

Multi-stage Robust Unit Commitment with Discrete Load Shedding Based on Partially Affine Policy and Two-stage Reformulation

Zhongjie Guo, *Member, IEEE*, Jiayu Bai, *Member, IEEE*, Wei Wei, *Senior Member, IEEE*, Haifeng Qiu, *Member, IEEE*, and Weihao Hu, *Senior Member, IEEE*

Abstract—This paper studies the problem of multi-stage robust unit commitment with discrete load shedding. In the day-ahead phase, the on-off status of thermal units is scheduled. During each period of real-time dispatch, the output of thermal units and the action of load shedding are determined, and the discrete choice of load shedding corresponds to the practice of tripping substation outlets. The entire decision-making process is formulated as a multi-stage adaptive robust optimization problem with mixed-integer recourse, whose solution takes three steps. First, we propose and apply partially affine policy, which is optimized ahead of the day and restricts intertemporal dispatch variables as affine functions of previous uncertainty realizations, leaving remaining continuous and binary dispatch variables to be optimized in real time. Second, we demonstrate that the resulting model with partially affine policy can be reformulated as a two-stage robust optimization problem with mixed-integer recourse. Third, we modify the standard nested column-and-constraint generation algorithm to accelerate the inner loops by warm start. The modified algorithm solves the two-stage problem more efficiently. Case studies on the IEEE 118-bus system verify that the proposed partially affine policy outperforms conventional affine policy in terms of optimality and robustness; the modified nested column-and-constraint generation algorithm significantly reduces the total computation time; and the proposed method balances well optimality and efficiency compared with state-of-the-art methods.

Index Terms—Unit commitment, robust optimization, uncertainty, affine policy, load shedding.

Manuscript received: February 26, 2024; revised: May 20, 2024; accepted: July 26, 2024. Date of CrossCheck: July 26, 2024. Date of online publication: January 23, 2025.

This work was supported in part by the Postdoctoral Fellowship Program of the Chinese Postdoctoral Science Foundation (No. GZB20240105), in part by the China Postdoctoral Science Foundation (No. 2024M750349), and in part by the National Natural Science Foundation of China (No. 52407084).

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

Z. Guo and W. Hu (corresponding author) are with the School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: guozhongjie@uestc.edu.cn; whu@uestc.edu.cn).

J. Bai is with the State Grid Sichuan Electric Power Research Institute, Chengdu 610041, China (e-mail: bai_jiayu01@163.com).

W. Wei is with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: wei-wei04@mails.tsinghua.edu.cn).

H. Qiu is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore (e-mail: haifeng.qiu@ntu.edu.sg).

DOI: 10.35833/MPCE.2024.000202

I. INTRODUCTION

RENEWABLE energy generation introduces unprecedented randomness to power system operations, so the grid operator should make decisions considering how to handle uncertainty. Unit commitment (UC) is one of the most important decision-making problems. Using robust optimization theory, two-stage robust UC has been widely investigated [1]–[3], which divides the variables into two groups. One group includes commitment variables, which are called here-and-now decisions and determined before uncertainty observation. The other group includes all the dispatch variables, which are called wait-and-see recourse variables and determined after uncertainty observation.

A shortcoming of two-stage formulation is pointed out via a counterexample in [4] that the non-anticipativity of dispatch decisions is violated. Non-anticipativity is a crucial issue in decision-making problems with uncertainty, which refers to that the decisions in any period (stage) cannot rely on the uncertainty realizations in the future [5]. In other words, the underlying decision rule of x_t should be $x_t = f_t(\xi_t)$ where ξ_t is the uncertainty realizations up to period t . The dispatch sequence behind the two-stage formulation is $\{(\xi_1, \xi_2, \dots, \xi_T), (x_1, x_2, \dots, x_T)\}$, which does not respect the non-anticipativity.

To address this issue, multi-stage robust UC (MRUC) has been studied, which formulates the real-time dispatch and uncertainty realization as two intertwined processes, i.e., $\{\xi_1, x_1, \xi_2, x_2, \dots, \xi_T, x_T\}$, which aligns with operation practice [6]. The MRUC problem has a nested structure and thus is intractable. Most existing research works consider continuous dispatch variables with three mainstream solution methods.

1) Policy/rule-based methods [7], [8]. These methods stipulate the dispatch variable x_t as a specific function (typically affine function) of previous uncertainty realizations. The affine policy is $x_t = a_t \xi_t + b_t$, where a_t and b_t are the coefficient variables that are determined in the day-ahead phase. Such methods naturally satisfy non-anticipativity because the rules have specified the dependence on previous information only. Using the affine policy, the resulting model is a static robust optimization problem and can be solved by its robust

counterpart or other techniques. However, the pre-selected affine structure may jeopardize the optimality.

2) Region-based methods, also named as implicit policy method [9]-[11]. These methods aim to find time-independent feasible regions/bounds for intertemporal dispatch variables, which are established ahead of the day by considering some vertices of uncertainty set and extreme ramping events between two consecutive periods. With these regions, the real-time dispatch resorts to optimal power flow problems that are decoupled over time.

3) Dynamic programming (DP)-based methods [12], [13]. Following the convention of DP, the optimization problem in each stage minimizes the sum of instant cost and future cost. The future cost is estimated by cost-to-go functions, which are trained ahead of the day. Due to the curses of dimensionality, DP-based methods suffer from high computation burden; besides, robust feasibility is a tough issue for DP-based methods, so they usually assume a relatively complete recourse, which is impossible in practice.

The above methods handle MRUC problem with continuous dispatch variables in their own ways. In practice, some real-time dispatch actions are discrete, such as load shedding. Existing research works treat load shedding as a continuous variable, which should be discrete [14]. The reason is that load shedding is usually realized by tripping substation outlets [15]. From the perspective of transmission system, the load cannot be continuously adjusted. As a result, we need to consider MRUC with mixed-integer dispatch variables.

With integer recourse, the problem becomes much more complicated. If we revisit the aforementioned three methods, the region-based methods do not work because they rely on the model linearity and convexity. The DP-based method is extended to consider integer variables in [16], but the computation burden remains a trouble due to curses of dimensionality and higher complexity. Besides, to use decision rules, one task is to tackle binary recourse, which requires piecewise constant policy. Finite adaptability technique makes sense in two-stage problems [17], [18], but no successful attempt is found in MRUC applications. The state-of-the-art rule-based methods are reported in [19]-[21]. The method in [19] and [20] uses a pre-defined partitioning function to formulate the piecewise constant rules, and the final problem is a scalable mixed-integer linear program (MILP). However, there is no specific method to design the optimal partitioning function. Instead, the method in [21] optimizes the partitioning function ahead of the day, but the scalability is jeopardized.

Therefore, this paper focuses on the MRUC problem with discrete load shedding and investigates how to handle binary recourse in a more tractable manner. The contributions are twofold below.

1) Problem modeling: we formulate the UC with discrete load shedding as a multi-stage adaptive robust optimization problem with mixed-integer recourse. The discrete choice of load shedding corresponds to the practice of tripping substation outlets. In the multi-stage decision-making sequence, the on-off status of thermal units is scheduled ahead of the day. During each period of real-time dispatch, the output of

thermal units and the action of load shedding are determined.

2) Solution methodology: we propose a systematic solution method to MRUC problem. The first step is to apply partially affine policy to restrict intertemporal dispatch variables as affine functions of previous uncertainty realizations. Non-state and binary dispatch variables are left to be optimized in real time. The second step is to derive a two-stage robust optimization problem with partially affine policy. The third step is to implement a modified nested column-and-constraint generation (M-nCCG) algorithm, which refines the standard nested column-and-constraint generation (S-nCCG) algorithm by warming up the inner loops.

The rest of this paper is organized as follows. The mathematical formulation is introduced in Section II. The multi-stage adaptive robust optimization model is presented in Section III, whose solution methodology is proposed in Section IV. Case studies are provided in Section V, and conclusions are drawn in Section VI.

II. MATHEMATICAL FORMULATION

A. Thermal Units

1) Coal-fired Units

$$b_t^c, u_t^c, v_t^c \in \{0, 1\} \quad \forall c \in \mathbb{C}, t \in \mathbb{T} \quad (1a)$$

$$-b_{t-1}^c + b_t^c - b_t^c \leq 0 \quad \forall c \in \mathbb{C}, t \in \mathbb{T}, \tau = (t+1): \min\{\beta_c^U + t - 1, T\} \quad (1b)$$

$$b_{t-1}^c - b_t^c + b_t^c \leq 1 \quad \forall c \in \mathbb{C}, t \in \mathbb{T}, \tau = (t+1): \min\{\beta_c^D + t - 1, T\} \quad (1c)$$

$$b_{t-1}^c - b_t^c + u_t^c - v_t^c = 0 \quad \forall c \in \mathbb{C}, t \in \mathbb{T} \quad (1d)$$

$$P_t^c - P_{t-1}^c \leq R_c^+(1 + b_{t-1}^c - b_t^c) + (2 - b_t^c - b_{t-1}^c)L^c \quad \forall c \in \mathbb{C}, t \in \mathbb{T} \quad (1e)$$

$$P_{t-1}^c - P_t^c \leq R_c^-(1 - b_{t-1}^c + b_t^c) + (2 - b_t^c - b_{t-1}^c)U^c \quad \forall c \in \mathbb{C}, t \in \mathbb{T} \quad (1f)$$

$$b_t^c P_{\min}^c \leq P_t^c \leq b_t^c P_{\max}^c \quad \forall c \in \mathbb{C}, t \in \mathbb{T} \quad (1g)$$

where c and \mathbb{C} are the index and set of coal-fired units, respectively; t and $\mathbb{T} = \{1, 2, \dots, T\}$ are the index and set of periods, respectively; b_t^c , u_t^c , v_t^c , and P_t^c are the decision variables, and if $b_t^c = 1$ ($b_t^c = 0$), unit c is on (off) during period t , and $u_t^c = 1$ ($v_t^c = 1$) means unit c is turned on (off) during period t ; β_c^U and β_c^D are the minimum-up time and minimum-down time, respectively; R_c^+ and R_c^- are the ramp-up and ramp-down rates, respectively, meanwhile, if unit c starts up during period t , its output P_t^c is no higher than the start-up rate L^c , and if it shuts down during period t , P_{t-1}^c is no higher than the shut-down rate U^c ; and P_{\min}^c and P_{\max}^c are the lower and upper bounds on P_t^c , respectively. Formulas (1b) and (1c) are the minimum-up and minimum-down time constraints, respectively; if unit c starts up during period t , i.e., $b_{t-1}^c = 0$ and $b_t^c = 1$, (1b) keeps this unit on during the next $\beta_c^U - 1$ periods, which is similar to the minimum-down time. According to (1d), we have $u_t^c = 1$ if $b_{t-1}^c = 0$ and $b_t^c = 1$, and $v_t^c = 1$ if $b_{t-1}^c = 1$ and $b_t^c = 0$. If $b_{t-1}^c = b_t^c$, both u_t^c and v_t^c will be zero since the start-up and shut-down costs are to be minimized. Formulas (1e) and (1f) are the improved ramping

constraints. Finally, (1g) imposes upper and lower bounds on P_i^c .

2) Gas-fired Units

Compared with coal-fired units, the gas-fired ones can reach the maximum output within a few minutes. Therefore, in an hourly model, the ramping limits can be discarded [22]. The operating model of gas-fired units is given below.

$$b_i^g, u_i^g, v_i^g \in \{0, 1\} \quad \forall g \in \mathbb{G}, t \in \mathbb{T} \quad (2a)$$

$$-b_{i-1}^g + b_i^g - b_\tau^g \leq 0 \quad \forall g \in \mathbb{G}, t \in \mathbb{T}, \tau = (t+1): \min\{\beta_g^U + t - 1, T\} \quad (2b)$$

$$b_{i-1}^g - b_i^g + b_\tau^g \leq 1 \quad \forall g \in \mathbb{G}, t \in \mathbb{T}, \tau = (t+1): \min\{\beta_g^D + t - 1, T\} \quad (2c)$$

$$b_{i-1}^g - b_i^g + u_i^g - v_i^g = 0 \quad \forall g \in \mathbb{G}, t \in \mathbb{T} \quad (2d)$$

$$b_i^g P_{\min}^g \leq P_i^g \leq b_i^g P_{\max}^g \quad \forall g \in \mathbb{G}, t \in \mathbb{T} \quad (2e)$$

where g and \mathbb{G} are the index and set of gas-fired units, respectively. We do not interpret model (2) in detail, because it can be easily understood by analogy with model (1). The main difference is that the gas-fired unit is free from ramping limits. The minimum-up and minimum-down constraints remain, but β_g^U and β_g^D may be smaller. In some research works, the ramping rate of gas-fired units is considered [23], [24]. Nevertheless, we can always divide all thermal units into ramping-constrained ones and ramping-free ones in an hourly UC problem.

B. Uncertain Wind Power

Let w and \mathbb{W} be the index and set of wind farms, respectively. Under time-varying weather conditions, the available wind power ξ_t^w is random. The utilized wind power is denoted by P_t^w , and the rest ΔP_t^w is curtailed. The operating model of wind generation is given below.

$$\begin{cases} 0 \leq P_t^w \leq \xi_t^w \\ P_t^w + \Delta P_t^w = \xi_t^w \end{cases} \quad \forall w \in \mathbb{W}, \forall t \in \mathbb{T} \quad (3)$$

To handle the randomness of wind power, we use the following box-type (polyhedral) uncertainty set:

$$\Xi_t^w = \left\{ \xi_t^w | \xi_t^{w,0} - \delta_t^w \leq \xi_t^w \leq \xi_t^{w,0} + \delta_t^w, 0 \leq \xi_t^w \leq P_{\max}^w \right\} \quad \forall w \in \mathbb{W}, \forall t \in \mathbb{T} \quad (4)$$

where $\xi_t^{w,0}$ is the wind power forecast; δ_t^w is the maximum forecast error; and P_{\max}^w is the installed capacity of wind generation. If necessary, spatial budget can be incorporated into the uncertainty set. The resulting uncertainty set remains a polyhedron, which can be handled by the method to be proposed. However, the temporal budget is usually not considered in the MRUC problem [10], because it contains global information that will violate the non-anticipativity of decisions.

C. Load

Let l and \mathbb{L} be the index and set of loads, respectively. We always hope that load P_l^l can be fully satisfied, but load shedding ΔP_l^l is sometimes inevitable. In the real world, load shedding is usually realized by tripping substation outlets [15]. As in Fig. 1, P_l^l is distributed in N_l outlets. In practice, the load on each outlet is measured. Without loss of general-

ty, we assume $P_l^{l,1} = P_l^{l,2} = \dots = P_l^{l,N_l} = P_l^l / N_l$.

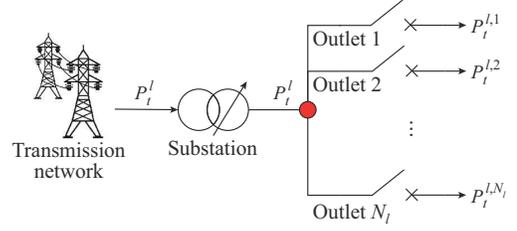


Fig. 1. Load shedding by tripping substation outlet.

We assume the former N_l^* outlets can be tripped. Outlets $(N_l^* + 1):N_l$ connecting to important loads, such as hospital, should not be tripped. Therefore, using binary variables $b_i^{l,1:N_l^*}$, the load model is:

$$b_i^{l,n} \in \{0, 1\} \quad n = 1:N_l^*, \forall l \in \mathbb{L}, \forall t \in \mathbb{T} \quad (5a)$$

$$\Delta P_l^l = \frac{P_l^l}{N_l} \sum_{n=1}^{N_l^*} b_i^{l,n} \quad \forall l \in \mathbb{L}, \forall t \in \mathbb{T} \quad (5b)$$

Like much UC research, load P_l^l is given as a parameter [25]. Nevertheless, the proposed method can handle load uncertainty in the same way as wind uncertainty.

Two remarks are given below.

1) Load shedding is usually a small probability event. However, making UC decisions with load shedding is meaningful in power systems with a high penetration of renewable generation. To respond the fluctuation of renewable power, strategic load shedding can be regarded as an emergency measure, ensuring the system to continue running with the minimum negative effects.

2) The substation in Fig. 1 interfaces the transmission network with downstream distribution networks. In other words, the n^{th} outlet is the inlet of the n^{th} distribution substation. If necessary, the outlets of distribution substations can be considered in model (5). As a result, N_l will be the total number of distribution outlets. Therefore, the proposed method can unify transmission and distribution systems, optimizing the power balance in a wider spatial range.

D. Network Constraints

The renowned direct-current power flow is employed in the form of power transfer distribution factor (PTDF):

$$\sum_{\forall c \in \mathbb{C}} P_c^c + \sum_{\forall g \in \mathbb{G}} P_g^g + \sum_{\forall w \in \mathbb{W}} P_w^w = \sum_{\forall l \in \mathbb{L}} (P_l^l - \Delta P_l^l) \quad \forall t \in \mathbb{T} \quad (6a)$$

$$\begin{aligned} -P_{\max}^f &\leq \sum_{\forall c \in \mathbb{C}} \pi_{cf} P_c^c + \sum_{\forall g \in \mathbb{G}} \pi_{gf} P_g^g + \sum_{\forall w \in \mathbb{W}} \pi_{wf} P_w^w - \\ &\sum_{\forall l \in \mathbb{L}} \pi_{lf} (P_l^l - \Delta P_l^l) \leq P_{\max}^f \end{aligned} \quad (6b)$$

where f and \mathbb{F} are the index and set of transmission lines, respectively; $\pi_{(c)g}$ is the PTDF from a certain facility to line f ; and P_{\max}^f is the transmission capacity. Equation (6a) is the system-wise power balance condition. Formula (6b) bounds the active power flow on each line f .

E. Cost Function

The costs in UC include the start-up/shun-down cost and fuel cost of thermal units, as well as the penalty of wind en-

ergy curtailment and load shedding. Hence, the cost function is expressed as:

$$\sum_{\forall t \in \mathbb{T}} \left[\sum_{\forall c \in \mathbb{C}} (C_c^{SU} u_t^c + C_c^{SD} v_t^c + C_c^{fixed} b_t^c) \sum_{\forall g \in \mathbb{G}} (C_g^{SU} u_t^g + C_g^{SD} v_t^g + C_g^{fixed} b_t^g) \right] + \sum_{\forall t \in \mathbb{T}} h \left(\sum_{\forall c \in \mathbb{C}} C_c^{Fuel} P_t^c + \sum_{\forall g \in \mathbb{G}} C_g^{Fuel} P_t^g + \sum_{\forall l \in \mathbb{L}} C_l^{Shed} \Delta P_t^l + \sum_{\forall w \in \mathbb{W}} C_w^{Curt} \Delta P_t^w \right) \quad (7)$$

where C_c^{SU} , C_g^{SU} , C_c^{SD} , and C_g^{SD} are the costs of start-up and shut-down, respectively; C_c^{fixed} and C_g^{fixed} are the fixed costs of running for one hour; C_c^{Fuel} and C_g^{Fuel} are the fuel cost coefficients; C_l^{Shed} and C_w^{Curt} are the penalty coefficients of load shedding and wind energy curtailment, respectively, and C_l^{Shed} is much larger; and constant h is 1 hour. One can also use quadratic fuel functions, which can be approximated by linear pieces [26].

III. MRUC WITH DISCRETE LOAD SHEDDING

A. Notations

For brevity, we define some notations. All decision variables are encapsulated into commitment vector $\boldsymbol{\varphi}$, continuous intertemporal (state) vector \mathbf{x}_t , continuous non-state vector \mathbf{y}_t , binary non-state vector \mathbf{z}_t , and random vector $\boldsymbol{\xi}_t$.

$$\boldsymbol{\varphi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_{|\mathbb{T}|}]^T \quad (8a)$$

$$\boldsymbol{\varphi}_{ta} = [b_t^{c_1}, u_t^{c_1}, v_t^{c_1}, b_t^{c_2}, u_t^{c_2}, v_t^{c_2}, \dots, b_t^{c_{|\mathbb{C}|}}, u_t^{c_{|\mathbb{C}|}}, v_t^{c_{|\mathbb{C}|}}] \quad (8b)$$

$$\boldsymbol{\varphi}_{ib} = [b_t^{g_1}, u_t^{g_1}, v_t^{g_1}, b_t^{g_2}, u_t^{g_2}, v_t^{g_2}, \dots, b_t^{g_{|\mathbb{G}|}}, u_t^{g_{|\mathbb{G}|}}, v_t^{g_{|\mathbb{G}|}}] \quad (8c)$$

$$\boldsymbol{\varphi}_t = [\boldsymbol{\varphi}_{ta}, \boldsymbol{\varphi}_{ib}] \quad (8d)$$

$$\mathbf{x}_t = [P_t^{c_1}, P_t^{c_2}, \dots, P_t^{c_{|\mathbb{C}|}}]^T \quad \forall t \in \mathbb{T} \quad (8e)$$

$$\left\{ \begin{array}{l} \mathbf{P}^g = [P_t^{g_1}, P_t^{g_2}, \dots, P_t^{g_{|\mathbb{G}|}}]^T \\ \mathbf{P}^w = [P_t^{w_1}, P_t^{w_2}, \dots, P_t^{w_{|\mathbb{W}|}}]^T \\ \Delta \mathbf{P}^w = [\Delta P_t^{w_1}, \Delta P_t^{w_2}, \dots, \Delta P_t^{w_{|\mathbb{W}|}}]^T \\ \Delta \mathbf{P}^l = [\Delta P_t^{l_1}, \Delta P_t^{l_2}, \dots, \Delta P_t^{l_{|\mathbb{L}|}}]^T \\ \mathbf{y}_t = [\mathbf{P}^g, \mathbf{P}^w, \Delta \mathbf{P}^w, \Delta \mathbf{P}^l]^T \end{array} \right. \quad \forall t \in \mathbb{T} \quad (8f)$$

$$\mathbf{b}_t^{l_1} = [b_t^{l_1,1}, b_t^{l_1,2}, \dots, b_t^{l_1, N_1}^*] \quad (8g)$$

$$\mathbf{b}_t^{l_2} = [b_t^{l_2,1}, b_t^{l_2,2}, \dots, b_t^{l_2, N_2}^*] \quad (8h)$$

$$\mathbf{b}_t^{l_{|\mathbb{L}|}} = [b_t^{l_{|\mathbb{L}|},1}, b_t^{l_{|\mathbb{L}|},2}, \dots, b_t^{l_{|\mathbb{L}|}, N_{|\mathbb{L}|}}^*] \quad (8i)$$

$$\mathbf{z}_t = [z_t^1, z_t^2, \dots, z_t^{|\mathbb{L}|}]^T \quad (8j)$$

$$\boldsymbol{\xi}_t = [\xi_t^1, \xi_t^2, \dots, \xi_t^{|\mathbb{W}|}]^T \quad \forall t \in \mathbb{T} \quad (8k)$$

Then, we define the feasible sets as:

$$\Phi = \{\boldsymbol{\varphi} | (1a) - (1d), (2a) - (2d)\} \quad (9a)$$

$$\Xi_t = \{\Xi_t^1, \Xi_t^2, \dots, \Xi_t^{|\mathbb{W}|}\} \quad \forall t \in \mathbb{T} \quad (9b)$$

$$\Omega_t(\boldsymbol{\varphi}, \boldsymbol{\xi}_t, \mathbf{x}_{t-1}) := \{(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t) | (1e), (1f), (1g), (2e), (3), (5), (6)\} \quad \forall t \in \mathbb{T} \quad (9c)$$

All the constraints in Ω_t are linear, so we can rewrite it as:

$$\Omega_t(\boldsymbol{\varphi}, \boldsymbol{\xi}_t, \mathbf{x}_{t-1}) := \{(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t) | \mathbf{A}_t \mathbf{x}_t + \mathbf{B}_t \mathbf{y}_t + \mathbf{C}_t \mathbf{z}_t \leq \mathbf{b}_t + \mathbf{D}_t \boldsymbol{\varphi} + \mathbf{E}_t \boldsymbol{\xi}_t + \mathbf{F}_t \mathbf{x}_{t-1}\} \quad (9d)$$

where \mathbf{A}_t , \mathbf{B}_t , \mathbf{C}_t , \mathbf{D}_t , \mathbf{E}_t , \mathbf{F}_t and \mathbf{b}_t are the constant matrices extracted from (1e), (1f), (1g), (2e), (3), (5), and (6), respectively.

Besides, the cost function in (7) is expressed as:

$$\mathbf{k}^T \boldsymbol{\varphi} + \sum_{t=1}^T (\mathbf{r}^T \mathbf{x}_t + \mathbf{s}^T \mathbf{y}_t)$$

where $\mathbf{k}^T \boldsymbol{\varphi}$ is the UC cost; $\sum_{t=1}^T (\mathbf{r}^T \mathbf{x}_t + \mathbf{s}^T \mathbf{y}_t)$ is the total dispatch cost; and \mathbf{k} , \mathbf{r} , and \mathbf{s} are the coefficient vectors.

B. MRUC

When randomness is absent, UC is a deterministic optimization problem, which minimizes the cost in (7) subject to the constraints in (1)-(3), (5), and (6). However, the wind power is uncertain. The entire decision-making sequence is:

$$\underbrace{\boldsymbol{\varphi}}_{\text{Day-ahead}}, \underbrace{\boldsymbol{\xi}_1, (\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1), \boldsymbol{\xi}_2, (\mathbf{x}_2, \mathbf{y}_2, \mathbf{z}_2), \dots, \boldsymbol{\xi}_T, (\mathbf{x}_T, \mathbf{y}_T, \mathbf{z}_T)}_{\text{Real-time}} \quad (10)$$

In the day-ahead phase, UC $\boldsymbol{\varphi}$ is determined. During each period t of real-time dispatch, decisions $(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t)$ are made after the wind power $\boldsymbol{\xi}_t$ is observed. Such a sequence is non-anticipated, i. e., dispatch decisions during any period t do not rely on the uncertainty realizations in the future.

To describe (10) and manage the uncertainty in a robust manner, we formulate the MRUC problem as:

$$\min_{\boldsymbol{\varphi} \in \Phi} \mathbf{k}^T \boldsymbol{\varphi} + \left(\max_{\boldsymbol{\xi}_1 \in \Xi_1} \min_{(\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1) \in \Omega_1(\boldsymbol{\varphi}, \boldsymbol{\xi}_1, \mathbf{x}_0)} \mathbf{r}^T \mathbf{x}_1 + \mathbf{s}^T \mathbf{y}_1 \right) + \left(\max_{\boldsymbol{\xi}_2 \in \Xi_2} \min_{(\mathbf{x}_2, \mathbf{y}_2, \mathbf{z}_2) \in \Omega_2(\boldsymbol{\varphi}, \boldsymbol{\xi}_2, \mathbf{x}_1)} \mathbf{r}^T \mathbf{x}_2 + \mathbf{s}^T \mathbf{y}_2 \right) + \dots + \left(\max_{\boldsymbol{\xi}_T \in \Xi_T} \min_{(\mathbf{x}_T, \mathbf{y}_T, \mathbf{z}_T) \in \Omega_T(\boldsymbol{\varphi}, \boldsymbol{\xi}_T, \mathbf{x}_{T-1})} \mathbf{r}^T \mathbf{x}_T + \mathbf{s}^T \mathbf{y}_T \right) \quad (11)$$

IV. SOLUTION METHODOLOGY

The MRUC problem (11) is intractable for two reasons. One is the nested structure coupled by intertemporal variables, and the other is the non-convexity and non-continuity caused by binary recourse. To solve (11), we take three steps: first, propose a partially affine policy and apply it to (11); second, establish an equivalent two-stage robust optimization problem with mixed-integer recourse; third, solve the equivalent problem by an M-nCCG algorithm.

A. Partially Affine Policy

Fully affine policy is proposed in [27], restricting continuous dispatch decisions during period t to be affine functions of previous uncertainty realizations. Define $\boldsymbol{\xi}_{[t]} = [\xi_1, \xi_2, \dots, \xi_t]$,

and the fully affine policy is:

$$[\mathbf{x}_t; \mathbf{y}_t] = \mathbf{U}_t \xi_{[t]} + \mathbf{V}_t \quad \forall t \in \mathbb{T} \quad (12)$$

where \mathbf{U}_t and \mathbf{V}_t are the decision matrices that will be optimized together with $\bar{\varphi}$ in the day-ahead phase. However, there are two main limitations:

1) Fully affine policy is criticized for suboptimality. Especially in the presence of gas-fired units which take non-state real-time actions, the affine relation jeopardizes the real-time flexibility.

2) Affine policy cannot handle binary recourse. The latest method in [28] establishes piecewise constant policies for binary recourse \mathbf{z}_t . However, this method is unfriendly for its complicated partitioning and lifting techniques, as well as the priori and arbitrary choice of a non-linear function.

To address the first limitation, we introduce partially affine policy to unleash non-state variables and impose affine relation only on intertemporal dispatch variables:

$$\mathbf{x}_t = \mathbf{U}_t \xi_{[t]} + \mathbf{V}_t \quad \forall t \in \mathbb{T} \quad (13)$$

During each period t of real-time dispatch, \mathbf{x}_t is determined based on (13) after $\xi_{[t]}$ is observed. How to determine \mathbf{y}_t and \mathbf{z}_t will be discussed later in Section IV-D.

B. Two-stage Equivalence

We define some new notations marked by overlines as: $\bar{\varphi} = [\varphi^\top, \mathbf{U}_1, \mathbf{V}_1, \mathbf{U}_2, \mathbf{V}_2, \dots, \mathbf{U}_{|\mathbb{T}|}, \mathbf{V}_{|\mathbb{T}|}]^\top$, $\bar{\Phi} = \{\bar{\varphi} | \varphi \in \Phi\}$ and $\bar{\mathcal{D}}_t(\bar{\varphi}, \xi_{[t]}) := \{(y_t, z_t) | \mathbf{B}_t y_t + \mathbf{C}_t z_t \leq \bar{\mathbf{b}}_t\}$, where $\bar{\mathbf{b}}_t = \mathbf{b}_t + \mathbf{D}_t \varphi + \mathbf{E}_t \xi_t + \mathbf{F}_t(\mathbf{U}_{t-1} \xi_{[t-1]} + \mathbf{V}_{t-1}) - \mathbf{A}_t(\mathbf{U}_t \xi_{[t]} + \mathbf{V}_t)$.

By applying partially affine policy (13) to problem (11), we derive:

$$\begin{aligned} & \min_{\bar{\varphi} \in \bar{\Phi}} \left(\mathbf{k}^\top \varphi + \sum_{t=1}^T \mathbf{r}^\top \mathbf{V}_t \right) + \left(\max_{\xi_1 \in \Xi_1} \min_{(y_1, z_1) \in \bar{\mathcal{D}}_1(\bar{\varphi}, \xi_{[1]})} \mathbf{r}^\top \mathbf{U}_1 \xi_{[1]} + \mathbf{s}^\top \mathbf{y}_1 \right) + \\ & \left(\max_{\xi_2 \in \Xi_2} \min_{(y_2, z_2) \in \bar{\mathcal{D}}_2(\bar{\varphi}, \xi_{[2]})} \mathbf{r}^\top \mathbf{U}_2 \xi_{[2]} + \mathbf{s}^\top \mathbf{y}_2 \right) + \dots + \\ & \left(\max_{\xi_T \in \Xi_T} \min_{(y_T, z_T) \in \bar{\mathcal{D}}_T(\bar{\varphi}, \xi_{[T]})} \mathbf{r}^\top \mathbf{U}_T \xi_{[T]} + \mathbf{s}^\top \mathbf{y}_T \right) \end{aligned} \quad (14)$$

Proposition 1 Problem (14) is equivalent to the following two-stage robust optimization problem with mixed-integer recourse:

$$\begin{aligned} & \min_{\bar{\varphi} \in \bar{\Phi}} \left(\mathbf{k}^\top \varphi + \sum_{t=1}^T \mathbf{r}^\top \mathbf{V}_t \right) + \\ & \max_{\xi_t \in \Xi_t, \forall t \in \mathbb{T}} \min_{(y_t, z_t) \in \bar{\mathcal{D}}_t(\bar{\varphi}, \xi_{[t]})} \sum_{t=1}^T \left(\mathbf{r}^\top \mathbf{U}_t \xi_{[t]} + \mathbf{s}^\top \mathbf{y}_t \right) \end{aligned} \quad (15)$$

Proof Below is the dispatch during periods $T-1$ and T .

$$\begin{aligned} & \left(\max_{\xi_{T-1} \in \Xi_{T-1}} \min_{(y_{T-1}, z_{T-1}) \in \bar{\mathcal{D}}_{T-1}(\bar{\varphi}, \xi_{[T-1]})} \mathbf{r}^\top \mathbf{U}_{T-1} \xi_{[T-1]} + \mathbf{s}^\top \mathbf{y}_{T-1} \right) + \\ & \left(\max_{\xi_T \in \Xi_T} \min_{(y_T, z_T) \in \bar{\mathcal{D}}_T(\bar{\varphi}, \xi_{[T]})} \mathbf{r}^\top \mathbf{U}_T \xi_{[T]} + \mathbf{s}^\top \mathbf{y}_T \right) \end{aligned} \quad (16)$$

Notice that the minimization over (y_{T-1}, z_{T-1}) is independent from the maximization over ξ_T and the minimization

over (y_T, z_T) , implying that the optimal solution of (16) remains the same if we exchange the inner minimum and maximum operators [29]. Therefore, (16) is equivalent to (17a) and by aggregation further equivalent to (17b).

$$\begin{aligned} & \left(\max_{\xi_{T-1} \in \Xi_{T-1}} \max_{\xi_T \in \Xi_T} \min_{(y_{T-1}, z_{T-1}) \in \bar{\mathcal{D}}_{T-1}(\bar{\varphi}, \xi_{[T-1]})} \mathbf{r}^\top \mathbf{U}_{T-1} \xi_{[T-1]} + \mathbf{s}^\top \mathbf{y}_{T-1} \right) + \\ & \left(\min_{(y_T, z_T) \in \bar{\mathcal{D}}_T(\bar{\varphi}, \xi_{[T]})} \mathbf{r}^\top \mathbf{U}_T \xi_{[T]} + \mathbf{s}^\top \mathbf{y}_T \right) \end{aligned} \quad (17a)$$

$$\max_{\xi_{T-1} \in \Xi_{T-1}} \min_{\xi_T \in \Xi_T} \min_{(y_{T-1}, z_{T-1}) \in \bar{\mathcal{D}}_{T-1}(\bar{\varphi}, \xi_{[T-1]})} \sum_{t=T-1}^T \left(\mathbf{r}^\top \mathbf{U}_t \xi_{[t]} + \mathbf{s}^\top \mathbf{y}_t \right) \quad (17b)$$

In a similar way, combining the dispatch during period $T-2$ with (17) will produce a maximum-minimum equivalence involving periods $T-2$, $T-1$, and T . By a backward induction to the first period, we can derive (15), which completes the proof.

C. M-nCCG Algorithm

According to Proposition 1, solving MRUC problem (14) is equivalent to solving (15), which is a two-stage robust optimization problem with mixed-integer recourse. The mainstream solution method is the nested column-and-constraint generation (CCG) algorithm proposed in [30]. The S-nCCG algorithm is a decomposition method for two-stage robust optimization problem with continuous recourse [31]. In the master problem (MP), the middle maximization is replaced by critical uncertainty scenarios that are strategically identified by feasibility and optimality check sub-problems (SPs). Each SP requires an inner CCG loop.

To concisely present the proposed M-nCCG, we will use the general model of two-stage robust optimization problem with mixed-integer recourse, which is:

$$\begin{cases} \min_{\gamma \in \Gamma} \left(\mathbf{p}^\top \gamma + \max_{\xi \in \Xi} \min_{y, z} \left(f(\gamma, \xi) + \mathbf{q}^\top \mathbf{y} \right) \right) \\ \text{s.t. } \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{z} \leq \mathbf{b}(\gamma, \xi) \quad \mathbf{z} \text{ is binary} \end{cases} \quad (18)$$

where γ is the vector of here-and-now variables; ξ is the uncertainty vector; \mathbf{y} and \mathbf{z} are the wait-and-see variables; \mathbf{A} , \mathbf{B} , \mathbf{b} , \mathbf{p} , and \mathbf{q} are the coefficient matrices; and $f(\gamma, \xi)$ is the cost function. Note that (18) is an independent and pure math problem where the variables have no physical meanings.

1) Outer CCG Loop

The MP in outer CCG loop (MP-outer) is:

$$\begin{cases} \min_{\gamma \in \Gamma, \theta} \left(\mathbf{p}^\top \gamma + \theta \right) \\ \text{s.t. } \mathbf{A}\mathbf{y}_i + \mathbf{B}\mathbf{z}_i \leq \mathbf{b}(\gamma, \xi_i) \quad \forall \xi_i \in \Xi^\dagger, \mathbf{z}_i \text{ is binary} \\ \theta \geq f(\gamma, \xi_i) + \mathbf{q}^\top \mathbf{y}_i \quad \forall \xi_i \in \Xi^\dagger \end{cases} \quad (19)$$

where Ξ^\dagger contains the critical uncertainty scenarios that have been identified so far indexed by i ; and θ is an auxiliary variable.

Given γ^* , there are two SPs. One is the feasibility check SP (SP1-outer), finding those uncertainty scenarios that cause the most severe infeasibility, i.e.,

$$\begin{cases} \max_{\xi \in \Xi} \min_{\mu \geq 0, y, z} \mathbf{1}^T \mu \\ \text{s.t. } \mathbf{A}y + \mathbf{B}z - \mu \leq \mathbf{b}(\gamma^*, \xi) \quad z \text{ is binary} \end{cases} \quad (20)$$

where μ is a slack vector; and $\mathbf{1}$ is an all-one column vector with compatible rows. If the optimal value of SP1-outer is zero, there is always a feasible (y, z) for any $\xi \in \Xi$; otherwise, the optimal ξ^* can cause infeasibility.

The other is the optimality check SP (SP2-outer), finding those uncertainty scenarios that cause the highest costs, i.e.,

$$\begin{cases} \max_{\xi \in \Xi} \min_{y, z} (f(\gamma, \xi) + \mathbf{q}^T y) \\ \text{s.t. } \mathbf{A}y + \mathbf{B}z \leq \mathbf{b}(\gamma, \xi) \quad z \text{ is binary} \end{cases} \quad (21)$$

2) Inner CCG Loop for SP1-outer

SP1-outer and SP2-outer are intractable due to the binary z , so the inner CCG loop is required. Considering SP1-outer first, we reformulate it as:

$$\begin{cases} \max_{\xi \in \Xi} \min_{z \text{ is binary}} \min_{\mu \geq 0, y} \mathbf{1}^T \mu \\ \text{s.t. } \mathbf{A}y - \mu \leq \mathbf{b}(\gamma^*, \xi) - \mathbf{B}z; \lambda \end{cases} \quad (22)$$

where λ after the colon is the vector of dual variables. Dualizing the inner minimization leads to:

$$\begin{cases} \max_{\xi \in \Xi} \min_{z \text{ is binary}} \max_{\lambda} \lambda^T (\mathbf{b}(\gamma^*, \xi) - \mathbf{B}z) \\ \text{s.t. } \mathbf{A}^T \lambda = \mathbf{0} \quad -1 \leq \lambda \leq 0 \end{cases} \quad (23)$$

Note that (23) is a two-stage robust optimization problem with continuous recourse, which can be handled by a CCG loop whose MP is:

$$\begin{cases} \max_{\xi \in \Xi, \theta, \lambda_i} \theta \\ \text{s.t. } \theta \leq \lambda_i^T (\mathbf{b}(\gamma^*, \xi) - \mathbf{B}z_i) \quad \forall z_i \in \mathcal{Z}_1^\dagger \\ \mathbf{A}^T \lambda_i = \mathbf{0} \quad -1 \leq \lambda_i \leq 0 \end{cases} \quad (24)$$

where \mathcal{Z}_1^\dagger is the set of critical binary recourse actions that have been identified so far in the inner loop of SP1-outer. Given ξ^* , the corresponding SP-inner1 is:

$$\begin{cases} \min_{z \text{ is binary}} \max_{\lambda} \lambda^T (\mathbf{b}(\gamma^*, \xi^*) - \mathbf{B}z) \\ \text{s.t. } \mathbf{A}^T \lambda = \mathbf{0} \quad -1 \leq \lambda \leq 0 \end{cases} \quad (25)$$

In (22), the inner minimization is always feasible because there is a slack vector μ . Hence, the feasible set of λ in (23) is non-empty. Since the feasible set of λ is independent from ξ and z , the inner maximization in (23) is also always feasible. Therefore, to solve (22) in the form of (23), we only need an SP-inner1. A feasibility check SP is unnecessary.

3) Inner CCG Loop for SP2-outer

In the similar way, we can reformulate SP2-outer as:

$$\begin{cases} \max_{\xi \in \Xi} \min_{z \text{ is binary}} \min_y (f(\gamma^*, \xi) + \mathbf{q}^T y) \\ \text{s.t. } \mathbf{A}y \leq \mathbf{b}(\gamma^*, \xi) - \mathbf{B}z; \zeta \end{cases} \quad (26)$$

where ζ is the vector of dual variables. Using duality theory to minimize the inner, we have:

$$\begin{cases} \max_{\xi \in \Xi} \min_{z \text{ is binary}} \max_{\zeta} [\zeta^T (\mathbf{b}(\gamma^*, \xi) - \mathbf{B}z) + f(\gamma^*, \xi)] \\ \text{s.t. } \mathbf{A}^T \zeta = \mathbf{q} \quad \zeta \leq 0 \end{cases} \quad (27)$$

Correspondingly, the MP-inner1 is:

$$\begin{cases} \max_{\xi \in \Xi, \theta, \zeta_i} \theta \\ \text{s.t. } \theta \leq \zeta_i^T (\mathbf{b}(\gamma^*, \xi) - \mathbf{B}z_i) + f(\gamma^*, \xi) \quad \forall z_i \in \mathcal{Z}_2^\dagger \\ \mathbf{A}^T \zeta_i = \mathbf{q} \quad \zeta_i \leq \mathbf{0} \end{cases} \quad (28)$$

where \mathcal{Z}_2^\dagger is the set of critical binary recourse actions that have been identified so far in the inner loop of SP2-outer.

The SP-inner2 is:

$$\begin{cases} \min_{z \text{ is binary}} \max_{\zeta} [\zeta^T (\mathbf{b}(\gamma^*, \xi^*) - \mathbf{B}z) + f(\gamma^*, \xi^*)] \\ \text{s.t. } \mathbf{A}^T \zeta = \mathbf{q} \quad \zeta \leq \mathbf{0} \end{cases} \quad (29)$$

For the same reason as in the last subsection, a feasibility check SP is not required here, as long as we can prove that the feasible set of dual vector ζ is non-empty by finding at least one feasible primal vector y . This can be achieved in the MP-outer by completing the feasibility check process before starting the optimality check process.

4) M-nCCG Algorithm

The M-nCCG algorithm is finalized in Algorithm 1. Two remarks are given below.

Algorithm 1: M-nCCG

Initialization: set $\Xi^\dagger = \{\xi^0\}$, $\mathcal{Z}_1^\dagger = \{z^0\}$, and $\mathcal{Z}_2^\dagger = \{z^0\}$ with arbitrary ξ^0 and z^0 ; a tolerance $\varepsilon = 10^{-6}$; a binary indicator $\kappa = 0$ ($\kappa = 0/1$ means feasibility check is incomplete/complete).

Step 1: solve MP-outer and save the optimal γ^* . The optimal value is LB^{outer} .

If $\kappa = 1$, go to Step 3.

Step 2: solve SP-outer by the following substeps.

Step 2-1: solve MP-inner1 and save the optimal solution ξ^* and value θ^* ; set the upper bound $UB^{\text{inner1}} = \theta^*$.

Step 2-2: solve SP-inner1 and save the optimal solution z^* . The optimal value is the lower bound, denoted as LB^{inner1} .

Step 2-3: if $UB^{\text{inner1}} - LB^{\text{inner1}} \geq \varepsilon$, let $\mathcal{Z}_1^\dagger = \mathcal{Z}_1^\dagger \cup \{z^*\}$ and return to Step 2-1; if $UB^{\text{inner1}} - LB^{\text{inner1}} < \varepsilon$ and $LB^{\text{inner1}} = 0$, set $\kappa = 1$ and go to Step 3; else if $UB^{\text{inner1}} - LB^{\text{inner1}} < \varepsilon$ and $LB^{\text{inner1}} > 0$, let $\Xi^\dagger = \Xi^\dagger \cup \{\xi^*\}$ and return to Step 1.

Step 3: solve SP2-outer by the following substeps.

Step 3-1: solve MP-inner2 and save the optimal solution ξ^* and value θ^* ; set the upper bound $UB^{\text{inner2}} = \theta^*$.

Step 3-2: solve SP-inner2 and save the optimal solution z^* . The optimal value is the lower bound, denoted as LB^{inner2} .

Step 3-3: if $UB^{\text{inner2}} - LB^{\text{inner2}} \geq \varepsilon$, let $\mathcal{Z}_2^\dagger = \mathcal{Z}_2^\dagger \cup \{z^*\}$ and return to Step 3-1; if $UB^{\text{inner2}} - LB^{\text{inner2}} < \varepsilon$, set $UB^{\text{outer}} = \mathbf{p}^T \gamma^* + LB^{\text{inner2}}$ and go to Step 4.

Step 4: If $UB^{\text{outer}} - LB^{\text{outer}} < \varepsilon$, terminate and report γ^* ; else, $\Xi^\dagger = \Xi^\dagger \cup \{\xi^*\}$ and return to Step 1.

Output: γ^* .

1) In the S-nCCG algorithm, the identified binary recourse actions (\mathcal{Z}_1^\dagger and \mathcal{Z}_2^\dagger) within each outer iteration are cleared when going to the next iteration. The proposed M-nCCG algorithm keeps them to give the next iteration a warm start, facilitating the convergence of inner loops.

A flowchart is given in Fig. 2, where the red boxes show the modification, which improves the performance by exploiting all the information in previous iterations.

2) Both SP-inner1 and SP-inner2 are bi-level min-max problems. For each of them, with a fixed outer level, the inner level is a linear program. Therefore, dualizing the inner level will lead to an MILP, which can be efficiently solved by commercial solvers like CPLEX.

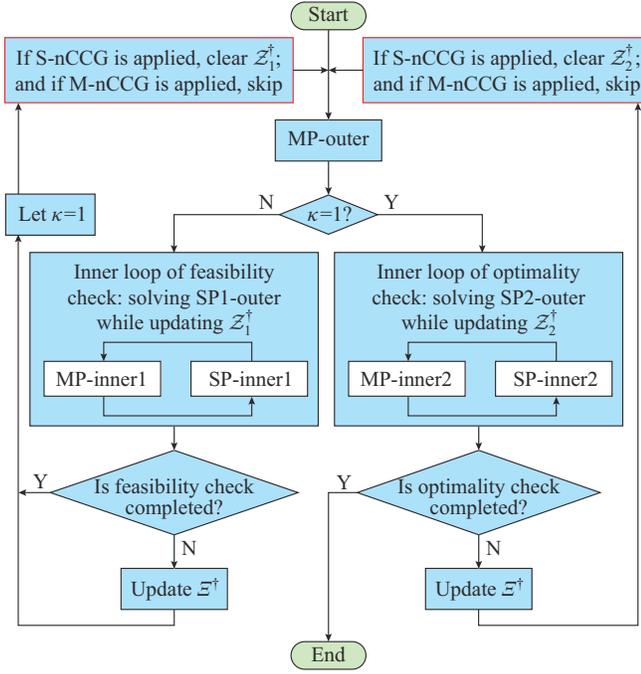


Fig. 2. Algorithm flowchart.

D. Implementation Procedure

Ahead of the day, problem (14) in the form of (15) is solved by Algorithm 1. The optimal solution $\bar{\varphi}^*$ includes UC strategy φ^* and partially affine policy (U_t^*, V_t^*) , $\forall t \in \mathbb{T}$, and the former is applied and the latter is passed to real-time dispatch.

During period t of real-time dispatch, we need to find x_t^* , y_t^* , and z_t^* . By partially affine policy, we have $x_t^* = U_t^* \xi_{[t]} + V_t^*$. Besides, we notice in (14) that all the optimization problems in the real-time dispatch are decoupled over periods. Therefore, we compute y_t^* and z_t^* by:

$$(y_t^*, z_t^*) = \arg \min_{y_t, z_t} (r^T U_t^* \xi_{[t]} + s^T y_t) \quad (30)$$

V. CASE STUDIES

The proposed method is tested on the modified IEEE 118-bus system with 186 transmission lines. This system consists of 15 coal-fired units (2800 MW), 12 gas-fired units (1400 MW), and 6 wind farms (3200 MW). The system-wide load curve is drawn in Fig. 3(a), where the peak load is 4434.2 MW. The forecast sequence $\{\xi_{[t]}^{w,0}\}$, $t = 1:T$ is drawn in Fig. 3(b), where the wind power is normalized with the installed capacity as base value. The uncertainty set is built with $\delta^w = 0.3$ p.u.. We set $C_l^{Shed} = 1000$ \$/MWh and $C_w^{Curt} = 15$ \$/MWh.

We choose 8 buses (substations) as candidates for load shedding. According to model (17), each of them has N_l outlets in total, and N_l^* outlets can be tripped. The load elsewhere must be satisfied. Table I presents the load shedding settings. The load at other buses must be satisfied. All the parameters and data used in this section are provided in an online archive [32] for interested readers. Optimization problems are solved by CPLEX 12.6 on a laptop with Intel i5-8250U CPU and 8 GB memory.

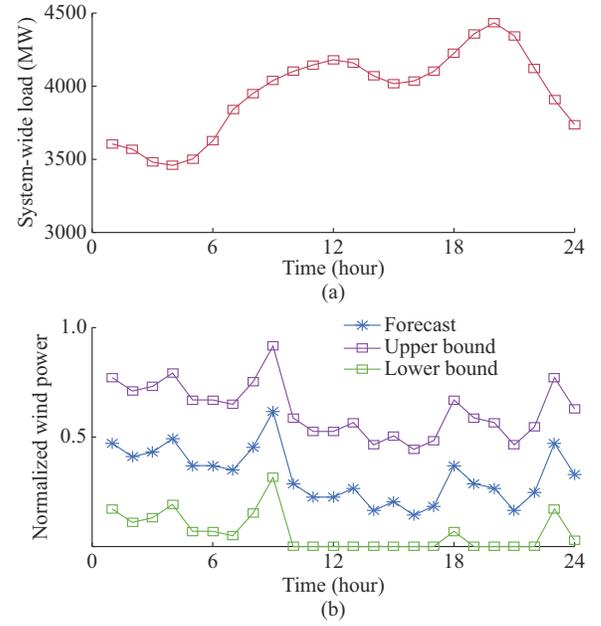


Fig. 3. Load and wind data. (a) System-wide load. (b) Forecast sequence.

 TABLE I
LOAD SHEDDING SETTINGS

Index	Bus	Peak load (MW)	N_l	N_l^*
1	15	90	2	1
2	42	96	2	1
3	49	87	2	2
4	54	113	3	2
5	59	277	6	4
6	80	130	3	2
7	90	163	4	3
8	116	184	4	3

A. UC Results

By Algorithm 1, the MRUC strategy is obtained, including the on-off status of coal-fired and gas-fired units. The UC cost, i.e., the first term in (7), is $\$1.942 \times 10^5$.

All the coal-fired units are on over the entire horizon. One reason is that coal is cheaper than natural gas, so the coal-fired units keep running to serve the base load. The other reason is that coal-fired units are less flexible than gas-fired units and cannot switch the on-off status frequently.

Table II exhibits the UC of gas-fired units, where 1 means the unit is on and 0 means the unit is off.

1) Most gas-fired units are committed from hour 10 to hour 22, since the load within this time interval is relatively high. Meanwhile, units 8 and 9 keep running all day long to help serve the base load. Besides, unit 12 is deployed only at noon and in the evening, providing power support in peak hours.

2) The flexibility of gas-fired units can be observed from their quick status switches, e.g., unit 7 is turned on in hour 17, turned off soon in hour 18, and turned on again in hour 19. Such a capability enables to respond to the fast fluctuation of wind power.

TABLE II
UC OF GAS-FIRED UNITS

Period (hour)	UC											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	1	1	0	0	0
2	0	0	0	1	0	0	0	1	1	0	0	0
3	0	0	0	0	0	0	0	1	1	0	0	0
4	0	0	0	0	0	0	0	1	1	0	0	0
5	0	0	0	0	1	0	0	1	1	0	0	0
6	0	1	0	0	1	0	0	1	1	0	0	0
7	0	1	0	0	1	0	0	1	1	1	0	0
8	0	0	0	0	1	0	0	1	1	1	0	0
9	0	1	0	0	1	0	0	1	1	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	0
11	1	1	1	1	1	1	1	1	1	1	1	0
12	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	0	1	1	1	1	1	1	1	0
15	1	1	1	0	1	1	0	1	1	1	1	0
16	1	1	1	0	1	1	0	1	1	1	1	0
17	1	1	1	1	1	1	1	1	1	1	1	0
18	1	1	0	1	1	1	0	1	1	1	1	0
19	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	0
23	0	1	0	0	1	0	0	1	1	0	0	0
24	0	1	0	0	1	0	0	1	1	1	0	0

These results show that gas-fired units are important for power supply, especially in terms of flexibility.

B. Dispatch Results in Worst-case Scenario

Besides UC strategy, Algorithm 1 also identifies the worst-case scenario, i.e., the one in \mathcal{E}^\dagger that causes the highest dispatch cost. The worst-case scenario here is exactly the lower bound in Fig. 3(b), where the wind energy is quite scarce.

The dispatch results with regard to this scenario are discussed below.

The coal fuel cost is $\$9.602 \times 10^5$ in total, and the gas fuel cost is $\$1.511 \times 10^6$. There is no penalty of wind energy curtailment because wind energy is poor in the worst-case scenario. The total penalty of load shedding is $\$5.581 \times 10^5$, so the amount of load shedding is 558.1 MWh.

Figure 4 shows the system-wide generation power curves of coal-fired units and gas-fired units. The output power of coal-fired units is stable and high. Since coal is cheaper than natural gas, the coal-fired units produce the majority of energy. Meanwhile, gas-fired units play a significant role in managing load peak and valley. The generation curve shares a similar trend with the load curve in Fig. 3(a). Around hour 4, the load is the lowest, so is the total output of gas-fired units. There exist two load peaks around hour 12 and hour 20 when the gas-fired units respond to enhance the output level.

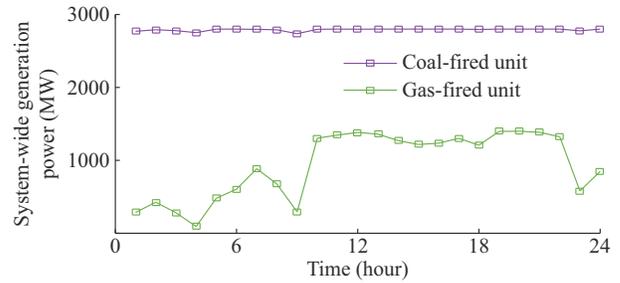


Fig. 4. System-wide generation power curves of coal-fired units and gas-fired units.

Furthermore, Fig. 5 shows the load shedding over time. Load shedding happens in hours 19, 20, and 21 and the total amount is 558.1 MWh. The total load is 9.502×10^4 MWh over the entire horizon, so the load shedding ratio is about 0.6%.

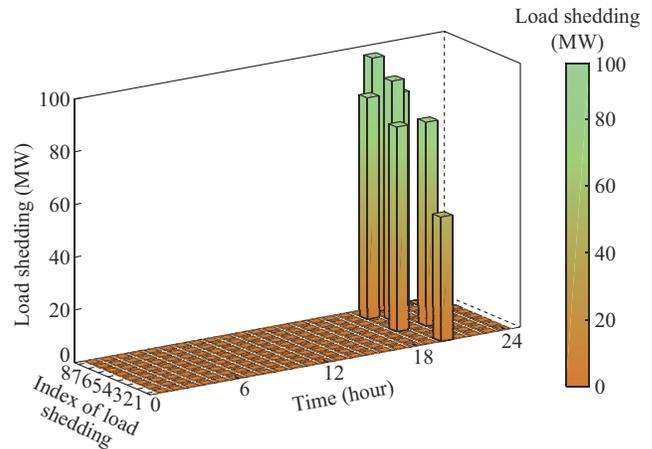


Fig. 5. Load shedding.

C. Sensitivity Analysis

The robustness of the proposed method is controlled by parameter δ_t^w , which can be regarded as the maximum forecast error of wind power. Therefore, this subsection investigates the impact of δ_t^w on UC and dispatch in the worst-case scenario.

The results of sensitivity analysis about δ_t^w are gathered in Table III. With the increase of δ_t^w , the wind energy in the worst-case scenario becomes less and less. Meanwhile, the possible fluctuation of wind power within the uncertainty set becomes more severe. Consequently, the UC cost goes up since the units are deployed for more time and with more status switches. The fuel consumption also increases to compensate the scarcity of wind energy, leading to a rising fuel cost. Especially, the gas fuel cost rises up by $\$0.9150 \times 10^6$ when δ_t^w changes from 0.1 to 0.5.

Furthermore, the wind curtailment is zero when $\delta_t^w \geq 0.3$. A larger δ_t^w implies that the wind energy in the worst-case scenario is rarer and thus can be fully utilized. On the contrary, load shedding is zero when $\delta_t^w = 0.1$, since the wind energy in the worst-case scenario is relatively rich.

TABLE III
SENSITIVITY ANALYSIS ABOUT δ_i^w

δ_i^w (p.u.)	UC cost (10 ⁵ \$)	Coal fuel cost (10 ⁵ \$)	Gas fuel cost (10 ⁶ \$)	Wind curtailment (MWh)	Load shedding (MWh)
0.1	1.358	9.214	0.880	141.90	0
0.2	1.699	9.504	1.233	84.73	204.2
0.3	1.942	9.602	1.511	0	558.1
0.4	2.044	9.627	1.707	0	595.8
0.5	2.101	9.635	1.795	0	595.8

D. Comparison with Fully Affine Policy

We propose the partially affine policy in (13) and the fully affine policy in (12) with $\delta_i^w=0.3$. For both of them, binary recourse is optimized during each period.

We compute the sample average of real-time dispatch cost instead of the results regarding the worst-case scenario. To this end, we collect 200 per-unit wind samples (trajectories) from a real wind farm in Ningxia Province, China [32]. Some samples fall within the uncertainty set Ξ and some not. For those within the uncertainty set, the real-time dispatch is always feasible. For those outside the uncertainty set, the infeasibility may occur.

Table IV summarizes the comparison results. The UC costs corresponding to these two policies are close.

TABLE IV
COMPARISONS BETWEEN PARTIALLY AND FULLY AFFINE POLICIES

Item	UC cost (10 ⁵ \$)	Average dispatch cost for 65 samples inside Ξ (10 ⁶ \$)	Number of feasible samples for 135 samples outside Ξ	Average dispatch cost for 135 samples outside Ξ (10 ⁶ \$)
Partially affine policy	1.666	1.593	86	1.541
Fully affine policy	1.682	1.827	69	1.800

In addition, we find that:

1) The proposed partially affine policy realizes a lower dispatch cost. According to (13), partially affine policy maintains the flexibility of non-state variables, which physically mean the outputs of gas-fired units; such flexibility can help find better real-time actions. With samples inside/outside the uncertainty, the dispatch cost of fully affine policy is significantly higher by 14.7%/16.8% than that of partially affine policy. Therefore, the advantage of the proposed method over optimality is verified.

2) The flexibility maintained by partially affine policy favors the robustness. Among 135 samples outside the uncertainty, the real-time dispatch by partially affine policy is feasible with 86 samples, while the number becomes 69 by fully affine policy. The reason is that unleashing the flexibility of gas-fired units enlarges the feasible region of dispatch actions, so the resulting system is more robust to uncertain wind power. Therefore, the advantage of the proposed method over dispatch robustness is verified.

E. Comparison with S-nCCG: Computation Performance

According to Section IV-C, in each outer iteration, both the proposed M-nCCG algorithm and the S-nCCG algorithm solve the same MP (MP-outer). The difference is how they solve SP1-outer and SP2-outer by inner loops. After the i^{th} outer iteration, M-nCCG algorithm takes the identified binary recourse actions (\mathcal{Z}_1^\dagger and \mathcal{Z}_2^\dagger) to the $(i+1)^{\text{th}}$ outer iteration. The inner loops are accelerated by these actions as a warm start. S-nCCG renews \mathcal{Z}_1^\dagger and \mathcal{Z}_2^\dagger when stepping into the $(i+1)^{\text{th}}$ outer iteration.

In Fig. 6, M-nCCG and S-nCCG share the same outer convergence curve, which indicates that the algorithm terminates after 6 outer iterations. Taking the third outer iteration as an example, M-nCCG, which maintains the previously identified scenarios, takes only 2 iterations to terminate the inner loop of optimality check SP while S-nCCG takes 5 iterations. Therefore, M-nCCG can accelerate the convergence.

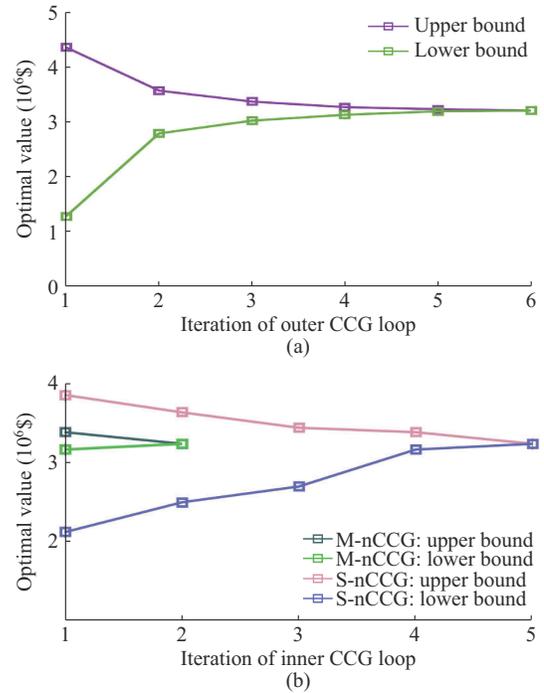


Fig. 6. Iteration of outer CCG loop and inner CCG loop. (a) Iteration of outer CCG loop. (b) Iteration of inner CCG loop.

The time to solve MP-outer is 103.6 s. To solve SP1-outer and SP2-outer, M-nCCG consumes 56.8 s while S-nCCG consumes 116.6 s. Therefore, M-nCCG reduces the total computation time by 27.2% compared with S-nCCG, which verifies the advantage over computation performance.

F. Comparison with State-of-the-art Methods

We compare the proposed method with two state-of-the-art methods mentioned in the literature review. The one in [19], [20] denoted by M-1 uses the affine policy for continuous recourse and the piecewise constant policy for binary recourse. The piecewise constant policy relies on a pre-defined partitioning function. M-1 resorts to an MILP. The other one denoted by M-2 is reported in [21], which optimizes the partitioning function. M-2 employs a customized CCG algo-

rithm. The total cost and computation time are compared, as shown in Table V. The total cost includes the UC cost and the sample average of dispatch cost.

TABLE V
COMPARISONS WITH STATE-OF-THE-ART METHODS

Method	Total cost (10^6 \$)	Computation time (s)
Proposed	1.730	160.4
M-1	1.941	148.6
M-2	1.755	210.2

Regarding dispatch economy, M-1 achieves the highest total cost $\$1.941 \times 10^6$. The total cost of M-2 is 9.58% lower since it optimizes the partitioning function instead of using a pre-defined one like M-1. The total cost of the proposed method is the lowest. The reason is the proposed method does not impose any decision structure for non-state continuous and binary dispatch variables, whose flexibility is exploited and retained.

Regarding computation time, M-1 is the fastest and the proposed method has a close efficiency to M-1. M-2 consumes much more time than M-1 by 41.5%, since it entails solving a large-scale MP, which is established to improve the optimality. In summary, the proposed method balances well the optimality and efficiency.

VI. CONCLUSION

This paper investigates the UC problem considering discrete load shedding, which is formulated as a multi-stage adaptive robust optimization problem with mixed-integer recourse. Partially affine policy and two-stage reformulation address this problem in a tractable way. We conclude that the proposed partially affine policy outperforms fully affine policy in terms of dispatch economy and robustness, the MnCCG algorithm can accelerate the convergence by warming up the inner loops, and the framework in this paper balances the optimality and efficiency compared with state-of-the-art methods.

The main weakness of the proposed method is that it can only handle linear problems. The exact power flow model is nonlinear and nonconvex. Besides, the power supply is affected by both balance and stability. The current formulation does not incorporate stability constraints. Future work will try to address these issues.

REFERENCES

- [1] D. Bertsimas, E. Litvinov, X. Sun *et al.*, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52-63, Feb. 2013.
- [2] B. Hu and L. Wu, "Robust SCUC considering continuous/discrete uncertainties and quick-start units: a two-stage robust optimization with mixed-integer recourse," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1407-1419, Mar. 2016.
- [3] J. Zhang, Z. Chen, N. Zhang *et al.*, "Frequency-constrained unit commitments with linear rules extracted from simulation results considering regulations from battery storage," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 4, pp. 1041-1052, Jul. 2023.
- [4] A. Lorca, X. Sun, E. Litvinov *et al.*, "Multistage adaptive robust optimization for the unit commitment problem," *Operations Research*, vol. 64, no. 1, pp. 32-51, Jan. 2016.
- [5] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, 2nd ed. New York, USA: Springer, 2011.
- [6] Z. Guo, W. Wei, L. Chen *et al.*, "Distribution system operation with renewables and energy storage: a linear programming based multistage robust feasibility approach," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 738-749, Jan. 2022.
- [7] A. Lorca and X. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1678-1688, May 2017.
- [8] Y. Zhou, M. Shahidehpour, Z. Wei *et al.*, "Multistage robust look-ahead unit commitment with probabilistic forecasting in multi-carrier energy systems," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 70-82, Jan. 2021.
- [9] N. G. Cobos, J. M. Arroyo, N. Alguacil *et al.*, "Robust energy and reserve scheduling considering bulk energy storage units and wind uncertainty," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5206-5216, Sept. 2018.
- [10] Y. Zhou, Q. Zhai, and L. Wu, "Multistage transmission-constrained unit commitment with renewable energy and energy storage: implicit and explicit decision methods," *IEEE Transactions Sustainable Energy*, vol. 12, no. 2, pp. 1032-1043, Apr. 2021.
- [11] X. Zheng, M. E. Khodayar, J. Wang *et al.*, "Distributionally robust multistage dispatch with discrete recourse of energy storage systems," *IEEE Transactions on Power Systems*, doi: 10.1109/TPWRS.2024.3369664
- [12] A. Georghiou, A. Tsoukalas, and W. Wiesemann, "Robust dual dynamic programming," *Operations Research*, vol. 67, no. 3, pp. 813-830, May 2019.
- [13] Y. Shi, S. Dong, C. Guo *et al.*, "Enhancing the flexibility of storage integrated power system by multi-stage robust dispatch," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 2314-2322, May 2021.
- [14] E. Nasrolahpour, J. Kazempour, H. Zareipour *et al.*, "A bilevel model for participation of a storage system in energy and reserve markets," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 582-598, Apr. 2018.
- [15] X. Xu, H. Zhang, C. Li *et al.*, "Optimization of the event-driven emergency load-shedding considering transient security and stability constraints," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2581-2592, Jul. 2017.
- [16] H. Qiu, W. Gu, W. Sheng *et al.*, "Resilience-oriented multistage scheduling for power grids considering nonanticipativity under tropical cyclones," *IEEE Transactions on Power Systems*, vol. 38, no. 4, pp. 3254-3267, Jul. 2023.
- [17] D. Bertsimas and C. Caramanis, "Finite adaptability in multistage linear optimization," *IEEE Transactions on Automatic Control*, vol. 55, no. 12, pp. 2751-2767, Dec. 2010.
- [18] A. Subramanyam, C. E. Gounaris, and W. Wiesemann, "K-adaptability in two-stage mixed-integer robust optimization," *Mathematical Programming Computation*, vol. 12, pp. 193-224, Dec. 2020.
- [19] D. Bertsimas and A. Georghiou, "Binary decision rules for multistage adaptive mixed-integer optimization," *Mathematical Programming*, vol. 167, pp. 395-433, Jan. 2018.
- [20] H. Qiu and H. B. Gooi, "A unified MILP solution framework for adaptive robust scheduling problems with mixed-integer recourse objective," *IEEE Transactions on Power Systems*, vol. 38, no. 1, pp. 952-955, Jan. 2023.
- [21] Z. Zhong, N. Fan, and L. Wu, "Multistage robust optimization for the day-ahead scheduling of hybrid thermal-hydro-wind-solar systems," *Journal of Global Optimization*, vol. 27, pp. 1-36, Nov. 2023.
- [22] Q. Chen, P. Zou, C. Wu *et al.*, "A Nash-Cournot approach to assessing flexible ramping products," *Applied Energy*, vol. 206, pp. 42-50, May 2017.
- [23] C. Wang, R. Gao, W. Wei *et al.*, "Risk-based distributionally robust optimal gas-power flow with Wasserstein distance," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2190-2204, May 2019.
- [24] A. Belderbos, T. Valkaert, K. Bruninx *et al.*, "Facilitating renewables and power-to-gas via integrated electrical power-gas system scheduling," *Applied Energy*, vol. 275, p. 115082, Jan. 2020.
- [25] P. Xiong and C. Singh, "A distributional interpretation of uncertainty sets in unit commitment under uncertain wind power," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 149-157, Jan. 2019.
- [26] C. Zhao and Y. Guan, "Unified stochastic and robust unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3353-3361, Aug. 2013.
- [27] A. Ben-Tal, A. Goryashko, E. Guslitzer *et al.*, "Adjustable robust solutions of uncertain linear programs," *Mathematical Programming*, vol. 99, pp. 351-376, Jun. 2004.

- [28] D. Bertsimas and A. Georghiou, "Binary decision rules for multistage adaptive mixed-integer optimization," *Mathematical Programming*, vol. 167, pp. 395-433, Jul. 2018.
- [29] C. Ning and F. You, "A transformation-proximal bundle algorithm for multistage adaptive robust optimization and application to constrained robust optimal control," *Automatica*, vol. 113, p. 108802, Dec. 2020.
- [30] L. Zhao and B. Zeng. (2012, Jan.). An exact algorithm for two-stage robust optimization with mixed integer recourse problems. [Online]. Available: <https://optimization-online.org/wpcontent/uploads/2012/01/3310.pdf>.
- [31] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, pp. 457-461, Sept. 2013.
- [32] Z. Guo. (2023, Feb.). Data MRUC integer recourse. [Online]. Available: <https://github.com/ZhongjieGuo/Papers>.

Zhongjie Guo received the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2023. He is currently a Postdoctoral Researcher with the School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China. His research interests include energy storage, renewable power generation, and power system operation.

Jiayu Bai received the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2023. She is currently with the State Grid Sichuan Electric Power Research Institute, Chengdu, China. Her research interests include compressed air energy storage, power system stability, and power system operation.

Wei Wei received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2008 and 2013, respectively. He was a Postdoctoral Research Associate with Tsinghua University from 2013 to 2015. He was a Visiting Scholar with Cornell University, Ithaca, USA, in 2014, and a Visiting Scholar with Harvard University, Cambridge, USA, in 2015. He is currently an Associate Professor with Tsinghua University. His research interests include applied optimization and energy system economics.

Haifeng Qiu received the B.Eng. degree from Nanjing Normal University, Nanjing, China, in 2015, and the Ph.D. degree from Southeast University, Nanjing, China, in 2021, both in electrical engineering. From 2019 to 2020, he was a Visiting Scholar in the Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, USA. He is currently a Postdoctoral Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include energy system planning, operation, and market trading.

Weihao Hu received the B.Eng. and M.Sc. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2007, respectively, and the Ph.D. degree from Aalborg University, Aalborg, Denmark, in 2012. He is currently a Full Professor and the Director of the Institute of Smart Power and Energy Systems, University of Electronic Science and Technology of China, Chengdu, China. He was an Associate Professor and the Vice Program Leader of Wind Power System Research Program, Department of Energy Technology, Aalborg University. His research interests include artificial intelligence in modern power system and renewable power generation.