Partition-combine Uncertainty Set for Robust Unit Commitment

Zhang, Menglin; Fang, Jiakun; Ai, Xiaomeng; Zhou, Bo; Yao, Wei; Wu, Qiuwei; Wen, Jinyu

Link back to DTU Orbit

Partition-combine Uncertainty Set for Robust Unit Commitment

Menglin Zhang, Jiakun Fang, *Senior Member, IEEE,* Xiaomeng Ai, *Member, IEEE,* Bo Zhou,
Wei Yao, *Senior Member, IEEE,* Qiuwei Wu, *Senior Member, IEEE,* and Jinyu Wen, *Member, IEEE*

Abstract—The construction of the uncertainty set is a key aspect of robust unit commitment (RUC) for the power systems with wind power integrated. The existing uncertainty sets based on the convex hull of historical data are limited to the specific shape of the historical data set or with high computational burden. In this letter, we propose the partition-combine method to build the minimal uncertainty set with the irregularly-distributed historical data. The partition of the box set is used to reduce the size of the uncertainty so that the set can properly adapt the data provided and is no longer limited to a specific shape. The combination of non-empty subsets is used to formulate the model of uncertainty set instead of the extreme points. The scale of uncertainty variables is reduced with quick identification of inner subsets which contain no extreme points. The simulation results validated the enhanced performances of the proposed method from both the conservativeness and computational burden perspectives.

Index Terms—Wind power, uncertainty set, convex hull, extreme points, robust unit commitment.

I. INTRODUCTION

ROBUST optimization (RO) is one of the most popular methods to make the day-ahead decision for a power system with fluctuating renewables, and some market has started to explore the possibility of using RO to incorporate uncertainties in the market clearing processes [1]. Robust unit commitment (RUC) gives priority to the reliability of power system operation while pursuing the economy of the operation. However, conservative decisions could be made if the decision is unnecessarily robust to those renewable cases with zero probabilities. Hence, recent researches have been mainly devoted to building the smallest uncertainty set covering all the possible scenarios.

Conventionally, RUC mainly adopt budget constraints to adjust the size of the uncertainty set [2] [3] [4]. However, the size adjustment is not based on historical data. If the uncertainty set is shrunk too much, the robustness cannot be guaranteed, otherwise conservative decision may be made. Refs. [5] and [6] propose the multi-band uncertainty set and the uncertainty set variant respectively, which are still based on budget constraints. Therefore, the dilemma for robustness and conservativeness remains. Ref. [7] utilizes the union of several separated regions instead of a single convex set to reduce the

M. Zhang, J. Fang, X. Ai, B. Zhou, W. Yao and J. Wen are with State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China.(Corresponding author: Xiaomeng Ai, Email: xiaomengai@hust.edu.cn)

Qiuwei Wu is with the Center for Electric Power and Energy (CEE), Department of Electrical Engineering, Technical University of Denmark (DTU), 2800 Kgs. Lyngby, Denmark.

conservativeness, which is only applicable for cases where the range of isolated regions are all known and fixed.

Some other works construct data-driven uncertainty sets. Refs. [8] and [9] utilize sufficient historical data to approach the precise probability distribution or determines the feasible dispatch range of uncertainty so as to guarantee the robustness of the decision. But these two works do not belong to the field of RO optimization.

Generally, the data-driven uncertainty sets in RO field can be classified into two categories. The first group models the correlation constraints in uncertainty sets to reduce the conservativeness, in which the estimated parameters have a great influence on the description accuracy of the uncertainty set [10] [11].

The second group utilizes the convex hull of historical data to construct the uncertainty set. Refs. [12] and [13] utilize the ellipse and diamond convex hulls respectively to cover the historical data in each period. But those two works only adapt to specifically shaped data set with high correlations. Ref. [14] construct a high-dimensional convex hull which considers temporal and spatial correlation simultaneously. But a large number of vertices are needed for the precise description of uncertainty, which in turn increases the computational burden.

In this letter, we aim to construct the minimal uncertainty set that can adapt to the irregularly distributed historical data while avoiding the computational burden in two-stage RUC. The main contributions are summarized as follows:

- A partition-combine method is proposed to build the uncertainty set. It reduces the conservativeness by partitioning the set and cutting off areas without historical data while maintaining the same level of robustness of the decision. The shape of the set can properly adapt to the data provided.
- The partition-combine uncertainty set is formulated with a combination of subsets, which can avoid finding the extreme points of the smallest convex hull. A fast identification method for the subsets that may contain extreme points is proposed to reduce the computational burden.

II. TWO-STAGE RUC FORMULATION

The general compact formulation of the two-stage RUC can be expressed as follows:

$$
\begin{aligned}
\min_{\mathbf{X}, \mathbf{Y}, \widetilde{\mathbf{w}}} \quad & \{f(\mathbf{x}) + \max_{\widetilde{\mathbf{w}}} \min_{\mathbf{Y}} \mathbf{c}^{\mathrm{T}} \mathbf{Y}\} \\
\text{s.t.} \quad & g(\mathbf{X}) \leq \mathbf{b} && \text{(a)} \\
& \mathbf{HX} + \mathbf{NY} \leq \mathbf{h} - \mathbf{E}\widetilde{w} && \text{(b)} \\
& \widetilde{w} \in U && \text{(c)}
\end{aligned} \quad (1)
$$

where $\mathbf{X}$ is a vector consisting of the first-stage decision variables, including the UC plan and the economic dispatch in the base case; $\mathbf{Y}$ is the second-stage decision vector, including the reserve deployment, wind power curtailment, and load

shedding; $\widetilde{\mathbf{w}}$ is the uncertainty parameter of wind power, belonging to the uncertainty set U. $\mathbf{H}$ and $\mathbf{N}$ are the coefficient matrices; $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{h}$ are the coefficient vectors. (1.a) and (1.b) are security constraints of day-ahead and intra-day UC, respectively.

The two-stage RUC is usually decomposed into a master problem (MP) and a sub-problem (SP), where MP is a mixed-integer linear programming. SP is a max-min optimization which can be solved by methods such as big-M and convex hull relaxation method [15]. By solving the MP and SP iteratively with Benders or column and constraint generation (C&CG) algorithms, the final decision can be obtained when the gap between the MP and SP is less than the predefined tolerance.

As can be seen from (1.c), the uncertainty set U is an essential part of the two-stage RUC, and it, therefore, has a direct and great influence on the decision results. The main objective of this work is to reduce the conservativeness of the uncertainty set while maintaining the same level of robustness of the decision in the two-stage RUC.

III. Uncertainty Set Construction

The construction of the partition-combine uncertainty set can be divided into three steps. Firstly, the box set enclosing all the historical data is partitioned into subsets with the same size, with the removal of empty subsets followed. Secondly, the remaining subsets with new derived boundaries are combined to reformulate the new uncertainty set, namely the partition-combine uncertainty set. Lastly, the inner subsets are identified to reduce the scale of uncertainty variables.

A. Set Partition and Combination

In this letter, only the spatial correlation between different wind farms are considered. Therefore, the uncertainty in different periods can be treated independently. Take the uncertainty set for two wind farms as an example. In period t, the output of the two wind farms is represented by the 2-dimensional vector $\widetilde{\mathbf{w}}_t$. The traditional box-like set is represented by the region surrounded by the blue rectangle in Fig. 1 (a). And its formulation can be described as follows:

$$U = \{\widetilde{\mathbf{w}}_t | \underline{\mathbf{W}}_t \leq \widetilde{\mathbf{w}}_t \leq \overline{\mathbf{W}}_t\} \tag{2}$$

where $\overline{\mathbf{W}}_t$ and $\underline{\mathbf{W}}_t$ are the upper and lower boundaries of two wind farms, respectively.

As the worst-case scenarios for RO always located at the extreme points of the uncertainty set [2] [3], the uncertainty parameter are often represented by the extreme points of the set, which is more tractable for computation. Then the uncertainty set in (2) can be replaced by

$$U = \{\widetilde{\mathbf{w}}_t^{ex} | \widetilde{\mathbf{w}}_t^{ex} = \underline{\mathbf{W}}_t \mathbf{Z}_t^- + \overline{\mathbf{W}}_t \mathbf{Z}_t^+, \mathbf{Z}_t^+ + \mathbf{Z}_t^- \leq 1\} \tag{3}$$

where $\widetilde{\mathbf{w}}_t^{ex}$ are the extreme points of the set, and for the box-like set, its extreme points are just the vertices; $\mathbf{Z}_t^+$ and $\mathbf{Z}_t^-$ are binary variables to decide which extreme point is reached.

In Fig. 1 (a), the box-like uncertainty set contains some blank region without historical data, which may result in a conservative decision.

(a) Box-like uncertainty set   (b) Partition the box set into equal subsets and remove empty subsets

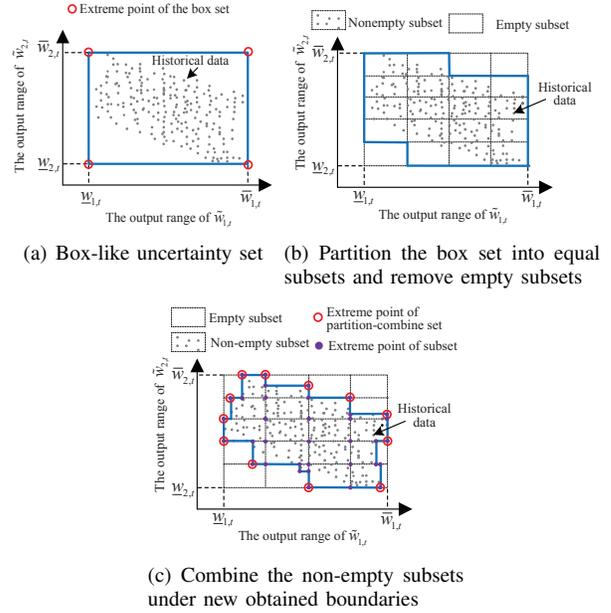(c) Combine the non-empty subsets under new obtained boundaries

Fig. 1. Box-like and partition-combine uncertainty sets

Fig. 1 (b) and Fig. 1 (c) show how to construct the partition-combine uncertainty set. Firstly, we partition the original box set into subsets with the same size according to the predefined number of divisions in each dimension, as shown in Fig.1 (b). All the subsets can be divided into two categories, the empty subsets, and non-empty subsets, depending on whether they contain historical data. As the empty subsets may result in a conservative solution, they are identified and removed to reduce the conservativeness.

Secondly, the new boundaries of each non-empty subsets are redefined according to the range of its covered historical data. Such a measure can further reduce the size of uncertainty set to decrease the conservativeness. After that, these remaining subsets with new boundaries are combined. Then, the partition-combine uncertainty set is obtained as shown in Fig. 1 (c).

Similar to (3), we utilize the extreme points of each subset to formulate the partition-combine uncertainty set U_{new}, which can be expressed as

$$U_{new} = \left\{ \widetilde{\mathbf{w}}_{t,new} \left| \begin{array}{ll} \widetilde{\mathbf{w}}_{t,new} = \sum_{i=1}^{N} \left(\underline{\mathbf{w}}_{t,i} \mathbf{Z}_{t,i}^- + \overline{\mathbf{w}}_{t,i} \mathbf{Z}_{t,i}^+ \right) X_{t,i} & \text{(a)} \\ \mathbf{Z}_{t,i}^- + \mathbf{Z}_{t,i}^+ \leq 1, \mathbf{Z}_{t,i}^-, \mathbf{Z}_{t,i}^+ \in \{\mathbf{0}, \mathbf{1}\} & \text{(b)} \\ \sum_{i=1}^{N} X_{t,i} = 1, X_{t,i} \in \{0, 1\} & \text{(c)} \end{array} \right. \right\} \tag{4}$$

where $\widetilde{\mathbf{w}}_{t,new}$ is the uncertainty parameter of the partition-combine uncertainty set in period t, and it can be any purple point in Fig. 1 (c); i is the index for subsets; N is the total number of subsets; $\overline{\mathbf{w}}_{t,i}/\underline{\mathbf{w}}_{t,i}$ are the upper and lower boundaries of the ith subset, respectively; $X_{t,i}$ is a binary variable, with $X_{t,i} = 1$ representing that $\widetilde{\mathbf{w}}_{t,new}$ is taken at one of the extreme points of the ith subset.

To make the solving procedure tractable, two auxiliary variables $\mathbf{Y}_{t,i}^-$ and $\mathbf{Y}_{t,i}^+$ are introduced to reformulate the two nonlinear terms $\mathbf{Z}_{t,i}^- X_{t,i}$ and $\mathbf{Z}_{t,i}^+ X_{t,i}$ in (4). Set $\mathbf{Y}_{t,i}^- = \mathbf{Z}_{t,i}^- X_{t,i}$ and $\mathbf{Y}_{t,i}^+ = \mathbf{Z}_{t,i}^+ X_{t,i}$, then introduce constraints concerning $\mathbf{Y}_{t,i}^-$ and $\mathbf{Y}_{t,i}^+$ to make an equivalent transformation as shown in (5).

After the equivalent transformation, the new uncertainty set in (5) is a mixed-integer linear model, which is similar to (3).

$$U_{new} = \left\{ \widetilde{\mathbf{w}}_{t,new} \middle| \begin{array}{ll} \widetilde{\mathbf{w}}_{t,new} = \sum_{i=1}^{N} \left(\underline{\mathbf{w}}_{t,i} \mathbf{Y}_{t,i}^- + \overline{\mathbf{w}}_{t,i} \mathbf{Y}_{t,i}^+ \right) & \text{(a)} \\ \mathbf{Y}_{t,i}^- + \mathbf{Y}_{t,i}^+ \leq 1, \mathbf{Y}_{t,i}^-, \mathbf{Y}_{t,i}^+ \in \{\mathbf{0},\mathbf{1}\} & \text{(b)} \\ \sum_{i=1}^{N} X_{t,i} = 1, X_{t,i} \in \{0,1\} & \text{(c)} \\ \mathbf{Y}_{t,i}^-, \mathbf{Y}_{t,i}^- \leq X_{t,i} & \text{(d)} \end{array} \right\} \quad (5)$$

It should be noted that (5) utilizes the collection of subsets instead of the extreme points of the smallest convex hull to represent the uncertainty set. This is because finding extreme points of the smallest convex hull is a challenging task, especially when the dimension of uncertainty is high.

Even though the collection of subsets contains non-extreme points which may increase the computational burden, this letter has proposed a method to reduce the size of variables to be solved in the following part.

B. Identification of Inner Subsets

Fig. 2 divides all the subsets into two categories, the periphery subsets, and the inner subsets. The periphery subsets are those of which at least one edge is on the blue boundary of the partition-combine set. On the contrary, for the inner sets, there is no edge on the boundaries.

As can be seen from Fig. 1 (c), only the vertices of the periphery subsets can become the extreme points of the partition-combine set. Then we consider identifying the inner subsets to simplify the description of U_{new}. Fig. 3 shows how to identify inner subsets.
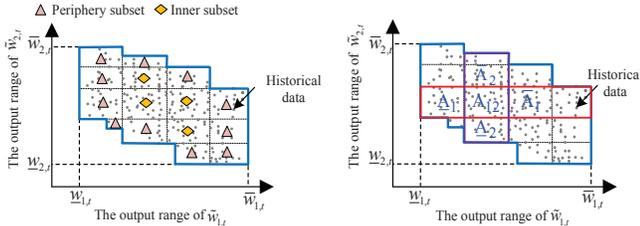

Fig. 2. Periphery and inner subsets Fig. 3. Inner subsets identification

As can be seen from Fig. 3, an inner subset always has two adjacent subsets in each dimension, which make its four edges would not fall on the blue boundary of the partition-combine uncertainty set. Here, we take the identification of subset A_{12} as an example. For the first dimension of A_{12}, which corresponds to the first wind farm, we lock the region surrounded by the red rectangle. In this region, there exist two adjacent subsets $\overline{A}_1$ and $\underline{A}_1$. Similarly, in the second dimension of A_{12}, which corresponds to the second wind farm, we lock the region surrounded by the purple rectangle, there also exist two adjacent subsets $\overline{A}_2$ and $\underline{A}_2$. Then, according to the adjacent subsets, we can judge A_{12} as an inner subset.

After finding out all the inner subsets, the following constraint is added to (5).

$$X_{t,i} = 0, i \in inner\ subsets \quad (6)$$

Then, the partition-combine uncertainty set will be only the combination of periphery subsets, and the number of uncertainty variables to be solved can be reduced.

IV. CASE STUDY

The modified IEEE-118 bus system with multiple wind farms is tested. The capacity of each wind farm is 300 MW. To validate the effectiveness of the proposed method in improving the economic efficiency, we compare the scheduling results between the box-like set and the partition-combine set in two cases with the different number of the wind farms. Besides, for the partition-combine uncertainty set, the scheduling results with the different number of subsets are also compared.

TABLE I
COMPARATIVE RESULTS BETWEEN THE BOX-LIKE SET AND THE
PARTITION-COMBINE UNCERTAINTY SET

Cases	No. of wind farms	Partition scheme $(i \times j)/(i \times j \times k)$	No. of subsets	Objective $(10^6\ \$)$
Case 1	2	1×1	1	1.619964
		2×2	4	1.613017
		3×3	9	1.611162
		4×4	16	1.610491
		5×5	25	1.609661
		2×18	36	1.608752
		3×12		1.608617
		4×9		1.609371
		6×6		1.609215
		7×7	49	1.609001
		8×8	64	1.608768
Case 2	3	$1 \times 1 \times 1$	1	1.588123
		$2 \times 1 \times 2$	4	1.578111
		$3 \times 1 \times 3$	9	1.578080
		$2 \times 4 \times 2$	16	1.577300
		$5 \times 5 \times 1$	25	1.575078
		$6 \times 1 \times 6$	36	1.573487
		$6 \times 2 \times 3$		1.574151
		$3 \times 2 \times 6$		1.573487
		$2 \times 1 \times 18$		1.573487
		$7 \times 1 \times 7$	49	1.573487
		$8 \times 2 \times 4$	64	1.573487

In Table 1, the number of subsets is equal to the product of partitions of each wind farm. Indices i, j, k represent the partitions of three wind farms, respectively. And when $i \times j = 1$ in Case 1 or $i \times j \times k = 1$ in Case 2, the uncertainty set is a box-like set. As can be seen from Table 1, when the number of subsets increases from 1 to 64, the optimal value of the objective function decreases gradually in both Case 1 and Case 2. This is because when the number of subsets increases, more regions without historical data will be removed, and the partition-combine uncertainty set has a smaller size to enclose the historical data. In Case 1, the objective has been decreased by 0.69%, when the number of subsets increases from 1 to 64. In Case 2, the objective has decreased by 0.92% with the same changes in the number of subsets.

For the same total number of subsets, we can have a different partition scheme. Take Case 1 as an example. When the total number of subsets is 36, the partition scheme for two wind farms can be 2×18, 3×12, 4×9 and 6×6. As can be seen from Table 1, a different partition scheme usually corresponds to a different objective. At present, we mainly rely on the simulation results to judge which partition scheme is better when the number of subsets is fixed.

To test the effects of historical data on the proposed method, Fig. 4 illustrates how the objective of Case 1 changes depending on the amount of historical data, with the partition scheme keeping as 88 unchanged.
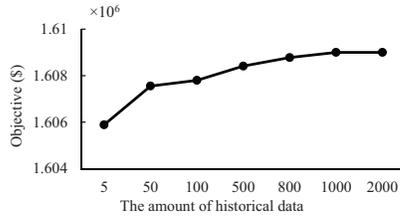
Fig. 4. Effects of historical data on the proposed method.

In Fig. 4, the objective increases with the amount of historical data firstly, then it tends to be constant after the number of historical data is large enough. This is because the partition-combine uncertainty set is data-driven. As the number of historical data increases, both the size and the worst-case scenario of the uncertainty set may be changed. Therefore, enough historical data is necessary to guarantee the effectiveness of the proposed method.

Fig. 5 gives the comparative results of calculation time in the different number of subsets and wind farms, as well as with and without the identification of inner subsets.

(a) The computation time with different subsets

(b) The computation time with and without identification of inner subsets for Case 1
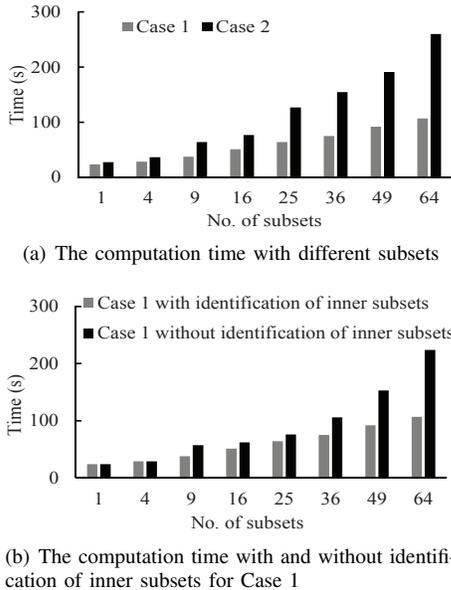
Fig. 5. Comparative results of computation time.

As can be seen from Fig. 5 (a), the computation time increases with the number of subsets both in Case 1 and Case 2. Besides, with the identification of the inner subset taken into consideration, the computation time can be decreased in Case 1. Besides, Case 2 spends more computation time than Case 1, due to the increase in the number of wind farms.

Similarly, when many wind farms are considered, the partition-combine uncertainty set will face a large scale of variables. This is because the scale of variables in the set increases with the number of wind farms when the divisions in each dimension are more than one. The RUC will therefore face a high computational burden. In the future, how to further reduce the scale of uncertainty variables and improve the computational efficiency will be considered.

In summary, although more subsets devote to a more economical decision, it is at the expense of increased computational complexity. To balance the economical and computational efficiencies, the number of subsets should be set by cases.

V. CONCLUSION

Constructing the smallest uncertainty set to enclose all the possible scenarios is important to guarantee the reliability and economic efficiency of the unit commitment. This letter proposes the partition-combine uncertainty set to make fewer conservative decisions. Numerical studies on the modified IEEE 118-bus system with multiple wind farms show that the proposed method can improve the economic efficiency of RUC. And a larger number of subsets means the more obvious effect in improving economic efficiency. However, the computation time will increase with the number of subsets. Therefore, the predefined number of subsets should be a compromise between economic efficiency and computation time. The temporal correlations and the computational efficiency improvement will be considered in the future work.

REFERENCES

[1] Y. Chen, Q. Wang, X. Wang, and Y. Guan, "Applying robust optimization to miso look-ahead commitment," in *2014 IEEE PES General Meeting— Conference & Exposition*. IEEE, 2014, pp. 1–5.
[2] 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.
[3] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, Feb. 2013.
[4] C. Zhao and Y. Guan, "Unified stochastic and robust unit commitment," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3353–3361, Aug. 2013.
[5] C. Dai, L. Wu, and H. Wu, "A multi-band uncertainty set based robust scuc with spatial and temporal budget constraints," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4988–5000, Nov. 2016.
[6] Y. An and B. Zeng, "Exploring the modeling capacity of two-stage robust optimization: Variants of robust unit commitment model," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 109–122, Jan. 2015.
[7] Y. Guan and J. Wang, "Uncertainty sets for robust unit commitment," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1439–1440, May 2014.
[8] C. Zhao and Y. Guan, "Data-driven stochastic unit commitment for integrating wind generation," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2587–2596, Jul. 2016.
[9] F. Qiu, Z. Li, and J. Wang, "A data-driven approach to improve wind dispatchability," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 421–429, Jan. 2017.
[10] A. Lorca and X. A. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1678–1688, May 2017.
[11] C. Li, J. Zhao, T. Zheng, and E. Litvinov, "Data-driven uncertainty sets: Robust optimization with temporally and spatially correlated data," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 2016, pp. 1–5.
[12] T. Ding, J. Lv, R. Bo, Z. Bie, and F. Li, "Lift-and-project mvee based convex hull for robust sced with wind power integration using historical data-driven modeling approach," *Renewable Energy*, vol. 92, pp. 415–427, Jul. 2016.
[13] Y. Zhang, X. Ai, J. Fang, J. Wen, 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.
[14] A. Velloso, A. Street, D. Pozo, J. M. Arroyo, and N. G. Cobos, "Two-stage robust unit commitment for co-optimized electricity markets: An adaptive data-driven approach for scenario-based uncertainty sets," *IEEE Trans. Sustain. Energy*, pp. 1–12, 2019, early access.
[15] M. Zhang, J. Fang, X. Ai, H. Shuai, W. Yao, H. He, Q. Wu, and J. Wen, "Feasibility identification and computational efficiency improvement for two-stage ruc with multiple wind farms," *IEEE Trans. Sustain. Energy*, pp. 1–10, 2019, early access.