

Distributed reinforcement learning to coordinate current sharing and voltage restoration for islanded DC microgrid



Zifa LIU¹ , Ya LUO¹, Ranqun ZHUO¹, Xianlin JIN²

Abstract A novel distributed reinforcement learning (DRL) strategy is proposed in this study to coordinate current sharing and voltage restoration in an islanded DC microgrid. Firstly, a reward function considering both equal proportional current sharing and cooperative voltage restoration is defined for each local agent. The global reward of the whole DC microgrid which is the sum of the local rewards is regarded as the optimization objective for DRL. Secondly, by using the distributed consensus method, the predefined pinning consensus value that will maximize the global reward is obtained. An adaptive updating method is proposed to ensure stability of the above pinning consensus method under uncertain communication. Finally, the proposed DRL is implemented along with the synchronization seeking process of the pinning reward, to maximize the global reward and achieve an optimal solution for a DC microgrid. Simulation studies with a typical DC microgrid demonstrate that the proposed DRL is computationally efficient and able to

provide an optimal solution even when the communication topology changes.

Keywords Distributed reinforcement learning (DRL), Distributed information discovery, DC microgrid, Local reward function

1 Introduction

DC microgrids attract increasing attention in recent years for two reasons: 1) on one hand, the emerging diversity of distributed generators (DGs) includes a majority of DC generators, such as photovoltaics (PVs), fuel cells (FCs) and energy storage systems (ESSs); 2) on the other hand, there will be more and more DC loads, such as electric vehicles (EVs), DC relays, in future smart grids [1–3]. Generally, a DC microgrid has several advantages compared with an AC microgrid: 1) it has less loss from power transformation because there are no AC/DC converters, 2) it avoids some problems often occurring in an AC microgrid, for instance, harmonics and synchronization; and 3) it can have improved power quality and reliability, because reactive power compensation is not needed from the power supply [4–6]. Therefore, much current research focusses on the control and management of DC microgrids.

Droop control is generally accepted as an effective solution for DC microgrids, and such applications of droop control have been investigated in many papers [7, 8]. However, it is hard to achieve predictive, accurate load sharing and voltage regulation by using the droop control without communication, and moreover, both line impedances and output impedances of DGs will affect the accuracy of load sharing [9, 10]. Therefore, hierarchical

CrossCheck Date: 26 July 2017

Received: 20 July 2016 / Accepted: 26 July 2017 / Published online: 23 September 2017

© The Author(s) 2017. This article is an open access publication

✉ Zifa LIU
tjubluesky@163.com

Ya LUO
luoyaturbo@163.com

Ranqun ZHUO
zrqSirius@163.com

Xianlin JIN
jinxianlin@guohua.com.cn

¹ North China Electric Power University, Beijing 102206, China

² Guohua Energy Investment Co., Ltd., Beijing 100007, China

control schemes which consist of primary and secondary control have been proposed and utilized to solve these problems.

The structures of hierarchical control schemes can be centralized or distributed [11–13]. A typical centralized control scheme was proposed in [14]; it collected global voltage and current information by low bandwidth communication and realized voltage restoration and enhanced current sharing accuracy. It is well known that a centralized control scheme requires a complicated communication network to collect global operating conditions and a powerful central controller to process the huge amount of information. Thus, centralized schemes are costly to implement and susceptible to single-point failures. Furthermore, taking the uncertainty of intermittent DGs into consideration, a generation fluctuation may result in unintentional structural changes in current flows, which will further increase the burden on centralized schemes [15, 16]. Advantages of a distributed scheme include the ability to survive unexpected disturbances and decentralized data updating, which leads to efficient information sharing and eventually a faster decision-making process and operation [8, 17–19].

Much research focuses on improving distributed control in multiple ways in AC or DC microgrids. [20] proposed a distributed cooperative control strategy based on a multi-agent system (MAS) that involves primary and secondary frequency control and multi-stage load shedding to achieve cooperative frequency recovery. [21] used input-output feedback linearization to convert secondary voltage control to a linear second-order tracker synchronization problem. A pinning-based scheme for microgrids is proposed to obviate the requirements for a central controller and a complex communication topology, and to achieve control under both fixed and uncertain communication topologies in [22]. With regard to DC microgrids, [23–25] proposed two kinds of distributed control schemes, which discover global current information and adjust the droop control gains using a distributed consensus algorithm, and implemented accurate load distribution in DC microgrids.

Hence, a distributed control scheme can be regarded as a feasible solution for DC microgrids in this study. Another problem that needs to be considered in a DC microgrid is to coordinate the following two objectives in which exist inherent contradictions: 1) to implement voltage restoration in DC buses; and 2) to realize accurate current or load sharing in a DC microgrid. For a DC microgrid, the average output current of each DG reflects the load fluctuation of the whole system, and can also reflect the voltage deviation caused by a load change. Thus, the average output current of the DGs is selected as the control input to simultaneously realize voltage adjustment and load proportional distribution of current.

To address the above problems, reinforcement learning (RL) has been introduced to the distributed control scheme, and this would be a possible solution [26]. RL is a simple iterative algorithm that learns to act in an optimal way through a reward signal evaluated by the performance of prior solutions obtained. Over the past few years, several multi-agent based RL algorithms have been proposed and applied to practical problems [27, 28]. The RL algorithm has major advantages. It is an online learning algorithm directly interacting with the environment, and it does not require an accurate model of the environment. It only needs a reward function to evaluate the quality of a solution instead of complicated mathematical operations. Finally, it has the ability to escape local minima because it performs stochastic optimization [29–32].

Inspired by distributed control and the RL algorithm, a novel distributed RL (DRL) approach for a DC microgrid is proposed and investigated in this study. It can achieve the same control performances as a centralized control scheme while overcoming some of its problems. It also can coordinate voltage restoration and load sharing during secondary control, and implement accurate current sharing while recovering the DC voltages. DRL with reward feedback and applying the distributed consensus method through pinning control are the distinguishing features of this work. More specifically, the main contributions of this study are as follows:

- 1) Proposal of a new DRL method, which combines RL and the distributed consensus method together to achieve an optimal solution for a DC microgrid.
- 2) Proposal of an evaluation method using a global reward discovered locally, which can be used to evaluate the control performance of DRL considering both equal proportional current sharing and cooperative voltage restoration for an islanded DC microgrid.
- 3) Proposal of a distributed consensus method through pinning control, which can be applied to discover global information or to achieve synchronization by seeking a pinning consensus value. Additionally, the corresponding adaptive updating method can adapt to changes of communication topology, including both exchanging coefficients and updating the identity of participating agents.

The rest of this paper is organized as follows: Section 2 presents a brief introduction to hierarchical control of a DC microgrid and the distributed consensus method through pinning control; Section 3 elaborates on the proposed DRL, including its reward function, the distributed consensus method through pinning control, and its detailed control process; the proposed DRL is simulated and investigated with a typical system in Section 4; and finally, conclusions are presented.



2 Preliminary

2.1 Hierarchical cooperative control of DC microgrid

Typically, a DC microgrid has a two-layered hierarchical control structure, comprising the primary control layer and the secondary control layer. Primary control, which is usually implemented by droop control, aims at quick response to maintain the stability of a DC microgrid. Whereas, secondary control has two control objectives: 1) to restore voltage and 2) to share the load in a suitable proportion.

In contrast to droop control in an AC microgrid, droop control in a DC microgrid is based on the predefined relationship between voltage and current as follows:

$$\begin{cases} U_i = U_{ref,i} - m_i I_i \\ U_{ref} = U_N - \lambda_V / 2 \\ m = \lambda_V / I_{max} \end{cases} \quad (1)$$

where $U_{ref,i}$ is the voltage reference of the i^{th} DG; m_i is the droop control gain of the i^{th} DG; I_i is the measured value of the output current of the i^{th} DG; λ_V indicates the maximum voltage deviation; U_N is the rated voltage; and I_{max} is the maximum current of the droop controller.

However, fully decentralized droop control may cause steady state deviations if there is no communication among droop-controlled DGs. To address this problem, secondary control is utilized to improve voltage restoration in DC buses and realize predictive load sharing in a DC microgrid. It is accomplished by controlling the voltage reference U_{ref} in (1) as follows:

$$\begin{cases} U_i = (U_{ref,i} + \Delta U_i) - m_i I_i \\ \Delta U_i = \Delta U_{C,i} + \Delta U_{V,i} \end{cases} \quad (2)$$

where the adjustment of U_{ref} can control both voltage and current. Thus, the control change of voltage reference ΔU_i is divided into the current adjustment term $\Delta U_{C,i}$ and the voltage adjustment term $\Delta U_{V,i}$; the control of $\Delta U_{C,i}$ aims at realizing proportional power dispatch, and $\Delta U_{V,i}$ aims at correcting the voltage deviation [19, 20].

2.2 Distributed consensus method through pinning control

2.2.1 Pinning-based distributed consensus method

Assume that r_i denotes the state variable of agent i . The distributed consensus method through pinning control can be expressed in a discrete form as follows:

$$r_i^{[k+1]}(t) = \sum_{j \in N_i} \alpha_{ij} \{ r_j^{[k]}(t) - r_j^{[k]}(t) \} - d_i \{ r_i^{[k]}(t) - r_p^* \} \quad (3)$$

where $i = 1, 2, \dots, n; j = 1, 2, \dots, n; n$ indicates the total number of participating agents; k is the discrete-time index; $r_i^{[k+1]}$ is the state of agent i at iteration $k + 1$, which corresponds to the local information defined in this study; $r_i^{[k]}, r_j^{[k]}$ are respectively the states of agents i and j at iteration k ; and α_{ij} is the connectivity coefficient between agents i and j . If agents i and j are connected through a communication line, $\alpha_{ij} \neq 0$, otherwise, $\alpha_{ij} = 0$. N_i expresses the neighboring agent set of the i^{th} agent; d_i is the pinning gain of the i^{th} agent, $d_i \geq 0$ with $d_i = 0$ when there is no pinning control over agent i ; and r_p^* is the preset pinning consensus value of the consensus method.

Generally, the method in (3) can be used to control all agents to the preset pinning consensus value using the connectivity coefficients among them. When $d_i = 0$, the method in (3) also can be used to discover global information as for other average consensus methods [19, 20, 22, 26].

For convenient analysis, define the control error as:

$$e_i^{[k]} = r_i^{[k]} - r_p^* \quad (4)$$

Then, the distributed consensus method based on pinning described in (3) can be rewritten in terms of (4) as follows:

$$e_i^{[k+1]}(t) = \sum_{j \in N_i} \alpha_{ij} \{ e_j^{[k]}(t) - e_j^{[k]}(t) \} - d_i e_i^{[k]}(t) \quad (5)$$

Hence, the consensus process of the whole DC microgrid based on pinning can be illustrated as

$$\begin{cases} E^{[k+1]}(t) = [A - (D \otimes I)]E^{[k]}(t) \\ A = [\alpha_{ij}] \\ D = [d_i] \end{cases} \quad (6)$$

where $E^{[k]}$ is the information matrix; A is the communication updating matrix that is determined according to the communication topology; D is the pinning matrix; I is the identity matrix; “ \otimes ” indicates Kronecker product of matrix.

2.2.2 Adaptive updating method

To adapt to communication link changes, a connectivity coefficient updating method is proposed in (7). Here, $\Delta(t)$ is utilized to express the communication topology changes in an DC microgrid; δ is the consensus constant, the value of which can affect the convergence characteristics of the two-layer algorithm, $0 < \delta < 2$; $n_{i,\Delta(t)}$ and $n_{j,\Delta(t)}$ respectively indicate the number of agents in the neighborhood of agents i and j according to the communication topology. Both $n_{i,\Delta(t)}$ and $n_{j,\Delta(t)}$ are local information which can be

detected by corresponding agents, so (7) can adapt locally to the communication link changes.

$$\alpha_{ij} = \begin{cases} \frac{\delta}{n_{i,\Delta(t)} + n_{j,\Delta(t)}} & j \in N_{i,\Delta(t)} \\ 1 - \sum_{j \in N_{i,\Delta(t)}} \frac{\delta}{n_{i,\Delta(t)} + n_{j,\Delta(t)}} & j = i \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Additionally, to adapt to changes in the number of agents and thereby meet the plug-and-play operation requirements for a DC microgrid, an agent identity updating method is proposed. If (3) is initialized with the predefined index i , and all d_i are set to 0, it will converge to the average value of total number of agents. Thus, the total number of agents can be determined by

$$\begin{cases} n_{a,i} = i/n_{\Delta(t)} \\ n_{\Delta(t)} = i/n_i = i/[i/n_{\Delta(t)}] \end{cases} \quad (8)$$

where $n_{a,i}$ is the average value discovered by agent i , and $n_{\Delta(t)}$ is the total number of participating agents in the DC microgrid, which will be adaptively adjusted when the number changes.

2.2.3 Stability proof

To verify the stability of the proposed information discovery method, a positive Lyapunov function L is defined, and the partial derivative of L with respect to $e^{[k]}$ is derived as follows:

$$\begin{aligned} L &= \frac{1}{2} \sum_{i=1}^n (e_i^{[k]})^T e_i^{[k]} \\ \Delta L &= \sum_{i=1}^n (e_i^{[k]})^T e_i^{[k]} \\ &= \sum_{i=1}^n (e_i^{[k]})^T \left[\sum_{j \in N_i} \alpha_{ij} (e_j^{[k]} - e_i^{[k]}) - d_i e_i^{[k]} \right] \\ &\leq \sum_{i=1}^n (e_i^{[k]})^T \sum_{j \in N_i} \alpha_{ij} (e_j^{[k]} - e_i^{[k]}) - \sum_{i=1}^n d_i \|e_i^{[k]}\|^2 \\ &\leq \sum_{i=1}^n \|e_i^{[k]}\| \left[\sum_{j \in N_i} \alpha_{ij} (\|e_j^{[k]}\| + \|e_i^{[k]}\|) - \sum_{i=1}^n d_i \|e_i^{[k]}\|^2 \right] \\ &= |E^{[k]}|^T (A - D \otimes I) |E^{[k]}| \end{aligned} \quad (9)$$

Therefore, to ensure the stability of the distributed consensus method, the stability condition can be finally expressed as

$$A - D \otimes I \leq 0 \Rightarrow \Delta L \leq 0 \quad (10)$$

where $\Delta L \leq 0$ implies that the stability of the proposed consensus method can be ensured and consensus will be reached asymptotically.

3 Distributed reinforcement learning control (DRLC) for a DC microgrid

In this study, a DC microgrid is considered as an MAS, which includes distributed generator agents (DGAs), energy storage system agents (ESSAs) and load agents (LAs). By implementing specific characteristics of agents in a MAS, such as autonomy, sociality, proactivity, and adaptability, the agents can provide greater functionality than traditional controls and cater to the special needs and difficulties of the proposed control [19–22]. The proposed DRL scheme can immediately take action in the event of disturbances and realize distributed decision-making to achieve cooperative recovery.

Furthermore, DRL for agents, which is a simple iterative algorithm by which optimal actions are learnt through rewards gained by exploring the unknown environment, can be applied to improve the control characteristics. As illustrated in Fig. 1, during the process of DRL, the solution is updated according to its performance as evaluated by the corresponding reward signal. Hence, each agent can optimize its control solution for the associated generator, storage, or load, while some elements of its solution can be communicated to other agents to arrive at a shared solution.

To implement such a distributed DRL, two related problems of defining the local reward function and achieving distributed consensus based on pinning are described in detail below.

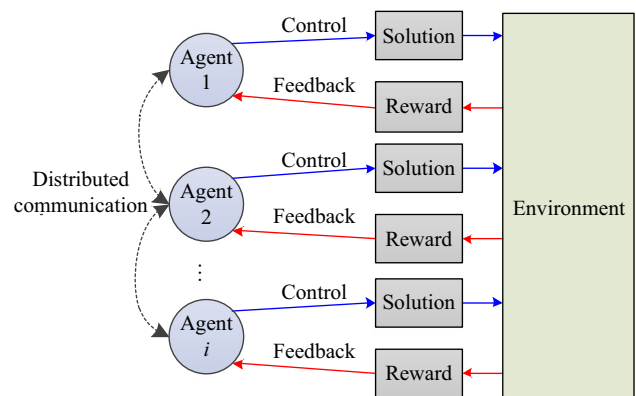


Fig. 1 Fundamental control structure of DRL

3.1 Definition of reward for DRL

For DRL the main challenge is finding the global reward of the entire system. It is hard to obtain the global reward directly under a distributed communication framework where each agent can exchange information only with its neighboring agents. Thus, a local reward function is designed to evaluate the performance of a candidate solution.

Firstly, to take into account the equal proportional current sharing in the DC microgrid, a proportional coefficient for the i^{th} agent is defined by

$$\kappa_i = \frac{I_i}{I_{N,i}} \quad (11)$$

where I_i is the measured current and $I_{N,i}$ is the rated current of the i^{th} agent. By using the distributed consensus method illustrated in (3), κ_i can be shared among the MAS as follows:

$$\kappa_i^{[k+1]}(t) = \sum_{j \in N_i} \alpha_{ij} \left\{ \kappa_j^{[k]}(t) - \kappa_j^{[k]}(t) \right\} \quad d_i = 0 \quad (12)$$

where, because all d_i are set to 0, all κ_i will converge to the average consensus value κ_p^* of the current proportional coefficients, which can be determined as

$$\kappa_p^* = \sum_i \kappa_i / n \quad (13)$$

Secondly, to take voltage restoration of the DC microgrid into consideration, voltage control should be coordinated with current control. The DC voltages need to be adjusted while maintaining the equal proportional current sharing. The voltage control adjustment $\Delta U_{V,i}$ defined in (2) can be calculated as

$$\Delta U_{V,i} = \lambda_i \kappa_p^* = \lambda_i \left(\sum_i \kappa_i / n \right) \quad (14)$$

where λ_i is the voltage control constant, which is set to bring the DC voltage to its new stable value.

Finally, the local reward function can be defined to solve the current sharing and voltage restoration problems as follows:

$$\eta_i = \frac{1}{\Delta U_{V,i} + |\kappa_i - \kappa_p^*|} = \frac{1}{\zeta_i \kappa_p^* + |e_{\kappa,i}|} \quad (15)$$

using (14), where η_i is the local reward defined for the i^{th} agent, and ζ_i is a constant set to decrease the sensitivity of η_i and avoid zero denominator.

Hence the global reward η is accordingly derived as the summation of all the local rewards:

$$\eta = \sum_i \eta_i = \sum_i \frac{1}{\zeta_i \kappa_p^* + |e_{\kappa,i}|} \quad (16)$$

This global reward η can be used to evaluate the performance of a candidate solution; generally, the larger the global reward, the better the current solution.

3.2 DRL based on pinning

Based on the above local and global rewards, the DRL can be recognized as an optimization method to maximize the global reward:

$$\max \left\{ \eta = \sum_i \frac{1}{\lambda_i \kappa_p^* + |e_{\kappa,i}|} \right\} \quad (17)$$

Accordingly, the optimal objective of (17) will be reached when all the local rewards of the DC microgrid converge to a common value η_p^* , which is the well-known solution of (16) as shown in [19, 33]. With regard to (16) and (17), it is clear that the global reward will reach its maximum value when the $|e_{\kappa,i}|$ become zero, therefore, the pinning consensus value η_p^* of the global reward can be predefined by

$$\eta_p^* = \frac{1}{\lambda_i \kappa_p^*} \quad (18)$$

Hence, by using the distributed consensus method illustrated in Section 2.2, DRL with respect to the local reward η_i can be accomplished as follows:

$$\eta_i^{[k+1]}(t) = \sum_{j \in N_i} \alpha_{ij} \left\{ \eta_j^{[k]}(t) - \eta_j^{[k]}(t) \right\} - d_i \left\{ \eta_j^{[k]}(t) - \eta_p^* \right\} \quad (19)$$

When all the local rewards converge to the pinning consensus value preset in (18), the global reward will reach its maximum value of

$$\eta_\infty = \sum_i \eta_{i,\infty} = \sum_i \frac{1}{\lambda_i \kappa_p^* + |e_{\kappa,i,\infty}|} = \sum_i \frac{1}{\lambda_i \kappa_p^*} = n \eta_p^* \quad (20)$$

Based on the synchronization process of the local reward described in (19), DC current and voltage control can be realized, and the control structure of the DC microgrid implemented through this process is shown in Fig. 2. The entire control process of the proposed DRL can be described in the following steps:

Step 1: To take into account the requirements for equal proportional current sharing and voltage restoration in the DC microgrid, the local reward function is defined for each agent as in (15), and the related current

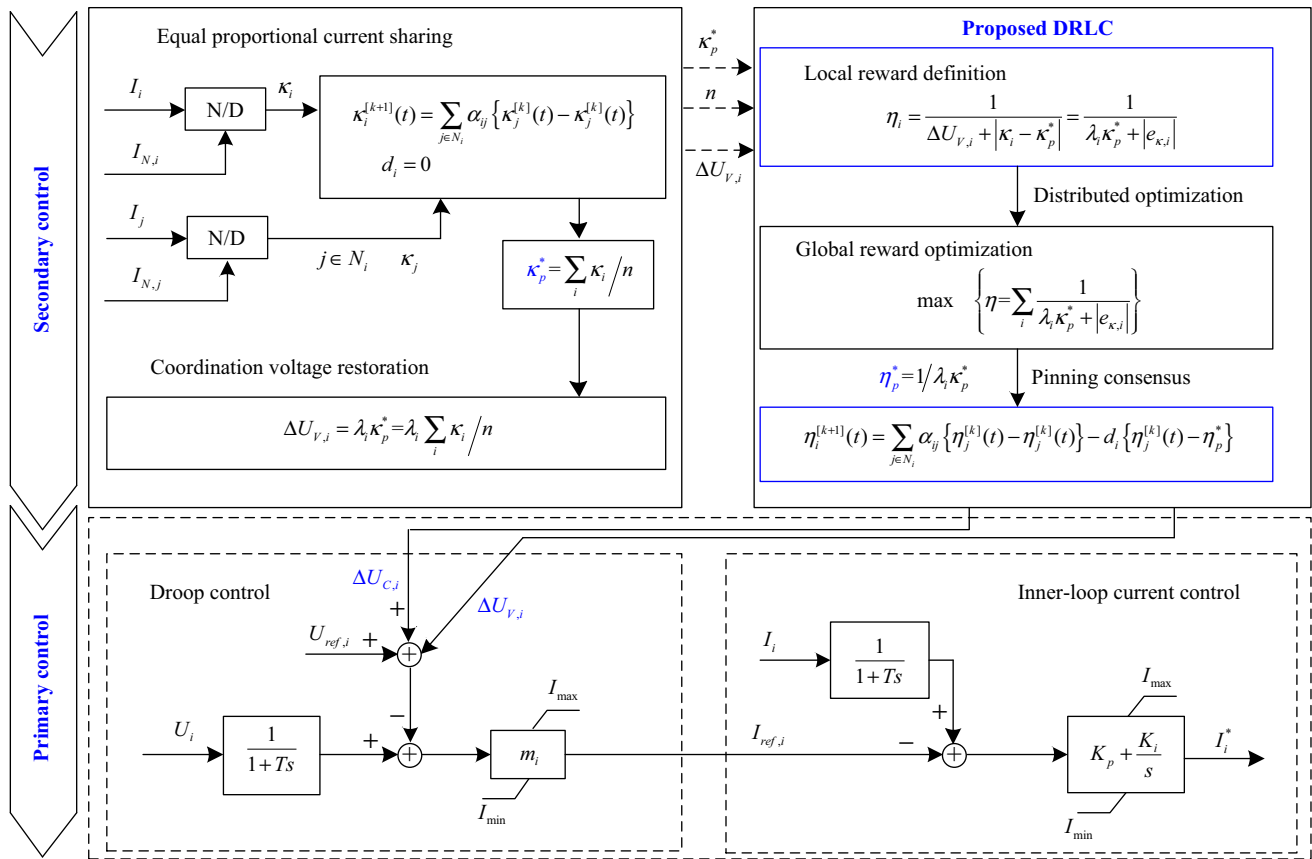


Fig. 2 Fundamental structure of DRL

proportional coefficients κ_i and the voltage control adjustments $\Delta U_{V,i}$ are calculated as in (11)–(14).

Step 2: Accordingly, maximizing the corresponding global reward of the whole DC microgrid is the optimization objective for DRL, as described in (17).

Step 3: The distributed consensus method based on pinning described in Section 2.2 is used to solve this optimization problem, and the pinning consensus value for the DRL is preset according to (18).

Step 4: The proposed DRL is finally implemented to achieve an optimal solution and control the DGs asymptotically, coordinating equal proportional current sharing and voltage restoration of the DC microgrid through the synchronization process of the global reward, as shown in (19) and (20).

4 Simulation studies

To investigate the effectiveness and adaptability of the proposed DRL, a typical DC microgrid containing 5 DGs is simulated in the PSCAD/EMTDC platform, and its configuration is shown in Fig. 3. The algorithms of the proposed DRL, including the calculation of local rewards, the

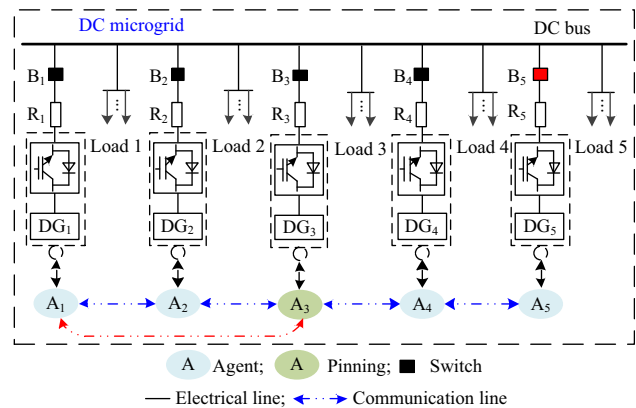


Fig. 3 Simulated DC microgrid

distributed consensus method for information sharing, and the distributed consensus method through pinning control, are compiled in MATLAB, making full use of the mathematical capabilities of this software. Then, the PSCAD model and MATLAB programs are connected together through a Fortran-language-based interface procedure [19]. The communication topology of the simulated DC microgrid is illustrated in Fig. 3.



Using this simulation model three case studies are presented in the following sections.

4.1 Case A: overload scenario

Initially, the DC microgrid works in a stable islanded mode, and all DGs are controlled by droop control. When $t = 1$ s an overload occurs. Consequently, the power balance between supply and demand is lost at that moment,

Table 1 Control parameters of DC microgrid in Case A

DG	m_i	κ_i (A)	$I_{N,i}$ (A)	R (Ω)	U_i (kV)
DG ₁	3	0.95	66.7	0.18	0.591
DG ₂	3	0.75	66.7	0.42	0.601
DG ₃	4	1.14	50	0.15	0.589
DG ₄	3	0.87	66.7	0.24	0.595
DG ₅	5	0.78	50	0.28	0.594

Note: R is the resistance of electrical line

and the proposed DRL is immediately implemented to maintain the DC microgrid.

The rated voltage of the DC microgrid is 0.6 kV and its control parameters are given in Table 1.

Firstly, the current proportional coefficients κ_i are collected by each agent, the average value of which κ_p^* is discovered by using the distributed consensus method described in (12). The synchronization process of the κ_i is shown in Fig. 4a.

Secondly, with the discovered average value κ_p^* , the associated values $\Delta U_{V,i}$ can be calculated by (14), and with $\lambda_i = 0.204$ the pinning consensus value of the local reward can be preset as $\eta_p^* = 5.576$ according to (18). Thus, the local reward η_i of the DRL defined in (15) is estimated by its corresponding agent, and the synchronization seeking process is shown in Fig. 4b.

Finally, the proposed DRL which coordinates the voltage restoration and equal proportional current sharing is implemented, and the current and voltage control per agent

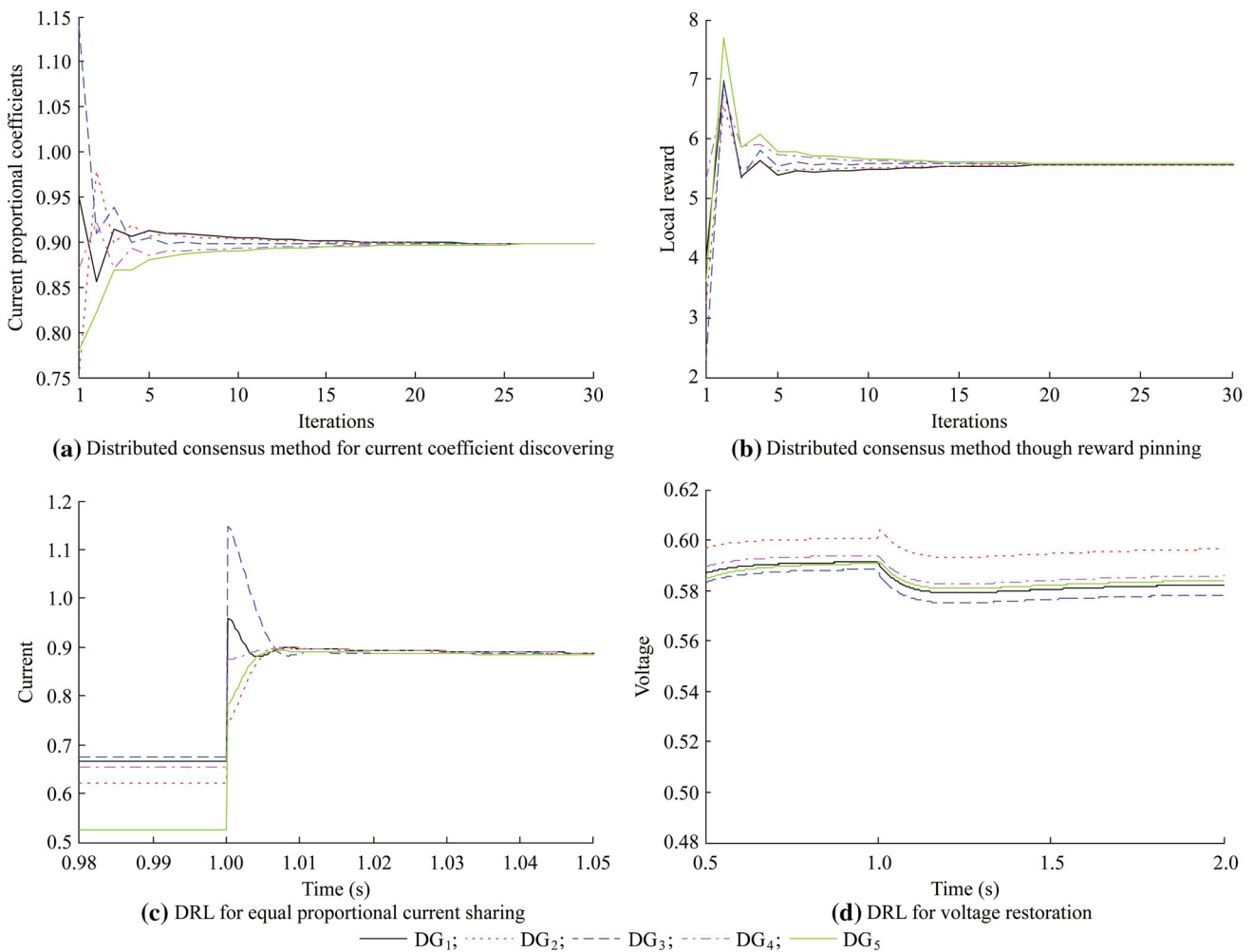


Fig. 4 Control performances of the proposed DRL in Case A

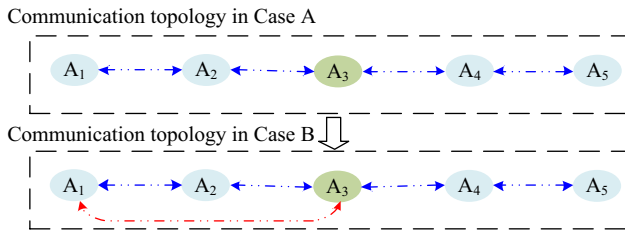


Fig. 5 Communication topology changes in Case B

Table 2 Control parameters of DC microgrid in Case B

DG	m_i	κ_i (A)	$I_{N,i}$ (A)	R (Ω)	U_i (kV)
DG ₁	3	0.80	66.7	0.18	0.591
DG ₂	3	0.68	66.7	0.42	0.600
DG ₃	4	0.89	50	0.15	0.587
DG ₄	3	0.75	66.7	0.24	0.593
DG ₅	5	0.65	50	0.28	0.590

are shown in Fig. 4c and d, where the consensus convergence process can be seen clearly.

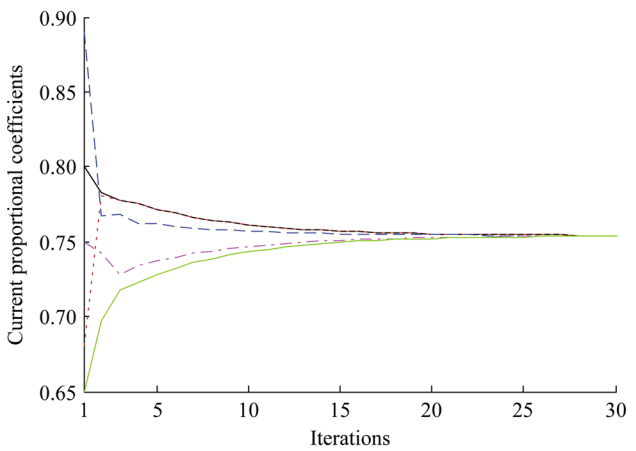
It can be observed in Fig. 4a and b that the distributed consensus method presented in Section 2.2 realizes two

functions in this case: 1) discovering global information based on average consensus and obtaining the averaged current proportional coefficient, as shown in Fig. 4a, and 2) implementing the distributed consensus method based on local reward pinning to coordinate equal proportional current sharing and voltage restoration, as shown in Fig. 4b. Thus, in the DRL for equal proportional current sharing and voltage restoration, illustrated in Fig. 4c and d respectively, the current proportional coefficients of all agents converge to equal consensus values, and the voltages reach a corresponding new state.

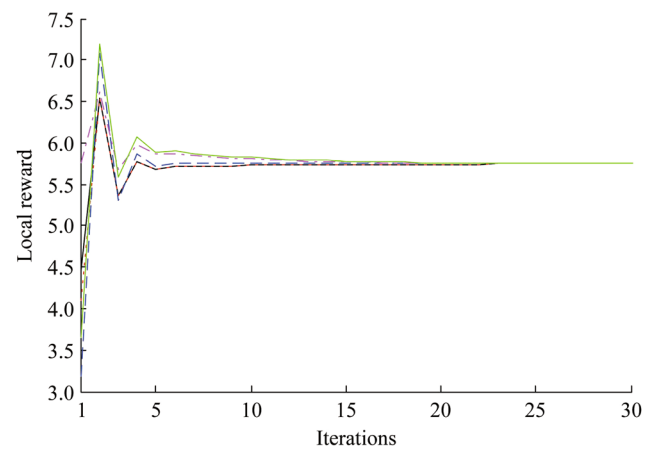
4.2 Case B: overload and communication line switches on

In this case, the overload accident occurs in the DC microgrid at $t = 1$ s, and at the same time a new communication link between agent 1 (A_1) and agent 3 (A_3) switches on, as illustrated in Fig. 5. The control parameters of the proposed DRL in Case B are shown in Table 2.

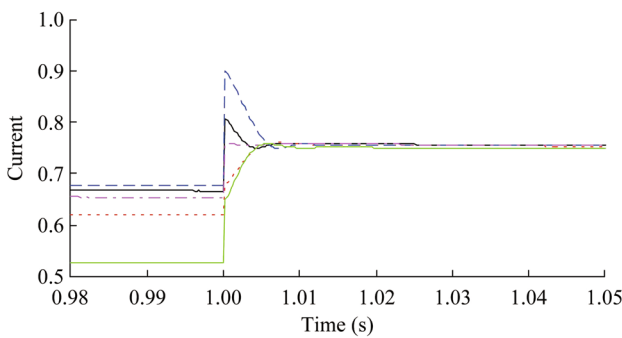
In contrast to Case A, to address the change in communication topology, the connectivity coefficients α_{ij} for the newly connected agents update as described in (7).



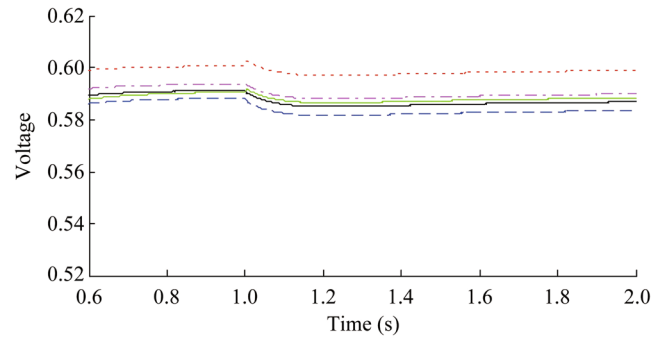
(a) Distributed consensus method for current coefficient discovering



(b) Distributed consensus method though reward pinning



(c) DRL for equal proportional current sharing



(d) DRL for voltage restoration

— DG₁; ···· DG₂; - - - DG₃; ···· DG₄; — DG₅

Fig. 6 Control performances of the proposed DRL in Case B



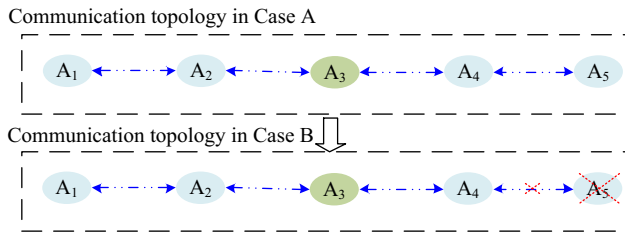


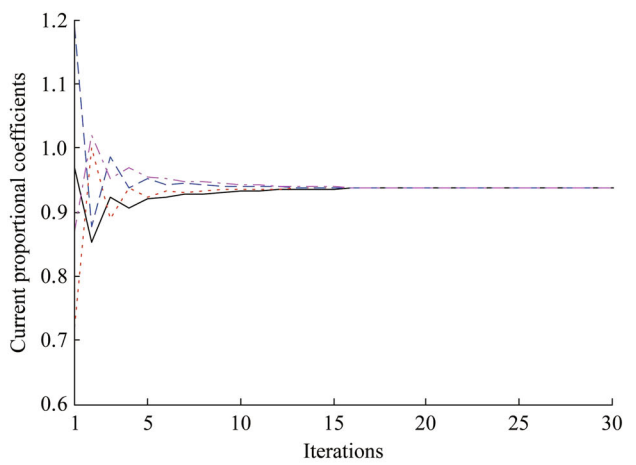
Fig. 7 Communication topology changes in Case C

Table 3 Parameters of DC microgrid in Case C

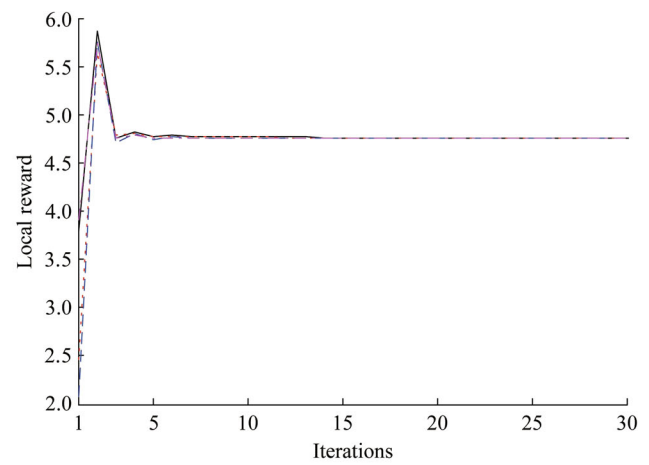
DG	m_i	κ_i (A)	$I_{N,i}$ (A)	R (Ω)	U_i (kV)
DG ₁	3	0.97	66.7	0.18	0.603
DG ₂	3	0.72	66.7	0.42	0.612
DG ₃	4	1.19	50	0.15	0.601
DG ₄	3	0.87	66.7	0.24	0.605

Then, by using the proposed distributed consensus method, the average value of the current proportional coefficients is discovered to be $\kappa_p^* = 0.75$. Additionally, with $\lambda_i = 0.232$, the pinning consensus value of the local reward can be preset as $\eta_p^* = 5.747$. The synchronization seeking process of the current proportional coefficients and the local rewards are shown in Fig. 6a and b respectively.

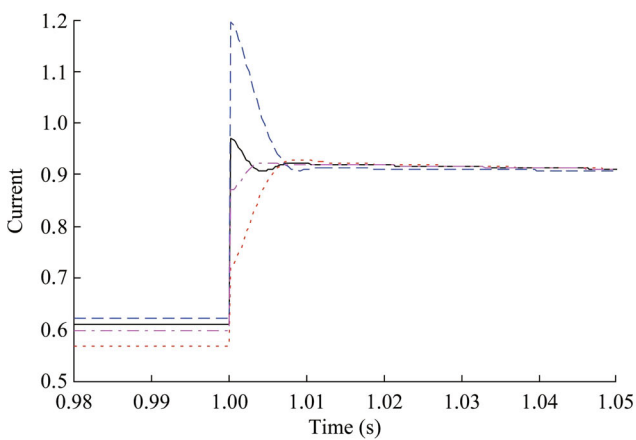
Through (17)–(20) the global reward is maximized when the pinning-based distributed consensus is reached. It can be seen in Fig. 6c that the current proportional coefficients of all DGs converge asymptotically to a new common value, and the synchronization seeking process is different from that of Case A because of the additional communication link. Similarly, DRL-based voltage restoration is also adjusted to adapt for changed communication topology, as can be seen in Fig. 6d.



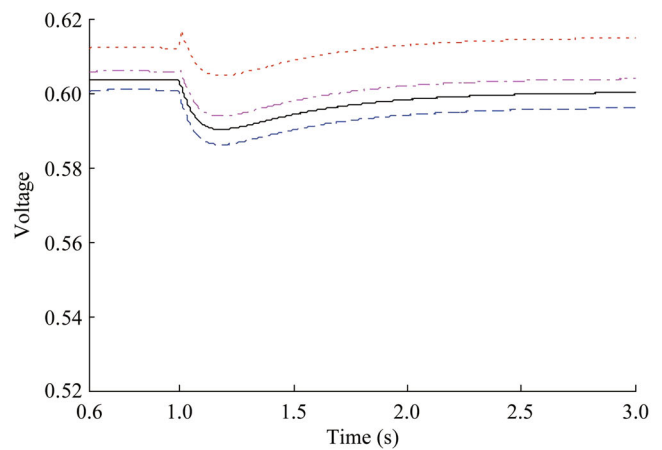
(a) Distributed consensus method for current coefficient discovering



(b) Distributed consensus method though reward pinning



(c) DRL for equal proportional current sharing



(d) DRL for voltage restoration

— DG₁; ···· DG₂; - - - DG₃; - · - · DG₄

Fig. 8 Control performances of the proposed DRL in Case C

4.3 Case C: overload and agent unplugs

In Case C, the agent 5 marked as A_5 unplugs from the DC microgrid and its corresponding communication link switches off accordingly at $t = 1$ s; as a result, the communication topology of the simulated DC microgrid changes, as shown in Fig. 7.

The control parameters are illustrated in Table 3 and the proposed control response is implemented as follows. Firstly, to adapt for the unplugging of A_5 , agent identities are updated according to the method described in (8). Only the neighboring agents of the faulted A_5 need to be updated. Secondly, after the adaptive updating, both the discovery of current proportional coefficients by distributed consensus and the pinning-based distributed consensus of the local reward can be implemented, as in Cases A and B.

The average value of the current proportional coefficients in Case C is discovered to be $\kappa_p^* = 0.916$. Additionally, with $\lambda_i = 0.236$, the pinning consensus value of the local reward can be preset as $\eta_p^* = 4.761$. The synchronization seeking process of the current proportional coefficients and the local rewards are shown in Fig. 8a and b.

In Fig. 8c and d it can be seen that the DC currents and voltages of all the DGs asymptotically converge to new common values through the proposed DRL, so the equal proportional current sharing and the voltage restoration problems are successfully coordinated, and the proposed DRL can be adaptively implemented when an agent is unplugged.

5 Conclusion

In this study, a novel DRL strategy has been proposed and investigated for an islanded DC microgrid. The implementation of this DRL strategy is achieved by integrating two methods, which are the distributed consensus method through pinning and the RL method.

The proposed distributed consensus method can be used to discover global information and implement pinning synchronization, and it can also meet the requirement to adapt to changes in the communication network, such as communication line switches or agent plug-and-play operations. The proposed DRL based on local and global rewards can be utilized to maximize the global reward and achieve an optimal solution for a DC microgrid. Hence, the proposed strategy can coordinate the equal proportional current sharing and the voltage restoration of an autonomous DC microgrid.

The effectiveness and advantages of this approach are demonstrated by simulating three representative cases of an overload condition, including addition of a new communication link and unplugging of a DG agent. The DRL method worked quickly and effectively in each case.

Acknowledgements This work was supported by National Key Research and Development Program of China (No. 2016YFB0900105).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- [1] Ding G, Gao F, Zhang S et al (2014) Control of hybrid AC/DC microgrid under islanding operational conditions. *J Mod Power Syst Clean Energy* 2(3):223–232. doi:10.1007/s40565-014-0065-z
- [2] Millar B, Jiang D, Me H (2015) Constrained coordinated distributed control of smart grid with asynchronous information exchange. *J Mod Power Syst Clean Energy* 3(4):512–525. doi:10.1007/s40565-015-0168-1
- [3] Elsayed A, Mohamed A, Mohammed O (2015) DC microgrids and distribution systems: an overview. *Electr Power Syst Res* 199:407–417
- [4] Gu W, Liu W, Wu Z (2013) Cooperative control to enhance the frequency stability of islanded microgrid with DFIG-SMES. *Energies* 6(8):3951–3971
- [5] Oureilidis KO, Bakirtzis EA, Demoulias CS (2016) Frequency-based control of islanded microgrid with renewable energy sources and energy storage. *J Mod Power Syst Clean Energy* 4(1):54–62. doi:10.1007/s40565-015-0178-z
- [6] Gu W, Wu Z, Bo R et al (2013) Modeling, planning and optimal energy management of combined cooling, heating and power microgrid: a review. *Int J Electr Power Energy Syst* 54(1):26–37
- [7] Shuai Z, Mo S, Wang J et al (2016) Droop control method for load share and voltage regulation in high-voltage microgrids. *J Mod Power Syst Clean Energy* 4(1):76–86. doi:10.1007/s40565-015-0176-1
- [8] Guo F, Wen C, Mao J et al (2015) Distributed secondary voltage and frequency restoration control of droop-controlled inverter-based microgrids. *IEEE Trans Ind Electron* 62(7):4355–4364
- [9] Khorsandi A, Ashourloo M, Mokhtari H (2014) A decentralized control method for a low-voltage dc microgrid. *IEEE Trans Energy Convers* 29(4):793–801
- [10] Ahmadi R, Ferdowsi M (2014) Improving the performance of a line regulating converter in a converter-dominated DC microgrid system. *IEEE Trans Smart Grid* 5(5):2553–2563
- [11] Guerrero JM, Vasquez JC, Matas J et al (2011) Hierarchical control of droop-controlled ac and dc microgrids—a general approach toward standardization. *IEEE Trans Ind Electron* 58(1):158–172
- [12] Papadimitriou C, Zountouridou E, Hatziaargyriou N (2015) Review of hierarchical control in dc microgrids. *Electr Power Syst Res* 122(2015):159–167



- [13] Gu W, Liu W, Shen C et al (2013) Multi-stage underfrequency load shedding for islanded microgrid with equivalent inertia constant analysis. *Int J Electr Power Energy Syst* 46(1):36–39
- [14] Lu X, Guerrero JM, Sun K et al (2014) An improved droop control method for dc microgrids based on low bandwidth communication with dc bus voltage restoration and enhanced current sharing accuracy. *IEEE Trans Power Electron* 29(4):1800–1812
- [15] Tan KT, Peng XY, So PL et al (2012) Centralized control for parallel operation of distributed generation inverters in microgrids. *IEEE Trans Smart Grid* 3(4):1977–1987
- [16] Tsikalakis AG, Hatziargyriou ND (2008) Centralized control for optimizing microgrids operation. *IEEE Trans Energy Convers* 23(1):241–248
- [17] Anand S, Fernandes BG, Guerrero JM (2013) Distributed control to ensure proportional load sharing and improve voltage regulation in low voltage dc microgrids. *IEEE Trans Power Electron* 28(4):1900–1913
- [18] Gu W, Liu W, Zhu J et al (2014) Adaptive decentralized underfrequency load shedding for islanded smart distribution networks. *IEEE Trans Sustain Energy* 5(3):886–895
- [19] Liu W, Gu W, Xu Y et al (2015) Improved average consensus algorithm based distributed cost optimization for loading shedding of autonomous microgrids. *Int J Electr Power Energy Syst* 73:89–96
- [20] Liu W, Gu W, Sheng W et al (2014) Decentralized multi-agent system-based cooperative frequency control for autonomous microgrids with communication constraints. *IEEE Trans Sustain Energy* 5(2):446–456
- [21] Bidram A, Davoudi A, Lewis FL et al (2013) Distributed cooperative secondary control of microgrids using feedback linearization. *IEEE Trans Power Syst* 28(3):3462–3470
- [22] Liu W, Gu W, Sheng W et al (2016) Pinning-based distributed cooperative control for autonomous microgrids under uncertainty communication topologies. *IEEE Trans Power Systems* 31(2):1620–1629
- [23] Nasirian V, Davoudi A, Lewis FL et al (2014) Distributed adaptive droop control for DC distribution systems. *IEEE Trans Energy Convers* 29(4):944–956
- [24] Wang P, Lu X, Yang X et al (2016) An improved distributed secondary control method for DC microgrids with enhanced dynamic current sharing performance. *IEEE Trans Power Electron* 31(9):6658–6673
- [25] Nasirian V, Moayedi S, Davoudi A et al (2015) Distributed cooperative control of dc microgrids. *IEEE Trans Power Electron* 30(4):2288–2303
- [26] Xu Y, Zhang W, Liu W et al (2012) Multiagent-based reinforcement learning for optimal reactive power dispatch. *IEEE Trans Syst Man Cybern C Appl Rev* 42(6):1742–1751
- [27] Fernandez F, Parker LE (2001) Learning in large cooperative multirobot systems. *Int J Robot Autom* 16(4):217–226
- [28] Lauer M, Riedmiller MA (2000) An algorithm for distributed reinforcement learning in cooperative multi-agent systems. *Proc Int Conf Mach Learn* 15(1):535–542
- [29] Plamondon P, Chaib-draa B, Benaskeur A (2007) A Q-decomposition and bounded RTDP approach to resource allocation. In: *Proceedings of the 6th international conference on autonomous agents and multiagent systems*, Honolulu, Hawaii, USA, 14–18 May 2007, pp 1212–1219
- [30] Russell SJ, Zimdars A (2003) Q-decomposition for reinforcement learning agents. In: *Proceedings of the international conference on machine learning*, Washington DC, USA, 21–24 August 2003, pp 656–663
- [31] Vlachogiannis JG, Hatziargyriou ND (2004) Reinforcement learning for reactive power control. *IEEE Trans Power Syst* 9(3):1317–1325
- [32] Martin HJA, de Lope J (2007) A distributed reinforcement learning control architecture for multi-link robots—experimental validation. In: *Proceedings of the 4th international conference on informatics in control, automation and robotics*, Angers, France, 9–12 May 2007, pp 192–197
- [33] Zhang Z, Chow MY (2012) Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid. *IEEE Trans Power Syst* 27(4):1761–1768

Zifa LIU received the B.Sc. and M.Sc. degree both in Electrical Engineering from Northeast Dianli University, Jilin, China, in 1995 and 2000. He received the Ph.D. degree in Electrical Engineering from Tianjin University in 2005. He has been working as Associate Professor in North China Electric Power University since 2008. His research interests include power system analysis, distribution and micro grid planning.

Ya LUO received the B.Sc. in Automation Engineering from North China Electric Power University in 2013 and is presently studying for the M.Sc. degree in North China Electric Power University. Her research interests include power system analysis and planning.

Ranqun ZHUO Received the B.Sc. in Automation Engineering from North China Electric Power University in 2015 and is presently studying for the M.Sc. degree in North China Electric Power University. His research interests include microgrid planning and renewable energy analysis.

Xianlin JIN is a senior engineer. He is now working on production management of wind farm in Guohua Energy Investment Co., Ltd.