The following publication X. Gao, K. W. Chan, S. Xia, X. Zhang, K. Zhang and J. Zhou, "A Multiagent Competitive Bidding Strategy in a Pool-Based Electricity Market With Price-Maker Participants of WPPs and EV Aggregators," in IEEE Transactions on Industrial Informatics, vol. 17, no. 11, pp. 7256-7268, Nov. 2021 is available at https://doi.org/10.1109/TII.2021.3055817

A Multi-agent Competitive Bidding Strategy in a Pool-Based Electricity Market with Price-Maker Participants of WPPs and EV Aggregators

Xiang Gao, Ka Wing Chan, *Member*, *IEEE*, and Shiwei Xia, *Senior Member*, *IEEE*,
Xiaoshun Zhang, *Member*, *IEEE,* Kuan Zhang, Jiahan Zhou

***Abstract*— Large-scale renewable energy suppliers and electric vehicles (EVs) are expected to become dominated participants in future electricity market. In this paper, a competitive bidding strategy is formulated for wind power plants (WPPs) and EV aggregators in a pool-based day-ahead electricity market. A bi-level multi-agent based model is proposed to study their bidding behaviors, with market clearing completion in the lower level and revenue maximization in the upper level. A stochastic framework is developed to incorporate the uncertainties in maximal power production of WPPs and EV aggregators and bid prices of other participants. The process of bidding decision is formulated as a stochastic game with incomplete information, in which electricity suppliers including WPPs and EV aggregators are considered as players of the game, their lack of information in this stochastic market environment is counterbalanced by a multi-agent reinforcement learning (MARL) algorithm named *win or learn fast policy hill climbing* (*WoLF-PHC*) with maximizing their own profits by self-game. The feasibility and effectiveness of the proposed model and the *WoLF-PHC* solution approach are successfully illustrated using a modified IEEE 6-bus system and a modified 118-bus system with different numbers of market players.**

Index Terms*— Bidding strategy, electricity market, renewable energy, multi-agent reinforcement learning (MARL), stochastic game, *WoLF-PHC

NOMENCLATURE

Indices:

l	Index of loads
g	Index of traditional generators
e	Index of EV aggregators
ω	Index of WPPs
n	Index of nodes
α	Index of scenarios
k	Index of iterations

This work was jointly supported by The Hong Kong Polytechnic University, the National Natural Science Foundation of China (52077075), National Natural Science Foundation of China (51907112), the Jiangsu Basic Research Project (BK20180284), Natural Science Foundation of Guangdong Province of China (2019A1515011671), Research and Development Start-Up Foundation of Shantou University (NTF19001). (Corresponding authors: Ka Wing Chan; Shiwei Xia.)

Xiang Gao, Ka Wing Chan and Jiahan Zhou are with the Department of Electrical Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China (e-mail: jocelyn.gao@connect.polyu.hk; eekwchan@polyu.edu.hk; jiahan.zhou@connect.polyu.hk). Shiwei Xia is with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, School of Electrical and Electronic Engineering, the North China Electric Power University, Beijing, China (e-mail: s.w.xia@ncepu.edu.cn). Xiaoshun Zhang is with the College of Engineering, Shantou University, Shantou, China (e-mail: xszhang1990@sina.cn). Kuan Zhang is with the College of Electrical and Information Engineering, Hunan University, Changsha, China (e-mail: zhangkuan1994@qq.com).

Sets:

Ω_n^L	Set of all nodes with loads
Ω_n^G	Set of all nodes with traditional generators
Ω_n^E	Set of all nodes with EV aggregators
Ω_n^W	Set of all nodes with WPPs
Ω_n^N	Set of the neighbour nodes of node *n*
N_α	Set of all scenarios for stochastic optimization (including bid prices of non-strategic participants as well as maximal power production of WPPs and EV aggregators)

Parameters:

P_g^{max}	Max power output of traditional generator *g*, in MW
$P_{\omega,\alpha,t}^{W,max}$	Max power production of WPP ω at time *t* in scenario α, in MW
$P_{e,\alpha}^{E,max}$	Max power production of EV aggregator *e* in scenario α, in MW
$P_{l,t}^{max}$	Max power consumption for load *l* at time *t*, in MW
E_e^{max}	Max energy capacity of EV aggregator *e*, in MWh
$SOC_{e,\alpha,end}$	Desired SOC of EV aggregator *e* in scenario α, in %
SOC^{max}	Upper limit of SOC for EV aggregators, in %
SOC^{min}	Lower limit of SOC for EV aggregators, in %
$\lambda_{gn,\alpha,t}^{DA}$	Bid price of traditional generator *g* at node *n* and time *t* in scenario α, in $/MW
$\lambda_{ln,\alpha,t}^{DA}$	Bid price of load *l* at node *n* and time *t* in scenario α, in $/MW
υ_c / υ_{dis}	Charging/discharging efficiency, in %
B_{nm}	Susceptance of the line connecting nodes *n* and *m*, in Siemens
f_{nm}^{max}	Thermal capacity of the line connecting nodes *n* and *m*, in MW
λ_{ev}	Price for EV aggregators traded with EV owners, in $/MW
τ_α	Weighting factor (probability) of scenario α, and the sum of probabilities for all scenarios is equal to 1, e.g., $\sum_{\alpha \epsilon N_\alpha} \tau_\alpha = 1$.
L_b	Battery lifetime, in MWh
C_b	Battery capital cost, in $
μ	Learning rate of *WoLF-PHC*
η	Discount factor of *WoLF-PHC*
δ_w	Win learning parameters for updating the policy of the algorithm
δ_l	Lose learning parameters for updating the policy of the algorithm

Variables:

$\varphi_{n,\alpha,t}^{DA}$ Locational Marginal Pricing (LMP) at node n and time t in scenario α, in \$/MW

$P_{e,\alpha,t}^{E,DA}$ Power dispatched for EV aggregator e at time t in scenario α, in MW

$P_{\omega,\alpha,t}^{W,DA}$ Power dispatched for WPP ω at time t in scenario α, in MW

$P_{g,\alpha,t}^{G,DA}$ Power dispatched for traditional generator g at time t in scenario α, in MW

$P_{l,\alpha,t}^{L,DA}$ Power dispatched for load l at time t in scenario α, in MW

$SOC_{e,\alpha,t}$ State of charge for EV aggregator e at time t in scenario α, in %

$\lambda_{\omega n,t}^{bid}$ Bid price of WPP ω at node n and time t, in \$/MW

$\lambda_{en,t}^{bid}$ Bid price of EV aggregator e at node n and time t, in \$/MW

$\theta_{n,\alpha,t}^{DA}$ Voltage angle at node n and time t in scenario α

I. Introduction

In oligopoly electricity market, large-scale energy suppliers have a vital influence on market prices and production. The share of renewable energy in gross electricity generation is expected to be increased to 80% by 2030 and 100% by 2050, of which 37% will be provided by wind energy [1]. Wind power producers (WPPs) have occupied a dominant position in some regions such as in Danish electricity markets [2]. Meanwhile, rapidly growing numbers of distributed electric vehicles (EVs) could be aggregated as a new demand response (DR) resource for providing energy through a coordinator called the EV aggregator, which dispatches EVs and exchanges information between independent system operator (ISO) and individual EV owners [3]. The green contribution of WPPs and EVs to energy conservation and environmental protection, as well as the rapid expansion of their scales, would lead them to be oligopolists in the wholesale market [4, 5]. Furthermore, it is worthy of developing an effective approach to model market behaviors for WPPs and EV aggregators.

In general, there are cooperative and competitive models for developing optimal bidding strategies in electricity market. Numerous cooperative models have been proposed based on collaborators' compensation for peak shaving [6-8]. Although the cooperative model in [9] has improved the common interests of EVs and WPPs on account of their energy coordination during different time slots, this model depends highly on the centralized control and scheduling of a central aggregator who owns lots of bidding information of EVs and WPPs. In this way, the communication requirement between EVs and WPPs is high, the privacy of EVs and WPPs cannot be guaranteed, and WPPs and EV aggregators have little flexibility to cope with any self-decisions. In future electricity market with great respect for personal privacy and autonomy of energy suppliers, a bidding strategy accustomed to a more flexible market without any central agent would be more attractive. Since energy suppliers would not share any personal data with each other for the sake of privacy, WPPs and EV aggregators would have the freedom to make their own bidding decisions to increase their respective profits in a competitive market.

For this purpose, equilibrium model and agent-based model (ABM) have been widely adopted in the modeling of a competitive market. Equilibrium models are formulated using mathematical programming approaches [10]. However, the equilibrium solution is often not easy to obtain or does not exist, especially in large systems. ABM is more flexible, in which all market players could be modeled as the artificial autonomous agents learning through repetitive interactions with a simulated market environment [11]. Thus it is more similar to a real electricity market. Model-based intuitive learning formulation and genetic algorithms are used separately to find the optimal bidding curves in [12] and [13]. However, these algorithms are designed to obtain the bidding strategy of a single agent, in which each agent makes the decision without regard to other rivals. Thus, they are inappropriate for the competitive market with each agent achieving its independent goal by adapting its behavior in the presence of other agents. As a class of reinforcement learning methods, various *Q-learning* algorithms have been widely used in the multi-agent electricity market to explore bidding strategies. A decentralized multi-agent model of EV owners bidding was developed based on a *Q-learning* algorithm without modeling the environment [14]. A deep reinforcement learning-based methodology was proposed to address bidding problems for energy suppliers, which has significant advantages in contrast with the traditional *Q-learning* algorithm [15]. However, there is no game process among these strategic participants. In [16] and [17], the game problem was considered to update multi-agent bidding strategies for energy suppliers respectively in a large power system and regionally integrated energy system, in which historical bidding decisions of rivals are essential. Although this model has some advantages compared to that with perfect information of other rivals' cost functions and market clearing mechanism, it is still not practical as the real electricity market is more likely a market with incomplete information, i.e. each player knows about only its own cost function and bidding strategy without any knowledge of other competitors [18]. A combined multi-agent model-based and learning-based approach for generators' bidding decision problems with incomplete information was presented in [19], but the calculation of the Nash equilibrium point is time-consuming and hard to solve for large systems. This is due to the complexity of the equilibrium calculation and storage pressure of state-action space. Further research shall therefore focus on exploring an optimal multi-agent approach without knowing the equilibrium point, suitable for the bidding strategy in a competitive market, which protects personal privacy with no bidding information communicated among agents and guarantees they could have the right to make their own bids.

In addition, existing researches on large-scale players bidding in a competitive environment refer to the WPP and demand response in [10], GENCOs and large consumers in [20], a demand-response virtual power plant in [21]. Especially in recent years, the decision making of the EV parking lots attracts numerous studies. Centralized dispatches are adopted in [22-24, 30], in which EVs have to provides lots of private information such as bidding and capacity to a central agent. Although [25-29] use decentralized EVs scheduling, it is essential to provide their own historical or rival information for decision making. Besides, [29, 30] need to calculate Nash equilibrium point, which is time consuming. Based on the above discussions and comparisons of the decision making of the EV parking lots in Table I, decentralized decision making of aggregated EVs relies on neither historical and rival information nor Nash equilibrium point is worth to study.

TABLE I
Comparison of the Decision Making of the EV Parking Lots

Reference	Decentralized	Independent of Historical and Rival Information	Independent of Nash Equilibrium Point
[22-24]	×	×	✓
[25-28]	✓	×	✓
[29]	✓	×	×
[30]	×	×	×
Proposed	✓	✓	✓

So far, to the best of authors' knowledge, there is few study considering the renewable WPPs and EV aggregators as oligopolists for developing a competitive multi-agent bidding strategy in a pool-based day-ahead (DA) market without either any information of opponents or calculating any equilibrium point. Based on the previous work to propose a bi-level stochastic optimization model of central controlled offering strategy for an aggregated WPP-EV hybrid power plant (HPP) as a price maker in the DA market considering the uncertainties of energy production and the spot price in real-time (RT) market [9], a new competitive DA market model for oligopoly players WPPs and EV aggregators with incomplete information to develop bidding strategy and consider uncertainties in their maximal power productions and bid prices of other players is first proposed in this paper, and then efficiently solved using a recently developed multi-agent decentralized *Win or Learn Fast Policy Hill Climbing* (*WoLF-PHC*) [31] method, based on an easy average policy instead of the equilibrium policy, which could not only meet the requirements of respecting individual privacy and the autonomy of energy providers but also accommodate the complex bidding behaviors of energy suppliers, to maximize revenues of WPPs and EV aggregators.

The main contributions of this paper are as follows.

(1) Newly developed a stochastic bi-level model, as compared to strategic participant WPPs in [2] and cooperative players including strategic WPPs and EV aggregators in [9], to explore a competitive bidding strategy for WPPs and EV aggregators in a pool-based oligopoly DA market based on multi-agent game system, considering the uncertainty of maximal power productions of WPPs and EV aggregators and bid prices for other participants. Compared with the centralized dispatch in [2, 9], WPPs and EV aggregators are neither controlled by an aggregator or central agent nor share any personal information with other rivals and make self-determined biddings to increase their respective profits in a competitive market.

(2) Comprehensively analyzed the proposed bidding problem of the market model as a multi-agent stochastic game with incomplete information, and applied a multi-agent decentralized *WoLF-PHC* to counterbalance their lack of information in this stochastic market environment for WPPs and EV aggregators and make their own bidding decisions by self-game. Compared with the existing methods in [16, 17] and [19], no rivals' information and time-consuming equilibrium calculation are needed.

(3) Successfully applied the *WoLF-PHC* method to solve the multi-supplier bidding strategies problem in the modified IEEE 6-bus and 118-bus systems.

The organization of this paper is as follows. Section II introduces the proposed bi-level electricity market model and the bidding mechanism. In Section III, the *WoLF-PHC* implementation for suppliers' bidding strategies is proposed. Simulation results and analysis of four cases are represented in Section IV. Section V concludes this paper.

II. Multi-agent Electricity Market Model

In the pool-based DA market, all market participants are required to provide their sale to or purchase offers from the ISO in each hour day-ahead [9]. In this paper, the market is modeled as a bi-level structure as shown in Fig.1. WPPs and

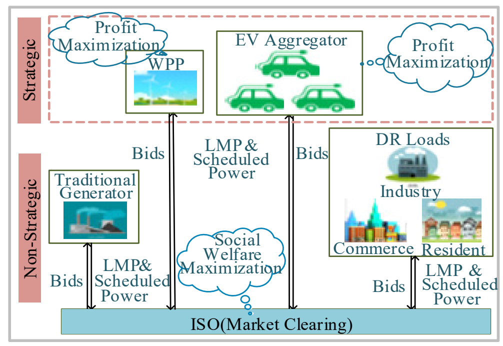


Fig. 1. Schematic representation of the proposed market model

EV aggregators are considered as two strategic oligopolies aimed to improve their respective revenues by solving bidding problems in the upper level. In the lower level, ISO collects bids from traditional generators, loads, and strategic WPPs and EV aggregators, and completes market clearing with maximizing the social welfare. Then ISO returns signals of Locational Marginal Pricing (LMP) and scheduled power to all participants. All the industrial, small residential and commercial loads are considered as a whole participant, i.e. a load aggregator, in this paper. Loads dispatched by the load aggregator consist of fixed and curtailed parts. The latter is considered as the elastic load and satisfied with the requirement of demand response. In this model, traditional generators and loads are assumed as non-strategic players. Their bid prices are their marginal cost prices and open to strategic suppliers [2]. Besides, the transmission network is represented by a DC model without losses.

In this paper, three sources of uncertainties including the maximal power productions of WPPs and EV aggregators, and bid prices for non-strategic participants are considered.

1) Stochastic model of WPP maximal power production

Weibull distribution is popularly used for modeling the probabilistic wind speed v_ω as (1).

$$f(v_\omega,\lambda,k)=\frac{k}{\lambda}\left(\frac{v_\omega}{\lambda}\right)^{k-1}e^{-\left(\frac{v_\omega}{\lambda}\right)^k} \tag{1}$$

where λ and k are the shape and scale parameters. As wind power at nearby locations may have related patterns due to the similar meteorological conditions, wind speed correlations of multiple wind farms should be generally considered for a stochastic optimization problem. In this paper, the correlation coefficient denoted as C_{WPP} with the range [-1, 1] is introduced to quantify how well wind speeds at two sites follow each other. With the probabilistic wind speed model in (1), the maximal power production of a WPP is determined from the speed-power curve by (2) [5].

$$P_{\omega}^{W,max} = \begin{cases} 0, & (v_{\omega} < v_{ci},\ v_{\omega} > v_{ct}) \\ P_{rated} \cdot (\dfrac{v_{\omega} - v_{ci}}{v_{rd} - v_{ci}}), & (v_{ci} \le v_{\omega} \le v_{rd}) \\ P_{rated}, & (v_{rd} < v_{\omega} < v_{ct}) \end{cases} \quad (2)$$

where P_{rated} is the rated power; v_{ci}, v_{rd} and v_{ct} are the cut-in, rated and cut-out wind speed, respectively.

2) Stochastic model of EV maximal power production

In this paper, the maximal charging/discharging power of an aggregator is calculated by multiplying the number of EVs N_{EV} with the EV average rated power. According to [32], the number of EVs connected at an aggregator could be described by Gaussian distribution, and therefore the stochastic maximal power production of an EV aggregator could be modeled as (3).

$$f(P_e^{E,\max}) = \frac{1}{\sqrt{2\pi}\sigma_p} e^{-(P_e^{E,\max} - \mu_p)^2/2\sigma_p^2} \quad (3)$$

where μ_p and σ_p are the mean and standard deviation of Gaussian distribution. In a general form, the maximal power production correlations of multiple EV aggregators are also represented by the correlation coefficients C_{EV} in this paper.

3) Stochastic model of bid prices for non-strategic participants

The bid price of non-strategic participants at each hour is characterized by the log-normal distribution as (4) [33].

$$f(\lambda_n^{DA}, \mu_q, \sigma_q) = \frac{1}{\lambda_n^{DA} \sigma_q \sqrt{2\pi}} e^{-(\ln \lambda_n^{DA} - \mu_q)^2/2\sigma_q^2} \quad (4)$$

where μ_q and σ_q are the mean and standard deviation of log-normal distribution, respectively. λ_n^{DA} is the bid price of the non-strategic participant. Similarly, the correlation coefficients C_{BID} is used to represent the correlated bid price of multiple non-strategic participants.

Scenario generation strategy: The above three stochastic models can be unified and denoted by the uncertain variables $\boldsymbol{X}$ with the correlations $\boldsymbol{R}_x$, and the Monte Carlo sampling method incorporated with the Cholesky decomposition strategy is introduced to generate representative scenarios for correlated uncertain variables via the following steps; 1) For the given probabilistic distributions of uncertain variables $\boldsymbol{X}$ with the correlations $\boldsymbol{R}_x$, build a correlation coefficient matrix $\boldsymbol{R}_y$ using $\boldsymbol{R}_y = G(\boldsymbol{R}_x)*\boldsymbol{R}_x$, where $G(\boldsymbol{R}_x)$ is the correlation coefficient shift function associated with the specific type of probability distributions, which could be obtained by using Tables 4-8 in [34]; 2) Apply the Cholesky decomposition to $\boldsymbol{R}_y$ to obtain an orthogonal matrix $\boldsymbol{B}$, namely $\boldsymbol{R}_y = \boldsymbol{B}*\boldsymbol{B}^T$ where T denotes the transpose of a matrix; 3) Generate a sample matrix $\boldsymbol{Z}$ of independent standard normal variables, for example using the statistical analysis function normrnd() in MATLAB, and afterward obtain the correlated standard normal matrix $\boldsymbol{Y}=\boldsymbol{B}^{-1}*\boldsymbol{Z}$; 4) Apply the transformation $\boldsymbol{S}=H^{-1}(\varphi(\boldsymbol{Y}))$ to generate the final scenarios $\boldsymbol{S}$ for the original uncertain variables $\boldsymbol{X}$, where φ is the cumulative distribution function (CDF) of the standard normal distribution, and H is the CDF of the input variables $\boldsymbol{X}$.

Scenario reduction strategy: Since computational burden would increase with the number of scenarios in consideration, associated optimization problems may become intractable as the cardinality of scenario sets increases. As a result, the forward selection method [35] is used as the scenario reduction technology in this paper to curtail the number of scenarios considered, while minimizing the inevitable dilution of stochastic information contained in the original set. By minimizing the Kantorovich distance between the initial scenario set and the reduced scenario set, the forward selection algorithm recursively adds scenarios from the initial set to the reduced set until the latter totals a desired number of constituent members.

Lastly, based on these scenario generation and reduction strategies, the stochastic features of the maximal power production of WPPs and EV aggregators, and the bid prices of non-strategic participants modeled in (1)-(4) would be represented by a proper set of scenarios. In specific, the maximal power production of the WPP and EV aggregator, the bid price of the traditional generator and load are respectively denoted by $P_{\omega,\alpha,t}^{W,max}$, $P_{e,\alpha}^{E,max}$, $\lambda_{gn,\alpha,t}^{DA}$ and $\lambda_{ln,\alpha,t}^{DA}$. And the expected profits of WPPs and EV aggregators and the social welfare from the perspective of ISO can also be expressed by these scenarios and their probabilities.

The uncertainties associated with stochastic resources may introduce risk into the optimized bidding problem. In such a condition, risk measuring categorized into risk-neutral and risk-averse can be used to provide guiding information to bidding decision makers. In general, the risk-averse bidding strategy is relatively conservative [36]. The risk-neutral decision maker has more opportunities to dynamically adapt the bidding strategy according to the information observed for increasing the expected profit, and thus is considered in this paper. In addition, the optimized objective function is represented as an expectation by selected scenarios and relevant weights, and all constraints should be satisfied for all scenarios. In this way, risk of profit caused by the volatility of uncertain resources could be partly reduced. More specific risk metrics in [36-39] would be studied in future work.

A. Clearing model (Lower level)

In the lower level, the ISO first collects bids from all market participants and then completes market clearing. The ISO aims to ensure the efficiency of electricity market, which is characterized by maximized social welfare. As a result, the objective function in the lower level is maximizing social welfare. With the help of traditional optimal power flow (OPF) method to clear the market, the dispatched power production $P_{\omega,\alpha,t}^{W,DA}$ and $P_{e,\alpha,t}^{E,DA}$ for WPPs and EV aggregators respectively and LMP $\varphi_{n,\alpha,t}^{DA}$ will be returned to the upper level for maximizing revenues of strategic WPPs and EV aggregators. The bid prices $\lambda_{\omega n,t}^{bid}$ and $\lambda_{en,t}^{bid}$ are bidding strategy for the WPP ω and the EV aggregator e at node n and time t respectively.

Maximize

$$\sum_{\alpha\in N_\alpha}\tau_\alpha\cdot(\sum_{l\in\Omega_n^L}\lambda_{ln,\alpha,t}^{DA}\cdot P_{l,\alpha,t}^{L,DA}-\sum_{g\in\Omega_n^G}\lambda_{gn,\alpha,t}^{DA}\cdot P_{g,\alpha,t}^{G,DA} -\sum_{\omega\in\Omega_n^W}\lambda_{\omega n,t}^{bid}\cdot P_{\omega,\alpha,t}^{W,DA}-\sum_{e\in\Omega_n^E}\lambda_{en,t}^{bid}\cdot P_{e,\alpha,t}^{E,DA}) \tag{5}$$

Subject to

$$P_{g,\alpha,t}^{G,DA}+P_{\omega,\alpha,t}^{W,DA}+P_{e,\alpha,t}^{E,DA}-\sum_{m\in\Omega_n^N}B_{nm}\cdot(\theta_{n,\alpha,t}^{DA}-\theta_{m,\alpha,t}^{DA}) =P_{l,\alpha,t}^{L,DA}:\varphi_{n,\alpha,t}^{DA},\forall n,t,\alpha \tag{6}$$

$$-f_{nm}^{\max}\le B_{nm}\cdot(\theta_{n,\alpha,t}^{DA}-\theta_{m,\alpha,t}^{DA})\le f_{nm}^{\max},\forall n,m\in\Omega_n^N,t,\alpha \tag{7}$$

$$0\le P_{g,\alpha,t}^{G,DA}\le P_g^{\max},\forall g,t,\alpha \tag{8}$$

$$P_{l,t}^{\min}\le P_{l,\alpha,t}^{L,DA}\le P_{l,t}^{\max},\forall l,t,\alpha \tag{9}$$

$$0\le P_{\omega,\alpha,t}^{W,DA}\le P_{\omega,\alpha,t}^{W,\max},\forall\omega,t,\alpha \tag{10}$$

$$T_{urgent}=\left\lceil (SOC_{e,\alpha,end}-SOC_{e,\alpha,t\text{-}1})E_e^{\max}/P_{e,\alpha}^{E,\max}\right\rceil, \forall e,t,\alpha \tag{11}$$

$$-P_{e,\alpha}^{E,\max}\le P_{e,\alpha,t}^{E,DA}\le P_{e,\alpha}^{E,\max},\forall e,\alpha,t<24-T_{urgent} \tag{12a}$$

$$P_{e,\alpha,t}^{E,DA}=-P_{e,\alpha}^{E,\max},\forall e,\alpha,t\ge 24-T_{urgent} \tag{12b}$$

$$\theta_{n,\alpha,t}^{DA}=0,\forall t,\alpha,n:ref \tag{13}$$

$$-\pi\le\theta_{n,\alpha,t}^{DA}\le\pi,\forall t,\alpha,n\setminus n:ref \tag{14}$$

$$SOC^{\min}\le SOC_{e,\alpha,t}\le SOC^{\max},\forall e,t,\alpha \tag{15}$$

$$SOC_{e,\alpha,t}=SOC_{e,\alpha,t\text{-}1}-(P_{e,\alpha,t}^{E,DA}/\upsilon_{dis}\cdot\Delta t/E_e^{\max}),\forall e,t,\alpha,P_{e,\alpha,t}^{E,DA}\ge 0 \tag{16a}$$

$$SOC_{e,\alpha,t}=SOC_{e,\alpha,t\text{-}1}-(P_{e,\alpha,t}^{E,DA}\cdot\upsilon_c\cdot\Delta t/E_e^{\max}),\forall e,t,\alpha,P_{e,\alpha,t}^{E,DA}<0 \tag{16b}$$

where $\lceil a\rceil$ is the round-up calculation of a.

The objective function (5) is to maximize social welfare. The first term is the revenues of selling electricity to load demands, while the other three terms represent the costs of purchasing electricity from traditional generators, WPPs and EV aggregators. The constraint (6) is the power production and consumption balance for node n with a dual variable $\varphi_{n,\alpha,t}^{DA}$ donating the LMP that would be provided to the upper-level. Inequality (7) limits the thermal capacity of the transmission line. Maximum and minimum energy dispatched for traditional units, load demand and WPPs are constrained in (8), (9) and (10) respectively. It is noted that a load demand consists of the fixed and curtailed parts represented by the minimum load $P_{l,t}^{min}$ and variable load $P_{l,\alpha,t}^{L,DA}$ respectively in (9). Equation (11) calculates the threshold hour of fully charging EV aggregator by the maximum charging power. Constraint (12a) applies when the EV aggregator is not in urgent charging period and could participate in the power market, while constraint (12b) applies when the EV aggregator is in urgent charging period so that EV aggregator is charged at the maximum power to satisfy the daily driving utilization. Equation (13) and inequality (14) set voltage angle limits at the slack bus and other buses respectively. Inequation (15) represents the SOC range of an EV aggregator at the present hour, while constraints (16a) and (16b) indicate time-series SOC formulation of an EV aggregator at present and the previous hours.

B. The suppliers' bidding problem (Upper level)

The proposed bidding problem considers the output capacity uncertainties of WPPs and EV aggregators as a set of scenarios. Strategic oligopolists' revenues for the ωth WPP and the eth EV aggregator are expressed by these scenarios with corresponding probabilities and represented in the upper level by (17) and (18) respectively, where LMP $\varphi_{n,\alpha,t}^{DA}$, scheduled WPP power $P_{\omega,\alpha,t}^{W,DA}$ and EV aggregator output $P_{e,\alpha,t}^{E,DA}$ are obtained from the market clearing in the lower level. While the revenue of WPP is presented in (17), the revenue of an EV aggregator in (18) includes the income of selling electricity to power market in the first term, the cost of buying electricity from EV owners in the second term and the battery degradation cost in the third term. Absolute-value function in (18) could be handled by a linear programming simplex method in [40].

$$R_\omega=\sum_{\alpha\in N_\alpha}\tau_\alpha\cdot(\varphi_{n,\alpha,t}^{DA}\cdot P_{\omega,\alpha,t}^{W,DA}) \tag{17}$$

$$R_e=\sum_{\alpha\in N_\alpha}\tau_\alpha\cdot(\varphi_{n,\alpha,t}^{DA}\cdot P_{e,\alpha,t}^{E,DA}-\lambda_{ev}\cdot P_{e,\alpha,t}^{E,DA} -(C_b/L_b)\cdot\left|P_{e,\alpha,t}^{E,DA}\right|) \tag{18}$$

In this paper, the game problem Ψ among strategic suppliers including WPPs and EV aggregators is defined by a set of nodes with strategic players, bid prices and revenues as (19).

$$\psi=\begin{Bmatrix}\Omega_n^W,\{\lambda_\omega^{bid}\}_{\omega\in\Omega_n^W},\{R_\omega\}_{\omega\in\Omega_n^W},\\ \Omega_n^E,\{\lambda_e^{bid}\}_{e\in\Omega_n^E},\{R_e\}_{e\in\Omega_n^E}\end{Bmatrix} \tag{19}$$

It is obvious that the revenues of other rivals are influenced by every strategic supplier's bid price through the clearing process in (5)-(16b). As there is no information of other strategic suppliers' profit functions and historical bidding decisions, the decision process of bidding strategies is a game problem rather than an optimization. As a result, these strategic players are required to learn for their optimal bids by repeated interaction with the market. In the following Section III, strategic bidding behaviors of WPPs and EV aggregators in a competitive market with incomplete information are modeled as stochastic games, where the market environment is stochastic and determined by both bid prices and characteristics of all players. Suppliers including WPPs and EV aggregators are considered as players of the game. Then the bidding strategy is explored using a multi-agent reinforcement learning (MARL) algorithm *WoLF-PHC* as a decision support tool.

III. Methodology

In this part, a brief summary of the RL theory and MARL is first introduced, followed by a stochastic game framework for a DA electricity market. Afterward, a MARL method *WoFL-PHC* and the utilization of *WoLF-PHC* for multi-agent bidding strategies are described in detail.

A. Description of RL theory and the MARL

RL is an area of machine learning developed to analyze animal and artificial behavioral systems [41]. The agent updates its decision relying on instant feedbacks and gradual learning through repetitive interaction with the external environment, which is quite similar to the relationship between strategic energy suppliers and the market. *Q-learning* is one of frequently used RL approaches. The MARL is developed based on the single-agent RL to decide the optimal action of every agent in multi-agent domains. It combines the single-agent RL, game theory and policy research techniques, and hence is more suitable for solving bidding problems of multi-supplier in electricity market. However, due to the alterability and unpredictability of multi-agent dynamic environments, it would be more difficult and challenging for each agent to learn with other agents as moving targets. Next, a stochastic game framework is introduced to describe this complicated problem of multi-agent bidding decisions in electricity market.

B. The stochastic game framework of proposed DA electricity market

The multi-agent stochastic game could be described as a tuple (*M, X, A, T, R*) where $M = \{1, 2, \dots, m\}$ denotes a set of agents, X is a set of game states $\{x_i\}$, A=$\{a_1, \dots, a_j, \dots, a_m\}$, in which a_j=$\{aa_{min}, \dots, aa_{max}\}$ represents the sects of actions available to any player a_j, T is the transition function represented by $X \times A \times X \rightarrow [0,1]$, and R=$\{R_1,\dots, R_j \dots, R_m\}$ shows the set of reward functions for all agents in which R_j: $(x_i, a_j) \rightarrow \mathcal{R}$ is the reward function of the *j*th agent in the state x_i with executing the action a_j. At each step, all agents observe the state $x_i \in X$ and choose to perform the action a_j according to an optimal action selection policy of learning algorithm, then go to the next state $x_i \in X$.

The proposed multi-agent market model is expressed as a stochastic game framework. There exist two types of agents, which are the WPP $\omega \in \Omega_n^W$ and the EV aggregator $e \in \Omega_n^E$. M=$\{\omega \in \Omega_n^W,\ e \in \Omega_n^E\}$ in the proposed model. States of the stochastic game consider different levels of WPP and EV aggregator suppliers' capacities represented as two sets $\{x_\omega\}_{\omega\in\Omega_n^W}$ and $\{x_e\}_{ej\in\Omega_n^E}$ respectively. The market clearing process, representing the interactive environment, would provide the signal of scheduled power production $P_{\omega,\alpha}^{W,DA}$ and $P_{e,\alpha}^{E,DA}$ to every agent for WPPs and EV aggregators respectively. In this way, a state would be chosen for each agent after every market clearing. Next, admissible actions A=$\{a_j\}$ are defined for agents of WPPs and EV aggregators to update bid prices $\{\lambda_{\omega j}^{bid}\}_{\omega j\in\{a_j\}}$ and $\{\lambda_{ej}^{bid}\}_{ej\in\{a_j\}}$ respectively. As to the returned reward functions in (13) and (14), they could be expressed by $\{R_{\omega j}\}_{\omega j\in\{a_j\}}$ and $\{R_{ej}\}_{ej\in\{a_j\}}$ where $R_{\omega j}$: $(x_\omega, a_{\omega j}) \rightarrow \mathcal{R}$ is the payoff of the *ω*th agent after clearing the market with bidding $\lambda_{\omega j}^{bid}$ in the WPP' capacity level x_ω, and R_{ej}: $(x_e, a_{ej}) \rightarrow \mathcal{R}$ is the *e*th agent' profit after providing bid price λ_{ej}^{bid} in the EV aggregators' capacity level x_e to the clearing market. The decision process involves the choice of bidding actions for each player can be defined as B=$\{\{\lambda_{\omega j}^{bid}\}_{\omega j\in\{a_j\}}, \{\lambda_{ej}^{bid}\}_{ej\in\{a_j\}}\}$. Then their own perception of states would be got through executing corresponding bidding decisions in the stochastic market, in which the set of states is described as X=$\{\{x_\omega\}_{\omega\in\Omega_n^W}, \{x_e\}_{ej\in\Omega_n^E}\}$. Consider the property of stochastic game and the decision vectors B of players are in a competitive environment, the combined process given by B and M is a competitive stochastic game with incomplete information.

In order to determine an action to update the bid price for every player in the market environment, a policy, p: $X \times A \rightarrow [0,1]$, is introduced, which states the probability of the algorithm choosing an available action a_j, based on the game is in the state x_i. As a result, the target in this multi-agent stochastic game framework is to find a suitable algorithm that every agent could learn a policy along with others learning simultaneously. Here, the *WoLF-PHC* would be introduced to model the learning process in stochastic games.

C. The MARL method WoLF-PHC

Policy Hill Climbing (*PHC*) is a simple extension of *Q-learning,* the policy of which is evolved by increasing the probability of selecting the action with the highest value with a learning rate $\delta \in (0,1]$ and performs hill-climbing in the space of mixed policies. However, its convergence is unclear [31]. *PHC* is then further developed to have a variable learning rate, consisting of two learning parameters δ_w and δ_l standing for the win and lose, respectively. The variable learning rate introduced aids in the convergence, with leaving more time for rivals to adjust when the player is gainful or compelling the player to adapt more quickly to rivals' strategy changes once its interest is damaged [31]. The unknown equilibrium policy can further be replaced by a more general average strategy to asymptotically approximate to the equilibrium, such that the agent could determine its win or lose by comparing the expected and average payoff. Specifically, δ_l should be greater than δ_w. The larger learning rate δ_l means agents could learn quickly to adjust their strategies after losing, while the smaller learning rate δ_w is used to remain caution if it wins. This is called the win or learn fast (*WoLF*) principle. The final algorithm is named as *WoLF-PHC* and described as follows.

The transition from the last state to the current state comes as a result of the market clearance. For a given agent *i*, Q-function after the *k*th market clearance based on an exploration action a_{jk} under the past state $x_{i(k-1)}$ and the current state x_{ik} with a reward function R_{jk} is updated as (20).

$$\begin{aligned} Q_i(x_{i(k-1)}, a_{jk}) \leftarrow\ & (1-\mu)Q_i(x_{i(k-1)}, a_{jk}) \\ & +\mu(R_{jk} + \eta \max_{a'_{jk}} Q_i(x_{ik}, a'_{jk})) \end{aligned} \tag{20}$$

The corresponding policy p_i is updated as (21)-(24). δ_{ik} is kept within (0,1] in (23), policy for the last state is updated in (21)-(22) by increasing the probability of the action with the maximum Q-value or decrease probabilities of other actions in the last state. The maximized Q-value is used to determine the probability distribution of actions for the last state, which is prepared for the next visit in this state. δ_{ik} refers to the variable learning rate for the agent *i* at the *k*th iteration, in which δ_ω is chosen and the agent is winning if the past expected Q-value is larger than the expected Q-value with the average policy $\overline{p_i}$, otherwise δ_l is selected.

$$p_i(x_{i(k-1)}, a_{jk}) \leftarrow p_i(x_{i(k-1)}, a_{jk}) + \Delta_{x_{i(k-1)}a_{jk}} \tag{21}$$

$$\Delta_{x_{i(k-1)}a_j} = \begin{cases} -\delta_{x_{i(k-1)}a_{jk}}, \\ if a_{jk} \neq \arg\max_{a'_{jk}} Q_i(x_{i(k-1)}, a'_{jk}) \\ \sum_{a'_{jk} \neq a_{jk}} \delta_{x_{i(k-1)}a'_{jk}}, otherwise \end{cases} \tag{22}$$

$$\delta_{x_{i(k-1)}a_{jk}} = \min(p_i(x_{i(k-1)}, a_{jk}), \delta_{ik} / (|A|-1)) \tag{23}$$

$$\delta_{ik} = \begin{cases} \delta_\omega, if \sum_{a'_{jk}} p_i(x_{i(k-1)}, a'_{jk}) Q_i(x_{i(k-1)}, a'_{jk}) \\ > \sum_{a'_{jk}} \overline{p_i}(x_{i(k-1)}, a'_{jk}) Q_i(x_{i(k-1)}, a'_{jk}) \\ \delta_l, otherwise \end{cases} \tag{24}$$

where $\overline{p_i}$ in (24) is the average policy updated as in (25)-(26), $c(x_i)$ is the total number of the state x_i from the initial state to the current state, and the Q-value, policy and average policy are required to be updated based on the previous state.

$$c(x_i) \leftarrow c(x_i) + 1 \tag{25}$$

$$\begin{aligned} \overline{p}_i(x_{i(k-1)}, a'_{jk}) \leftarrow \overline{p}_i(x_{i(k-1)}, a'_{jk}) + \frac{1}{c(x_i)} (p_i(x_{i(k-1)}, a'_{jk}) \\ - \overline{p}_i(x_{i(k-1)}, a'_{jk})), \forall a'_{jk} \in A_i \end{aligned} \tag{26}$$

D. *Implementation of WoLF-PHC for Suppliers' Bidding Strategies*

In the proposed model, WPPs and EV aggregators would not share any private information with other rivals and make self-determined biddings to increase their respective profits. The proposed model thereby is a multi-agent stochastic game problem with incomplete information. Considering that *WoLF-PHC* is a decentralized self-game algorithm which could well counterbalance the lack of information, a multi-agent decentralized *WoLF-PHC* is therefore applied in this stochastic market environment for WPPs and EV aggregators to make bidding decisions. The specific learning procedures for the ωth WPP and eth EV aggregator strategically bidding through *WoLF-PHC* could be described as follows. Each agent would repeat steps (a)-(h) in (2) until the count is met.

(1) Set learning rate $\mu \in (0,1]$, discount factor $\eta \in (0,1]$ and learning parameters used to update the policy $\delta_l > \delta_w \in (0,1]$. Initialize Q-value Q_ω and Q_e, policy p_ω and p_e, and the count of states $c(x_\omega)$ and $c(x_\omega)$ as (27) and (28).

$$Q_\omega(x,a) \leftarrow 0,\ p_\omega(x,a) \leftarrow 1/|A|,\ c(x_\omega) \leftarrow 0 \tag{27}$$

$$Q_e(x,a) \leftarrow 0,\ p_e(x,a) \leftarrow 1/|A|,\ c(x_e) \leftarrow 0 \tag{28}$$

(2) Repeat, in the kth episode,

(a) According to policy $p_\omega(x_{\omega(k-1)}, a_{\omega jk})$ and $p_e(x_{e(k-1)}, a_{ejk})$, choose corresponding actions $a_{\omega jk}$ and a_{ejk} respectively.

(b) The ωth WPP and eth EV aggregator' bid prices $\lambda^{bid}_{\omega j}$ and λ^{bid}_{ej} are updated in terms of actions $a_{\omega jk}$ and a_{ejk} selected in step (a).

(c) Bid prices $\lambda^{bid}_{\omega j}$ and λ^{bid}_{ej} updated in step (b) are sent to the ISO in the lower level of the proposed market model, with LMPs and scheduled WPP's and EV aggregator's productions are obtained from the clearing process (5)-(16b) and are provided to the upper level. Then reward functions $R_{\omega jk}$ and R_{ejk} are calculated based on (17) and (18).

(d) Q-functions $Q_\omega(x_{\omega(k-1)}, a_{\omega jk})$, $Q_e(x_{e(k-1)}, a_{ejk})$ are updated by observing scalar rewards $R_{\omega jk}$, R_{ejk} in step (c), and the current state $x_{\omega k}$, x_{ek} as (20).

(e) Observe every action $a'_{\omega jk}$ and a'_{ejk} for states $x_{\omega(k-1)}$ and $x_{e(k-1)}$ respectively, Q_ω and Q_e related to each pair $(x_{\omega(k-1)}, a'_{\omega jk})$ and $(x_{e(k-1)}, a'_{ejk})$.

(f) Update average polices $\bar{p}_\omega(x_{\omega(k-1)}, a'_{\omega jk})$ and $\bar{p}_e(x_{e(k-1)}, a'_{ejk})$ as (25) and (26).

(g) Update the variable learning rate $\delta_{\omega k}$ and δ_{ek} relied on $Q_\omega(x_{\omega(k-1)}, a'_{\omega jk})$ and $Q_e(x_{e(k-1)}, a'_{ejk})$ in step (e) and $\bar{p}_\omega(x_{\omega(k-1)}, a'_{\omega jk})$ and $\bar{p}_e(x_{e(k-1)}, a'_{ejk})$ in step (f) as (24). Then policies $p_\omega(x_{\omega(k-1)}, a_{\omega jk})$, $p_e(x_{e(k-1)}, a_{ejk})$ are updated according to as (21)-(23).

(h) Set k=k+1, return to step (a).

The above learning process for a WPP and an EV aggregator developing respective bidding strategies through their interactions with the market is shown in Fig.2, and it could be extended and scaled to have more strategic players. The *WoLF-PHC* to solve the proposed bidding problem for the agent i is represented as Fig.3.

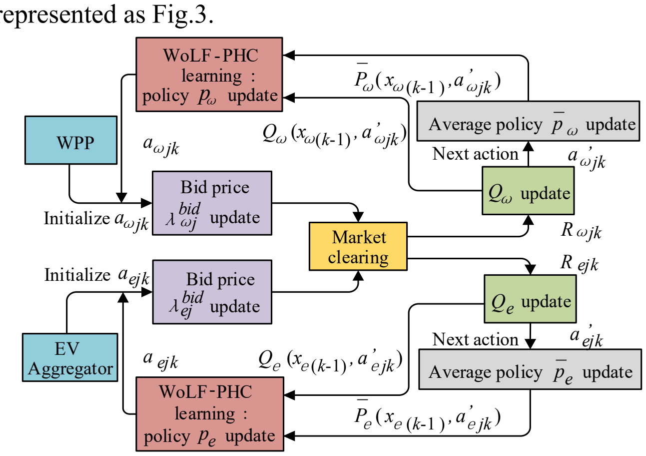


Fig. 2. The specific learning process for a WPP and an EV aggregator strategically bidding through *WoLF-PHC*

Remarks: The proposed model is for a single hour of the DA market and can be considered as an offline approach. Since the model involves a large number of scenarios, the calculation process could be time consuming. The *WoLF-PHC* pre-learning process [42] could therefore be firstly introduced to boost the computation efficiency, which uses the optimized Q-value and bid price in the previous hour as the initial point of the current *WoLF-PHC* computation such that the convergence rate can be accelerated [42]. In addition, there are numerous scenarios in solving the DA market clearing, scenario-parallel computing could further be used to reduce the computation time.

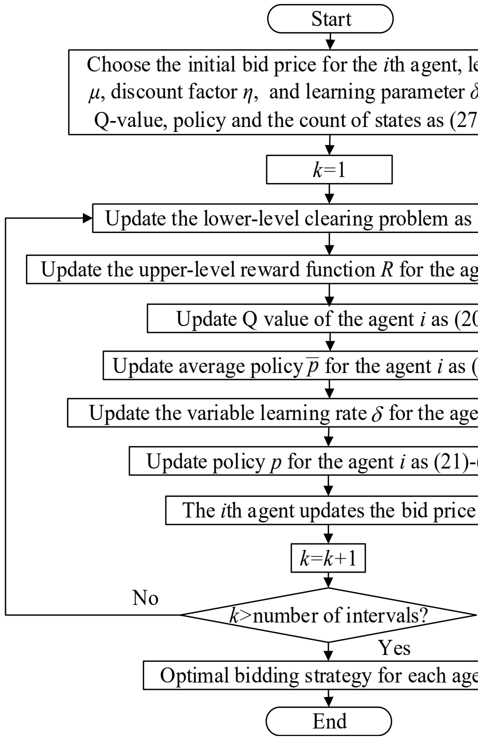


Fig. 3 Algorithm to solve the proposed bidding problem

IV. Case Studies

Parameters of the three sources of uncertainties are configured as below. The shape and scale parameters of wind speed are set as λ=2 and k=12 with correlations of 0.3, while the cut-in, cut-out and rated wind speed are v_{ci}=3m/s, v_{ct}=25m/s and v_{rd}=12 m/s, respectively [43]. The parameters for the normal distribution of the EV aggregator are set as μ_p=120MW and σ_p =0.1μ_p with correlations of 0.1. The parameters of the log-normal distribution for the stochastic bid prices of non-strategic participants are set as μ_q=40$/MW and σ_q=0.1μ_q with correlations of 0.2. Besides, the price for the EV aggregator traded with EV owners is set as 20$/MW [5]. The lower limit of bid price is the cost price of each energy supplier. Other data for the EV aggregator is shown in Table II and parameters for the *WoLF-PHC* are listed in Table III.

TABLE II
EV Aggregator Parameters

EV Aggregator	SOC^{max} (pu)	SOC^{min} (pu)	υ_c	υ_{dis}	E_{ev}^{max} (MWh)
Value	0.9	0.1	0.9	0.85	1000

TABLE III
Data for the *WOLF-PHC*

Parameter	μ	η	δ_w	δ_l
Value	0.1	0.5	0.01	0.02

Four different cases are studied in this section. Case 1 would first simulate the model proposed in Section II, in which a WPP and an EV aggregator represent two strategic players, and then the proposed model is compared with the cooperative model in [9]. In Case 2, the number of strategic players increases to three to allow a traditional generator to strategically bid and compete with the WPP and EV aggregator. Loads are changed to inelastic. Bidding results would be analyzed according to the relationship between generation supply and load demand. In Case 3, strategic players of the proposed model are expanded to cover loads as well. Four strategic players including a load, a traditional generator, a WPP and an EV aggregator are simulated in the model. In Case 4, the four players model in Case 3 is applied to a modified IEEE 118-bus system with the number of strategic players tripled to twelve, and their bidding results will be fully studied and analyzed. While Case 1 would provide the solutions for both a single hour 20:00 and the successive 24-hour operations to show the time-series performance of the proposed model, Cases 2-4 would only present the results of a single hour 20:00 for illustrating the convergence of bi-level model. All the cases are simulated in MATLAB running on a 1.6 GHz Intel Core i5 -5250U computer for Case 1-3 and a 3.2 GHz Intel Core i7-8700 computer for Case 4 with 8 GB of RAM.

A. Case 1

The proposed model is first tested on the IEEE 6-bus system. There are three traditional generators locating in buses 1-3 and three loads connecting to buses 4-6. A WPP and an EV aggregator representing two strategic participants are set in bus 4 and bus 5, respectively.

In order to generate the appropriate number of scenarios for the stochastic maximized power outputs of WPPs/EV aggregators and bid prices of non-strategic participants, the convergence theory in [43, 44] is used to determine the number of scenarios by formulas (29)-(31).

$$\varepsilon_{\mu} = 100(|\,\mu_{MC} - \mu_{*}\,|) / \mu_{*}[\%] \tag{29}$$

$$\varepsilon_{\sigma} = 100(|\,\sigma_{MC} - \sigma_{*}\,|) / \sigma_{*}[\%] \tag{30}$$

$$\varepsilon_{\lambda} = 100(|\,\lambda_{MC} - \lambda_{*}\,|) / \lambda_{*}[\%] \tag{31}$$

where μ_{MC}, σ_{MC} and λ_{MC} are the mean value, standard deviation and skewness calculated based on the generated scenarios; μ_*, σ_* and λ_* are the mean value, standard deviation and skewness derived theoretically from the proposed model in Section II; ε_μ, ε_σ and ε_λ are the preset accuracy threshold. In this paper, the number of generated scenarios is increased from 100 by a step of 100, and the number of scenarios with ε_μ, ε_σ and ε_λ just smaller than the settled accuracy 1% is deemed as the finally determined numbers. By using this strategy, the uncertainties of the maximized power outputs of WPPs/EV aggregators, and bid prices of non-strategic participants are respectively represented by 1000, 1000, and 500 scenarios in this paper.

In order to reduce the computational burden, the aforementioned forward selection method is used to curtail the number of scenarios and thus a set of the reduced number of scenarios is obtained. In the following, for investigating the impacts of the number of reduced scenarios on the accuracy of the scenario-based market model, comparative studies with different numbers of reduced scenarios are performed. Table IV shows the profits of the WPP/EV aggregator, social welfare and the solution time for different numbers of reduced scenarios at hour 20:00. It can be found that, with the number of reduced scenarios increasing from 125 to 500, profits of the WPP and the EV aggregator as well as the social welfare would be slightly enhanced; whereas as the number of reduced scenarios further increasing beyond 500, there will be very little further improvement. Regarding to the solution time, it would monotonically increase with the growing number of reduced scenarios. Consequently, to make a trade-off between the profits of the WPP/EV aggregator, social welfare and time consumption, 500 reduced scenarios are adopted as an appropriate choice for the following case study.

TABLE IV
Profits of the WPP/EV Aggregator/Social Welfare and Solution Time under Different Numbers of Reduced Scenarios at hour 20:00

Number of Reduced Scenarios	Profit ($)		Social Welfare ($)	Solution Time(s)
	WPP	EV Aggregator		
125	7630.9	2934.1	7352.5	105.68
250	7651.3	2949.6	7529.3	196.87
500	**7688.1**	**2961.2**	**7830.7**	**365.72**
1000	7686.7	2962.7	7832.9	1165.59
2000	7688.5	2961.5	7830.9	2508.94

Based on the parameter set above, the optimized revenues of EV aggregator and WPP in all scenarios are used to display the randomness of the stochastic model. Fig.4 shows the relationship of revenues of the EV aggregator or WPP, as formulated in (17) and (18) respectively, LMPs and power output of the EV aggregator or WPP. The relationship of social welfare as represented in (5) and the EV aggregator/WPP power output is illustrated in Fig.5. The revenues of WPP or EV aggregator and the social welfare are calculated based on 500 scenarios with probabilities derived by the scenario-generation and scenario-reduction algorithm.

In the model, WPP and EV aggregator develop respective bidding strategies through the learning process of *WoLF-PHC* and interact with the market, while their revenues and the social welfare are influenced by constraints of the market clearing model, i.e. the power output limits of WPP and EV aggregator in (10)-(12b) and the SOC limit of EV aggregator in (15)-(16b). As shown in Fig.6, bidding results converge nicely to equilibrium after 50 iterations. The computation time is 365.72s. In this process, there is no complex KKT transformation nor complex calculations for exploring the Nash equilibrium point, which demonstrates that flexible and concise *WoLF-PHC* could be used in the oligopoly electricity market to respectively optimize bid prices for the competitive WPP and the EV aggregator. Besides, each agent makes self-decision on bidding for maximizing its own interest only through the bidding information obtained from the ISO in the market without the cost functions or bidding data from the competitor. In this way, the privacy of personal data is protected.

Comparison with the cooperative model in [9]: In the cooperative model in [9], the WPP and the EV aggregator as a strategic Hybrid Power Plant (HPP) and the objective is to maximize the overall benefit of the HPP model. In order to improve the WPP and EV aggregator's common interests, the HPP owns their bidding information and cost function to strategically bid and dispatch their power as a central market player. The WPP and EV aggregator are fully controlled by the HPP and do not have any self-determination on bidding. Besides, the cooperative model would break if the WPP and EV aggregator do not provide their personal data to the HPP. Whereas, the proposed competitive model could adapt to a more flexible market environment and allows for self-bidding to increase individual revenue of the WPP and EV aggregator separately, which ensures better privacy of strategic players. Table V shows comparative results of the cooperative and competitive models for the two strategic participants in hour 20:00. Compared with the cooperative model, both revenues of the EV aggregator and social welfare are increased while the profit of the WPP is slightly declined. This is due to the WPP's profits are increased by the EV aggregator with the help of the HPP's centralized dispatching in the cooperative model. In the competitive market, the EV aggregator could make self-decision on bidding and reduce its bid price to lower than the WPP to sell more and earn more in return. As a result, the EV aggregator's profit has an increase and the WPP' interest is reduced. Meanwhile, social revenue is improved as competition brings the reduction of bid prices.

In addition, the bidding strategy of the proposed model is also simulated for 24 hours. Scenarios of maximum power production of the WPP and bid prices of non-strategic participants are generated on an hourly basis as in [2, 45]. While the number of EVs connected to the EV aggregator is a normal distribution for 24 hours, and the initial and desired SOC are set as 0.6 and 0.8, respectively [46]. Comparative results of the cooperative and competitive models for 24 hours are provided in Table VI. The revenues of the WPP and EV aggregator in the proposed model outperformance those of the cooperative model, while the social welfare is also higher. Among the 500 reduced scenarios, Fig. 7 shows the SOC curve and scheduled output of the EV aggregator for 24 hours for scenario 1 as an example. As shown in Fig. 7, the EV aggregator is not in urgent charging period before 23:00 and its power output could be scheduled in the power market (refer to the constraint (12a)); whereas for the last several hours, the EV aggregator is fixed at the maximum charging power for preparing the daytime driving utilization (refer to the constraint (12b)). In this way, the EV aggregator strategically participates in the pool-based electricity market as well as satisfy the traffic energy demand.

TABLE V
Revenues Comparison of Cooperative and Competitive Models at hour 20:00

Market Type	WPP Revenues ($)	EV Aggregator Revenues ($)	Social Welfare ($)
Cooperate Model	7724.9	2489.8	4157.6
Competitive Model	7688.1	2961.2	7830.7

TABLE VI
Revenues Comparison of Cooperative and Competitive Models for 24 hours

Market Type	WPP Revenues ($)	EV Aggregator Revenues ($)	Social Welfare ($)
Cooperate Model	131943.8	44519.2	79943.1
Competitive Model	133720.3	45664.8	133674.2

B. *Case 2*

In this case, the number of strategic players increases to three suppliers, with a traditional generation unit, a WPP and an EV aggregator located at bus 1, bus 4 and bus 6, respectively. The total generation capacity is 497.1MW consisting of 174.6MW, 122.5MW and 200MW for the traditional generator, WPP and EV aggregator respectively. The cost price of the traditional unit is set as 40$/MW. Loads are inelastic and set by the following four levels with bidding results analyzed as below.

(1) The total load demand is set equal to the total capacity of three suppliers. Fig.8 shows the market clearing bid prices,

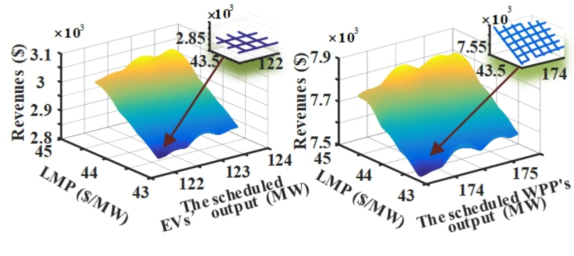


Fig. 4 Revenues of EV and WPP for all scenarios

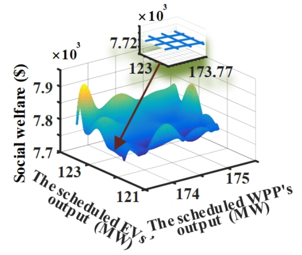


Fig. 5 Social welfare for all scenarios

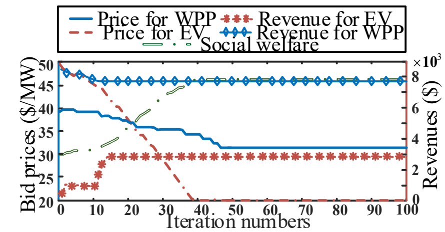


Fig. 6. Market clearing results for the WPP and EV

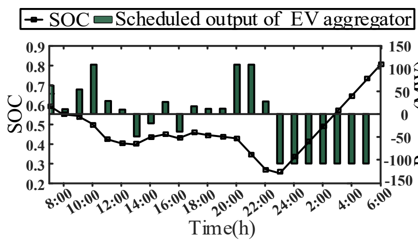


Fig. 7. SOC and scheduled output of the EV aggregator for scenario 1 during 24 hours

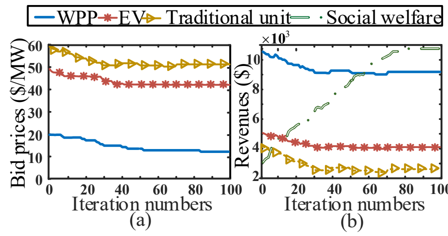


Fig. 8. Market clearing results of 3 players for total demand equal to total generation capacity in 6-bus system

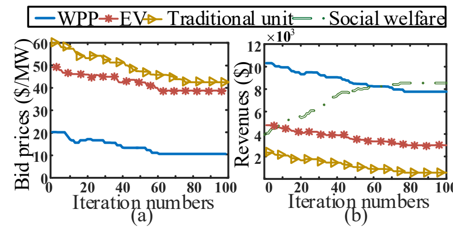


Fig. 9. Market clearing results of 3 players for total demand less than total generation capacity in 6-bus system

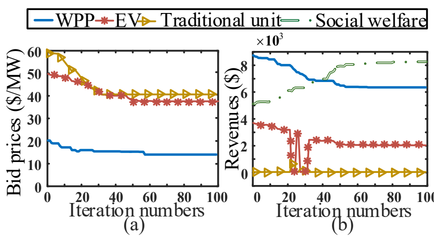


Fig. 10. Market clearing results of 3 players for total demand equal to total generation capacity of WPP and EV aggregator in 6-bus system

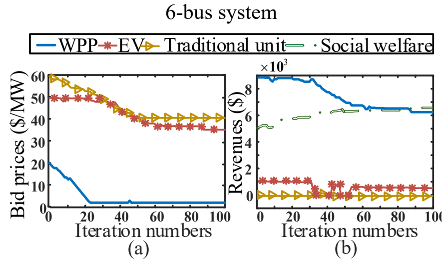


Fig. 11. Market clearing results of 3 players for total demand less than total generation capacity of WPP and EV aggregator in 6-bus system

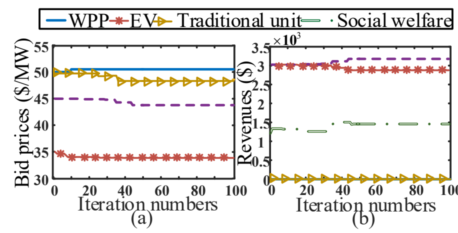


Fig. 12. Market clearing results for 4 players in 6-bus system

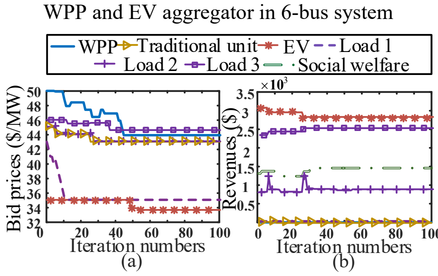


Fig. 13. Market clearing results for 6 players in 6-bus system

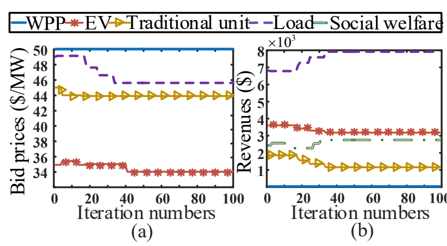


Fig. 14. Market clearing results for 4 players in 118-bus system

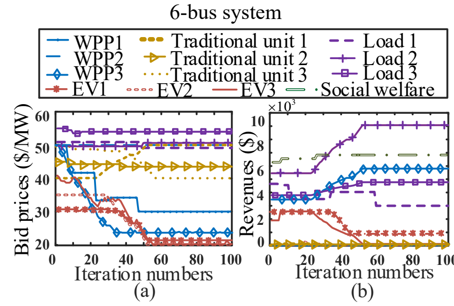


Fig. 15. Market clearing results for 12 players in 118-bus system

revenues of the three suppliers and the overall social welfare. The computation time is 250.87s. It is observed that the game for these three strategic players converges to its equilibrium, in which their bid prices are much higher than their cost prices. This is because there is no competitive relationship among players in this case, they have no incentive to decrease the bid prices and every supplier raises its bid price to gain more profits.

(2) The total load demand is reduced and set to 420 MW, and in this condition the total power supply is higher than the total load demand. Hence, the relationship among these three suppliers becomes competitive. Fig.9 shows curves of bid prices and revenues as well as the social welfare obtained in 100 iterations of WoLF-PHC with the computation time of 238.08s. The curves converge towards the equilibrium with ever decreasing amplitude. It is obvious that three players' bid prices are reduced compared to those in Fig.8, which shows the competition forces every player to cut down its prices for selling more product.

(3) The total load demand is further reduced to 322.5MW, which is equal to the total capacity of the WPP and the EV aggregator. The converged bid prices and revenues of these three players as well as the social welfare are shown in Fig.10. In this way, the total load demand can be covered just by two generation units with lower cost prices. The traditional supplier with relatively higher cost price has little chances to sell products and earn profits. It is obvious in Fig.10 that none has interest to buy from the traditional supplier although its bid price has been reduced to its cost price. Meanwhile, WPP and EV aggregator have raised their prices to earn more with the WPP being the first to run out due to its lower bid prices. Afterward, if the bid price of the EV aggregator exceeds the traditional generator, the market becomes competitive. As a result, the EV aggregator would ensure its bid price slightly lower than that of the traditional supplier, which is consistent with curves in Fig.10. The computation time is 255.78s.

(4) The total load demand is further reduced to below total

capacity of the WPP and EV aggregator to form a competitive relationship between these two suppliers. Bid prices combined with the social welfare are shown in Fig.11 and the computation time is 251.38s. Similar to the previous case shown in Fig.10, the traditional generator bids according to its cost price and has no revenue. The competition drives the EV aggregator and the WPP to lower its bid price than that in Fig.10.

C. *Case 3*

Loads are reset as elastic with curtailed parts and considered as a load aggregator. In order to expand the model for more participants, four strategic players consisting of an elastic load, a traditional generator, a WPP and an EV aggregator are investigated in this case. Fig.12 shows the bid prices and revenues of these four strategic players as well as the social welfare. The computation time is 265.23s and the convergence performance is good. These imply that the proposed model is suitable for traditional units to participate in strategically bidding and works well for different typical strategic players participated in the market. Next, three energy suppliers remain the same as before, and the load demand is split into three agents with the capacity equal to 1/3 of total capacity of the original load. The computation time is 289.72s. Bidding results of these six players and the social welfare are plotted in Fig.13. The resulted bid prices are much lower compared with the ones with four players in Fig.12. This implies that players would adopt a relatively conservative behavior in a more competitive environment. Also noted is that WPP and traditional generator with higher bid prices will have little chance to sell their power generation and the load demand can be preferentially supported by the EV aggregator with a lower price.

D. *Case 4*

The model of four strategic players in Case 3 is applied to a modified IEEE 118-bus system, in which the WPP, EV aggregator and traditional generator are located at nodes 32, 49 and 94, respectively, and loads are considered as a load aggregator. Both the bid prices and the social welfare are nicely converged as shown in Fig.14. It shows that *WoLF-PHC* could successfully obtain the optimal bid price for every strategic player in a larger fully competitive electricity market. Each of these four strategic players is then duplicated to become three identical players at the same location while loads are equally assigned to three load aggregators, and thereby there are twelve strategic players in total. The game of these twelve strategic players in the modified IEEE 118-bus system converges to its equilibrium with bid prices and the social welfare plotted in Fig.15. It confirms that the model with *WoLF-PHC* could be successfully used for more players to strategically bid in a large scale market. In addition, the overall bid prices of twelve strategic makers are kept at lower level than the case with four strategic players, which is consistent with Case 3 that bid prices would be reduced in a more competitive environment. With the use of the afore mentioned *WoLF-PHC* pre-learning process and 6 parallel computation threads running on a 6-core 3.2 GHz Intel Core i7-8700 computer, the computation times for market clearing with four and twelve strategic players are reduced to 735.09s and 946.72s respectively, and would satisfactorily meet the offline market clearing requirement for the proposed DA market model.

The objective of the proposed model is to maximize the social welfare as the priority, while with the LMP $\varphi_{n,\alpha,t}^{DA}$ interacted by the market-clearing in the lower level, the WPP and EV aggregator participants in the upper level make strategic biddings for increasing their respective revenues. As can be observed from Fig.6-15, the social welfare is increased on different levels, while the revenues of WPP and EV aggregator participants in some case studies would increase during the iteration process while others would reduce. Furthermore, the revenue of a particular market player would go up or down depending on the position of the final solution relative to the initial point of the *WoLF-PHC* algorithm. And the revenues of a group of market players, such as WPPs, EV aggregators, generators and loads, are not necessary always going up or down together, as demonstrated in Fig.15(b) of Case 4 where the revenue of Load 1 is reduced while those of Load 2 and 3 are increased.

V. Conclusion

A new competitive bidding market model with incomplete information for considering the uncertainties in bid prices of non-strategic participants and maximal power productions of WPPs and EV aggregators is presented in this paper. A recently developed MARL algorithm named *WoLF-PHC* is adopted to successfully solve the proposed model for strategic players to optimize their bidding in an oligopoly electricity market with personal privacy protection and respecting the autonomy of strategic suppliers. The market is simulated as a multi-agent based system, with three test cases built on a modified IEEE 6-bus system and a larger case study based on a modified IEEE 118-bus system, the bid result in four cases is nicely converged to the equilibrium. Promising conclusions drawn from these case studies include 1) multiple participants could respectively optimize their bids by learning using the *WoLF-PHC* algorithm in competitive electricity markets; 2) compared with the cooperative model of the WPP and EV aggregator in a previous study, the proposed competitive model is able to adapt to a more flexible market environment in which every strategic player has full autonomy in biddings with incomplete information to maximize its own profit; 3) bid prices of market players would be reduced with more competition brought from either the decreased demand or the increased number of strategic participants.

References

[1] M. Z. Jacobson, "100% Clean and renewable wind, water, and sunlight all-sector energy roadmaps for 139 countries of the world," *Joule*, vol. 1, no. 1, pp. 108-121, 2017.

[2] L. Baringo and A. J. Conejo, "Strategic offering for a wind power producer," *IEEE Transactions on Power Systems,* vol. 28, no. 4, pp. 4645-4654, 2013.

[3] X. Lu, K. W. Chan, S. Xia, X. Zhang, G. Wang, and F. Li, "A model to mitigate forecast uncertainties in distribution systems using the temporal flexibility of electric vehicle aggregators," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 2212-2221, 2020.

[4] A. A. S. d. l. Nieta, J. Contreras, J. I. Muñoz, and M. O. Malley, "Modeling the impact of a wind power producer as a price-maker," *IEEE Transactions on Power Systems,* vol. 29, no. 6, pp. 2723-2732, 2014.

[5] M. Shafie-khah and J. P. S. Catalao, "A stochastic multi-layer agent-based model to study electricity market participants behavior," *IEEE Transactions on Power Systems,* vol. 30, no. 2, pp. 867-881, 2015.

[6] J. Aghaei, M. Barani, M. Shafie-khah, A. A. Sanchez de la Nieta, and J. P. S. Catalao, "Risk-constrained offering strategy for aggregated hybrid power plant including wind power producer and demand response provider," *IEEE Transactions on Sustainable Energy,* vol. 7, no. 2, pp. 513-525, 2016.

[7] H. Ding, P. Pinson, Z. Hu, J. Wang, and Y. Song, "Optimal offering and operating strategy for a large wind-storage system as a price maker," *IEEE Transactions on Power Systems,* vol. 32, no. 6, pp. 4904-4913, 2017.

[8] Y. Li, W. Liu, M. Shahidehpour, F. Wen, K. Wang, and Y. Huang, "Optimal operation strategy for integrated natural gas generating unit and power-to-gas conversion facilities," *IEEE Transactions on Sustainable Energy,* vol. 9, no. 4, pp. 1870-1879, 2018.

[9] X. Gao, K. W. Chan, S. Xia, and B. Zhou, "Risk-constrained offering strategy for a hybrid power plant consisting of wind power producer and electric vehicle aggregator," *Energy,* vol. 177, pp. 183-191, 2019.

[10] M. Shafie-khah, E. Heydarian-Forushani, M. E. H. Golshan, M. P. Moghaddam, M. K. Sheikh-El-Eslami, and J. P. S. Catalão, "Strategic offering for a price-maker wind power producer in oligopoly markets considering demand response exchange," *IEEE Transactions on Industrial Informatics,* vol. 11, no. 6, pp. 1542-1553, 2015.

[11] C. Dou, D. Yue, X. Li, and Y. Xue, "MAS-based management and control strategies for integrated hybrid energy system," *IEEE Transactions on Industrial Informatics,* vol. 12, no. 4, pp. 1332-1349, 2016.

[12] J. Bower and D. W. Bunn, "Model-based comparisons of pool and bilateral markets for electricity," *The energy journal,* pp. 1-29, 2000.

[13] A. Weidlich and D. Veit, "A critical survey of agent-based wholesale electricity market models," *Energy Economics,* vol. 30, no. 4, pp. 1728-1759, 2008.

[14] S. Najafi, M. Shafie-khah, P. Siano, W. Wei, and J. P. S. Catalão, "Reinforcement learning method for plug-in electric vehicle bidding," *IET Smart Grid,* vol. 2, no. 4, pp. 529-536, 2019.

[15] Y. Ye, D. Qiu, M. Sun, D. Papadaskalopoulos, and G. Strbac, "Deep reinforcement learning for strategic bidding in electricity markets," *IEEE Transactions on Smart Grid, vol. 11, no. 2, pp. 1343-1355, 2020.*

[16] N. Rashedi, M. A. Tajeddini, and H. Kebriaei, "Markov game approach for multi-agent competitive bidding strategies in electricity market," *IET Generation, Transmission & Distribution,* vol. 10, no. 15, pp. 3756-3763, 2016.

[17] R. Hao, Q. Ai, and Z. Jiang, "Bilayer game strategy of regional integrated energy system under multi-agent incomplete information," *The Journal of Engineering,* vol. 2019, no. 16, pp. 1285-1291, 2019.

[18] I. Praça, C. Ramos, Z. Vale, and M. Cordeiro, "MASCEM: a multiagent system that simulates competitive electricity markets," *IEEE Intelligent Systems,* vol. 18, no. 6, pp. 54-60, 2003.

[19] H. Kebriaei, A. Rahimi-Kian, and M. N. Ahmadabadi, "Model-based and learning-based decision making in incomplete information cournot games: a state estimation approach," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 45, no. 4, pp. 713-718, 2015.

[20] R. K. Mallick, R. Agrawal, and P. K. Hota, "Bidding strategies of Gencos and large consumers in competitive electricity market based on TLBO," in *2016 IEEE 6th International Conference on Power Systems (ICPS)*, 2016, pp. 1-6.

[21] J. Xiao, X. Kong, Q. Jin, H. You, K. Cui, and Y. Zhang, "Demand-responsive virtual power plant optimization scheduling method based on competitive bidding equilibrium," *Energy Procedia,* vol. 152, pp. 1158-1163, 2018.

[22] M. Mozaffari, H. Askarian Abyaneh, M. Jooshaki and M. Moeini-Aghtaie, "Joint Expansion Planning Studies of EV Parking Lots Placement and Distribution Network," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6455-6465, 2020.

[23] W. Yang, W. Liu, C. Y. Chung and F. Wen, "Joint Planning of EV Fast Charging Stations and Power Distribution Systems with Balanced Traffic Flow Assignment," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1795-1809, 2021.

[24] C. Chen, J. Chen, Y. Wang, S. Duan, and T. Cai, "A Price Optimization Method for Microgrid Economic Operation Considering Across-Time-and-Space Energy Transmission of Electric Vehicles," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1873-1884, 2020.

[25] B. Zeng, J. Feng, N. Liu and Y. Liu, "Co-optimized Public Parking Lot Allocation and Incentive Design for Efficient PEV Integration Considering Decision-Dependent Uncertainties," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1863-1872, 2021.

[26] S. Wang, S. Bi and Y. J. Angela Zhang, "Reinforcement Learning for Real-time Pricing and Scheduling Control in EV Charging Stations," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 849-859, 2021.

[27] D. A. Chekired, L. Khoukhi and H. T. Mouftah, "Fog-Computing-Based Energy Storage in Smart Grid: A Cut-Off Priority Queuing Model for Plug-In Electrified Vehicle Charging," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3470-3482, 2020.

[28] K. Chaudhari, N. K. Kandasamy, A. Krishnan, A. Ukil and H. B. Gooi, "Agent-Based Aggregated Behavior Modeling for Electric Vehicle Charging Load," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 856-868, 2019.

[29] A. Alsabbagh, X. Wu and C. Ma, "Distributed Electric Vehicles Charging Management Considering Time Anxiety and Customer Behaviors," *IEEE Transactions on Industrial Informatics*, Early Access, 2020, doi: 10.1109/TII.2020.3003669.

[30] Z. Yang, T. Guo, P. You, Y. Hou and S. J. Qin, "Distributed Approach for Temporal–Spatial Charging Coordination of Plug-in Electric Taxi Fleet," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3185-3195, 2019.

[31] M. Bowling and M. Veloso, "Multiagent learning using a variable learning rate," *Artificial Intelligence,* vol. 136, no. 2, pp. 215-250, 2002.

[32] B. Zhou, K. Zhang and K. W. Chan, "Optimal coordination of electric vehicles for Virtual Power Plants with dynamic communication spectrum allocation," *IEEE Transactions on Industrial Informatics*, Early Access, 2020. doi: 10.1109/TII.2020.2986883.

[33] E. Heydarian-Forushani, M. P. Moghaddam, M. K. Sheikh-El-Eslami, M. Shafie-khah, and J. P. Catalão, "Risk-constrained offering strategy of wind power producers considering intraday demand response exchange," *IEEE Transactions on sustainable energy*, vol. 5, no. 4, pp. 1036-1047, 2014.

[34] P. L. Liu, A. Der Kiureghian. "Multivariate distribution models with prescribed marginals and covariances," *Probabilistic Engineering Mechanics*, vol. 1, no. 2, pp. 105-112, 1986.

[35] J. M. Morales, S. Pineda, A. J. Conejo and M. Carrion, "Scenario reduction for futures market trading in electricity markets," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 878-888, 2009.

[36] M. Ghamkhari, H. Mohsenian-Rad and A. Wierman, "Optimal risk-aware power procurement for data centers in day-ahead and real-time electricity markets," in *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2014, pp. 610-615.

[37] S. Moazeni, A. H. Miragha and B. Defourny, "A Risk-Averse Stochastic Dynamic Programming Approach to Energy Hub Optimal Dispatch," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2169-2178, 2019.

[38] 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 Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2620-2630, 2017.

[39] J. Kettunen, A. Salo and D. W. Bunn, "Optimization of Electricity Retailer's Contract Portfolio Subject to Risk Preferences," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 117-128, 2010.

[40] T.W. Hill, jr and A. Ravindran, “On programming with absolute-value functions,” *Optimization Theory and Application*, vol. 17, no. 1, pp. 181-183, 1975.

[41] K. Hwang, W. Jiang, and Y. Chen, "Pheromone-based planning strategies in dyna-Q learning," *IEEE Transactions on Industrial Informatics,* vol. 13, no. 2, pp. 424-435, 2017.

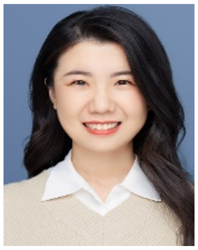
[42] L. Xi, T. Yu, B. Yang, and X. Zhang, "A novel multi-agent decentralized win or learn fast policy hill-climbing with eligibility trace algorithm for smart generation control of interconnected complex power grids," *Energy Conversion and Management,* vol. 103, pp. 82-93, 2015.

[43] S. W Xia, X. Luo, K.W. Chan, M. Zhou, and G. Y. Li, "Probabilistic transient stability constrained optimal power flow for power systems with multiple correlated uncertain wind generation," *IEEE Transactions on Sustainable Energy,* vol.7, no.3, pp.1133-1144, 2016.

[44] A. M. Leite Da Silva, R. A. G. Fernandez, and C. Singh, "Generating capacity reliability evaluation based on Monte Carlo simulation and Cross-Entropy methods," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 129-137, 2010.

[45] D. Xu, Q. Wu, B. Zhou, C. Li, and L. Bai, "Distributed multi-energy operation of coupled electricity, heating and natural gas networks. " *IEEE Transactions on Sustainable Energy*, vol. 11, no. 4, pp. 2457-2469, 2020.

[46] Y. Guo, J. Xiong, S. Xu and W. Su, "Two-Stage Economic Operation of Microgrid-Like Electric Vehicle Parking Deck," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1703-1712, 2016.



Xiang Gao received the B.Eng. degree in electrical engineering from Taiyuan University of Science and Technology, Taiyuan, China, in 2010, the M.Sc. degrees from Tianjin University of Technology, Tianjin, China and Newcastle University, Newcastle upon Tyne, UK, both in electrical engineering in 2014 and 2015 respectively. She is currently pursuing the Ph.D. degree in electrical engineering at the Hong Kong Polytechnic University, Hong Kong. Her research interests include renewable energy resources, stochastic optimization and electricity market.

Ka Wing Chan (M'98) received the B.Sc. (Hons) and Ph.D. degrees in electronic and electrical engineering from the University of Bath, U.K., in 1988 and 1992, respectively. He currently is an Associate Professor and Associate Head in the Department of Electrical Engineering of the Hong Kong Polytechnic University. His general research interests include power system stability, analysis and control, power grid integration, security, resilience and optimization, demand response management, etc.

Shiwei Xia (S'11-M'15-SM'20) received the Ph.D. degree in power systems from The Hong Kong Polytechnic University, Hung Hom, Hong Kong, in 2015. He then stayed with the department and worked as a Research Associate and subsequently as a Postdoctoral Fellow in 2015 and 2016. He was also a Visiting Faculty with the Robert W. Galvin Center for Electricity Innovation, Illinois Institute of Technology, Chicago, IL, USA, in 2019. Currently, he is with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, China. His general research interests include security and risk analysis for power systems with renewable energies, distributed optimization and control of multiple sustainable energy sources in smart grid.

Xiaoshun Zhang (M'17) received the B.Eng. and Ph.D. degrees in electrical engineering from South China University of Technology, Guangzhou, China, in 2012 and 2017 respectively. Now, he is an Associate Professor in College of Engineering, Shantou University, Shantou, China. His major research interests include intelligent optimization, machine learning based optimization in operation of power systems.

Kuan Zhang received the B.Sc. degree in electrical engineering from Hainan University, Haikou, China, in 2017. He is currently working toward the Ph.D. degree in electrical engineering with the College of Electrical and Information Engineering in Hunan University, Changsha, China. His major research interests include integrated energy system operation and optimization, renewable energy generation.

Jiahan Zhou received the B.Sc. degree from Harbin Engineering University, Harbin, China, the M.Sc. degree from University of Southern California, Los Angeles, CA, USA, both in Electrical Engineering in 2015 and 2017 respectively. Currently, he is a Ph.D. candidate in Electrical Engineering at the Hong Kong Polytechnic University, Hong Kong. His research interests include applications of machine learning and data analytics in smart grids.