Data-driven Planning for Renewable Distributed Generation in Distribution Networks

Abolhassan Mohammadi Fathabad; Jianqiang Cheng, *Member, IEEE;* Kai Pan, *Member, IEEE;* Feng Qiu, *Senior Member, IEEE.*

Abstract—Many countries are setting ambitious goals to integrate large amounts of sustainable, low-emission renewable energy into their electricity supply. These to-be-installed renewable sites, mainly wind farms or solar panels, are highly uncertain sources of electricity. Thus meeting these ambitious goals will require careful planning and operational strategies. In this paper, we first propose a two-stage data-driven distributionally robust optimization model (O-DDSP) for the optimal placement of renewable distributed generation (RDG). In our model, we consider both load and generation uncertainties through data-driven ambiguity set, which enables more flexibility than stochastic optimization and allows less conservative solutions than robust optimization. The objective is to minimize the total cost of RDG installation plus the total operational cost on the planning horizon. Next, we introduce a tight approximation of O-DDSP based on principal component analysis (leading to a model denoted by P-DDSP), which reduces the original problem size by keeping the most valuable data in the ambiguity set. The performance of O-DDSP and P-DDSP is compared with stochastic optimization (SO) and robust optimization (RO) on the IEEE 33-Bus radial network with real data set, where we show that P-DDSP significantly speeds up the solution procedure, especially when the number of periods increases. Indeed, as compared to SO and RO, which become computationally impractical solving the siting and sizing problem with large sample sizes, our P-DDSP formulation can increase the accuracy of its solutions by utilizing larger sample sizes. Finally, extensive numerical experiments demonstrate that the optimal RDG planning decisions lead to significant savings as well as increasing penetration of intermittent renewables in the distribution network.

Index Terms—Distributionally robust optimization, renewable distributed generation, principal component analysis, semidefinite programming, delayed constraint generation algorithm.

I. NOMENCLATURE

A. Indices and Sets

$n/k/t$	Index of buses / RDGs / periods
$\mathcal{N}/\mathcal{E}$	Set of buses / power distribution lines.
$\mathcal{N}_0$	Set of buses connected to bus 0.
$\mathcal{N}_1$	Set of buses with dispatchable DG units.
$\mathcal{N}_2$	Set of buses with reactive power sources.
$\mathcal{X}$	Feasibility set of the first-stage decision x.
$\mathcal{Y}^t$	Set of all the power flow variables at period t.
$\mathcal{D}^t/\mathcal{D}_r^t$	Original/projected ambiguity set of ξ^t at period t.
$\mathcal{F}^t$	Feasibility region of the dual of the second-stage problem at period t.
$\mathcal{V}_{\mathcal{F}^t}$	Set of selected vertices of $\mathcal{F}^t$ used to solve O-DDSP and P-DDSP.

Abolhassan Mohammadi Fathabad and Jianqiang Cheng are with the Department of Systems and Industrial Engineering, University of Arizona, Tucosn, AZ. E-mail: jqcheng@email.arizona.edu

Kai Pan is with the Department of Logistics and Maritime Studies, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: kai.pan@polyu.edu.hk.

Feng Qiu is with the Energy Systems Division, Argonne National Laboratory, Lemont, IL 60439, USA. E-mail: fqiu@anl.gov.

$\mathcal{G}_K$	Set of location alternatives for installing K RDGs.

B. Parameters

$\hat{A}/\hat{B}/\hat{e}$ $\hat{C}_{(x)}/\hat{d}$	Auxiliary matrices and vectors used to represent the second-stage problem in compact form.
c_{kn}^0	Setup cost of placing the kth RDG at bus n.
c_k^1/c_k^2	Size-based investment / maintenance cost of the kth RDG.
c_p^t/c_q^t	Electricity price of purchasing active / reactive power from the main grid.
c_n^f/c_n^e	Fuel / emission price for the dispatchable DGs at bus n.
ω	Emission factor of the dispatchable DGs (kg/kWh).
K	The total number of RDGs to be installed.
T	The total number of time intervals in the planning horizon.
$\bar{c}_{mn}$	Capacity of the power transmission line (m,n).
$(\bar{p}_n^t, \underline{p}_n^t)/$ $(\bar{q}_n^t, \underline{q}_n^t)$	Active / Reactive power output bounds of dispatchable DGs at period t.
$\mathfrak{R}_{mn}/\mathfrak{X}_{mn}$	Electrical resistance / reactance of line (m,n).
δ_n	Binary indicator if bus n has a dispatchable unit.
τ_{kn}	Binary indicator if kth reactive power source is installed at bus n.
$\boldsymbol{\mu}^t/\boldsymbol{\Sigma}^t/$ $\mathcal{S}^t$	Mean vector / covariance matrix / support set of random variable $\boldsymbol{\xi}^t$.
A, b	Coefficients used to represent $\mathcal{S}^t$.
U^t/Λ^t	Matrix of eigenvectors / eigenvalues of $\boldsymbol{\Sigma}^t$.
$U_{m\times m_1}^t/$ $\Lambda_{m_1}^t$	The dominant m_1 elements in matrix of eigenvectors / eigenvalues of $\boldsymbol{\Sigma}^t$.
$\bar{x}_k/\underline{x}_k$	The max / min capacity of the kth RDG.
z_{kn}	Binary indicator if kth RDG is located at bus n.
$\mathfrak{u}$	Number of pieces in the polyhedral ϵ-approximation of the SOCP constraint.
$\bar{v}/\underline{v}$	The upper / lower bound of voltage magnitude.

C. Random Variables

d_{pn}^t/d_{qn}^t	Active / reactive load at bus n at period t.
s_k^t	Active power output of the kth RDG at period t.
$\boldsymbol{\xi}^t$	Vector of uncertainty in compact form at period t, i.e., $[s_1^t, ..s_k^t, d_{p1}^t, ...d_{pN}^t, d_{q1}^t...., d_{qN}^t]^T$.
$\boldsymbol{\xi}_r^t$	The PCA projection of the $\boldsymbol{\xi}^t$ into m_1 dimensions.

D. Decision Variables

x	Size (i.e., capacity) of RDG units, i.e., $[x_1, ..x_K]^T$.
p_0^t/q_0^t	Active / reactive power purchased from the main grid at bus 0 at period t.
$P_{mn}^t/$ Q_{mn}^t	Active / reactive power flow from buses m to n at period t.
$V_n^t/$ $|V_n^t|^2$	Complex voltage at bus n at period t / its magnitude.
$I_{mn}^t/$ $|I_{mn}^t|^2$	Complex current from buses m to n at period t / its magnitude.

p_n^t / q_n^t	Active / reactive power output of the dispatchable DG units (reactive DG units) at bus n at period t.
$\boldsymbol{W}^t$	Auxiliary variable denoting the second-stage OPF variables at period t in compact form.
$\boldsymbol{\lambda}^t$	Dual variable of the second-stage constraints at period t.
$\boldsymbol{\beta}^t$	Dual variable corresponding to the polyhedral support constraint at period t.
s^t / s_r^t	Scalar dual variable used in reformulation process for O-DDSP / P-DDSP.
$\boldsymbol{q}^t / \boldsymbol{q}_r^t$	Vector dual variable used in reformulation process for O-DDSP / P-DDSP.
$\boldsymbol{Q}^t / \boldsymbol{Q}_r^t$	Matrix dual variable used in reformulation process for O-DDSP / P-DDSP.

E. Auxilary variables for ϵ-approx. of SOCP cons.

$\mathfrak{Y}_{1,mn}^t$ /$\mathfrak{Y}_{2,mn}^t$/ $\mathfrak{Y}_{3,mn}^t$/ $\varepsilon_1{}^0,..,\varepsilon_3{}^u$/ $\eta_1{}^0,..,\eta_3{}^u$

F. Functions

$Q(\boldsymbol{x}, \boldsymbol{\xi}^t)$	Optimal value of the second-stage problem at period t when $\boldsymbol{x}$ and $\boldsymbol{\xi}^t$ are given.
$\mathbb{E}_{F_{\boldsymbol{\xi}^t}}$	Expectation over distribution $F_{\boldsymbol{\xi}^t}$.
$\mathbf{P}$	Probability of an event.
"$\bullet$"	The trace of two conformal matrices.
$vert(.)$	The set of vertices of a polytope.

II. INTRODUCTION

Renewable energy is becoming an increasingly prevalent source of electricity throughout the world. Recent technological advances along with supportive government policies have provided the means for significant investments in exploiting wind and solar energy. For instance, the total installed capacity for renewable energy was higher than that of traditional fossil fuels in 2017 [1]. This implies a tremendous integration of renewable distributed generation (RDG) at the distribution level of the power grid. However, the electrical outputs from RDGs are extremely variable and uncertain over different time windows, from minutes to hours to days –because of weather patterns, diurnal nature of the sun, changes in sun angles, etc. This variability makes it extremely harder to maintain the balance between electricity supply and demand at all times, and it can lead to blackouts or other cascading problems. Therefore, there is a limit to safe development of RDGs in an existing distribution system. In brief, adding RDGs complicates the configuration of circuit breakers and other protection systems and leads to more sophisticated distribution networks. Moreover, it further raises new challenges including managing the distribution voltage and managing the contingency plan for power generators and distribution lines, all of which could lead to higher costs and faults on the electricity network [2]. To overcome these challenges, improved analysis of distribution system such a careful siting and sizing of RDGs within a distribution feeder is proven to be an accessible and low-cost strategy that can increase the penetration of renewable energy while guaranteeing the smooth operation of an existing distribution system [3].

Indeed, if RDG units (RDGs) are installed efficiently (with optimal sizing and siting decisions), then they can not only increase the penetration of clean and cheap renewable energy, but also boost the overall efficiency and reliability of a microgrid as well as the whole power grid by providing reactive power support, minimizing the network power loss, and improving the voltage stability and energy security [4, 5]. On the contrary, an improper installation of the RDGs can cause various malfunctions in the system such as unbalanced supply and demand, power quality decay, voltage increase at the end of the feeder, more power losses, and low system reliability [6, 7]. Therefore, the planning of the RDGs is of outstanding importance to ensure the aforementioned benefits as well as to maximize the expansion of renewable energy penetration while respecting the installation and operation costs of the RDGs. Consequently, the problem of siting and sizing of new RDGs has received much attention from both academia and industry practice [8, 9]. For instance, [10] suggests that the siting and sizing of RDGs have a large affect on the system losses and proposes an ant lion optimization algorithm that minimizes power losses. An environmentally committed short-term planning of electrical distribution systems considering RDGs siting and sizing is proposed in [11]. [12] proposes a mixed-integer linear programming technique to find the optimal short-term siting and sizing plan of RDGs and reactive power sources.

The RDG sizing and siting problems are difficult to solve in general because 1) the distribution systems are extremely complex as the mathematical model includes nonlinear expressions, complex numbers, and large number of variables and constraints, and 2) there is a significant amount of uncertainty incorporated in the model. Therefore, implementation of meta-heuristic methods (e.g., genetic algorithms, particle swarm optimization, tabu search) [13, 14] and analytical techniques (e.g., eigen-value based analysis, index method) [6] are usually preferred to the exact mathematical programming methods. While the meta-heuristic techniques are efficient and provide a fast search of the solution space, they may not be able to obtain global optimality and there is no guarantee about the quality of the solution. The analytical approaches, on the other hand, are valuable as they ensure the convergence of RDG planning solution. However, the assumptions used for over-simplifying the problem come at the expense of the accuracy of the solution.

Among few papers that use mathematical programming methods to solve the sizing and siting of RDGs, the uncertainty has been mostly modeled by stochastic optimization (SO) [15, 16, 17] and robust optimization (RO) [18, 19, 20]. Although both of these techniques have been studied to support decision making under uncertainty, their shortcomings for the RDG planning problem are also straightforward. In particular, SO requires exact probability distribution of random parameters, which is usually impractical to accurately estimate in real practice. Thus, Sample Average Approximation (SAA) is often used to solve SO because its advantage in guaranteeing convergence to the optimal solution with large enough scenarios. Nevertheless, the SAA is proven to be extremely computationally expensive when the sample size is large. Meanwhile, RO is too conservative because it finds solutions by considering the worst case of uncertain parameters, which rarely happens in practice. Distributionally robust optimization (DRO), on the other hand, considers both distributional information and

the range of uncertain parameters, and can help address the limitations of RO and SO [21, 22]. Other advantages of applying the DRO method to the sizing and siting of RDGs are 1) we can use historical data to estimate certain information (e.g., support, mean, variance, etc.) of random parameters, which can be further used to support decision making, leading to a data-driven approach. 2) Increasing the sample size in DRO problems will further improve the accuracy of its ambiguity set without adversely affecting the size of the problem, whereas for the classic optimization methods, i.e., SO and RO, the computation time grows exponentially with larger sample sizes. This is especially important as the sizing and siting decisions are medium-term planning decisions that are made by considering the big dataset of distribution system's load and generation data over a long planning horizon (6-48 months). Moreover, 3) these data usually do not follow a single distribution as stochastic optimization assumes. In contrast, DRO provides a lot of flexibility as it allows to incorporate information about the estimation errors into the optimization problem and its solutions are guaranteed for all the distributions in the ambiguity set [23].

Therefore, DRO has been recently used to determine the optimal size of wind farms in power system [24], however, this work only considers a low dimension uncertainty vector (with five random variables), and it solves the DRO model using linear decision rule which fails for large-scale problems (i.e., with a lot of uncertain parameters). A DRO based data-driven RDG sizing model is proposed in [25] which considers both uncertain RDG outputs and uncertain load demands, however, it uses the CVaR (conditional value at rist) objective function to maximize the penetration of RDGs for active distribution networks. Note that a DRO problem with first- and second-moment information can be reformulated exactly as a semidefinite programming problem (SDP), and if the size of the uncertainty vector increases, size of SDP constraints will increase as a result and the corresponding problem becomes quickly intractable dealing with large SDP constraints. Indeed, most of the DRO papers in the power system literature study only the case when the size of the uncertainty vectors is low. For instance, [26] considers a maximum of two wind outputs (with 2 random variables) and [27] considers a maximum of four wind outputs (with 4 random variables) in their numerical tests. Also, [25] considers CVaR objective function which limits the number of SDP constraints to very few. In general, while showing the benefits of using the DRO vs. RO or SO, none of DRO papers in literature introduce any means to deal with large-scale uncertainty which is usually the case in power system problems. Moreover, to the best of our knowledge, few DRO studies have focused on planning strategies regarding the placement and capacity of intermittent RDGs in the power distribution network, and even less have studied the efficient computational approaches to solve practical large-scale planning problems.

Therefore, in this paper, we aim to determine the sizing and siting plan for the RDGs in an active distribution network by using historical generation and load data and constructing the corresponding DRO model. The main contributions of this paper can be summarized as follows:

- A novel two-stage data-driven distributionally robust optimization model (O-DDSP) to make the planning decision for the RDGs in an active distribution network considering multiple periods is proposed. In our model, the probability distribution function (PDF) of uncertain renewable generation, as well as uncertain demand of each bus, belongs to a set of multivariate distributions with known support, first- and second-moment information. The information is estimated by using historical data from real industry. The historical load data is obtained from the Pecanstreet project [28], which provides access to a large set of real electricity usage data for academic use. The renewable generation data is from ERCOT [29]. The objective function is to minimize the expected total cost, including strategic installation cost in the first-stage and operational cost against the worst-case distribution in the second stage.

- We present an exact semidefinite programming (SDP) reformulation of the proposed multi-period O-DDSP model, and it is solved by a delayed constraint generation algorithm. The size of the O-DDSP problem, however, can be very large-scale in practice so that it cannot be handled by the solvers. To overcome the significant computational difficulty of the SDP reformulation, we propose a tight approximation of the O-DDSP based on principal component analysis (PCA), leading to a model denoted by P-DDSP, by keeping the most dominant information in the ambiguity set. This P-DDSP formulation has significantly lower CPU times, and it allows for investigating practical sizing and siting optimization problems that have large-sized uncertainty vectors. The efficiency of P-DDSP, as compared to O-DDSP, SO method and RO method, has been demonstrated with extensive numerical results.

- A general framework to decentralize planning strategies is proposed. By performing numerical tests and sensitivity analyses on the detailed second-stage cost components, we confirm the efficiency of our data-driven results such that the RDG sizing and siting decisions can have a conclusive impact on the total cost. The joint RDG siting and sizing problem requires several instances of the RDG sizing problem to be run quickly, and our P-DDSP allows several instances to be run within several minutes. Note that, it would take several days instead of minutes to run the several instances of the sizing problem with RO or SO. Therefore, our proposed method makes it possible to practically expand the RDGs towards the minimum cost and high penetration of renewable energy as compared to RO which is too conservative, and SO which is too computationally expensive for large sample sizes.

The remaining parts of this paper are organized as follows. The mathematical modeling of the two-stage multi-period planning problem and its DRO reformulation is presented in Section II. The SDP reformulation of the DRO problem, a PCA approximation of the original DRO problem, and also the constraint generation algorithm are presented in Section IV. Numerical results are provided in Section V. Finally, we conclude this paper in Section VI.

III. MATHEMATICAL MODEL

In this section, we first present a two-stage planning model for RDG placement and then introduce its DRO counterpart.

A. Two-stage Model Formulation

We consider a common topology of distribution networks, i.e., a radial network. The load demand in the distribution network is primarily satisfied by the power internally generated from the DG units. When the internal power supply is not enough, the system will buy power from the main grid via the Point of Common Coupling (PCC), thus incurring operating costs. This paper investigates the optimal siting and sizing of RDGs at each bus in $\mathcal{N}$ to minimize the total cost during the whole time horizon, including investment, maintenance, and operational costs. In the following, we formulate the planning model as a two-stage optimization model and describe the corresponding first-stage and second-stage objectives and constraints.

1) **First-stage planning model**: The objective is to minimize the total cost, including a set-up cost and size-based investment and maintenance costs for the RDGs as follows:
$$\mathcal{C}_1(\boldsymbol{x}) = \sum_{k=1}^{K} \left(\sum_{n \in \mathcal{N}} c_{kn}^0 z_{kn} + (c_k^1) x_k + (c_k^2 \times T) x_k \right).$$
The support of first-stage decision variable $\boldsymbol{x}$ is defined as
$$\mathcal{X} := \left\{ \boldsymbol{x} \in \mathbb{R}^K : \underline{x}_k \leq x_k \leq \bar{x}_k, \ \forall k = 1, 2, \cdots, K \right\}.$$

2) **Second-stage operational model**: The objective is to minimize the costs of the power generated by the dispatchable DG units as well as the power purchased from the main grid, while respecting physical constraints such as optimal power flow (OPF) constraints. Note that the OPF problem is non-convex and thus is hard to solve. Recently, [30] shows that when the network is radial, the OPF constraints can be exactly reformulated to a set of Second Order Cone Programming (SOCP) constraints in a branch flow model. Hence, we adopt the branch flow model to construct the OPF constraints. Therefore, given the first-stage decision $\boldsymbol{x}$ and the actual realization of uncertainty $\boldsymbol{\xi}^t$ at period t, the minimum operating cost (denoted by $Q(\boldsymbol{x}, \boldsymbol{\xi}^t)$) is obtained by solving the following problem:

$$\min_{\mathcal{Y}, p} c_p^t p_0^t + c_q^t q_0^t + \sum_{n \in \mathcal{N}_1} c_n^f p_n^t + \sum_{n \in \mathcal{N}_1} c_n^e \omega p_n^t \tag{1a}$$

$$s.t. \quad \underline{p}_n^t \leq p_n^t \leq \bar{p}_n^t, \qquad \forall n \in \mathcal{N}_1 \tag{1b}$$

$$\underline{q}_n^t \leq q_n^t \leq \bar{q}_n^t, \qquad \forall n \in \mathcal{N}_2 \tag{1c}$$

$$\underline{v} \leq |V_n^t|^2 \leq \bar{v}, \qquad \forall n \in \mathcal{N} \setminus \{0\} \tag{1d}$$

$$p_0^t = \sum_{n \in \mathcal{N}_0} P_{0n}^t, \quad q_0^t = \sum_{n \in \mathcal{N}_0} Q_{0n}^t, \tag{1e}$$

$$P_{mn}^t - \mathfrak{R}_{mn}|I_{mn}^t|^2 = d_{pn}^t - \sum_{k=1}^{K} z_{kn} s_k^t x_k - \delta_n p_n^t + \sum_{l \in \mathcal{N}_n} P_{nl}^t, \quad \forall (m,n) \in \mathcal{E} \tag{1f}$$

$$Q_{mn}^t - \mathfrak{X}_{mn}|I_{mn}^t|^2 = d_{qn}^t - \sum_{k \in \mathcal{N}_2} \tau_{kn} q_{kn}^t + \sum_{l \in \mathcal{N}_n} Q_{nl}^t, \quad \forall (m,n) \in \mathcal{E} \tag{1g}$$

$$|V_m^t|^2 - |V_n^t|^2 = 2\mathfrak{R}_{mn}P_{mn}^t + 2\mathfrak{X}_{mn}Q_{mn}^t - (\mathfrak{R}_{mn}^2 + \mathfrak{X}_{mn}^2)|I_{mn}^t|^2, \quad \forall (m,n) \in \mathcal{E} \tag{1h}$$

$$\|[2P_{mn}^t, 2Q_{mn}^t, |V_m^t|^2 - |I_{mn}^t|^2]\|_2 \leq |V_m^t|^2 + |I_{mn}^t|^2, \\ \forall (m,n) \in \mathcal{E} \tag{1i}$$

$$(P_{mn}^t)^2 + (Q_{mn}^t)^2 \leq (\bar{c}_{mn})^2, \qquad \forall (m,n) \in \mathcal{E}, \tag{1j}$$

where $|V_m^t|^2$ represents the voltage magnitude of bus m and $|I_{mn}^t|^2$ represents the current magnitude on line (m,n). Constraints (1b) and (1c) restrict the active and reactive generations of each unit to their bounds, respectively. (1d) sets the bounds on voltage of each bus. Constraints (1e) represent the active and reactive power balance equations at bus 0, respectively. Constraints (1f) and (1g) are active and reactive power balance equations from the Kirchhoff's current law, respectively. (1h) represents the voltage drop on each line. (1i) is the branch power flow constraint and the capacity of each transmission line is limited by (1j).

Note that constraints (1i) and (1j) are nonlinear SOCP constraints. Moreover, in order to obtain a tractable formulation that can be used practically in large-scale settings, we approximate these constraints with linear constraints. First, we build a polyhedral ϵ-approximation for (1i) according to [31] with the following constraints:

$$\begin{cases} \varepsilon_1{}^0 \geq \left|2P_{mn}^t\right|, & \eta_1{}^0 \geq \left||V_m^t|^2 - |I_{mn}^t|^2\right|, \\ \varepsilon_2{}^0 \geq \left|\mathfrak{Y}_{3,mn}^t\right|, & \eta_2{}^0 \geq \left|2Q_{mn}^t\right|, \\ \varepsilon_3{}^0 \geq \left|\mathfrak{Y}_{1,mn}^t\right|, & \eta_3{}^0 \geq \left|\mathfrak{Y}_{2,mn}^t\right|, \end{cases} \tag{2a}$$

$$\begin{cases} \varepsilon_1{}^j = \cos(\frac{\pi}{2^{j+1}})\varepsilon_1{}^{j-1} + \sin(\frac{\pi}{2^{j+1}})\eta_1{}^{j-1}, \\ \eta_1{}^j \geq \left|-\sin(\frac{\pi}{2^{j+1}})\varepsilon_1{}^{j-1} + \cos(\frac{\pi}{2^{j+1}})\eta_1{}^{j-1}\right|, \\ \varepsilon_2{}^j = \cos(\frac{\pi}{2^{j+1}})\varepsilon_2{}^{j-1} + \sin(\frac{\pi}{2^{j+1}})\eta_2{}^{j-1}, \\ \eta_2{}^j \geq \left|-\sin(\frac{\pi}{2^{j+1}})\varepsilon_2{}^{j-1} + \cos(\frac{\pi}{2^{j+1}})\eta_2{}^{j-1}\right|, \\ \varepsilon_3{}^j = \cos(\frac{\pi}{2^{j+1}})\varepsilon_3{}^{j-1} + \sin(\frac{\pi}{2^{j+1}})\eta_3{}^{j-1}, \\ \eta_3{}^j \geq \left|-\sin(\frac{\pi}{2^{j+1}})\varepsilon_3{}^{j-1} + \cos(\frac{\pi}{2^{j+1}})\eta_3{}^{j-1}\right|, \end{cases} j = 1, 2, ..., \mathfrak{u}, \tag{2b}$$

$$\begin{cases} \varepsilon_1{}^{\mathfrak{u}} \leq \mathfrak{Y}_{1,mn}^t, & \eta_1{}^{\mathfrak{u}} \leq \tan(\frac{\pi}{2^{v+1}})\varepsilon_1{}^{\mathfrak{u}}, \\ \varepsilon_2{}^{\mathfrak{u}} \leq \mathfrak{Y}_{2,mn}^t, & \eta_2{}^{\mathfrak{u}} \leq \tan(\frac{\pi}{2^{v+1}})\varepsilon_2{}^{\mathfrak{u}}, \\ \varepsilon_3{}^{\mathfrak{u}} \leq |V_m^t|^2 + |I_{mn}^t|^2, & \eta_3{}^{\mathfrak{u}} \leq \tan(\frac{\pi}{2^{v+1}})\varepsilon_3{}^{\mathfrak{u}}, \end{cases} \tag{2c}$$

Note that the system of constraints (3) are a polyhedral ϵ approximation of constraint (1i). Thus, the error ϵ of this approximation depends on the parameter of the construction $\mathfrak{u}$, and it is calculated by $\frac{1}{\cos(\pi/2^{\mathfrak{u}+1})} - 1$. For example, for $\mathfrak{u}$ equal to 2, 5, and 10, the ϵ is 0.41, 10^{-3} and 10^{-6}, respectively. Next, note that the equation (1j) describes a circle, and it can be approximated by several squares as depicted in Figure 1. Generally, the more squares considered, the higher the accuracy. Here, we used 2 squares as it already provides a tight approximation appropriate for engineering applications:

$$- \bar{c}_{mn} \leq P_{mn}^t \leq \bar{c}_{mn}, \tag{3a}$$

$$- \bar{c}_{mn} \leq Q_{mn}^t \leq \bar{c}_{mn}, \tag{3b}$$

$$- \sqrt{2}\bar{c}_{mn} \leq P_{mn}^t + Q_{mn}^t \leq \sqrt{2}\bar{c}_{mn}, \tag{3c}$$

$$- \sqrt{2}\bar{c}_{mn} \leq P_{mn}^t - Q_{mn}^t \leq \sqrt{2}\bar{c}_{mn}, \tag{3d}$$

For notational brevity, we use a compact matrix form to represent the linearized model (1), as shown in the following:

$$Q(\boldsymbol{x}, \boldsymbol{\xi}^t) = \min_{\boldsymbol{W}^t} \hat{\boldsymbol{e}}^T \boldsymbol{W}^t \tag{4a}$$

$$s.t. \quad \hat{\boldsymbol{A}}\boldsymbol{W}^t + \hat{\boldsymbol{B}}\boldsymbol{x} + \hat{\boldsymbol{C}}_{(\boldsymbol{x})}\boldsymbol{\xi}^t \leq \hat{\boldsymbol{d}} \tag{4b}$$

The dual of model (4) can be represented as follows:
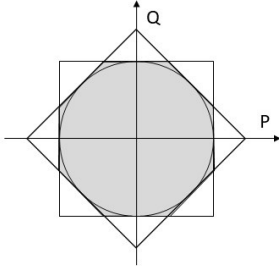
Fig. 1. Linearization of circular constraints: the intersection of the two squares, i.e., grey area, is used to approximate the circle.

$$Q(\boldsymbol{x}, \boldsymbol{\xi}^t) = \max_{\boldsymbol{\lambda}^t} \quad \boldsymbol{\lambda}^{t^T}(\hat{\boldsymbol{B}}\boldsymbol{x} + \hat{\boldsymbol{C}}_{(\boldsymbol{x})}\boldsymbol{\xi}^t - \hat{\boldsymbol{d}}) \tag{5a}$$

$$\text{s.t.} \quad \mathcal{F}^t = \left\{ \boldsymbol{\lambda}^t \,\middle|\, \hat{\boldsymbol{A}}^T\boldsymbol{\lambda}^t + \hat{\boldsymbol{e}} = 0, \boldsymbol{\lambda}^t \geq \boldsymbol{0} \right\}. \tag{5b}$$

where $\boldsymbol{\lambda}^t$ are the dual variables corresponding to constraints (4b).

B. O-DDSP Model

In reality, when we make the RDG planning decisions, we do not know the real value of the uncertainty vector $\boldsymbol{\xi}^t$ in advance. More specifically, the multivariate PDF of ξ^t is unknown in many cases in practice, which prevents us from performing a stochastic programming approach. Accordingly by adopting the DRO method, we assume that the PDF of $\boldsymbol{\xi}^t$ belongs to a moment-based ambiguity set $\mathcal{D}^t(\mathcal{S}^t, \boldsymbol{\mu}^t, \boldsymbol{\Sigma}^t)$. It consists of all the possible probability distributions that share the mean vector $\boldsymbol{\mu}^t$ and covariance matrix $\boldsymbol{\Sigma}^t$ and are on the support set $\mathcal{S}^t$ as follows:

$$\mathcal{D}^t(\mathcal{S}^t, \boldsymbol{\mu}^t, \boldsymbol{\Sigma}^t) = \left\{ F \,\middle|\, \begin{array}{l} \mathbf{P}(\boldsymbol{\xi}^t \in \mathcal{S}^t) = 1 \\ \mathbb{E}_F[\boldsymbol{\xi}^t] = \boldsymbol{\mu}^t \\ \mathbb{E}_F[(\boldsymbol{\xi}^t - \boldsymbol{\mu}^t)(\boldsymbol{\xi}^t - \boldsymbol{\mu}^t)^T] \preceq \boldsymbol{\Sigma}^t \end{array} \right\}, \tag{6}$$

where these parameters $(\mathcal{S}^t, \boldsymbol{\mu}^t, \boldsymbol{\Sigma}^t)$ are usually calculated empirically from the historical data.

The DRO approach looks for the solutions that perform the best under the worst-case distribution within the ambiguity set. Since uncertain vector $\boldsymbol{\xi}^t$ follows an independent PDF with specific ambiguity set $\mathcal{D}^t$ for each period t, we can present the multi-period second-stage problem as follows:

$$\sup_{F_{\boldsymbol{\xi}^t} \in \mathcal{D}^t} \mathbb{E}_{F_{\boldsymbol{\xi}^t}}\left[\sum_{t=1}^{T} Q(\boldsymbol{x}, \boldsymbol{\xi}^t)\right] = \sum_{t=1}^{T} \sup_{F_{\boldsymbol{\xi}^t} \in \mathcal{D}^t} \mathbb{E}_{F_{\boldsymbol{\xi}^t}}[Q(\boldsymbol{x}, \boldsymbol{\xi}^t)]. \tag{7}$$

Combined with the first-stage problem, the complete DRO model, denoted by O-DDSP, can be described as the following two-stage optimization problem:

$$\min_{\boldsymbol{x} \in \mathcal{X}} \left[\mathcal{C}_1(\boldsymbol{x}) + \sum_{t=1}^{T} \sup_{F_{\boldsymbol{\xi}^t} \in \mathcal{D}^t} \mathbb{E}_{F_{\boldsymbol{\xi}^t}}[Q(\boldsymbol{x}, \boldsymbol{\xi}^t)] \right]. \tag{8}$$

By replacing the inner maximization problem in (8) under ambiguity set $\mathcal{D}^t$ with its conic dual, we obtain the following equivalent reformulation:

$$\min_{\substack{\boldsymbol{x}, s^t, \\ \boldsymbol{q}^t, \boldsymbol{Q}^t}} \quad \mathcal{C}_1(\boldsymbol{x}) + \sum_{t=1}^{T} \left[s^t + (\boldsymbol{\mu}^t)^\top \boldsymbol{q}^t + (\boldsymbol{\Sigma}^t + \boldsymbol{\mu}^t(\boldsymbol{\mu}^t)^\top) \bullet \boldsymbol{Q}^t \right] \tag{9a}$$

$$\text{s.t.} \quad s^t + \boldsymbol{\xi}^{t^\top} \boldsymbol{q}^t + \boldsymbol{\xi}^{t^\top} \boldsymbol{Q}^t \boldsymbol{\xi}^t \geq Q(\boldsymbol{x}, \boldsymbol{\xi}^t), \forall \boldsymbol{\xi}^t \in \mathcal{S}^t, t = 1, ..T, \tag{9b}$$

$$\boldsymbol{x} \in \mathcal{X}, \boldsymbol{Q}^t \succeq 0, \tag{9c}$$

where $s^t \in \mathbb{R}$, $\boldsymbol{q}^t \in \mathbb{R}^m$, $\boldsymbol{Q}^t \in \mathbb{R}^{m \times m}$. Note that strong duality holds because $Q(\boldsymbol{x}, \boldsymbol{\xi}^t)$ is F-integrable for any $F \in \mathcal{D}^t$.

Note that, when the ambiguity set is singleton and the distribution of uncertainty is assumed to be known beforehand, then the DRO model becomes the classic stochastic optimization. Alternatively, when the ambiguity set only includes support information, then the DRO model becomes robust optimization.

IV. SOLUTION METHOD

In this section, we first obtain the equivalent SDP reformulation of O-DDSP (9). Next, we introduce P-DDSP, which is a relaxation of O-DDSP (8) based on PCA. Finally, we present the algorithm that we will use to solve both O-DDSP and P-DDSP.

A. O-DDSP in SDP form

Note that we cannot solve O-DDSP (9) directly because (9b) includes an infinite number of constraints and also the function $Q(\boldsymbol{x}, \boldsymbol{\xi}^t)$ is not given in a closed form. To resolve the latter issue, we replace $Q(\boldsymbol{x}, \boldsymbol{\xi}^t)$ with the dual of the second-stage problem (5). Thus, constraints (9b) can be equivalently reformulated with the following constraints:

$$s^t + \boldsymbol{\xi}^{t^\top} \boldsymbol{q}^t + \boldsymbol{\xi}^{t^\top} \boldsymbol{Q}^t \boldsymbol{\xi}^t \geq \boldsymbol{\xi}^{t^\top} \boldsymbol{\lambda}^{t^\top}(\hat{\boldsymbol{B}}\boldsymbol{x} + \hat{\boldsymbol{C}}_{(\boldsymbol{x})}\boldsymbol{\xi}^t - \hat{\boldsymbol{d}}),$$
$$\forall \boldsymbol{\lambda}^t \in \mathcal{F}^t, \quad \boldsymbol{\xi}^t \in \mathcal{S}^t, \ t = 1, .., T. \tag{10}$$

Here we assume the support set $\mathcal{S}^t$ are in the form of a polytope with at least one interior point, i.e. $\mathcal{S}^t = \{\boldsymbol{\xi}^t | \boldsymbol{A}\boldsymbol{\xi}^t \leq \boldsymbol{b}\} \neq \emptyset, \ \forall t = 1, 2, ..T$. Thus model (8) can be reformulated as the following SDP problem:

$$\min_{\substack{\boldsymbol{x}, s^t, \boldsymbol{\beta}^t, \\ \boldsymbol{q}^t, \boldsymbol{Q}^t}} \quad \mathcal{C}_1(\boldsymbol{x}) + \sum_{t=1}^{T} \left[s^t + (\boldsymbol{\mu}^t)^\top \boldsymbol{q}^t + (\boldsymbol{\Sigma}^t + \boldsymbol{\mu}^t(\boldsymbol{\mu}^t)^\top) \bullet \boldsymbol{Q}^t \right], \tag{11}$$

$$\text{s.t.} \quad \boldsymbol{x} \in \mathcal{X},$$
$$\boldsymbol{\beta}^t \geq \boldsymbol{0}, \boldsymbol{Q}^t \succeq 0,$$
$$\begin{bmatrix} M_{11}^t & M_{21}^{t\ \top} \\ M_{21}^t & M_{22}^t \end{bmatrix} \succeq 0, \quad \forall \boldsymbol{\lambda}^t \in vert(\mathcal{F}^t), \ t = 1, .., T,$$
$$M_{11}^t = s^t + \boldsymbol{\xi}^{t^\top} \boldsymbol{\lambda}^{t^\top}(\hat{\boldsymbol{d}} - \hat{\boldsymbol{B}}\boldsymbol{x}) - \boldsymbol{\beta}^{t^\top}\boldsymbol{b},$$
$$M_{21}^t = \left[\frac{1}{2}(\boldsymbol{q}^t - \hat{\boldsymbol{C}}_{(\boldsymbol{x})}^\top \boldsymbol{\lambda}^t + \boldsymbol{A}^\top \boldsymbol{\beta}^t) \right], \quad M_{22}^t = [\boldsymbol{Q}^t],$$

where set $vert(\mathcal{F}^t)$ includes all the vertices of the feasibility region $\{\boldsymbol{\lambda}^t | \boldsymbol{\lambda}^t \geq \boldsymbol{0}, \hat{\boldsymbol{A}}^\top \boldsymbol{\lambda}^t + \hat{\boldsymbol{e}} = 0\}$.

B. Low-rank approximation for O-DDSP

It is important to notice that O-DDSP (11) still has an exponential number of SDP constraints. Each of these constraints are of size $m \times m$, leading to a heavy computational burden. Thus, to improve the computational performance, we implement a relaxation technique based on PCA [32] to decrease the size of the SDP constraints to a lower size of $m_1 \times m_1$ with $m_1 < m$. To that end, we excavate the dominant and accordingly useful information in the ambiguity set by projecting the random variable $\boldsymbol{\xi}^t$ with original dimension m to a m_1-dimensional vector $\boldsymbol{\xi}_r^t$.

Consider the ambiguity set $\mathcal{D}^t(\mathcal{S}^t, \boldsymbol{\mu}^t, \boldsymbol{\Sigma}^t)$. Given a positive definite matrix $\boldsymbol{\Sigma}^t$, its eigendecomposition can be presented as $\boldsymbol{\Sigma}^t = \boldsymbol{U}^t \boldsymbol{\lambda}^t \boldsymbol{U}^{t\top} = \boldsymbol{U}^t \boldsymbol{\Lambda}^{t\frac{1}{2}} (\boldsymbol{U}^t \boldsymbol{\Lambda}^{t\frac{1}{2}})^\top$, where $\boldsymbol{U}^t \in \mathbb{R}^{m \times m}$ and $\boldsymbol{\Lambda}^t \in \mathbb{R}^{m \times m}$ is a diagonal matrix with the eigenvalues on the diagonal. By sorting the eigenvalues in $\boldsymbol{\Lambda}^t$ in a decreasing order and picking the largest m_1 elements and their corresponding eigenvectors, $\boldsymbol{U}_{m \times m_1}^t \boldsymbol{\Lambda}^{t\frac{1}{2}}_{m_1} \boldsymbol{\xi}_r^t$ will carry the highest variability of the uncertainty in the data in m_1 dimension. Thus, the distributional ambiguity set for random variable $\boldsymbol{\xi}_r^t$ centered around zero can be represented as:

$$\mathcal{D}_r^t = \left\{ F_r \left| \begin{array}{l} \mathbf{P}(\boldsymbol{\xi}_r^t \in \mathcal{S}_r^t) = 1 \\ \mathbf{E}_{F_r}[\boldsymbol{\xi}_r^t] = \mathbf{0}_{m_1} \\ \mathbf{E}_{F_r}[(\boldsymbol{\xi}_r^t)(\boldsymbol{\xi}_r^t)^T] \preceq \mathbf{I}_{m_1} \end{array} \right. \right\}, \text{ with} \tag{12}$$

$$\mathcal{S}_r^t := \{\boldsymbol{\xi}_r^t \in \mathbb{R}^{m_1} : \boldsymbol{U}_{m \times m_1}^t \boldsymbol{\Lambda}^{t\frac{1}{2}}_{m_1} \boldsymbol{\xi}_r^t + \boldsymbol{\mu}^t \in \mathcal{S}^t\}, \tag{13}$$

where $\mathbf{0}_{m_1}$ is a vector of zeros with dimension m_1 and $\mathbf{I}_{m_1}$ is an identity matrix with dimension m_1.

Therefore, problem (8) with the projected uncertainty vector $\boldsymbol{\xi}_r^t$ instead of the $\boldsymbol{\xi}^t$ and under the new ambiguity set $\mathcal{D}_r^t$ is an approximation of O-DDSP and we refer to it as P-DDSP. Similar to Subsection (IV-A), we can derive an equivalent SDP reformulation of P-DDSP as follows:

$$\min_{\substack{\boldsymbol{x}, s^t, \boldsymbol{\beta}^t, \\ \boldsymbol{q}_r^t, \boldsymbol{Q}_r^t}} \mathcal{C}_1(\boldsymbol{x}) + \sum_{t=1}^T \left[s^t + \mathbf{I}_{m_1} \bullet \boldsymbol{Q}_r^t \right], \tag{14}$$

$$\text{s.t. } \boldsymbol{x} \in \mathcal{X}$$

$$\boldsymbol{\beta}^t \geq \mathbf{0}, s^t \in \mathbb{R}, \boldsymbol{q}_r^t \in \mathbb{R}^{m_1}, \boldsymbol{Q}_r^t \in \mathbb{R}^{m_1 \times m_1}$$

$$\begin{bmatrix} M'_{11} & M'_{21}{}^\top \\ M'_{21} & M'_{22} \end{bmatrix} \succeq 0, \qquad \forall \boldsymbol{\lambda}^t \in vert(\mathcal{F}^t), \ t = 1, 2, ..T$$

$$M'_{11} = s^t + \boldsymbol{\lambda}^{t\top}(\hat{\boldsymbol{d}} - \hat{\boldsymbol{B}}\boldsymbol{x}) - \boldsymbol{\beta}^{t\top}\boldsymbol{b} + (\boldsymbol{\beta}^{t\top}A - \boldsymbol{\lambda}^{t\top}\hat{\boldsymbol{C}}_{(\boldsymbol{x})})\boldsymbol{\mu}^t$$

$$M'_{21} = \left[\frac{1}{2}(\boldsymbol{q}_r^t + (\boldsymbol{U}_{m \times m_1}^t \boldsymbol{\Lambda}^{t\frac{1}{2}}_{m_1})(\boldsymbol{A}^\top \boldsymbol{\beta}^t - \hat{\boldsymbol{C}}_{(\boldsymbol{x})}^\top \boldsymbol{\lambda}^t)) \right]$$

$$M'_{22} = \left[\boldsymbol{Q}_r^t \right].$$

C. Delayed Constraint Generation Algorithm

Since enumerating all the vertices of the feasibility region $\mathcal{F}^t$ in both O-DDSP (11) and P-DDSP (14) is not practical, we present an efficient algorithm (as shown in Algorithm 1) to solve both models in this subsection. In fact, SDP constraints are highly computationally intense and the global optimization techniques would quickly become computationally expensive, especially for such large-scale problems. In Algorithm 1, we present the details of the delayed constraint generation

algorithm [33] to overcome this challenge and solve (11) efficiently. According to this algorithm, we start to solve a relaxed version of the SDP problem by considering a subset of $vert(\mathcal{F}^t)$. We then check if the SDP constraint in (11) is satisfied. If yes, then the optimal solution is found. Otherwise, we add the corresponding feasibility cuts.

Algorithm 1: Delayed Constraint Generation Algorithm

Step1: Find initial sets of vertices of the feasibility region $\{\boldsymbol{\lambda}^t | \boldsymbol{\lambda}^t \geq 0, \hat{\boldsymbol{A}}^\top \boldsymbol{\lambda}^t + \hat{\boldsymbol{e}} = 0\}$ for each t, as the starting point for $\mathcal{V}_f^t$.

Step2: Solve the problem (11) considering the vertices only in $\mathcal{V}_f^t$, namely the **master problem**. Then, save the optimal value R^* and optimal solution $\mathcal{X}^*$.

Step3: Solve the following biconvex **subproblem** on each period t using the optimal values of s, $\boldsymbol{q}^t$, $\boldsymbol{Q}^t$, $\boldsymbol{x}$ from ***Step2***:

$$\min_{\boldsymbol{\lambda}^t, \boldsymbol{\xi}^t} \quad s^t + \boldsymbol{\xi}^{t\top}\boldsymbol{q}^t + \boldsymbol{\xi}^{t\top}\boldsymbol{Q}^t\boldsymbol{\xi}^t - \boldsymbol{\lambda}^{t\top}(\hat{\boldsymbol{B}}\boldsymbol{x} + \hat{\boldsymbol{C}}_{(\boldsymbol{x})}\boldsymbol{\xi}^t - \hat{\boldsymbol{d}})$$

$$\text{s.t.} \quad \forall \boldsymbol{\lambda}^t \in \mathcal{F}^t, \boldsymbol{\xi}^t \in \mathcal{S},$$

and save the optimal value as r^{t*} and optimal solution as $\boldsymbol{\lambda}^{t*}$. To do so, we use the alternating direction search algorithms.

Step4: If $r^{t*} \geq 0$ for all t, then (9b) is feasible and the optimal solution R^* obtained in ***Step2*** is optimal for O-DDSP. Otherwise, if $r^{t*} < 0$ for some t, then add the optimal $\boldsymbol{\lambda}^{t*}$ obtained in ***Step3*** to set $\mathcal{V}_f^t$ and go to ***Step2***.

Note that in the first step of Algorithm 1 and to find the vertices of the feasibility region $\{\boldsymbol{\lambda}^t | \boldsymbol{\lambda}^t \geq 0, \hat{\boldsymbol{A}}^\top \boldsymbol{\lambda}^t + \hat{\boldsymbol{e}} = 0\}$, we solve a linear program with different objective functions each time subject to this feasibility region. In this way, the obtained solution is a vertex of the feasibility region. In Step3, the optimization problem is nonlinear due to the bilinear term $\boldsymbol{\lambda}^{t\top}\boldsymbol{\xi}^t$ and it can be solved by the nonlinear optimization solvers to the exact optimal point for small-size problems. For the large-scale problems, it is possible to solve the problem in Step3 to the near-optimal solutions using alternating direction search techniques. In addition, note that Algorithm 1 is specially designed to solve O-DDSP (11), while similarly one can be developed to solve P-DDSP (14) and thus we omit it due to limited space.

V. NUMERICAL RESULTS

In this section, first we briefly introduce the IEEE 33-Bus system and the historical real data that we acquired. Next, we investigate the effectiveness of Algorithm 1 to solve the proposed multi-period model O-DDSP (11) and its PCA relaxation P-DDSP (14). Finally, sensitivity analyses are carried out by using in-sample and out-of-sample tests to further validate the effectiveness of the obtained planning solutions. All numerical tests were executed on a PC with an Intel Core i7-7700 CPU and 16 GB RAM. The master problem and subproblem were solved by MOSEK solver on CVX software.

A. Data

We consider a modified IEEE 33-Bus radial distribution network [34]. All generators and DGs are connected to the grid in three phases. Two active power DGs are considered at buses 15 and 29, and three reactive power sources are considered at buses 11, 13, and 32. The reactive power sources are of the hybrid (capacitive and inductive) compensator type which have the ability to both generate and absorb reactive power in order to stabilize the voltage. Three RDGs are to be placed in the network, i.e., $K = 3$. The optimal capacity of each of these RDGs (between $\underline{x}_k = 0.2$ MW and $\bar{x}_k = 2.5$ MW for k=1,2,3) should be decided by O-DDSP (11), while the siting decision will be analyzed later. We consider two folds of uncertainties: 1) the hourly active/reactive loads at each bus, i.e., d_{pn}^t/d_{qn}^t, at period t, and 2) the renewable generation outputs of each RDG, i.e., s_k^t, at period t. Our dataset includes hourly load data (from Pecanstreet project [28]) and hourly wind generation coefficient data (from ERCOT [29]) for 4 years. As we consider our planning periods to be seasonal, we have 2160 (90 days $\times$ 24 hours) sample points available at each season.

B. Performance Analyses of O-DDSP and P-DDSP

First, in Table I, we report the performance of O-DDSP by considering different number of seasons T (labeled as "T") and two different candidate location alternatives to install the RDGs, which are labeled as Alte. **A** at buses (32, 17, 21) and Alte. **B** at buses (24, 29, 13). In addition, we consider two different cases regarding the support set $\mathcal{S}$, in which the results inside the parenthesis are obtained when increasing the support set from $[-2\sigma, +2\sigma]$ to $[-3\sigma, +3\sigma]$. From the table, we can observe that the penetration of the RDGs (labeled as "Pene. (MWh)") increases when we consider longer planning horizons, and the grid will have higher costs (labeled as "Tot. Cost (K$)") as the planning horizon increases. However, it is easy to infer that the average second-stage costs per season (labeled as "Season OPT (K$)") is significantly lower when we consider a longer planning horizon. This clearly shows the long-term cost-wise benefits of the RDG penetration. Moreover, we can see that adopting a less accurate ambiguity set, which may be due to insufficient data, usually leads to longer computational time (labeled as "Time (s)") and more iterations (labeled as "# ITE") towards optimality. Indeed, a broader support set means a larger solution space to search for the optimality, and thus potentially longer computational time. In addition, comparing the results for alternatives **A** and **B**, we can see that the siting decision will have a decisive effect on the outcomes, especially on the renewable penetration, and we will investigate it in detail in Subsection V-C.

Next, we compare the performance of O-DDSP with its approximation, P-DDSP, on solution quality and time in Tables II and III. In Table II, we consider $m_1 = 67, 40, 30, 20$, and 10, which corresponds to 100%, 60%, 45%, 30%, and 15% of the size of $\boldsymbol{\xi}$, respectively. In Table II, the optimality gap between P-DDSP and O-DDSP (labeled as "GAP(%)") is defined as $\frac{\pi^*(m) - \pi^*(m_1)}{\pi^*(m)} \times 100\%$, where $\pi^*(m)$ and $\pi^*(m_1)$ are the optimal values of O-DDSP and P-DDSP, respectively. We can observe that P-DDSP will provide very high-quality

TABLE I
THE O-DDSP PERFORMANCE

Alte.	T	Tot. Cost (K$)	Season Cost (K$)	Pene. (MWh)	Time (s)	# ITE
A	2	6988 (7159)	1747 (1790)	0.60 (0.60)	59 (192)	4 (9)
	4	14768 (9558)	923 (597)	0.60 (0.60)	112 (223)	11 (12)
	6	19624 (14975)	545 (415)	2.77 (1.75)	611 (757)	22 (13)
	8	24305 (17147)	379 (268)	3.44 (2.57)	606 (461)	6 (15)
B	2	7228 (7286)	1807 (1821)	0.60 (0.60)	3275 (1370)	6 (9)
	4	14919 (9666)	932 (705)	0.60 (0.60)	1196 (2271)	11 (9)
	6	19837 (19889)	551 (553)	2.63 (2.76)	2353 (2334)	13 (17)
	8	26177 (26023)	409 (407)	3.98 (4.07)	3033 (3627)	19 (19)

solutions in dramatically shorter CPU times (labeled as "Time (s)"). For instance, for the case with $T = 8$ and support $[-2\sigma, +2\sigma]$, solving the original reformulation O-DDSP takes 7227 seconds, while P-DDSP obtains a high-quality solution (less than 1% optimality gap) in only 247 seconds when $m_1 = 10$. This is because the dimension of SDP constraints for O-DDSP is 68×68 while for P-DDSP with $m_1 = 10$ the dimension is 11×11. The O-DDSP, having to deal with large number of big SDP constraints, becomes very slow while the P-DDSP is able to solve the problem and provide high quality solution in a fraction of O-DDSP's computational time. Table III shows that such improvements get amplified by further increasing the number of periods to $T = 16$, as it leads to larger number of SDP constraints (e.g., P-DDSP is 2000 times faster). Therefore, we show that when the size of the problem is very large-scale so that it cannot be handled by the solvers, our proposed P-DDSP maintains a very tight optimality gap, thus, it brings about practical usefulness. Note that the results in Tables II and III is based on a given candidate location to install RDGs at buses 17, 32, and 24. To better show the efficiency of model P-DDSP, we try to obtain average performance of Algorithm 1 by running instances that consider different candidate location alternatives.

TABLE II
O-DDSP VS. P-DDSP

	m_1	T=2 seasons			T=8 seasons		
		GAP(%)	Time (s)	# ITE	GAP(%)	Time (s)	# ITE
P-DDSP	10	1.36	24	6	0.92	247	12
	20	0.17	40	8	0.26	251	11
	30	0.88	45	8	0.70	296	11
	40	1.22	125	13	0.48	556	13
	67	0.59	822	18	0.13	4939	20
O-DDSP		**0**	**721**	**15**	**0**	**7227**	**24**

TABLE III
O-DDSP VS. P-DDSP: $T = 16$

	T	Tot. Cost (K$)	Time (s)	# ITE
O-DDSP	16	17015.1	31925	65
P-DDSP	16	17097.4	63	4

In particular, we consider different values of m_1 and two different support sets. For given m_1 and the support set, models O-DDSP and P-DDSP run forty times, with each time considering a different location alternative, and we obtain the average optimality gap and computational time, as shown in Fig. 2. From Fig. 2, we summarize the following observations: 1) the gap between O-DDSP and P-DDSP is very small, which

is as low as 5% when $m_1 = 10$; 2) the computational time of P-DDSP is largely reduced when m_1 decreases; 3) more accurate ambiguity sets with tighter supports will bring not only less conservative solutions but also shorter computational times; 4) even when less accurate ambiguity set is used (e.g., due to lack of enough data), our proposed P-DDSP will lead to high percentage reduction of the computational time.
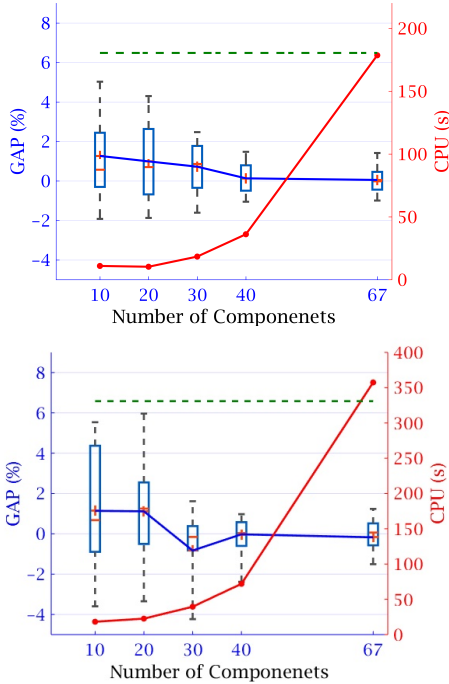

Fig. 2. Performance of P-DDSP (with m_1=10, 20, 30, 40, and 67) on two different support sets: $[-\sigma, +\sigma]$ on top, and $[-2\sigma, +2\sigma]$ in below. Green line indicates the average O-DDSP time. The red line shows the run time in seconds and the blue line shows the optimality gap.

In table IV, we compare the out-of-sample performance between the two proposed methods, i.e., O-DDSP and P-DDSP, versus the conventional SO and RO methods. The data-driven SO problem is solved by SAA method, and we use the scenario-based polyhedral uncertainty sets introduced in [35, 36] to solve the RO problem. Here, we randomly divide the N samples into a training dataset (75% of the samples) and a validation dataset (25% of the samples). We first solve the sizing problem on the training dataset using each method, and then use the validation dataset to obtain the out-of-sample performance of each method. We can observe that the proposed P-DDSP formulation reduces the computational times dramatically the sample size N increases. This is especially important for our data-driven problem because the accuracy of the results will increase by including more samples. However, the disadvantage of SO and RO methods is that its computational time severely increases with large sample sizes while increasing the sample size in our proposed DRO based methods, i.e., O-DDSP and P-DDSP, will only increase the accuracy of the mean vectors and covariance matrices of uncertainty without adversely affecting the size of the problem. Especially, by using the PCA concept, our P-DDSP formulation outperform SO and RO in all cases with its extra-ordinary shorter CPU times.

TABLE IV
OUT-OF-SAMPLE PERFORMANCE

T	N	OPT (K$)				Time (s)			
		RO	SO	O-DDSP	P-DDSP	RO	SO	O-DDSP	P-DDSP
2	200	3664.8	3532.6	3545.8	3586.7	109	73	464	18
	600	4047.5	3688.7	3752.6	3704.0	707	423	1355	100
	1000	3359.8	2997.5	2988.3	2992.7	1992	1031	1171	41
	2000	1570.2	1570.2	1570.2	1570.2	7028	3329	1177	40
4	200	11090.4	10599.2	11125.5	10911.8	306	194	896	86
	600	12400.0	11067.9	11084.2	11088.2	1910	1003	1628	111
	1000	8435.8	7520.7	7523.6	7523.6	7081	3107	2246	208
	2000	3426.4	3426.4	3426.4	3426.4	28089	12190	3880	75

C. RDGs Siting Planning Analyses

From Section V-B, we can observe that the sizing decision is highly dependent on the predetermined RDG locations. Mathematically, it is possible to consider the siting decision in model (8) by turning the parameter z_{kn} to a binary decision variable and solve the mixed-integer sizing and siting problem. Unfortunately, due to the weakness of optimization solvers and the CPU limits to this date, the corresponding problem becomes quickly intractable dealing with SDP constraints integrated with large-scale mixed-integer variables, not to mention that the current model O-DDSP is already difficult to solve. Nevertheless, considering that usually the options for RDG placement are limited due to natural barriers and constraints, enumerating these different candidate location alternatives and comparing their corresponding costs is a plausible alternative approach that we investigate hereafter.

For each given candidate location alternative, we will assess the corresponding first-stage investment cost and second-stage operational cost on an 8-season planning horizon (2 years) in order to identify the elements of a successful RDG placement as general siting strategies. Note that solving the SDP formulation of P-DDSP (or O-DDSP) does not tell each individual cost component of the second-stage operational cost, i.e., the costs of purchasing active/reactive electricity from the main grid, the fuel and emission costs of dispatchable DGs, because all of them are aggregated by the dual of the second-stage problem. Thus, to numerically retrieve these cost components in the second stage, we perform the following procedure and thereby validate the first-stage solution, where we use the historical data in years 2014-2015 and 2016-2017 as the primary and secondary sets of historical data, respectively.

a) Use the primary set of historical data to construct the ambiguity set $\mathcal{D}^t$ at each period t.
b) Consider a set of different location alternatives for the placement of K RDGs in set $\mathcal{G}_K$.
c) For each candidate location alternative $g^s \in \mathcal{G}_K$:
   1) Solve P-DDSP and report the first-stage costs, then fix the first-stage decision.
   2) To retrieve the second-stage cost components, sample 1000 scenarios from the primary (resp. secondary) set of historical data on each period t to conduct in-sample (resp. out-of-sample) test. That is, for each scenario, given the obtained first-stage decision, solve the deterministic second-stage problem, i.e., model (1).

Here, we consider the IEEE 33-Bus network as shown in Fig. 3 and let $K = 3$ and $|\mathcal{G}_K| = 60$ to show the main results.

To begin with, we employ the in-sample test results to present the detailed cost components in the second stage, more specifically the variance of the electricity purchased from the main grid under these different location alternatives in $\mathcal{G}_K$, as shown in Fig. 4. We can observe that the candidate location alternatives with higher penetration of renewables lead to higher variance of electricity purchase, which indicates higher pressure towards the main grid and further implies a trade-off between renewable penetration and the pressure towards the main grid. This trade-off is one of the major drawbacks of renewable expansion and confirms the complexity of the problem and the necessity of developing optimal planning strategies. In the following part, we will show how we can address this trade-off by inferring some general rules for the optimal planning of the RDGs. The rules will help the distribution network operators and utility companies validate different candidate location alternatives quickly and reasonably.
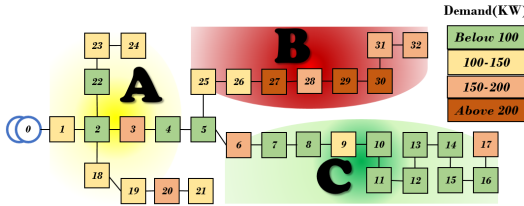

Fig. 3. The Radial 33-Bus Network with varying range of demand on each bus. Region A representing medium load. Region B has few but highly demanded buses. Region C has a larger number of low demanded buses.
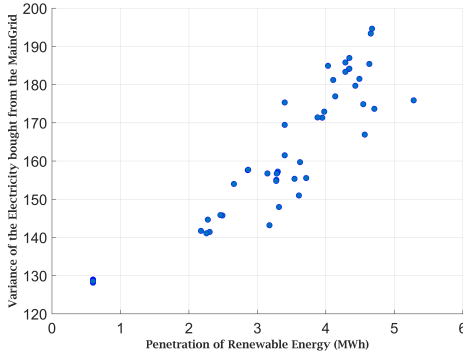

Fig. 4. The variance of the electricity purchased from the main grid is plotted against the penetration of renewable energy for 60 different location alternatives.

To clearly derive some RDG siting strategies, we report numerical results corresponding to 9 representative location alternatives in Fig. 5 and the numerical results corresponding to another 9 alternatives in Fig. 6. Here, we use the out-of-sample test results to validate the obtained first-stage decision and better present the second-stage operational performance against those future scenarios (e.g., in the secondary set of historical data). Fig. 5 shows the results including the first-stage investment (labeled as RDG Inv.) and maintenance costs and the average second-stage operational cost of individual location alternatives. We can see that the placement of the RDGs at buses that are further away from bus 0 will reduce the

pressure on the main grid, the traditional and costly electricity generation at buses 15 and 29, and the total cost in the network.
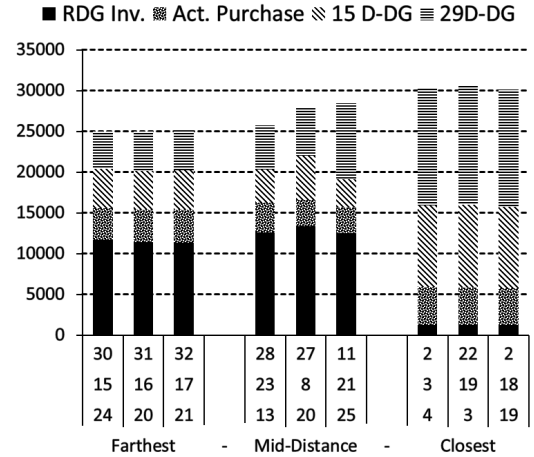

Fig. 5. The distribution network costs ($\$\times 10^3$) for 9 different location alternatives of 3 RDGs; long (resp. short) distance from bus 0 on the left (resp. right) side of the horizontal axis.

Moreover, the distance between a bus and bus 0 is not the only decisive factor for optimal RDG planning strategy. Note that in Fig. 3, we use different colors to represent the load intensity at each bus as well as at different demand regions, and our numerical results show that these load intensities are an important factor of successful RDG placement. In particular, demand regions B and C are both far away from bus 0 so they are potential regions for RDG placement, and it is clear to see from Fig. 6 that ignoring both demand regions B and C (candidate location alternatives labeled as "Ignores both Regions", i.e., candidate location alternatives that does not place any RDGs in regions B and C) does not have economic justification, as it leads to minimum RDG penetration and highest total cost. More specifically, we observe that candidate location alternatives that ignore region B (i.e., a highly demanded region) bring the highest costs of purchasing electricity from bus 0 and lead to high pressure on the dispatchable DG at bus 29 (which is in region B) to produce electricity. Thus, a sparse arrangement of RDGs (e.g., candidate location alternatives labeled as "Covers both regions") within the distribution network with consideration of the regional demand is the best strategy, which will lead to the maximum RDG penetration and minimum total cost. By merely comparing the two location alternatives that place RDGs at buses (24,32,17) and at buses (18,22,2) as shown in Fig. 6, there is a $5,000,000 decrease in total cost on a 2-year horizon when picking optimal planning decisions. This means the utility companies can have as much as 20% decrease in the total cost (selecting the candidate location alternatives sparsely to include high-demand regions) in the microgrid while following their renewable expansion goals.

Finally, in Table V, the out-of-sample performance of our proposed planning methodology is compared with SO and RO. Here, we randomly divide 500 scenarios in two parts; we first use the training dataset (includes 400 samples) to solve the sizing and siting problem, and then we use the validation dataset (includes 100 samples) to examine the performance
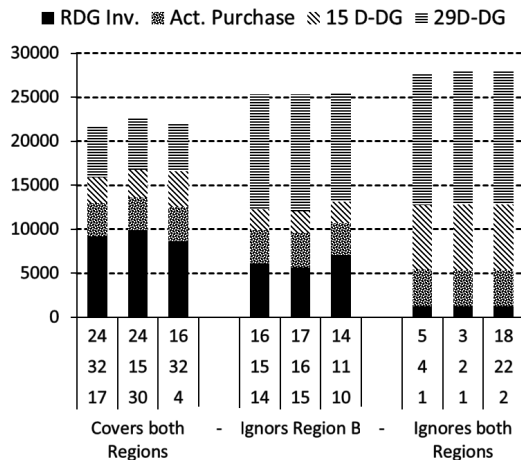
Fig. 6. The distribution network costs ($\times 10^3$) for 9 different location alternatives in 3 groups of strategies regarding demand regions B and C.

TABLE V
OUT OF SAMPLE PERFORMANCE FOR SITING AND SIZING

Method	N	OPT (K$)		Time (s)	
		T=2	T=8	T=2	T=8
SAA	5	3620.4	24895.2	13,432	***
	10	3651.3	26782.4	28,957	***
	100	***	***	***	***
RO	5	3611.0	25225.2	21,202	***
	10	3816.4	26814.4	***	***
	100	***	***	***	***
DRO	**400**	**3542.5**	**22964.4**	**120**	**3481**

of the solution obtained by each method. The time used by MOSEK to reach optimality is reported in column labeled "Time (s)," where "***" means MOSEK cannot solve the problem within ten hours time limit, if so we use the best integer solution retrieved from the solver to examine the out-of-sample performance. The out-of-sample performance is reported in the column labeled "OPT (K$)," where "***" means MOSEK cannot obtain any feasible solutions within the time limit. We can observe that our proposed planning methodology outperforms SO and RO dramatically. Indeed, for the case of $T = 8$ and $N = 5$, MOSEK could not solve the SO and RO problems in more than 2 days, so that's why we set a time limit. When we choose sample sizes larger than 100 for the SO and RO methods, the sizing and siting problem becomes completely impractical computationally. Our proposed methodology, on the other hand, is based on capturing the first- and second-moment information in the ambiguity set, and increasing the sample size does not increases the size of our problem but only make the solution more accurate. The extremely short computational times of the P-DDSP is also promising, as the P-DDSP problem can be further extended to include the siting decision as well, which is our ongoing research that deals with mixed-inter SDP problems.

VI. CONCLUSION AND REMARKS

The recent technological advances together with governments supportive policies have progressively promoted the penetration of renewable electricity generation. The renewable energy, though, is a highly uncertain source of electricity and its integration into the grid is extremely challenging and requires novel planning techniques. In this paper, a data-driven DRO approach with first- and second-moment information is proposed for RDG sizing problem to safely increase the penetration of low-carbon-emission renewable energy while we avoid damaging the network equipment, blackouts, more power losses, etc. that happens by over-sizing and under-sizing the renewable sites. Additionally, a tight PCA based approximation of the DRO model is proposed which speeds up the solution procedure thus enable practical usefulness in industry. Unlike conventional RO and SO methods that become impractically time-consuming to solve the sizing problems with large sample sizes, the proposed DRO models use the large sets of historical data in their advantage to increase the accuracy of their ambiguity sets thus their solutions, and they are solved quickly. In addition, unlike other DRO works in literature, we do not limit our paper to only problems with small uncertainty vectors, or a small number of SDP constraints. Moreover, a general framework for siting strategies was proposed through sensitivity analyses, and it is compared to conventional SO and RO methods.

Through numerical results we showed that while the conventional methods become intractable to solve the siting and sizing problems with even 5 samples, our method is able to use huge sample sizes to make planning decisions. Moreover, we show that our optimal planning of the RDGs will lead to minimum planning cost, less pressure on the main grid, maximum penetration of renewables, and decreasing greenhouse gas production from the dispatchable DGs. More specifically, our numerical results indicate that a sparse arrangement of RDGs with optimal sizing and siting decisions will lead up to a 20% reduction in costs. This is achieved while guaranteeing the performance of the distribution system through our proposed detailed mathematical model based on the optimal power flow equations. Overall, we provided an optimal planning tool for the management of utility companies when RDG placement decision is not trivial due to the uncertainties in data and the large-scale size of the problem.

Future research will consider developing tractable mixed-integer DRO methods that can integrate the sizing and siting of RDGs in one DRO model instead of using sensitivity analyses for siting. In addition, alternative ambiguity set techniques, such as Phi-divergence and Wasserstein ambiguity sets will be analyzed. Another relevant avenue of research is the extension of this work to include battery storage in the model.

REFERENCES

[1] "Renewable energy policies in a time of transition," https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2018/Apr/IRENA_IEA_REN21_Policies_2018.pdf.

[2] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan, "Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy," *IEEE POWER ENERGY M*, vol. 15, no. 2, pp. 61–73, 2017.

[3] P. Prakash and D. K. Khatod, "Optimal sizing and siting techniques for distributed generation in distribution systems: A review," *RENEW SUST ENERG REV*, vol. 57, pp. 111–130, 2016.

[4] Z. Abdmouleh, A. Gastli, L. Ben-Brahim, M. Haouari, and N. A. Al-Emadi, "Review of optimization techniques applied for the integration of distributed generation from renewable energy sources," *RENEW ENERG*, vol. 113, pp. 266–280, 2017.

[5] F. Qiu and J. Wang, "Chance-constrained transmission switching with guaranteed wind power utilization," *IEEE T POWER SYST*, vol. 30, no. 3, pp. 1270–1278, 2014.

[6] A. Ehsan and Q. Yang, "Optimal integration and planning of renewable distributed generation in the power distribution networks: A review of analytical techniques," *APPL ENERG*, vol. 210, pp. 44–59, 2018.

[7] V. Oree, S. Z. S. Hassen, and P. J. Fleming, "Generation expansion planning optimisation with renewable energy integration: A review," *RENEW SUST ENERG REV*, vol. 69, pp. 790–803, 2017.

[8] T. Bahar, O. Singh, and V. Yadav, "Optimal planning strategies of DG in distribution systems," in *Applications of Computing, Automation and Wireless Systems in Electrical Engineering*. Springer, 2019, pp. 333–345.

[9] B. Palmintier, R. Broderick, B. Mather, M. Coddington, K. Baker, F. Ding, M. Reno, M. Lave, and A. Bharatkumar, "On the path to sunshot. emerging issues and challenges in integrating solar with the distribution system," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2016.

[10] E. Ali, S. A. Elazim, and A. Abdelaziz, "Ant lion optimization algorithm for optimal location and sizing of renewable distributed generations," *RENEW ENERG*, vol. 101, pp. 1311–1324, 2017.

[11] O. D. M. Dominguez, M. Pourakbari-Kasmaei, J. R. S. Mantovani, and M. Lavorato, "Environmentally committed short-term planning of electrical distribution systems considering renewable based dg siting and sizing," in *EEEIC/I&CPS Europe*. IEEE, 2017, pp. 1–6.

[12] O. D. Melgar-Dominguez, M. Pourakbari-Kasmaei, and J. R. S. Mantovani, "Adaptive robust short-term planning of electrical distribution systems considering siting and sizing of renewable energy based dg units," *IIEEE T SUSTAIN ENERG*, vol. 10, no. 1, pp. 158–169, 2018.

[13] H. Su, "Siting and sizing of distributed generators based on improved simulated annealing particle swarm optimization," *ENVIRON SCI POLLUT R*, pp. 1–12, 2017.

[14] E. Ali, S. A. Elazim, and A. Abdelaziz, "Optimal allocation and sizing of renewable distributed generation using ant lion optimization algorithm," *ELECTR ENG*, vol. 100, no. 1, pp. 99–109, 2018.

[15] H. Xing, H. Fan, X. Sun, S. Hong, and H. Cheng, "Optimal siting and sizing of distributed renewable energy in an active distribution network," *CSEE J POWER ENERGY SYST*, vol. 4, no. 3, pp. 380–387, 2018.

[16] H. Cetinay, F. A. Kuipers, and A. N. Guven, "Optimal siting and sizing of wind farms," *RENEW ENERG*, vol. 101, pp. 51–58, 2017.

[17] D. A. Quijano, J. Wang, M. R. Sarker, and A. Padilha-Feltrin, "Stochastic assessment of distributed generation hosting capacity and energy efficiency in active distribution networks," *IET GENER TRANSM DIS*, vol. 11, no. 18, pp. 4617–4625, 2017.

[18] Z. Wang, B. Chen, J. Wang, J. Kim, and M. M. Begovic, "Robust optimization based optimal dg placement in microgrids," *IEEE T SMART GRID*, vol. 5, no. 5, pp. 2173–2182, 2014.

[19] O. D. Melgar-Dominguez, M. Pourakbari-Kasmaei, and J. R. S. Mantovani, "Adaptive robust short-term planning of electrical distribution systems considering siting and sizing of renewable energy based DG units," *IEEE T SUSTAIN ENERG*, vol. 10, no. 1, pp. 158–169, 2019.

[20] R. H. Zubo, G. Mokryani, H.-S. Rajamani, J. Aghaei, T. Niknam, and P. Pillai, "Operation and planning of distribution networks with integration of renewable distributed generators considering uncertainties: A review," *RENEW SUST ENERG REV*, vol. 72, pp. 1177–1198, 2017.

[21] D. Bertsimas, X. V. Doan, K. Natarajan, and C.-P. Teo, "Models for minimax stochastic linear optimization problems with risk aversion," *MATH OPER RES*, vol. 35, no. 3, pp. 580–602, 2010.

[22] 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.

[23] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *OPER RES*, vol. 62, no. 6, pp. 1358–1376, 2014.

[24] F. Alismail, P. Xiong, and C. Singh, "Optimal wind farm allocation in multi-area power systems using distributionally robust optimization approach," *IEEE T POWER SYST*, vol. 33, no. 1, pp. 536–544, 2017.

[25] X. Chen, W. Wu, B. Zhang, and C. Lin, "Data-driven dg capacity assessment method for active distribution networks," *IEEE T POWER SYST*, vol. 32, no. 5, pp. 3946–3957, 2016.

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

[27] Y. Yang and W. Wu, "A distributionally robust optimization model for real-time power dispatch in distribution networks," *IEEE T SMART GRID*, 2018.

[28] "More than the largest source of energy data and water data," https://dataport.cloud/.

[29] "The Electric Reliability Council of Texas (ERCOT)," http://www.ercot.com/gridinfo/generation.

[30] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification—part I," *IEEE T POWER SYST*, vol. 28, no. 3, pp. 2554–2564, 2013.

[31] A. Ben-Tal and A. Nemirovski, "On polyhedral approximations of the second-order cone," *MATH OPER RES*, vol. 26, no. 2, pp. 193–205, 2001.

[32] J. Cheng, R. Li-Yang Chen, H. N. Najm, A. Pinar, C. Safta, and J.-P. Watson, "Distributionally robust optimization with principal component analysis," *SIAM J OPTIM.*, vol. 28, no. 2, pp. 1817–1841, 2018.

[33] W. Wei, F. Liu, and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE T SUSTAIN ENERG*, vol. 7, no. 1, pp. 289–300, 2016.

[34] A. M. Fathabad, "Test data of modified ieee 33-bus system," Click for the DropBox Link.

[35] A. Nemirovski, "Lectures on robust convex optimization," *Lecture notes. Georgia Institute of Technology*, 2012.

[36] A. Velloso, A. Street, D. Pozo, J. M. Arroyo, and N. G. Cobos, "Two-stage robust unit commitment for co-optimized electricity markets: an adaptive data-driven approach for scenario-based uncertainty sets," *IIEEE T SUSTAIN ENERG*, 2019.