A Data-Driven Model of Virtual Power Plants in Day-Ahead Unit Commitment

Sadra Babaei, *Student Member, IEEE*, Chaoyue Zhao, *Member, IEEE*, and Lei Fan, *Member, IEEE*

***Abstract*—Due to the increasing penetration of distributed energy resources (DERs), power system operators face significant challenges of ensuring the effective integration of DERs. The virtual power plant (VPP) enables DERs to provide their valuable services by aggregating them and participating in the wholesale market as a single entity. However, the available capacity of VPP depends on its DER outputs, which is time varying and not exactly known when the independent system operator runs the day-ahead unit commitment engine. In this study, we develop a model to evaluate the physical characteristics of the VPP, i.e., its maximum capacity and ramping capabilities, given the uncertainty in wind power output and load consumption. The proposed model is based on a distributionally robust optimization approach that utilizes moment information (e.g., mean and covariance) of the unknown parameter. We reformulate the model as a binary second-order conic program and develop a separation framework to address it. We first solve a two-stage problem and then benchmark it with a multi-stage case. Case studies are conducted to show the performance of the proposed approach.**

***Index Terms*—Electricity market, virtual power plant, distributionally robust optimization.**

NOMENCLATURE

A. Sets

$\mathcal{N}^g$	Set of generators.
$\mathcal{T}$	Set of time periods.

B. Parameters

SU_i	Start-up cost for generator i.
SD_i	Shut-down cost for generator i.
NL_i	No load cost for generator i.
CU_i^g	Maximal generation capacity for generator i.
CD_i^g	Minimal generation capacity for generator i.
UT_i	Minimum up-time for generator i.
DT_i	Minimum down-time for generator i.
RU_i	Ramp-up limit for generator i.
$\overline{\mathrm{RU}}_i$	Start-up ramp-up limit for generator i.
RD_i	Ramp-down limit for generator i.
$\overline{\mathrm{RD}}_i$	Shut-down ramp-down limit for generator i.
L_t^{in}	Load level inside the VPP in time period t.
$C_i^g(.)$	Fuel cost function of generator i.
x_0^s	Initial storage level.
η^{s+}	Discharging efficiency of storage unit.
η^{s-}	Charging efficiency of storage unit.
CU_t^{s+}	Maximum discharging level of storage unit in time period t.
CU_t^{s-}	Maximum charging level of storage unit in time period t.
CU^s	Capacity of storage unit.
λ_t^C	Capacity price in time period t.

C. First Stage Decision Variables

y_{it}^+	Binary variable to indicate if generator i is started up in time period t.
y_{it}^-	Binary variable to indicate if generator i is shut down in time period t.
y_{it}^o	Binary variable to indicate if generator i is on in time period t.
y_t^C	Capacity offered by a VPP to the ISO for time period t.
y_t^{RU}	Ramp-Up limit offered by a VPP to the ISO for time period t.
y_t^{RD}	Ramp-Down limit offered by a VPP to the ISO for time period t.

D. Second Stage Decision Variables

x_{it}^g	Power produced by generator i in time period t.
x_t^{s+}	Amount of power discharged by storage unit in time period t.
x_t^{s-}	Amount of power absorbed by storage unit in time period t.
x_t^{vpp}	Total power generation of a VPP in time period t.

E. Random parameter

ξ_t	Virtual net load consumption in time period t.

Manuscript received March 23, 2018; revised July 22, 2018 and October 5, 2018; accepted December 19, 2018. Date of publication January 1, 2019; date of current version October 24, 2019. This work of C. Zhao is supported by NSF 1662589 and NSF 1610935. Paper no. TPWRS-00426-2018. *(Corresponding author: Lei Fan.)*

S. Babaei and C. Zhao are with the School of Industrial Engineering and Management, Oklahoma State University, Stillwater, OK 74074 USA (e-mail: sadra.babaei@okstate.edu; chaoyue.zhao@okstate.edu).

L. Fan is with the Siemens Industry, Inc., Minneapolis, MN 55414 USA (e-mail: lei.fan@ieee.org).

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.2018.2890714

I. INTRODUCTION

WITH the influx of distributed energy resources (DERs), passive power networks are going through a transformation, from a centralized to a decentralized scheme, to enhance the flexibility and reliability of the system by offering more resources to the grid operator [1], [2]. However, since each individual distributed energy resource has small capacities and is lack of controllability over its outputs, small distributed energy

resource is either excluded from participating in the wholesale energy market or it is not able to participate in a cost-effective manner. Currently, Independent System Operators (ISOs) have limited control over the DERs connected to the grid, because most of these DERs are invisible to ISOs. Moreover, due to the computational limitation, ISOs are not able to simultaneously co-optimize schedules of conventional generators and a huge number of DERs across the grid. Accordingly, the development of an entity in the wholesale market as a market participant to represent and operate these DERs, becomes an important approach to facilitate the utilization of renewable energy resources and modernization of electricity grid. As an aggregator, the virtual power plant (VPP) [3] acts as an intermediary between the DERs and ISO, and allows small DERs to be pooled and actively participate in the wholesale energy markets. Such a bundled entity alleviates the ISO from having to obtain additional reserves or other ancillary service products for mitigating renewables intermittency. The concept of VPP is successfully implemented in Belgium and Netherlands [4], and is exploited in Germany and United Kingdom [5]. In July 2018, the first trial of VPP initiated by Tesla, has been successfully set up in South Australia, which is expected to generate 250 MW of solar power and 650 MWh of battery storage capacity [6]. Moreover, [7] investigates the method to estimate the operation and cost of VPP and report the study of a VPP project on an 11 KV system in Brixton by using the data from UK Power Network, and [8] reports the studies of VPPs across cities in the iURBAN project. However, although via aggregation, the VPP enlarges the visibility of DER units to ISOs in consideration of market participation and operation, it is still challenging to transform all the information of aggregated DERs into one bid with standard attributes such as maximum capacity, ramping restrictions, so that it can fit ISOs' bidding offer schema and the wholesale market's mechanism. In addition, submitting inaccurate parameters in the ISOs' market clear engine can jeopardize the grid operation. In this study, we provide a framework to characterize and evaluate the standard attributes/parameters in the VPP's bid submitted to the ISO that can optimize its entire portfolio as an aggregation of distributed energy resources.

Existing literature on the operation and scheduling of the VPP mostly seeks to optimize the dispatching of resources within the VPP so as to maximize its profitability. References [9], [10] consider the self-scheduling operation of the VPP, which means that the VPP only submits energy quantities to ISO, instead of quantity-price bid pairs. Price-based unit commitment models are proposed in [11], [12] to develop the bidding strategy for a VPP that behaves as a price-taker in the market, since the VPP is assumed to be a relatively small entity that has less impact on the market clearing price compared to the other market participants. In [13], a large scale of wind farm-energy storage system is studied as a price maker. The behavior of rival participants is taken into account using the residual demand curve. In [14] and [15], a bilevel problem for the optimal bidding strategy of VPP is addressed, in which the upper level aims to maximize the profit of VPP, and the lower level calculates the ISO day-ahead market clearing price. In [9], VPPs are allowed to establish bilateral contracts to hedge against the volatility of the electricity market. The idea of introducing VPPs as ancillary service providers is investigated in [16], [17]. A detailed literature review about the scheduling problem of VPP is presented in [18].

Another key direction of studying VPP focuses on managing the uncertain parameters like renewable generation output and load consumption. Stochastic programming (SP) has been extensively utilized for this purpose [19], [20]. Using this approach, the uncertain parameters are characterized by a set of scenarios based on the estimated probability distribution. For example, [21] studies a two-stage stochastic programming model for the optimal offering strategy of a VPP in the day-ahead and balancing markets. The uncertain wind power and market price are represented by a set of equi-probable scenarios. A major obstacle of SP is that fixing a particular probability distribution of the uncertain parameter may yield to biased solutions with unreliable out-of-sample performance. Moreover, SP usually suffers from the curse of dimensionality, which means that the computational difficulty surges exponentially in the number of scenarios, and makes it impractical to solve large scale problems. As an alternative approach, robust optimization (RO) aims at constructing an uncertainty set to characterize the uncertain parameter, and allows the uncertain parameter to run adversely within the constructed uncertainty set to guarantee the feasibility of the optimal solution. For instance, in [22], [23], confidence bounds are constructed for the uncertain wind and market price, and robust bidding strategy models are proposed for a VPP consisting of price-responsive demands, wind power plants, and storage units. However, the RO approach is criticized as its over-conservativeness since it ignores the probabilistic nature of unknown parameters and the solution is solely based on the worst-case scenario.

To cope with the limitations of stochastic and robust optimization approaches, distributionally robust (DR) optimization models have been developed (see e.g., [24], [25]). According to this approach, the probability distribution of the uncertain parameters is itself subject to uncertainty. In fact, the probability distribution is merely known to be within an ambiguity set, which can be characterized using certain statistical properties (e.g., estimation of mean and covariance). To guarantee the robustness of the approach, DR approach finds a solution that minimizes the worst-case expected cost over the ambiguity set. Unlike the traditional SP that exploits a collection of representative scenarios based on an estimated probability distribution to characterize the uncertain parameters, and thereby has no robustness to the error of distribution estimation, DR models release the assumption on any particular distribution. Therefore, this approach can accommodate the estimation error on the distribution due to the noisiness and incompleteness of the data, and also avoid the computational prohibition of scenario enumerations. Furthermore, contrary to the classical RO that is basically a distribution-free approach and minimize the total cost based on a worst-case scenario, DR models account for distributional knowledge through the ambiguity set, and minimizes the total expected cost based on a worst-case distribution over a set of probability distributions (ambiguity set). Hence, DR models trigger to less conservative solutions.

Because of these advantages, DR models have been successfully applied in power system problems under uncertainties, including contingency-constrained unit commitment [26], unit commitment with wind power integration [27], reserve scheduling [28], and optimal power flow [29]. However, most of the works fail to consider the physical limits (i.e., the support space) for the uncertain parameters, which is essentially critical in many applications. For example, renewable generators (e.g., wind turbines) have a limited capacity that cannot be exceeded from. The resulted solutions in these works, though restricted by the set of the distributions, are more conservative since the uncertain parameter itself can take very large positive or negative value even with its distribution still being within the range. Modeling of dependencies among the uncertain parameters (e.g., renewable energy outputs) is another crucial feature that is usually captured by including covariance matrix in the ambiguity set. Despite importance, there are limited studies in this regard. [28] stipulates that the covariance of uncertain renewable outputs exactly matches the empirical covariance obtained from data and formulates a semi-definite program, and [30] bounds the correlation between pairs of wind farms generations and proposes a second-order conic program.

This paper presents a two-stage DR model for the VPPs participation in the wholesale market. The VPP is a profit-driven entity that participates in the day-ahead wholesale market by submitting the cost curve and other related parameters like its maximum capacity and ramping limits to ISO. The goal is to find these standard attributes in the VPP's bid submitted to the ISO that can optimize its entire portfolio. The ISO collects such bidding information from all participants and run the reliability unit commitment to decide about the generation amount of each participant. Thus, the optimal solutions, i.e., the bidding parameter information of VPP will be served as an input for the ISO's unit commitment run. Furthermore, we consider two uncertainties in the model, i.e., the uncertain renewable energy output, and the unknown energy cleared by the ISO. We represent these two uncertainties with one parameter: the virtual net load, which is defined as the energy cleared by the ISO minus the renewable generation. It is assumed that in the first stage (day-ahead market), the VPP determines its total capacity and ramping limits to be reported to ISO that can optimize its entire portfolio by considering the physical constraints and the uncertain virtual net load. The second stage will give the first-stage a recourse, so that in the real-time operating day, after knowing the virtual net load, the VPP is able to supply enough power as it reported, by controlling its generation level of its conventional generators and power storage level. Using available moment information such as the empirical mean and covariance matrix of VPP's virtual net load that are learned from the data, we construct a second-order conic (SOC) representable ambiguity set for the unknown probability distribution, and reformulate the DR problem as a second-order conic programming (SOCP), which is efficiently solvable by off-the-shelf solvers like CPLEX. The objective of the model is to minimize the worst-case expected total cost over all probability distributions of the virtual net load in the ambiguity set. The conservativeness of the model can be adjusted based on the preference of the VPP operator. That is, if the VPP operator utilizes more information about the virtual net load data, the ambiguity set becomes smaller and the model becomes less conservative accordingly. On the contrary, if the VPP operator ignores the moment information on probability distribution and just utilizes the boundary information of the virtual net load (i.e., upper and lower bound), the proposed model is reduced to a traditional robust optimization model. Moreover, the proposed ambiguity set is able to capture the temporal dependencies among different time periods in the virtual net load profile by considering the covariance matrix.

A more realistic approach to model the VPP's problem is to allow the sequential revelation of the uncertain virtual net load and restrict the dispatch decisions to only hinge on the virtual net load observed up to the current time period. This restriction is called the non-anticipativity of dispatch decisions and the resulting formulation describes a multi-stage model. The main advantage of the multi-stage model over the two-stage approach is that the former framework caters for a dynamic decision making, where the VPP operator has the opportunity to update its knowledge about uncertain outcomes as they unfold in periods. Using linear decision rules, we extend the two-stage model to the multi-stage case, where recourse decisions take the form of a linear function of uncertain virtual net loads and a set of auxiliary variables.

The contributions of this study are summarized as follows:

1) We propose an innovative distributionally robust optimization model to help VPPs to optimally characterize the parameters in their bidding offers to ISO for the reliability unit commitment run.
2) The DR model can effectively manage the intrinsic uncertainty arising from virtual net load consumption. An tractable reformulation of the proposed DR model is derived, which can be implemented effectively by off-the-shelf solvers.
3) To better capture the nonanticipativity of the uncertainty, we extend the two-stage case to a multi-stage DR problem by using linear decision rules, and benchmark it with the two-stage DR model.

The rest of the paper is organized as follows. In Section II, we formulate the two-stage and multi-stage DR models, and describe the ambiguity set of virtual net load probability distributions. In Section III, we derive a separation framework to address the DR models. Finally, in Section IV, we demonstrate the effectiveness of two-stage and multi-stage models through several case studies.

II. Mathematical Formulation

We consider a VPP consisting of conventional generators, a wind farm, an energy storage unit, and non-flexible customers. Before submitting its bids to the day-ahead market, the VPP utilizes the moment information of virtual net load to construct an ambiguity set of virtual net load probability distribution. Using the ambiguity set and the physical characters of each conventional generators, the VPP determines the offering information needed to be submitted to the ISO, i.e., capacity, ramping limits, and cost curve in such a way that it minimizes the worst-case

expected cost. We present the DR model as follows:

$$\min_{\mathbf{y}} \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}^g}(\mathrm{SU}_i y_{it}^{+} + \mathrm{SD}_i y_{it}^{-} + \mathrm{NL}_i y_{it}^{o}) - \sum_{t\in\mathcal{T}} \lambda_t^C y_t^C + \max_{\mathbb{P}\in\mathcal{D}} E_{\mathbb{P}}[Q(\mathbf{y},\boldsymbol{\xi})], \tag{1a}$$

$$\text{s.t. } y_{it}^{o} - y_{i,t-1}^{o} = y_{it}^{+} - y_{it}^{-},\ \forall i\in\mathcal{N}^g,\ t\in\mathcal{T}, \tag{1b}$$

$$-y_{i,t-1}^{o} + y_{it}^{o} - y_{ik}^{o} \le 0, \quad \forall i\in\mathcal{N}^g,\ t\in\mathcal{T},\ 1\le k-(t-1)\le \mathrm{UT}_i, \tag{1c}$$

$$y_{i,t-1}^{o} - y_{it}^{o} + y_{ik}^{o} \le 1, \quad \forall i\in\mathcal{N}^g,\ t\in\mathcal{T},\ 1\le k-(t-1)\le \mathrm{DT}_i, \tag{1d}$$

$$y_t^C \le \sum_{i\in\mathcal{N}^g} \mathrm{CU}_i^g y_{it}^{o} - L_t^{in},\ \forall t\in\mathcal{T}, \tag{1e}$$

$$y_t^{RU} \le \sum_{i\in\mathcal{N}^g} (\mathrm{RU}_i y_{i,t-1}^{o} + \overline{\mathrm{RU}}_i y_{it}^{+}),\ \forall t\in\mathcal{T}, \tag{1f}$$

$$y_t^{RD} \le \sum_{i\in\mathcal{N}^g} (\mathrm{RD}_i y_{it}^{o} + \overline{\mathrm{RD}}_i y_{it}^{-}),\ \forall t\in\mathcal{T}, \tag{1g}$$

$$y_{it}^{o}, y_{it}^{+}, y_{it}^{-} \in \{0,1\},\ \ \forall i\in\mathcal{N}^g, t\in\mathcal{T}, \tag{1h}$$

where $\mathbf{y} := (\mathbf{y}^o, \mathbf{y}^+, \mathbf{y}^-, \mathbf{y}^C, \mathbf{y}^{RU}, \mathbf{y}^{RD})$ denotes first-stage decisions, $\mathcal{D}$ indicates the ambiguity set of virtual net load probability distribution, and $Q(\mathbf{y},\boldsymbol{\xi})$ represents the operating cost for a given first-stage decision $\mathbf{y}$ and realized virtual net load $\boldsymbol{\xi}$, and it can be calculated as follows:

$$Q(\mathbf{y},\boldsymbol{\xi}) = \min_{\mathbf{x}} \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}^g} C_i^g(x_{it}^g), \tag{2a}$$

$$\text{s.t. } x_t^{\text{vpp}} = \sum_{i\in\mathcal{N}^g} x_{it}^g + x_t^{s+} - x_t^{s-},\ \forall t\in\mathcal{T}, \tag{2b}$$

$$\xi_t \le x_t^{\text{vpp}} \le y_t^C,\ \ \forall t\in\mathcal{T}, \tag{2c}$$

$$x_t^{\text{vpp}} - x_{t-1}^{\text{vpp}} \le y_t^{RU},\ \ \forall t\in\mathcal{T}, \tag{2d}$$

$$x_{t-1}^{\text{vpp}} - x_t^{\text{vpp}} \le y_t^{RD},\ \ \forall t\in\mathcal{T}, \tag{2e}$$

$$\mathrm{CD}_i^g y_{it}^{o} \le x_{it}^g \le \mathrm{CU}_i^g y_{it}^{o},\ \forall i\in\mathcal{N}^g, t\in\mathcal{T}, \tag{2f}$$

$$x_{it}^g - x_{i,t-1}^g \le \mathrm{RU}_i y_{i,t-1}^{o} + \overline{\mathrm{RU}}_i y_{it}^{+},\ \forall i\in\mathcal{N}^g, t\in\mathcal{T}, \tag{2g}$$

$$x_{i,t-1}^g - x_{it}^g \le \mathrm{RD}_i y_t^{o} + \overline{\mathrm{RD}}_i y_{it}^{-},\ \forall i\in\mathcal{N}^g, t\in\mathcal{T}, \tag{2h}$$

$$0 \le \eta^{s+} x_t^{s+} \le \mathrm{CU}_t^{s+},\ \ \forall t\in\mathcal{T}, \tag{2i}$$

$$0 \le \eta^{s-} x_t^{s-} \le \mathrm{CU}_t^{s-},\ \ \forall t\in\mathcal{T}, \tag{2j}$$

$$0 \le x_0^s + \sum_{j\in[1:t]} \left(\eta^{s-} x_j^{s-} - \frac{1}{\eta^{s+}} x_j^{s+}\right) \le \mathrm{CU}^s, \forall t\in\mathcal{T}, \tag{2k}$$

where $\mathbf{x} := (\mathbf{x}^g, \mathbf{x}^{s+}, \mathbf{x}^{s-}, \mathbf{x}^{\text{vpp}})$ denotes the second-stage decisions, including the conventional generation amount, storage charging/discharging amount and total generation level. In the above formulation (1)–(2), the objective function (1a) is to minimize the worst-case expected total net cost, i.e., total cost minus the potential revenue by selling the capacity to the grid. The total cost includes start-up, shut-down, no load, and fuel costs. The potential revenue is assumed to be linearly depending on the value of offered capacity to ISO. Note that we do not put the potential revenue from selling power to the customers in the objective since it is a constant value in our model. Constraints (1b) indicate the commitment relationship among y_{it}^{+}, y_{it}^{-}, and y_{it}^{o}. Constraints (1c) and (1d) represent minimum up-time and minimum down-time limits, respectively. Constraints (1e) define an upper bound for the total capacity offered by VPP to ISO, which is the total available capacity, minus the load inside the VPP. Constraints (1f) and (1g) characterize the ramp-up and ramp-down limits offered by VPP to ISO. The objective of formulation (2) is to minimize the total operating cost, while respecting dispatch and storage related constraints. More precisely, $C_i^g(x_{it}^g)$ represents the quadratic fuel cost function corresponding to the generation level x_{it}^g, and it can be estimated by a N-piece-wise linear function as follows:

$$x_{it}^c \ge \delta_i^n y_{it}^{o} + \varrho_i^n x_{it}^g,\ \forall n = 1,\ldots,N,\ i\in\mathcal{N}^g,\ t\in\mathcal{T}. \tag{3}$$

Constraints (2b) calculate the total power generation by the VPP. Constraints (2c) describe that the total generation amount of VPP should satisfy the virtual net load, and also should not exceed the capacity level. Constraints (2d) and (2e) describe ramp-up and ramp-down limits for the VPP. Constraints (2f)-(2h) describe generation as well as ramping limits for the conventional generators inside the VPP. Constraints (2i)-(2j) describe upper bounds for the amount of discharge and charge levels. Finally, constraints (2k) enforce energy storage limits for the storage unit.

After solving the above models (1)–(2), the first-stage decisions including the capacity y_t^C and ramping limits y_t^{RU} and y_t^{RD} of the VPP can be found. As for the cost curve, since the storage unit is only responsible for smoothing out the scheduled generation levels by absorbing power when the demand is low and generating power in peak demand hours, it does not impact on the cost curve. Therefore, the cost curve submitted to the ISO will have the similar shape as the cost curve of conventional generators, except a shift because of the wind power output, as illustrated in Fig. 1. The reason for shifting the curve is that the conventional generators will be started up only when the wind power output is not enough to met the demand.

A. Abstract Formulation

For notation brevity, we recast the problem into the following compact matrix formulation:

$$\min_{\mathbf{y}\in\mathcal{Y}} \mathbf{c}^T\mathbf{y} + \max_{\mathbb{P}\in\mathcal{D}} E_{\mathbb{P}}[Q(\mathbf{y},\boldsymbol{\xi})], \tag{4a}$$

with

$$Q(\mathbf{y},\boldsymbol{\xi}) = \min_{\mathbf{x}} \{\mathbf{d}^T\mathbf{x} :\ \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{x} \ge \mathbf{h} - \mathbf{M}\boldsymbol{\xi}\}, \tag{4b}$$

where $\mathbf{c}^T\mathbf{y}$ indicates commitment costs minus capacity value, $\mathcal{Y}$ represents the first-stage constraints (1b)–(1h), $\mathbf{d}^T\mathbf{x}$ indicates the operating cost, and $\mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{x} \ge \mathbf{h} - \mathbf{M}\boldsymbol{\xi}$ represents the

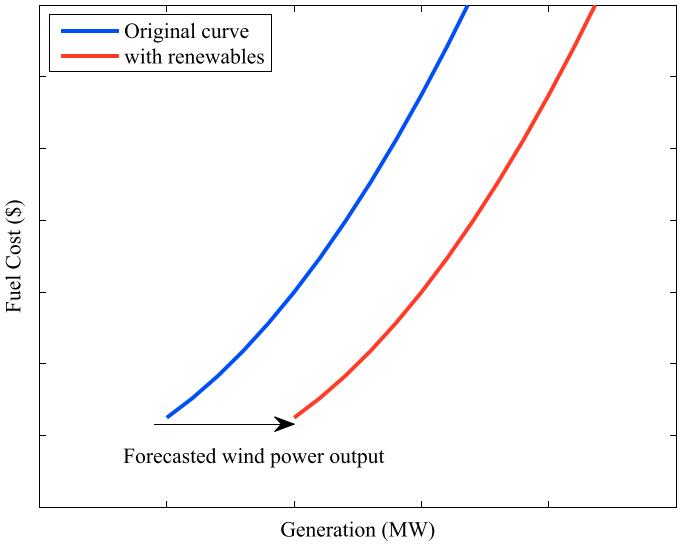


Fig. 1. An example of cost curve.

second-stage constraints (2a)–(2k) for a fixed commitment and offering decision $\mathbf{y}$ and virtual net load realization $\boldsymbol{\xi}$.

B. Ambiguity Set Construction

We construct an ambiguity set $\mathcal{D}$ of the virtual net load probability distribution by using its moment information. This ambiguity set can capture the dynamics of virtual net load evolution over time periods. Another main feature of this ambiguity set is that since it is a second-order conic representable set [31], it can be reformulated as a second-order conic program, which can be solved by many commercial solvers. More precisely, the ambiguity set is constructed as:

$$\mathcal{D} = \left\{ \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{\mathtt{T}}) \;\middle|\; \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \Omega) = 1, \\ E_{\mathbb{P}}[\boldsymbol{\xi}] = \boldsymbol{\mu}, \\ E_{\mathbb{P}}[(\boldsymbol{\xi} - \boldsymbol{\mu})^2] \leq \boldsymbol{\gamma}, \\ E_{\mathbb{P}}\left[\left(\sum_{i=k}^{t}(\xi_i - \mu_i)\right)^2\right] \leq \bar{\gamma}_{kt}, \\ \qquad\qquad \forall k \leq t, \end{array} \right\} \quad (5)$$

where $\mathcal{P}_0(.)$ is the set of all probability distributions, $\mathtt{T} = |\mathcal{T}|$ is the number of time periods, Ω is the support space of $\boldsymbol{\xi}$ and defined by $\Omega = [\underline{\boldsymbol{\xi}}, \overline{\boldsymbol{\xi}}]$, $\boldsymbol{\mu}$ is the estimation of the mean value of uncertain virtual net load $\boldsymbol{\xi}$, and parameters $\boldsymbol{\gamma}$ and $\bar{\gamma}$ are used to adjust the conservativeness of the optimal solution, which can capture the VPP operator's risk attitude. A risk-averse VPP operator may select larger values for these parameters to enrich the ambiguity set with more distributions, and thus arrive at a more robust solution. On the other hand, a risk-prone VPP operator may tend to choose smaller values of $\boldsymbol{\gamma}$ and $\bar{\gamma}$ to exclude pathological distributions with the moment information far away from the sample ones, and thus obtain less conservative solutions. In the above set (5), the first constraint incorporates a range for the virtual net load. The second constraint ensures that $\boldsymbol{\xi}$, given by any distribution in $\mathcal{D}$, has the same mean as the empirical mean, while the third constraint bound the variance of $\boldsymbol{\xi}$. Finally, the last constraint captures the virtual net load correlations across time periods. Indeed, this constraint ensures that the variance of sum of the virtual net loads during the time window $[k, t]$ is bounded by $\bar{\gamma}_{kt}$. In practice, the ambiguity set (5) is characterized by parameters $\underline{\boldsymbol{\xi}}, \overline{\boldsymbol{\xi}}, \boldsymbol{\mu}, \boldsymbol{\gamma}$, and $\bar{\gamma}$, which can be learned from historical data. Consider N data samples $\{\boldsymbol{\xi}^\ell\}_{\ell=1}^N$ of $\boldsymbol{\xi}$ such that $\boldsymbol{\xi}^\ell = [\xi_1, \xi_2, ..., \xi_{\mathtt{T}}]$, $\underline{\boldsymbol{\xi}}$ and $\overline{\boldsymbol{\xi}}$ can be set to be .05- and .95-quantiles of random virtual net load, respectively, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ can be chosen as the sample mean $\boldsymbol{\mu} = N^{-1}\sum_{\ell=1}^N \boldsymbol{\xi}^\ell$ and sample variance $\boldsymbol{\gamma} = N^{-1}\sum_{\ell=1}^N (\boldsymbol{\xi}^\ell - \boldsymbol{\mu})^2$. $\bar{\gamma}_{kt}$ can be determined by summing up specific elements in the covariance matrix of $\boldsymbol{\xi}$, i.e., $\bar{\gamma}_{kt} = \mathbf{f}_{kt}^T \boldsymbol{\Sigma} \mathbf{f}_{kt}$, where $\mathbf{f}_{kt}$ is a vector with i^{th} element equals to one if i is within the time window $[k, t]$, and zero otherwise, and $\boldsymbol{\Sigma}$ denotes sample covariance matrix $\boldsymbol{\Sigma} = N^{-1}\sum_{\ell=1}^N (\boldsymbol{\xi}^\ell - \boldsymbol{\mu})(\boldsymbol{\xi}^\ell - \boldsymbol{\mu})^T$.

As explained in [32], in order to derive a tractable reformulation for the DR problems with the ambiguity set (5), it is equivalent to work with the following lifted ambiguity set:

$$\overline{\mathcal{D}} = \left\{ \mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^{\mathtt{T}\times\mathtt{T}\times\frac{\mathtt{T}(\mathtt{T}+1)}{2}}\right) \;\middle|\; \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \overline{\Omega}) = 1, \\ E_{\mathbb{P}}[\boldsymbol{\xi}] = \boldsymbol{\mu}, \\ E_{\mathbb{P}}[\mathbf{u}] \leq \boldsymbol{\gamma} \\ E_{\mathbb{P}}[v_{kt}] \leq \bar{\gamma}_{kt}, \; \forall k \leq t, \end{array} \right\} \quad (6)$$

where $\overline{\Omega}$ is the lifted support set and defined as follows:

$$\overline{\Omega} = \left\{ (\boldsymbol{\xi}, \mathbf{u}, \mathbf{v}) \;\middle|\; \begin{array}{l} \underline{\boldsymbol{\xi}} \leq \boldsymbol{\xi} \leq \overline{\boldsymbol{\xi}}, \\ (\boldsymbol{\xi} - \boldsymbol{\mu})^2 \leq \mathbf{u}, \\ \left(\sum_{i=k}^{t}(\boldsymbol{\xi}_i - \boldsymbol{\mu}_i)\right)^2 \leq v_{kt}, \; \forall k \leq t, \end{array} \right\} \quad (7)$$

That is, $\mathcal{D}$ includes set of marginal distributions of $\boldsymbol{\xi}$, where the joint distribution of $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v})$ is in $\overline{\mathcal{D}}$.

C. Multi-Stage Formulation

The main assumption of the two-stage model is that all of the real-time dispatch decisions are made simultaneously at the beginning of the operating day. However, the unit commitment problem is inherently sequential. That means the uncertain virtual net load is revealed as the time progresses, and dispatch decisions at each period are made after knowing the realization of uncertain parameters up to that period (non-anticipativity enforcement). In other words, the VPP operator first observes the uncertain virtual net load of the first period $\boldsymbol{\xi}_1$, and then takes real-time decisions $\mathbf{x}_1(\boldsymbol{\xi}_1)$ of the first period. Subsequently, the virtual net load of second period $\boldsymbol{\xi}_2$ is realized, and then the VPP operator takes real-time decisions $\mathbf{x}_2(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ of the second period accordingly. This alternating process continues over the entire $\mathtt{T}$ periods. We present the multi-stage DR model as follows:

$$\min_{\mathbf{y}\in\mathcal{Y}, \mathbf{x}(.)} \left(\mathbf{c}^T\mathbf{y} + \max_{\mathbb{P}\in\overline{\mathcal{D}}} E_{\mathbb{P}}\left[\mathbf{d}^T\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]})\right] \right) \quad (8a)$$

$$\text{s.t. } \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]}) \geq \mathbf{h} - \mathbf{M}\boldsymbol{\xi}_{[t]}, \; \forall (\boldsymbol{\xi}, \mathbf{u}) \in \overline{\Omega}, \quad (8b)$$

where $\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]})$ denotes that the dispatch decision at time period t is a function of the uncertain virtual net load as well as the auxiliary random variable associated with the lifted ambiguity set realized up to time period t, and $\min_{\mathbf{x}(.)}$ can be interpreted as

optimizing over the policies, i.e., functions of random variables. Solving problem (8) is challenging since dispatch decisions are general functions of all past uncertain parameter realizations, instead of a finite vector of decision variables [33]. An effective approach on addressing this fully adaptive problem is to apply the linear decision rule technique, which restricts the dispatch decisions to be linear functions of the uncertain parameters ([27], [34], [35]). More precisely, we define the following policies for decision variables $x^g_{it}, x^{s+}_{it}, x^{s-}_{it}, x^{\text{vpp}}_{it}$:

$$x_{it}\left(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]}\right) = x^0_{it} + x^1_{it}\xi_t + x^2_{it}u_t. \tag{9}$$

Notice that linear coefficients $(\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2)$ are considered as decision variables in the problem and they are defined separately for decision variable $x^g_{it}, x^{s+}_{it}, x^{s-}_{it}, x^{\text{vpp}}_{it}$. Furthermore, in these policies it is assumed that the second-stage decision variables are affine functions of both primary and auxiliary random parameters. [32] has shown that this enhanced linear decision rule can significantly improve the computational results compared with the one restricting to the primary random parameters only. In addition, with the decision rule defined in (9), the nonanticipativity requirement can be met automatically.

III. Solution Methodology

In this section, we develop solution methods for the two-stage DR problem (4) and multi-stage DR problem (8), respectively.

A. Solution Method for Two-Stage DR

First, we dualize the second-stage worst-case expectation problem $\max_{\mathbb{P}\in\overline{\mathcal{D}}} E_{\mathbb{P}}[Q(\mathbf{y},\boldsymbol{\xi})]$ and obtain the following dual problem:

$$\min_{\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\alpha} \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{10a}$$

$$\text{s.t. } \eta \geq F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}), \tag{10b}$$

with

$$F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) = \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} Q(\mathbf{y},\boldsymbol{\xi}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{10c}$$

where $\eta, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha}$ are dual variables associated with the constraints in the ambiguity set (6). Note that we can add a slack variable in constraint (2c) to ensure the feasibility of the second-stage problem and in return consider a penalty cost for under-generation in the objective function. Since the second-stage linear problem $Q(\mathbf{y},\boldsymbol{\xi})$ is feasible and bounded, strong duality holds and $Q(\mathbf{y},\boldsymbol{\xi})$ can be replaced by its dual formulation. Hence, the subproblem (10c) can be equivalently formulated as follows:

$$F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) = \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \max_{\pi\in\Pi} \boldsymbol{\pi}^T(\mathbf{h} - \mathbf{M}\boldsymbol{\xi} - \mathbf{B}\mathbf{y}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{11}$$

where $\Pi = \{\boldsymbol{\pi} \geq 0 : \boldsymbol{\pi}^T\mathbf{B} = \mathbf{d}\}$ is the set of dual variables of $Q(\mathbf{y},\boldsymbol{\xi})$. Since the feasible regions Π and $\overline{\Omega}$ are separable, the optimal solution of problem (11) occurs at extreme points of these regions. Therefore, if we denote all the extreme points of Π as $\{\boldsymbol{\pi}^*_i\}_{i=1}^I$ and exchange the order of two maximization operations, i.e., $\max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}}$ and $\max_{\pi\in\Pi}$, $F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha})$ can be further reformulated as follows:

$$F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) = \max_{\pi_i,\forall i} \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \boldsymbol{\pi}^{*T}_i(\mathbf{h} - \mathbf{M}\boldsymbol{\xi} - \mathbf{B}\mathbf{y}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}. \tag{12}$$

Since the inner maximization problem in (12) is bounded with non-empty interior, conic duality can be applied. Taking dual of inner maximization problem in (12) leads to (details are provided in the Appendix A):

$$\begin{aligned} F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha}) = \max_{\pi_i,\forall i}\ & \boldsymbol{\pi}^{*T}_i(\mathbf{h} - \mathbf{B}\mathbf{y}) \\ & + \min_{\boldsymbol{\psi}}\ \bar{\boldsymbol{\xi}}^T\bar{\boldsymbol{\tau}} - \underline{\boldsymbol{\xi}}^T\tilde{\boldsymbol{\tau}} - \boldsymbol{\mu}^T\bar{\boldsymbol{\theta}} + \frac{1}{2}\mathbf{1}^T(\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}) \\ & - \sum_{t\in\mathcal{T}}\sum_{k\in[1:t]}\sum_{i\in[k:t]} \mu_i\bar{\phi}_{kt} + \frac{1}{2}\mathbf{1}^T(\hat{\boldsymbol{\phi}} - \tilde{\boldsymbol{\phi}}), \end{aligned} \tag{13a}$$

$$\text{s.t. } \tilde{\tau}_t - \bar{\tau}_t + \bar{\theta}_t + \bar{\phi} = \mathbf{e}_t{}^T\left(\mathbf{M}^T\boldsymbol{\pi}^*_i + \boldsymbol{\lambda}\right),\ \forall t\in\mathcal{T}, \tag{13b}$$

$$\frac{1}{2}(\tilde{\theta}_t + \hat{\theta}_t) = \beta_t,\ \forall t\in\mathcal{T}, \tag{13c}$$

$$\frac{1}{2}(\tilde{\phi}_{kt} + \hat{\phi}_{kt}) = \alpha_{kt},\ \forall t\in\mathcal{T}, k\leq t, \tag{13d}$$

$$\sqrt{\bar{\theta}^2_t + \tilde{\theta}^2_t} \leq \hat{\theta}_t,\ \forall t\in\mathcal{T}, \tag{13e}$$

$$\sqrt{\bar{\phi}^2_{kt} + \tilde{\phi}^2_{kt}} \leq \hat{\phi}_{kt},\ \forall t\in\mathcal{T}, k\leq t, \tag{13f}$$

where $\mathbf{1}$ is a vector whose elements are all one, $\boldsymbol{e}_t$ is an unit vector whose t^{th} component is one and zero otherwise, and $\boldsymbol{\psi} := \{\bar{\boldsymbol{\tau}}, \tilde{\boldsymbol{\tau}}, \bar{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}, \bar{\boldsymbol{\phi}}, \tilde{\boldsymbol{\phi}}, \hat{\boldsymbol{\phi}}\}$ are dual variables associated with the constraints in the inner maximization, i.e., constraints in $\overline{\Omega}$ (7). We rewrite the inner minimization problem in (13) in a compact form: $\min\{\mathbf{b}^T\boldsymbol{\psi} : \mathbf{E}\boldsymbol{\psi} \succeq_\kappa \mathbf{f}\boldsymbol{\alpha} + \mathbf{g}\boldsymbol{\beta} + \mathbf{p}\boldsymbol{\lambda} + \mathbf{q}\boldsymbol{\pi}^*\}$, where $\succeq_\kappa$ denotes the generalized inequality respect to some cone κ. Now, by considering the second-stage dual formulation (10) and replacing $F(\mathbf{y},\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha})$ in (10c) with formulation (13), the original DR optimization problem (4) is equivalent to:

$$\min_{\mathbf{y},\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\psi}_i} \mathbf{c}^T\mathbf{y} + \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{14a}$$

$$\text{s.t. } \eta \geq \boldsymbol{\pi}^{*T}_i(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}_i,\ \forall i = 1,\ldots,I, \tag{14b}$$

$$\mathbf{E}\boldsymbol{\psi}_i \succeq_\kappa \mathbf{f}\boldsymbol{\alpha} + \mathbf{g}\boldsymbol{\beta} + \mathbf{p}\boldsymbol{\lambda} + \mathbf{q}\boldsymbol{\pi}^*_i,\ \forall i = 1,\ldots,I, \tag{14c}$$

$$\mathbf{y}\in\mathcal{Y}, \tag{14d}$$

where $\boldsymbol{\psi}_i$ is a vector of decisions corresponding to the extreme point $\boldsymbol{\pi}^*_i$. Notice here that in (14b), we release the maximization operation in (13a) since it is equivalent to restrict the constraint (14b) holds for every extreme point $\boldsymbol{\pi}^*_i$, and we release the minimization operation in (13a) since it is equivalent to the existence of a feasible solution of constraint (14b) with constraint (14c). The above problem (14) has an appropriate structure to apply a two-level decomposition algorithm. More precisely, in the kth

Algorithm 1: Solution Procedure for Problem (14).

1: Initialize iteration index $i = 0$, and set $\mathcal{I} = \emptyset$.
2: **repeat**
3: Solve the master problem (15). Let $(\mathbf{y}, \eta, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ be the optimal solution.
4: Solve the subproblem problem $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ in (11) using the Algorithm 2. Let $\boldsymbol{\pi}^*_{i+1}$ be the optimal solution.
5: $\mathcal{I} = \mathcal{I} \cup \{\boldsymbol{\pi}^*_{i+1}\}$.
6: Define the new variable $\boldsymbol{\psi}_{i+1}$ and the corresponding constraints (15b)–(15c).
7: $i = i + 1$.
8: **until** $\eta \geq F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$.

Output: $(\mathbf{y}, \eta, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ is the optimal solution.

iteration, we solve the following master problem:

$$\min_{\mathbf{y},\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\psi}_i} \ \mathbf{c}^T\mathbf{y} + \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{15a}$$

$$\text{s.t. } \eta \geq \boldsymbol{\pi}_i^{*T}(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}_i, \ \forall i \leq k, \tag{15b}$$

$$\mathbf{E}\boldsymbol{\psi}_i \succeq_\kappa \mathbf{f}\boldsymbol{\alpha} + \mathbf{g}\boldsymbol{\beta} + \mathbf{p}\boldsymbol{\lambda} + \mathbf{q}\boldsymbol{\pi}^*_i, \ \forall i \leq k, \tag{15c}$$

$$\mathbf{y} \in \mathcal{Y}, \tag{15d}$$

to obtain the first-stage decisions. Then, the master problem (15) can be augmented iteratively (i.e., adding new variables $\boldsymbol{\psi}_i$ and the corresponding cuts) using the information provided by the subproblem (11). The more detailed procedure for solving the problem (14) is summarized in Algorithm 1.

Observe that since the feasible region Π includes finite number of extreme points, the Algorithm 1 converges in finite number of steps. Details are provided in the Appendix B. Also, notice that the Algorithm 1 is a natural extension of the column-and-constraint generation algorithm proposed in [36], where the polyhedron uncertainty set is extended to a SOC set. A similar extension is proposed in [37] to solve the robust AC optimal power flow problem.

To solve the subproblem, we need to evaluate $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ at each iteration of Algorithm 1, which is a second-order conic problem with a quadratic objective function. We employ the Alternative Separation Heuristic (ASH) to solve this problem. The basic idea behind this framework is to obtain the optimal $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v})$ of $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ with a fixed $\boldsymbol{\pi}$, and then fixed the obtained $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v})$ to find the optimal $\boldsymbol{\pi}$ of $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$. This back and forth procedure continues until the optimality gap is no more than a predefined level. Since the feasible region Π is a polyhedron, the ASH converges to a KKT point of (11) in a finite number of iterations [37], [38]. The ASH framework is summarized in Algorithm 2.

B. Solution Method for Multi-Stage DR

In order to solve the multi-stage DR model (8) using linear decision rules (9), we can apply a duality-based approach to reformulate the problem. More specifically, similar to the previous approach we start from dualizing the worst-case expectation problem $\max_{\mathbb{P}\in\overline{\mathcal{D}}} E_{\mathbb{P}}[Q(\mathbf{y}, \boldsymbol{\xi})]$. Thus, (8) is equivalent to:

Algorithm 2: Alternative Separation Heuristic for Solving Subproblem (11).

1: Pick a $\widehat{\boldsymbol{\pi}} \in \Pi$, and optimality gap $\widehat{\epsilon}$
2: **repeat**
3: Fix $\boldsymbol{\pi} = \widehat{\boldsymbol{\pi}}$ and solve (11). Let $(\widehat{\boldsymbol{\xi}}, \widehat{\mathbf{u}}, \widehat{\mathbf{v}})$ be the optimal solution with objective value Υ_1.
4: Fix $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v}) = (\widehat{\boldsymbol{\xi}}, \widehat{\mathbf{u}}, \widehat{\mathbf{v}})$ and solve (11). Let $\widehat{\boldsymbol{\pi}}$ be the optimal solution with objective value Υ_2.
5: **until** $|\Upsilon_1 - \Upsilon_2| \leq \widehat{\epsilon}$.

Output: Υ_1 is the estimation of $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ with solution $\widehat{\boldsymbol{\pi}}$.

$$\min_{\mathbf{y}\in\mathcal{Y},\mathbf{x}^0,\mathbf{x}^1,\mathbf{x}^2} \mathbf{c}^T\mathbf{y} + \min_{\eta,\boldsymbol{\lambda},\boldsymbol{\beta},\alpha} \eta + \boldsymbol{\mu}^T\boldsymbol{\lambda} + \boldsymbol{\gamma}^T\boldsymbol{\beta} + \bar{\boldsymbol{\gamma}}^T\boldsymbol{\alpha} \tag{16a}$$

$$\text{s.t. } \eta \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \ \mathbf{d}^T\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]}) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{16b}$$

$$\mathbf{A}\mathbf{y} - \mathbf{h} \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} -\mathbf{B}\mathbf{x}(\boldsymbol{\xi}_{[t]}, \mathbf{u}_{[t]}) - \mathbf{M}\boldsymbol{\xi}. \tag{16c}$$

By substituting (9) into the above formulation, constraints (16b) and (16c) become:

$$\eta \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} \mathbf{d}^T(\mathbf{x}^0 + \boldsymbol{\xi}^T\mathbf{x}^1 + \mathbf{u}^T\mathbf{x}^2) - \boldsymbol{\lambda}^T\boldsymbol{\xi} - \boldsymbol{\beta}^T\mathbf{u} - \boldsymbol{\alpha}^T\mathbf{v}, \tag{17a}$$

$$\mathbf{A}\mathbf{y} - \mathbf{h} \geq \max_{(\boldsymbol{\xi},\mathbf{u},\mathbf{v})\in\overline{\Omega}} -\mathbf{B}(\mathbf{x}^0 + \boldsymbol{\xi}^T\mathbf{x}^1 + \mathbf{u}^T\mathbf{x}^2) - \mathbf{M}\boldsymbol{\xi}. \tag{17b}$$

Note that constraints (17a) and (17b) have the same structure. In the following, we derive a reformulation for the constraint (17a). Similar approach can be done for constraint (17b). Applying strong duality, (17a) can be written as:

$$\eta - \mathbf{d}^T\mathbf{x}^0 \geq \min_{\boldsymbol{\chi}\in\Xi(\mathbf{d},\mathbf{x}^1,\mathbf{x}^2)} \mathbf{g}^T\boldsymbol{\chi}, \tag{18}$$

where $\boldsymbol{\chi}$ and $\Xi(\mathbf{d}, \mathbf{x}^1, \mathbf{x}^2)$ are the corresponding dual variables and dual feasible region of the constraints in $\overline{\Omega}$ with respect to $(\boldsymbol{\xi}, \mathbf{u}, \mathbf{v})$, and $\mathbf{g}$ is also derived from the definition of $\overline{\Omega}$. Note that since the second-order cone is self-dual, feasible region Ξ has a similar structure to $\overline{\Omega}$.

Relation (18) is further equivalent to the existence of $\boldsymbol{\chi} \in \Xi(\mathbf{d}, \mathbf{x}^1, \mathbf{x}^2)$ such that $\eta - \mathbf{d}^T\mathbf{x}^0 \geq \mathbf{g}^T\boldsymbol{\chi}$. Thus, using duality-based approach, constraints (17a)–(17b) can be reformulated as a finite number of linear and second-order conic constraints. As a result, the original problem (16) can be reformulated as a single minimization problem involving new added variables and constraints.

IV. Case Study

In this section, we first apply the two-stage DR model based on a real data set to provide the operation of the VPP. Second, we test the performance of two-stage and multi-stage DR models

TABLE I
GENERATOR DATA

Gen	Lower (MW)	Upper (MW)	Min. down (h)	Min. up (h)	Ramp (MW/h)
G1	5	20	1	1	10
G2	5	50	2	2	25
G3	50	100	2	3	50

TABLE II
FUEL DATA

Gen	a (MBtu)	b (MBtu/ MWh)	c (MBtu/ MW^2h)	Start Up Fuel (MBtu)	Fuel Price ($/MBtu)
G1	31.67	29.24	0.0697	40	1
G2	58.81	22.94	0.0098	60	1
G3	50	6	0.0004	100	1

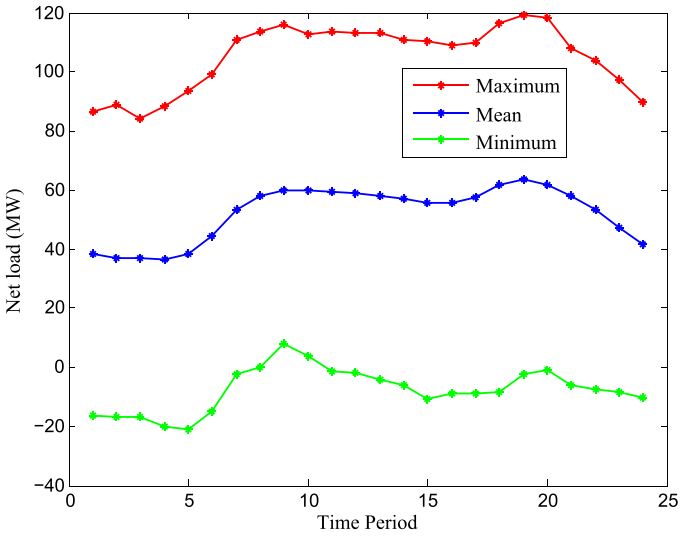


Fig. 2. Virtual net load profile.

for various simulated data sets. All the experiments are implemented in the C++ language with CPLEX 12.6 on a computer with Intel Xeon 3.2 GHz and 8 GB memory.

A. Data Preparation

We consider a VPP including three conventional generators, a wind farm, a storage unit, and a set of loads that all are located in a single bus in the system. The characteristics of the generators are shown in Tables I–II. The load and wind power outputs are collected from PJM market website [39] and scaled down to fit the test system.

The mean value together, upper and lower limits of the virtual net load profile are depicted in Fig. 2. We assume that the capacity and efficiency of the storage unit on both absorbing and generating electricity are 100 MW and 0.9, respectively. The load consumption inside the VPP and the capacity price are randomly simulated from intervals $[4, 20]$MWs and \$[5,30]/MW, respectively. The penalty cost for load shedding is set to be \$6000/MWh. In addition, for the comparative studies in Sections IV-D and IV-C, we consider various numerical settings to evaluate the performance of DR models. More precisely, we randomly generate 5,000 samples of virtual net load consumption $\boldsymbol{\xi}$ such that each ξ_t follows a normal distribution with mean

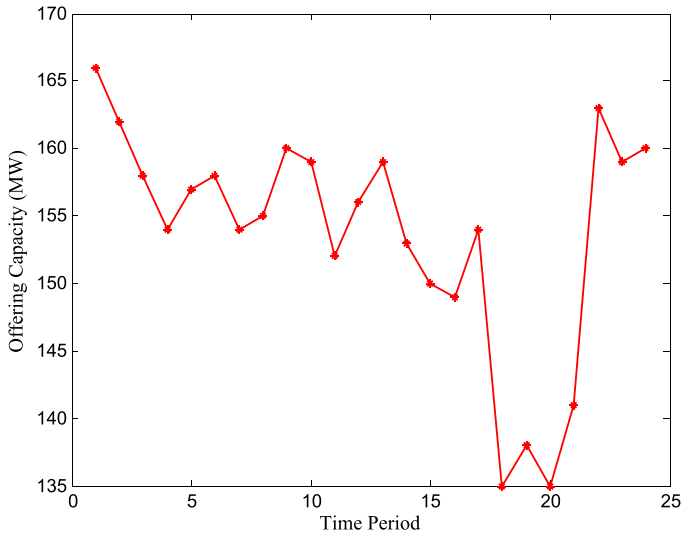


Fig. 3. Capacity offered to ISO.

$\mu_t \in [50, 80]$, standard deviation $\sigma_t = \mu_t \varepsilon$ with $\varepsilon \in \{0.1, 0.2\}$, and correlation coefficient $\zeta \in \{0, 0.25, 0.5, 0.75, 1\}$. These 5,000 samples comprise the support set Ω, which will be used later to build the ambiguity set.

B. VPP Offering Parameters

We solve the two-stage DR model to determine the VPP offering parameters. The offering capacity at each time period is reported in Fig. 3. The optimal ramp-up and ramp-down parameters are obtained 75 MW/h for periods 19, 20, and 21, and 85 MW/h for other periods. As it can be seen in Fig. 3, there is a significant reduction in offering capacity at time periods 18 and 20. The reason is that the capacity price at these periods suddenly drops to its minimum level. To recover this unexpectedly low capacity price, the generator $G1$, which is more expensive than $G2$ and $G3$, is turned off at these periods. On the other hand, the VPP offers higher level of capacity at time periods like 1,2, and 22 since the inside load is low at these periods. In order to investigate the performance of the DR model in the real time operations, we fixed the optimal unit commitment and offering decisions and solve the dispatch problem for a realized virtual net load. The total power generation by conventional generators, storage operation, and the realized virtual net load are shown in Fig. 4. We can observe that the storage unit absorbs the power during the periods that the virtual net load is low (e.g., periods 5–6), and generates the power during the periods that the virtual net load is high (e.g., 8–10). More specifically, the storage unit absorbs 70MW power at period 5 (valley demand hour) when the wind power output is higher than the grid demand, and generates 43MW at period 17 to preclude from load shedding when the expensive generator $G1$ is offline at this period. Thus, we can conclude that the storage unit significantly contributes to the VPP flexibility.

C. Comparing With Robust Optimization

In this subsection, we compare the two-stage DR (TDR) model with the RO approach. For this purpose, since the VPP

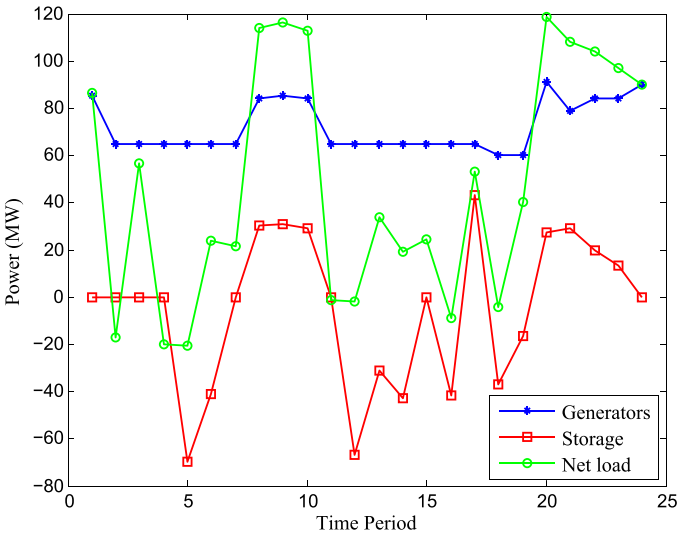


Fig. 4. Optimal dispatch decisions.

TABLE III
TWO-STAGE DR VS. TWO-STAGE RO

	$\varepsilon = 0.1$		$\varepsilon = 0.2$	
ζ	TDR	RO	TDR	RO
$\zeta = 0$	19169	19701	24114	24461
$\zeta = 0.25$	19367	19715	24949	24956
$\zeta = 0.5$	19215	19730	24701	25041
$\zeta = 0.75$	19290	19801	24168	24512
$\zeta = 1$	19378	19887	23876	24378

operator has limited information of Ω, we build the ambiguity set of virtual net load probability distribution by choosing 50 samples from Ω, and calculating the parameters $\underline{\boldsymbol{\xi}}, \overline{\boldsymbol{\xi}}, \boldsymbol{\mu}, \boldsymbol{\gamma}, \bar{\boldsymbol{\gamma}}$, as explained in Section II-B. After solving the corresponding TDR and RO models, we fix the optimal first-stage decision $\mathbf{y}$, and run the second-stage problem $Q(\mathbf{y}, \boldsymbol{\xi})$ for all 5,000 samples in Ω. We report the average operating cost in Table III. The results verify that TDR model provides less conservative solutions compared to the RO approach. In particular, TDR model can save up to \$532. Moreover, as the value of ε increases, i.e., the size of support set is expanded, the difference between TDR and RO models reduces, since the TDR model becomes more conservative.

D. Comparing With Multi-Stage Model

In this subsection, we benchmark the multi-stage DR model (MDR) with the two-stage DR model (TDR), and report the profit values (i.e., the opposite value of the objective) and CPU time in Table IV. We can observe that the MDR model provides a lower profit, which is because that the MDR model enforces nonanticipativity constraints. That is, as compared with MDR model, the TDR model assumes the virtual net load consumption information throughout the entire time periods is known at the beginning of the operating day, which offers more flexibility in the dispatch decisions.

TABLE IV
MULTI-STAGE DR VS. TWO-STAGE DR

		$\varepsilon = 0.1$		$\varepsilon = 0.2$	
ζ		TDR	MDR	TDR	MDR
$\zeta = 0$	profit(\$)	34370	28421	27470	25200
	time(s)	71	39	82	44
$\zeta = 0.25$	profit(\$)	34364	28485	16583	14625
	time(s)	92	35	101	59
$\zeta = 0.5$	profit(\$)	31189	25435	18550	17431
	time(s)	165	66	144	72
$\zeta = 0.75$	profit(\$)	34548	28382	28688	23723
	time(s)	89	56	80	59
$\zeta = 1$	profit(\$)	32090	26952	19795	11006
	time(s)	109	79	111	85

TABLE V
TWO-STAGE DR VS. TWO-STAGE RO FOR A 6-CONVENTIONAL-GENERATOR CASE

Time series model	TDR	RO
ARIMA$(1,1,1)(1,0,0)_{24}$	30492	32152
ARIMA$(2,1,2)(0,1,1)_{24}$	30844	32297
ARIMA$(2,1,2)(1,1,1)_{24}$	30579	32348

TABLE VI
MULTI-STAGE DR VS. TWO-STAGE DR FOR A 6-CONVENTIONAL-GENERATOR CASE

Time series model		TDR	MDR
ARIMA$(1,1,1)(1,0,0)_{24}$	profit(\$)	49708	42890
	time(s)	265	104
ARIMA$(2,1,2)(0,1,1)_{24}$	profit(\$)	47993	41344
	time(s)	201	110
ARIMA$(2,1,2)(1,1,1)_{24}$	profit(\$)	48770	42351
	time(s)	234	131

E. Computational Results for a Complicated System

In this subsection, we evaluate the performance of our framework for the case that the random samples do not necessarily follow a standard distribution such as the multivariate normal distribution. For this purpose, the stochastic process describing the virtual net load behavior is captured through parsimonious time series models. Following the Box-Jenkins' procedure of the model identification [40] and using Bayesian information criterion, we choose three competing seasonal ARIMA models and simulate samples accordingly. In the case studies, we also assume that the VPP includes six conventional generators to represent a more complicated system. We follow the procedure in the Section IV-C to compare the TDR with RO using samples generated from time series models. As shown in Table V, TDR model leads to less conservative results than those of RO in the out-of-sample simulation. This demonstrates efficiency of our approach even when the underlying uncertainty distribution is not normal. Furthermore, Table VI compares profit values obtained in TDR and MDR approaches for the complicated system. We can also see that TDR model yields to more profits at the expense of having a full knowledge of realized virtual net

load throughout the entire scheduling horizon. This confirms our observation from the previous Section IV-D.

APPENDIX A

Consider the following second-order conic program:

$$\begin{aligned} &\min\ f^T x \\ &\text{s.t.}\ \ ||A_i x + b_i|| \le c_i^T + d_i,\ \ i = 1, \ldots, N, \end{aligned}$$

where $x \in \mathbb{R}^n$ is the optimization variable, and the problem parameters are $f \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{(n_i-1)n}$, $b_i \in \mathbb{R}^{n_i-1}$, $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, and the norm we use in the constraints is the Euclidean norm, i.e., $||y|| = \sqrt{y^T y}$. The dual of this problem is as follows [41]:

$$\begin{aligned} &\max -\sum_{i=1}^{N}(b_i^T z_i + d_i w_i) \\ &\text{s.t.}\ \ \sum_{i=1}^{N}(A_i^T z_i + c_i w_i) = f, \\ &||z_i|| \le w_i,\ \ i = 1, \ldots, N, \end{aligned}$$

where $z_i \in \mathbb{R}^{n_i-1}$ and $w \in \mathbb{R}^N$ are dual optimization variables. By applying this primal-dual relationship, we can get the dual formulation of (12). In order to apply it, it suffices that we reformulate constraints (7) as a set of second-order conic constraints. Consider the following constraint:

$$(\boldsymbol{\xi} - \boldsymbol{\mu})^2 \le \mathbf{u}$$

We can rewrite it as follow:

$$\sqrt{(\boldsymbol{\xi} - \boldsymbol{\mu})^2 + \left(\frac{\mathbf{u}-1}{2}\right)^2} \le \frac{\mathbf{u}+1}{2}$$

or equivalently:

$$\left|\left|\begin{pmatrix} \boldsymbol{\xi} - \boldsymbol{\mu} \\ \frac{\mathbf{u}-1}{2} \end{pmatrix}\right|\right| \le \frac{\mathbf{u}+1}{2},$$

which is a second-order conic constraint. We can apply the same technique for the other quadratic constraints in (7) and get the dual formulation of (12).

APPENDIX B

Proposition 1: Algorithm 1 converges to the optimal solution of problem (14) in a finite number of iterations.

Proof: We rewrite the inner minimization problem in (13) as: $\min\{\mathbf{b}^T\boldsymbol{\psi} : \psi \in \Psi\}$, where $\Psi = \{\psi : \mathbf{E}\boldsymbol{\psi} \succeq_\kappa \mathbf{f}\boldsymbol{\alpha} + \mathbf{g}\boldsymbol{\beta} + \mathbf{p}\boldsymbol{\lambda} + \mathbf{q}\boldsymbol{\pi}^*\}$. Accordingly, F can be written as:

$$F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha}) = \max_{\pi_i, \forall i}\ \min_{\psi \in \Psi} \boldsymbol{\pi}_i^{*T}(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}, \quad (19)$$

where $\{\boldsymbol{\pi}_i^*\}_{i=1}^I$ represents the set of extreme points of Π. Since Π is a polyhedron, $\{\boldsymbol{\pi}_i^*\}_{i=1}^I$ is a finite set, i.e., its cardinality is equal to I. In the kth iteration of Algorithm 1, we have that:

$$\eta = \max_{\pi_i, i \le k}\ \min_{\psi \in \Psi} \boldsymbol{\pi}_i^{*T}(\mathbf{h} - \mathbf{B}\mathbf{y}) + \mathbf{b}^T\boldsymbol{\psi}. \quad (20)$$

Comparing (19) and (20), it is obvious that $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha}) \ge \eta$. Therefore, the stopping criterion for the algorithm is $F(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha}) \le \eta$. We will report $(\mathbf{y}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ as the optimal first-stage decisions when this criterion is achieved. In the worst-case condition, the Algorithm explores all the extreme points of Π, i.e., it reaches at iteration I. Therefore, the stopping criterion will be achieved for at most I iterations, which means that Algorithm 1 converges to the optimal solution in a finite number of iterations. ■

REFERENCES

[1] S. Burger, J. P. Chaves-Ávila, C. Batlle, and I. J. Pérez-Arriaga, "A review of the value of aggregators in electricity systems," *Renewable Sustain. Energy Rev.*, vol. 77, pp. 395–405, 2017.

[2] H. Nosair and F. Bouffard, "Energy-centric flexibility management in power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 5071–5081, Nov. 2016.

[3] D. Pudjianto, C. Ramsay, and G. Strbac, "Virtual power plant and system integration of distributed energy resources," *IET Renewable Power Gener.*, vol. 1, no. 1, pp. 10–16, Mar. 2007.

[4] B. Willems, "Physical and financial virtual power plants," 2005. [Online]. Available: https://ssrn.com/abstract=808944

[5] IPA Consulting, Econnect Ltd & Martin Energy, "Reducing the cost of system intermittency using demand side control measures," 2006, [Online]. Available: https://webarchive.nationalarchives.gov.uk/20100211000651/http://www.ensg.gov.uk/index.php?article=31

[6] G. Parkinson, "Tesla builds case for 250 mw virtual power plant after first trial success," 2018. [Online]. Available: https://reneweconomy.com.au/tesla-builds-case-for-250mw-virtual-power-plant-after-first-trial-success-63950/

[7] D. Pudjianto, G. Strbac, and D. Boyer, "Virtual power plant: Managing synergies and conflicts between transmission system operator and distribution system operator control objectives," *CIRED-Open Access Proc. J.*, vol. 2017, no. 1, pp. 2049–2052, 2017.

[8] M. Oates, A. Melia, and V. Ferrando, "Energy balancing accross cities: Virtual power plant prototype and iURBAN case studies," *Entrepreneurship Sustainability Issues*, vol. 4, no. 3, pp. 351–363, 2017.

[9] H. Pandžić, I. Kuzle, and T. Capuder, "Virtual power plant mid-term dispatch optimization," *Appl. Energy*, vol. 101, pp. 134–141, 2013.

[10] R. M. Lima, A. Q. Novais, and A. J. Conejo, "Weekly self-scheduling, forward contracting, and pool involvement for an electricity producer. An adaptive robust optimization approach," *Eur. J. Oper. Res.*, vol. 240, no. 2, pp. 457–475, 2015.

[11] E. Mashhour and S. M. Moghaddas-Tafreshi, "Bidding strategy of virtual power plant for participating in energy and spinning reserve markets part I: Problem formulation," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 949–956, May 2011.

[12] M. Peik-herfeh, H. Seifi, and M. K. Sheikh-El-Eslami, "Two-stage approach for optimal dispatch of distributed energy resources in distribution networks considering virtual power plant concept," *Int. Trans. Elect. Energy Syst.*, vol. 24, no. 1, pp. 43–63, 2014.

[13] H. Ding, P. Pinson, Z. Hu, J. Wang, and Y. Song, "Optimal offering and operating strategy for a large wind-storage system as a price maker," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4904–4913, Nov. 2017.

[14] E. G. Kardakos, C. K. Simoglou, and A. G. Bakirtzis, "Optimal offering strategy of a virtual power plant: A stochastic bi-level approach," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 794–806, Mar. 2016.

[15] L. Ju, Z. Tan, J. Yuan, Q. Tan, H. Li, and F. Dong, "A bi-level stochastic scheduling optimization model for a virtual power plant connected to a wind–photovoltaic–energy storage system considering the uncertainty and demand response," *Appl. Energy*, vol. 171, pp. 184–199, 2016.

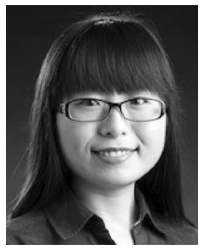
[16] P. Moutis, P. S. Georgilakis, and N. D. Hatziargyriou, "Voltage regulation support along a distribution line by a virtual power plant based on a center of mass load modeling," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3029–3038, Jul. 2018.

[17] D. Koraki and K. Strunz, "Wind and solar power integration in electricity markets and distribution networks through service-centric virtual power plants," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 473–485, Jan. 2018.

[18] S. M. Nosratabadi, R.-A. Hooshmand, and E. Gholipour, "A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems," *Renewable Sustain. Energy Rev.*, vol. 67, pp. 341–363, 2017.

[19] R. M. Lima, A. J. Conejo, S. Langodan, I. Hoteit, and O. M. Knio, "Risk-averse formulations and methods for a virtual power plant," *Comput. Oper. Res.*, vol. 96, no. 2, pp. 350–373, 2018.
[20] S. R. Dabbagh and M. K. Sheikh-El-Eslami, "Risk assessment of virtual power plants offering in energy and reserve markets," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3572–3582, Sep. 2016.
[21] H. Pandžić, J. M. Morales, A. J. Conejo, and I. Kuzle, "Offering model for a virtual power plant based on stochastic programming," *Appl. Energy*, vol. 105, pp. 282–292, 2013.
[22] M. Rahimiyan and L. Baringo, "Strategic bidding for a virtual power plant in the day-ahead and real-time markets: A price-taker robust optimization approach," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2676–2687, Jul. 2016.
[23] M. Shabanzadeh, M.-K. Sheikh-El-Eslami, and M.-R. Haghifam, "The design of a risk-hedging tool for virtual power plants via robust optimization approach," *Appl. Energy*, vol. 155, pp. 766–777, 2015.
[24] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Oper. Res.*, vol. 58, no. 3, pp. 595–612, 2010.
[25] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Oper. Res.*, vol. 62, no. 6, pp. 1358–1376, 2014.
[26] C. Zhao and R. Jiang, "Distributionally robust contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 94–102, Jan. 2018.
[27] 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.
[28] W. Wei, F. Liu, and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 289–300, Jan. 2016.
[29] Y. Zhang, S. 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.
[30] F. Alismail, P. Xiong, and C. Singh, "Optimal wind farm allocation in multi-area power systems using distributionally robust optimization approach," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 536–544, Jan. 2018.
[31] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: SIAM, 2001, vol. 2.
[32] D. Bertsimas, M. Sim, and M. Zhang, "Adaptive distributionally robust optimization," *Manag. Sci.*, May 9, 2018.
[33] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Math. Program.*, vol. 99, no. 2, pp. 351–376, 2004.
[34] A. Lorca, X. A. Sun, E. Litvinov, and T. Zheng, "Multistage adaptive robust optimization for the unit commitment problem," *Oper. Res.*, vol. 64, no. 1, pp. 32–51, 2016.
[35] 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.
[36] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Oper. Res. Lett.*, vol. 41, no. 5, pp. 457–461, 2013.
[37] A. Lorca and X. A. Sun, "The adaptive robust multi-period alternating current optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1993–2003, Mar. 2018.
[38] H. Konno, "A cutting plane algorithm for solving bilinear programs," *Math. Program.*, vol. 11, no. 1, pp. 14–27, 1976.
[39] PJM. [Online]. Available: http://www.pjm.com/markets-and-operations/ops-analysis/historical-load-data.aspx
[40] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "Arima-based time series model of stochastic wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 667–676, May 2010.
[41] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl.*, vol. 284, no. 1–3, pp. 193–228, 1998.

Sadra Babaei (S'17) received the B.S. degree in industrial engineering from the Iran University of Science and Technology, Tehran, Iran, in 2009, and the M.S. degree in industrial engineering from the Amirkabir University of Technology, Tehran, Iran, in 2012. He is currently working toward the Ph.D. degree at the School of Industrial Engineering and Management, Oklahoma State University, Stillwater, OK, USA. His research interests include decision making under uncertainty and data-driven optimization with their applications in power system operations and renewable energy integration.



Chaoyue Zhao (M'12) received the B.S. degree in information and computing sciences from Fudan University, Shanghai, China, in 2010, and the Ph.D. degree in industrial and systems engineering from the University of Florida, Gainsville, FL, USA, in 2014. She is currently an Assistant Professor in industrial engineering and management with the Oklahoma State University, Stillwater, OK, USA. Her research interests include data-driven stochastic optimization and stochastic integer programming with their applications in power grid security and renewable energy management. She was with the Pacific Gas and Electric Company in 2013.

Lei Fan (S'13–M'15) received the B.S. degree in electrical engineering from the Hefei University of Technology, Hefei, China, in 2009, and the Ph.D. degree in industrial and systems engineering from the University of Florida, Gainesville, FL, USA, in 2015. He was an Application Engineer with General Electric from 2015 to 2017. He is currently a Software Engineer with the Siemens Industry, Inc., Minneapolis, MN, USA. His research interests include the optimization of complex system operations, planning, and energy market analysis.