A Wasserstein distributionally robust planning model for renewable sources and energy storage systems under multiple uncertainties

Junkai Li, *Student Member IEEE*, Hong Liu *Member IEEE*, Zhengyang Xu, Chengshan Wang, *Senior Member IEEE*, Liyong Wang, Chenghong Gu

***Abstract*—Nowadays, electricity markets and carbon trading mechanisms can promote investment in renewable sources but also generate new uncertainties in decision-making. In this paper, a two-stage planning model is presented, considering supply-side and demand-side uncertainties in the distribution network and the interaction uncertainty from the main grid. First, these uncertain factors are depicted by the ambiguity sets based on the Wasserstein metric and historical data. Then, a two-stage Wasserstein distributionally robust optimization (WDRO) model is formulated to decide the optimal planning strategy for renewable energy generators (REGs) and energy storage systems (ESSs). Both 1-norm and ∞-norm Wasserstein metric constraints are considered to satisfy the decision-maker preference in different aspects. Furthermore, a systematic solution methodology with a three-step process is developed to solve the proposed WDRO. Numerical results from a modified IEEE 33-node system and a practical 130-node system demonstrate the advantages of the two-stage WDRO model and the effectiveness of the solution method.**

***Index Terms*—Two-stage planning model, Wasserstein metric, WDRO, decision maker's preference.**

NOMENCLATURE

A. Indices(*Sets*)

a	Index for different uncertain factors.
i , j , k	Index for distribution network buses.
l	Index for data samples.
t	Index for hourly intervals.
Ξ	Uncertainty Set.

B. Parameters

d	Discount rate.
N	Number of data samples.
N_{REG} , N_{ES}	Number of REGs and ESSs.
$\hat{P}_N$	Empirical distribution on basis of historical data.
$\tilde{P}_{\mathrm{L}j,t}$, $\tilde{Q}_{\mathrm{L}j,t}$	Active and reactive load demand in bus j at time t.
$\tilde{P}_{\mathrm{REG}j,t}$	REG outputs of in bus j at time t.
T_{REG} , T_{ES}	Lifetime of REG and ESS.
T_d , T_y	Operation time.
r_{ij} , x_{ij}	Resistance and reactance of branch ij.
$U_{\min}$, $U_{\max}$	Lower bound and upper bound of voltage in the distribution network.
$\hat{\boldsymbol{\xi}}$	Dataset containing all historical data.
$\hat{\xi}_{1t}^{(l)}$, $\hat{\xi}_{2t}^{(l)}$, $\hat{\xi}_{3t}^{(l)}$, $\hat{\xi}_{4t}^{(l)}$	The l-th data sample about different uncertain factors at time t.
κ , κ_1 , κ_∞	The radius of Wasserstein ball.
β , β_1 , β_∞	The confidence of Wasserstein metric.
η_{ESD} , η_{ESR}	Charging and discharging efficiency of ESSs.
η_{ope}	Coefficient of annual operational cost.
λ_{REG} , λ_{ESC} , λ_{ESP}	Capital cost of REGs and ESSs.
λ_{CO2} , $\tilde{e}_t$	Carbon trading price and nodal carbon intensity from the main grid.
$\tilde{\lambda}_{\mathrm{LMP},t}$	Location marginal price from the main grid at time t.
$E_{\mathrm{ES}i,0}$, $E_{\mathrm{ES}i,24}$	Initial state of charge ESSs.

C. Variables

$\tilde{C}_{\mathrm{CO2},t}$, $C_{\mathrm{CO2},t}$	Carbon trading cost.
$\tilde{C}_{\mathrm{E},t}$, $C_{\mathrm{E},t}$	Power trading cost.
C_{REG} , C_{ES} , C_{ope}	Annual capital cost of REGs, ESSs and devices operation.
$E_{\mathrm{ES}i,t}$	Operation strategy of ESSs.
$I_{ij,t}$, $U_{i,t}$	Branch current and node voltage.
$P_{j,t}$, $Q_{j,t}$	Active and reactive power injected in node j at time t.
$P_{ij,t}$, $P_{jk,t}$, $Q_{ij,t}$, $Q_{jk,t}$	Active and reactive power in different branches.
ξ_{1t} , ξ_{2t} , ξ_{3t} , ξ_{4t}	Four kinds of uncertainties at time t.

This work was supported by the National Natural Science Foundation of China under Grant U1866207. *(Corresponding author: Hong Liu)*

Junkai Li, Hong Liu, Zhengyang Xu and Chengshan Wang are with the School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: lijunkai6666@tju.edu.cn; liuhong@tju.edu.cn; xuzhengyang@tju.edu.cn; cswang@tju.edu.cn).

Liyong Wang was with the Xi'an Jiaotong University, Xi'an 710049, China. He is now with the State Grid Beijing Electric Power Company, Beijing 100031, China (e-mail: 13810007912@163.com).

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

$\Delta\xi_{1t}^{(l)}, \Delta\xi_{2t}^{(l)}, \Delta\xi_{3t}^{(l)}, \Delta\xi_{4t}^{(l)}$	Uncertain variables robust deviation of original samples.
$S_{\mathrm{REG},i}, S_{\mathrm{ESC},i}, S_{\mathrm{ESP},i}$	Capacity of REGs and ESSs.

I. Introduction

NOWADAYS, excessive greenhouse gas emissions have caused numerous environmental issues [1], such as global warming and frequent extreme weather events. To address this concern, investing in REGs and ESSs in the power industry attracts great attention and leads to an investment boom [2]. The distributed system operator (DSO) usually installs REGs and ESSs on the grid side. They can reduce power purchase fees and carbon tax under power and carbon trading mechanisms, thereby creating more economic benefits. However, these mechanisms also bring more uncertainties that do not have typical probability distributions, resulting in difficulties and challenges in decision-making. Hence, it is significant to formulate a credible and tractable mathematical model to determine the planning strategy for REGs and ESSs under multiple uncertainties.

In the past few years, many researchers have concentrated on the coordinated planning of REGs and ESSs in the distribution network. For instance, Conte *et al*. [3] proposed an optimal planning method for the integrated battery and photovoltaic power plant, which can produce photovoltaic power and provide droop-based primary frequency regulations to the main grid. In [4],[5], a two-stage planning of REGs and ESSs was formulated to minimize the comprehensive costs of the DSO. Oskouei *et al*. [6] presented an unified decision-making structure that consists of network partitioning and optimal planning of REGs and ESSs. However, most research does not consider economic profits generated by carbon emission reductions when the DSO invests in REGs and ESSs. To fill this knowledge gap, a deterministic planning model of REGs and ESSs considering energy price and transmission access fee was formulated in [7] preliminarily, but no uncertain factors were considered.

Generally, decision-makers should lay stress on the uncertain factors in finding the optimal planning strategy, which satisfies different scenarios comprehensively. In the coordinated planning of REGs and ESSs, most literature expounds supply-side and demand-side uncertainties, which refer to the uncertain outputs of REGs and the randomness in demand respectively [8]-[11]. In addition, the DSO also has interaction uncertainties with the main grid under the power and carbon trading mechanisms. It is analogous to the uncertain transaction fee of integral energy, which has been considered in the optimal scheduling of distributed energy systems [12]. However, these uncertain factors do not attract much attention in the coordinated planning of REGs and ESSs, particularly the uncertainties of locational marginal price (LMP) and nodal carbon intensity (NCI) in the main grid [13],[14].

To address the aforementioned uncertainties, numerous studies have been conducted in this area. Stochastic optimization (SO) is a simple but practical modeling paradigm that has been widely applied in decision-making. Hemmati *et al*. [15] developed a stochastic expansion planning in microgrids, incorporating renewable energy resources and energy storage systems. Zhu *et al*. [16] proposed an energy storage capacity sizing method by using numerous solar power outputs and load samples. Generally, SO often makes the optimal decision through known probability distributions or a family of typical scenarios. However, the decision-maker usually does not know the probability distributions hiding behind historical samples in advance. Without a large number of data samples, the distribution is often partially observed through limited data samples and the out-of-sample performance is usually disappointing [17]. Even if there are adequate samples, the performance of SO is also unsatisfying because of high computational complexity [18]. Robust optimization (RO) is another common method for decision-making under uncertain factors. Golpira *et al*. [9] expounded a multi-objective risk-based robust optimization approach for energy management under demand and supply uncertainties. RO often models uncertain variables by intervals and is only concerned with the worst-case scenario. Compared to SO, RO has the advantages of improving decision robustness and reducing computational complexity. However, RO often leads to the excessive conservativeness of the scheme because the worst-case scenario appears with an extremely low probability.

In recent years, DRO has gained increasing popularity by overcoming the disadvantages of SO and RO [19]. The traditional DRO can be divided into two categories. The first category of DRO is based on the moment ambiguity set which contains all distributions satisfying moment constraints [19]. With increasing uncertainty variables, the optimized results might deviate from the true probability distributions [20]. The second category is based on several worst-case scenarios clustering from historical data [21]. However, the characteristics of the worst-case scenario have to be known in advance for extracting them from clustering results. It might be difficult when several uncertainties are considered simultaneously. To address these issues, Wasserstein metric-based DRO has attracted many attentions. In [22]-[25], a WDRO model was proposed for economic dispatch and unit commitment in the power system with variable renewable energy resources. Y. Zhou *et al*. [26] and Y. Wang *et al*. [27] extended this approach to the resilient operation and economic scheduling in the integrated energy system respectively. However, these methods only utilize one type of Wasserstein metric, which is not adequate enough to limit the ambiguity set from different dimensions. Furthermore, very few papers propose a two-stage planning model to decide the optimal allocation of REGs and ESSs in the distribution network under multiple uncertainties.

Based on the abovementioned analysis, this paper proposes a two-stage WDRO model for the coordinated planning of REGs and ESSs. The main contributions of this work can be summarized as follows.

1) Both supply-side and demand-side uncertainties from the distribution network and interaction uncertainty from the main

grid are considered in this paper. These uncertainties are depicted by ambiguity sets based on Wasserstein metric and historical data, which are more proper to model the interaction uncertainties without any typical distributions, such as LMP and NCI.

2) A two-stage WDRO planning model with 1-norm and ∞-norm Wasserstein metric constraints is formulated to decide the planning and operation schemes for REGs and ESSs. This model can comprehensively depict the characteristics of ambiguity sets from general differences and local deviations on basis of decision-maker's preference.

3) A systematic solution methodology is developed to solve the proposed two-stage WDRO planning model with second-order cone constraints. Specifically, this model is converted into a tractable reformulation with linearized power flow, from which the worst-case distribution of uncertainty variables can be obtained. Through the known distribution, the original model is transformed into a mixed-integer second-order cone programming (MISOCP) and solved by a commercial optimizer efficiently.

The rest of this paper is organized as follows. Section 2 presents a Wasserstein ambiguity set to describe the multiple uncertainties in the decision-making. Section 3 formulates a two-stage WDRO model to obtain the optimal planning strategies for REGs and ESSs. The solution methodology for this model is developed in Section 4. Section 5 validates the advantages of the proposed two-stage WDRO model and the performance of the solution methodology by a modified IEEE 33-node system and a practical 130-node system. Section 6 provides a summary and concluding remarks.

II. Wasserstein Ambiguity Set for Multiple Uncertainties in Decision-Making

To decide the optimal planning strategies, DSO should lay stress on multiple uncertain factors, including supply-side uncertainty, demand-side uncertainty and interaction uncertainty. The former two correspond to the output of REGs and demand characteristics in the distribution system. The latter one refers to the uncertainties generated by the power and carbon trading mechanisms in the main grid, such as LMP and NCI. Notably, carbon trading prices determined by the government [28]are often comparatively stable, and thereby they are assumed to be deterministic in this paper.

Assuming that the decision-maker has collected hourly historical data in N days about the aforementioned information. Then, a dataset A_{train}= { $\hat{\boldsymbol{\xi}}^{(1)}$, $\hat{\boldsymbol{\xi}}^{(2)}$,…, $\hat{\boldsymbol{\xi}}^{(l)}$,…, $\hat{\boldsymbol{\xi}}^{(N)}$ } is constructed by the LMP, NCI, renewable output and demand characteristics in each hour. $\hat{\boldsymbol{\xi}}^{(l)}$ denotes four kinds of historical data respectively, i.e. $\hat{\boldsymbol{\xi}}^{(l)}$ = { $\hat{\xi}_1^{(l)}$, $\hat{\xi}_2^{(l)}$, $\hat{\xi}_3^{(l)}$, $\hat{\xi}_4^{(l)}$ }. The empirical distribution from the historical data can be defined as follows.

$$\hat{P}_N = \frac{1}{N}\sum_{l=1}^{N}\delta_{\hat{\boldsymbol{\xi}}^{(l)}} \tag{1}$$

To quantitatively measure the distance between the empirical distribution and any other distributions, the Wasserstein metric is defined as follows [24].

$$d_{\mathrm{W}}(\hat{P}_N, R) = \min_{\Phi}\int_{\Xi^2}\left\|\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}\right\|\Phi(\mathrm{d}\hat{\boldsymbol{\xi}},\mathrm{d}\boldsymbol{\xi}) \tag{2}$$

By analyzing this definition in details, the essence of the Wasserstein metric is an optimization problem, where the decision variable Φ is a joint distribution of $\hat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}$. It can be viewed as the minimum cost of the transportation plan Φ for moving one probability distribution to another.

Based on the Wasserstein metric, a Wasserstein ambiguity set that contains all distributions analogous to the empirical distribution is constructed. This set can be described by the Wasserstein ball of radius κ centered at the empirical distribution.

$$B_\kappa(\hat{P}_N) = \{R : R(\boldsymbol{\xi}\in\Xi)=1\}\bigcap\{R : d_{\mathrm{W}}(\hat{P}_N, R)\le\kappa\} \tag{3}$$

Notably, radius κ is closely related to the performance of the Wasserstein ambiguity set and the following formulas provide the relationship between the confidence level and the radius on the basis of some statistical methods [29].

$$P\{d_{\mathrm{W}}(R,\hat{P}_N)\le\kappa\}\ge\beta = 1-\exp(-N\frac{\kappa^2}{D^2}) \tag{4}$$

$$\kappa = D\sqrt{\frac{2}{N}\ln(\frac{1}{1-\beta})} \tag{5}$$

where D is a constant that can be derived by the following optimization problem.

$$D = \min_{\mu\ge0} 2\sqrt{\frac{1}{2\mu}(1+\ln(\frac{1}{N}\sum_{l=1}^{N}e^{\mu\left\|\hat{\boldsymbol{\xi}}^{(l)}-\hat{\boldsymbol{\gamma}}\right\|}))} \tag{6}$$

Where $\hat{\boldsymbol{\gamma}}$ is the sample mean.

Note that there are several advantages of this Wasserstein ambiguity set. First, it is centered on the empirical distribution derived from finite data samples, while restricting the distributional uncertainty on the basis of the Wasserstein metric. This feature is more appropriate to depict multiple uncertainties in the decision-making where the distribution is often partially observable through a finite amount of historical data. Then, there is a statistical guarantee that the true probability distribution is contained in the Wasserstein ball with a certain confidence level. Lastly, the decision-maker can change the radius of the Wasserstein ball to control the conservatism of optimized results. Note that if the radius of the Wasserstein ball is reduced to 0, it is identical to the original empirical distribution.

III. Two-stage WDRO Model for Renewable Sources and Energy Storage Systems Planning

In this section, a data-driven Wassterstein DRO is formulated as a two-stage optimization, in which the planning strategy is determined in the first stage and the operation strategy is optimized in the second stage. The location and capacity of devices resulted from the first stage will affect the energy purchase and operation of ESSs in the second stage.

A. Objective Function

The objective is to minimize the comprehensive cost of the DSO, including the annual capital cost, and operation cost of REGs and ESSs, and the annual energy trading cost from the power consumption and equivalent carbon emissions.

$$F = \min C_{\mathrm{REG}} + C_{\mathrm{ES}} + C_{\mathrm{ope}} + T_y \max_{\boldsymbol{\xi}} \min_{P_{\mathrm{ES},t}, P_{0,t}} \sum_{t=0}^{T_d} (\tilde{C}_{\mathrm{E},t} + \tilde{C}_{\mathrm{CO2},t}) \quad (7)$$

$$C_{\mathrm{REG}} = \frac{d(1+d)T_{\mathrm{REG}}}{(1+d)^{T_{\mathrm{REG}}} - 1} \sum_{i=1}^{N_{\mathrm{REG}}} \lambda_{\mathrm{REG}} S_{\mathrm{REG},i} \quad (8)$$

$$C_{\mathrm{ES}} = \frac{d(1+d)T_{\mathrm{ES}}}{(1+d)^{T_{\mathrm{ES}}} - 1} \sum_{i=1}^{N_{\mathrm{ES}}} (\lambda_{\mathrm{ESC}} S_{\mathrm{ESC},i} + \lambda_{\mathrm{ESP}} S_{\mathrm{ESP},i}) \quad (9)$$

$$C_{\mathrm{ope}} = \eta_{\mathrm{ope}} (C_{\mathrm{REG}} + C_{\mathrm{ES}}) \quad (10)$$

$$\tilde{C}_{\mathrm{E},t} = \tilde{\lambda}_{\mathrm{LMP},t} P_{0.t} \quad (11)$$

$$\tilde{C}_{\mathrm{CO2},t} = \lambda_{\mathrm{CO2}} \tilde{e}_t P_{0.t} \quad (12)$$

Eqs. (8)-(10) represent the capital cost of REGs, the capital cost of ESSs and related operation cost. Eqs. (11)-(12) indicate the energy trading costs.

B. Constraints

1) *Wassertein metric constraints*

$$d_{\mathrm{W1}}(\hat{P}_N, R_t(\xi)) = \min_{\Phi} \int_{\Xi^2} \left\| \hat{\boldsymbol{\xi}} - \boldsymbol{\xi} \right\|_1 \Phi(\mathrm{d}\hat{\boldsymbol{\xi}}, \mathrm{d}\boldsymbol{\xi}) \le \kappa_1 \quad \forall t \in T_d \quad (13)$$

$$P\{d_{\mathrm{W1}}(R, \hat{P}_N) \le \kappa_1\} \ge \beta_1 \quad (14)$$

$$d_{\mathrm{W}\infty}(\hat{P}_N, R_t(\xi)) = \min_{\Phi} \int_{\Xi^2} \left\| \hat{\boldsymbol{\xi}} - \boldsymbol{\xi} \right\|_\infty \Phi(\mathrm{d}\hat{\boldsymbol{\xi}}, \mathrm{d}\boldsymbol{\xi}) \le \kappa_\infty \quad \forall t \in T_d \quad (15)$$

$$P\{d_{\mathrm{W}\infty}(R, \hat{P}_N) \le \kappa_\infty\} \ge \beta_\infty \quad (16)$$

Both 1-norm Wasserstein metric and ∞-norm Wasserstein metric are considered in this paper. The former one is related to the general difference between probability distributions and the latter one represents the local deviation between probability distributions. If κ_1 or κ_∞ reduces to 0, this problem will be equivalent to the SO.

2) *Power flow constraints*

$$P_{j,t} = \tilde{P}_{\mathrm{L}j,t} - P_{\mathrm{ESD}j,t} + P_{\mathrm{ESR}j,t} - \tilde{P}_{\mathrm{REG}j,t} \quad \forall t, \forall j \quad (17)$$

$$Q_{j,t} = \tilde{Q}_{\mathrm{L}j,t} \quad \forall t, \forall j \quad (18)$$

$$\sum_{i \in \mathrm{end}(j)} (P_{ij,t} - r_{ij} I_{ij,t}^2) = \sum_{k \in \mathrm{head}(j)} (P_{jk,t} + P_{j,t}) \quad \forall t, \forall j \quad (19)$$

$$\sum_{i \in \mathrm{end}(j)} (Q_{ij,t} - x_{ij} I_{ij,t}^2) = \sum_{k \in \mathrm{head}(j)} (Q_{jk,t} + Q_{j,t}) \quad \forall t, \forall j \quad (20)$$

$$U_{j,t}^2 = U_{i,t}^2 - 2(r_{ij} P_{ij,t} + x_{ij} Q_{ij,t}) + (r_{ij}^2 + x_{ij}^2) I_{ij,t}^2 \quad \forall t, \forall ij \quad (21)$$

$$(P_{ij,t})^2 + (Q_{ij,t})^2 = (I_{ij,t} U_{i,t})^2 \quad \forall t, \forall ij \quad (22)$$

$$\left\| 2P_{ij,t} \quad 2Q_{ij,t} \quad I_{ij,t}^2 - U_{i,t}^2 \right\|_2 \le I_{ij,t}^2 + U_{i,t}^2 \quad \forall t, \forall ij \quad (23)$$

Eqs. (17)-(18) represent the power injection and Eqs. (19)-(20) indicate the law of power conversation. Eq. (21) is Ohm's law over branch *ij* and Eq. (22) is converted into a second-order cone formulation (23).

3) *Operation constraints of energy storage*

$$0 \le \left| P_{\mathrm{ESD}i,t} \right| \le S_{\mathrm{ESP}i} \quad (24)$$

$$0 \le \left| P_{\mathrm{ESR}i,t} \right| \le S_{\mathrm{ESP}i} \quad (25)$$

$$E_{\mathrm{ES}i,t} \le S_{\mathrm{ESC}i} \quad \forall t, \forall i \quad (26)$$

$$E_{\mathrm{ES}i,t+1} = E_{\mathrm{ES}i,t} - P_{\mathrm{ESD}i,t} / \eta_{\mathrm{ESD}} \Delta t + P_{\mathrm{ESR}i,t} \eta_{\mathrm{ESR}} \Delta t \quad \forall t, \forall i \quad (27)$$

$$E_{\mathrm{ES}i,0} = E_{\mathrm{ES}i,24} \quad \forall j \quad (28)$$

The charge/discharge constraints of the energy storage system are represented by Eqs. (24)-(26). Eq. (27) denotes the relationship between charge power, discharge power and the state of charge. Eq. (28) indicates that the initial and final states of the energy storage are the same.

4) *Additional constraints*

$$U_{\min}^2 \le U_{j,t}^2 \le U_{\max}^2 \quad \forall t, \forall j \quad (29)$$

$$P_{0.t} \ge 0 \quad \forall t \quad (30)$$

$$P_{ij}^2 + Q_{ij}^2 \le \chi S_{\mathrm{L}}^2 \quad (31)$$

Eq. (29) represents the security constraint of the distribution network. Moreover, the DSO is cast as a price taker in the trading mechanism and the reverse power flow injected into the main grid is not allowed in (30). Eq. (31) is the load factor constraint and χ is the load factor requirement in the network.

IV. Solution Methodology of Two-stage WDRO

By analyzing the WRO model in details, the dual problem of the second-stage optimization is difficult to obtain. On one hand, if the dual problem is based on the minimum problem about $P_{\mathrm{ES},t}$ and $P_{0,t}$, the second-order constraint (23) in the distribution network has quadratic terms, such as P^2, which will increase the difficulty in formulating the dual problem. On the other hand, if the dual problem is based on the maximum problem about $\boldsymbol{\xi}$, new difficulty will occur when both 1-norm Wasserstein metric and ∞-norm Wasserstein metric are considered simultaneously.

Therefore, a systematic solution method is developed to solve the two-stage WDRO model in this section. As shown in Fig.1, this method contains the following three steps. First, the two-stage WDRO is simplified by the linearized power flow and transformed into a tractable formulation. Then, the Column and Constraint Generation (CCG) algorithm is used to solve the reformulation model [25] and the probability distribution of uncertain variables can be obtained. Finally, the original model is transformed into the MISOCP with known distribution and the optimal planning strategy is obtained from the MISOCP model easily.

A. Original Model Tractable Transformation

Based on the linearized power flow, the power flow constraints (19)-(23) can be simplified as follows [30].

$$\sum_{i \in \mathrm{end}(j)} P_{ij,t} = \sum_{k \in \mathrm{head}(j)} (P_{jk,t} + P_{j,t}) \quad \forall t, \forall j \quad (32a)$$

$$\sum_{i \in \mathrm{end}(j)} Q_{ij,t} = \sum_{k \in \mathrm{head}(j)} (Q_{jk,t} + Q_{j,t}) \quad \forall t, \forall j \quad (32b)$$

$$U_{j,t} = U_{i,t} - (r_{ij} P_{ij,t} + x_{ij} Q_{ij,t}) \quad \forall t, \forall ij \quad (32c)$$

Meanwhile, (31) is a typical quadratic constraint, which can be linearized by the polygonal approximation method [30].

$$\gamma_{b,0} P_{ij} + \gamma_{b,1} Q_{ij} + \gamma_{b,2} S_{\mathrm{L}} \le 0 \quad \forall b \in Z \quad (33)$$

where $\gamma_{a,0}$, $\gamma_{a,1}$ and $\gamma_{a,2}$ are the coefficients of each linear term.

Therefore, the original DRO model can be transformed into a linear programming with Wassertein ambiguity set constraint. In the following formulas, $\boldsymbol{y}$ represents the decision variables in the first-stage, i.e. location and capacity of REGs and ESSs. $\boldsymbol{\xi}$ and $\boldsymbol{x}$

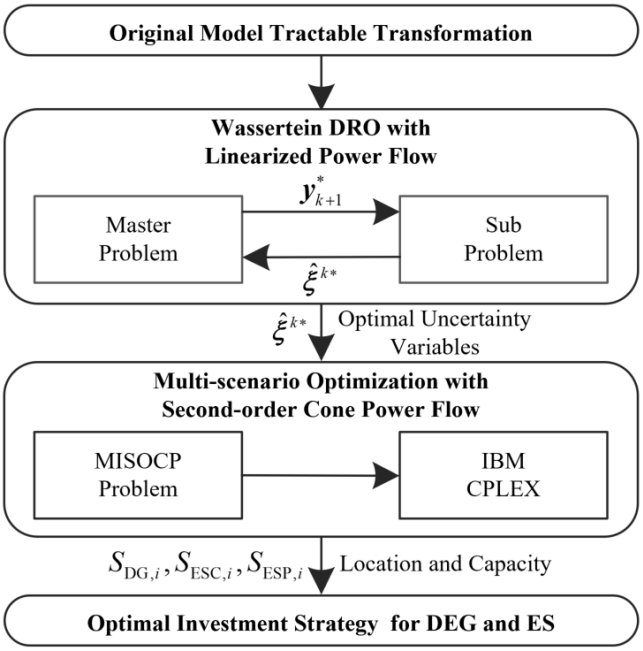

Fig. 1. Flow chart of solving the two-stage WDRO.

denote the uncertain factors and operation strategies of ESSs optimized in the second-stage.

$$F=\min_{\boldsymbol{y}}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{y}+\max_{\boldsymbol{\xi}}\min_{\boldsymbol{x}}\sum_{t=0}^{T_d}(\xi_{1t}\boldsymbol{F}_1^{\mathrm{T}}\boldsymbol{x}+\xi_{2t}\boldsymbol{F}_2^{\mathrm{T}}\boldsymbol{x}) \tag{34a}$$

s.t.

$$\boldsymbol{y}\geq 0 \tag{34b}$$

$$\boldsymbol{y}_{\mathrm{REG}}=\boldsymbol{F}_3 \tag{34c}$$

$$\boldsymbol{G}_1\boldsymbol{x}\geq\boldsymbol{G}_2-\boldsymbol{G}_3\boldsymbol{y} \tag{34d}$$

$$\boldsymbol{J}_1\boldsymbol{x}=\boldsymbol{J}_2-\xi_{3t}\boldsymbol{F}_3-\xi_{4t}\boldsymbol{F}_4\quad\forall t\in T_d \tag{34e}$$

$$\text{Eqs. (13)-(16)} \tag{34f}$$

Eq. (34b) denotes the linear inequality constraint in the first stage, and Eqs. (34c)-(34e) indicate the linear equality and inequality constraints in the second stage.

If the convexity of the optimization problem can be ensured, then Eqs. (34a)-(34e) coincide with the following tractable formulation [17].

$$F=\min_{\boldsymbol{y}}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{y}+\max_{\boldsymbol{\xi}}\min_{\boldsymbol{x}}\frac{1}{N}\sum_{t=0}^{T_d}\sum_{l=1}^{N}[(\hat{\xi}_{1t}^{(l)}-\Delta\xi_{1t}^{(l)})\boldsymbol{F}_1^{\mathrm{T}}\boldsymbol{x} +(\hat{\xi}_{2t}^{(l)}-\Delta\xi_{2t}^{(l)})\boldsymbol{F}_2^{\mathrm{T}}\boldsymbol{x}] \tag{35a}$$

s.t.

$$\boldsymbol{y}\geq 0 \tag{35b}$$

$$\boldsymbol{y}_{\mathrm{REG}}=\boldsymbol{F}_3 \tag{35c}$$

$$\boldsymbol{G}_1\boldsymbol{x}\geq\boldsymbol{G}_2-\boldsymbol{G}_3\boldsymbol{y} \tag{35d}$$

$$\boldsymbol{J}_1\boldsymbol{x}=\boldsymbol{J}_2-\frac{1}{N}\sum_{l=1}^{N}(\hat{\xi}_{3t}^{(l)}-\Delta\xi_{3t}^{(l)})\boldsymbol{F}_3 -\frac{1}{N}\sum_{l=1}^{N}(\hat{\xi}_{4t}^{(l)}-\Delta\xi_{4t}^{(l)})\boldsymbol{F}_4\quad\forall t\in T_d \tag{35e}$$

$$\frac{1}{N}\sum_{l=1}^{N}\left(\left|\Delta\xi_{1t}^{(l)}\right|+\left|\Delta\xi_{2t}^{(l)}\right|+\left|\Delta\xi_{3t}^{(l)}\right|+\left|\Delta\xi_{4t}^{(l)}\right|\right)\leq\kappa_1\quad\forall t\in T_d \tag{35f}$$

$$\frac{1}{N}\sum_{l=1}^{N}\left|\Delta\xi_{at}^{(l)}\right|\leq\kappa_\infty\quad\forall t\in T_d,\forall a=1,2,3,4 \tag{35g}$$

B. Solution Method for Reformulation Model

The reformulation model also has a two-stage structure, which can be solved by the wildly used CCG algorithm. According to different objectives, the reformulation model is decomposed to the following master problem and subproblem, and the optimal values of decision variables are derived by iteration.

1) *Master problem of the reformulation model*

$$\min_{\boldsymbol{y}}\boldsymbol{E}^{\mathrm{T}}\boldsymbol{y}+\eta \tag{36a}$$

s.t.

$$\eta\geq\frac{1}{N}\sum_{t=0}^{T_d}\sum_{l=1}^{N}[\hat{\xi}_{1t}^{(l)k*}\boldsymbol{F}_1^{\mathrm{T}}\boldsymbol{x}_k+\hat{\xi}_{2t}^{(l)k*}\boldsymbol{F}_2^{\mathrm{T}}\boldsymbol{x}_k]\quad\forall k \tag{36b}$$

$$\boldsymbol{y}\geq 0 \tag{36c}$$

$$\boldsymbol{G}_1\boldsymbol{x}_k\geq\boldsymbol{G}_2-\boldsymbol{G}_3\boldsymbol{y}\quad\forall k \tag{36d}$$

$$\boldsymbol{J}_1\boldsymbol{x}_k=\boldsymbol{J}_2-\frac{1}{N}\sum_{l=1}^{N}\hat{\xi}_{3t}^{(l)k*}\boldsymbol{F}_3-\frac{1}{N}\sum_{l=1}^{N}\hat{\xi}_{4t}^{(l)k*}\boldsymbol{F}_4\quad\forall t,\forall k \tag{36e}$$

The master problem is related to the decision-making of planning strategy in the first stage. In this problem, the values of uncertainty variables (i.e. $\hat{\xi}_1^{(l)k*}$, $\hat{\xi}_2^{(l)k*}$, $\hat{\xi}_3^{(l)k*}$, $\hat{\xi}_4^{(l)k*}$) are input parameters, obtained from the optimal solution of the subproblem.

2) *Subproblem of the reformulation model*

$$\max_{\boldsymbol{\xi}}\min_{\boldsymbol{x}}\frac{1}{N}\sum_{t=0}^{T_d}\sum_{l=1}^{N}[(\hat{\xi}_{1t}^{(l)}-\Delta\xi_{1t}^{(l)})\boldsymbol{F}_1^{\mathrm{T}}\boldsymbol{x}+(\hat{\xi}_{2t}^{(l)}-\Delta\xi_{2t}^{(l)})\boldsymbol{F}_2^{\mathrm{T}}\boldsymbol{x}] \tag{37a}$$

s.t.

$$\boldsymbol{G}_1\boldsymbol{x}\geq\boldsymbol{G}_2-\boldsymbol{G}_3\boldsymbol{y}_{k+1}^{*}\quad(\boldsymbol{\pi}) \tag{37b}$$

$$\boldsymbol{J}_1\boldsymbol{x}=\boldsymbol{J}_2-\frac{1}{N}\sum_{l=1}^{N}(\hat{\xi}_{3t}^{(l)}-\Delta\xi_{3t}^{(l)})\boldsymbol{F}_3 -\frac{1}{N}\sum_{l=1}^{N}(\hat{\xi}_{4t}^{(l)}-\Delta\xi_{4t}^{(l)})\boldsymbol{F}_4\quad\forall t\in T_d\quad(\boldsymbol{\mu}_t) \tag{37c}$$

$$\frac{1}{N}\sum_{l=1}^{N}\left(\left|\Delta\xi_{1t}^{(l)}\right|+\left|\Delta\xi_{2t}^{(l)}\right|+\left|\Delta\xi_{3t}^{(l)}\right|+\left|\Delta\xi_{4t}^{(l)}\right|\right)\leq\kappa_1\quad\forall t\in T_d \tag{37d}$$

$$\frac{1}{N}\sum_{l=1}^{N}\left|\Delta\xi_{at}^{(l)}\right|\leq\kappa_\infty\quad\forall t\in T_d,\forall a=1,2,3,4 \tag{37e}$$

$$\boldsymbol{y}_{\mathrm{REG},k+1}^{*}=\boldsymbol{F}_3 \tag{37f}$$

The subproblem represents the decision-making of operation strategies in the second stage. This problem is a max-min optimization, which is difficult to solve directly. By the strong duality theorem, it is reformulated to the following monolithic form. $\boldsymbol{\pi}$ and $\boldsymbol{\mu}_t$ are the vectors of dual variables.

$$f_S=\max_{\boldsymbol{\xi}}\boldsymbol{\pi}^{\mathrm{T}}(\boldsymbol{G}_2-\boldsymbol{G}_3\boldsymbol{y}_{k+1}^{*})+\sum_{t=0}^{T_d}\boldsymbol{\mu}_t^{\mathrm{T}}[\boldsymbol{J}_2-\frac{1}{N}\sum_{l=1}^{N}\boldsymbol{F}_3(\hat{\xi}_{3t}^{(l)}-\Delta\xi_{3t}^{(l)})-\frac{1}{N}\sum_{l=1}^{N}\boldsymbol{F}_4(\hat{\xi}_{4t}^{(l)}-\Delta\xi_{4t}^{(l)})] \tag{38a}$$

s.t.

$$\boldsymbol{y}^{*}_{\mathrm{REG},k+1}=\boldsymbol{F}_3 \tag{38b}$$

$$\frac{1}{N}\sum_{t=0}^{T_d}\sum_{l=1}^{N}[(\hat{\xi}^{(l)}_{1t}-\Delta\xi^{(l)}_{1t})\boldsymbol{F}_1^{\mathrm{T}}+(\hat{\xi}^{(l)}_{2t}-\Delta\xi^{(l)}_{2t})\boldsymbol{F}_2^{\mathrm{T}}]-\boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{G}_1 -\sum_{t=0}^{T_d}\boldsymbol{\mu}_t^{\mathrm{T}}\boldsymbol{J}_1=\boldsymbol{0}\quad\forall t\in T_d \tag{38c}$$

$$\frac{1}{N}\sum_{l=1}^{N}\left(\left|\Delta\xi^{(l)}_{1t}\right|+\left|\Delta\xi^{(l)}_{2t}\right|+\left|\Delta\xi^{(l)}_{3t}\right|+\left|\Delta\xi^{(l)}_{4t}\right|\right)\le\kappa_1\quad\forall t\in T_d \tag{38d}$$

$$\frac{1}{N}\sum_{l=1}^{N}\left|\Delta\xi^{(l)}_{at}\right|\le\kappa_\infty\quad\forall t\in T_d,\forall a=1,2,3,4 \tag{38e}$$

3) *CCG algorithm*

CCG algorithm is an efficient method to solve the two-stage optimization. The master problem and subproblem provide the upper bound and lower bound for the iteration respectively. The detailed process is shown as follows.

Step 1: Set LB=0, UB=+∞ and initialize the index k=0.

Step 2: Solve the master problem and update the lower bound LB=max{LB, $\boldsymbol{E}^{\mathrm{T}}\boldsymbol{y}^{*}_{k+1}+\eta^{*}_{k+1}$}.

Step 3: Solve the subproblem and update the upper bound LB=min{LB, $\boldsymbol{E}^{\mathrm{T}}\boldsymbol{y}^{*}_{k+1}+f_S$}

Step 4: Check the convergence: If UB−LB$\leqslant\gamma$, the iteration terminates. Otherwise, update k=k+1, add new constraints and go back to step 2.

C. MISOCP Model with Known Distribution

To modify the optimal results based on the linearized power flow, a MISOCP model with the worst-case probability distribution $\boldsymbol{\xi}^{*}$ from Section IV-B is constructed. In this model, the worst-case distribution of uncertainty is known so that the original Wassertein DRO can be simplified to the following MISOCP.

$$\min_{S_{\mathrm{DEG}},S_{\mathrm{ESC}},S_{\mathrm{ESP}},P_{\mathrm{ES},t},P_{0,t}} C_{\mathrm{REG}}+C_{\mathrm{ES}}+C_{\mathrm{ope}}+T_y\sum_{t=0}^{T_d}(C_{\mathrm{E},t}+C_{\mathrm{CO2},t}) \tag{39a}$$

$$C_{\mathrm{E},t}=\frac{1}{N}\sum_{l=1}^{N}\hat{\lambda}^{(l)*}_{\mathrm{LMP},t}P_{0.t} \tag{39b}$$

$$C_{\mathrm{CO2},t}=\frac{1}{N}\sum_{l=1}^{N}\lambda_{\mathrm{CO2}}\hat{e}^{(l)*}_{t}P_{0.t} \tag{39c}$$

$$P_{j,t}=\hat{P}^{(l)*}_{\mathrm{L}j,t}-P_{\mathrm{ESD}j,t}+P_{\mathrm{ESR}j,t}-\hat{P}^{(l)*}_{\mathrm{REG}j,t}\quad\forall t,\forall j,\forall l \tag{39d}$$

$$Q_{j,t}=\hat{Q}^{(l)*}_{\mathrm{L}j,t}\quad\forall t,\forall j,\forall l \tag{39e}$$

$$\text{Eqs. (8)-(10), (19)-(23), (24)-(30), (33)} \tag{39f}$$

In Eqs. (39b)-(39e), all uncertain variables are replaced by the optimized probability distribution. Essentially, this MISOCP model is a multi-scenario optimization, which can be solved by many commercial optimizers, such as IBM CPLEX.

V. Case Studies

A. System Parameters

In this section, the modified IEEE 33-node system is chosen to verify the proposed model. The specific parameters of this system are shown in [31]. As a typical representative of renewables, photovoltaic panels (PVs) are chosen in this paper. Assume that the proper candidate locations for PVs and ESSs of the IEEE-33-node system are illustrated in Fig.2. and the specific

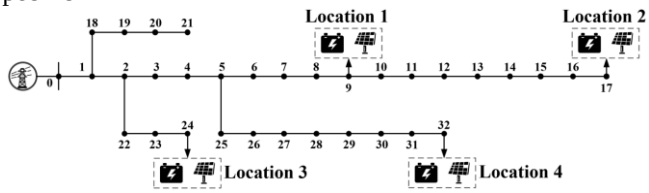

Fig. 2. Modified IEEE 33-node system.

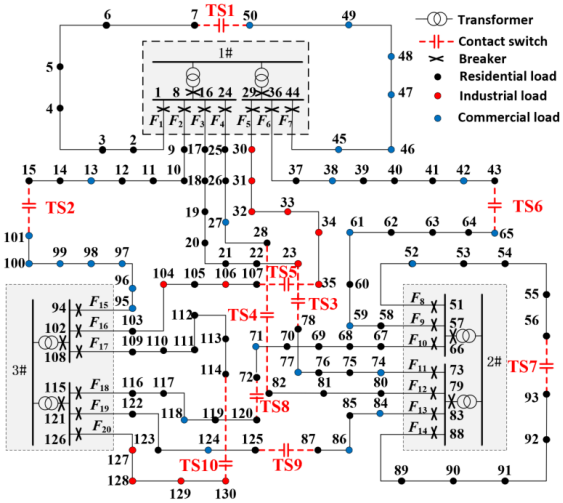

Fig. 3. 130-node system from northern China.

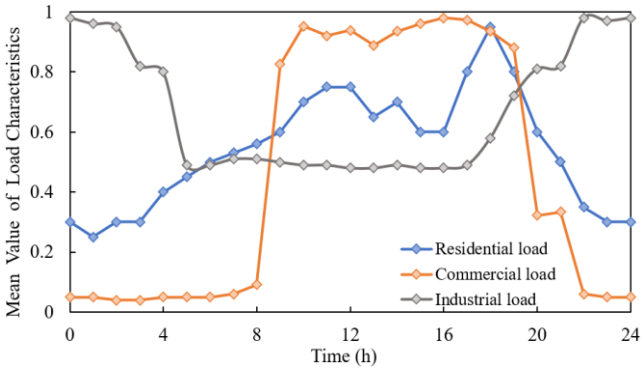

Fig. 4. Mean value of different load characteristics.

capital cost is presented in [15]. Meanwhile, in Fig.3, a practical system configuration with 130 nodes from northern China [32] is selected to discuss the applicability of the proposed model in the real world. The line capacity of this system is 2MW. The historical samples about PVs outputs are derived from [33] and the mean values of different load characteristics are presented in Fig.4. The carbon trading price is set as 15 \$/tonCO_2 [28].

All experiment is conducted on an Intel-i7 computer with 16GB RAM and 3.4 GHz basic frequency. All simulations are performed by the CPLEX solver on the YALMIP platform. The confidence levels β_1 and β_∞ of the Wasserstein metric constraints are set to 90% and 70%.

B. 33-node System Test

1) *Planning Strategies under Different External Factors*

To demonstrate the significance of external factors such as LMP and NCI, the following two districts are considered.

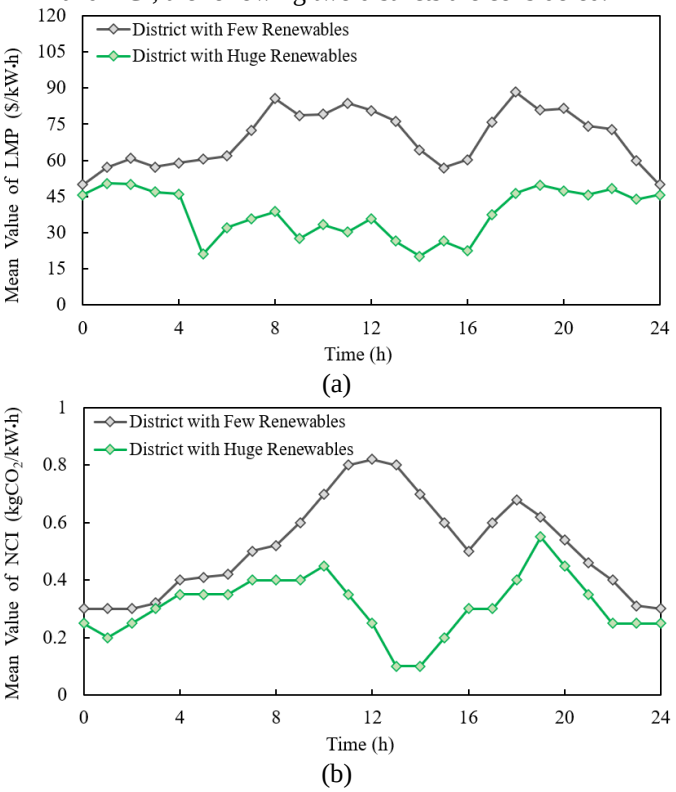

Fig. 5. Mean value of LMP and NCI in different time and districts.

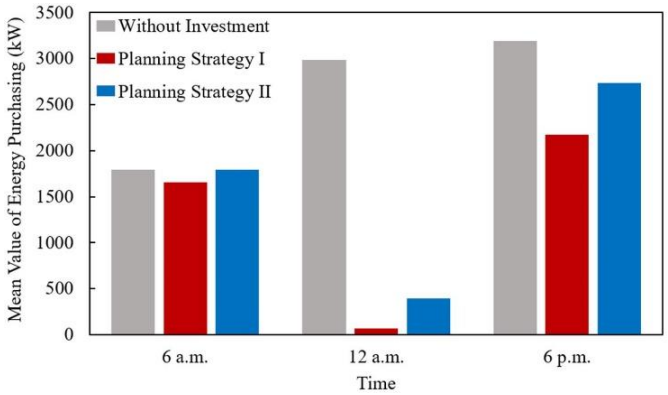

Fig. 6. Mean value of energy purchasing with different planning strategies.

Data samples from the residential load are chosen in this test.

Planning strategy I: Install REGs and ESSs in District I which has few renewables so the LMP and NCI are relatively higher.

Planning strategy II: Install REGs and ESSs in District II which has huge renewable so that the LMP and NCI are relatively lower.

Fig.5(a) and (b) illustrate the mean values of LMP and NCI in District I and District II respectively. The optimal planning strategies for PVs and ESSs in the two districts are shown in Table I, and the annual costs of different planning strategies are shown in Table II.

Comparing the planning strategies for PVs and ESSs in Table I, the total investment capacity of PVs in District I is 850 kW more than that in District II, and the optimal scheme

TABLE I
PLANNING STRATEGIES IN DIFFERENT DISTRICTS

District	Location number	PV capacity (kW)	ESS capacity (kW·h)	ESS power (kW)
District I	9, 17	1300, 450	400, 300	100, 100
	24, 32	1100, 900	400, 400	100, 100
District II	9, 17	800, 400	0, 0	0, 0
	24, 32	900, 800	0, 0	0, 0

TABLE II
ANNUAL COST OF DIFFERENT PLANNING STRATEGIES

District	Planning strategy	Annual cost ($10^4$$)				
		C_{DG}	C_{ES}	C_E	C_{CO2}	F
District I	Without planning	0	0	136.3	12.9	149.2
	Planning strategy I	27.5	3.5	71.7	7.6	110.3
District II	Without planning	0	0	70.4	8.9	79.3
	Planning strategy II	21.4	0	50.1	5.8	77.3

in District I also needs further investment in ESSs. Meanwhile, more benefits will be obtained by DSO in district I if the planning strategy for PVs and ESSs is adopted. Essentially, these phenomena have a close relationship with the development of local renewables, reflected in the LMP and NCI from the main grid. If there are few renewables, such as District I, the LMP and NCI are relatively higher and the DSO will create more benefits by investing in PVs and ESSs. By contrast, only few benefits can be obtained by the investor. Different planning schemes and economic profits convincingly demonstrate that the DSO should attach great importance to the interaction uncertainty from the main grid, such as LMP and NCI, which have an enormous impact on the optimal decision-making.

Furthermore, Fig.6 presents the mean value of energy purchased from the main grid in typical time when different planning strategies are implemented in the distribution network. At 12 am, energy purchase from the main grid significantly decreases in strategies I and II because of PVs outputs. At 6 a.m. and 6 p.m., strategy I can realize less energy purchase than strategy II for the effectiveness of ESSs. Therefore, the coordinated planning of REGs and ESSs can reduce the reliance on the main grid and prompt the autonomy capacity of the distribution network.

2) *Benefits Analysis of Wasserstein DRO*

In this section, District I with few renewables is chosen to verify the effectiveness of the Wasserstein DRO. The original problem is optimized by four different methods, SO, RO, traditional DRO and Wasserstein DRO. The optimal comprehensive costs of these methods are presented in Fig.7. In SO, the decision-maker utilizes the finite data samples directly to obtain the optimal planning strategy. RO and traditional DRO

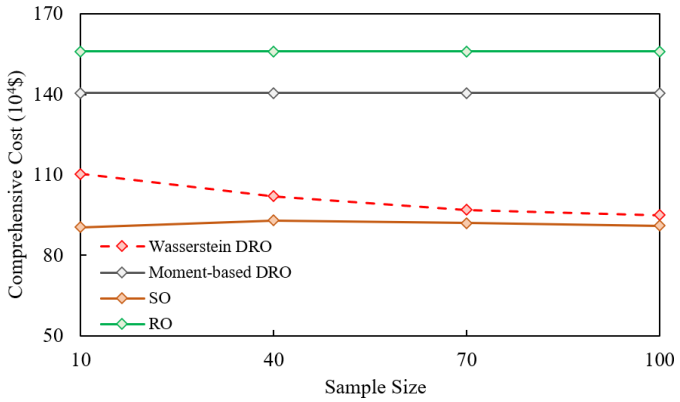

Fig. 7. Comprehensive cost of four methods with different sample size.

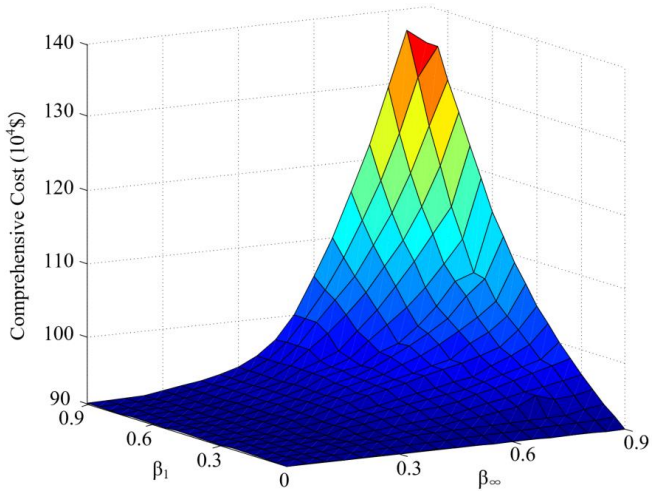

Fig. 8. Comprehensive cost of different confidence level

often models uncertain factors by the uncertain interval and moment-based ambiguity set respectively. In Wasserstein DRO, a novel ambiguity set with Wasserstein metric is formulated on basis of data samples.

Apparently, the optimal value of SO is the least conservative and it tends to be stable with the increase in historical data. However, when data samples are not adequate enough, such as N=10, underlying scenarios of uncertain variables are not fully considered. In this case, although this method has the least comprehensive cost, the robustness of the planning scheme and security of network operation cannot be guaranteed. On the contrary, RO has great robustness for it only considers the worst-case scenario and then generates the most conservative planning strategy. Both traditional DRO and Wasserstein DRO are less conservative than RO because these methods consider the probability of the worst-case scenario. However, the objective value of the traditional DRO with the moment ambiguity set is constant with the increase in historical samples. It is because the moment information is not adequate enough to describe the true probability distributions. By comparison, the Wasserstein DRO can reduce the optimal value of decision-making gradually with the increase in sample size. The true probability distribution hiding behind the historical data can be well-characterized by the Wasserstein ambiguity set. Simultaneously, improbable distributions can be excluded with the increasing data samples, resulted in a high confidence level. In a nutshell, the Wasserstein DRO not only combines the advantages of SO and RO, but also overcomes the defects of traditional DRO, especially when multiple uncertainties occur in the decision-making.

Furthermore, both 1-norm Wasserstein metric and ∞-norm Wasserstein metric are considered in the proposed WDRO model. As is shown in Fig.8, the objective value increases with the rising confidence level of both Wasserstein metric constraints. It is because that higher confidence level will expand the size of the Wasserstein ambiguity set and enhance the robustness of the planning strategy. Note that if the confidence level of ∞-norm Wasserstein metric is quite low, the objective value is almost the same as the variation in the confidence level of 1-norm Wassertein metric. This phenomenon demonstrates that ∞-norm Wasserstein metric is less conservative than 1-norm Wasserstein metric under the same confidence level. Essentially, 1-norm Wasserstein metric is related to the general difference between probability distributions and ∞-norm Wasserstein metric represents the local deviation between probability distributions. The proposed method combines them together rather than considering only one Wasserstein metric. Hence, it can depict the ambiguity set from two dimensions comprehensively to satisfy the decision maker's preference in different aspects and obtain a less conservative result.

3) *Effectiveness Analysis of Solution Methodology*

In this section, the performance of the proposed solution methodology is validated. The following different methods are used by comparing the absolute error, relative error and solution time. For simplification, assuming that only the interaction uncertainty from the main grid and 1-norm Wasserstein metric constraint is considered.

Method I: The dual problem is formulated on basis of uncertain variables. Then, the original model is transformed into a min-min optimization problem with second-order cone constraint which can be solved by the commercial optimizer directly. Note that this dual problem is much easier to construct because it only has 1-norm Wasserstein metric constraint about uncertain variables.

Method II: Through the linearized power flow, the dual problem is formulated on basis of operation variables and then the original model is transformed into a min-max optimization which can be solved by the CCG algorithm.

Method III: Utilizing the optimal probability distribution from Method II, a multi-scenario optimization with the known distribution is formulated. Essentially, it is also a MISOCP model without any difficulties to solve.

In Table III, the performance of the aforementioned methods is presented. Although the simplification in Method II reduces the difficulty in solving the original Wasserstein DRO effectively, the accuracy of Method II cannot be guaranteed without calculating the network loss. By comparison, Method III overcomes the defect of Method II by reformulating a MISOCP model with the known distribution. Hence, the solution methodology proposed in this paper performs well in solving the Wasserstein DRO with second-order cone constraint because it enhances the accuracy of the model solution without sacrificing efficiency.

TABLE III
PERFORMANCE COMPARISON IN SOLUTION METHODS

Solution method	Objective value (10^4\$)	Absolute error (10^4\$)	Relative error	Time (s)
Method I	102.50	0	0	10s
Method II	100.26	2.24	2.19%	33s
Method III	102.47	0.03	0.03%	35s

TABLE IV
OPTIMAL PLANNING SCHEME OF TYPICAL LINES IN 130-NODE SYSTEM

Line Number	Original load factor	Load Type	Location and capacity (kW)	
			PV	ESS
F_1	50%	Resident	5, 800	5, 300
F_5	50%	Industry	32, 600	32, 300
F_7	50%	Commerce	47, 1000	47, 200
F_{20}	60%	Industry	129, 700	129, 500
F_2	50%	Combined resident and commerce	13, 850	13, 300

C. 130-node System Test

In this section, a practical distribution system with 130 nodes is selected to verify the effectiveness of the proposed method in the real world. The optimal planning schemes for PVs and ESSs in several typical lines are shown in Table IV and specific analysis is as follows.

First, PVs and ESSs often have the same optimal location which is approximate to the middle of each distribution network. It is because the same location can not only avoid the network loss in storing PV outputs, but also reduce the excessive PV power injection which may lead to voltage violation in the network. Meanwhile, siting in the middle of each distribution network is beneficial to make a further reduction in the energy loss.

Then, comparing F_1, F_2, F_5 and F_7, there is a huge difference in planning schemes for PVs and ESSs under different load characteristics. More specifically, if the load characteristic is more analogous to PV outputs, such as commercial load, the optimal capacity for PVs will increase while that for ESSs will decrease. Generally speaking, DSO's investment for PVs can decline energy consumption cost in distribution systems obviously due to the extremely low marginal cost in PVs. The decision-maker tends to allocate as many PVs as possible until it reaches the limits of system security constraints. By contrast, ESSs are hardly ever allocated alone because of the relatively higher cost. Hence, the planning scheme can generate more benefits if the load characteristic is more appropriate for taking advantage of PV outputs.

In addition, comparing planning schemes in line F_5 and F_{20}, the optimal allocation capacity of PVs and ESSs is larger when the original loads are heavier. It is easy to imagine that the decision-maker will allocate more PVs under the heavier loads for economic benefits. However, load factor constraint tends to be the main factor impacting ESS capacity rather than the economy. To support the first circuit outage in each line with one contact switch, the maximum load factor cannot exceed 50%, while the original load factor of line F_{20} has reached 60%. Note that the peak load of industry occurs at night so that PVs allocation cannot tackle this issue. In this case, a larger capacity of ESS is needed to satisfy the load factor constraint.

VI. CONCLUSION

This paper expounded a two-stage WDRO model for REGs and ESSs planning in the distribution network where both internal uncertainty and interaction uncertainty were considered. This model adopts 1-norm and ∞-norm Wasserstein metric constraints comprehensively to reflect the decision maker's preference in different aspects. Then, a three-step solution method was developed to solve this WDRO with the second-order cone constraint. Case studies are performed in a modified IEEE 33-node system and a practical 130-node system. Numerical results represent the significance of considering the internal and external influence comprehensively in the decision-making since load characteristic, LMP and NCI vary in different districts and these factors often lead to different optimal planning schemes. Furthermore, results also validate that the proposed WDRO can combine the advantages of SO and RO and provide a less conservative planning scheme on basis of decision-maker's preference from different dimensions. The future work will focus on different REG and ESS investors in the distribution network and introduce game theory in the proposed model to describe the interaction between different investors.

REFERENCES

[1] Y. Wang, J. Qiu, Y. Tao and J. Zhao, "Carbon-oriented operational planning in coupled electricity and emission trading markets," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3145-3157, Jul. 2020.
[2] K. Zou, A. P. Agalgaonkar, K. M. Muttaqi and S. Perera, "Distribution system planning with incorporating DG reactive capability and system uncertainties," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 112-123, Jan. 2012.
[3] F. Conte, S. Massucco, G. Schiapparelli and F. Silvestro, "Day-ahead and intra-day planning of integrated BESS-PV systems providing frequency regulation," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1797-1806, Jul. 2020.
[4] S. Mahdavi, R. Hemmati and M. A. Jirdehi, "Two-level planning for coordination of energy storage systems and wind-solar-diesel units in active distribution networks," *Energy*, vol. 151, pp. 954-965, May. 2018.
[5] Y. Li, B. Feng, G. Li, J. Qi, D. Zhao and Y. Mu, "Optimal distributed generation planning in active distribution networks considering integration of energy storage," *Appl. Energy*, vol. 210, pp. 1073-1081, Jan. 2018.
[6] M. Z. Oskouei, B. Mohammadi-Ivatloo, O. Erdinç and F. G. Erdinç, "Optimal allocation of renewable sources and energy storage systems in partitioned power networks to create supply-sufficient areas," *IEEE Trans. Sustain. Energy*, vol. 12, no. 2, pp. 999-1008, Apr. 2021.
[7] M. R. Jannesar, A. Sedighi, M. Savaghebi and J. M. Guerrero, "Optimal placement, sizing, and daily charge/discharge of battery energy storage in low voltage distribution network with high photovoltaic penetration," *Appl. Energy*, vol. 226, pp. 957-966, Sep. 2018.
[8] T. Adefarati, R.C. Bansal, M. Bettayeb and R. Naidoo, "Optimal energy management of a PV-WTG-BSS-DG microgrid system," *Energy*, vol. 217, Feb. 2021.
[9] H. Golpîra and S. A. R. Khan, "A multi-objective risk-based robust optimization approach to energy management in smart residential buildings under combined demand and supply uncertainty," *Energy*, vol. 170, pp. 1113-1129, Mar. 2019.

[10] A. Ahmadian, M. Sedghi, H. Fgaier, B. Mohammadi-ivatloo, M. A. Golkar and A. Elkamel, "PEVs data mining based on factor analysis method for energy storage and DG planning in active distribution network: Introducing S2S effect," *Energy*, vol. 175, pp. 265-277, May. 2019.
[11] J. Cervantes and F. Choobineh, "Optimal sizing of a nonutility-scale solar power system and its battery storage," *Appl. Energy*, vol. 216, pp. 105-115, Apr. 2018.
[12] L. Li, X. Cao and P. Wang, "Optimal coordination strategy for multiple distributed energy systems considering supply, demand, and price uncertainties," *Energy*, vol. 227, Jul. 2021.
[13] C. Kang, T. Zhou, Q. Chen, J. Wang, Y. Sun, Q. Xia and H. Yan, "Carbon emission flow from generation to demand: a network-based model," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2386-2394, Sep. 2015.
[14] Y. Wang, J. Qiu and Y. Tao, "Robust energy systems scheduling considering uncertainties and demand side emission impacts," *Energy*, vol. 239, Jan. 2022.
[15] R. Hemmati, H. Saboori and P. Siano, "Coordinated short-term scheduling and long-term expansion planning in microgrids incorporating renewable energy resources and energy storage systems," *Energy*, vol 134, pp. 699-708, Sep. 2017.
[16] X. Zhu, J. Yan and N. Lu, "A graphical performance-based energy storage capacity sizing method for high solar penetration residential feeders," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 3-12, Jan. 2017.
[17] E. P. Mohajerin and K. Daniel, "Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, no. 1-2, pp. 115-166, Sep. 2018.
[18] C. Ning and F. You, "Data-driven Wasserstein distributionally robust optimization for biomass with agricultural waste-to-energy network design under uncertainty," *Appl. Energy*, vol. 255, Dec. 2019.
[19] W. Wei, F. Liu and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 289-300, Jan. 2016.
[20] Y. Zhou, Z. Wei, M. Shahidehpour and S. Chen, "Distributionally robust resilient operation of integrated energy systems using moment and wasserstein metric for contingencies," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3574-3584, Jul. 2021.
[21] T. Ding, Q. Yang, Y. Yang, C. Li, Z. Bie and F. Blaabjerg, "A data-driven stochastic reactive power optimization considering uncertainties in active distribution networks and decomposition method," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4994-5004, Sep. 2018.
[22] R. Zhu, H. Wei and X. Bai, "Wasserstein metric based distributionally robust approximate framework for unit commitment," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2991-3001, Jul. 2019.
[23] 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. 33, no. 5, pp. 4924-4936, Sept. 2018.
[24] A. Zhou, M. Yang, M. Wang and Y. Zhang, "A linear programming approximation of distributionally robust chance-constrained dispatch with Wasserstein distance," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3366-3377, Sep. 2020.
[25] X. Zheng and H. Chen, "Data-driven distributionally robust unit commitment with Wasserstein metric: tractable formulation and efficient solution method," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4940-4943, Nov. 2020.
[26] Y. Zhou, Z. Wei, M. Shahidehpour and S. Chen, "Distributionally robust resilient operation of integrated energy systems using moment and Wasserstein metric for contingencies," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3574-3584, Jul. 2021.
[27] Y. Wang, Y. Yang, H. Fei, M. Song and M. Jia, "Wasserstein and multivariate linear affine based distributionally robust optimization for CCHP-P2G scheduling considering multiple uncertainties," *Appl. Energy*, vol. 306, Jan. 2022, Art. no. 118034.
[28] A. Bartolini, S. Mazzoni, G. Comodi and A. Romagnoli, "Impact of carbon pricing on distributed energy systems planning," *Appl. Energy*, vol. 301, Nov. 2021.
[29] F. Nicolas and G. Arnaud, "On the rate of convergence in Wasserstein distance of the empirical measure," *Probab. Theory Relat. Fields*, vol. 162, no. 3-4, pp. 707-738, Aug. 2015.
[30] X. Chen, W. Wu and B. Zhang, "Robust capacity assessment of distributed generation in unbalanced distribution networks incorporating ANM techniques," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 651-663, Apr. 2018.
[31] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Trans. Power Del.*, vol. 4, no. 2, pp. 1401-1407, Apr. 1989.
[32] H. Liu, J. Li, S. Zhang, S. Ge, B. Yang and C. Wang " Distributionally robust co-optimization of the demand-side resources and soft open points allocation for the high penetration of renewable energy," *IET Renew. Power Gener.*, vol. 16, pp. 713-725, March 2022.
[33] C. Wang, G. Song, P. Li, H. Ji, J. Zhao and J. Wu, "Optimal siting and sizing of soft open points in active electrical distribution networks," *Appl. Energy*, vol. 189, pp. 301-309, Mar. 2017.

Junkai Li (Student Member IEEE) received the M.S. degree in 2021 from the School of Electrical and Information Engineering, Tianjin University, Tianjin, China, where he is currently working toward the Ph.D. degree with the School of Electrical and Information Engineering in Tianjin University, Tianjin, China. His researh focuses on the optimal planning and design of distribution system.

Hong Liu (Member, IEEE) received M.S. and Ph.D. degrees from the School of electric automation engineering in Tianjin University, Tianjin, China, in 2005 and 2009, respectively. He is currently a Professor with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China. His research interests include the optimal planning and operation of distribution system and integrated energy system.

Zhengyang Xu received B.S., M.S. and Ph.D. degrees from the School of Electrical and Information Engineering in Tianjin University, Tianjin, China. Currently, he is a lecturer with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China. His research interests is in distribution system evaluation and planning.