Risk Based Distributionally Robust Real-Time Dispatch Considering Voltage Security

Li, Peng; Wang, Chengfu; Wu, Qiuwei; Yang, Ming

Published in:
IEEE Transactions on Sustainable Energy

Link to article, DOI:
10.1109/TSTE.2020.2964949

Publication date:
2020

Document Version
Peer reviewed version

Risk Based Distributionally Robust Real-Time Dispatch Considering Voltage Security

Peng Li, *Student Member, IEEE,* Chengfu Wang, *Member, IEEE,* Qiuwei Wu, *Senior Member, IEEE,* and Ming Yang, *Senior Member, IEEE*

***Abstract*—A risk-based distributionally robust real-time dispatch approach is proposed to strike a balance between the operational costs and risk as well as considering the nodal voltage security even when the distribution of uncertainties cannot be precisely estimated. To model uncertainties, a data-drive ambiguity set is developed based on the imprecise probability theory without any presumption on the probability distribution of uncertainties. The obtained ambiguity set is based on the confidence interval, which can be constructed according to actual needs by choosing the point-wise or family-wise confidence interval. Two strategies for wind farms to provide voltage support are incorporated in the proposed approach so that the nodal voltage security can be ensured. By incorporating the risk tractable estimation method and sequential convex optimization method, an efficient algorithm is developed to release the computational burden. Numerical results show that compared with the regular robust optimization, the proposed approach reduces the total operational cost to achieve statistically optimal dispatch decisions. The proposed approach also outperformances the stochastic programming in term of the operational risk reliability. Meanwhile, the nodal voltage security issues can be mitigated with the voltage support from wind farms. In conclusion, under the uncertainty of probability distributions, the proposed approach can efficiently strike a balance between the operational cost and risk while ensuring the nodal voltage security.**

***Index Terms*—Confidence interval, distributionally robust optimization, real-time dispatch, risk, voltage security.**

NOMENCLATURE

A. Sets and Indices:

$g \in \mathcal{G}$	Conventional thermal units.
$FSWT$	Fixed speed wind turbine.
$VSWT$	Variable speed wind turbine.
$w \in \mathcal{W}$	All wind farms, i.e., $\mathcal{W}1 \cup \mathcal{W}2 = \mathcal{W}$.
$w1 \in \mathcal{W}1$	FSWT-based wind farms.
$w2 \in \mathcal{W}2$	VSWT-based wind farms.
$l_e \in \mathcal{L}_e$	Power transmission lines.
$\mathcal{R}, \mathcal{S}, \mathcal{L}$	Reference bus, PV bus, and PQ bus sets.
$\mathcal{A}$	Ambiguity set.
i, j	Indices for power system buses.

B. Vectors:

$\boldsymbol{P}, \boldsymbol{Q}$	Active and reactive power vectors.
$\boldsymbol{V}, \boldsymbol{\theta}$	Voltage and phase angle vectors.
$\underline{\boldsymbol{P}}, \overline{\boldsymbol{P}}$	Active power capacity limit vectors.
$\underline{\boldsymbol{Q}}, \overline{\boldsymbol{Q}}$	Reactive power capacity limit vectors.
$\underline{\boldsymbol{V}}, \overline{\boldsymbol{V}}$	Voltage magnitude limit vectors.
$\boldsymbol{\Delta\tilde{P}}, \boldsymbol{\Delta\tilde{Q}}, \boldsymbol{\Delta\tilde{V}}$	Random disturbance vectors of active power, reactive power, and voltage magnitudes.

C. Parameters:

G_{ij}, B_{ij}	$G_{ij} + jB_{ij}$ is the admittance of line ij.
$B_{ij}^{'}$	The susceptance without shunt elements.
c_g	Prices of generators to provide energy.
c_g^{up}, c_g^{dn}	Prices of generators to provide upward and downward reserve.
ρ^l, ρ^u	Prices of wind power curtailment and power shortage.
$\overline{w}$	Installed capacity of wind farms.
F_μ	CDF of the probability distribution μ.
$(x)^+$	$\max\{x, 0\}$.
$\mathbb{E}_\mu$	Expectation with probability distribution μ.
p_w, p_l	Expected output of Wind Farm and Load.
p_i, q_i	Nodal active and reactive power.
Z^r	Operational risk limitation.

D. Decision Variables:

α_g	Participation factors of AGC units.
β_{w2}	Wind farm reactive power linear decision factor.
p_g	Base point of generators.
$\Delta p_g^{up}, \Delta p_g^{dn}$	Upward and downward reserve of generators.
p_w^u, p_w^l	Lower and upper output bound of wind farms.
p_{w1}^u, p_{w1}^l	Output bounds of FSWT-based wind farms.
p_{w2}^u, p_{w2}^l	Output bounds of VSWT-based wind farms.
$\Delta p_{w1}, \Delta p_{w2}$	Deviation of FSWT-based and VSWT-based wind farms, satisfying $\Delta p_w = \Delta p_{w1} + \Delta p_{w2}$.
v_j, θ_j	Nodal voltage magnitude and phase angle.

E. Random Variables:

$\tilde{p}_w$	Actual output of wind farms.
$\tilde{p}_{w1}$	Actual output of FSWT-based wind farms.
$\tilde{p}_{w2}$	Actual output of VSWT-based wind farms.

This work was supported by the National Key Research and Development Program of China under Grant 2018YFB0904200 and the Science and Technology Project of State Grid Corporation of China under Grant 52060018000X.

P. Li, C. Wang, and M. Yang (Corresponding Author) are with Key Laboratory of Power System Intelligent Dispatch and Control, Shandong University, Jinan 250061, China (e-mail: pengli121121@yahoo.com; wangcf@sdu.edu.cn; myang@sdu.edu.cn).

Q. Wu is with the Center for Electric Power and Energy (CEE), Department of Electrical Engineering, Technical University of Denmark (DTU), Kgs. Lyngby 2800, Denmark, and also with the School of Electrical Engineering, Shandong University, Jinan 260061, China (e-mail: qw@elektro.dtu.dk).

I. INTRODUCTION

REAL-TIME power dispatch is one of the crucial parts in the operation dispatch of power systems, which targets at balancing the fluctuating supply and demand resulted from load and renewable energy (RE) at a time-scale of minutes [1], [2]. With large-scale RE penetration, great benefits have

been brought to environment, yet the existence of challenges is undeniable [3]. Real-time power dispatch methods should be improved to account for the challenges brought by RE, such as uncertainties due to the intermittent nature of RE [4].

Robust optimization (RO) and stochastic programming (SP) are two widely used methods to account for uncertainties in power dispatch. In SP, uncertainties are generally assumed to follow known probability distributions (PDs) and then the deterministic problems can be obtained based on the assumed PDs [5], [6]. However, the assumed PD may not be the true one, because in practice the true PD is hard to be exactly estimated due to insufficient historical sample data. Therefore, SP may obtain incorrect decisions. In contrast, RO uses an uncertainty set to model the uncertainty, which does not require the PD information, and thus the modeling can be simplified [7]. RO can guarantee the operational feasibility for every possible scenario in the prescribed uncertainty set, leading to more reliable decisions than SP [8], [9]. Meanwhile, in general, RO is more efficient than SP, and is especially suitable for the time-sensitive applications [9]. RO, however, may increase conservativeness of decisions in return. To address the problem, the budget of uncertainty technique is proposed to compromise the robustness and conservativeness by adding different kinds of constraints to uncertainty sets [10]–[13]. For example, the total allowed number of wind power output cases can be varied [12]; the linear coefficients of the polyhedral sets are varied [13]. However, in the aforementioned RO-based approaches, the PD information of uncertainties is not well regarded in the optimization, which may cause sub-optimal decisions. To overcome the shortage, some literatures, e.g., [14], [15], try to incorporate PD information into RO. However, they all share the similar drawback with SP.

To overcome the disadvantages of SP and RO, distributionally robust optimization (DRO) is developed [16]–[18]. It utilizes the PD information without assumption on the exact PD, and thus statistically optimal solutions can be obtained. DRO also has the data-exploiting ability, leading to less conservative solutions. Recently, DRO has been successfully applied to power system operation problems [19]–[26]. In [19]–[21], a moment based-DRO model is established for the reserve schedule or optimal power flow (OPF) problem, where the moment of uncertainties is used to construct the ambiguity set. Therefore, only the moment information is utilized in the above moment-based DRO, whereas much more information can be extracted from historical data. Meanwhile, the moment-based models are usually cast into semidefinite programmings, which are very computationally intensive. To address the shortcoming, distance-based DRO methods are proposed, such as the KL-based DRO [22] and Wasserstein-based DRO (W-DRO) [23]–[25]. However, KL-based ambiguity sets cannot be applied to stochastic optimization models with heavy tail random functions because the worst-case expectation will be infinite. And the out-of sample performances cannot be guaranteed. For the W-DRO, [23] tackles the reserve and energy joint dispatch problem considering wind power uncertainties based on the W-DRO. A W-DRO model for OPF problems is established in [24], which considers dynamic line rating and line overload risk. [25] studies a W-DRO-based optimal gas-power flow problem with uncertain wind generation. Though the above W-DRO-based studies make a great contribution to the power system dispatch problems under uncertainties, on one hand, to obtain trackable models, relaxation transformation is usually used, which may lead to a infeasible solution. And the computational burden grows heavily with the number of employed data [27]. On the other hand, they all apply the DC power flow, ignoring the voltage security. In that case, the obtained operation strategy may result in voltage issues at some critical buses. Based on [28], [26] provides a possible way to consider the voltage security in the DRO-based OPF. However, in [26], only conventional units is utilized to ensure voltage security, ignoring the reactive power support from wind farms, which may lead to reactive power reserve shortage. On the other hand, the effect of voltage security constraints on the admissible range of wind power (ARWP) is also ignored, which influences not only the power dispatch strategy but also the solvability of the problem, especially when there is no sufficient reactive power reserve in the system. When high-level wind power is integrated into a grid and there is no sufficient reactive power reserve to ensure the voltage security when accommodating the wind power entirely, the problem established in [26] will be infeasible.

In this paper, a novel risk-based DR real-time dispatch (R-DRTD) model is proposed for transmission networks. It is distinguished by the strategies to ensure voltage security and the ambiguity set. We also direct attention to decreasing the computational burden for application to large-scale systems. The main advantages of the proposed approach are as follows:

1) According to the imprecise probability theory, a novel ambiguity set is developed to model the uncertainties of wind power PDs. On the one hand, based on the novel ambiguity set, the R-DRTD model is proposed to strike a balance between the operational cost and risk, considering the uncertainties of wind power PDs. On the other hand, the constructed ambiguity set is data-driven, which makes no reference to the exact PD of uncertainties, and has the data-exploiting ability. Besides, using the proposed ambiguity set, the computational efficiency is irrelevant to the amount of data, leading to a highly scalable approach.
2) The nodal voltage security is considered in DRO-based transmission network real-time dispatch problems. To ensure the voltage security, on the one hand, two reactive power compensation strategies for VSWT-based wind farms are incorporated into the proposed DRO-based transmission network dispatch model. On the other hand, the admissible region of wind power can also be optimized to further ensure the voltage security. To the best of our knowledge, we are the first to consider the above two issues in the DRO-based transmission network real-time dispatch problems, which can improve the operational economy and security of transmission networks under uncertainties.
3) An efficient algorithm is also developed to transform the proposed R-DRTD model into a tractable mixed integer linear programming (MILP) problem. By using the proposed algorithm, not only a global optimal approximate solution can be obtained, but also the computational burden can be

significantly reduced, which makes the proposed approach suitable for large-scale power system applications.

II. Mathematical Formulation

A. Deterministic Real-time Dispatch

The mathematical formulation of the deterministic transmission network real-time dispatch problem is given as below:

$$Z' = \min \sum_{g\in\mathcal{G}} c_g p_g \tag{1}$$

$$s.t. \qquad \sum_{g\in\mathcal{G}} p_g + \sum_{w\in\mathcal{W}} p_w = \sum_{l\in\mathcal{L}} p_l, \tag{2}$$

$$\underline{\boldsymbol{P}}_{\mathcal{G}} \le \boldsymbol{P}_{\mathcal{G}} \le \overline{\boldsymbol{P}}_{\mathcal{G}}, \tag{3}$$

$$\underline{\boldsymbol{P}}_{\mathcal{L}_e} \le \boldsymbol{P}_{\mathcal{L}_e} \le \overline{\boldsymbol{P}}_{\mathcal{L}_e}. \tag{4}$$

The objective function (1) is the total generation cost. (2) ensures the active power balance at the base case where wind power exactly matches the forecasted value. (3) denotes the generator capacity limits. (4) denotes line MW flow limits.

B. Active and Reactive Power Compensation Strategy

In the real-time dispatch, the automatic generation control (AGC) is usually applied to maintain the real-time active power balance, where the total imbalance is distributed to AGC units according to the following affine function [29]:

$$\tilde{p}_g = p_g + \Delta\tilde{p}_g = p_g - \alpha_i \boldsymbol{e}^{\mathrm{T}} \Delta\tilde{\boldsymbol{p}}_{\mathcal{W}} \tag{5}$$

where $\Delta\tilde{\boldsymbol{p}}_{\mathcal{W}}$ is active power imbalance vector. When p_g and α_g are determined, this affine mechanism can be used to estimate the system status with respect to any realization of the wind power. Therefore, in this paper, the AGC of conventional units is applied as the active power compensation strategy.

Active power fluctuations are usually accompanied by reactive power fluctuations. For example, in FSWTs with advantages of low costs and simple structures, any fluctuation in active power generation results in similar behavior in reactive power absorption [30]. The reactive consumption of the turbines is generally offset using capacitor banks to achieve a constant power factor [31]. The relationship between the reactive consumption and active power output of FSWT-based wind farms can be expressed as

$$\cos\theta_{w1} = \tilde{p}_{w1} / \sqrt{(\tilde{p}_{w1})^2 + (\tilde{q}_{w1})^2} \tag{6}$$

where $\cos\theta_{w1}$ denotes the constant power factor, which is a given parameter in this paper; $\tilde{q}_{w1}$ and $\tilde{p}_{w1}$ are the reactive power consumption and active power output of FSWT-based wind farms, respectively. Both of active and reactive power fluctuations result in voltage fluctuations. To ensure voltage security, the automatic voltage control (AVC) strategy and the linear decision rule [32] are applied in this paper. The AVC is to keep nodal voltage magnitudes at the set values by reactive power compensation. And in the linear decision rule, the reactive power compensation is a linear function of the total reactive power consumption, as shown below

$$\tilde{q}_{com} = -\beta_{com} \boldsymbol{e}^{\mathrm{T}} \tilde{\boldsymbol{q}}_{\mathcal{C}} \tag{7}$$

where $\tilde{q}_{com}$ and $\tilde{\boldsymbol{q}}_{\mathcal{C}}$ are the reactive power support and the reactive power consumption vector, respectively; β_{com} is the reactive power compensation linear factor; $\boldsymbol{e}$ is all one vector.

Note, in the real-time power dispatch, except for thermal units, RE generators, e.g., VSWTs, can also provide voltage support. VSWTs achieve fast and flexible decoupling control of active and reactive power, suitable for applications in the real-time power dispatch. Therefore, in this paper, both thermal units and VSWT-based wind farms are assumed to provide voltage support. Specifically, conventional units apply the AVC and VSWT-based wind farms apply the AVC or linear decision rule as the reactive compensation strategy. When the linear decision rule is applied, the reactive power output $\tilde{q}_{w2}$ of VSWT-based wind farms is computed according to the following linear function

$$\tilde{q}_{w2} = -\beta_{w2} \boldsymbol{e}^{\mathrm{T}} \tilde{\boldsymbol{q}}_{\mathcal{W}1} \tag{8}$$

where $\tilde{\boldsymbol{q}}_{\mathcal{W}1}$ is the reactive power consumption vector.

C. Linear Power Flow Model

Reference [28] proposes a decoupled linearized power flow model with respect to voltage magnitude and phase angle. The linear power flow model is as follows:

$$\begin{cases} P_i = \sum\limits_{j=1}^{n} G_{ij} V_j - \sum\limits_{j=1}^{n} B'_{ij} \theta_j, \\ Q_i = -\sum\limits_{j=1}^{n} B_{ij} V_j - \sum\limits_{j=1}^{n} G_{ij} \theta_j, \end{cases} \tag{9}$$

$$P_{ij}^{l_e} = g_{ij}(V_i - V_j) - b_{ij}(\theta_i - \theta_j). \tag{10}$$

With the aid of the known information of reference bus, PV buses, and PQ buses, the following incremental matrix can be obtained from Eq. (9). More details can be found in [28].

$$\begin{bmatrix} \Delta\boldsymbol{\theta}_{\mathcal{S}\cup\mathcal{L}} \\ \Delta\boldsymbol{V}_{\mathcal{L}} \\ \Delta\boldsymbol{Q}_{\mathcal{R}\cup\mathcal{S}} \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} \Delta\boldsymbol{P}_{\mathcal{S}} \\ \Delta\boldsymbol{P}_{\mathcal{L}} \\ \Delta\boldsymbol{Q}_{\mathcal{L}} \end{bmatrix} + \boldsymbol{B} \begin{bmatrix} \Delta\boldsymbol{\theta}_{\mathcal{R}} \\ \Delta\boldsymbol{V}_{\mathcal{R}} \\ \Delta\boldsymbol{V}_{\mathcal{S}} \end{bmatrix} \tag{11}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are constant coefficient matrices. Note, for reference bus, the phase angle is kept constant, i.e., $\Delta\boldsymbol{\theta}_{\mathcal{R}} = 0$. When implementing the AGC, AVC, and linear decision rule in the previous section, $\Delta\boldsymbol{V}_{\mathcal{R}} = \Delta\boldsymbol{V}_{\mathcal{S}} = 0$, the active power of conventional units changes affinely to the total wind power active power uncertainty, and the reactive power outputs of VSWT-based wind farms change according to the AVC or linear decision rule. Thus, Eq. (11) can be reformulated as

$$\begin{bmatrix} \Delta\boldsymbol{\theta}_{\mathcal{S}\cup\mathcal{L}} \\ \Delta\boldsymbol{V}_{\mathcal{L}} \\ \Delta\boldsymbol{Q}_{\mathcal{R}\cup\mathcal{S}} \end{bmatrix} = \boldsymbol{C}_1 \begin{bmatrix} -\boldsymbol{e}^{\mathrm{T}}(\Delta\boldsymbol{P}_{\mathcal{W}1} + \Delta\boldsymbol{P}_{\mathcal{W}2})\boldsymbol{\alpha}_{\mathcal{S}} \\ \Delta\boldsymbol{P}_{\mathcal{W}2} \\ \Delta\boldsymbol{P}_{\mathcal{W}1} \\ (\sin\theta/\cos\theta)\Delta\boldsymbol{P}_{\mathcal{W}1} \end{bmatrix} \tag{12a}$$

$$\begin{bmatrix} \Delta\boldsymbol{\theta}_{\mathcal{S}\cup\mathcal{L}} \\ \Delta\boldsymbol{V}_{\mathcal{L}} \\ \Delta\boldsymbol{Q}_{\mathcal{R}\cup\mathcal{S}} \end{bmatrix} = \boldsymbol{C}_2 \begin{bmatrix} -\boldsymbol{e}^{\mathrm{T}}(\Delta\boldsymbol{P}_{\mathcal{W}1} + \Delta\boldsymbol{P}_{\mathcal{W}2})\boldsymbol{\alpha}_{\mathcal{S}} \\ \Delta\boldsymbol{P}_{\mathcal{W}1} + \Delta\boldsymbol{P}_{\mathcal{W}2} \\ (\sin\theta/\cos\theta)\Delta\boldsymbol{P}_{\mathcal{W}1} \\ -\boldsymbol{e}^{\mathrm{T}}\Delta\boldsymbol{P}_{\mathcal{W}1}\boldsymbol{\beta}_{\mathcal{W}2} \end{bmatrix} \tag{12b}$$

where (12a) and (12b) are obtained by applying the AVC and linear decision rule, respectively; $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are constant matrices obtained from Eq. (6)-(9) and (11); $(\Delta\boldsymbol{P}_{\mathcal{W}1}, \Delta\boldsymbol{P}_{\mathcal{W}2})$

are wind power uncertainty vectors; $(\boldsymbol{\alpha}_{\mathcal{S}}, \boldsymbol{\beta}_{\mathcal{W}2})$ are decision variable vectors. According to Eq. (10), (12a) and (12b), the changes in line MW flow can also be represented as follows:

$$\begin{aligned}\Delta \boldsymbol{P}_{\mathcal{L}_e} = &(\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{D}_1^{\mathrm{T}}\boldsymbol{\alpha}_{\mathcal{S}}) + \boldsymbol{e}^{\mathrm{T}}(\boldsymbol{D}_2^{\mathrm{T}}\boldsymbol{\beta}_{\mathcal{W}2}) + \boldsymbol{D}_3^{\mathrm{T}})\Delta \boldsymbol{P}_{\mathcal{W}1} \\ &+ (\boldsymbol{e}^{\mathrm{T}}(\boldsymbol{D}_4^{\mathrm{T}}\boldsymbol{\alpha}_{\mathcal{S}}) + \boldsymbol{D}_5^{\mathrm{T}})\Delta \boldsymbol{P}_{\mathcal{W}2}\end{aligned} \tag{13}$$

where $\boldsymbol{D}_i, i = 1,2,3,4,5$ are all constant matrices that can be obtained from Eq. (10), (12a) and (12b).

D. Risk-Based Distributionally Robust Optimization Model

In practice, there are uncertainties resulting from RE and thus reserves should be committed. Because of limited system reserve capacity, RE may not be entirely accommodated in extreme conditions, resulting in according operational risk. In this paper, the expected operational risk is penalized in the objective function [33]. The proposed R-DRTD aims to strike a balance between the operational costs and risk, considering the PD uncertainty and the nodal voltage security:

$$Z = \min \sum_{g\in\mathcal{G}} (c_g p_g + c_g^{up}\Delta p_g^{up} + c_g^{dn}\Delta p_g^{dn}) + \sum_{w\in\mathcal{W}} Z_w^r, \tag{14}$$

$$s.t. \qquad p_i = \sum\nolimits_{j=1}^{N} G_{ij}v_j - \sum\nolimits_{j=1}^{N} B_{ij}^{'}\theta_j, \forall i, \tag{15}$$

$$q_i = -\sum\nolimits_{j=1}^{N} B_{ij}v_j - \sum\nolimits_{j=1}^{N} G_{ij}\theta_j, \forall i, \tag{16}$$

$$Z^r \le Risk_{dh}, \tag{17}$$

$$\underline{\boldsymbol{V}} \le \boldsymbol{V}_{\mathcal{S}\cup\mathcal{L}} \le \overline{\boldsymbol{V}}, \tag{18}$$

$$\Delta \boldsymbol{P}_{\mathcal{G}}^{up} \ge \boldsymbol{e}^{\mathrm{T}}(\boldsymbol{P}_{\mathcal{W}} - \boldsymbol{P}_{\mathcal{W}}^{l})\boldsymbol{\alpha}_{\mathcal{G}}, \tag{19}$$

$$\Delta \boldsymbol{P}_{\mathcal{G}}^{dn} \ge \mathbf{e}^{\mathrm{T}}(\boldsymbol{P}_{\mathcal{W}}^{u} - \boldsymbol{P}_{\mathcal{W}})\boldsymbol{\alpha}_{\mathcal{G}}, \tag{20}$$

$$\boldsymbol{e}^{\mathrm{T}}\boldsymbol{\alpha}_{\mathcal{G}} = 1, \tag{21}$$

$$\underline{\boldsymbol{P}}_{\mathcal{G}} + \Delta \boldsymbol{P}_{\mathcal{G}}^{dn} \le \boldsymbol{P}_{\mathcal{G}} \le \overline{\boldsymbol{P}}_{\mathcal{G}} - \Delta \boldsymbol{P}_{\mathcal{G}}^{up}, \tag{22}$$

$$\underline{\boldsymbol{V}} \le \boldsymbol{V}_{\mathcal{L}} + \Delta\tilde{\boldsymbol{V}}_{\mathcal{L}} \le \overline{\boldsymbol{V}}, \tag{23}$$

$$\underline{\boldsymbol{Q}}_{\mathcal{R}\cup\mathcal{S}} \le \boldsymbol{Q}_{\mathcal{R}\cup\mathcal{S}} + \Delta\tilde{\boldsymbol{Q}}_{\mathcal{R}\cup\mathcal{S}} \le \overline{\boldsymbol{Q}}_{\mathcal{R}\cup\mathcal{S}}, \tag{24}$$

$$\underline{\boldsymbol{P}}_{\mathcal{L}_e} \le \boldsymbol{P}_{\mathcal{L}_e} + \Delta\tilde{\boldsymbol{P}}_{\mathcal{L}_e} \le \overline{\boldsymbol{P}}_{\mathcal{L}_e}. \tag{25}$$

In the objective function (14), the first term is the operational cost, including the generation cost and reserve capacity cost, and the second term Z^r is the worst-case expectation operational risk cost. (15) and (16) ensure the active and reactive power balance at the base case where wind power exactly matches the expected value. (17) is the operational risk limit. (18) and (23) are voltage limits for all buses under the base case and disturbance cases, respectively. (19) and (20) denote the reserve capacity constraints. (21) ensures that wind power uncertainties in the ambiguity set are fully mitigated. (22) denotes the generator capacity limits. (24) denotes reactive capacity limits of wind farms and conventional units. (25) denotes line MW flow limits under disturbance cases.

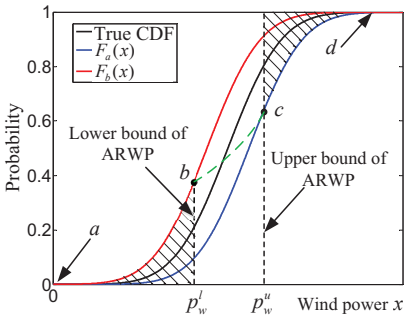

Fig. 1. Diagram of operational risk under the worst-case distribution.

The objective function (14) and constraints (12a)-(13), (15)-(25) form the proposed R-DRTD model, which is a nonlinear optimization problem with uncertain variables.

Remark 1: In this paper, we take the condition that the linear decision rule is applied to control reactive outputs of VSWT-based wind farms, for example to formulate the proposed R-DRTD model and develop the corresponding algorithm.

E. Imprecise Dirichlet Model

Imprecise probability theory that bloomed in the 1990s, is a generalization of the classical probability theory allowing partial probability specifications [34]. Typically, it quantifies the uncertainties of a random event by a probability interval. The more the employed sample data, the narrower the probability interval. To estimate the interval, the imprecise Dirichlet model (IDM) that is an extension of the deterministic Dirichlet model, is proposed [35]. Consider a random variable with K types of possible outputs, whose occurrence probabilities are $P_k, k = 1,2,...,K$. To estimate P_k, a set of density functions are used to model prior ignorance, which can be described as

$$\begin{aligned}f(P_k) = &\frac{\Gamma(s')}{\prod_{k=1}^{K}\Gamma(s'\cdot r_k)}\prod\nolimits_{k=1}^{K} P_k^{s'\cdot r_k - 1}, \\ &\forall r_k \in [0,1], \sum\nolimits_{k=1}^{K} r_k = 1,\end{aligned} \tag{26}$$

where r_k is the kth weight factor; Γ is the Gamma function; s' is the equivalent sample size that determines how quickly upper and lower probabilities converge as statistical data accumulate and $s \ge 1$; $s'\cdot r_k$ is the positive parameters of the Dirichlet distribution. Using the above set, the prejudice of the prior can be avoided, because all of the possible density functions given are included in the above set.

Then according to Bayesian rules, the corresponding estimation density function set can be obtained as

$$\begin{aligned}f(P_k) = &\frac{\Gamma(s'+n)}{\prod_{k=1}^{K}\Gamma(s'\cdot r_k + n_k)}\prod\nolimits_{k=1}^{K} P_k^{s'\cdot r_k + n_k - 1}, \\ &\forall r_k \in [0,1], \sum\nolimits_{k=1}^{K} r_k = 1,\end{aligned} \tag{27}$$

where n_k and n are the numbers of kth output samples and total samples, respectively.

Thus, the imprecise probability interval of P_k can be obtained from the estimation density function set as

$$\tilde{P}_k = [\underline{P}_k, \overline{P}_k] = [n_k/(n+s'), (n_k+s')/(n+s')], \forall k, \tag{28}$$

where $\underline{P}_k$ and $\overline{P}_k$ are calculated with respect to bounds of r_k. The above interval intends to include all the possible probabilities corresponding to different density functions to model the prior ignorance.

F. Construction of the Ambiguity Set

The CIAS method is provided here, which can directly establish the confidence bands (CBs) for the CDF of wind power based on historical data. All PDs, whose CDFs are within the CBs, construct the ambiguity set.

As is well known, the CDF of a real-valued random variable x can be defined as $F_x(X) = P(x \leq X)$, indicating the probability of $x \leq X$. When we have infinite samples of x, based on the Law of Large Numbers, a precise cumulative probability can be obtained for all points in the support of x and then F_x is obtained. However, in practice, only finite samples are available and thus the true CDF cannot be exactly estimated. To indicate the uncertainties of estimating the true CDF based on finite samples, the CBs of the CDF are explicitly estimated with respect to the available sample data.

According to the definition of CDF, two key steps are designed to estimate the CBs.

Step 1: For each value point of x, say A, the probability interval of $x \leq A$ at a specified confidence level γ is estimated. Here, the interval is referred to as a confidence interval (CI) and the CIAS applies the IDM in [36] to estimate the CI as

$$\begin{cases} a_k = 0, \quad b_k = G^{-1}(\frac{1+\gamma}{2}), & n_k = 0, \\ a_k = H^{-1}(\frac{1-\gamma}{2}), \quad b_k = G^{-1}(\frac{1+\gamma}{2}), & 0<n_k<n, \\ a_k = H^{-1}(\frac{1-\gamma}{2}), \quad b_k = 1, & n_k = n, \end{cases} \tag{29}$$

where a_k and b_k are the lower and upper bounds of the CI, respectively; H and G are the CDFs of beta distributions $B(n_k, s' + n - n_k)$ and $B(s' + n_k, n - n_k)$, respectively. Then, the CI of the cumulative probability at each point can be obtained.

Step 2: The CBs of the whole CDF are constructed according to the bounds at each point. Since only finite samples are available in practice, CIs are just calculated at the sample points, and the CBs can then be estimated using the stair-step interpolation method in [37]. Here, the CBs are expressed as:

$$\begin{cases} \underline{P}(x) = a_k, & x \in (X_k, X_{k+1}], \\ \overline{P}(x) = b_{k+1}, & x \in [X_k, X_{k+1}). \end{cases} \tag{30}$$

Finally, the ambiguity set $\mathcal{A}$ can be constructed as

$$\mathcal{A} = \{F_x | F_x(X) \in [\underline{P}(X), \overline{P}(X)]\}. \tag{31}$$

G. Point-wise and Family-wise Confidence Intervals

There are two kinds of CIs, i.e., point-wise confidence level based confidence interval (PW-CI) that is employed in CIAS, and family-wise confidence level based confidence interval (FW-CI) that can be estimated by the method in [37]. The PW-CI ensures that the true cumulative probability is within the CI at a credibility for each point in the support of x, while using the FW-CI, the true CDF is wholly within the constructed ambiguity set at a credibility. The two kinds of CIs differ in their meanings and can be chosen based on actual needs.

In fact, CIAS can also be used to estimate FW-CIs. It only needs to map the specified family-wise confidence level $\tilde{\gamma}$ to the point-wise confidence level $\bar{\gamma}$ using the mapping function $\bar{\gamma}(\tilde{\gamma}, n)$ obtained via simulations [37]. Then, the $\tilde{\gamma}$-FW-CI can be obtained by estimating the corresponding $\bar{\gamma}$-PW-CI.

The FW-CI is utilized in [38] to develop another similar ambiguity set. However, in the ambiguity set estimation method mentioned in [38], all employed samples have to be different from each other [39]. In other words, any two samples in the employed sample set are not the same, which may not be met in practice. Because, in practice, it is likely that there are two or more identical samples in the available sample set, especially when the sample size is large. Under the circumstances, the method mentioned in ref. [38] may not make full use of all available samples, leading to conservative results. In contrast, in the CIAS, all available samples can be made full use of, regardless of whether there are identical samples in the available sample set. Therefore, the CIAS is more general compared with the method utilized in [38].

III. Solution Methodology

A. Tractable Approximation of the Operational Risk

According to the proposed ambiguity set, the worst-case expectation operational risk Z_w^r in (14) can be reformulated as follows:

$$\max_{F_\mu(\tilde{p}_w)\in\mathcal{A}} \mathbb{E}_\mu \left(\rho^l (p_w^l - \tilde{p}_w)^+ + \rho^u (\tilde{p}_w - p_w^u)^+\right) \tag{32}$$

Equation (32) can be rewritten as (32a), where $F_\mu(\tilde{p}_w)$ is the CDF of $\tilde{p}_w$ and $P_\mu(\tilde{p}_w)$ is the PDF of $\tilde{p}_w$.

$$\begin{aligned} \max_{F_\mu(\tilde{p}_w)\in\mathcal{A}} \Big(&\rho^l \int_0^{p_w^l} (p_w^l - \tilde{p}_w) P_\mu(\tilde{p}_w) d\tilde{p}_w + \\ &\rho^u \int_{p_w^u}^{\bar{w}} (\tilde{p}_w - p_w^u) P_\mu(\tilde{p}_w) d\tilde{p}_w \Big) \end{aligned} \tag{32a}$$

According to the partial integration method, Eq. (32a) can be reformulated as follows:

$$\begin{aligned} \max_{F_\mu(\tilde{p}_w)\in\mathcal{A}} \Bigg\{ &\rho^l \Big((p_w^l - \tilde{p}_w) F_\mu(\tilde{p}_w) \Big|_0^{p_w^l} + \int_0^{p_w^l} F_\mu(\tilde{p}_w) d\tilde{p}_w \Big) + \\ &\rho^u \Big((\tilde{p}_w - p_w^u) F_\mu(\tilde{p}_w) \Big|_{p_w^u}^{\bar{w}} - \int_{p_w^u}^{\bar{w}} F_\mu(\tilde{p}_w) d\tilde{p}_w \Big) \Bigg\} \end{aligned} \tag{32b}$$

Then, Eq. (32b) can be rewritten as follows:

$$\begin{aligned} \max_{F_\mu(\tilde{p}_w)\in\mathcal{A}} \Bigg\{ &\rho^l \int_0^{p_w^l} F_\mu(\tilde{p}_w) d\tilde{p}_w + \\ &\rho^u \Big(\bar{w} - p_w^u - \int_{p_w^u}^{\bar{w}} F_\mu(\tilde{p}_w) d\tilde{p}_w \Big) \Bigg\} \end{aligned} \tag{33}$$

By dividing the maximum term into two terms, Eq. (34) provides an upper bound on (33), as shown below.

$$\begin{aligned} &\max_{F_{\mu 1}(\tilde{p}_w)\in\mathcal{A}} \rho^l \textstyle\int_0^{p_w^l} F_{\mu 1}(\tilde{p}_w) d\tilde{p}_w + \\ &\max_{F_{\mu 2}(\tilde{p}_w)\in\mathcal{A}} \rho^u \left(\bar{w} - p_w^u - \textstyle\int_{p_w^u}^{\bar{w}} F_{\mu 2}(\tilde{p}_w) d\tilde{p}_w \right) \end{aligned} \tag{34}$$

As shown in Fig. 1, the integral value of the first term in (34) equals to the area between $F_{\mu 1}(x)$ and the horizontal axis over $[0, p_w^l]$. Obviously, for any $F_{\mu 1}(x) \in \mathcal{A}$, if $F_{\mu 1}(x) = F_b(x)$, the area is the largest. Then, the first term is simplified to

$$Cost_{Risk}^{1,w} = \max_{F_{\mu 1} \in \mathcal{A}} \int_0^{p_w^l} F_{\mu 1}(\tilde{p}_w) d\tilde{p}_w = \int_0^{p_w^l} F_b(\tilde{p}_w) d\tilde{p}_w \quad (35)$$

Similarly, for the last term in (34), the worst-case CDF in the ambiguity set $\mathcal{A}$ is $F_a(x)$. Thus, it can be simplified to

$$Cost_{Risk}^{2,w} = \bar{w} - p_w^u - \int_{p_w^u}^{\bar{w}} F_a(\tilde{p}_w) d\tilde{p}_w. \quad (36)$$

To handle bilinear terms, the support of random wind power is discretized in Section III.C. And then (35) and (36) can be exactly approximated by the cumulative method.

We take (35) for example. Discrete the support of $\tilde{p}_{w1}$ by a series of points, i.e., $\{o_1^{w1}, o_2^{w1}, ..., o_s^{w1}, ..., o_S^{w1}, \forall w1\}$. And thus p_{w1}^l can only be at these points. Let m_{w1}^1, $m_{w1}^2, ..., m_{w1}^n, ..., m_{w1}^N$ be the ascending ordered samples of $\tilde{p}_{w1}$, where $m_{w1}^0 = 0$ and $m_{w1}^{k_s} \le o_s^{w1} \le m_{w1}^{k_s+1}$. When $p_{w1}^l = o_s^{w1}$, (35) can be exactly approximated as follows:

$$\sum_{n=1}^{k_s} \overline{P}(m_{w1}^n)(m_{w1}^n - m_{w1}^{n-1}) + \overline{P}(m_{w1}^{k_s+1})(o_s^{w1} - m_{w1}^{k_s}). \quad (37)$$

where $\overline{P}(m_{w1}^n)$ and $\overline{P}(m_{w1}^{k_s+1})$ are the upper bounds of the CDF confidence band at o_s^{w1} and $m_{w1}^{k_s+1}$, respectively.

Remark 2: Estimating (33) by (34) is conservative, because the worst pair of CDFs, i.e., F_a and F_b, instead of the worst single CDF is employed. However, when $F_a(p_w^u) > F_b(p_w^l)$, the risk estimated using the worst pair of CDFs is exactly equal to that estimated using the worst CDF, since in this regard the worst CDF can always be constructed by connecting the points $F_a(p_w^u)$ and $F_b(p_w^l)$ with a non-decreasing curve, such as the CDF a-b-c-d in Fig. 1. Moreover, the more the available data, the narrower the CDF CBs, and thus the larger the probability of $F_a(p_w^u) > F_b(p_w^l)$. Therefore, the proposed approach can usually obtain an exact risk estimation because wind power usually has a considerable amount of data.

B. Determinate Reformulation of Uncertain Constraints

In the proposed R-DRTD, all constraints are determinate except constraints (23)-(25) that can be represented in form of

$$\begin{cases} \underline{\boldsymbol{M}} \le \boldsymbol{M} + \boldsymbol{J}^{\mathrm{T}} \Delta \tilde{\boldsymbol{P}}_{\mathcal{W}1} + \boldsymbol{S}^{\mathrm{T}} \Delta \tilde{\boldsymbol{P}}_{\mathcal{W}2} \le \overline{\boldsymbol{M}} \\ \boldsymbol{J}^{\mathrm{T}} = \boldsymbol{e}^{\mathrm{T}} (\boldsymbol{H}_1^{\mathrm{T}} \boldsymbol{\alpha}_{\mathcal{S}}) + \boldsymbol{e}^{\mathrm{T}} (\boldsymbol{H}_2^{\mathrm{T}} \boldsymbol{\beta}_{\mathcal{W}2}) + \boldsymbol{H}_3^{\mathrm{T}} \\ \boldsymbol{S}^{\mathrm{T}} = \boldsymbol{e}^{\mathrm{T}} (\boldsymbol{H}_4^{\mathrm{T}} \boldsymbol{\alpha}_{\mathcal{S}}) + \boldsymbol{H}_5^{\mathrm{T}} \end{cases} \quad (38)$$

where $\boldsymbol{H}_i, i = 1, 2, 3, 4, 5$ are all constant matrices that can be obtained from Eq. (11)-(12b). Further, the first inequality in (38) can be safely replaced by the following constraints [24].

$$\begin{cases} \min\limits_{\Delta \tilde{\boldsymbol{P}}_{\mathcal{W}1}} \left(\boldsymbol{J}^{\mathrm{T}} \Delta \tilde{\boldsymbol{P}}_{\mathcal{W}1} \right) + \min\limits_{\Delta \tilde{\boldsymbol{P}}_{\mathcal{W}2}} \left(\boldsymbol{S}^{\mathrm{T}} \Delta \tilde{\boldsymbol{P}}_{\mathcal{W}2} \right) \ge \underline{\boldsymbol{M}} - \boldsymbol{M} \\ \max\limits_{\Delta \tilde{\boldsymbol{P}}_{\mathcal{W}1}} \left(\boldsymbol{J}^{\mathrm{T}} \Delta \tilde{\boldsymbol{P}}_{\mathcal{W}1} \right) + \max\limits_{\Delta \tilde{\boldsymbol{P}}_{\mathcal{W}2}} \left(\boldsymbol{S}^{\mathrm{T}} \Delta \tilde{\boldsymbol{P}}_{\mathcal{W}2} \right) \le \overline{\boldsymbol{M}} - \boldsymbol{M} \end{cases} \quad (39)$$

The first maximization term in the second equation of (39) can be equivalently expressed using:

$$\begin{aligned} \max_{\Delta \tilde{\boldsymbol{P}}_{\mathcal{W}1}} \left(\boldsymbol{J}^{\mathrm{T}} \Delta \tilde{\boldsymbol{P}}_{\mathcal{W}1} \right) &= \sum_{w1 \in \mathcal{W}1} \max_{\Delta \tilde{P}_{w1}} \left(\left(\boldsymbol{J}^{\mathrm{T}} \right)_{w1} \Delta \tilde{P}_{w1} \right) \\ &= \sum_{w1 \in \mathcal{W}1} \max \left(\left(\boldsymbol{J}^{\mathrm{T}} \right)_{w1} \Delta \underline{P}_{w1}, \left(\boldsymbol{J}^{\mathrm{T}} \right)_{w1} \Delta \overline{P}_{w1} \right) \end{aligned} \quad (40)$$

Then, the second equation of (39) can be reformulated as:

$$\begin{cases} \sum\limits_{w1 \in \mathcal{W}1} \boldsymbol{N}_{w1}^1 + \sum\limits_{w2 \in \mathcal{W}2} \boldsymbol{N}_{w2}^2 \le \overline{\boldsymbol{M}} - \boldsymbol{M} \\ \boldsymbol{N}_{w1}^1 \ge \left(\boldsymbol{J}^{\mathrm{T}} \right)_{w1} \Delta \underline{P}_{w1}, \boldsymbol{N}_{w1}^1 \ge \left(\boldsymbol{J}^{\mathrm{T}} \right)_{w1} \Delta \overline{P}_{w1}, \\ \boldsymbol{N}_{w2}^2 \ge \left(\boldsymbol{S}^{\mathrm{T}} \right)_{w2} \Delta \underline{P}_{w2}, \boldsymbol{N}_{w2}^2 \ge \left(\boldsymbol{S}^{\mathrm{T}} \right)_{w2} \Delta \overline{P}_{w2}, \end{cases} \quad (41)$$

Similarly, the first equation of (39) can also be reformulated as a determinate equation. As a result, uncertain constraint (39) is transformed into a determinate constraint.

C. The Big-M Method for Bilinear Constraints

The resulting model forms a bilinear programming problem. To solve the problem, the big-M method in [20] is applied to handle bilinear terms $\alpha_g \Delta p_w$ and $\beta_{w2} \Delta p_{w1}$. To apply the big-M method, the wind power deviation variables Δp_{w1} and Δp_{w2} are discretized. We take the bilinear term $\beta_{w2} \Delta \underline{p}_{w1}$ for example. The specific procedures are as follows:

Step 1: Discrete the support of $\tilde{p}_{w1}$ by a series of points, e.g., $\{o_1^{w1}, o_2^{w1}, ..., o_s^{w1}, ..., o_S^{w1}, \forall w1\}$. And then it is approximately supposed that $\Delta \underline{p}_{w1}$ can only be at the discretization points, i.e., $\Delta \underline{p}_{w1}$ is discretized as follows:

$$\begin{cases} \Delta \underline{p}_{w1} = \sum\limits_{s=1}^{S} p_{w1,s}^u - \overline{p}_{w1}, \\ p_{w1,s}^u = \eta_{w1,s} o_s^{w1}, \\ \sum\limits_{s=1}^{S} \eta_{w1,s} = 1. \end{cases} \quad (42)$$

where $\eta_{w1,s}$ is the 0-1 variable; $p_{w1,s}^u$ is the discretization point for p_{w1}^u; $\overline{p}_{w1}$ is the expected wind power output.

Step 2: Based on Step 1, a new bilinear term is obtained. Then, the big-M method is applied to transform the new bilinear term into mixed integer linear constraints as below.

$$\begin{cases} \beta_{w2} \eta_{w1,s} \le M_b \eta_{w1,s}, \\ \beta_{w2} \eta_{w1,s} \ge \beta_{w2} - M_b (1 - \eta_{w1,s}), \\ \sum\limits_{s=1}^{S} \eta_{w1,s} = 1. \end{cases} \quad (43)$$

where M_b is a sufficient large positive real number.

IV. Case Studies

In this section, to illustrate the performance of the proposed approach, case studies are conducted on a 6-bus test system, IEEE 118-bus system, and a real 445-bus system. All tests were carried out on a PC with an Intel Xeon E5-1620 CPU and 64 GB RAM. And all established models are solved using CPLEX of GAMS 23.8.2.

Unless otherwise specified, the confidence level in (29) is set to $\gamma = 0.95$, the installed capacity of every wind farm is set to 50 MW, and prices for the wind power curtailment and power shortage are set to \$300/(MWh) and \$3000/(MWh) [24], respectively. In practice, the prices can be chosen according to the historical data or the long-term electricity contract. Wind power samples are assumed to be generated from a Normal distribution, where the standard deviation of wind power is set to 20% of the forecasted value [40].

TABLE I
TEST RESULTS UNDER DIFFERENT CASES

	Size of ARWPs (MW)	Total cost ($)	Risk cost ($)	under- or over-voltage ?
Case 1	40.68	9110.4	29.4	Yes
Case 2	22.8	9355.7	232.1	No
Case 3	25.3	9294.1	166.7	No
Case 4	31.2	9218.1	103.6	No
Case 5	28.6	9242.5	116.1	No

TABLE II
ERRORS UNDER DIFFERENT FORECAST ERRORS

RE forecast error (MW)	-10	-5	0	5	10
Total Cost Errors	0.46%	0.51%	0.55%	0.62%	0.69%
Voltage Average Errors	0.23%	0.29%	0.18%	0.31%	0.27%

A. The Six-Bus System

The structure of the 6-bus system is shown in [14]. All conventional units in the system are assumed to be AVC units, whose parameters can be found in [14]. And G2 and G3 are assumed to be AGC units. The loads on bus 3, 5, 6 are 70MW, 150MW, 90 MW, respectively, and their power factor is 0.95. The parameters of transmission lines can be found in [29]. Two wind farms are installed at bus 4 and 5, which are equipped with VSWTs (W2) and FSWTs (W1), respectively. And their expected outputs are 30.3MW and 22.6MW, respectively.

To illustrate the necessity and validity of the proposed approach, the following five cases are studied. In each case, except case 1, voltage limits are set to $\underline{V} = 0.96$, $\overline{V} = 1.05$. The comparison results are listed in Table I.

Case 1: Reactive power and voltage constraints (19), (24), (25) are excluded from the proposed model;

Case 2: No reactive power support is obtained from W2;

Case 3: W2 provides fixed reactive power support, regardless of wind power realizations;

Case 4: W2 applies the linear decision rule to provide reactive power support;

Case 5: W2 uses AVC strategy to provide reactive support, i.e., the nodal voltage of bus 4 is maintained at a set value.

Firstly, it can be seen that if voltage constraints are excluded from the optimization, the voltage security cannot be guaranteed, indicating the necessity to consider the voltage security in the system real-time power dispatch. This is because on one hand, as active power fluctuates, reactive power usually fluctuates, which will result in nodal voltage fluctuations. On the other hand, large active power fluctuations, i.e., wind generation disturbances, also lead to nodal voltage fluctuations.

Secondly, compared with Case 2, more optimal results, i.e., lower total cost and larger ARWP, can be obtained in Cases 3-5. This is because when W2 provides reactive power support, more reactive power capacity can be utilized to eliminate the voltage fluctuations resulting from the wind power integretion. As a result, on one hand, the generation costs of conventional units can be decreased, which will improve the economic performance. On the other hand, the nodal voltage security can be ensured within larger wind power disturbance ranges, which decreases the system operational risk under uncertainties.

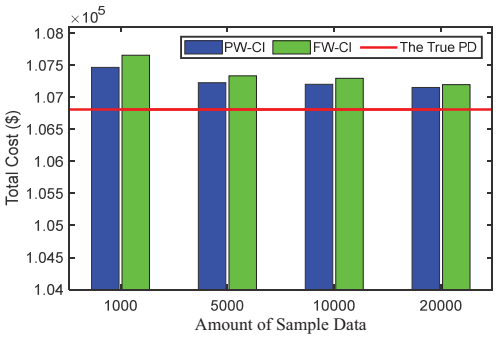

Fig. 3. Total cost obtained using different CIs.

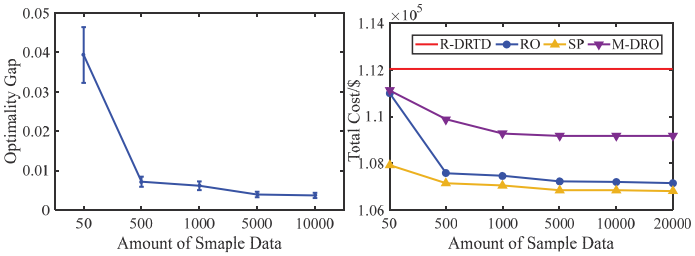

Fig. 4. Optimality gap of the proposed method (left); Evolution of total cost under different methods (right).

In addition, although the size of ARWP is similar, compared with the AVC, the linear decision rule can obtain more economic results. This is because the AVC is a local regulation method, which can only utilize local information to obtain a local optimal solution, while using the linear decision rule, global information is used to obtain the global optimal solution. On the other hand, by using the linear decision rule, the voltage magnitude at W2 can change within a certain range, which is more flexible than the AVC strategy. Of course, the AVC also has some advantages compared with the linear decision rule. For example, it does not require the global information. On the one hand, the privacy of different operators can be ensured. On the other hand, the local voltage control can be simplified, leading to a more timely voltage control. Therefore, in practice, we can chose suitable voltage control strategies according to actual needs.

Table II shows the accuracy of the employed linear power flow compared with the exact AC model. The total cost errors are computed under different power flow model. And the voltage average errors are computed by checking the voltage magnitude a posterior. It shows that the accuracy of the employed power flow model is high enough to ensure the operational economy and security in the proposed transmission network real-time dispatch problem. In fact, if more accurate results are required, we can incorporate the exact AC model into the base case to improve the accuracy.

B. Modified IEEE 118-Bus System

A modified IEEE 118-bus system is also tested to demonstrate the features of the proposed DRO approach compared

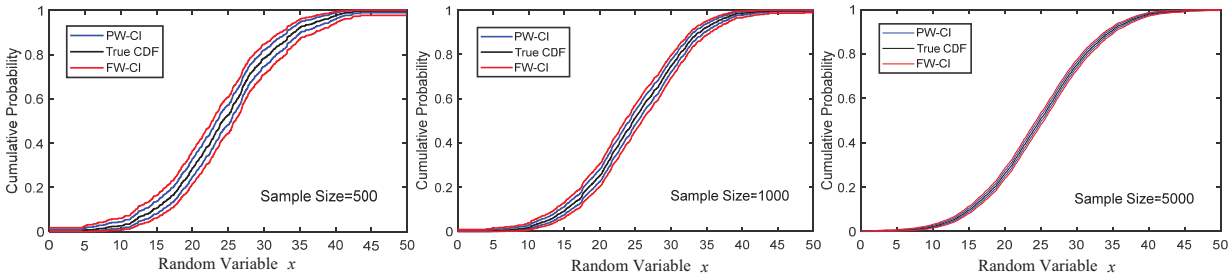

Fig. 2. Confidence bands under different sample sizes.

TABLE III
TEST RESULTS UNDER DIFFERENT APPROACHES

Method	Sample Size	Total cost ($)	CPU Time (s)
R-DRTD	1000	107462	1.64
	5000	107224	1.355
	10000	107200	1.245
	20000	107150	1.52
M-DRO	20000	109372	1.32
SP	20000	106807	2.79
RO	20000	112041	0.95

to other methods on dealing with uncertainties. The 118-bus system has 54 units and 186 lines and ten generators of 100 MW are selected as the AGC and AVC units. The system data can be found in [18]. Three wind farms are installed at buses 17 (FSWT), 66 (VSWT), and 99 (FSWT), respectively. The forecasted wind power and total load in the test system are selected from [29]. The linear decision rule is selected to control the reactive outputs of the VSWT based wind farm.

To test the performance of different CIs, the proposed R-DRTD based on PW-CIs and FW-CIs are compared and the comparison results are shown in Fig. 2 and Fig. 3. As the sample data becomes richer, both of the established ambiguity sets gradually shrink and then the total cost obtained using the established ambiguity sets gradually approaches that obtained using the true PD. Therefore, in practice, to obtain better solutions, we should incorporate as many samples as possible. Of course, if only few samples are available, although the conservativeness of the solution obtained using the proposed method will increase, the reliability is ensured. In this case, larger ambiguity sets are established and thus more possible PDs are considered to ensure the robustness. Besides, with the same data, the obtained FW-CIs are usually wider than the obtained PW-CIs, which lead to a more conservative result.

To illustrate the features of the proposed DRO approach on dealing with uncertainties, three other methods are chosen for benchmarking: 1) the RO method where the support of the random variable is used as the uncertainty set; 2) the SP method which assumes wind power uncertainties follow the Laplace distribution; 3) the moment-based DRO method (M-DRO), where the ambiguity set is established with given mean and covariance of samples. After running all the aforementioned methods under the same sample set, monte carlo simulation with another 10^6 samples generated from the true PD is employed to test the practical performance. Fig. 4 shows the optimality gap of the proposed method. And the total cost comparison results can also be found in Fig. 4 and Table III, where some valuable observations can be obtained: 1) The conservativeness of R-DRTD can be decreased by incorporating more data, indicating the data-exploiting ability of the proposed method. The reason is that when more data are available, according the Law of Large Number, the true probability at the value points can be estimated more accurately and thus the CIs will be shrank. More impossible PDs can be excluded from the ambiguity set. The proposed method based on the worst-case PD will obtain less conservative solutions. 2) The total cost of R-DRTD is upper bounded by RO and lower bounded by SP. This is because in R-DRTD, the PD information is involved in the optimization, which is ignored in RO. On the other, R-DRTD utilizes the ambiguity set to model the uncertainty of the PD while SP directly uses the empirical PD, ignoring the uncertainty of the PD, leading to over-optimistic solutions. Although SP can obtain more economic solutions, this is at the cost of decreasing the reliability, leading to more unexpected risk, as shown in Table IV. 3) When the size of data is low, the total cost of M-DRO is similar to that of R-DRTD. However, as the size of data is increasing, the different between R-DRTD and M-DRO gradually enlarge. This is because M-DRO just uses moment information of the distribution. When more data are at hand, just more accurate moment information can be obtained in M-DRO. Even if M-DRO obtains the true moment information, all PDs with the true moment have to be considered. In contrast, R-DRTD involves more PD information in the optimization and thus fewer PDs have to be considered.

Table IV shows the operational risk reliability, i.e., the reliability of risk constraint (17), where the required reliability is 95%. Obviously, the operational risk reliability of is higher than that of SP, especially when the size of samples is low. This is because when the size of samples is low, the parameters of SP, which is estimated from samples, is much uncertain while the proposed DRO method can utilize the ambiguity set to model the uncertainties and thus a more reliable solution. Therefore, although the total cost of SP is less than that of the proposed method, we need to perform DRO methods to ensure the reliability. Of course, if we prefer the operational economic, SP is a more suitable method to be performed.

TABLE IV
RISK RELIABILITY UNDER DIFFERENT METHODS

Sample Number	50	500	5000	10000
R-DRTD	99.1%	98.2%	97.4%	96.9%
SP	80.5%	86.9%	89.2%	90.3%

TABLE V
TEST RESULTS UNDER DIFFERENT DISCRETIZATION NUMBER

Number	6	8	10	12	14
Time(s)	1.12	1.31	1.52	1.69	1.92
Total cost ($)	107826	107265	107150	107093	107083

The computing efficiency of different methods can be found in Table III. Obvilusly, the computing time of R-DRTD is less than SP and is similar to RO and M-DRO, indicating the high computing efficiency of R-DRTD. On the other hand, as the size of the sample data increases, the computing time of R-DRTD fluctuates slightly, which ensures the feasibility of incorporating more data to decrease the conservatism.

The sensitivity test of the discretization point number is also conducted on the 118-bus system. From the test results shown in Table V, it can be observed that when the number increases from 10 to 14, the total cost becomes almost constant. This indicates that the calculation precision has been high enough with 10 discretization points for the test system. Of course, as the number increasing, the calculation time also increases. But the time under 10 discretization points is acceptable.

The scalability of the proposed approach is tested on the real 445-bus system of Shandong Province, China. The 445-bus system has 48 generators and 693 transmission lines. Even for 50 wind farms with 10 discretization points, the time is still fast enough for on-line application, as shown in Table VI, indicating the high scalability of the proposed approach.

TABLE VI
TIME UNDER DIFFERENT NUMBERS OF WIND FARMS

Number	5	10	15	20	25	30	40	50
Time(s)	4.9	5.5	6.7	7.4	9.8	13.1	20.2	27.3

V. CONCLUSIONS

This paper studies a data-driven distributionally robust real-time dispatch problem, where both wind power uncertainties and the voltage security are considered. According to IDM, a CI-based ambiguity set is constructed to model the distributional uncertainties of wind power. Meanwhile, to ensure the nodal voltage security, except conventional units, VSWT based wind farms are also considered to provide reactive power support using two strategies, i.e., the AVC and linear decision rule. Based on the risk tractable estimation method and sequential convex optimization method, an efficient algorithm for the proposed R-DRTD model is also designed to ensure the scalability of the proposed approach in large-scale systems. Numerical results on a 6-bus and IEEE 118-bus systems verify the effectiveness and efficiency of the proposed approach.

REFERENCES

[1] N. Yorino, H. M. Hafiz, Y. Sasaki, and Y. Zoka, "High-speed real-time dynamic economic load dispatch," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 621–630, May 2012.

[2] A. Papavasiliou, Y. Mou, L. Cambier, and D. Scieur, "Application of stochastic dual dynamic programming to the real-time dispatch of storage under renewable supply uncertainty," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 547–558, Apr. 2018.

[3] Q. Wang, A. R. Florita, and et al, "Quantifying the economic and grid reliability impacts of improved wind power forecasting," *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1525–1537, Oct. 2016.

[4] Z. Li, W. Wu, B. Zhang, and B. Wang, "Adjustable robust real-time power dispatch with large-scale wind power integration," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 357–368, Apr. 2015.

[5] U. Ozturk, M. Mazumdar, and B. A. Norman, "A solution to the stochastic unit commitment problem using chance constrained programming," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1589–1598, Aug. 2004.

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

[7] R. Jiang, J. Wang, M. Zhang, and Y. Guan, "Two-stage minimax regret robust unit commitment," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2271–2282, Apr. 2013.

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

[9] R. A. Jabr, S. Karaki, and J. A. Korbane, "Robust multi-period OPF with storage and renewables," *IEEE Trans. Sustain. Energy*, vol. 30, no. 5, pp. 2790–2799, Nov. 2015.

[10] N. Amjady, A. Attarha, S. Dehghan, and A. J. Conejo, "Adaptive robust expansion planning for a distribution network with ders," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1698–1715, Mar. 2018.

[11] E. Samani and F. Aminifar, "Tri-level robust investment planning of DERs in distribution networks with AC constraints," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3749–3757, Sept. 2019.

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

[13] D. Bertsimas, E. Litvinov, X. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, Feb. 2013.

[14] P. Li, D. Yu, M. Yang, and J. Wang, "Flexible look-ahead dispatch realized by robust optimization considering CVaR of wind power," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5330–5340, Sept. 2018.

[15] A. Lorca and X. A. Sun, "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1702–1713, Jul. 2015.

[16] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 595–612, 2010.

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

[18] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Math. Program.*, vol. 158, no. 1, pp. 291–327, 2016.

[19] Q. Bian, H. Xin, Z. Wang, and et al., "Distributionally robust solution to the reserve scheduling problem with partial information of wind power," *IEEE Trans. Power App. Syst.*, vol. 30, no. 5, pp. 2822–2823, Sept. 2015.

[20] Z. Wang, Q. Bian, H. Xin, and D. Gan, "A distributionally robust coordinated reserve scheduling model considering cvar-based wind power reserve requirements," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 625–636, Apr. 2016.

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

[22] Y. Chen, Q. Guo, H. Sun, and et al., "A distributionally robust optimization model for unit commitment based on Kullback-Leibler Divergence," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5147–5160, Sept. 2018.

[23] L. Yao, X. Wang, and et al., "Data-driven distributionally robust reserve and energy scheduling over Wasserstein balls," *IET Gener. Transm. Distrib.*, vol. 12, no. 1, pp. 178–189, Jan. 2018.

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

[25] C. Wang, R. Gao, W. Wei, and et al., "Risk-based distributionally robust optimal gas-power flow with wasserstein distance," *IEEE Trans. Power Syst.*, 2018 in press.

[26] C. Duan, W. Fang, L. Jiang, L. Yao, and J. Liu, "Distributionally robust chance-constrained approximate AC-OPF with wasserstein metric," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 4924–4936, Sept. 2018.

[27] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Math. Programming*, pp. 1–52, 2015.

[28] J. Yang, N. Zhang, C. Kang, and Q. Xia, "A state-independent linear power flow model with accurate estimation of voltage magnitude," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3607–3617, Sept. 2016.

[29] M. Yang, M. Q. Wang, F. L. Cheng, and W. J. Lee, "Robust economic dispatch considering automatic generation control with affine recourse process," *Electr. Power Energy Syst.*, vol. 81, pp. 289–298, Oct. 2016.

[30] X. Zeng, J. Yao, Z. Chen, W. Hu, and T. Zhou, "Co-ordinated control strategy for hybrid wind farms with PMSG and FSIG under unbalanced grid voltage condition," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1100–1110, Jul. 2016.

[31] J. Yao, L. Guo, T. Zhou, D. Xu, and R. Liu, "Capacity configuration and coordinated operation of a hybrid wind farm with FSIG-based and PMSG-based wind farms during grid faults," *IEEE Trans. Energy Conv.*, vol. 32, no. 3, pp. 1188–1199, Sept. 2017.

[32] R. A. Jabr, "Linear decision rules for control of reactive power by distributed photovoltaic generators," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2165–2174, Mar. 2018.

[33] Y. Wang, S. Zhao, Z. Zhou, A. Botterud, Y. Xu, and R. Chen, "Risk adjustable day-ahead unit commitment with wind power based on chance constrained goal programming," *IEEE Trans. Sustain. Energy*, vol. 8, no. 2, pp. 530–541, Apr. 2017.

[34] P. Walley, *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, 1991.

[35] J. M. Bernard, "An introduction to the imprecise dirichlet model for multinomial data," *Int. J. Approx. Reason.*, vol. 39, no. 2, pp. 123–150, Jun. 2005.

[36] M. Yang, J. Wang, H. Diao, J. Qi, and X. Han, "Interval estimation for conditional failure rates of transmission lines with limited samples," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2752–2763, Jul. 2018.

[37] M. Goldman and D. M. Kaplan, "Evenly sensitive KS-type inference on distributions," *Tech. Rep.*, 2015.

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

[39] S. S. Wilks, *Mathematical Statistics*. Wiley series in probability and mathematical statistics, Wiley, 1962.

[40] F. Bouffard and F. D. Galiana, "Stochastic security for operations planning with significant wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 306–316, May 2008.

Peng Li (S'17) received the B.S. degree in electrical engineering from the School of Electrical Engineering, Shandong University, Jinan, China, in 2015. He is currently pursuing the Ph.D. degree in the School of Electrical Engineering, Shandong University, Jinan, China. In 2018, he was a Visiting Ph.D Student with the Department of Electrical Engineering at Southern Methodist University, Dallas, Texas, USA, for five months. He was also a Visiting Ph.D Student with the Center for Electric Power and Energy, Department of Electrical Engineering, Technical University of Denmark, from July 2019 to January 2020.

His research interests include renewable energy integration and power system optimal operation.

Chengfu Wang (M'15) received his B.Eng. degree in electrical engineering from Northeast Dianli University, Jilin, China, in 2006, the M.Eng. degree in electrical engineering from Harbin Institute of Technology, Heilongjiang, China, in 2008, lnd Ph.D. degree from the School of Electrical Engineering at Shandong University, Jinan, China, in 2012

He is currently a Research Associate at Shandong University. His research interests include wind power system operation and control, integrated energy system planning and operation.

Qiuwei Wu (M'08–SM'15) received the Ph.D. degree in power system engineering from Nanyang Technological University, Singapore, in 2009. He was a Senior Research and Development Engineer with VESTAS Technology R&D Singapore Pte, Ltd., from March 2008 to October 2009. He has been an Assistant Professor with the Department of Electrical Engineering, Technical University of Denmark since 2009. In 2012, he was a Visiting Scholar with the Department of Industrial Engineering and Operations Research, University of California, Berkeley, for three months funded by Danish Agency for Science, Technology and Innovation, Denmark. He was a Visiting Scholar with the Harvard China Project, School of Engineering and Applied Sciences, Harvard University from 2017 to 2018.

His research interests are operation and control of power systems with high penetration of renewables, including wind power modeling and control, active distribution networks, and operation of integrated energy systems. He is the Deputy-Editor-In-Chief of the International Journal of Electrical Power and Energy Systems. He is an Editor of the IEEE Transactions on Smart Grid and IEEE Power Engineering Letters. He is the Regional Editor for Europe of the IET Renewable Power Generation, a Subject Editor of IET Generation, Transmission and Distribution and an Associate Editor of the Journal of Modern Power Systems and Clean Energy.

Ming Yang (M'09-SM'18) received the B.Eng. and Ph.D degrees in electrical engineering from Shandong University, Jinan, China, in 2003 and 2009, respectively. He was an exchange Ph.D. student at Energy System Research Center, The University of Texas at Arlington, Arlington, TX, USA, from October 2006 to October 2007. He conducted postdoctoral research at School of Mathematics of Shandong University from July 2009 to July 2011. From November 2015 to October 2016 he was a Visiting Scholar at the Energy Systems Division, Argonne National Laboratory, Argonne, IL, USA.

Currently, he is an Professor at Shandong University. He is an Associate Editor of IEEE Transactions on Industry Applications and IET Renewable Power Generation. He is also an Editor of Protection and Control of Modern Power Systems. His research interests are power system optimal operation and control.