A Distributed Two-Layer Frequency Compensation for Islanded Microgrids Based on Q-learning and PI Controllers

Li, Sijia; Gao, Xian; Blaabjerg, Frede; Anvari-Moghaddam, Amjad

Published in:
8th IEEE Workshop on the Electronic Grid (eGRID)

DOI (link to publication from Publisher):
10.1109/eGrid58358.2023.10380864

Publication date:
2023

Document Version
Accepted author manuscript, peer reviewed version

A Distributed Two-Layer Frequency Compensation for Islanded Microgrids Based on Q-learning and PI Controllers

Sijia Li, Xian Gao, Frede Blaabjerg, Amjad Anvari-Moghaddam
Department of Energy (AAU Energy)
Aalborg University
Aalborg, Denmark
sili, xiga, fbl, aam@energy.aau.dk

***Abstract*—Frequency instability generates significant challenges to the stability of the system. To solve the frequency deviation problem, the traditional secondary control uses PID controller to achieve frequency compensation for the primary control, but simultaneously the traditional PID controller has disadvantages such as poor dynamic performance and the need for manual tuning of parameters. The problems mentioned above will lead to poor compensation accuracy. To address such issue, this paper proposes a new frequency compensation scheme that divides the traditional frequency secondary control into two layers, the first layer uses an improved PID controller that considers the average value of output frequency of all distributed generators, and the second layer is based on Q-learning technology to compensate again. The proposed scheme shortens the response time and improves the control accuracy, and effectiveness is verified in MATLAB/Simulink.**

***Index Terms*—Secondary Control, Frequency Compensation, Q-Learning, PID Controller, Islanded Microgrid**

I. Introduction

With the development of renewable energy, the emergence of a large number of distributed generators (DGs) in the power system has changed the control and operation scheme of the previous system, creating new challenges. For that, Consortium for Electric Reliability Technology Solutions (CERTS) came out with the concept of microgrid [1]. Meanwhile, to solve the problems of power distribution, and frequency compensation in microgrids, distributed hierarchical control schemes are extensively accepted by scholars [2]. In islanded microgrids, droop control as a primary control scheme to achieve power distribution is one of the main control options [3-4]. Some progress has been made in the proposed improvement schemes based on droop control. In [5] the authors proposed a virtual multi-relaxation droop control to achieve power distribution by regulating the bus voltage. In [6], the authors improved the conventional droop control to adaptive droop control, so that the control gain can be changed with the line conditions to improve the power distribution accuracy. At the same time, the system will have frequency deviations, which are caused by various factors. A sudden increase or decrease in load can result in a deviation of the system output frequency from the reference value. Unbalanced microgrid generations can also result in frequency deviations, and system faults or environmental factors can lead to frequency instability [7]. Therefore, frequnency compensation by secondary control is required. In [8], frequency compensation is proposed at multiple time levels and the system output frequency partition is compensated by priority. In [9], the authors proposed secondary control based on small AC signal injection (SACS-SFC), which achieves frequency compensation by injecting power signals on additional buses for each DG. However, the aforementioned control schemes suffer from poor dynamic characteristics and the traditional control is less robust under complex operating conditions. To solve the above problems, machine learning (ML) based-control schemes seem to be more future-oriented [10]. ML can be mainly classified as supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, deep learning and transfer learning [11]. Among them, reinforcement learning techniques without a priory data have been paid more attention by scholars in the secondary control of microgrids. A deep reinforcement learning scheme based on deep deterministic policy gradient (DDPG) was proposed in [12] for the secondary control of a microgrid, instead of using PID controller. The multi-intelligence quantum deep reinforcement learning was proposed in [13], which uses an algorithm that combines a neural network with a quantum simulator for distributed compensation of frequency deviations. In [14], a deep reinforcement learning technique with multiple intelligences was proposed to achieve secondary compensation of frequency for complex control environments with multiple regions. In [15], the authors proposed an improved reinforcement learning-based scheme i.e., a brain emotional learning-based intelligent controller (BELBIC), for secondary control.

Based on previous studies, in this paper, a distributed two-layer frequency compensation scheme is proposed for islanded microgrids. In the first layer, besides considering individual DG frequency output values, the conventional PID controller compensation scheme is replaced by an improved PID controller solution using the average value of the overall DG output frequency. In the second layer, the frequency is compensated again on top of the compensation output of the first layer by employing a Q-learning based technique. Two layers of control are applied to ensure the accuracy of the control and to improve the dynamic performance for optimal compensation.

II. Primary Control of Islanded Microgrids

Droop control has excellent "plug-and-play" characteristics and does not require a communication link to the neighboring control units. In this paper, droop control is used at the primary level for the studied islanded microgrid to enable proportional power distribution as shown in (1) and (2).

$$f_i = f^* - m_i(P_i - P_i^*) \tag{1}$$

$$u_i = u^* - n_i(Q_i - Q_i^*) \tag{2}$$

where f_i and u_i,represent the frequency and voltage, f^* and u^* are corresponding rated values and m_i, n_i are the droop factors of the ith DG output. Where P_i and Q_i ,represent the active power and reactive power outputs by the ith DG, while P^*,Q^* are rated values.

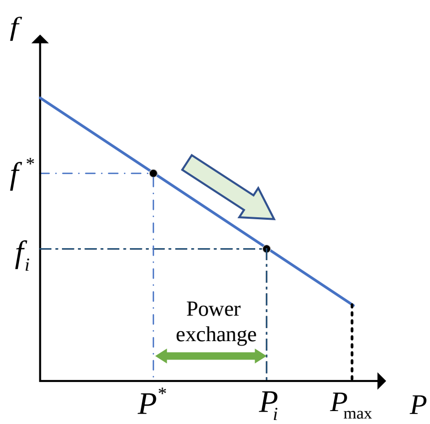

Fig. 1. Frequency droop control diagram.

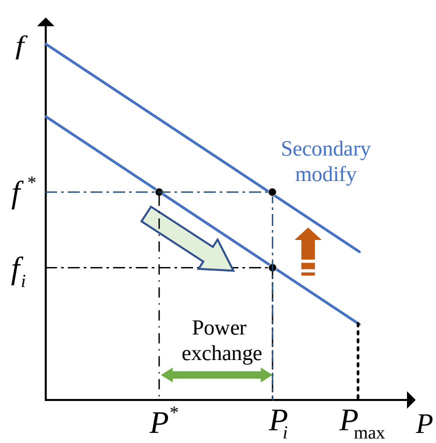

Fig. 2. Frequency droop control with secondary compensation diagram.

Whenever the load changes, the active power output will change, and the output frequency will deviate according to the droop control curve as shown in Fig. 1. In this paper, the secondary control compensation is realized and carried out for the islanded microgrid as shown in Fig. 2, where the system can achieve stable operation following a disturbance.

III. Proposed Novel Secondary Control

In this paper, an islanded microgrid hierarchical control system with four DGs is built. In order to compensate the frequency deviation of primary control with droop control, this paper proposes a new distributed secondary control with a two-layer control structure. The first layer adopts an improved PID-based controller, and the second layer uses Q-learning based control to compensate on the basis of the first layer to improve the compensation accuracy and the response speed.

A. First-layer control based on improved PID control

PID controllers are accepted by industry because of their ease of implementation, low cost, robustness and stability. The secondary control method based on PID control is also one of the common compensation methods for microgrids. In this paper, the traditional PID control structure is improved as shown in Fig. 3.

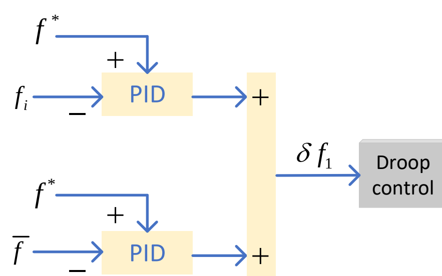

Fig. 3. Improved PID control diagram.

In this structure, the original single PID controller is replaced with a dual PID controller. The first PID controller uses the i-th DG output measured frequency value f_i for compensation, and the compensation signal can be obtained from (3). The second PID controller uses the average value of the output frequency $\overline{f}$ of the all DGs during stable operation to compensate as stated in (4).

$$f_{PI}^1 = k_p^1(f^* - f_i) + k_i^1 \int (f^* - f_i)\mathrm{d}x \tag{3}$$

$$f_{PI}^2 = k_p^2(f^* - \overline{f}) + k_i^2 \int (f^* - \overline{f})\mathrm{d}x \tag{4}$$

where k_p^1, k_i^1, k_p^2, k_i^2 are proportional and integral gains of the PID controllers. The sum of these two PID controller outputs, δf_1(as shown in (5)) will be entered into the droop control loop as a compensation term.

$$\delta f_1 = f_{PI}^1 + f_{PI}^2 \tag{5}$$

B. second-layer control based on Q-learning technology

Q-learning is a reinforcement machine learning-type algorithm that trains agents to make optimal decisions in dynamic and complex environments by accumulating reward signals through reward functions or penalty terms. It is a model-free, offline policy algorithm that does not require a prior knowledge about transition probabilities or environmental rewards, the control agent learns the optimal policy by interacting with the environment. Therefore, Q-learning is widely used in the fields of robot control, and autonomous driving. In this paper, a Q-learning based controller is developed to recompense the frequency output received from the first control layer. The Q function, also known as the state action value function, is certainly a value-based learning technique. In this paper, the Bellman equation is used to update the value function. Q-value function is shown in (6) where s is state value, and a is action value chosen in s under strategy π.

$$Q_{t+1}^{\pi}(s,a) = Q_t^{\pi}(s,a) + \alpha[r(s,a) + \gamma max(Q_t^{\pi}(s^{'},a^{'}) - Q_t^{\pi}(s,a))] \tag{6}$$

Also, a,s represent the current value of action and state, a', s' represent the next moment value, and α represents the learning rate. γ represents the discount factor, which is the weight that the agent places on obtaining the reward at the time of making the action. It is chosen between 0 and 1. γ values close to nearly 1 imply that the subject is more influenced by future long-term rewards, while gamma values close to 0 imply that the subject only considers short-term immediate rewards. The choice of γ depends on the specific problem and the agent's goal. $r(s,a)$ is the reward function obtained after taking action a. Since Q-learning requires the expected cumulative reward for following the optimal strategy, the agent starts from the current state, accumulates the rewards of subsequent states, and will obtain the expected cumulative reward, as shown in (7).

$$V_{\pi}(s) = E[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + ...] \tag{7}$$

In Q-learning, the approach of greedy strategy is proposed in order to implement the action of choosing the action that maximizes the Q-value of the current state. However, the greedy strategy sometimes leads the agent to fall into the misconception of suboptimal strategies. If the agent only focuses on the actions that can produce high Q-values in the current state without considering other actions that may have lower Q-values but will bring higher returns in the long term, it will fall into the misconception. Therefore, in this paper, we use the ϵ-greedy strategy to balance the exploration of the current value with the consideration of the future state. ϵ-greedy strategy is shown in (8), where ϵ is the probability of making an arbitrary action, while 1-ϵ is the probability of making the action with the highest Q-value among the existing actions.

$$\pi(s,a) = \begin{cases} 1-\varepsilon+\frac{\varepsilon}{|A(s)|}, & a = argmaxQ(s,a) \\ \frac{\varepsilon}{|A(s)|} & otherwise \end{cases} \tag{8}$$

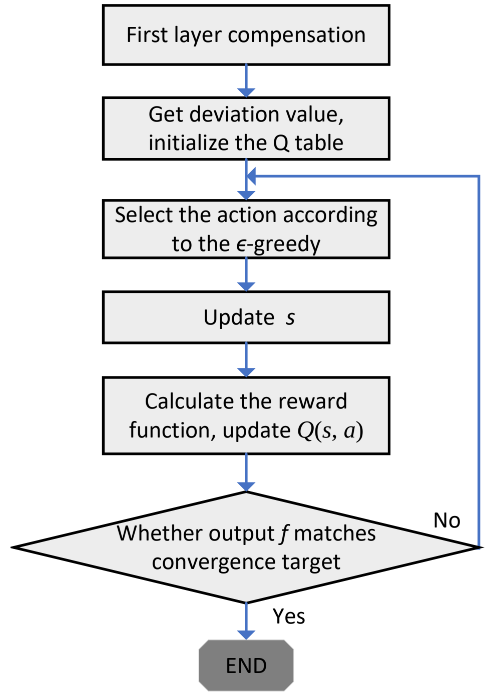

Fig. 4. Q-learning based second layer frequency compensation process diagram.

The process diagram of the second layer control algorithm is shown in Fig. 4. In this paper, considering that the frequency deviation has been compensated in the first layer of control, according to [16], the obtained discrete frequency state quantity is set as: S={(-∞,-0.06), [-0.06,-0.04), [-0.04,-0.02), [-0.02,0.02], (-0.02,0.04], (0.04,0.06], (0.06,+∞)}. The action is set to A={-0.07, -0.035, -0.05, -0.03, -0.005, 0, 0.005, 0.03, 0.05, 0.035, 0.07}. The rewards is set as r={-120.244, -100, -1000, 0, -1000, -100, -120.244}. The learning rate is set to 0.1, the discount factor is set to 0.9.

After implementing the second layer of Q-learning based secondary compensation again, the droop control frequency output can be realized as (9).

$$f_i = f^* - m_i(P_i - P_i^*) + \delta f_1 + \delta f_2 \tag{9}$$

δf_1 is the first layer frequency compensation value, and δf_2, i.e. a, is the action taken by Q-learning, which is also the second layer of frequency compensation amount.The entire system control structure diagram is shown in Fig. 5.

IV. Simulation results

To verify the effectiveness of the proposed control scheme, an islanded microgrid consisting of four DGs is modeled in Matlab 2023a with a number of load change scenarios. The simulation structure is shown in Fig. 6. The initial value of the load is 46kW, when t=1s, the load changes to 62kW, and finally at t=2s, the load is decreased to 54kW. The PID parameters in the first layer of secondary control are selected as $k_p^1 = 0.8, k_i^1 = 10, k_p^2 = 0.08, k_i^2 = 10$. The droop

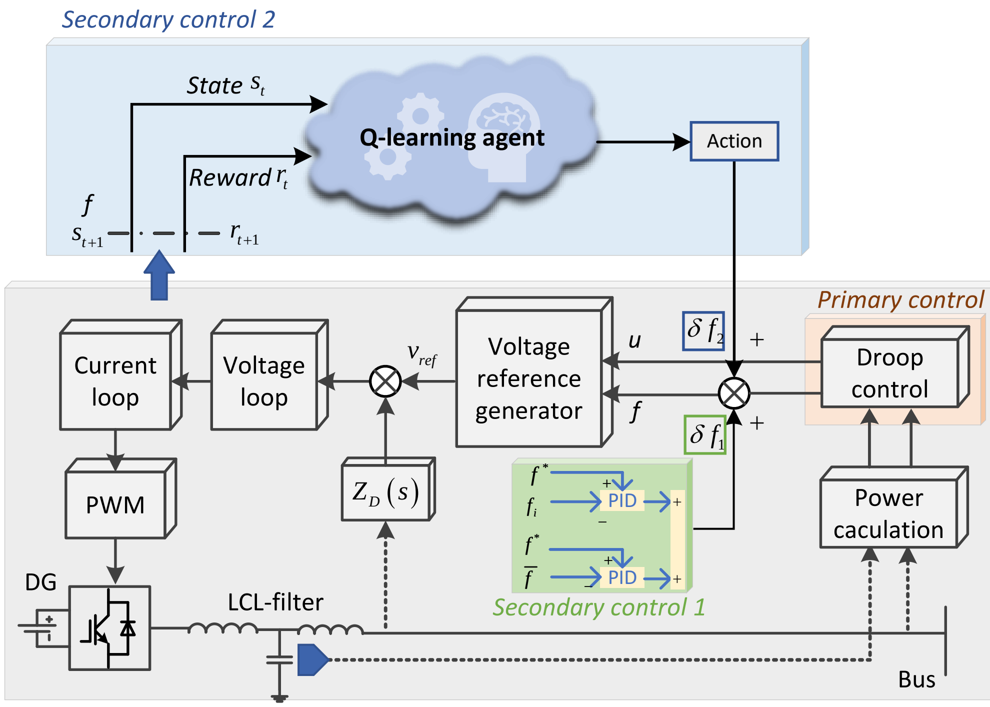

Fig. 5. Single DG hierarchical control diagram.

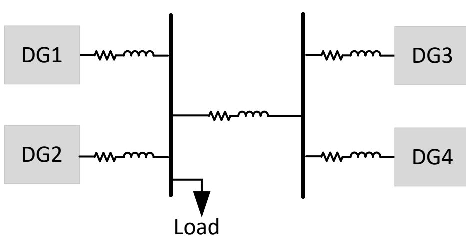

Fig. 6. Simulation structure diagram consisting of 4 DGs.

coefficients of active power is selected as shown in Table I. The droop coefficients of reactive power is 2e-4$V/VAr.s$.

TABLE I
DROOP GAIN OF FOUR DGS(rad/W)

DG1	DG2	DG3	DG4
0.552e-4	0.373e-4	0.674e-3	1e-3

Firstly, the distributed secondary control with only a single PID controller is used for frequency compensation, and the simulation results are shown in Fig. 7. Secondly, the first layer of improved PID control secondary control proposed in this paper is used for compensation, and the results are shown in Fig. 8. Finally, the compensation is carried out using the improved PID controller and Q-learning proposed in this paper as shown in Fig. 9 and the active power output results are given as shown in Fig. 10.

The above three frequency compensation results are summarized in Table II. From the simulation results, it can be seen that the control proposed in this paper achieves the best results for frequency compensation, reduces the overshoot and shortens the time required to reach the reference values. The settling time is defined as the time required for the system to reach the steady value [49.99, 50.01].

TABLE II
SIMULATION RESULTS COMPARISON

Method	Settling time	Max overshoot	Mean
PID	1s	0.25Hz	49.9Hz
Improved PID	0.9s	0.23Hz	49.9Hz
Improved PID+Q-learning	0.45s	0.23Hz	50Hz

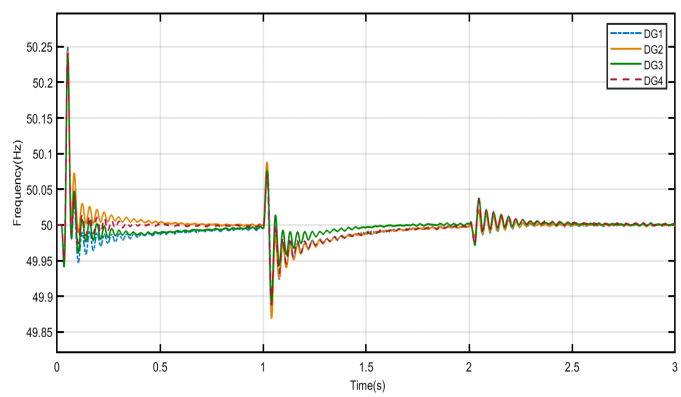

Fig. 7. Secondary compensation results based on PID.

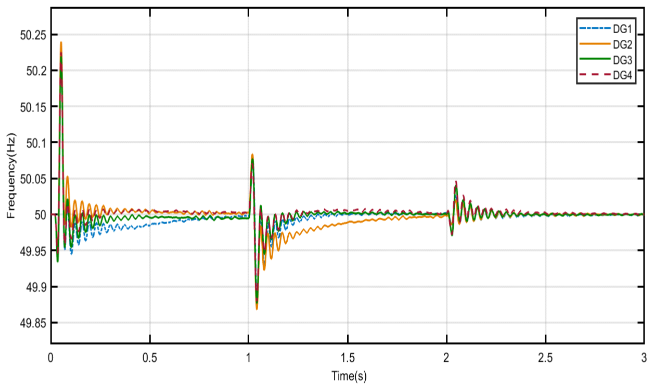

Fig. 8. Secondary compensation results based on improved PID.

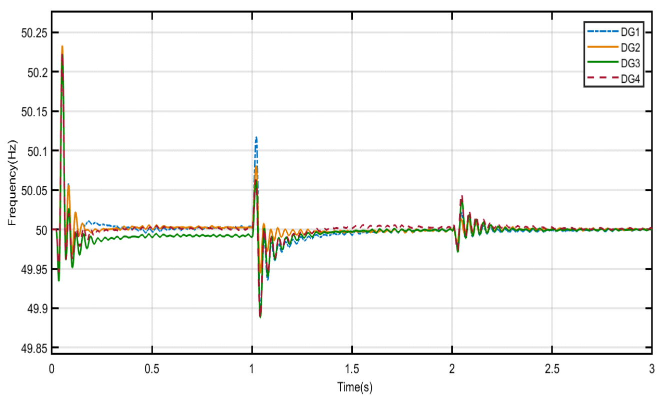

Fig. 9. Secondary compensation results based on improved PID and Q-learning double-layer compensation.

V. Conclusion

In this paper, a distributed two-layer control structure for frequency secondary control was proposed. The first layer adopted an improved PID control that considered the average value of output frequency when all DGs operate stably, and the second layer adopted a Q-learning-based control scheme to enable accurate compensation of DG output frequency. Compared to the conventional PID quadratic compensation, the proposed strategy in this paper gave more accurate compensation results with less settling time. The Q-learning compensation was used after a layer of improved PID control. This scheme reduced the impact of Q-learning uncertainty on the system and ensured that the system has the ability to operate properly. How to reduce the overshoot, with the implementation of voltage compensation will be further studied in the future.

References

[1] A. Anvari-Moghaddam, H. Abdi, B. Mohammadi-Ivatloo, and N. Hatziargyriou, "Microgrids - Advances in Oper-ation, Control, and Protection", *Springer*, 2021, doi: 10.1007/978-3-030-59750-4.

[2] N. Vafamand, M.M. Arefi, A. Anvari-Moghaddam, "Advanced Kalman Filter-based Backstepping Control of AC Microgrids: A Command Filter Approach", *IEEE Systems Journal*, vol. 17, no.1, pp. 1060-1070, 2023.

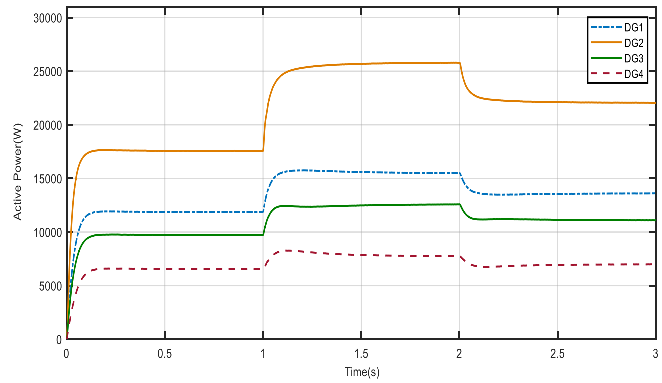

Fig. 10. Active power output results based on improved PID and Q-learning double-layer compensation.

[3] M. Mohiti, H. Monsef, A. Anvari-Moghaddam, H. Lesani, "Two-Stage Robust Optimization for Resilient Operation of Microgrids Considering Hierarchical Frequency Control Structure", *IEEE Trans. Industrial Electronics*, vol. 67, no. 11, pp. 9439-9449, 2020.

[4] Y. Song, S. Sahoo, Y. Yang and F. Blaabjerg, "Quantitative Mapping of Modeling Methods for Stability Validation in Microgrids", *IEEE Open Journal of Power Electronics*, vol. 3, pp. 679-688, 2022.

[5] D. Choi, J. -W. Park and S. H. Lee, "Virtual Multi-Slack Droop Control of Stand-Alone Microgrid With High Renewable Penetration Based on Power Sensitivity Analysis," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3408-3417, May 2018.

[6] B. Alghamdi and C. A. Cañizares, "Frequency Regulation in Isolated Microgrids Through Optimal Droop Gain and Voltage Control", *IEEE Transactions on Smart Grid,* vol. 12, no. 2, pp. 988-998, March 2021.

[7] M. Farrokhabadi et al., "Microgrid Stability Definitions, Analysis, and Examples", *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 13-29, Jan. 2020.

[8] Z. Zhao, P. Yang, J. M. Guerrero, Z. Xu and T. C. Green, "Multiple-Time-Scales Hierarchical Frequency Stability Control Strategy of Medium-Voltage Isolated Microgrid", *IEEE Transactions on Power Electronics*, vol. 31, no. 8, pp. 5974-5991, Aug. 2016.

[9] B. Liu, T. Wu, Z. Liu and J. Liu, "A Small-AC-Signal Injection-Based Decentralized Secondary Frequency Control for Droop-Controlled Islanded Microgrids", *IEEE Transactions on Power Electronics*, vol. 35, no. 11, pp. 11634-11651, Nov. 2020.

[10] E. Mohammadi, M. Alizadeh, M. Asgarimoghaddam, X. Wang and M. G. Simões, "A Review on Application of Artificial Intelligence Techniques in Microgrids", in *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 3, no. 4, pp. 878-890, Oct. 2022.

[11] W. Dong, Q. Yang, W. Li and A. Y. Zomaya, "Machine-Learning-Based Real-Time Economic Dispatch in Islanding Microgrids in a Cloud-Edge Computing Environment", *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13703-13711, 1 Sept.1, 2021.

[12] Y. Xia, Y. Xu, Y. Wang and S. Dasgupta, "A Distributed Control in Islanded DC Microgrid based on Multi-Agent Deep Reinforcement Learning", *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, Singapore, pp. 2359-2363, 2020.

[13] R. Yan, Y. Wang, Y. Xu and J. Dai, "A Multiagent Quantum Deep Reinforcement Learning Method for Distributed Frequency Control of Islanded Microgrids", *IEEE Transactions on Control of Network Systems*, vol. 9, no. 4, pp. 1622-1632, Dec. 2022.

[14] Z. Yan and Y. Xu, "A Multi-Agent Deep Reinforcement Learning Method for Cooperative Load Frequency Control of a Multi-Area Power System", *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4599-4608, Nov. 2020.

[15] M. S. O. Yeganeh, A. Oshnoei, N. Mijatovic, T. Dragicevic and F. Blaabjerg, "Intelligent Secondary Control of Islanded AC Microgrids: A Brain Emotional Learning-Based Approach", *IEEE Transactions on Industrial Electronics*, vol. 70, no. 7, pp. 6711-6723, July 2023.

[16] W. Liu et al., "Distributed Secondary Control Strategy Based on Q - learning and Pinning Control for Droop-controlled Microgrids", *Journal*

of Modern Power Systems and Clean Energy, vol. 10, no. 5, pp. 1314-1325, Sept. 2022.