IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 7, 2020, accepted May 14, 2020, date of publication May 19, 2020, date of current version June 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2995614

A Virtual Generation Ecosystem Control Strategy for Automatic Generation Control of Interconnected Microgrids

LEI XI[1], LE ZHANG[1], JIANCHU LIU[2], YUDAN LI[3], XI CHEN[1], (Member, IEEE), LIUQING YANG[4], (Fellow, IEEE), AND SHOUXIANG WANG[5], (Senior Member, IEEE)

[1]College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China
[2]Super Vision Optical Technology Company, Ltd., Tongling 244000, China
[3]Changshou Branch of State Grid Chongqing Electric Power Company, Chongqing 401220, China
[4]Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80309, USA
[5]Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin 300072, China

Corresponding author: Lei Xi (xilei2014@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51707102.

ABSTRACT The continuous access of new energy and distributed energy, as well as the random disturbance of load power, affect the security and stability of microgrids. A virtual generation ecosystem control (VGEC) strategy is proposed in this paper, which adopts the idea of time tunnel and the principle of a new win-loss criterion to achieve a fast automatic generation control power dispatch, optimal coordinated control of microgrids. A two-layer dynamic power dispatch structure is introduced in the proposed strategy, which combines hierarchical Q-learning with consensus theory to improve the adaptability of the consistency algorithm in complex random environments. Both the IEEE standard two-area load frequency control model and the interconnected microgrids model are used in simulation for comparison and verification. The results show that, by using the VGEC strategy, the control performance of microgrids can be improved, while can reduce power generation cost, and obtain faster convergence speed and stronger robustness compared with other algorithms.

INDEX TERMS Automatic generation control, virtual generation ecosystem control, time tunnel idea, interconnected microgrids.

I. INTRODUCTION

The microgrids can integrate all kinds of distributed generation technologies effectively, and provide an effective technical way for the large-scale access of new energy and distributed energy, which has become an important part of smart grid research and construction [1]–[4]. However, because of environmental uncertainties and load fluctuations, the energy management system (EMS) of microgrids faces many challenges [5]–[7]. Automatic generation control (AGC) [8], as one of the most important control functions in EMS, can effectively improve the frequency quality and economic efficiency.

The traditional AGC strategies are usually composed of two categories: a) track the total power references of AGC usually controlled; b) assign the total power references to each unit by a fixed allocation method. In practice, proportional-integral (PI) controller is widely used in the total power references tracking of AGC in microgrids. Moreover, bacterial foraging optimization (BFO) [9], particle swarm optimization (PSO) [10], genetic algorithm (GA) [11], and conventional gradient descent algorithm were applied to simultaneously optimize all the control parameters of microgrids. In previous studies of the authors, the reinforcement learning (RL) has been applied to the traditional AGC [12]–[14] of the interconnected power grid to solve the random disturbance caused by massive integration of distributed energy. However, the aforementioned studies are all based on centralized control structure, which is an inefficient cooperation with economical dispatch due to the ignorance of power grid topology. Particularly, the overall AGC command of provincial dispatch centre is assigned through a fixed

The associate editor coordinating the review of this manuscript and approving it for publication was Huai-Zhi Wang.

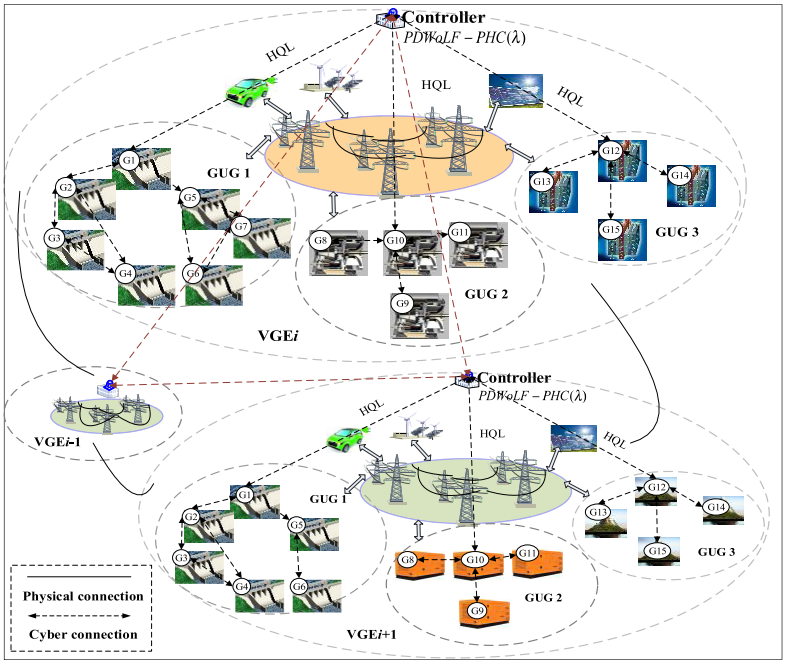


FIGURE 1. The control framework of interconnected microgrids based on VGEC strategy.

proportion of the adjustable capacity rather than a dynamic optimization, and it cannot effectively cooperate with the AGC of interconnected microgirds.

Therefore, a wolf pack hunting (WPH) [15] strategy based on the multi-agent system stochastic consensus game (MAS-SCG) [16] framework is presented to obtain optimal coordinated control of the islanded distribution network model with a large amount of distributed energy. However, the control method based on the multi-agent system stochastic game (MAS-SG) [17] in the WPH strategy is the DWoLF-PHC(λ), which cannot accurately calculate the win-and-loss criterion and quickly converge to the Nash. Hence the author proposes an ecological population cooperative control (EPCC) [18] to solve above problems, and its control method is PDWoLF-PHC(λ) which has new win-and-loss criterion and integrates time tunnel ideas. The power allocation algorithms in both WPH and EPCC adopt simple first-order consensus algorithm based on multi-agent system collaborative consensus (MAS-CC) [19], which relies heavily on the model and falls easily into the local optimal solution.

To improve the adaptability of the consensus algorithm in a dynamic random environment, a multi-robot behavior consensus Q-learning algorithm [20] that combined consensus algorithm and RL is proposed to realize robot behavior control. Inspired by this, a two-layer model for dynamic power dispatch is constructed. Based on this model, a novel hierarchical Q-learning consensus (HQC) strategy is proposed to obtain the optimal solution of power allocation and solve the dimension disaster. Then a virtual generation ecosystem control (VGEC) strategy is proposed to achieve a fast automatic generation control power dispatch, optimal coordinated control of microgrids, which is a hybrid of control (PDWoLF-PHC(λ) based on MAS-SG) and optimization (HQC based on MAS-CC). In particular, PDWoLF-PHC(λ) is used under MAS-SG to rapidly obtain the overall power reference (control), then HQC is firstly adopted under MAS-CC to optimally distribute the obtained overall power command into each unit (optimization). Both the IEEE standard two-area load frequency control (LFC) model and the interconnected microgrids model are used for simulation comparison and verification. The results show that the proposed strategy can improve the control performance of interconnected microgrids, reduce power generation cost, and obtain faster convergence speed and stronger robustness compared with other algorithms.

II. FRAMEWROK DEVELOPMENT

The interconnected microgrids integrated with a large number of distributed energy resources can be divided into several small area grid virtually according to the graph theory cut method, and the control framework is shown in Fig. 1. Each small area micro-network is considered as a VGE. Each VGE exchanges power through the regional boundary to maintain the frequency stability of the interconnected microgrid. When a serious fault occurs in the microgrid, each VGE is automatically separated as an island operation. At this time, the frequency of each VGE needs to be maintained by autonomous control.

The PDWoLF-PHC(λ) is adopted to obtain the total power command of each VGE through a multi-agent dynamic game since each VGE can be regarded as an agent. Each VGE contains multiple types of generator unit groups (GUGs),

which will select the generator unit with the largest capacity as the leader and the other units as the followers. Total power command of different GUGs is obtained through hierarchical Q-learning (HQL) [21], meanwhile power command of each unit in GUGs is assigned using the consensus algorithm.

III. VEGC

The VGEC strategy is a mixed strategy based on MAS-SCG framework. The PDWoLF-PHC(λ) based on the MAS-SG principle is adopted to obtain the total power command in the AGC control part of each VGE. An HQC strategy is proposed to distribute the power in an optimal dynamic way, which combines HQL algorithm with a consensus method based on the MAS-CC principle.

A. AGC CONTROL ALGORITHM

In the VGE, the PDWoLF-PHC(λ) is used as the AGC control algorithm, and its corresponding controller is equivalent to an agent, which can frequently exchange information with other agents. The key idea of the proposed PDWoLF-PHC(λ) algorithm is to explicitly use the time tunnel idea with backward multi-step prediction function to effectively backtrack the online reinforcement information of future multi-step decision-making. At the same time, each agent uses experience sharing to update the Q-function table, through the dynamic competition or cooperation appropriately adjust its own control strategy, and maximize the overall learning efficiency of the multi-agent system, so that the proposed algorithm can obtain the optimal cooperative control.

The control system is a multiple-agent system, and the actions of other agents can change the state of the entire system. At this time, the agent uses the product of the decision change rate and the decision space slope value to be negative to design the variable learning rate, and agent changes the learning rate as the state changes.

In the case of the state s_k and the reward function R_1, the agent executes the search action a_k according to the mixed table $\pi(s_k, a_k)$, and its state transits to the next state s_{k+1}. The update rule of $\pi(s_k, a_k)$ is as follows:

$$\pi(s_k, a_k) \leftarrow \pi(s_k, a_k) + \Delta_{s_k a_k} \tag{1}$$

$$\Delta_{s_k a_k} = \begin{cases} -\delta_{s_k a_k}, & a_k \neq \arg\max_{a_{k+1}} Q(s_k, a_{k+1}) \\ \sum \delta_{s_k a_{k+1}}, & \text{otherwise} \end{cases} \tag{2}$$

$$\delta_{s_k a_k} = \min\left(\pi(s_k, a_k), \varphi_i / (|A_i| - 1)\right) \tag{3}$$

where $\Delta_{s_k a_k}$ is the change of policy update at kth step iteration, $|A_i|$ is the number of optional actions, and φ is the variable learning rate. Its update law is as follows:

$$\phi = \begin{cases} \phi_{\text{win}}, & \Delta(s_{k-1}, a_{k-1}) \times \Delta^2(s_{k-1}, a_{k-1}) < 0 \\ \phi_{\text{lose}}, & \text{otherwise} \end{cases} \tag{4}$$

where $\Delta(s_{k-1}, a_{k-1})$ represents the decision change rate at $k-1$th step iteration, and $\Delta^2(s_{k-1}, a_{k-1})$ is the decision space slope. If the product of the decision change rate $\Delta(s_{k-1}, a_{k-1})$ and the decision space slope $\Delta^2(s_{k-1}, a_{k-1})$ is negative, ϕ_{win} is selected as the variable learning rate, otherwise ϕ_{lose} is selected, where $\phi_{\text{lose}} > \phi_{\text{win}}$. In the next step iteration, $\Delta(s_k, a_k)$ and $\Delta^2(s_k, a_k)$ will be updated according to (5) and (6):

$$\Delta^2(s_k, a_k) \leftarrow \Delta_{s_k a_k} - \Delta(s_{k-1}, a_{k-1}) \tag{5}$$

$$\Delta(s_k, a_k) \leftarrow \Delta_{s_k a_k} \tag{6}$$

This paper selects the eligibility trace based on SARSA(λ) [19]:

$$e_{k+1}(s, a) = \begin{cases} \gamma\lambda e_k(s, a) + 1, & (s, a) = (s_k, a_k) \\ \gamma\lambda e_k(s, a), & \text{otherwise} \end{cases} \tag{7}$$

where $e_k(s, a)$ is eligibility trace at kth step iteration under state s and action a, γ is the discount factor, and λ is the attenuation factor.

The agent uses the current reward R to calculate the evaluation value of Q function error:

$$\rho_k = R(s_k, s_{k+1}, a_k) + \gamma Q_k(s_{k+1}, a_g) - Q_k(s_k, a_k) \tag{8}$$

$$M_k = R(s_k, s_{k+1}, a_k) + \gamma Q_k(s_{k+1}, a_g) + Q_k(s_k, a_g) \tag{9}$$

In (8) and (9), $R(s_k, s_{k+1}, a_k)$ is the agent reward function at the state from s_k to s_{k+1} under the selected action a_k, a_g is the greedy action, ρ_k is the Q function error at kth step iteration, and M_k is the evaluation of Q function error.

The $Q(\lambda)$ [22] is iteratively updated as follows:

$$Q_{k+1}(s, a) = Q_k(s, a) + \alpha M_k e_k(s, a) \tag{10}$$

$$Q_{k+1}(s_k, a_k) = Q_k(s_k, a_k) + \alpha\rho_k \tag{11}$$

where α is the learning rate.

After experiencing enough trial and error iterations, the state value function $Q_k(s, a)$ will converge to the Q^* matrix with a probability of 1, and finally obtain an optimal control strategy represented by the Q^* matrix.

In general, area control error (ACE) can maximize the long-term benefits of CPS and avoid large power fluctuations. Meanwhile, generation cost takes into account the economic impact of the energy management system. Therefore, the weighted sum of ACE and C_{total} is chosen as a reward function, in which a larger weighted sum will result in a smaller reward. The reward function is expressed as follows:

$$R_1(s_k, s_{k+1}, a_k) = -\rho[ACE(k)]^2 - \frac{(1-\rho)C_{total}(k)}{50000} \tag{12}$$

where ACE and C_{total} represent the instantaneous absolute value of the ACE and the actual generation cost of all units at the kth step iteration, ρ and $1-\rho$ are the weight ratio of ACE and C_{total}, respectively, and $\rho = 0.5$ is chosen.

After several trial and error tests, the parameters of control algorithm are set in Table 1.

B. AGC POWER ALLOCATION ALGORITHM

1) MATHEMATICAL MODEL OF POWER ALLOCATION

Power deviation, regulation cost, and ramp time are chosen as three objective functions. The power command allocation

TABLE 1. PDWoLF-PHC(λ) parameters setting.

Parameters	Value
λ	0.9
γ	0.9
α	0.5
ϕ	0.06

process for each VGE is described by the following mathematical model:

$$\begin{cases} \min f_1 = \Delta P_{\text{error}}^2 + \sum_{i=1}^{m} C_{\text{total}-i} \\ \min f_2 = \sum_{i=1}^{m} \sum_{w=1}^{W_i} \Delta P_{iw} / \Delta P_{iw}^{\text{rate}} \\ \text{s.t. } \Delta P_{\text{error}} = \Delta P_{\Sigma} - \sum_{i=1}^{m} \Delta P_i \\ \Delta P_{\Sigma} = \sum_{i=1}^{m} \sum_{w=1}^{W_i} \Delta P_{iw} \\ \Delta P_i = \eta_i \Delta P_{\Sigma} \\ \sum_{i=1}^{m} \eta_i = 1, \quad 0 \le \eta_i \le 1 \end{cases} \tag{13}$$

where f_1 is the linear weighted multiple-objective function of power deviation and regulation cost;

f_2 is the objective function of ramp time;

ΔP_{error} is the deviation of the total power command calculated by the AGC controller and the total power of all GUGs;

$C_{total-i}$ is the regulation cost of the whole generator units in GUG_i;

ΔP_{iw} and ΔP_{iw}^{rate} are the power command and ramp rate of the wth generator unit of GUG_i respectively;

ΔP_{Σ} is the total power command calculated by the AGC controller;

ΔP_i is the power command of GUG_i;

η_i is the power allocation factor;

m is the number of GUG in each territorial power grid;

W_i is the number of generator units in GUG_i.

2) HQC STRATEGY

In the AGC allocation process, a novel HQC strategy that combines HQL with consensus algorithm is adopted to allocate power command. Each generator unit is regarded as an agent, the unit with the largest capacity is selected as the leader, and others are the followers. The leader interacts with the environment and gets an environment state s. The reward R of the environment and the next state $s+1$ are obtained by the leader through taking an action, and the self-learning process is completed. Through the consensus algorithm, leader–follower and follower–follower can frequently interact. Therefore, the optimal allocation can be obtained through self-learning and collaborative learning.

a: HQL ALGORITHM

HQL can realize the self-learning processes to obtain power command of each GUG. HQL is based on the $Q(\lambda)$, and the eligibility trace is iteratively updated by (7), in which $\gamma = 0.9$, and $\lambda = 0.5$.

The agent obtains reward value R_2 through the current exploration. $R_2(s_k, s_{k+1}, a_k)$ is the agent reward function through executing action a_k from state s_k to s_{k+1}. Reward function $R_2(s_k, s_{k+1}, a_k)$ is designed as follows:

$$R_2(s_k, s_{k+1}, a_k) = -\Delta P_{\text{error}}^2 - \frac{C_{\text{total}}}{1000} \tag{14}$$

The iteration of HQL is updated according to (10) and (11). Assuming that the probability of the action occurring in the initial state is the same, the action a_k is executed, and the state is transferred to s_{k+1}. The action probability function P is updated as follows:

$$\begin{cases} P_{k+1}(s_{k+1}, a_g) = P_k(s_k, a_g) + \theta(1 - \pi_k(s_k, a_g)) \\ P_{k+1}(s_{k+1}, a_k) = P_k(s_k, a_g)(1-\theta), \quad \forall a_k \in A, \ a_k \ne a_g \\ P_{k+1}(s_{k+1}, a_{k+1}) = P_{k+1}(s_k, a_k) \end{cases} \tag{15}$$

where θ is the action search speed with $0 \le \theta \le 1$, and $\theta = 0.9$ is chosen.

The total power command is taken as a state variable, which is discretized into $(-\infty, -650)$, $[-650, 20]$, $[20, 850]$, and $(850, +\infty)$. The action set is $A_i = [\eta_1, \eta_2, \ldots, \eta_i] = [(\eta_{11}, \eta_{12}, \ldots, \eta_{1i}), (\eta_{21}, \eta_{22}, \ldots, \eta_{1i}), \ldots (\eta_{n1}, \eta_{n2}, \ldots, \eta_{ni})]$, where η_i is a distribution factor of GUG.

b: MAS-CC ALGORITHM

Each GUG is treated as a hierarchical multiple-agent system network. Suppose that the GUG has a network of p agents, and the agents are represented by $p(p = 1, \ldots, n)$, respectively. The relationship of the interactions between agents is represented by a graph $G = (V, E, A)$. $V = (V_p, p = 1, \ldots, n)$ is a node set, and each node represents an agent; $E \in V \times V$ is the edge set, and its element represents the relationship between agents through a directed or undirected connection.

It is assumed that communication between agents v_p and v_q is determined by probability $b_{pq} (0 \le b_{pq} \le 1)$ and independent from other agents. Communication between agents means there is an information connection. Laplace matrix $L = [l_{pq}]$ can reflect the topology of multi-agent network [23], which is expressed as follows:

$$\begin{cases} l_{pp} = \sum_{q=1, p \ne q}^{p} b_{pq} \\ l_{pp} = -b_{pq}, \quad \forall p \ne q \end{cases} \tag{16}$$

The ramp time is chosen as a consensus variable for the GUG. The leader with a higher ramp rate undertakes more disturbances. The ramp time of the wth unit of GUG$_i$ is expressed as follows:

$$t_{iw} = \Delta P_{iw}/\Delta P_{iw}^{\text{rate}} \tag{17}$$

where ΔP_{iw} and ΔP_{iw}^{rate} are the power command and ramp rate of the wth unit of GUG$_i$, respectively. And ΔP_{iw}^{rate} is expressed as follows:

$$\Delta P_{iw}^{\text{rate}} = \begin{cases} \Delta P_{iw}^{\text{rate}+}, & \Delta P_i > 0 \\ \Delta P_{iw}^{\text{rate}-}, & \Delta P_i < 0 \end{cases} \tag{18}$$

where ΔP_{iw}^{rate+} and ΔP_{iw}^{rate-} are the upper and lower bounds of the ramp rate, respectively.

The ramp time for each follower in GUG is updated as (19).

$$t_{iw}[k+1] = \sum_{v=1}^{W_i} d_{wv}[k] t_{iv}[k] \tag{19}$$

where W_i is the number of generator units of GUG$_i$, and $d_{wv}[k]$ represents the term $[w, v]$ of the discrete time k of the row random matrix $\boldsymbol{D} = d_{wv}[k] \in R^{Wi} \times {}^{Wi}$. It can be expressed as follows.

$$d_{wv}[k] = \frac{|l_{wv}|}{\sum_{v=1}^{W_i} |l_{wv}|}, \quad w = 1, 2, \ldots, W_i \tag{20}$$

The collaborative consensus of the agents can be achieved under the condition of frequent information interaction among the agents and constant gain b_{wv} if and only if the directed graph is connected strongly [24].

The ramp time of the leader can be updated according to Reference [25] as follows:

$$t_{iw}[k+1] = \begin{cases} \sum_{v=1}^{W_i} d_{wv}[k] t_{iv}[k] + \sigma_i \Delta P_{\text{error-}i}, & if\ \Delta P_i > 0 \\ \sum_{v=1}^{W_i} d_{wv}[k] t_{iv}[k] - \sigma_i \Delta P_{\text{error-}i}, & if\ \Delta P_i < 0 \end{cases} \tag{21}$$

where σ_i is the power regulation factor for the GUG$_i$, $\boldsymbol{D} = d_{wv}[k] \in R^{Wi} \times {}^{Wi}$ is a row random matrix, and $\Delta P_{\text{error}-i}$ is the power deviation for the GUG$_i$, which is expressed as follows.

$$\Delta P_{\text{error}-i} = \Delta P_i - \sum_{w=1}^{W_i} \Delta P_{iw} \tag{22}$$

Similarly, the power generation commands ΔP_{iw} and the maximum ramp time t_{iw} are expressed in (23) and (24) as the boundary conditions are achieved.

$$\Delta P_{iw} = \begin{cases} \Delta P_{iw}^{\max}, & \Delta P_{iw} > \Delta P_{iw}^{\max} \\ \Delta P_{iw}^{\min}, & \Delta P_{iw} < \Delta P_{iw}^{\min} \end{cases} \tag{23}$$

For $s \in S$, $a \in$ A, initialize all parameters, and set the parameters s_0, $k = 0$;
Repeat
1. Select an action a_k according to the mixed strategy $\pi(s_k, a_k)$;
2. Obtain short-term reward $R_1(k)$ via (12);
3. Calculate the single-step Q function error ρ_k and value function error M_k through (10) and (11);
4. For each state-action pair (s, a), execute the following steps:
   (i) let $e_{k+1}(s, a) = \gamma\lambda e_k(s, a)$;
   (ii) update the Q function $Q_k(s, a)$ from $Q_{k+1}(s, a)$ using (8);
5. Solve the mixed strategy $\pi(s_k, a_k)$ via (1), (2), and (3);
6. Update the Q function $Q_k(s_k, a_k)$ to $Q_{k+1}(s_k, a_k)$ by (11); update the eligibility trace element using (10), let $e(s_k, a_k) \leftarrow e(s_k, a_k)+1$; choose variable learning rate ϕ by (4); update decision change rate $\Delta(s,a)$ and policy space gradient value $\Delta^2(s,a)$ according to (5) and (6);
7. Output total power command value ΔP_Σ;
8. Obtain reward $R_2(k)$ via (14);
9. Calculate the single-step Q function via (10) and (11);
10. Solve greedy action a_g according to the maximum Q value, and update the P function via (15);
11. Update the eligibility trace element using (10), let $e(s_k, a_k) \leftarrow e(s_k, a_k)+1$;
12. Output each GUG power command value $\Delta P_i (i=1,2,\ldots,m)$;
13. Determine the ramp rate via (18);
14. Apply a consensus algorithm according by (19) or (21);
15. If the generation constraint is not exceeded, then execute step 19;
16. Calculate regulation power ΔP_{iw} and ramp time by (21) and (24); correct d_{wv} via (25), (20), and (16);
17. Calculate the power deviation $\Delta P_{\text{error}-i}$ using (22)
18. If formula $|\Delta P_{\text{error}-i}| < \sigma_i$ is not satisfied, then execute step 14;
19. Output unit adjustment power $\Delta P_{iw} (w=1,2,\ldots,W_i)$;
20. Let $k = k + 1$ and return to step 1.
End

FIGURE 2. Execution steps of the VGEC strategy.

$$t_{iw} = t_{iw}^{\max} = \begin{cases} \dfrac{\Delta P_{iw}^{\max}}{\Delta P_{iw}^{\text{rate}+}}, & \Delta P_{iw} > \Delta P_{iw}^{\max} \\ \dfrac{\Delta P_{iw}^{\min}}{\Delta P_{iw}^{\text{rate}-}}, & \Delta P_{iw} < \Delta P_{iw}^{\min} \end{cases} \tag{24}$$

where ΔP_{iw}^{max} and ΔP_{iw}^{min} are the maximum and minimum reserve capacity of the wth units of GUG$_i$, respectively.

Moreover, the weighed factor will be changed as shown in Eq. (25) if the power command ΔP_{iw} of the wth unit of GUG$_i$ exceeds the limitation.

$$b_{wv} = 0, \quad v = 1, 2 \cdots, W_i \tag{25}$$

C. VGEC PROCEDURE

The execution steps of the VGEC are shown as in Fig. 2.

IV. CASE STUDIES

In the section of case studies, two-area LFC power system model and interconnected microgrids model are built to analyze the performance of the proposed strategy. All the simulation cases in this paper are run in MATLAB R2016b environment and the total instruction control period of the dispatcher is 4 seconds. Meanwhile, Simulink is used

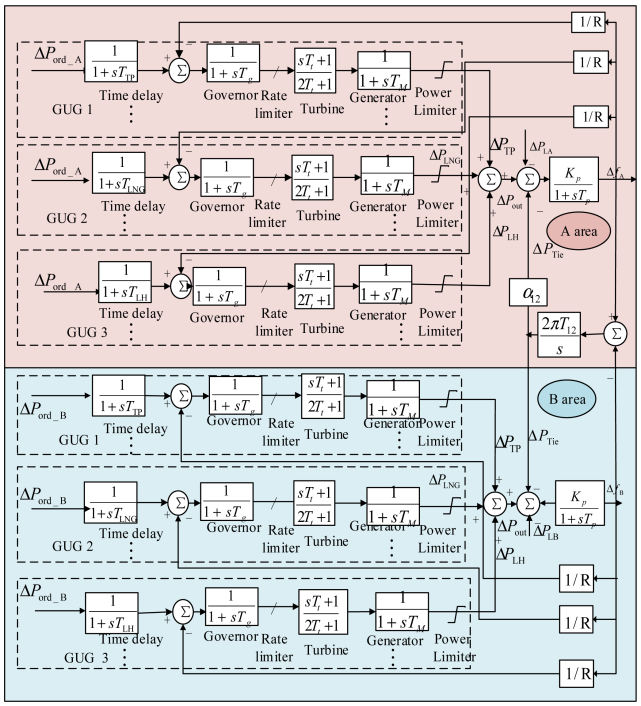


FIGURE 3. IEEE standard two-area LFC power system.

for modeling and simulation, and the proposed algorithm and controller of VGEC strategy are written by S-function module.

A. TWO-AREA LFC MODEL

Based on the IEEE standard two-area LFC power system model [26], one equivalent unit in area A and B is replaced by GUGs that contain thermal power (TP) plant, liquefied natural gas (LNG) plant, and large hydropower (LH) plant. The framework structure of the two-area LFC power system model is shown in Fig. 3, and the system parameters are selected from [27].

Before running on-line, numerous explorations are needed to identify the optimal action strategy through off-line trial-and-error that optimizes Q function and state value function [28], and achieves sufficient pre-learning. Sinusoidal load disturbance with period of 5000 s and amplitude of 1000 MW is also introduced in area A and B, respectively. The pre-learning process of two areas produced by continuous sinusoidal disturbance is shown in Fig. 4.

As shown in Fig. 4(a), the strategy can quickly track load disturbance in two areas. The AGC control performance is evaluated by control performance standard (CPS) and ACE. Fig. 4(b) illustrates that CPS1 in areas A and area B are maintained within the range of 185% to 200% and 150% to 200%, respectively. Moreover, Fig. 4(c) shows that the ACE in area A and area B remains in the range of −88 MW to 0 MW and −158 MW to 0 MW and finally reaches a stable value. The CPS standards are as follows:

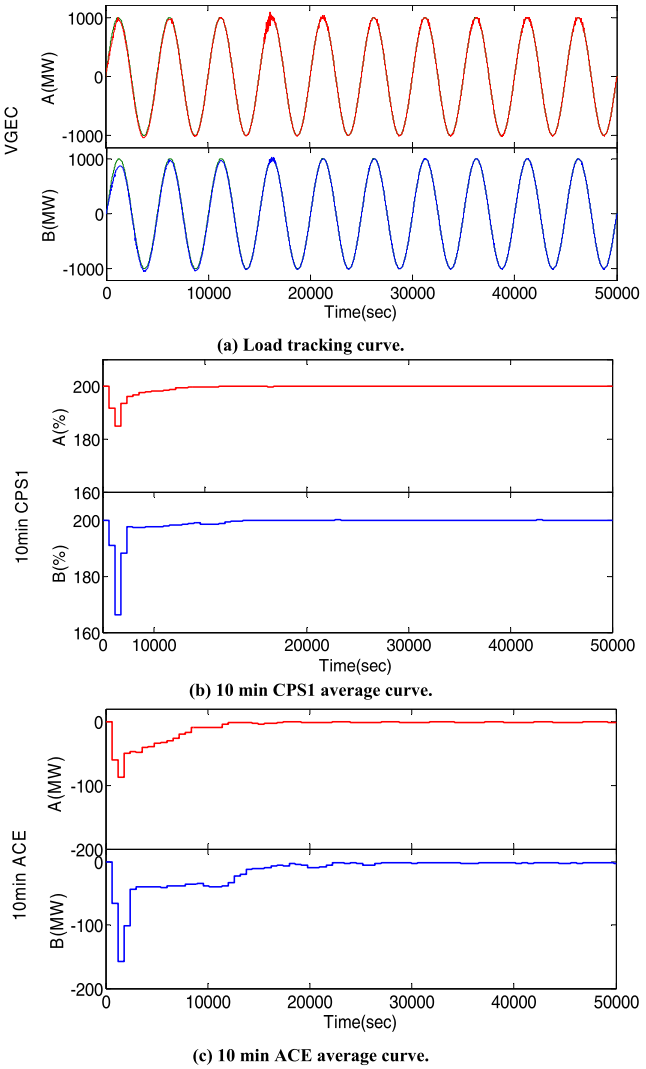


FIGURE 4. Pre-learning of VGEC strategy in each area.

(1) If CPS1 $\geq$ 200%, and CPS2 is an arbitrary value, the CPS is qualified;

(2) If 100% $\leq$ CPS1 $<$200% and CPS2 $\geq$ 90%, the CPS is qualified;

(3) If CPS1 $<$ 100%, the CPS is unqualified.

B. INTERCONNECTED MICROGRIDS MODEL

In this paper, a interconnected microgrids model is established that includes three microgrids with the communication topology shown in Fig. 1. The model integrates a large number of new energy and distributed energy including photovoltaics (PV), wind farms (WF), small hydro-powers (SH), micro-gas turbines (MT), diesel generators (DG), biomass energy (BE), and fuel cells (FC). The model is simplified to some extent because it does not include PV, WF, and electric vehicles (EVs) as participants in the system frequency modulation. The corresponding PV model [29] was established by simulating the change of the full-day light intensity; the output model of the WF [30], and other generator set models are established in previous studies [31]–[35].

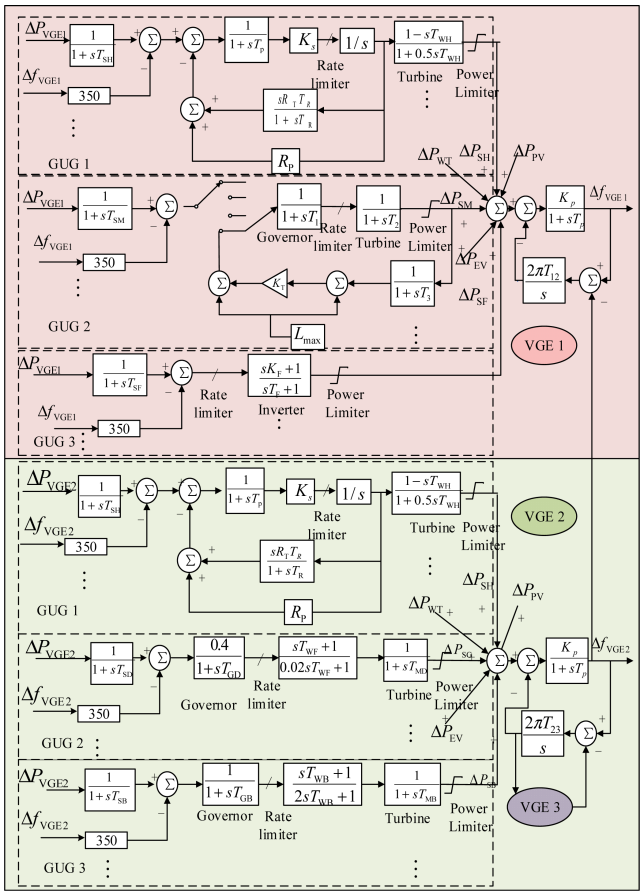


FIGURE 5. Three areas interconnected microgrid model.

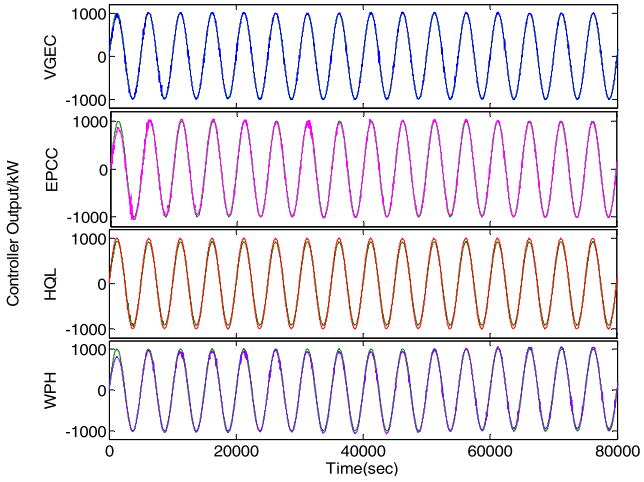


FIGURE 6. Pre-learning of different algorithms in VGE1.

Fig. 5 shows the structure of a three-area interconnected microgrid cluster including VGE1, VGE2, and VGE3. The model parameters and unit parameters of VGE1 and VGE3 are the same, so only VGE1 and VGE2 are analyzed. The regulation power of the three areas is 2350, 2590, and 2350 kW, respectively, and the non-adjustable units (PV, WF, and EV) are regarded as load disturbances. Each adjustable unit (SH, MT, FC, DG, and BE) is treated as a different agent,

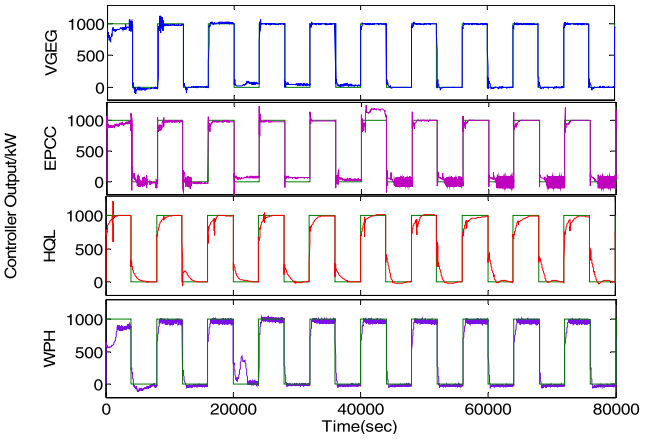


FIGURE 7. Impulsive load tracking curve of different algorithms in VGE1.

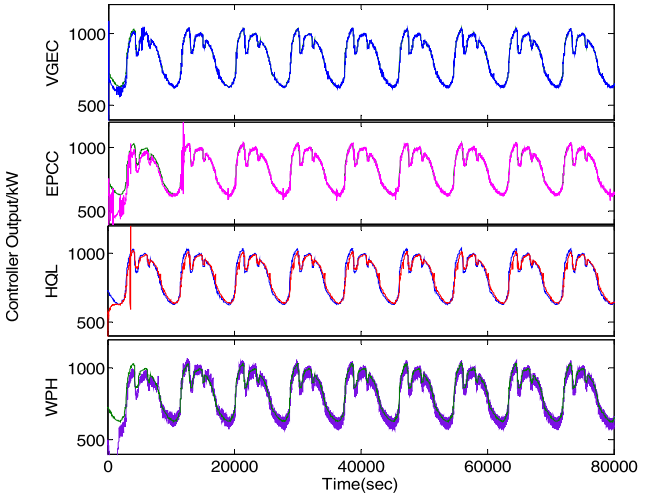


FIGURE 8. White noise load tracking curve of different algorithms in VGE1.

and the connection weight b_{wv} between the agents is chosen to be 1. The parameters of the AGC unit [36], [37] are shown in Table 2. C_i represents the cost of microgrid generation, and its formula is $C_i(P_{\mathrm{Gi,actual}}) = C_i(P_{\mathrm{Gi,plan}} + \Delta P_{\mathrm{Gi}}) = \alpha_i \Delta P_{\mathrm{Gi}}^2 + \beta_i \Delta P_{\mathrm{Gi}} + \gamma_i$, where $P_{\mathrm{Gi,actual}}$ is actual active power for the ith unit, $P_{\mathrm{Gi,plan}}$ is the planned power generation of the ith unit, ΔP_{Gi} is the AGC regulation power of the ith unit, and positive constants α_i, β_i, and γ_i are dynamic coefficients under load disturbance with $\alpha_i = a_i$, $\beta_i = 2a_i P_{\mathrm{Gi,plan}} + b_i$, and $\gamma_i = P_{\mathrm{Gi,plan}}^2 + b_i P_{\mathrm{Gi,plan}} + c_i$.

1) IMPULSIVE DISTURBANCE AND WHITE NOISE DISTURBANCE

In the pre-learning stage, the sinusoidal load disturbance with a period of 5000 s and an amplitude of 1000 kW is introduced in the VGE1, and the VGEC strategy is contrasted with three algorithms, namely EPCC [18], HQL [21], and WPH [15]. Fig. 6 shows the load disturbance tracking curves for different algorithms during the pre-learning process. Compared with other algorithms, the VGEC strategy can quickly track load disturbances with faster convergence.

TABLE 2. AGC unit parameters.

Unit cluster	Type unit	ΔP^{max} /kW	ΔP^{min} /kW	ΔP^{rate+} /(kW/s)	ΔP^{rate-} /(kW/s)	t^{max} /s	C_i($/h) a_i	b_i	c_i
VGE 1 and VGE 3									
GUG1 SH	G1	250	-250	15	-15	16.67	0.0001	0.0346	8.5957
	G2	250	-250	15	-15	16.67	0.0001	0.0346	8.5957
	G3	150	-150	8	-8	18.75	0.0001	0.0335	8.0643
	G4	150	-150	8	-8	18.75	0.0001	0.0335	8.0643
	G5	150	-150	8	-8	18.75	0.0001	0.0335	8.0643
	G6	100	-100	7	-7	14.29	0.0001	0.0314	7.6248
	G7	100	-100	7	-7	14.29	0.0001	0.0314	7.6248
GUG2 MT	G8	100	-100	1.2	-1.2	83.33	0.0002	0.1088	5.2164
	G9	100	-100	1.2	-1.2	83.33	0.0002	0.1088	5.2164
	G10	150	-150	1.8	-1.8	83.33	0.0002	0.1164	5.4976
	G11	150	-150	1.8	-1.8	83.33	0.0002	0.1164	5.4976
GUG3 FC	G12	200	-200	7	-7	28.57	0.0002	0.1164	5.4976
	G13	200	-200	7	-7	28.57	0.0002	0.1164	5.4976
	G14	150	-150	6	-6	25	0.0003	0.1189	3.5442
	G15	150	-150	6	-6	25	0.0003	0.1189	3.5442
VGE 2									
GUG1 SH	G1	250	−250	15	−15	16.67	0.0001	0.0346	8.5957
	G2	250	−250	15	−15	16.67	0.0001	0.0346	8.5957
	G3	150	−150	8	−8	18.75	0.0001	0.0335	8.0643
	G4	150	−150	8	−8	18.75	0.0001	0.0335	8.0643
	G5	150	−150	8	−8	18.75	0.0001	0.0335	8.0643
	G6	100	−100	7	−7	14.29	0.0001	0.0314	7.6248
GUG2 DG	G7	250	−250	2	−2	125	0.0004	0.2348	10.9952
	G8	250	−250	2	−2	125	0.0004	0.2348	10.9952
	G9	120	−120	1	−1	120	0.0004	0.2348	10.9952
	G10	120	−120	1	−1	120	0.0004	0.2348	10.9952
GUG3 BE	G11	200	−200	3	−3	66.67	0.0004	0.0656	8.7657
	G12	200	−200	3	−3	66.67	0.0004	0.0656	8.7657
	G13	200	−200	3	−3	66.67	0.0004	0.0656	8.7657
	G14	200	−200	3	−3	66.67	0.0004	0.0656	8.7657

TABLE 3. Control index values of different algorithms under impulsive and white noise disturbance.

Disturbance	Algorithm	$\lvert\Delta f\rvert$ (Hz)	|ACE| (kW)	CPS1 (%)	CPS2 (%)
Impulsive disturbance	VGEC	0.0104	19.32	198.57	97.22
	EPCC	0.0108	19.89	197.57	95.14
	HQL	0.0178	33.78	195.01	96.53
	WPH	0.0254	40.16	191.87	85.46
White noise disturbance	VGEC	0.0018	3.35	199.87	100
	EPCC	0.0030	5.48	199.65	100
	HQL	0.0053	9.70	199.69	99.67
	WPH	0.0194	15.67	199.65	100

After the full pre-learning process, impulsive disturbance and white noise are introduced into the VGE1 to simulate different types of disturbances in the power system. The long-term performance of the VGEC strategy was statistically evaluated using 24 h experimental results under given impulsive disturbance and white noise disturbance. Fig. 7 and 8 are the load tracing curves of different algorithms. It can be seen from Fig. 7 and 8 that the VGEC strategy has smoother adjustment commands and can quickly track load disturbances.

Table 3 shows the control index values of different algorithms under impulsive disturbance and white noise disturbance. In this case, $|\Delta f|$ is the average of the absolute values of the frequency deviation, and all the indicators are the average values in the simulation time. CPS1 evaluates the effect of ACE changes on system frequency, and CPS2 evaluates the ACE amplitude. The CPS index considers the distribution of CPS1 and CPS2 indicators, which are mainly used to evaluate the control performance of the entire AGC system.

Values shown in Table 3 indicate that in the VGE1 area impulsive disturbance, compared with other methods, the VGEC strategy can reduce $|\Delta f|$ by 0.0004 Hz to 0.0104 Hz, |ACE| by 0.57 kW to 20.84 kW, while it can increase CPS1 by 1.01% to 6.72%, CPS2 by 0.69% to 11.76%. Under white noise disturbance, the VGEC strategy can reduce $|\Delta f|$ by 0.0012 Hz to 0.0176 Hz, |ACE| by 2.13 kW to 12.32 kW, while it can increase CPS1 0.18% to 0.22% and increase CPS2 0.00% to 0.33%.

2) RANDOM DISTURBANCE

Random disturbances are applied to the VGE1, VGE2, and VGE3 over a simulation time of 24h to verify the robustness of the VGEC strategy. Three algorithms are introduced for comparison. As shown in Fig. 9, under the VGEC strategy,

TABLE 4. Statistics of AGC control performance indexes of different algorithms under random disturbance.

Area	Algorithm	$\lvert\Delta f\rvert$ (Hz)	|ACE| (kW)	CPS1 (%)	CPS2 (%)	Generation cost ($\times 10^4$ US$)
VGE1	VGEC	0.0096	13.40	199.31	99.31	12.6295
	EPCC	0.0111	20.55	198.24	97.22	12.7149
	HQL	0.0127	20.57	196.65	93.20	12.6509
	WPH	0.0183	13.82	197.81	97.22	29.2641
VGE2	VGEC	0.0065	12.12	198.62	99.31	12.6344
	EPCC	0.0111	20.55	198.62	98.61	12.7129
	HQL	0.0116	20.66	197.34	93.37	12.6396
	WPH	0.0114	21.84	196.11	97.22	29.1401

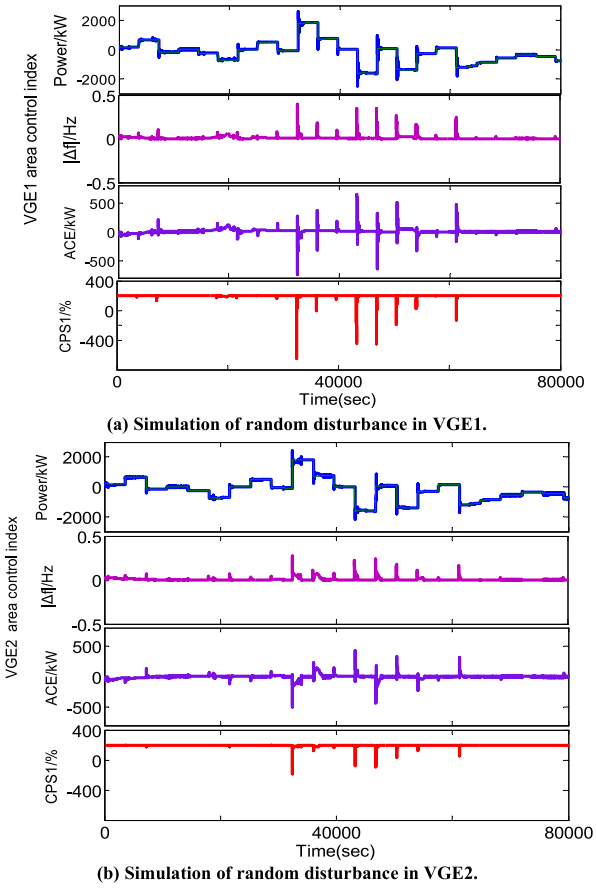


FIGURE 9. The control performance of VGEC algorithm under random disturbance in VGE1 and VGE2.

the total power command output of the unit can track the load disturbance well, so that the AGC dynamic control index can quickly return to its ideal value after the disturbance occurs, including the Δf, ACE, CPS1, which also illustrates the robustness of the VGEC strategy.

Table 4 shows the AGC control performance indicators of the four algorithms under random disturbances, where the cost is the sum of the total regulation costs of all units within 24 h. VGEC has the highest value of CPS1, the smallest value of ACE and Δf, and the lowest total control cost.

Compared with EPCC strategy, the HQC strategy is adopted in the AGC power allocation part, so VGEC can improve the convergence speed through the interaction between the agents and the self-learning of the agent. Compared with HQL, VGEC is not affected by the size of the AGC unit, so the global optimal search capability is stronger, and the optimal solution can be obtained. Compared with WPH, VGEC uses the product of decision rate of change and the value of slope of decision space to be negative to design a variable learning rate, which has faster convergence speed and can reduce the total adjustment cost of AGC.

V. CONCLUSION

The contribution of this paper can be summarized as follows:

(1) Based on VGE framework, a novel VGEC strategy is proposed to obtain the optimal cooperative control and fast power allocation of the interconnected microgrids, such that the energy autonomy of the microgrids is realized.

(2) Using PDWoLF-PHC(λ) as the control algorithm of VGEC strategy, an accurate calculation of the winning and losing criteria is possible, and can more quickly converge to Nash equilibrium. HQC strategy with interactive coordination and self-learning is used as the power allocation algorithm of VGEC strategy by constructing a hierarchical power allocation mode. This strategy improves the adaptability of the consistency algorithm in complex random environments and effectively alleviates ''dimension disaster'' that typically results from the large scale of the unit.

(3) The validity of the VGEC strategy is verified by the IEEE standard two-area LFC model. An interconnected

microgrids model with a large amount of distributed energy is used for simulation comparison and verification. The results show that the proposed strategy can improve the control performance of the interconnected microgrids, reduce the power generation cost, and realize the energy autonomy of the microgrids cluster. Compared with other algorithms, the VGEC strategy is more robust and has faster convergence.

REFERENCES

[1] A. Bellini, S. Bifaretti, and F. Giannini, "A robust synchronization method for centralized microgrids," in *Proc. IEEE Energy Convers. Congr. Exposit.*, Sep. 2013, pp. 4587–4594.

[2] Y. Xu, W. Zhang, W. Liu, X. Wang, F. Ferrese, C. Zang, and H. Yu, "Distributed subgradient-based coordination of multiple renewable generators in a microgrid," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 23–33, Jan. 2014.

[3] D. Xu, B. Zhou, Q. Wu, C. Y. Chung, C. Li, S. Huang, and S. Chen, "Integrated modelling and enhanced utilization of power-to-ammonia for high renewable penetrated multi-energy systems," *IEEE Trans. Power Syst.*, early access, Apr. 22, 2020, doi: 10.1109/TPWRS.2020.2989533.

[4] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, "A review of deep learning for renewable energy forecasting," *Energy Convers. Manage.*, vol. 198, Oct. 2019, Art. no. 111799.

[5] G. Dehnavi and H. L. Ginn, "Distributed load sharing among converters in an autonomous microgrid including PV and wind power units," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4289–4298, Jul. 2019.

[6] C. Wang, T. Tian, Z. Xu, S. Cheng, S. Liu, and R. Chen, "Optimal management for grid-connected three/single-phase hybrid multimicrogrids," *IEEE Trans. Sustain. Energy*, early access, Oct. 7, 2019, doi: 10.1109/TSTE.2019.2945924.

[7] M. Xie, X. Ji, X. Hu, P. Cheng, Y. Du, and M. Liu, "Autonomous optimized economic dispatch of active distribution system with multi-microgrids," *Energy*, vol. 153, pp. 479–489, Jun. 2018.

[8] 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 Convers. Manage.*, vol. 103, pp. 82–93, Oct. 2015.

[9] Y. Arya and N. Kumar, "Design and analysis of BFOA-optimized fuzzy PI/PID controller for AGC of multi-area traditional/restructured electrical power systems," *Soft Comput.*, vol. 21, no. 21, pp. 6435–6452, 2016.

[10] A. D. Falehi, "Optimal design of fuzzy-AGC based on PSO & RCGA to improve dynamic stability of interconnected multi area power systems," *Int. J. Autom. Comput.*, vol. 4, pp. 1–11, Apr. 2017.

[11] A. Demiroren and H. L. Zeynelgil, "GA application to optimization of AGC in three-area power system after deregulation," *Int. J. Electr. Power Energy Syst.*, vol. 29, no. 3, pp. 230–240, Mar. 2007.

[12] L. Xi, J. Chen, Y. Huang, Y. Xu, L. Liu, Y. Zhou, and Y. Li, "Smart generation control based on multi-agent reinforcement learning with the idea of the time tunnel," *Energy*, vol. 153, pp. 977–987, Jun. 2018.

[13] H. Z. Wang, G. B. Wang, G. Q. Li, J. C. Peng, and Y. T. Liu, "Deep belief network based deterministic and probabilistic wind speed forecasting approach," *Appl. Energy*, vol. 182, pp. 80–93, Nov. 2016.

[14] T. Yu, Y. M. Wang, W. J. Ye, B. Zhou, and K. W. Chan, "Stochastic optimal generation command dispatch based on improved hierarchical reinforcement learning approach," *IET Gener., Transmiss. Distrib.*, vol. 5, no. 8, pp. 789–797, Aug. 2011.

[15] L. Xi, Z. Zhang, B. Yang, L. Huang, and T. Yu, "Wolf pack hunting strategy for automatic generation control of an islanding smart distribution network," *Energy Convers. Manage.*, vol. 122, pp. 10–24, Aug. 2016.

[16] Y. Zhang, R. Li, W. Zhao, and X. Huo, "Stochastic leader-following consensus of multi-agent systems with measurement noises and communication time-delays," *Neurocomputing*, vol. 282, pp. 136–145, Mar. 2018.

[17] T. Tatarenko, "Stochastic payoff-based learning in multi-agent systems modeled by means of potential games," in *Proc. IEEE 55th Conf. Decis. Control (CDC)*, Dec. 2016, pp. 5298–5303.

[18] L. Xi, Y. Li, Y. Huang, L. Liu, and J. Chen, "A novel automatic generation control method based on the ecological population cooperative control for the islanded smart grid," *Complexity*, vol. 18, Aug. 2018, Art. no. 2456963.

[19] A. Nedic and J. Liu, "On convergence rate of weighted-averaging dynamics for consensus problems," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 766–781, Feb. 2017.

[20] A. K. Sadhu, A. Konar, B. Banerjee, and A. K. Nagar, "Multi-robot cooperative planning by consensus Q-learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 4158–4164.

[21] X. Zhang, T. Yu, and J. Tang, "Optimal CPS command dispatch based on hierarchically correlated equilibrium reinforcement learning," *Autom. Electr. Power Syst.*, vol. 39, no. 8, pp. 80–86, 2015.

[22] C. J. Watkins and P. Dayan, "Technical note: Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, May 1992.

[23] L. Ji, Y. Tang, and Q. Liu, "On hybrid adaptive and pinning consensus for multiagent networks," *Math. Problems Eng.*, vol. 2016, pp. 1–11, Nov. 2016.

[24] C. Wang, X. Li, T. Tian, Z. Xu, and R. Chen, "Coordinated control of passive transition from grid-connected to islanded operation for three/single-phase hybrid multimicrogrids considering speed and smoothness," *IEEE Trans. Ind. Electron.*, vol. 67, no. 3, pp. 1921–1931, Mar. 2020.

[25] Y. Shang, "Leader-following consensus problems with a time-varying leader under measurement noises," *Adv. Dyn. Syst. Appl.*, vol. 2, pp. 255–270, Sep. 2012.

[26] G. Ray, A. N. Prasad, and G. D. Prasad, "A new approach to the design of robust load-frequency controller for large scale power systems," *Electr. Power Syst. Res.*, vol. 51, no. 1, pp. 13–22, Jul. 1999.

[27] X. Zhang and T. Yu, "Virtual generation tribe based collaborative consensus algorithm for dynamic generation dispatch of AGC in interconnected power grids," *Proc. CSEE*, vol. 35, no. 15, pp. 3750–3759, 2015.

[28] B. Zhou, K. W. Chan, and T. Yu, "Q-learning approach for hierarchical AGC scheme of interconnected power grids," *Energy Procedia*, vol. 12, pp. 43–52, Jan. 2011.

[29] H. Wang, Y. Liu, B. Zhou, C. Li, G. Cao, N. Voropai, and E. Barakhtenko, "Taxonomy research of artificial intelligence for deterministic solar power forecasting," *Energy Convers. Manage.*, vol. 214, Jun. 2020, Art. no. 112909.

[30] Y.-T. Wu and F. Porté-Agel, "Modeling turbine wakes and power losses within a wind farm using LES: An application to the horns rev offshore wind farm," *Renew. Energy*, vol. 75, pp. 945–955, Mar. 2015.

[31] P. Kundur, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill, 1994.

[32] B. Awad, J. Ekanayake, and N. Jenkins, "Intelligent load control for frequency regulation in microgrids," *Intell. Autom. Soft Comput.*, vol. 16, no. 2, pp. 303–318, Jan. 2010.

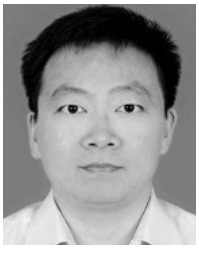
[33] D. Xu, Q. Wu, B. Zhou, C. Li, L. Bai, and S. Huang, "Distributed multi-energy operation of coupled electricity, heating and natural gas networks," *IEEE Trans. Sustain. Energy*, early access, Dec. 23, 2020, doi: 10.1109/TSTE.2019.2961432.

[34] A. K. Saha and S. P. Chowdhury, "Modelling and simulation of microturbine in islanded and grid-connected mode as distributed energy resource," in *Proc. Power Energy Soc. Gen. Meeting-Convers. Del. Elect. Energy Century*, Jul. 2008, pp. 1–7.

[35] F. Gao, B. Blunier, D. Chrenko, D. Bouquain, and A. Miraoui, "Multirate fuel cell emulation with spatial reduced real-time fuel cell modeling," *IEEE Trans. Ind. Appl.*, vol. 48, no. 4, pp. 1127–1135, Jul. 2012.

[36] C. L. Moreira and J. A. P. Lopes, "Microgrids operation and control under emergency conditions," in *Smart Power Grids*. Berlin, Germany: Springer, 2012.

[37] L. Xi, T. Yu, B. Yang, X. Zhang, and X. Qiu, "A wolf pack hunting strategy based virtual tribes control for automatic generation control of smart grid," *Appl. Energy*, vol. 178, pp. 198–211, Sep. 2016.



LEI XI received the M.S. degree in control theory and control engineering from the Harbin University of Science and Technology, and the Ph.D. degree from the School of Electric Power, South China University of Technology, China, in 2016. He is currently an Associate Professor with the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include load frequency control, artificial intelligence techniques, and automatic generation control.

LE ZHANG received the bachelor's degree in electrical engineering from the Southwest University of Science and Technology, China, in 2018. He is currently pursuing the M.S. degree with the College of Electrical Engineering and New Energy, China Three Gorges University. His research interests include smart generation control and artificial intelligence techniques.

JIANCHU LIU received the M.S. degree in control theory and control engineering from the Harbin University of Science and Technology, China, in 2009. He is currently working with Super Vision Optical Technology Company, Ltd. His research interests include system modeling, image recognition, and data analysis.

YUDAN LI received the master's degree in electrical engineering from China Three Gorges University, China, in 2019. She is currently working with Changshou Branch of State Grid Chongqing Electric Power Company. Her major research interest includes artificial intelligence techniques in automatic generation control.

XI CHEN (Member, IEEE) received the M.S. degree in pattern recognition and intelligent systems from the College of Electronic and Information Engineering, Henan University of Science and Technology, Luoyang, China, in 2013, and the Ph.D. degree in power electronics from the School of Electric Power, South China University of Technology, Guangzhou, China, in 2018. His research interest includes the modeling and control of nonlinear systems.

LIUQING YANG (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Minnesota, Minneapolis, in 2004. She is currently a Professor with Colorado State University. Her general interests are in signal processing with applications to communications, networking and power systems – subjects on which she has published more than 310 journal and conference papers, four book chapters, and five books.

SHOUXIANG WANG (Senior Member, IEEE) received the B.S. and M.S. degrees from Shandong University, Jinan, China, in 1995 and 1998, respectively, and the Ph.D. degree from Tianjin University, Tianjin, China, in 2001, all in electrical engineering. He is currently a Professor with the School of Electrical and Information Engineering, Tianjin University. His main research interests are distributed generation, microgrid, and smart distribution systems.