Integrated Stochastic Optimal Self-Scheduling for Two-Settlement Electricity Markets

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

Yongpei Guan

Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA, guan@ise.ufl.edu

The complexity of current electricity wholesale markets and the increased volatility of electricity prices due to the intermittent nature of renewable generation make independent power producers (IPPs) face significant challenges to submit offers. This challenge increases for those owning traditional coal-fired thermal generators and renewable generation. In this paper, an integrated stochastic optimal strategy is proposed for an IPP using the self-scheduling approach through its participation in both day-ahead and real-time markets (i.e., two-settlement electricity markets) as a price taker. In the proposed approach, the IPP submits an offer for all periods to the day-ahead market, for which a multistage stochastic programming setting is explored for providing real-time market offers for each period as a recourse. This strategy has the advantage of achieving overall maximum profits for both markets in the given operational time horizon. Such a strategy is theoretically proved to be more profitable than alternative self-scheduling strategies, as it takes advantage of the continuously realized scenario information of the renewable energy output and real-time prices over time. To improve computational efficiency, we explore polyhedral structures to derive strong valid inequalities, including convex hull descriptions for certain special cases, thus strengthening the formulation of our proposed model. Polynomial-time separation algorithms are then established for the derived exponential-sized inequalities to speed up the branch-and-cut process. Finally, both numerical and real case studies demonstrate the potential of the proposed strategy.

Key words: innovative formulation; renewable generation; self-scheduling; stochastic optimization
History:

1. Introduction

Environmentally friendly policies and government incentives, such as tax benefits and renewable energy certificate programs, have resulted in renewable energy becoming increasingly important in global power systems (U.S. DOE 2008, Agora Energiewende and Sandbag 2019). This increase in renewable generation can help significantly reduce emissions and has profoundly influenced electricity market economics and operations. This trend brings challenges for market participants, such as independent power producers (IPPs), whose generation portfolios consist of coal-fired generators and renewable generation, as they must well utilize their renewable generation resources when participating in wholesale electricity markets.

The wholesale electricity markets in the U.S. and Europe mainly consist of day-ahead and real-time markets for electricity trading, plus an ancillary service/reserve market to ensure that the system remains reliable. The day-ahead market is cleared the day before the actual operating day, and an independent system operator (ISO) (or transmission system operator (TSO)) takes offers from sellers (e.g., IPPs) and buyers (e.g., utilities) and clears the market, thus ensuring the power system's balance and security. In this financial commitment, day-ahead locational marginal prices (LMPs) are calculated as the basis to compensate the sellers (Conejo et al. 2005) and charge the buyers (Ott 2003). The real-time market is cleared 10 to 15 minutes before the actual operating time. The ISO takes the real-time generation offers to cover the load discrepancy between the day-ahead commitment and the real-time load. Real-time LMPs are calculated for financial transactions. Within this framework, a market participant (e.g., an IPP) is allowed to participate in one, two or all of the day-ahead, real-time, and ancillary service markets by submitting offers to the ISO.

For the current U.S. wholesale markets, an IPP owning a coal-fired generator can submit a three-part offer, a self-commitment offer or a self-scheduling offer. For the three-part offer, the IPP submits a start-up offer, a minimum-energy offer, and an energy offer curve for ISOs to make unit commitment decisions and economic dispatch amounts (Anderson and Philpott 2002, Fleten and Pettersen 2005, Kwon and Frances 2012). For the self-commitment offer, the IPP submits the unit commitment status of a generator for each operating hour, and the ISO decides the economic dispatch amounts (Papavasiliou et al. 2015). For the self-scheduling offer, the IPP offers the generation amount for each period for day-ahead and/or real-time markets as a price taker, whatever the LMPs are. The IPP decides unit commitment and economic dispatch to meet the generation amount committed in the offer (California ISO 2018, MISO 2020, Conejo et al. 2002, Li et al. 2007). After the market clearing procedure is completed, (i) for the three-part offer submissions, the IPP whose offer is selected is notified of the awards, including the unit commitment status and generation amount at each period, (ii) the self-commitment IPPs are notified of the generation amounts at each committed operating hour, and (iii) the self-scheduling IPPs are notified of the price at each operating hour for financial settlement.

In practice, a significant number of IPPs submit self-commitment and self-scheduling offers. Midcontinent ISO (MISO), the largest ISO in the U.S., reports that approximately 76% of coal-fired and 33% of gas-fired power plants in MISO are running self-scheduling (UtilityDive 2020). The Southwest Power Pool (SPP) conducts analysis and reports the detailed reasons (SPP 2019). Some reasons are unavoidable such as periodic power plant emission performance tests, cogeneration processes for electricity and others, and specific requirements following long-term maintenance agreements with original equipment manufacturers. Other reasons provided in the industry include

fulfilling long-term (some are over 40 years) bilateral contracts, accommodating the gas market (the forward gas market is cleared before the day-ahead electricity market and the gas amount to be purchased must be decided before the electricity offer is awarded), and smooth operation transitions between two operating days considering the commitments of generators at the beginning and end of the operational horizon (Papavasiliou et al. 2015).

On top of these practical considerations, the non-convexities in the day-ahead market, due to commitment decisions in the electricity market clearing problem, result in energy prices not covering all generator operating costs. That is, day-ahead LMPs may not cover start-up and no-load (or minimum-load) costs, because these costs are not counted in the calculation to obtain LMPs. Thus, a generator's revenue obtained through following the ISO's instruction may not reach that obtained through self-scheduling (Gribik et al. 2007) (see an illustrative example in Online Supplement A.1 for the reader's reference). This indicates that self-scheduling can be helpful for an IPP even in the deterministic setting. Furthermore, in a multi-interval real-time market without non-convexities, as shown in Hua et al. (2019), due to dynamics of the intervals used in market clearance and ramping constraint of generators, following the ISO's instructions could lead to a loss that cannot be compensated by energy prices. When the uncertainty is considered, the problem becomes more challenging to be solved. Although a multistage equilibrium is provided in Philpott et al. (2016) to address hydro-thermal electricity systems under uncertainty, the system optimization problem is a convex program, and commitment decisions are not considered.

Meanwhile, the recent increased penetration of renewable energy and the fluctuating electricity demand, along with the traditional unexpected outages of generation and transmission components, have led to both the day-ahead and real-time electricity prices being significantly volatile (Valenzuela and Mazumdar 2003). Many IPPs owning coal-fired generators have invested in renewable energy and are investing in energy storage (Dunn et al. 2011, Eyer and Corey 2010), partially with the consideration of obtaining renewable energy certificates to meet state renewable portfolio standard requirements. For example, in the U.S., the law in California requires 33% of electricity production to be from renewable sources by 2020, i.e., electric companies must hold renewable energy certificates equivalent to 33% of their electricity sales (Evomarkets 2015). Considering that the day-ahead offer amount for each operating hour must be fulfilled in real-time, it accordingly affects the real-time offer amounts due to the thermal generator's physical constraints, uncertain renewable generation output, and uncertain prices. Thus, making day-ahead offer amount decisions needs to consider real-time decisions as a recourse. Specifically, it is important for coal-fired generator owners with renewable generation, to derive efficient approaches to participate in both day-ahead and real-time markets (Löhndorf et al. 2010, 2013). The invested energy storage, together with flexible ramping capability of the thermal generator, can accommodate the intermittency of

renewable generation to obtain relatively stable profit. Thus, submitting offers combining thermal, renewable, and energy storage sources is more profitable than submitting each one separately by taking advantage of flexible ramping and energy storage.

In this paper, we investigate the optimal offer submission strategy for an IPP using the self-scheduling option to maximize its total profit under both price and renewable generation output uncertainties, by participating in both day-ahead and real-time markets. The IPP owns a coal-fired generator and a wind farm, with energy storage (e.g., compressed air energy storage with a large capacity of approximately several hundred megawatts and a lead acid battery with a relatively low capacity of below 100 MW; see Eyer and Corey (2010), IEC (2011)). We utilize stochastic programming approaches (Birge and Louveaux 2011, Wallace and Fleten 2003) to set up the self-scheduling model and eventually provide the optimal offer amounts.

As the day-ahead market is cleared one day ahead of the real-time market, a two-stage stochastic programming model can reflect the reality of the offer submission process. Under this setting, an IPP submits the offer amounts to the day-ahead market (i.e., the first-stage decisions) considering possible offer amounts to the real-time market (i.e., the second-stage decisions) under different realizations of uncertain real-time market prices and renewable generation outputs as a recourse. Accordingly, two-stage stochastic programming approaches (Carøe and Schultz 1998) have been extensively utilized to enhance the offer submission process by solving self-scheduling unit commitment problems. Studies of different market settings (Plazas et al. 2005, Heredia et al. 2010), risk-constrained bidding strategies (Li et al. 2007), and hydrothermal scheduling (Wu et al. 2008) have been conducted, and various solution approaches, including Benders decomposition (Wang et al. 2008, Zheng et al. 2013) and Lagrangian relaxation (Papavasiliou and Oren 2013), have been elaborated to solve the corresponding models.

In the above two-stage framework, the second-stage decisions, in terms of offer amounts to the real-time market corresponding to each period, are assumed to be made at the beginning of the operating day (e.g., before midnight). In practice, the real-time market is cleared every 15 minutes, and the real-time offer amounts can be submitted dynamically by adapting to uncertainty realization at each period, thus utilizing the partial uncertainty that has been realized in the real-time markets. For example, the real-time market offer amount for period t can be decided at period $t-1$, instead of before midnight. In this way, the renewable generation amounts and electricity prices from periods 1 to $t-1$ are realized. Uncertain information from period t to the end of the operational horizon will also be more accurately presented, as the forecast becomes further accurate when approaching the actual specific time slot. Thus, making an offer for period t based on the historical information from periods 1 to $t-1$, with the anticipation of uncertainty from period t to the end of the operational horizon, fits well in practice and leads to an improved

recourse for the first-stage decision with an increased overall expected profit. Therefore, in the second stage, we take a scenario-tree-based multistage stochastic programming approach (Cerisola et al. 2009) to decide the real-time offer amounts submitted to the real-time market corresponding to each period, where the scenario tree can also help express the uncertain parameter dependencies among different periods. Thus, this approach fits the real-time market operations better than the two-stage approach, as multistage decisions are made by capturing the uncertainty dynamics over the periods and incorporating multistage forecasting information with varying accuracy and by modeling the renewable generation and electricity price dependencies among consecutive periods, as reflected in practice. Multistage stochastic programming approaches were originally proposed in the 1990s for power system operators (Carpentier et al. 1996, Takriti et al. 1996) to address load uncertainty for vertically integrated utilities. Transmission constraints were later incorporated in Wu et al. (2007). A multistage stochastic self-scheduling model for IPPs participating in the day-ahead and ancillary service markets is proposed in Cerisola et al. (2009) and that for the real-time market only is proposed in Pan and Guan (2016b). In Morales et al. (2010), a three-stage stochastic programming model is developed for a wind power producer to participate in day-ahead, adjustment, and balancing markets to address various uncertainties, where each stage corresponds to day-ahead, adjustment, and balancing markets, respectively.

To accommodate both day-ahead and real-time markets, our innovative model includes a two-stage stochastic programming framework with a multistage stochastic program embedded in the second stage as a recourse. This leads to a large-scale deterministic equivalent formulation, for which the computational complexity increases as the size of the scenario tree grows. To reduce the computational burden and ensure the approach is practically useful, we explore strong valid inequalities, which are customized for our innovative formulation and derived through the polyhedral study of the corresponding polytope, as cutting planes (Guan et al. 2009) to strengthen the formulation of our proposed model and speed up the branch-and-cut algorithm used to solve the problem. Our main contributions are as follows:

1) We offer an innovative self-scheduling strategy for IPPs to participate in both day-ahead and real-time markets, instead of the traditional approaches to participate in each market separately, to ensure a high overall profit, by coordinating generation assets and day-ahead and real-time markets. The proposed strategy can well capture current electricity market practices, including the renewable generation amount and electricity price dependencies among different periods.

2) We develop a new framework that combines two-stage and multistage stochastic programming modeling approaches, in which the overall two-stage framework corresponds to the one offer submission process in the day-ahead market and overall real-time market operations, and the

multistage model corresponds to the multiple real-time offer submissions within the operating day. This framework enables the IPPs to develop an optimal self-scheduling strategy to exploit the synergy of renewable generation, the ramping capability of coal-fired generation, and energy storage in a way that cannot be accomplished separately by two-stage and multistage stochastic programming models.

3) We explore strong valid inequalities to efficiently solve the proposed model. Specifically, we derive several families of strong valid inequalities, including convex hull descriptions for certain special cases and multi-period strong valid inequalities for general cases, by exploring the structures of the specific model and the scenario tree, to strengthen the formulation.

4) We show that theoretically our innovative model outperforms alternatives in terms of profit. Our case studies confirm this finding and also show the effectiveness of our proposed strong valid inequalities.

The remainder of this paper is organized as follows. We first introduce the notation and model for the optimal self-scheduling strategy problem in Section 2, and then compare our proposed model with alternatives to demonstrate that ours offers the best objective value in Section 3. Subsequently, we explore strong valid inequalities to tighten the derived formulation and the corresponding separation algorithms, thus improving the computational performance of our proposed model, in Section 4. We conduct numerical and real case studies in Section 5 to verify the strength of our model and the effectiveness of our strong valid inequalities in speeding up the branch-and-cut algorithm used to solve the problem. The paper concludes in Section 6.

2. Offer Submission Process and the Model

Under the current electricity market framework, an IPP can submit offers at different periods to day-ahead and real-time electricity markets to an ISO, as shown in Figure 1. To participate in the day-ahead market, for each operating day (e.g., day m), the IPP is required to submit its day-ahead offer (e.g., the offer amount for each period of day m) at noon the day before (e.g., day $m-1$). At some time in the afternoon (depending on different markets), the ISO runs the market-clearing procedure (i.e., day-ahead unit commitment runs) and publishes the day-ahead LMPs for each period of day m (MISO 2020). During the operating day (e.g., day m), the IPP can submit its real-time offer by participating in the real-time market for each period (e.g., each operating hour) of day m. The ISO publishes the real-time LMPs for each period after running security-constrained economic dispatch problems. By using the self-scheduling strategy, the IPP can submit the offer amounts to participate in both day-ahead and real-time markets. Under this policy, each offer is guaranteed to be taken.

To model the offer submission process for an IPP participating in both markets, we propose a two-stage framework, where the IPP submits day-ahead and real-time offers in the first and second
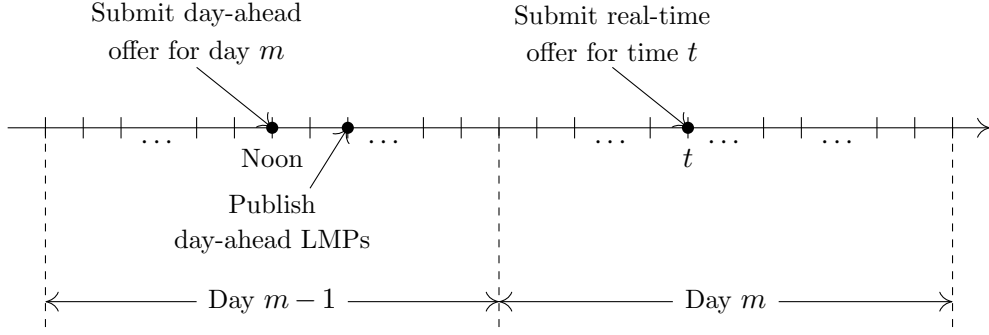
Figure 1   Offer submission process for day-ahead and real-time markets

stages, respectively. We assume there are T periods in a given day (denoted by day m) and use Figure 2 to illustrate this framework when there is no uncertainty involved. We use a solid node t ($t \in \{1, 2, \ldots, T\}$) in Figure 2 to represent the deterministic system state at period t of day m. We use d_t and r_t to represent the day-ahead and real-time offer amounts, respectively, for each period t of day m. In the first stage (on day $m-1$), the IPP submits day-ahead offers $(d_1, d_2, \ldots, d_T)$ in one shot before noon, and in the second stage (on day m), the IPP submits real-time offers $(r_1, r_2, \ldots, r_T)$ individually at each period.
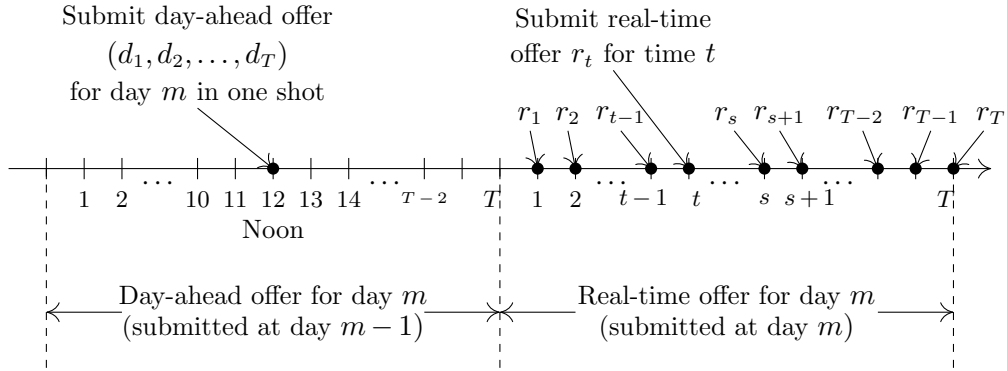

Figure 2   Brief two-stage offer submission process

During the real-time offer submission process where uncertainty is generally involved, the IPP can submit its offer when close to the actual operating period (e.g., one period ahead), so an adaptive offer submission approach can be utilized by observing the uncertain real-time prices and renewable generation outputs stage by stage as time goes on. Thus, to determine the real-time offer amounts, we apply the scenario tree to represent the possible realizations of the real-time prices and renewable generation outputs, and accordingly consider multistage decisions. We assume that the uncertain parameters (i.e., real-time price and renewable generation output) during the real-time operation follow a discrete-time stochastic process that evolves in a finite probability space, and use a scenario tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ with T stages to describe the possible realizations, as shown in

Figure 3. Note here that $T = 24$ for most industry practices. In addition, we add a parent node (denoted as node 0, including T periods) to the root node of $\mathcal{V}$ (denoted as node 1) to indicate the offer submission to the day-ahead market and let $\hat{\mathcal{V}} = \mathcal{V} \cup \{0\}$. Each scenario node $i \in \mathcal{V}$ on stage t of the tree provides the system state that can be distinguished by information available up to stage t. Accordingly, corresponding to each node $i \in \mathcal{V}$, we let $t(i)$ represent its time period in the real-time operation, $\mathcal{P}(i)$ the set of nodes along the path from the root node (i.e., node 1) to node i, and p_i the probability associated with the state represented by node i. We use $\mathcal{S}_t$ to denote the set of nodes in stage t. The decisions corresponding to each node i are assumed to be made after observing the realizations of the problem parameters along the path from the root node to this node i, but are nonanticipative with respect to future realizations.
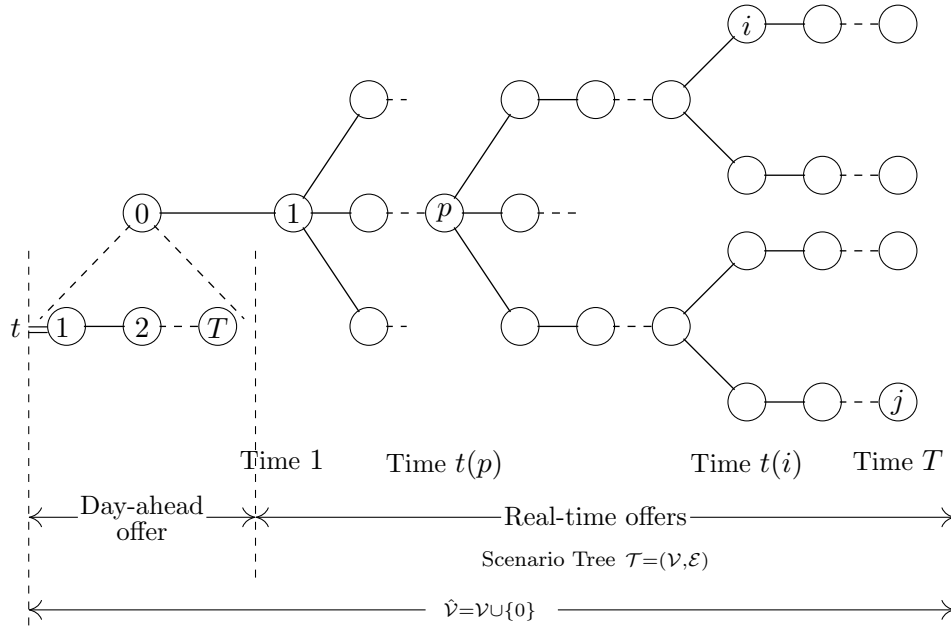
Figure 3   Proposed two-stage offer submission process

Based on the above two-stage framework and scenario tree setting, we develop an optimal self-scheduling strategy model for the IPP in this framework, including day-ahead offer amounts as the first-stage decisions and real-time multistage offer amounts as a recourse for the first-stage decisions, and denote the model as TMS.

To mathematically describe model TMS, we first introduce the physical characteristics of the thermal generator as follows. We let L (ℓ) denote its minimum-up (-down) time limit, $\overline{C}$ $(\underline{C})$ its generation upper (lower) bound if the generator is online, V its ramp-up/-down rate limits, $\overline{V}$ its start-up/shut-down ramp rate limits, U (D) its start-up (shut-down) cost, and a nondecreasing convex function $f(\cdot)$ the fuel cost as a convex function of its electricity generation amount. In

addition, we let $Q_t^d(\xi) > 0$ denote the day-ahead electricity price at period t corresponding to scenario ξ and $Q_i^r > 0$ denote the real-time electricity price at node i for each $i \in \mathcal{V}$. In addition, we let W_i denote the renewable generation at node i for each $i \in \mathcal{V}$, λ_c (λ_d) $\in (0,1)$ the storage charge (discharge) efficiency, H_c (H_d) the unit storage charge (discharge) cost, and Λ the energy storage capacity.

For the decision variables, the first-stage decisions have been denoted by d_t for any $t \in \{1,\dots,T\}$ in Figure 2, i.e., the offer amount for each period t to the day-ahead market. In the second stage, for each node $i \in \mathcal{V}$, we let r_i represent the offer amount to the real-time market, x_i the actual electricity generation amount, h_i^c (h_i^d) the electricity charged to (discharged from) the energy storage, and s_i the storage level at node i. We also let the binary variables (y_i, u_i) represent the unit commitment status of the generator, with y_i indicating the online/offline status and u_i indicating the start-up decision.

Therefore, model TMS can be formulated as follows:

$$\mathcal{Z}^{\mathrm{TMS}} = \max\bigg\{ g^{\mathrm{TMS}}(d,y,u,x,r,s,h^c,h^d) := \mathrm{E}\bigg(\sum_{t=1}^{T} Q_t^d(\xi) d_t \bigg) + \sum_{i\in\mathcal{V}} p_i \bigg(Q_i^r r_i - \bigg(U u_i$$

$$+ D(y_{i-} - y_i + u_i) + f(x_i) \bigg) - H_c h_i^c - H_d h_i^d \bigg) : (d,y,u,x,r,s,h^c,h^d) \in \mathcal{X}^{\mathrm{TMS}} \bigg\},$$

where $\mathcal{X}^{\mathrm{TMS}}$ is defined by the following constraints:

$$y_i - y_{i-} \leq y_k, \quad \forall i \in \mathcal{V}\setminus\{1\}, \forall k \in \mathcal{H}_L(i), \tag{1a}$$

$$y_{i-} - y_i \leq 1 - y_k, \quad \forall i \in \mathcal{V}\setminus\{1\}, \forall k \in \mathcal{H}_\ell(i), \tag{1b}$$

$$y_i - y_{i-} \leq u_i, \quad \forall i \in \mathcal{V}\setminus\{1\}, \tag{1c}$$

$$u_i \leq \min\{y_i,\ 1 - y_{i-}\}, \quad \forall i \in \mathcal{V}\setminus\{1\}, \tag{1d}$$

$$\underline{C} y_i \leq x_i \leq \overline{C} y_i, \quad \forall i \in \mathcal{V}, \tag{1e}$$

$$x_i - x_{i-} \leq V y_{i-} + \overline{V}(1 - y_{i-}), \quad \forall i \in \mathcal{V}\setminus\{1\}, \tag{1f}$$

$$x_{i-} - x_i \leq V y_i + \overline{V}(1 - y_i), \quad \forall i \in \mathcal{V}\setminus\{1\}, \tag{1g}$$

$$s_i = s_{i-} + \lambda_c h_i^c - h_i^d/\lambda_d, \quad \forall i \in \mathcal{V}, \tag{1h}$$

$$d_{t(i)} + r_i = x_i + W_i - h_i^c + h_i^d, \quad \forall i \in \mathcal{V}, \tag{1i}$$

$$s_0 = 0;\ y_i \in \{0,1\}, 0 \leq s_i \leq \Lambda, r_i \geq 0, h_i^c \geq 0, h_i^d \geq 0,\ \forall i \in \mathcal{V},$$

$$u_i \in \{0,1\},\ \forall i \in \mathcal{V}\setminus\{1\};\ d_t \geq 0,\ \forall t = 1,\dots,T, \tag{1j}$$

where $\mathcal{H}_r(i) = \{k \in \mathcal{V}(i) : 0 \leq t(k) - t(i) \leq r - 1\}$ represents all of the nodes that occur no more than r periods after the occurrence of node i, including itself, $\mathcal{V}(i)$ represents the set of all of the descendants of node i, including itself, and $f(\cdot)$ is a quadratic function, i.e., $f(x_i) = ax_i^2 + bx_i +$

cy_i, and approximated by a piece-wise linear function throughout the paper for computational convenience.

The objective of TMS is to maximize the expected total profit, which is equal to the revenue obtained from the day-ahead and real-time markets minus the start-up, shut-down, fuel, and storage charge/discharge costs. Constraints (1a) (resp. (1b)) represent the minimum-up (resp. minimum-down) time requirements. That is, if the generator starts up (resp. shuts down) at node i, then it should stay online (resp. offline) for all of the nodes in the set $\mathcal{H}_L(i)$ (resp. $\mathcal{H}_\ell(i)$). Constraints (1c) and (1d) define the relationships between y and u. Constraints (1e) describe the upper and lower bounds of the electricity generation amount if the generator is online at node i. Constraints (1f) (resp. (1g)) describe the ramp rate limits, including the start-up/shut-down ramp rate. Finally, constraints (1h) and (1i) represent the storage balance and power balance at each node, respectively. In this model, for each node $i \in \mathcal{V}$, h_i^c and h_i^d cannot be strictly positive simultaneously in the optimal solution due to the energy loss in the charging/discharging activities, as shown in the following Proposition 1. We delay all the proofs in the paper to the Online Supplement.

PROPOSITION 1. *For any optimal solution $(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*})$ of model TMS, we have $h_i^{c*} h_i^{d*} = 0$ for any $i \in \mathcal{V}$.*

In summary, model TMS is a stochastic programming with recourse model in which the multistage stochastic programming decisions for the real-time market provide the recourse for the day-ahead offer decisions.

REMARK 1. Although TMS is a risk-neutral model with the expected profit as the objective function, different risk measures can be applied to formulate its risk-averse counterparts. We provide a risk-averse model that minimizes a weighted sum of the negative expected profit and conditional value-at-risk (CVaR) of the loss function (i.e., negative profit) (Rockafellar and Uryasev 2000). When uncertain parameters (i.e., real-time price and renewable generation output) have finite support, which is the case for our scenario-tree based setting, the CVaR-based risk-averse model can be reformulated using linear constraints, as shown in Online Supplement B.2. It leads to a model with the same computational tractability as model TMS.

3. Alternative Models and Comparisons

In this section, we compare model TMS with alternatives studied and implemented by IPPs.

3.1. Day-Ahead Self-Scheduling and Real-Time Adaptive Offer Submission Model

In this model, we provide an alternative two-stage stochastic programming framework, with a multistage stochastic real-time offer submission process embedded in the second stage as a recourse, denoted as TMSR. This model is similar to TMS, but it makes the unit commitment decisions

only in the first stage. That is, TMSR makes decisions on the day-ahead offer amounts and unit commitment status in the first stage before the real-time market operation starts, with the adaptive real-time offer amounts in the real-time market following different scenarios explored by the scenario tree $\mathcal{T}$ as a recourse. The detailed formulation is provided in Online Supplement C.1. We refer to a similar model (Garces and Conejo 2010), in which an IPP decides the generator status and weekly power generation quantity before the price is realized and submits real-time offers to adjust the power quantity when the prices are realized using a scenario tree.

3.2. Two-Phase Offer Submission Models

In this part, we provide two two-phase offer submission models. For both, we obtain the day-ahead offer amounts and unit commitment decisions in the first phase. Then, using the first-phase decisions as inputs, in the second phase we take two different approaches to decide the real-time offer amounts, resulting in two strategies.

First Phase: Two-Stage Stochastic Programming Model. To obtain the first-phase decisions, we develop a traditional two-stage stochastic programming model, denoted as TS, in which the day-ahead offer amounts and unit commitment decisions are made in the first stage with the second-stage offer amounts corresponding to each scenario as a recourse, as shown in Figure 4, where set $\mathcal{S}$ is defined as the collection of all of the scenarios (i.e., all of the possible paths from the root node 1 to each leaf node at period T) in the scenario tree $\mathcal{T}$. As compared to TMSR described in Section 3.1, each scenario in TS represents the possible realizations of renewable generation amounts and electricity prices from period 1 to the end of the operational horizon. There are no nonanticipativity constraints among scenarios as illustrated in the scenario-tree setting for the multistage case in TMSR. In an example that reflects the TS model (Garcia-Gonzalez et al. 2008), a pumped-storage unit and wind farm submit the day-ahead offer in the first stage and adjust the real-time power output in the second stage as a recourse after observing a set of possible price scenarios.

Based on the obtained day-ahead offer amounts and unit commitment decisions following model TS, real-time offers are traditionally submitted immediately before the real-time market operations begin by respecting one scenario (in set $\mathcal{S}$) for the operational horizon. However, as mentioned in Section 1, in practice the real-time market is cleared every period and the real-time offer amounts can be submitted dynamically at each period. Meanwhile, at the beginning of the real-time market operations, it is not accurate to estimate the price or renewable generation output scenario that will happen in the entire operational horizon. Therefore, using the first-phase decisions (i.e., the day-ahead offer amounts and unit commitment decisions from model TS) as inputs, we decide the second-phase decisions (i.e., real-time offer amounts) through two realistic methods:
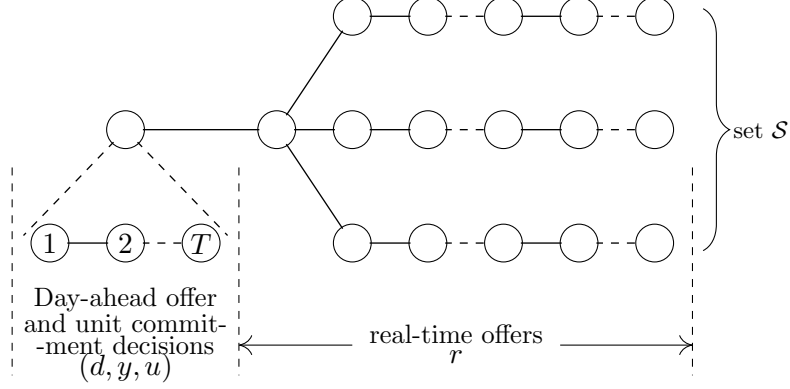
Figure 4   Offer submission process for TS

1. A multistage stochastic real-time offer submission by respecting the scenario tree $\mathcal{T}$,
2. An offer submission with respect to an arbitrary scenario in which the outcome for each period is set as the mean value of all of the scenarios,

which lead to two second-phase models. We first describe model TS as follows, to obtain the day-ahead offer amounts and unit commitment decisions for the first phase.

$$
\mathcal{Z}^{\mathrm{TS}} = \max \Bigg\{ g^{\mathrm{TS}}(d,y,u,x,r,s,h^c,h^d) := \mathtt{E}\bigg(\sum_{t=1}^{T} Q_t^d(\xi) d_t \bigg) - \sum_{t=2}^{T} \bigg(U u_t + D(y_{t-1} - y_t + u_t) \bigg)
$$

$$
+ \mathtt{E}\bigg(\sum_{t=1}^{T} \Big(Q_t^r(\eta) r_t(\eta) - f(x_t(\eta)) - H_c h_t^c(\eta) - H_d h_t^d(\eta) \Big) \bigg) : (d,y,u,x,r,s,h^c,h^d) \in \mathcal{X}^{\mathrm{TS}} \Bigg\}, \quad (2)
$$

where $\mathcal{X}^{\mathrm{TS}}$ is described in Online Supplement C.2, $Q_t^r(\eta)$ denotes the real-time price at t corresponding to scenario $\eta \in \mathcal{S}$, and $W_t(\eta)$ denotes the renewable generation output at t corresponding to scenario $\eta \in \mathcal{S}$. By solving model TS (i.e., (2)), we obtain an optimal solution (d^*, y^*, u^*) and use this as an input in the second-phase models to decide the offer amounts for the real-time markets. The corresponding two second-phase models are as follows.

Second Phase: Multistage Stochastic Real-Time Offer Submission Model. In this model, denoted as MSV, based on the obtained first-phase decision (d^*, y^*, u^*) and following the scenario tree $\mathcal{T}$, we obtain the real-time offer amounts for the real-time markets corresponding to each node in the scenario tree. The corresponding second-phase model can be formulated as:

$$
\mathcal{Z}^{\mathrm{MSV}} = \max \Bigg\{ g^{\mathrm{MSV}}(x,r,s,h^c,h^d) := \mathtt{E}\bigg(\sum_{t=1}^{T} Q_t^d(\xi) d_t^* \bigg) - \sum_{t=2}^{T} \bigg(U u_t^* + D(y_{t-1}^* - y_t^* + u_t^*) \bigg)
$$

$$
+ \sum_{i \in \mathcal{V}} p_i \bigg(Q_i^r r_i - f(x_i) - H_c h_i^c - H_d h_i^d \bigg) : (x,r,s,h^c,h^d) \in \mathcal{X}^{\mathrm{MSV}} \Bigg\},
$$

where $\mathcal{X}^{\mathrm{MSV}}$ is described in Online Supplement C.2. This model respects actual real-time operations, because real-time markets are cleared stage by stage and uncertain real-time electricity prices and

renewable generation amounts are realized continuously as time passes. The real-time offer amounts obtained from this model that explores the uncertainty evolution are more realistic than those obtained by respecting one scenario in set $\mathcal{S}$ in model TS.

Second Phase: Mean-Value Scenario Model. From this model, denoted as MEV, we obtain the real-time offer amounts based on the first-phase decision (d^*, y^*, u^*), by only considering the scenario that represents the mean value of all possible scenarios in set $\mathcal{S}$, as shown in Figure 5. Thus, the corresponding second-phase model can be formulated as:

$$\mathcal{Z}^{\text{MEV}} = \max\left\{ g^{\text{MEV}}(x, r, s, h^c, h^d) := \mathbb{E}\left(\sum_{t=1}^{T} Q_t^d(\xi)d_t^*\right) - \sum_{t=2}^{T}\left(Uu_t^* + D(y_{t-1}^* - y_t^* + u_t^*)\right) \right.$$
$$\left. + \sum_{t=1}^{T}\left(\bar{Q}_t^r r_t - f(x_t) - H_c h_t^c - H_d h_t^d\right) : (x, r, s, h^c, h^d) \in \mathcal{X}^{\text{MEV}} \right\}, \tag{3}$$

where $\bar{Q}_t^r$ (resp. $\bar{W}_t$) represents the mean value of uncertain real-time prices (resp. renewable generation amounts) at period t, i.e., $\bar{Q}_t^r = \sum_{i \in \mathcal{V}_t} p_i Q_i^r$ and $\bar{W}_t = \sum_{i \in \mathcal{V}_t} p_i W_i$ with $\mathcal{V}_t = \{i \in \mathcal{V} : t(i) = t\}$ for any $t \in \{1, \ldots, T\}$, and $\mathcal{X}^{\text{MEV}}$ is described in Online Supplement C.2.
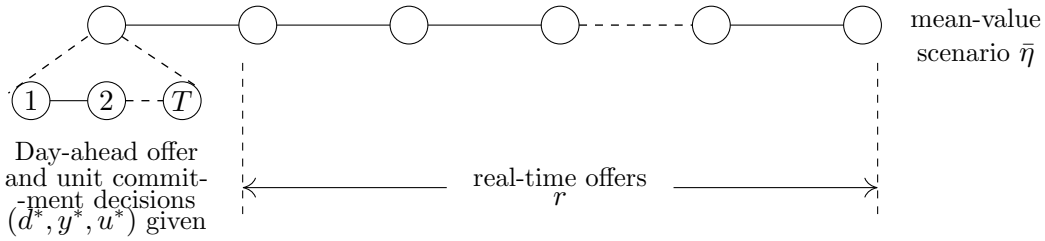
Figure 5   Offer submission process for MEV

Considering the mean-value scenario and then deciding the real-time offer at the beginning of the real-time market operations is straightforward. However, this model may not be profitable enough, as the mean-value scenario is not sufficient to capture the actual realizations of the uncertain parameters.

3.3.   Day-Ahead Only Model

Finally, we present an offer submission model for the IPP by only participating in the day-ahead market, denoted as DA. This model is widely used by IPPs (Arroyo and Conejo 2000, Conejo et al. 2002, Simoglou et al. 2012), as they only need to submit the day-ahead offer amounts to maximize their total expected profits. There are no recourse decisions to be made for the real-time market. Accordingly, the IPP runs this offer submission model (see details in Online Supplement C.3) with the day-ahead price.

3.4. Model Comparison

We present four alternative models in total to realistically submit offers by participating in day-ahead and/or real-time markets, and assess their differences. Plus model TMS, all of the five models measure the same thing, i.e., the expected total profit for the day-ahead and real-time offer submissions, under the same underlying evolved discrete-time stochastic process. It can be observed that Proposition 1 also applies to all models, i.e., h_i^c and h_i^d cannot be strictly positive simultaneously in the optimal solution for each of them. These models demonstrate that our proposed model TMS allows flexibility in unit commitment decisions in either day-ahead or real-time stages, as long as the offer amounts can be satisfied during the real-time operations. Thus, TMS produces a larger profit than TMSR does, which requires the unit commitment decisions made during the day-ahead stage. In addition, as the optimal solution of MSV is a feasible solution of TMSR, MSV leads to a smaller profit than TMSR does. MEV only considers the mean-value scenario, so it may produce a smaller profit than that provided by MSV. The theoretical relationships among different models can be shown in the following proposition.

PROPOSITION 2. *The optimal objective values for different offer submission strategies satisfy* $\mathcal{Z}^{TMS} \geq \mathcal{Z}^{TMSR} \geq \mathcal{Z}^{MSV}$, *and* $\mathcal{Z}^{MSV} \geq \mathcal{Z}^{MEV}$ *if* $W_i = W_j$ *when* $t(i) = t(j)$ *for each pair* $(i,j) \in \mathcal{V}$. *Furthermore, if* $\bar{Q}_t^d = \bar{Q}_t^r$ *for any* t *with* $\bar{Q}_t^d = \mathtt{E}(Q_t^d(\xi))$, *we have* $\mathcal{Z}^{DA} \geq \mathcal{Z}^{MEV}$, *and* $\mathcal{Z}^{TMSR} \geq \mathcal{Z}^{DA}$ *if* $W_i = W_j$ *when* $t(i) = t(j)$ *for each pair* $(i,j) \in \mathcal{V}$.

Proposition 2 shows that TMS provides the highest profit for an IPP. However, the curse of dimensionality prevents it from being solved quickly, as the numbers of scenarios and binary variables increase significantly with the scenario tree size. To obtain the benefits from TMS in a computationally efficient way for the IPPs, we focus on deriving its strong formulations to improve the computational performance in the next section. More specifically, we consider strengthening the model TMS by studying the polyhedral structure of the generator's physical constraints (1a) - (1g). We define this set of constraints as $P := \{(x, y, u) \in \mathbb{R}^{|\mathcal{V}|} \times \mathbb{B}^{(2|\mathcal{V}|-1)} : (1a) - (1g)\}$ with $\mathbb{B} := \{0, 1\}$ and use conv(P) to denote the convex hull of the set of feasible points in P. Before exploring the details, we describe the following observation at the end of this section.

REMARK 2. The polyhedral results of conv(P) can be applied to strengthen related problems with set P embedded, such as the problems considering two or more units in the portfolio.

4. Strengthening Model TMS

In this section, we study the polytope conv(P) to strengthen the model TMS in two steps by deriving convex hull results for certain special cases in Section 4.1 and strong valid inequalities for the general case in Section 4.2.

4.1. Convex Hull Results

We focus on two-period and certain three-period cases by deriving their convex hull representations, which provide the perfect linear programming formulations for the corresponding two- and three-period problems. The derived strong valid inequalities used to describe the convex hulls can be applied to strengthen the original model TMS effectively.

For the two-period case, there is only one root node with several scenarios in the second period (i.e., leaf nodes), each with a corresponding given probability. We let n represent the total number of leaf nodes and $\mathcal{N} = \{1, 2, \ldots, n\}$ represent the set of leaf nodes, i.e., the scenario nodes in the second period, which share the same parent node, denoted as i^-, $\forall i \in \mathcal{N}$. Because there are only two periods, without loss of generality, we assume $L = \ell = 1$; accordingly, the minimum-up/-down time constraints (1a) and (1b) are not needed here. Thus, the corresponding original set P can be described as $P_2 = \Big\{ (x, y, u) \in \mathbb{R}^{n+1} \times \mathbb{B}^{n+1} \times \mathbb{B}^n :$

$$y_i - y_{i^-} - u_i \leq 0, \quad u_i - y_i \leq 0, \quad u_i + y_{i^-} \leq 1, \quad \forall i \in \mathcal{N}, \tag{4a}$$

$$-x_i + \underline{C} y_i \leq 0, \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{4b}$$

$$x_i - \overline{C} y_i \leq 0, \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{4c}$$

$$x_i - x_{i^-} \leq V y_{i^-} + \overline{V}(1 - y_{i^-}), \quad x_{i^-} - x_i \leq V y_i + \overline{V}(1 - y_i), \quad \forall i \in \mathcal{N} \Big\}. \tag{4d}$$

Note that in P_2, we have x and y variables defined for all of the nodes and u variables defined for each node in the second period. That is, there is no start-up decision for the root node. In this way, the derived inequalities can be applied recursively for each node (not only the root node) in the scenario tree $\mathcal{T}$, e.g., we can apply them to each non-leaf node $i \in \mathcal{V}$ and its child nodes. Meanwhile, considering the industrial practices for coal-fired generators, we assume $\overline{C} - \underline{C} - 2V \geq 0$ and $\overline{C} - \overline{V} - V \geq 0$. Considering the relationships among different scenarios in model TMS, we can derive the following strong valid inequalities.

PROPOSITION 3. *The inequalities*

$$x_{i^-} \leq \overline{V} y_{i^-} + (\overline{C} - \overline{V})(y_i - u_i), \quad \forall i \in \mathcal{N}, \tag{5}$$

$$x_i \leq (\overline{V} + V) y_i - V u_i + (\overline{C} - \overline{V} - V)(y_j - u_j), \quad \forall i, j \in \mathcal{N}, \tag{6}$$

$$x_i - x_{i^-} \leq \overline{V} y_i - \underline{C} y_{i^-} + (\underline{C} + V - \overline{V})(y_i - u_i), \quad \forall i \in \mathcal{N}, \tag{7}$$

$$x_{i^-} - x_i \leq \overline{V} y_{i^-} - \underline{C} y_i + (\underline{C} + V - \overline{V})(y_j - u_j), \quad \forall i, j \in \mathcal{N}, \tag{8}$$

$$x_i - x_j \leq (\overline{V} + V) y_i - V u_i - \underline{C} y_j + (\underline{C} + V - \overline{V})(y_k - u_k), \quad \forall i, j, k \in \mathcal{N}, i \neq j, \tag{9}$$

are valid for conv(P_2).

Based on inequalities (5) - (9), we can further obtain a stronger result.

THEOREM 1. *The linear programming description of conv(P_2) can be described as*

$$Q_2 := \left\{ (x,y,u) \in \mathbb{R}^{3n+2} : (4a)-(4b),(5)-(9) \right.$$

$$\left. u_i \geq 0, \quad \forall i \in \mathcal{N} \right\}. \tag{10}$$

To make the formulation of model TMS stronger, we extend our study to the three-period case in which the uncertainty is explored in the third period and there is only one scenario node in each of the first and second periods. We denote the scenario nodes in the third period as set $\mathcal{N} = \{1,2,\ldots,n\}$, where n represents the total number of leaf nodes (i.e., the scenario nodes in the third period). For notational brevity, for any $i \in \mathcal{N}$, we let i_k^- represent the kth-fold parent of node i for any integer $k \geq 0$ with $i_0^- = i$ and $i_1^- = i^-$. For this case, the convex hull results are different depending on the minimum-up/-down time limits. We illustrate the results for the case in which $L = \ell = 1$ here with other cases described in Online Supplement D.6 for brevity. The corresponding P can be described as $P_3^1 := \left\{ (x,y,u) \in \mathbb{R}^{n+2} \times \mathbb{B}^{n+2} \times \mathbb{B}^{n+1} : \right.$

$$y_i - y_{i^-} - u_i \leq 0, \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{11a}$$

$$u_i - y_i \leq 0, \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{11b}$$

$$u_i + y_{i^-} \leq 1, \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{11c}$$

$$-x_i + \underline{C}y_i \leq 0, \quad \forall i \in \mathcal{N} \cup \{i_2^-, i^-\}, \tag{11d}$$

$$x_i - \overline{C}y_i \leq 0, \quad \forall i \in \mathcal{N} \cup \{i_2^-, i^-\}, \tag{11e}$$

$$x_i - x_{i^-} \leq V y_{i^-} + \overline{V}(1 - y_{i^-}), \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{11f}$$

$$x_{i^-} - x_i \leq V y_i + \overline{V}(1 - y_i), \quad \forall i \in \mathcal{N} \cup \{i^-\} \left.\right\}. \tag{11g}$$

As compared to P_2, P_3^1 contains one more node: the parent node of the root node of P_2. Accordingly, the linear programming description of conv(P_3^1) can be described as follows:

THEOREM 2. *For a three-period problem in which $L = \ell = 1$, conv(P_3^1) can be described as $Q_3^1 = \left\{ (x,y,u) \in \mathbb{R}^{3n+5} : (11a)-(11d), \right.$*

$$u_i \geq 0, \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{12a}$$

$$x_{i_2^-} \leq \overline{V}y_{i_2^-} + (\overline{C} - \overline{V})(y_{i^-} - u_{i^-}), \tag{12b}$$

$$x_{i_2^-} \leq \overline{V}y_{i_2^-} + V(y_{i^-} - u_{i^-}) + (\overline{C} - \overline{V} - V)(y_i - u_i), \quad \forall i \in \mathcal{N}, \tag{12c}$$

$$x_{i^-} \leq \overline{V}y_{i^-} + (\overline{C} - \overline{V})(y_{i^-} - u_{i^-}), \tag{12d}$$

$$x_{i^-} \leq \overline{V}y_{i^-} + (\overline{C} - \overline{V})(y_i - u_i), \quad \forall i \in \mathcal{N}, \tag{12e}$$

$$x_i \leq (\overline{V} + V)y_i - Vu_i + (\overline{C} - \overline{V} - V)(y_j - u_j), \quad \forall i \in \mathcal{N}, j \in \mathcal{N} \cup \{i^-\}, \tag{12f}$$

$$x_{i^-} - x_{i_2^-} \leq \overline{V}y_{i^-} - \underline{C}y_{i_2^-} + (\underline{C} + V - \overline{V})(y_i - u_i), \quad \forall i \in \mathcal{N}, \tag{12g}$$

$$x_{i_2^-} - x_{i^-} \leq \overline{V} y_{i_2^-} - \underline{C} y_{i^-} + (\underline{C} + V - \overline{V})(y_{i^-} - u_{i^-}), \tag{12h}$$

$$x_i - x_{i^-} \leq \overline{V} y_i - \underline{C} y_{i^-} + (\underline{C} + V - \overline{V})(y_i - u_i), \quad \forall i \in \mathcal{N} \cup \{i^-\}, \tag{12i}$$

$$x_{i^-} - x_i \leq \overline{V} y_{i^-} - \underline{C} y_i + (\underline{C} + V - \overline{V})(y_j - u_j), \quad \forall i \in \mathcal{N}, j \in \mathcal{N} \cup \{i^-\}, \tag{12j}$$

$$x_i - x_{i_2^-} \leq (\overline{V} + V)y_i - V u_i - \underline{C} y_{i_2^-} + (\underline{C} + V - \overline{V})(y_j - u_j), \quad \forall i \in \mathcal{N}, j \in \mathcal{N} \cup \{i^-\}, \tag{12k}$$

$$x_{i_2^-} - x_i \leq \overline{V} y_{i_2^-} - \underline{C} y_i + V(y_{i^-} - u_{i^-}) + (\underline{C} + V - \overline{V})(y_j - u_j), \quad \forall i, j \in \mathcal{N}, \tag{12l}$$

$$x_{i_2^-} - x_i \leq \overline{V} y_{i_2^-} - \underline{C} y_i + (\underline{C} + 2V - \overline{V})(y_{i^-} - u_{i^-}), \quad \forall i \in \mathcal{N}, \tag{12m}$$

$$x_i - x_j \leq (\overline{V} + V)y_i - V u_i - \underline{C} y_j + (\underline{C} + V - \overline{V})(y_k - u_k),$$
$$\forall i, j \in \mathcal{N}, i \neq j, k \in \mathcal{N} \cup \{i^-\}\Big\}. \tag{12n}$$

The proofs are similar to those for Theorem 1 and are thus omitted here.

REMARK 3. Note that P_3^1, in which $L = \ell = 1$, is equivalent to the problem without minimum-up/-down time constraints. It then can be considered as a relaxation of the cases in which the minimum-up/-down times are larger than 1 in model TMS. Thus, the derived inequalities in Q_3^1 are valid for the original model TMS.

REMARK 4. In the convex hull representations, there are no inequalities containing variables corresponding to four or more scenario nodes. It follows that the number of inequalities to describe the convex hulls is in the order of n^3, i.e., $\mathcal{O}(n^3)$, and thus no separation algorithms are required.

4.2. General Multi-Period Strong Valid Inequalities

We further strengthen the model TMS by exploring strong valid inequalities for conv(P) covering multi-period nodes. For notational brevity, we let $[a, b]_{\mathbb{Z}}$ represent the set of integers $\{a, a+1, \ldots, b\}$ with $[a, b]_{\mathbb{Z}} = \emptyset$ if $b < a$.
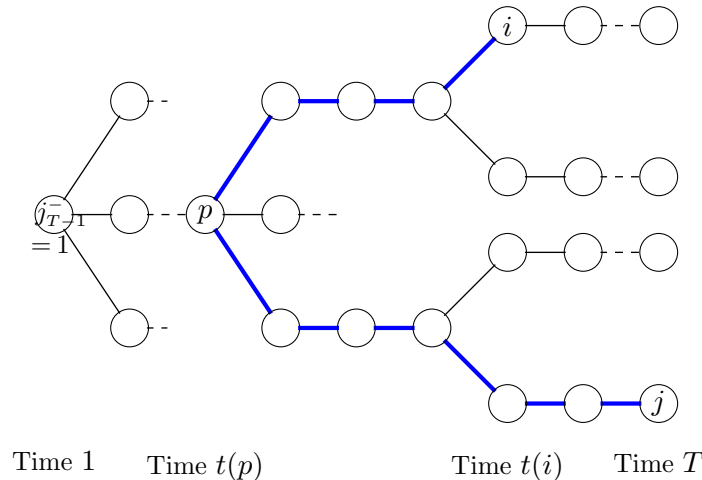

Figure 6   **Multistage stochastic scenario tree**

We consider a general multi-period scenario tree in model TMS as described in Figure 6. For any two nodes i and j in the figure, we let $p = \arg\max\{t(k) : k \in \mathcal{P}(i) \cap \mathcal{P}(j)\}$. This indicates a path from nodes i to j passing through node p as the node with the smallest period, i.e., the bold solid line path in Figure 6. We define $\mathcal{P}(i,p) = \mathcal{P}(i) \setminus \mathcal{P}(p)$ and the distance between nodes i and j $dist(i,j) = |\mathcal{P}(i,p)| + |\mathcal{P}(j,p)|$.

We first present a family of strong valid inequalities incorporating the nodes on the same scenario (Pan and Guan 2016a) by utilizing the physical thermal generator constraints in model TMS as follows.

PROPOSITION 4. *For each $k \in \{[2, T-2]_{\mathbb{Z}} : \overline{C} - \overline{V} - (k-1)V > 0\}$, the inequality*

$$x_{i_k^-} \leq \overline{V} y_{i_k^-} + V \sum_{h=1}^{k-1} \left(y_{i_h^-} - \sum_{n=h}^{\min\{k,h+L-1\}} u_{i_n^-} \right) + (\overline{C} - \overline{V} - (k-1)V) \left(y_i - \sum_{h=0}^{\min\{k,L-1\}} u_{i_h^-} \right), \quad (13)$$

is valid for conv(P) for each $i \in \mathcal{V}$ such that $t(i) \in [\min\{k+2, L+1\}, T]_{\mathbb{Z}}$. Furthermore, it is facet-defining for conv(P) when one of the following conditions is satisfied: (1) $L \leq 3$ and $t(i) = T$; (2) $L \leq 3$ and $k = \lfloor (\overline{C} - \overline{V})/V \rfloor + 1$ for each $i \in \mathcal{V}$ such that $t(i) \in [\min\{k+2, L+1\}, T]_{\mathbb{Z}}$.

Inequality (13) tightens the generation upper bound for each node in the tree. Now, we propose several families of strong valid inequalities (i.e., (14) - (16)) incorporating nodes covering multiple scenarios in the scenario tree by utilizing the specific structures of model TMS and the scenario tree.

PROPOSITION 5. *For each pair $(i,j) \in \mathcal{V}$ with $i \notin \mathcal{P}(j)$, $j \notin \mathcal{P}(i)$, and i being a leaf node of $\mathcal{V}$, the inequality*

$$x_i - x_j \leq (\underline{C} + kV)y_i - \underline{C}y_j - \sum_{h=0}^{\min\{L-1,k-1\}} \left(\underline{C} + (k-h)V - \overline{V} \right) u_{i_h^-}, \quad (14)$$

where $k = dist(i,j)$ such that $k \in \{[2, T-1]_{\mathbb{Z}} : \overline{C} - \underline{C} - kV > 0\}$, is valid and facet-defining for conv(P).

Inequality (14) describes the relationships between the generation amounts of any two nodes (e.g., i and j) on different scenarios in the scenario tree. The condition $\overline{C} - \underline{C} - kV > 0$ guarantees that the generator is able to ramp up k times starting from its generation lower bound $\underline{C}$ or ramp down k times starting from its generation upper bound $\overline{C}$. With respect to the scenario tree, this condition implies that the generator is able to ramp up m times based on the generation amount at node p (i.e., x_p) along one path to node i and meanwhile ramp down n times based on x_p along another path to node j such that $m + n = k$ and $p = \arg\max\{t(s) : s \in \mathcal{P}(i) \cap \mathcal{P}(j)\}$. Thus, the maximum difference in the generation amounts between nodes i and j is kV if $y_i = y_j = 1$. Similarly, we can propose the following inequality (15) by linking nodes i^- and j that are not on the same scenario.

PROPOSITION 6. *For each pair $(i,j) \in \mathcal{V}$ with $i^- \notin \mathcal{P}(j)$ and $j \notin \mathcal{P}(i^-)$, the inequality*

$$x_{i^-} - x_j \leq \overline{V}y_{i^-} - \underline{C}y_j + (\underline{C}+V-\overline{V})\left(y_j - \sum_{h=0}^{L-1} u_{j_h^-}\right) + (k-1)V(y_i - u_i) - \sum_{h=1}^{\min\{L-1,k\}} \left(k-h\right)Vu_{i_h^-},$$
(15)

where $k = dist(i^-,j)$ such that $k \in \{[2, T-2]_{\mathbb{Z}} : \overline{C} - \underline{C} - kV > 0\}$, is valid and facet-defining for conv(P) when $t(i^-) = t(j)$ or $t(i) = t(j)$.

The proof is similar to that for Proposition 5 and is thus omitted here.

All of the proposed inequalities until now in this section are in polynomial size in terms of the input size of the scenario tree and are at most in the order of $|\mathcal{V}|^2$, i.e., $\mathcal{O}(|\mathcal{V}|^2)$. Therefore, we do not need separation procedures for them to speed up the branch-and-cut algorithm. In the following, we further explore and discover a more general family of inequalities in exponential size for model TMS to bound the generation difference between $x_{i_k^-}$ and x_j for any two nodes i and j in $\mathcal{V}$.

PROPOSITION 7. *For each pair $(i,j) \in \mathcal{V}$ with $i_k^- \notin \mathcal{P}(j)$, $j \notin \mathcal{P}(i_k^-)$, and $\min\{t(i_k^-), t(j)\} \geq 2$, $S_0 = [1, \hat{n}-1]_{\mathbb{Z}}$, and $S \subseteq [\hat{n}+1, k-1]_{\mathbb{Z}}$ with $\hat{n} = \min\{t(i_k^-)-2, L-2\}$ if $\min\{t(i_k^-)-2, L-2\} \geq L/2$ and $\hat{n} = \max\{1, L+1-t(i_k^-)\}$ otherwise, the inequality*

$$x_{i_k^-} - x_j \leq \overline{V}y_{i_k^-} - \underline{C}y_j + V\sum_{n \in S_0}\left(y_{i_{k-n}^-} - \sum_{m=0}^{\min\{L-1,n+w\}} u_{i_{k-n+m}^-}\right)$$

$$+ V\sum_{n \in S \cup \{\hat{n}\}}\left(g_n - n\right)\left(y_{i_{k-n}^-} - \sum_{m=0}^{L-1} u_{i_{k-n+m}^-}\right) + \psi(y,u) + \phi(u),$$
(16)

is valid and facet-defining for conv(P), where $k = dist(i_k^-, j)$ such that $k \in \{[2, 2T-2]_{\mathbb{Z}} : \overline{C} - \underline{C} - kV > 0\}$, $g_n = \min\{a \in S \cup \{k\} : a > n\}$, w is a nonnegative integer for which $t(i_{k+w}^-) = 2$, $\psi(y,u) = (\underline{C}+V-\overline{V})(y_i - \sum_{m=0}^{L-1} u_{i_m^-})$ or $(\underline{C}+V-\overline{V})(y_j - \sum_{m=0}^{L-1} u_{j_m^-})$, and $\phi(u) = V\sum_{m:m\geq 1, t(i_{k+m}^-)\geq T-L+1} mu_{i_{k+m}^-} + V\sum_{m:2\leq t(i_{k+m}^-)\leq T-L, m\leq L-1} \min\{L-m-1, m\}u_{i_{k+m}^-}$.

Separation: Because the size of inequalities (16) is exponential, we explore a separation scheme to find the most violated inequality (correspondingly the set S in (16)) in polynomial time. For a given point $(\hat{x}, \hat{y}, \hat{u}) \in \mathbb{R}_+^{3|\mathcal{V}|-1}$, to find the most violated inequality (16) corresponding to each combination of (i, j, k), we construct a shortest path problem on a directed acyclic graph $\mathbb{G} = (\mathbb{V}, \mathbb{A})$, as shown in Figure 7, with the node and arc sets described as follows:

(i) Node set $\mathbb{V} = \{s, t\} \cup \mathbb{V}'$ with s representing the source node, t representing the sink node, and $\mathbb{V}' = \{\hat{n}, \hat{n}+1, \ldots, k-1, k\}$ representing a set of nodes from $i_{k-\hat{n}}^-$ to i in (16).

(ii) Arc set $\mathbb{A} = \{(s, \hat{n}), (k, t)\} \cup \mathbb{A}'$ with $\mathbb{A}' = \cup_{\hat{n}\leq n_1 < n_2 \leq k}\{(n_1, n_2)\}$. We let w_{ij} represent the cost of arc (i, j) for each $(i, j) \in \mathbb{A}$ and provide the details as follows:

(1) $w_{s\hat{n}} = \overline{V}\hat{y}_{i_k^-} - \underline{C}\hat{y}_j + V\sum_{n\in S_0}(\hat{y}_{i_{k-n}^-} - \sum_{m=0}^{\min\{L-1,n+w\}}\hat{u}_{i_{k-n+m}^-}) - \hat{x}_{i_k^-} + \hat{x}_j$;

(2) $w_{kt} = \psi(\hat{y},\hat{u}) + \phi(\hat{u})$;

(3) $w_{n_1 n_2} = V(n_2 - n_1)(\hat{y}_{i_{k-n_1}^-} - \sum_{j=0}^{L-1}\hat{u}_{i_{k-n_1-j}^-})$ for $n_1, n_2 \in [\hat{n},k]_{\mathbb{Z}}$ and $n_1 < n_2$.

The shortest path from nodes s to t represents the maximum violation of inequality (16) if the value is negative, and the visited nodes in $\mathbb{V}'$ on the shortest path determine the set S. Because it is an acyclic graph and there are $\mathcal{O}(T^2)$ arcs and $\mathcal{O}(T)$ nodes, the shortest path can be found in $\mathcal{O}(T^2)$ time for each combination of (i,j,k). Therefore, there is an $\mathcal{O}(|\mathcal{V}|^2 T^2)$ time algorithm to solve the separation problem for all (i,j) with $dist(i_k^-, j) = k$, considering the fact that k is bounded above by a constant number, i.e., $(\overline{C} - \overline{V})/V$.
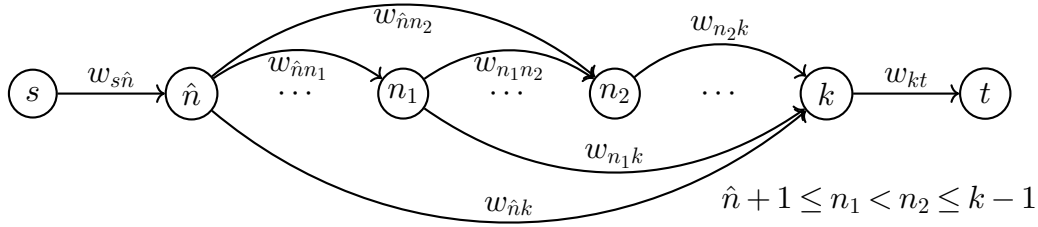
Figure 7   **Directed acyclic graph**

PROPOSITION 8. *Given a point $(\hat{x},\hat{y},\hat{u}) \in \mathbb{R}_+^{3|\mathcal{V}|-1}$, there exists an $\mathcal{O}(|\mathcal{V}|^2 T^2)$ time separation algorithm to find the most violated inequality (16), if one exists.*

At the end of this section, we would like to point out that our study not only strengthens the formulation of our proposed model TMS but also enriches the existing approaches for related polyhedral structures, e.g., Lee et al. (2004), Rajan and Takriti (2005), Ostrowski et al. (2012), Morales-España et al. (2013), Pan et al. (2015), Damcı-Kurt et al. (2016), Gentile et al. (2017), Queyranne and Wolsey (2017), among others. Our polytope is more general in terms of the model setting, polytope structure, and derived inequalities. First, these existing approaches consider deterministic problems, which can be considered as a special case for our stochastic programming setting with only one scenario with probability one. All of our derived results for a specific scenario could be applied to the corresponding deterministic cases. We have also derived inequalities, e.g., (14)-(16), linking different scenarios (by incorporating the nodes at different scenarios in the scenario tree), which cannot be generated from the deterministic case. Second, for the embedded deterministic problem corresponding to each scenario, we consider all of the physical constraints of each generator (e.g., generation upper/lower bounds, minimum-up-/down time, ramping rate, and logical constraints) in one single set (i.e., P), as compared to these existing approaches considering

a part of the physical constraints. For instance, Damcı-Kurt et al. (2016) consider the ramping polytope without the minimum-up/-down time constraints. In general, for a MILP formulation, if we consider more constraints when deriving the strong valid inequalities, the formulation will be tighter.

5. Numerical and Case Studies

In this section, we provide numerical results and case studies that demonstrate the benefits of our proposed self-scheduling strategy model presented in Section 2. We numerically compare our proposed model TMS with alternative models. We also verify the effectiveness of our proposed strong valid inequalities in a branch-and-cut framework by comparing our approach with default CPLEX. We finally illustrate the performance of our model TMS through empirical studies using real data. All of the numerical experiments and case studies were conducted on a computer node with two AMD Opteron 2378 Quad-Core Processors at 2.4 GHz with 4 GB of memory. IBM ILOG CPLEX 12.3 with a single thread was utilized as the MIP solver, and the time limit was set at one hour per run except when specified.

5.1. Problem Settings

The first set of instances are based on a coal-fired generator from a modified IEEE 118-bus system, available online at `motor.ece.iit.edu/data/SCUC_118`, with $\underline{C} = 61\text{MW}, \overline{C} = 300\text{MW}, V = 58\text{MW/h}, \overline{V} = 97\text{MW/h}$, and $U = \$300$. We let $H_c = \$8, H_d = \$10, \Lambda = 30\text{MW}$, and $\lambda_c = \lambda_d = 0.99$ for the energy storage settings, similar to those described in (Eyer and Corey 2010, IEC 2011). To obtain the electricity price, we follow the historical electricity price published by PJM (PJM 2018). For the renewable generation output, we consider $W_i \in [0, \overline{C}/2], \forall i \in \mathcal{V}$. To test the variations of the proposed instances, we consider different types of minimum-up/-down time limits (i.e., $L(\ell)$), scenario-tree structures, and total periods (i.e., T). We let $L = \ell = 2, 3$, and 4. We denote K as the number of branches for each non-leaf node in the scenario tree $\mathcal{T}$, with each branch assigned the same probability. We set $K = 2, 3$, and 4. We let $T \in \{10, 11, 12\}$ for the case $K = 2$, $T \in \{6, 7, 8\}$ for the case $K = 3$, and $T \in \{5, 6, 7\}$ for the case $K = 4$. Thus, various instances are generated based on different combinations of $L(\ell), K$, and T. For each combination, we test three randomly generated instances[1] and report the average result.

5.2. Comparison of Offer Submission Models

In this section, we compare the five models presented in Sections 2 and 3 and report the results in Table 1. The CPLEX under default settings (henceforth denoted as "default CPLEX") is used to solve them. For each combination of $L(\ell), K$, and T, we report the objective value and CPU time for each model. When TMS cannot be solved to default optimality (i.e., 0.01%) within the

Table 1   Model Comparisons

$L(\ell)$	K	T	TMS		TMSR		MSV		MEV		DA	
			$\mathcal{Z}^{\mathrm{TMS}}$	CPU secs (TGap(%))	$\mathcal{Z}^{\mathrm{TMSR}}$	CPU secs	$\mathcal{Z}^{\mathrm{MSV}}$	CPU secs	$\mathcal{Z}^{\mathrm{MEV}}$	CPU secs	$\mathcal{Z}^{\mathrm{DA}}$	CPU secs
2	2	10	12623.1	1256.2 (0.05) [1]	11938.2	46.0	11915.0	1141.4	9816.6	1117.9	10289.2	0.02
		11	14795.8	3600 (0.08) [3]	14052.7	235.1	14051.9	5083.5	12336.2	4998.9	12868.4	0.02
		12	14040.8	3600 (0.10) [3]	13114.8	787.9	12171.6	36059.1	10321.6	35674.8	11673.1	0.01
	3	6	7283.1	31.2	6957.3	611.6	6955.0	114.7	6242.7	59.4	6463.1	0.01
		7	10033.8	2513.4 (0.04) [2]	9545.9	53.4	9545.1	898.9	8376.0	872.9	8635.2	0.01
		8	9329.2	3600 (0.73) [3]	8828.5	503.5	8827.0	15980.8	7399.3	15736.1	7848.7	0.01
	4	5	6633.8	46.2	6288.9	498.5	6229.6	79.5	5086.7	52.6	5653.7	0.01
		6	7482.2	3600 (0.06) [3]	7014.9	76.5	7014.8	1227.2	5957.9	1187.9	6273.8	0.01
		7	9500.6	3600 (0.76) [3]	9115.6	1450.8	9073.6	36033.7	7741.0	35654.2	8114.1	0.01
3	2	10	12434.0	434.7	11981.1	51.6	11981.1	993.0	9803.7	969.1	10636.8	0.01
		11	14854.0	3600 (0.06) [3]	14464.4	184.4	14463.8	4951.2	11737.6	4864.2	12185.7	0.02
		12	14737.7	3600 (0.11) [3]	14176.7	793.6	14169.0	33770.0	11575.3	33412.8	12358.3	0.01
	3	6	6557.7	942.7	6268.7	958.2	6266.3	124.7	5751.5	70.4	5966.6	0.01
		7	8650.6	1062 (0.03) [1]	8327.9	54.8	8323.6	906.2	7106.4	880.3	7420.8	0.01
		8	8654.7	3600 (0.20) [3]	8336.1	672.2	8333.1	22211.7	7060.4	21971.2	7260.8	0.01
	4	5	8550.8	41.9	8347.7	175.8	7973.5	52.4	7714.6	48.2	7793.4	0.01
		6	6858.4	3600 (0.08) [3]	6602.9	80.3	6600.8	1204.7	5309.9	1168.9	5652.6	0.01
		7	9407.9	3600 (0.63) [3]	9100.1	1481.7	9098.2	44469.7	8073.2	43804.5	8390.8	0.01
4	2	10	12197.3	3600 (0.05) [3]	11658.2	48.5	11656.3	938.2	9032.4	915.5	9884.3	0.02
		11	14370.6	3600 (0.10) [3]	13843.8	175.2	13822.5	5862.3	11524.1	5788.1	12013.3	0.02
		12	14559.5	3600 (0.62) [3]	13876.8	770.7	13294.8	48482.2	10397.0	34294.6	12074.1	0.02
	3	6	9664.3	29.1	9415.2	411.6	9410.3	89.9	8480.3	65.4	8757.4	0.01
		7	9293.3	3600 (0.06) [3]	8908.5	55.3	8907.0	966.4	7565.2	940.6	7922.5	0.01
		8	8917.1	3600 (0.39) [3]	8443.0	690.5	8425.2	13855.3	7150.1	13667.7	7682.2	0.01
	4	5	7336.5	48.7	7079.4	838.5	7078.7	117.4	6493.8	96.3	6628.8	0.01
		6	6805.3	2963.6 (0.03) [2]	6442.8	80.2	6442.8	1304.9	5116.5	1268.7	5773.5	0.01
		7	9054.9	3600 (0.43) [3]	8679.5	1527.1	8677.1	46692.4	7603.6	46379.7	7758.1	0.01

time limit, we report the terminating gap (i.e., TGap(%)). The numbers in brackets indicate the number of instances (out of three) not solved to default optimality within the time limit. When all three instances cannot be solved to default optimality within the time limit, we report the CPU time as the time limit (i.e., 3600 seconds). The table shows that the numerical results verify the theoretical results described in Proposition 2. For example, our proposed model TMS provides higher profits than the other models, as the multistage stochastic real-time offer submission process takes advantage of the continuously realized scenario information as time evolves. Although the optimal solution of MSV is only a feasible solution of TMSR, it obtains good unit commitment decisions from model TS and accordingly produces profits similar to $\mathcal{Z}^{\mathrm{TMSR}}$, but our approach is far better than both alternatives. In addition, $\mathcal{Z}^{\mathrm{MSV}}$ is much larger than $\mathcal{Z}^{\mathrm{DA}}$ and $\mathcal{Z}^{\mathrm{MEV}}$ is far smaller than the profits of the other models, which implies that intuitive stochastic programming approaches may perform poorly. Furthermore, solving most of the instances for model TMS within the time limit is difficult.

5.3. Branch-and-Cut Algorithm

In this section, we show the effectiveness of the convex hull results and strong valid inequalities in solving our proposed model TMS. We add strong valid inequalities as cutting planes to speed up the branch-and-cut algorithm to solve TMS. In our branch-and-cut framework, the strong valid inequalities for the convex hull results in Section 4.1 (i.e., inequalities (5)-(9), (12b)-(12n), and (EC.16a)-(EC.16l)) are added to the model as constraints. The strong valid inequalities for the multi-period cases in Section 4.2 are added as cutting planes (i.e., the user cuts through the callback function of CPLEX). Considering the trade-off between the reduction in the number of branch-and-bound nodes and the size increment of the problem corresponding to each node, as shown in Achterberg (2007), inequalities (13)-(15) are added as user cuts in all of the possible branch-and-bound nodes and inequalities (16) in the first 50 nodes.

Table 2   **Effectiveness of Proposed Inequalities in Tightening the LP Relaxation**

K	T	$L = \ell = 2$			$L = \ell = 3$			$L = \ell = 4$		
		LP Gap (%)	Cut (%)	Percent-age (%)	LP Gap (%)	Cut (%)	Percent-age (%)	LP Gap (%)	Cut (%)	Percent-age (%)
	10	2.77	0.21	92.56	3.30	0.05	98.55	4.63	0.19	95.84
2	11	2.29	0.19	91.93	2.78	0.04	98.39	2.98	0.16	94.64
	12	2.68	0.18	93.19	3.06	0.04	98.77	3.60	0.10	97.19
	6	2.61	0.17	93.65	2.32	0.02	99.29	2.56	0.07	97.20
3	7	2.73	0.18	93.52	3.14	0.05	98.43	3.21	0.14	95.74
	8	4.40	0.18	95.91	3.79	0.04	99.04	4.15	0.07	98.31
	5	3.38	0.10	96.97	1.55	0.06	96.01	1.83	0.11	93.81
4	6	2.92	0.17	94.10	3.83	0.06	98.51	4.63	0.05	98.83
	7	2.69	0.11	95.82	3.72	0.07	98.22	3.54	0.12	96.58

We tested the same instances used by TMS in Section 5.2. The results are reported in Tables 2 and 3. In Table 2, the effectiveness of our strong valid inequalities in tightening the linear programming (LP) relaxation of model TMS (a maximization problem) is reported for different instances. The column "LP Gap (%)" reports the relative LP relaxation gap of the original formulation with respect to a feasible solution leading to the largest objective value obtained within the time limit (referred to as the best feasible solution) from default CPLEX and our branch-and-cut framework. The value "LP Gap (%)" is defined as $(Z_{\mathrm{LP}} - Z_{\mathrm{MILP}})/Z_{\mathrm{LP}}$, where Z_{LP} is the optimal objective value of the LP relaxation of model TMS without adding any strong valid inequalities, and Z_{MILP} is the objective value corresponding to the best feasible solution. The column "Cut(%)" reports the LP relaxation gap after adding our strong valid inequalities. The value "Cut(%)" is defined as $(Z_{\mathrm{LP}}^{\mathrm{Cut}} - Z_{\mathrm{MILP}})/Z_{\mathrm{LP}}^{\mathrm{Cut}}$, where $Z_{\mathrm{LP}}^{\mathrm{Cut}}$ represents the objective value of the LP relaxation of model TMS with our derived inequalities added as constraints. The table shows that the LP relaxation gap decreases dramatically after adding the proposed strong valid inequalities to tighten model TMS.

The degree of the reduction is shown in the column "Percentage (%)," where Percentage (%) = (LP Gap (%) - Cut (%))/LP Gap (%). A reduction of more than 95% can be achieved, indicating that the proposed strong valid inequalities are extremely effective in tightening TMS.

Table 3   Results for the Branch-and-Cut Framework

$L(\ell)$	K	T	Default CPLEX			Branch-and-Cut			
			CPU Time	TGap (%)	# Nodes	CPU Time	TGap (%)	# Nodes	# Cuts
2	2	10	1256.2	0.05 [1]	80486	82.3	0.00	2988	2848
		11	3600.0	0.08 [3]	88505	441.3	0.00	7074	3569
		12	3600.0	0.10 [3]	35537	1688.5	0.00	12389	7541
	3	6	131.2	0.00	3803	33.4	0.00	821	2695
		7	2513.4	0.04 [2]	141526	127.5	0.00	5386	3801
		8	3600.0	0.73 [3]	38272	1455.7	0.00	13600	7857
	4	5	146.2	0.00	1979	87.1	0.00	842	3007
		6	3600.0	0.06 [3]	145486	201.4	0.00	4505	6076
		7	3600.0	0.76 [3]	22488	3598.3	0.05 [2]	18246	17041
3	2	10	434.7	0.00	18321	66.9	0.00	2620	5471
		11	3600.0	0.06 [3]	86388	321.7	0.00	7255	4858
		12	3600.0	0.11 [3]	36174	3468.2	0.02 [1]	28198	8371
	3	6	942.7	0.00	39143	189.9	0.00	7622	4203
		7	1602.0	0.03 [1]	75680	117.1	0.00	2584	5206
		8	3600.0	0.20 [3]	39205	1880.5	0.00	22636	9382
	4	5	41.9	0.00	1299	27.4	0.00	925	5310
		6	3600.0	0.08 [3]	112913	177.5	0.00	5378	5991
		7	3600.0	0.63 [3]	18717	3036.3	0.01 [1]	16914	16698
4	2	10	3600.0	0.05 [3]	150761	69.0	0.00	2572	4658
		11	3600.0	0.10 [3]	77072	548.9	0.00	12924	4289
		12	3600.0	0.21 [3]	32555	1662.1	0.00	15646	8562
	3	6	99.1	0.00	3471	61.7	0.00	3433	3233
		7	3600.0	0.06 [3]	168636	101.7	0.00	4266	4233
		8	3600.0	0.39 [3]	35915	1669.4	0.00	19675	11189
	4	5	48.7	0.00	1556	46.8	0.00	899	4996
		6	2963.6	0.03 [2]	91189	163.7	0.00	4545	7737
		7	3600.0	0.43 [3]	19098	3279.7	0.00	15464	16991

In Table 3, we report the performance of our branch-and-cut framework ("Branch-and-Cut") in solving model TMS and compare it with the default CPLEX ("Default CPLEX"). The derived inequalities are added to model TMS only because they are developed based on the specific structure of the model. The column "CPU Time" provides the average computational time in seconds needed to solve the corresponding instances, with 3600 given for cases in which all three instances cannot be solved to default optimality. The column "TGap (%)" provides the final optimality gap obtained within the time limit. The numbers in brackets indicate the number of instances (out of three) that are not solved to default optimality (i.e., 0.01%) within the time limit. The column "# Nodes" provides the number of nodes explored in the branch-and-bound searching tree. The last

column "# Cuts" represents the number of multi-period strong valid inequalities utilized in the branch-and-cut framework to solve the instances.

Table 3 shows that our branch-and-cut framework has much better performance than the default CPLEX does, and can solve most instances in a short time, while the default CPLEX cannot solve most of them to optimality within the time limit. Our framework explores fewer branch-and-bound nodes than the default CPLEX does. For the case in which both approaches cannot solve the problem to optimality within the time limit, our framework provides a much smaller terminating gap. In summary, our proposed model TMS provides both the highest profit (as indicated in Table 1) and can be solved in a shorter time than most of the alternative models (see Tables 1 and 3).

Table 4   Results for Long Horizon Instances

L (ℓ)	# Sce.	Model Comparisons					Branch-and-Cut							
							IGap (%)		Pct. (%)	CPU Time (TGap (%))		# Nodes		# Cuts
		TMS	TMSR	MSV	MEV	DA	TMS	Cut		TMS	Cut	TMS	Cut	
4	250	30883.5	29650.2	29649.7	23806.7	24198.2	2.56	0.14	94.53	93.3	51.0	2137	928	4203
	500	30966.8	29922.4	29922.3	22625.9	23511.3	2.78	0.22	92.09	*** (0.05) [3]	4158.6 (0.03) [1]	38605	46282	8315
	1000	31915.5	30775.1	30770.5	24296.5	24816.6	2.21	0.11	95.02	*** (0.02) [3]	3416.5 (0.02) [1]	37945	19036	10638
	2000	30103.8	29031.6	29031.5	22872.5	23308.7	2.58	0.21	91.86	*** (0.11) [3]	*** (0.09) [3]	9737	11533	19376
6	250	29695.7	28698.4	28696.8	22571	23415.9	3.33	0.31	90.69	2470.5	1373.9	17899	24212	4278
	500	31146.6	30237.5	30237.1	23328.4	23876.7	2.69	0.22	91.82	*** (0.04) [3]	4808.1 (0.02) [1]	31267	35517	8311
	1000	32053.3	31117.5	31097.1	25133	25541	2.73	0.18	93.41	*** (0.04) [3]	6537.2 (0.03) [2]	17214	19901	10083
	2000	29514.5	28436	28436	21549.5	22645.5	3.04	0.23	92.43	*** (7.31) [3]	*** (0.06) [3]	2047	1423	19753
8	250	30039	29173.5	29160.7	23167.3	23871.2	3.09	0.32	89.64	221.0	107.3	477	155	3668
	500	31135.9	30351.4	30341.7	23780.2	24364.7	3.02	0.33	89.07	*** (0.02) [3]	5416.8	16621	22497	6680
	1000	32299.9	31625.7	31625.6	24735	24897.4	2.75	0.27	90.18	*** (0.06) [3]	*** (0.04) [3]	8342	13286	10194
	2000	30319.6	29446.5	29446.4	22644	23535.8	2.95	0.26	91.19	6618.1 (18.24) [2]	5782.1 (0.15) [2]	0	49	20665
10	250	30791.9	30141.8	30141.4	23941.7	24526.7	2.96	0.33	88.85	4610.0 [1]	2944.0 [1]	17520	22008	3001
	500	30594.1	29922.4	29922.3	22625.9	23511.3	3.26	0.32	90.18	*** (0.04) [3]	*** (0.03) [3]	8599	11919	5910
	1000	30509.9	29708.6	29705.3	23470.3	24136	2.98	0.28	90.6	*** (0.04) [3]	4816.1 (0.09) [1]	9220	7181	9656
	2000	30352.3	29601.9	29601.5	22826.1	23702.1	3.01	0.34	88.7	*** (32.80) [3]	7088.0 (0.09) [2]	0	71	21738

To further demonstrate the benefits of our proposed model TMS and the effectiveness of our proposed strong valid inequalities in a setting with a long horizon, we conduct experiments over various 24-time-period instances based on the generator data described in Section 5.1. We generate four scenario trees with 24 periods and each tree has a different number of scenarios, with the number of branches at each node between 1 and 4 (Heitsch and Römisch 2009). For each scenario tree, we consider the minimum-up-/-down time limit (i.e., $L(\ell)$) varies in set $\{4, 6, 8, 10\}$. We set the time limit per run as two hours as we are solving larger instances. Similar to the above experiments, given $L(\ell)$ and a scenario tree, we test three randomly generated instances and report the average result in Table 4. The column "# Sce." provides the number of scenarios in each scenario tree. The column "Model Comparisons" compares our model TMS with the other four alternative models

(i.e., TMSR, MSV, MEV, and DA). We can observe similar results as those in Table 1. The column "Branch-and-Cut" reports the performance of our strong valid inequalities in speeding up the solution process of model TMS. We use columns "TMS" and "Cut" to represent the original model solved by default CPLEX and that by adding our inequalities as user cuts, respectively. The column "IGap (%)" reports the relative LP relaxation gap and the column "Pct. (%)" provides the gap reduction in percentage after adding our inequalities to tighten model TMS. We can observe that a reduction of around 90% can be obtained. The column "CPU Time (TGap (%))" reports the computational time in seconds to solve the model. We use "***" to indicate that none of three randomly generated instances are solved into optimality within the time limit, and accordingly, we report the terminating gap labeled "TGap (%)", which indicates the relative gap between the objective value corresponding to the best integer solution and the best upper bound when the time limit is reached. It is clear that our proposed inequalities significantly reduce computational time and terminating gap.

5.4. Arbitrage vs No-arbitrage Cases

Considering the fact that optimization under uncertainty models that deal with price uncertainty tend to be extremely sensitive to the uncertainty model. In this section, we conduct case studies in which each scenario tree is constructed based on explicit no-arbitrage assumptions. The second-stage real-time tree is a martingale. The expected values of uncertain parameters for the nodes at a time period in the first-stage day-ahead tree are equal to the expected values of uncertain parameters for the corresponding nodes in the real-time tree at the same time period. Such a scenario tree does not create arbitrage opportunities for the IPPs. To that end, we perform additional experiments following the same problem settings in Section 5.1 except that we randomly generate the uncertain electricity price and wind generation satisfying the no-arbitrage conditions (denoted as "No-arbitrage Case"). In addition, to perform a comparison, based on the scenario tree under no-arbitrage conditions, we perturb the day-ahead price at each time period by 10% and keep the real-time prices the same, leading to a scenario tree with arbitrage opportunities (denoted as "Arbitrage Case"). For both cases, we compare the five models presented in Sections 2 and 3 and report the results in Table 5. We observe similar results as in Table 1. That is, our proposed model TMS provides higher profits than the other alternative models under both cases. Note here that as the "Arbitrage Case" provides arbitrage opportunities between the day-ahead and real-time markets, the benefits of our proposed model TMS can be slightly exaggerated. However, since we primarily focus on the self-scheduling for an IPP's assets with limited capacities, although the "Arbitrage Case" does provide arbitrage opportunities, the capacity limits of generation resources prevent the IPP from selling and buying electricity in extremely large or small values.

Table 5   Model Comparisons under No-arbitrage vs. Arbitrage Cases

$L(\ell)$	K	T	No-arbitrage Case					Arbitrage Case				
			$\mathcal{Z}^{\text{TMS}}$	$\mathcal{Z}^{\text{TMSR}}$	$\mathcal{Z}^{\text{MSV}}$	$\mathcal{Z}^{\text{MEV}}$	$\mathcal{Z}^{\text{DA}}$	$\mathcal{Z}^{\text{TMS}}$	$\mathcal{Z}^{\text{TMSR}}$	$\mathcal{Z}^{\text{MSV}}$	$\mathcal{Z}^{\text{MEV}}$	$\mathcal{Z}^{\text{DA}}$
2	2	10	12580.0	9468.9	9468.9	4976.2	6600.0	12602.9	9800.9	9517.2	5025.0	6632.7
		11	14615.4	11042.5	10799.1	5500.1	7235.1	14722.2	11325.3	10825.9	5527.2	7210.1
		12	15660.8	11957.3	11957.3	5971.4	7920.0	15714.2	12427.5	12070.2	6084.7	7953.7
	3	6	7880.4	6207.8	6207.8	2985.7	3960.0	7895.9	6418.4	6313.6	3093.3	3907.1
		7	7360.3	5638.2	5638.2	3483.3	4620.0	7452.0	5915.5	5756.8	3605.1	4547.2
		8	10034.7	7610.3	7610.3	3980.9	5280.0	10051.5	7798.5	7700.5	4072.8	5071.4
	4	5	4776.3	3702.1	3702.1	2488.1	3300.0	4800.4	3787.8	3721.3	2508.2	3135.7
		6	7728.5	6000.1	6000.1	2985.7	3960.0	7873.4	6334.3	6109.9	3098.6	4018.1
		7	7871.4	6134.8	6134.8	3483.3	4620.0	7947.1	6494.1	6291.0	3642.7	4631.6
3	2	10	12918.8	9908.3	9908.3	4976.2	6600.0	13012.3	10246.1	9952.1	5020.2	6411.3
		11	19497.9	15728.9	15476.1	5499.9	7229.3	19593.6	16017.4	15546.8	5571.0	7198.6
		12	23420.6	17986.3	17986.3	5971.4	7920.0	23510.1	18547.9	18121.9	6107.1	8003.9
	3	6	7457.5	5826.6	5826.6	2985.7	3960.0	7475.4	5983.3	5916.2	3076.3	3931.2
		7	8704.1	6777.6	6777.6	3483.3	4620.0	8766.8	6917.0	6795.4	3501.1	4385.8
		8	12616.7	10012.2	10012.2	3980.9	5280.0	12654.1	10205.3	10030.4	3999.3	5053.7
	4	5	4590.7	3411.7	3411.7	2488.1	3300.0	4658.3	3514.3	3417.9	2494.3	3247.8
		6	6014.0	4597.5	4597.5	2985.7	3960.0	6165.1	4959.3	4798.7	3190.3	4090.4
		7	9139.5	7209.9	7209.9	3483.3	4620.0	9322.0	7706.8	7332.6	3608.3	4737.0
4	2	10	13941.5	10814.4	10814.4	4976.2	6600.0	14080.3	11176.9	10876.4	5038.6	6519.9
		11	15786.6	12812.4	12512.1	5505.3	7367.0	15878.4	13159.9	12567.2	5560.7	7474.0
		12	19474.2	15013.7	15013.7	5971.4	7920.0	19500.2	15412.0	15083.4	6041.2	7771.7
	3	6	7555.7	5868.4	5868.4	2985.7	3960.0	7654.5	6206.7	6001.6	3120.3	4172.0
		7	8399.0	6595.9	6595.9	3483.3	4620.0	8490.4	6983.1	6726.2	3616.8	4673.5
		8	12653.6	10081.8	10081.8	3980.9	5280.0	12868.9	10465.5	10116.7	4016.1	5139.4
	4	5	5038.9	3811.7	3811.7	2488.1	3300.0	5157.1	4111.2	4049.9	2742.1	3377.1
		6	7406.1	5952.8	5952.8	2985.7	3960.0	7496.5	6257.6	6101.5	3138.4	4001.0
		7	9442.2	7279.5	7279.5	3483.3	4620.0	9510.6	7710.3	7409.8	3616.3	4797.4

5.5.   The Risk-Averse Model

We report the computational results of the risk-averse model of TMS based on CVaR, as described in Online Supplement B.2. We follow the settings in Section 5.1 to generate instances corresponding to a combination with $(L(\ell), K, T) = (4, 2, 9)$ and conduct the experiments by varying the values of $\lambda \in [0, 1]$ for $\alpha = 0.9$ and $\alpha = 0.95$, and report the results in Figures 8 and 9, respectively. Note
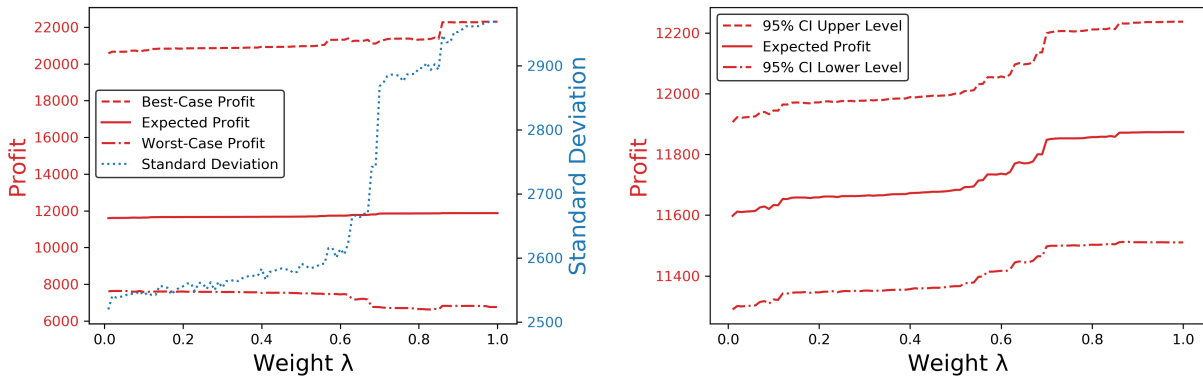
(a) Profit Information   (b) 95% Confidence Interval

Figure 8   Risk-Averse Model with $\alpha = 0.9$

that when $\lambda = 1$, the risk-averse model (EC.1) is equivalent to the risk-neutral model TMS. For model (EC.1), we solve it to obtain the risk-averse solution and evaluate the solution using the objective function of our proposed risk-neutral model TMS, leading to the expected profit. We also check the profits corresponding to the worst and best scenarios in the scenario tree $\mathcal{T}$ evolved in the real-time market, leading to the worst-case and best-case profits, respectively, and calculate the standard deviation of the profits corresponding to all the scenarios. The results are reported in Figure 8a, where the left and right vertical axes specify the profit and standard deviation, respectively. The dashed, solid, dash-dot, and dotted lines represent the best-case profit, expected profit, worst-case profit, and standard deviation, respectively. In addition, based on the results on the expected profit and standard deviation, we construct a 95% confidence interval (CI) for each given λ and report the results in Figure 8b, where the dashed, solid, and dash-dot lines represent the CI upper level, expected profit, and CI lower level, respectively. From Figure 8, we can observe that when λ increases, i.e., the model becomes more risk-neutral, the expected profit increases, while the profit variance increases with respect to possible realizations of the uncertainty, and the confidence interval length slightly increases as well. The results are due to the well-known fact that a risk-averse model protects the decision-maker with low risk but may lead to a meager profit. Similar results can be obtained when we set $\alpha = 0.95$, as shown in Figure 9. We also note that the magnitudes of both the profit and variance differences are not very significant, and in the following, we accordingly continue to conduct the experiments using our risk-neutral model TMS.
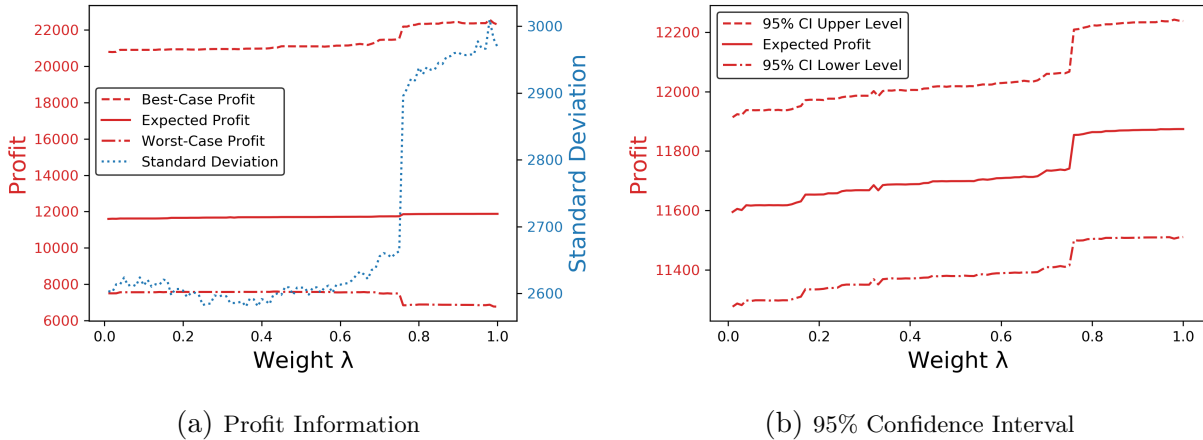

(a) Profit Information   (b) 95% Confidence Interval

Figure 9   **Risk-Averse Model with** $\alpha = 0.95$

5.6. Empirical Validation

In this section, we use out-of-sample tests to implement our proposed approach using real data, to further show its potential benefits to real-world practitioners.

5.6.1.   Data Instances We select four thermal generators that are currently used in practice for numerical testing to show that the results are robust for generators with different physical characteristics. They are from regions covered by a U.S. wholesale market (i.e., PJM). The detailed physical data, with minimal perturbation from the real data of these four generators, are shown in Table 6, where parameters a, b, and c are the coefficients of function $f(x_i) = ax_i^2 + bx_i + cy_i$.

Table 6   Generator Data

Generator	$\underline{C}$ (MW)	$\overline{C}$ (MW)	$L(\ell)$	$\overline{V}$ (MW/h)	V (MW/h)	U ($)	a ($/MW²h)	b ($/MWh)	c ($/h)
1	17.88	44.7	5	23.38	11	45	0.00977	22.9423	58.81
2	40.8	170	3	61.8	42	400	0.01088	12.8875	6.78
3	44.2	260	6	76.7	65	100	0.0024	12.3299	28
4	22	87	2	32.5	21	300	0.0068	5.8645	230

For the uncertain electricity price, we collect the historical day-ahead and real-time electricity prices published by PJM at its Zone 1 from July 14, 2013 to July 10, 2017. For the uncertain renewable generation, we collect the historical wind outputs generated from one region (i.e., the Mid-Atlantic) covered by PJM within the same period. The electricity price and wind output data are divided into two groups with respect to the time periods they cover for the out-of-sample test purpose: group one covers July 14, 2013 to November 9, 2016 (1215 days in total) and group two covers November 10, 2016 to July 10, 2017 (243 days in total).

5.6.2.   Out-of-Sample Tests The out-of-sample tests are performed in two steps: First, we use the group one data to generate scenarios to construct the optimization models described in Sections 2 and 3 and thus obtain the corresponding day-ahead solutions. Second, for each model, we validate the effectiveness of the obtained day-ahead solutions by using the group two data.

We perform the tests for the TMS model and two common models in practice: the MEV and DA models. In the first step, the historical day-ahead and real-time prices in group one are used to generate the day-ahead and real-time price scenarios, respectively. The historical wind output data in the same group are utilized to generate the renewable generation output scenarios. Following this approach, since we have one observed sample that combines day-ahead prices, real-time prices, and wind output amounts for each particular operating day, there are 1215 samples. Following the procedure described in Heitsch and Römisch (2009), we use these 1215 samples to generate a scenario tree $\mathcal{T}$ with 240 scenarios to build the TMS model. For the MEV model, we aggregate these 1215 samples into 240 independent scenarios for its first phase and one mean-value scenario for its second phase. The historical wind outputs represent the generation amounts in a large region, so we scale them down to the level at which the maximum wind generation amount is around one

half of the generation upper bound (i.e., $\overline{C}$) of the thermal generator (detailed percentages shown in the column "Wind Scale" in Table 7).

After obtaining the day-ahead offer amounts for the TMS, MEV, and DA models through solving the optimization problems above, in the second step, we take the group two samples as simulation scenarios (one day as a scenario and 243 days in total) to perform the validation (Fleten et al. 2002). We define set $\mathcal{U}$ as the collection of these 243 scenarios with each scenario representing the day-ahead price, real-time price, and wind output for each specific day. The average value of the day-ahead prices in the simulation scenarios in $\mathcal{U}$ is used to calculate the day-ahead market profit for each model. For the real-time market profit, the calculations are as follows: For the DA model, there is no real-time market profit. For the MEV model, we take the average real-time price and wind power output amount of the simulation scenarios in $\mathcal{U}$ as the input in the second-stage mean-value scenario model to obtain the real-time market offer amounts and the corresponding real-time market profit. For the TMS model, we validate the obtained day-ahead solutions with respect to each simulation scenario in $\mathcal{U}$ through a rolling horizon approach. Corresponding to each simulation scenario, we solve an optimization model at each period in the real-time market operation. Specifically, given a simulation scenario $s \in \mathcal{U}$ and at each period t, we perform the following procedure:

1. Generate a scenario tree that covers the following six periods from periods t to $t + 5$. Based on the observed realization of this scenario s from periods 1 to t, we select 240 scenarios from the 1215 historical samples used in the first step, which have similar values from periods 1 to t as those in this simulation scenario s, to construct the scenario tree. Based on this scenario tree, we build a multistage stochastic programming model (covering periods t to $t+5$) corresponding to this simulation scenario s and this period t.

2. Solve the corresponding multistage stochastic programming problem and obtain the root node decision as the offer amount corresponding to this scenario s and this period t with the corresponding profit for this scenario and this period denoted as z_{st}.

Accordingly, the total profit for scenario s is equal to $\sum_{t=1}^{T} z_{st}$. The average of profits among all scenarios in $\mathcal{U}$, i.e., $\sum_{s \in \mathcal{U}} \sum_{t=1}^{T} z_{st}/|\mathcal{U}|$, represents the real-time market profit from the TMS model. Note that in Step 2, each problem is solved within five seconds with most instances within one second, which is fast enough for the real-time operations in practice.

The results are reported in Table 7. The columns "TMS", "MEV", and "DA" provide the total profits obtained from the TMS, MEV, and DA models under different cases. The fifth column labelled "PR_MEV" indicates the percentage of the profit reduction from TMS to MEV, i.e., $\text{PR\_MEV} = (\mathcal{Z}^{\text{TMS}} - \mathcal{Z}^{\text{MEV}})/\mathcal{Z}^{\text{TMS}}$. The last column "PR_DA" indicates the percentage of profit reduction from TMS to DA. The table shows that our approach achieves a significantly higher profit than the MEV and DA models do for all combinations in the case study.

Table 7   Out-of-Sample Test Results

Generator	Wind Scale	TMS	MEV	PR_MEV	DA	PR_DA
1	45%	1688.7	1372.3	18.74%	1480.3	12.34%
	55%	3152.2	2717.2	13.80%	2960.5	6.08%
2	45%	6647.0	5136.3	22.73%	5523.2	16.91%
	55%	7194.0	5630.9	21.73%	6074.8	15.56%
3	45%	6989.5	5762.0	17.56%	5921.1	15.29%
	55%	8306.0	7073.8	14.83%	7401.3	10.89%
4	45%	1642.5	1488.6	9.37%	1586.0	3.44%
	55%	1980.5	1757.7	11.25%	1903.2	3.90%

6.   Conclusion

In this paper, we proposed an integrated optimal self-scheduling strategy for an IPP that participates in both day-ahead and real-time electricity markets. It leads to a practical innovation that allows IPPs to better manage their generation assets and coordinate the offers between the two markets. To effectively model this practical innovation, we developed a new stochastic programming framework that combines two-stage and multistage stochastic programs. The IPP submits the day-ahead offer amounts in the first stage considering the possible real-time offer amounts under uncertain real-time price and renewable generation output as a recourse in the second stage, which is a scenario-tree-based multistage stochastic program in itself. We showed that our approach outperforms alternative models through both theoretical proofs and numerical experiments. To efficiently solve the resulting formulation of our proposed model, we performed an extensive study to derive strong valid inequalities, including convex hull representations for certain special cases, and corresponding polynomial-time separation algorithms to speed up the branch-and-cut process, to solve the formulation. Finally, real case studies confirmed the advantage of our proposed approach and the strength of our proposed strong valid inequalities. In future research, we will explore the integration of our approach with decomposition algorithms, e.g., progressive hedging (Rockafellar and Wets 1991, Watson and Woodruff 2011) and Lagrangian relaxation (Fisher 1981), to solve large-scale problems by decomposing the scenario tree into subtrees and making use of our derived cutting planes in this paper to help solve the subproblem corresponding to each subtree or scenario.

Acknowledgments

Endnotes

1. All the instance data are publicly available at `https://github.com/letsplayagamee/Integrated-Self-Scheduling/`.

References

Achterberg, T. (2007). *Constraint Integer Programming*. PhD thesis, Technische Universität Berlin.

Agora Energiewende and Sandbag (2019). The European power sector in 2018. *State of Affairs and Review of Current Developments*, `https://static.agora-energiewende.de/fileadmin/Projekte/2018/EU-Jahresauswertung_2019/Agora-Energiewende_European-Power-Sector-2018_WEB.pdf` Accessed on October 25, 2021.

Anderson, E. J. and Philpott, A. B. (2002). Using supply functions for offering generation into an electricity market. *Operations Research*, 50(3):477–489.

Arroyo, J. M. and Conejo, A. J. (2000). Optimal response of a thermal unit to an electricity spot market. *IEEE Transactions on Power Systems*, 15(3):1098–1104.

Birge, J. R. and Louveaux, F. (2011). *Introduction to Stochastic Programming*. Springer Science & Business Media.

California ISO (2018). Fifth replacement electronic tariff. `https://www.caiso.com/Documents/Section30_Bid-Self-ScheduleSubmission_CAISOMarkets_asof_Feb15_2018.pdf`, Accessed on October 25, 2021.

Carøe, C. C. and Schultz, R. (1998). A two-stage stochastic program for unit commitment under uncertainty in a hydro-thermal power system. In *Preprint SC 98-11, Konrad-Zuse-Zentrum Fur Informationstechnik*, pages 1–17.

Carpentier, P., Gohen, G., Culioli, J. C., and Renaud, A. (1996). Stochastic optimization of unit commitment: a new decomposition framework. *IEEE Transactions on Power Systems*, 11(2):1067–1073.

Cerisola, S., Baíllo, Á., Fernández-López, J. M., Ramos, A., and Gollmer, R. (2009). Stochastic power generation unit commitment in electricity markets: A novel formulation and a comparison of solution methods. *Operations Research*, 57(1):32–46.

Conejo, A. J., Arroyo, J. M., Contreras, J., and Villamor, F. A. (2002). Self-scheduling of a hydro producer in a pool-based electricity market. *IEEE Transactions on Power Systems*, 17(4):1265–1272.

Conejo, A. J., Castillo, E., Mínguez, R., and Milano, F. (2005). Locational marginal price sensitivities. *IEEE Transactions on Power Systems*, 20(4):2026–2033.

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

Dunn, B., Kamath, H., and Tarascon, J.-M. (2011). Electrical energy storage for the grid: a battery of choices. *Science*, 334(6058):928–935.

Evomarkets (2015). Compliance markets. `http://www.evomarkets.com`, Accessed on October 25, 2021.

Eyer, J. and Corey, G. (2010). Energy storage for the electricity grid: Benefits and market potential assessment guide. *Sandia National Laboratories*, 20(10):5.

Fisher, M. L. (1981). The Lagrangian relaxation method for solving integer programming problems. *Management Science*, 27(1):1–18.

Fleten, S.-E., Høyland, K., and Wallace, S. W. (2002). The performance of stochastic dynamic and fixed mix portfolio models. *European Journal of Operational Research*, 140(1):37–49.

Fleten, S.-E. and Pettersen, E. (2005). Constructing bidding curves for a price-taking retailer in the norwegian electricity market. *IEEE Transactions on Power Systems*, 20(2):701–708.

Garces, L. P. and Conejo, A. J. (2010). Weekly self-scheduling, forward contracting, and offering strategy for a producer. *IEEE Transactions on Power Systems*, 25(2):657–666.

Garcia-Gonzalez, J., de la Muela, R. M. R., Santos, L. M., and Gonzalez, A. M. (2008). Stochastic joint optimization of wind generation and pumped-storage units in an electricity market. *IEEE Transactions on Power Systems*, 23(2):460–468.

Gentile, C., Morales-España, G., and 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.

Gribik, P. R., Hogan, W. W., and Pope, S. L. (2007). Market-clearing electricity prices and energy uplift. *Cambridge, MA*.

Guan, Y., Ahmed, S., and Nemhauser, G. L. (2009). Cutting planes for multistage stochastic integer programs. *Operations Research*, 57(2):287–298.

Heitsch, H. and Römisch, W. (2009). Scenario tree modeling for multistage stochastic programs. *Mathematical Programming*, 118(2):371–406.

Heredia, F.-J., Rider, M. J., and Corchero, C. (2010). Optimal bidding strategies for thermal and generic programming units in the day-ahead electricity market. *IEEE Transactions on Power Systems*, 25(3):1504–1518.

Hua, B., Schiro, D. A., Zheng, T., Baldick, R., and Litvinov, E. (2019). Pricing in multi-interval real-time markets. *IEEE Transactions on Power systems*, 34(4):2696–2705.

IEC (2011). Electrical energy storage. http://www.iec.ch/whitepaper/pdf/iecWP-energystorage-LR-en.pdf. International Electrotechnical Commission, Accessed on October 25, 2021.

Kwon, R. H. and Frances, D. (2012). Optimization-based bidding in day-ahead electricity auction markets: A review of models for power producers. In *Handbook of Networks in Power Systems I*, pages 41–59. Springer.

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

Li, T., Shahidehpour, M., and Li, Z. (2007). Risk-constrained bidding strategy with stochastic unit commitment. *IEEE Transactions on Power Systems*, 22(1):449–458.

Löhndorf, N., Wozabal, D., and Minner, S. (2010). Optimizing the day-ahead bidding strategy of electricity storage using approximate dynamic programming. *Preprint, University of Vienna*.

Löhndorf, N., Wozabal, D., and Minner, S. (2013). Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming. *Operations Research*, 61(4):810–823.

MISO (2020). Energy and operating reserve markets. `https://www.misoenergy.org/legal/business-practice-manuals/`, Accessed on October 25, 2021.

Morales, J. M., Conejo, A. J., and Pérez-Ruiz, J. (2010). Short-term trading for a wind power producer. *IEEE Transactions on Power Systems*, 25(1):554–564.

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

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

Ott, A. L. (2003). Experience with PJM market operation, system design, and implementation. *IEEE Transactions on Power Systems*, 18(2):528–534.

Pan, K. and Guan, Y. (2016a). A polyhedral study of the integrated minimum-up/-down time and ramping polytope. [Online] `http://arxiv.org/abs/1604.02184`.

Pan, K. and 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., and Wang, J. (2015). Strengthened MILP formulation for certain gas turbine unit commitment problems. *IEEE Transactions on Power Systems*, 31(2):1440–1448.

Papavasiliou, A., He, Y., and Svoboda, A. (2015). Self-commitment of combined cycle units under electricity price uncertainty. *IEEE Transactions on Power Systems*, 30(4):1690–1701.

Papavasiliou, A. and Oren, S. S. (2013). Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network. *Operations Research*, 61(3):578–592.

Philpott, A., Ferris, M., and Wets, R. (2016). Equilibrium, uncertainty and risk in hydro-thermal electricity systems. *Mathematical Programming*, 157(2):483–513.

PJM (2018). Daily LMP. `http://www.pjm.com/markets-and-operations/energy.aspx` Accessed on October 25, 2021.

Plazas, M. A., Conejo, A. J., and Prieto, F. J. (2005). Multimarket optimal bidding for a power producer. *IEEE Transactions on Power Systems*, 20(4):2041–2050.

Queyranne, M. and 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. and Takriti, S. (Jun. 2005). Minimum up/down polytopes of the unit commitment problem with start-up costs. *IBM, Research Report RC23628*.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42.

Rockafellar, R. T. and Wets, R. J.-B. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147.

Simoglou, C. K., Biskas, P. N., and Bakirtzis, A. G. (2012). Optimal self-scheduling of thermal units during commissioning. *IEEE Transactions on Power Systems*, 27(1):181–188.

SPP (2019). Self-committing in SPP markets: Overview, impacts, and recommendations. `https://www.spp.org/documents/61118/spp%20mmu%20self-commit%20whitepaper.pdf` Accessed on October 25, 2021.

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

U.S. DOE (2008). 20% wind energy by 2030: Increasing wind energy's contribution to U.S. electricity supply. `http://eere.energy.gov/wind/pdfs/41869.pdf`. U.S. Department of Energy, Accessed on October 25, 2021.

UtilityDive (2020). MISO: Majority of coal is self-committed, 12% was uneconomic over 3-year period. `https://www.utilitydive.com/news/miso-majority-of-coal-is-self-committed-12-was-uneconomic-over-3-year-pe/577508/#:~:text=The%20majority%20of%20coal%2Dfired,analysis%20from%20the%20grid%20operator`. Accessed on October 25, 2021.

Valenzuela, J. and Mazumdar, M. (2003). Commitment of electric power generators under stochastic market prices. *Operations Research*, 51(6):880–893.

Wallace, S. W. and Fleten, S.-E. (2003). Stochastic programming models in energy. *Handbooks in Operations Research and Management Science*, 10:637–677.

Wang, J., Shahidehpour, M., and Li, Z. (2008). Security-constrained unit commitment with volatile wind power generation. *IEEE Transactions on Power Systems*, 23(3):1319–1327.

Watson, J.-P. and Woodruff, D. L. (2011). Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science*, 8(4):355–370.

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

Wu, L., Shahidehpour, M., and Li, Z. (2008). GENCO's risk-constrained hydrothermal scheduling. *IEEE Transactions on Power Systems*, 23(4):1847–1858.

Zheng, Q. P., Wang, J., Pardalos, P. M., and Guan, Y. (2013). A decomposition approach to the two-stage stochastic unit commitment problem. *Annals of Operations Research*, 210(1):387–410.

Online Supplement for "Integrated Stochastic Optimal Self-Scheduling for Two-Settlement Electricity Markets"

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

Yongpei Guan

Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA, guan@ise.ufl.edu

Appendix A: Supplement to Section 1

A.1.   An Illustrative Example

We create an example showing the incentive for an IPP to conduct self-scheduling due to the non-convexity issue. In particular, we consider two generators (denoted by G1 and G2) initially off in a single-bus system to serve system loads in two operational periods (with each period as one hour), which are 50MW and 120MW in the first and second periods, respectively. The first generator G1 has generation lower and upper bounds at 55MW and 100MW, respectively, start-up/shut-down ramp rate limits at 100MW/h, general ramp-up/-down rate limits at 5MW/h, and minimum-up/-down time limits at 1 hour. It has a start-up cost of \$100 and its unit generation cost is \$1/MWh. The second generator G2 has generation lower and upper bounds at 30MW and 200MW, respectively, start-up/shut-down ramp rate limits at 200MW/h, general ramp-up/-down rate limits at 10MW/h, and minimum-up/-down time limits at 1 hour. It has a start-up cost of \$200 and its unit generation cost is \$10/MWh. Both generators have shut-down costs at \$0 and only linear generation costs. By using our notation in the paper, we summarize the above generator parameters in the following table:

Table EC.1   Generator Parameters

Generator	$\underline{C}$ (MW)	$\overline{C}$ (MW)	$\overline{V}$ (MW/h)	V (MW/h)	$L(\ell)$	U (\$)	b (\$/MWh)
G1	55	100	100	5	1	100	1
G2	30	200	200	10	1	200	10

The system optimization model for market clearance schedules G1 to produce 0MW and 80MW and G2 to produce 50MW and 40MW in the first and second periods, respectively, to satisfy system loads towards the minimum total generation cost at \$1280 ($= 80 \times 1 + 100 + (50 + 40) \times 10 + 200$). The corresponding locational marginal prices (LMPs, i.e., the optimal dual values corresponding to the load balance constraints in solving the economic dispatch problem when the generators' online/offline statuses are fixed) are \$19/MWh and \$1/MWh in the first and second periods, respectively.

On the one hand, if these two generators follow the system dispatch, then the individual profits (i.e., revenue minus generation cost) for G1 and G2 are -100 ($= 80 \times (1-1) - 100$) and -110 ($= 50 \times (19 - 10) + 40 \times (1-10) - 200$), respectively. On the other hand, if they perform self-scheduling based on the LMPs with the objectives of maximizing the profits, then generator G1 produces 100MW in both the first and second periods, leading to the total profit at $1700 ($= 100 \times (19-1) + 100 \times (1-1) - 100$) and generator G2 produces nothing, leading to the total profit at $0. For both generators, we see profit differences (i.e., $1700 - (-\$100) = \$1800$ for G1 and $0 - (-\$110) = \$110$ for G2) due to the non-convexity issue.

Appendix B: Supplement to Section 2

B.1.   Proof of Proposition 1

Proof of Proposition 1   We prove it by contradiction. Suppose $h_i^{c*}h_i^{d*} \neq 0$ for some $i \in \mathcal{V}$. Then $h_i^{c*} > 0$ and $h_i^{d*} > 0$ by nonnegativity restrictions (1j). We define $\mathcal{V}^> = \{i \in \mathcal{V} : h_i^{c*}h_i^{d*} \neq 0\}$ and let $\mathcal{Z}^{\text{TMS}*}$ denote the objective value corresponding to $(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*})$. In the following, we create another feasible solution $(d', y', u', x', r', s', h^{c'}, h^{d'})$ that achieves an objective value $\mathcal{Z}^{\text{TMS}'}$ such that $\mathcal{Z}^{\text{TMS}'} > \mathcal{Z}^{\text{TMS}*}$.

We let $(d', y', u', x', s') = (d^*, y^*, u^*, x^*, s^*)$. For the values of $(r', h^{c'}, h^{d'})$, we first define $\mathcal{V}_+^> = \{i \in \mathcal{V}^> : h_i^{c*} - h_i^{d*}/(\lambda_c \lambda_d) \geq 0\}$ and $\mathcal{V}_-^> = \mathcal{V}^> \setminus \mathcal{V}_+^>$. Then for any $i \in \mathcal{V}_+^>$, we let $h_i^{c'} = h_i^{c*} - h_i^{d*}/(\lambda_c \lambda_d)$, $h_i^{d'} = 0$, and $r_i' = r_i^* + (1/(\lambda_c \lambda_d) - 1) h_i^{d*}$. For any $i \in \mathcal{V}_-^>$, where $h_i^{c*} - h_i^{d*}/(\lambda_c \lambda_d) < 0$, we let $h_i^{c'} = 0$, $h_i^{d'} = h_i^{d*} - \lambda_c \lambda_d h_i^{c*}$, and $r_i' = r_i^* + (1 - \lambda_c \lambda_d) h_i^{c*}$. For any $i \in \mathcal{V} \setminus \mathcal{V}^>$, where $h_i^{c*}h_i^{d*} = 0$, we let $h_i^{c'} = h_i^{c*}$, $h_i^{d'} = h_i^{d*}$, and $r_i' = r_i^*$.

Since $(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*})$ is a feasible solution of model TMS, we have $(d', y', u', x', r', s', h^{c'}, h^{d'})$ satisfies (1b) - (1g) and (1j) because $(d', y', u', x', s') = (d^*, y^*, u^*, x^*, s^*)$. Further, it is easy to check that $(d', y', u', x', r', s', h^{c'}, h^{d'})$ satisfies (1h) - (1i), and thus it is also a feasible solution of model TMS. It follows that

$$\mathcal{Z}^{\text{TMS}'} = \mathcal{Z}^{\text{TMS}*} + \sum_{i \in \mathcal{V}_+^>} p_i \left(Q_i^r \left(1/(\lambda_c \lambda_d) - 1 \right) h_i^{d*} + \left(H_c/(\lambda_c \lambda_d) + H_d \right) h_i^{d*} \right)$$
$$+ \sum_{i \in \mathcal{V}_-^>} p_i \left(Q_i^r (1 - \lambda_c \lambda_d) h_i^{c*} + \left(H_c + \lambda_c \lambda_d H_d \right) h_i^{c*} \right)$$
$$> \mathcal{Z}^{\text{TMS}*},$$

where the inequality holds because $\lambda_c \lambda_d < 1$, $h^{c*} > 0$, $h^{d*} > 0$, $Q_i^r > 0$ ($\forall i \in \mathcal{V}$), and $|\mathcal{V}^>| > 0$. This contradicts the fact that $(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*})$ is an optimal solution. Thus, our original claim $h_i^{c*}h_i^{d*} = 0$, $\forall i \in \mathcal{V}$ holds. $\square$

B.2.   A Risk-Averse Model of TMS

The conditional value-at-risk (CVaR) is a risk measure applied to cope with uncertain loss function. We consider the negative profit in the second stage of TMS as the loss function and propose a risk-averse model of TMS using CVaR. Given the first-stage decision d and realized uncertain parameters (Q^r, W), we let $\Omega^{\text{TMS}}(d, Q^r, W)$ denote the second-stage loss function. Thus, we have

$$\mathbb{E}_{\mathbb{P}} \left[\Omega^{\text{TMS}}(d, Q^r, W) \right] = - \sum_{i \in \mathcal{V}} p_i \left(Q_i^r r_i - \left(U u_i + D(y_{i-} - y_i + u_i) + f(x_i) \right) - H_c h_i^c - H_d h_i^d \right),$$

where $\mathbb{E}_{\mathbb{P}}$ represents the expectation with respect to the joint probability distribution $\mathbb{P}$ of the underlying uncertain electricity price and renewable generation. We consider the α-CVaR of the loss function, with $\alpha \in (0,1)$ being the given risk parameter, and introduce a parameter $\lambda \in [0,1]$ to allow the decision maker to control her risk attitude. We accordingly consider $\lambda\mathbb{E}_{\mathbb{P}}[\Omega^{\mathrm{TMS}}(d,Q^r,W)] + (1-\lambda)\mathrm{CVaR}_{\alpha}(\Omega^{\mathrm{TMS}}(d,Q^r,W))$ as the second-stage objective, where $\lambda = 1$ reflects the risk-neutral attitude and $\lambda = 0$ focuses on risk-averse attitude only. Based on the reformulation described in Rockafellar and Uryasev (2000), we have

$$\mathrm{CVaR}_{\alpha}\left(\Omega^{\mathrm{TMS}}(d,Q^r,W)\right) = \min_{\zeta \geq 0}\left\{\zeta + \frac{1}{1-\alpha}\mathbb{E}_{\mathbb{P}}\left[\left(\Omega^{\mathrm{TMS}}(d,Q^r,W) - \zeta\right)^+\right]\right\}.$$

Therefore, by introducing auxiliary variable ω_k for any $k \in \mathcal{S}_T$, the risk-averse model of TMS based on CVaR can be described as follows:

$$\min \ -\mathbb{E}\left(\sum_{t=1}^{T} Q_t^d(\xi)d_t\right) - \lambda\sum_{i \in \mathcal{V}}p_i\left(Q_i^r r_i - \left(Uu_i + D(y_{i^-} - y_i + u_i) + f(x_i)\right) - H_c h_i^c - H_d h_i^d\right)$$

$$+(1-\lambda)\left(\zeta + \frac{1}{1-\alpha}\sum_{k \in \mathcal{S}_T}p_k\omega_k\right) \tag{EC.1a}$$

$$\text{s.t. } (1a) - (1j),$$

$$\omega_k \geq -\sum_{i \in \mathcal{P}(k)}\left(Q_i^r r_i - \left(Uu_i + D(y_{i^-} - y_i + u_i) + f(x_i)\right) - H_c h_i^c - H_d h_i^d\right) - \zeta, \ \forall k \in \mathcal{S}_T, \tag{EC.1b}$$

$$\omega_k \geq 0, \ \forall k \in \mathcal{S}_T; \ \zeta \geq 0, \tag{EC.1c}$$

where $\mathcal{S}_T$ represents the set of all scenario nodes in stage T.

Appendix C: Supplement to Section 3

C.1. Day-Ahead Self-Scheduling and Real-Time Adaptive Offer Submission Model

The corresponding model can be described as follows.

$$\mathcal{Z}^{\mathrm{TMSR}} = \max\left\{g^{\mathrm{TMSR}}(d,y,u,x,r,s,h^c,h^d) := \mathbb{E}\left(\sum_{t=1}^{T}Q_t^d(\xi)d_t\right) - \sum_{t=2}^{T}\left(Uu_t + D(y_{t-1} - y_t + u_t)\right)\right.$$

$$\left. + \sum_{i \in \mathcal{V}}p_i\left(Q_i^r r_i - f(x_i) - H_c h_i^c - H_d h_i^d\right) : (d,y,u,x,r,s,h^c,h^d) \in \mathcal{X}^{\mathrm{TMSR}}\right\},$$

where $\mathcal{X}^{\mathrm{TMSR}}$ is defined by the following constraints: (1h) - (1i),

$$y_t - y_{t-1} \leq y_k, \ \ \forall t \in [2,T]_{\mathbb{Z}}, \forall k \in [t,\min\{t+L-1,T\}]_{\mathbb{Z}}, \tag{EC.2a}$$

$$y_{t-1} - y_t \leq 1 - y_k, \ \ \forall t \in [2,T]_{\mathbb{Z}}, \forall k \in [t,\min\{t+\ell-1,T\}]_{\mathbb{Z}}, \tag{EC.2b}$$

$$y_t - y_{t-1} \leq u_t \leq \min\{y_t, \ 1 - y_{t-1}\}, \ \ \forall t \in [2,T]_{\mathbb{Z}}, \tag{EC.2c}$$

$$\underline{C}y_{t(i)} \leq x_i \leq \overline{C}y_{t(i)}, \ \ \forall i \in \mathcal{V},$$

$$x_i - x_{i^-} \leq Vy_{t(i^-)} + \overline{V}(1 - y_{t(i^-)}), \ \ \forall i \in \mathcal{V}\setminus\{1\},$$

$$x_{i^-} - x_i \leq Vy_{t(i)} + \overline{V}(1 - y_{t(i)}), \ \ \forall i \in \mathcal{V}\setminus\{1\},$$

$$s_0 = 0; \ 0 \leq s_i \leq \Lambda, r_i \geq 0, h_i^c \geq 0, h_i^d \geq 0, \ \forall i \in \mathcal{V},$$

$$d_t \geq 0, y_t \in \{0,1\}, \ \forall t \in [1,T]_{\mathbb{Z}}; \ u_t \in \{0,1\}, \ \forall t \in [2,T]_{\mathbb{Z}}.$$

C.2. Two-Phase Offer Submission Models

The constraint set $\mathcal{X}^{\mathrm{TS}}$ is defined by the following constraints: (EC.2a) - (EC.2c),

$$\underline{C}y_t \leq x_t(\eta) \leq \overline{C}y_t, \quad \forall t \in [1,T]_{\mathbb{Z}}, \forall \eta \in \mathcal{S},$$

$$x_t(\eta) - x_{t-1}(\eta) \leq Vy_{t-1} + \overline{V}(1 - y_{t-1}), \quad \forall t \in [2,T]_{\mathbb{Z}}, \forall \eta \in \mathcal{S},$$

$$x_{t-1}(\eta) - x_t(\eta) \leq Vy_t + \overline{V}(1 - y_t), \quad \forall t \in [2,T]_{\mathbb{Z}}, \forall \eta \in \mathcal{S},$$

$$s_t(\eta) = s_{t-1}(\eta) + \lambda_c h_t^c(\eta) - h_t^d(\eta)/\lambda_d, \quad \forall t \in [1,T]_{\mathbb{Z}}, \forall \eta \in \mathcal{S},$$

$$d_t + r_t(\eta) = x_t(\eta) + W_t(\eta) - h_t^c(\eta) + h_t^d(\eta), \quad \forall t \in [1,T]_{\mathbb{Z}}, \forall \eta \in \mathcal{S},$$

$$s_0 = 0;\ 0 \leq s_t(\eta) \leq \Lambda, r_t(\eta) \geq 0, h_t^c(\eta) \geq 0, h_t^d(\eta) \geq 0, \ \forall t \in [1,T]_{\mathbb{Z}}, \forall \eta \in \mathcal{S},$$

$$y_t \in \{0,1\}, d_t \geq 0, \ \forall t \in [1,T]_{\mathbb{Z}}; u_t \in \{0,1\}, \ \forall t \in [2,T]_{\mathbb{Z}},$$

where set $\mathcal{S}$ collects all of the scenarios in the scenario tree $\mathcal{T}$.

The constraint set $\mathcal{X}^{\mathrm{MSV}}$ is defined by the following constraints: (1h),

$$\underline{C}y^*_{t(i)} \leq x_i \leq \overline{C}y^*_{t(i)}, \quad \forall i \in \mathcal{V},$$

$$x_i - x_{i^-} \leq Vy^*_{t(i^-)} + \overline{V}(1 - y^*_{t(i^-)}), \quad \forall i \in \mathcal{V} \setminus \{1\},$$

$$x_{i^-} - x_i \leq Vy^*_{t(i)} + \overline{V}(1 - y^*_{t(i)}), \quad \forall i \in \mathcal{V} \setminus \{1\},$$

$$d^*_{t(i)} + r_i = x_i + W_i - h_i^c + h_i^d, \quad \forall i \in \mathcal{V},$$

$$s_0 = 0;\ 0 \leq s_i \leq \Lambda, r_i \geq 0, h_i^c \geq 0, h_i^d \geq 0, \ \forall i \in \mathcal{V}.$$

The constraint set $\mathcal{X}^{\mathrm{MEV}}$ is defined by the following constraints:

$$\underline{C}y^*_t \leq x_t \leq \overline{C}y^*_t, \quad \forall t \in [1,T]_{\mathbb{Z}},$$

$$x_t - x_{t-1} \leq Vy^*_{t-1} + \overline{V}(1 - y^*_{t-1}), \quad \forall t \in [2,T]_{\mathbb{Z}},$$

$$x_{t-1} - x_t \leq Vy^*_t + \overline{V}(1 - y^*_t), \quad \forall t \in [2,T]_{\mathbb{Z}},$$

$$s_t = s_{t-1} + \lambda_c h_t^c - h_t^d/\lambda_d, \quad \forall t \in [1,T]_{\mathbb{Z}}, \tag{EC.3a}$$

$$d^*_t + r_t = x_t + \bar{W}_t - h_t^c + h_t^d, \quad \forall t \in [1,T]_{\mathbb{Z}},$$

$$s_0 = 0;\ 0 \leq s_t \leq \Lambda, h_t^c \geq 0, h_t^d \geq 0, \ \forall t \in [1,T]_{\mathbb{Z}}.$$

C.3. Day-Ahead Only Model

The corresponding model can be described as follows.

$$\mathcal{Z}^{\mathrm{DA}} = \max\left\{ g^{\mathrm{DA}}(d,y,u,x,s,h^c,h^d) := \mathrm{E}\left(\sum_{t=1}^{T} Q_t^d(\xi)d_t\right) - \sum_{t=2}^{T}\left(Uu_t + D(y_{t-1} - y_t + u_t)\right)\right.$$

$$\left. - \sum_{t=1}^{T}\left(f(x_t) + H_c h_t^c + H_d h_t^d\right) : (d,y,u,x,s,h^c,h^d) \in \mathcal{X}^{\mathrm{DA}}\right\},$$

where $\mathcal{X}^{\mathrm{DA}}$ is defined by the following constraints: (EC.2a) - (EC.2c), (EC.3a),

$$\underline{C}y_t \leq x_t \leq \overline{C}y_t, \quad \forall t \in [1,T]_{\mathbb{Z}},$$

$$x_t - x_{t-1} \le V y_{t-1} + \overline{V}(1 - y_{t-1}), \quad \forall t \in [2, T]_{\mathbb{Z}},$$

$$x_{t-1} - x_t \le V y_t + \overline{V}(1 - y_t), \quad \forall t \in [2, T]_{\mathbb{Z}},$$

$$d_t = x_t + \bar{W}_t - h_t^c + h_t^d, \quad \forall t \in [1, T]_{\mathbb{Z}}, \tag{EC.4a}$$

$$s_0 = 0; \; y_t \in \{0, 1\}, 0 \le s_t \le \Lambda, d_t \ge 0, h_t^c \ge 0, h_t^d \ge 0, \; \forall t \in [1, T]_{\mathbb{Z}},$$

$$u_t \in \{0, 1\}, \; \forall t \in [2, T]_{\mathbb{Z}}.$$

C.4. Proof of Proposition 2

Proof of Proposition 2   First, we show $\mathcal{Z}^{\mathrm{TMS}} \ge \mathcal{Z}^{\mathrm{TMSR}} \ge \mathcal{Z}^{\mathrm{MSV}}$ by proving each inequality holds as follows.

(i) $\mathcal{Z}^{\mathrm{TMS}} \ge \mathcal{Z}^{\mathrm{TMSR}}$. For each optimal solution $(d_t^*, y_t^*, u_t^*, \forall t; x_i^*, r_i^*, s_i^*, h_i^{c*}, h_i^{d*}, \forall i)$ of model TMSR, we can always construct a feasible solution $(\bar{d}_t, \forall t; \bar{y}_i, \bar{u}_i, \bar{x}_i, \bar{r}_i, \bar{s}_i, \bar{h}^c, \bar{h}^d, \forall i)$ of model TMS by letting 1) $\bar{d}_t = d_t^*$ for each t; 2) $\bar{y}_i = y_t^*$ and $\bar{u}_i = u_t^*$ when $t(i) = t$, $\bar{x}_i = x_i^*$, $\bar{r}_i = r_i^*$, $\bar{s}_i = s_i^*$, $\bar{h}_i^c = h_i^{c*}$, and $\bar{h}_i^d = h_i^{d*}$ for each i. It is easy to observe that $\mathcal{Z}^{\mathrm{TMS}} \ge g^{\mathrm{TMS}}(\bar{d}, \bar{y}, \bar{u}, \bar{x}, \bar{r}, \bar{s}, \bar{h}^c, \bar{h}^d) = g^{\mathrm{TMSR}}(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*}) = \mathcal{Z}^{\mathrm{TMSR}}$.

(ii) $\mathcal{Z}^{\mathrm{TMSR}} \ge \mathcal{Z}^{\mathrm{MSV}}$. Because each optimal solution $(d_t^*, y_t^*, u_t^*, \forall t; x_i^*, r_i^*, s_i^*, h_i^{c*}, h_i^{d*}, \forall i)$ based on the two-phase model MSV is always a feasible solution of model TMSR, it follows that $\mathcal{Z}^{\mathrm{TMSR}} \ge g^{\mathrm{TMSR}}(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*}) = g^{\mathrm{MSV}}(x^*, r^*, s^*, h^{c*}, h^{d*}) = \mathcal{Z}^{\mathrm{MSV}}$.

(iii) $\mathcal{Z}^{\mathrm{MSV}} \ge \mathcal{Z}^{\mathrm{MEV}}$ if $W_i = W_j$ for each $i, j \in \mathcal{V}$ with $t(i) = t(j)$, i.e., any deterministic sequence of wind power generation. Model MEV can be rewritten as

$$\mathcal{Z}^{\mathrm{MEV}} = \max \Big\{ g^{\mathrm{MSV}}(x, r, s, h^c, h^d) : (x, r, s, h^c, h^d) \in \mathcal{X}^{\mathrm{MSV}},$$

$$x_i = x_j, \; r_i = r_j, \; s_i = s_j, \; h_i^c = h_j^c, \; h_i^d = h_j^d, \; \forall i, j \in \mathcal{V} \text{ such that } t(i) = t(j) \Big\}.$$

That means, model MSV is a relaxation of model MEV under the same objective function of a maximization problem. It follows that $\mathcal{Z}^{\mathrm{MSV}} \ge \mathcal{Z}^{\mathrm{MEV}}$.

Next, we prove the conclusion for the case when $\bar{Q}_t^d = \bar{Q}_t^r$ for any t. We show each inequality holds as follows.

(i) $\mathcal{Z}^{\mathrm{DA}} \ge \mathcal{Z}^{\mathrm{MEV}}$. We split variable d_t in model DA into d_t' and r_t' by letting $d_t = d_t' + r_t'$ with $d_t' \ge 0$ and $r_t' \ge 0$ for each t. It follows that

$$\mathcal{Z}^{\mathrm{DA}} = \max \Big\{ g^{\mathrm{DA}}(d, y, u, x, s, h^c, h^d) := \mathsf{E}\Big(\sum_{t=1}^{T} Q_t^d(\xi) d_t \Big) - \sum_{t=2}^{T} \Big(U u_t + D(y_{t-1} - y_t + u_t) \Big)$$

$$- \sum_{t=1}^{T} \Big(f(x_t) + H_c h_t^c + H_d h_t^d \Big) : (d, y, u, x, s, h^c, h^d) \in \mathcal{X}^{\mathrm{DA}} \Big\}$$

$$= \max \Big\{ g^{\mathrm{DA}}(d', y, u, x, r', s, h^c, h^d) := \mathsf{E}\Big(\sum_{t=1}^{T} Q_t^d(\xi)(d_t' + r_t') \Big) - \sum_{t=2}^{T} \Big(U u_t + D(y_{t-1} - y_t + u_t) \Big)$$

$$- \sum_{t=1}^{T} \Big(f(x_t) + H_c h_t^c + H_d h_t^d \Big) : (d', y, u, x, r', s, h^c, h^d) \in \hat{\mathcal{X}}^{\mathrm{DA}} \Big\}$$

$$= \max \Big\{ g^{\mathrm{DA}}(d', y, u, x, r', s, h^c, h^d) := \mathsf{E}\Big(\sum_{t=1}^{T} Q_t^d(\xi) d_t' \Big) - \sum_{t=2}^{T} \Big(U u_t + D(y_{t-1} - y_t + u_t) \Big)$$

$$+ \sum_{t=1}^{T} \Big(\bar{Q}_t^r r_t' - f(x_t) - H_c h_t^c - H_d h_t^d \Big) : (d', y, u, x, r', s, h^c, h^d) \in \hat{\mathcal{X}}^{\mathrm{DA}} \Big\}, \tag{EC.5}$$

where the first equation is due to the definition of $\mathcal{Z}^{\mathrm{DA}}$ in Section C.3, the second equation is because $d_t = d'_t + r'_t$ for each t ($\hat{\mathcal{X}}^{\mathrm{DA}}$ collects all of the constraints in $\mathcal{X}^{\mathrm{DA}}$ except that d_t in (EC.4a) is replaced with $d'_t + r'_t$), and the third equation holds because $\bar{Q}^r_t = \bar{Q}^d_t$ for each t. Through comparing (EC.5) with model MEV in Section C.2, we can observe that any optimal solution $(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*})$ of the two-phase model MEV is a feasible solution of $\hat{\mathcal{X}}^{\mathrm{DA}}$ in model DA because (y^*, u^*) satisfies constraints (EC.2a) - (EC.2c) and $(d^*, x^*, r^*, s^*, h^{c*}, h^{d*})$ satisfies the remaining constraints in $\hat{\mathcal{X}}^{\mathrm{DA}}$. It follows that

$$
\begin{aligned}
\mathcal{Z}^{\mathrm{DA}} &\geq g^{\mathrm{DA}}(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*}) \\
&= \mathbb{E}\left(\sum_{t=1}^{T} Q^d_t(\xi)d^*_t\right) - \sum_{t=2}^{T}\left(Uu^*_t + D(y^*_{t-1} - y^*_t + u^*_t)\right) + \sum_{t=1}^{T}\left(\bar{Q}^r_t r^*_t - f(x^*_t) - H_c h^{c*}_t - H_d h^{d*}_t\right) \\
&= g^{\mathrm{MEV}}(x^*, r^*, s^*, h^{c*}, h^{d*}) = \mathcal{Z}^{\mathrm{MEV}},
\end{aligned}
$$

where the first inequality holds because $(d^*, y^*, u^*, x^*, r^*, s^*, h^{c*}, h^{d*})$ is a feasible solution of model DA, the first equation holds due to (EC.5), and the second and third equations hold due to the definition of $\mathcal{Z}^{\mathrm{MEV}}$ in (3) in Section 3.2.

(ii) $\mathcal{Z}^{\mathrm{TMSR}} \geq \mathcal{Z}^{\mathrm{DA}}$ if $W_i = W_j$ for each $i, j \in \mathcal{V}$ with $t(i) = t(j)$, i.e., any deterministic sequence of wind power generation. Following (EC.5), we have

$$
\begin{aligned}
\mathcal{Z}^{\mathrm{DA}} = \max \Big\{ &g^{\mathrm{TMSR}}(d, y, u, x, r, s, h^c, h^d) : (d, y, u, x, r, s, h^c, h^d) \in \mathcal{X}^{\mathrm{TMSR}}, \\
&x_i = x_j,\ r_i = r_j,\ s_i = s_j,\ h^c_i = h^c_j,\ h^d_i = h^d_j,\ \forall i, j \in \mathcal{V} \text{ such that } t(i) = t(j) \Big\}.
\end{aligned}
$$

The above formulation shows that any optimal solution for model DA is a feasible solution for model TMSR, and corresponding to each optimal solution for model DA, we have the corresponding objective of TMSR is the same as that of model DA. Thus, model TMSR is a relaxation of model DA and $\mathcal{Z}^{\mathrm{TMSR}} \geq \mathcal{Z}^{\mathrm{DA}}$. $\quad\square$

Appendix D: Supplement to Section 4.1

D.1.   Proof of Proposition 3

Proof of Proposition 3   To prove the validity of inequality (5), which essentially tightens $x_{i-} \leq \overline{C}y_{i-}$ in (4c), we show how (5) is obtained and accordingly illustrate the corresponding insights. Since $\overline{C}y_{i-} = \overline{V}y_{i-} + (\overline{C} - \overline{V})y_{i-}$, the following inequality

$$
x_{i-} \leq \overline{V}y_{i-} + (\overline{C} - \overline{V})(y_i - u_i) \tag{EC.6}
$$

tightens $x_{i-} \leq \overline{C}y_{i-}$ because $0 \leq y_i - u_i \leq y_{i-}$ due to (4a). It is easy to observe that (EC.6) is valid when $y_i - u_i = y_{i-}$, which reduces back to $x_{i-} \leq \overline{C}y_{i-}$. We only need to verify the case in which $y_i - u_i < y_{i-}$, which only happens when $y_i - u_i = 0$ and $y_{i-} = 1$ (because $y_i - u_i \geq 0$). This further implies that $y_i = u_i = 0$ because $y_{i-} = 1$. Thus, for this case, (EC.6) reduces to $x_{i-} \leq \overline{V}$, which is valid due to ramp-down constraints in (4d) as the generator shuts down from node i^- (online) to node i (offline). Since inequality (EC.6) is the same as inequality (5), the validity proof of (5) is done.

For inequality (6), which tightens $x_i \leq \overline{C}y_i$ in (4c), we can similarly show how it is obtained as follows. Since $\overline{C}y_i = \overline{V}y_i + (\overline{C} - \overline{V})y_i$, the following inequality

$$
x_i \leq \overline{V}y_i + (\overline{C} - \overline{V})(y_i - u_i) \tag{EC.7}
$$

tightens $x_i \leq \overline{C} y_i$ because $0 \leq y_i - u_i \leq y_i$ as $0 \leq u_i \leq 1$. It is easy to observe that (EC.7) is valid when $y_i - u_i = y_i$, which reduces back to $x_i \leq \overline{C} y_i$. We only need to verify the case in which $y_i - u_i = 0$ and $y_i = 1$, i.e., $y_i = u_i = 1$, from which (EC.7) reduces to $x_i \leq \overline{V}$, which is valid due to ramp-up constraints in (4d) as the generator starts up at node i (note that $u_i = 1$ implies $y_{i^-} = 0$).

Furthermore, in (EC.7) we can rewrite $\overline{V} y_i + (\overline{C} - \overline{V})(y_i - u_i) = \overline{V} y_i + V(y_i - u_i) + (\overline{C} - \overline{V} - V)(y_i - u_i)$. Thus, $x_i \leq \overline{C} y_i$ can be further tightened to be

$$x_i \leq \overline{V} y_i + V(y_i - u_i) + (\overline{C} - \overline{V} - V)(y_j - u_j) \tag{EC.8}$$

by replacing $(\overline{C} - \overline{V} - V)(y_i - u_i)$ with $(\overline{C} - \overline{V} - V)(y_j - u_j)$. It is easy to observe that (EC.8) is valid when $j = i$, which reduces back to (EC.7). For the case in which $j \neq i$, (EC.8) is clearly valid when $y_{i^-} = 0$ since it leads to $y_i = u_i$ and $y_j = u_j$ due to (4a), indicating ramp-up restrictions. When $y_{i^-} = 1$, (EC.8) is also valid because (EC.8) converts to $x_i \leq (\overline{V} + V) y_i + (\overline{C} - \overline{V} - V) y_j$ (This conclusion is obvious when $y_i = 0$ or $y_i = y_j = 1$. When $y_i = 1$ and $y_j = 0$, we have $x_{i^-} \leq \overline{V}$, which implies $x_i \leq \overline{V} + V$). Since inequality (EC.8) is the same as inequality (6), the validity proof of (6) is done.

To prove the validity of (7), which essentially tightens $x_i - x_{i^-} \leq V y_{i^-} + \overline{V}(1 - y_{i^-})$ in (4d), we show how (7) is obtained and accordingly illustrate the corresponding insights. Due to ramp-up process characteristics and $\overline{V} < \underline{C} + V$, we have

$$x_i - x_{i^-} \leq (\underline{C} + V) y_i - \underline{C} y_{i^-}. \tag{EC.9}$$

Since $(\underline{C} + V) y_i = \overline{V} y_i + (\underline{C} + V - \overline{V}) y_i$ and $y_i \geq y_i - u_i \geq 0$ due to $0 \leq u_i \leq 1$, we can try to tighten (EC.9) to be

$$x_i - x_{i^-} \leq \overline{V} y_i - \underline{C} y_{i^-} + (\underline{C} + V - \overline{V})(y_i - u_i), \tag{EC.10}$$

which is clearly valid when $y_i - u_i = y_i$. When $y_i - u_i < y_i$, i.e., $y_i = u_i = 1$, it follows that $y_{i^-} = 0$ and therefore (EC.10) is valid because it reduces to $x_i \leq \overline{V}$, which is valid due to ramp-up constraints in (4d) as the generator starts up at node i. Since inequality (EC.10) is the same as inequality (7), the validity proof of (7) is done.

Similar argument as above for (6) and (7) can be applied to prove that inequality (8) is valid and thus we omit the corresponding proofs here.

For inequality (9), we can show how it is obtained as follows. Due to ramp-up process characteristics and $\overline{V} < \underline{C} + V$, we have

$$x_i - x_j \leq (\underline{C} + 2V) y_i - \underline{C} y_j, \quad \forall i \neq j. \tag{EC.11}$$

The validity of inequality (EC.11) holds obviously when $y_i = y_j = 0$ or $y_i = 0$, $y_j = 1$. When $y_i = y_j = 1$, due to ramp-up restriction, we have $x_i - x_j \leq 2V$. When $y_i = 1, y_j = 0$, the generator shuts down from node i^- to node j. Then, we have $x_{i^-} \leq \overline{V}$ due to shut-down ramping and accordingly $x_i \leq \overline{V} + V \leq \underline{C} + 2V$. The inequality (EC.11) can further be tightened to be

$$x_i - x_j \leq \overline{V} y_i - \underline{C} y_j + (\underline{C} + 2V - \overline{V})(y_i - u_i), \tag{EC.12}$$

following the similar logic of obtaining (EC.10). Since $(\underline{C} + 2V - \overline{V})(y_i - u_i) = V(y_i - u_i) + (\underline{C} + V - \overline{V})(y_i - u_i)$, we can also tighten (EC.11) to be

$$x_i - x_j \leq \overline{V} y_i - \underline{C} y_j + V(y_i - u_i) + (\underline{C} + V - \overline{V})(y_k - u_k), \tag{EC.13}$$

by replacing $(\underline{C} + V - \overline{V})(y_i - u_i)$ with $(\underline{C} + V - \overline{V})(y_k - u_k)$ following the similar logic of obtaining (EC.8). Therefore, inequality (EC.13) is the same as inequality (9) and it follows that the corresponding validity proof is done. □

D.2. Proof of Theorem 1

Proof of Theorem 1　　Based on Proposition 3, this conclusion holds immediately if every inequality in Q_2 is facet-defining for conv(P_2) and every extreme point of Q_2 is integral in y and u and feasible for P_2, which are proved in the following two propositions respectively.

PROPOSITION EC.1.　*Each inequality in Q_2 is facet-defining for conv(P_2).*

Proof of Theorem EC.1　　See Online Supplement D.3. □

Now we prove that all of the extreme points of Q_2 are integral in y and u. To show this, we first provide the following Lemma.

LEMMA EC.1.　*For the following two-period multistage stochastic self-scheduling problem*

$$z^* = \max \left\{ \sum_{i=0}^{n} a_i x_i + \sum_{i=0}^{n} b_i y_i + \sum_{i=1}^{n} c_i u_i : (x, y, u) \in P_2 \right\}, \tag{EC.14}$$

where $(a, b, c) \in \mathbb{R}^{3n+2}$, there exists at least one optimal solution satisfying one of the following six conditions: (1) $x_0 = 0$, $x_i \in \{0, \underline{C}, \overline{V}\}$, $\forall i = 1, \ldots, n$, (2) $x_0 = \underline{C}$, $x_i \in \{0, \underline{C}, \underline{C} + V\}$, $\forall i = 1, \ldots, n$, (3) $x_0 = \overline{V}$, $x_i \in \{0, \underline{C}, \overline{V} + V\}$, $\forall i = 1, \ldots, n$, (4) $x_0 = \underline{C} + V$, $x_i \in \{\underline{C}, \underline{C} + 2V\}$, $\forall i = 1, \ldots, n$, (5) $x_0 = \overline{C} - V$, $x_i \in \{\overline{C}, \overline{C} - 2V\}$, $\forall i = 1, \ldots, n$, and (6) $x_0 = \overline{C}$, $x_i \in \{\overline{C} - V, \overline{C}\}$, $\forall i = 1, \ldots, n$, where binary variables y and u are uniquely decided following the constraints in P_2 in each condition.

Proof of Lemma EC.1　　See Online Supplement D.4. □

PROPOSITION EC.2.　*All of the extreme points of Q_2 are integral in y and u.*

Proof of Proposition EC.2　　See Online Supplement D.5. □

Finally, by the formulation representations, we have both P_2 and Q_2 bounded. Meanwhile, it is easy to observe that Q_2 is full-dimensional. In addition, because all of the inequalities in Q_2 are valid and facet-defining for conv(P_2) based on Propositions 3 and EC.1, we have $Q_2 \supseteq$ conv(P_2). Moreover, we have that any extreme point in Q_2 is binary in y and u and is feasible for P_2 based on Proposition EC.2. Thus $Q_2 =$ conv(P_2). □

D.3. Proof of Proposition EC.1

Proof of Proposition EC.1　　For each inequality, we generate $3n+2$ affinely independent points in conv(P_2) that satisfy the inequality at equality. We sort the scenario nodes in the second period in the order of $1, 2, \ldots, n$ and label node i^- as index 0, so that we have $n+1$ nodes, i.e., $0, 1, 2, \ldots, n$. Because $0 \in$ conv(P_2), it is sufficient to generate other $3n+1$ linearly independent points in conv(P_2) for each inequality. In the following proofs, we use the superscript of (x, y, u), e.g., r in (x^r, y^r, u^r), to represent the index of different points in conv(P_2) and let ϵ be an arbitrarily small positive real number.

For inequality (4a) $y_i - y_{i^-} - u_i \leq 0, \quad \forall i \in \mathcal{N}$: We create five groups of points as follows:

(i) We create a point $(\acute{x}, \acute{y}, \acute{u}) \in \text{conv}(P_2)$ (totally one point) such that

$$\acute{x}_s = \begin{cases} \underline{C} + \epsilon, & s = 0 \\ \underline{C}, & s = i \\ 0, & s \in [0, n]_{\mathbb{Z}} \setminus \{0, i\} \end{cases} , \quad \acute{y}_s = \begin{cases} 1, & s \in \{0, i\} \\ 0, & s \in [0, n]_{\mathbb{Z}} \setminus \{0, i\} \end{cases} , \quad \text{and} \quad \acute{u}_s = 0, \quad s \in [1, n]_{\mathbb{Z}} .$$

(ii) For each $r \in [0, n]_{\mathbb{Z}} \setminus \{i\}$ (totally n points), we create $(\bar{x}^r, \bar{y}^r, \bar{u}^r) \in \text{conv}(P_2)$ such that

$$\bar{x}_s^r = \begin{cases} \underline{C}, & s \in [0, r]_{\mathbb{Z}} \cup \{i\} \\ 0, & \text{o.w.} \end{cases} , \quad \bar{y}_s^r = \begin{cases} 1, & s \in [0, r]_{\mathbb{Z}} \cup \{i\} \\ 0, & \text{o.w.} \end{cases} , \quad \text{and} \quad \bar{u}_s^r = 0, \quad s \in [1, n]_{\mathbb{Z}} .$$

(iii) For each $r \in [0, n]_{\mathbb{Z}} \setminus \{i\}$ (totally n points), we create $(\hat{x}^r, \hat{y}^r, \hat{u}^r) \in \text{conv}(P_2)$ such that

$$\hat{x}_s^r = \begin{cases} \overline{V}, & s \in [0, r]_{\mathbb{Z}} \cup \{i\} \\ 0, & \text{o.w.} \end{cases} , \quad \hat{y}_s^r = \begin{cases} 1, & s \in [0, r]_{\mathbb{Z}} \cup \{i\} \\ 0, & \text{o.w.} \end{cases} , \quad \text{and} \quad \hat{u}_s^r = 0, \quad s \in [1, n]_{\mathbb{Z}} .$$

(iv) We create a point $(\dot{x}, \dot{y}, \dot{u}) \in \text{conv}(P_2)$ (totally one point) such that

$$\dot{x}_s = \begin{cases} 0, & s = 0 \\ \underline{C}, & \text{o.w.} \end{cases} , \quad \dot{y}_s = \begin{cases} 0, & s = 0 \\ 1, & \text{o.w.} \end{cases} , \quad \text{and} \quad \dot{u}_s = 1, \quad s \in [1, n]_{\mathbb{Z}} .$$

(v) For each $r \in [1, n]_{\mathbb{Z}} \setminus \{i\}$ (totally $n - 1$ points), we create $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r) \in \text{conv}(P_2)$ such that

$$\tilde{x}_s^r = \begin{cases} \underline{C}, & s \in [r, n]_{\mathbb{Z}} \setminus \{i\} \\ 0, & \text{o.w.} \end{cases} , \quad \text{and} \quad \tilde{y}_s^r = \tilde{u}_s^r = \begin{cases} 1, & s \in [r, n]_{\mathbb{Z}} \setminus \{i\} \\ 0, & \text{o.w.} \end{cases} .$$

Finally, these five groups of points are collected in Table EC.2, from which we can observe that $(\hat{x}, \hat{y}, \hat{u})$, $(\dot{x}, \dot{y}, \dot{u})$, and $(\tilde{x}, \tilde{y}, \tilde{u})$ are linearly independent because they can construct a lower-triangular matrix based on the values of y and u after Gaussian elimination on the u part. Moreover, $(\acute{x}, \acute{y}, \acute{u})$ and $(\bar{x}, \bar{y}, \bar{u})$ are further linearly independent with them because all of them can construct a whole lower-triangular matrix after Gaussian elimination between $(\bar{x}, \bar{y}, \bar{u})$ and $(\hat{x}, \hat{y}, \hat{u})$ on the x part because $\overline{V} > \underline{C}$. Thus, we have created $1 + n + n + 1 + (n - 1) = 3n + 1$ linearly independent points in $\text{conv}(P_2)$ as desired.

Similarly, we can show that $u_i - y_i \leq 0$ and $u_i + y_{i^-} \leq 1, \quad \forall i \in \mathcal{N}$, in inequalities (4a) are facet-defining, so is inequality (10). In the following proofs, we follow the similar way described above to create linearly independent points in $\text{conv}(P_2)$ by firstly generating several groups of points that can construct a lower-triangular matrix in terms of the values of x and y (e.g., $(\acute{x}, \acute{y}, \acute{u})$, $(\bar{x}, \bar{y}, \bar{u})$, and $(\hat{x}, \hat{y}, \hat{u})$ for $y_i - y_{i^-} - u_i \leq 0$ above) and then generating several groups of points that can construct an upper-triangular matrix in terms of the value of u (e.g., $(\dot{x}, \dot{y}, \dot{u})$ and $(\tilde{x}, \tilde{y}, \tilde{u})$ for $y_i - y_{i^-} - u_i \leq 0$ above).

For inequality (4b) $x_i \geq \underline{C} y_i, \quad \forall i \in \mathcal{N} \cup \{i^-\}$: For the root node i^- indexed as 0, we create $3n + 1$ independent points in $\text{conv}(P_2)$ in the following three groups:

(i) For each $r \in [0, n]_{\mathbb{Z}}$ (totally $n + 1$ points), we create $(\bar{x}^r, \bar{y}^r, \bar{u}^r) \in \text{conv}(P_2)$ such that

$$\bar{x}_s^r = \begin{cases} \underline{C}, & s \in [0, r]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \quad \bar{y}_s^r = \begin{cases} 1, & s \in [0, r]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \quad \text{and} \quad \bar{u}_s^r = 0, \quad s \in [1, n]_{\mathbb{Z}} .$$

(ii) For each $r \in [1, n]_{\mathbb{Z}}$ (totally n points), we create $(\hat{x}^r, \hat{y}^r, \hat{u}^r) \in \text{conv}(P_2)$ such that

$$\hat{x}_s^r = \begin{cases} \underline{C}, & s \in [r, n]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \quad \text{and} \quad \hat{y}_s^r = \hat{u}_s^r = \begin{cases} 1, & s \in [r, n]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} .$$

(iii) For each $r \in [1, n]_{\mathbb{Z}}$ (totally n points), we create $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r) \in \text{conv}(P_2)$ such that

$$\tilde{x}_s^r = \begin{cases} \overline{V}, & s \in [r, n]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \quad \text{and} \quad \tilde{y}_s^r = \tilde{u}_s^r = \begin{cases} 1, & s \in [r, n]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} .$$

In addition, for a fixed node $i \in \mathcal{N}$ of inequality (4b) and for inequalities (5) - (8), we can similarly create $3n + 1$ linearly independent points in $\text{conv}(P_2)$ for each inequality and thus we omit them here.

Group	x_{i-}	x_i	$x_1\cdots x_{i-1}$	$x_{i+1}\cdots x_n$	y_{i-}	y_i	$y_1\cdots y_{i-1}$	$y_{i+1}\cdots y_n$	u_i	$u_1\cdots u_{i-1}$	$u_{i+1}\cdots u_n$
$(\acute{x},\acute{y},\acute{u})$	$\underline{C}+\epsilon$	$\underline{C}$	$0\cdots 0$	$0\cdots 0$	1	1	$0\cdots 0$	$0\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\underline{C}$	$\underline{C}$	$0\cdots 0$	$0\cdots 0$	1	1	$0\cdots 0$	$0\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\underline{C}$	$\underline{C}$	$\underline{C}\cdots 0$	$0\cdots 0$	1	1	$1\cdots 0$	$0\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(\bar{x},\bar{y},\bar{u})$	$\underline{C}$	$\underline{C}$	$\underline{C}\cdots\underline{C}$	$0\cdots 0$	1	1	$1\cdots 1$	$0\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\underline{C}$	$\underline{C}$	$\underline{C}\cdots\underline{C}$	$\underline{C}\cdots 0$	1	1	$1\cdots 1$	$1\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$\underline{C}$	$\underline{C}$	$\underline{C}\cdots\underline{C}$	$\underline{C}\cdots\underline{C}$	1	1	$1\cdots 1$	$1\cdots 1$	0	$0\cdots 0$	$0\cdots 0$
	$\overline{V}$	$\overline{V}$	$0\cdots 0$	$0\cdots 0$	1	1	$0\cdots 0$	$0\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\overline{V}$	$\overline{V}$	$\overline{V}\cdots 0$	$0\cdots 0$	1	1	$1\cdots 0$	$0\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(\hat{x},\hat{y},\hat{u})$	$\overline{V}$	$\overline{V}$	$\overline{V}\cdots\overline{V}$	$0\cdots 0$	1	1	$1\cdots 1$	$0\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\overline{V}$	$\overline{V}$	$\overline{V}\cdots\overline{V}$	$\overline{V}\cdots 0$	1	1	$1\cdots 1$	$1\cdots 0$	0	$0\cdots 0$	$0\cdots 0$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$\overline{V}$	$\overline{V}$	$\overline{V}\cdots\overline{V}$	$\overline{V}\cdots\overline{V}$	1	1	$1\cdots 1$	$1\cdots 1$	0	$0\cdots 0$	$0\cdots 0$
$(\dot{x},\dot{y},\dot{u})$	0	$\underline{C}$	$\underline{C}\cdots\underline{C}$	$\underline{C}\cdots\underline{C}$	0	1	$1\cdots 1$	$1\cdots 1$	1	$1\cdots 1$	$1\cdots 1$
	0	0	$\underline{C}\cdots\underline{C}$	$\underline{C}\cdots\underline{C}$	0	0	$1\cdots 1$	$1\cdots 1$	0	$1\cdots 1$	$1\cdots 1$
	0	0	$0\cdots\underline{C}$	$\underline{C}\cdots\underline{C}$	0	0	$0\cdots 1$	$1\cdots 1$	0	$0\cdots 1$	$1\cdots 1$
$(\tilde{x},\tilde{y},\tilde{u})$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	0	0	$0\cdots 0$	$\underline{C}\cdots\underline{C}$	0	0	$0\cdots 0$	$1\cdots 1$	0	$0\cdots 0$	$1\cdots 1$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	0	0	$0\cdots 0$	$0\cdots\underline{C}$	0	0	$0\cdots 0$	$0\cdots 1$	0	$0\cdots 0$	$0\cdots 1$

Table EC.2 $3n+1$ **linearly independent points for** $y_i - y_{i-} - u_i \le 0,\ \forall i \in \mathcal{N}$

For inequality (9) $x_i - x_j \le (\overline{V}+V)y_i - Vu_i - \underline{C}y_j + (\underline{C}+V-\overline{V})(y_k-u_k)$, $\forall i,j,k \in \mathcal{N}, i \neq j$:   Due to the scenario node symmetry, without loss of generality we assume $j \neq i$ and $i=1, j=2, k=n$. Now, we create $3n+1$ linearly independent points in $\mathrm{conv}(P_2)$ in the following six groups:

(i) We create a point $(\dot{x},\dot{y},\dot{u}) \in \mathrm{conv}(P_2)$ (totally one point) such that

$$\dot{x}_s = \begin{cases} \underline{C}, & s=0 \\ 0, & \text{o.w.} \end{cases}, \quad \dot{y}_s = \begin{cases} 1, & s=0 \\ 0, & \text{o.w.} \end{cases}, \quad \text{and } \begin{matrix} \dot{u}_s = 0, \\ s \in [1,n]_{\mathbb{Z}} \end{matrix}.$$

(ii) For each $r \in [0,n-1]_{\mathbb{Z}}$ (totally n points), we create $(\bar{x}^r,\bar{y}^r,\bar{u}^r) \in \mathrm{conv}(P_2)$ such that

$$\bar{x}^r_s = \begin{cases} \overline{V}, & s=0 \\ \overline{V}+V, & s=[0,r]_{\mathbb{Z}}\cap\{1\} \\ \underline{C}, & s\in[2,r]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases}, \quad \bar{y}^r_s = \begin{cases} 1, & s\in[0,r]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases}, \quad \text{and } \begin{matrix} \bar{u}^r_s = 0, \\ s\in[1,n]_{\mathbb{Z}} \end{matrix}.$$

(iii) For $r=n$ (totally one point), we create $(\bar{x}^r,\bar{y}^r,\bar{u}^r) \in \mathrm{conv}(P_2)$ such that

$$\bar{x}^r_s = \begin{cases} \underline{C}+V, & s=0 \\ \underline{C}+2V, & s=1 \\ \underline{C}, & s\in[2,n]_{\mathbb{Z}} \end{cases}, \quad \begin{matrix} \bar{y}^r_s=1, \\ s\in[0,n]_{\mathbb{Z}} \end{matrix}, \quad \text{and } \begin{matrix} \bar{u}^r_s=0, \\ s\in[1,n]_{\mathbb{Z}} \end{matrix}.$$

(iv) We create a point $(\ddot{x},\ddot{y},\ddot{u}) \in \mathrm{conv}(P_2)$ (totally one point) such that

$$\ddot{x}_s = \begin{cases} \overline{C}-V, & s=0 \\ \overline{C}, & s=1 \\ \overline{C}-2V, & s\in[2,n]_{\mathbb{Z}} \end{cases}, \quad \begin{matrix} \ddot{y}_s=1, \\ s\in[0,n]_{\mathbb{Z}} \end{matrix}, \quad \text{and } \begin{matrix} \ddot{u}_s=0, \\ s\in[1,n]_{\mathbb{Z}} \end{matrix}.$$

(v) For each $r \in [1,n]_\mathbb{Z}$ (totally n points), we create $(\hat{x}^r, \hat{y}^r, \hat{u}^r) \in \text{conv}(P_2)$ such that

$$\hat{x}^r_s = \begin{cases} \overline{V}, & s \in [r,n]_\mathbb{Z} \cap \{1\} \\ \underline{C}, & s \in [r,n]_\mathbb{Z} \setminus \{1\} \\ 0, & \text{o.w.} \end{cases}, \text{ and } \hat{y}^r_s = \hat{u}^r_s = \begin{cases} 1, & s \in [r,n]_\mathbb{Z} \\ 0, & \text{o.w.} \end{cases}.$$

(vi) For each $r \in [3,n]_\mathbb{Z}$ (totally $n-2$ points), we create $(\tilde{x}^r, \tilde{y}^r, \tilde{u}^r) \in \text{conv}(P_2)$ such that

$$\tilde{x}^r_s = \begin{cases} \overline{V}, & s \in [r,n]_\mathbb{Z} \\ 0, & \text{o.w.} \end{cases}, \quad \tilde{y}^r_s = \begin{cases} 1, & s \in [r,n]_\mathbb{Z} \\ 0, & \text{o.w.} \end{cases}, \text{ and } \tilde{u}^r_s = \begin{cases} 1, & s \in [r,n]_\mathbb{Z} \\ 0, & \text{o.w.} \end{cases}.$$

We collect these $3n+1$ linearly independent points in Table EC.3.

Group	x						y	u
	x_{i^-} $\quad x_1 \quad\ x_2 \quad\ x_3 \quad \cdots\ x_{n-1} \quad x_n$						$y_{i^-}\, y_1 y_2 y_3 \cdots y_{n-1} y_n$	$u_1 u_2 u_3 \cdots u_{n-1} u_n$
$(\hat{x},\hat{y},\hat{u})$	$\underline{C}$ $\quad\ 0 \quad\ 0 \quad\ 0 \quad \cdots\ 0 \quad\ 0$						$1\ 0\ 0\ 0\cdots\ 0\ \ 0$	$0\ 0\ 0\cdots\ 0\ \ 0$
	$\overline{V}$ $\quad\ 0 \quad\ 0 \quad\ 0 \quad \cdots\ 0 \quad\ 0$						$1\ 0\ 0\ 0\cdots\ 0\ \ 0$	$0\ 0\ 0\cdots\ 0\ \ 0$
	$\overline{V}$ $\quad \overline{V}+V \quad 0 \quad\ 0 \quad \cdots\ 0 \quad\ 0$						$1\ 1\ 0\ 0\cdots\ 0\ \ 0$	$0\ 0\ 0\cdots\ 0\ \ 0$
	$\overline{V}$ $\quad \overline{V}+V \quad \underline{C} \quad\ 0 \quad \cdots\ 0 \quad\ 0$						$1\ 1\ 1\ 0\cdots\ 0\ \ 0$	$0\ 0\ 0\cdots\ 0\ \ 0$
$(\bar{x},\bar{y},\bar{u})$	$\overline{V}$ $\quad \overline{V}+V \quad \underline{C} \quad \underline{C} \quad \cdots\ 0 \quad\ 0$						$1\ 1\ 1\ 1\cdots\ 0\ \ 0$	$0\ 0\ 0\cdots\ 0\ \ 0$
	$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad\quad \vdots \qquad \vdots$						$\vdots\ \vdots\ \vdots\ \vdots \quad \vdots\ \vdots$	$\vdots\ \vdots\ \vdots \quad \vdots\ \vdots$
	$\overline{V}$ $\quad \overline{V}+V \quad \underline{C} \quad \underline{C} \quad \cdots\ \underline{C} \quad 0$						$1\ 1\ 1\ 1\cdots\ 1\ \ 0$	$0\ 0\ 0\cdots\ 0\ \ 0$
	$\underline{C}+V\,\underline{C}+2V \quad \underline{C} \quad \underline{C} \quad \cdots\ \underline{C} \quad \underline{C}$						$1\ 1\ 1\ 1\cdots\ 1\ \ 1$	$0\ 0\ 0\cdots\ 0\ \ 0$
$(\ddot{x},\ddot{y},\ddot{u})$	$\overline{C}-V\ \ \overline{C}\ \ \overline{C}-2V\ \overline{C}-2V\cdots\overline{C}-2V\ \overline{C}-2V$						$1\ 1\ 1\ 1\cdots\ 1\ \ 1$	$0\ 0\ 0\cdots\ 0\ \ 0$
	0 $\quad\ \overline{V} \quad \underline{C} \quad \underline{C} \quad \cdots\ \underline{C} \quad \underline{C}$						$0\ 1\ 1\ 1\cdots\ 1\ \ 1$	$1\ 1\ 1\cdots\ 1\ \ 1$
	0 $\quad\ 0 \quad \underline{C} \quad \underline{C} \quad \cdots\ \underline{C} \quad \underline{C}$						$0\ 0\ 1\ 1\cdots\ 1\ \ 1$	$0\ 1\ 1\cdots\ 1\ \ 1$
	0 $\quad\ 0 \quad\ 0 \quad \underline{C} \quad \cdots\ \underline{C} \quad \underline{C}$						$0\ 0\ 0\ 1\cdots\ 1\ \ 1$	$0\ 0\ 1\cdots\ 1\ \ 1$
$(\hat{x},\hat{y},\hat{u})$	$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad\quad \vdots \qquad \vdots$						$\vdots\ \vdots\ \vdots\ \vdots \quad \vdots\ \vdots$	$\vdots\ \vdots\ \vdots \quad \vdots\ \vdots$
	0 $\quad\ 0 \quad\ 0 \quad\ 0 \quad \cdots\ \underline{C} \quad \underline{C}$						$0\ 0\ 0\ 0\cdots\ 1\ \ 1$	$0\ 0\ 0\cdots\ 1\ \ 1$
	0 $\quad\ 0 \quad\ 0 \quad\ 0 \quad \cdots\ 0 \quad \underline{C}$						$0\ 0\ 0\ 0\cdots\ 0\ \ 1$	$0\ 0\ 0\cdots\ 0\ \ 1$
	0 $\quad\ 0 \quad\ 0 \quad \overline{V} \quad \cdots\ \overline{V} \quad \overline{V}$						$0\ 0\ 0\ 1\cdots\ 1\ \ 1$	$0\ 0\ 1\cdots\ 1\ \ 1$
$(\tilde{x},\tilde{y},\tilde{u})$	$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad\quad \vdots \qquad \vdots$						$\vdots\ \vdots\ \vdots\ \vdots \quad \vdots\ \vdots$	$\vdots\ \vdots\ \vdots \quad \vdots\ \vdots$
	0 $\quad\ 0 \quad\ 0 \quad\ 0 \quad \cdots\ \overline{V} \quad \overline{V}$						$0\ 0\ 0\ 0\cdots\ 1\ \ 1$	$0\ 0\ 0\cdots\ 1\ \ 1$
	0 $\quad\ 0 \quad\ 0 \quad\ 0 \quad \cdots\ 0 \quad \overline{V}$						$0\ 0\ 0\ 0\cdots\ 0\ \ 1$	$0\ 0\ 0\cdots\ 0\ \ 1$

Table EC.3 $\quad 3n+1$ **linearly independent points for** (9)

In summary, all of the inequalities in Q_2 are facet-defining for $\text{conv}(P_2)$. $\qquad\square$

D.4. Proof of Lemma EC.1

Proof of Lemma EC.1 Let $A^+ = \{i \in [1,n]_\mathbb{Z} : a_i \geq 0\}$, $A^- = \{i \in [1,n]_\mathbb{Z} : a_i < 0\}$. We discuss two different cases based on the online/offline status of the generator at the root node.

(1) The generator is offline at the root node, i.e., $x_0 = y_0 = 0$. For this case, we further discuss the following two situations based on if $i \in A^+$ or $i \in A^-$ for each $i \in [1,n]_\mathbb{Z}$:

(i) If $i \in A^+$, to maximize the objective function (EC.14), the generator at node i should be scheduled online at its start-up ramp rate limit $\overline{V}$ following constraints (4c) and (4d) if $a_i\overline{V} + b_i + c_i \geq 0$ or offline otherwise. It follows that $(x_i, y_i, u_i) = (\overline{V}, 1, 1)$ if the generator is online at node i or $(x_i, y_i, u_i) = (0, 0, 0)$ otherwise.

(ii) If $i \in A^-$, to maximize the objective function (EC.14), the generator at node i should be scheduled online at its minimum generation amount $\underline{C}$ if $a_i \underline{C} + b_i + c_i \geq 0$ or offline otherwise. It follows that $(x_i, y_i, u_i) = (\underline{C}, 1, 1)$ if the generator is online at node i or $(x_i, y_i, u_i) = (0, 0, 0)$ otherwise.

From the above (i) and (ii), we verified Claim (1).

(2) The generator is scheduled online at the root node, i.e., $y_0 = 1$. It follows that $u_i = 0$ for all $i = 1, \ldots, n$. For notational brevity, we let $\bar{A}^+(x_0) = \{i \in [1, n]_{\mathbb{Z}} : a_i \geq 0, \quad a_i \min\{\overline{C}, x_0 + V\} + b_i \geq 0\}$ and $\bar{A}^-(x_0) = \{i \in [1, n]_{\mathbb{Z}} : a_i < 0, \quad a_i \max\{\underline{C}, x_0 - V\} + b_i \geq 0\}$. We further discuss the following two cases in terms of the value of x_0:

(i) If $\underline{C} \leq x_0 \leq \overline{V}$, similar to (1) above, we further discuss the following two cases based on if $i \in A^+$ or $i \in A^-$ for each $i \in [1, n]_{\mathbb{Z}}$:

  1) If $i \in A^+$, to maximize the objective function (EC.14), the generator at node i should be scheduled online at $\min\{\overline{C}, x_0 + V\}$ following constraints (4c) and (4d) if $i \in \bar{A}^+(x_0)$ or offline otherwise. It follows that $(x_i, y_i, u_i) = (\min\{\overline{C}, x_0 + V\}, 1, 0)$ if the generator is online at node i or $(x_i, y_i, u_i) = (0, 0, 0)$ otherwise.

  2) If $i \in A^-$, to maximize the objective function (EC.14), the generator at node i should be scheduled to be online at $\max\{\underline{C}, x_0 - V\}$ following constraints (4b) and (4d) if $i \in \bar{A}^-(x_0)$ or offline otherwise. It follows that $(x_i, y_i, u_i) = (\max\{\underline{C}, x_0 - V\}, 1, 0)$ if the generator is online at node i or $(x_i, y_i, u_i) = (0, 0, 0)$ otherwise.

Based on the above 1) and 2), we can write the optimal objective value of (EC.14) for a given set of $(a_0, b_0, a_i, b_i, c_i), i = 1, \ldots, n$, as a function of x_0. Denote it as $g(x_0) = (a_0 x_0 + b_0) + \sum_{i \in \bar{A}^+(x_0)} (a_i \min\{\overline{C}, x_0 + V\} + b_i) + \sum_{i \in \bar{A}^-(x_0)} (a_i \max\{\underline{C}, x_0 - V\} + b_i)$, which is a continuous function with respect to x_0 on $[\underline{C}, \overline{V}]$. Thus, $z^* = \max\{g(x_0) : \underline{C} \leq x_0 \leq \overline{V}\}$. Because $\underline{C} \leq x_0 \leq \overline{V}$, it follows that $\min\{\overline{C}, x_0 + V\} = x_0 + V$ and $\max\{\underline{C}, x_0 - V\} = \underline{C}$. Then we have $g(x_0) = (a_0 x_0 + b_0) + \sum_{i \in A^+} [a_i(x_0 + V) + b_i]^+ + \sum_{i \in A^-} [a_i \underline{C} + b_i]^+$ where we define $[t]^+ = \max\{0, t\}$ for $\forall t \in \mathbb{R}$. Thus, $g(x_0)$ is a convex function with respect to x_0 on $[\underline{C}, \overline{V}]$. It follows that the optimal solutions happen at the points where $x_0 = \underline{C}$ or $\overline{V}$. Now we discuss these two scenarios as follows.

- When $x_0 = \underline{C}$, $x_i, i = 1, \ldots, n$, can be obtained based on (i) and (ii) right above. Thus, Claim (2) is verified.

- When $x_0 = \overline{V}$, there exists at least one x_k for some $k \in \{1, \ldots, n\}$ such that $x_k = 0$. This can be proved by contradiction. If no such x_k exists, then x_i can only be either $\underline{C}$ or $\underline{C} + V$ based on the calculation (i) and (ii) right above. Without loss of generality, we let $x_i = \underline{C} + V$ for each $i \in \mathcal{N}_1 \subseteq \mathcal{N}$ and $x_i = \underline{C}$ for each $i \in \mathcal{N} \setminus \mathcal{N}_1$. It is easy to observe that this solution (denoted as $(x; y; u)$) can be written as a linear combination of the following two solutions:

$$(x; y; u) = (1/2)(\hat{x}; \hat{y}; \hat{u}) + (1/2)(\tilde{x}; \tilde{y}; \tilde{u}),$$

where $\hat{y} = \tilde{y} = y, \hat{u} = \tilde{u} = u, \hat{x}_0 = x_0 + \epsilon, \tilde{x}_0 = x_0 - \epsilon, \hat{x}_i = x_i + \epsilon$ and $\tilde{x}_i = x_i - \epsilon$ for each $i \in \mathcal{N}_1$, and $\hat{x}_i = \tilde{x}_i = x_i$ for each $i \in \mathcal{N} \setminus \mathcal{N}_1$, with ϵ representing an arbitrarily small positive real number. This is a contradiction because $(x; y; u)$ should be an extreme point of $\text{conv}(P_2)$ if there is only one optimal solution for (EC.14). Thus, Claim (3) is verified.

(ii) If $\overline{V} < x_0 \le \overline{C}$, similar to (1) above, we further discuss the following two cases based on if $i \in A^+$ or $i \in A^-$ for each $i \in [1, n]_{\mathbb{Z}}$:

    1) If $i \in A^+$, to maximize the objective function (EC.14), the generator at node i should be scheduled online at $\min\{\overline{C}, x_0 + V\}$ following constraints (4c) and (4d). It follows that $(x_i, y_i, u_i) = (\min\{\overline{C}, x_0 + V\}, 1, 0)$.

    2) If $i \in A^-$, to maximize the objective function (EC.14), the generator at node i should be scheduled to be online at $\max\{\underline{C}, x_0 - V\}$ following constraints (4b) and (4d). It follows that $(x_i, y_i, u_i) = (\max\{\underline{C}, x_0 - V\}, 1, 0)$.

Note that because $x_0 > \overline{V}$, the generator at node i (for each $i \in \mathcal{N}$) should be scheduled online so that ramp-down rate constraints (4d) hold. Based on the above 1) and 2), we can write the optimal objective value of (EC.14) for a given set of $(a_0, b_0, a_i, b_i, c_i), i = 1, \ldots, n$, as a function of x_0. Denote it as $g(x_0) = (a_0 x_0 + b_0) + \sum_{i \in A^+} (a_i \min\{\overline{C}, x_0 + V\} + b_i) + \sum_{i \in A^-} (a_i \max\{\underline{C}, x_0 - V\} + b_i)$, which is a continuous function with respect to x_0 on $(\overline{V}, \overline{C}]$. Thus, $z^* = \max\{g(x_0) : \overline{V} < x_0 \le \overline{C}\}$.

To obtain an explicit formula of z^*, we continue considering the following three situations:

    i) If $\overline{V} < x_0 \le \underline{C} + V$, it follows that $\min\{\overline{C}, x_0 + V\} = x_0 + V$ and $\max\{\underline{C}, x_0 - V\} = \underline{C}$. Then we have $g(x_0) = (a_0 x_0 + b_0) + \sum_{i \in A^+} [a_i (x_0 + V) + b_i] + \sum_{i \in A^-} (a_i \underline{C} + b_i)$, which is a convex function with respect to x_0 on $(\overline{V}, \underline{C} + V]$. Thus, the optimal solutions happen at the points where $x_0 = \overline{V} + \epsilon$ or $\underline{C} + V$. A similar contradiction argument as that in (i) can be applied to show the point with $x_0 = \overline{V} + \epsilon$ is not an optimal solution. Thus we only need to discuss the case when $x_0 = \underline{C} + V$. Similarly, we can follow the contradiction argument in (i) right above and calculation in 1) and 2) to verify Claim (4).

    ii) If $\underline{C} + V \le x_0 \le \overline{C} - V$, it follows that $\min\{\overline{C}, x_0 + V\} = x_0 + V$ and $\max\{\underline{C}, x_0 - V\} = x_0 - V$. Then we have $g(x_0) = (a_0 x_0 + b_0) + \sum_{i \in A^+} [a_i (x_0 + V) + b_i] + \sum_{i \in A^-} [a_i (x_0 - V) + b_i]$, which is a convex function with respect to x_0 on $[\underline{C} + V, \overline{C} - V]$. Thus, the optimal solutions happen at the points where $x_0 = \underline{C} + V$ or $\overline{C} - V$. Here we only need to discuss the case when $x_0 = \overline{C} - V$. Similarly, we can follow the contradiction argument in (i) right above and calculation in 1) and 2) to verify Claim (5).

    iii) If $\overline{C} - V \le x_0 \le \overline{C}$, it follows that $\min\{\overline{C}, x_0 + V\} = \overline{C}$ and $\max\{\underline{C}, x_0 - V\} = x_0 - V$. Then we have $g(x_0) = (a_0 x_0 + b_0) + \sum_{i \in A^+} (a_i \overline{C} + b_i) + \sum_{i \in A^-} [a_i (x_0 - V) + b_i]$, which is a convex function with respect to x_0 on $[\overline{C} - V, \overline{C}]$. Thus, the optimal solutions happen at the points where $x_0 = \overline{C} - V$ or $\overline{C}$. We only need to consider the case when $x_0 = \overline{C}$. For this case, $x_i, i = 1, \ldots, n$, can be defined based on 1) and 2) right above. Hence, Claim (6) is verified.

This completes the proof. $\square$

D.5. Proof of Proposition EC.2

Proof of Proposition EC.2 We prove the claim by showing that every point in the six groups of extreme points described in Lemma EC.1 satisfies $3n + 2$ linearly independent inequalities at equation, which indicates that they are also extreme points of Q_2. Thus, the desired statement holds because of the facet-defining conditions shown in Proposition EC.1. In the following, we prove the claim in the following six possible cases:

1) For Group (1) points, we have $x_0 = y_0 = 0$ and let $x_i = 0$ for $i \in [1, r]_{\mathbb{Z}}$, $x_i = \underline{C}$ for $i \in [r+1, s]_{\mathbb{Z}}$, and $x_i = \overline{V}$ for $i \in [s+1, n]_{\mathbb{Z}}$ for some given r and s. It follows that $y_i = u_i = 0$ for $i \in [1, r]_{\mathbb{Z}}$, $y_i = u_i = 1$ for $i \in [r+1, n]_{\mathbb{Z}}$. Without loss of generality, we only consider the case in which $r \geq 1$, $s \geq r+1$, and $n \geq s+1$. That is, there exists at least one scenario corresponding to each possible generation amount x_i of 0, $\underline{C}$, or $\overline{V}$. The following $3n+2$ linearly independent inequalities, (4b) (for each $i \in [0, s]_{\mathbb{Z}}$), (6) (for each $i \in [s+1, n]_{\mathbb{Z}}$, and some $j \in [1, r]_{\mathbb{Z}}$), $y_i - y_{i-} - u_i \leq 0$ (some $i \in [1, n]_{\mathbb{Z}}$), $u_i - y_i \leq 0$ (for each $i \in [1, n]_{\mathbb{Z}}$), (10) (for each $i \in [1, r]_{\mathbb{Z}}$), and $u_i + y_{i-} \leq 1$ (for each $i \in [r+1, n]_{\mathbb{Z}}$) are tight.

2) For Group (2) points, we have $x_0 = \underline{C}$, $y_0 = 1$ and let $x_i = 0$ for $i \in [1, r]_{\mathbb{Z}}$, $x_i = \underline{C}$ for $i \in [r+1, s]_{\mathbb{Z}}$, and $x_i = \underline{C} + V$ for $i \in [s+1, n]_{\mathbb{Z}}$. It follows that $u_i = 0$ for $i \in [1, n]_{\mathbb{Z}}$, $y_i = 0$ for $i \in [1, r]_{\mathbb{Z}}$, and $y_i = 1$ for $i \in [r+1, n]_{\mathbb{Z}}$. Without loss of generality, we only consider the case in which $r \geq 1$, $s \geq r+1$, and $n \geq s+1$. That is, there exists at least one scenario corresponding to each possible generation amount x_i of 0, $\underline{C}$, or $\underline{C} + V$. The following $3n+2$ linearly independent inequalities, (4b) (for each $i \in [0, s]_{\mathbb{Z}}$), (7) (for each $i \in [s+1, n]_{\mathbb{Z}}$), $u_i - y_i \leq 0$ (for each $i \in [1, r]_{\mathbb{Z}}$), $y_i - y_{i-} - u_i \leq 0$ (for each $i \in [r+1, n]_{\mathbb{Z}}$), (10) (some $i \in [1, n]_{\mathbb{Z}}$), and $u_i + y_{i-} \leq 1$ (for each $i \in [1, n]_{\mathbb{Z}}$) are tight.

3) For Group (3) points, we have $x_0 = \overline{V}$, $y_0 = 1$ and let $x_i = 0$ for $i \in [1, r]_{\mathbb{Z}}$, $x_i = \underline{C}$ for $i \in [r+1, s]_{\mathbb{Z}}$, and $x_i = \overline{V} + V$ for $i \in [s+1, n]_{\mathbb{Z}}$. It follows that $u_i = 0$ for $i \in [1, n]_{\mathbb{Z}}$, $y_i = 0$ for $i \in [1, r]_{\mathbb{Z}}$, and $y_i = 1$ for $i \in [r+1, n]_{\mathbb{Z}}$. Without loss of generality, we only consider the case in which $r \geq 1$, $s \geq r+1$, and $n \geq s+1$. That is, there exists at least one scenario corresponding to each possible generation amount x_i of 0, $\underline{C}$, or $\overline{V} + V$. The following $3n+2$ linearly independent inequalities, (5) (some $i \in [1, n]_{\mathbb{Z}}$), (4b) (for each $i \in [1, s]_{\mathbb{Z}}$), (9) (for each $i \in [s+1, n]_{\mathbb{Z}}$, some $j \in [r+1, s]_{\mathbb{Z}}$, and some $k \in [1, r]_{\mathbb{Z}}$), $u_i - y_i \leq 0$ (for each $i \in [1, r]_{\mathbb{Z}}$), $y_i - y_{i-} - u_i \leq 0$ (for each $i \in [r+1, n]_{\mathbb{Z}}$), (10) (some $i \in [1, n]_{\mathbb{Z}}$), and $u_i + y_{i-} \leq 1$ (for each $i \in [1, n]_{\mathbb{Z}}$) are tight.

4) For Group (4) points, we have $x_0 = \underline{C} + V$, $y_0 = 1$ and let $x_i = \underline{C}$ for $i \in [1, r]_{\mathbb{Z}}$ and $x_i = \underline{C} + 2V$ for $i \in [r+1, n]_{\mathbb{Z}}$. It follows that $y_i = 1, u_i = 0$ for $i \in [1, n]_{\mathbb{Z}}$. Without loss of generality, we only consider the case in which $r \geq 1$ and $n \geq r+1$. That is, there exists at least one scenario corresponding to each possible generation amount x_i of $\underline{C}$ or $\underline{C} + 2V$. The following $3n+2$ linearly independent inequalities, (8) (some $i \in [1, r]_{\mathbb{Z}}$ and some $j \in [1, n]_{\mathbb{Z}}$), (4b) (for each $i \in [1, r]_{\mathbb{Z}}$), (7) (for each $i \in [r+1, n]_{\mathbb{Z}}$), $y_i - y_{i-} - u_i \leq 0$ (for each $i \in [1, n]_{\mathbb{Z}}$), (10) (some $i \in [1, n]_{\mathbb{Z}}$), and $u_i + y_{i-} \leq 1$ (for each $i \in [1, n]_{\mathbb{Z}}$) are tight.

5) For Group (5) points, we have $x_0 = \overline{C} - V$, $y_0 = 1$ and let $x_i = \overline{C}$ for $i \in [1, r]_{\mathbb{Z}}$ and $x_i = \overline{C} - 2V$ for $i \in [r+1, n]_{\mathbb{Z}}$. It follows that $y_i = 1, u_i = 0$ for $i \in [1, n]_{\mathbb{Z}}$. Without loss of generality, we only consider the case in which $r \geq 1$ and $n \geq r+1$. That is, there exists at least one scenario corresponding to each possible generation amount x_i of $\overline{C}$ or $\overline{C} - 2V$. The following $3n+2$ linearly independent inequalities, (7) (some $i \in [1, r]_{\mathbb{Z}}$), (6) (for each $j = i \in [1, r]_{\mathbb{Z}}$), (9) (for each $j \in [r+1, n]_{\mathbb{Z}}$, some $i \in [1, r]_{\mathbb{Z}}$, and some $k \in [1, n]_{\mathbb{Z}}$), $y_i - y_{i-} - u_i \leq 0$ (for each $i \in [1, n]_{\mathbb{Z}}$), (10) (some $i \in [1, n]_{\mathbb{Z}}$), and $u_i + y_{i-} \leq 1$ (for each $i \in [1, n]_{\mathbb{Z}}$) are tight.

6) For Group (6) points, we have $x_0 = \overline{C}$, $y_0 = 1$ and let $x_i = \overline{C}$ for $i \in [1, r]_{\mathbb{Z}}$ and $x_i = \overline{C} - V$ for $i \in [r+1, n]_{\mathbb{Z}}$. It follows that $y_i = 1, u_i = 0$ for $i \in [1, n]_{\mathbb{Z}}$. Without loss of generality, we only consider the case in which $r \geq 1$ and $n \geq r + 1$. That is, there exists at least one scenario corresponding to each possible generation amount x_i of $\overline{C}$ or $\overline{C} - V$. The following $3n + 2$ linearly independent inequalities, (5) (some $i \in [1, n]_{\mathbb{Z}}$), (6) (for each $j = i \in [1, r]_{\mathbb{Z}}$), (8) (for each $j = i \in [r+1, n]_{\mathbb{Z}}$), $y_i - y_{i-} - u_i \leq 0$ (for each $i \in [1, n]_{\mathbb{Z}}$), (10) (some $i \in [1, n]_{\mathbb{Z}}$), and $u_i + y_{i-} \leq 1$ (for each $i \in [1, n]_{\mathbb{Z}}$) are tight.

In summary, we show that all of the extreme points of Q_2 are integral in y and u. $\qquad\square$

D.6. Convex Hull Results for Three-Period Cases

For the case in which the minimum-up/-down times are 2 ($L = \ell = 2$), the mathematical formulation for the original set P can be described as $P_3^2 := \{(x, y, u) \in \mathbb{R}^{n+2} \times \mathbb{B}^{n+2} \times \mathbb{B}^{n+1} : (11a), (11d) - (11g),$

$$u_{i-} + u_i - y_i \leq 0, \quad \forall i \in \mathcal{N}, \tag{EC.15a}$$

$$y_{i_2^-} + u_{i-} + u_i \leq 1, \quad \forall i \in \mathcal{N}\}. \tag{EC.15b}$$

THEOREM EC.1. *For a three-period problem in which the minimum-up/-down times are 2, $conv(P_3^2)$ can be described as* $Q_3^2 = \{(x, y, u) \in \mathbb{R}^{3n+5} : (11a), (EC.15a) - (EC.15b), (11d), (12a),$

$$x_{i_2^-} \leq \overline{V} y_{i_2^-} + V(y_{i-} - u_{i-}) + (\overline{C} - \overline{V} - V)(y_i - u_i - u_{i-}), \quad \forall i \in \mathcal{N}, \tag{EC.16a}$$

$$x_{i-} \leq \overline{V} y_{i-} + (\overline{C} - \overline{V})(y_i - u_i - u_{i-}), \quad \forall i \in \mathcal{N}, \tag{EC.16b}$$

$$x_i \leq (\overline{V} + V) y_i - V u_i + (\overline{C} - \overline{V} - V)(y_j - u_j - u_{i-}), \quad \forall i, j \in \mathcal{N}, \tag{EC.16c}$$

$$x_{i-} - x_{i_2^-} \leq \overline{V} y_{i-} - \underline{C} y_{i_2^-} + (\underline{C} + V - \overline{V})(y_i - u_i - u_{i-}), \quad \forall i \in \mathcal{N}, \tag{EC.16d}$$

$$x_{i_2^-} - x_{i-} \leq \overline{V} y_{i_2^-} - \underline{C} y_{i-} + (\underline{C} + V - \overline{V})(y_{i-} - u_{i-}), \tag{EC.16e}$$

$$x_i - x_{i-} \leq \overline{V} y_i - \underline{C} y_{i-} + (\underline{C} + V - \overline{V})(y_i - u_i), \quad \forall i \in \mathcal{N}, \tag{EC.16f}$$

$$x_{i-} - x_i \leq \overline{V} y_{i-} - \underline{C} y_i + (\underline{C} + V - \overline{V})(y_j - u_j - u_{i-}), \quad \forall i, j \in \mathcal{N}, \tag{EC.16g}$$

$$x_i - x_{i_2^-} \leq (\overline{V} + V) y_i - V u_i - \underline{C} y_{i_2^-} + (\underline{C} + V - \overline{V})(y_j - u_j - u_{i-}), \quad \forall i, j \in \mathcal{N}, \tag{EC.16h}$$

$$x_{i_2^-} - x_i \leq \overline{V} y_{i_2^-} - \underline{C} y_i + V(y_{i-} - u_{i-}) + (\underline{C} + V - \overline{V})(y_j - u_j - u_{i-}), \quad \forall i, j \in \mathcal{N}, \tag{EC.16i}$$

$$x_i - x_j \leq (\overline{V} + V) y_i - V u_i - \underline{C} y_j + (\underline{C} + V - \overline{V})(y_k - u_k - u_{i-}), \forall i, j, k \in \mathcal{N}, i \neq j, \tag{EC.16j}$$

$$x_{i_2^-} - x_{i-} + x_i \leq \overline{V} y_{i_2^-} - [(\overline{V} - V) y_{i-} + V u_{i-}] + [(\overline{V} + V) y_i - V u_i]$$
$$+ (\overline{C} - \overline{V} - V)(y_j - u_j - u_{i-}), \quad \forall i, j \in \mathcal{N}, \tag{EC.16k}$$

$$x_{i_2^-} - x_{i-} + x_i - x_j \leq \overline{V} y_{i_2^-} - [(\overline{V} - V) y_{i-} + V u_{i-}] + [(\overline{V} + V) y_i - V u_i] - \underline{C} y_j$$
$$+ (\underline{C} + V - \overline{V})(y_k - u_k - u_{i-}), \quad \forall i, j, k \in \mathcal{N}, i \neq j\}. \tag{EC.16l}$$

The proofs are similar to those for Theorem 1 and thus are omitted here.

THEOREM EC.2. *For the case in which $L = 1$ and $\ell = 2$, the convex hull representation of the original set (e.g., denoted as $P_3^{1,2}$) can be described as $Q_3^{1,2} = conv(P_3^{1,2}) = \{(x, y, u) \in \mathbb{R}^{3n+5} : (11a) - (11b), (EC.15b), (11d), (12a) - (12n)\}$.*

THEOREM EC.3. *For the case in which $L = 2$ and $\ell = 1$, the convex hull representation of the original set (e.g., denoted as $P_3^{2,1}$) can be described as $Q_3^{2,1} = conv(P_3^{2,1}) = \{(x, y, u) \in \mathbb{R}^{3n+5} : (11a), (EC.15a), (11c) - (11d), (12a), (EC.16a) - (EC.16l)\}$.*

Appendix E: Supplement to Section 4.2

To show that one inequality is facet-defining for $\text{conv}(P)$, we create $3|\mathcal{V}| - 1$ affinely independent points in $\text{conv}(P)$ that satisfy the inequality at equality. Because $0 \in \text{conv}(P)$, it is sufficient to create the remaining $3|\mathcal{V}| - 2$ linearly independent points. Meanwhile, for each node i in $\mathcal{V}$, we denote $\mathcal{C}(i)$ as the set of immediate children and $\mathcal{V}(i)$ as the set of all descendants of node i, including itself, respectively. For the convenience of generating points, we label the nodes in the tree as follows. Due to the symmetry of the scenario tree as shown in Figure EC.2 with each node having n children scenario nodes, we label the nodes in $\mathcal{V}$ following the breadth-first search rule as we did for the nodes of a three-period case in Figure EC.1, i.e., root node 0 is labelled as 0, the first node at $t(1)$ is labelled as 1, the first node at $t(2)$ is labelled as $n+1$, etc. Moreover, in Figure EC.2, without loss of generality, we assume that node p is the first node in the stage $t(p)$, node i is the first node in the stage $t(i)$, node j is the first node in set $\mathcal{C}(j^-) \setminus \{j^-\}$ (i.e., the set of scenario nodes following node j^-). Similarly, we assume that every node along the path from node i to node j passing through node p is the first scenario node among those following their corresponding parent node. In addition, we let $k_1 = dist(i,p) = |\mathcal{P}(i,p)|$, $k_2 = dist(j,p) = |\mathcal{P}(j,p)|$, $\mathcal{P}(i,j) = \mathcal{P}(i,p) \cup \mathcal{P}(j,p) \cup \{p\}$. Meanwhile, we use the superscript of (x,y,u), e.g., r in (x^r, y^r, u^r), to indicate the index of different points in $\text{conv}(P)$.

E.1. Proof of Proposition 4

Proof of Proposition 4 Here we only prove inequality (13) is facet-defining under condition (2), from which we have $\overline{C} \leq \overline{V} + kV$, as the case under condition (1) can be proved similarly. Due to the symmetry of the scenario nodes in Figure EC.2, without loss of generality we assume every node i_h^- along the path $\mathcal{P}(i)$ is the first node at period $t(i_h^-)$, as shown in Figure EC.2. We collect all of the nodes at period $t(i_h^-)$ $(\forall h \in [0,k]_{\mathbb{Z}})$ in $\Gamma(t(i_h^-))$, collect all of the nodes between period $[t(0), t(i_{k+1}^-)]$ in $\Gamma(0)$, and collect all of the nodes between period $[t(i)+1, T]$ in $\Gamma(T)$. In addition, we denote the second node at period $t(i_h^-)$ $(\forall h \in [0,k]_{\mathbb{Z}})$ as $\sigma(t(i_h^-))$ and the last node as $\tau(t(i_h^-))$.

In general, we create the points in two main steps. First, we create two groups of points, G_1 and G_2, which form a lower-triangular matrix in terms of the values y and x. Second, we create another two groups of points, G_3 and G_4, which form an upper-triangular matrix in terms of the value u. Finally, all of the points can form a similar structure like the points in Table EC.2.

Now we explain the details to construct the points in G_1 and G_2 for the first main step. First, we can easily observe that 1) the y part for G_1 and G_2 can be easily transformed to a lower-triangular matrix as shown in Table EC.4, with each row corresponding to one point, and 2) $u_s = 0$ for $\forall s \in \mathcal{V} \setminus \{0\}$, because y_s is ordered as the way we label the node above, i.e., the first one is y_0 for the root node. Then, we assign the value x for each row in Table EC.4. For each row from row 0 to row $|\Gamma(0)| - 1$, two groups of value are assigned to x, and for each remaining row, one group of value is assigned to x, to make the inequality tight at the points as follows. Eventually we will obtain $|\mathcal{V}| + |\Gamma(0)|$ linearly independent points in $\text{conv}(P)$. Because when $y_s = 0$, the corresponding $x_s = 0$, we only assign the value x when $y_s = 1$.

(i) For each row $r \in [0, |\Gamma(0)|-1]$, let $x_s = \overline{V}$ $(\forall s : y_s = 1)$ and assign this point to G_1; let $x_s = \underline{C}$ $(\forall s : y_s = 1)$ and assign this point to G_2.
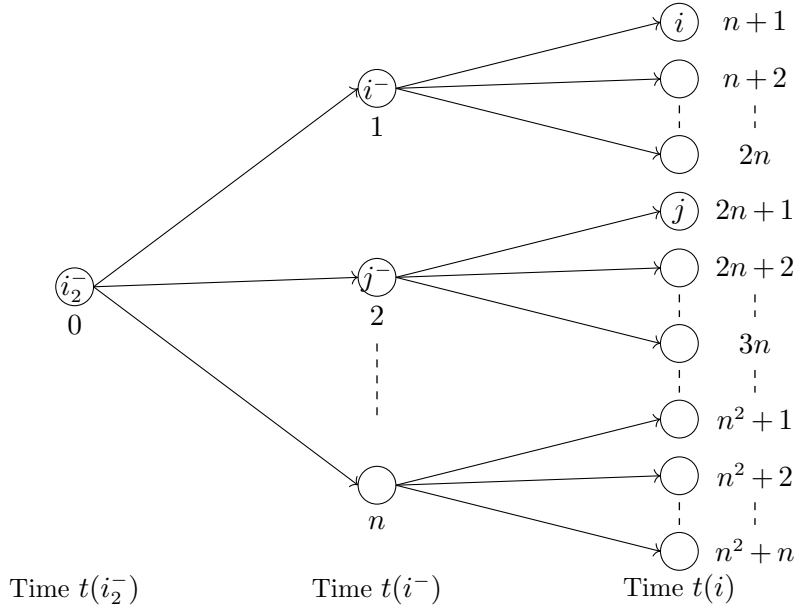
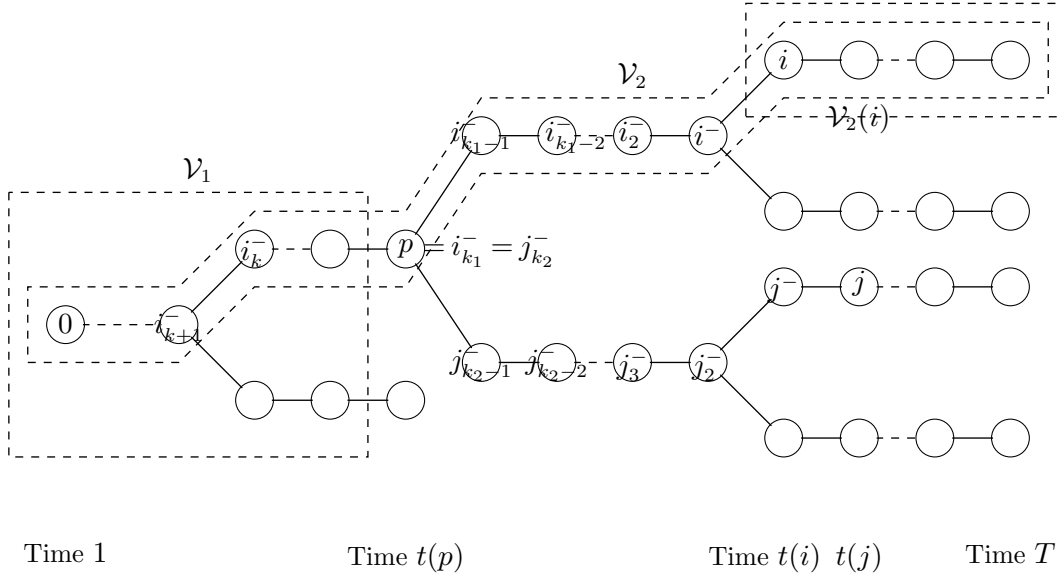Figure EC.1   Breadth-first search rule

Figure EC.2   Complete scenario tree

(ii) For each row $r \in [|\Gamma(0)|, |\Gamma(0)| + \sum_{h=1}^{k} |\Gamma(t(i_h^-))| - 1]$ such that $y_{i_h^-} = 1, y_{i_{h-1}^-} = 0$ $(\forall h \in [1, k]_{\mathbb{Z}})$, let $x_\theta = \overline{V} + (s-h)V$ for $\forall s \in [h, k]_{\mathbb{Z}}, \theta \in \Gamma(t(i_s^-))$ and $x_s = x_{i_k^-} = \overline{V} + (k-h)V$ for $\forall s \in \Gamma(0)$. We assign this point to G_1. This point is valid because $\overline{V} + (k-1)V \le \overline{C}$.

(iii) For each row $r \in [|\Gamma(0)| + \sum_{h=1}^{k} |\Gamma(t(i_h^-))|, |\mathcal{V}| - 1]$ where $y_i = 1$, let $x_s = \overline{V}$ for $\forall s \in \{\Gamma(t(i)) \cup \Gamma(T) : y_s = 1\}$, $x_\theta = \overline{V} + sV$ for $\forall s \in [1, k-1]_{\mathbb{Z}}, \theta \in \Gamma(t(i_s^-))$, and $x_s = x_{i_k^-} = \overline{C}$ for $\forall s \in \Gamma(0) \cup \Gamma(t(i_k^-))$. We assign this point to G_1. This point is valid because $\overline{V} + kV \ge \overline{C}$.

Next, we create another $2|\mathcal{V}| - |\Gamma(0)| - 2$ linearly independent points in two groups, G_3, G_4, while the value u for each group of points constructs an upper-triangular matrix, as shown in Table EC.5, which are

Table EC.4 Matrix in terms of y

row	$\Gamma(0)$	$\Gamma(t(i_k^-))$	⋯	$\Gamma(t(i^-))$	$\Gamma(t(i))$	$\Gamma(T)$
	$y_0 \cdots y_{i_{k+1}^-} \cdots y_{\tau(t(i_{k+1}^-))}$	$y_{i_k^-} \cdots y_{\tau(t(i_k^-))}$	⋯	$y_{i^-} \cdots y_{\tau(t(i^-))}$	$y_i \;\; \cdots y_{\tau(t(i))}$	$y_n, \forall n \in \Gamma(T)$
0	1 ⋯ 0 ⋯ 0	0 ⋯ 0	⋯	0 ⋯ 0	0 ⋯ 0	0 ⋯ 0
⋮	1 ⋯ 1 ⋯ 0	0 ⋯ 0	⋯	0 ⋯ 0	0 ⋯ 0	0 ⋯ 0
$\|\Gamma(0)\|-1$	1 ⋯ 1 ⋯ 1	0 ⋯ 0	⋯	0 ⋯ 0	0 ⋯ 0	0 ⋯ 0
$\|\Gamma(0)\|$	1 ⋯ 1 ⋯ 1	1 ⋯ 0	⋯	0 ⋯ 0	0 ⋯ 0	0 ⋯ 0
$\|\Gamma(0)\|+\|\Gamma(t(i_k^-))\|-1$	1 ⋯ 1 ⋯ 1	1 ⋯ 1	⋯	0 ⋯ 0	0 ⋯ 0	0 ⋯ 0
$\|\Gamma(0)\|+\sum\limits_{h=2}^{k}\|\Gamma(t(i_h^-))\|$	1 ⋯ 1 ⋯ 1	1 ⋯ 1	⋯	1 ⋯ 0	0 ⋯ 0	0 ⋯ 0
$\|\Gamma(0)\|+\sum\limits_{h=1}^{k}\|\Gamma(t(i_h^-))\|-1$	1 ⋯ 1 ⋯ 1	1 ⋯ 1	⋯	1 ⋯ 1	0 ⋯ 0	0 ⋯ 0
$\|\Gamma(0)\|+\sum\limits_{h=1}^{k}\|\Gamma(t(i_h^-))\|$	1 ⋯ 1 ⋯ 1	1 ⋯ 1	⋯	1 ⋯ 1	1 ⋯ 0	0 ⋯ 0
$\|\Gamma(0)\|+\sum\limits_{h=0}^{k}\|\Gamma(t(i_h^-))\|-1$	1 ⋯ 1 ⋯ 1	1 ⋯ 1	⋯	1 ⋯ 1	1 ⋯ 1	0 ⋯ 0
$\|\Gamma(0)\|+\sum\limits_{h=0}^{k}\|\Gamma(t(i_h^-))\|$	1 ⋯ 1 ⋯ 1	1 ⋯ 1	⋯	1 ⋯ 1	1 ⋯ 1	1 ⋯ 0
⋮						
$\|\mathcal{V}\|-1$	1 ⋯ 1 ⋯ 1	1 ⋯ 1	⋯	1 ⋯ 1	1 ⋯ 1	1 ⋯ 1

immediately linearly independent with G_1 and G_2 above. In Table EC.5, y_s ($\forall s \in [0,|\mathcal{V}|-1]_\mathbb{Z}$) is ordered as the way we label the node above, i.e., the first one is y_0 for the root node and $y_{|\mathcal{V}|-1}$ corresponds to the last node at the last period. The same order is applied to u_s ($\forall s \in [1,|\mathcal{V}|-1]_\mathbb{Z}$). In addition, we define $\hat{\mathcal{V}} = \mathcal{V} \setminus (\Gamma(0) \cup \Gamma(t(i_k^-)))$. For each row $r \in [1,|\mathcal{V}|-1]_\mathbb{Z}$, $y_s = 0$ ($\forall s \in [1,r]_\mathbb{Z}$). And at this row r, for those nodes with $u_s = 1$ (i.e., the generator starts up at the node s), the generator keeps online on the scenario node set $\mathcal{H}_L(s)$ for the minimum-up (L) periods.

(i) For each row $r \in [1,|\Gamma(0)|-1]_\mathbb{Z}$, we let $x_{i_k^-} = \max\{\underline{C},\overline{V}+(k-h)V\}$ if $y_{i_k^-} = 1$ and the generator shuts down at i_h^- along the path from i_k^- to i_h^-. The value of x for other nodes can be assigned easily. We assign this point to G_3. For the validity of this point, we let the generator start up at i_{k+1}^- and keep online for L periods, $x_{i_k^-}$ will be $\max\{\underline{C},\overline{V}+(L-2)V\}$, which should be less than $\overline{V}+V$ because of ramp-up constraints (1f). Because we assume $L \leq 3$, it is automatically valid. For other cases, it can be easily checked.

(ii) For each row $r \in [|\Gamma(0)|,|\mathcal{V}|-1]_\mathbb{Z}$, we let $x_s = \overline{V}$ for $\forall s$ such that $y_s = 1$ and assign this point to G_3.

(iii) For each row $r \in [|\Gamma(0)|+1,|\mathcal{V}|-1]_\mathbb{Z}$, we let $x_s = \underline{C}$ for $\forall s$ such that $y_s = 1$ and assign this point to G_4.

□

Table EC.5 Upper-triangular matrix in terms of u

row	$\Gamma(0)$		$\Gamma(t(i_k^-))$			$\hat{\mathcal{V}}$	$\Gamma(0)$		$\Gamma(t(i_k^-))$			$\hat{\mathcal{V}}$
	$y_0 y_1 \cdots$	$y_{\tau(t(i_{k+1}^-))}$	$y_{i_k^-}$ $y_{\sigma(t(i_k^-))}$	$\cdots$	$y_{\tau(t(i_k^-))}$	$y_n, n \in \hat{\mathcal{V}}$	$u_1 \cdots u_{\tau(t(i_{k+1}^-))}$		$u_{i_k^-}$ $u_{\sigma(t(i_k^-))} \cdots u_{\tau(t(i_k^-))}$			$u_n, n \in \hat{\mathcal{V}}$
1	$0 \; 1 \cdots$		$y_{\hat{s}} = 1, \forall \hat{s} \in \mathcal{H}_L(s), \forall s \in \Gamma(t(1))$			$\cdots \quad 0$	$u_s = 1, \forall s \text{ s.t. } t(s) = 1$		$0 \quad \cdots \quad 0$			$0 \quad \cdots \quad 0$
$\vdots$	$\vdots \; \vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots \; \vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots \qquad \vdots$
$|\Gamma(0)| - 1$	$0 \; 0 \cdots$	1	$y_{\hat{s}} = 1, \forall \hat{s} \in \mathcal{H}_L(s), \forall s \in \Gamma(t(i_k^-))$		0	$0 \cdots \qquad 1$		$u_s = 1, \forall s \in \Gamma(t(i_k^-)) \setminus \bar{\mathcal{H}}_1(\tau(t(i_{k+1}^-)))$			0
$|\Gamma(0)|$	$0 \; 0 \cdots$	0	$y_{\hat{s}} = 1, \forall \hat{s} \in \mathcal{H}_L(s), \forall s \in \Gamma(t(i_k^-))$		0	$0 \cdots \qquad 0$	1	$1 \quad \cdots \quad 1$			$0 \quad \cdots \quad 0$
$|\Gamma(0)| + 1$	$0 \; 0 \cdots$	0	$0 \quad 1$	$\cdots$	1	$1 \quad \cdots \quad 0$	$0 \cdots \qquad 0$	0	$1 \quad \cdots \quad 1$		$1 \quad \cdots \quad \cdots$
$\vdots$	$\vdots \; \vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots \; \vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots \qquad \vdots$
$\vdots$	$0 \; 0 \cdots$	0	$0 \quad 0$	$\cdots$	1	$1 \quad \cdots \quad 0$	$0 \cdots \qquad 0$	0	$0 \quad \cdots \quad 1$		$1 \quad \cdots \quad 0$
	$0 \; 0 \cdots$	0	$0 \quad 0$	$\cdots$	0	$1 \quad \cdots \quad 0$	$0 \cdots \qquad 0$	0	$0 \quad \cdots \quad 0$		$1 \quad \cdots \quad 0$
$\vdots$	$\vdots \; \vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots \; \vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots \qquad \vdots$
$|\mathcal{V}| - 1$	$0 \; 0 \cdots$	0	$0 \quad 0$	$\cdots$	0	$0 \quad \cdots \quad 1$	$0 \cdots \qquad 0$	0	$0 \quad \cdots \quad 0$		$0 \quad \cdots \quad 1$

For the proofs in the remaining part of this paper, we continue to apply this way to create linearly independent points, i.e., first create points constructing a lower-triangular matrix in terms of y, then create points constructing an upper-triangular matrix in terms of u, and in the meantime assign corresponding value to x. Meanwhile, from the construction of the linearly independent points here, we only need to consider the scenario tree with two different scenarios including nodes i and j, respectively, due to the symmetry of the scenario tree.

E.2. Proof of Proposition 5

Proof of Proposition 5 Following the proof in Online Supplement E.1, we apply the same way to create linearly independent points in $\mathrm{conv}(P)$ to prove inequality (14) is facet-defining for $\mathrm{conv}(P)$. Note that in this way of creating points, we only need to prove inequality (14) is facet-defining for $\mathrm{conv}(\bar{P})$, where $\bar{P}$ is constructed with the same constraints in P that are applied to the scenario structure $\bar{\mathcal{V}}$ in Figure EC.3, as shown in the proof described in Online Supplement E.1.
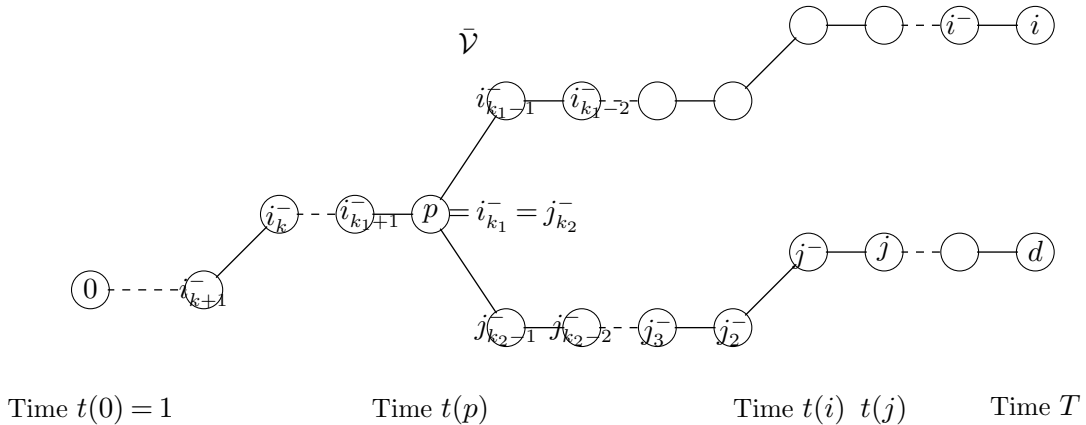

Figure EC.3 Complete scenario tree

Because node i is a leaf node, we have $k_1 \geq k_2$ and $t(i) = T$, where k_1 and k_2 are defined in Figure EC.3, $dist(i,j) = dist(i,p) + dist(j,p) = k_1 + k_2 = k$. We let node d be a leaf node in which $j \in \mathcal{P}(p,d)$. To simplify the process of creating linearly independent points, we reindex the nodes in $\bar{\mathcal{V}}$ as follows:

(i) The nodes $0, \ldots, i^-_{\bar{k}+1}, i^-_{\bar{k}}, \ldots, i^-_{\bar{k}_1+1}, p$ are reindexed as $0, \ldots, n - k_2 - 1, n - k_2, \ldots, n - 1, n$ with $n = t(p) - 1$,

(ii) The nodes $j^-_{\bar{k}_2-1}, j^-_{\bar{k}_2-2}, \ldots, j^-, j, \ldots, d$ are reindexed as $n+1, n+2, \ldots, n+k_2-1, n+k_2, \ldots, n+k_1$,

(iii) The nodes $i^-_{\bar{k}_1-1}, i^-_{\bar{k}_1-2}, \ldots, i^-, i$ are reindexed as $n+k_1+1, n+k_1+2, \ldots, n+2k_1-1, n+2k_1$.

In total, there are $n + 2k_1 + 1$ nodes.

Now we create $3(n + 2k_1 + 1) - 1$ affinely independent points in $\mathrm{conv}(\bar{P})$ that satisfy inequality (14) at equality. Because $0 \in \mathrm{conv}(\bar{P})$, we generate the remaining $3n + 6k_1 + 1$ linearly independent points in the following groups.

First, we create two groups of points based on a lower-triangular matrix in terms of y and assign corresponding value to x.

(i) For each $r \in [0, n + 2k_1 - 1]_{\mathbb{Z}}$ (totally $n + 2k_1$ points), we create $(\bar{x}^r, \bar{y}^r, \bar{u}^r) \in \mathrm{conv}(\bar{P})$ such that

$$\bar{x}^r_s = \begin{cases} \underline{C}, & s \in [0,r]_{\mathbb{Z}} \\ 0, & s \in [r+1, n+2k_1]_{\mathbb{Z}} \end{cases}, \quad \bar{y}^r_s = \begin{cases} 1, & s \in [0,r]_{\mathbb{Z}} \\ 0, & s \in [r+1, n+2k_1]_{\mathbb{Z}} \end{cases}, \quad \text{and} \quad \begin{matrix} \bar{u}^r_s = 0, \\ \forall s \in [1, n+2k_1]_{\mathbb{Z}} \end{matrix}.$$

(ii) For $r = n + 2k_1$ (totally one point), we create $(\bar{x}^r, \bar{y}^r, \bar{u}^r) \in \mathrm{conv}(\bar{P})$ such that

$$\bar{x}^r_s = \begin{cases} \underline{C} + k_2 V, & s \in [0,n]_{\mathbb{Z}} \\ \underline{C} + (n + k_2 - s)V, & s \in [n+1, n+k_2-1]_{\mathbb{Z}} \\ \underline{C}, & s \in [n+k_2, n+k_1]_{\mathbb{Z}} \\ \underline{C} + (k_2 + s - n - k_1)V, & s \in [n+k_1+1, n+2k_1]_{\mathbb{Z}} \end{cases}, \quad \begin{matrix} \bar{y}^r_s = 1, \\ \forall s \end{matrix}, \quad \text{and} \quad \begin{matrix} \bar{u}^r_s = 0, \\ \forall s \end{matrix}.$$

(iii) For each $r \in [0, n + k_2 - 1]_{\mathbb{Z}}$ (totally $n + k_2$ points), we create $(\hat{x}^r, \hat{y}^r, \hat{u}^r) \in \mathrm{conv}(\bar{P})$ such that

$$\hat{x}^r_s = \begin{cases} \overline{V}, & s \in [0,r]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases}, \quad \hat{y}^r_s = \begin{cases} 1, & s \in [0,r]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases}, \quad \text{and} \quad \begin{matrix} \hat{u}^r_s = 0, \\ \forall s \end{matrix}.$$

(iv) For each $r \in [n + k_2 + 1, n + 2k_1 - 1]_{\mathbb{Z}}$ (totally $(2k_1 - k_2 - 1)^+$ points), we create $(\hat{x}^r, \hat{y}^r, \hat{u}^r) \in \mathrm{conv}(\bar{P})$ such that

$$\hat{x}^r_s = \begin{cases} \overline{V}, & s \in [0,r]_{\mathbb{Z}} \setminus \{n+k_2\} \\ \underline{C}, & s = n + k_2 \\ 0, & \text{o.w.} \end{cases}, \quad \hat{y}^r_s = \begin{cases} 1, & s \in [0,r]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases}, \quad \text{and} \quad \begin{matrix} \hat{u}^r_s = 0, \\ \forall s \end{matrix}.$$

(v) For $r = n + 2k_1$ (totally one point), we create $(\hat{x}^r, \hat{y}^r, \hat{u}^r) \in \mathrm{conv}(\bar{P})$ such that

$$\hat{x}^r_s = \begin{cases} \underline{C} + k_2 V + \epsilon, & s \in [0,n]_{\mathbb{Z}} \\ \underline{C} + (n + k_2 - s)V + \epsilon, & s \in [n+1, n+k_2-1]_{\mathbb{Z}} \\ \underline{C} + \epsilon, & s \in [n+k_2, n+k_1]_{\mathbb{Z}} \\ \underline{C} + (k_2 + s - n - k_1)V + \epsilon, & s \in [n+k_1+1, n+2k_1]_{\mathbb{Z}} \end{cases}, \quad \begin{matrix} \hat{y}^r_s = 1, \\ \forall s \end{matrix}, \quad \text{and} \quad \begin{matrix} \hat{u}^r_s = 0, \\ \forall s \end{matrix}.$$

Next, we create a group of points based on an upper-triangular matrix in terms of u and assign corresponding value to x. Here we assume $\min\{L - 1, k - 1\} \geq k_1$, otherwise linearly independent points can be generated easily. In addition, without loss of generality, we let $k \geq L$ and then $\min\{L - 1, k - 1\} = L - 1$, as for the case in which $k \leq L - 1$ the linearly independent points can be generated similarly.

(vi) For each $r \in [1, n + k_1 - L]_{\mathbb{Z}}$ (totally $n + k_1 - L$ points), we create $(\acute{x}^r, \acute{y}^r, \acute{u}^r) \in \mathrm{conv}(\hat{P})$ such that

$$\acute{x}^r_s = \begin{cases} \underline{C}, & \text{s.t. } \acute{y}^r_s = 1 \\ 0, & \text{o.w.} \end{cases}, \quad \acute{y}^r_s = \begin{cases} 1, & s \in [r, r+L-1]_{\mathbb{Z}} \\ & \cup [n+k_1+1, k_1+r+L-1]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases}, \quad \text{and} \quad \acute{u}^r_s = \begin{cases} 1, & s = r \\ 0, & \text{o.w.} \end{cases}$$

(vii) For each $r \in [n+k_1-L+1, n]_{\mathbb{Z}}$ (totally $(L-k_1)^+$ points), i.e., the generator starts up at node $i^-_{n+k_1-r}$, we create $(\acute{x}^r, \acute{y}^r, \acute{u}^r) \in \text{conv}(\hat{P})$ such that

$$\acute{x}^r_s = \begin{cases} \overline{V} + (s-r)V, & s \in [r,n]_{\mathbb{Z}} \\ \min\{\overline{V} + (2n-s-r)V, \underline{C}\}, \\ \quad s \in [n+1, n+k_1]_{\mathbb{Z}} \\ \overline{V} + (s-k_1-r)V, \\ \quad s \in [n+k_1+1, n+2k_1]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \quad \acute{y}^r_s = \begin{cases} 1, & s \in [r, n+2k_1]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \text{ and } \acute{u}^r_s = \begin{cases} 1, & s=r \\ 0, & \text{o.w.} \end{cases}$$

(viii) For each $r \in [n+1, n+k_1]_{\mathbb{Z}}$ (totally k_1 points), we create $(\acute{x}^r, \acute{y}^r, \acute{u}^r) \in \text{conv}(\hat{P})$ such that

$$\acute{x}^r_s = \begin{cases} \underline{C}, & s \in [r, n+k_1]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \quad \acute{y}^r_s = \begin{cases} 1, & s \in [r, n+k_1]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \text{ and } \acute{u}^r_s = \begin{cases} 1, & s=r \\ 0, & \text{o.w.} \end{cases}$$

(ix) For each $r \in [n+k_1+1, n+2k_1]_{\mathbb{Z}}$ (totally k_1 points), we create $(\acute{x}^r, \acute{y}^r, \acute{u}^r) \in \text{conv}(\hat{P})$ such that

$$\acute{x}^r_s = \begin{cases} \overline{V} + (s-r)V, & s \in [r, n+2k_1]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \quad \acute{y}^r_s = \begin{cases} 1, & s \in [r, n+2k_1]_{\mathbb{Z}} \\ 0, & \text{o.w.} \end{cases} , \text{ and } \acute{u}^r_s = \begin{cases} 1, & s=r \\ 0, & \text{o.w.} \end{cases}$$

Finally, these linearly independent points, i.e., $(\bar{x}^r, \bar{y}^r, \bar{u}^r)_{r=0}^{n+2k_1}$, $(\hat{x}^r, \hat{y}^r, \hat{u}^r)_{r=0, r \neq n+k_2}^{n+2k_1}$, and $(\acute{x}^r, \acute{y}^r, \acute{u}^r)_{r=1}^{n+k_2}$, can construct a table similar to Table EC.2 and can be transformed to be a lower-triangular matrix easily. Thus we created $2(n+2k_1+1) - 1 + n + 2k_1 = 3n + 6k_1 + 1$ linearly independent points and thus the statement holds. $\qquad \square$

E.3. Proof of Proposition 7

Proof of Proposition 7 Here we only provide the proof for the case in which $\psi = (\underline{C} + V - \overline{V})(y_j - \sum_{m=0}^{L-1} u_{j_m^-})$ because the case in which $\psi = (\underline{C} + V - \overline{V})(y_i - \sum_{m=0}^{L-1} u_{i_m^-})$ can be proved similarly.

(Validity) It is clear that inequality (16) is valid when $y_{i_k^-} = 0$ due to constraints (1a). In the following, we continue to prove the validity by discussing the cases in which $y_{i_k^-} = 1$.

First, we consider the case in which $y_j = 0$. We let the last start-up node (denoted as i^-_{k+s}, $s \geq 0$) before i_k^- in the following two possible cases.

1) $s \geq L - 1$. It follows that $\phi = 0$ and $i^-_{k+s-L+1} \in \mathcal{P}(i_k^-)$. We further discuss the following three possible cases in terms of the first shut-down node (denoted as h, $h \notin \mathcal{P}(i_k^-)$) after i_k^-, i.e., $y_h = 0$. We observe that $y_n - \sum_{m=0: t(n_m^-) \geq 2}^{L-1} u_{n_m^-} = 1$ for each $n \in \mathcal{P}(h^-) \setminus \mathcal{P}(i_k^-)$.

   (1) $h \in \mathcal{P}(i^-_{k-\hat{n}})$. In this case, we have $h = i^-_{k-\hat{s}}$ for some $\hat{s} \in [1, \hat{n}]_{\mathbb{Z}}$ and $x_{i_k^-} \leq \overline{V} + \min\{k-1, s, \hat{s}-1\}V \leq \overline{V} + (\hat{s}-1)V = \overline{V}y_{i_k^-} + V \sum_{n=1}^{\hat{s}-1}(y_{i^-_{k-n}} - \sum_{m=0}^{\min\{L-1, n+w\}} u_{i^-_{k-n+m}})$, which is clearly less than the RHS of (16).

   (2) $h \in \mathcal{P}(i) \setminus \mathcal{P}(i^-_{k-\hat{n}})$. In this case, we have $h = i^-_{k-\hat{s}}$ for some $\hat{s} \in [\hat{n}+1, k]_{\mathbb{Z}}$ and $x_{i_k^-} \leq \overline{V} + \min\{k-1, s, \hat{s}-1\}V \leq \overline{V} + (\hat{s}-1)V = \overline{V} + (\hat{t}-1)V + (\hat{s}-\hat{t})V \leq V \sum_{n \in S_0}(y_{i^-_{k-n}} - \sum_{m=0}^{\min\{L-1, n+w\}} u_{i^-_{k-n+m}}) + V \sum_{n \in (S \cap [\hat{n}+1, k-\hat{s}+1]_{\mathbb{Z}}) \cup \{\hat{n}\}}(g_n - n)(y_{i^-_{k-n}} - \sum_{m=0}^{L-1} u_{i^-_{k-n+m}})$, which is clearly less than the RHS of (16).

   (3) $h \in \mathcal{V}(i) \setminus \mathcal{P}(i)$. Inequality (16) converts to $x_{i_k^-} \leq \underline{C} + kV \leq \overline{V} + (k-1)V$, which is clearly valid due to $y_j = 0$ and constraints (1f) and (1g).

2) $s \in [0, L-2]_{\mathbb{Z}}$. It follows that $i_{k+s-L+1}^- \in \mathcal{V}(i_k^-) \setminus \mathcal{P}(i_k^-)$ and $y_{i_{k-n}^-} - \sum_{m=0}^{\min\{L-1,n+w\}} u_{i_{k-n+m}^-}$ for each $n \in [1, L-1-s]_{\mathbb{Z}}$. We further discuss the following three possible cases in terms of the first shut-down node (denoted as h, $h \notin \mathcal{P}(i_{k+s-L+1}^-)$) after i_k^-, i.e., $y_h = 0$. We observe that $y_n - \sum_{m=0:t(n_m)\geq 2}^{L-1} u_{n_m} = 1$ for each $n \in \mathcal{P}(h^-) \setminus \mathcal{P}(i_{k+s-L+1}^-)$.

(1) $L-1-s \in [1, \hat{n}-1]_{\mathbb{Z}}$. We further discuss the following three possible cases in terms of the value h.

(a) $h \in \mathcal{P}(i_{k-\hat{n}}^-)$. In this case, we have $h = i_{k-\hat{s}}^-$ for some $\hat{s} \in [L-1-s, \hat{n}]_{\mathbb{Z}}$ and $x_{i_k^-} \leq \overline{V} + \min\{k-1, s, \hat{s}-1\}V \leq \overline{V} + \min\{s, \hat{s}-1\}V \leq \overline{V} + (\hat{s}-1-(L-1-s))V + \phi$ for $\phi = sV$ or $(L-1-s)V$, which is clearly the RHS of (16).

(b) $h \in \mathcal{P}(i) \setminus \mathcal{P}(i_{k-\hat{n}}^-)$. In this case, we have $h = i_{k-\hat{s}}^-$ for some $\hat{s} \in [\hat{n}+1, k]_{\mathbb{Z}}$ and $x_{i_k^-} \leq \overline{V} + \min\{k-1, s, \hat{s}-1\}V$. Inequality (16) converts to $x_{i_k^-} \leq \overline{V} + (\hat{n}-1-(L-1-s))V + \tilde{f} + \phi$, where $\tilde{f} = V \sum_{n \in (S \cup \{\hat{n}\}) \cap [\hat{n}, \hat{s}-1]_{\mathbb{Z}}} (g_n - n)$. It is easy to observe that $\tilde{f} \geq \hat{s} - \hat{t}$. Now we only need to show

$$\overline{V} + \min\{s, \hat{s}-1\}V \leq \overline{V} + (\hat{n}-1-(L-1-s))V + \tilde{f} + \phi. \qquad (\text{EC.17})$$

If $\phi = sV$, then (EC.17) holds clearly; otherwise, $\phi = (L-1-s)V$, then the RHS of (EC.17) becomes $\overline{V} + (\hat{n}-1+\tilde{f})V \geq \overline{V} + (\hat{s}-1)V$ due to $\tilde{f} \geq \hat{s}-\hat{t}$, indicating (EC.17) holds.

(c) $h \in \mathcal{V}(i) \setminus \mathcal{P}(i)$. In this case, we have $x_{i_k^-} \leq \overline{V} + \min\{k-1, s, \hat{s}-1\}V$. Inequality (16) converts to $x_{i_k^-} \leq \overline{V} + (\hat{n}-1-(L-1-s))V + V \sum_{n \in S \cup \{\hat{n}\}} (g_n - n) + \phi = \overline{V} + (\hat{n}-1-(L-1-s))V + (k-\hat{n}) + \phi$. Now we only need to show

$$\overline{V} + \min\{k-1, s, \hat{s}-1\}V \leq \overline{V} + (\hat{n}-1-(L-1-s))V + (k-\hat{n}) + \phi. \qquad (\text{EC.18})$$

If $\phi = sV$, then clearly (EC.18) holds; otherwise, $\phi = (L-1-s)V$, then the RHS of (EC.18) becomes $\overline{V} + (k-1)V$, indicating (EC.18) holds.

(2) $L-1-s \geq \hat{n}$. It follows that $s \leq L-1-\hat{n} \leq L-2$ Note that if $L-1-s \geq k$, we have $i_{k+s-L+1}^- \in \mathcal{V}(i) \setminus \mathcal{P}(i)$ We further discuss the following three possible cases in terms of the value h. Similar to the argument above, we only need to show

$$\overline{V} + \min\{k-1, s, \hat{s}-1\}V \leq \overline{V} + \tilde{f} + \phi, \qquad (\text{EC.19})$$

where $\tilde{f} = V \sum_{n \in (S \cup \{\hat{n}\}) \cap [0, L-s]_{\mathbb{Z}}} (g_n - n)$. If $\phi = sV$, then clearly (EC.19) holds; otherwise, $\phi = (L-1-s)V$, i.e., $L-1-s \leq s-1$. In the following, we assume $s \geq L-s$ and try to obtain the contradiction, which indicates that $s \geq L-s$ is not possible to happen.

(a) If $t(i_k^-) \geq L$, then $\min\{t(i_k^-)-2, L-2\} = L-2 \geq L/2$ and therefore $\hat{n} = L-2$ by the definition of $\hat{n}$. It follows that $L-1-s \leq L/2-1 \leq (L-2)-1 = \hat{n}-1$ (the first and second inequalities follow because $L-2 \leq s \leq L/2$), which contradicts to the condition $L-1-s \geq \hat{n}$.

(b) If $t(i_k^-) \leq L-1$, then $t(i_{k+s-L+1}^-) = t(i_k^-) + L-1-s \leq 2(L-1) - s \leq -2$ because $s \geq L/2$, which contradicts to the condition $L-1-s \geq \hat{n}$.

Next, we consider the case in which $y_j = 1$. We let $p = \arg\max\{t(k) : k \in \mathcal{P}(i_k^-) \cap \mathcal{P}(j)\} = i_{k+k_1}^- = j_{k_2}^-$. If there is a start-up between $i_{k+k_1}^-$ and i_k^- or(and) a start-up between $j_{k_2}^-$ and j, then the proof is similar to the discussion above and thus is omitted here. Therefore, we consider the case in which $y_n = 0$ for each $n \in (\mathcal{P}(i_k^-) \cup \mathcal{P}(j)) \setminus \mathcal{P}(j_{k_2}^-)$. If $k_1 \geq L-1$, then it is clear that inequality (EC.19) is valid following the similar argument above, because $\phi = 0$ in (EC.19); otherwise, we consider the case in which $k_1 \leq L-2$ as follows. We let the last start-up node (denoted as i_{k+s}^-, $s \geq 0$) before i_k^- in the following two possible cases.

1) $s \geq L-1$. It follows that $\phi = 0$ and $i_{k+s-L+1}^- \in \mathcal{P}(i_k^-)$. We further discuss the following three possible cases in terms of the first shut-down node (denoted as h, $h \notin \mathcal{P}(i_k^-)$) after i_k^-, i.e., $y_h = 0$. We observe that $y_n - \sum_{m=0}^{L-1} u_{n_m^-} = 1$ for each $n \in \mathcal{P}(h^-) \setminus \mathcal{P}(i_k^-)$.

  (1) $h \in \mathcal{P}(i_{k-\hat{n}}^-)$. In this case, we have $h = i_{k-\hat{s}}^-$ for some $\hat{s} \in [1, \hat{n}]_{\mathbb{Z}}$ and $x_{i_k^-} - x_j \leq \min\{kV, \overline{V} + \min\{k - 1, s, \hat{s}-1\}V - \underline{C} + (\underline{C}+V-\overline{V})\} \leq \overline{V} + (\hat{s}-1)V - \underline{C} + (\underline{C}+V-\overline{V}) = \overline{V}y_{i_k^-} + V\sum_{n=1}^{\hat{s}-1}(y_{i_{k-n}^-} - \sum_{m=0}^{\min\{L-1,n+w\}} u_{i_{k-n+m}^-}) - \underline{C}y_j + (\underline{C}+V-\overline{V})(y_j - \sum_{m=0}^{L-1} u_{j_m^-})$, which is clearly less than the RHS of (16).

  (2) $h \in \mathcal{P}(i) \setminus \mathcal{P}(i_{k-\hat{n}}^-)$. In this case, we have $h = i_{k-\hat{s}}^-$ for some $\hat{s} \in [\hat{n}+1, k]_{\mathbb{Z}}$ and $x_{i_k^-} - x_j \leq \min\{kV, \overline{V} + \min\{k-1, s, \hat{s}-1\}V - \underline{C} + (\underline{C}+V-\overline{V})\} \leq \overline{V} + (\hat{s}-1)V - \underline{C} + (\underline{C}+V-\overline{V}) = \overline{V} + (\hat{t}-1)V + (\hat{s}-\hat{t})V - \underline{C} + (\underline{C}+V-\overline{V}) \leq V\sum_{n \in S_0}(y_{i_{k-n}^-} - \sum_{m=0}^{\min\{L-1,n+w\}} u_{i_{k-n+m}^-}) + V\sum_{n \in (S \cap [\hat{n}+1, k-\hat{s}+1]_{\mathbb{Z}}) \cup \{\hat{n}\}}(g_n - n)(y_{i_{k-n}^-} - \sum_{m=0}^{L-1} u_{i_{k-n+m}^-}) - \underline{C}y_j + (\underline{C}+V-\overline{V})(y_j - \sum_{m=0}^{L-1} u_{j_m^-})$, which is clearly less than the RHS of (16).

  (3) $h \in \mathcal{V}(i) \setminus \mathcal{P}(i)$. Inequality (16) converts to $x_{i_k^-} - x_j \leq kV$, which is clearly valid due to constraints (1f) and (1g).

2) $s \in [k_1, L-2]_{\mathbb{Z}}$. The discussion is similar to subcase 2) in the case in which $y_j = 0$ and thus is omitted here.

(Facet-defining) The facet-defining proof is similar to that for Proposition 4 in Online Supplement E.1 and thus is omitted here. $\qquad\square$