A Customer-Centric Distributed Data-Driven Stochastic Coordination Method for Residential PV and BESS

Huichuan Liu, Jing Qiu, *Senior Member, IEEE*, Junhua Zhao, *Senior Member, IEEE*, Yuechuan Tao, and Zhao Yang Dong, *Fellow, IEEE*

***Abstract*—An aggregation scheme is an effective transactive manner of Distributed Energy Resources (DER) spreading across distribution networks. Distributed approach locally achieves cost minimization of an aggregator and customers. The uncertainties of wholesale market price and rooftop PV output will impact on aggregator's scheduling decision and each customer's cost, while solar energy fluctuation can cause an overvoltage problem in distribution networks. However, the probability distributions of these uncertainties always have errors, even in emerging data-based methods. There is no stochastic method using real data with an out-of-sample guarantee suitable for this distributed approach so far to help an aggregator avoid price risk and manage customers' energy against solar energy fluctuation. To address these unsolved issues, we propose a data-driven Wasserstein distributionally robust formulation of the aggregator's agent and customer's agent respectively. The Wasserstein metric is employed to construct the Wasserstein ambiguity set. The mathematical models are then reformulated equivalently to convex programming respectively so that the operating model can be solved by the off-the-shelf solver. To improve the efficiency of the distributed solving framework, an alternating optimization procedure (AOP) process is proposed to overcome the issue caused by binary variables in the alternating direction method of multipliers (ADMM). The proposed operation framework is verified on the modified IEEE 33-bus distribution network and realistic single-feeder LV network.**

***Index Terms*—Data-driven, residential PV and BESS, stochastic optimization, wasserstein metric.**

Manuscript received 17 June 2022; revised 27 September 2022 and 18 November 2022; accepted 28 November 2022. Date of publication 6 December 2022; date of current version 20 October 2023. This work was supported in part by the Australian Research Council Research Hub under Grant IH180100020, in part by the ARC Training Centre under Grant IC200100023, in part by the ARC Linkage Project under Grant LP200100056, in part by the ARC under Grant DP220103881, in part by the Shenzhen Institute of Artificial Intelligence and Robotics for Society, in part by the National Natural Science Foundation of China (Key Program), under Grants 71931003 and 72061147004, and in part by the National Natural Science Foundation of China under Grant 72171206. Paper no. TPWRS-00892-2022. *(Corresponding authors: Jing Qiu; Junhua Zhao.)*

Huichuan Liu, Jing Qiu, and Yuechuan Tao are with the School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: huichuan.liu@sydney.edu.au; qiujing0322@gmail.com; ytao5383@uni.sydney.edu.au).

Junhua Zhao is with the School of Science and Engineering and the Shenzhen Institute of Artificial Intelligence and Robotics for Society, The Chinese University of Hong Kong, Shenzhen 518100, China (e-mail: junhua.zhao@outlook.com).

Zhao Yang Dong is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: zydong@ieee.org).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRS.2022.3227178.

Digital Object Identifier 10.1109/TPWRS.2022.3227178

Nomenclature

A. Indices and Sets

$\mathcal{A}$	Set of all customers.
$\mathcal{V}$	Set of all buses in the distribution network.
$\mathcal{E}$	Set of all branches in the distribution network.
$\mathcal{T}$	Set of time slots.
$\mathcal{X}_a$	Set of all variables of customer agent.
$\mathcal{Z}$	Set of all variables of aggregator agent.

B. Parameters

c^{FiT}	Feed-in tariff.
c^{ToU}	Tariff.
$\boldsymbol{\mu}^{Price}$	Price vector of wholesale market.
$\widehat{\boldsymbol{\mu}}^{Price}$	Sample of price vector of wholesale market.
Δ^t	Time-step resolution.
$\overline{V}_i, \underline{V}_i$	Max/Min limit of voltage magnitude at the bus i.
$\overline{p}_t^{agg}, \underline{p}_t^{agg}$	Max/min limit of aggregator's output.
p_a^L, q_a^L	Active and reactive load of the customer a at bus *i*.
$\overline{p}_a^g$	Max limit of the active power of customer a's feeder.
$\eta_a^{ch}, \eta_a^{dis}$	Charging/discharging efficiency of user a's battery.
η_a^{inv}	Inverter efficiency of customer a's PV system.
$\overline{p}_a^{ch}, \overline{p}_a^{dis}$	Max charge/discharge power of customer a's battery.
$\overline{e}_a^b, \underline{e}_a^b$	Max/min energy state of customer a's battery.
$p_{a,t}^{PV,f}$	Forecasting power of customer a's PV.
$\overline{p}_{a,t}^{PV,f}$	Forecasted maximum dispatchable power of customer a's PV.
r_{ij}, x_{ij}	Resistance/reactance of branch (*i, j*).
ρ	ADMM penalty parameter.
σ	The weighting factor for the incorporation of risk.
ε	Wasserstein radius.
T	Time horizon.

C. Variables

$\boldsymbol{p}^{agg}$	Power output vector of aggregator.
p_a^{PV}	Dispatch power of customer a's PV.
p_a^{grid}	Net active power injection of the customer a at bus *i*.
$\widehat{p}_i^{grid}$	Duplicate variable of p_a^{grid} at bus *i*.
p_a^{g+}, p_a^{g-}	Power flowing from/to grid of customer a.
p_a^{b+}, p_a^{b-}	Battery charge/discharge power of customer a.
$p_a^{b,g}$	Power flowing from battery to grid of customer a.
$p_a^{b,d}$	Power flowing from the battery to the demand of the customer a.
$e_{a,t}^b$	Energy state of customer a's battery.
$e_{a,initial}^b, e_{a,final}^b$	Initial and final energy state of customer a's battery over the optimization horizon.
κ_a^{pv}	Proportion of PV power flow.
d_a^g	Binary variable of power flowing direction of customer a.
χ_a^b	Binary value of battery state of customer a.
p_{ij}, q_{ij}	Active/reactive power flow from the bus *i* to *j*.
v_j	Squared voltage magnitude of bus *i*.
l_{ij}	Squared current magnitude of branch (*i, j*).
λ_a	Dual variables.
τ	Auxiliary variable in CVaR.
$\lambda_0, g_i, \boldsymbol{\gamma}_{ij}$	Dual variables.

D. Functions

$C_a(\boldsymbol{x}_a)$	The cost of customer a
$C_0(\boldsymbol{z})$	The cost of aggregator.
$\mathbb{E}^{\mathbb{P}}[]$	Mathematical expectation.
$\delta_{(\boldsymbol{\cdot})}$	Dirac measure.
$CVaR[]$	Conditional Value at Risk.
$\max\{\cdot,\cdot\}$	Taking the larger one of two values.
$\max\{\cdot,0\}$	Taking the larger one compared with zero.
$\mathbb{P}$	Probability distribution.
$\widehat{\mathbb{P}}_N$	Empirical distribution.
$\mathbb{B}_\varepsilon(\widehat{\mathbb{P}}_N)$	Ambiguity set.
$\mathbb{B}_{\varepsilon_{pv}}(\widehat{\mathbb{P}}_N^e)$	Ambiguity set of PV forecast error distribution.
$\mathcal{D}(\widehat{\mathbb{P}}_N, \mathbb{P})$	Wasserstein metric between $\widehat{\mathbb{P}}_N$ and $\mathbb{P}$.
$\Gamma()$	Joint probability distribution.
$\widehat{P}_{Ni}$	Marginal distribution of $\widehat{\boldsymbol{\mu}}_i^{Price}$.
P_h	Marginal distribution of $\boldsymbol{\mu}_h^{Price}$.

I. Introduction

EXTREME climate change caused by greenhouse gas emissions has become a global challenge for humans. To cope with the issue, the power sectors in the world are experiencing an energy transition by increasing the penetration of sustainable energy to substitute for fossil fuels. Distributed energy resources (DER) in distribution networks (i.e., rooftop PV and battery storage systems) have huge potential to reduce emissions. Many countries are positively integrating these DER into energy markets. For example, Australia is constructing the distribution level market to integrate DER into the wholesale energy market and it is expected to be completed after 2022 [1]. As a result of the characteristics of a large number and spatial dispersion, each customer's DER can hardly participate in the wholesale energy market. One way to effectively prompt the integration is to aggregate DER in a coordination approach as a single entity traded in the market, which is so-called transactive energy [2]. The typical aggregation paradigm comprises one aggregator and a set of end-customers, each with DER managed by an individual home energy management system (HEMS) [3]. The aggregator usually plays a role in interacting with each HEMS to schedule aggregated port power in offers to the market based on the information about market prices and each customer's DER output forecast.

However, the operation of DER aggregation brings challenges to all stakeholders. When the aggregator schedules the aggregated port power with its objective for trade, the decision must be simultaneously aligned with each customer's objective of minimizing individual costs through HEMS. On the other hand, the market price and PV forecast error are uncertain for the aggregator and all customers. The uncertain market price will affect the aggregator's profit earned from the market while the PV fluctuation not only affects the customer's cost but also causes an overvoltage issue due to excessive PV feed-in. Thus, an efficient and flexible operation approach is desperately required for the DER aggregation problem.

In response to these challenges, the virtual power plant is initially explored and implemented as an efficient means of aggregating DER regardless of the network constraints [4]. [5] and [6] cast the aggregated DER scheduling problem as a centralized single problem that aligns the objective of individual customers with the objective of the aggregator. The aggregator aims to optimize the market-based interface [7] while end-customers pursue some properties as their objective such as satisfaction and comfort [8], end-customer preferences [9] or the customer's electricity cost [10]. Considering the overvoltage problem, the optimal power flow (OPF) is introduced into the aggregation model with the network constraints. To reduce the computational burden, a distributed fashion is proposed to efficiently solve the aggregation scheme [11]. The coordination in a distributed manner is completed through the interaction between the aggregator and customer agents by individually solving the decomposed subproblems [12]. Some distributed optimization techniques have been applied in OPF [13], [14], [15]. ADMM as an augmented Lagrangian-based method has been the most popular tool applied in the distributed coordination of the DER aggregation problem [16], [17], [18].

To address the impacts caused by uncertainties, stochastic and robust models are adopted in the centralized coordinated DER aggregation. According to the different types of impacts caused by price and forecast error, the stochastic optimization (SO) model is employed to cope with price uncertainty [19] while the robust optimization (RO) model or the chance constraints based SO is adopted against the PV fluctuation [20], [21]. A joint method of the stochastic adaptive robust model is proposed in

[22] for deciding the VPP offering decisions. Nevertheless, the defects of the stochastic and robust models are obvious. The SO relays on the accurate description of the probability distribution, but to reduce the computation burden, the scenarios reduction technique is adopted [23] in the stochastic model. Even though the full empirical distribution is used, the decision still has a poor out-of-sample performance [24]. Furthermore, the robust model leads to an overly conservative solution. The distributionally robust optimization (DRO) model as an alternative stochastic paradigm deals with uncertainties to overcome the defects of above both models, and it was recently applied in OPF [25], [26], [27], economics dispatch [28] and scheduling [29], [30] problem. The idea of DRO aims to solve the decisions by optimizing the expected objective under the worst-case distribution within an ambiguity set. According to the existing applications of DRO in the literature, three types of ambiguity sets are adopted in the renewable energy field: moment-based [31], [32], [33], KL divergence-based [34] and data-driven Wasserstein metric-based [35], [36]. Among these ambiguity sets, the data-driven Wasserstein ambiguity set performs more advantages than others. The moment-based and KL divergence-based DRO models must be approximately reformulated. The final reformulation of data-driven Wasserstein DRO (DWDRO) is not only more concise but also linear. Better yet, the reformulation of DWDRO is fully equivalent to the original problem. In addition, since the stochastic model including the DWDRO depends on the probability distribution, the Vale-at-risk (VaR) and Conditional Vale-at-risk (CVaR) are adopted as risk measures to control the amount of probability tail risk of the optimization objective or violating the chance constraints [37], [38], [39]. The CVaR is also be applied in stochastic game theory for DER trading. The payoff function of each player is formulated as the sum of the corresponding expected form and the CVaR [40], [41].

Through reviewing the previous works, all the SO and RO methods are applied in a single centralized DER aggregation problem. Even though the distributed manner is successfully adopted in the scheduling of the DER aggregation problem to reduce the solution burden, the impacts of uncertainties are neglected in the transformation from the single centralized form into the distributed fashion. After directly introducing the existing SO or RO methods into the distributed framework, the efficiency of the distributed manner is running in the opposite direction to the original purpose of reducing the computation burden. If the SO model is adopted, the chance constraints in each customer's subproblem need to complex approximate reformulation and add more constraints corresponding to many scenarios, whereas the RO model requires iterative algorithms such as Benders Decomposition [42] or Column-and-Constraint Generation [43] to solve the subproblem locally by the agent. It is expected to find an efficient, easily tractable and flexibly adjustable stochastic method suitable for the distributed manner in the DER aggregation scheme.

In this paper, to address the above challenges, a customer-centric distributed data-driven stochastic coordination method is proposed for the decision-making of residential PV and BESS aggregation. According to the type of uncertainties in agents of the aggregator and customers, corresponding data-driven Wasserstein distributionally robust models are proposed respectively. The customer's DER are aggregated in a stochastically distributed fashion. The detailed contributions of this paper are outlined as follows:

1) Data-driven distributionally robust models with Wasserstein ambiguity set of the aggregator agent and customer ones are constructed respectively for the aggregation scheduling problem. For the aggregator agent, a distributionally robust objective is proposed in the form of a piecewise affine function combining expected aggregator operation cost and corresponding CVaR. The distributionally robust chance constraint of PV output is proposed for each customer agent.
2) The aggregator agent is equivalently reformulated into a linear objective with extra linear constraints. Then, a value-at-risk (VaR) based reformulation is used in the customer agent to determine the maximum available PV power output with a confidence level by utilizing a historical dataset of the PV forecast error percentage and the upper limit of constraint is obtained by optimization offline.
3) Based on the existing distributed optimization framework of the aggregation scheduling problem via replication of the power import on the wire between the distribution network and customers to form equality boundary constraints, an ADMM algorithm with AOP process is employed to solve the problem with a guarantee of convergence.

The remainder of this paper is organized as follows. Section II describes some fundamental mathematical models. Section III presents the detailed DWDRO agent models with their reformulations and the AOP process. Section IV carries out numerical simulations of the proposed scheduling method and demonstrates the results. At last, Section V concludes the whole paper.

II. Fundamental Mathematical Models

In this paper, all the market entities submit the next trading day's bids and offers before 12:30 pm in the day ahead, and the market keeps being pre-cleared by rolling until real-time dispatch to generate real-time Locational Marginal Prices (LMP) [44]. The aggregator participates in the wholesale market via the Distribution System Operator (DSO) as shown in Fig. 1. Each aggregator is the corresponding LMP taker and the price is uncertain for the aggregator when it determines the power output in the day ahead. The aggregator schedules its port power limited by network constraints which are given by DSO while each customer's minimum electricity costs are pursued simultaneously.

A. Deterministic Customer Agent Model

The customers registered in energy aggregation are the end-users with the PV and/or the battery storage system, which are usually residential users in distribution networks. The agent of a customer is responsible for locally optimizing the charging profile of the battery so that the electricity bill cost is minimized, and communicating with the aggregator agent (i.e., Agent 0

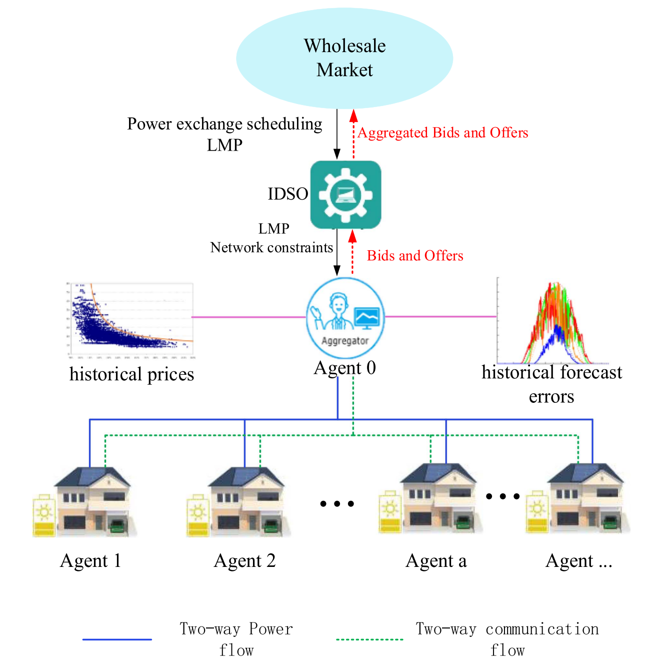

Fig. 1. Market participation mode of customers with DER.

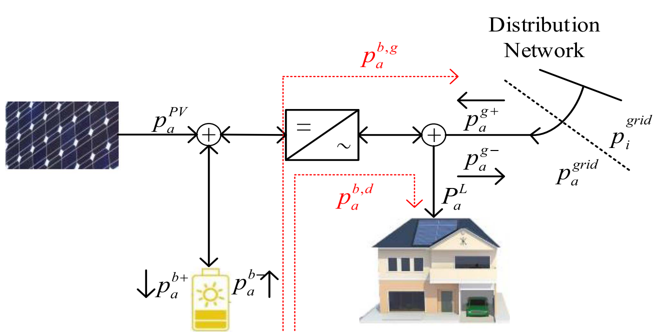

Fig. 2. The agent model of a customer with PV and battery storage.

in Fig. 1) in distributed coordinated approach. In the corresponding distributed coordination manner, the customer agent locally optimizes the decomposed subproblem and iteratively shares the decisions of export with the aggregator agent for the update. In this section, the deterministic customer agent model in the uncoordinated approach is shown. The DWDRO customer agent in distributed coordinated approach will be given in the following sections. Assume that each customer has a certain demand for a particular time slot.

The PV output of each customer can be centrally forecasted by DSO because the end-users in the local distribution network share the same solar radiation. The maximum dispatchable output of each customer's PV can be evaluated according to their capacity and they share the same forecast error percentage.

Let's define an agent $a \in \mathcal{A}\backslash 0$ on behalf of a customer, which will communicate with an aggregator agent ($a = 0 \in \mathcal{A}$) in a distributed manner as shown in Section III-D. The uncoordinated agent model of a customer is an optimization-oriented individual home energy management model [10] as shown in Fig. 2 and it is described as follows:

$$\min C_a = c^{FiT} p^{g+}_{a,t} \Delta^t - c^{ToU} p^{g-}_{a,t} \Delta^t \tag{1}$$

$$s.t. \quad p^{g+}_{a,t} = p^{L}_{a,t} - \eta^{inv}_a \left(p^{b,d}_{a,t} + p^{PV,d}_{a,t} - p^{b+}_{a,t}\right) \tag{2}$$

$$p^{g-}_{a,t} = \eta^{inv}_a \left(p^{b,g}_{a,t} + p^{PV,g}_{a,t}\right) \tag{3}$$

$$p^{PV}_{a,t} = p^{PV,g}_{a,t} + p^{PV,d}_{a,t} \tag{4}$$

$$p^{b-}_{a,t} = p^{b,g}_{a,t} + p^{b,d}_{a,t} \tag{5}$$

$$p^{g}_{a,t} = p^{g+}_{a,t} - p^{g-}_{a,t} \tag{6}$$

$$0 \leq p^{PV}_{a,t} \leq \overline{p}^{PV,f}_{a,t} \tag{7}$$

$$0 \leq p^{g+}_{a,t} \leq \overline{p}^{g}_a d^{g}_{a,t}, 0 \leq p^{g-}_{a,t} \leq \overline{p}^{g}_a \left(1 - d^{g}_{a,t}\right) \tag{8}$$

$$e^{b}_{a,t} = e^{b}_{a,t-\Delta^t} + \eta^{ch}_a p^{b+}_{a,t} \Delta^t - \frac{p^{b-}_{a,t}}{\eta^{dis}_a} \Delta^t \tag{9}$$

$$0 \leq p^{b+}_{a,t} \leq \overline{p}^{ch}_a \chi^{b}_{a,t}, 0 \leq p^{b-}_{a,t} \leq \overline{p}^{dis}_a \left(1 - \chi^{b}_{a,t}\right) \tag{10}$$

$$\underline{e}^{b}_a \leq e^{b}_{a,t} \leq \overline{e}^{b}_a \tag{11}$$

$$e^{b}_{a,t=0} = e^{b}_{a,initial}, e^{b}_{a,t=T} \geq e^{b}_{a,final} \tag{12}$$

For each customer agent $a \in \mathcal{A}\backslash 0$ during the time-slot $t \in \mathcal{T}$, the reference flow directions of all power variables are given in Fig. 2. (1) gives the objective of the agent to minimize the customer's electricity cost. (2)-(5) describe the power balance of a household by defining the power imported $p^{g+}_{a,t}$ and exported $p^{g-}_{a,t}$, and the battery discharging power $p^{b-}_{a,t}$ covers the customer's demand $p^{b,d}_{a,t}$ and the power fed back into the grid $p^{b,g}_{a,t}$; $p^{PV,g}_{a,t}$ and $p^{PV,b}_{a,t}$ denote the proportion of PV output flowing to grid and battery; (6) shows the net active power injection to grid $p^{g}_{a,t}$ derived by the power $p^{g+}_{a,t}$ and $p^{b-}_{a,t}$; (7) denotes the PV output should be dispatched lower than it maximum dispatchable power during time-slot t; the binary variable $d^{g}_{a,t}$ in (8) determines the single power flow state (import or export) between customer and grid; (9)-(12) describes the customer a's battery storage model; (9) shows that relation of energy stored $e^{b}_{a,t}$ in a battery at time t to the charging and discharging power $p^{b+}_{a,t}$ /$p^{b-}_{a,t}$; (10) denotes $p^{b+}_{a,t}$ and $p^{b-}_{a,t}$ should be controlled lower than charging/discharging limitation $\overline{p}^{ch}_a$ and $\overline{p}^{dis}_a$, and the charging/discharging states are denoted by binary variable $\chi^{b}_{a,t}$; (11) and (12) gives the limitation of energy stored in a battery at time t.

B. Deterministic Aggregator Agent Model

The aggregator as an independent entity trades the port power in the pool of the wholesale market and ensures the distribution network voltage within the operating envelope. All customers are physically connected to the AC distribution network via the electricity feeder. The power imported from or exported to the grid is governed by network constraints: the bus voltage magnitudes of distribution networks are strictly required within a specific range [45], [46].

$$\begin{cases} -p_{j,t}^{grid} = \sum_{jk\in\mathcal{E}} \left(p_{jk,t} + l_{jk,t} r_{jk}\right) - \sum_{ij\in\mathcal{E}} p_{ij,t} \\ -q_{j,t}^{L} = \sum_{jk\in\mathcal{E}} \left(q_{jk,t} + l_{jk,t} x_{jk}\right) - \sum_{ij\in\mathcal{E}} q_{ij,t} \end{cases} \tag{13}$$

$$v_{i,t} = v_{j,t} + 2\left(p_{ij,t} r_{ij} + q_{ij,t} x_{ij}\right) + l_{ij,t}\left(r_{ij}^2 + x_{ij}^2\right) \tag{14}$$

$$\left\| \begin{matrix} 2p_{ij,t} \\ 2q_{ij,t} \\ l_{ij,t} - v_{j,t} \end{matrix} \right\|_2 \le l_{ij,t} + v_{j,t} \tag{15}$$

$$\underline{V}^2 \le v_{i,t} \le \overline{V}^2 \tag{16}$$

$$p_{j,t}^{grid} = p_{a,t}^{grid}, q_{j,t}^{L} = q_{a,t}^{L} \tag{17}$$

Equation (13) are the active and reactive power balance constraints of bus j at time-slot t. (14) is the relation between bus j voltage and its connecting branch current. Constraint (15) indicates the second-order cone relaxation equality power flow constraints. (16) gives the lower and upper limits of squared bus voltage magnitude. (17) shows the active and reactive power consumptions at bus j are the power consumed by customer a.

Let's define the agent $a = 0$ on behalf of the aggregator to interface with customer agents to coordinate their DER and make decisions on power output $\boldsymbol{p}^{agg} = [p_1^{agg}, p_2^{agg}, \ldots p_t^{agg} \ldots p_T^{agg}]^\top \in \mathbb{R}^{\mathrm{T}}$ over the scheduling horizon. The power can be bought by the aggregator from the grid for $p_t^{\mathrm{agg}} > 0$ and sold to gird for $p_t^{\mathrm{agg}} < 0$. The aggregator output is limited as shown in constraints (18) and (19).

$$\underline{p}^{\mathrm{agg}} \le p_t^{agg} \le \overline{p}^{agg} \tag{18}$$

$$\sum_{i=1,\forall j} p_{ij,t} = p_t^{\mathrm{agg}} \tag{19}$$

The objective cost function C_0 of the aggregator agent can be derived by the inner production of $\boldsymbol{p}^{agg}$ and $\boldsymbol{\mu}^{Price}$ as shown in (20), where $\boldsymbol{\mu}^{Price} = [\mu_1^{Price}, \mu_2^{Price}, \ldots \mu_t^{Price} \ldots, \mu_T^{Price}]^\top \in \mathbb{R}^T$ is the vector of LMP where the aggregator locates over the scheduling horizon T.

$$\min_{(13)-(19)} C_0 = \sum_{t\in\mathcal{T}} \mu_t^{Price} p_t^{agg} \Delta^t = \left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg} \Delta^t \tag{20}$$

C. Wasserstein Ambiguity Set

The wholesale market price vector $\boldsymbol{\mu}^{Price} \in \mathbb{R}^T$ is uncertain for the aggregator in the day ahead. We can construct an ambiguity set to contain a true distribution of a random vector. Assuming that the true probability distribution $\mathbb{P}$ is supported on $\Xi = \{\boldsymbol{\mu}_1^{Price}, \boldsymbol{\mu}_2^{Price}, \ldots \boldsymbol{\mu}_h^{Price} \ldots, \boldsymbol{\mu}_M^{Price}\}$ and the sample set of $\boldsymbol{\mu}^{Price}$ is $\widehat{\Xi}_N = \{\widehat{\boldsymbol{\mu}}_1^{Price}, \widehat{\boldsymbol{\mu}}_2^{Price}, \ldots \widehat{\boldsymbol{\mu}}_i^{Price} \ldots, \widehat{\boldsymbol{\mu}}_N^{Price}\}$. Based on N independent and identically distributed training samples in $\widehat{\Xi}_N$, the uniformly discrete empirical distribution $\widehat{\mathbb{P}}_N$ can be obtained as follows:

$$\widehat{\mathbb{P}}_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\widehat{\boldsymbol{\mu}}_i^{Price}} \tag{21}$$

where $\delta_{\widehat{\boldsymbol{\mu}}_i^{Price}}$ is Dirac measure on $\widehat{\boldsymbol{\mu}}_i^{Price}$. $\widehat{\mathbb{P}}_N$ becomes similar to $\mathbb{P}$ for large N. The dataset $\widehat{\Xi}_N$ can be regarded to be governed by $\widehat{\mathbb{P}}_N$. Therefore, the ambiguity set can be constructed as a ball in probability distributions space employing a probability metric as follows:

$$\mathbb{B}_\varepsilon\left(\widehat{\mathbb{P}}_N\right) = \left\{\mathbb{P} \mid \mathcal{D}\left(\widehat{\mathbb{P}}_N, \mathbb{P}\right) \le \varepsilon\right\} \tag{22}$$

The ambiguity set $\mathbb{B}_\varepsilon(\widehat{\mathbb{P}}_N)$ is a family of all distributions centered at $\widehat{\mathbb{P}}_N$ within radius ε with a confidence. In this paper, we take the Wasserstein metric $\mathcal{D}(\widehat{\mathbb{P}}_N, \mathbb{P})$as the distance function between two distributions, and its formulation is given as follows:

$$\mathcal{D}\left(\widehat{\mathbb{P}}_N, \mathbb{P}\right) = \min_{\Gamma} \int_{\Xi^2} \left\|\widehat{\mathbb{P}}_N - \mathbb{P}\right\| \Gamma\left(d\widehat{\mathbb{P}}_N, d\mathbb{P}\right) \tag{23}$$

where Γ is the joint distribution of $\boldsymbol{\mu}^{Price}$ and $\widehat{\boldsymbol{\mu}}^{Price}$ with respect to their marginals $\mathbb{P}$ and $\widehat{\mathbb{P}}_N$. For discrete support set, $\mathcal{D}(\widehat{\mathbb{P}}_N, \mathbb{P})$ can be reformulated as follows:

$$\mathcal{D}\left(\widehat{\mathbb{P}}_N, \mathbb{P}\right) = \min_{\Gamma} \left\{ \sum_{h=1}^{M} \sum_{i=1}^{N} \left\|\widehat{\boldsymbol{\mu}}_i^{Price} - \boldsymbol{\mu}_h^{Price}\right\| \cdot \Gamma_{i,h} \;\middle|\; \begin{matrix} \sum_{h=1}^{M} \Gamma_{i,h} = \widehat{P}_{Ni}, \forall i \\ \sum_{i=1}^{N} \Gamma_{i,h} = P_h, \forall h \end{matrix} \right\} \tag{24}$$

Replacing the $\mathcal{D}(\widehat{\mathbb{P}}_N, \mathbb{P})$ in (22) by (24), the Wasserstein ambiguity set can be formulated as follows:

$$\mathbb{B}_\varepsilon\left(\widehat{\mathbb{P}}_N\right) = \left\{ \mathbb{P} \;\middle|\; \begin{matrix} \min_{\Gamma} \sum_{h=1}^{M} \sum_{i=1}^{N} \left\|\widehat{\boldsymbol{\mu}}_i^{Price} - \boldsymbol{\mu}_h^{Price}\right\| \cdot \Gamma_{i,h} \le \varepsilon \\ \sum_{h=1}^{M} \Gamma_{i,h} = \widehat{P}_{Ni}, \forall i \\ \sum_{i=1}^{N} \Gamma_{i,h} = P_h, \forall h \end{matrix} \right\} \tag{25}$$

The radius ε offers a high confidence level to guarantee that the unknown $\mathbb{P}$ resides inside of $\mathbb{B}_\varepsilon(\widehat{\mathbb{P}}_N)$ [24]. Similarly, the Wasserstein ambiguity set of the distribution $\mathbb{P}^e$ of the PV forecast error $\widetilde{p}_{a,t}^e$ at time t for a customer is represented by $\mathbb{B}_{\varepsilon_{pv}}(\widehat{\mathbb{P}}_N^e)$.

III. DWDRO Agent Model and Scheduling Method

A. Centralized DER Aggregation Problem

The scheduling problem of transactive energy aggregation aims to minimize the total cost of the aggregator and customers, and coordinate its internal customers' DER efficiently subject to the operating envelopes simultaneously. The aggregation problem can be centrally solved by the following deterministic optimization problem:

$$\begin{aligned} &\min \left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg} \Delta^t + \sum_{a\in\mathcal{A}\setminus 0} \sum_{t\in\mathcal{T}} \left(c^{FiT} p_{a,t}^{g+} \Delta^t - c_t^{ToU} p_{a,t}^{g-} \Delta^t\right) \\ &\text{subject to } (1)-(11), (13)-(19) \end{aligned} \tag{26}$$

where the equality constraint (19) denotes the power sold or bought by the aggregator via the port of the distribution network. To formulate compactly, let's define a feasible set $\mathcal{X}_a$ of the agent composed of constraints (2)–(12) for corresponding customer agent $a \in \mathcal{A}\backslash 0$ variables $\boldsymbol{x}_a = \{p_{a,t}^{grid}, p_{a,t}^{g+}, p_{a,t}^{g-}, p_{a,t}^{b+}, p_{a,t}^{b-}, p_{a,t}^{b,g}, p_{a,t}^{b,d}, e_{a,t}^{b}, p_{a,t}^{PV}, p_{a,t}^{PV,g}, p_{a,t}^{PV,d}, d_{a,t}^{g}, \chi_{a,t}^{b}\} \in \mathcal{X}_a$ and a feasible set $\mathcal{Z}$ composed of constraints (13)–(19) for aggregator agent variables $\boldsymbol{z} = \{p_t^{agg}, p_{ij,t}, q_{ij,t}, v_{i,t}, l_{ij,t}, p_{a,t}^{grid}\} \in \mathcal{Z}$. The cost functions of customers and an aggregator in (1) and (20) are expressed by $C_a(\boldsymbol{x}_a)$ and $C_0(\boldsymbol{z})$. The deterministic aggregation problem is shown in the resulting compact form as follows:

$$\min_{\boldsymbol{x}_a \in \mathcal{X}_a, \boldsymbol{z} \in \mathcal{Z}} C_0(\boldsymbol{z}) + \sum_{a \in \mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a) \tag{27}$$

B. Distributionally Robust Form of Aggregator and Customer Agents

The problem (27) is the deterministic form of aggregation scheduling to make decisions of its power output and each DER coordination over the time horizon in the day ahead if the LMP μ_t^{Price} and each customer's maximum dispatchable PV output $\overline{p}_a^{PV,f}$ have been known accurately. However, they both are uncertain when the aggregator makes decisions for bidding. Moreover, the LMP can hardly be forecasted in practice due to unavailable information from all market participants, but the maximum dispatchable PV power can be forecasted based on the environment input information from the weather forecast. On the other hand, the uncertain p_a^{PV} power fed into the distribution network might cause the overvoltage issue while the uncertainty of LMP only impacts the aggregator's cost. In terms of predictability and their corresponding impacts, distributionally robust forms of the aggregation problem are given concerning the aggregator and customers separately.

1) Distributionally Robust Form of Aggregator Agent: Considering the uncertainty of LMP, the objective of the aggregator agent $C_0(\boldsymbol{z})$ is replaced by minimizing the expected cost function $\mathbb{E}^{\mathbb{P}}[C_0(\boldsymbol{z})]$. To avoid the risk raised by price uncertainty in aggregator cost, the conditional value at risk (*CVaR*) is employed together with $\mathbb{E}^{\mathbb{P}}[C_0(\boldsymbol{z})]$ in the objective. The objective of the aggregation problem is thereby shown in (28):

$$\min_{\boldsymbol{z}} \ \mathbb{E}^{\mathbb{P}}\left[\left(\boldsymbol{\mu}^{Price}\right)^{\top} \boldsymbol{p}^{agg} \Delta^t\right] + \sigma \times CVaR\left[\left(\boldsymbol{\mu}^{Price}\right)^{\top} \boldsymbol{p}^{agg} \Delta^t\right] \tag{28}$$

where σ is the parameter denoting a weighting factor for the risk control [47]. To make the following distributionally robust form of aggregation objective tractable, the objective function (28) can be constructed equivalently into a piecewise affine function as shown in (29).

$$\min_{\boldsymbol{z},\tau} \ \mathbb{E}^{\mathbb{P}}\left[\max_{j \le 2}\left\{a_j\left(\boldsymbol{\mu}^{Price}\right)^{\top} \boldsymbol{p}^{agg} \Delta^t + b_j \tau\right\}\right] \tag{29}$$

where $a_1 = \left(1 + \frac{\sigma}{\alpha}\right), a_2 = 1, b_1 = \left(\sigma - \frac{\sigma}{\alpha}\right), b_2 = \sigma$. The detailed derivation of (29) is given in **Appendix I**. The distributionally robust optimization model aims to optimize the expectation with respect to the random vector with the worst-case distribution. The distributionally robust form of aggregator agent is hereby shown in (30). The worst-case distribution of $\boldsymbol{\mu}^{Price}$ is covered by the Wasserstein ambiguity set $\mathbb{B}_\varepsilon(\widehat{\mathbb{P}}_N)$.

$$\min_{\boldsymbol{z},\tau} \left(\sup_{\mathbb{P} \in \mathbb{B}_\varepsilon(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{P}}\left[\max_{j \le 2}\left\{a_j\left(\boldsymbol{\mu}^{Price}\right)^{\top} \boldsymbol{p}^{agg} \Delta^t + b_j \tau\right\}\right]\right) \tag{30}$$

2) Distributionally Robust Form of Customer Agent: Considering the uncertainty of maximum dispatchable PV outputs $\overline{p}_{a,t}^{PV,f}$ of each customer during different time-slots, the value of $\overline{p}_{a,t}^{PV,f}$ hardly is forecasted accurately. The actual maximum dispatchable output power fluctuates centering on $\overline{p}_{a,t}^{PV,f}$ with a forecast error $\widetilde{p}_{a,t}^{e}$. The right side of the inequality constraint (7) can be rewritten into Wasserstein distributionally robust (WDR) chance constraint as follows:

$$\inf_{\mathbb{P}^e \in \mathbb{B}_{\varepsilon_{pv}}(\widehat{\mathbb{P}}_N)} \mathbb{P}^e\left(p_{a,t}^{PV} \le \overline{p}_{a,t}^{PV,f} + \widetilde{p}_{a,t}^{e}\right) \ge 1 - \beta_{pv} \tag{31}$$

The dataset of $\widetilde{p}_{a,t}^{e}$ can be calculated by the production of $\overline{p}_{a,t}^{PV,f}$ and the historical forecast error percentage which is commonly shared by all customers.

C. Reformulation of Aggregator and Customer Agents

The DWDRO form of the aggregator and customer agent models (30) and (31) is intractable by using the off-the-shelf solver. The objective function of the aggregator agent in (30) can be equivalently reformulated into a linear form as shown in (32)-(35):

$$\min_{\boldsymbol{z} \in \mathcal{Z}, \tau \in \mathbb{R}, \lambda_0, g_i, \boldsymbol{\gamma}_{ij}} C_0(\boldsymbol{z}) = \lambda_0 \varepsilon + \frac{1}{N} \sum_{i=1}^{N} g_i \tag{32}$$

subject to:

$$b_j \tau + a_j\left(\widehat{\boldsymbol{\mu}}_i^{Price}\right)^{\top} \boldsymbol{p}^{agg} \Delta^t + \boldsymbol{d}^{\top} \boldsymbol{\gamma}_{ij} - \left(\boldsymbol{W}\, \widehat{\boldsymbol{\mu}}_i^{Price}\right)^{\top} \boldsymbol{\gamma}_{ij} \le g_i \tag{33}$$

$$\left\|\boldsymbol{W}^{\top} \boldsymbol{\gamma}_{ij} - a_j \boldsymbol{p}^{agg} \Delta^t\right\|_{\infty} \le \lambda_0 \tag{34}$$

$$\boldsymbol{\gamma}_{ij} \ge \boldsymbol{0}, \lambda_0 \ge 0 \tag{35}$$

where the LMP vector is limited by $\Xi = \{\boldsymbol{\mu}^{Price}: \boldsymbol{W}\boldsymbol{\mu}^{Price} \le \boldsymbol{d}\}$ since the price in the wholesale market is restricted by price ceiling and floor and $\lambda_0, g_i, \boldsymbol{\gamma}_{ij}$ are dual variables in reformulation. The detailed reformulation is given in **Appendix II**. To ensure that the problem is tractable, the norm can be taken to be 1-norm so that the dual norm in (34) is the infinite norm. Accordingly, the aggregation problem becomes a second-order cone programming problem.

As for the DWDRO customer agent model, a VaR-based reformulation is proposed to separate the uncertainty $\widetilde{p}_{a,t}^{e}$ from decision variables in WDR chance constraints (31) as follows [48]:

$$(31) \Leftrightarrow \sup_{\mathbb{P}^e \in \mathbb{B}_{\varepsilon_{pv}}(\widehat{\mathbb{P}}_N)} \mathbb{P}^e\left(0 \le -\overline{p}_{a,t}^{PV,f} - \widetilde{p}_{a,t}^{e} + p_{a,t}^{PV}\right) \le \beta_{pv} \tag{36}$$

$$\Leftrightarrow \sup_{\mathbb{P}^e \in \mathbb{B}_{\varepsilon_{pv}}(\widehat{\mathbb{P}}_N)} \mathrm{VaR}_{1-\beta_{pv}}\left(-\overline{p}_{a,t}^{PV,f} - \widetilde{p}_{a,t}^{e} + p_{a,t}^{PV}\right) \leq 0 \tag{37}$$

$$\Leftrightarrow \sup_{\mathbb{P}^e \in \mathbb{B}_{\varepsilon_{pv}}(\widehat{\mathbb{P}}_N)} \mathrm{VaR}_{1-\beta_{pv}}\left(-\widetilde{p}_{a,t}^{e}\right) \leq \overline{p}_{a,t}^{PV,f} - p_{a,t}^{PV} \tag{38}$$

Let's define $L_{a,t} = \sup_{\mathbb{P}^e \in \mathbb{B}_{\varepsilon_{pv}}(\widehat{\mathbb{P}}_N^e)} \mathrm{VaR}_{1-\beta_{pv}}(-\widetilde{p}_{a,t}^{e})$, the value of $L_{a,t}$ can be calculated offline by the N_{pv}-sample dataset of $\hat{\widetilde{p}}_{a,t}^{e}$ as shown in the problem (39).

$$\begin{aligned} &\min_{\pi_1 \in \mathbb{R}^{N_{pv}}, \pi_2 \in \mathbb{R}^{N_{pv}}, \pi_3 \in \mathbb{R}} L_{a,t} \\ &s.t. \quad \pi_3 \varepsilon_{pv} + \frac{1}{N_{pv}} \sum_{n_{pv}} \pi_{1,n_{pv}} \leq \beta_{pv} \\ &\pi_{1,n_{pv}} \geq 1 - \pi_{2,n_{pv}}\left(L_{a,t} - \hat{\widetilde{p}}_{a,t,n_{pv}}^{e}\right) \qquad \forall n_{pv} \\ &\pi_3 \geq \pi_{2,n_{pv}}, \pi_{1,n_{pv}} \geq 0, \pi_{2,n_{pv}} \geq 0 \qquad \forall n_{pv} \end{aligned} \tag{39}$$

Therefore, the constraint (7) can be substituted by (40) to maximize the unitality of PV power.

$$0 \leq p_{a,t}^{PV} \leq \overline{p}_{a,t}^{PV,f} - L_{a,t} \tag{40}$$

D. Distributed Coordination Scheduling Manner

Through the above reformulation, the aggregation scheduling problem (27) is a large-scale mixed-integer second-order conic optimization problem. It is an NP-hard problem that is usually intractable when solved centrally. In addition, the centrally solved problem (27) needs to collect private information in $\mathcal{X}_a$. To overcome these challenges above, the aggregation scheduling problem is proposed to be implemented in a decomposition fashion by using the alternating direction method of multipliers (ADMM) [49]. After reformulation of the aggregator agent and customer agent models, let's redefine the feasible set $\mathcal{Z}$ by adding the constraints (33)-(35) and substituting (40) for the constraint (7). It is noticed that the problem (27) is not separatable due to the comment variables $p_{a,t}^{grid}$ existing in both $\boldsymbol{z}$ and $\boldsymbol{x}_a$. To make the problem (27) decomposable and suitable for ADMM, the boundary coupling variables $p_{a,t}^{grid}$ are duplicated to produce additional equality constraints as shown in (41) and in Fig. 2 to achieve the resulting component-based decomposition.

$$p_{a,t}^{grid} = \widehat{p}_{a,t}^{grid} \tag{41}$$

The variables $p_{a,t}^{grid}$ belong to customer agents and $\widehat{p}_{a,t}^{grid}$ belong to the aggregator agent. Let's rewrite aggregator and costumer agent variables vectors $\boldsymbol{z}$ and $\boldsymbol{x}_a$:

$$\boldsymbol{x}_a = \begin{Bmatrix} p_{a,t}^{grid}, p_{a,t}^{g+}, p_{a,t}^{g-}, p_{a,t}^{b+}, p_{a,t}^{b-}, p_{a,t}^{b,g}, \\ p_{a,t}^{b,d}, e_{a,t}^{b}, p_{a,t}^{PV,g}, p_{a,t}^{PV,b}, d_{a,t}^{g}, \chi_{a,t}^{b} \end{Bmatrix} \in \mathcal{X}_a \forall a \in \mathcal{A}\backslash 0 \tag{42}$$

$$\boldsymbol{z} = \left\{p_t^{agg}, p_{ij,t}, q_{ij,t}, v_{i,t}, l_{ij,t}, \widehat{p}_{a,t}^{grid}, \tau, \lambda_0, g_i, \gamma_{ij}\right\} \in \mathcal{Z} \tag{43}$$

The resulting form of aggregation scheduling problem is shown as follows:

$$\begin{aligned} &\min_{\boldsymbol{z}, \boldsymbol{x}_a} C_0(\boldsymbol{z}) + \sum_{\forall a \in \mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a) \\ &s.t. \ \boldsymbol{z} \in \mathcal{Z}, \boldsymbol{x}_a \in \mathcal{X}_a \\ &p_{a,t}^{grid} = \widehat{p}_{a,t}^{grid} \end{aligned} \tag{44}$$

where $C_0(\boldsymbol{z})$ is the reformulated form in (32). The problem (44) is a solvable structure for standard ADMM. The augmented Lagrangian form of each agent in the distributed framework is shown as follows:

Aggregator agent a = 0:

$$\begin{aligned} L_0 = C_0(\boldsymbol{z}) + \sum_{a \in \mathcal{A}\backslash 0} \sum_{t \in \mathcal{T}} \lambda_{a,t}\left(p_{a,t}^{grid} - \widehat{p}_{a,t}^{grid}\right) \\ + \frac{\rho}{2}\left(p_{a,t}^{grid} - \widehat{p}_{a,t}^{grid}\right)^2 \end{aligned} \tag{45}$$

Each customer agent $\forall a \in \mathcal{A}\backslash 0$:

$$\begin{aligned} L_a = C_a(\boldsymbol{x}_a) + \sum_{t \in \mathcal{T}} \Bigg(\lambda_{a,t}\left(p_{a,t}^{grid} - \widehat{p}_{a,t}^{grid}\right) \\ + \frac{\rho}{2}\left(p_{a,t}^{grid} - \widehat{p}_{a,t}^{grid}\right)^2\Bigg) \end{aligned} \tag{46}$$

where $\lambda_{a,t}$ is the dual variable and ρ is the penalty parameter in ADMM. Then, the ADMM solving steps for the problem (44) are given below via iteration and information update.

Step 1: Update primal each customer variable in parallel by:

$$\begin{aligned} \boldsymbol{x}_a^{k+1} = \arg\min_{\boldsymbol{x}_a \in \mathcal{X}_a} C_a(\boldsymbol{x}_a) + \sum_{t \in \mathcal{T}} \Bigg(\lambda_{a,t}^{k}\left(p_{a,t}^{grid} - \widehat{p}_{a,t}^{grid,k}\right) \\ + \frac{\rho}{2}\left(p_{a,t}^{grid} - \widehat{p}_{a,t}^{grid,k}\right)^2\Bigg) \forall a \in \mathcal{A}\backslash 0 \end{aligned} \tag{47}$$

Step 2: Update primal aggregator variables by:

$$\begin{aligned} \boldsymbol{z}^{k+1} = \arg\min_{\boldsymbol{z} \in \mathcal{Z}} C_0(\boldsymbol{z}) + \sum_{\forall a \in \mathcal{A}\backslash 0} \sum_{t \in \mathcal{T}} \lambda_{a,t}^{k}\Big(p_{a,t}^{grid,k+1} \\ - \widehat{p}_{a,t}^{grid}\Big) + \frac{\rho}{2}\left(p_{a,t}^{grid,k+1} - \widehat{p}_{a,t}^{grid}\right)^2 \end{aligned} \tag{48}$$

Step 3: Aggregator updates dual variables $\lambda_{a,t}$ by:

$$\lambda_{a,t}^{k+1} = \lambda_{a,t}^{k} + \rho\left(p_{a,t}^{grid,k+1} - \widehat{p}_{a,t}^{grid,k+1}\right) \quad \forall a \in \mathcal{A}\backslash 0, \forall t \in \mathcal{T} \tag{49}$$

The primal and dual residuals in (50) and (51) are used as termination criteria of iteration in ADMM:

$$\boldsymbol{r}^k = \left(\boldsymbol{p}^{grid,k} - \widehat{\boldsymbol{p}}^{grid,k}\right) \tag{50}$$

$$\boldsymbol{s}^k = -\rho\left(\widehat{\boldsymbol{p}}^{grid,k} - \widehat{\boldsymbol{p}}^{grid,k-1}\right) \tag{51}$$

where primal residual $\boldsymbol{r}^k$ denotes violation of the constraint (41) at the current iteration; dual residual $\boldsymbol{s}^k$ means the violation of the Karush-Kuhn-Tucker (KKT) stationarity constraints at the current iteration. The iteration of Steps 1-3 continues until the

Algorithm 1: AOP with ADMM Algorithm.

1: Initialization: Set $\varphi = 0$ as the initial iteration index for AOP and relax the binary variables in the problem (44) to be continuous variables between 0 and 1. The relaxed (44) is solved by ADMM according to (47)-(51), which ensures the convergence since each agent problem is linear. The continuous relaxed problem is solved to get the initial values of boundary variables $\widehat{p}_{a,t}^{grid,0}$ for AOP iteration.

2: Solve the $p_{a,t}^{grid}$−fixed **problem:** Optimize the problem (44) with $p_{a,t}^{grid}$−fixedaccording to ADMM (47)–(51). The boundary variables $p_{a,t}^{grid}$ are fixed at the $\widehat{p}_{a,t}^{grid,(\varphi-1)}$ given at last solution.

$$p_{a,t}^{grid} = \widehat{p}_{a,t}^{grid,(\varphi-1)} \tag{52}$$

As a result of equality (52), the augmented item in (47) and (49) is reduced into regional subproblems and they are individually solved once as follows:
For customer agents $\forall a \in \mathcal{A} \backslash 0$:

$$\min_{\boldsymbol{x}_a \in \mathcal{X}_a} C_a\left(\boldsymbol{x}_a\right) \tag{53}$$

The optimal solution of binary variables is denoted by $d_{a,t}^{g,(\varphi)}, \chi_{a,t}^{b,(\varphi)}$.

3: Check convergence: If the binary variables are the same as the values in the last iteration, i.e., $d_{a,t}^{g,(\varphi)} == d_{a,t}^{g,(\varphi-1)}, \chi_{a,t}^{b,(\varphi)} == \chi_{a,t}^{b,(\varphi-1)}$, then take the $d_{a,t}^{g,(\varphi)}, \chi_{a,t}^{b,(\varphi)}$ as the final solution and iteration stops; otherwise, set $\varphi = \varphi + 1$ and go to **step 4**.

4: Solve $d_{a,t}^{g}, \chi_{a,t}^{b}$−fixed **problem:** Solve the problem (44) with fixed binary variables given in the last step by ADMM according to (47)-(51), and then go to **step 2**.

residuals converge to the feasibility tolerances ε^{pri} and ε^{dual}. Step 1 is implemented by each customer agent individually; Step 2 is done by the aggregator agent. The formulas on the right side of the equation in (47) and (48) are the final DWDRO form of agents in the distributed framework.

The problem, whereas, may not be effectively solved with a convergence guarantee by using the original ADMM due to the binary variables $d_{a,t}^{g}, \chi_{a,t}^{b}$ in (8) and (10). A tractable heuristic algorithm, AOP is proposed to ensure convergence of distributed scheduling framework [50]. The key idea is to optimize the aggregation problem over binary variables $d_{a,t}^{g}, \chi_{a,t}^{b}$ with fixed boundary variables $p_{a,t}^{grid}$, then over $p_{a,t}^{grid}$ with fixed $d_{a,t}^{g}, \chi_{a,t}^{b}$, and repeat the process until the convergence condition is satisfied. The detailed process of AOP with ADMM is given as shown in Algorithm 1

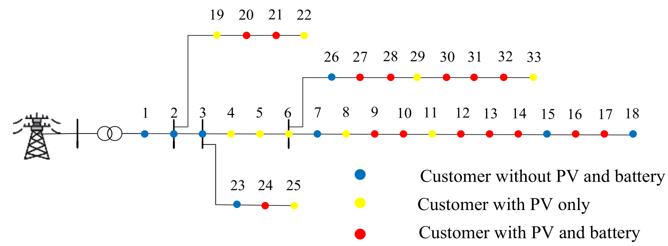

Fig. 3. Three types of customers on the modified IEEE 33-bus radial distribution network.

TABLE I
THE TIME OF USE (TOU) PRICE AND FEED-IN-TARIFF (FIT)

	10pm~7am	7am~2pm	2pm~8pm	8pm~10pm
ToU	0.12$/kwh	0.17$/kwh	0.49$/kwh	0.17$/kwh
FiT	0.075$/kwh			

IV. NUMERICAL EXPERIMENT

The numerical experiment of the proposed aggregation framework is presented on the modified IEEE 33-bus single-phase radial distribution network including 33 customers as shown in Fig. 3. The PV system in the simulation has a capacity of 400kwp and the battery has a capacity of 850kwh/450kw. The Time of Use (ToU) price and Feed-in-Tariff (FiT) are given in Table I. The historical data of LMP and the solar forecast error percentage can be collected from websites [51] and [52]. A solver MOSEK is applied to solve the distributed framework on the YALMIP in MATLAB by a PC with an Intel Core (TM) i5-6200U CPU @ 2.40 GHZ with 8.00 GB RAM. The value $L_{a,t}$ in (39) is solved offline by a nonlinear solver IPOPT.

A. Market Risk-Aversion for Aggregator

To first demonstrate the validation of the proposed data-driven DRO scheduling method for aggregator decision-making, the maximum available PV output of each customer is set to be its forecast value. The PV has a power output during the period between 6 am to 7 pm.

The impacts of Wasserstein radius ε and risk preference factor σ on day-ahead scheduling decisions are demonstrated in Figs. 4 and 5. The Wasserstein radius determines the size of the ambiguity set. Furthermore, it determines the conservatism of the DWDRO method. As shown in Fig. 4, the power purchased by the aggregator at night is nearly the same with respect to different ε, and the impact of ε on decision-making is reflected during the period of daytime (7 am $\sim$ 7 pm) when the aggregator submits bids to sell the electricity in the wholesale market. When $\varepsilon = 1$ (light pink shaded area in Fig. 4), the power sold keeps the level at 2115kw from 8 am to 6 pm. When ε drops to 0.5, the power sold increases to 3001kw and it still has a flat output decision from 7 am to 7 pm. However, when $\varepsilon = 0.1$, the power export increases to 3908 kW during daytime except 2929kw at 4:30 pm and 2958kw at 5 pm. As ε continues to drop, the power

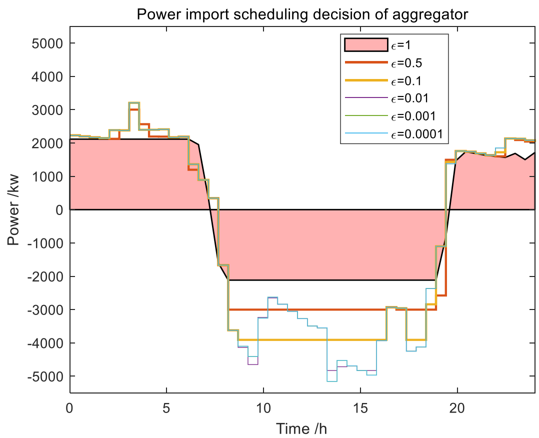

Fig. 4. Day-ahead scheduling decisions with respect to the Wasserstein radius.

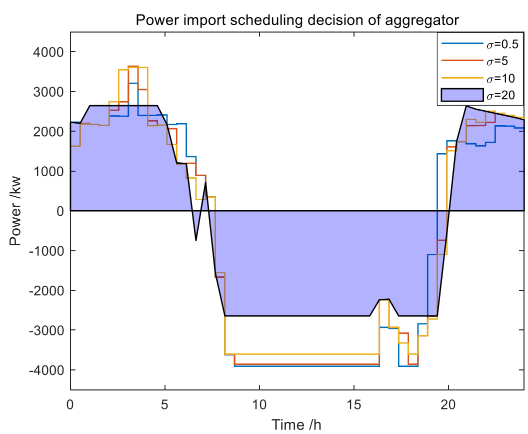

Fig. 5. Day-ahead scheduling decisions with respect to risk preference factor.

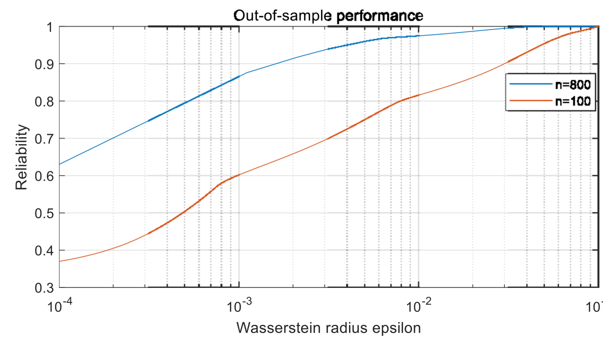

Fig. 6. Out-of-sample performance.

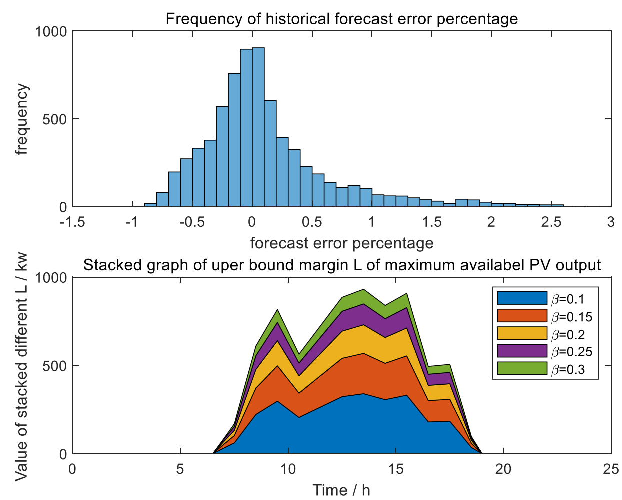

Fig. 7. Frequency of historical dataset and the relation of β_{pv} and $L_{a,t}$.

output of each time interval during daytime does not keep the same. The aggregator's bid at some intervals has more power export to pursue more revenue and the aggregator is willing to undertake more price risks. Therefore when $\varepsilon = 0.0001$, the day-ahead scheduling decisions at each interval during the period of 7 am ~7 pm are various. The power export between 10:30 am and 11 am is 2929 kW while the one is 5155kW between 4:30 pm and 5 pm. The smaller the Wasserstein radius is, the less conservative the decision is. As a result, the scheduling decision of $\varepsilon = 0.0001$ is solved by the proposed method to bid more power at some intervals to pursue higher revenue while bidding less power at other intervals. The impacts of the risk preference factor on the scheduling decision are shown in Fig. 5. The larger σ is, the more conservative the decision is, which is reflected during the period of daytime to sell electricity. When $\sigma = 20$, the power export in the bid only levels at 2643kw between 8:30 am and 7 pm (purple shaded area in Fig. 5).

The out-of-sample performance of the scheduling decision solved by the proposed method is validated by implementing 100 runs to compare the expected cost and the averaged cost realized by the given decision and out-of-sample market prices. The occurrence frequency of the out-of-sample average cost lower than the expected cost is denoted as reliability. As we can see in Fig. 6, a large size of the dataset in the construction of the Wasserstein ambiguity set can help the aggregator improve the reliability of the power scheduling decision in offering and bidding.

B. DRO Model of Customer's PV Power and Cost

The dataset for solving the upper bound margin $L_{a,t}$ in (40) is constructed by using the historical forecast error percentage. The dataset of the forecast error at each time interval is calculated by the production of PV forecast power and the historical forecast error percentage. The frequency distribution of the historical forecast error percentage of the dataset with 7000 samples and the impact of the confidence level β_{pv} on $L_{a,t}$ are shown in Fig. 7. Almost error percentages distribute between -1 to 1. A small value of β_{pv} corresponds to the day-ahead dispatch of PV with high conservativeness, i.e., the PV output cannot exceed its maximum available realization with high probability. As a result, when β_{pv} varies from 0.1 (blue segment) to 0.3 (green segment), $L_{a,t}$ at each interval tends to become smaller. The upper limit $\overline{p}_{a,t}^{PV,f} - L_{a,t}$ in (40) increases correspondingly. The expected cost of customers can be adjusted through the value of β_{pv} as shown in Table II. The expected costs of two customers are selected to show relations of parameters β_{pv} and the customer's cost. A negative value means the customers get profits by exporting excessive solar energy. As for customer

TABLE II
EXPECTED CUSTOMER COST WITH RESPECT TO β_{pv}

β_{pv}	Customer 9	Customer 11
0.1	$293.59	$141.89
0.15	$42.11	-$7.42
0.2	-$59.28	-$86.59
0.25	-$145.31	-$139.18
0.3	-$173.83	-$180.88

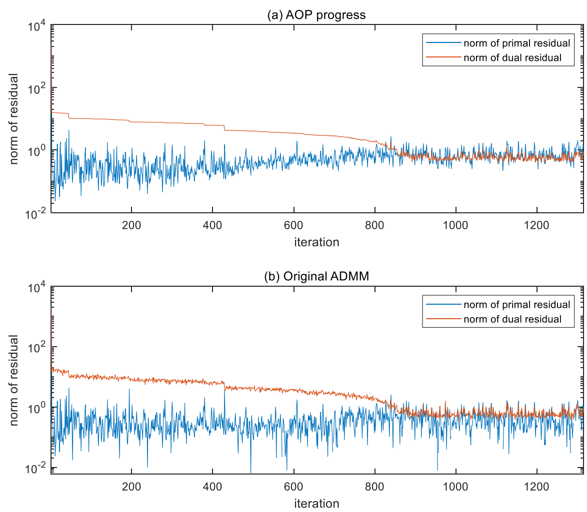

Fig. 8. Iteration process of AOP and original ADMM.

9, when $\beta_{pv} = 0.3$, the customer can earn profits of $173.83 during the daytime; when $\beta_{pv} = 0.2$, the customer can earn profits of $59.28 during the daytime. The deviation between the feed-in power decision of the day head and the real-time is not encouraged for customers. The punishment might be adopted by DSO to reduce the deviation. The customer can adjust the parameter β_{pv} to reduce the expected PV output according to their acceptance of the deviation penalty.

C. Efficiency of AOP With ADMM Algorithm

To check the validation of the AOP process, the AOP solution process of the distributed framework on the modified 33-bus network is shown in Fig. 8(a). It is compared with the original ADMM algorithm to solve the proposed distributed aggregation problem directly, and the solution process is shown in Fig. 8(b) as well for the sake of comparison. The AOP process in Fig. 8(a) corresponds to step 4 in the proposed algorithm when $\varphi = 1$, which means the aggregation problem is solved by ADMM with fixed binary variables after the initialization is implemented. The whole solving process stops at $\varphi = 2$. Compared with the original ADMM, the AOP process stops at the 1316th iteration, and the norm of primal and dual residuals have relevantly smaller fluctuation as the solving process goes forward. Whereas the iteration process solved by the original ADMM does not stop at the 1316th iteration and the norm of primal residual has more obvious fluctuation than AOP. As a result, through solving the $p^{grid}_{a,t}-$fixed problem and $d^g_{a,t}, \chi^b_{a,t}-$fixed alternatively, the convergence of ADMM is ensured due to the guarantee of convexity.

Fig. 9. Realistic single feeder LV network in U.K. and customers allocation.

TABLE III
RESULTS COMPARISON OF TWO NETWORKS

Network	AOP processes	Iterations in AOP	Computational time (min)	Error%
33-bus network	$\varphi = 2$	1316	9.65	0.211%
Realistic LV network	$\varphi = 3$	2189	17.61	0.384%

D. Scalability on a Realistic LV Network in the U.K.

To further verify the scalability of the proposed AOP process, a numerical experiment is implemented on a single-phase LV network extracted from a realistic distribution system in the U.K. with 75 customers (45 customers with both PV and battery). The residential house as a single-phase customer is connected to the single-phase LV distribution network. The allocation of different customers in the network is shown in Fig. 9. The PV forecast and load profile are collected from the website [53]. The comparison of two experiments on 33-bus and the LV distribution network is given in Table III. As the number of customers increases from 33 to 75, the number of agents increases accordingly. The number of AOP processes rises from $\varphi = 2$ to $\varphi = 3$, while the number of iterations in the first process rises from 1316 to 2189. Even though the higher iteration number means much more time consumption (i.e., the computational time in an AOP process increases from 9.65min to 17.61min)due to more customers participating in the aggregation scheme, the time consumption will not affect the application of the proposed method since the scheduling decisions are made in the day ahead. Moreover, the convergence speedup of ADMM has been extensively researched in the existing literature, which can be adopted directly. The optimum objective value of the aggregation problem solved by the proposed distributed manner is compared with the one solved in a centralized manner. The error percentages are much less than 1%. Fig. 10 demonstrates the 45-customer scheduling results of the power export. Each customer feeds the excess power back into the grid. Each customer's power export profile is different but they have a similar characteristic

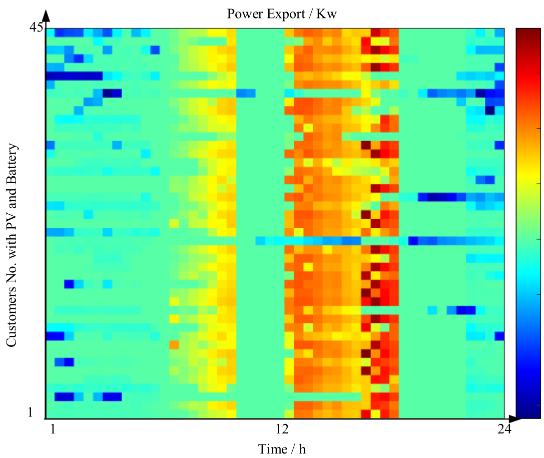

Fig. 10. Power export of customers with PV and battery.

that most customers do not export power from 10 am – 12 pm. The reason is that the objective of the aggregator agent existing in the aggregation problem influences each customer's power export since the aggregator needs to decide the port power in trade considering the electricity market price.

V. Conclusion

The proposed DWDRO agent models of the aggregator and customers reflect the conservativeness of aggregation scheduling decisions under the risks of uncertain market prices and PV forecast errors. The aggregator can adjust the parameters including the dataset size n, radius ε of the Wasserstein ambiguity set and risk factor σ to change the conservative performance of transactive energy import or export decisions to minimize the expected cost under the worst-cast distribution of uncertainties, especially during the period of daytime when aggregator exports the excessive energy generated by PV. The larger size of the dataset the agent has, the higher reliability the scheduling decision has. As for the customer agent DWDRO model with a given size of historical PV forecast error dataset, the confidence level $1-\beta_{pv}$ can be adjusted to determine the upper bound margin $L_{a,t}$ of the maximum available PV output to reduce the uncertainty of PV forecast. As the β_{pv} increases, the $L_{a,t}$ decreases, then each customer's expected cost decreases accordingly from a positive value to a negative one, which means customer energy status becomes from import to export. The proposed AOP process with ADMM ensures the solution convergence of the data-driven distributed framework when the customer agents have the binary variables representing battery charging/discharging and energy import/export.

Appendix I

The aggregation scheduling problem below is formulated composed of aggregator cost and customers' cost as shown in (26):

$$\min_{\boldsymbol{z}\in\mathcal{Z},\boldsymbol{x}_a\in\mathcal{X}_a} \left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t + \sum_{\forall a\in\mathcal{A}\backslash 0}\sum_{t\in\mathcal{T}}\left(c^{FiT}p^{g+}_{a,t}\Delta^t - c^{ToU}_t p^{g-}_{a,t}\Delta^t\right)$$

Let's formulate the aggregator cost in the form of expectation with *CVaR* item and denote each customer's cost with $C_a(\boldsymbol{x}_a) = \sum_{t\in\mathcal{T}}(c^{FiT}p^{g+}_{a,t}\Delta^t - c^{ToU}_t p^{g-}_{a,t}\Delta^t)$. The following process is derived only to formulate the equivalently piecewise affine form of expected aggregator cost with *CVaR*. According to the definition of *CVaR*, the *CVaR* of $\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t$ is given in (54):

$$CVaR\left[\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t\right] = \min_{\tau\in\mathbb{R}}\mathbb{E}^{\mathbb{P}}\left[\tau + \frac{1}{\alpha}\max\left\{\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t - \tau, 0\right\}\right] \tag{54}$$

Substitute the (54) in (55) and move the "min" sign to the left side together to get (56).

$$\min_{\boldsymbol{z}\in\mathcal{Z},\boldsymbol{x}_a\in\mathcal{X}_a} \mathbb{E}^{\mathbb{P}}\left[\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t\right] + \sigma \times CVaR\left[\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t\right] + \sum_{\forall a\in\mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a) \tag{55}$$

$$\min_{\boldsymbol{z}\in\mathcal{Z},\boldsymbol{x}_a\in\mathcal{X}_a,\tau\in\mathbb{R}} \mathbb{E}^{\mathbb{P}}\left[\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t + \sigma\tau + \frac{\sigma}{\alpha} \times \max\left\{\left[\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t - \tau\right], 0\right\}\right] + \sum_{\forall a\in\mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a) \tag{56}$$

Then, continue to formulate by moving the item $\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t$ into the max function $\max\{\cdot,0\}$ in (56) to get (57) as follows:

$$\min_{\substack{\boldsymbol{z}\in\mathcal{Z},\boldsymbol{x}_a\in\mathcal{X}_a,\\ \tau\in\mathbb{R}}} \mathbb{E}^{\mathbb{P}}\left[\max\left\{\begin{array}{l}\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t + \sigma\tau + \frac{\sigma}{\alpha}\\ \times\left[\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t - \tau\right],\\ \left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t + \sigma\tau\end{array}\right\}\right] + \sum_{\forall a\in\mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a) \tag{57}$$

Then, formulate the items in max function in the form of expression with respect to $\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t$ and τ to get (58).

$$\min_{\substack{\boldsymbol{z}\in\mathcal{Z},\boldsymbol{x}_a\in\mathcal{X}_a,\\ \tau\in\mathbb{R}}} \mathbb{E}^{\mathbb{P}}\left[\max\left\{\begin{array}{l}\left(1+\frac{\sigma}{\alpha}\right)\times\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t\\ +\left(\sigma-\frac{\sigma}{\alpha}\right)\tau,\\ \left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t + \sigma\tau\end{array}\right\}\right] + \sum_{\forall a\in\mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a) \tag{58}$$

Finally, the equivalently piecewise affine form of expected aggregator cost with *CVaR* is given as shown in (59) with parameter a_1, a_2, b_1, b_2, which is the form shown in (29).

$$\min_{\boldsymbol{z},\boldsymbol{x}_a,\tau\in\mathbb{R}} \mathbb{E}^{\mathbb{P}}\left[\max_{j\leq 2}\left\{a_j\left(\boldsymbol{\mu}^{Price}\right)^\top \boldsymbol{p}^{agg}\Delta^t + b_j\tau\right\}\right] + \sum_{\forall a\in\mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a)$$
$$a_1 = \left(1+\frac{\sigma}{\alpha}\right), a_2 = 1, b_1 = \left(\sigma-\frac{\sigma}{\alpha}\right), b_2 = \sigma \tag{59}$$

■

APPENDIX II

The distributional robust form of the objective function of the aggregation problem is shown in (60). To derive the reformulation of the distributionally robust model of aggregator agent, only (61) will be discussed in this section.

$$\min_{\substack{\boldsymbol{z}\in\mathcal{Z},\mathbf{x}_a\in\mathcal{X}_a,\\ \tau\in\mathbb{R}}} \sup_{\mathbb{B}_\varepsilon(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{P}}\left[\max_{j\le 2}\left\{a_j(\boldsymbol{\mu}^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + b_j\tau\right\}\right] + \sum_{\forall a\in\mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a) \tag{60}$$

$$\sup_{\mathbb{P}\in\mathbb{B}_\varepsilon(\widehat{\mathbb{P}}_N)} \mathbb{E}^{\mathbb{P}}\left[\max_{j\le 2}\left\{a_j(\boldsymbol{\mu}^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + b_j\tau\right\}\right] \tag{61}$$

According to [24], due to (61) is linear with respect to the random vector $\boldsymbol{\mu}^{Price}$, the strong dual holds. Based on the Wasserstein ambiguity set in (25), the problem (61) can be reformulated equivalently into (62), where λ_0 and g_i are dual variables, and $\widehat{\boldsymbol{\mu}}_i^{Price}$ is i$^{\text{th}}$ sample in the historical dataset.

$$\begin{aligned}\Leftrightarrow &\min_{\lambda_0\ge 0, g_i} \lambda_0\varepsilon + \frac{1}{N}\sum_{i=1}^N g_i\\ \text{s.t.} &\max_{\boldsymbol{\mu}^{Price}\in\Xi}\max_{j\le 2}\left\{a_j(\boldsymbol{\mu}^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + b_j\tau\right\}\\ &-\lambda_0\left\|\boldsymbol{\mu}_h^{Price} - \widehat{\boldsymbol{\mu}}_i^{Price}\right\| < g_i \forall i\end{aligned} \tag{62}$$

Then, merge the two max signs in constraint of (62) to get:

$$\begin{aligned}\Leftrightarrow &\min_{\lambda_0\ge 0, g_i} \lambda_0\varepsilon + \frac{1}{N}\sum_{i=1}^N g_i\\ \text{s.t.} &\max_{\boldsymbol{\mu}^{Price}\in\Xi}\left\{a_j(\boldsymbol{\mu}^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + b_j\tau\right\} - \lambda_0\left\|\boldsymbol{\mu}_h^{Price}\right.\\ &\left. - \widehat{\boldsymbol{\mu}}_i^{Price}\right\| < g_i \ \ \forall i, \forall j\end{aligned} \tag{63}$$

Then, substitute the norm $\|\bullet\|$ in (63) by the dual norm $\|\bullet\|_*$ and introduce the auxiliary variables $\boldsymbol{\nu}_{ij}$ to get (64):

$$\begin{aligned}\Leftrightarrow &\min_{\tau\in\mathbb{R},\lambda_0,g_i,\boldsymbol{\nu}_{ij}} \lambda_0\varepsilon + \frac{1}{N}\sum_{i=1}^N g_i\\ \text{s.t.} &\max_{\boldsymbol{\mu}^{Price}\in\Xi} a_j(\boldsymbol{\mu}^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + b_j\tau - \boldsymbol{\nu}_{ij}^\top\left(\boldsymbol{\mu}_h^{Price}\right.\\ &\left.- \widehat{\boldsymbol{\mu}}_i^{Price}\right) \le g_i \ \forall i, j \|\boldsymbol{\nu}_{ij}\|_* \le \lambda_0 \ \forall i, j\end{aligned} \tag{64}$$

According to the market rules, there are a ceiling price and a floor price to limit prices when bidding. As a result, the random vector $\boldsymbol{\mu}^{Price}$ is constrained within a set Ξ in (65).

$$\Xi = \left\{\boldsymbol{\mu}^{Price}:\boldsymbol{W}\boldsymbol{\mu}^{Price} \le \boldsymbol{d}\right\} \tag{65}$$

The constraints (64) become (66):

$$\begin{aligned}\Leftrightarrow &\max_{\boldsymbol{\mu}^{Price}\in\Xi} a_j(\boldsymbol{\mu}^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + b_j\tau - \boldsymbol{\nu}_{ij}^\top\left(\boldsymbol{\mu}^{Price} - \widehat{\boldsymbol{\mu}}_n^{Price}\right) \le g_i\\ &\text{s.t. } \boldsymbol{W}\boldsymbol{\mu}^{Price} \le \boldsymbol{d}\\ &\|\boldsymbol{\nu}_{ij}\|_* \le \lambda_0\end{aligned} \tag{66}$$

Since (66) is a linear problem with respect to $\boldsymbol{\mu}^{Price}$ and strong dual holds again, (67) is got.

$$\begin{aligned}\Leftrightarrow &\min_{\tau,\lambda_0,g_i,\boldsymbol{\gamma}_{ij}} b_j\tau + a_j(\widehat{\boldsymbol{\mu}}_i^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + \boldsymbol{d}^\top\boldsymbol{\gamma}_{ij}\\ &-\left(\boldsymbol{W}\widehat{\boldsymbol{\mu}}_i^{Price}\right)^\top\boldsymbol{\gamma}_{ij} \le g_i\\ &\text{s.t.}\left\|\boldsymbol{W}^\top\boldsymbol{\gamma}_{ij} - a_j\boldsymbol{p}^{agg}\Delta^t\right\|_* \le \lambda_0\\ &\boldsymbol{\gamma}_{ij} \ge \mathbf{0}, \lambda_0 \ge 0\end{aligned} \tag{67}$$

Finally, take (67) back into (64), the equivalent reformulation of DRO aggregator agent is derived below in (68) as shown in (32) to (35).

$$\begin{aligned}\Leftrightarrow &\min_{\substack{\boldsymbol{z}\in\mathcal{Z},\boldsymbol{x}_a\in\mathcal{X}_a,,\\ \tau,\lambda_0,g_i,\boldsymbol{\gamma}_{ij}}} \lambda_0\varepsilon + \frac{1}{N}\sum_{i=1}^N g_i\\ &\text{s.t. } b_j\tau + a_j(\widehat{\boldsymbol{\mu}}_i^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + \boldsymbol{d}^\top\boldsymbol{\gamma}_{ij} - \left(\boldsymbol{W}\widehat{\boldsymbol{\mu}}_i^{Price}\right)^\top\boldsymbol{\gamma}_{ij} \le g_i\\ &\left\|\boldsymbol{W}^\top\boldsymbol{\gamma}_{ij} - a_j\boldsymbol{p}^{agg}\Delta^t\right\|_* \le \lambda_0\\ &\boldsymbol{\gamma}_{ij} \ge \mathbf{0}, \lambda_0 \ge 0\end{aligned} \tag{68}$$

Let $C_0(\boldsymbol{z}) = \lambda_0\varepsilon + \frac{1}{N}\sum_{i=1}^N g_i$, the equivalent reformulation of the distributionally robust form of aggregation problem is given as follows:

$$\begin{aligned}\Leftrightarrow &\min_{\substack{\boldsymbol{z},\boldsymbol{x}_a,\\ \tau,\lambda_0,g_i,\boldsymbol{\gamma}_{ij}}} C_0(\boldsymbol{z}) + \sum_{a\in\mathcal{A}\backslash 0} C_a(\boldsymbol{x}_a)\\ &\text{s.t. } \boldsymbol{z}\in\mathcal{Z}, \boldsymbol{x}_a\in\mathcal{X}_a\\ &b_j\tau + a_j(\widehat{\boldsymbol{\mu}}_i^{Price})^\top \boldsymbol{p}^{agg}\Delta^t + \boldsymbol{d}^\top\boldsymbol{\gamma}_{ij} - \left(\boldsymbol{W}\widehat{\boldsymbol{\mu}}_i^{Price}\right)^\top\boldsymbol{\gamma}_{ij} \le g_i\\ &\left\|\boldsymbol{W}^\top\boldsymbol{\gamma}_{ij} - a_j\boldsymbol{p}^{agg}\Delta^t\right\|_* \le \lambda_0\\ &\boldsymbol{\gamma}_{ij} \ge \mathbf{0}, \lambda_0 \ge 0\end{aligned}$$

■

REFERENCES

[1] AEMO and Energy Networks Australia, "Interim report: Required capabilities and recommended actions," 2019.
[2] F. Rahimi, A. Ipakchi, and F. Fletcher, "The changing electrical landscape: End-to-end power system operation under the transactive energy paradigm," *IEEE Power Energy Mag.,* vol. 14, no. 3, pp. 52–62, May/Jun. 2016.
[3] K. O. Adu-Kankam and L. M. Camarinha-Matos, "Towards collaborative virtual power plants: Trends and convergence," *Sustain. Energy, Grids Netw.,* vol. 16, pp. 217–230, 2018.
[4] E. Mashhour and S. M. Moghaddas-Tafreshi, "Bidding strategy of virtual power plant for participating in energy and spinning reserve markets—Part I: Problem formulation," *IEEE Trans. Power Syst.,* vol. 26, no. 2, pp. 949–956, May 2011.
[5] N. Ruiz, I. Cobelo, and J. Oyarzabal, "A direct load control model for virtual power plant management," *IEEE Trans. Power Syst.,* vol. 24, no. 2, pp. 959–966, May 2009.
[6] M. Giuntoli and D. Poli, "Optimized thermal and electrical scheduling of a large scale virtual power plant in the presence of energy storages," *IEEE Trans. Smart Grid,* vol. 4, no. 2, pp. 942–955, Jun. 2013.
[7] F. Moret and P. Pinson, "Energy collectives: A community and fairness based approach to future electricity markets," *IEEE Trans. Power Syst.,* vol. 34, no. 5, pp. 3994–4004, Sep. 2019.

[8] S. Mhanna, A. C. Chapman, and G. Verbič, "A fast distributed algorithm for large-scale demand response aggregation," *IEEE Trans. Smart Grid,* vol. 7, no. 4, pp. 2094–2107, Jul. 2016.

[9] T. Morstyn and M. D. McCulloch, "Multiclass energy management for peer-to-peer energy trading driven by prosumer preferences," *IEEE Trans. Power Syst.,* vol. 34, no. 5, pp. 4005–4014, Sep. 2019.

[10] J. Guerrero et al., "Towards a transactive energy system for integration of distributed energy resources: Home energy management, distributed optimal power flow, and peer-to-peer energy trading," *Renewable Sustain. Energy Rev.,* vol. 132, 2020, Art. no. 110000.

[11] D. K. Molzahn et al., "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Nov. 2017.

[12] P. Chavali, P. Yang, and A. Nehorai, "A distributed algorithm of appliance scheduling for home energy management system," *IEEE Trans. Smart Grid,* vol. 5, no. 1, pp. 282–290, Jan. 2014.

[13] J.-Y. Joo and M. D. Ilić, "Multi-layered optimization of demand resources using Lagrange dual decomposition," *IEEE Trans. Smart Grid,* vol. 4, no. 4, pp. 2081–2088, Dec. 2013.

[14] S. Kar, G. Hug, J. Mohammadi, and J. M. F. Moura, "Distributed state estimation and energy management in smart grids: A consensus+ innovations approach," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 6, pp. 1022–1038, Dec. 2014.

[15] A. J. Conejo et al., "A decomposition procedure based on approximate newton directions," *Math. Program.,* vol. 93, no. 3, pp. 495–515, 2002.

[16] A. Attarha, P. Scott, and S. Thiébaux, "Affinely adjustable robust ADMM for residential DER coordination in distribution networks," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1620–1629, Mar. 2020.

[17] X. Kou et al., "A scalable and distributed algorithm for managing residential demand response programs using alternating direction method of multipliers (ADMM)," *IEEE Trans. Smart Grid,* vol. 11, no. 6, pp. 4871–4882, Nov. 2020.

[18] P. Scott, D. Gordon, E. Franklin, L. Jones, and S. Thiébaux, "Network-aware coordination of residential distributed energy resources," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6528–6537, Nov. 2019.

[19] H. Pandžić, J. M. Morales, A. J. Conejo, and I. Kuzle, "Offering model for a virtual power plant based on stochastic programming," *Appl. Energy,* vol. 105, pp. 282–292, 2013.

[20] X. Cao, J. Wang, and B. Zeng, "A chance constrained information-gap decision model for multi-period microgrid planning," *IEEE Trans. Power Syst.,* vol. 33, no. 3, pp. 2684–2695, May 2018.

[21] Z. Liang and Y. Guo, "Robust optimization based bidding strategy for virtual power plants in electricity markets," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2016, pp. 1–5.

[22] A. Baringo and L. Baringo, "A stochastic adaptive robust optimization approach for the offering strategy of a virtual power plant," *IEEE Trans. Power Syst.,* vol. 32, no. 5, pp. 3492–3504, Sep. 2017.

[23] H. Heitsch and W. Römisch, "Scenario reduction algorithms in stochastic programming," *Comput. Optim. Appl.,* vol. 24, no. 2, pp. 187–206, 2003.

[24] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, no. 1, pp. 115–166, 2018.

[25] L. Yang, Y. Xu, H. Sun, and W. Wu, "Tractable convex approximations for distributionally robust joint chance-constrained optimal power flow under uncertainty," *IEEE Trans. Power Syst.,* vol. 37, no. 3, pp. 1927–1941, May 2022.

[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. 33, no. 5, pp. 4924–4936, Sep. 2018.

[27] Y. Guo, K. Baker, E. Dall'Anese, Z. Hu, and T. H. Summers, "Data-based distributionally robust stochastic optimal power flow—Part I: Methodologies," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1483–1492, Mar. 2019.

[28] B. K. Poolla, A. R. Hota, S. Bolognani, D. S. Callaway, and A. Cherukuri, "Wasserstein distributionally robust look-ahead economic dispatch," *IEEE Trans. Power Syst.,* vol. 36, no. 3, pp. 2010–2022, May 2021.

[29] P. Xiong, P. Jirutitijaroen, and C. Singh, "A distributionally robust optimization model for unit commitment considering uncertain wind power generation," *IEEE Trans. Power Syst.,* vol. 32, no. 1, pp. 39–49, Jan. 2017.

[30] Z. Chu, N. Zhang, and F. Teng, "Frequency-constrained resilient scheduling of microgrid: A distributionally robust approach," *IEEE Trans. Smart Grid,* vol. 12, no. 6, pp. 4914–4925, Nov. 2021.

[31] H. Yang et al., "Distributionally robust optimal bidding of controllable load aggregators in the electricity market," *IEEE Trans. Power Syst.,* vol. 33, no. 1, pp. 1089–1091, Jan. 2018.

[32] Z. Shi, H. Liang, S. Huang, and V. Dinavahi, "Distributionally robust chance-constrained energy management for islanded microgrids," *IEEE Trans. Smart Grid,* vol. 10, no. 2, pp. 2234–2244, Mar. 2019.

[33] Y. Ding, T. Morstyn, and M. D. McCulloch, "Distributionally robust joint chance-constrained optimization for networked microgrids considering contingencies and renewable uncertainty," *IEEE Trans. Smart Grid,* vol. 13, no. 3, pp. 2467–2478, May 2022.

[34] Y. Chen, Q. Guo, H. Sun, Z. Li, W. Wu, and Z. Li, "A distributionally robust optimization model for unit commitment based on Kullback–Leibler divergence," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5147–5160, Sep. 2018.

[35] 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.

[36] C. Wang, R. Gao, W. Wei, M. Shafie-khah, T. Bi, and J. P. S. Catalão, "Risk-based distributionally robust optimal gas-power flow with Wasserstein distance," *IEEE Trans. Power Syst.,* vol. 34, no. 3, pp. 2190–2204, May 2019.

[37] R. A. Jabr, "Distributionally robust CVaR constraints for power flow optimization," *IEEE Trans. Power Syst.,* vol. 35, no. 5, pp. 3764–3773, Sep. 2020.

[38] F. Pourahmadi and J. Kazempour, "Distributionally robust generation expansion planning with unimodality and risk constraints," *IEEE Trans. Power Syst.,* vol. 36, no. 5, pp. 4281–4295, Sep. 2021.

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

[40] C. Li, Y. Xu, X. Yu, C. Ryan, and T. Huang, "Risk-averse energy trading in multienergy microgrids: A two-stage stochastic game approach," *IEEE Trans. Ind. Inform.*, vol. 13, no. 5, pp. 2620–2630, Oct. 2017.

[41] S. R. Etesami, W. Saad, N. B. Mandayam, and H. V. Poor, "Stochastic games for the smart grid energy management with prospect prosumers," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 2327–2342, Aug. 2018.

[42] W. Zhang, Y. Xu, Z. Dong, and K. P. Wong, "Robust security constrained-optimal power flow using multiple microgrids for corrective control of power systems under uncertainty," *IEEE Trans. Ind. Inform.,* vol. 13, no. 4, pp. 1704–1713, Aug. 2017.

[43] C. Zhang, Y. Xu, Z. Y. Dong, and J. Ma, "Robust operation of microgrids via two-stage coordinated energy storage and direct load control," *IEEE Trans. Power Syst.,* vol. 32, no. 4, pp. 2858–2868, Jul. 2017.

[44] AEMO, "Fact sheet: The national electricity market," 2020.

[45] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification—Part I," *IEEE Trans. Power Syst.,* vol. 28, no. 3, pp. 2554–2564, Aug. 2013.

[46] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification—Part II," *IEEE Trans. Power Syst.,* vol. 28, no. 3, pp. 2554–2564, Aug. 2013.

[47] K. Leong and R. J. E. R. M. Siddiqi, "Value at risk for power markets," *Energy Risk Manage.*, pp. 157–178, 1998.

[48] G. Chen, H. Zhang, H. Hui, and Y. Song, "Fast Wasserstein-distance-based distributionally robust chance-constrained power dispatch for multi-zone HVAC systems," *IEEE Trans. Smart Grid,* vol. 12, no. 5, pp. 4016–4028, Sep. 2021.

[49] S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations Trends Mach. Learn.,* vol. 3, no. 1, pp. 1–122, 2011.

[50] Z. Li, M. Shahidehpour, W. Wu, B. Zeng, B. Zhang, and W. Zheng, "Decentralized multiarea robust generation unit and tie-line scheduling under wind power uncertainty," *IEEE Trans. Sustain. Energy,* vol. 6, no. 4, pp. 1377–1388, Oct. 2015.

[51] "Aggregated price and demand data," Accessed: Aug. 2021. [Online]. Available: https://aemo.com.au/energysystems/electricity/national-electricity-market-nem/data-nem/aggregateddata

[52] "Solar home electricity data," Accessed: Aug. 2021. [Online]. Available: https://www.tennet.eu/electricitymarket/transparency-pages/transparency-germany/networkfigures/actual-and-forecast-solar-energy-feed-in/

[53] "Solar home electricity data," Accessed: Aug. 2022. [Online]. Available: https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Solar-home-electricity-data

Huichuan Liu received the master's degree in electrical engineering in 2018 from The University of Sydney, Sydney, NSW, Australia, where he is currently working toward the Ph.D. degree in electrical engineering. His research interests include power system operation and planning, distributed energy resources integration, electricity market, converter-dominated power system stability, and AI application in energy systems.

Jing Qiu (Senior Member, IEEE) received the B.Eng. degree in control engineering from Shandong University, Jinan, China, in 2008, the M.Sc. degree in environmental policy and management, majoring in carbon financing in the power sector from The University of Manchester, Manchester, U.K., in 2010, and the Ph.D. degree in electrical engineering from The University of Newcastle, Callaghan, NSW, Australia, in 2014. He is currently a Senior Lecturer in electrical engineering with The University of Sydney, Sydney, NSW, Australia. His research interests include power system operation and planning, energy economics, electricity markets, and risk management. He is also the Editorial Board Member of *IET Energy Conversion and Economics.*

Junhua Zhao (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from The University of Queensland, Brisbane, QLD, Australia, in 2007. He was a Senior Lecturer with the University of Newcastle, Callaghan NSW, Australia, and Center for Intelligent Electricity Networks (CIEN), University of Newcastle. He is currently the Assistant Dean and an Associate Professor with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China. He is also the Director of the Energy Market and Finance Lab, Shenzhen Finance Institute, The Chinese University of Hong Kong. His research interests include power system analysis and computation, smart grid, electricity market, data mining, and artificial intelligence. Dr. Zhao was the recipient of Young Scientist of the Future Award from ADC Forum in 2016, IEEE PES General Meeting Best Paper Award in 2014, and Science and Technology Progress Award (second prize) from the Hunan Provincial Government, China, in 2013. He is the Editorial Board Member of *IET Energy Conversion and Economics* and *Electric Power Components and Systems.*

Yuechuan Tao received the B.Eng. degree in electrical engineering and automation from Shanghai Normal University, Shanghai, China, in 2017, and the M.E. degree in electrical engineering, in 2019 from The University of Sydney, Sydney, NSW, Australia, where he is currently working toward the Ph.D. degree. His research interests include power system operation and planning, electric vehicles, data-driven, and smart grid.

Zhao Yang Dong (Fellow, IEEE) received the Ph.D. degree from The University of Sydney, Sydney, NSW, Australia. He is currently with Nanyang Technological University, Singapore. His previous roles include a SHARP Professor and the Director of the UNSW Digital Grid Futures Institute, University of New South Wales, Sydney, NSW, Australia, and Ausgrid Chair and Director of the Ausgrid Centre for Intelligent Electricity Networks providing support for the Smart Grid, Smart City National Demonstration Project. His research interests include smart grid, smart cities, power system planning and stability, renewable energy systems, electricity market, and computational methods for power engineering applications. He was/is the editor of a number of IEEE Transactions and IET journals.