021). Learning to solve large-scale security-constrained unit commitment problems. *INFORMS Journal on Computing*, 33(2), 739–756.

Zhao, B., Conejo, A. J., & Sioshansi, R. (2017). Unit commitment under gas-supply uncertainty and gas-price variability. *IEEE Transactions on Power Systems*, 32(3), 2394–2405.

Zou, J., Ahmed, S., & Sun, X. A. (2019a). Multistage stochastic unit commitment using stochastic dual dynamic integer programming. *IEEE Transactions on Power Systems*, 34(3), 1814–1823.

Zou, J., Ahmed, S., & Sun, X. A. (2019b). Stochastic dual dynamic integer programming. *Mathematical Programming*, 175(1–2), 461–502.

Online Appendix

Appendix A: Mathematical Proofs

A.1. Proof of Proposition 1

We transform the Equal-size Partition problem, a variant of the Partition problem, to Problem (1). Given a set $A = \{1, \ldots, 2n\}$ and a positive integer size $s(i)$ for each $i \in A$, the Equal-size Partition problem asks if there is a subset $A' \subseteq A$ such that $|A'| = |A|/2$ and $\sum_{i\in A'} s(i) = \sum_{i\in A\setminus A'} s(i)$. The Equal-size Partition problem is known to be NP-hard (Garey & Johnson 1979, p. 223).

Consider an arbitrary given instance of Equal-size Partition with $n \geq 2$. We let $B = (1/2)\sum_{i\in A} s(i)$ and $R = (2n-3)B + n - 1$, and construct a corresponding instance of Problem (1) as follows: $T = 4n$; $Q = nR + B$; $\overline{C} = \overline{V} = V = R + 2B$; $\underline{C} = R$; $L = \ell = 1$; $\overline{U} = \underline{U} = 0$; and $x_0 = y_0 = 0$. For $t = 1, \ldots, 4n$,

$$f_t(x_t) = \begin{cases} \max\big\{0,\, Q\big(x_t - R - s(t/2)\big)\big\}, & \text{if } t \text{ is even;} \\ 0, & \text{if } t \text{ is odd;} \end{cases}$$

$$c_t = \begin{cases} R + s(t/2), & \text{if } t \text{ is even;} \\ Q + 1, & \text{if } t \text{ is odd;} \end{cases}$$

and

$$\xi_t = \begin{cases} 2, & \text{if } t \text{ is even;} \\ 0, & \text{if } t \text{ is odd.} \end{cases}$$

Clearly, this construction can be done in polynomial time. Note that $f_t(\cdot)$ is a non-decreasing convex piecewise linear function for each $t = 1, \ldots, 4n$. We will show that there exists a feasible solution of this constructed instance of Problem (1) with a total cost no greater than $-Q$ if and only if there exists a solution to the given instance of Equal-size Partition.

Suppose that in the given instance of Equal-size Partition, there is a subset $A' \subseteq A$ such that $|A'| = |A|/2$ and $\sum_{i\in A'} s(i) = \sum_{i\in A\setminus A'} s(i)$. Then, consider the following solution of the constructed instance of Problem (1): For $t = 1, \ldots, 4n$, if t is even and $t/2 \in A'$, then $x_t = R + s(t/2)$ and $y_t = u_t = 1$, otherwise let $x_t = y_t = u_t = 0$. It is easy to check that this solution satisfies constraints (1b)–(1h) and (1j). This solution also satisfies constraint (1i), because $\sum_{t=1}^{T} x_t = \sum_{i=1}^{2n} x_{2i} = \sum_{i\in A'} (R + s(i)) = |A'| \cdot R + \sum_{i\in A'} s(i) = nR + B = Q$. Hence, this solution is feasible for Problem (1). The total cost of this solution is

$$\begin{aligned}
&\sum_{t=1}^{T} \Big[f_t(x_t) + c_t y_t + \overline{U} u_t + \underline{U}(y_{t-1} - y_t + u_t) - \xi_t x_t\Big] \\
&= \sum_{i=1}^{2n} \Big[f_{2i}(x_{2i}) + c_{2i} y_{2i} + \overline{U} u_{2i} + \underline{U}(y_{2i-1} - y_{2i} + u_{2i}) - \xi_{2i} x_{2i}\Big] \\
&= \sum_{i\in A'} \Big[f_{2i}(R + s(i)) + c_{2i}\cdot 1 + \overline{U}\cdot 1 + \underline{U}(0 - 1 + 1) - \xi_{2i}(R + s(i))\Big] \\
&= \sum_{i\in A'} \Big[0 + (R + s(i))(1) + (0)(1) + (0)(0) - (2)(R + s(i))\Big] \\
&= -\sum_{i\in A'} (R + s(i))
\end{aligned}$$

$$
\begin{aligned}
&= -|A'| \cdot R - \sum_{i \in A'} s(i) \\
&= -nR - B \\
&= -Q.
\end{aligned}
$$

Conversely, suppose that the constructed instance of Problem (1) has a feasible solution $(x_1^*, \ldots, x_T^*; y_1^*, \ldots, y_T^*; u_1^*, \ldots, u_T^*)$ with a total cost no greater than $-Q$. This feasible solution has the following properties:

(i) The generator is offline in periods $1, 3, \ldots, 4n-1$. This property holds, since otherwise a fixed cost of $Q+1$ would be incurred in some odd period, implying that the total cost of this solution will be at least $(Q+1) - \sum_{t=1}^T \xi_t x_t^* \geq (Q+1) - 2Q > -Q$, which is a contradiction. Note that this property implies that for $i = 1, \ldots, 2n$, if the generator is online in period $2i$, then it starts up in period $2i$.

(ii) For any $i = 1, \ldots, 2n$, if the generator is online in period $2i$, then the cost incurred in period $2i$ is at least $-x_{2i}^*$, and this cost equals $-x_{2i}^*$ if and only if $x_{2i}^* = R + s(i)$. To show the validity of this property, we consider three different cases. If $x_{2i}^* = R + s(i)$, then the cost incurred in period $2i$ is $f_{2i}(x_{2i}^*) + c_{2i} y_{2i}^* + \overline{U} u_{2i}^* + \underline{U}(y_{2i-1}^* - y_{2i}^* + u_{2i}^*) - \xi_{2i} x_{2i}^* = 0 + (R + s(i)) + 0 + 0 - 2(R + s(i)) = -x_{2i}^*$. If $0 < x_{2i}^* < R + s(i)$, then the cost incurred in period $2i$ is $f_{2i}(x_{2i}^*) + c_{2i} y_{2i}^* + \overline{U} u_{2i}^* + \underline{U}(y_{2i-1}^* - y_{2i}^* + u_{2i}^*) - \xi_{2i} x_{2i}^* = 0 + (R + s(i)) + 0 + 0 - 2x_{2i}^* > -x_{2i}^*$. If $x_{2i}^* > R + s(i)$, then the cost incurred in period $2i$ is $f_{2i}(x_{2i}^*) + c_{2i} y_{2i}^* + \overline{U} u_{2i}^* + \underline{U}(y_{2i-1}^* - y_{2i}^* + u_{2i}^*) - \xi_{2i} x_{2i}^* = Q(x_{2i}^* - R - s(i)) + (R + s(i)) + 0 + 0 - 2x_{2i}^* = (Q-1)(x_{2i}^* - R - s(i)) - x_{2i}^* > -x_{2i}^*$.

(iii) For any $t = 1, \ldots, T$, the cost incurred in period t is at least $-x_t^*$. Clearly, this property is valid if the generator is offline in period t. If the generator is online in period t, then by property (i), t must be even, and the validity of this property follows from property (ii).

(iv) The total generation amount is Q (i.e., $\sum_{t=1}^T x_t^* = Q$). This property holds because by property (iii), the total cost of this solution is least $-\sum_{t=1}^T x_t^*$. Thus, if $\sum_{t=1}^T x_t^* < Q$, then the total cost of the solution would be greater than $-Q$.

(v) The generator is online for at most n even periods. This property holds, since otherwise $\sum_{t=1}^T x_t^* \geq (n+1)\underline{C} = (n+1)R > nR + B = Q$, which violates constraint (1i).

(vi) The generator is online for at least n even periods. To show the validity of this property, suppose the opposite—that the generator is online for no more than $n-1$ even periods. Then, by properties (i) and (iv), there exists an even period t' in which the generation amount is at least $Q/(n-1)$. The total cost of this solution is at least $f_{t'}(Q/(n-1)) - \sum_{t=1}^T \xi_t x_t^* \geq f_{t'}(Q/(n-1)) - 2Q = f_{t'}((nR+B)/(n-1)) - 2Q = f_{t'}(R + 2B + 1) - 2Q = Q(2B + 1 - s(t'/2)) - 2Q = Q(2B - s(t'/2)) - Q > -Q$, which is a contradiction.

From properties (v) and (vi), in the solution $(x_1^*, \ldots, x_T^*; y_1^*, \ldots, y_T^*; u_1^*, \ldots, u_T^*)$, the generator is online in exactly n even periods. Consider the following solution of the given instance of Equal-size Partition: For each $i = 1, \ldots, 2n$, $i \in A'$ if and only if the generator is online in period $2i$. Clearly, $|A'| = |A|/2$. In the following, we show that $\sum_{i \in A'} s(i) = \sum_{i \in A \setminus A'} s(i)$. By property (i), the generator is offline in periods $1, 3, \ldots, 4n-1$. Thus, the generator is online in period t only if t is even and $t/2 \in A'$. Hence, by property (iv), $\sum_{i \in A'} x_{2i}^* = Q$, or equivalently, $\sum_{i \in A'} (-x_{2i}^*) = -Q$. By property (ii), the cost incurred in period $2i$ is at least $-x_{2i}^*$ for all

$i \in A'$. Suppose that, on the contrary, the cost incurred in period $2i$ is greater than $-x^*_{2i}$ for some $i \in A'$. Then, the total cost of this solution must be greater than $\sum_{i\in A'}(-x^*_{2i}) = -Q$, which is a contradiction. Thus, the cost incurred in period $2i$ is exactly $-x^*_{2i}$ for all $i \in A'$. By property (ii), $x^*_{2i} = R + s(i)$ if $i \in A'$. Hence, $\sum_{i\in A'} x^*_{2i} = \sum_{i\in A'}(R + s(i))$, which implies that $Q = nR + \sum_{i\in A'} s(i)$. Therefore, $\sum_{i\in A'} s(i) = B$, or equivalently, $\sum_{i\in A'} s(i) = \sum_{i\in A\setminus A'} s(i)$. □

A.2. Proof of Proposition 2

Because there are $3T-1$ variables in $\mathcal{D}$, it suffices to show that $dim(\text{conv}(\mathcal{D})) = 3T-1$. Note that $\mathbf{0} \in \text{conv}(\mathcal{D})$. Thus, it suffices to show that there exist $3T-1$ linearly independent non-zero points in $\text{conv}(\mathcal{D})$. We create these $3T-1$ points and divide them into three groups, namely groups (A1), (A2), and (A3). We use $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r)$, $(\hat{x}^r, \hat{y}^r, \hat{u}^r)$, and $(\bar{x}^r, \bar{y}^r, \bar{u}^r)$ to denote point r in groups (A1), (A2), and (A3), respectively, and we refer to r as the index of the point within the group. Let

$$\epsilon = \min\left\{\overline{V} - \underline{C},\ \overline{C} - \underline{C},\ \frac{Q}{L+1} - \underline{C}\right\}.$$

Because $\underline{C} < \overline{V}$, $\underline{C} < \overline{C}$, and $(L+1)\underline{C} < Q$, we have $\epsilon > 0$. The $3T-1$ points are created as follows:

(A1) For each $r \in \mathcal{T}$, we create a point $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r) \in \text{conv}(\mathcal{D})$ as follows: (a) For each $s \in \mathcal{T}$, set $\tilde{x}^r_s = \underline{C} + \epsilon$ and $\tilde{y}^r_s = 1$ if $s \in [\max\{1, r-L\}, r]_{\mathbb{Z}}$, and set $\tilde{x}^r_s = \tilde{y}^r_s = 0$ otherwise. (b) If $r < L+2$, then set $\tilde{u}^r_s = 0$ for each $s \in \mathcal{T}\setminus\{1\}$. (c) If $r \geq L+2$, then for each $s \in \mathcal{T}\setminus\{1\}$, set $\tilde{u}^r_s = 1$ if $s = r-L$, and set $\tilde{u}^r_s = 0$ otherwise. There are T points created in this group.

(A2) For each $r \in \mathcal{T}$, we create a point $(\hat{x}^r, \hat{y}^r, \hat{u}^r) \in \text{conv}(\mathcal{D})$ as follows: (a) For each $s \in \mathcal{T}$, set $\hat{x}^r_s = \underline{C}$ and $\hat{y}^r_s = 1$ if $s \in [\max\{1, r-L\}, r]_{\mathbb{Z}}$, and set $\hat{x}^r_s = \hat{y}^r_s = 0$ otherwise. (b) If $r < L+2$, then set $\hat{u}^r_s = 0$ for each $s \in \mathcal{T}\setminus\{1\}$. (c) If $r \geq L+2$, then for each $s \in \mathcal{T}\setminus\{1\}$, set $\hat{u}^r_s = 1$ if $s = r-L$, and set $\hat{u}^r_s = 0$ otherwise. There are T points created in this group.

(A3) For each $r \in \mathcal{T}\setminus\{1\}$, we create a point $(\bar{x}^r, \bar{y}^r, \bar{u}^r) \in \text{conv}(\mathcal{D})$ as follows: (a) For each $s \in \mathcal{T}$, set $\bar{x}^r_s = \underline{C}$ and $\bar{y}^r_s = 1$ if $s \in [r, \min\{r+L-1, T\}]_{\mathbb{Z}}$, and set $\bar{x}^r_s = \bar{y}^r_s = 0$ otherwise. (b) For each $s \in \mathcal{T}\setminus\{1\}$, set $\bar{u}^r_s = 1$ if $s = r$, and set $\bar{u}^r_s = 0$ otherwise. There are $T-1$ points created in this group.

Table EC.1 shows a matrix in which each row represents a point created by the above process. It is easy to check that each of these $3T-1$ points satisfies constraints (2a)–(2h). Thus, these points are in $\text{conv}(\mathcal{D})$. In the following, we show that the matrix in Table EC.1 can be transformed into a lower triangular matrix via Gaussian elimination. The transformed matrix is shown in Table EC.2, where the rows are divided into Groups 1, 2, and 3. The Gaussian elimination process is as follows:

(i) For each $r \in \mathcal{T}$, point r of Group 1, denoted $(\underline{\tilde{x}}^r, \underline{\tilde{y}}^r, \underline{\tilde{u}}^r)$, is obtained by setting $(\underline{\tilde{x}}^r, \underline{\tilde{y}}^r, \underline{\tilde{u}}^r) = (\tilde{x}^r, \tilde{y}^r, \tilde{u}^r) - (\hat{x}^r, \hat{y}^r, \hat{u}^r)$. Here, $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r)$ is the point with index r in group (A1), and $(\hat{x}^r, \hat{y}^r, \hat{u}^r)$ is the point with index r in group (A2).

(ii) For each $r \in \mathcal{T}$, point r of Group 2, denoted $(\underline{\hat{x}}^r, \underline{\hat{y}}^r, \underline{\hat{u}}^r)$, is obtained by setting $(\underline{\hat{x}}^r, \underline{\hat{y}}^r, \underline{\hat{u}}^r) = (\hat{x}^r, \hat{y}^r, \hat{u}^r)$ if $r \leq L+1$, and setting $(\underline{\hat{x}}^r, \underline{\hat{y}}^r, \underline{\hat{u}}^r) = (\hat{x}^r, \hat{y}^r, \hat{u}^r) - (\bar{x}^{r-L}, \bar{y}^{r-L}, \bar{u}^{r-L})$ if $r > L+1$. Here, $(\hat{x}^r, \hat{y}^r, \hat{u}^r)$ is the point with index r in group (A2), and $(\bar{x}^{r-L}, \bar{y}^{r-L}, \bar{u}^{r-L})$ is the point with index $r-L$ in group (A3).

Table EC.1 A matrix with the rows representing $3T-1$ linearly independent points in $\mathrm{conv}(\mathcal{D})$

Grp	Index	x: 1	2	3	…	$L+1$	$L+2$	$L+3$	…	$T-L$	$T-L+1$	…	$T-1$	T
(A1)	1	$\underline{C}+\epsilon$	0	0	…	0	0	0	…	0	0	…	0	0
	2	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	0	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$L+1$	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	…	$\underline{C}+\epsilon$	0	0	…	0	0	…	0	0
	$L+2$	0	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	…	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	0	…	0	0	…	0	0
	$L+3$	0	0	$\underline{C}+\epsilon$	…	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	0	…	0	0	0	…	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$	…	$\underline{C}+\epsilon$	$\underline{C}+\epsilon$
(A2)	1	$\underline{C}$	0	0	…	0	0	0	…	0	0	…	0	0
	2	$\underline{C}$	$\underline{C}$	0	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$L+1$	$\underline{C}$	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	0	0	…	0	0	…	0	0
	$L+2$	0	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	$\underline{C}$	0	…	0	0	…	0	0
	$L+3$	0	0	$\underline{C}$	…	$\underline{C}$	$\underline{C}$	$\underline{C}$	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	0	…	0	0	0	…	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	$\underline{C}$
(A3)	2	0	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	0	0	…	0	0	…	0	0
	3	0	0	$\underline{C}$	…	$\underline{C}$	$\underline{C}$	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$T-L$	0	0	0	…	0	0	0	…	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	0
	$T-L+1$	0	0	0	…	0	0	0	…	0	$\underline{C}$	…	$\underline{C}$	$\underline{C}$
	$T-L+2$	0	0	0	…	0	0	0	…	0	0	…	$\underline{C}$	$\underline{C}$
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	0	…	0	0	0	…	0	0	…	0	$\underline{C}$

Grp	Index	y: 1	2	3	…	$L+1$	$L+2$	$L+3$	…	$T-L$	$T-L+1$	…	$T-1$	T
(A1)	1	1	0	0	…	0	0	0	…	0	0	…	0	0
	2	1	1	0	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$L+1$	1	1	1	…	1	0	0	…	0	0	…	0	0
	$L+2$	0	1	1	…	1	1	0	…	0	0	…	0	0
	$L+3$	0	0	1	…	1	1	1	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	0	…	0	0	0	…	1	1	…	1	1
(A2)	1	1	0	0	…	0	0	0	…	0	0	…	0	0
	2	1	1	0	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$L+1$	1	1	1	…	1	0	0	…	0	0	…	0	0
	$L+2$	0	1	1	…	1	1	0	…	0	0	…	0	0
	$L+3$	0	0	1	…	1	1	1	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	0	…	0	0	0	…	1	1	…	1	1
(A3)	2	0	1	1	…	1	0	0	…	0	0	…	0	0
	3	0	0	1	…	1	1	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$T-L$	0	0	0	…	0	0	0	…	1	1	…	1	0
	$T-L+1$	0	0	0	…	0	0	0	…	0	1	…	1	1
	$T-L+2$	0	0	0	…	0	0	0	…	0	0	…	1	1
	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	0	…	0	0	0	…	0	0	…	0	1

Grp	Index	u: 2	3	…	$L+1$	$L+2$	$L+3$	…	$T-L$	$T-L+1$	…	$T-1$	T
(A1)	1	0	0	…	0	0	0	…	0	0	…	0	0
	2	0	0	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$L+1$	0	0	…	0	0	0	…	0	0	…	0	0
	$L+2$	1	0	…	0	0	0	…	0	0	…	0	0
	$L+3$	0	1	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	…	0	0	0	…	1	0	…	0	0
(A2)	1	0	0	…	0	0	0	…	0	0	…	0	0
	2	0	0	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$L+1$	0	0	…	0	0	0	…	0	0	…	0	0
	$L+2$	1	0	…	0	0	0	…	0	0	…	0	0
	$L+3$	0	1	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	…	0	0	0	…	1	0	…	0	0
(A3)	2	1	0	…	0	0	0	…	0	0	…	0	0
	3	0	1	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	$T-L$	0	0	…	0	0	0	…	1	0	…	0	0
	$T-L+1$	0	0	…	0	0	0	…	0	1	…	0	0
	$T-L+2$	0	0	…	0	0	0	…	0	0	…	0	0
	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮	⋮	⋱	⋮	⋮
	T	0	0	…	0	0	0	…	0	0	…	0	1

Table EC.2 Lower-triangular matrix obtained by Gaussian elimination

x

Grp	Index	1	2	3	…	$L+1$	$L+2$	$L+3$	…	$T-L$	$T-L+1$	…	$T-1$	T
1	1	ϵ	0	0	…	0	0	0	…	0	0	…	0	0
	2	ϵ	ϵ	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$L+1$	ϵ	ϵ	ϵ	…	ϵ	0	0	…	0	0	…	0	0
	$L+2$	0	ϵ	ϵ	…	ϵ	ϵ	0	…	0	0	…	0	0
	$L+3$	0	0	ϵ	…	ϵ	ϵ	ϵ	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	0	…	0	0	0	…	ϵ	ϵ	…	ϵ	ϵ
2	1	$\underline{C}$	0	0	…	0	0	0	…	0	0	…	0	0
	2	$\underline{C}$	$\underline{C}$	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$L+1$	$\underline{C}$	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	0	0	…	0	0	…	0	0
	$L+2$	0	0	0	…	0	$\underline{C}$	0	…	0	0	…	0	0
	$L+3$	0	0	0	…	0	0	$\underline{C}$	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	0	…	0	0	0	…	0	0	…	0	$\underline{C}$
3	2	0	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	0	0	…	0	0	…	0	0
	3	0	0	$\underline{C}$	…	$\underline{C}$	$\underline{C}$	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$T-L$	0	0	0	…	0	0	0	…	$\underline{C}$	$\underline{C}$	…	$\underline{C}$	0
	$T-L+1$	0	0	0	…	0	0	0	…	0	$\underline{C}$	…	$\underline{C}$	$\underline{C}$
	$T-L+2$	0	0	0	…	0	0	0	…	0	0	…	$\underline{C}$	$\underline{C}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	0	…	0	0	0	…	0	0	…	0	$\underline{C}$

y

Grp	Index	1	2	3	…	$L+1$	$L+2$	$L+3$	…	$T-L$	$T-L+1$	…	$T-1$	T
1	1	0	0	0	…	0	0	0	…	0	0	…	0	0
	2	0	0	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$L+1$	0	0	0	…	0	0	0	…	0	0	…	0	0
	$L+2$	0	0	0	…	0	0	0	…	0	0	…	0	0
	$L+3$	0	0	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	0	…	0	0	0	…	0	0	…	0	0
2	1	1	0	0	…	0	0	0	…	0	0	…	0	0
	2	1	1	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$L+1$	1	1	1	…	1	0	0	…	0	0	…	0	0
	$L+2$	0	0	0	…	0	1	0	…	0	0	…	0	0
	$L+3$	0	0	0	…	0	0	1	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	0	…	0	0	0	…	0	0	…	0	1
3	2	0	1	1	…	1	0	0	…	0	0	…	0	0
	3	0	0	1	…	1	1	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$T-L$	0	0	0	…	0	0	0	…	1	1	…	1	0
	$T-L+1$	0	0	0	…	0	0	0	…	0	1	…	1	1
	$T-L+2$	0	0	0	…	0	0	0	…	0	0	…	1	1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	0	…	0	0	0	…	0	0	…	0	1

u

Grp	Index	2	3	…	$L+1$	$L+2$	$L+3$	…	$T-L$	$T-L+1$	…	$T-1$	T
1	1	0	0	…	0	0	0	…	0	0	…	0	0
	2	0	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$L+1$	0	0	…	0	0	0	…	0	0	…	0	0
	$L+2$	0	0	…	0	0	0	…	0	0	…	0	0
	$L+3$	0	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	…	0	0	0	…	0	0	…	0	0
2	1	0	0	…	0	0	0	…	0	0	…	0	0
	2	0	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$L+1$	0	0	…	0	0	0	…	0	0	…	0	0
	$L+2$	0	0	…	0	0	0	…	0	0	…	0	0
	$L+3$	0	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	…	0	0	0	…	0	0	…	0	0
3	2	1	0	…	0	0	0	…	0	0	…	0	0
	3	0	1	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$T-L$	0	0	…	0	0	0	…	1	0	…	0	0
	$T-L+1$	0	0	…	0	0	0	…	0	1	…	0	0
	$T-L+2$	0	0	…	0	0	0	…	0	0	…	0	0
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	T	0	0	…	0	0	0	…	0	0	…	0	1

(iii) For each $r \in \mathcal{T} \setminus \{1\}$, point r of Group 3, denoted $(\underline{\bar{x}}^r, \underline{\bar{y}}^r, \underline{\bar{u}}^r)$, is obtained by setting $(\underline{\bar{x}}^r, \underline{\bar{y}}^r, \underline{\bar{u}}^r) = (\bar{x}^r, \bar{y}^r, \bar{u}^r)$, where $(\bar{x}^r, \bar{y}^r, \bar{u}^r)$ is the point with index r in group (A3).

For each $r \in \mathcal{T}$, point $(\underline{\tilde{x}}^r, \underline{\tilde{y}}^r, \underline{\tilde{u}}^r)$ is a row vector, where the last non-zero component is $\underline{\tilde{x}}_r^r$ (i.e., $\underline{\tilde{x}}_r^r \neq 0$; $\underline{\tilde{x}}_i^r = 0$ for $i > r$; $\underline{\tilde{y}}_i^r = 0$ for $i \in \mathcal{T}$; and $\underline{\tilde{u}}_i^r = 0$ for $i \in \mathcal{T} \setminus \{1\}$). For each $r \in \mathcal{T}$, point $(\underline{\hat{x}}^r, \underline{\hat{y}}^r, \underline{\hat{u}}^r)$ is a row vector, where the last non-zero component is $\underline{\hat{y}}_r^r$ (i.e., $\underline{\hat{y}}_r^r \neq 0$; $\underline{\hat{y}}_i^r = 0$ for $i > r$; and $\underline{\hat{u}}_i^r = 0$ for $i \in \mathcal{T} \setminus \{1\}$). For each $r \in \mathcal{T} \setminus \{1\}$, point $(\underline{\bar{x}}^r, \underline{\bar{y}}^r, \underline{\bar{u}}^r)$ is a row vector, where the last nonzero component is $\underline{\bar{u}}_r^r$ (i.e., $\underline{\bar{u}}_r^r \neq 0$ and $\underline{\bar{u}}_i^r = 0$ for $i > r$). Hence, the points in Groups 1, 2, and 3 form a lower triangular matrix with nonzero entries on the diagonal (see Table EC.2), and thus they are linearly independent. This implies that the $3T - 1$ points in groups (A1), (A2), and (A3) are linearly independent. This completes the proof of the proposition. $\square$

A.3. Proof of Lemma 1

Recall that $\mathcal{D}_{sc} = \{x \in \mathbb{R}^T : \text{(2h)};\ x_t \in \{0\} \cup [\underline{C}, \overline{C}]\ \forall t \in \mathcal{T}\}$. Clearly, given any extreme point $(x_1^*, \ldots, x_T^*)$ of $\text{conv}(\mathcal{D}_{\text{sc}})$, it remains an extreme point of $\text{conv}(\mathcal{D}_{\text{sc}})$ if we permute the components of $(x_1^*, \ldots, x_T^*)$. In other words, to characterize an extreme point of $\text{conv}(\mathcal{D}_{\text{sc}})$, the order of the x_t^*'s is irrelevant. Recall that $\lambda = \lfloor Q/\overline{C} \rfloor$, $\lambda' = \lceil (\underline{C} + \lambda\overline{C} - Q)/(\overline{C} - \underline{C}) \rceil$, and $\theta^* = Q - \lambda\overline{C} + \lambda'(\overline{C} - \underline{C})$. Note that

$$\theta^* \geq Q - \lambda\overline{C} + \frac{\underline{C} + \lambda\overline{C} - Q}{\overline{C} - \underline{C}} \cdot (\overline{C} - \underline{C}) = \underline{C}$$

and

$$\theta^* < Q - \lambda\overline{C} + \left(\frac{\underline{C} + \lambda\overline{C} - Q}{\overline{C} - \underline{C}} + 1\right)(\overline{C} - \underline{C}) = \overline{C}.$$

Thus, $\theta^* \in [\underline{C}, \overline{C})$.

Let x be any extreme point of $\text{conv}(\mathcal{D}_{\text{sc}})$. Note that $x \in \mathcal{D}_{sc}$. Because $\mathcal{D}_{sc}$ is a knapsack set, at most one component of x belongs to the interval $(\underline{C}, \overline{C})$, while all other components belong to $\{0, \underline{C}, \overline{C}\}$. Hence, one of the following two cases holds:

(a) $x_i \in \{0, \overline{C}\}\ \forall i \in \mathcal{T}$.

(b) x contains one component with value in the interval $[\underline{C}, \overline{C})$, while each of the other $T - 1$ components is equal to 0, $\underline{C}$, or $\overline{C}$.

Note that the definition of λ implies that $(\lambda + 1)\overline{C} > Q$. Thus, in case (a), no more than λ components are equal to $\overline{C}$. This case belongs to the category 1 in Lemma 1. Hence, it suffices to show that the lemma is valid in case (b). In this case, we let θ denote the value of the component in the interval $[\underline{C}, \overline{C})$. Among the remaining $T - 1$ components, we let τ_u, τ_l, and τ_0 denote the number of components with values $\overline{C}$, $\underline{C}$, and 0, respectively, where $\tau_u, \tau_l \geq 0$ and $\tau_0 = T - \tau_u - \tau_l - 1 \geq 0$. It suffices to show that $\theta = \theta^*$, $\tau_u = \lambda - \lambda'$, and $\tau_l = \lambda'$. Note that x must satisfy the knapsack inequality (2h) with equality, since otherwise $x = \frac{x_i}{x_i + \epsilon}(x + \epsilon e_i) + \frac{\epsilon}{x_i + \epsilon}(x - x_i e_i)$ and $x + \epsilon e_i, x - x_i e_i \in \text{conv}(\mathcal{D}_{\text{sc}})$ for some small positive value ϵ, where x_i is a component of x with $x_i < \overline{C}$, and $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the unit vector with the ith component equal to 1, contradicting that x is an extreme point of $\text{conv}(\mathcal{D}_{\text{sc}})$. This implies that $\tau_u\overline{C} + \tau_l\underline{C} + \theta = Q$.

We first prove that $\tau_u + \tau_l = \lambda$. Note that

$$\tau_u + \tau_l \geq \frac{\tau_u\overline{C} + \tau_l\underline{C}}{\overline{C}} > \frac{\tau_u\overline{C} + \tau_l\underline{C} + \theta - \overline{C}}{\overline{C}} = \frac{Q - \overline{C}}{\overline{C}} \geq \left\lfloor \frac{Q}{\overline{C}} \right\rfloor - 1 = \lambda - 1,$$

which implies that $\tau_u+\tau_l \geq \lambda$. Thus, to prove $\tau_u+\tau_l=\lambda$, it suffices to show that $\tau_u+\tau_l \leq \lambda$. By contradiction, suppose $\tau_u+\tau_l \geq \lambda+1$. Note that under this supposition, $\tau_l \geq 1$, since otherwise the sum of all components of x exceeds $(\lambda+1)\overline{C}=(\lfloor Q/\overline{C}\rfloor+1)\overline{C}>Q$. Let x' denote the extreme point of $\operatorname{conv}(\mathcal{D}_{\rm sc})$ obtained by permuting the components of x such that the components are arranged in non-increasing order of their values; that is,

$$x'=(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},\theta,\underbrace{\underline{C},\ldots,\underline{C}}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}}).$$

We consider two different cases, namely the case with $\tau_l \geq \overline{C}/(\overline{C}-\underline{C})$ and the case with $\tau_l < \overline{C}/(\overline{C}-\underline{C})$.

Case (i): $\tau_l \geq \overline{C}/(\overline{C}-\underline{C})$. Let $C'=\frac{\tau_l}{\tau_l-1}\cdot\underline{C}$. It is easy to verify that $\underline{C}<C'\leq\overline{C}$. Then,

$$\begin{aligned}
x' = &\ \frac{1}{\tau_l}\cdot(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},\theta,\underbrace{C',\ldots,C',0}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}})\\
&+\frac{1}{\tau_l}\cdot(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},\theta,\underbrace{C',\ldots,C',0,C'}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}})\\
&+\cdots\\
&+\frac{1}{\tau_l}\cdot(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},\theta,\underbrace{0,C',\ldots,C'}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}}),
\end{aligned}$$

and the τ_l vectors on the right-hand side of this equation are distinct elements of $\mathcal{D}_{sc}$. Hence, x' is a convex combination of τ_l distinct elements of $\operatorname{conv}(\mathcal{D}_{\rm sc})$.

Case (ii): $\tau_l < \overline{C}/(\overline{C}-\underline{C})$. Let $\theta'=\theta-(\tau_l-1)\overline{C}+\tau_l\underline{C}$ and $C''=\frac{\theta}{\tau_l}+\underline{C}$. Note that

$$\theta'=\theta+(\overline{C}-\underline{C})\left(\frac{\overline{C}}{\overline{C}-\underline{C}}-\tau_l\right)>\theta$$

and

$$\theta'=\tau_u\overline{C}+\tau_l\underline{C}+\theta-(\tau_u+\tau_l-1)\overline{C}=Q-(\tau_u+\tau_l-1)\overline{C}\leq Q-\lambda\overline{C}=Q-\left\lfloor\frac{Q}{\overline{C}}\right\rfloor\cdot\overline{C}\leq\overline{C},$$

which together imply that $\theta'\in[\underline{C},\overline{C}]$ and $\theta'\geq\theta$. Note also that $C''>\underline{C}$ and

$$\begin{aligned}
C''&=\frac{1}{\tau_l}\left[\tau_u\overline{C}+\tau_l\underline{C}+\theta-(\tau_u+\tau_l)\overline{C}\right]+\overline{C}=\frac{1}{\tau_l}\left[Q-(\tau_u+\tau_l)\overline{C}\right]+\overline{C}\leq\frac{1}{\tau_l}\left[Q-(\lambda+1)\overline{C}\right]+\overline{C}\\
&=\frac{1}{\tau_l}\left[Q-\left(\left\lfloor\frac{Q}{\overline{C}}\right\rfloor+1\right)\overline{C}\right]+\overline{C}\leq\overline{C},
\end{aligned}$$

which together imply that $C''\in[\underline{C},\overline{C}]$. It is easy to verify that

$$\begin{aligned}
x'=\frac{\theta'-\theta}{\theta'}\cdot(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},0,\underbrace{C'',\ldots,C''}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}})+\frac{\theta}{\theta'}\Bigg[&\frac{1}{\tau_l}\cdot(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},\theta',\underbrace{\underline{C},\ldots,\underline{C},0}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}})\\
&+\frac{1}{\tau_l}\cdot(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},\theta',\underbrace{\underline{C},\ldots,\underline{C},0,\underline{C}}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}})\\
&+\cdots\\
&+\frac{1}{\tau_l}\cdot(\underbrace{\overline{C},\ldots,\overline{C}}_{\tau_u \text{ terms}},\theta',\underbrace{0,\underline{C},\ldots,\underline{C}}_{\tau_l \text{ terms}},\underbrace{0,\ldots,0}_{\tau_0 \text{ terms}})\Bigg]
\end{aligned}$$

(note: if $\tau_l=1$, then the right-hand side of this equation becomes $\frac{\theta'-\theta}{\theta'}(\overline{C},\ldots,\overline{C},0,C'',0,\ldots,0)+\frac{\theta}{\theta'}(\overline{C},\ldots,\overline{C},\theta',0,0,\ldots,0)$). Hence, x' is a convex combination of τ_l+1 distinct elements of $\operatorname{conv}(\mathcal{D}_{\rm sc})$.

In both cases, x' is a convex combination of two or more distinct elements of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$, which contradicts that x' is an extreme point of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$. Therefore, $\tau_u + \tau_l = \lambda$.

Note that

$$\tau_l = \frac{\theta + (\tau_u + \tau_l)\overline{C} - (\tau_u \overline{C} + \tau_l \underline{C} + \theta)}{\overline{C} - \underline{C}} = \frac{\theta + \lambda \overline{C} - Q}{\overline{C} - \underline{C}}.$$

Because $\underline{C} \leq \theta < \overline{C}$, we have

$$\frac{\underline{C} + \lambda \overline{C} - Q}{\overline{C} - \underline{C}} \leq \frac{\theta + \lambda \overline{C} - Q}{\overline{C} - \underline{C}} < \frac{\underline{C} + \lambda \overline{C} - Q}{\overline{C} - \underline{C}} + 1.$$

Thus,

$$\tau_l = \left\lceil \frac{\underline{C} + \lambda \overline{C} - Q}{\overline{C} - \underline{C}} \right\rceil = \lambda'.$$

This implies that $\tau_u = \lambda - \lambda'$, and that $\theta = Q - \tau_u \overline{C} + \tau_l \underline{C} = Q - (\lambda - \lambda')\overline{C} + \lambda' \underline{C} = \theta^*$. This completes the proof of the Lemma. □

A.4. Proof of Proposition 3

Consider any $\mathcal{T}_1 \subseteq \mathcal{T}$ such that $\lambda - \lambda' + 1 \leq |\mathcal{T}_1| \leq \lambda$. We show that inequality (3) is valid for $\mathrm{conv}(\mathcal{D})$ and is a semi-continuous inequality. To show that inequality (3) is valid for $\mathrm{conv}(\mathcal{D})$, because $\mathcal{D}_{sc} \subseteq \mathcal{D}$, it suffices to show that inequality (3) is valid for $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$. To do so, we show that every extreme point of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ satisfies inequality (3). To show that inequality (3) is a semi-continuous inequality, it suffices to show that inequality (3) is facet-defining for $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$.

First, we show that every extreme point of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ satisfies inequality (3). Note that because

$$Q - \lambda \overline{C} = Q - \left\lfloor \frac{Q}{\overline{C}} \right\rfloor \cdot \overline{C} \geq 0$$

and

$$\underline{C} - (\lambda - |\mathcal{T}_1|)(\overline{C} - \underline{C}) \geq \underline{C} - (\lambda' - 1)(\overline{C} - \underline{C}) > \underline{C} - \frac{\underline{C} + \lambda \overline{C} - Q}{\overline{C} - \underline{C}} \cdot (\overline{C} - \underline{C}) = Q - \lambda \overline{C},$$

we have $0 < \rho \leq 1$. Let x^* be any extreme point of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$. Suppose x^* belongs to category 1 in Lemma 1. Let s_u be the number of components of x^* with value $\overline{C}$ that belong to $\mathcal{T}_1$, and let s'_u be the number of components of x^* with value $\overline{C}$ that belong to $\mathcal{T} \setminus \mathcal{T}_1$, where $s_u + s'_u \leq \lambda$. Then,

$$\begin{aligned}
\sum_{t \in \mathcal{T}_1} x_t + \rho \sum_{t \in \mathcal{T} \setminus \mathcal{T}_1} x_t &= s_u \overline{C} + \rho s'_u \overline{C} \\
&\leq s_u \overline{C} + \rho s'_u \overline{C} + (1 - \rho)(|\mathcal{T}_1| - s_u)\overline{C} \quad (\text{as } s_u \leq |\mathcal{T}_1| \text{ and } \rho \leq 1) \\
&= |\mathcal{T}_1|\overline{C} + \rho(s'_u + s_u - |\mathcal{T}_1|)\overline{C} \\
&\leq \left[|\mathcal{T}_1| + \rho(\lambda - |\mathcal{T}_1|)\right]\overline{C} \\
&= \zeta,
\end{aligned}$$

and thus x^* satisfies inequality (3).

Now, suppose x^* belongs to category 2 in Lemma 1. Let $s_m = 1$ and $s'_m = 0$ if the component of x^* with value θ^* belongs to $\mathcal{T}_1$, and let $s_m = 0$ and $s'_m = 1$ otherwise. Among the other $T - 1$ components of x^*, let s_u and s_l be the number of components with value $\overline{C}$ and $\underline{C}$, respectively, that belong to $\mathcal{T}_1$, and let s'_u and s'_l

be the number of components with value $\overline{C}$ and $\underline{C}$, respectively, that belong to $\mathcal{T}\setminus\mathcal{T}_1$, where $s_u+s_u'=\lambda-\lambda'$ and $s_l+s_l'=\lambda'$. Then,

$$\begin{aligned}
&\sum_{t\in\mathcal{T}_1} x_t+\rho\sum_{t\in\mathcal{T}\setminus\mathcal{T}_1} x_t\\
&=(s_u\overline{C}+s_l\underline{C}+s_m\theta^*)+\rho(s_u'\overline{C}+s_l'\underline{C}+s_m'\theta^*)\\
&\le(s_u\overline{C}+s_l\underline{C}+s_m\theta^*)+\rho(s_u'\overline{C}+s_l'\underline{C}+s_m'\theta^*)+(1-\rho)\big[s_u'(\overline{C}-\underline{C})\\
&\quad+s_m'(\theta^*-\underline{C})+(|\mathcal{T}_1|-s_u-s_l-s_m)\underline{C}\big]\quad(\text{as }\overline{C}>\underline{C},\ \theta^*\ge\underline{C},\ s_u+s_l+s_m\le|\mathcal{T}_1|,\text{ and }\rho\le 1)\\
&=(s_u+s_u')\overline{C}+(s_l+s_l')\underline{C}+(s_m+s_m')\theta^*+(1-\rho)(|\mathcal{T}_1|-s_u-s_u'-s_l-s_l'-s_m-s_m')\underline{C}\\
&=(\lambda-\lambda')\overline{C}+\lambda'\underline{C}+(1)\big[Q-(\lambda-\lambda')\overline{C}-\lambda'\underline{C}\big]+(1-\rho)(|\mathcal{T}_1|-\lambda-1)\underline{C}\\
&=\big[|\mathcal{T}_1|+\rho(\lambda-|\mathcal{T}_1|)\big]\overline{C}+(Q-\lambda\overline{C})-(1-\rho)\big[\underline{C}-(\lambda-|\mathcal{T}_1|)(\overline{C}-\underline{C})\big]\\
&=\big[|\mathcal{T}_1|+\rho(\lambda-|\mathcal{T}_1|)\big]\overline{C}\\
&=\zeta,
\end{aligned}$$

and thus x^* satisfies inequality (3).

Next, we show that inequality (3) is facet-defining for $\operatorname{conv}(\mathcal{D}_{\rm sc})$. To do so, we construct T affinely independent and feasible points on the face of $\operatorname{conv}(\mathcal{D}_{\rm sc})$ defined by inequality (3). Note that $\lambda<T$ (because $\lambda\overline{C}\le Q<T\overline{C}$). Note also that $\lambda'\ge 1$, since otherwise $\mathcal{T}_1$ does not exist. Without loss of generality, we assume that $\mathcal{T}_1=\{1,\dots,|\mathcal{T}_1|\}$. We consider two different cases.

Case 1: $\underline{C}<\theta^*<\overline{C}$. For $j=1,\dots,T$, define

$$x^j=\begin{cases}(\underbrace{\overline{C},\dots,\theta^*,\dots,\overline{C}}_{\lambda-\lambda'+1\text{ terms}},\underbrace{\underline{C},\dots,\underline{C}}_{\lambda'\text{ terms}},0,\dots,0), & \text{if } j\le\lambda-\lambda'+1,\ \text{where }\theta^*\text{ is in the }j\text{th position;}\\
(\underbrace{\overline{C},\dots,\overline{C}}_{\lambda-\lambda'\text{ terms}},\underbrace{\underline{C},\dots,\theta^*,\dots,\underline{C}}_{\lambda'+1\text{ terms}},0,\dots,0), & \text{if } \lambda-\lambda'+2\le j\le|\mathcal{T}_1|,\ \text{where }\theta^*\text{ is in the }j\text{th position;}\\
(\underbrace{\overline{C},\dots,0,\dots,\overline{C}}_{\lambda+1\text{ terms}},0,\dots,0), & \text{if } |\mathcal{T}_1|+1\le j\le\lambda+1,\ \text{where the first 0 is in the }j\text{th position;}\\
(\underbrace{\overline{C},\dots,\overline{C}}_{\lambda-1\text{ terms}},0,\dots,0,\overline{C},0,\dots,0), & \text{if } j\ge\lambda+2,\ \text{where the last }\overline{C}\text{ is in the }j\text{th position.}\end{cases}$$

It is easy to verify that x^j is an element of $\mathcal{D}_{sc}$, and it satisfies inequality (3) with equality. For $j=2,\dots,T$, define

$$\bar{x}^j=x^j-x^1.$$

We consider two subcases.

Case 1.1: $|\mathcal{T}_1|=\lambda$. In this case, for each $j=2,\dots,T$, the jth component of $\bar{x}^j$ is nonzero, while the ith component of $\bar{x}^j$ is zero for all $i=j+1,\dots,T$. Thus, $\bar{x}^2,\dots,\bar{x}^T$ are linearly independent.

Case 1.2: $|\mathcal{T}_1|<\lambda$. For $j=2,\dots,T$, define

$$\hat{x}^j=\begin{cases}\bar{x}^j, & \text{if } j\le|\mathcal{T}_1| \text{ or } j\ge\lambda+2;\\ \bar{x}^j-\bar{x}^{j-1}, & \text{if } |\mathcal{T}_1|+2\le j\le\lambda+1;\\ \bar{x}^{|\mathcal{T}_1|+1}+\varphi(\bar{x}^{|\mathcal{T}_1|+2}+\dots+\bar{x}^{\lambda+1}), & \text{if } j=|\mathcal{T}_1|+1;\end{cases}$$

where $\varphi=(\overline{C}-\underline{C})/[\overline{C}-(\lambda-|\mathcal{T}_1|)(\overline{C}-\underline{C})]$. Note that the ith component of $\bar{x}^j$ is zero for all $i=j+1,\dots,T$. It is easy to see that if $j\ne|\mathcal{T}_1|+1$, then the jth component of $\hat{x}^j$ is non-zero. The $(|\mathcal{T}_1|+1)$st component of $\hat{x}^{|\mathcal{T}_1|+1}$ is $-\underline{C}+\varphi(\lambda-|\mathcal{T}_1|)(\overline{C}-\underline{C})$. Note that $\lambda-|\mathcal{T}_1|\neq \underline{C}/(\overline{C}-\underline{C})$ (since otherwise $|\mathcal{T}_1|$ is less than $\lambda-\lambda'+1$). Thus, $\varphi\neq 1$ and the $(|\mathcal{T}_1|+1)$st component of $\hat{x}^{|\mathcal{T}_1|+1}$ is non-zero. Hence, $\hat{x}^2,\ldots,\hat{x}^T$ are linearly independent. This implies that $\bar{x}^2,\ldots,\bar{x}^T$ are linearly independent.

Therefore, for both Cases 1.1 and 1.2, $\bar{x}^2,\ldots,\bar{x}^T$ are linearly independent, which implies that $x^1,\ldots,x^T$ are affinely independent.

Case 2: $\theta^*=\underline{C}$. For $j=1,\ldots,T$, define

$$x^j=\begin{cases}(\underbrace{\overline{C},\ldots,\underline{C},\ldots,\overline{C}}_{\lambda-\lambda'+1\text{ terms}},\underbrace{\underline{C},\ldots,\underline{C}}_{\lambda'\text{ terms}},0,\ldots,0), & \text{if } j\le\lambda-\lambda'+1, \text{ where the first } \underline{C} \text{ is in the } j\text{th position;}\\ (\underbrace{\overline{C},\ldots,\overline{C}}_{\lambda-\lambda'-1\text{ terms}},\underbrace{\underline{C},\ldots,\overline{C},\ldots,\underline{C}}_{\lambda'+2\text{ terms}},0,\ldots,0), & \text{if } \lambda-\lambda'+2\le j\le|\mathcal{T}_1|, \text{ where the last } \overline{C} \text{ is in the } j\text{th position;}\\ (\underbrace{\overline{C},\ldots,0,\ldots,\overline{C}}_{\lambda+1\text{ terms}},0,\ldots,0), & \text{if } |\mathcal{T}_1|+1\le j\le\lambda+1, \text{ where the first } 0 \text{ is in the } j\text{th position;}\\ (\underbrace{\overline{C},\ldots,\overline{C}}_{\lambda-1\text{ terms}},0,\ldots,0,\overline{C},0,\ldots,0), & \text{if } j\ge\lambda+2, \text{ where the last } \overline{C} \text{ is in the } j\text{th position.}\end{cases}$$

It is easy to verify that x^j is an element of $\mathcal{D}_{sc}$, and it satisfies inequality (3) with equality. For $j=2,\ldots,T$, define

$$\bar{x}^j=x^j-x^1.$$

We consider two subcases.

Case 2.1: $|\mathcal{T}_1|=\lambda$. In this case, for each $j=2,\ldots,T$, the jth component of $\bar{x}^j$ is non-zero, while the ith component of $\bar{x}^j$ is zero for all $i=j+1,\ldots,T$. Thus, $\bar{x}^2,\ldots,\bar{x}^T$ are linearly independent.

Case 2.2: $|\mathcal{T}_1|<\lambda$. For $j=2,\ldots,T$, define

$$\hat{x}^j=\begin{cases}\bar{x}^j, & \text{if } j\le|\mathcal{T}_1| \text{ or } j\ge\lambda+2;\\ \bar{x}^j-\bar{x}^{j-1}, & \text{if } |\mathcal{T}_1|+2\le j\le\lambda+1;\\ \bar{x}^{|\mathcal{T}_1|+1}+\varphi(\bar{x}^{|\mathcal{T}_1|+2}+\cdots+\bar{x}^{\lambda+1}), & \text{if } j=|\mathcal{T}_1|+1;\end{cases}$$

where $\varphi=(\overline{C}-\underline{C})/[\overline{C}-(\lambda-|\mathcal{T}_1|)(\overline{C}-\underline{C})]$. Following the same argument as in Case 1.2, the jth component of $\hat{x}^j$ is non-zero, while the ith component of $\hat{x}^j$ is zero for all $i=j+1,\ldots,T$. Hence, $\hat{x}^2,\ldots,\hat{x}^T$ are linearly independent. Thus, $\bar{x}^2,\ldots,\bar{x}^T$ are linearly independent.

Therefore, for both Cases 2.1 and 2.2, $\bar{x}^2,\ldots,\bar{x}^T$ are linearly independent, which implies that $x^1,\ldots,x^T$ are affinely independent. $\square$

A.5. Proof of Proposition 4

We first state the following property, which will be used in the proof.

Property 1. Consider any two valid inequalities "$\sum_{t\in\mathcal{T}}\alpha_{1t}x_t\le\beta_1$" and "$\sum_{t\in\mathcal{T}}\alpha_{2t}x_t\le\beta_2$" of $\text{conv}(\mathcal{D}_{\text{sc}})$. If the inequality "$\sum_{t\in\mathcal{T}}\alpha_{1t}x_t\le\beta_1$" is facet-defining for $\text{conv}(\mathcal{D}_{\text{sc}})$ and

$$\left\{x\in\mathcal{D}_{sc}:\sum_{t\in\mathcal{T}}\alpha_{1t}x_t=\beta_1\right\}\subseteq\left\{x\in\mathcal{D}_{sc}:\sum_{t\in\mathcal{T}}\alpha_{2t}x_t=\beta_2\right\},$$

then the inequalities "$\sum_{t\in\mathcal{T}}\alpha_{1t}x_t\le\beta_1$" and "$\sum_{t\in\mathcal{T}}\alpha_{2t}x_t\le\beta_2$" are equivalent; that is, there exists a positive scalar k such that $\alpha_{1t}=k\alpha_{2t}$ for all $t\in\mathcal{T}$ and that $\beta_1=k\beta_2$.

To show the validity of this property, we note that $\mathbf{0} \in \text{conv}(\mathcal{D}_{\text{sc}})$ and that $\overline{C}e_i \in \mathcal{D}_{sc} \subseteq \text{conv}(\mathcal{D}_{\text{sc}})$ for any $i \in \mathcal{T}$, where $e_i = (0,\ldots,0,1,0,\ldots,0)$ is a unit vector with the ith component equal to 1. Thus, $\text{conv}(\mathcal{D}_{\text{sc}})$ is full dimensional, because $\mathbf{0}, \overline{C}e_1, \overline{C}e_2, \ldots, \overline{C}e_T$ are $T+1$ affinely independent points. For $i = 1, 2$, denote

$$F_i = \left\{ x \in \mathcal{D}_{sc} : \sum_{t \in \mathcal{T}} \alpha_{it} x_t = \beta_i \right\}.$$

Because $\text{conv}(\mathcal{D}_{\text{sc}})$ is full dimensional and the inequality "$\sum_{t \in \mathcal{T}} \alpha_{1t} x_t \leq \beta_1$" is facet-defining, F_1 is a $(T-1)$-dimensional face of $\text{conv}(\mathcal{D}_{\text{sc}})$. In addition, because the inequality "$\sum_{t \in \mathcal{T}} \alpha_{2t} x_t \leq \beta_2$" is valid for $\text{conv}(\mathcal{D}_{\text{sc}})$ and $F_1 \subseteq F_2$, the inequality "$\sum_{t \in \mathcal{T}} \alpha_{2t} x_t \leq \beta_2$" is also a $(T-1)$-dimensional face of $\text{conv}(\mathcal{D}_{\text{sc}})$. Since F_1 and F_2 are both $(T-1)$-dimensional faces of $\text{conv}(\mathcal{D}_{\text{sc}})$ and $F_1 \subseteq F_2$, the inequalities "$\sum_{t \in \mathcal{T}} \alpha_{1t} x_t \leq \beta_1$" and "$\sum_{t \in \mathcal{T}} \alpha_{2t} x_t \leq \beta_2$" are equivalent. Hence, Property 1 is valid.

To prove the proposition, we refer to inequality (2h) and the inequality "$0 \leq x_t \leq \overline{C}$" as trivial inequalities of $\text{conv}(\mathcal{D}_{\text{sc}})$, as they are obviously facet-defining for $\text{conv}(\mathcal{D}_{\text{sc}})$. Without loss of generality, we assume that any non-trivial facet-defining inequality for $\text{conv}(\mathcal{D}_{\text{sc}})$ is expressed in the following form:

$$a \sum_{t \in \mathcal{N}_1} x_t + \sum_{t \in \mathcal{N}_2} \alpha_t x_t \leq \beta, \tag{EC.1}$$

where $\mathcal{N}_1$ and $\mathcal{N}_2$ are subsets of $\mathcal{T}$ such that $\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$, $\mathcal{N}_1 \cup \mathcal{N}_2 = \mathcal{T}$, $\mathcal{N}_1 \neq \emptyset$, and $a > \max\{\alpha_t : t \in \mathcal{N}_2\}$. Note that $\mathcal{N}_2 \neq \emptyset$, since otherwise (EC.1) becomes (2h) which is a trivial facet-defining inequality.

We first show that $\alpha_t \geq 0$ for all $t \in \mathcal{N}_2$. Suppose that, on the contrary, $\alpha_s < 0$ for some $s \in \mathcal{N}_2$. Then, for any $x \in \text{conv}(\mathcal{D}_{\text{sc}})$ that satisfies (EC.1) with equality, we must have $x_s = 0$, since otherwise the point $x' \in \text{conv}(\mathcal{D}_{\text{sc}})$ obtained by setting $x'_s = 0$ and $x'_t = x_t$ $\forall t \in \mathcal{T} \setminus \{s\}$ will violate (EC.1). Hence, the facet of $\text{conv}(\mathcal{D}_{\text{sc}})$ defined by (EC.1) is one of the facets defined by the trivial inequalities "$x_t \geq 0$" for $t \in \mathcal{N}_2$. This contradicts that (EC.1) is a non-trivial inequality. Therefore, $\alpha_t \geq 0$ for all $t \in \mathcal{N}_2$. This also implies that $a > 0$, because $a > \max\{\alpha_t : t \in \mathcal{N}_2\}$.

Next, we show that $\alpha_t > 0$ for all $t \in \mathcal{N}_2$. Define $\mathcal{N}_2^+ = \{t \in \mathcal{N}_2 : \alpha_t > 0\}$ and $\mathcal{N}_2^0 = \mathcal{N}_2 \setminus \mathcal{N}_2^+$. Note that $\mathcal{N}_2^+ \neq \emptyset$, since otherwise $\alpha_t = 0$ for all $t \in \mathcal{N}_2$, and thus (EC.1) is dominated by (2h). We show that $\mathcal{N}_2^0 = \emptyset$. To do so, we first show that for any $s \in \mathcal{N}_2^0$, we must have $x_s = 0$ for all $x \in \text{conv}(\mathcal{D}_{\text{sc}})$ that satisfies (EC.1) with equality. By contradiction, suppose $x_s > 0$ for some $x \in \text{conv}(\mathcal{D}_{\text{sc}})$ that satisfies (EC.1) with equality. Then, $x_t = \overline{C}$ for all $t \in \mathcal{N}_1 \cup \mathcal{N}_2^+$, since otherwise there exists some $t' \in \mathcal{N}_1 \cup \mathcal{N}_2^+$ with $x_{t'} < \overline{C}$, and thus we can construct an $\hat{x} \in \text{conv}(\mathcal{D}_{\text{sc}})$ such that $\hat{x}$ violates inequality (EC.1) by letting $\hat{x}_s = 0$, $\hat{x}_{t'} = \min\{\overline{C}, x_{t'} + x_s\}$, and $\hat{x}_t = x_t$ for all $t \in \mathcal{T} \setminus \{s, t'\}$ (because $a, \alpha_{t'} > 0$). Since (EC.1) is a facet-defining inequality for $\text{conv}(\mathcal{D}_{\text{sc}})$, there exist T affinely independent points that satisfy $x_t = \overline{C}$ for all $t \in \mathcal{N}_1 \cup \mathcal{N}_2^+$. This is possible only if $|\mathcal{N}_1 \cup \mathcal{N}_2^+| \leq 1$, which contradicts that $\mathcal{N}_1 \neq \emptyset$ and $\mathcal{N}_2^+ \neq \emptyset$. Hence, for any $s \in \mathcal{N}_2^0$, $x_s = 0$ for all $x \in \text{conv}(\mathcal{D}_{\text{sc}})$ that satisfies (EC.1) with equality. This implies that the facet of $\text{conv}(\mathcal{D}_{\text{sc}})$ defined by (EC.1) is one of the facets defined by the trivial inequalities "$x_t \geq 0$" with $t \in \mathcal{N}_2^0$. This is impossible unless $\mathcal{N}_2^0 = \emptyset$. Therefore, $\alpha_t > 0$ for all $t \in \mathcal{N}_2$.

Consider any extreme point x^* of $\text{conv}(\mathcal{D}_{\text{sc}})$ that satisfies (EC.1) with equality. We will show that x^* also satisfies (3) with equality when $\mathcal{T}_1 = \mathcal{N}_1$. To do so, we first show that $\lambda - \lambda' + 1 \leq |\mathcal{N}_1| \leq \lambda$. Suppose that, on the contrary, $|\mathcal{N}_1| \leq \lambda - \lambda'$. Then we consider three different cases.

Case (i): x^* belongs to category 1 of Lemma 1 and $x^*_t = 0$ for all $t \in \mathcal{N}_2$. In this case, $x^*_s = \overline{C}$ for all $s \in \mathcal{N}_1$, since otherwise $x^*_s < \overline{C}$ for some $s \in \mathcal{N}_1$, which implies that there exists $\hat{x} \in \mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ with

$$\hat{x}_i = \begin{cases} x^*_i, & \text{if } i \in \mathcal{T} \setminus \{s\}; \\ \overline{C}, & \text{if } i = s; \end{cases}$$

that violates inequality (EC.1), as $a > 0$ (note: $\hat{x} \in \mathcal{D}_{sc} \subseteq \mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$, because $|\mathcal{N}_1| \le \lambda - \lambda' \le \lambda$ and $x^*_t = 0$ for all $t \in \mathcal{N}_2$, and thus $\hat{x}$ satisfies (2h)).

Case (ii): x^* belongs to category 1 of Lemma 1 and $x^*_t = \overline{C}$ for some $t \in \mathcal{N}_2$. In this case, $x^*_s = \overline{C}$ for all $s \in \mathcal{N}_1$, since otherwise $x^*_s < \overline{C}$ for some $s \in \mathcal{N}_1$, which implies that there exists $\hat{x} \in \mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ with

$$\hat{x}_i = \begin{cases} x^*_i, & \text{if } i \in \mathcal{T} \setminus \{s, t\}; \\ x^*_t, & \text{if } i = s; \\ x^*_s, & \text{if } i = t; \end{cases}$$

that violates inequality (EC.1) (note: $\hat{x} \in \mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ because any extreme point of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ remains an extreme point of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ if we permute its components, and $\hat{x}$ violates (EC.1) because $a > \alpha_t$).

Case (iii): x^* belongs to category 2 of Lemma 1. In this case, at least $\lambda - \lambda'$ components of x^* are equal to $\overline{C}$. Since $|\mathcal{N}_1| \le \lambda - \lambda'$, we have $x^*_t = \overline{C}$ for some $t \in \mathcal{N}_2$. Thus, similar to Case (ii), $x^*_s = \overline{C}$ for all $s \in \mathcal{N}_1$, since otherwise $x^*_s < \overline{C}$ for some $s \in \mathcal{N}_1$, which implies that there exists $\hat{x} \in \mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ with

$$\hat{x}_i = \begin{cases} x^*_i, & \text{if } i \in \mathcal{T} \setminus \{s, t\}; \\ x^*_t, & \text{if } i = s; \\ x^*_s, & \text{if } i = t; \end{cases}$$

that violates inequality (EC.1).

In all three cases, $x^*_s = \overline{C}$ for all $s \in \mathcal{N}_1$. This implies that the facet of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ defined by (EC.1) is one of the facets defined by the trivial inequalities "$x_t \le \overline{C}$" for $t \in \mathcal{N}_1$. This contradicts that (EC.1) is a non-trivial inequality. Hence, $|\mathcal{N}_1| \ge \lambda - \lambda' + 1$.

Now, suppose, on the contrary, that $|\mathcal{N}_1| \ge \lambda + 1$. By Lemma 1, x^* has at most $\lambda + 1$ non-zero components. Thus, $x^*_t = 0$ for all $t \in \mathcal{N}_2$, since otherwise $x^*_t > 0$ for some $t \in \mathcal{N}_2$ and $x^*_s = 0$ for some $s \in \mathcal{N}_1$, which implies that there exists $\hat{x} \in \mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ with

$$\hat{x}_i = \begin{cases} x^*_i, & \text{if } i \in \mathcal{T} \setminus \{s, t\}; \\ x^*_t, & \text{if } i = s; \\ x^*_s, & \text{if } i = t; \end{cases}$$

that violates inequality (EC.1) (because $a > \alpha_t$). This implies that the facet of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ defined by (EC.1) is one of the facets defined by the trivial inequalities "$x_t \ge 0$" for $t \in \mathcal{N}_2$. This contradicts that (EC.1) is a non-trivial inequality. Hence, $\lambda - \lambda' + 1 \le |\mathcal{N}_1| \le \lambda$.

Recall that (EC.1) is an arbitrary non-trivial facet-defining inequality for $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$, and x^* is an arbitrary extreme point of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ satisfying (EC.1) with equality. We show that x^* also satisfies (3) with equality when $\mathcal{T}_1 = \mathcal{N}_1$. We consider two different cases.

Case (i): x^* belongs to category 1 of Lemma 1. In this case, $|\{t \in \mathcal{T} : x^*_t = \overline{C}\}| = \lambda$, since otherwise $x^*_s = 0$ for some $s \in \mathcal{T}$, which implies that there exists $x' \in \mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ with

$$x'_i = \begin{cases} x^*_i, & \text{if } i \in \mathcal{T} \setminus \{s\}; \\ \overline{C}, & \text{if } i = s; \end{cases}$$

that violates inequality (EC.1) (because $a>0$ and $\alpha_t>0$ for all $t\in\mathcal{N}_2$). Suppose that, on the contrary, $x_s^*=0$ for some $s\in\mathcal{N}_1$. Then, because $|\mathcal{N}_1|\le\lambda$, there exists $t\in\mathcal{N}_2$ such that $x_t^*=\overline{C}$. Consider $\hat{x}\in\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ with

$$\hat{x}_i=\begin{cases} x_i^*, & \text{if } i\in\mathcal{T}\setminus\{s,t\};\\ x_t^*, & \text{if } i=s;\\ x_s^*, & \text{if } i=t.\end{cases}$$

Because $a>\alpha_t$, $\hat{x}$ violates inequality (EC.1), which is a contradiction. Hence, $x_s^*=\overline{C}$ for all $s\in\mathcal{N}_1$. Letting $\mathcal{T}_1=\mathcal{N}_1$, we have

$$\sum_{t\in\mathcal{T}_1}x_t^*+\rho\sum_{t\in\mathcal{T}\setminus\mathcal{T}_1}x_t^*=|\mathcal{T}_1|\overline{C}+\rho(\lambda-|\mathcal{T}_1|)\overline{C}=\zeta.$$

Case (ii): x^* belongs to category 2 of Lemma 1. In this case, $x_s^*>x_t^*$ for any $s\in\mathcal{N}_1$ and $t\in\mathcal{N}_2$, since otherwise there exists $\hat{x}\in\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ with

$$\hat{x}_i=\begin{cases} x_i^*, & \text{if } i\in\mathcal{T}\setminus\{s,t\};\\ x_t^*, & \text{if } i=s;\\ x_s^*, & \text{if } i=t;\end{cases}$$

that violates inequality (EC.1) because $a>\alpha_t$. Because $\lambda-\lambda'+1\le|\mathcal{N}_1|\le\lambda$, we have $\overline{C},\theta^*\in\{x_s^*:s\in\mathcal{N}_1\}$ and $0\notin\{x_s^*:s\in\mathcal{N}_1\}$. Letting $\mathcal{T}_1=\mathcal{N}_1$, we have

$$\begin{aligned}\sum_{t\in\mathcal{T}_1}x_t^*+\rho\sum_{t\in\mathcal{T}\setminus\mathcal{T}_1}x_t^*&=(\lambda-\lambda')\overline{C}+\theta^*+(|\mathcal{T}_1|-\lambda+\lambda'-1)\underline{C}+\rho(\lambda+1-|\mathcal{T}_1|)\underline{C}\\&=Q-\lambda'\underline{C}+(|\mathcal{T}_1|-\lambda+\lambda'-1)\underline{C}+\rho(\lambda+1-|\mathcal{T}_1|)\underline{C}\\&=Q-\lambda\overline{C}-\big[\underline{C}-(\lambda-|\mathcal{T}_1|)(\overline{C}-\underline{C})\big]+|\mathcal{T}_1|\overline{C}+\rho\big[\underline{C}-(\lambda-|\mathcal{T}_1|)(\overline{C}-\underline{C})\big]+\rho(\lambda-|\mathcal{T}_1|)\overline{C}\\&=\zeta.\end{aligned}$$

In both cases, x^* satisfies (3) with equality when $\mathcal{T}_1=\mathcal{N}_1$. Hence, by setting $\mathcal{T}_1=\mathcal{N}_1$, any extreme point x^* of $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ satisfying (EC.1) with equality also satisfies (3) with equality. By Property 1, any non-trivial facet-defining inequality for $\mathrm{conv}(\mathcal{D}_{\mathrm{sc}})$ of the form (EC.1) is equivalent to (3). This completes the proof of the proposition. □

A.6. Proof of Proposition 5

Consider any given LP relaxation optimum (x^*,y^*,u^*) of Problem (1) with $x^*=(x_1^*,\ldots,x_T^*)$. Let $(\pi(1),\ldots,\pi(T))$ be a permutation of $(1,\ldots,T)$ such that $x_{\pi(1)}^*\ge\cdots\ge x_{\pi(T)}^*$. Let $\upsilon(|\mathcal{T}_1|)=\zeta-\left(\sum_{t=1}^{|\mathcal{T}_1|}x_{\pi(t)}^*+\rho\sum_{t=|\mathcal{T}_1|+1}^{T}x_{\pi(t)}^*\right)$ for $|\mathcal{T}_1|=\lambda-\lambda'+1,\ldots,\lambda$. Note that $\rho\le1$. Thus, for any given value of $|\mathcal{T}_1|$, $\sum_{t\in\mathcal{T}_1}x_t^*+\rho\sum_{t\in\mathcal{T}\setminus\mathcal{T}_1}x_t^*$ is the largest possible when $\mathcal{T}_1=\{\pi(1),\ldots,\pi(|\mathcal{T}_1|)\}$. Hence, $\upsilon(|\mathcal{T}_1|)$ measures by how much the given solution violates (3). Let $\upsilon^*=\max_{|\mathcal{T}_1|\in[\lambda-\lambda'+1,\lambda]}\{\upsilon(|\mathcal{T}_1|)\}$. If $\upsilon^*<0$, then υ^* is the largest possible violation of inequality (3). The permutation $(\pi(1),\ldots,\pi(T))$ can be obtained in $O(T\log T)$ time. The values of $\upsilon(\lambda-\lambda'+1),\ldots,\upsilon(\lambda)$ can be obtained in $O(T)$ time. Therefore, the most violated inequality, if exists, can be identified in $O(T\log T)$ time. □

A.7. Proof of Proposition 6

Define

$$\psi_0(s) = \underline{C}(2s+1) + s^2 V$$

for $s \in \mathbb{R}$. Let s_0 be the positive root of the quadratic equation "$\psi_0(s) = Q$"; that is,

$$s_0 = \frac{1}{V}\left[\sqrt{\underline{C}^2 + V(Q - \underline{C})} - \underline{C}\right] > 0.$$

Define

$$\psi_1(s) = \underline{C}(s + T - t + 1) + \left[\frac{s(s+1) - (T-t)(T-t+1)}{2} + s(T-t)\right]V$$

for $s \in \mathbb{R}$. Let s_1 be the positive root of the quadratic equation "$\psi_1(s) = Q$"; that is,

$$s_1 = \frac{1}{V}\left[\sqrt{\left[\underline{C} + \frac{V}{2} + (T-t)V\right]^2 + (T-t+1)V\left[(T-t)V - 2\underline{C}\right] + 2QV} - \left[\underline{C} + \frac{V}{2} + (T-t)V\right]\right] > 0.$$

Define

$$\psi_2(s) = \underline{C}(t+s) + \left[\frac{s(s+1) - (t-1)t}{2} + (t-1)s\right]V$$

for $s \in \mathbb{R}$. Let s_2 be the positive root of the quadratic equation "$\psi_2(s) = Q$"; that is,

$$s_2 = \frac{1}{V}\left[\sqrt{\left[\underline{C} + \frac{V}{2} + (t-1)V\right]^2 + tV\left[(t-1)V - 2\underline{C}\right] + 2QV} - \left[\underline{C} + \frac{V}{2} + (t-1)V\right]\right] > 0.$$

We now consider α_t. Consider the case where $Q < Q_{B1}$. In this case, $\omega_1(n,n) = \underline{C}(2n+1) + n^2 V = \psi_0(n)$, which is increasing in n when $n \geq 0$. Thus, $\alpha_t = \max\{n \in \mathbb{Z} : \psi_0(n) \leq Q\} = \lfloor s_0 \rfloor$. Now, consider the case where $Q_{B1} \leq Q < Q_{B2}$ and $t - 2 \geq T - t$. Note that in this case $\omega_2(t-1, T-t) = Q_{B2} > Q$. Note also that because $t - 2 \geq T - t$, we have

$$\omega_2(T-t, T-t) = \omega_1(T-t, T-t) = \omega_1(\min\{t-1, T-t\}, \min\{t-1, T-t\}) = Q_{B1} \leq Q.$$

It is easy to see that $\omega_2(n, T-t)$ is increasing in n when $n \geq T - t$. Thus,

$$\alpha_t = \max\left\{n \in \mathbb{Z} : T - t \leq n < t - 1 \text{ and } \omega_2(n, T-t) \leq Q\right\}. \tag{EC.2}$$

When $T - t \leq n < t - 1$,

$$\omega_2(n, T-t) = \underline{C}(n + T - t + 1) + \left[\frac{n(n+1) - (T-t)(T-t+1)}{2} + n(T-t)\right]V = \psi_1(n).$$

Hence, in this case, $\alpha_t = \lfloor s_1 \rfloor$. Thus,

$$\alpha_t = \begin{cases} \lfloor s_0 \rfloor, & \text{if } Q < Q_{B1}; \\ \lfloor s_1 \rfloor, & \text{if } Q_{B1} \leq Q < Q_{B2} \text{ and } t - 2 \geq T - t; \\ t - 1, & \text{otherwise.} \end{cases}$$

Therefore, $\alpha_t \geq 0$, and α_t can be determined in $\mathcal{O}(1)$ time.

Next, we consider β_t. If $Q < Q_{B1}$, then $\beta_t = \alpha_t = \lfloor s_0 \rfloor$. Consider the case where $Q_{B1} \leq Q < Q_{B2}$ and $t \leq T - t$. Note that in this case $\omega_2(t-1, T-t) = Q_{B2} > Q$. Note also that because $t \leq T - t$, we have

$$\omega_2(t-1, t-1) = \omega_1(t-1, t-1) = \omega_1(\min\{t-1, T-t\}, \min\{t-1, T-t\}) = Q_{B1} \leq Q.$$

It is easy to see that $\omega_2(t-1,n)$ is increasing in n when $n \geq t-1$. Thus,

$$\beta_t = \max\left\{n \in \mathbb{Z} : t-1 \leq n < T-t \text{ and } \omega_2(t-1,n) \leq Q\right\}. \tag{EC.3}$$

When $t-1 \leq n < T-t$,

$$\omega_2(t-1,n) = \underline{C}(t+n) + \left[\frac{n(n+1)-(t-1)t}{2} + (t-1)n\right]V = \psi_2(n).$$

Hence, in this case, $\beta_t = \lfloor s_2 \rfloor$. Thus,

$$\beta_t = \begin{cases} \lfloor s_0 \rfloor, & \text{if } Q < Q_{B1}; \\ \lfloor s_2 \rfloor, & \text{if } Q_{B1} \leq Q < Q_{B2} \text{ and } t \leq T-t; \\ T-t, & \text{otherwise.} \end{cases}$$

Therefore, $\beta_t \geq 0$, and β_t can be determined in $\mathcal{O}(1)$ time.

Obviously, once α_t and β_t are determined, τ_{t1} and τ_{t2} can be obtained in $\mathcal{O}(1)$ time. This completes the proof of the proposition. $\square$

A.8. Proof of Proposition 7

(i) Consider the case where $Q < Q_{B1}$. The validity of the equation "$\alpha_t = \beta_t$" follows directly from the definitions of α_t and β_t. By Definition 2, $\omega_1(\alpha_t,\alpha_t) \leq Q < Q_{B1} = \omega_1(\min\{t-1,T-t\},\min\{t-1,T-t\})$. Because $\omega_1(n,n)$ is strictly increasing in n, this implies that $\alpha_t < \min\{t-1,T-t\}$. Thus, $\alpha_t \neq t-1$. Because $\alpha_t = \beta_t$, we have $\beta_t \neq T-t$. Hence, by the definitions of τ_{t1} and τ_{t2}, we have $\tau_{t1} = \tau_{t2}$.

(ii) Consider the case where $Q_{B1} \leq Q < Q_{B2}$. We first show that $t-1 \neq T-t$. By contradiction, suppose $t-1 = T-t$. Then, $Q_{B1} = \omega_1(\min\{t-1,T-t\},\min\{t-1,T-t\}) = \omega_1(t-1,t-1) = \omega_2(t-1,t-1) = \omega_2(t-1,T-t) = Q_{B2}$, which contradicts the condition "$Q_{B1} \leq Q < Q_{B2}$." Thus, either $t-2 \geq T-t$ or $t \leq T-t$. If $t-2 \geq T-t$, then $\omega_2(\alpha_t,T-t) \leq Q$ (by the definition of α_t) and $\beta_t = T-t$ (by the definition of β_t), which together imply that $Q \geq \omega_2(\alpha_t,\beta_t)$. If $t \leq T-t$, then $\alpha_t = t-1$ (by the definition of α_t) and $\omega_2(t-1,\beta_t) \leq Q$ (by the definition of β_t), which together imply that $Q \geq \omega_2(\alpha_t,\beta_t)$. Hence, in both cases, $Q \geq \omega_2(\alpha_t,\beta_t)$. If $t-2 \geq T-t$, then because $\beta_t = T-t$, we have $\tau_{t2} = 0$ (by the definition of τ_{t2}). If $t \leq T-t$, then because $\alpha_t = t-1$, we have $\tau_{t1} = 0$ (by the definition of τ_{t1}). Hence, we have $\tau_{t1} = 0$ or $\tau_{t2} = 0$.

(iii) We first show that $\alpha_t + \tau_{t1} \leq t-1$. Consider the case where $Q < Q_{B1}$. In this case, by the definition of α_t, $\omega_1(\alpha_t,\alpha_t) \leq Q < Q_{B1} = \omega_1(\min\{t-1,T-t\},\min\{t-1,T-t\})$. Note that $\omega_1(n,n) = (2n+1)\underline{C} + n^2V$, which is increasing in n when $n \geq 0$. Thus, $\alpha_t < \min\{t-1,T-t\} \leq t-1$, which implies that $\alpha_t + \tau_{t1} \leq t-1$. Next, consider the case where $Q_{B1} \leq Q < Q_{B2}$ and $t-2 \geq T-t$. In this case, by (EC.2) in the Proof of Proposition 6, we have $\alpha_t < t-1$, which implies that $\alpha_t + \tau_{t1} \leq t-1$. Finally, consider the case where the conditions "$Q < Q_{B1}$" and "$Q_{B1} \leq Q < Q_{B2}$ and $t-2 \geq T-t$" do not hold. In this case, by Definition 2, $\alpha_t = t-1$ and $\tau_{t1} = 0$, which implies that $\alpha_t + \tau_{t1} \leq t-1$.

Next, we show that $\beta_t + \tau_{t2} \leq T-t$. Consider the case where $Q < Q_{B1}$. In this case, by the definition of β_t, $\omega_1(\beta_t,\beta_t) \leq Q < Q_{B1} = \omega_1(\min\{t-1,T-t\},\min\{t-1,T-t\})$. Since $\omega_1(n,n)$ is increasing in n when $n \geq 0$, we have $\beta_t < \min\{t-1,T-t\} \leq T-t$, which implies that $\beta_t + \tau_{t2} \leq T-t$. Next, consider the case where $Q_{B1} \leq Q < Q_{B2}$ and $t \leq T-t$. In this case, by (EC.3) in the Proof of Proposition 6, $\beta_t < t-1$, which implies that $\beta_t + \tau_{t2} \leq T-t$. Finally, consider the case where the conditions "$Q < Q_{B1}$" and "$Q_{B1} \leq Q < Q_{B2}$ and $t \leq T-t$" do not hold. In this case, by Definition 2, $\beta_t = T-t$ and $\tau_{t2} = 0$, which implies that $\beta_t + \tau_{t2} \leq T-t$. $\square$

A.9. Proof of Proposition 8

We show that inequality (4) holds for any element of $\mathcal{D}$. To do so, we first prove that for any $t \in \mathcal{T}$ and any element of $\mathcal{D}$,

$$\frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1} \geq (\max\{\alpha_t, \beta_t\} - 1)V + \overline{V} \tag{EC.4}$$

and

$$x_t \leq \frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1}. \tag{EC.5}$$

To prove inequality (EC.4), we first note that for any $t \in \mathcal{T}$,

$$\begin{aligned}
&\left[\frac{|\alpha_t(\alpha_t+1) - \beta_t(\beta_t+1)|}{2} + \alpha_t\beta_t\right] V + \Delta(\alpha_t, \beta_t)V \\
&= \left[\frac{|\alpha_t(\alpha_t+1) - \beta_t(\beta_t+1)|}{2} + \alpha_t\beta_t\right] V + \left[\frac{\alpha_t(\alpha_t+1)}{2} + \frac{\beta_t(\beta_t+1)}{2}\right] V \\
&= [\max\{\alpha_t(\alpha_t+1), \beta_t(\beta_t+1)\} + \alpha_t\beta_t]\, V \\
&= (\alpha_t + \beta_t + 1)\max\{\alpha_t, \beta_t\} V.
\end{aligned} \tag{EC.6}$$

Then, we consider three different cases.

Case 1: $Q < Q_{B1}$. In this case, by Proposition 7, $\alpha_t = \beta_t$ and $\tau_{t1} = \tau_{t2}$. By the definition of α_t, $Q \geq \omega_1(\alpha_t, \alpha_t)$. If $\tau_{t1} = \tau_{t2} = 0$, then $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} = Q + \Delta(\alpha_t, \alpha_t)V \geq \omega_1(\alpha_t, \alpha_t) + \Delta(\alpha_t, \alpha_t)V = \underline{C}(2\alpha_t + 1) + [\alpha_t(\alpha_t - 1) + \alpha_t]V + \Delta(\alpha_t, \alpha_t)V = \underline{C}(2\alpha_t + 1) + (2\alpha_t^2 + \alpha_t)V = (\underline{C} + V)(2\alpha_t + 1) + (2\alpha_t + 1)(\alpha_t - 1)V > \overline{V}(2\alpha_t + 1) + (2\alpha_t + 1)(\alpha_t - 1)V$. If $\tau_{t1} = \tau_{t2} = 1$, then by the definitions of τ_{t1} and τ_{t2}, $Q \geq \omega_3(\alpha_t, \beta_t)$, and thus $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} \geq \omega_3(\alpha_t, \beta_t) + \Delta(\alpha_t, \beta_t)V - 2\underline{C} = \omega_3(\alpha_t, \alpha_t) + \Delta(\alpha_t, \alpha_t)V - 2\underline{C} = \overline{V}(2\alpha_t + 1) + [\alpha_t(\alpha_t - 1) + \alpha_t]V + \alpha_t(\alpha_t + 1)V > \overline{V}(2\alpha_t + 1) + (2\alpha_t + 1)(\alpha_t - 1)V$. Hence, in both scenarios, $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} > \overline{V}(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)(\max\{\alpha_t, \beta_t\} - 1)V$.

Case 2: $Q_{B1} \leq Q < Q_{B2}$. In this case, by Proposition 7, $Q \geq \omega_2(\alpha_t, \beta_t)$, and $\tau_{t1} + \tau_{t2}$ equals either 0 or 1. If $\tau_{t1} + \tau_{t2} = 0$, then $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} \geq \omega_2(\alpha_t, \beta_t) + \Delta(\alpha_t, \beta_t)V = \underline{C}(\alpha_t + \beta_t + 1) + [|\alpha_t(\alpha_t + 1) - \beta_t(\beta_t + 1)|/2 + \alpha_t\beta_t]V + \Delta(\alpha_t, \beta_t)V = (\underline{C} + V)(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)(\max\{\alpha_t, \beta_t\} - 1)V$, where the last equality follows from (EC.6). Thus, $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} > \overline{V}(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)(\max\{\alpha_t, \beta_t\} - 1)V$. If $\tau_{t1} + \tau_{t2} = 1$, then by the definitions of τ_{t1} and τ_{t2}, $Q \geq \omega_4(\alpha_t, \beta_t)$, and thus $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} \geq \omega_4(\alpha_t, \beta_t) + \Delta(\alpha_t, \beta_t)V - \underline{C} = \overline{V}(\alpha_t + \beta_t + 1) + [|\alpha_t(\alpha_t + 1) - \beta_t(\beta_t + 1)|/2 + \alpha_t\beta_t]V + \Delta(\alpha_t, \beta_t)V = \overline{V}(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)\max\{\alpha_t, \beta_t\}V$, where the last equality follows from (EC.6). Hence, in both scenarios, $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} > \overline{V}(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)(\max\{\alpha_t, \beta_t\} - 1)V$.

Case 3: $Q_{B2} \leq Q$. In this case, by Definition 2, $\alpha_t = t - 1$, $\beta_t = T - 1$, and $\tau_{t1} = \tau_{t2} = 0$. Hence, $Q \geq Q_{B2} = \omega_2(\alpha_t, \beta_t)$. Thus, $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} \geq \omega_2(\alpha_t, \beta_t) + \Delta(\alpha_t, \beta_t)V = \underline{C}(\alpha_t + \beta_t + 1) + [|\alpha_t(\alpha_t + 1) - \beta_t(\beta_t + 1)|/2 + \alpha_t\beta_t]V + \Delta(\alpha_t, \beta_t)V = (\underline{C} + V)(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)(\max\{\alpha_t, \beta_t\} - 1)V$, where the last equality follows from (EC.6). Hence, $Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} \geq \overline{V}(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)(\max\{\alpha_t, \beta_t\} - 1)V$.

In all three cases,

$$Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C} \geq \overline{V}(\alpha_t + \beta_t + 1) + (\alpha_t + \beta_t + 1)(\max\{\alpha_t, \beta_t\} - 1)V.$$

Thus, inequality (EC.4) holds.

To prove inequality (EC.5), we define

$\hat{\mathcal{D}} = \{(x,y,u) \in \mathbb{R}^T \times \{0,1\}^T \times \{0,1\}^{T-1} : (x,y,u)$ satisfies (2c), (2d), (2f), (2g), (2h), and $x_s \le Qy_s\ \forall s \in \mathcal{T}\}$

which is a superset of $\mathcal{D}$, and we first show that for any $t \in \mathcal{T}$, there exists $(\hat{x},\hat{y},\hat{u}) \in \hat{\mathcal{D}}$ such that $\hat{x}_t = \max\{x_t : (x,y,u) \in \hat{\mathcal{D}}\}$ and $\hat{y}_s = 1$ for all $s \in [t-\alpha_t-\tau_{t1}, t+\beta_t+\tau_{t2}]_{\mathbb{Z}}$ (note: by Proposition 7(iii), $\alpha_t+\tau_{t1} \le t-1$ and $\beta_t+\tau_{t2} \le T-t$, and thus $[t-\alpha_t-\tau_{t1}, t+\beta_t+\tau_{t2}]_{\mathbb{Z}} \subseteq \mathcal{T}$). Let (x^*,y^*,u^*) be any element of $\hat{\mathcal{D}}$ with $x_t^* = \max\{x_t : (x,y,u) \in \hat{\mathcal{D}}\}$. If $y_s^* = 1$ for all $s \in [t-\alpha_t-\tau_{t1}, t+\beta_t+\tau_{t2}]_{\mathbb{Z}}$, then $(\hat{x},\hat{y},\hat{u})$ exists. Thus, it suffices to consider the situation where $y_r^* = 0$ for some $r \in [t-\alpha_t-\tau_{t1}, t+\beta_t+\tau_{t2}]_{\mathbb{Z}}$. For the case where $r < t$, we let r' be the largest integer such that $t-\alpha_t-\tau_{t1} \le r' < t$ and $y_{r'}^* = 0$. Hence, $y_{r'+1}^* = \cdots = y_{t-1}^* = 1$ and $x_{r'}^* = 0$. By (2f), $\sum_{s=r'+1}^{t}(x_s^* - x_{s-1}^*) \le \sum_{s=r'+1}^{t}(Vy_{s-1}^* + \overline{V}(1-y_{s-1}^*))$, which implies that $x_t^* \le \overline{V} + (t-r'-1)V$. For the case where $r > t$, we let r' be the smallest integer such that $t < r' \le t+\beta_t+\tau_{t2}$ and $y_{r'}^* = 0$. Hence, $y_{t+1}^* = \cdots = y_{r'-1}^* = 1$ and $x_{r'}^* = 0$. By (2g), $\sum_{s=t+1}^{r'}(x_{s-1}^* - x_s^*) \le \sum_{s=t+1}^{r'}(Vy_s^* + \overline{V}(1-y_s^*))$, which implies that $x_t^* \le \overline{V} + (r'-t-1)V$. For the case where $r = t$, we have $x_t^* = 0 \le \overline{V} - V$, and we let $r' = t$. Thus, in all three cases, $x_t^* \le \overline{V} + (|r'-t|-1)V$ and $|r'-t| \le \max\{\alpha_t+\tau_{t1}, \beta_t+\tau_{t2}\}$. Therefore,

$$x_t^* \le \overline{V} + (\max\{\alpha_t+\tau_{t1}, \beta_t+\tau_{t2}\} - 1)V. \tag{EC.7}$$

Consider $(\hat{x},\hat{y},\hat{u}) \in \mathbb{R}^T \times \{0,1\}^T \times \{0,1\}^{T-1}$, where

$$\hat{x}_s = \begin{cases} \overline{V} + (\max\{\alpha_t+\tau_{t1}, \beta_t+\tau_{t2}\} - 1 - |s-t|)V, & \text{if } |s-t| \le \max\{\alpha_t+\tau_{t1}, \beta_t+\tau_{t2}\} - 1; \\ \underline{C}, & \text{if } |s-t| = \max\{\alpha_t+\tau_{t1}, \beta_t+\tau_{t2}\}; \\ 0, & \text{if } |s-t| \ge \max\{\alpha_t+\tau_{t1}, \beta_t+\tau_{t2}\} + 1; \end{cases}$$

$$\hat{y}_s = \begin{cases} 1, & \text{if } s \in [t-\alpha_t-\tau_{t1}, t+\beta_t+\tau_{t2}]_{\mathbb{Z}}; \\ 0, & \text{otherwise}; \end{cases}$$

$$\hat{u}_s = \begin{cases} 1, & \text{if } s = t-\alpha_t-\tau_{t1}; \\ 0, & \text{otherwise}. \end{cases}$$

It is easy to see that $(\hat{x},\hat{y},\hat{u})$ satisfies (2c), (2d), and that for any $s \in \mathcal{T}$, $\hat{x}_s$ is non-zero only when $\hat{y}_s = 1$. In the following, we show that $(\hat{x},\hat{y},\hat{u})$ also satisfies (2f), (2g), and (2h) (which implies that $(\hat{x},\hat{y},\hat{u}) \in \hat{\mathcal{D}}$). We consider four different cases.

Case 1: $Q < Q_{B1}$. In this case, by Proposition 7(i), $\alpha_t = \beta_t$ and $\tau_{t1} = \tau_{t2}$. We consider two subcases. Case 1.1: $\tau_{t1} = \tau_{t2} = 0$. In this subcase, $\hat{x}_{t-\alpha_t} = \hat{x}_{t+\beta_t} = \underline{C}$ and $|\hat{x}_s - \hat{x}_{s-1}| \le V$ for $s \in [t-\alpha_t+1, t+\beta_t]_{\mathbb{Z}}$, which imply that $(\hat{x},\hat{y},\hat{u})$ satisfies (2f) and (2g). Furthermore, if $\alpha_t = \beta_t = 0$, then $\sum_{s=1}^{T}\hat{x}_s = \underline{C} \le Q$. Otherwise,

$$\begin{aligned}
\sum_{s=1}^{T}\hat{x}_s &= 2\underline{C} + (\alpha_t+\beta_t-1)\overline{V} + \sum_{s=t-\alpha_t+1}^{t-1}(\alpha_t-1-t+s)V + \sum_{s=t}^{t+\beta_t-1}(\beta_t-1+t-s)V \\
&= 2\underline{C} + (\alpha_t+\beta_t-1)\overline{V} + \left[\frac{(\alpha_t-2)(\alpha_t-1)}{2} + \frac{(\beta_t-2)(\beta_t-1)}{2} + (\beta_t-1)\right]V \\
&\le 2\underline{C} + (\alpha_t+\beta_t-1)(\underline{C}+V) + \left[\frac{(\alpha_t-2)(\alpha_t-1)}{2} + \frac{(\beta_t-2)(\beta_t-1)}{2} + (\beta_t-1)\right]V \\
&= \omega_1(\alpha_t,\alpha_t) \le Q,
\end{aligned}$$

where the last inequality follows from the definition of α_t. Case 1.2: $\tau_{t1} = \tau_{t2} = 1$. In this subcase, $\hat{x}_{t-\alpha_t-1} = \hat{x}_{t+\beta_t+1} = \underline{C}$ and $|\hat{x}_s - \hat{x}_{s-1}| \le V$ for $s \in [t-\alpha_t, t+\beta_t+1]_{\mathbb{Z}}$, which imply that $(\hat{x},\hat{y},\hat{u})$ satisfies (2f) and (2g). Furthermore,

$$\sum_{s=1}^{T}\hat{x}_s = 2\underline{C} + (\alpha_t+\beta_t+1)\overline{V} + \sum_{s=t-\alpha_t}^{t-1}(\alpha_t-t+s)V + \sum_{s=t}^{t+\beta_t}(\beta_t+t-s)V$$

$$= 2\underline{C} + (\alpha_t + \beta_t + 1)\overline{V} + \left[\frac{(\alpha_t - 1)\alpha_t}{2} + \frac{(\beta_t - 1)\beta_t}{2} + \beta_t\right] V$$
$$= \omega_3(\alpha_t, \beta_t) \le Q,$$

where the inequality follows from the condition "$\tau_{t1} = \tau_{t2} = 1$" and the definitions of τ_{t1} and τ_{t2}. Hence, in both Cases 1.1 and 1.2, $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2h).

Case 2: $Q_{B1} \le Q < Q_{B2}$ and $t - 2 \ge T - t$. In this case, by Definition 2, $\alpha_t = \max\{n \in \mathbb{Z} : \omega_2(n, T - t) \le Q\}$ and $\beta_t = T - t$. By Proposition 7(iii), $\tau_{t2} = 0$. Note that $\omega_2(T - t, T - t) = \omega_1(T - t, T - t) = Q_{B1} \le Q$ and $\omega_2(t - 1, T - t) = Q_{B2} > Q$. Note also that $\omega_2(n, T - t)$ increases as n increases when $n > T - t$. Thus, by the definition of α_t, $T - t \le \alpha_t < t - 1$. Hence, $\alpha_t + \tau_{t1} \ge \beta_t + \tau_{t2}$. In this case, $\hat{x}_{t-\alpha_t-\tau_{t1}} = \underline{C}$ and $|\hat{x}_s - \hat{x}_{s-1}| \le V$ for $s \in [t - \alpha_t - \tau_{t1} + 1, T]_{\mathbb{Z}}$, which together imply that $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2f) and (2g). We consider three subcases. Case 2.1: $\tau_{t1} = 1$. In this subcase, $\alpha_t + \tau_{t1} > \beta_t + \tau_{t2}$. We have

$$\sum_{s=1}^{T} \hat{x}_s = \underline{C} + (\alpha_t + \beta_t + 1)\overline{V} + \sum_{s=t-\alpha_t}^{t} (\alpha_t - t + s)V + \sum_{s=t+1}^{t+\beta_t} (\alpha_t + t - s)V$$
$$= \underline{C} + (\alpha_t + \beta_t + 1)\overline{V} + \left[\frac{\alpha_t(\alpha_t + 1)}{2} - \frac{\beta_t(\beta_t + 1)}{2} + \alpha_t\beta_t\right] V$$
$$= \omega_4(\alpha_t, \beta_t) \le Q,$$

where the inequality follows from the definition of τ_{t1}. Case 2.2: $\tau_{t1} = 0$ and $\alpha_t = \beta_t$. In this subcase, $\alpha_t + \tau_{t1} = \beta_t + \tau_{t2}$. We have

$$\sum_{s=1}^{T} \hat{x}_s = 2\underline{C} + (2\beta_t - 1)\overline{V} + \sum_{s=t-\beta_t+1}^{t-1} (\beta_t - 1 - t + s)V + \sum_{s=t}^{t+\beta_t-1} (\beta_t - 1 + t - s)V$$
$$= 2\underline{C} + (2\beta_t - 1)\overline{V} + \left[\frac{(\beta_t - 2)(\beta_t - 1)}{2} + \frac{(\beta_t - 2)(\beta_t - 1)}{2} + (\beta_t - 1)\right] V$$
$$\le 2\underline{C} + (2\beta_t - 1)(\underline{C} + V) + \left[\frac{(\beta_t - 2)(\beta_t - 1)}{2} + \frac{(\beta_t - 2)(\beta_t - 1)}{2} + (\beta_t - 1)\right] V$$
$$= \omega_1(\beta_t, \beta_t) = \omega_1(T - t, T - t) = Q_{B1} \le Q.$$

Case 2.3: $\tau_{t1} = 0$ and $\alpha_t > \beta_t$. In this subcase, $\alpha_t + \tau_{t1} > \beta_t + \tau_{t2}$. We have

$$\sum_{s=1}^{T} \hat{x}_s = \underline{C} + (\alpha_t + \beta_t)\overline{V} + \sum_{s=t-\alpha_t+1}^{t} (\alpha_t - 1 - t + s)V + \sum_{s=t+1}^{t+\beta_t} (\alpha_t - 1 + t - s)V$$
$$= \underline{C} + (\alpha_t + \beta_t)\overline{V} + \left[\frac{(\alpha_t - 1)\alpha_t}{2} - \frac{\beta_t(\beta_t + 1)}{2} + (\alpha_t - 1)\beta_t\right] V$$
$$\le \underline{C} + (\alpha_t + \beta_t)(\underline{C} + V) + \left[\frac{(\alpha_t - 1)\alpha_t}{2} - \frac{\beta_t(\beta_t + 1)}{2} + (\alpha_t - 1)\beta_t\right] V$$
$$= \omega_2(\alpha_t, \beta_t) \le Q,$$

where the last inequality follows from Proposition 7(ii). Hence, in Cases 2.1–2.3, $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2h).

Case 3: $Q_{B1} \le Q < Q_{B2}$ and $t \le T - t$. In this case, by Definition 2, $\alpha_t = t - 1$ and $\beta_t = \max\{n \in \mathbb{Z} : \omega_2(t - 1, n) \le Q\}$. By Proposition 7(iii), $\tau_{t1} = 0$. Note that $\omega_2(t - 1, t - 1) = \omega_1(t - 1, t - 1) = Q_{B1} \le Q$ and $\omega_2(t - 1, T - t) = Q_{B2} > Q$. Note also that $\omega_2(t - 1, n)$ increases as n increases when $n > t - 1$. Thus, by the definition of β_t, $t - 1 \le \beta_t < T - t$. Hence, $\beta_t + \tau_{t2} \ge \alpha_t + \tau_{t1}$. In this case, $\hat{x}_{t+\beta_t+\tau_{t2}} = \underline{C}$ and $|\hat{x}_s - \hat{x}_{s-1}| \le V$

for $s \in [2, t+\beta_t+\tau_{t2}]_{\mathbb{Z}}$, which together imply that $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2f) and (2g). We consider three subcases. Case 3.1: $\tau_{t2} = 1$. In this subcase, $\beta_t + \tau_{t2} > \alpha_t + \tau_{t1}$. We have

$$\begin{aligned}\sum_{s=1}^{T} \hat{x}_s &= \underline{C} + (\alpha_t + \beta_t + 1)\overline{V} + \sum_{s=t-\alpha_t}^{t-1} (\beta_t - t + s)V + \sum_{s=t}^{t+\beta_t} (\beta_t + t - s)V \\ &= \underline{C} + (\alpha_t + \beta_t + 1)\overline{V} + \left[\frac{\beta_t(\beta_t+1)}{2} - \frac{\alpha_t(\alpha_t+1)}{2} + \alpha_t\beta_t\right]V \\ &= \omega_4(\alpha_t, \beta_t) \le Q,\end{aligned}$$

where the inequality follows from the definition of τ_{t2}. Case 3.2: $\tau_{t2} = 0$ and $\beta_t = \alpha_t$. In this subcase, $\beta_t + \tau_{t2} = \alpha_t + \tau_{t1}$. We have

$$\begin{aligned}\sum_{s=1}^{T} \hat{x}_s &= 2\underline{C} + (2\alpha_t - 1)\overline{V} + \sum_{s=t-\alpha_t+1}^{t-1} (\alpha_t - 1 - t + s)V + \sum_{s=t}^{t+\alpha_t-1} (\alpha_t - 1 + t - s)V \\ &= 2\underline{C} + (2\alpha_t - 1)\overline{V} + \left[\frac{(\alpha_t-2)(\alpha_t-1)}{2} + \frac{(\alpha_t-2)(\alpha_t-1)}{2} + (\alpha_t - 1)\right]V \\ &\le 2\underline{C} + (2\alpha_t - 1)(\underline{C} + V) + \left[\frac{(\alpha_t-2)(\alpha_t-1)}{2} + \frac{(\alpha_t-2)(\beta_t-1)}{2} + (\alpha_t - 1)\right]V \\ &= \omega_1(\alpha_t, \alpha_t) = \omega_1(t-1, t-1) = Q_{B1} \le Q.\end{aligned}$$

Case 3.3: $\tau_{t2} = 0$ and $\beta_t > \alpha_t$. In this subcase, $\beta_t + \tau_{t2} > \alpha_t + \tau_{t1}$. We have

$$\begin{aligned}\sum_{s=1}^{T} \hat{x}_s &= \underline{C} + (\alpha_t + \beta_t)\overline{V} + \sum_{s=t-\alpha_t}^{t-1} (\beta_t - 1 - t + s)V + \sum_{s=t}^{t+\beta_t-1} (\beta_t - 1 + t - s)V \\ &= \underline{C} + (\alpha_t + \beta_t)\overline{V} + \left[\frac{(\beta_t-1)\beta_t}{2} - \frac{\alpha_t(\alpha_t+1)}{2} + \alpha_t(\beta_t - 1)\right]V \\ &\le \underline{C} + (\alpha_t + \beta_t)(\underline{C} + V) + \left[\frac{(\beta_t-1)\beta_t}{2} - \frac{\alpha_t(\alpha_t+1)}{2} + \alpha_t(\beta_t - 1)\right]V \\ &= \omega_2(\alpha_t, \beta_t) \le Q,\end{aligned}$$

where the last inequality follows from Proposition 7(ii). Hence, in Cases 3.1–3.3, $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2h).

Case 4: ($Q_{B1} \le Q < Q_{B2}$ and $t - 1 = T - t$) or $Q \ge Q_{B2}$. By Definition 2, $\alpha_t = t - 1$ and $\beta_t = T - t$. By Proposition 7(iii), $\tau_{t1} = \tau_{t2} = 0$. In this case, $|\hat{x}_s - \hat{x}_{s-1}| \le V$ for $s \in [2, T]_{\mathbb{Z}}$, which together imply that $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2f) and (2g). We consider three subcases. Case 4.1: $\alpha_t < \beta_t$. In this subcase, $\beta_t + \tau_{t2} > \alpha_t + \tau_{t1}$. We have

$$\begin{aligned}\sum_{s=1}^{T} \hat{x}_s &= \underline{C} + (\alpha_t + \beta_t)\overline{V} + \sum_{s=t-\alpha_t}^{t-1} (\beta_t - 1 - t + s)V + \sum_{s=t}^{t+\beta_t-1} (\beta_t - 1 + t - s)V \\ &= \underline{C} + (\alpha_t + \beta_t)\overline{V} + \left[\frac{(\beta_t-1)\beta_t}{2} - \frac{\alpha_t(\alpha_t+1)}{2} + \alpha_t(\beta_t - 1)\right]V \\ &\le \underline{C} + (\alpha_t + \beta_t)(\underline{C} + V) + \left[\frac{(\beta_t-1)\beta_t}{2} - \frac{\alpha_t(\alpha_t+1)}{2} + \alpha_t(\beta_t - 1)\right]V \\ &= \omega_2(\alpha_t, \beta_t).\end{aligned}$$

Case 4.2: $\alpha_t = \beta_t$. In this subcase, $\beta_t + \tau_{t2} = \alpha_t + \tau_{t1}$. We have

$$\sum_{s=1}^{T} \hat{x}_s = 2\underline{C} + (2\beta_t - 1)\overline{V} + \sum_{s=t-\beta_t+1}^{t-1} (\beta_t - 1 - t + s)V + \sum_{s=t}^{t+\beta_t-1} (\beta_t - 1 + t - s)V$$

$$
\begin{aligned}
&= 2\underline{C} + (2\beta_t - 1)\overline{V} + \left[\frac{(\beta_t-2)(\beta_t-1)}{2} + \frac{(\beta_t-2)(\beta_t-1)}{2} + (\beta_t - 1)\right] V \\
&\le 2\underline{C} + (2\beta_t - 1)(\underline{C} + V) + \left[\frac{(\beta_t-2)(\beta_t-1)}{2} + \frac{(\beta_t-2)(\beta_t-1)}{2} + (\beta_t - 1)\right] V \\
&= \omega_1(\beta_t, \beta_t) = \omega_2(\alpha_t, \beta_t).
\end{aligned}
$$

Case 4.3: $\alpha_t > \beta_t$. In this subcase, $\alpha_t + \tau_{t1} > \beta_t + \tau_{t2}$. We have

$$
\begin{aligned}
\sum_{s=1}^{T} \hat{x}_s &= \underline{C} + (\alpha_t + \beta_t)\overline{V} + \sum_{s=t-\alpha_t+1}^{t} (\alpha_t - 1 - t + s)V + \sum_{s=t+1}^{t+\beta_t} (\alpha_t - 1 + t - s)V \\
&= \underline{C} + (\alpha_t + \beta_t)\overline{V} + \left[\frac{(\alpha_t - 1)\alpha_t}{2} - \frac{\beta_t(\beta_t+1)}{2} + (\alpha_t - 1)\beta_t\right] V \\
&\le \underline{C} + (\alpha_t + \beta_t)(\underline{C} + V) + \left[\frac{(\alpha_t - 1)\alpha_t}{2} - \frac{\beta_t(\beta_t+1)}{2} + (\alpha_t - 1)\beta_t\right] V \\
&= \omega_2(\alpha_t, \beta_t).
\end{aligned}
$$

Thus, in Cases 4.1–4.3, $\sum_{s=1}^{T} \hat{x}_s \le \omega_2(\alpha_t, \beta_t)$. If $Q_{B1} \le Q < Q_{B2}$ and $t - 1 = T - t$, then by Proposition 7(ii), $\omega_2(\alpha_t, \beta_t) \le Q$. If $Q \ge Q_{B2}$, then $\omega_2(\alpha_t, \beta_t) = \omega_2(t-1, T-t) = Q_{B2} \le Q$. Hence, in both scenarios, $\sum_{s=1}^{T} \hat{x}_s \le Q$. Therefore, $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2h).

In Cases 1–4, $(\hat{x}, \hat{y}, \hat{u})$ satisfies (2f), (2g), and (2h). Hence, $(\hat{x}, \hat{y}, \hat{u}) \in \hat{\mathcal{D}}$. Note that

$$
\hat{x}_t = \begin{cases} \underline{C}, & \text{if } \alpha_t + \tau_{t1} = \beta_t + \tau_{t2} = 0; \\ \overline{V} + (\max\{\alpha_t + \tau_{t1}, \beta_t + \tau_{t2}\} - 1)V, & \text{otherwise.} \end{cases}
$$

Thus, by (EC.7), $\hat{x}_t \ge x_t^*$, which implies that $\hat{x}_t = \max\{x_t : (x, y, u) \in \hat{\mathcal{D}}\}$; that is, $(\hat{x}, \hat{y}, \hat{u})$ is an element of $\hat{\mathcal{D}}$ where the x_t value is the largest possible. Note that

$$
\begin{aligned}
&\hat{x}_t - \hat{x}_{t-i} \le iV, \quad \forall i \in [1, \alpha_t]_{\mathbb{Z}}; \\
&\hat{x}_t - \hat{x}_{t+i} \le iV, \quad \forall i \in [1, \beta_t]_{\mathbb{Z}}; \\
&-\tau_{t1}\hat{x}_{t-\alpha_t-\tau_{t1}} \le -\tau_{t1}\underline{C}; \\
&-\tau_{t2}\hat{x}_{t-\beta_t-\tau_{t2}} \le -\tau_{t2}\underline{C}; \\
&\sum_{i=t-\alpha_t-\tau_{t1}}^{t+\beta_t+\tau_{t2}} \hat{x}_i \le Q.
\end{aligned}
$$

Summing up these inequalities, we have

$$
(\alpha_t + \beta_t + 1)\hat{x}_t \le Q + \sum_{i=1}^{\alpha_t} iV + \sum_{i=1}^{\beta_t} iV - (\tau_{t1} + \tau_{t2})\underline{C},
$$

or equivalently,

$$
\hat{x}_t \le \frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1}.
$$

Because $\mathcal{D} \subseteq \hat{\mathcal{D}}$ and $(\hat{x}, \hat{y}, \hat{u})$ is an element of $\hat{\mathcal{D}}$ where the x_t value is the largest possible,

$$
x_t \le \frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1}
$$

for all $(x, y, y) \in \mathcal{D}$. Therefore, inequality (EC.5) holds for any $t \in \mathcal{T}$ and any element of $\mathcal{D}$.

Next, we show that inequality (4) holds for any element of $\mathcal{D}$. For any $t \in \mathcal{T}$, if the generator is online in period t (i.e., $y_t = 1$), then we refer to $\max\{r \mid y_{r-1} = 0;\, y_r = 1;\, r \in [2, t]_{\mathbb{Z}}\}$ as the "latest start-up period before period $t+1$" when $y_{r-1}=0$ and $y_r=1$ for some $r\in[2,t]_{\mathbb{Z}}$, and we refer to $\min\{s \mid y_{s-1}=1;\, y_s=0;\, s\in[t+1,T]_{\mathbb{Z}}\}$ as the "earliest shut-down period after period t" when $y_{s-1}=1$ and $y_s=0$ for some $s\in[t+1,T]_{\mathbb{Z}}$. For ease of exposition, we let $-(\varrho-1)$ be the latest start-up period before period $t+1$ if $y_r=1$ for all $r\in[1,t]_{\mathbb{Z}}$, and let $T+\varrho+1$ be the earliest shut-down period after period t if $y_s=1$ for all $s\in[t+1,T]_{\mathbb{Z}}$, where $\varrho=\lceil(\overline{C}-\overline{V})/V\rceil$.

By (2a), $y_{t+k}-\sum_{j=0}^{L-1}u_{t+k-j}\ge 0$. Because $k\in[2,\beta_t]_{\mathbb{Z}}$, inequality (EC.4) implies that

$$\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}\ge(k-1)V+\overline{V}. \tag{EC.8}$$

Thus, the third term on the right-hand side of inequality (4) is non-negative. By (2a), $y_i-\sum_{j=0}^{L-1}u_{i-j}\ge 0$ for all $i\in S$. Thus, the second term on the right-hand side of inequality (4) is non-negative. It is easy to see that the other terms on the right-hand side of (4) are also non-negative. Hence, all four terms on the right-hand side of (4) are non-negative. Thus, inequality (4) holds when $x_t=0$. Therefore, to prove the validity of (4), it suffices to show that inequality (4) holds for any element of $\mathcal{D}$ such that $x_t>0$ (i.e., $y_t=1$). Consider such an element of $\mathcal{D}$. Let $t-p$ and $t+q+1$ denote the latest start-up period before period $t+1$ and the earliest shut-down period after period t, respectively, where $p,q\ge 0$. From (2g), we have $x_{i-1}-x_i\le Vy_i+\overline{V}(1-y_i)$ for $i\in[t+1,t+q+1]_{\mathbb{Z}}\cap\mathcal{T}$. If $t+q+1\le T$, then $y_{t+1}=y_{t+2}=\cdots=y_{t+q}=1$ and $y_{t+q+1}=x_{t+q+1}=0$, and thus $x_t=\sum_{i=t+1}^{t+q+1}(x_{i-1}-x_i)\le\sum_{i=t+1}^{t+q+1}[Vy_i+\overline{V}(1-y_i)]=qV+\overline{V}$. If $t+q+1=T+\varrho+1$, then $y_{t+1}=y_{t+2}=\cdots=y_T=1$, and thus $x_t=\sum_{i=t+1}^{T}(x_{i-1}-x_i)+x_T\le\sum_{i=t+1}^{T}[Vy_i+\overline{V}(1-y_i)]+\overline{C}=(T-t)V+\overline{C}=(q-\varrho)V+\overline{C}\le qV+\overline{V}$. From (2f), we have $x_i-x_{i-1}\le Vy_{i-1}+\overline{V}(1-y_{i-1})$ for $i\in[t-p,t]_{\mathbb{Z}}\cap\mathcal{T}$. If $t-p\ge 2$, then $y_{t-p}=y_{t-p+1}=\cdots=y_{t-1}=1$ and $y_{t-p-1}=x_{t-p-1}=0$, and thus $x_t=\sum_{i=t-p}^{t}(x_i-x_{i-1})\le\sum_{i=t-p}^{t}[Vy_{i-1}+\overline{V}(1-y_{i-1})]=pV+\overline{V}$. If $t-p=-(\varrho-1)$, then $y_1=y_2=\cdots=y_{t-1}=1$, and thus $x_t=\sum_{i=2}^{t}(x_i-x_{i-1})+x_1\le\sum_{i=2}^{t}[Vy_{i-1}+\overline{V}(1-y_{i-1})]+\overline{C}=(t-1)V+\overline{C}=(p-\varrho)V+\overline{C}\le pV+\overline{V}$. Hence,

$$x_t\le\min\{q,p\}V+\overline{V}. \tag{EC.9}$$

Note that for any $r,s\in\mathbb{Z}$, if $r\in S$ and $r-1\le s\le t+k-1$, then

$$\sum_{i\in S\cap[r,s]_{\mathbb{Z}}}(d_i-i)\ge s-r+1. \tag{EC.10}$$

This is because (i) if $s=r-1$, then both sides of this inequality equal 0; and (ii) if $s\ge r$, then when the elements of $S\cap[r,s]_{\mathbb{Z}}$ are arranged in increasing order, the jth term of the summation "$\sum_{i\in S\cap[r,s]_{\mathbb{Z}}}d_i$" equals the $(j+1)$st term of the summation "$\sum_{i\in S\cap[r,s]_{\mathbb{Z}}}i$" for any j, except that the last term of the summation "$\sum_{i\in S\cap[r,s]_{\mathbb{Z}}}d_i$" is at least $s+1$, while the first term of the summation "$\sum_{i\in S\cap[r,s]_{\mathbb{Z}}}i$," equals r. When $r=t+1$ and $s=t+k-1$, we have

$$\sum_{i\in S}(d_i-i)=k-1. \tag{EC.11}$$

Let RHS denote the right-hand side of inequality (4). We consider five different cases.

Case (1): $p\le\max\{L-k-1,0\}$. In this case,

$$RHS\ge V\left(\sum_{j=[L-k]^+}^{L-1}\min\{L-1-j,j\}u_{t-j}+\sum_{j=0}^{L-k-1}ju_{t-j}\right)+\overline{V}y_t$$

$$\geq V \sum_{j=0}^{L-k-1} j u_{t-j} + \overline{V}$$
$$\geq pV + \overline{V}$$
$$\geq x_t,$$

where the third inequality holds because $u_{t-p}=1$ and $p \leq \max\{L-k-1,0\}$, and the last inequality is due to (EC.9). Therefore, inequality (4) holds.

Case (2): $p \geq L-1$ and $q \leq k-1$. Note that the condition "$q \leq k-1$" implies that $t+q+1 \leq t+k \leq t+\beta_t \leq T$, where the last inequality follows from Proposition 7(iii). Thus, the earliest shut-down period after period t is at most T. In this case, $y_j=1$ and $u_j=0$ for all $j \in [t-L+2, t+q]_{\mathbb{Z}}$. Hence, $y_i - \sum_{j=0}^{L-1} u_{i-j} = 1$ for all $i \in [t+1, t+q]_{\mathbb{Z}}$. This, together with (EC.10), implies that $\sum_{i \in S \cap [t+1,t+q]_Z} (d_i - i)V(y_i - \sum_{j=0}^{L-1} u_{i-j}) = \sum_{i \in S \cap [t+1,t+q]_Z} (d_i - i)V \geq qV$. Thus,

$$\begin{aligned}
RHS &\geq \sum_{i\in S}(d_i - i)V\left(y_i - \sum_{j=0}^{L-1} u_{i-j}\right) + \overline{V}y_t \\
&\geq \sum_{i \in S\cap[t+1,t+q]_Z}(d_i - i)V\left(y_i - \sum_{j=0}^{L-1} u_{i-j}\right) + \overline{V} \\
&\geq qV + \overline{V} \\
&\geq x_t,
\end{aligned}$$

where the last inequality is due to (EC.9). Therefore, inequality (4) holds.

Case (3): $p \geq L-1$ and $q \geq k$. In this case, $t+k \leq \min\{t+q, t+\beta_t\} \leq \min\{t+q, T\}$, which implies that $t+k \in [t+1, \min\{t+q,T\}]_{\mathbb{Z}}$ and $S \subseteq [t+1, \min\{t+q,T\}]_{\mathbb{Z}}$. Note that in this case, $y_j = 1$ and $u_j = 0$ for all $j \in [t-L+2, \min\{t+q,T\}]_{\mathbb{Z}}$. Thus, $y_t = 1$, $y_{t+k} - \sum_{j=0}^{L-1} u_{t+k-j} = 1$, and $y_i - \sum_{j=0}^{L-1} u_{i-j} = 1$ for all $i \in S$. Hence, by (EC.5) and (EC.11),

$$\begin{aligned}
RHS \geq& \sum_{i\in S}(d_i - i)V\left(y_i - \sum_{j=0}^{L-1} u_{i-j}\right) \\
&+ \left(\frac{Q+\Delta(\alpha_t,\beta_t)V - (\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1} - (k-1)V - \overline{V}\right)\left(y_{t+k} - \sum_{j=0}^{L-1} u_{t+k-j}\right) + \overline{V}y_t \\
=& \sum_{i\in S}(d_i - i)V + \frac{Q+\Delta(\alpha_t,\beta_t)V - (\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1} - (k-1)V \\
=& \frac{Q+\Delta(\alpha_t,\beta_t)V - (\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1} \\
\geq& x_t.
\end{aligned}$$

Therefore, inequality (4) holds.

Case (4): $\max\{L-k,1\} \leq p \leq L-2$ and $q \leq k-1$. The condition "$p \leq L-2$" implies that $t-p \geq t-(L-2) \geq 2$, and the condition "$q \leq k-1$" implies that $t+q+1 \leq t+k \leq t+\beta_t \leq T$. Thus, the latest start-up period before period $t+1$ is at least 2, and the earliest shut-down period after period t is at most T. Note that

$$RHS \geq V\left(\sum_{j=[L-k]^+}^{L-1} \min\{L-1-j, j\} u_{t-j} + \sum_{j=0}^{L-k-1} j u_{t-j}\right) + \sum_{i\in S}(d_i - i)V\left(y_i - \sum_{j=0}^{L-1} u_{i-j}\right) + \overline{V}y_k$$

$$\geq V\min\{L-1-p,p\}u_{t-p}+\sum_{i\in S}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)+\overline{V}y_k$$

$$=\min\{L-1-p,p\}V+\sum_{i\in S}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)+\overline{V}, \tag{EC.12}$$

where the second inequality holds because $[L-k]^+\leq p\leq L-1$, and the equality holds because $u_{t-p}=1$. By inequality (2a), $\sum_{i=(t+q+1)-L+1}^{t+q+1}u_i\leq y_{t+q+1}=0$, which implies that $u_i=0$ for all $i\in[(t+q+1)-L+1,t+q+1]_{\mathbb{Z}}$. This in turn implies that $(t+q+1)-L+1\geq t-p+1$, or equivalently, $q\geq L-p-1$. If $q=L-p-1$, then $\min\{L-1-p,p\}=\min\{q,p\}$, and thus by (EC.9) and (EC.12), $RHS\geq x_t$. Hence, it suffices to consider the situation where $q\geq L-p$. In this situation, $t-p+L\in[t+1,t'-1]_{\mathbb{Z}}\subseteq S$. This is because (i) if $k\geq L$, then $t'=t+L$, and thus $t-p+L\in[t+1,t+L-1]_{\mathbb{Z}}=[t+1,t'-1]_{\mathbb{Z}}$; and (ii) if $k<L-1$, then $t'=t+k$, and thus $t-p+L\in[t+1,t+k-1]_{\mathbb{Z}}=[t+1,t'-1]_{\mathbb{Z}}$ (as the condition "$p\leq L-2$" implies that $t-p+L\geq t+1$, while the conditions "$q\leq k-1$" and "$q\geq L-p$" imply that $t-p+L\leq t+k-1$). Hence, by (EC.10), $\sum_{i\in S\cap[t-p+L,t+q]_{\mathbb{Z}}}(d_i-i)\geq p+q+1-L$. Note that $y_j=1$ and $u_j=0$ for all $j\in[t-p+1,t+q]_{\mathbb{Z}}$. Thus, $y_i-\sum_{j=0}^{L-1}u_{i-j}=1$ for all $i\in[t-p+L,t+q]_{\mathbb{Z}}$. Hence,

$$\begin{aligned}\sum_{i\in S}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)&\geq\sum_{i\in S\cap[t-p+L,t+q]_{\mathbb{Z}}}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)\\&=\sum_{i\in S\cap[t-p+L,t+q]_{\mathbb{Z}}}(d_i-i)V\\&\geq(p+q+1-L)V.\end{aligned} \tag{EC.13}$$

By (EC.9), (EC.12), and (EC.13), we have

$$\begin{aligned}RHS&\geq\min\{L-1-p,p\}V+(p+q+1-L)V+\overline{V}\\&\geq\min\{(L-1-p)+(p+q+1-L),p\}V+\overline{V}=\min\{q,p\}V+\overline{V}\geq x_t.\end{aligned}$$

Therefore, inequality (4) holds.

Case (5): $\max\{L-k,1\}\leq p\leq L-2$ and $q\geq k$. The condition "$p\leq L-2$" implies that $t-p\geq t-(L-2)\geq 2$. Thus, the latest start-up period before period $t+1$ is at least 2. We first show that

$$\sum_{i\in S\cap[t-p+L,t+k-1]_{\mathbb{Z}}}(d_i-i)\geq k+p-L. \tag{EC.14}$$

Note that $t-p+L\leq t+k$. If $t-p+L=t+k$, then inequality (EC.14) holds because both the left-hand side and right-hand side are zero. If $t-p+L\leq t+k-1$, then $t-p+L\in[t+1,t'-1]_{\mathbb{Z}}\subseteq S$ (because (i) if $k\geq L$, then $t'=t+L$ and $t-p+L\in[t+1,t+L-1]_{\mathbb{Z}}=[t+1,t'-1]_{\mathbb{Z}}$; and (ii) if $k<L$, then $t'=t+k$ and $t-p+L\in[t+1,t+k-1]_{\mathbb{Z}}=[t+1,t'-1]_{\mathbb{Z}}$). Thus, by (EC.10), inequality (EC.14) holds. Note that $y_j=1$ and $u_j=0$ for all $j\in[t-p+1,t+k-1]_{\mathbb{Z}}$. Hence, $y_i-\sum_{j=0}^{L-1}u_{i-j}=1$ for all $i\in[t-p+L,t+k-1]_{\mathbb{Z}}$. Thus, by (EC.14),

$$\begin{aligned}\sum_{i\in S}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)&\geq\sum_{i\in S\cap[t-p+L,t+k-1]_{\mathbb{Z}}}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)\\&=\sum_{i\in S\cap[t-p+L,t+k-1]_{\mathbb{Z}}}(d_i-i)V\end{aligned}$$

$$\geq (k+p-L)V.$$

This implies that

$$\begin{aligned}
&V\left(\sum_{j=[L-k]^+}^{L-1}\min\{L-1-j,j\}u_{t-j}+\sum_{j=0}^{L-k-1}ju_{t-j}\right)+\sum_{i\in S}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)\\
&\geq V\left(\sum_{j=[L-k]^+}^{L-1}\min\{L-1-j,j\}u_{t-j}+\sum_{j=0}^{L-k-1}ju_{t-j}\right)+(k+p-L)V\\
&\geq V\min\{L-1-p,p\}u_{t-p}+(k+p-L)V\\
&=[\min\{L-1-p,p\}+(k+p-L)]V\\
&\geq \min\{k-1,p\}V,
\end{aligned}$$

where the second inequality holds because $[L-k]^+\leq p\leq L-1$. Note that $y_{t+k}=1$ and $u_{t+k-j}=0$ for $j=0,1,\ldots,L-1$, which implies that $y_{t+k}-\sum_{j=0}^{L-1}u_{t+k-j}=1$. Hence,

$$\begin{aligned}
RHS &= V\left(\sum_{j=[L-k]^+}^{L-1}\min\{L-1-j,j\}u_{t-j}+\sum_{j=0}^{L-k-1}ju_{t-j}\right)+\sum_{i\in S}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)\\
&\quad+\left(\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V-\overline{V}\right)\left(y_{t+k}-\sum_{j=0}^{L-1}u_{t+k-j}\right)+\overline{V}\\
&\geq \min\{k-1,p\}V+\left(\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V-\overline{V}\right)\left(y_{t+k}-\sum_{j=0}^{L-1}u_{t+k-j}\right)+\overline{V}\\
&=\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V+\min\{k-1,p\}V. \qquad \text{(EC.15)}
\end{aligned}$$

If $p\geq k-1$, then $\min\{k-1,p\}V=(k-1)V$, and thus by (EC.5) and (EC.15), we have $RHS\geq x_t$. If $p<k-1$, then by (EC.9), $\min\{k-1,p\}V=pV\geq\min\{q,p\}V\geq x_t-\overline{V}$, and thus from (EC.8) and (EC.15), we have $RHS\geq x_t$. Therefore, inequality (4) holds.

Summarizing Cases (1)–(5), we conclude that inequality (4) holds for any element of $\mathcal{D}$. □

A.10. Proof of Remark 4

First, we show that under the condition "$k\geq(L+1)/2$," inequality (4) can be rewritten as

$$\begin{aligned}
x_t\leq\ &V\sum_{j=1}^{L-1}\min\{L-1-j,j\}u_{t-j}+\sum_{i\in S}(d_i-i)V\left(y_i-\sum_{j=0}^{L-1}u_{i-j}\right)\\
&+\left(\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V-\overline{V}\right)\left(y_{t+k}-\sum_{j=0}^{L-1}u_{t+k-j}\right)+\overline{V}y_t. \qquad \text{(EC.16)}
\end{aligned}$$

If $L\leq k$, then clearly, $\sum_{j=[L-k]^+}^{L-1}\min\{L-1-j,j\}u_{t-j}+\sum_{j=0}^{L-k-1}ju_{t-j}=\sum_{j=1}^{L-1}\min\{L-1-j,j\}u_{t-j}$, and thus inequality (4) can be rewritten as (EC.16). If $L>k$, then because $k\geq(L+1)/2$, we have $j\leq(L-1)/2$ for all $j=0,1,\ldots,L-k-1$, which implies that $\min\{L-1-j,j\}=j$ for all $j=0,1,\ldots,L-k-1$. Thus, in this case, inequality (4) can also be rewritten as (EC.16).

To prove that the look forward inequalities are facet-defining, we show that there exist $3T-1$ affinely independent points in conv($\mathcal{D}$) that satisfy (EC.16) with equality when $Q\geq Q_{B2}$, $\overline{C}\geq[Q+\Delta(t-1,T-$

$t)V]/T$, $t+k=T$, $k \geq (L+1)/2$, and $S' = [t', t+k-1]_{\mathbb{Z}}$. Because $\mathbf{0} \in \text{conv}(\mathcal{D})$ and this point satisfies (EC.16) with equality, it suffices to create the remaining $3T-2$ non-zero linearly independent points. We denote these $3T-2$ points as $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r)$ for $r \in \mathcal{T} \setminus \{T\}$, $(\bar{x}^r, \bar{y}^r, \bar{u}^r)$ for $r \in \mathcal{T} \setminus \{1\}$, and $(\hat{x}^r, \hat{y}^r, \hat{u}^r)$ for $r \in \mathcal{T}$, and we divide these $3T-2$ points into nine groups. Let $\epsilon = \min\{\overline{V} - \underline{C}, \overline{C} - \underline{C}\} > 0$. The $3T-2$ points are created as follows:

(A1) For each $r \in [1, t-1]_{\mathbb{Z}}$, we create a point $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r)$ as follows: For each $s \in \mathcal{T}$, set $\tilde{x}_s^r = \underline{C}$ and $\tilde{y}_s^r = 1$ if $s \leq r-1$, set $\tilde{x}_s^r = \underline{C} + \epsilon$ and $\tilde{y}_s^r = 1$ if $s = r$, and set $\tilde{x}_s^r = \tilde{y}_s^r = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\tilde{u}_s^r = 0$. There are $t-1$ points in this group. It is easy to verify that these $t-1$ points satisfy (2a)–(2g). Note that $\sum_{s=1}^T \tilde{x}_s^r = r\underline{C} + \epsilon < (T-1)\underline{C} + \overline{V} < T\underline{C} + V \leq \omega_2(t-1, T-t) = Q_{B2} \leq Q$ for all $r \in [1, t-1]_{\mathbb{Z}}$. Thus, these points also satisfy (2h) and are in $\mathcal{D}$. It is easy to verify that these points satisfy (EC.16) with equality.

(A2) For each $r \in [t, T-1]_{\mathbb{Z}}$, we create a point $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r)$ as follows: For each $s \in \mathcal{T}$, set $\tilde{x}_s^r = \underline{C} + \epsilon$ and $\tilde{y}_s^r = 1$ if $s = r+1$, set $\tilde{x}_s^r = \underline{C}$ and $\tilde{y}_s^r = 1$ if $r+2 \leq s \leq r+L$, and set $\tilde{x}_s^r = \tilde{y}_s^r = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\tilde{u}_s^r = 1$ if $s = r+1$, and set $\tilde{u}_s^r = 0$ otherwise. There are $T-t$ points in this group. It is easy to verify that these $T-t$ points satisfy (2a)–(2g). Note that $\sum_{s=1}^T \tilde{x}_s^r \leq (L+1)\underline{C} < Q$ for all $r \in [t, T-1]_{\mathbb{Z}}$. Thus, these points also satisfy (2h) and are in $\mathcal{D}$. It is easy to verify that these points satisfy (EC.16) with equality.

(A3) For each $r \in [2, t-L]_{\mathbb{Z}}$, we create a point $(\bar{x}^r, \bar{y}^r, \bar{u}^r)$ as follows: For each $s \in \mathcal{T}$, set $\bar{x}_s^r = \underline{C}$ and $\bar{y}_s^r = 1$ if $r \leq s \leq r+L-1$, and set $\bar{x}_s^r = \bar{y}_s^r = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\bar{u}_s^r = 1$ if $s = r$, and set $\bar{u}_s^r = 0$ otherwise. There are $t-L-1$ points in this group. It is easy to verify that these $t-L-1$ points satisfy (2a)–(2g). Note that $\sum_{s=1}^T \bar{x}_s^r < (L+1)\underline{C} < Q$ for all $r \in [2, t-L]_{\mathbb{Z}}$. Thus, these points also satisfy (2h) and are in $\mathcal{D}$. It is easy to verify that these points satisfy (EC.16) with equality.

(A4) For each $r \in [t-L+1, t]_{\mathbb{Z}}$, we create a point $(\bar{x}^r, \bar{y}^r, \bar{u}^r)$ as follows: For each $s \in \mathcal{T}$, set $\bar{x}_s^r = \overline{V} + \min\{s-r, r+L-1-s\}V$ and $\bar{y}_s^r = 1$ if $r \leq s \leq r+L-1$, and set $\bar{x}_s^r = \bar{y}_s^r = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\bar{u}_s^r = 1$ if $s = r$, and set $\bar{u}_s^r = 0$ otherwise. There are L points in this group. It is easy to verify that these L points satisfy (2a)–(2d) and (2f)–(2g). Note that under the condition "$Q \geq Q_{B2}$," we have $\alpha_t = t-1$, $\beta_t = T-t$, and $\tau_{t1} = \tau_{t2} = 0$ (by Definition 2 and Proposition 7). Thus, by inequality (EC.8) in the proof of Proposition 8, $[Q + \Delta(t-1, T-t)]/T \geq (k-1)V + \overline{V}$, which implies that $\overline{C} \geq (k-1)V + \overline{V} \geq \frac{L-1}{2} \cdot V + \overline{V}$. Thus, $\bar{x}_s^r \leq \overline{C}$ for any $s \in \mathcal{T}$. Hence, these L points satisfy (2e). Note that

$$
\begin{aligned}
\sum_{s=1}^T \bar{x}_s^r &\leq L\overline{V} + \left\lfloor \frac{L-1}{2} \right\rfloor \left\lfloor \frac{L}{2} \right\rfloor V \\
&< L(\underline{C} + V) + \left\lfloor \frac{L-1}{2} \right\rfloor \left\lfloor \frac{L}{2} \right\rfloor V \\
&< T\underline{C} + \frac{L^2}{2} \cdot V \\
&\leq T\underline{C} + \frac{(t-1)^2}{2} \cdot V \\
&\leq T\underline{C} + \left[\frac{|(t-1)t - (T-t)(T-t+1)|}{2} + (t-1)(T-t) \right] V
\end{aligned}
$$

$$= \omega_2(t-1, T-t) = Q_{B2} \le Q$$

for any $r \in [t-L+1, t]_{\mathbb{Z}}$. Thus, these points also satisfy (2h) and are in $\mathcal{D}$. Note that $\bar{y}_i^r - \sum_{j=0}^{L-1} \bar{u}_{i-j}^r = 0$ for any $i \in [t+1, t+k]_{\mathbb{Z}}$. Note also that for any $r \in [t-L+1, t]_{\mathbb{Z}}$, $\bar{x}_t^r = \overline{V} + \min\{t-r, r+L-1-t\}V$, $\bar{y}_t^r = 1$, and $\sum_{j=0}^{L-1} \min\{L-1-j, j\}\bar{u}_{t-j}^r = \min\{L-1-t+r, t-r\}$. Hence, $(\bar{x}^r, \bar{y}^r, \bar{u}^r)$ satisfies (EC.16) with equality for any $r \in [t-L+1, t]_{\mathbb{Z}}$.

(A5) For each $r \in [t+1, T]_{\mathbb{Z}}$, we create a point $(\bar{x}^r, \bar{y}^r, \bar{u}^r)$ as follows: For each $s \in \mathcal{T}$, set $\bar{x}_s^r = \underline{C}$ and $\bar{y}_s^r = 1$ if $r \le s \le r+L-1$, and set $\bar{x}_s^r = \bar{y}_s^r = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\bar{u}_s^r = 1$ if $s = r$, and set $\bar{u}_s^r = 0$ otherwise. There are $T-t$ points in this group. It is easy to verify that these $T-t$ points satisfy (2a)–(2g). Note that $\sum_{s=1}^T \bar{x}_s^r < (L+1)\underline{C} < Q$ for all $r \in [t+1, T]_{\mathbb{Z}}$. Thus, these points also satisfy (2h) and are in $\mathcal{D}$. It is easy to verify that these points satisfy (EC.16) with equality.

(A6) For each $r \in [1, t-1]_{\mathbb{Z}}$, we create a point $(\hat{x}^r, \hat{y}^r, \hat{u}^r)$ as follows: For each $s \in \mathcal{T}$, set $\hat{x}_s^r = \underline{C}$ and $\hat{y}_s^r = 1$ if $s \le r$, and set $\hat{x}_s^r = \hat{y}_s^r = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\hat{u}_s^r = 0$. There are $t-1$ points in this group. It is easy to verify that these $t-1$ points satisfy (2a)–(2g). Note that $\sum_{s=1}^T \hat{x}_s^r = r\underline{C} < (T-1)\underline{C} + \overline{V} < T\underline{C} + V \le \omega_2(t-1, T-t) = Q_{B2} \le Q$ for all $r \in [1, t-1]_{\mathbb{Z}}$. Thus, these points also satisfy (2h) and are in $\mathcal{D}$. It is easy to verify that these points satisfy (EC.16) with equality.

(A7) We create a point $(\hat{x}^t, \hat{y}^t, \hat{u}^t)$ as follows: For each $s \in \mathcal{T}$, set $\hat{x}_s^t = \underline{C}$ and $\hat{y}_s^t = 1$ if $s \le t-1$, set $\hat{x}_s^t = \overline{V}$ and $\hat{y}_s^t = 1$ if $s = t$, and set $\hat{x}_s^t = \hat{y}_s^t = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\hat{u}_s^t = 0$. There is one point in this group. It is easy to verify that this point satisfies (2a)–(2g). Note that $\sum_{s=1}^T \hat{x}_s^t \le (T-1)\underline{C} + \overline{V} < T\underline{C} + V \le \omega_2(t-1, T-t) = Q_{B2} \le Q$. Thus, this point also satisfies (2h) and is in $\mathcal{D}$. It is easy to verify that this point satisfies (EC.16) with equality.

(A8) For each $r \in [t+1, T-1]_{\mathbb{Z}}$, we create a point $(\hat{x}^r, \hat{y}^r, \hat{u}^r)$ as follows: For each $s \in \mathcal{T}$, set $\hat{x}_s^r = \max\{\underline{C}, \overline{V} + (s+r-2t)V\}$ and $\hat{y}_s^t = 1$ if $s \le t$, set $\hat{x}_s^r = \overline{V} + (r-s)V$ and $\hat{y}_s^t = 1$ if $t < s \le r$, and set $\hat{x}_s^r = \hat{y}_s^t = 0$ otherwise. For each $s \in \mathcal{T} \setminus \{1\}$, set $\hat{u}_s^r = 0$. There are $T-t-1$ points in this group. It is easy to verify that these $T-t-1$ points satisfy (2a)–(2d) and (2f)–(2g). Under the condition "$Q \ge Q_{B2}$," we have $\alpha_t = t-1$, $\beta_t = T-t$, and $\tau_{t1} = \tau_{t2} = 0$. Then, by inequality (EC.8) in the proof of Proposition 8, $[Q + \Delta(t-1, T-t)]/T \ge (k-1)V + \overline{V} = (T-t-1)V + \overline{V} \ge \overline{V} + (s+r-2t)V$ for any $s \le t$ and $r \in [t+1, T-1]_{\mathbb{Z}}$. Thus, $\bar{x}_s^r \le \overline{C}$ for any $s \in \mathcal{T}$. Hence, these L points satisfy (2e). Consider the quantity "$\sum_{s=1}^T \hat{x}_s^r$," which increases as r increases. If $t-1 \ge T-t$, then for any $r \in [t+1, T-1]_{\mathbb{Z}}$,

$$\begin{aligned}
\sum_{s=1}^T \hat{x}_s^r \le \sum_{s=1}^T \hat{x}_s^{T-1} \\
&= (2t-T)\underline{C} + (2T-2t-1)\overline{V} + (T-t-1)^2 V \\
&< (2t-T)\underline{C} + (2T-2t-1)(\underline{C} + V) + (T-t-1)^2 V \\
&= (T-1)\underline{C} + (T-t)^2 V \\
&< T\underline{C} + (t-1)(T-t)V \\
&\le \omega_2(t-1, T-t).
\end{aligned}$$

If $t-1<T-t$, then for any $r\in[t+1,T-1]_{\mathbb{Z}}$,

$$\begin{aligned}\sum_{s=1}^{T}\hat{x}_s^r &\le \sum_{s=1}^{T}\hat{x}_s^{T-1}\\ &=(T-1)\overline{V}+(T-t-1)^2V-\frac{1}{2}(T-2t-1)(T-2t)V\\ &<(T-1)(\underline{C}+V)+(T-t-1)^2V-\frac{1}{2}(T-2t-1)(T-2t)V\\ &<T\underline{C}+\frac{1}{2}(T-t)(T-t+1)V-\frac{1}{2}(t-1)tV+(t-1)(T-t)V\\ &=\omega_2(t-1,T-t).\end{aligned}$$

Thus, in both cases, $\sum_{s=1}^{T}\hat{x}_s^r<\omega_2(t-1,T-t)=Q_{B2}\le Q$ for all $r\in[t+1,T-1]_{\mathbb{Z}}$. Hence, these points also satisfy (2h) and are in $\mathcal{D}$. We consider the left-hand side and right-hand side values of (EC.16) when $(x,y,z)=(\hat{x}^r,\hat{y}^r,\hat{u}^r)$. The left-hand side equals $\overline{V}+(r-t)V$. The first and third terms of the right-hand side of (EC.16) equal 0. The second and fourth terms of the right-hand side equal $(r-t)V$ and $\overline{V}$, respectively. Hence, $(\bar{x}^r,\bar{y}^r,\bar{u}^r)$ satisfies (EC.16) with equality for any $r\in[t+1,T-1]_{\mathbb{Z}}$.

(A9) We create a point $(\hat{x}^T,\hat{y}^T,\hat{u}^T)$ as follows: For each $s\in\mathcal{T}$, set $\hat{x}_s^T=[Q+\Delta(t-1,T-t)V]/T-|s-t|V$ and $\hat{y}_s^T=1$. For each $s\in\mathcal{T}\setminus\{1\}$, set $\hat{u}_s^T=0$. There is one point in this group. It is easy to verify that this point satisfies (2a)–(2c) and (2e)–(2g). For any $s\in\mathcal{T}$,

$$\begin{aligned}\hat{x}_s^T &\ge \frac{Q_{B2}+\Delta(t-1,T-t)V}{T}-\max\{t-1,T-t\}V\\ &=\frac{\omega_2(t-1,T-t)+\Delta(t-1,T-t)V}{T}-\max\{t-1,T-t\}V=\underline{C}.\end{aligned}$$

Thus, this point satisfies (2d). Note that $\sum_{s=1}^{T}\hat{x}_s^T=Q+\Delta(t-1,T-t)V-\sum_{s=1}^{T}|s-t|V=Q$. Hence, this point satisfies (2h). Therefore, this point is in $\mathcal{D}$. We consider the left-hand side and right-hand side values of (EC.16) when $(x,y,z)=(\hat{x}^T,\hat{y}^T,\hat{u}^T)$. Recall that $\alpha_t=t-1$, $\beta_t=T-t$, and $\tau_{t1}=\tau_{t2}=0$. Thus, the left-hand side of (EC.16) equals $[Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}]/(\alpha_t+\beta_t+1)$. The first term of the right-hand side of (EC.16) equals 0. The second, third, and fourth terms of the right-hand side equal $(k-1)V$, $[Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}]/(\alpha_t+\beta_t+1)-(k-1)V-\overline{V}$, and $\overline{V}$, respectively. Therefore, $(\bar{x}^T,\bar{y}^T,\bar{u}^T)$ satisfies (EC.16) with equality.

Let $\mathcal{L}(p,q)$ denote the T-dimensional row vector, in which the jth component equals 1 if $j\in[p,q]_{\mathbb{Z}}\cap\mathcal{T}$, and it equals 0 otherwise. Let $\mathcal{I}(q)$ denote the $(T-1)$-dimensional row vector, in which the $(j-1)$st component equals 1 if $j=q$ and $j\in\mathcal{T}\setminus\{1\}$, and it equals 0 otherwise. Let $\mathbf{0}$ denote the row vector with all components equal to 0 ($\mathbf{0}$ is T-dimensional if it appears in the y column, and is $(T-1)$-dimensional if it appears in the u column). Table EC.3 shows a matrix in which each row represents a point created by the above process. In the following, we show that the matrix in Table EC.3 can be transformed into a lower triangular matrix (i.e., a $(3T-2)\times(3T-1)$ matrix in which the element in ith row and jth column is zero if $j>i+1$) via Gaussian elimination. The transformed matrix is shown in Table EC.4, where the rows are divided into Groups 1–4. The Gaussian elimination process is as follows:

(i) For each $r\in[1,t-1]_{\mathbb{Z}}$, point r of Group 1, denoted $(\underline{\tilde{x}}^r,\underline{\tilde{y}}^r,\underline{\tilde{u}}^r)$, is obtained by setting $(\underline{\tilde{x}}^r,\underline{\tilde{y}}^r,\underline{\tilde{u}}^r)=(\tilde{x}^r,\tilde{y}^r,\tilde{u}^r)-(\hat{x}^r,\hat{y}^r,\hat{u}^r)$. Here, $(\tilde{x}^r,\tilde{y}^r,\tilde{u}^r)$ is the point with index r in group (A1), and $(\hat{x}^r,\hat{y}^r,\hat{u}^r)$ is the point with index r in group (A6).

Table EC.3 A matrix with the rows representing $3T-2$ points that satisfy the Look Forward Inequality with equality*

Group	Point	Index r	x																	y	u
			1	2	3	$\ldots$	$t-2$	$t-1$	t	$t+1$	$t+2$	$\ldots$	$t+L$	$t+L+1$	$\ldots$	$T-1$	T			1 2 $\ldots$ $T-1$ T	2 3 $\ldots$ $T-1$ T
(A1)		1	$\underline{C}+\epsilon$	0	0	$\ldots$	0	0	0	0	0	$\ldots$	0	0	$\ldots$	0	0			$\mathcal{L}(1,1)$	$\mathbf{0}$
		2	$\underline{C}$	$\underline{C}+\epsilon$	0	$\ldots$	0	0	0	0	0	$\ldots$	0	0	$\ldots$	0	0			$\mathcal{L}(1,2)$	$\mathbf{0}$
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$			$\vdots$	$\vdots$
	$(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r)$	$t-1$	$\underline{C}$	$\underline{C}$	$\underline{C}$	$\ldots$	$\underline{C}$	$\underline{C}+\epsilon$	0	0	0	$\ldots$	0	0	$\ldots$	0	0			$\mathcal{L}(1,t-1)$	$\mathbf{0}$
(A2)		t	0	0	0	$\ldots$	0	0	0	$\underline{C}+\epsilon$	$\underline{C}$	$\ldots$	$\underline{C}$	0	$\ldots$	0	0			$\mathcal{L}(t+1,t+L)$	$\mathcal{I}(t+1)$
		$t+1$	0	0	0	$\ldots$	0	0	0	0	$\underline{C}+\epsilon$	$\ldots$	$\underline{C}$	$\underline{C}$	$\ldots$	0	0			$\mathcal{L}(t+2,t+L+1)$	$\mathcal{I}(t+2)$
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$			$\vdots$	$\vdots$
		$T-1$	0	0	0	$\ldots$	0	0	0	0	0	$\ldots$	0	0	$\ldots$	0	$\underline{C}+\epsilon$			$\mathcal{L}(T,T+L-1)$	$\mathcal{I}(T)$
(A3)		2	(omitted)																	$\mathcal{L}(2,L+1)$	$\mathcal{I}(2)$
		3																		$\mathcal{L}(3,L+2)$	$\mathcal{I}(3)$
		$\vdots$																		$\vdots$	$\vdots$
		$t-L$																		$\mathcal{L}(t-L,t-1)$	$\mathcal{I}(t-L)$
(A4)	$(\bar{x}^r, \bar{y}^r, \bar{u}^r)$	$t-L+1$	(omitted)																	$\mathcal{L}(t-L+1,t)$	$\mathcal{I}(t-L+1)$
		$t-L+2$																		$\mathcal{L}(t-L+2,t+1)$	$\mathcal{I}(t-L+2)$
		$\vdots$																		$\vdots$	$\vdots$
		t																		$\mathcal{L}(t,t+L-1)$	$\mathcal{I}(t)$
(A5)		$t+1$	0	0	0	$\ldots$	0	0	0	$\underline{C}$	$\underline{C}$	$\ldots$	$\underline{C}$	0	$\ldots$	0	0			$\mathcal{L}(t+1,t+L)$	$\mathcal{I}(t+1)$
		$t+2$	0	0	0	$\ldots$	0	0	0	0	$\underline{C}$	$\ldots$	$\underline{C}$	$\underline{C}$	$\ldots$	0	0			$\mathcal{L}(t+2,t+L+1)$	$\mathcal{I}(t+2)$
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$			$\vdots$	$\vdots$
		T	0	0	0	$\ldots$	0	0	0	0	0	$\ldots$	0	0	$\ldots$	0	$\underline{C}$			$\mathcal{L}(T,T+L-1)$	$\mathcal{I}(T)$
(A6)		1	$\underline{C}$	0	0	$\ldots$	0	0	0	0	0	$\ldots$	0	0	$\ldots$	0	0			$\mathcal{L}(1,1)$	$\mathbf{0}$
		2	$\underline{C}$	$\underline{C}$	0	$\ldots$	0	0	0	0	0	$\ldots$	0	0	$\ldots$	0	0			$\mathcal{L}(1,2)$	$\mathbf{0}$
		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$			$\vdots$	$\vdots$
		$t-1$	$\underline{C}$	$\underline{C}$	$\underline{C}$	$\ldots$	$\underline{C}$	$\underline{C}$	0	0	0	$\ldots$	0	0	$\ldots$	0	0			$\mathcal{L}(1,t-1)$	$\mathbf{0}$
(A7)	$(\hat{x}^r, \hat{y}^r, \hat{u}^r)$	t	(omitted)																	$\mathcal{L}(1,t)$	$\mathbf{0}$
(A8)		$t+1$	(omitted)																	$\mathcal{L}(1,t+1)$	$\mathbf{0}$
		$t+2$																		$\mathcal{L}(1,t+2)$	$\mathbf{0}$
		$\vdots$																		$\vdots$	$\vdots$
		$T-1$																		$\mathcal{L}(1,T-1)$	$\mathbf{0}$
(A9)		T	(omitted)																	$\mathcal{L}(1,T)$	$\mathbf{0}$

*The x values of the points in (A3)–(A4) and (A7)–(A9) are not shown in this table.

Table EC.4 Lower triangular matrix obtained by Gaussian elimination*

Group	Point	Index r	x: 1	2	…	$t-1$	t	$t+1$	$t+2$	…	T	y: 1 2 … $T-1$ T	u: 2 3 … $T-1$ T
1	$(\underline{\tilde{x}}^r, \underline{\tilde{y}}^r, \underline{\tilde{u}}^r)$	1	ϵ	0	…	0	0	0	0	…	0	$\mathbf{0}$	$\mathbf{0}$
		2	0	ϵ	…	0	0	0	0	…	0	$\mathbf{0}$	$\mathbf{0}$
		⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮
		$t-1$	0	0	…	ϵ	0	0	0	…	0	$\mathbf{0}$	$\mathbf{0}$
2		t	0	0	…	0	0	ϵ	0	…	0	$\mathbf{0}$	$\mathbf{0}$
		$t+1$	0	0	…	0	0	0	ϵ	…	0	$\mathbf{0}$	$\mathbf{0}$
		⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮
		$T-1$	0	0	…	0	0	0	0	…	ϵ	$\mathbf{0}$	$\mathbf{0}$
3	$(\underline{\hat{x}}^r, \underline{\hat{y}}^r, \underline{\hat{u}}^r)$	1	(omitted)									$\mathcal{L}(1,1)$	$\mathbf{0}$
		2										$\mathcal{L}(1,2)$	$\mathbf{0}$
		⋮										⋮	⋮
		T										$\mathcal{L}(1,T)$	$\mathbf{0}$
4	$(\underline{\bar{x}}^r, \underline{\bar{y}}^r, \underline{\bar{u}}^r)$	2	(omitted)									(omitted)	$\mathcal{I}(2)$
		3											$\mathcal{I}(3)$
		⋮											⋮
		T											$\mathcal{I}(T)$

*The x values in Groups 3–4 and the y values in Group 4 are not shown in this table.

(ii) For each $r \in [t, T-1]_{\mathbb{Z}}$, point r of Group 2, denoted $(\underline{\tilde{x}}^r, \underline{\tilde{y}}^r, \underline{\tilde{u}}^r)$, is obtained by setting $(\underline{\tilde{x}}^r, \underline{\tilde{y}}^r, \underline{\tilde{u}}^r) = (\tilde{x}^r, \tilde{y}^r, \tilde{u}^r) - (\bar{x}^{r+1}, \bar{y}^{r+1}, \bar{u}^{r+1})$. Here, $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r)$ is the point with index r in group (A2), and $(\bar{x}^{r+1}, \bar{y}^{r+1}, \bar{u}^{r+1})$ is the point with index $r+1$ in group (A5).

(iii) For each $r \in [1, T]_{\mathbb{Z}}$, point r of Group 3, denoted $(\underline{\hat{x}}^r, \underline{\hat{y}}^r, \underline{\hat{u}}^r)$, is obtained by setting $(\underline{\hat{x}}^r, \underline{\hat{y}}^r, \underline{\hat{u}}^r) = (\hat{x}^r, \hat{y}^r, \hat{u}^r)$, which is the point with index r in groups (A6), (A7), (A8), and (A9).

(iv) For each $r \in [2, T]_{\mathbb{Z}}$, point r of Group 4, denoted $(\underline{\bar{x}}^r, \underline{\bar{y}}^r, \underline{\bar{u}}^r)$, is obtained by setting $(\underline{\bar{x}}^r, \underline{\bar{y}}^r, \underline{\bar{u}}^r) = (\bar{x}^r, \bar{y}^r, \bar{u}^r)$, which is the point with index r in groups (A3), (A4), and (A5).

As shown in Table EC.4, the points in Groups 1–4 form a lower triangular matrix in which the column index of the last nonzero entry of the ith row is greater than that of the $(i-1)$st row, for $i = 2, \ldots, 3T-2$. Thus, these points are linearly independent. Hence, the $3T-2$ points in groups (A1)–(A9) are linearly independent. □

A.11. Proof of Proposition 9

We show that inequality (5) holds for any element of $\mathcal{D}$. From the proof of Proposition 8, for any $t \in \mathcal{T}$ and any element of $\mathcal{D}$,

$$\frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1} \geq (\max\{\alpha_t, \beta_t\} - 1)V + \overline{V} \tag{EC.17}$$

and

$$x_t \leq \frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1}. \tag{EC.18}$$

Using (2a), it is easy to verify that $y_{t-k} - \sum_{j=1}^{L} w_{t-k-j} \geq 0$. Because $k \in [2, \alpha_t]_{\mathbb{Z}}$, inequality (EC.17) implies that

$$\frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1} \geq (k-1)V + \overline{V}. \tag{EC.19}$$

Thus, the third term on the right-hand side of inequality (5) is non-negative. Similarly, $y_i - \sum_{j=1}^{L} w_{i+j} \geq 0$ for all $i \in S$. Thus, the second term on the right-hand side of inequality (5) is non-negative. It is easy to see that the other terms on the right-hand side of (5) are also non-negative. Hence, all four terms on the right-hand side of (5) are non-negative. Thus, inequality (5) is valid when $x_t = 0$. Therefore, to prove the validity of (5), it suffices to show that inequality (5) is valid for any element of $\mathcal{D}$ such that $x_t > 0$ (i.e., $y_t = 1$). Consider such an element of $\mathcal{D}$. We define the "latest start-up period before period $t+1$" and the "earliest shut-down period after period t" in the same way as in the proof of Proposition 8. Let $t-p$ and $t+q+1$ denote the latest start-up period before period $t+1$ and the earliest shut-down period after period t, respectively, where $p,q \geq 0$. Following the same argument as in the proof of Proposition 8, we have

$$x_t \leq \min\{q,p\}V + \overline{V}. \tag{EC.20}$$

Note that for any $r,s \in \mathbb{Z}$, if $s \in S$ and $t-k+1 \leq r \leq s+1$, then

$$\sum_{i \in S \cap [r,s]_{\mathbb{Z}}} (i - d_i) \geq s - r + 1. \tag{EC.21}$$

This is because (i) if $r = s+1$, then both sides of this inequality equal 0; and (ii) if $r \leq s$, then when the elements of $S \cap [r,s]_{\mathbb{Z}}$ are arranged in increasing order, the jth term of the summation "$\sum_{i \in S \cap [r,s]_{\mathbb{Z}}} i$" equals the $(j+1)$st term of the summation "$\sum_{i \in S \cap [r,s]_{\mathbb{Z}}} d_i$" for any j, except that the last term of the summation "$\sum_{i \in S \cap [r,s]_{\mathbb{Z}}} i$" equals s, while the first term of the summation "$\sum_{i \in S \cap [r,s]_{\mathbb{Z}}} d_i$" is at most $r-1$. When $r = t-k+1$ and $s = t-1$, we have

$$\sum_{i \in S} (i - d_i) = k - 1. \tag{EC.22}$$

Let RHS denote the right-hand side of (5). We consider five different cases.

Case (1): $q \leq \max\{L-k-1, 0\}$. In this case,

$$\begin{aligned}
RHS &\geq V \left(\sum_{j=[L-k]^+ + 1}^{L-1} \min\{L-j, j-1\} w_{t+j} + \sum_{j=1}^{L-k} (j-1) w_{t+j} \right) + \overline{V} y_t \\
&\geq V \sum_{j=1}^{L-k} (j-1) w_{t+j} + \overline{V} \\
&\geq qV + \overline{V} \\
&\geq x_t,
\end{aligned}$$

where the third inequality holds because $w_{t+q+1} = 1$ and $q+1 \leq \max\{L-k, 1\}$, and the last inequality is due to (EC.20). Therefore, inequality (5) holds.

Case (2): $q \geq L-1$ and $p \leq k-1$. Note that the condition "$p \leq k-1$" implies that $t-p \geq t-k+1 \geq t-\alpha_t+1 \geq 2$, where the last inequality follows from Proposition 7(iii). Thus, the latest start-up period before period $t+1$ is at least 2. In this case, $y_j = 1$ and $w_j = 0$ for all $j \in [t-p, t+L-1]_{\mathbb{Z}}$. Hence, $y_i - \sum_{j=1}^{L} w_{i+j} = 1$ for all $i \in [t-p, t-1]_{\mathbb{Z}}$. This, together with (EC.21), implies that $\sum_{i \in S \cap [t-p,t-1]_{\mathbb{Z}}} (i-d_i) V (y_i - \sum_{j=1}^{L} w_{i+j}) = \sum_{i \in S \cap [t-p,t-1]_{\mathbb{Z}}} (i-d_i) V \geq pV$. Thus,

$$\begin{aligned}
RHS &\geq \sum_{i \in S} (i - d_i) V \left(y_i - \sum_{j=1}^{L} w_{i+j} \right) + \overline{V} y_t \\
&\geq \sum_{i \in S \cap [t-p,t-1]_{\mathbb{Z}}} (i - d_i) V \left(y_i - \sum_{j=1}^{L} w_{i+j} \right) + \overline{V}
\end{aligned}$$

$$\geq pV + \overline{V}$$
$$\geq x_t,$$

where the last inequality is due to (EC.20). Therefore, inequality (5) holds.

Case (3): $q \geq L-1$ and $p \geq k$. In this case, $t-k+1 \geq \max\{t-p+1, t-\alpha_t+1\} \geq \max\{t-p+1, 2\}$, which implies that $t-k+1 \in [\max\{t-p+1,2\}, t]_{\mathbb{Z}}$ and $S \subseteq [\max\{t-p+1,2\}, t]_{\mathbb{Z}}$. Note that in this case, $y_{t-k} = 1$, $y_j = 1$, and $w_j = 0$ for all $j \in [\max\{t-p+1,2\}, t+L-1]_{\mathbb{Z}}$. Thus, $y_t = 1$, $y_{t-k} - \sum_{j=1}^{L} w_{t-k+j} = 1$, and $y_i - \sum_{j=1}^{L} w_{i+j} = 1$ for all $i \in S$. Hence, by (EC.18) and (EC.22),

$$\begin{aligned}
RHS \geq & \sum_{i\in S}(i-d_i)V\left(y_i - \sum_{j=1}^{L} w_{i+j}\right) \\
& + \left(\frac{Q+\Delta(\alpha_t,\beta_t)V - (\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1} - (k-1)V - \overline{V}\right)\left(y_{t-k} - \sum_{j=1}^{L} w_{t-k+j}\right) + \overline{V}y_t \\
= & \sum_{i\in S}(i-d_i)V + \frac{Q+\Delta(\alpha_t,\beta_t)V - (\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1} - (k-1)V \\
= & \frac{Q+\Delta(\alpha_t,\beta_t)V - (\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1} \\
\geq & \, x_t.
\end{aligned}$$

Therefore, inequality (5) holds.

Case (4): $\max\{L-k, 1\} \leq q \leq L-2$ and $p \leq k-1$. The condition "$p \leq k-1$" implies that $t-p \geq t-k+1 \geq t-\alpha_t+1 \geq 2$, and the condition "$q \leq L-2$" implies that $t+q+1 \leq t+L-1 \leq T$. Thus, the latest start-up period before period $t+1$ is at least 2, and the earliest shut-down period after period t is at most T. Note that

$$\begin{aligned}
RHS \geq & \, V\left(\sum_{j=[L-k]^{+}+1}^{L} \min\{L-j, j-1\}w_{t+j} + \sum_{j=1}^{L-k}(j-1)w_{t+j}\right) + \sum_{i\in S}(i-d_i)V\left(y_i - \sum_{j=1}^{L} w_{i+j}\right) + \overline{V}y_k \\
\geq & \, V\min\{L-q-1, q\}w_{t+q+1} + \sum_{i\in S}(i-d_i)V\left(y_i - \sum_{j=1}^{L} w_{i+j}\right) + \overline{V}y_k \\
= & \min\{L-q-1, q\}V + \sum_{i\in S}(i-d_i)V\left(y_i - \sum_{j=1}^{L} w_{i+j}\right) + \overline{V},
\end{aligned} \tag{EC.23}$$

where the second inequality holds because $[L-k]^{+}+1 \leq q+1 \leq L$, and the equality holds because $w_{t+q+1} = 1$. By inequality (2a), $\sum_{i=(t+q+1)-L+1}^{t+q+1} u_i \leq y_{t+q+1} = 0$, which implies that $u_i = 0$ for all $i \in [(t+q+1)-L+1, t+q+1]_{\mathbb{Z}}$. This in turn implies that $(t+q+1)-L+1 \geq t-p+1$, or equivalently, $p \geq L-q-1$. If $p = L-q-1$, then $\min\{L-q-1, q\} = \min\{q, p\}$, and thus by (EC.20) and (EC.23), $RHS \geq x_t$. Hence, it suffices to consider the situation where $p \geq L-q$. In this situation, $t+q-L \in [t'+1, t-1]_{\mathbb{Z}} \subseteq S$. This is because (i) if $k \geq L$, then $t' = t-L$, and thus $t+q-L \in [t-L+1, t-1]_{\mathbb{Z}} = [t'+1, t-1]_{\mathbb{Z}}$; and (ii) if $k < L-1$, then $t' = t-k$, and thus $t+q-L \in [t-k+1, t-1]_{\mathbb{Z}} = [t'+1, t-1]_{\mathbb{Z}}$ (as the condition "$q \leq L-2$" implies that $t+q-L \leq t-1$, while the conditions "$p \leq k-1$" and "$p \geq L-q$" imply that $t+q-L \geq t-k+1$). Hence, by (EC.21), $\sum_{i\in S\cap[t-p,t+q-L]_{\mathbb{Z}}}(i-d_i) \geq p+q+1-L$. Note that $y_j = 1$ and $w_j = 0$ for all $j \in [t-p, t+q]_{\mathbb{Z}}$. Thus, $y_i - \sum_{j=1}^{L} w_{i+j} = 1$ for all $i \in [t-p, t+q-L]_{\mathbb{Z}}$. Hence,

$$\sum_{i\in S}(i-d_i)V\left(y_i - \sum_{j=1}^{L} w_{i+j}\right) \geq \sum_{i\in S\cap[t-p,t+q-L]_{\mathbb{Z}}}(i-d_i)V\left(y_i - \sum_{j=1}^{L} w_{i+j}\right)$$

$$
\begin{aligned}
&= \sum_{i \in S \cap [t-p, t+q-L]_{\mathbb{Z}}} (i - d_i) V \\
&\geq (p + q + 1 - L) V. && \text{(EC.24)}
\end{aligned}
$$

By (EC.20), (EC.23), and (EC.24), we have

$$
\begin{aligned}
RHS &\geq \min\{L - q - 1, q\} V + (p + q + 1 - L) V + \overline{V} \\
&\geq \min\{(L - q - 1) + (p + q + 1 - L), q\} V + \overline{V} = \min\{p, q\} V + \overline{V} \geq x_t.
\end{aligned}
$$

Therefore, inequality (5) holds.

Case (5): $\max\{L - k, 1\} \leq q \leq L - 2$ and $p \geq k$. The condition "$q \leq L - 2$" implies that $t + q + 1 \leq t + L - 1 \leq T$. Thus, the earliest shut-down period after period t is at most T. We first show that

$$
\sum_{i \in S \cap [t-k+1, t+q-L]_{\mathbb{Z}}} (i - d_i) \geq k + q - L. \tag{EC.25}
$$

Note that $t + q - L \geq t - k$. If $t + q - L = t - k$, then inequality (EC.25) holds because both the left-hand side and right-hand side are zero. If $t + q - L \geq t - k + 1$, then $t + q - L \in [t' + 1, t - 1]_{\mathbb{Z}} \subseteq S$ (because (i) if $k \geq L$, then $t' = t - L$ and $t + q - L \in [t - L + 1, t - 1]_{\mathbb{Z}} = [t' + 1, t - 1]_{\mathbb{Z}}$; and (ii) if $k < L$, then $t' = t - k$ and $t + q - L \in [t - k + 1, t - 1]_{\mathbb{Z}} = [t' + 1, t - 1]_{\mathbb{Z}}$). Thus, by (EC.21), inequality (EC.25) holds. Note that $y_j = 1$ and $w_j = 0$ for all $j \in [t - k + 1, t + q]_{\mathbb{Z}}$. Hence, $y_i - \sum_{j=1}^{L} w_{i+j} = 1$ for all $i \in [t - k + 1, t + q - L]_{\mathbb{Z}}$. Thus, by (EC.25),

$$
\begin{aligned}
\sum_{i \in S} (i - d_i) V \left(y_i - \sum_{j=1}^{L} w_{i+j} \right) &\geq \sum_{i \in S \cap [t-k+1, t+q-L]_{\mathbb{Z}}} (i - d_i) V \left(y_i - \sum_{j=1}^{L} w_{i+j} \right) \\
&= \sum_{i \in S \cap [t-k+1, t+q-L]_{\mathbb{Z}}} (i - d_i) V \\
&\geq (k + q - L) V.
\end{aligned}
$$

This implies that

$$
\begin{aligned}
& V \left(\sum_{j=[L-k]^+ + 1}^{L} \min\{L - j, j - 1\} w_{t+j} + \sum_{j=1}^{L-k} (j - 1) w_{t+j} \right) + \sum_{i \in S} (i - d_i) V \left(y_i - \sum_{j=1}^{L} w_{i+j} \right) \\
&\geq V \left(\sum_{j=[L-k]^+ + 1}^{L} \min\{L - j, j - 1\} w_{t+j} + \sum_{j=1}^{L-k} (j - 1) w_{t+j} \right) + (k + q - L) V \\
&\geq V \min\{L - q - 1, q\} w_{t+q+1} + (k + q - L) V \\
&= [\min\{L - q - 1, q\} + (k + q - L)] V \\
&\geq \min\{k - 1, q\} V,
\end{aligned}
$$

where the second inequality holds because $[L - k]^+ + 1 \leq q + 1 \leq L$. Note that $y_{t-k} = 1$ and $w_{t-k+j} = 0$ for $j = 1, \ldots, L$, which implies that $y_{t-k} - \sum_{j=1}^{L} w_{t-k+j} = 1$. Hence,

$$
RHS = V \left(\sum_{j=[L-k]^+ + 1}^{L} \min\{L - j, j - 1\} w_{t+j} + \sum_{j=1}^{L-k} (j - 1) w_{t+j} \right) + \sum_{i \in S} (i - d_i) V \left(y_i - \sum_{j=1}^{L} w_{i+j} \right)
$$

$$
\begin{aligned}
&+\left(\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V-\overline{V}\right)\left(y_{t-k}-\sum_{j=1}^{L}w_{t-k+j}\right)+\overline{V}\\
&\geq \min\{k-1,q\}V+\left(\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V-\overline{V}\right)\left(y_{t-k}-\sum_{j=1}^{L}w_{t-k+j}\right)+\overline{V}\\
&=\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V+\min\{k-1,q\}V. \qquad \text{(EC.26)}
\end{aligned}
$$

If $q\geq k-1$, then $\min\{k-1,q\}V=(k-1)V$, and thus by (EC.18) and (EC.26), we have $RHS\geq x_t$. If $q<k-1$, then by (EC.20), $\min\{k-1,q\}V=qV\geq\min\{q,p\}V\geq x_t-\overline{V}$, and thus from (EC.19) and (EC.26), we have $RHS\geq x_t$. Therefore, inequality (5) holds.

Summarizing Cases (1)–(5), we conclude that inequality (5) holds for any element of $\mathcal{D}$. □

A.12. Proof of Proposition 10

Consider any given LP relaxation optimum (x^*,y^*,u^*) of Problem (1) with $x^*=(x_1^*,\ldots,x_T^*)$, $y^*=(y_1^*,\ldots,y_T^*)$, and $u^*=(u_1^*,\ldots,u_T^*)$. For any $t\in[L+1,T]_{\mathbb{Z}}$ and $k\in[2,\beta_t]_{\mathbb{Z}}$, let

$$
\begin{aligned}
\Psi(t,k)=&\ V\left(\sum_{j=[L-k]^+}^{L-1}\min\{L-1-j,j\}u_{t-j}^*+\sum_{j=1}^{L-k-1}ju_{t-j}^*\right)+\min_{S'\subseteq[t',t+k-1]_{\mathbb{Z}}}\sum_{i\in S}(d_i-i)V\left(y_i^*-\sum_{j=0}^{L-1}u_{i-j}^*\right)\\
&+\left(\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V-\overline{V}\right)\left(y_{t+k}^*-\sum_{j=0}^{L-1}u_{t+k-j}^*\right)+\overline{V}y_t^*-x_t^*,
\end{aligned}
$$

where $t'=\min\{t+L,t+k\}$, $S=[t+1,t'-1]_{\mathbb{Z}}\cup S'$, and $d_i=\min\{a\in S\cup\{t+k\}:a>i\}$ for any $i\in S$. Given any $t\in[L+1,T]_{\mathbb{Z}}$ and $k\in[2,\beta_t]_{\mathbb{Z}}$, if $\Psi(t,k)<0$, then $\Psi(t,k)$ measures how much (x^*,y^*,u^*) violates inequality (4).

For any $t\in[L+1,T]_{\mathbb{Z}}$, $k\in[2,\beta_t]_{\mathbb{Z}}$, and $S'\subseteq[t',t+k-1]_{\mathbb{Z}}$, let

$$
\Phi(t,k,S')=\sum_{i\in\{t'-1\}\cup S'}(d_i-i)V\left(y_i^*-\sum_{j=0}^{L-1}u_{i-j}^*\right).
$$

Then,

$$
\begin{aligned}
\Psi(t,k)=&\ V\left(\sum_{j=[L-k]^+}^{L-1}\min\{L-1-j,j\}u_{t-j}^*+\sum_{j=1}^{L-k-1}ju_{t-j}^*\right)\\
&+\sum_{i=t+1}^{t'-2}V\left(y_i^*-\sum_{j=0}^{L-1}u_{i-j}^*\right)+\min_{S'\subseteq[t',t+k-1]_{\mathbb{Z}}}\Phi(t,k,S')\\
&+\left(\frac{Q+\Delta(\alpha_t,\beta_t)V-(\tau_{t1}+\tau_{t2})\underline{C}}{\alpha_t+\beta_t+1}-(k-1)V-\overline{V}\right)\left(y_{t+k}^*-\sum_{j=0}^{L-1}u_{t+k-j}^*\right)+\overline{V}y_t^*-x_t^*.
\end{aligned}
$$

The quantity "$\min_{S'\subseteq[t',t+k-1]_{\mathbb{Z}}}\Phi(t,k,S')$" can be obtained by solving a shortest path problem on a directed acyclic network $\mathbb{G}=(\mathbb{V},\mathbb{A})$. The nodes and arcs of $\mathbb{G}$ are as follows:

(i) Node set $\mathbb{V}=\{t'-1,t',\ldots,t+k\}$, which is a set of time indices from $t'-1$ to $t+k$.

(ii) Arc set $\mathbb{A}=\{(t_1,t_2):t'-1\leq t_1<t_2\leq t+k\}$, where the length of (t_1,t_2) is $(t_2-t_1)V(y_{t_1}^*-\sum_{j=0}^{L-1}u_{t_1-j}^*)$.

The shortest path from node $t'-1$ to node $t+k$ represents the optimal choice of the set S', where a node $t_1 \in [t', t+k-1]_{\mathbb{Z}}$ is on the shortest path if and only if t_1 is included in the set S'. The shortest distance from node $t'-1$ to node $t+k$ is equal to $\min_{S' \subseteq [t', t+k-1]_{\mathbb{Z}}} \Phi(t,k,S')$ (see Damcı-Kurt et al. (2016) for a similar shortest path approach to separation algorithm development).

For each t and k, solving the shortest path problem requires $\mathcal{O}(T^2)$ time. Hence, $\Psi(t,k)$ can be determined in $\mathcal{O}(T^2)$ time. Given (x^*, y^*, u^*), the most violated inequality can be obtained in $\mathcal{O}(T^4)$ time by selecting t and k with the smallest $\Psi(t,k)$ value. $\square$

A.13. Proof of Proposition 11

Consider any given LP relaxation optimum (x^*, y^*, u^*) of Problem (1) with $x^* = (x_1^*, \ldots, x_T^*)$, $y^* = (y_1^*, \ldots, y_T^*)$, and $u^* = (u_1^*, \ldots, u_T^*)$. For any $t \in [2, T-L]_{\mathbb{Z}}$ and $k \in [2, \alpha_t]_{\mathbb{Z}}$, let

$$\Psi(t,k) = V\left(\sum_{j=[L-k]^+}^{L} \min\{L-j, j-1\} w_{t+j}^* + \sum_{j=1}^{L-k-1} (j-1) w_{t+j}^*\right) + \min_{S' \subseteq [t+k-1, t']_{\mathbb{Z}}} \sum_{i \in S} (i-d_i) V \left(y_i^* - \sum_{j=1}^{L} w_{i+j}^*\right)$$
$$+ \left(\frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1} - (k-1)V - \overline{V}\right)\left(y_{t-k}^* - \sum_{j=1}^{L} w_{t-k+j}^*\right) + \overline{V} y_t^* - x_t^*,$$

where $w_t^* = y_{t-1}^* - y_t^* + u_t^*$, $t' = \max\{t-L, t-k\}$, $S = S' \cup [t'+1, t-1]_{\mathbb{Z}}$, and $d_i = \max\{a \in \{t-k\} \cup S : a < i\}$ for any $i \in S$. Given any $t \in [2, T-L]_{\mathbb{Z}}$ and $k \in [2, \alpha_t]_{\mathbb{Z}}$, if $\Psi(t,k) < 0$, then $\Psi(t,k)$ measures how much (x^*, y^*, u^*) violates inequality (5).

For any $t \in [2, T-L]_{\mathbb{Z}}$, $k \in [2, \alpha_t]_{\mathbb{Z}}$, and $S' \subseteq [t+k-1, t']_{\mathbb{Z}}$, let

$$\Phi(t,k,S') = \sum_{i \in S' \cup \{t'+1\}} (i-d_i) V \left(y_i^* - \sum_{j=1}^{L} w_{i+j}^*\right).$$

Then,

$$\begin{aligned}
\Psi(t,k) = {} & V\left(\sum_{j=[L-k]^+}^{L} \min\{L-j, j-1\} w_{t+j}^* + \sum_{j=1}^{L-k-1} (j-1) w_{t+j}^*\right) \\
& + \sum_{i=t'+2}^{t-1} V \left(y_i^* - \sum_{j=1}^{L} w_{i+j}^*\right) + \min_{S' \subseteq [t+k-1, t']_{\mathbb{Z}}} \Phi(t,k,S') \\
& + \left(\frac{Q + \Delta(\alpha_t, \beta_t)V - (\tau_{t1} + \tau_{t2})\underline{C}}{\alpha_t + \beta_t + 1} - (k-1)V - \overline{V}\right)\left(y_{t-k}^* - \sum_{j=1}^{L} w_{t-k+j}^*\right) + \overline{V} y_t^* - x_t^*.
\end{aligned}$$

The quantity "$\min_{S' \subseteq [t+k-1, t']_{\mathbb{Z}}} \Phi(t,k,S')$" can be obtained in $\mathcal{O}(T^2)$ time by solving a shortest path problem on a directed acyclic network similar to that in the proof of Proposition 10. Hence, $\Psi(t,k)$ can be determined in $\mathcal{O}(T^2)$ time for each combination of t and k. Given (x^*, y^*, u^*), the most violated inequality can be obtained in $\mathcal{O}(T^4)$ time by selecting t and k with the smallest $\Psi(t,k)$ value. $\square$