Robust Unit Commitment: Multi-State Uncertainty Set Formulation and Efficient Solution Method

First A. Author, *Fellow, IEEE*, Second B. Author, and Third C. Author Jr., *Member, IEEE*

APPENDIX A
COMPLETE MODEL OF TSRUCM

Objective of two-stage adaptive robust unit commitment problem:

$$\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p}\in\mathcal{X}} \sum_t \sum_g \left(C_g^{\mathrm{SU}} y_{g,t} + C_g^0 x_{g,t} + C_g p_{g,t}\right) + \max_{\boldsymbol{w}\in U} \min_{\Delta\boldsymbol{p},\Delta\boldsymbol{w},\Delta\boldsymbol{d}\in\mathcal{Y}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p},\boldsymbol{w})} \sum_t \left(\sum_r C^{\mathrm{WS}} \Delta w_{r,t} \sum_j C^{\mathrm{LS}} \Delta d_{j,t}\right) \quad \text{(A.1a)}$$

First-stage feasibility region:

$$\mathcal{X} = \left\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p}\right| \quad \text{(A.1b)}$$

$$\sum_g p_{g,t} + \sum_r p_{r,t} = \sum_j d_{j,t}, \forall t \quad \text{(A.1c)}$$

$$x_{g,t} - x_{g,t-1} = y_{g,t} - z_{g,t}, \forall g, \forall t \quad \text{(A.1d)}$$

$$\sum_{\tau=t-TU_g+1}^{t} y_{g,\tau} \le x_{g,t}, \forall g, \forall t\,. \quad \text{(A.1e)}$$

$$\sum_{\tau=t-TD_g+1}^{t} z_{g,\tau} \le 1 - x_{g,t}, \forall g, \forall t \quad \text{(A.1f)}$$

$$P_g^{\min} x_{g,t} \le p_{g,t} \le P_g^{\max} x_{g,t}, \forall g, \forall t \quad \text{(A.1g)}$$

$$p_{g,t} - p_{g,t-1} \le RU_g^{1\mathrm{h}} x_{g,t-1} + SU_g^{1\mathrm{h}} \left(1 - x_{g,t-1}\right), \forall g, \forall t \quad \text{(A.1h)}$$

$$p_{g,t-1} - p_{g,t} \le RD_g^{1\mathrm{h}} x_{g,t} + SD_g^{1\mathrm{h}} \left(1 - x_{g,t}\right), \forall g, \forall t \quad \text{(A.1i)}$$

$$0 \le p_{r,t} \le \bar{\omega}_{r,t}, \forall r, \forall t \quad \text{(A.1j)}$$

$$-F_l^{\max} \le \sum_m SF_{l,m} \left(\sum_{g\in \mathrm{G}(m)} p_{g,t} + \sum_{r\in R(m)} p_{r,t} - \sum_{j\in \mathrm{D}(m)} d_{j,t}\right) \le F_l^{\max}, \forall l, \forall t\Big\} \quad \text{(A.1k)}$$

Objective of second-stage recourse problem for a given first-stage solution $\left(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p}\right)$ and WPO scenario $\boldsymbol{w}$:

$$Q\left(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p},\boldsymbol{w}\right) = \min_{\Delta\boldsymbol{p},\Delta\boldsymbol{w},\Delta\boldsymbol{d}\in\mathcal{Y}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p},\boldsymbol{w})} \sum_t \sum_r C^{\mathrm{WS}} \Delta w_{r,t} + \sum_t \sum_j C^{\mathrm{LS}} \Delta d_{j,t} \quad \text{(A.1l)}$$

Second-stage feasibility region:

$$\mathcal{Y}\left(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p},\boldsymbol{w}\right) = \left\{\Delta\boldsymbol{p},\Delta\boldsymbol{w},\Delta\boldsymbol{d}\right| \quad \text{(A.1m)}$$

$$\hat{p}_{g,t} = p_{g,t} + \Delta p_{g,t}, \forall g, \forall t \quad \text{(A.1n)}$$

$$\hat{p}_{r,t} = w_{r,t} + \Delta w_{r,t}, \forall r, \forall t \quad \text{(A.1o)}$$

$$\hat{d}_{j,t} = d_{j,t} + \Delta d_{j,t}, \forall j, \forall t \quad \text{(A.1p)}$$

$$\sum_g \hat{p}_{g,t} + \sum_r \hat{p}_{r,t} = \sum_j \hat{d}_{j,t}, \forall t \quad \text{(A.1q)}$$

$$P_g^{\min} x_{g,t} \le \hat{p}_{g,t} \le P_g^{\min} x_{g,t}, \forall g, \forall t \quad \text{(A.1r)}$$

$$\hat{p}_{g,t} - \hat{p}_{g,t-1} \le RU_g^{1\mathrm{h}} x_{g,t-1} + SU_g^{1\mathrm{h}} \left(1 - x_{g,t-1}\right), \forall g, \forall t \quad \text{(A.1s)}$$

$$\hat{p}_{g,t-1} - \hat{p}_{g,t} - \le RD_g^{1\mathrm{h}} x_{g,t} + SD_g^{1\mathrm{h}} \left(1 - x_{g,t}\right), \forall g, \forall t \quad \text{(A.1t)}$$

$$0 \le \Delta p_{g,t} \le R_g^{10\min} x_{g,t}, \forall g, \forall t \quad \text{(A.1u)}$$

$$0 \le \Delta w_{r,t} \le w_{r,t}, \forall r, \forall t \quad \text{(A.1v)}$$

$$0 \le \Delta d_{j,t} \le d_{j,t}, \forall j, \forall t \quad \text{(A.1w)}$$

$$-F_l^{\max} \le \sum_m SF_{l,m} \left(\sum_{g\in \mathrm{G}(m)} \hat{p}_{g,t} + \sum_{w\in \mathrm{W}(m)} \hat{p}_{r,t} - \sum_{j\in \mathrm{D}(m)} \hat{d}_{j,t}\right) \le F_l^{\max}, \forall l, \forall t \quad \text{(A.1x)}$$

where $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$, $\boldsymbol{p}$, $\Delta\boldsymbol{p}$, $\Delta\boldsymbol{w}$, $\Delta\boldsymbol{d}$, $x_{g,t}$, $y_{g,t}$, $z_{g,t}$, $p_{g,t}$, $p_{r,t}$, $\hat{p}_{g,t}$, $\hat{p}_{r,t}$, $\hat{d}_{j,t}$, $\Delta p_{g,t}$, $\Delta w_{r,t}$, $\Delta d_{j,t}$ are decision variables, C_g^{SU}, C_g^0, C_g, C^{WS}, C^{LS}, TU_g, TD_g, $P_g^{\max}$, $P_g^{\min}$, $RU_g^{1\mathrm{h}}$, $RD_g^{1\mathrm{h}}$, $SU_g^{1\mathrm{h}}$, $SD_g^{1\mathrm{h}}$, $R_g^{10\min}$, $SF_{l,m}$, $F_l^{\max}$, $d_{j,t}$ are constants, $\boldsymbol{w}$, $w_{r,t}$ are uncertain variables, subscript g denotes the g-th thermal unit, subscript r denotes the r-th wind farm, subscript j denotes the j-th load, subscript l denotes the l-th transmission line, m denotes the m-th bus, subscript t denotes the t-th time period. The first-stage decision variables are as follows. $\boldsymbol{x}$ denotes the vector of unit on/off status, $\boldsymbol{y}$ denotes the vector of unit start-up operation, $\boldsymbol{z}$ denotes the vector of unit shut-down operation, $\boldsymbol{p}$ denotes the vector of the dispatch of thermal unit and wind farm, $x_{g,t}$ denotes if the unit is on $(x_{g,t}=1)$ or off $(x_{g,t}=0)$, $y_{g,t}$ denotes if the unit is started up $(y_{g,t}=1)$ or not $(y_{g,t}=0)$, $z_{g,t}$ denotes if the unit is shut down $(z_{g,t}=1)$ or not $(z_{g,t}=0)$, $p_{g,t}$ denotes the dispatch of thermal unit, $p_{r,t}$ denotes the dispatch of wind farm. The second-stage decision variables are as follows. $\Delta\boldsymbol{p}$ denotes the vector of adjustment of thermal unit dispatch, $\Delta\boldsymbol{w}$ denotes the vector of wind curtailment, $\Delta\boldsymbol{d}$ denotes the vector of load shedding, $\Delta p_{g,t}$ denotes the adjustment of thermal unit dispatch, $\Delta w_{r,t}$ denotes the wind curtailment, $\Delta d_{j,t}$ denotes the load shedding, $\hat{p}_{g,t}$ denotes the redispatch of thermal unit, $\hat{p}_{r,t}$ denotes the redispatch of wind

farm, $\hat{d}_{r,t}$ denotes the left load demand after load shedding. The constants are as follows. C_g^{SU} denotes the start-up cost, C_g^0 denotes no-load cost, C_g denotes generation cost, C^{WC} denotes wind curtailment penalty cost, C^{LS} denotes load shedding penalty cost, TU_g / TD_g denotes the minimum on/off time of thermal unit, $P_g^{\max}$ / $P_g^{\min}$ denotes maximum/minimum capacity of thermal unit, $RU_g^{1\mathrm{h}}$ / $RD_g^{1\mathrm{h}}$ denotes hourly ramp up/down rate of thermal unit, $SU_g^{1\mathrm{h}}$ / $SD_g^{1\mathrm{h}}$ denotes hourly start-up/shut-down rate of thermal unit, $R_g^{10\,\mathrm{min}}$ denotes the corrective capacity of thermal unit, i.e., 10-minute spinning reserve of thermal unit, $SF_{l,m}$ denotes the shift factor, $F_l^{\max}$ denotes the maximum capacity of transmission line, $d_{j,t}$ denotes the load demand.

The objective of two-stage adaptive robust unit commitment problem (A.1a) is to minimize the day-ahead start-up cost, no-load cost, generation cost under nominal condition and the wind curtailment penalty cost, load shedding penalty cost under the uncertain WPO scenarios. Constraints (A.1b) - (A.1k) denotes the first-stage feasibility region under nominal condition, including system load balance (A.1c), start-up/shut-down operations of thermal units (A.1d), minimum on/off time limits of thermal units (A.1e) - (A.1f), maximum /minimum capacity limits of thermal units (A.1g), hourly ramp-up and ramp-down rate limits of thermal units (A.1h) - (A.1i), available WPO limits of wind farms (A.1j), DC power flow limits of transmission lines (A.1k). The objective of second-stage recourse problem is to minimize the wind curtailment and load shedding penalty cost for a given first-stage solution $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p})$ and WPO scenario $\boldsymbol{w}$ (A.1l). Condition (A.1m)-(A.1x) denotes the second feasibility region under the given first-stage solution $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{p})$ and WPO scenario $\boldsymbol{w}$, including redispatch of thermal unit and wind farm (A.1n) - (A.1o), load shedding of load (A.1p), system load balance (A.1q), maximum/minimum capacity limits of thermal units (A.1r), hourly ramp-up and ramp-down rate limits of thermal units (A.1s) - (A.1t), corrective capacity limits of thermal units (A.1u), wind curtailment limits of wind farm (A.1v), load shedding limits of load (A.1w), DC power flow limits of transmission lines (A.1x).

APPENDIX B
PROOF OF PROPOSITION 1

In light of the extensiveness of the proof of Proposition 1, it has been partitioned into several smaller sub-propositions, namely Proposition 4 - 9, to enhance clarity and readability. In this section, we will first introduce the prerequisite knowledge, followed by the definition of several auxiliary uncertainty sets. Next, we will subsequently introduce and prove Proposition 4 - 9. At last, we will demonstrate Proposition 1 through the conclusions drawn from Proposition 4 - 9.

A. Prerequisite Knowledge

1) *Definition of extreme point:* Consider a polyhedron P defined in terms of linear equality and inequality constraints

$$\boldsymbol{a}_i^T\boldsymbol{x} \ge b_i, i \in M_1 \tag{B.1a}$$

$$\boldsymbol{a}_i^T\boldsymbol{x} \le b_i, i \in M_2 \tag{B.1b}$$

$$\boldsymbol{a}_i^T\boldsymbol{x} = b_i, i \in M_3 \tag{B.1c}$$

where M_1, M_2 and M_3 are finite index sets, each $\boldsymbol{a}_i$ is a vector in $\mathbb{R}^n$, and each b_i is a scalar. A vector $\boldsymbol{x} \in P$ is an extreme point of P if we can't find two vectors $\boldsymbol{y},\boldsymbol{z} \in P$, both different from $\boldsymbol{x}$, and a scalar $\lambda \in [0,1]$, such that $\boldsymbol{x} = \lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{z}$ [1].

2) *Definition of active constraint:* Consider a polyhedron P defined by linear equality and inequality constraints (B.1a) - (B.1c). If a vector $\boldsymbol{x}^* \in \mathbb{R}^n$ satisfies $\boldsymbol{a}_i^T\boldsymbol{x}^* = b_i$ for some i in M_1, M_2 or M_3, then the corresponding constraint is active or binding at $\boldsymbol{x}^*$.

3) *Definition of linearly independent constraints:* Consider a polyhedron P defined by linear equality and inequality constraints (B.1a) - (B.1c). Let $I = \{i \mid \boldsymbol{a}_i^T\boldsymbol{x}^* = b_i\}$ be the set of indices of constraints that are active at a given vector $\boldsymbol{x}^*$. Then there exist n linearly independent constraints active at $\boldsymbol{x}^*$ if there exist n vectors in the set $\{\boldsymbol{a}_i \mid i \in I\}$, which are linearly independent [1].

4) *Definition of basic feasible solution:* Consider a polyhedron P defined by linear equality and inequality constraints (B.1a) - (B.1c). The vector $\boldsymbol{x}^* \in \mathbb{R}^n$ is a basic solution if all equality constraints are active. Meanwhile, out of the constraints that are active at $\boldsymbol{x}^*$, there are n of them are linearly independent. If $\boldsymbol{x}^*$ is a basic solution that satisfies all of the constraints, it is a basic feasible solution [1].

5) *Definition of projection of polyhedron:* Let $\boldsymbol{x} = (x_1,\cdots,x_n)$ be a vector in $\mathbb{R}^n$ and $k \le n$. The projection mapping $\pi_k : \mathbb{R}^n \mapsto \mathbb{R}^k$ projects $\boldsymbol{x}$ onto its first k coordinates:

$$\pi_k(\boldsymbol{x}) = \pi_k(x_1,\cdots,x_n) = (x_1,\cdots,x_k) \tag{B.2}$$

Consider a polyhedron P defined by linear equality and inequality constraints (B.1a) - (B.1c). Then, the projection $\Pi_k(P)$ of the polyhedron $P \subset \mathbb{R}^n$ is defined as [1]

$$\Pi_k(P) = \{\pi_k(\boldsymbol{x}) \mid \boldsymbol{x} \in P\} \tag{B.3}$$

6) *Equivalence of Basic Feasible Solution and Extreme Point:* Consider a polyhedron P defined by linear equality and inequality constraints (B.1a) - (B.1c). Then, if the vector $\boldsymbol{x}^* \in P$ is a basic feasible solution, it is also an extreme point and a vertex of P [1].

B. Auxiliary Uncertainty Sets

Throughout the proof, 5 auxiliary uncertainty sets U_2', U_0, U_0', U_0'', U_0''' are involved, playing a significant role in the demonstration.

The mathematical formulation of U_2' can be expressed as

$$U_2' = \left\{ \boldsymbol{w} \in \mathbb{R}^{N_w \times N_t} : \right. \tag{B.4a}$$

$$w_{r,t} = \sum\nolimits_{i=0}^{N_s} \omega_{r,t}^i y_{r,t}^i, \forall r, \forall t \tag{B.4b}$$

$$\sum\nolimits_{i=0}^{N_s} y_{r,t}^i = 1, \forall r, \forall t \tag{B.4c}$$

$$\sum\nolimits_{t=1}^{N_t} \sum\nolimits_{i=0}^{N_s} \tau_{r,t}^i y_{r,t}^i \le \Gamma_r, \forall r \tag{B.4d}$$

$$\left. y_{r,t}^i, y_{r,t}^0 \in \{0,1\}, \forall r, \forall t, \forall i \right\} \tag{B.4e}$$

where the same symbols in U_2' as in U_1 and U_2 have the same meaning, $y_{r,t}^0$ denotes if the uncertain WPO takes the 0-th state values ($y_{r,t}^0 = 1$), $\omega_{r,t}^0$ denotes the 0-th state values, which takes value of the forecasted value $\bar{\omega}_{r,t}$, $\tau_{r,t}^0$ denotes the relative deviation of the 0-th state values, which takes the value of 0 according to (4). *a*

The mathematical formulation of U_0 can be expressed as

$$U_0 = \{ \boldsymbol{w} \in \mathbb{R}^{N_w \times N_t} : \tag{B.5a}$$

$$w_{r,t} = \bar{\omega}_{r,t} + \delta_{r,t}^+ y_{r,t}^+ + \delta_{r,t}^- y_{r,t}^-, \forall r, \forall t \tag{B.5b}$$

$$y_{r,t}^+ + y_{r,t}^- \le 1, \forall r, \forall t \tag{B.5c}$$

$$\sum\nolimits_{t=1}^{N_t} \left(y_{r,t}^+ + y_{r,t}^- \right) = \Gamma_r, \forall r \tag{B.5d}$$

$$\left. y_{r,t}^+, y_{r,t}^- \in \{0,1\}, \forall r, \forall t \right\} \tag{B.5e}$$

where the same symbols in U_0 as in U_1 have the same meaning.

The mathematical formulation of U_0' can be expressed as

$$U_0' = \left\{ \boldsymbol{w} \in \mathbb{R}^{N_w \times N_t} : \right. \tag{B.6a}$$

$$w_{r,t} = \bar{\omega}_{r,t} + \delta_{r,t}^+ z_{r,t}^+ + \delta_{r,t}^- z_{r,t}^-, \forall r, \forall t \tag{B.6b}$$

$$z_{r,t}^+ + z_{r,t}^- \le 1, \forall r, \forall t \tag{B.6c}$$

$$\sum\nolimits_{t=1}^{N_t} \left(z_{r,t}^+ + z_{r,t}^- \right) \le \Gamma_r, \forall r \tag{B.6d}$$

$$\left. 0 \le z_{r,t}^+, z_{r,t}^- \le 1, \forall r, \forall t \right\} \tag{B.6e}$$

where the same symbols in U_0' as in U_1 have the same meaning, $z_{r,t}^+$ and $z_{r,t}^-$ are variables, $z_{r,t}^+$ and $z_{r,t}^-$ denotes the positive and negative relative deviation of $w_{r,t}$ from the forecasted value $\bar{\omega}_{r,t}$.

The mathematical formulation of U_0'' can be expressed as

$$U_0'' = \left\{ \left(\boldsymbol{w}, \boldsymbol{z}^+, \boldsymbol{z}^- \right) \in \mathbb{R}^{3 \times N_w \times N_t} : \right. \tag{B.7a}$$

$$w_{r,t} = \bar{\omega}_{r,t} + \delta_{r,t}^+ z_{r,t}^+ + \delta_{r,t}^- z_{r,t}^-, \forall r, \forall t \tag{B.7b}$$

$$z_{r,t}^+ + z_{r,t}^- \le 1, \forall r, \forall t \tag{B.7c}$$

$$\sum\nolimits_{t=1}^{N_t} \left(z_{r,t}^+ + z_{r,t}^- \right) \le \Gamma_r, \forall r \tag{B.7d}$$

$$\left. 0 \le z_{r,t}^+, z_{r,t}^- \le 1, \forall r, \forall t \right\} \tag{B.7e}$$

where the same symbols in U_0'' as in U_0' have the same meaning.

The mathematical formulation of U_0''' can be expressed as

$$U_0''' = \left\{ \left(\boldsymbol{w}, \boldsymbol{z}^+, \boldsymbol{z}^- \right) \in \mathbb{R}^{3 \times N_w \times N_t} : \right. \tag{B.8a}$$

$$w_{r,t} = \bar{\omega}_{r,t} + \delta_{r,t}^+ z_{r,t}^+ + \delta_{r,t}^- z_{r,t}^-, \forall r, \forall t \tag{B.8b}$$

$$z_{r,t}^+ + z_{r,t}^- \le 1, \forall r, \forall t \tag{B.8c}$$

$$\sum\nolimits_{t=1}^{N_t} \left(z_{r,t}^+ + z_{r,t}^- \right) \le \Gamma_r, \forall r \tag{B.8d}$$

$$\left. z_{r,t}^+, z_{r,t}^- = \{0,1\}, \forall r, \forall t \right\} \tag{B.8e}$$

where the same symbols in U_0''' as in U_0'' have the same meaning.

C. Proposition 4 and Its Proof

Proposition 4: $U_2 \subseteq U_2'$.

Proof: To prove that U_2 is a subset of U_2', we need to prove that for any element (WPO scenario) in U_2, it also belongs to U_2'. Here are the steps for the proof:

1) Let's assume $\tilde{\boldsymbol{w}} \in U_2$ to be arbitrary. According to the definition of U_2 , this implies that there exist binary variables $\tilde{\boldsymbol{y}} = (\tilde{y}_{r,t}^i)$, such that $\tilde{\boldsymbol{w}}$ satisfies the constraints (3a) - (3f).
2) We need to prove that there always exist binary variables $\boldsymbol{y} = (y_{r,t}^i, y_{r,t}^0)$ such that $\tilde{\boldsymbol{w}}$ satisfies the constraints defining U_2' , namely (B.4a) - (B.4e). Obviously, when $\boldsymbol{y}$ takes value at $(\tilde{y}_{r,t}^i, 0)$, $\tilde{\boldsymbol{w}}$ satisfies (B.4a) - (B.4e).
3) It is evident that any element in U_2 also belongs to U_2'. Thus $U_2 \subseteq U_2'$.

D. Proposition 5 and Its Proof

Proposition 5: $U_0 \subseteq U_1 \subseteq U_2' \subseteq U_0'$

Proof: First, to prove that U_0 is a subset of U_1, we need to prove that for any element (WPO scenario) in U_0, it also belongs to U_1. Here are the steps for the proof:

1) Let's assume $\tilde{\boldsymbol{w}} \in U_0$ to be arbitrary. According to the definition of U_0 , this implies that there exist binary variables $\tilde{\boldsymbol{y}} = (\tilde{y}_{r,t}^+, \tilde{y}_{r,t}^-)$, such that $\tilde{\boldsymbol{w}}$ satisfies the constraints (B.5a) - (B.5e).
2) We need to prove that there always exist binary variables $\boldsymbol{y} = (y_{r,t}^+, y_{r,t}^-)$ such that $\tilde{\boldsymbol{w}}$ satisfies the constraints defining U_1 , namely (2a) - (2e). Obviously, when $\boldsymbol{y}$ takes value at $(\tilde{y}_{r,t}^+, \tilde{y}_{r,t}^-)$, $\tilde{\boldsymbol{w}}$ satisfies (2a) - (2e).
3) It is evident that any element in U_0 also belongs to U_1. Thus $U_0 \subseteq U_1$.

Second, to prove that U_1 is a subset of U_2', we need to prove that for any element (WPO scenario) in U_1, it also belongs to U_2'. Here are the steps for the proof:

1) Let's assume $\tilde{\boldsymbol{w}} \in U_1$ to be arbitrary. According to the definition of U_1 , this implies that there exist binary

variables $\tilde{\boldsymbol{y}}=(\tilde{y}^{+}_{r,t},\tilde{y}^{-}_{r,t})$, such that $\tilde{\boldsymbol{w}}$ satisfies the constraints (2a) - (2e).

2) We need to prove that there always exist binary variables $\boldsymbol{y}=(y^{i}_{r,t},y^{0}_{r,t})$ such that $\tilde{\boldsymbol{w}}$ satisfies the constraints defining U_1, namely (B.4a) - (B.4e). Obviously, under the given 2 conditions in Proposition 1, when the elements of $\boldsymbol{y}$ are taken as follows, $\tilde{\boldsymbol{w}}$ satisfies (B.4a) - (B.4e). If $\tilde{y}^{+}_{r,t}=1$ and $\tilde{y}^{-}_{r,t}=0$, then $y^{N_s}_{r,t}=1$. If $\tilde{y}^{+}_{r,t}=0$ and $\tilde{y}^{-}_{r,t}=1$, then $y^{1}_{r,t}=1$. If $\tilde{y}^{+}_{r,t}=0$ and $\tilde{y}^{-}_{r,t}=0$, then $y^{0}_{r,t}=1$. The other elements $y^{i}_{r,t}=0, i=2,\cdots,N_s-1$.

3) It is evident that any element in U_1 also belongs to U_2'. Thus $U_1\subseteq U_2'$.

Third, to prove that U_2' is a subset of U_0', we need to prove that for any element (WPO scenario) in U_2', it also belongs to U_0'. Here are the steps for the proof:

1) Let's assume $\tilde{\boldsymbol{w}}\in U_2'$ to be arbitrary. According to the definition of U_2', this implies that there exist binary variables $\tilde{\boldsymbol{y}}=(\tilde{y}^{i}_{r,t},\tilde{y}^{0}_{r,t})$, such that $\tilde{\boldsymbol{w}}$ satisfies the constraints (2a) - (2e).

2) We need to prove that there always exist variables $\boldsymbol{z}=(z^{+}_{r,t},z^{-}_{r,t})$ such that $\tilde{\boldsymbol{w}}$ satisfies the constraints defining U_0', namely (B.6a) - (B.6e). Obviously, when the elements of $\boldsymbol{z}$ are taken as follows, $\tilde{\boldsymbol{w}}$ satisfies (B.6a) - (B.6e).: If $\tilde{y}^{i}_{r,t}=1$ and $\omega^{i}_{r,t}>\bar{\omega}_{r,t}$, then $z^{+}_{r,t}=(\omega^{i}_{r,t}-\bar{\omega}_{r,t})/\delta^{+}_{r,t}$ and $z^{-}_{r,t}=0$. If $\tilde{y}^{i}_{r,t}=1$ and $\omega^{i}_{r,t}<\bar{\omega}_{r,t}$, then $z^{+}_{r,t}=0$ and $z^{-}_{r,t}=(\bar{\omega}_{r,t}-\omega^{i}_{r,t})/\delta^{-}_{r,t}$. If $\tilde{y}^{0}_{r,t}=1$, then $z^{+}_{r,t}=z^{-}_{r,t}==0$.

3) It is evident that any element in U_2' also belongs to U_0'. Thus $U_2'\subseteq U_0'$.

Since $U_0\subseteq U_1$, $U_1\subseteq U_2'$ and $U_2'\subseteq U_0'$ hold true, and thus $U_0\subseteq U_1\subseteq U_2'\subseteq U_0'$ holds true.

E. Proposition 6 and Its Proof

Proposition 6: $U_0\subseteq \mathrm{vert}(U_0')$

Proof: To prove that U_0 is a subset of the set of extreme points of U_0', we need to prove that all the elements in U_0 are the extreme points of U_0'. Here are the steps for the proof.

a) Proof of Contradiction

Assume $\tilde{\boldsymbol{w}}\in U_0$ to be an arbitrary. According to the definition of U_0, this implies that there exist binary variables $\tilde{\boldsymbol{y}}=(\tilde{y}^{+}_{r,t},\tilde{y}^{-}_{r,t})$, such that $\tilde{\boldsymbol{w}}$ satisfies the constraints (B.5a) - (B.5e). Assume that $\tilde{\boldsymbol{w}}$ is not an extreme point of U_0', which implies that there exist two elements $\boldsymbol{w}_1,\boldsymbol{w}_2\in U_0'$ and a scalar $\lambda\in[0,1]$, such that $\boldsymbol{w}_1\neq\boldsymbol{w}_2\neq\tilde{\boldsymbol{w}}$ and $\tilde{\boldsymbol{w}}=\lambda\boldsymbol{w}_1+(1-\lambda)\boldsymbol{w}_2$. According to the definition of U_0', this implies that there exist variables $\boldsymbol{z}_1=(z^{+}_{r,t,1},z^{-}_{r,t,1})$ and $\boldsymbol{z}_2=(z^{+}_{r,t,2},z^{-}_{r,t,2})$, such that $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ satisfies the constraints (B.6a) - (B.6e). According to the equation $\tilde{\boldsymbol{w}}=\lambda\boldsymbol{w}_1+(1-\lambda)\boldsymbol{w}_2$, each element of the vector on both sides is equal, which means the following $N_w\times N_t$ equations hold.

$$\begin{aligned}&\bar{\omega}_{r,t}+\delta^{+}_{r,t}\tilde{y}^{+}_{r,t}+\delta^{-}_{r,t}\tilde{y}^{-}_{r,t}=\lambda\left(\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,1}+\right.\\&\left.\delta^{-}_{r,t}z^{-}_{r,t,1}\right)+(1-\lambda)\left(\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,2}+\delta^{-}_{r,t}z^{-}_{r,t,2}\right)\\&,r=1,\cdots,N_w,t=1,\cdots,N_t\end{aligned}\tag{B.9}$$

b) Classification Discussion

When $\lambda=1$: All the equations hold true for $\boldsymbol{z}_1=(z^{+}_{r,t,1},z^{-}_{r,t,1})=(\tilde{y}^{+}_{r,t},\tilde{y}^{-}_{r,t})$. However, in this case, $\boldsymbol{w}_1=\tilde{\boldsymbol{w}}$, which contradicts the assumption that $\boldsymbol{w}_1$ is different from $\tilde{\boldsymbol{w}}$.

When $\lambda=0$: All the equations hold true for $\boldsymbol{z}_2=(z^{+}_{r,t,2},z^{-}_{r,t,2})=(\tilde{y}^{+}_{r,t},\tilde{y}^{-}_{r,t})$. However, in this case, $\boldsymbol{w}_2=\tilde{\boldsymbol{w}}$, which contradicts the assumption that $\boldsymbol{w}_2$ is different from $\tilde{\boldsymbol{w}}$.

When $\lambda\in(0,1)$: We denote the set of indices for which the elements of vector $\tilde{\boldsymbol{y}}$ satisfy $\tilde{y}^{+}_{r,t}=1$ and $\tilde{y}^{-}_{r,t}=0$ as I^{+}, i.e.,

$$I^{+}=\left\{r,t\mid\tilde{y}^{+}_{r,t}=1,\tilde{y}^{-}_{r,t}=0\right\}\tag{B.10a}$$

Then, the equations where the left-hand side equals $\bar{\omega}_{r,t}+\delta^{-}_{r,t}$, $(r,t)\in I^{+}$ hold true only for $z^{+}_{r,t,1}=z^{+}_{r,t,2}=1,(r,t)\in I^{+}$ and $z^{-}_{r,t,1}=z^{-}_{r,t,2}=0,(r,t)\in I^{+}$. This is due to the fact that $\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,1}+\delta^{-}_{r,t}z^{-}_{r,t,1}\le\bar{\omega}_{r,t}+\delta^{+}_{r,t}$ and equality is achieved exclusively when $z^{+}_{r,t,1}=1$ and $z^{-}_{r,t,1}=0$, $\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,2}+\delta^{-}_{r,t}z^{-}_{r,t,2}\le\bar{\omega}_{r,t}+\delta^{+}_{r,t}$ and equality is achieved exclusively when $z^{+}_{r,t,2}=1$ and $z^{-}_{r,t,2}=0$.Thus, $\lambda(\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,1}+\delta^{-}_{r,t}z^{-}_{r,t,1})+(1-\lambda)(\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,2}+\delta^{-}_{r,t}z^{-}_{r,t,2})\le\bar{\omega}_{r,t}+\delta^{+}_{r,t}$ and equality is achieved exclusively when $z^{+}_{r,t,1}=z^{+}_{r,t,2}=1$ and $z^{-}_{r,t,1}=z^{-}_{r,t,2}=0$.

We denote the set of indices for which the elements of vector $\tilde{\boldsymbol{y}}$ satisfy $\tilde{y}^{+}_{r,t}=0$ and $\tilde{y}^{-}_{r,t}=1$ as I^{-}, i.e.,

$$I^{-}=\left\{r,t\mid\tilde{y}^{+}_{r,t}=0,\tilde{y}^{-}_{r,t}=1\right\}\tag{B.10b}$$

Then, the equations where the left side equals $\bar{\omega}_{r,t}+\delta^{-}_{r,t}$, $(r,t)\in I^{-}$ hold true only for $z^{+}_{r,t,1}=z^{+}_{r,t,2}=0,(r,t)\in I^{-}$ and $z^{-}_{r,t,1}=z^{-}_{r,t,2}=1,(r,t)\in I^{-}$. This is due to the fact that $\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,1}+\delta^{-}_{r,t}z^{-}_{r,t,1}\ge\bar{\omega}_{r,t}+\delta^{-}_{r,t}$ and equality is achieved exclusively when $z^{+}_{r,t,1}=0$ and $z^{-}_{r,t,1}=1$, $\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,2}+\delta^{-}_{r,t}z^{-}_{r,t,2}\ge\bar{\omega}_{r,t}+\delta^{-}_{r,t}$ and equality is achieved exclusively when $z^{+}_{r,t,2}=0$ and $z^{-}_{r,t,2}=1$.Thus, $\lambda(\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,1}+\delta^{-}_{r,t}z^{-}_{r,t,1})+(1-\lambda)(\bar{\omega}_{r,t}+\delta^{+}_{r,t}z^{+}_{r,t,2}+\delta^{-}_{r,t}z^{-}_{r,t,2})\ge\bar{\omega}_{r,t}+\delta^{-}_{r,t}$ and equality is achieved exclusively when $z^{+}_{r,t,1}=z^{+}_{r,t,2}=0$ and $z^{-}_{r,t,1}=z^{-}_{r,t,2}=1$.

We denote the set of indices for which the elements of vector $\tilde{\boldsymbol{y}}$ satisfy $\tilde{y}^{+}_{r,t}=0$ and $\tilde{y}^{-}_{r,t}=0$ as I^{0}, i.e.,

$$I^0 = \left\{ r,t \mid \tilde{y}^+_{r,t} = 0, \tilde{y}^-_{r,t} = 0 \right\} \tag{B.10c}$$

The equations where the left-hand side equals $\bar{\omega}_{r,t}, (r,t) \in I^0$ hold true for many values of $(z^+_{r,t,1}, z^-_{r,t,1}), (r,t) \in I^0$ and $(z^+_{r,t,2}, z^-_{r,t,2}), (r,t) \in I^0$. However, to ensure that $\boldsymbol{w}_1, \boldsymbol{w}_2 \in U_0'$, $(z^+_{r,t,1}, z^-_{r,t,1}), (r,t) \in I^0$ and $(z^+_{r,t,2}, z^-_{r,t,2}), (r,t) \in I^0$ can only take the value of $(0,0)$. This is due the fact that to ensure all those equations $(r,t) \in I^+ \cup I^-$ hold true, $\sum_{t=1}^{N_t} \left(z^+_{r,t,1} + z^-_{r,t,1} \right) = \sum_{t=1}^{N_t} \left(z^+_{r,t,2} + z^-_{r,t,2} \right) = \sum_{t=1}^{N_t} \left(\tilde{y}^+_{r,t} + \tilde{y}^-_{r,t} \right) = \Gamma_r, (r,t) \in I^+ \cup I^-$ according to the constraints (B.5d). Thus, $\sum_{t=1}^{N_t} \left(z^+_{r,t,1} + z^-_{r,t,1} \right)$ and $\sum_{t=1}^{N_t} \left(z^+_{r,t,2} + z^-_{r,t,2} \right)$ have already reach their maximum value of Γ_r according to the constraints (B.6d). Thus, to ensure that $\boldsymbol{w}_1, \boldsymbol{w}_2 \in U_0'$, the rest element $(z^+_{r,t,1}, z^-_{r,t,1}), (r,t) \in I^0$ and $(z^+_{r,t,2}, z^-_{r,t,2}), (r,t) \in I^0$ can only take the value of $(0,0)$.

In this case, $\boldsymbol{w}_1 = \boldsymbol{w}_2 = \tilde{\boldsymbol{w}}$, which contradicts the assumption that $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ is different from $\tilde{\boldsymbol{w}}$.

c) Conclusion

It is evident that there is no element $\boldsymbol{w}_1, \boldsymbol{w}_2 \in U_0'$ and scalar $\lambda \in [0,1]$, such that $\boldsymbol{w}_1 \neq \boldsymbol{w}_2 \neq \tilde{\boldsymbol{w}}$ and $\tilde{\boldsymbol{w}} = \lambda \boldsymbol{w}_1 + (1-\lambda)\boldsymbol{w}_2$. Thus, all the element in U_0 are extreme points of U_0'. And thus, $U_0 \subseteq \text{vert}(U_0')$ holds true.

F. Proposition 7 and Its Proof

Proposition 7: $\text{vert}\left(U_0'\right) \subseteq \Pi_{N_w \times N_t} \left(\text{vert}\left(U_0''\right) \right)$

Proof: It is evident that U_0' is the projection $\Pi_{N_w \times N_t}(U_0'')$ of U_0'', i.e., $U_0' = \Pi_{N_w \times N_t}(U_0'')$, and thus the projection of the set of extreme points of U_0'' belongs to U_0', i.e., $\Pi_{N_w \times N_t}(\text{vert}(U_0'')) \subseteq U_0'$. Let $\{(\boldsymbol{w}, \boldsymbol{z}^+, \boldsymbol{z}^-)^{(1)}, \cdots, (\boldsymbol{w}, \boldsymbol{z}^+, \boldsymbol{z}^-)^{(N_k)}\}$ be the set of the extreme points of U_0''. Then, the projection of the set of the extreme points of U_0'' is $\{\boldsymbol{w}^{(1)}, \cdots, \boldsymbol{w}^{(N_k)}\}$, i.e., $\Pi_{N_w \times N_t}(\text{vert}(U_0'')) = \{\boldsymbol{w}^{(1)}, \cdots, \boldsymbol{w}^{(N_k)}\} \subseteq U_0'$.

To prove that the set of extreme points of U_0' is a subset of the projection of the set of the extreme points of U_0'', i.e., $\text{vert}(U_0') \subseteq \{\boldsymbol{w}^{(1)}, \cdots, \boldsymbol{w}^{(N_k)}\}$, we can equivalently prove that any element in U_0' can be represented as a convex combination of $\{\boldsymbol{w}^{(1)}, \cdots, \boldsymbol{w}^{(N_k)}\}$. Here are the steps for the proof:

1) Assume $\tilde{\boldsymbol{w}} \in U_0'$ to be arbitrary. We need to prove that there always exist non-negative scalars $\lambda_k, k = 1, \cdots, N_k$ such that $\tilde{\boldsymbol{w}} = \sum_{k=1}^{N_k} \lambda_k \boldsymbol{w}^{(N_k)}$ and $\sum_{k=1}^{N_k} \lambda_k = 1$.
2) According to the definition of projection, there exist $(\tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$ such that $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-) \in U_0''$. According to the Kerin-Milman Theorem, $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$ can be represented as a convex combination of the extreme points of U_0'', i.e.,

$$(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-) = \sum_{k=1}^{N_k} \tilde{\lambda}_k (\boldsymbol{w}, \boldsymbol{z}^+, \boldsymbol{z}^-)^{(N_k)} \tag{B.11a}$$

$$\sum_{k=1}^{N_k} \tilde{\lambda}_k = 1 \tag{B.11b}$$

$$\lambda_k \geq 0, k = 1, \cdots, N_k \tag{B.11c}$$

Obviously, when the scalar λ_k takes the value of $\tilde{\lambda}_k$, $\tilde{\boldsymbol{w}} = \sum_{k=1}^{N_k} \tilde{\lambda}_k \boldsymbol{w}^{(N_k)}$ and $\sum_{k=1}^{N_k} \tilde{\lambda}_k = 1$.

3) It is evident that any element in U_0' can be represented as a convex combination of $\{\boldsymbol{w}^{(1)}, \cdots, \boldsymbol{w}^{(N_k)}\}$. Thus, the extreme points of U_0' are contained in $\{\boldsymbol{w}^{(1)}, \cdots, \boldsymbol{w}^{(N_k)}\}$, namely, $\text{vert}(U_0') \subseteq \Pi_{N_w \times N_t} \left(\text{vert}\left(U_0''\right) \right)$

G. Proposition 8 and Its Proof

Proposition 8: $\text{vert}\left(U_0''\right) = U_0'''$

Proof: To prove that the set of extreme points of U_0'' is U_0''', we can equivalently prove that any element in U_0''' is a basic feasible solution of U_0''. Here are the steps for the proof:

We divide the entire feasible region of U_0'' into two subregions, and then categorically discuss the basic feasible solutions existing in each of the two subregions.

a) Subregion 1 U_0'''

Assume $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-) \in U_0'''$ to be arbitrary. According to the definition of U_0''', all the equality constraints (B.7b) are active at $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$. Meanwhile, there are $3 \times N_w \times N_t$ linearly independent constraints active at $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$, namely $N_w \times N_t$ constraints from all of the constraint (B.7b) and $2 \times N_w \times N_t$ constraints from half of the constraints (B.7d). Since $U_1 \subseteq U_0''$, $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$ is feasible for the constraints (B.7a) - (B.7e). Thus, $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$ is a basic solution of U_0''. It's evident that all the elements of U_1 are basic feasible solutions of U_0''.

b) Subregion 2 S

The second subregion S is the set difference between U_0'' and U_0''', i.e., $S = U_0'' - U_0'''$. Assume $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-) \in S$ to be arbitrary. Then, we can find that there is at least one element of $(\tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$ is a decimal between 0 and 1.

Assume that $\tilde{z}^+_{1,1}$ is the only one element of $(\tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$ that is a decimal between 0 and 1. Then, $\tilde{z}^-_{1,1}$ must be 0. Otherwise, $\tilde{z}^-_{1,1} = 1$ and $\tilde{z}^+_{1,1} + \tilde{z}^-_{1,1} > 1$, which contradict the definition of $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$. In this case, there are only $3 \times N_w \times N_t - 1$ linearly independent constraints active at $(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}^+, \tilde{\boldsymbol{z}}^-)$, namely $N_w \times N_t$ constraints from all of the constraint (B.7b) and $2 \times N_w \times N_t - 1$ constraints from the constraints (B.7d). This is

due to the fact that the constraint $0<\tilde{z}_{1,1}^{+}+\tilde{z}_{1,1}^{-}<1$ is not active. And even if the other constraints $\tilde{z}_{r,t}^{+}+\tilde{z}_{r,t}^{-}=1,(r,t)\neq(1,1)$ are active, they are not linearly independent from the active constraints $\tilde{z}_{r,t}^{+}+\tilde{z}_{r,t}^{-}=\{0,1\}$ $,(r,t)\neq(1,1)$. Moreover, the constraint $\sum_{t=1}^{N_t}\tilde{z}_{1,t}^{+}+\tilde{z}_{1,t}^{-}<\Gamma_1$ is not active since Γ_1 is an integer. And even if the other constraints $\sum_{t=1}^{N_t}\tilde{z}_{r,t}^{+}+\tilde{z}_{r,t}^{-}=\Gamma_r, r\neq 1$ are active, they are not linearly independent from the active constraints $\tilde{z}_{r,t}^{+},\tilde{z}_{r,t}^{-}=\{0,1\}$, $(r,t)\neq(1,1)$. Thus, if $\tilde{z}_{1,1}^{+}$ is the only one element of $(\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$ that is a decimal between 0 and 1, then $(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$ is not a basic solution of U_0'' .

Assume that $\tilde{z}_{1,1}^{+}$ and $\tilde{z}_{1,1}^{-}$ are the only two elements of $(\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$ that are decimals between 0 and 1. In this case, there are at most $3\times N_w\times N_t-1$ linearly independent constraints active at $(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$, namely $N_w\times N_t$ constraints from all of the constraint (B.7b) and $2\times N_w\times N_t-2$ constraints from the constraints (B.7d) and 1 constraints from the constraints (B.7c). Only if the constraint $\tilde{z}_{1,1}^{+}+\tilde{z}_{1,1}^{-}=1$ are active, there are $3\times N_w\times N_t-1$ linearly independent constraints active at $(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$. If $0<\tilde{z}_{1,1}^{+}+\tilde{z}_{1,1}^{-}<1$, then there are only $3\times N_w\times N_t-2$ linearly independent constraints active at $(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$. Thus, if $\tilde{z}_{1,1}^{+}$ and $\tilde{z}_{1,1}^{-}$ are only two elements of $(\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$ that are decimals between 0 and 1, then $(\tilde{\boldsymbol{w}},\tilde{\boldsymbol{z}}^{+},\tilde{\boldsymbol{z}}^{-})$ is also not a basic solution of U_0'' .

Other cases can be generalized from the above two cases, and then we can find that none of the element in S is a basic solution of U_0'' . Therefore, U_0''' is the set of basic feasible solutions of U_0'' as well as the set of extreme points of U_0'' , i.e., $\text{vert}(U_0'')=U_0'''$.

H. Proposition 9 and Its Proof

Proposition 9: $\text{vert}(U_0')\subseteq U_0$

Proof: Followed by Proposition 7 and 8, the set of the extreme points of U_0' is a subset of the projection of U_0''' , i.e., $\text{vert}(U_0')\subseteq\Pi_{N_w\times N_t}(U_0''')$. It is evident that U_1 is projection of U_0''' , i.e., $\Pi_{N_w\times N_t}(U_0''')=U_1$ .And thus, the set of the extreme points of U_0' is a subset of U_1 , i.e., $\text{vert}(U_0')\subseteq U_1$.

It is evident that U_0 is a subset of U_1 , i.e., $U_0\subseteq U_1$. Thus, to prove that the set of extreme points of U_0' is a subset of U_0 , we need to prove that the none of the extreme points belongs to the set difference between U_0 and U_1 , i.e., $V=U_1-U_0$, which means that none of the elements in V is an extreme points of U_0' . Here are the steps for the proof:

Assume $\tilde{\boldsymbol{w}}\in V$. In the WPO scenario $\tilde{\boldsymbol{w}}$, the WPO of wind farm 1 reaches the upper/ lower bound in the $1,\cdots,\Gamma_1-1$ time periods and reaches the forecasted value in the $\Gamma_1,\cdots,N_t$ time periods, while the WPO of other wind farm $r\neq 1$ reaches the upper/lower bound in the $1,\cdots,\Gamma_r$ time periods and reaches the forecasted value in the $\Gamma_r+1,\cdots,N_t$ time periods. In this case, we can easily find two elements $\boldsymbol{w}_1\in U_0'$ and $\boldsymbol{w}_2\in U_0'$. $\boldsymbol{w}_1=\tilde{\boldsymbol{w}}$ except for the WPO of wind farm 1 reaches the upper bound at the Γ_1 time periods. $\boldsymbol{w}_2=\tilde{\boldsymbol{w}}$ except for the WPO of wind farm 1 reaches the lower bound at the Γ_1 time periods. Therefore, $\tilde{\boldsymbol{w}}$ can be easily expressed as the convex combination of $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$

$$\tilde{\boldsymbol{w}}=\frac{-\delta_{1,\Gamma_1}^{-}}{\delta_{1,\Gamma_1}^{+}-\delta_{1,\Gamma_1}^{-}}\boldsymbol{w}_1+\frac{\delta_{1,\Gamma_1}^{+}}{\delta_{1,\Gamma_1}^{+}-\delta_{1,\Gamma_1}^{-}}\boldsymbol{w}_2 \tag{B.12}$$

Thus, $\tilde{\boldsymbol{w}}$ is not an extreme point of U_0' .

Other cases can be generalized from the above case, and then we can find that if any element in U_1 that is not active at all the constraints (2d), then this element can't not be an extreme point of U_0' . And thus, none of the elements in V is an extreme point of U_0' . And thus, $\text{vert}(U_0')\subseteq U_1-V=U_0$.

I. Proof for Proposition 1

Proof: Followed by Proposition 6 and 9, $U_0\subseteq\text{vert}(U_0')$ and $\text{vert}(U_0')\subseteq U_0$, and thus $\text{vert}(U_0')=U_0$. Moreover, followed by Proposition 5, we can conclude that $\text{vert}(U_0')=U_0\subseteq U_1\subseteq U_2'\subseteq U_0'$. Since that $\text{conv}(\text{vert}(U_0'))=\text{conv}(U_0')$, thus $\text{conv}(U_0')=\text{conv}(U_0)\subseteq\text{conv}(U_1)\subseteq\text{conv}(U_2')\subseteq\text{conv}(U_0')$. Equivalently, $\text{conv}(U_1)=\text{conv}(U_2')=\text{conv}(U_0')$. Furthermore, followed by Proposition 4, $U_2\subseteq U_2'$, and thus the convex hull of U_2 is a subset of the convex hull of U_2' , i.e., $\text{conv}(U_2)\subseteq\text{conv}(U_2')$. Ultimately, we conclude that he convex hull of U_2 is a subset of the convex hull of U_1 , i.e., $\text{conv}(U_2)\subseteq\text{conv}(U_1)$.

APPENDIX C
PROOF OF PROPOSITION 2

Proof: The master problem (9a)-(9c) and subproblem (10a)-(10f) at iteration j , along with the best feasible first-stage and second-stage solutions found within the relative optimality gap tolerance ε_{MP} and ε_{SP} are expressed as follows. The master problem (9a)-(9c) at iteration j can be equivalently expressed as

$$\upsilon_j^{*}=\min\boldsymbol{c}^T\boldsymbol{x}+\delta \tag{C.1a}$$

$$\text{s.t. } \boldsymbol{x}\in\mathcal{X} \tag{C.1b}$$

$$\delta\geq\boldsymbol{q}^T\boldsymbol{y}_k,\ k=1,\cdots,|\mathcal{S}| \tag{C.1c}$$

$$\boldsymbol{T}\boldsymbol{x}+\boldsymbol{W}\boldsymbol{y}_k+\boldsymbol{C}\boldsymbol{w}_k\geq\boldsymbol{h},\ k=1,\cdots,|\mathcal{S}| \tag{C.1d}$$

$$\boldsymbol{c}^T\boldsymbol{x}+\delta\geq\bar{L} \tag{C.1e}$$

where the same symbols in (C.1a)-(C.1e) as in (9a)-(9c) have the same meaning, $\boldsymbol{w}_k$ denotes the k-th WPO scenario

contained in $\mathcal{S}$, $\boldsymbol{y}_k$ denotes the recourse variables corresponding to the k-th WPO scenario of $\mathcal{S}$, $|\mathcal{S}|$ denotes the number of WPO scenarios contained in $\mathcal{S}$. The best feasible solution of the first-stage variables found at iteration j can be expressed as $(\boldsymbol{x}^j, \delta^j)$. The best feasible solution of the recourse variables found at iteration j can be expressed as $(\boldsymbol{y}_k^j)$ $k=1,\cdots,|\mathcal{S}|$.

The subproblem (10a)-(10f) at iteration j can be equivalently expressed as

$$D_j^* = \max_{\boldsymbol{w}\in U} \min_{\boldsymbol{y}\in\mathcal{Y}(\boldsymbol{x}^j,\boldsymbol{w})} \boldsymbol{q}^T\boldsymbol{y} \tag{C.2a}$$

where the same symbols in (C.2a) as in (10a)-(10f) have the same meaning. The best feasible second-stage solution found at iteration j can be expressed as $(\boldsymbol{w}^j, \boldsymbol{y}^j)$.

First, we show that when $\varepsilon_{MP}=0$ and $\varepsilon_{SP}=0$, the scenario $\boldsymbol{w}^j=\boldsymbol{\omega}^T\boldsymbol{y}_\omega^j$ reported from the subproblem (10a)-(10f) at any iteration j $(j\geq 2)$ will never duplicate a scenario already included in the scenario set $\mathcal{S}$, unless the upper and lower bounds converge. In the master problem at iteration j , since that υ_j^* is a lower bound to the optimal value υ^* of (1), i.e., $\upsilon_j^*\leq\upsilon^*$, the optimal solution of the first-stage variables $\boldsymbol{x}^j$ ensures that for every scenario $\boldsymbol{w}_k, k=1,\cdots,|\mathcal{S}|$, there exists a solution of the recourse variables $\boldsymbol{y}_k^j$ such that $\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}_k^j\leq\upsilon_j^*\leq\upsilon^*$. Therefore, the minimum scheduling cost based on the optimal first-stage solution $\boldsymbol{x}^j$ is less than or equal to υ^* for every scenario $\boldsymbol{w}_k, k=1,\cdots,|\mathcal{S}|$, i.e., $\boldsymbol{c}^T\boldsymbol{x}^j+\min_{\boldsymbol{y}\in\mathcal{Y}(\boldsymbol{x}^j,\boldsymbol{w}_k)}\boldsymbol{q}^T\boldsymbol{y}\leq\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}_k^j\leq\upsilon_j^*\leq\upsilon^*, k=1,\cdots,|\mathcal{S}|$. Meanwhile, in the subproblem at iteration j , since that $\boldsymbol{c}^T\boldsymbol{x}^j+D_j^*$ is an upper bound to υ^* , i.e., $\boldsymbol{c}^T\boldsymbol{x}^j+D_j^*=\boldsymbol{c}^T\boldsymbol{x}^j+\max_{\boldsymbol{w}\in U}\min_{\boldsymbol{y}\in\mathcal{Y}(\boldsymbol{x}^j,\boldsymbol{w})}\boldsymbol{q}^T\boldsymbol{y}\geq\upsilon^*$, the minimum scheduling costs based on the optimal first-stage solution $\boldsymbol{x}^j$ is larger than or equal to υ^* for the worst-case scenario $\boldsymbol{w}^j$, i.e., $\boldsymbol{c}^T\boldsymbol{x}^j+\max_{\boldsymbol{w}\in U}\min_{\boldsymbol{y}\in\mathcal{Y}(\boldsymbol{x}^j,\boldsymbol{w})}\boldsymbol{q}^T\boldsymbol{y}=\boldsymbol{c}^T\boldsymbol{x}^j+\min_{\boldsymbol{y}\in\mathcal{Y}(\boldsymbol{x}^j,\boldsymbol{w}^j)}\boldsymbol{q}^T\boldsymbol{y}=\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}^j\geq\upsilon^*$. Thus, the minimum scheduling cost based on the optimal first-stage solution $\boldsymbol{x}^j$ for every scenario $\boldsymbol{w}_k, k=1,\cdots,|\mathcal{S}|$ is less than or equal to that for the worst-case scenario $\boldsymbol{w}^j$, i.e., $\min_{\boldsymbol{y}\in\mathcal{Y}(\boldsymbol{x}^j,\boldsymbol{w}_k)}\boldsymbol{q}^T\boldsymbol{y}\leq\boldsymbol{q}^T\boldsymbol{y}^j, k=1,\cdots,|\mathcal{S}|$. Consequently, the worst-case scenario $\boldsymbol{w}^j$ duplicate to one of the scenarios $\boldsymbol{w}_{k'}\in\mathcal{S}$ if and only if $\min_{\boldsymbol{y}\in\mathcal{Y}(\boldsymbol{x}^j,\boldsymbol{w}_{k'})}\boldsymbol{q}^T\boldsymbol{y}=\boldsymbol{q}^T\boldsymbol{y}^j=\upsilon^*$, in which case the upper and lower bounds converge.

Next, we show that when $\varepsilon_{MP}>0$ and $\varepsilon_{SP}>0$, the scenario $\boldsymbol{w}^j=\boldsymbol{\omega}^T\boldsymbol{y}_\omega^j$ reported from the subproblem (10a)-(10f) at any iteration j $(j\geq 2)$ may duplicate a scenario already included in the scenario set $\mathcal{S}$ before the upper and lower bounds converging. Noting that when $\varepsilon_{MP}=0$ and $\varepsilon_{SP}=0$, a key issue for $\boldsymbol{w}^j$ not to duplicate $\boldsymbol{w}_k, k=1,\cdots,|\mathcal{S}|$ is that the minimum scheduling cost based on the optimal first-stage solution $\boldsymbol{x}^j$ for every scenario $\boldsymbol{w}_k, k=1,\cdots,|\mathcal{S}|$ is always less than or equal to that for the worst-case scenario $\boldsymbol{w}^j$. However, this condition does not necessarily hold when $\varepsilon_{MP}>0$ and $\varepsilon_{SP}>0$. In the master problem at iteration j , the best feasible solution of the first-stage variables $\boldsymbol{x}^j$ ensures that for every scenario $\boldsymbol{w}_k, k=1,\cdots,|\mathcal{S}|$, there exists a solution of the recourse variables $\boldsymbol{y}_k^j$ such that $\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}_k^j\leq U_{MP}^j$. Noting that U_{MP}^j is not necessarily a lower bound to υ^* . Meanwhile, in the subproblem at iteration j , the scheduling costs based on the best feasible first-stage solution $\boldsymbol{x}^j$ is equal to L_{SP}^j for the scenario $\boldsymbol{w}^j$, i.e., $\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}^j=L_{SP}^j$. Noting that L_{SP}^j is not necessarily an upper bound to υ^* . Consequently, $\boldsymbol{w}^j$ is possible to duplicate to one of the scenarios $\boldsymbol{w}_{k'}\in\mathcal{S}$. For example, $(\boldsymbol{w}_{k'}, \boldsymbol{y}_{k'}^j)$ can be the best feasible solution of the subproblem found at iteration j when $L_{SP}^j\leq U_{MP}^j$, since that $\boldsymbol{y}_k^j$ satisfies all the constraints $\boldsymbol{y}_{k'}^j\in\left\{\boldsymbol{y}\middle|\boldsymbol{T}\boldsymbol{x}^j+\boldsymbol{W}\boldsymbol{y}+\boldsymbol{C}\boldsymbol{w}_k^j\geq\boldsymbol{h}\right\}$ and the scheduling cost based on the best feasible first-stage solution $\boldsymbol{x}^j$ for the scenario $\boldsymbol{w}_{k'}$ satisfies the constraints $\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}_{k'}^j\leq U_{MP}^j$, i.e., $\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}_{k'}^j=\boldsymbol{c}^T\boldsymbol{x}^j+\boldsymbol{q}^T\boldsymbol{y}^j=L_{SP}^j\leq U_{MP}^j$.

APPENDIX D
PROOF OF PROPOSITION 3

In this section, we first introduce an auxiliary proposition, namely Proposition 10. Then, we will demonstrate Proposition 3 through the conclusions drawn from Proposition 10.

A. Proposition 10 and Its Proof

Proposition 10: If the II-C&CG method reaches the exploitation step at iteration j , i.e., the valid relative gap satisfies $(\bar{U}-L_{MP}^\ell)/\bar{U}\geq\varepsilon$, but the inexact relative gap satisfies $(\bar{U}-\bar{L})/\bar{U}<\tilde{\varepsilon}$, then the valid relative gap is at most $(1-\tilde{\varepsilon})^{-1}(1-\varepsilon_{MP})^{-(j-\ell+1)}-1$.

Proof: According to the deductions in [2], we have the following inequalities hold for $j\geq\ell$

$$\frac{U_{MP}^j-L_{MP}^\ell}{L_{MP}^\ell}\leq(1-\varepsilon_{MP})^{-(j-\ell+1)}-1=:\varepsilon'>0 \tag{D.1}$$

Then we have

$$\frac{U_{MP}^j}{1-\tilde{\varepsilon}}>\bar{U}\geq\upsilon^*\geq L_{MP}^\ell\geq\frac{U_{MP}^j}{1+\varepsilon'} \tag{D.2}$$

where the first inequality follows from $(\bar{U}-\bar{L})/\bar{U}<\tilde{\varepsilon}$, the second inequality follows since $\bar{U}$ is a valid upper bound on υ^* followed by Proposition 2, the third equality follows since L_{MP}^ℓ is a valid lower bound on υ^* [2], and the last inequality follows from (D.1). Hence, we can derive

$$\frac{\bar{U}-L_{MP}^{\ell}}{\bar{U}}<\frac{U_{MP}^{j}}{\bar{U}}\left(\frac{1}{1-\tilde{\varepsilon}}-\frac{1}{1+\varepsilon'}\right)\le\frac{1+\varepsilon'}{1-\tilde{\varepsilon}}-1 \qquad \text{(D.3)}$$

where the first inequality follows from the fact that (D.2) shows $\bar{U}<U_{MP}^{j}/(1-\tilde{\varepsilon})$ and $L_{MP}^{\ell}>U_{MP}^{j}/(1+\varepsilon')$, and the second inequality follows from the fact that (D.2) shows $U_{MP}^{j}/\bar{U}\le 1+\varepsilon'$. Hence, we can derive that

$$\frac{\bar{U}-L_{MP}^{\ell}}{\bar{U}}\le(1-\tilde{\varepsilon})^{-1}(1-\varepsilon_{MP})^{-(j-\ell+1)}-1 \qquad \text{(D.4)}$$

B. Proof for Proposition 3

Proof: In step 3, if the termination criterion is not satisfied, then the algorithm proceeds to either the exploration or the exploitation step. Therefore, to verify the finite convergence of the algorithms, it suffices to show that:

The algorithm can proceed to the exploration step only a finite number of times.

Followed by Proposition 2, we notice that during the iterations of Algorithm 1, the scenario $\boldsymbol{w}^j=\boldsymbol{\omega}^T\boldsymbol{y}_{\omega}^j$ reported from the subproblem (10a)-(10f) might duplicate a scenario already included in the scenario set $\mathcal{S}$. Therefore, if the exploration step adds $\boldsymbol{w}^j$ to the scenario set $\mathcal{S}$ without considering whether $\boldsymbol{w}^j$ already belongs to $\mathcal{S}$, then the algorithm may fall into an infinite repetitive loop, repeatedly adding the same duplicate scenario $\boldsymbol{w}^j$ to the $\mathcal{S}$. However, since that the criterion of the exploration step in Algorithm 1 requires that $\boldsymbol{w}^j$ can only be added to $\mathcal{S}$ when $\boldsymbol{w}^j\notin\mathcal{S}$, thus avoiding the aforementioned issue. Since that each time the algorithm visits the exploration step, a new scenario $\boldsymbol{w}^j\in U$ $(\mathcal{S}\subseteq U)$ and $\boldsymbol{w}^j\notin\mathcal{S}$ is added to $\mathcal{S}$, thus the algorithm can only proceed to the exploration step at most as many times as there are elements in U. Since that MSUS U_2 is discrete finite set, the algorithm can proceed to the exploration step only a finite number of times.

The algorithm can proceed to the exploitation step only a finite number of times.

According to the analysis in **Error! Reference source not found.**, if the selection range of $\tilde{\varepsilon}$ is not explicitly defined, then the algorithm may fall into another infinite loop, repeatedly backtracking at a certain iteration j $(j\ge 1)$ without progressing to the next iteration j or terminating the iterations altogether. This implies that for the iteration $j=\ell$, the conditions $(\bar{U}-L_{MP}^{\ell})/\bar{U}\ge\varepsilon$ and $(\bar{U}-\bar{L})/\bar{U}<\tilde{\varepsilon}$ remain satisfied regardless of how many times the algorithm backtracks. However, followed by Proposition 10, the valid relative gap is bounded by $(1-\tilde{\varepsilon})^{-1}(1-\varepsilon_{MP}^{k})^{-1}-1$. Since the value of ε_{MP}^{ℓ} is reduced in each backtracking, ε_{MP}^{ℓ} will converge to zero, and in turn the valid relative gap $(1-\tilde{\varepsilon})^{-1}(1-\varepsilon_{MP})^{-1}-1$ will converge $(1-\tilde{\varepsilon})^{-1}-1$, after a large enough number of backtrackings. Therefore, if $\tilde{\varepsilon}<\varepsilon/(1+\varepsilon)$ holds, then the valid relative gap $(1-\tilde{\varepsilon})^{-1}-1$ will inevitably be less than ε after a large enough number of backtrackings. And thus, the algorithm can proceed to the exploitation step only a finite number of times.

APPENDIX E
SYSTEM CONFIGURATION AND PARAMETERS

A. 6-Bus System

The IEEE 6-bus test system contains 6 buses, 7 transmission lines, 3 conventional generating units with a total capacity of 420 MW and 1 wind farm (No. 6389 wind farm) with a total capacity of 283 MW. No. 6389 wind farm is connected to No. 4 bus. Network and generator configurations and data for the IEEE 6-bus test system are described in [3].

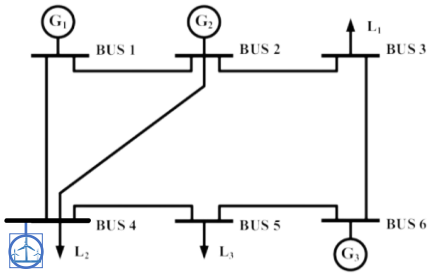

Fig. 1. IEEE 6-Bus Test System One-Line Diagram

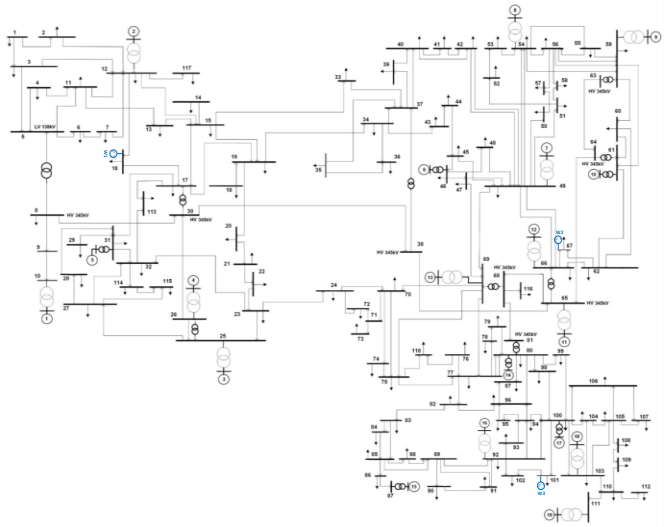

Fig. 2. IEEE 118-Bus Test System One-Line Diagram

B. 118-Bus System

The IEEE 118-bus test system contains 118 buses, 186 transmission lines, 54 conventional generating units with a total capacity of 7220 MW and 3 wind farms (No. 6389, 3902, 3945 wind farm) with a total capacity of 1130 MW. No. 6389, 3902, 3945 wind farms are connected to No. 16, 67, 101 bus respectively. Network and generator configurations and data for the IEEE 118-bus test system are described in [4].

APPENDIX F
PARAMETER SETTINGS AND SUPPLEMENTARY COMPARISONS

A. Parameter Settings of U_1, U_3, U_4, U_5

The parameters of U_1, U_3, U_4, U_5 are set based on the

same historical WPO data as U_2 using the parameter setting strategies given in [5], [6], [7], [8], respectively. The calculation of boundaries in U_1 employs the same conditional quantile regression technique as used in U_2. The prespecified parameters of U_1 are set as $\mathrm{Pr_I}$=0.95 . The prespecified parameters of U_3 are set as $\alpha_1 = 0.3$, $\Lambda = 5$. The prespecified parameters of U_4 are set as $N_b = 4$, $\alpha_1 = 0.3$, $\alpha_2 = 0.08$, $\Lambda = 5$. Among them, N_b denotes the number of bands of U_4, when $N_s = 2N_b + 1$, the interval of U_2 and U_4 between upper and lower bounds are divided into the same number of sub-intervals, α_1 denotes that the upper and lower bounds of U_3 and U_4 are set as $(1+\alpha_1)\times 100\%$ and $(1-\alpha_1)\times 100\%$ of the forecasted values, α_2 denotes that when $N_b > 1$, each new band is generated by reducing band interval of the most inner band by $\alpha_2 \times 100\%$, Λ denotes the variation budget of U_3 and U_4 . Moreover, the weight coefficients of U_4 are calculated based on prediction error distributions [7]. U_5 is constructed as a convex hull of daily historical WPO scenarios [8].

B. Supplementary Comparison Experiments

We conduct supplementary experiments to compare the effectiveness of U_1 , U_2 , U_3 , U_4 , U_5 in reducing the conservatism across the 6-node and 118-node systems, as well as with various settings of uncertainty set parameters, following the comparative methodology outlined in Section V-B.

In addition to the average day-ahead scheduling cost and the average intraday redispatch cost, we also supplement with another indicator, the number of days that the daily day-ahead scheduling schemes fail to pass safety verification (denoted as n), as 3 indicators to assess the effectiveness of U_1, U_2, U_3, U_4, U_5. We say that a day-ahead scheduling scheme does not pass the safety verification, if the load shedding penalty cost computed from the redispatch program is larger than 0. Table I, II, III, IV, V shows the above 3 indicators calculated under different parameters of U_1, U_2, U_3, U_4, U_5 respectively.

TABLE I
THREE INDICATORS CALCULATED FOR U_1

Bus	$\mathrm{Pr_I}$	Γ	c_1^1 (\$)	c_2^1 (\$)	n
6	0.95	6	172,438	129,227	0
		12	201,690	133,052	0
		18	224,734	133,409	0
		24	239,891	134,144	0
118		12	1,369,757	1,333,735	0
		24	1,409,315	1,336,933	0

TABLE II
THREE INDICATORS CALCULATED FOR U_2

Bus	N_s	$\mathrm{Pr_I}$	M	$\mathrm{Pr_\Gamma}$	Γ	c_1^2 (\$)	c_2^2 (\$)	n
6	9	0.95	20	0.9	6	164,234	122,423	0
					12	186,636	127,008	0
					18	211837	127,469	0
					24	224,651	128,807	0
				0.95	6	166,810	125,479	0
					12	192,528	129,935	0
					18	217,101	130,134	0
					24	233,007	131,319	0
				1	6	168,945	128,465	0
					12	194,897	132,772	0
					18	220,689	132,475	0
					24	238,973	130,109	0
118				0.9	12	1,351,435	1,446,563	1
					24	1,408,466	1,335,365	0
				0.95	12	1,359,268	1,329,049	0
					24	1,408,526	1,335,819	0
				1	12	1,366,584	1,333,185	0
					24	1,409,101	1,336,678	0

TABLE III
THREE INDICATORS CALCULATED FOR U_3

Bus	α_1	Λ	Γ	c_1^3 (\$)	c_2^3 (\$)	n
6	0.3	1	6	75,932	2,468,000	5
			12	97,202	1,397,854	5
			18	105,419	1,369,054	5
			24	111,292	1,389,548	5
		5	6	86,370	1,404,527	5
			12	99,950	1,363,314	5
			18	107,463	1,353,739	5
			24	112,248	1,359,221	5
	1	1	6	165,238	1,80,6170	5
			12	270,804	155,534	0
			18	307,500	156,598	0
			24	332,184	157,132	0
		5	6	228,531	154,564	0
			12	277,509	154,469	0
			18	310,605	154,932	0
			24	331,940	156,318	0
118	0.3	1	12	1,325,446	8,418,346	1
			24	1,336,473	8,425,195	1
		5	12	1,325,446	8,418,346	1
			24	1,321,293	8,719,817	1
	1	1	12	1,548,564	1,355,843	0
			24	1,626,314	1,365,031	0
		5	12	1,555,804	1,345,337	0
			24	1,627,783	1,356,596	0

TABLE IV
THREE INDICATORS CALCULATED FOR U_4

Bus	N_b	α_1	α_2	Λ	Γ	c_1^4 (\$)	c_2^4 (\$)	n
6	4	0.3	0.08	1	6	65,112	2,833,062	6
					12	79,799	1,993,041	6
					18	85,304	1,922,467	6
					24	89,731	1,867,714	6
				5	6	71,951	1,934,815	6
					12	82,067	1,898,995	6
					18	86,984	1,922,299	6
					24	90,653	1,925,514	6
		1	0.25	1	6	128056	2,382,343	5
					12	203916	591,514	5
					18	244867	152,125	0
					24	262731	152,556	0
				5	6	197782	154,686	0
					12	222899	154,535	0
					18	249022	154,379	0
					24	265526	154,293	0

118		0.3	0.08	1	12	1,305,930	8,382,304	1
					24	1,316,179	8,443,149	1
				5	12	1,304,202	8,401,256	1
					24	1,304,040	8,549,558	1
		1	0.25	1	12	1,332,122	1,366,817	0
					24	1,497,036	1,352,937	0
				5	12	1,441,116	1,346,558	0
					24	1,503,226	1,350,410	0

TABLE V
THREE INDICATORS CALCULATED FOR U_5

Bus	c_1^5 ($)	c_2^5 ($)	n
6	254,840	164,220	0
118	1,413,425	1341,772	0

From Table I, II, III, IV, V, we can draw the same conclusions as in Section V-B. Due to the lack of comprehensive parameter setting strategies for U_3, U_4, the scheduling schemes based on U_3, U_4 are either unreliable or conservative. We can observe that when the prespecified parameters of U_3, U_4 are set the same to the parameters used in the numerical experiments of [6], [7] respectively, i.e. the prespecified parameters of U_3 are set as $\alpha_1 = 0.3$, $\Lambda = 5$ and the prespecified parameters of U_4 are set as $N_b = 4$, $\alpha_1 = 0.3$, $\alpha_2 = 0.08$, $\Lambda = 5$, the scheduling schemes based on U_3, U_4 lead to serve load shedding in the intraday redispatch program. Meanwhile, we can observe that when the boundary value and interval ranges of U_3, U_4 are expanded, i.e., $\alpha_1 = 1$, $\alpha_2 = 0.25$, the scheduling schemes based on U_3, U_4 don't lead to load shedding in the intraday redispatch program, which validates the robustness of the scheduling scheme based on U_3, U_4. However, the day-ahead scheduling costs and intraday cost under U_3, U_4 are much higher than that under U_1, U_2 at the same Γ, which suggests that the scheduling schemes based on U_3, U_4 are over-conservative than U_1, U_2. Thus, we can find that the robustness or economics of the scheduling scheme based on U_3, U_4 can't be guaranteed due the empirical parameter setting strategies of U_3, U_4.

In contrast, the scheduling schemes based on U_1, U_2, U_5 don't lead to load shedding in the intraday redispatch program, which validates the robustness of the scheduling schemes based on U_1, U_2, U_5. The boundary value parameterized based on the quantile regression technique and the transition indicator parameterized based on the Markov chain analysis scientifically ensure the robustness of the scheduling scheme based on U_1, U_2 with confidence level $\mathrm{Pr_I}$ and $\mathrm{Pr_T}$. Convex hull analysis ensures the robustness of U_5. Meanwhile, the day-ahead scheduling cost and intraday redispatch cost based on U_2 is lower than that based on U_1, U_5, which demonstrates the economics of the scheduling scheme based on U_2. A more refined formulation of the WPO uncertainties and a more refined parameter setting strategy ensure the economics of U_2.

APPENDIX G
VALIDATION OF THE PROPOSITIONS

A. *Validation of Proposition 1*

Based on the MSUS parameters calculated in Section V-A, we construct a 2-dimensional MSUS U_2 using the parameters from the first 2 time periods. In accordance with the assumptions of Proposition 1, we further construct a 2-dimensional SB-DBUS U_1 under the same upper and lower bounds, forecasted values, and budgets. Fig. 3 shows the convex hull of U_1, U_2, U_2 not considering (3d), denoted as U_2', U_2 not considering (3d), (3e), denoted as U_2''. The dots in Fig. 3 represent the 2-dimensional WPO scenarios in the uncertainty sets, and the polygon formed by the connected line segments denotes the 2-dimensional convex hull composed of the WPO scenarios.

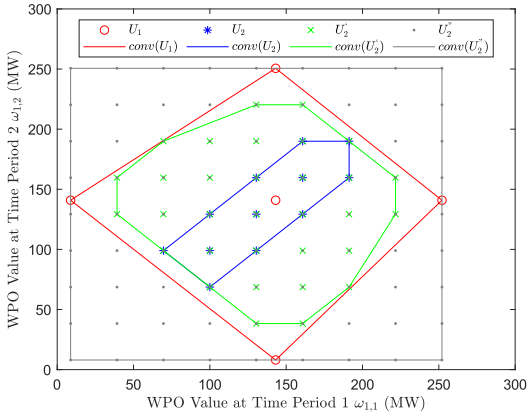

Fig. 3. The convex hull of U_1, U_2, U_2', U_2'' at $\Gamma = 1$

From Fig. 3, we can observe that $\mathrm{conv}(U_2) \subseteq \mathrm{conv}(U_2') \subseteq \mathrm{conv}(U_1) \subseteq \mathrm{conv}(U_2'')$. Among them, $\mathrm{conv}(U_2'')$ is a square, which, after considering constraints (3d), (3e), is divided into smaller polyhedrons $\mathrm{conv}(U_2)$ and $\mathrm{conv}(U_2')$. Additionally, both $\mathrm{conv}(U_2)$ and $\mathrm{conv}(U_2')$ are subsets of $\mathrm{conv}(U_1)$, verifying Proposition 1.

Moreover, we further construct 3-dimensional U_1, U_2 U_2', U_2'' using the parameters from the first 3time periods. Fig. 4 shows the convex hull of U_2, U_2''. Fig. 5 shows the convex hull of U_2, U_2', U_2''. Fig. 6 shows the convex hull of U_1, U_2, U_2', U_2' U_2''. Additionally, in Fig. 4, 5, 6, there are 2-dimensional projection of the 3-dimensional polyhedron onto the $\omega_{1,3} = 0$ plane.

From Fig. 4, 5, 6, we can observe that $\mathrm{conv}(U_2) \subseteq \mathrm{conv}(U_2') \subseteq \mathrm{conv}(U_1) \subseteq \mathrm{conv}(U_2'')$. Among them, $\mathrm{conv}(U_2'')$ is a cube, which, after considering constraints (3d), (3e), is divided into smaller polyhedrons $\mathrm{conv}(U_2)$ and $\mathrm{conv}(U_2')$.

Additionally, both conv(U_2) and conv(U_2') are subsets of conv(U_1), verifying Proposition 1.

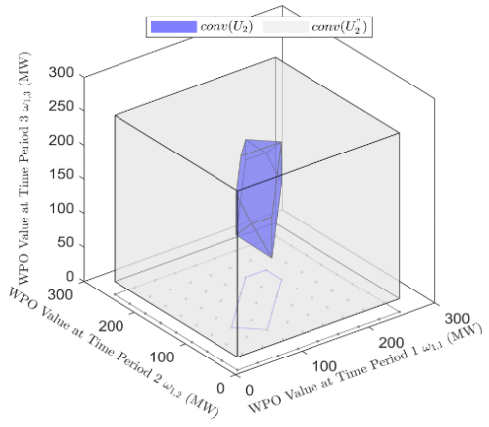

Fig. 4. The convex hull of U_2, U_2'' at $\Gamma = 1$

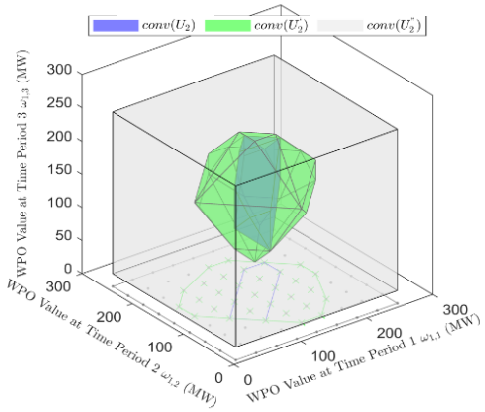

Fig. 5. The convex hull of U_2, U_2', U_2'' at $\Gamma = 1$

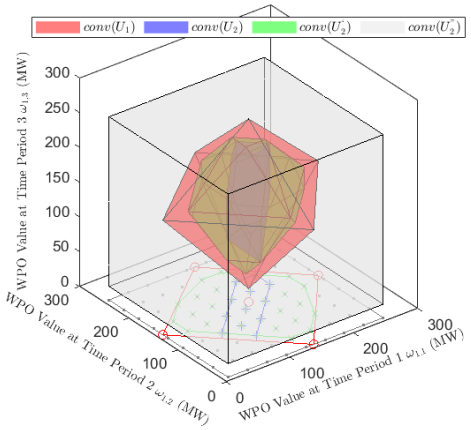

Fig. 6. The convex hull of U_1, U_2, U_2', U_2'' at $\Gamma = 1$

Furthermore, the results from the Appendix F indicate that, across various parameter settings and in both the 6-node and 118-node systems, when the assumptions of Proposition 1 are met, the day-ahead costs based on U_1 are consistently higher than those based on U_2. In conjunction with the analysis in [9], the larger the convex hull of a finite US, the greater the uncertainty it describes, leading to a higher cost of robust optimization decisions needed to hedge against such uncertainty. Thus, the results from the Appendix F also confirms the validity of Proposition 1.

Finally, we validate whether Proposition 1 still holds if $\tau_{r,t}^i$ is not set according to (4), i.e.,

$$\tau_{r,t}^i = \frac{|i-(N_s+1)/2|}{(N_s+1)/2}, i = 1, \cdots, N_s \tag{G.1}$$

This method of setting $\tau_{r,t}^i$ is computationally simple and independent of the specific values of the WPO state values $\omega_{r,t}^i$, depending solely on the state i of $\omega_{r,t}^i$. Under this setting, $\tau_{r,t}^i$ for the upper bounds ($y_{r,t}^{N_s} = 1$) and lower bounds ($y_{r,t}^1 = 1$) equals 1, while $\tau_{r,t}^i$ for intermediate state decreases as their states i approach the center $(N_s+1)/2$. Combined with the budget constraints (3d), this setting of $\tau_{r,t}^i$ can also limit the realization of WPO uncertainties deviating from the central value to some extent. However, under this setting, the convex hull of U_2 may not necessarily be a subset of the convex hull of U_1. Fig. 7 shows the convex hulls of U_1, U_2, U_2', U_2'' when $\tau_{r,t}^i$ is set according to (G.1).

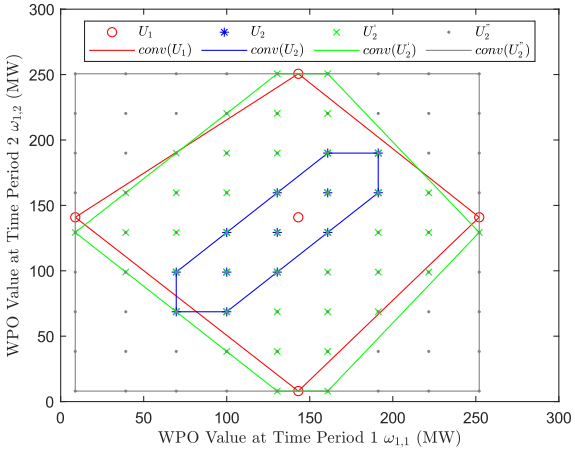

Fig. 7. The convex hull of U_1, U_2, U_2', U_2'' at $\Gamma = 1$ when $\tau_{r,t}^i$ is set according to (G.1).

Evidently, when $\tau_{r,t}^i$ is set according to (G.1), the convex hull of U_2 no longer exhibits a containment relationship with that of U_1. Therefore, when $\tau_{r,t}^i$ is not set according to (4), Proposition 1 may not necessarily hold.

B. Validation of Proposition 2 and 3

First, we verify the accuracy and convergence of the II-C&CG method. We apply the C&CG, I-C&CG and II-C&CG method to solve TSRUCM based on U_2 on the 6-bus system. We set $\varepsilon = 1\times10^{-4}$ for the 3 method, set $\tilde{\varepsilon} = 0.8\times10^{-4}$, $\varepsilon_{MP} = 1\times10^{-2}$ and $\alpha = 0.8$ for the I-C&CG and II-C&CG method, set $\varepsilon_{SP} = 1\times10^{-2}$ and $\beta = 0.8$ for the II-C&CG method.

Fig. 7 shows the computational results during the iterations using the II-C&CG method. The yellow curve with circles

denotes the objective values of the master problem $U_{MP}^{j} = \boldsymbol{c}^T \boldsymbol{x}^j + \delta^j$ calculated at each iteration. The blue curve with diamonds denotes the extended objective values of the subproblem $L_{SP}^{j} = \boldsymbol{c}^T \boldsymbol{x}^j + \boldsymbol{q}^T \boldsymbol{y}^j$. The green and red dashed lines denote the valid upper bound $\bar{U}$ and valid lower bound L_{MP}^{ℓ}, respectively. The algorithm converges after 4 backtracks and 77 iterations, which verifies the convergence of our proposed method.

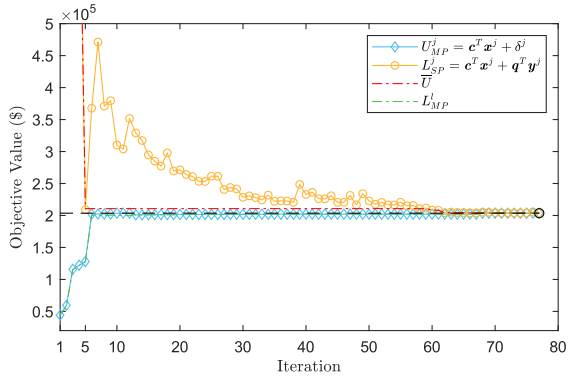

Fig. 7. The computational results of the I-C&CG method

Table VI shows the final computational results under 3 methods. We can observe that the final relative gap $\left(\bar{U} - L_{MP}^{\ell}\right)/\bar{U}$ of the II-C&CG method converges to zero and the optimal value calculated by our proposed method is the same as that calculated by the C&CG and I-C&CG method, which verifies the accuracy of our proposed method. We can also find that the computation time of the IIC&CG method is 32% and 45% shorter than that of the C&CG and I-C&CG method, although the total number of iterations of the II-C&CG method is greater than that of the C&CG method, verifying the computational advantage of our proposed method.

TABLE VI
COMPUTATION RESULTS UNDER THE 3 METHODS

Method	c_1 ($)	Final Relative Gap	Iteration	Computation Time
C&CG	203,591	0	67	1572
I-C&CG	203,591	0	82	1276
II-C&CG	203,591	0	77	865

Second, to validate the correctness of Proposition 2 and the effectiveness of criterion $\boldsymbol{w}^j \notin \mathcal{S}$ in the exploration step, we remove this criterion from Algorithm 1 and apply the modified algorithm to solve TSRUCM based on U_2 on the 6-bus system. The computational parameters are set as above.

Fig. 8 shows the computational results during the iterations using the II-C&CG method without the criterion $\boldsymbol{w}^j \notin \mathcal{S}$ in the exploration step. During the 69th iteration, the scenario calculated was identical to the one obtained in the 68th iteration. This causes the algorithm to repeatedly recalculate the master and subproblems of the 69th iteration, producing the same result after the 69th iteration. Consequently, the algorithm enters an infinite loop, resulting in convergence failure. Comparing to the successful convergence of the original algorithm presented in Fig. 7, the correctness of Proposition 2 and the necessity of the criterion $\boldsymbol{w}^j \notin \mathcal{S}$ is verified.

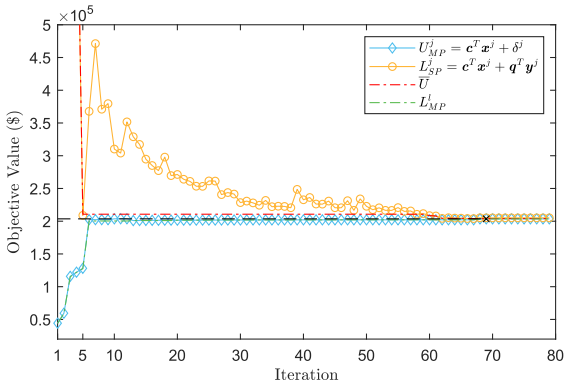

Fig. 8. The iteration results of the II-C&CG method without criterion $\boldsymbol{w}^j \notin \mathcal{S}$ in the exploration step

At last, to validate the correctness of Proposition 3 and the effectiveness of appropriate selection range of $\tilde{\varepsilon}$, we select a $\tilde{\varepsilon}$ out of the given range $\tilde{\varepsilon} < \varepsilon/(1+\varepsilon)$, i.e., $\tilde{\varepsilon} = 0.5$. The remaining computational parameters are set the same as above.

Fig. 9 shows the computational results during the iterations using the II-C&CG method when $\tilde{\varepsilon}$ does not satisfy the condition $\tilde{\varepsilon} < \varepsilon/(1+\varepsilon)$. The algorithm conducts the first backtrack during the 5th iteration. However, since $\tilde{\varepsilon}$ does not satisfy the condition $\tilde{\varepsilon} < \varepsilon/(1+\varepsilon)$, the algorithm falls into an infinite loop, continuously backtracking at the 5th iteration. In subsequent calculations, the algorithm repeatedly recalculates the master and subproblems of the 5th iteration, with progressively decreasing value of ε_{MP} and ε_{SP}. Consequently, the algorithm can't enter the exploration step to explore new scenarios or meet the convergence criterion to terminate the iterations. Comparing to the successful convergence of the original algorithm presented in Fig. 7, the correctness of Proposition 3 and the necessity of the selection range of $\tilde{\varepsilon}$ is verified.

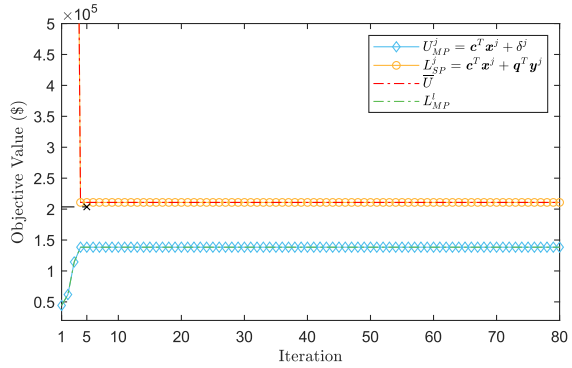

Fig. 9. The iteration results of the II-C&CG method when $\tilde{\varepsilon}$ does not satisfy the condition $\tilde{\varepsilon} < \varepsilon/(1+\varepsilon)$

APPENDIX H
VERIFICATION OF THE EFFECTIVENESS OF MSUS IN MSRUCM

First, U_1, U_2, U_3, U_4, U_5 for each day of the week gained at Section V-B are utilized. Next, we solve the MSRUCM [10] based on U_1, U_2, U_3, U_4, U_5, respectively, to generate daily day-ahead scheduling schemes and costs for this week. MSRUCM based on U_1, U_2, U_3, U_4, U_5 are solved using affine policy and constraint generation method [10].

Afterwards, we substitute each day's day-ahead scheduling schemes of this week into an intraday redispatch program to generate intraday redispatch schemes and costs. The cost coefficients are consistent with those set in Section V-B. We take the average day-ahead scheduling cost of the week (denoted as d_1), and the average intraday redispatch cost of the week (denoted as d_2) and the number of days that the daily day-ahead scheduling schemes fail to pass safety verification (denoted as n) as 3 indicators to assess the effectiveness and advantage of U_1, U_2, U_3, U_4, U_5. d_1^i, d_2^i ($i=1,2,3,4,5$) represent the specific cost under the corresponding U_i ($i=1,2,3,4,5$).

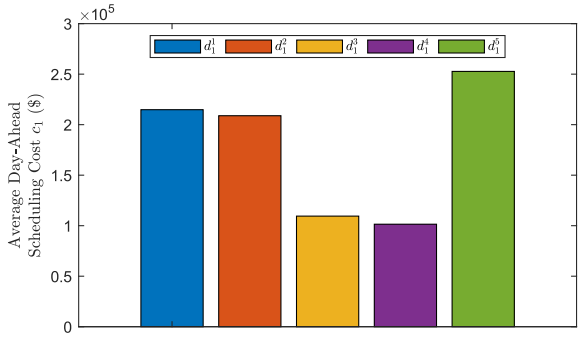

Fig. 10. Average day-ahead scheduling cost d_1^i at $\Gamma=12$

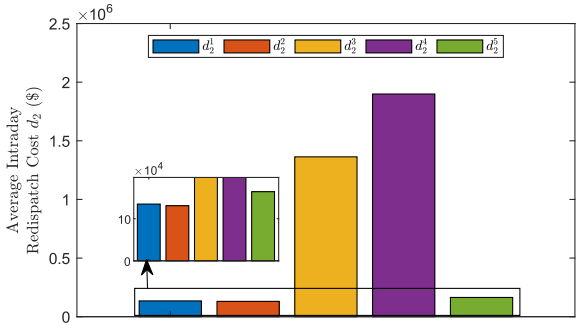

Fig. 11. Average intraday redispatch cost d_2^i at $\Gamma=12$

From Fig. 10, 11, we can observe that the day-ahead and intra-day computational results of the MSRUCM based on U_1, U_2, U_3, U_4, U_5 show similar outcomes to those of the TSRUCM. Due to the lack of comprehensive parameter setting strategies for U_3, U_4, the scheduling schemes based on U_3, U_4 are not robust enough, leading to the occurrence of load shedding in the intra-day redispatch and expensive intra-day costs. In contrast, the scheduling schemes based on U_1, U_2, U_5 don't lead to load shedding in the intraday redispatch program, which validates the robustness of the scheduling schemes based on U_1, U_2, U_5. Meanwhile, the day-ahead scheduling cost and intraday redispatch cost based on U_2 is lower than that based on U_1, U_5, which demonstrates the economics of the scheduling schemes based on U_2. Therefore, MSUS is also effective and advanced in reducing the conservatism of the scheduling schemes generated from the MSRUCM.

APPENDIX I
VERIFICATION OF THE EFFECTIVENESS OF TSDRUCM

We construct a Wasserstein-matric based AS [11]. First, we generated the original support scenario set for AS by sampling based on the probability error matrix obtained from historical data statistics. Subsequently, we used a Wasserstein distance-based scenario reduction method to reduce the original support scenario set, creating a typical support scenario set [12]. Finally, we constructed AS based on this typical support scenario set. Fig. 12 shows the forecasted WPO scenario for Jan. 1, 2006, along with 100 WPO scenarios from the original support scenario set generated based on the prediction error probability sampling.

Next, we solve the TSDRUCM [11] based on the Wasserstein-metric based AS to generate daily day-ahead scheduling schemes and costs for the first week of January 2006 on the 6-bus system. Afterwards, we substitute each day's day-ahead scheduling schemes of this week into an intraday redispatch program to generate intraday redispatch schemes and costs. The cost coefficients are consistent with those set in Section V-B. We take the average day-ahead scheduling cost of the week (denoted as e_1), the average intraday redispatch cost of the week (denoted as e_2) and the number of days that the daily day-ahead scheduling schemes fail to pass safety verification (denoted as n) as 3 indicators to assess the effectiveness and advantage of TSDRUCM based on the Wasserstein-metric based AS. Table VII presents the computational results of TSDRUCM under different numbers of typical support scenario sets.

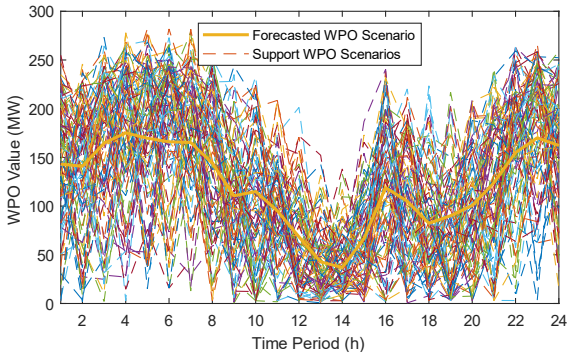

Fig. 12. The forecasted WPO scenario and support WPO scenarios

TABLE VII
THREE INDICATORS CALCULATED FOR TSDRUCM

Bus	number	e_1 ($)	e_2 ($)	n
6	10	126,119	4,289,485	4
	50	184,643	136,827	0
	100	189,032	138,456	0

From Table VII, we can find that when the number of scenarios in the support scenario set is small, the day-ahead scheduling scheme of TSDRUCM tends to be less robust, often leading to load shedding incidents during intraday redispatch and resulting in high intraday scheduling costs. As the number of scenarios increases, the reliability of the day-ahead scheduling scheme gradually improves, and load shedding incidents no longer occur during intraday redispatch. However, the selection of an appropriate number of scenarios remains an unresolved issue in current research, and the reliability assessment of the TSDRUCM scheduling scheme still relies on empirical judgment. When the number of wind farms is small, AS can ensure the economic efficiency and robustness of the scheduling schemes based on TSDRUCM with a relatively small number of WPO scenario samples. However, as the number of wind farms increases, the required sample size of the combinatorial WPO scenarios for multiple farms surges, making it difficult for DRUCM to guarantee the robustness of the scheduling schemes. In such cases, the scheduling schemes based on TSRUCM exhibit superior robustness and provide greater security and reliability.

REFERENCES

[1] Dimitris Bertsimas and John Tsitsiklis, “The geometry of linear programming,” in *Introduction to Linear Optimization*, 1st. ed. Massachusetts, USA: Athena Scientific, 1997, pp. 41-79.

[2] M. Tsang, K. Shehadeh and F. Curtis, “An inexact column-and-constraint generation method to solve two-stage robust optimization problems,” *Oper. Res. Lett.*, vol. 51, no.1, pp. 92-98, Jan. 2023, doi: 10.1016/j.orl.2022.12.002.

[3] H. Wu, X. Guan, Q. Zhai, F. Gao and Y. Yang, “Security-constrained generation scheduling with feasible energy delivery,” *IEEE PES General Meeting*, Calgary, AB, Canada, 2009, pp. 1-6.

[4] P. Dehghanian, “Power System Topology Control for Enhanced Resilience of Smart Electricity Grids,” PhD Dissertation, Texas A&M University, Texas , USA, 2017.

[5] R. Jiang, J. Wang and Y. Guan, “Robust unit commitment with wind power and pumped storage hydro,” *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 800-810, May 2012, doi: 10.1109/TPWRS.2011.2169817.

[6] L. Fan, K. Wang, and G. Li, “Robust unit commitment considering temporal correlation of wind power,” Autom. Elect. Power Syst., vol. 42, no. 18, pp. 91–97, Sep. 2018, doi: 10.7500/AEPS20170917007.

[7] Y. B. Chen, Z. Zhang, H. Chen, and H. P. Zheng, “Robust UC model based on multi-band uncertainty set considering temporal correlation of wind/load prediction errors,” *IET Gener. Transm. Distrib.*, vol. 14, no. 2, pp. 180-190, Jan. 2020, doi: 10.1049/iet-gtd.2019.1439.

[8] Y. Zhang, X. Ai, J. Wen, J. Fang and H. He, “Data-Adaptive robust optimization method for the economic dispatch of active distribution networks,” *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3791-3800, Jul. 2019, doi: 10.1109/TSG.2018.2834952.

[9] A. Ben-Tal, A. Nemirovski, “Robust optimization methodology and applications,” *Math Program.*, vol. 92, pp. 453–480, May 2002, doi: 10.1007/s101070100286.

[10] A. Lorca, X. Sun, E. Litvinov and T. Zheng, “Multistage adaptive robust optimization for the unit commitment problem,” *Oper. Res.*, vol. 64, no. 1, pp. 32-51, Jan 2016, doi: 10.1287/opre.2015.1456.

[11] S. Wang, C. Zhao, L. Fan and R. Bo, “Distributionally robust unit commitment with flexible generation resources considering renewable energy uncertainty,” *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4179-4190, Nov. 2022, doi: 10.1109/TPWRS.2022.3149506.

[12] X. Dong, Y. Sun, T. Pu, N. Chen, K. Sun, “An Optimal Scenario Reduction Method Based on Wasserstein Distance and Validity Index,” Proceedings of the CSEE, vol. 39, no. 16, pp. 4650-4658, Aug. 2019, doi: 10.13334/j.0258-8013.pcsee.181494.