

Distributionally Robust Optimal Reactive Power Dispatch with Wasserstein Distance in Active Distribution Network

Jun Liu, *Member, IEEE*, Yefu Chen, Chao Duan, *Member, IEEE*, Jiang Lin, *Senior Member, IEEE*, and Jia Lyu

Abstract—The uncertainties from renewable energy sources (RESs) will not only introduce significant influences to active power dispatch, but also bring great challenges to the analysis of optimal reactive power dispatch (ORPD). To address the influence of high penetration of RES integrated into active distribution networks, a distributionally robust chance constraint (DRCC)-based ORPD model considering discrete reactive power compensators is proposed in this paper. The proposed ORPD model combines a second-order cone programming (SOCP)-based model at the nominal operation mode and a linear power flow (LPF) model to reflect the system response under certainties. Then, a distributionally robust optimization (WDRO) method with Wasserstein distance is utilized to solve the proposed DRCC-based ORPD model. The WDRO method is data-driven due to the reason that the ambiguity set is constructed by the available historical data without any assumption on the specific probability distribution of the uncertainties. And the more data is available, the smaller the ambiguity would be. Numerical results on IEEE 30-bus and 123-bus systems and comparisons with the other three-benchmark approaches demonstrate the accuracy and effectiveness of the proposed model and method.

Index Terms—Active distribution network, chance constraint, renewable energy source, optimal reactive power dispatch (ORPD).

I. INTRODUCTION

OPTIMAL reactive power dispatch (ORPD), also known as steady-state voltage control, is important for the secure and economic operation of power systems [1]. It aims at finding the optimal control variables of the power

system to minimize a certain objective function [2], [3]. Besides, ORPD satisfies a bunch of physical and operation constraints such as branch currents and bus voltage magnitudes to be within their reasonable ranges. The optimal control variables consist of both continuous variables, i.e., voltage magnitudes, and discrete variables, i.e., ratios of transformers and the number of switchable capacitors/reactors. Considering the nonlinearity of power flow equations together with the numerous continuous and discrete decision variables, ORPD problem is a rather complex optimization problem.

Traditionally, ORPD is subject to a series of nonlinear constraints that make it a mixed-integer nonlinear programming problem. Many methods have been proposed to solve the ORPD model, which can be generally divided into two categories: artificial intelligence methods (such as particle swarm optimization [4], simulated annealing [5]) and conventional methods (such as linear programming [6], gradient-based optimization [7], interior-point method [8]). In these optimization methods, the discrete variables are treated as continuous variables firstly and then rounded off to the nearest integer value, which may lead to unnecessary deviation in the objective function and more constraint violations. Nowadays, mixed-integer programming approaches have been frequently used, and they can be efficiently solved by branch-and-cut, branch-and-bound or cutting plane methods [9]. Besides, the global optimal solution can be guaranteed under the Karush-Kuhn-Tucker (KKT) conditions when the models are relaxed to be a convex problem.

Meanwhile, since the penetration of renewable energy sources (RESs) increases sharply in recent years [10], they will probably bring great challenges to the ORPD due to their stochastic nature. For example, the RES uncertainties could result in frequent voltage fluctuations due to the delay of system reaction, especially in distribution systems [11], [12]. Therefore, the integration of such volatile renewable energy into power system requires more considerations on the planning and scheduling [13]-[16]. Presently, there are several approaches to tackle the uncertainties of RES in power systems. One is stochastic programming (SP) [17] - [19], which supposes that the uncertainties follow a presumed probability distribution, and then it is feasible to be transformed into a deterministic problem. For example, a chance-constrained programming method is proposed in [18], in

Manuscript received: January 1, 2019; accepted: August 21, 2019. Date of CrossCheck: August 21, 2019. Date of online publication: April 29, 2020.

This work was supported in part by National Key Research and Development Program of China (No. 2018YFB0905000), in part by Key Research and Development Program of Shaanxi (No. 2017ZDCXL-GY-02-03).

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

J. Liu (corresponding author), C. Duan, and J. Lyu are with Shaanxi Key Laboratory of Smart Grid, School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (e-mail: eeliujun@mail.xjtu.edu.cn; duanchao@stu.xjtu.edu.cn; z15878351771@163.com).

Y. Chen is with the Electric Power Dispatch and Control Center of Guangdong Power Grid Corporation, Guangzhou, Guangdong 510050, China (e-mail: cheneyefu@stu.xjtu.edu.cn).

L. Jiang is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, U.K. (e-mail: ljiang@liverpool.ac.uk).

DOI: 10.35833/MPCE.2019.000057



which the uncertain nodal power injections and random branch outages are considered as uncertainty sources. A stochastic multi-objective ORPD problem is presented in [19], wherein a wind-integrated power system with both loads and wind power generation uncertainties is considered. The second approach is robust optimization (RO) [9], [20], [21]. For example, a two-stage robust ORPD model is proposed in [9], which can achieve a more robust solution than traditional deterministic approaches, although it has longer computation time. The RO-related methods have been demonstrated to give out rather conservative solutions, because they ignore the exact probability information of uncertainties and only search for the solution that performs best in the worst-case scenarios.

To bridge the gap between SP and RO, the third approach called distributionally robust method (DRO) has been brought up [22]-[30]. DRO assumes that the true probability distribution lies in an ambiguity set, and it minimizes the worst-case expected cost over this ambiguity set. The most popular ambiguity set is the moment-based, e.g., the set of probability distribution with the first- and second-order moments [28]-[30]. However, the information of first- and second-order moments cannot cover all the probability information of the true distribution. We do not even know the moment information of the true probability distribution. All that we have are the available historical data. Therefore, a desired ambiguity set should contain the true probability distribution. The ambiguity set will become smaller with the increasing historical data.

Currently, most existing references are using two-stage model for ORPD problem under uncertainties, and the solutions are usually complicated due to their multi-level structures [9]. Besides, non-anticipative constraints on ORPD decisions are always not considered in two-stage model formulations [31]. Moreover, the two-stage distributionally robust optimization model is much more time-consuming and difficult to be solved when more historical data are available. To address the above concerns, a new distributionally robust chance-constraint (DRCC) ORPD model with Wasserstein distance is proposed for active distribution networks in this paper. The main contributions are as follows:

1) In the ORPD model formulation, we propose an approximate second-order cone programming (SOCP) power flow model, which combines an exact SOCP model at the nominal operation mode and a linear power flow (LPF) model to express the system response under uncertainties. It largely inherits the accuracy of the exact SOCP model. Moreover, it is a single-level mixed-integer programming formulation rather than multi-level formulation. Therefore, it can be directly solved by popular commercial solvers like Gurobi, and no other complex algorithms for minimum-maximum structure problems are required.

2) We firstly apply Wasserstein distance to the DRCC-based ORPD model to construct the ambiguity set, so as not to presume any true probability distribution for the uncertain RESs. The Wasserstein-distance-based distributionally robust optimization (WDRO) method is a data-driven method, and larger quantity of data will lead to smaller ambiguity set and

less conservative solution.

3) We then reformulate the original DRCC-based ORPD model to be a mixed-integer convex programming model, according to the Wasserstein-distance-based ambiguity set. Simulations are performed on IEEE standard test systems, and optimal solutions of the proposed WDRO method are compared to those of other three benchmark approaches. The proposed WDRO is able to guarantee fast computation performance as RO, which is better than moment-based distributionally robust optimization (MDRO) and SP approaches. Moreover, the proposed WDRO method is also effective when a large number of historical data are available, which benefits from the unique reformulation of the proposed DRCC-based ORPD model.

The structure of the paper is as follows. Section II introduces the formulation of DRCC-based ORPD model. A WDRO method is then proposed in Section III to solve the special optimization problem. In Section IV, numerical results on IEEE 30-bus and 123-bus systems and comparisons with another three benchmark approaches are presented to demonstrate the accuracy and effectiveness of the proposed model and method. Finally, conclusions are drawn in Section V.

II. ORPD MODEL IN DISTRIBUTION NETWORKS

In this section, a novel DRCC-based ORPD model is proposed considering multiple continuous and discrete decision variables as well as the power flow constraints and uncertainties from RES. The proposed model combines an SOCP model at the nominal operation mode and an LPF model to reflect the system response under uncertainties.

A. ORPD Model Based on SOCP Relaxation

Recently, the conic relaxation technique has been deeply studied [32]-[35] to relax the nonconvex power flow equations by the use of SOCP, and it has been recognized that the conic relaxation has no gap or small gap to the original exact power flow equations in most distribution networks. An ORPD model based on SOCP has been formulated in [9]. In an RES-integrated power system, the SOCP-based ORPD model can be given as follows [9]:

$$\min f(P_i^g, Q_i^g, v_i) \quad (1)$$

s.t.

$$\begin{cases} P_i^g + P_i^w - P_i^l = \sum_{j \in \delta(i)} P_{ij} - \sum_{k \in \pi(i)} (P_{ki} - r_{ki} l_{ki}) \quad \forall i \in B \\ Q_i^g + Q_i^w - Q_i^l = \sum_{j \in \delta(i)} Q_{ij} - \sum_{k \in \pi(i)} (Q_{ki} - x_{ki} l_{ki}) + b_{s,i} v_i^2 \quad \forall i \notin \Omega \\ Q_i^g + Q_i^w - Q_i^l + \frac{1}{2} [C_i^{\min} + S_i (2^0 \sigma_{i,0} + 2^1 \sigma_{i,1} + \dots + 2^{\tau_i} \sigma_{i,\tau_i})] = \\ \sum_{j \in \delta(i)} Q_{ij} - \sum_{k \in \pi(i)} (Q_{ki} - x_{ki} l_{ki}) + b_{s,i} v_i^2 \quad \forall i \in \Omega \end{cases} \quad (2)$$

$$\begin{cases} v_j^2 = v_i^2 - 2(r_{ij} P_{ij} + x_{ij} Q_{ij}) + (r_{ij}^2 + x_{ij}^2) l_{ij} \quad \forall (i,j) \in E \setminus T \\ \sum_{k=1}^{n_{ij}} h_{j,k} / t_{ij,k}^2 = v_i^2 - 2(r_{ij} P_{ij} + x_{ij} Q_{ij}) + (r_{ij}^2 + x_{ij}^2) l_{ij} \quad \forall (i,j) \in T \end{cases} \quad (3)$$

$$\begin{cases} -M(1-z_{ij,1})+v_j^2 \leq h_{j,k} \leq v_j^2 + M(1-z_{ij,1}) & \forall (i,j) \in T \\ -Mz_{ij,1} \leq h_{j,k} \leq Mz_{ij,1} & \forall (i,j) \in T, z_{ij,1}, z_{ij,2}, \dots, z_{ij,n_{ij}} \in \{0,1\} \\ \sum_{k=1}^{n_{ij}} z_{ij,k} = 1 & \forall (i,j) \in T \end{cases} \quad (4)$$

$$\begin{cases} v_i^2 - M(1-y_{i,k}) \leq \sigma_{i,k} \leq v_i^2 + M(1-y_{i,k}) \\ -My_{i,k} \leq \sigma_{i,k} \leq My_{i,k} & \forall i \in \Omega \cap \Omega_D, k=1, 2, \dots, \tau_i \\ 0 \leq 2^0 y_{i,0} + 2^1 y_{i,1} + \dots + 2^{\tau_i} y_{i,\tau_i} \leq (C_i^{\max} - C_i^{\min})/s_i \\ y_{i,0}, y_{i,1}, \dots, y_{i,\tau_i} \in \{0,1\} & \forall i \in \Omega \cap \Omega_D \end{cases} \quad (5)$$

$$\left\| \begin{matrix} 2P_{ij} \\ 2Q_{ij} \\ I_{ij} - v_i^2 \end{matrix} \right\|_2 \leq I_{ij} + v_i^2 \quad \forall (i,j) \in E \quad (6)$$

$$\begin{cases} \underline{v}^2 \leq v_i^2 \leq \bar{v}^2 & \forall i \in B \\ I_{ij} \leq (I_{ij}^{\max})^2 & \forall (i,j) \in E \\ \underline{P}^g \leq P_i^g \leq \bar{P}^g & \forall i \in B \\ \underline{Q}^g \leq Q_i^g \leq \bar{Q}^g & \forall i \in B \end{cases} \quad (7)$$

where the subscripts i or j and ij represent the specific bus and branch, respectively; $P_i^g, Q_i^g, P_i^w, Q_i^w, P_i^l, Q_i^l$ are the nominal active and reactive power for the generation from generators, injection from RES and power consumption of load, respectively; B, E are the sets of buses and branches, respectively; T is the set of branches with transformers; Ω is the set of buses for reactive power compensators; $\pi(i), \delta(i)$ are the sets of all parents and children of bus i , respectively; r_{ij}, x_{ij} are the resistance and reactance of branch (i,j) , respectively; $b_{s,i}$ is the shunt susceptance from bus i to the ground; C_i is the value of shunt capacitors/reactors at bus i ; s_i is the step size of shunt capacitors/reactors at bus i ; I_{ij}^{\max} is the current capacity limit of branch (i,j) ; P_{ij}, Q_{ij} are the active and reactive power flows from bus i to j , respectively; t_{ij} is the tap ratio of the transformer of branch (i,j) ; v_i is the nominal bus voltage magnitude; \bar{v}, \underline{v} are the upward and downward bus voltage magnitude, respectively; $\bar{P}^g, \underline{P}^g$ are the upward and downward active power of generators, respectively; and $\bar{Q}^g, \underline{Q}^g$ are the upward and downward reactive power of generators, respectively.

The constraint (2) denotes the power balance at each bus; constraints (3)-(6) denote the Ohm's law for each branch; constraints (4) and (5) are the constraints for discrete compensators; constraint (7) is the bounds for bus voltages, branch currents and the generator outputs. For any fixed forecasted load power (P^l, Q^l) and RES (P^w, Q^w), traditional ORPD aims to perform an optimal dispatch of reactive power while guaranteeing the power balance and security constraints of the system, i.e., (1)-(7). With the increasing growth of RES, the power system operation is influenced more deeply by the uncertainties of RES. By using automatic generation control (AGC) and automatic voltage regulation (AVR), the adverse impact of the RES uncertainties need to be considered when arranging optimal operation modes.

B. Response Model of Voltage-concerned System and Its Control Under Uncertainties

Since LPF model is convenient in dealing with the uncertainties of RES, a linear approximation of the AC power flow equation can be obtained as follows [36]:

$$\begin{cases} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} -B' & G'' \\ -G' & -B'' \end{bmatrix} \begin{bmatrix} \theta \\ v^2 \end{bmatrix} \\ f = \frac{1}{2} G^f v^2 - B^f \theta \end{cases} \quad (8)$$

where p and q are the vectors of the nodal injected active and reactive power, respectively; $Y = G + jB$ is the admittance matrix of power system; $Y' = G' + jB'$ is the admittance matrix without shunt elements; $Y'' = G'' + jB''$ is the modified admittance matrix of power system $G''_{ii} = G_{ii}, G''_{ij} = 0.5G_{ij}$; f is the nominal line flow; and G^f, B^f are the $n_l \times n_b$ matrices with $G^f_{ki} = -G^f_{kj} = G_{ij}, B^f_{ki} = -B^f_{kj} = B_{ij}$, and other elements as zeros, respectively. It is well-known that the state variables consist of three sub-vectors corresponding to the $V\theta, PV, PQ$ types of buses. Let θ, v denote the nominal bus voltage angles and magnitudes, respectively; R, L, S denote the sets for reference bus, PV buses, and PQ buses, respectively; then the state variables can be written as $\theta = (\theta_R, \theta_S, \theta_L)$, $v^2 = (v_R^2, v_S^2, v_L^2)$, and other variables and coefficient matrix can also be partitioned in the same manner. Generally, with the AVR control system, the exciters of the generator are able to maintain the pre-scheduled voltage magnitudes so that $\Delta v_R^2 = \Delta v_S^2 = 0$, and the reference bus has fixed phase angle, typically as 0. Then, (8) can be transformed partially into an incremental form as (9)-(11). In (9), v_L^2 is the voltage calculated by the SOCP-based model.

$$\begin{bmatrix} \Delta \theta_S \\ \Delta \theta_L \\ \Delta v_L^2 \end{bmatrix} = N^{-1} \begin{bmatrix} \Delta p_S \\ \Delta p_L \\ \Delta q_L \end{bmatrix} - N^{-1} H \begin{bmatrix} 0 \\ \Delta v_R^2 \\ \Delta v_S^2 \end{bmatrix} \quad (9)$$

$$\begin{bmatrix} \Delta q_R \\ \Delta q_S \end{bmatrix} = LN^{-1} \begin{bmatrix} \Delta p_S \\ \Delta p_L \\ \Delta q_L \end{bmatrix} + (-LN^{-1}H + M) \begin{bmatrix} 0 \\ \Delta v_R^2 \\ \Delta v_S^2 \end{bmatrix} \quad (10)$$

$$\begin{cases} H = \begin{bmatrix} -B''_{SR} & G''_{SR} & G''_{SS} \\ -B''_{LR} & G''_{LR} & G''_{LS} \\ -G'_{LR} & -B''_{LR} & -B''_{LS} \end{bmatrix} \\ N = \begin{bmatrix} -B'_{SS} & -B'_{SL} & G''_{SL} \\ -B'_{LS} & -B'_{LL} & G''_{LL} \\ -G'_{LS} & -G'_{LL} & -B''_{LL} \end{bmatrix} \\ M = \begin{bmatrix} -G'_{RR} & -B''_{RR} & -B''_{RS} \\ -G'_{SR} & -B''_{SR} & -B''_{SS} \end{bmatrix} \\ L = \begin{bmatrix} -G'_{RS} & -G'_{RL} & -B''_{RL} \\ -G'_{SS} & -G'_{SL} & -B''_{SL} \end{bmatrix} \end{cases} \quad (11)$$

Without the loss of generality, only wind power generation is considered as the fluctuating RES in this study, and any other types of RES can be integrated and modelled similarly. The actual active power outputs are considered as a combination of the forecasted power P^w plus a random forecasting error ξ . The wind farm is assumed to maintain a

fixed power factor $\cos \phi$. In order to address the ORPD problem under uncertainties, the traditional ORPD model should be modified. Firstly, we use the SOCP model from Section II-A to get a nominal operation mode $\mathbf{x} = (\mathbf{v}^2, \mathbf{P}^g, \mathbf{Q}^g, \mathbf{z}, \mathbf{y})$, and \mathbf{z}, \mathbf{y} denote sets of discrete variables. Then, (9) and (10) can be used to calculate incremental system response under RES uncertainties. Besides, with the regulation of AGC, generators can respond affinely to the total forecasting errors of active wind power. In such circumstance, the incremental part of active and reactive nodal power injections in (9) and (10) are random variables as:

$$\begin{cases} \Delta \tilde{\mathbf{p}}_S = -(\mathbf{1}^T \tilde{\boldsymbol{\xi}}) \boldsymbol{\alpha} + \tilde{\boldsymbol{\xi}}_S \\ \Delta \tilde{\mathbf{p}}_L = \tilde{\boldsymbol{\xi}}_L \\ \Delta \tilde{\mathbf{q}}_L = \tilde{\boldsymbol{\xi}}_L \tan \phi \end{cases} \quad (12)$$

where $\boldsymbol{\alpha}$ is the AGC participation factor of generation units; and $\tilde{\boldsymbol{\xi}} = (\tilde{\boldsymbol{\xi}}_R + \tilde{\boldsymbol{\xi}}_S + \tilde{\boldsymbol{\xi}}_L)$ is the random forecasting errors of RES generation.

Then, the incremental state variables in (9) and (10) will also become random variables. By substituting (12) into (9) and (10), the random state variables can be written as:

$$\begin{cases} \tilde{\mathbf{v}}_L^2 = (\mathbf{1}^T \tilde{\boldsymbol{\xi}}) \mathbf{C}^v \boldsymbol{\alpha} + \mathbf{D}^v \tilde{\boldsymbol{\xi}} + \mathbf{v}_L^2 \\ \tilde{\mathbf{f}} = (\mathbf{1}^T \tilde{\boldsymbol{\xi}}) \mathbf{C}^f \boldsymbol{\alpha} + \mathbf{D}^f \tilde{\boldsymbol{\xi}} + \mathbf{f} \\ \tilde{\mathbf{q}}_{RUS} = (\mathbf{1}^T \tilde{\boldsymbol{\xi}}) \mathbf{C}^q \boldsymbol{\alpha} + \mathbf{D}^q \tilde{\boldsymbol{\xi}} + \mathbf{q}_{RUS} \end{cases} \quad (13)$$

where $\mathbf{C}^v, \mathbf{D}^v, \mathbf{C}^f, \mathbf{D}^f, \mathbf{C}^q, \mathbf{D}^q$ are the constant matrices decided by the network parameters; and \mathbf{q}_{RUS} is the reactive power injection in the nominal operation mode, which is calculated by the SOCP-based model.

C. Formulation of DRCC-based ORPD Model

Based on the developed SOCP model and LPF model, a DRCC-based ORPD model can be formulated as (14)-(20) with constraints (2)-(7):

$$\min_x \mathbf{E}_{\mathbb{P}} \sup_{\mathbb{P} \in \hat{\mathbb{P}}_N} \left\{ \sum_{i=1}^{n_b} f_i^P \left[\mathbf{P}_i^g - (\mathbf{1}^T \tilde{\boldsymbol{\xi}}) \boldsymbol{\alpha}_i \right] + \sum_{i=1}^{n_b} f_i^Q (Q_i^g) + \bar{\mathbf{c}}^T \bar{\mathbf{r}} + \underline{\mathbf{c}}^T \underline{\mathbf{r}} \right\} \quad (14)$$

$$\mathbf{1}^T \boldsymbol{\alpha} = 1 \quad (15)$$

$$\begin{cases} \mathbf{P}^g - \underline{\mathbf{r}} \geq \underline{\mathbf{P}}^g & \underline{\mathbf{r}} \geq \mathbf{0} \\ \mathbf{P}^g + \bar{\mathbf{r}} \leq \bar{\mathbf{P}}^g & \bar{\mathbf{r}} \geq \mathbf{0} \end{cases} \quad (16)$$

$$\inf_{\mathbb{P} \in \hat{\mathbb{P}}_N} \mathbb{P} \left\{ -\underline{\mathbf{r}} \leq -(\mathbf{1}^T \tilde{\boldsymbol{\xi}}) \boldsymbol{\alpha} \leq \bar{\mathbf{r}} \right\} \geq 1 - \rho_1 \quad (17)$$

$$\inf_{\mathbb{P} \in \hat{\mathbb{P}}_N} \mathbb{P} \left\{ \underline{\mathbf{v}}_i^2 \leq \tilde{\mathbf{v}}_i^2 \leq \bar{\mathbf{v}}_i^2 \right\} \geq 1 - \rho_2 \quad \forall i \in L \quad (18)$$

$$\inf_{\mathbb{P} \in \hat{\mathbb{P}}_N} \mathbb{P} \left\{ \underline{\mathbf{f}}_j \leq \tilde{\mathbf{f}}_j \leq \bar{\mathbf{f}}_j \right\} \geq 1 - \rho_3 \quad \forall j = 1, 2, \dots, n_l \quad (19)$$

$$\inf_{\mathbb{P} \in \hat{\mathbb{P}}_N} \mathbb{P} \left\{ \underline{\mathbf{q}}_k \leq \tilde{\mathbf{q}}_k \leq \bar{\mathbf{q}}_k \right\} \geq 1 - \rho_4 \quad \forall k \in RUS \quad (20)$$

where \mathbb{P} is a probability distribution (measure); $\mathbf{E}_{\mathbb{P}}$ is the expectation with probability distribution \mathbb{P} ; $\bar{\mathbf{r}}, \underline{\mathbf{r}}$ are the upward and downward regulating reserves of generation units, respectively; $\bar{\mathbf{c}}, \underline{\mathbf{c}}$ are the upward and downward regulating re-

serve prices, respectively; n_b, n_l are the numbers of buses and lines, respectively; $\hat{\mathbb{P}}_N$ is the ambiguity set constructed from the historical data of forecasting error $\tilde{\boldsymbol{\xi}}$; and ρ_1 to ρ_4 are the tolerable violation probabilities.

According to [37], the objective function for the DRCC-based ORPD model is chosen for minimizing the worst-case expectation of operation costs in this study. Only minimizing the network losses without considering generator costs may conflict with the economic operation principles [38]. The operation costs consist of the reserve costs and the production costs of both the active and reactive power of generator whose cost functions are obtained from MATPOWER [37]. Actually, other alternative objectives such as minimizing network losses and the operation costs of tap ratio (TR) and switchable capacitors/reactors (SCRs), can also be incorporated into the proposed model easily, as all control variables in the proposed DRCC-based ORPD model are expressed explicitly [39]. Constraints (2)-(7) ensure the power balance and security limits of state variables in the nominal operation mode. Constraint (15) denotes the basic requirement for participation factors of AGC systems; constraints (16) and (17) ensure the adequacy of the upward and downward generation reserves; chance constraint (18) is for the voltage magnitudes at PQ buses under uncertainties; chance constraint (19) is for the line flow; and chance constraint (20) is for the reactive power outputs of generators.

III. SOLVING DRCC-BASED ORPD MODEL WITH WASSERSTEIN DISTANCE

A. Ambiguity Set with Wasserstein Distance

In the real-life application, the real probability distribution \mathbb{P} for random variables $\tilde{\boldsymbol{\xi}}$ is usually unknown, so we construct an empirical distribution $\hat{\mathbb{P}}_N = \sum_{k=1}^N \delta_{\tilde{\boldsymbol{\xi}}^{(k)}} / N$ as an estimation of the true \mathbb{P} , using the historical sample set $\{\tilde{\boldsymbol{\xi}}^{(1)}, \tilde{\boldsymbol{\xi}}^{(2)}, \dots, \tilde{\boldsymbol{\xi}}^{(N)}\}$ without any assumption of \mathbb{P} . Here, $\delta_{\tilde{\boldsymbol{\xi}}^{(k)}}$ is the Dirac distribution concentrating unit mass at $\tilde{\boldsymbol{\xi}}^{(k)}$, and these samples are the forecasting errors of the wind power outputs. Then the Wasserstein distance can be used to measure the distance between the empirical probability $\hat{\mathbb{P}}_N$ and the true probability \mathbb{P} , given by Definition 1.

Definition 1 (Wasserstein distance) [22], [23]: for any probability distribution $\mathbb{Q}_1, \mathbb{Q}_2 \in P(\mathcal{E})$, where $P(\mathcal{E})$ denotes the set of all probability distributions with support \mathcal{E} , the Wasserstein distance can be defined as:

$$W(\mathbb{Q}_1, \mathbb{Q}_2) = \inf_{\Pi} \left(\int_{\mathcal{E}^2} \|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\| \Pi(d\boldsymbol{\xi}_1, d\boldsymbol{\xi}_2) \right) \quad (21)$$

where Π is a joint distribution of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ with marginals \mathbb{Q}_1 and \mathbb{Q}_2 ; and $\|\cdot\|$ is a norm operator in \mathbf{R}^n used in this paper. Accordingly, we have $W(\hat{\mathbb{P}}_N, \mathbb{P}) \leq \varepsilon$, where ε is some sample-dependent monotone function. In our data-driven frame, given a historical sample set with N samples, the true distribution \mathbb{P} will be included in the following ambiguity set:

$$\hat{P}_N = \left\{ \mathbb{P} \in P(\mathcal{E}) : W(\mathbb{P}, \hat{\mathbb{P}}_N) \leq \varepsilon(N) \right\} \quad (22)$$

As shown in (22), the performance of the DRCC-based ORPD will heavily rely on the radius $\varepsilon(N)$ of the Wasserstein ball. Several possible choices for the radius are given in [40]-[42], and a radius from [42] is selected in this study, which can be represented as:

$$\varepsilon(N) = C \sqrt{\frac{1}{N} \log_{10} \left(\frac{1}{1-\beta} \right)} \quad (23)$$

where β is the confidence level; and C is the diameter of the support of the random variable that can be written as (24), where $\hat{\boldsymbol{\mu}}$ is the sample mean and the minimization over α_1 can be denoted by bisection search method.

$$\begin{aligned} C &= 2 \inf_{\xi_0 \in \Xi, \alpha > 0} \left(\frac{1}{2\alpha_1} \left(1 + \ln E_{\mathbb{P}} \left(e^{\alpha \|\tilde{\xi} - \xi_0\|_1^2} \right) \right) \right)^{1/2} \leq \\ & 2 \inf_{\alpha > 0} \left(\frac{1}{2\alpha_1} \left(1 + \ln E_{\mathbb{P}} \left(e^{\alpha \|\tilde{\xi} - \hat{\boldsymbol{\mu}}\|_1^2} \right) \right) \right)^{1/2} \approx \\ & 2 \inf_{\alpha > 0} \left(\frac{1}{2\alpha_1} \left(1 + \ln \left(\frac{1}{N} \sum_{k=1}^N e^{\alpha \|\tilde{\xi}^{(k)} - \hat{\boldsymbol{\mu}}\|_1^2} \right) \right) \right)^{1/2} \end{aligned} \quad (24)$$

B. Convex Reformulation of Chance Constraints

Considering a more general form of (17)-(20) as:

$$\inf_{\mathbb{P} \in \hat{\mathcal{P}}_N} \mathbb{P} \left(\mathbf{g}(\mathbf{x}, \tilde{\xi}) \leq \mathbf{0} \right) \geq 1 - \rho \quad (25)$$

where \mathbf{g} is linear with both the state variable \mathbf{x} and random variable $\tilde{\xi}$; and ρ is the tolerable violation probability.

In fact, the chance constraint is usually non-convex so that it is hard to find an equivalent formulation which is solvable. A feasible way is to find a convex conservative formulation of (25) as follows.

Firstly, find a deterministic ambiguity set that meets the robust constraint, for the random variable $\tilde{\xi}$ as:

$$\mathbf{g}(\mathbf{x}, \tilde{\xi}) \leq \mathbf{0} \quad \forall \tilde{\xi} \in \mathbf{U} \quad (26)$$

For a given historical sample set $\{\tilde{\xi}^{(1)}, \tilde{\xi}^{(2)}, \dots, \tilde{\xi}^{(N)}\}$, it is easy to compute the sample mean $\hat{\boldsymbol{\mu}}$ and sample covariance $\hat{\boldsymbol{\Sigma}}$. Then the standardised version $\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\Sigma}}^{-1/2}(\tilde{\xi} - \hat{\boldsymbol{\mu}})$ of random variable $\tilde{\xi}$ with sample set $\{\tilde{\boldsymbol{\theta}}^{(k)} = \hat{\boldsymbol{\Sigma}}^{-1/2}(\tilde{\xi}^{(k)} - \hat{\boldsymbol{\mu}})\}_{k=1,2,\dots,N}$ can be obtained. Thus $\tilde{\boldsymbol{\theta}}$ has the sample mean as $\mathbf{0}$, sample covariance as $\mathbf{1}$, and support as $\Theta = \{-\sigma_{\max} \mathbf{1} \leq \tilde{\boldsymbol{\theta}} \leq \sigma_{\max} \mathbf{1}\}$. Then, another set of $\mathbf{v} \subseteq \mathbf{R}^m$ needs to be found for random variable $\tilde{\boldsymbol{\theta}}$ that meets:

$$\sup_{\mathbb{Q} \in \hat{\mathcal{Q}}_N} \mathbb{Q} \left(\tilde{\boldsymbol{\theta}} \notin \mathbf{V} \right) \leq \rho \quad (27)$$

where \mathbb{Q} is the true distribution of $\tilde{\boldsymbol{\theta}}$, with ambiguity set $\hat{\mathcal{Q}}_N$ constructed using (22). As a result, $\mathbf{U} = \hat{\boldsymbol{\Sigma}}^{1/2} \mathbf{V} + \hat{\boldsymbol{\mu}}$ can be utilised as a possible ambiguity set for (26). Considering the rules of sample independence and equal variance among different components of $\tilde{\boldsymbol{\theta}}$, we restrict \mathbf{V} within the hypercube (28) so as to find out \mathbf{V} more efficiently.

$$\mathbf{V}(\sigma) = \{\boldsymbol{\theta} \in \mathbf{R}^m \mid -\sigma \mathbf{1} < \boldsymbol{\theta} < \sigma \mathbf{1}\} \quad (28)$$

Secondly, to reduce conservatism, σ needs to be as small as possible:

$$\begin{cases} \min_{0 \leq \sigma \leq \sigma_{\max}} \sigma \\ \text{s.t. } \sup_{\mathbb{Q} \in \hat{\mathcal{Q}}_N} \mathbb{Q} \left(\tilde{\boldsymbol{\theta}} \notin \mathbf{V}(\sigma) \right) \leq \rho \end{cases} \quad (29)$$

Using Lemma 1 in Appendix A, (29) is equivalent to:

$$\begin{cases} \min_{0 \leq \lambda, 0 \leq \sigma \leq \sigma_{\max}} \sigma \\ \text{s.t. } h(\sigma, \lambda) \leq \rho \end{cases} \quad (30)$$

where $h(\sigma, \lambda) = \lambda \varepsilon + \frac{1}{N} \sum_{k=1}^N \left[1 - \lambda \left(\sigma - \|\hat{\boldsymbol{\theta}}^{(k)}\|_{\infty} \right)^+ \right]^+$, $(x)^+ = \max(x, 0)$.

As shown in Lemma 1, (A1) is non-decreasing in σ . Therefore, (30) has a unique solution. The optimal solution can be found quickly by a nested bisection search method shown in Appendix B.

After determining the optimal σ , the hypercube $\mathbf{V}(\sigma)$ can be expressed as the convex hull of its vertices. Then the hypercube (28) can be obtained as $\mathbf{V}(\sigma) = \text{conv} \left(\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(2^m)}\} \right)$, specially, $\mathbf{V}(\sigma) = \text{conv}(\{-\sigma, \sigma\})$ for 1-dimensional random variable, and $\mathbf{V}(\sigma) = \text{conv}(\{\pm\sigma, \pm\sigma\})$ for 2-dimensional random variable. Accordingly, the constituted ambiguity set can be expressed as:

$$\mathbf{U}_g = \text{conv} \left(\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(2^m)}\} \right) \quad (31)$$

where $\mathbf{u}^{(i)} = \hat{\boldsymbol{\Sigma}}^{1/2} \mathbf{v}^{(i)} + \hat{\boldsymbol{\mu}}$, $1 \leq i \leq 2^m$. Then, (25) is equivalent to a deterministic expression:

$$\mathbf{g}(\mathbf{x}, \mathbf{u}^{(i)}) \leq \mathbf{0} \quad 1 \leq i \leq 2^m \quad (32)$$

Equation (32) is a set of linear constraints with the state variable \mathbf{x} . Thus, (17)-(20) can all be replaced by their corresponding deterministic linear forms using (32).

C. Reformulation of Objective Function

After the chance constraints being reformulated to a solvable form, there is still an obstacle $\sup_{\mathbb{P} \in \hat{\mathcal{P}}_N} E_{\mathbb{P}}(\cdot)$ in the objective function of the DRCC-based ORPD model. The objective function (14) inside the $\sup_{\mathbb{P} \in \hat{\mathcal{P}}_N} E_{\mathbb{P}}(\cdot)$ can be rewritten as:

$$l(\mathbf{x}, \tilde{\omega}) = c_2 \tilde{\omega}^2 + c_1 \tilde{\omega} + c_0 \quad \tilde{\omega} = \mathbf{1}^T \tilde{\xi} \quad (33)$$

Note that the objective function would be a convex quadratic function of \mathbf{x} and that (33) is also a convex quadratic function of $\tilde{\omega}$. Given a sample set $\{\hat{\omega}_1, \hat{\omega}_2, \dots, \hat{\omega}_N\}$ and its support $[\underline{\omega}, \bar{\omega}]$, use Lemma 2 in the Appendix C and let $\lambda = \max\{l'(\mathbf{x}, \bar{\omega}), -l'(\mathbf{x}, \underline{\omega})\}$ with $l'(\mathbf{x}, \tilde{\omega}) = 2c_2 \tilde{\omega} + c_1$, then:

$$\begin{cases} l(\mathbf{x}, \underline{\omega}) + \lambda(\underline{\omega} - \omega) \leq l(\mathbf{x}, \omega) \quad \forall \omega \in [\underline{\omega}, \bar{\omega}] \\ l(\mathbf{x}, \bar{\omega}) - \lambda(\bar{\omega} - \omega) \leq l(\mathbf{x}, \omega) \quad \forall \omega \in [\underline{\omega}, \bar{\omega}] \end{cases} \quad (34)$$

A close upper approximate of the worst-case evaluation of the cost in (14) can be rewritten as:

$$\begin{cases} \sup_{\mathbb{P} \in \hat{\mathcal{P}}_N} E_{\mathbb{P}} \{l(\mathbf{x}, \tilde{\omega})\} \leq \inf_{\lambda \in \mathbf{R}} (\lambda \varepsilon) + \frac{1}{N} \sum_{k=1}^N l(\mathbf{x}, \hat{\omega}_k) \\ \text{s.t. } l'(\mathbf{x}, \underline{\omega}) \leq \lambda \\ -l'(\mathbf{x}, \underline{\omega}) \leq \lambda \end{cases} \quad (35)$$

Equation (35) can be used as the worst-case cost with good computation performance, because the number of constraints and decision variables of (35) remain unchanged when using a larger historical data set in our WDRO method.

IV. NUMERICAL TESTS

A. Description of Two Test Systems

In the case study, the proposed DRCC-based ORPD model and the WDRO method are tested on both IEEE 30-bus system and IEEE 123-bus distribution system [43]. YALMIP is used as the modelling tool and Gurobi 7.5.2 as the solver, running on a 12-core 2.4 GHz Workstation.

The IEEE 30-bus system consists of 6 generators and 41 branches. The total load of the system is 189.2 MW. Five wind farms are connected to buses 3, 7, 17, 20, 24, respectively, and the capacity of each wind farm is set to 30 MW. The forecasted values of wind power output are assumed to be 50% of their capacities. Two transformers are installed at branches 11 and 16, respectively, whose ranges of TR are both $[0.95, 1.05]$ with the step size of 0.01. And two SCRs are installed at buses 7 and 26, respectively, whose capacities are both $[-0.12, 0.12]$ p.u. with the step size of 0.01 p.u.. The base is 100 MVA.

For the IEEE 123-bus distribution system, the total load is 3490 kW. Ten wind farms are connected to buses 5, 16, 29, 33, 46, 59, 64, 71, 75, 79, respectively, with 240 kW wind power capacity for each wind farm. The forecasted values of wind power output are also assumed to be 50% of their capacities. There are one transformer installed at the substation whose range of tap ratio is $[0.95, 1.05]$ with the step size of 0.01. And there are four SCRs installed at buses 12, 35, 54 and 108, whose capacities are $[-0.006, 0.006]$ p.u. with the step size of 0.002 p.u.. The base is 1000 kVA. Besides, numerical tests on other three-benchmark approaches are performed to compare with the proposed WDRO method.

1) RO: it requires (17)-(20) to be satisfied for all possible scenarios of the random variable.

2) SP: it presumes that the random variable follows the Gaussian distribution with a pre-given mean and covariance, and (17)-(20) are formulated as SOCP constraints by using inequality $\mathbb{P}\left\{\left|\tilde{\xi}-\mu\right|\geq\Phi^{-1}\left(1-\rho/2\right)\sigma\right\}\leq\rho$, where Φ is the cumulative distribution function of standard Gaussian random variable.

3) MDRO: the ambiguity set of MDRO is a set of probability distribution with a pre-given mean and variance, then (17)-(20) are formulated as SOCP constraints by using inequality $\mathbb{P}\left\{\left|\tilde{\xi}-\mu\right|\geq\sqrt{1/\rho}\sigma\right\}\leq\rho$.

In this paper, the tolerable violation probability in (17)-(20) are set to $\rho_1=\rho_2=\rho_3=\rho_4=0.05$, and the confidence level in (23) is set to be $\beta=0.9$. According to [44], Laplace distribution is used to generate realistic historical data whose sizes are ranging from 10^3 to 10^6 , with typical standard forecasting error variance [45]. Then, we use a set of 10^7 samples to estimate the simulated costs of the true distribution, so as to test the practical and out-of-sample performance of the corresponding method.

B. IEEE 30-bus System

1) The accuracy of the proposed model: the proposed DRCC-based ORPD model combines an SOCP model in the nominal operation mode and an LPF model to reflect the system response under uncertainties. Therefore, the accuracy of the proposed model should lie between the SOCP-based model and the LPF-based model with deterministic power flow. The comparisons of operation costs in different models are shown in Fig. 1, when the total wind power output forecasting error varies deterministically from -30 MW to 30 MW. The results demonstrate that the operation costs of the proposed model lie exactly between the SOCP model and the LPF model. And when the total forecasting error is -30 MW, the percentage difference between the proposed model and the SOCP model reaches a maximum value of -0.43% . Besides, when the total forecasting error is zero, the proposed model coincides with the SOCP model.

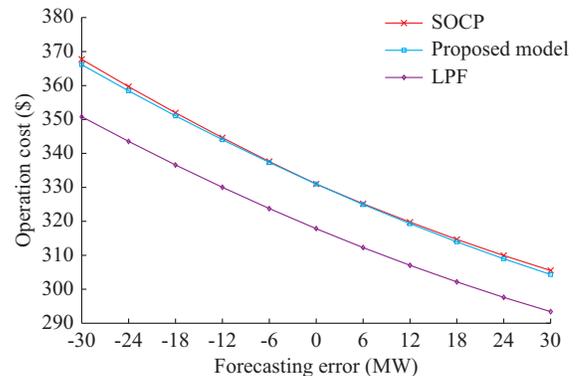


Fig. 1. Comparisons of objective operation costs for different models with different forecasting errors.

In addition, the proposed model has much higher accuracy than the LPF model in minimizing the operation costs with different total forecasting errors, which are depicted clearly in Fig. 1. Besides, the comparisons in Fig. 2 and Table I further presents the accuracy of the proposed model in voltage magnitudes and reactive power outputs, with the total forecasting error of -30 MW. Therefore, the proposed model has a considerably high accuracy as the SOCP model and a mathematical tractability as the LPF model, which makes it more attractive for ORPD under uncertainties.

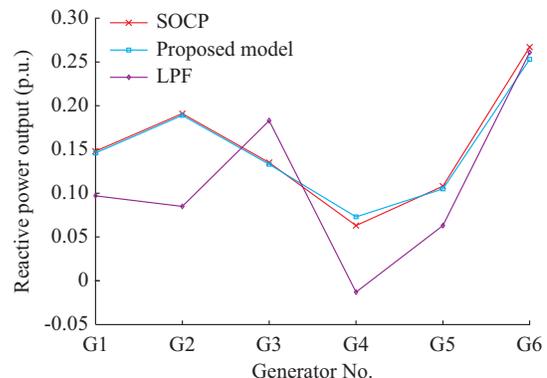


Fig. 2. Comparisons of reactive power outputs for different models.

TABLE I
COMPARISONS OF VOLTAGE MAGNITUDES OF PQ BUSES FOR DIFFERENT MODELS

Bus No.	SOCP (p.u.)	Proposed model (p.u.)	Error (p.u.)	LPF (p.u.)	Error (p.u.)
3	0.974	0.974	0.0000	0.978	0.0033
4	0.969	0.969	-0.0000	0.973	0.0033
5	0.975	0.976	0.0005	0.982	0.0073
6	0.963	0.963	0.0000	0.967	0.0031
7	0.967	0.967	0.0002	0.972	0.0048
8	0.951	0.951	0.0000	0.955	0.0041
9	0.972	0.972	-0.0004	0.983	0.0108
10	0.967	0.967	-0.0007	0.971	0.0038
11	0.972	0.972	-0.0004	0.983	0.0108
12	0.973	0.973	0.0000	0.984	0.0111
14	0.964	0.962	-0.0017	0.965	0.0014
15	0.970	0.970	0.0004	0.978	0.0079
16	0.965	0.964	-0.0004	0.974	0.0087
17	0.965	0.964	-0.0007	0.970	0.0056
18	0.960	0.960	-0.0000	0.966	0.0058
19	0.959	0.958	-0.0002	0.964	0.0050
20	0.964	0.963	-0.0003	0.968	0.0048
21	0.969	0.969	-0.0001	0.970	0.0011
24	0.989	0.989	-0.0002	0.991	0.0016
25	1.025	1.025	0.0001	1.026	0.0014
26	1.045	1.045	0.0002	1.047	0.0022
28	0.968	0.967	-0.0002	0.972	0.0048
29	1.015	1.015	0.0002	1.016	0.0012
30	1.003	1.003	0.0002	1.005	0.0014

2) Effectiveness of reactive power-related chance constraints: the optimization results of the proposed model with different chance constraints are shown in Table II.

TABLE II
COMPARISONS OF PROPOSED MODEL UNDER DIFFERENT CHANCE CONSTRAINTS

Model	Generator No.	P_g (p.u.)	Q_g (p.u.)	Voltage (p.u.)	SCR1 capacity (p.u.)	SCR2 capacity (p.u.)	TR1 (p.u.)	TR2 (p.u.)
Complete proposed model	G1	0.308	0.186	0.999	0.09	0.09	1.02	0.96
	G2	0.434	0.219	0.989				
	G3	0.079	0.155	0.961				
	G4	0.186	0.055	0.977				
	G5	0.073	0.110	0.997				
	G6	0.087	0.195	1.040				
Without voltage chance constraint	G1	0.307	0.183	0.996	0.10	0.09	1.00	1.05
	G2	0.433	0.215	0.986				
	G3	0.078	0.148	1.043				
	G4	0.185	0.055	0.967				
	G5	0.074	0.110	0.989				
	G6	0.090	0.200	1.041				
Without reactive power chance constraint	G1	0.308	0.177	0.997	0.09	0.10	1.01	0.95
	G2	0.434	0.208	0.988				
	G3	0.079	0.151	0.951				
	G4	0.186	0.065	0.977				
	G5	0.076	0.136	1.001				
	G6	0.085	0.170	1.037				

The results show that compared with the complete proposed model, the active power of generators in the model without (18) does not have much difference. But the reactive power related variables, i. e., the discrete control devices TR1, TR2, SCR1 and the continuous reactive power output of generators varies a lot, which will make the voltages at generator buses closer to their upper or lower limits. For example, the voltage magnitudes of G3 and G6 increase to 1.043 p.u. and 1.041 p.u., respectively, which are under potential risk of over-voltage; and the voltage magnitude of G4 decreases to 0.967 p.u., which may violate the under-excitation limit. For another comparison, if the reactive power chance constraints (20) is excluded from the complete proposed ORPD model, the active power does not have much difference. However, the reactive power of generators, e. g., G1, G2, G3, G6 decreases, which might lead to the decrease of voltage magnitudes at the corresponding generator buses. This change makes the voltage magnitude of G3 under a risk of violating the under-excitation limit. The results show that (18) and (20) are significant in the DRCC-based ORPD model to ensure a safe operation under uncertainties.

3) Comparison of discrete reactive power variables at different penetration levels of wind power: the optimized results of the discrete reactive power variables at different penetration levels of wind power are shown in Table III and Table IV. Under low wind power penetration condition (wind power output is $P^w = 5$ MW, with a maximum load of 189.2 MW), the reactive power compensators, i. e., SCR1, SCR2, TR1, stay nearly unchanged as the forecasting error variance increases from 0.03 p.u. to 0.12 p.u., because such low wind power penetration would not induce much voltage fluctuations in the system. However, the situation varies with higher penetration of wind power. For higher wind power penetration ($P^w = 50$ MW), the transformer taps change significantly with the a small step size of 0.01 p.u.. The operation statuses of SCRs also vary significantly with small step size, whereas SCRs with bigger step size of 0.02 p.u. only change when the forecasting error variance increases to a big enough value. In summary, no matter with high or low penetration of wind power, the reactive power compensators respond accurately under uncertainties, which verifies the effectiveness and accuracy of the proposed model.

TABLE III
COMPARISONS OF DISCRETE REACTIVE POWER VARIABLES WITH STEP SIZE OF SCR AS 0.01 P.U.

Forecasting error	Reactive power with $P^w=5$ MW				Reactive power with $P^w=50$ MW			
	TR1 (p.u.)	TR2 (p.u.)	SCR1 capacity (p.u.)	SCR2 capacity (p.u.)	TR1 (p.u.)	TR2 (p.u.)	SCR1 capacity (p.u.)	SCR2 capacity (p.u.)
0.03	1.01	0.95	0.10	0.09	0.99	0.96	0.10	0.09
0.06	1.01	1.03	0.10	0.09	1.00	1.05	0.10	0.10
0.09	1.01	0.96	0.10	0.09	1.01	1.05	0.09	0.08
0.12	1.01	1.05	0.10	0.09	1.02	1.03	0.09	0.09

TABLE IV
COMPARISONS OF DISCRETE REACTIVE POWER VARIABLES WITH STEP SIZE OF SCR AS 0.02 P.U.

Forecasting error	Reactive power with $P^w=5$ MW				Reactive power with $P^w=50$ MW			
	TR1 (p.u.)	TR2 (p.u.)	SCR1 capacity (p.u.)	SCR2 capacity (p.u.)	TR1 (p.u.)	TR2 (p.u.)	SCR1 capacity (p.u.)	SCR2 capacity (p.u.)
0.03	1.01	0.95	0.10	0.08	1.01	0.96	0.10	0.10
0.06	1.01	1.05	0.10	0.08	1.00	1.05	0.10	0.10
0.09	1.01	0.96	0.10	0.08	1.01	1.05	0.10	0.10
0.12	1.01	0.95	0.10	0.08	1.02	0.96	0.08	0.08

C. IEEE 123-bus Distribution System

The performances and advantages of the WDRO method on the proposed DRCC-based ORPD model are further compared with other three-benchmark approaches on the IEEE 123-bus distribution system.

1) Method conservatism and reliability comparison: the comparisons of objective operation costs on IEEE 123-bus distribution system are shown in Table V.

Among the simulated costs of different approaches, we can sort the method conservatism as: RO > WDRO (10^3) > MDRO > WDRO (10^4 , 10^5 , 10^6) > SP. In addition, it can be seen clearly from Fig. 3 that the simulated cost of RO is the highest and that of SP is the lowest, due to the fact that RO ignores most of the probabilistic information while SP assumes the precise knowledge about the presumed true probability distribution. In other words, RO outputs the most conservative solutions and SP provides the most optimistic solu-

tions. Since MDRO assumes partial knowledge, namely the first- and second-order moments of the probability distribution, it certainly reduces conservatism to some extent compared with RO. However, MDRO is still too conservative.

TABLE V
COMPARISON OF OPERATION COSTS

Method	Objective cost (\$)	Simulated cost (\$)
RO	6.8754	6.8683
Proposed method (10^3)	5.8452	5.8134
Proposed method (10^4)	5.4835	5.4748
Proposed method (10^5)	5.3238	5.3207
Proposed method (10^6)	5.2511	5.2498
MDRO	5.7035	5.6964
SP	5.1704	5.1634

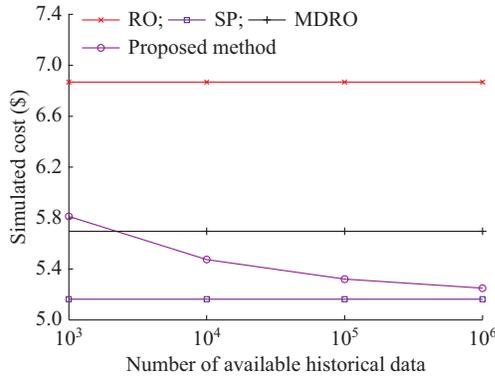


Fig. 3. Comparisons of simulated costs among different methods.

All the simulated costs of the three conventional approaches are nearly irrelevant with the quantity of data, except the proposed WDRO method. Since the proposed WDRO fully relies on the available historical data, the WDRO will provide a conservative solution as RO when it is short of data. And it can get closer to the result of the SP approach when more historical data is available. Although SP obtains the lowest simulated costs, it fails to guarantee the reliability level of the security constraints, which can be seen in Fig. 4. This is because the true probability distribution would always be different from the presumed Gaussian distribution used in the SP approach, while all the other tested approaches, including RO, MDRO and the proposed WDRO, are able to ensure higher reliability level (>95%) due to the robust nature.

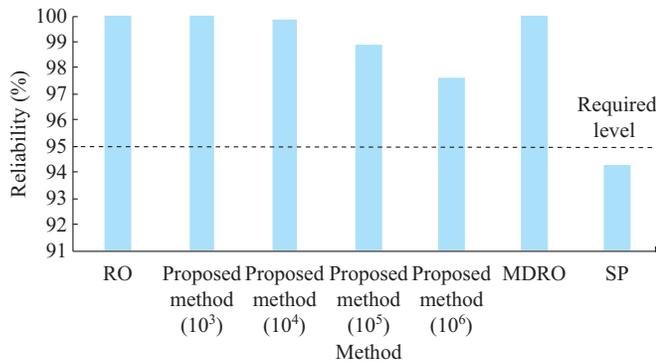


Fig. 4. Comparisons of reliability among different methods.

2) Data-driven characteristic of the WDRO method: as analyzed above, all the simulated costs of the other three approaches are nearly irrelevant with the quantity of the data samples, except the proposed WDRO method. In fact, the WDRO method would safely reduce the reliability level to a slight extent when more data is available. The objective operation cost will be lower with the increasing quantity of available data, which demonstrates that the WDRO is a data-driven method, and it can be less conservative with more historical data. It can also be seen that the simulated costs are always less than the objective values in Table V, because the true probability distribution usually differs from the conservative worst-case one. Moreover, the percentage difference between the objective values and simulated costs for the proposed WDRO method is also plotted in Fig. 5. The difference

may be large when the quantity of available historical data is as small as 1000. If more data is obtained, the difference will decrease and eventually approach 0, which demonstrates the data-driven nature of the proposed WDRO method.

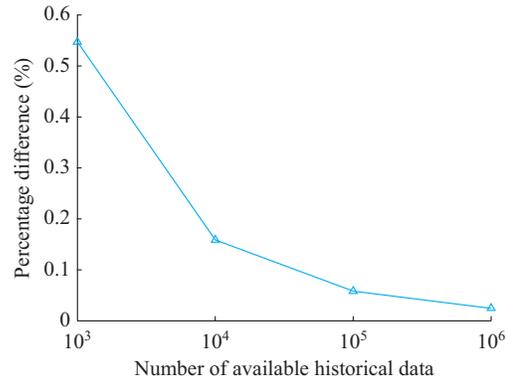


Fig. 5. Percentage differences between objective function and simulated costs with increasing data samples.

3) Comparison of computation performance: the comparison of the computation performance for all the methods is shown in Table VI, and the wind capacity is 10×240 kW. The calculation of the whole WDRO method consists of two parts. One part is the construction of U in (26) with available historical data, which can be completed before optimization. It can be concluded from the IEEE 123-bus distribution system case that the computation time of uncertainty set does not increase with the order of magnitude of the available dataset. This characteristic guarantees that the proposed WDRO method stays effective when more data is available.

TABLE VI
COMPARISON OF COMPUTATIONAL PERFORMANCE ON IEEE 123-BUS
DISTRIBUTION SYSTEM

Method	Construction time (s)	Solver time (s)
RO		0.76
WDRO (10^3)	3.04	0.65
WDRO (10^4)	3.34	0.89
WDRO (10^5)	14.36	0.73
WDRO (10^6)	269.42	0.77
MDRO		1.12
SP		1.18

The other part of the calculation is the optimization process. The computation time of WDRO is almost the same as RO approach, and much less than those of SP and MDRO methods. The proposed WDRO method reformulates the chance constraints into a bunch of linear constraints, which is similar to RO method. However, MDRO and SP methods both reformulate the chance constraints into SOCP constraints that induce larger computation burden. Most importantly, the computation time of WDRO is not sensitive to the quantity of historical data, namely that the solver time would not increase when more historical data are used, which further demonstrates better computation performance of the proposed WDRO method.

V. CONCLUSION

This paper proposes a DRCC-based ORPD model under uncertainties, whose ambiguity set is constructed by Wasserstein distance. Different from the conventional ORPD model, the proposed model is a combination of an exact SOCP model in nominal operation mode and an LPF model to reflect the system response under uncertainties.

1) Numerical case studies on IEEE 30-bus system demonstrate that the proposed model largely inherits both the accuracy of exact SOCP model and the tractability of LPF model. The model is also able to deal with discrete control variables such as transformer tap ratios and switchable capacitors/reactors, which ensures a reliable and efficient ORPD under uncertainties from volatile RES.

2) A WDRO method is proposed only based on historical data without any assumption on the specific probability distribution of the uncertainties. When more historical data are available, the solution will become less conservative to guarantee the required reliability level of security constraints.

3) Numerical studies on IEEE 123-bus distribution system further verifies the DRCC-based ORPD model and the WDRO method, by comparing the conservatism, reliability, data-driven characteristic and computation performance to those of the other three approaches. Compared with RO and SP approaches, the proposed WDRO method can robustly give a less conservative solution. Compared to MDRO, WDRO can extract more information of the true probability distribution by directly using the historical dataset. Besides, the proposed WDRO is able to guarantee fast computation performance as RO, which is better than MDRO and SP approaches, and stays effective when a large number of historical data are available.

APPENDIX A

The underlying formulation is developed in [26].

Lemma 1:

$$\sup_{\mathcal{Q} \in \hat{\mathcal{Q}}_N} \mathbb{Q}(\tilde{\mathbf{g}} \notin \mathcal{V}(\sigma)) = \inf_{\lambda \geq 0} \left\{ \lambda \varepsilon + \frac{1}{N} \sum_{k=1}^N \left[1 - \lambda \left(\sigma - \|\hat{\mathbf{g}}^{(k)}\|_{\infty} \right)^+ \right]^+ \right\} \quad (\text{A1})$$

APPENDIX B

The method is summarised in Algorithm 1 in which the function $\text{bisearch}(f(\cdot), a, b)$ returns the minimum of $f(\cdot)$ in the interval $[a, b]$ by performing a bisection search. Note that $h(\sigma, \lambda)$ is convex in λ for a fixed σ , so the bisection search in Step 4 of Algorithm 1 is well-defined. Since Algorithm 1 only involves function evaluations, it efficiently solves the problem (30).

APPENDIX C

The underlying formulation is developed in [24].

Lemma 2: given a random variable $\tilde{\omega} \in \mathbf{R}^m$ with closed and convex support \mathcal{E} , the Wasserstein ball $\mathcal{B}_\varepsilon(\hat{\mathbb{P}}_\varepsilon)$ is constructed from sample set $\{\hat{\omega}_1, \hat{\omega}_2, \dots, \hat{\omega}_N\}$. If the loss function $l(\mathbf{x}, \tilde{\omega})$ is upper semi-continuous, the worst-case expectation is as:

Algorithm 1 Nested bisection search

```

1: Initialize  $\underline{\sigma} = 0, \bar{\sigma} = \sigma_{\max}$ ;
2: while  $(\bar{\sigma} - \underline{\sigma} > 10^{-4})$  do
3:  $\sigma = (\bar{\sigma} + \underline{\sigma})/2$ ;
4:  $\gamma = \text{bisearch}(h(\sigma, \cdot), 0, 100)$ ;
5: if  $\gamma > \rho$  then
6:  $\underline{\sigma} = \sigma$ ;
7: else
8:  $\bar{\sigma} = \sigma$ ;
9: end if
10: end while
11: Output  $\sigma = \bar{\sigma}$ .

```

$$\begin{cases} \sup_{\mathbb{P} \in \hat{\mathbb{P}}_N} E_{\mathbb{P}} \{l(\mathbf{x}, \tilde{\omega})\} = \inf_{\lambda \geq 0, s \in \mathbf{R}^N} \lambda \varepsilon + \frac{1}{N} \sum_{k=1}^N s_k \\ \text{s.t. } l(\mathbf{x}, \underline{\omega}) + \lambda(\underline{\omega} - \hat{\omega}_k) \leq s_k \quad \forall k \leq N \\ l(\mathbf{x}, \bar{\omega}) - \lambda(\bar{\omega} - \hat{\omega}_k) \leq s_k \quad \forall k \leq N \\ l(\mathbf{x}, \hat{\omega}_k) \leq s_k \quad \forall k \leq N \end{cases} \quad (\text{C1})$$

REFERENCES

- [1] Q. Wu and J. Ma, "Power system optimal reactive dispatch using evolutionary programming," *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1243-1249, Aug. 1995.
- [2] S. Yang, W. Wang, C. Liu *et al.*, "Optimal reactive power dispatch of wind power plant cluster considering static voltage stability for low-carbon power system," *Journal of Modern Power Systems and Clean Energy*, vol. 3, no. 1, pp. 114-122, Aug. 2015.
- [3] Y. Li, M. Li, Q. Wu *et al.*, "Optimal reactive power dispatch with wind power integrated using group search optimizer with intraspecific competition and lévy walk," *Journal of Modern Power Systems and Clean Energy*, vol. 2, no. 4, pp. 308-318, Jun. 2014.
- [4] B. Zhao, C. Guo, Y. Cao *et al.*, "A multiagent-based particle swarm optimization approach for optimal reactive power dispatch," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1070-1078, May 2005.
- [5] W. Jwo, C. Liu, C.-C. Liu *et al.*, "Hybrid expert system and simulated annealing approach to optimal reactive power planning," *IET Generation, Transmission & Distribution*, vol. 142, no. 4, pp. 381-385, Jul. 1995.
- [6] N. Deeb and S. M. Shahidehpour, "Linear reactive power optimization in a large power network using the decomposition approach," *IEEE Transactions on Power Systems*, vol. 5, no. 2, pp. 428-438, May 1990.
- [7] K. Y. Lee, Y. M. Park, J. L. Ortiz *et al.*, "A united approach to optimal real and reactive power dispatch," *IEEE Transactions on Power Apparatus & Systems*, vol. PAS-104, no. 5, pp. 1147-1153, May 1985.
- [8] S. Granville, "Optimal reactive dispatch through interior point methods," *IEEE Transactions on Power Systems*, vol. 9, no. 1, pp. 136-146, Feb. 1994.
- [9] T. Ding, S. Liu, W. Yuan *et al.*, "A two-stage robust reactive power optimization considering uncertain wind power integration in active distribution networks," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 301-311, Nov. 2015.
- [10] J. Liu, W. Fang, X. Zhang *et al.*, "An improved photovoltaic power forecasting model with the assistance of aerosol index data," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 2, pp. 434-442, Feb. 2015.
- [11] A. Samimi and M. Nikzad, "Complete active-reactive power resource scheduling of smart distribution system with high penetration of distributed energy resources," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 6, pp. 863-875, Sept. 2017.
- [12] J. Liu, H. Gao, Z. Ma *et al.*, "Review and prospect of active distribution system planning," *Journal of Modern Power Systems and Clean Energy*, vol. 3, no. 4, pp. 457-467, Aug. 2015.
- [13] D. Thukaram and G. Yesuratnam, "Optimal reactive power dispatch in a large power system with AC-DC and FACTS controllers," *IET Generation, Transmission & Distribution*, vol. 2, no. 1, pp. 71-81, Jan. 2008.
- [14] C. Duan, W. Fang, L. Jiang *et al.*, "FACTS device allocation via

- sparse optimization," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1308-1319, May 2016.
- [15] J. Liu, X. Hao, X. Wang *et al.*, "Application of thyristor controlled phase shifting transformer excitation impedance switching control to suppress short-circuit fault current level," *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 1, pp. 1-12, Jul. 2018.
- [16] W. Qin, P. Wang, X. Han *et al.*, "Risk analysis of power systems for both real and reactive power," *Journal of Modern Power Systems and Clean Energy*, vol. 1, no. 2, pp. 150-158, Jun. 2013.
- [17] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 206-215, Jul. 2012.
- [18] Z. Hu, X. Wang, G. Taylor *et al.*, "Stochastic optimal reactive power dispatch: formulation and solution method," *International Journal of Electrical Power & Energy Systems*, vol. 32, no. 6, pp. 615-621, Jul. 2010.
- [19] S. M. Mohseni-Bonab, A. Rabiee, and B. Mohammadi-Ivatloo, "Voltage stability constrained multi-objective optimal reactive power dispatch under load and wind power uncertainties: a stochastic approach," *Renewable Energy*, vol. 85, pp. 598-609, Jan. 2016.
- [20] R. Jabr, "Adjustable robust opf with renewable energy sources," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4742-4751, Aug. 2013.
- [21] W. Wei, F. Liu, S. Mei *et al.*, "Robust energy and reserve dispatch under variable renewable generation," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp. 369-380, Jan. 2015.
- [22] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, no. 1-2, pp. 115-166, Sept. 2018.
- [23] R. Gao and A. J. Kleywegt, "Distributionally robust stochastic optimization with Wasserstein distance," *Arxiv Preprint*, arXiv: 1604.02199, Apr. 2016.
- [24] L. Yao, X. Wang, C. Duan *et al.*, "Data-driven distributionally robust reserve and energy scheduling over Wasserstein balls," *IET Generation, Transmission & Distribution*, vol. 12, no. 1, pp. 178-189, Feb. 2018.
- [25] C. Duan, L. Jiang, W. Fang *et al.*, "Data-driven distributionally robust energy-reserve-storage dispatch," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 2826-2836, Nov. 2018.
- [26] C. Duan, L. Jiang, W. Fang *et al.*, "Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4924-4936, Feb. 2018.
- [27] C. Duan, L. Jiang, W. Fang *et al.*, "Data-driven affinely adjustable distributionally robust unit commitment," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1385-1398, Mar. 2018.
- [28] W. Wei, F. Liu, and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 289-300, Jan. 2016.
- [29] W. Wei, J. Wang, and S. Mei, "Dispatchability maximization for co-optimized energy and reserve dispatch with explicit reliability guarantee," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3276-3288, Jul. 2016.
- [30] Z. Wang, Q. Bian, H. Xin *et al.*, "A distributionally robust coordinated reserve scheduling model considering CVaR-based wind power reserve requirements," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 2, pp. 625-636, Apr. 2016.
- [31] Q. Zhai, X. Li, X. Lei *et al.*, "Transmission constrained UC with wind power: an all-scenario-feasible MILP formulation with strong nonanticipativity," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1805-1817, May 2017.
- [32] M. Farivar and S. H. Low, "Branch flow model: relaxations and convexification-part I," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554-2564, Aug. 2013.
- [33] M. Farivar and S. H. Low, "Branch flow model: relaxations and convexification-part II," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2565-2572, Aug. 2013.
- [34] S. H. Low, "Convex relaxation of optimal power flow - part II: exactness," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177-189, Jun. 2014.
- [35] S. Huang, Q. Wu, J. Wang *et al.*, "A sufficient condition on convex relaxation of AC optimal power flow in distribution networks," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1359-1368, Mar. 2017.
- [36] Z. Yang, H. Zhong, A. Bose *et al.*, "A linearized OPF model with reactive power and voltage magnitude: a pathway to improve the MW-only DC OPF," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1734-1745, Mar. 2018.
- [37] Z. Yang, H. Zhong, A. Bose *et al.*, "Optimal reactive power dispatch with accurately modelled discrete control devices: a successive linear approximation approach," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2435-2444, May 2017.
- [38] M. B. Cain, R. P. O'Neill, A. Castillo *et al.*, "History of optimal power flow and formulations," *Federal Energy Regulatory Commission*, vol. 1, pp. 1-36, Sept. 2012.
- [39] Y. J. Zhang and Z. Ren, "Optimal reactive power dispatch considering costs of adjusting the control devices," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1349-1356, Aug. 2005.
- [40] C. Zhao and Y. Guan, "Data-driven risk-averse stochastic optimization with Wasserstein metric," *Operations Research Letters*, vol. 46, no. 2, pp. 262-267, Mar. 2018.
- [41] C. Zhao and Y. Guan, "Data-driven risk-averse two-stage stochastic program with-structure probability metrics," *Optimization Online*, vol. 2, no. 5, pp. 1-40, Apr. 2015.
- [42] F. Bolley and C. Villani, "Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities," *Annales de la Faculté des Sciences de Toulouse Math*, vol. 14, no. 3, pp. 331-352, Sept. 2005.
- [43] W. H. Kersting, "Radial distribution test feeders," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 975-985, Aug. 1991.
- [44] D. Lee and R. Baldick, "Probabilistic wind power forecasting based on the laplace distribution and golden search," in *Proceedings of IEEE/PES Transmission and Distribution Conference and Exposition*, Dallas, USA, Jul. 2016, pp. 1-5.
- [45] B. M. Hodge, D. Lew, M. Milligan *et al.*, "Wind power forecasting error distributions: an international comparison," in *Proceedings of 11th Annual International Workshop on Large-scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants Conference*, Lisbon, Portugal, Nov. 2012, pp. 1-6.

Jun Liu received his B.S. and Ph.D. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2012, respectively. He is now an associate professor at the Department of Electrical Engineering, Xi'an Jiaotong University, and he was a visiting scholar at Texas A&M University, College Station, USA, from 2008 to 2010. His research interests include renewable energy integration, power system operation and control, power system stability, high-voltage direct current, flexible alternating current transmission systems and smart grids.

Yefu Chen received the B.S. and M.S. degrees from the School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 2016 and 2019, respectively. He is currently working at Electric Power Dispatch and Control Center of Guangdong Power Grid Corporation, Guangzhou, China. His research interests include power system optimization, operation and control.

Chao Duan received the B.S. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2012. He is currently working toward the Ph.D. degree at Xi'an Jiaotong University, and the University of Liverpool, Liverpool, U.K. His research interests include stochastic optimization, stability analysis, and robust control of power systems.

Lin Jiang received the B.Sc. and M.Sc. degrees from Huazhong University of Science and Technology (HUST), Wuhan, China, in 1992 and 1996, respectively, and the Ph.D. degree from the University of Liverpool, Liverpool, U.K., in 2001, all in electrical engineering. He was a Postdoctoral Research Assistant in the University of Liverpool from 2001 to 2003, and a Postdoctoral Research Associate in the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, U.K., from 2003 to 2005. He was a Senior Lecturer with the University of Glamorgan, Wales, U.K., from 2005 to 2007 and moved to the University of Liverpool in 2007. He is currently a Reader in the University of Liverpool. His current research interests include control and analysis of power system, smart grid, and renewable energy.

Jia Lyu received the B.S. degree from the School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 2017, where he is currently working toward the M.S. degree. His research interests include power system operation and control.