Data-Driven Adaptive Robust Unit Commitment Under Wind Power Uncertainty: A Bayesian Nonparametric Approach

Chao Ning and Fengqi You, *Member, IEEE*

***Abstract*—This paper proposes a novel data-driven adaptive robust optimization (ARO) framework for the unit commitment (UC) problem integrating wind power into smart grids. By leveraging a Dirichlet process mixture model, a data-driven uncertainty set for wind power forecast errors is constructed as a union of several basic uncertainty sets. Therefore, the proposed uncertainty set can flexibly capture a compact region of uncertainty in a nonparametric fashion. Based on this uncertainty set and wind power forecasts, a data-driven adaptive robust UC problem is then formulated as a four-level optimization problem. A decomposition-based algorithm is further developed. Compared to conventional robust UC models, the proposed approach does not presume single mode, symmetry, or independence in uncertainty. Moreover, it not only substantially withstands wind power forecast errors, but also significantly mitigates the conservatism issue by reducing operational costs. We also compare the proposed approach with the state-of-the-art data-driven ARO method based on principal component analysis and kernel smoothing to assess its performance. The effectiveness of the proposed approach is demonstrated with the six-bus and IEEE 118-bus systems. Computational results show that the proposed approach scales gracefully with problem size and generates solutions that are more cost effective than the existing data-driven ARO method.**

***Index Terms*—Data-driven optimization, Dirichlet process mixture model, renewable energy, uncertainty, unit commitment.**

NOMENCLATURE

A. Sets and Indices

- b Index for buses.
- i Index for generators.
- l Index for transmission lines.
- t Index for time periods.

B. Parameters

- C_i^{SU} Startup cost of generator *i*.
- C_i^{SD} Shutdown cost of generator.
- C_i^{F} No-load cost of generator *i*.
- C_i^{V} Fuel cost of generator *i*.
- C^{+} Penalty cost for the violation of energy balance constraint.
- C^{-} Penalty cost for the violation of energy balance constraint.
- C^{f} Penalty cost for the violation of transmission line constraint.
- D_{bt} Load demand located at bust *b* at period *t*.
- DT_i Minimum down time of generator *i*.
- F_l Capacity of transmission line *l*.
- f_{bt} Wind power forecast at bus *b* at period *t*.
- K_{bl} Power flow distribution factor for the transmission line *l* due to the net injection at bus *b*.
- $P_i^{\min}$ Minimum power output of generator *i*.
- $P_i^{\max}$ Maximum power output of generator *i*.
- RU_i Ramp up rate of generator *i*.
- RD_i Ramp down rate of generator *i*.
- SU_i Start up rate of generator *i*.
- SD_i Shutdown rate of generator *i*.
- UT_i Minimum up time of generator *i*.
- ξ_{bt} Uncertain wind power generation at bus *b* at period *t*.
- θ_{bt} Wind power forecast error at bus *b* at period *t*.

C. Binary Decision Variables

- u_{it} Startup status that is equal to 1 if generator *i* starts up at period *t*.
- v_{it} Shutdown status that is equal to 1 if generator *i* shuts down at period *t*.
- z_{it} Commitment status that is equal to 1 if generator *i* is online at period *t*.

D. Continuous Decision Variables

- p_{it} Power output of generator *i* at period *t*.
- q_{lt}^{f} Slack variable in capacity constraint of transmission line *l* at period *t*.
- q_t^{+} Slack variable in the power balance constraint at period *t*.
- q_t^{-} Slack variable in the power balance constraint at period *t*.
- w_{bt} Power dispatch of wind farm at bus *b* at period *t*.

Manuscript received August 1, 2018; revised November 11, 2018; accepted December 31, 2018. Date of publication January 7, 2019; date of current version April 17, 2019. This work was partially supported by the National Science Foundation under Grant CBET-1643244. Paper no. TPWRS-01158-2018. (*Corresponding author: Fengqi You.*)

The authors are with the Robert Frederick Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: cn293@cornell.edu; fengqi.you@cornell.edu).

This paper has supplementary material available at http://ieeexplore.ieee.org (File size: 20 MB).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2019.2891057

I. INTRODUCTION

WITH growing concerns over climate change and a global energy crisis, the utilization of renewable energy sources, including wind and solar power, is growing rapidly around the globe [1]. Because of its economic incentives and sustainability benefits, wind power generation is aimed to meet 20% of U.S electricity needs by 2030 [2]. However, the intermittent nature and limited predictability of wind power generation poses a significant challenge to power systems operations. The unit commitment (UC) decision, one of the most critical decisions for the day-ahead electricity market, is highly susceptible to wind power forecast errors [3]. Such wind power uncertainty could lead to significant economic loss and even jeopardize the security of power grids [4]. Therefore, it becomes essential for independent system operators (ISOs) to make cost-effective and robust unit commitment decisions in the face of highly variable wind power.

Over the past few years, extensive research efforts have been made on methodologies of optimization under uncertainty in support of large-scale renewable energy integration into power systems operations [5]. The UC problem is typically formulated following the two-stage optimization structure due to the day-ahead market and real-time economic dispatch characteristics [6]. Within such structure, the first-stage decisions are the commitments of generating units made 24 hours ahead, and the power dispatch decisions are made at the second stage after uncertain wind power realizations.

Stochastic UC models utilize a set of scenarios as a description of uncertain parameters and aims to optimize the expected operational cost based on the stochastic programming approach [5], [7]. In [8], the UC problem with variable wind power generation was cast as a chance-constrained stochastic program to guarantee the wind power utilization. To account for the risk aversion, a stochastic UC model based on conditional value at risk (CVaR) was proposed for isolated systems with high penetration of wind power [9]. Despite its attractive features, the stochastic UC problem inevitably suffers from prohibitive computational burdens [10], [11].

As an alternative paradigm, two-stage adaptive robust optimization (ARO) has attracted increasing attention in power systems operation [12], [13]. It not only immunizes power systems with operational security, but also enjoys enormous computational benefits [14]. Two-stage robust UC problems considering network constraints were studied in [15], [16]. A solution methodology combining the cutting plane algorithm and a heuristic approach was developed to solve two-stage robust UC problem [17]. In [18], the asymmetric characteristic of uncertain wind power was effectively captured in robust security constrained UC problem through a multi-band uncertainty set. Recently, robust risk-constrained UC model based on an adjustable uncertainty set was developed to cope with the large-scale penetration of wind power [19]. To further promote the deregulation of power industries, energy storage was incorporated into the robust UC problem subject to renewable energy uncertainty [20]. In [21], the authors discussed various ways to construct uncertainty sets, including how to combine uncertainty sets of heterogeneous uncertainties. In [22], uncertainty sets of wind generation and load demand were constructed by incorporating their variability correlations. A robust security-constrained UC framework with flexible uncertainty set was proposed, in which the size and position of uncertainty set are treated as decision variables [23]. A method to devise uncertainty sets for distribution parameters characterizing wind generation variability was developed in [24]. The existing uncertainty sets typically have parametric expressions involving only a finite number of parameters, which could limit their capability of characterizing uncertainty using historical data.

Recently, distributionally robust optimization (DRO) emerges as a new data-driven optimization paradigm which hedges against the worst-case distribution in an ambiguity set [25], and has various applications in power systems, such as UC problems [26]–[31], optimal power flow (OPF) [32]–[37], distribution systems [38], reserve dispatch [39], and congestion management [40]. Rather than relying on a known distribution of uncertainty, DRO leverages available uncertainty data to construct a family of distributions based on either moment information or statistical distance. A distributionally robust UC problem was proposed to minimize the worst-case expected cost over a moment-based ambiguity set and was further solved using the linear decision rule [26]. By incorporating moment information and support of contingency uncertainty, a distributionally robust contingency-constrained UC model generated less conservative solutions compared with conventional robust optimization approaches. Another popular way to construct ambiguity sets is to include a family of distributions which are close to an empirical distribution in terms of some distance. In [28], a distance-based DRO was proposed for UC problems utilizing Kullback-Leibler divergence. An ambiguity set was developed for UC based on a confidence region for cumulative distribution function and could converge to the true distribution with increasing amount of uncertainty data [29]. In [30], the authors studied a data-driven risk-averse stochastic UC problem under wind uncertainty, and proposed an ambiguity set derived from data based on norms.

In the power industry, a large amount of data is routinely archived, and this data has potential values to support power systems operations [29], [30], [41], [42]. Data-driven robust optimization is an emerging paradigm for decision-making under uncertainty by leveraging recent advances in machine learning methods [43]–[47]. As a paramount ingredient in robust UC [21], uncertainty sets endogenously determine the UC solutions and therefore should be devised with special care. The uncertainty sets in existing robust UC models, however, typically admit a parametric expression involving only a finite number of parameters, thereby leading to limited data-adaptive flexibility. This conventional “one-set-fits-all” approach falls short of appropriately capturing intricate yet useful uncertainty information for UC decisions. Additionally, data outliers, which are ubiquitous in data acquisition [48], could drastically enlarge conventional uncertainty sets and lead to unfavorable conservatism. Therefore, it remains a research challenge to develop a holistic UC optimization framework that leverages state-of-the-art machine learning methods and the value of large renewable generation datasets for the UC optimization.

To fill the knowledge gap, we propose a novel data-driven adaptive robust UC (DDARUC) optimization framework for high penetration of wind power in smart grids. First, a systematic data-driven approach to construct an uncertainty set for wind power forecast error is proposed by taking advantage of a Dirichlet process mixture model (DPMM). We devise the uncertainty set based on the posterior predictive distribution of future wind power forecast errors given their historical dataset. Moreover, the proposed uncertainty set manifests itself as a union of multiple basic uncertainty sets, the number of which is automatically derived from uncertainty data. Accordingly, it could capture a compact, high-density region of uncertain parameters in a nonparametric fashion. Combining this uncertainty set with day-ahead wind power forecasts, the DDARUC problem is then formulated as a four-level optimization problem. The proposed DDARUC approach not only hedges against wind power uncertainty, but also ameliorates the conservatism issue by slashing operational costs. The illustrative six-bus and IEEE 118-bus systems are used to demonstrate the effectiveness of the proposed approach. Comparisons with the existing data-driven ARO method are also made to verify the superior performance of the proposed DDARUC approach using the DPMM.

The major contributions of this paper are summarized as follows.

- A novel data-driven adaptive robust UC optimization framework that leverages the Bayesian nonparametric method for smart power systems operations;
- A systematic data-driven approach based on DPMM to construct an uncertainty set that is self-adaptive to the underlying structure and complexity of wind forecast error data;
- Comparisons with the state-of-the-art data-driven ARO method and the data-driven distributionally robust optimization method for the UC optimization in terms of solution quality and computational time.

The remainder of this paper is organized as follows. Section II presents a detailed formulation of the two-stage robust UC problem. A data-driven robust UC optimization framework is proposed in Section III. In Section IV, an efficient solution algorithm is developed. Section V presents two case studies to demonstrate the effectiveness of the proposed approach. Conclusions are drawn in the last section.

II. Mathematical Formulation

In this section, a two-stage adaptive robust UC model formulation considering economic dispatch is presented. Within this model, the first-stage decisions are the commitment status of generators (i.e., on/off status and startup/shutdown schedules of generation units). The second-stage decisions involve the dispatch decisions of conventional generators and renewable wind power, which are made after knowing uncertainty realizations [49], [50].

The two-stage adaptive robust UC model is formulated as follows. The objective of the UC problem is to minimize the total operating cost in (1). The total cost includes startup cost, shutdown cost, no-load cost, fuel cost and penalty costs for electricity demand shedding and transmission line overload. Constraints (2), (3) describe the logic relationships between the generator commitment and start-up/shut-down decisions. Constraints (4), (5) enforce the minimum up and down times, respectively. The minimum and maximum outputs of each generation unit are specified in Constraint (6). Constraints (7), (8) describe the ramping rate limits of generators for each hour. Constraint (9) enforces the energy balance [51]. The capacity constraints for transmission lines are described in Constraints (10), (11). Constraint (12) enforces that generation outputs of wind farms cannot exceed the available wind power [20], [26]. Constraint (13) specifies the integrity restriction of first-stage decisions, while Constraint (14) specifies the non-negativity of second-stage decisions. Note that the wind power generation is subject to uncertainty, and U is its uncertainty set. Indices/sets, parameters and variables is given in the Nomenclature section. Throughout this paper, vectors are denoted using bold lower-case letters and matrices are denoted using bold upper-case letters, while scalars are written with italic letters.

$$\min_{u_{it},v_{it},z_{it}} \sum_i \sum_t \left(C_i^{SU} u_{it} + C_i^{SD} v_{it} + C_i^F z_{it}\right) + \max_{\xi_{bt} \in U} \min_{\substack{p_{it},w_{bt} \\ q_t^+,q_t^-,q_{lt}^f}}$$

$$\left[\sum_i \sum_t C_i^V p_{it} + \sum_t \left(C^+ q_t^+ + C^- q_t^-\right) + \sum_l \sum_t C^f q_{lt}^f\right] \quad (1)$$

$$\text{s.t. } u_{it} - v_{it} = z_{it} - z_{i,t-1}, \quad \forall i,t \quad (2)$$

$$u_{it} + v_{it} \le 1, \quad \forall i,t \quad (3)$$

$$\sum_{t'=t-UT_i+1}^{t} u_{it'} \le z_{it}, \quad \forall i, t \in [UT_i, T] \quad (4)$$

$$\sum_{t'=t-DT_i+1}^{t} v_{it'} \le 1 - z_{it}, \quad \forall i, t \in [DT_i, T] \quad (5)$$

$$P_i^{\min} z_{it} \le p_{it} \le P_i^{\max} z_{it}, \quad \forall i,t \quad (6)$$

$$p_{it} - p_{i,t-1} \le RU_i z_{i,t-1} + SU_i u_{it}, \quad \forall i,t \quad (7)$$

$$p_{i,t-1} - p_{it} \le RD_i z_{i,t} + SD_i v_{it}, \quad \forall i,t \quad (8)$$

$$\sum_b \left(\sum_{i \in G_b} p_{it} + w_{bt}\right) + q_t^+ - q_t^- = \sum_b D_{bt}, \quad \forall t \quad (9)$$

$$\sum_b K_{bl} \left(\sum_{i \in G_b} p_{it} + w_{bt} - D_{bt}\right) - q_{lt}^f \le F_l, \quad \forall l,t \quad (10)$$

$$\sum_b K_{bl} \left(\sum_{i \in G_b} p_{it} + w_{bt} - D_{bt}\right) + q_{lt}^f \ge -F_l, \quad \forall l,t \quad (11)$$

$$w_{bt} \le \xi_{bt}, \quad \forall b,t \quad (12)$$

$$u_{it}, v_{it}, z_{it} \in \{0,1\}, \quad \forall i,t \quad (13)$$

$$p_{it}, w_{bt}, q_t^+, q_t^-, q_{lt}^f \ge 0, \quad \forall i,b,l,t \quad (14)$$

The adaptive robust UC is cast as the above tri-level optimization problem, which captures the sequential decision-making process in power systems operations.

Existing robust UC optimization methods can neither make full use of the complex uncertainty data information, nor can they effectively account for issues on correlation, asymmetry, and the multimodal nature of wind power forecast errors. More specifically, they typically provide limited modeling flexibility to be adaptive to the underlying data structure and complexity. Therefore, the knowledge gap to fill is a data-driven robust UC optimization framework that is self-adaptive to the ever-increasing data complexity and leverages the value of large datasets available for wind power forecast for more efficient power systems operations.

III. Data-Driven Adaptive Robust Unit Commitment Optimization Framework

In this section, we propose a novel data-driven adaptive robust UC optimization framework. We first present a brief introduction to the DPMM. We then propose the data-driven uncertainty set of wind power forecast errors. Finally, a data-driven robust UC model is developed.

Hourly wind power at wind farms is routinely measured, and data is collected and archived in datasets [52]. Suppose ISOs have access to dataset $D^w = \{\xi^{(1)}, \ldots, \xi^{(N_t)}\}$, in which $\xi^{(t)}$ is a vector of wind power measurements in different wind farms at period t, and N_t denotes the number of data samples. These data typically exhibit both spatial and temporal correlations. Thus, they are time series, rather than independent and identically distributed (i.i.d.) data points. As a consequence, traditional unsupervised machine learning methods, such as DPMM and kernel density estimation, cannot be applied directly. Motivated by the above observation, we take advantage of wind power forecasting data in addition to the wind power measurements. Various wind power forecasting techniques using weather information are well developed and widely used in power systems operations [53]. Suppose $D^f = \{\mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(N_t)}\}$ is a dataset of N_t-hour wind power forecasts that capture the temporal dynamics, where $\mathbf{f}^{(t)}$ denotes a vector of wind power forecast values in different wind farms at period t.

The wind forecast error is defined as the difference between actual wind power measurement and its corresponding forecast value at a specific time, which is given by,

$$\theta_b^{(t)} = \xi_b^{(t)} - f_b^{(t)}, \quad \forall b, t \tag{15}$$

where $\theta_b^{(t)}$ is the forecast error for wind farm at bus b for time period t, $\xi_b^{(t)}$ denotes the actual wind power measurement at bus b for time period t, and $f_b^{(t)}$ represents the wind power forecast at bus b for period t.

Accordingly, we obtain the dataset $D^e = \{\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N_t)}\}$. Suggested by the literature [54, 55], the forecast error data in D^e can be regarded as i.i.d. samples of $\boldsymbol{\theta}$. Thus, instead of considering the time series modeling, we directly employ a DPMM for such i.i.d. case for wind power forecast errors. Notably, these forecast error data distribution is typically not Gaussian, and can be highly dependent on the forecasting technique adopted by ISOs. Moreover, these data could exhibit strong spatial correlations among wind farms due to geographical features [56]. Note that the paper considers the day-ahead unit commitment problem, and the forecast horizon is 24 hours. If the forecast error does not follow the assumption regarding i.i.d. data, some time series models could be employed to address this issue.

A. Dirichlet Process Mixture Model

The Dirichlet process (DP) constitutes a fundamental building block for the DPMM that relies on mixtures to characterize data distribution. The DP is technically a probability distribution over distributions. Suppose a random distribution G follows a DP parameterized by a concentration parameter α and a base measure G_0 over space Θ, denoted as $G \sim DP(\alpha, G_0)$. For any fixed partitions $(A_1, \ldots, A_r)$ of Θ, we have the following

$$(G(A_1), \ldots, G(A_r)) \sim Dir\,(\alpha G_0(A_1), \ldots, \alpha G_0(A_r)) \tag{16}$$

where *Dir* represents the Dirichlet distribution.

Following the stick-breaking procedure [57], a random draw from the DP can be expressed by $G = \sum_{k=1}^{\infty} \rho_k \delta(\varphi_k)$, where $\rho_k = \beta_k \sum_{j=1}^{k-1} (1 - \beta_j)$ is the weight, φ_k is sampled from G_0, and $\delta(\varphi_k)$ denotes the Dirac delta function at φ_k. β_j represents the proportion being broken from the remaining stick, and follows a Beta distribution, denoted as $\beta_j \sim Beta\,(1, \alpha)$.

Note that the Dirichlet process is essentially a distribution on distributions. Accordingly, a random draw from a Dirichlet process $DP(\alpha, G_0)$ is a distribution G. One can still generate a random draw from this distribution G. The Bayesian nonparametric model, i.e., DPMM, adds one more level to the hierarchy by using φ_k as the parameters of some data distribution. Based on the DP, we summarize the basic form of a DPMM as follows [58], [59]:

$$\begin{aligned} &\{\rho_k, \varphi_k\}_{k=1}^{\infty} \sim DP\,(\alpha, G_0) \\ &l_n \sim Mult\,(\rho) \\ &o_n \sim F(\varphi_{l_n}) \end{aligned} \tag{17}$$

where *Mult* denotes a multinomial distribution, l_n is the label indicating the component or cluster of observation o_n, n is an index ranging from 1 to N, and data $o_1, \ldots, o_N$ are distributed according to a family of distributions F. Based on the stick-breaking process, G is discrete with probability one. Such discreteness further induces the clustering of data.

Due to its computational efficiency, variational inference has become a method of choice for approximating the conditional distribution of latent variables in the DPMM [59]. In the variational inference, a factorized variational distribution is used to approximate the true posterior in terms of Kullback–Leibler divergence. Following [58], we use a mixture of Gaussians in this paper. Therefore, we choose $\varphi_k \sim NW(\boldsymbol{\mu}_k, \lambda_k, \omega_k, \boldsymbol{\Psi}_k^{-1})$, where $\varphi_k = (\boldsymbol{\eta}_k, \mathbf{H}_k)$ includes mean vector and precision matrix and *NW* represents the normal Wishart distribution.

Determining the number of clusters is critical for wind forecast error data. More specifically, too many clusters could lead to over-fitting, while too few clusters could greatly compromise the modeling flexibility. The issue of under-fitting is avoided because the DPMM allows for a countably infinite number of clusters. The over-fitting problem is handled by the prior of DP within a Bayesian framework.

The DPMM enjoys some unique features that make it suitable for modeling wind power forecast error data. Unlike other machine learning methods, such as Gaussian mixture model and K-means, the DPMM can determine the number of clusters systematically and automatically rather than specifying this number *a priori*. As a result, its model complexity is adaptive to the complexity of the uncertainty data, which in turn facilitates extracting useful uncertainty information for the robust UC optimization. Moreover, as a Bayesian approach, the DPMM also provides a platform for the posterior inference, which characterizes uncertainty accurately.

B. Data-Driven Uncertainty Set

The data-driven uncertainty set is constructed based on posterior predictive distribution $\Pr(\boldsymbol{\theta}_{\text{new}}|D^{\text{e}})$, where random vector $\boldsymbol{\theta}_{\text{new}}$ represents future wind forecast errors. The conditional distribution of $\boldsymbol{\theta}_{\text{new}}$ given forecast error dataset D^{e} is a mixture of Student's *t*-distributions [58], [60], shown as follows.

$$\boldsymbol{\theta}_{\text{new}} \sim \sum_j \gamma_j St_{\kappa_j+1-\dim(\boldsymbol{\theta})}\left(\boldsymbol{\mu}_j, \frac{\boldsymbol{\Psi}_j^{-1}}{s_j^2}\right) \tag{18}$$

where γ_j is the weight of *j* th component, $\gamma_j = \frac{\tau_j}{\tau_j+\upsilon_j}\prod_{r=1}^{j-1}\frac{\upsilon_r}{\tau_r+\upsilon_r}$, $j = 1,\ldots,M-1$, $\gamma_M = 1-\sum_{j=1}^{M-1}\gamma_j$, and *M* is the truncation level [59]. $\tau_j, \nu_j, \boldsymbol{\mu}_j, \lambda_j, \kappa_j, \psi_j$ are the inference results of *j* th component using the variational inference algorithm [59], and $s_j = \sqrt{(\lambda_j+1)/[\lambda_j(\kappa_j+1-\dim(\boldsymbol{\theta}))]}$ [58].

Based on the posterior predictive distribution, we propose a data-driven uncertainty set Ξ for wind forecast errors below.

$$\Xi = \bigcup_{j:\,\gamma_j\geq\gamma^*}\Xi_j = \Xi_1\bigcup\Xi_2\cdots\bigcup\Xi_m \tag{19}$$

where γ^* is a threshold value, and *m* denotes the number of basic uncertainty sets. The number of basic uncertainty sets is determined by the number of clusters within uncertainty data, which is obtained by the variational inference algorithm. Basic uncertainty set Ξ_j for component *j* is explicitly defined as follows [61], and manifests itself as an intersection of l_1 and l_∞ norm balls.

$$\Xi_j = \left\{\boldsymbol{\theta}\middle|\ \boldsymbol{\theta} = \boldsymbol{\mu}_j + s_j\boldsymbol{\Psi}_j^{1/2}\Lambda_j\,\boldsymbol{\delta},\ \|\boldsymbol{\delta}\|_\infty \leq 1, \|\boldsymbol{\delta}\|_1 \leq \Phi_j\right\} \tag{20}$$

where δ is the latent uncertainty, parameter Φ_j is used to control the size of the uncertainty set, and Λ_j is a scaling factor. The proposed uncertainty set is variable, and its size can be adjusted using the tuning parameter Φ_j. Increasing the value of parameter Φ_j enlarges the size of the uncertainty set. Note that (17) is the generative model for DPMM. To get the posterior of latent variables, variational inference is employed. After that, the predictive distribution (18) is calculated using the inference results. For each component in the predictive mixture model, we construct a basic uncertainty set in (20). The proposed uncertainty set is a union of these basic uncertainty sets, which is expressed in (19).

The proposed data-driven uncertainty set for the wind forecast error is cast as a union of multiple basic uncertainty sets. Its geometric shape is flexible enough to capture the high-density region of uncertain forecast error data. There are several highlights of the proposed uncertainty set. First, the proposed uncertainty set is constructed based on the posterior predictive distribution of forecast errors, i.e., $\Pr(\boldsymbol{\theta}_{\text{new}}|D^{\text{e}})$, which guarantees its out-of-sample performance. Second, the complexity of the proposed data-driven uncertainty set is self-adaptive to the underlying complexity and structure of wind forecast error data. This feature implies that our proposed method is general enough to be synthesized with any types of wind forecasting techniques or software adopted by ISOs. Third, each basic uncertainty set is devised as a polyhedral set, which endows the resulting robust UC problem with enormous computational benefits.

C. Data-Driven Robust Unit Commitment Model

For the ease of exposition, we present (DDARUC) model based on the data-driven uncertainty set in the following abstract form, in which vectors and matrices are introduced.

$$\min_{\mathbf{x}}\mathbf{c}^T\mathbf{x} + \max_{j\in\{1,\ldots,m\}}\max_{\boldsymbol{\theta}\in\Xi_j}\min_{(\mathbf{p},\mathbf{q},\mathbf{w})\in\Omega(\mathbf{x},\boldsymbol{\theta})}\left(\mathbf{b}^T\mathbf{p}+\mathbf{d}^T\mathbf{q}\right) \tag{21}$$

$$\text{s.t. } \mathbf{Ax}\geq\mathbf{g},\ \mathbf{x}\in\{0,1\}^{3\times|I|\times|T|} \tag{22}$$

$$\Omega(\mathbf{x},\boldsymbol{\theta}) = \left\{(\mathbf{p},\mathbf{q},\mathbf{w})\middle|\begin{array}{l}\mathbf{Ep}+\mathbf{Fq}+\mathbf{Gw}\geq\mathbf{h}-\mathbf{Tx}-\mathbf{M}\boldsymbol{\theta}\\ \mathbf{p},\mathbf{q},\mathbf{w}\geq\mathbf{0}\end{array}\right\} \tag{23}$$

where $\mathbf{x}$ denotes the vector of all first-stage decisions including u_{it}, v_{it}, and z_{it}; $\mathbf{p}$ is the vector of generator dispatch decision p_{it}; $\mathbf{q}$ represents the vector of slack decision variables q_{lt}^f, q_t^+, and q_t^-; $\mathbf{w}$ denotes the vector of wind power dispatch decision w_{bt}; and $\boldsymbol{\theta}$ is the vector of uncertain wind power forecast errors. Note that the second-stage decisions include vectors $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{w}$. Constraint (22) represents constraints (2)–(5), and expression (23) encompasses constraints (6)–(14). Note that the first "max" in the DDARUC model represents the maximization problem over the index of basic uncertainty sets.

IV. Solution Methodology

Due to the multilevel optimization structure coupled with the nonconvex nature of the proposed uncertainty set, the resulting data-driven robust UC problem cannot be solved directly by any off-the-shelf optimization solvers. Thus, we develop an efficient solution algorithm in this section.

The optimal solution for the maximization problem over $\boldsymbol{\theta}$ in (DDARUC) is one extreme point of basic polyhedral uncertainty set Ξ_j. In view of this, we can reduce the four-level optimization problem into a single-level full master problem through the enumeration of all the extreme points. From a computational perspective, the resulting full master problem is prohibitively challenging to solve due to the large number of the induced UC constraints. Therefore, a partial enumeration scheme of extreme points is employed to generate significant uncertainty scenarios on the fly [62], [63], which yields the following problem (MP).

$$\text{(MP)}\quad\begin{aligned}\min_{\mathbf{x},\eta,\mathbf{p}^k,\mathbf{q}^k,\mathbf{w}^k}\ & \mathbf{c}^T\mathbf{x}+\eta\\ \text{s.t.}\ & \mathbf{Ax}\geq\mathbf{g},\mathbf{x}\in\{0,1\}^{3\times|I|\times|T|}\\ & \eta\geq\mathbf{b}^T\mathbf{p}^k+\mathbf{d}^T\mathbf{q}^k, k\in K\\ & \mathbf{Tx}+\mathbf{Ep}^k+\mathbf{Fq}^k+\mathbf{Gw}^k\geq\mathbf{h}-\mathbf{M}\hat{\boldsymbol{\theta}}^k, k\in K\\ & \mathbf{p}^k,\mathbf{q}^k,\mathbf{w}^k\geq\mathbf{0}, k\in K\end{aligned}$$

where η is an auxiliary variable, *K* is the index set of extreme points. $\mathbf{p}^k$, $\mathbf{q}^k$, and $\mathbf{w}^k$ are the introduced second-stage decisions associated with wind forecast error scenario $\hat{\boldsymbol{\theta}}^k$.

The master problem (MP) is solved to generate the unit commitment decision $\mathbf{x}^*$. Since (MP) is only a relaxation of the full master problem, its objective value yields a lower bound of problem (DDARUC). For a given unit commitment schedule,

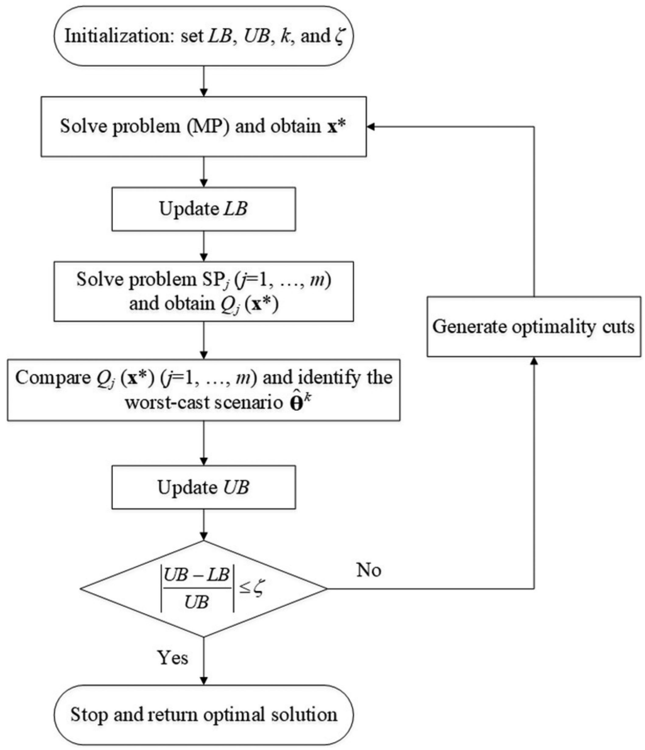


Fig. 1. Flowchart of the proposed solution algorithm.

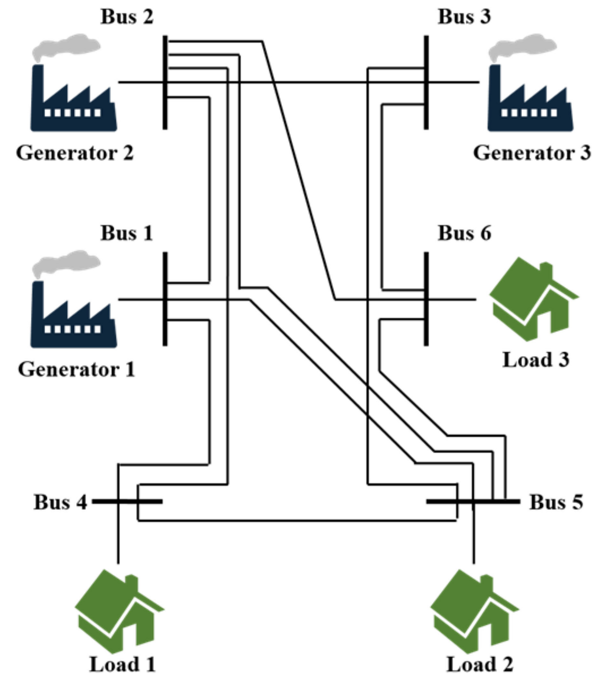


Fig. 2. The schematic representation of the six-bus system.

the inner "max-max-min" problem in (DDARUC) is then used to identify the worst-case wind forecast error scenario. The optimality cuts associated with the identified uncertainty scenario are then fed back to problem (MP) for the next iteration. To address the "max-max-min" problem, we further decompose it into a set of subproblems $(\mathrm{SP}_j,\ j = 1, \ldots, m)$, and each subproblem (SP_j) corresponds to the basic uncertainty set Ξ_j. The optimal objective values of these subproblems are compared to obtain the largest one such that a worst-case uncertainty scenario can be identified and then used to generate optimality cuts.

$$(\mathrm{SP}_j) \qquad \begin{aligned} Q_j\left(\mathbf{x}^*\right) = \max_{\boldsymbol{\theta}\in\Xi_j} \min_{\mathbf{p},\mathbf{q},\mathbf{w}\geq\mathbf{0}} & \left(\mathbf{b}^T\mathbf{p} + \mathbf{d}^T\mathbf{q}\right) \\ \text{s.t. } & \mathbf{Ep} + \mathbf{Fq} + \mathbf{Gw} \geq \mathbf{h} - \mathbf{Tx}^* - \mathbf{M}\boldsymbol{\theta} \end{aligned}$$

The flowchart of the solution algorithm for problem (DDARUC) is summarized in Fig. 1. Note that *LB* denotes the lower bound, *UB* is the upper bound, and ζ represents the tolerance of relative optimality gap. The decomposition based algorithmic framework iteratively solves (MP) and a set of (SP_j) to obtain the global optimal UC decision. The DDARUC problem can be reformulated as a single-level optimization problem by enumerating all the extreme points of uncertainty set [62]. Moreover, a different extreme point is guaranteed to emerge at each iteration if the termination criterion is not satisfied. Since the number of extreme points of basic uncertainty sets is finite, this solution algorithm is guaranteed to converge to the global optimum within a finite number of iterations.

V. Computational Experiments

In this section, case studies on the six-bus and IEEE 118-bus systems are presented. We also implement the conventional adaptive robust UC with a budget uncertainty set and data-driven adaptive robust UC using principal component analysis (PCA) and kernel smoothing. The budget uncertainty set [64] and data-driven uncertainty set using PCA and kernel smoothing [65] are used in these UC problems. All optimization problems are solved with CPLEX 12.8.0, implemented on a computer with an Intel (R) Core (TM) i7-6700 CPU @ 3.40 GHz and 32 GB RAM. The optimality tolerance for CPLEX 12.8.0 is set to 0, and the relative optimality gap tolerance for the algorithm is 10^{-4}. Penalty costs C^+, C^-, C^f are set to 50,000$/MWh.

A. Illustrative Six-Bus System

The system diagram of six-bus is shown in Fig. 2. Moreover, there are three wind farms located at Buses 1–3.

1) Comparison With Adaptive Robust UC Methods: Following the literature [66], historical wind forecast error data of 500 hours are generated based on a Gaussian mixture model. Fig. 3(a) shows the scatter plot of these data points, which are clearly correlated, asymmetric and multimodal. The conventional gamma or budget uncertainty set merely utilizes the uncertainty deviation information, and adjusts the size of uncertainty set via an uncertainty budget parameter [64]. From Fig. 3(b), we can observe that the budget uncertainty set is non-adaptive to the underlying data distribution, and inappropriately covers much unnecessary region, which could potentially lead to over-conservative robust UC solutions. The uncertainty set budget for conventional robust UC is set to be two. Fig. 3(c) presents the uncertainty set based on PCA and kernel smoothing. It accurately captures the correlations among renewable energy forecast errors in different wind farms via the identified principal components (denoted as PC1, PC2, and PC3 in the figure). However, due to its parametric nature, it still does not correctly reflect the multimodal characteristics. By contrast, the data-driven approach based on DPMM yields a compact non-convex uncertainty set that is a union of three green polytopes.

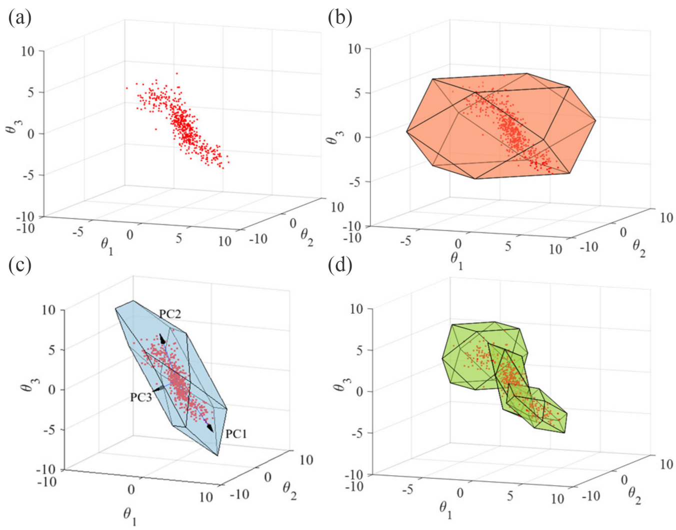


Fig. 3. (a) Scatter plot of forecast error data points. (b) Budget or gamma uncertainty set. (c) Data-driven uncertainty set based on PCA and kernel smoothing. (d) Data-driven uncertainty set based on DPMM.

TABLE I
COMPARISONS OF PROBLEM SIZES AND COMPUTATIONAL RESULTS FOR THE SIX-BUS TEST SYSTEM

Parameter	ARUC	DDARUC-PCA	DDARUC-DPMM
Bin. Var.	216	216	216
Cont. Var.	241	241	247
Constraints	1,275	1,287	1,289
Min. cost ($)	46,850	44,984	43,879
CPU time (s)	1.2	1.4	2.7

We note that the number of polytopes, which correspond to basic uncertainty sets in the proposed data-driven uncertainty set, is automatically determined by the machine learning algorithm based on the data structure and complexity. The number of basic uncertainty sets derived from the uncertainty data is three by using DPMM. As a result, the complicated forecast error data distribution is truthfully captured through the Bayesian nonparametric approach. Note that parameters in data-driven uncertainty sets are set such that levels of data coverage are the same.

The computational results are given in Table I. In terms of computational time, all the methods are comparable since the relatively small size of the test system allows all optimization problems to be solved within five seconds. The numbers of variables and constraints listed in the table are for the original multi-level adaptive robust UC problems. When solving the DDARUC problem based on DPMM, we need to decompose it into one master problem and three subproblems. Therefore, the computational time for DDARUC based on DPMM is the highest. The conventional method with the budget uncertainty set, ARUC, generates the highest total cost of $46,850. By contrast, the DDARUC method based on PCA and kernel smoothing enjoys a lower cost of $44,984 by leveraging the correlation information. By taking advantage of the powerful Bayesian nonparametric approach, the proposed DDARUC with DPMM significantly ameliorates the conservatism issue through

Fig. 4. The schedule of generators determined by three different methods.

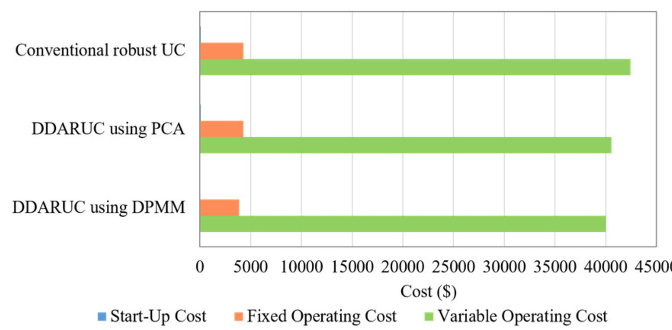


Fig. 5. The cost breakdowns of different adaptive robust UC methods.

the incorporation of intricate yet useful uncertainty information into the UC optimization process. More specifically, the total cost of the proposed DDARUC approach using DPMM is 6.34% and 2.46% less than those of conventional robust UC approach and DDARUC based PCA and kernel smoothing, respectively.

The robust day-ahead schedules of generators are summarized in Fig. 4. Observed from the schedule determined by the conventional adaptive robust UC approach, thermal generator G2 starts up at period 18, and then shuts down at period 20, which inevitably incurs additional costs. The DDARUC approach based on PCA and kernel smoothing commits generator G2 to provide power from period 17 to period 18. In contrast to the first two solutions, the schedule determined by the DDARUC method using DPMM keeps generator G2 off from period 17 to period 19.

The details on cost breakdowns are shown in Fig. 5. From the bar charts, we can readily see that most of the total cost comes from the variable operating cost. Moreover, the start-up, no-load, and fuel costs determined by the DDARUC approach using DPMM are all less than those of conventional UC method and DDARUC method based on PCA and kernel smoothing. The values of startup costs are very small so that it is difficult to read from Fig. 5. The startup costs for conventional adaptive robust UC, DDARUC using PCA, and DDARUC using DPMM are $100, $100, and $0, respectively. Compared with the conventional method, the two data-driven approaches lead to more cost-effective solutions by excavating more uncertainty information from data for UC decisions.

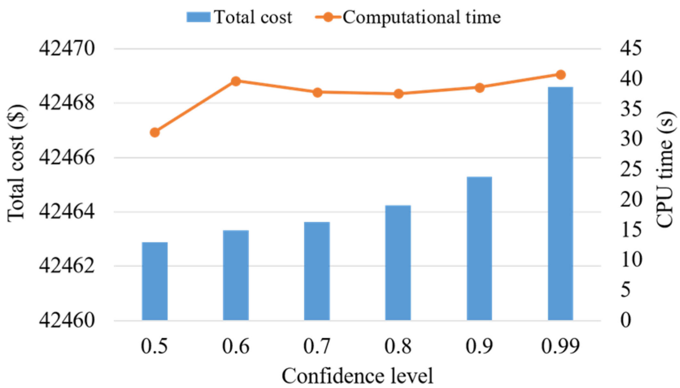


Fig. 6. The computational results of the DRO-based UC method.

To fully evaluate the effectiveness of the proposed approach, another 500 wind forecast error data are used as testing data. In the testing experiment or the out-of-sample simulation, the average cost of the conventional adaptive robust UC method is \$43,286, while its worst-case cost is \$44,456. In this case study, the DDARUC using PCA has the same average and worst-case costs as the conventional method in the testing experiment due to their similar first-stage decisions. By contrast, the average and worst-case costs of the proposed DDARUC approach using DPMM are \$42,453 and \$43,623, respectively. Therefore, the proposed approach compares favorably against the other adaptive robust UC methods for the test dataset, and the resulting UC decision leads to a better out-of-sample performance. The robustness of the solution or the system security is measured by the performance under the worst case among unrevealed uncertainty scenarios. The DDARUC approach using DPMM has a lower worst-case cost in the testing experiment compared with the other two methods, thus demonstrating its robustness.

2) Comparisons With Distributionally Robust Optimization-Based UC Method: The DRO method, also known as data-driven stochastic optimization, is recognized as an emerging data-driven optimization method for UC problems. In this subsection, we further compare the proposed DDARUC approach using DPMM with a DRO-based UC method [30]. In [30], the authors proposed two types of ambiguity sets based on L_1 norm and L_∞ norm. The DRO method using L_1 norm is more conservative than the DRO using L_∞ norm with the same amount of uncertainty data [30]. Therefore, we focus on the comparison with the DRO method using L_∞ norm to guarantee a fair comparison. In the DRO method, the decision maker needs to specify a confidence level, which is the probability of the true distribution being in the ambiguity set.

We present the computational results of the DRO-based UC method under various confidence levels in Fig. 6. As can be observed from the figure, the computational time varies from 31.2 s to 40.8 s, while the proposed DDARUC method only consumes 2.7 s. Accordingly, the proposed approach is more computationally efficient than the DRO-based method. From the bar chart, we can see that objective values of DRO-based UC problems are lower than that of DDARUC using DPMM. The reason is that the DRO method minimizes the worst-case expected cost, while DDARUC aims for the lowest worst-case cost. To evaluate their performance, we perform an out-of-sample simulation with the same test data. In this case study, the first-stage decisions generated by DRO remain the same for different confidence levels. The average cost is \$42,453, and the worst-case cost is \$43,623. Therefore, the proposed DDARUC method generates the same out-of-sample performance as the DRO-based approach in this case study. For general cases, the proposed DDARUC approach is superior in terms of worst-case cost, while the DRO method enjoys a better expected cost compared with the proposed approach.

TABLE II
COMPARISONS OF PROBLEM SIZES AND COMPUTATIONAL RESULTS FOR THE IEEE 118-BUS TEST SYSTEM

	ARUC	DDARUC-PCA	DDARUC-DPMM
Bin. Var.	3,888	3,888	3,888
Cont. Var.	4,591	4,591	4,709
Constraints	22,014	22,250	22,252
Min. cost (\$)	845,039	832,037	828,888
CPU time (s)	4,085.9	1,593.2	5,471.1

B. IEEE 118-Bus System

In this subsection, we consider a large-scale IEEE 118-bus test system to demonstrate the scalability and effectiveness of the proposed data-driven approach using DPMM. This system consists of 118 buses, 54 thermal generators, 186 transmission lines, and 91 loads. Moreover, three wind farms are installed at Buses 36, 69 and 77. As in the previous case study, a set of 5,000 wind forecast error data points is used to construct uncertainty sets. The other details can be found in [67], [68].

The problem sizes and computational results of different robust UC methods are summarized in Table II. The total cost determined by the conventional robust UC method is the highest among the three methods (\$845,039). Meanwhile, the DDARUC approach based on PCA and kernel smoothing generated a less conservative schedule by having a lower cost (\$832,037 vs \$845,039). In terms of solution quality, the DDARUC based on DPMM demonstrates a superior performance than the other two approaches consistently.

By comparing the CPU times, we can see the DDARUC approach using DPMM is less efficient compared with the other methods due to the increased number of subproblems, which are challenging to solve for large-scale UC problems [69]. It is worth noting that the DDARUC approach using DPMM scales gracefully with the problem size by solving the large-scale UC problem within a reasonable amount of time (less than 2 hours). Note that the solution time is dependent on the numerical structure of the mathematical programming problems. That is why the DDARUC approach using PCA consumes less computational time compared with the conventional robust UC in this case study.

To take a close look at the solution of DDARUC with DPMM, we present its 24-hour thermal generator output and wind power generation profile in Fig. 7. The total energy generation by thermal generators and wind farms changes according to the dynamic pattern of electricity loads across the scheduling horizon. Additionally, the percentages of renewable energy are more than 30% in the early morning (e.g., 1 am ~ 6am). It can be concluded from the economic dispatch result in Fig. 7 that the proposed approach promotes a high penetration of renewable wind power and guarantees more sustainable power systems operations.

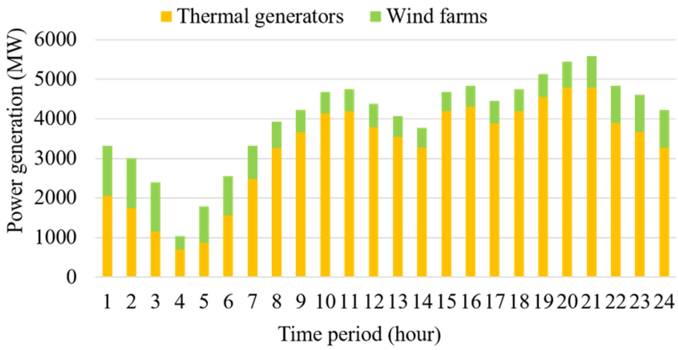


Fig. 7. The power generation profile of traditional thermal generators and wind power determined by the DDARUC method with DPMM.

To demonstrate computational efficiency of the proposed approach under larger number of wind farms, we consider a case with 60 wind farms. As the number of wind farms increases, the total cost is reduced to $210,276 because more wind energy can be utilized. The proposed data-driven approach works efficiently under larger number of wind farms, and the computational time is 9,250.8s. Solar power is quite different from wind power in terms of forecasting and frequency. Therefore, we focus on the wind power uncertainty.

VI. Conclusion

In this paper, we proposed a novel DDARUC optimization framework. A data-driven uncertainty set for wind forecast errors was developed by taking advantage of the DPMM. The proposed uncertainty set was a union of multiple basic uncertainty sets, which granted more flexibility in capturing the underlying intricate distribution. Experiments were performed on the six-bus system and IEEE 118-bus system. By exploiting available uncertainty data for UC decision making, both the DDARUC approach with DPMM and the DDARUC method using PCA and kernel smoothing compared favorably against the conventional robust UC approach in terms of total cost.

Acknowledgment

The authors would like to thank Natalia Luján Juncua for her effort in organizing data of the six-bus, and IEEE 118-bus systems.

References

[1] I. Dincer, "Renewable energy and sustainable development: A crucial review," *Renewable Sustain. Energy Rev.*, vol. 4, pp. 157–175, 2000.

[2] US Department of Energy, "20% wind energy by 2030. Increasing wind energy's contribution to U. S. electricity supply," U.S. Department of Energy, Oak Ridge, TN, USA, Rep. DOE/GO-102008-2567, 2008.

[3] L. Xie *et al.*, "Wind integration in power systems: Operational challenges and possible solutions," *Proc. IEEE*, vol. 99, no. 1, Jan. 2011, pp. 214–232.

[4] J. H. Wang, M. Shahidehpour, and Z. Y. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1319–1327, Aug. 2008.

[5] Q. P. P. Zheng, J. H. Wang, and A. L. Liu, "Stochastic optimization for unit commitment—A review," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1913–1924, Jul. 2015.

[6] 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.

[7] S. Takriti, J. R. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE Trans. Power Syst.*, vol. 11, no. 3, pp. 1497–1508, Aug. 1996.

[8] Q. F. Wang, Y. P. Guan, and J. H. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 206–215, Feb. 2012.

[9] M. Asensio and J. Contreras, "Stochastic unit commitment in isolated systems with renewable penetration under CVaR assessment," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1356–1367, May 2016.

[10] A. Papavasiliou, S. S. Oren, and R. P. O'Neill, "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2197–2206, Nov. 2011.

[11] L. Wu, M. Shabidehpour, and T. Li, "Stochastic security-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 800–811, May 2007.

[12] A. Street, A. Moreira, and J. M. Arroyo, "Energy and reserve scheduling under a joint generation and transmission security criterion: An adjustable robust optimization approach," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 3–14, Jan. 2014.

[13] A. Gupta and L. Anderson, "Statistical bus ranking for flexible robust unit commitment," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 236–245, Jan. 2019.

[14] X. A. Sun and Á. Lorca, "Robust optimization in electric power systems operations," in *Integration of Large-Scale Renewable Energy into Bulk Power Systems: From Planning to Operation*, P. Du, R. Baldick, and A. Tuohy, Eds. Cham, Switzerland: Springer, 2017, pp. 227–258.

[15] R. W. Jiang, M. H. Zhang, G. Li, and Y. P. Guan, "Two-stage network constrained robust unit commitment problem," *Eur. J. Oper. Res.*, vol. 234, pp. 751–762, 2014.

[16] N. Amjady, S. Dehghan, A. Attarha, and A. J. Conejo, "Adaptive robust network-constrained AC unit commitment," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 672–683, Jan. 2017.

[17] C. Lee, C. Liu, S. Mehrotra, and M. Shahidehpour, "Modeling transmission line constraints in two-stage robust unit commitment problem," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1221–1231, May 2014.

[18] C. X. Dai, L. Wu, and H. Y. 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.

[19] C. Wang *et al.*, "Robust risk-constrained unit commitment with large-scale wind generation: An adjustable uncertainty set approach," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 723–733, Jan. 2017.

[20] 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.

[21] Y. P. Guan and J. H. Wang, "Uncertainty sets for robust unit commitment," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1439–1440, May 2014.

[22] B. Hu, L. Wu, and M. Marwali, "On the robust solution to SCUC with load and wind uncertainty correlations," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2952–2964, Nov. 2014.

[23] C. Shao, X. Wang, M. Shahidehpour, X. Wang, and B. Wang, "Security-constrained unit commitment with flexible uncertainty set for variable wind power," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1237–1246, Jul. 2017.

[24] Y. Dvorkin, M. Lubin, S. Backhaus, and M. Chertkov, "Uncertainty sets for wind power generation," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 3326–3327, Jul. 2016.

[25] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Res.*, vol. 58, pp. 595–612, 2010.

[26] P. Xiong, P. Jirutitijaroen, and C. Singh, "A distributionally robust optimization model for unit commitment considering uncertain wind power generation," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 39–49, Jan. 2017.

[27] C. Y. Zhao and R. W. Jiang, "Distributionally robust contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 94–102, Jan. 2018.

[28] Y. W. Chen, Q. L. Guo, H. B. Sun, Z. S. Li, W. C. Wu, and Z. H. Li, "A distributionally robust optimization model for unit commitment based on Kullback-Leibler divergence," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5147–5160, Sep. 2018.

[29] C. Duan, L. Jiang, W. L. Fang, and J. Liu, "Data-driven affinely adjustable distributionally robust unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1385–1398, Mar. 2018.

[30] C. Y. Zhao and Y. P. Guan, "Data-driven stochastic unit commitment for integrating wind generation," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2587–2596, Jul. 2016.

[31] C. Shang and F. You, “Distributionally robust optimization for planning and scheduling under uncertainty,” *Comput. Chem. Eng.*, vol. 110, pp. 53–68, 2018.

[32] Y. L. Zhang, S. Q. Shen, and J. L. Mathieu, “Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads,” *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1378–1388, Mar. 2017.

[33] C. Duan, W. L. Fang, L. Jiang, L. Yao, and J. Liu, “Distributionally robust chance-constrained approximate AC-OPF with wasserstein metric,” *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4924–4936, Sep. 2018.

[34] W. J. Xie and S. Ahmed, “Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation,” *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1860–1867, Mar. 2018.

[35] C. Wang, R. Gao, F. Qiu, J. Wang, and L. Xin, “Risk-based distributionally robust optimal power flow with dynamic line rating,” *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6074–6086, Nov. 2018.

[36] Y. Guo, K. Baker, E. Dall’Anese, Z. Hu, and T. Summers, “Data-based distributionally robust stochastic optimal power flow, Part I: Methodologies,” *IEEE Trans. Power Syst*.

[37] Y. Guo, K. Baker, E. Dall’Anese, Z. Hu, and T. Summers, “Data-based distributionally robust stochastic optimal power flow, Part II: Case studies,” *IEEE Trans. Power Syst*.

[38] A. Zare, C. Y. Chung, J. Zhan, and S. O. Faried, “A distributionally robust chance-constrained MILP model for multistage distribution system planning with uncertain renewables and loads,” *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5248–5262, Sep. 2018.

[39] Q. Bian, H. Xin, Z. Wang, D. Gan, and K. P. Wong, “Distributionally robust solution to the reserve scheduling problem with partial information of wind power,” *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2822–2823, Sep. 2015.

[40] F. Qiu and J. H. Wang, “Distributionally robust congestion management with dynamic line ratings,” *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2198–2199, Jul. 2015.

[41] 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*.

[42] P. Xiong and C. Singh, “A distributional interpretation of uncertainty sets in unit commitment under uncertain wind power,” *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 149–157, Jan. 2019.

[43] P. M. Esfahani, and D. Kuhn, “Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations,” *Math. Program.*, vol. 171, no. 1–2, pp. 115–166, 2017.

[44] D. Bertsimas, V. Gupta, and N. Kallus, “Data-driven robust optimization,” *Math. Program.*, vol. 167, pp. 235–292, 2018.

[45] C. Shang, X. Huang, and F. You, “Data-driven robust optimization based on kernel learning,” *Comput. Chem. Eng.*, vol. 106, pp. 464–479, 2017.

[46] C. Ning and F. You, “Data-driven stochastic robust optimization: General computational framework and algorithm leveraging machine learning for optimization under uncertainty in the big data era,” *Comput. Chem. Eng.*, vol. 111, pp. 115–133, 2018.

[47] C. Ning and F. You, “A data-driven multistage adaptive robust optimization framework for planning and scheduling under uncertainty,” *AIChE J.*, vol. 63, pp. 4343–4369, 2017.

[48] J. Kiviluoma *et al.*, “Variability in large-scale wind power generation,” *Wind Energy*, vol. 19, pp. 1649–1665, 2016.

[49] G. Morales-Espana, A. Lorca, and M. M. de Weerdt, “Robust unit commitment with dispatchable wind power,” *Electric Power Syst. Res.*, vol. 155, pp. 58–66, 2018.

[50] R. W. Jiang, J. H. Wang, and Y. P. Guan, “Robust unit commitment with wind power and pumped storage hydro,” *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 800–810, May 2012.

[51] A. Sun and A. Lorca, “Adaptive robust optimization for daily power system operation,” in *Proc. Power Syst. Comput. Conf.*, 2014, pp. 1–9.

[52] C. L. Archer, H. P. Simao, W. Kempton, W. B. Powell, and M. J. Dvorak, “The challenge of integrating offshore wind power in the US electric grid. Part I: Wind forecast error,” *Renewable Energy*, vol. 103, pp. 346–360, 2017.

[53] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, “Wind power forecasting: State-of-the-art 2009,” Argonne National Laboratory (ANL)2009, Argonne, IL, USA, Rep. ANL/DIS-10-1, 2009.

[54] V. Gupta, “Data-driven models for uncertainty and behavior,” Ph.D. dissertation, Operations Research Center, MIT, Cambridge, MA, USA, 2014.

[55] H. Bludszuweit, J. A. Dominguez-Navarro, and A. Llombart, “Statistical analysis of wind power forecast error,” *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 983–991, Aug. 2008.

[56] C. Li, J. Y. Zhao, T. X. Zheng, and E. Litvinov, “Data-driven uncertainty sets: Robust optimization with temporally and spatially correlated data,” in *Proc. IEEE Power Energy Soc. General Meeting*, 2016, pp. 1–5.

[57] J. Sethuraman, “A constructive definition of dirichlet priors,” *Statistica Sinica*, vol. 4, pp. 639–650, 1994.

[58] T. Campbell and J. P. How, “Bayesian nonparametric set construction for robust optimization,” in *Proc. Amer. Control Conf.*, 2015, pp. 4216–4221.

[59] D. M. Blei and M. I. Jordan, “Variational inference for dirichlet process mixtures,” *Bayesian Anal.*, vol. 1, pp. 121–143, 2006.

[60] M. C. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York, NY, USA: Springer-Verlag, 2006.

[61] C. Ning and F. You, “Data-driven adaptive nested robust optimization: General modeling framework and efficient computational algorithm for decision making under uncertainty,” *AIChE J.*, vol. 63, pp. 3790–3817, 2017.

[62] A. Takeda, S. Taguchi, and R. H. Tutuncu, “Adjustable robust optimization models for a nonlinear two-period system,” *J. Optim. Theory Appl.*, vol. 136, pp. 275–295, 2008.

[63] B. Zeng and L. Zhao, “Solving two-stage robust optimization problems using a column-and-constraint generation method,” *Operations Res. Lett.*, vol. 41, pp. 457–461, 2013.

[64] D. Bertsimas and M. Sim, “The price of robustness,” *Operations Res.*, vol. 52, pp. 35–53, 2004.

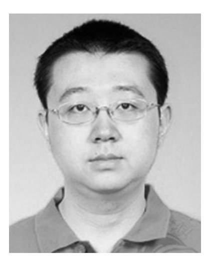
[65] C. Ning and F. You, “Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods,” *Comput. Chem. Eng.*, vol. 112, pp. 190–210, 2018.

[66] Y. Wang, Q. H. Hu, D. Y. Meng, and P. F. Zhu, “Deterministic and probabilistic wind power forecasting using a variational Bayesian-based adaptive robust multi-kernel regression model,” *Appl. Energy*, vol. 208, pp. 1097–1112, 2017.

[67] C. Sahin, M. Shahidehpour, and I. Erkmen, “Allocation of hourly reserve versus demand response for security-constrained scheduling of stochastic wind energy,” *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 219–228, Jan. 2013.

[68] G. Morales-España, “Unit commitment: Computational performance, system representation and wind uncertainty management,” Doctoral thesis, Pontifical Comillas University, Madrid, Spain, 2014.

[69] H. X. Ye, J. H. Wang, and Z. Y. Li, “MIP reformulation for max-min problems in two-stage robust SCUC,” *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1237–1247, Mar. 2017.



Chao Ning received the B.Eng. degree in automation from the University of Electronic Science and Technology of China, Chengdu, China, in 2012, and the M.S. degree in control science and engineering from Tsinghua University, Beijing, China, in 2015. He is currently working toward the Ph.D. degree at Cornell University, Ithaca, NY, USA. His research interests include data-driven optimization under uncertainty and machine learning.

Fengqi You (M’17) received the B.Eng. degree from Tsinghua University, Beijing, China, in 2005, and the Ph.D. degree from Carnegie Mellon University, Pittsburg, PA, USA, in 2009. He worked with the Argonne National Laboratory, Argonne, IL, USA, as an Argonne Scholar from 2009 to 2011. He was with the faculty of Northwestern University from 2011 to 2016. He is the Roxanne E. and Michael J. Zak Professor with Cornell University, Ithaca, NY, USA, and is affiliated with the graduate fields of chemical engineering, electrical and computer engineering, operations research and information engineering, civil and environmental engineering, mechanical engineering, Center of Applied Mathematics, and systems engineering. His research focuses on novel computational models, optimization algorithms, statistical machine learning methods, and multi-scale systems analysis tools for smart manufacturing, digital agriculture, energy systems, and sustainability.