

DOI: 10.19783/j.cnki.pspc.220104

基于协同强化学习的微电网分布式两级电压优化控制

汪超¹, 赵婵娟², 程志友^{1,2}, 宋俊材²

(1. 安徽大学电子信息工程学院, 安徽 合肥 230601; 2. 安徽大学互联网学院, 安徽 合肥 230039)

摘要: 由于微电网中分布式电源组成复杂, 运行模式多样, 孤岛微电网的电压恢复控制面临着不确定性干扰的影响。为此, 针对不确定性干扰下微电网的二级电压恢复控制问题, 提出了一种基于协同强化学习的微电网分布式两级电压优化控制方法, 实现孤岛模式下微电网的电压调节控制。首先构建孤岛微电网分布式一致性协同电压控制算法, 并建立李雅普诺夫函数稳定性判定方法。其次根据控制器性能与控制器增益参数的关系, 求解孤岛微电网电压控制器增益上界, 并根据控制器增益参数上界限制强化学习智能体动作集。随后, 采用强化学习算法优化二级控制器增益参数, 给出相应的强化学习智能体状态集、协同全局奖励函数。最后在 Matlab/Simulink 上通过仿真实验验证了所提出的控制方法的有效性和适应性。

关键词: 微电网; 协同强化学习; 分布式两级优化控制; 稳定性判定; 全局奖励函数

Distributed secondary voltage optimization control for a microgrid based on cooperative reinforcement learning

WANG Chao¹, ZHAO Chanjuan², CHENG Zhiyou^{1,2}, SONG Juncai²

(1. School of Electronic and Information Engineering, Anhui University, Hefei 230601, China;

2. School of Internet, Anhui University, Hefei 230039, China)

Abstract: Because of the complex composition of distributed generators and various operational modes in the microgrid, the voltage restoration control of an islanded microgrid will be affected by uncertain disturbances. In this paper, a distributed secondary microgrid voltage optimization control method based on cooperative reinforcement learning algorithm applied to the secondary voltage restore control problem of the microgrid under uncertain disturbances is proposed. This can realize the voltage regulation control of an islanded microgrid. First, a distributed consensus cooperative voltage control algorithm of the islanded microgrid is constructed, and a stability judgement method based on the Lyapunov function is established. Second, from the relationship between controller performance and gain parameters, the upper bound of the voltage controller gain of the islanded microgrid is derived. This can be used to constrain the action space of the reinforcement learning agent. Third, the reinforcement learning algorithm is applied to optimize the secondary voltage controller gain parameters, and the corresponding reinforcement learning state space and cooperative global reward function are presented. Finally, the effectiveness and adaptability of the proposed control are verified by simulation experiments on Matlab/Simulink.

This work is supported by the Natural Science Foundation of Anhui Province (No. 2108085QE237).

Key words: microgrid; cooperative reinforcement learning; distributed secondary optimization control; stability judgement; global reward function

0 引言

微电网是由分布式电源(distributed generator,

DG)、储能装置、本地负载以及监控通信装置等组成的小型分布式发电系统^[1], 具有提高能源利用率, 有效管理各分布式电源的优点^[2-4]。通常微电网可以和主电网连接并网运行, 当主电网发生故障或检修时, 微电网与主电网分离运行于孤岛模式, 此时, 微电网需要协调各分布式电源维持自身频率、电压稳定^[5-6]。传统的下垂控制受线路阻抗的影响, 难以

基金项目: 安徽省自然科学基金项目资助(2108085QE237); 国家自然科学基金项目资助(51877060); 安徽省高等学校自然科学基金项目资助(KJ2021A0018)

保证微电网输出电压、频率稳定在参考值^[7]。针对这一问题,文献[8]提出了一种基于线路阻抗辨识的改进下垂控制策略,依据 DG 的等效线路阻抗辨识结果,在传统无功下垂控制中增加无功偏差补偿,实现分布式电源无功功率的合理分配。文献[9]提出了一种基于自适应虚拟阻抗的改进下垂控制方法,根据 DG 输出的无功信息构造虚拟阻抗,提高了分布式电源的无功功率分配精度。上述方法虽然在一定程度上提高了分布式电源功率分配精度,但在引入虚拟阻抗的同时会引起微电网系统的电压跌落,研究微电网的输出电能质量调节控制具有重要意义。

为实现孤岛微电网的输出电压、频率调节,进一步提高微电网的稳定性,近年来,很多学者对微电网的分层控制展开了相关研究。文献[10]提出了基于事件触发通信机制的孤岛微电网分布式两级控制方案,在保证微电网输出电压、频率调节性能的同时,能有效减少邻节点间的通信次数。文献[11]提出了一种基于对等稀疏网络的微电网分布式两级控制,解决了负荷不平衡及线路阻抗未知情况下的微电网电压补偿问题。文献[12]提出了一种通信延时情况下孤岛微电网的分布式两级电压控制方法。这些方法均采用分布式通信方式,不仅能确保微电网输出电压/频率的调节性能,而且可以有效避免微电网两级控制过程中的单点故障问题,文献[13]提出了一种基于分布式一致性积分控制的孤岛微电网两级电压控制方法,采用传统的积分控制方式实现孤岛微电网的电压恢复,并通过事件触发的方式减少孤岛微电网的通信负担。但是,上述方法在孤岛微电网控制过程中都很少涉及到微电网的输出控制性能优化问题,研究微电网的输出电压优化控制,进一步提高微电网的输出控制性能,具有重要意义。

随着人工智能技术的兴起与发展,强化学习(reinforcement learning, RL)逐渐在各种领域中发挥重要作用。目前强化学习方法在电力系统中已经应用于负荷预测^[14]、能源管理^[15]和经济调度^[16]等方面。文献[17]采用强化学习解决多场景下综合能源系统优化调度问题。文献[18]提出了基于强化学习的微电网能源管理。文献[19]采用强化学习解决电力系统中的经济调度问题。在系统控制性能优化方面,文献[20]考虑旋翼无人机模型的不确定性,通过引入强化学习算法有效消除了不确定性干扰和未建模动态特性对四旋翼无人机控制系统的影响,提高了旋翼无人机对外界干扰的鲁棒性。文献[21]针对复杂环境下无人船的精确轨迹航行问题,提出了基于强化学习方法的无人船高性能轨迹跟踪控制,

实现无人船在复杂环境下稳定轨迹跟踪的同时确保动态性能最优。以上控制方法都是基于复杂应用场景环境,并且易受到多种不确定性外界干扰影响。强化学习优化控制方法通常不需要系统具有精确的数学模型,通过与环境不断交互,根据环境反馈的奖励改进决策能力,因而能够很好地解决复杂环境下的系统控制优化问题^[22]。

考虑微电网面对的环境复杂多变,易受到多种不确定性干扰的影响,因此本文提出了一种基于协同强化学习的分布式两级电压优化控制策略,实现孤岛模式下微电网的电压恢复控制。该策略在原有的分布式一致性二级控制中引入强化学习方法,采用强化学习优化控制器增益参数,克服了传统二级控制中控制器增益参数单一,系统性能受参数限制的影响,可以有效地提高微电网在面对不确定性干扰时输出电压的动态控制性能。首先,建立孤岛微电网分布式一致性协同电压控制算法,并建立李雅普诺夫函数证明系统的稳定性。然后,根据控制器增益参数与控制器性能之间的关系求解出孤岛微电网电压控制器增益上界,在此基础上通过求解出的控制器增益上界限制强化学习动作集边界,并给出智能体的动作集和状态集。随后提出了一种协同全局奖励函数促进智能体收敛获得最优动作序列。最后,对孤岛微电网在负荷投切以及即插即用情况下的输出性能进行仿真分析,验证本文所提出方法的有效性。

1 微电网分布式两级协同优化控制系统结构设计

为了解决孤岛微电网分布式电源输出电压和频率偏差补偿问题,本文将分布式两级协同控制方法应用于微电网内的变流器控制,提出一种基于分布式两级协同控制策略的孤岛微电网网络化控制结构,如图 1 所示。

由图 1 可知,孤岛微电网网络化控制结构由初级控制和二级控制两部分组成。其中初级控制是由传统的功率计算、下垂控制、电压外环控制器和电流内环控制器组成,实现微电网的输出电压、频率的稳定控制。为了弥补初级下垂控制引起的电压、频率偏差,在二级控制部分利用通信网络将各个分布式电源输出的电压幅值 E 和频率 ω 在相邻节点间传输,通过分布式协同控制将各分布式电源的输出电压和频率稳定在参考值。

孤岛微电网输出电压为非全局变量,容易受到外界干扰的影响。考虑强化学习方法对多种不确定

性因素具有自适应能力, 特别适用于复杂系统的优化控制。因此, 本文考虑将强化学习方法与传统控制结构相结合, 采用协同强化学习算法优化孤岛微电网分布式二级电压控制器增益参数, 提出一种基于协同强化学习的孤岛微电网分布式两级优化控制算法, 以实现不确定性干扰下孤岛微电网输出电压的自适应快速稳定调节。

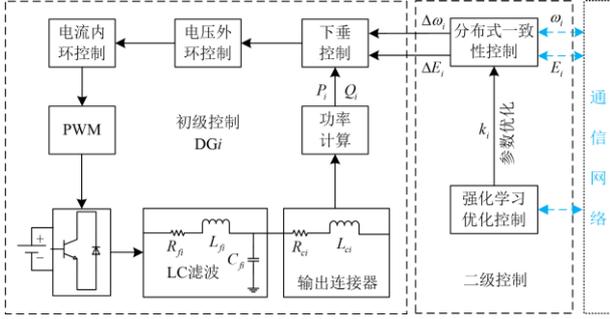


图 1 基于分布式强化学习的两级控制的孤岛微电网控制结构图

Fig. 1 Control structure diagram of islanded microgrid based on distributed reinforcement learning with secondary control

2 孤岛微电网分布式两级控制策略研究

2.1 下垂控制原理

考虑微电网各分布式电源变流器输出线路阻抗通常呈现感性特性, 因此对分布式电源输出信号的有功功率与频率、无功功率与电压幅值进行解耦, 得到第 i 个分布式电源的下垂特性如式(1)^[13]。

$$\begin{cases} \omega_i = \omega^* - k_{pi} P_i \\ E_i = E^* - k_{qi} Q_i \end{cases} \quad (1)$$

式中: ω_i 和 E_i 分别为第 i 个分布式电源的输出频率和电压; ω^* 和 E^* 分别为第 i 个分布式电源的参考频率和参考电压; k_{pi} 和 k_{qi} 分别为第 i 个分布式电源的频率、电压下垂控制系数; P_i 和 Q_i 分别为第 i 个分布式电源输出的有功和无功功率。

2.2 孤岛微电网分布式二级协同电压控制算法

2.2.1 分布式一致性协同电压控制器

考虑微电网初级控制具有较快的动态特性, 在微电网的二级控制建模时^[13], 可忽略初级控制的影响。因此, 建立孤岛微电网电压控制的状态空间模型如式(2)所示。

$$\begin{cases} \dot{E}_i(t) = u_i(t) \\ y_i(t) = E_i(t) \end{cases} \quad (2)$$

式中: $E_i(t)$ 表示第 $i(i=1, L, n)$ 个 DG 输出的电压幅值, n 表示智能体的个数; $u_i(t)$ 表示第 i 个 DG 的

二级控制器输出控制变量; $y_i(t)$ 表示第 i 个 DG 的输出变量。

微电网二级控制目标是将各分布式电源的输出电压调节到参考值, 其本质可以看成是一个同步跟踪控制问题。分布式一致性控制算法仅需要邻节点之间进行信息交互, 并且当邻节点间通信网络拓扑满足生成树条件时, 所有节点的状态信息可在有限时间内收敛到同一值。因此, 本文采用分布式一致性协同控制策略来实现对孤岛微电网的电压调节控制。孤岛微电网分布式一致性协同跟踪控制器定义如式(3)。

$$u_i(t) = k_i \left[\sum_{j \in N_i} (E_j(t) - E_i(t)) + g_i (E_{ref} - E_i(t)) \right] \quad (3)$$

式中: k_i 表示第 i 个 DG 的控制器增益参数; N_i 为第 i 个智能体的相邻智能体的集合; E_{ref} 表示电压幅值的参考值; g_i 为定位增益, 当 DG $_i$ 与参考点直接相连时, $g_i = 1$, 否则 $g_i = 0$ 。

2.2.2 稳定性分析与性能研究

为了验证上述分布式一致性协同控制方法的有效性, 本节分析了该分布式控制方法的收敛性与动态性能。定义 $\eta_i(t) = E_i(t) - E_{ref}$, 结合式(2)、式(3)可以推出式(4)。

$$\begin{aligned} \dot{\eta}(t) &= \dot{E}(t) - \dot{E}_{ref} = \\ & k_i \left[\sum_{j \in N_i} (E_j(t) - E_i(t)) + g_i (E_{ref} - E_i(t)) \right] = \\ & k_i \left[\sum_{j \in N_i} (\eta_j(t) - \eta_i(t)) - g_i \eta_i(t) \right] \end{aligned} \quad (4)$$

定义 $\eta(t) = [\eta_1(t) \ \eta_2(t) \ \dots \ \eta_n(t)]^T$, $K = \text{diag}(k_i) \in R^{n \times n}$, $g = \text{diag}(g_i) \in R^{n \times n}$, 则

$$\dot{\eta}(t) = -K(L + g)\eta(t) \quad (5)$$

式中: $L = D_G - A_G$ 为孤岛微电网通信网络拓扑图的拉普拉斯矩阵; $D_G = \text{diag} \left[\sum_{j \in N_i} a_{ij} \right] \in R^{n \times n}$ 为入度矩阵; $A_G = [a_{ij}] \in R^{n \times n}$ 为邻接矩阵, a_{ij} 为 A_G 第 i 行第 j 列的元素, 若节点 i 与节点 j 之间有信息交互, 则 $a_{ij} = 1$, 否则 $a_{ij} = 0$ 。

在上述系统模型建立的基础上, 以下给出系统模型的稳定性分析。设 X 为正定对称矩阵, 选取 Lyapunov 方程如下:

$$V = \frac{1}{2} \eta^T X \eta \quad (6)$$

$V \geq 0$ 成立, 对式(6)进行微分得到

$$\begin{aligned} \mathbf{V} &= \frac{1}{2} (\boldsymbol{\eta}^T \mathbf{X} \boldsymbol{\eta} + \boldsymbol{\eta}^T \mathbf{X} \boldsymbol{\eta}) = \\ &= -\frac{1}{2} \boldsymbol{\eta}^T \{ [\mathbf{K}(\mathbf{L} + \mathbf{g})]^T \mathbf{X} + \mathbf{X} \mathbf{K}(\mathbf{L} + \mathbf{g}) \} \boldsymbol{\eta} \end{aligned} \quad (7)$$

定义 $\mathbf{H} = \mathbf{L} + \mathbf{g}$, 则当孤岛微电网通信网络拓扑图存在生成树, 且满足至少有一个分布式电源与参考点相连时, 可以得到矩阵 $\mathbf{L} + \mathbf{g}$ 为对称正定矩阵^[18], 即 $(\mathbf{L} + \mathbf{g})^T = (\mathbf{L} + \mathbf{g})$ 。此时有

$$\mathbf{V} = -\frac{1}{2} \boldsymbol{\eta}^T [(\mathbf{K}\mathbf{H})^T \mathbf{X} + \mathbf{X} \mathbf{K}\mathbf{H}] \boldsymbol{\eta} \quad (8)$$

式中: \mathbf{K} 为对角正定矩阵; \mathbf{H} 为正定矩阵。因此, $(\mathbf{K}\mathbf{H})^T \mathbf{X} + \mathbf{X} \mathbf{K}\mathbf{H}$ 正定。可以推出式(9)成立。

$$\mathbf{V} \leq 0 \quad (9)$$

因此, 可以得出微电网分布式一致性协同控制器跟踪误差能渐近收敛到 0, 系统是稳定的。为了进一步分析该分布式一致性控制系统的性能, 定义如下。

$$\mathbf{O} = (\mathbf{K}\mathbf{H})^T \mathbf{X} + \mathbf{X} \mathbf{K}\mathbf{H} \quad (10)$$

则式(8)满足如下不等式:

$$\mathbf{V} \leq -\frac{1}{2} \sigma_{\min}(\mathbf{O}) \|\boldsymbol{\eta}\|^2 \quad (11)$$

式中, $\sigma_{\min}(\mathbf{O})$ 为矩阵 \mathbf{O} 的最小奇异值。又由于

$$\mathbf{V} = \frac{1}{2} \boldsymbol{\eta}^T \mathbf{X} \boldsymbol{\eta} \leq \frac{1}{2} \sigma_{\max}(\mathbf{X}) \|\boldsymbol{\eta}\|^2 \quad (12)$$

可以推导出

$$\|\boldsymbol{\eta}\|^2 \geq \frac{2\mathbf{V}}{\sigma_{\max}(\mathbf{X})} \quad (13)$$

结合式(11)、式(13)可得

$$\mathbf{V} \leq -\frac{1}{2} \sigma_{\min}(\mathbf{O}) \|\boldsymbol{\eta}\|^2 \leq \frac{-\sigma_{\min}(\mathbf{O})}{\sigma_{\max}(\mathbf{X})} \mathbf{V} \quad (14)$$

定义 $\alpha = \frac{\sigma_{\min}(\mathbf{O})}{\sigma_{\max}(\mathbf{X})}$, 则有

$$\mathbf{V} \leq e^{-\alpha t} \mathbf{V}(t_0) \quad (15)$$

由式(6)和式(15)可得

$$\frac{1}{2} \sigma_{\min}(\mathbf{X}) \|\boldsymbol{\eta}(t)\|^2 \leq \frac{1}{2} \sigma_{\max}(\mathbf{X}) e^{-\alpha t} \|\boldsymbol{\eta}(t_0)\|^2 \quad (16)$$

$$\|\boldsymbol{\eta}\| \leq \sqrt{\frac{\sigma_{\max}(\mathbf{X})}{\sigma_{\min}(\mathbf{X})}} e^{-\frac{\alpha}{2} t} \|\boldsymbol{\eta}(t_0)\| \quad (17)$$

由式(17)可知, 孤岛微电网分布式协同控制器跟踪误差以时间常数 $-\frac{\alpha}{2}$ 渐近收敛到 0。由于

$\alpha = \frac{\sigma_{\min}(\mathbf{O})}{\sigma_{\max}(\mathbf{X})}$, 其中 $\mathbf{O} = (\mathbf{K}\mathbf{H})^T \mathbf{X} + \mathbf{X} \mathbf{K}\mathbf{H}$, 孤岛微

电网分布式一致性协同控制器的跟踪收敛速度可以由控制器增益 \mathbf{K} 来调整。并且可以得出: 随着 \mathbf{K} 增大, $\sigma_{\min}(\mathbf{O})$ 逐渐增大, $-\frac{\alpha}{2}$ 逐渐减小, 此时孤岛微

电网跟踪收敛速度加快, 但当 \mathbf{K} 过大时, 系统动态特性变差, 因此本文通过自适应调节控制器增益 \mathbf{K} 进一步改善孤岛微电网分布式控制器控制性能。

3 协同强化学习优化算法

由于孤岛微电网输出电压容易受到多种不确定因素的干扰, 导致输出电压的动态性能不理想。本文提出采用协同强化学习算法实现孤岛模式下微电网的电压恢复控制。

3.1 协同强化学习模型

本文采用图 2 所示的协同控制架构, 其中 S_1 、 S_2 、 S_3 和 S_4 分别表示相应智能体获得的状态, 虚线表示通信链路。各分布式电源智能体通过分布式通信链路相互连接并获取邻节点信息。节点 i 的电压跟踪偏差为

$$\Delta U_i^t = \frac{1}{N} \sum_{j \in N_i} (E_j(t) - E_i(t)) + g_i (E_{\text{ref}} - E_i(t)) \quad (18)$$

式中: N 为智能体的邻节点总数; ΔU_i^t 包含了第 i 个分布式电源的输出电压与其相邻分布式电源或者参考值之间的电压偏差。

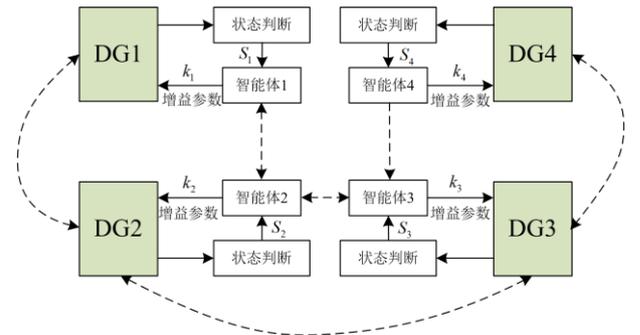


图 2 孤岛微电网系统协同强化学习控制框图

Fig. 2 Block diagram of cooperative reinforcement learning control for islanded microgrid system

3.2 状态、动作、奖励函数的定义

强化学习智能体的参数主要包括输入状态、输出动作以及奖励。状态是指有利于智能体进行决策的信息, 本文选择状态集为各分布式电源的输出电压偏差。动作是指智能体所做的决策, 本文中人为预定义一系列控制器的增益参数。奖励则是评判智能体所执行动作的价值标准。本文采用基于强化学习的微电网两级电压优化控制, 通过优化微电网分布式二级控制器增益参数, 有效提高了微电网分布

式电源输出电压调节性能。

假设孤岛微电网系统协同强化学习控制拓扑图如图 2 所示, 系统参数如表 1 所示。在系统中将智能体的状态信息 ΔU_i^t 分解为 12 个状态量。状态集为 $S = [0 \ 0.2 \ 0.3 \ 0.75 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8]$, 其中 S 表示在允许出现的范围内分布式电源输出的电压偏差状态。

表 1 微电网系统参数

Table 1 Parameters of microgrid system

参数大小	
负荷 1	$P_L: 20 \text{ kW}, Q_L: 8 \text{ kvar}$
负荷 2	$P_L: 12 \text{ kW}, Q_L: 5 \text{ kvar}$
负荷 3	$P_L: 16 \text{ kW}, Q_L: 8 \text{ kvar}$
R_s/Ω	0.03(DG1&DG2), 0.04(DG3&DG4)
L_s/mH	0.35(DG1& DG2), 0.46(DG3&DG4)
R_f/Ω	0.1
L_f/mH	1.35
$C_f/\mu\text{F}$	50
Z_{11}/Ω	0.23+j0.1
Z_{12}/Ω	0.35+j0.58
Z_{13}/Ω	0.23+j0.1

智能体的奖励是用来评估该智能体动作的价值, 考虑到系统控制的目标是恢复微电网输出电压至参考值并使其动态性能最好。为了避免智能体训练的过程奖励发散, 加快收敛速度, 经过多次实验最终选择基于势能的奖励函数(potential-based reward shaping, PBRS)促进智能体的奖励收敛, 本文使用的具有势能的局部奖励定义为

$$r_i^t = \begin{cases} 0, & |\Delta U_i^t| = 0 \\ -\lambda_1 |\Delta U_i^t|, & 0 < |\Delta U_i^t| < 1 \\ -\lambda_2 |\Delta U_i^t|, & 1 \leq |\Delta U_i^t| < 4 \\ -\lambda_3 |\Delta U_i^t|, & |\Delta U_i^t| \geq 4 \end{cases} \quad (19)$$

式中: r_i^t 表示在 t 时刻第 i 个智能体本身所获得的局部奖励; λ_1 、 λ_2 、 λ_3 为奖励系数。本实验中, λ_1 、 λ_2 、 λ_3 分别设置为 5、10、50。在智能体到达前述所设置的离散状态 0 时给予奖励值为 0, 在智能体到达前述状态空间中所列其他离散状态时给予不同程度的负奖励。为了使多智能体之间协同选取动作, 定义了协同全局奖励公式为

$$R_i^t = r_i^t + \beta \sum_{j \in M_i} r_j^t \quad (20)$$

式中: M_i 表示除第 i 个智能体本身之外其他所有智

能体的集合; β 为平均加权系数。通过协同全局奖励, 每个智能体所获得的动作价值中同样包含其他智能体的局部奖励, 这种协同奖励函数使各分布式电源的智能体在考虑自身状态的同时, 将其他智能体的状态纳入自己的考虑因素之内, 进而选择出全局最优的动作序列。

3.3 协同强化学习算法

本实验中, 设定系统从 0 s 开始运行到经历负荷投切等事件之后系统恢复为稳定值的过程, 称为一幕(Episode), 那么强化学习模型的训练目标是寻找最优策略使得每一个 Episode 结束后, 强化学习智能体获得的奖励(reward)最大化。

Q 学习是一种基于价值的强化学习算法, 智能体在训练初期会创建一个状态动作映射表 Q-table, 该表格中的数值表示在离散状态 s 下对应的某个动作 a 的价值, 智能体通过该数值来评估一个策略的优劣。智能体在学习过程中, 通过式(21)更新 Q 表的内容。

$$Q(s, a) = Q(s, a) + \theta [R + \gamma (\max_{a \in A} Q(s', a) - Q(s, a))] \quad (21)$$

式中: θ 为学习速率; γ 为折扣因子, $\gamma \in [0, 1]$ 表示未来回报相当于当前回报的重要程度; s 表示当前时刻的环境状态, s' 表示下一时刻的状态且 $s \in S$ 、 $s' \in S$ 、 $a \in A$, S 为环境状态集, A 为智能体动作集。Q 学习智能体通常选择贪婪策略以 $(1 - \varepsilon)$ 的概率从 Q 表里面选取价值最高的动作执行或按照概率 ε 从动作集中随机选取一个动作执行, 这种选择随机动作的行为可以有效避免智能体陷入局部最优状态, 该学习探索的训练过程直到智能体的全局奖励逐渐收敛趋于平稳才结束训练。为了加快智能体对状态动作对(即状态与动作的映射)的探索速率, 在本文中智能体训练初期将 ε 的值初始化为一个较大的数, 在智能体训练过程中通过 $\varepsilon = \varepsilon(1 - \varepsilon_{\text{decay}})$ 来逐渐减小 ε 的值, 其中 $\varepsilon_{\text{decay}}$ 是衰减系数。具体参数如表 2 所示。

表 2 Q 学习智能体参数

Table 2 Parameters of Q-learning agents

参数	数值
θ	0.01
γ	0.5
ε	0.9
$\varepsilon_{\text{decay}}$	0.05

本文提出的微电网协同强化学习优化算法流程图如图 3 所示。

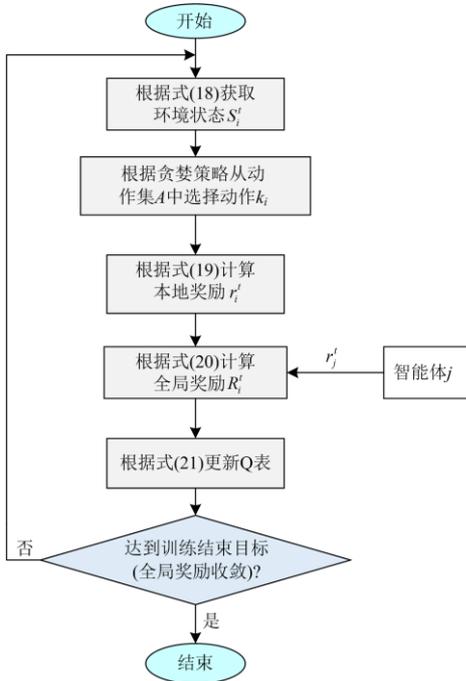


图3 协同强化学习优化算法流程图

Fig. 3 Flow chart of cooperative reinforcement learning optimization algorithm

4 仿真验证

为了验证本文提出的方法的有效性, 在 Matlab 2020a/Simulink 上搭建了包含 4 个并联分布式电源的孤岛微电网仿真模型, 仿真系统拓扑图如图 4 所示, 系统参数见表 1 所示。

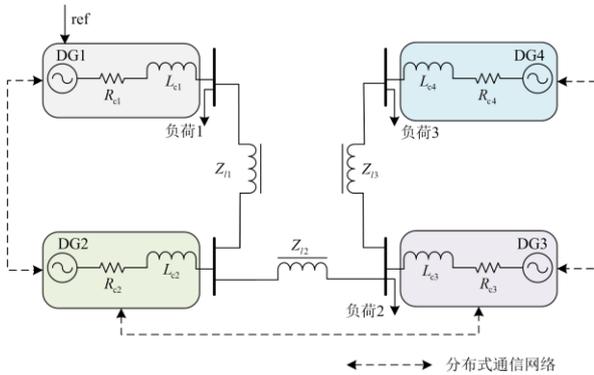


图4 仿真系统拓扑图

Fig. 4 Topology of simulation system

根据式(10)和式(15)可知, 孤岛微电网输出电压控制性能受控制器增益限制。一般微电网分布式控制器增益参数越大, 微电网输出电压调节速度越快, 但相应的输出电压动态性能可能越差。因此, 结合孤岛微电网对输出电压控制性能的要求, 根据式(17)可做如下推算。

考虑微电网外界干扰对输出电压影响造成的电压偏差情况, 假设初始输出跟踪误差 $\|\eta(t_0)\|$ 为 10, 根据电网通信控制实时性要求^[23], 本实验设定输出电压的调节时间为 0.2 s, 由于 $e^{-5.81} = 0.003$, 可以满足孤岛微电网输出电压的调节误差精度要求, 结合式(17)得到 $-\frac{\alpha}{2} \times 0.2 = -5.81$, $\alpha = 58.1 = \frac{\sigma_{\min}(\mathbf{O})}{\sigma_{\max}(\mathbf{X})}$ 。

本文设定 $\mathbf{X} = \mathbf{I}$ 为单位阵, 则有

$$\begin{cases} \sigma_{\min}(\mathbf{O}) = \sigma_{\min}[(\mathbf{KH})^T + \mathbf{KH}] = 58.1 \\ (\mathbf{KH})^T = \mathbf{H}^T \mathbf{K}^T = \mathbf{H}^T \mathbf{K} \end{cases} \quad (22)$$

可以推导出 $\sigma_{\min}(\mathbf{HK}) = 29.05$ 。

本文假设每个 DG 的强化学习动作集限制作用域相同, 即设定控制器增益上界相同且均为 k_{\max} , 则有 $\mathbf{K} = k_{\max} \mathbf{I}$ 。考虑:

$$\sigma_{\min}(\mathbf{H}^T \mathbf{K} + \mathbf{KH}) = 2k_{\max} \sigma_{\min}(\mathbf{H}) = 58.1 \quad (23)$$

$$\text{且 } \mathbf{H} = \mathbf{L} + \mathbf{g} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \text{ 因此, 有}$$

$$k_{\max} = \frac{29.05}{\sigma_{\min}(\mathbf{H})} = \frac{29.05}{0.1206} = 240.88 \quad (24)$$

综上所述, 本文设定孤岛微电网分布式二级控制器增益上界为 240。由于本文选用基于 Q 学习的强化学习方法, 通过离散化动作空间, 设定孤岛微电网协同强化学习动作集为 $A = [30 \ 60 \ 90 \ 120 \ 150 \ 180 \ 210 \ 240]$ 。

4.1 控制性能分析

图 5 给出了孤岛微电网在负荷投切时各分布式电源电压恢复过程。在 0~1 s 时孤岛微电网只采用初级控制带动负荷运行。此时各分布式电源的输出电压均在低于参考值 $V_{\text{ref}} = 311 \text{ V}$ 附近保持稳定。仅有下垂控制, 无法使微电网的电压稳定在参考值。为了消除初级控制产生的偏差, 在 1 s 时将本文提出的基于协同强化学习的分布式二级控制加入微电网中, 从仿真结果可以明显看出各分布式电源电压可以很快恢复至参考值。在 2 s 时将负荷 3 接入微电网中, 3 s 时将负荷 3 从微电网断开, 从图中可以看出, 当孤岛微电网受到负荷投切的干扰时, 基于强化学习的分布式协同二级控制可使电压快速恢复到参考值, 且具有良好的电压跟踪性能。

图 6 为训练过程中各 Q 学习智能体的平均奖励情况, 可以看出在训练的初始阶段智能体的平均奖励随训练次数增加而迅速增大, 经过数百次训练后增幅变缓并开始逐渐收敛。以智能体 4 的累计奖励为例, 经过大约 100 次训练之后, 从最初 -220 逐渐

收敛到-130 左右, 仿真结果表明本文所提出的强化学习方法是收敛有效的。

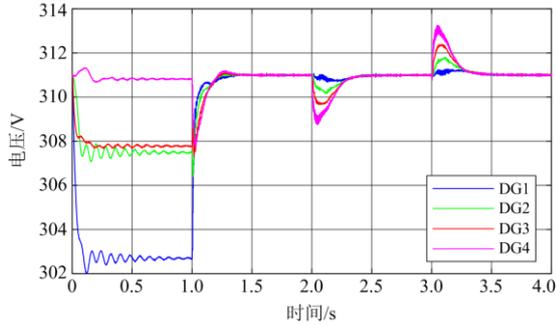


图 5 负荷投切情景孤岛微电网电压恢复性能
Fig. 5 Island microgrid voltage recovery performance in load switching scenario

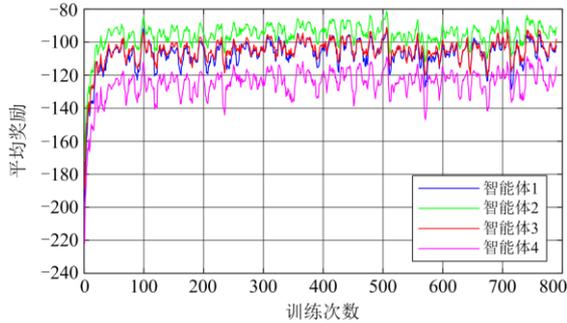


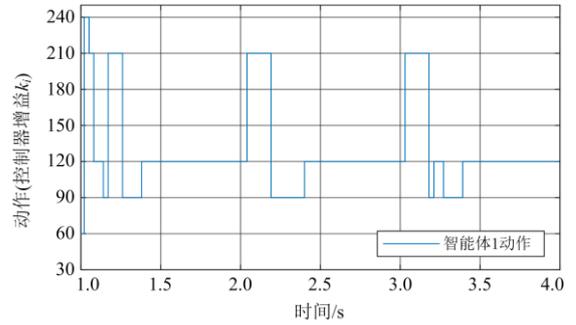
图 6 Q 学习智能体训练奖励
Fig. 6 Training episode rewards of Q-learning agents

图 7(a)一图 7(d)分别给出了各个 Q 学习智能体在训练奖励收敛情况下不同时刻动作图。从图 7 中可以看出智能体在不同运行环境下选择了不同的动作, 结合图 5 微电网各分布式电源良好的输出电压跟踪性能, 得出强化学习方法对不确定性干扰具有良好的自适应能力。

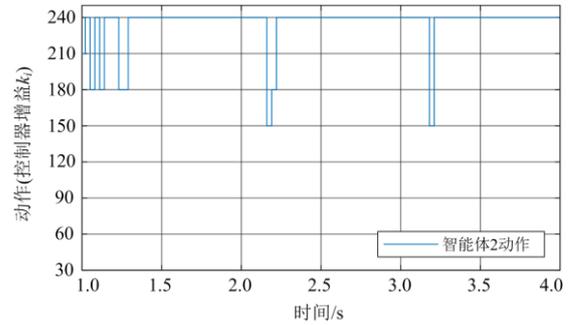
4.2 负荷投切控制性能

为了验证本文所提出的协同强化学习优化方法在负荷投切扰动情况下对电压控制性能的优越性, 考虑到文献[13]中同样采用分布式一致性积分控制方法对孤岛微电网进行电压恢复控制, 本节将本文的控制方法和文献[13]的控制方法进行对比。为了不失一般性, 列出了文献[13]的控制方法在不同控制器增益情况下各分布式电源的电压恢复情况, 如图 8 所示。

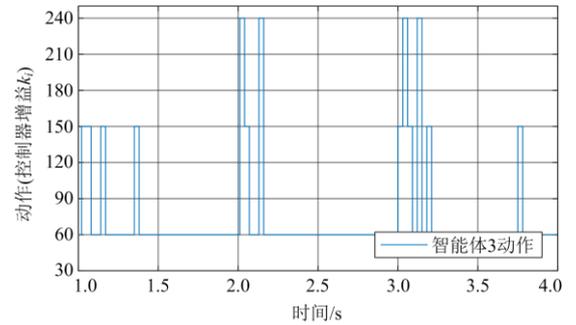
由 2.2.2 节可知, 控制器增益影响微电网的收敛速度, 控制器增益越大, 受到干扰时孤岛微电网电压恢复时间越短。从图 8 可以看出当控制器增益 $k_i = 30$ 时, 微电网电压恢复时间接近 1 s, 且最大超调量超过 5 V; 随着控制器增益 k_i 逐渐增大, 电压



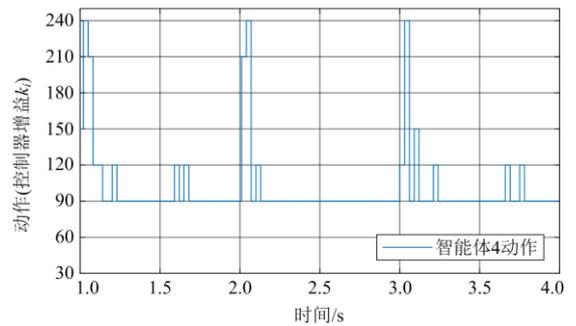
(a) 分布式电源1



(b) 分布式电源2



(c) 分布式电源3



(d) 分布式电源4

图 7 Q 学习智能体动作

Fig. 7 Action of Q-learning agents

调节时间明显缩短, 且超调量逐渐减小; 当控制器增益增加到 $k_i = 240$ 时, 采用分布式一致性积分控制仍然能将微电网输出电压调节到参考值, 但此时微电网的输出电压振荡特性明显, 动态性能变差。

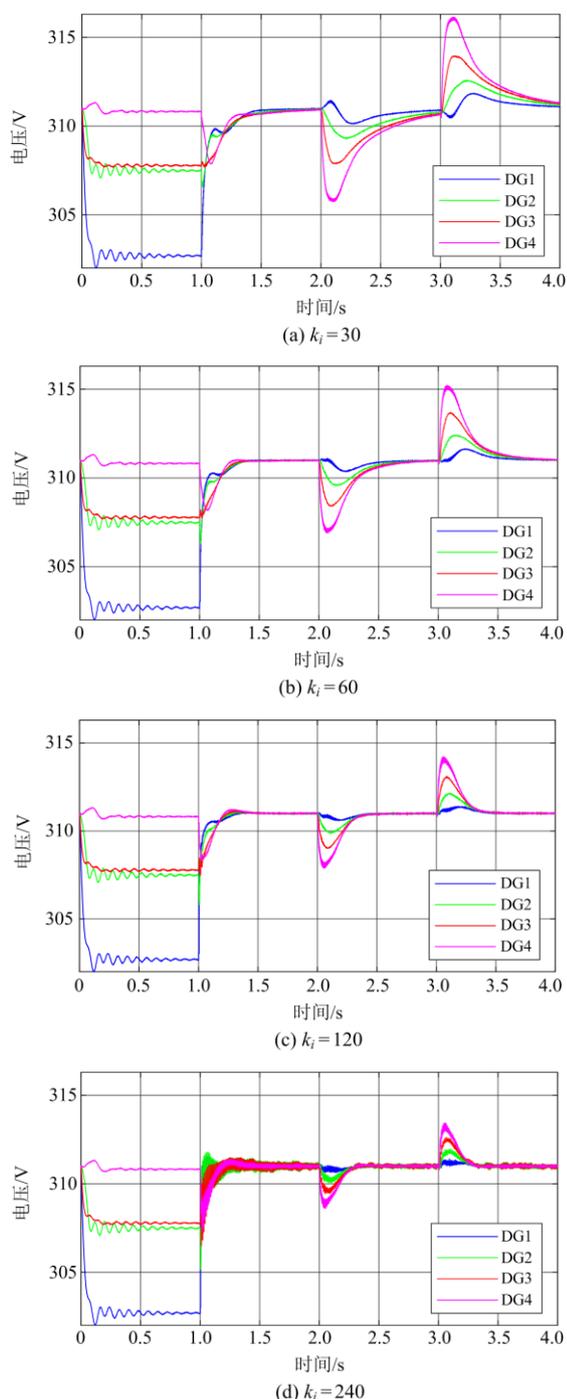
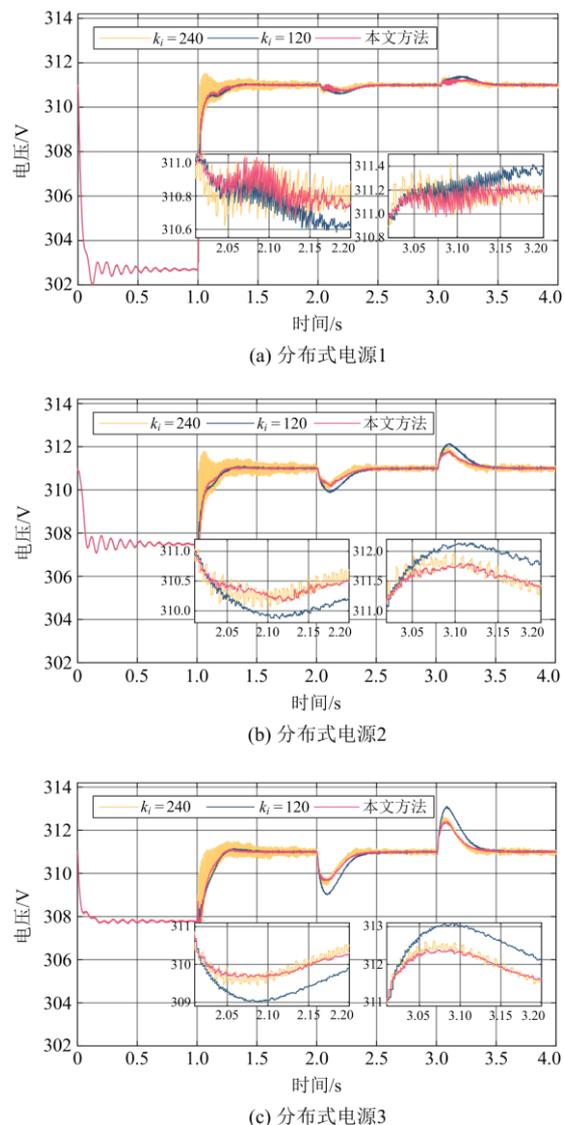


图8 不同控制器增益下微电网电压恢复性能

Fig. 8 Voltage recovery performance of microgrid under different controller gains

为了不失一般性, 图9给出了本文控制方法和文献[13]中 $k_i = 120$ 、 $k_i = 240$ 两种控制器增益情形下的孤岛微电网输出电压性能对比图, 图10给出了这三种控制情形下的控制器增益动作图。从图9可以看出, 当 $k_i = 120$ 时, 文献[13]中的控制方法相较

于本文的控制方法, 在负荷投切扰动下输出电压最大超调量明显更大, 且调节时间更长; 当 $k_i = 240$ 时, 文献[13]中的控制方法输出电压超调量与本文的控制方法大体相同, 但是其输出电压振荡特性明显, 动态性能差。由图10可知, 文献[13]中的传统控制方法由于控制器增益 k_i 固定不变, 在负荷投切扰动情况下, 虽然控制器增益增大可以减小电压恢复时间, 但与此同时微电网输出电压的动态性能变差。相较于文献[13]中传统的控制方法, 本文所提出的基于强化学习的优化方法可以根据负荷投切扰动情况灵活、自适应地选择控制器增益参数, 具有更快的调节速度以及更优的动态控制性能。本文所提出的控制方法, 相比文献[13]中的方法, 可以有效地提高孤岛微电网在面对负荷投切干扰时输出电压的动态控制性能。



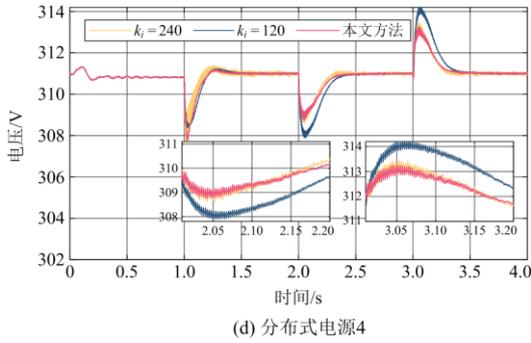


图 9 文献[15]中的方法和本文提出的协同强化学习优化控制方法在负荷投切情况下的性能对比

Fig. 9 Performance comparison between the method in Reference [15] and the proposed cooperative reinforcement learning optimization control in load switching scenario

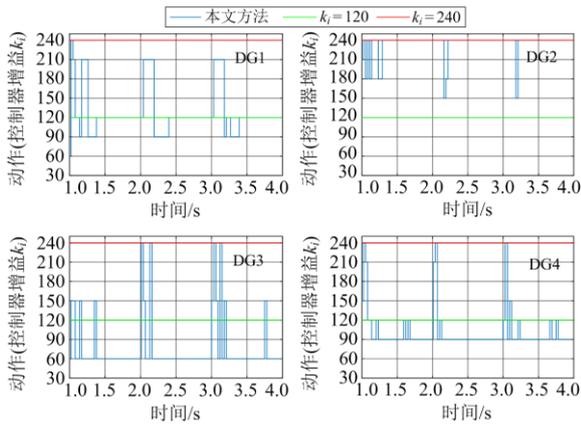


图 10 负荷投切情景控制器增益动作对比
Fig. 10 Controller gain action comparison in load switching scenario

4.3 即插即用控制性能

为了验证本文所提出的协同强化学习优化方法在即插即用扰动情况下其电压控制性能的优越性，在 2 s 时将 DG4 与微电网断开连接，在 3 s 时将 DG4 重新接入微电网。图 11 给出了即插即用情况下孤岛微电网电压恢复性能。从图 11 中可以看出，采用本文所提出的控制策略可以快速实现孤岛微电网电压恢复并具有良好的电压跟踪效果。由此可见本文所提出的方法针对孤岛微电网分布式电源的即插即用具有良好的控制性能。

为了进一步验证本文所提出的协同强化学习方法在即插即用扰动情况下对电压控制的优越性，将本文的控制方法和文献[13]中的控制方法进行对比。为了简化分析过程，图 12 列举了本文控制方法和文献[13]中控制器增益 $k_i = 120$ 情形下的孤岛微电网输出电压性能对比图，图 13 给出了这两种控制情形下的控制器增益动作图。从图 12 中可以看出，

本文所提出的控制方法在即插即用扰动下可以实现较好的电压跟踪效果，当分布式电源接入微电网，或分布式电源与微电网断开时，文献[13]中的控制方法相较于本文的控制方法，其输出电压具有相对更大的超调量和更长的调节时间。因此，本文提出的基于协同强化学习的孤岛微电网分布式两级电压优化方法具有更优的电压调节性能。

由图 13 可知，文献[13]中的传统控制方法在即插即用扰动情况下，由于控制器增益 k_i 固定不变，无法实现控制器增益参数的自适应变化。相较于文献[13]中传统的控制方法，本文所提出的基于协同强化学习的优化方法可以根据即插即用扰动情况自适应地选择控制器增益参数，因而具有更快的调节速度以及更优的动态控制性能。

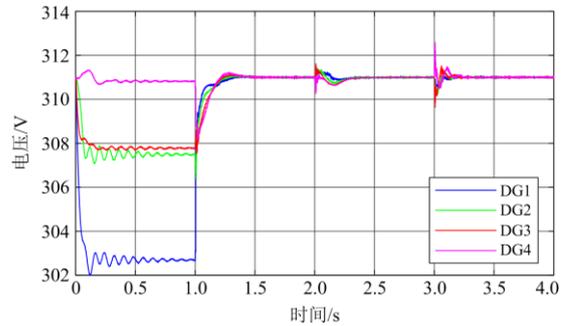
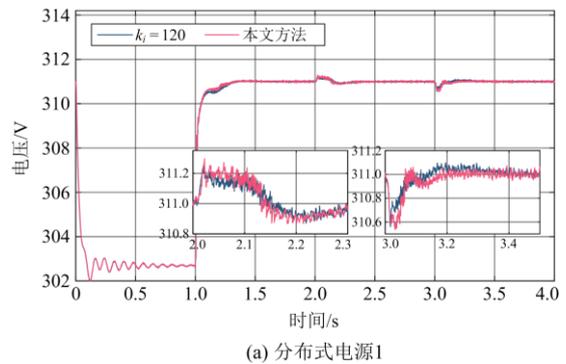
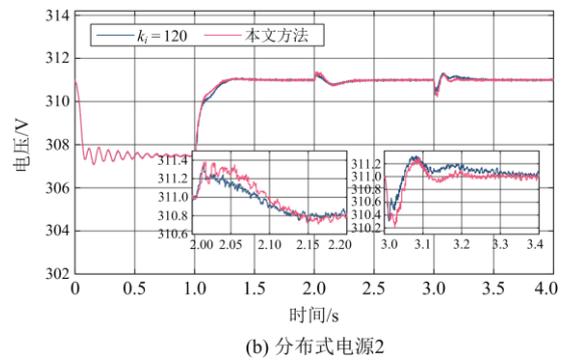


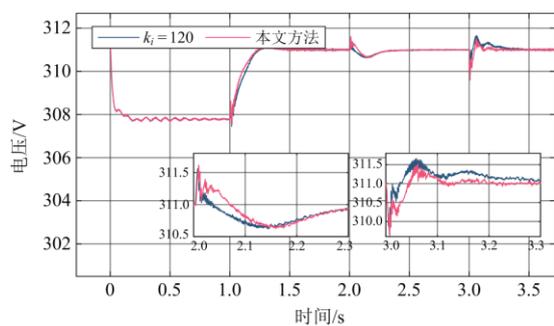
图 11 即插即用情景孤岛微电网电压恢复性能
Fig. 11 Island mode microgrid voltage restoration performance in plug-and-play scenario



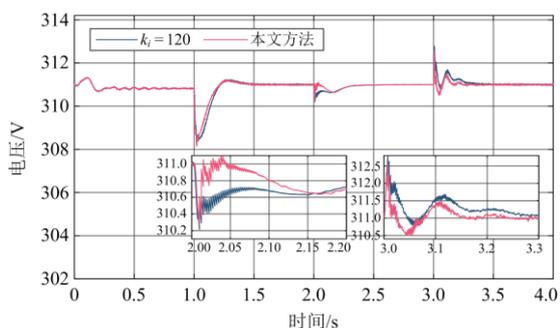
(a) 分布式电源1



(b) 分布式电源2



(c) 分布式电源3



(d) 分布式电源4

图 12 文献[13]中的方法和本文提出的协同强化学习优化控制方法在即插即用情况下的性能对比

Fig. 12 Performance comparison between the method in Reference [13] and the proposed cooperative reinforcement learning optimization control in plug-and-play scenario

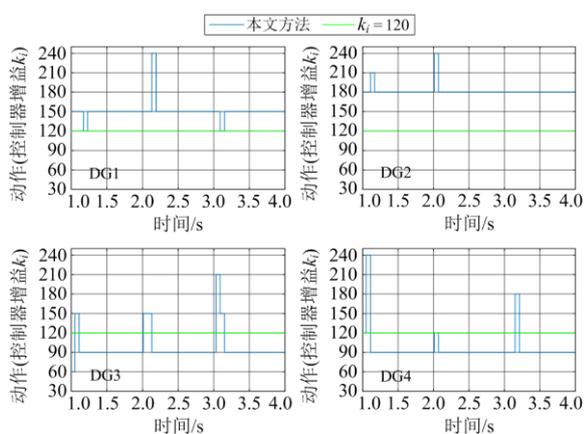


图 13 即插即用情景下控制器增益动作对比

Fig. 13 Controller gain comparison in plug-and-play scenario

5 结语

本文提出了一种基于协同强化学习的分布式两级电压优化控制策略, 实现孤岛模式下微电网的电压恢复控制。首先构建孤岛微电网分布式一致性协同电压控制算法, 并建立李雅普诺夫函数稳定性判

定方法, 通过进一步研究控制器性能与控制器增益之间的关系, 求解孤岛微电网电压控制器增益上界。然后在分布式一致性电压控制中引入协同强化学习算法优化二级控制器增益参数, 并用求解出的控制器增益上界限强化学习动作集边界, 有效提高微电网输出电压的动态控制性能。最后通过仿真对本文所提出的方法进行了验证, 与传统的二级控制方法相比, 本文所提出的方法在负荷投切以及分布式电源即插即用情况下可以自适应调节控制器增益参数, 本文控制方法具有更好的电压恢复性能。

参考文献

- [1] 许鹏, 黄鑫, 汪可友, 等. 适用于含多并联逆变器微电网网络功率控制策略[J]. 电力系统保护与控制, 2021, 49(16): 50-59.
XU Peng, HUANG Xin, WANG Keyou, et al. Tie-line power control strategy for a microgrid with multi-parallel connected inverters[J]. Power System Protection and Control, 2021, 49(16): 50-59.
- [2] 张良, 闫凯宏, 冷祥彪, 等. 基于 SOC 下垂控制的独立直流微电网协调控制策略研究[J]. 电力系统保护与控制, 2021, 49(12): 87-97.
ZHANG Liang, YAN Kaihong, LENG Xiangbiao, et al. Research on coordinated control strategy of an autonomous DC microgrid based on SOC droop control[J]. Power System Protection and Control, 2021, 49(12): 87-97.
- [3] 米阳, 王鹏, 邓锦, 等. 孤岛交直流混合微电网群分层协调控制[J]. 电力系统保护与控制, 2021, 49(20): 1-8.
MI Yang, WANG Peng, DENG Jin, et al. Hierarchical coordinated control of island AC/DC hybrid microgrids[J]. Power System Protection and Control, 2021, 49(20): 1-8.
- [4] 何黎君, 程杉, 陈梓铭, 等. 考虑交互功率控制和双边竞价交易的多微电网双层优化调度[J]. 电力系统保护与控制, 2020, 48(11): 10-17.
HE Lijun, CHENG Shan, CHEN Ziming, et al. A scheduling model of a multi-microgrid system based on bi-layer optimization with consideration of PCC power control and bilateral bidding[J]. Power System Protection and Control, 2020, 48(11): 10-17.
- [5] CHEN Feixiong, CHEN Minyou, LI Qiang, et al. Cost-based droop schemes for economic dispatch in islanded microgrids[J]. IEEE Transactions on Smart Grid, 2017, 8(1): 63-74.
- [6] HAQUE Y Y U, ISLAM R, AHMED T, et al. Improved voltage tracking of autonomous microgrid technology using a combined resonant controller with lead-lag compensator adopting negative imaginary theorem[J]. Protection and Control of Modern Power Systems, 2022, 7(1): 149-164.
- [7] ZHAO Changjuan, SUN Wei, WANG Jianping, et al. Distributed cooperative secondary control for islanded microgrid with Markov time-varying delays[J]. IEEE Transactions on Energy Conversion, 2019, 34(4): 2235-2247.

- [8] 陈晓祺, 贾宏杰, 陈硕翼, 等. 基于线路阻抗辨识的电网无功均分改进下垂控制策略[J]. 高电压技术, 2017, 43(4): 1271-1279.
CHEN Xiaoqi, JIA Hongjie, CHEN Shuoyi, et al. Improved droop control strategy based on line impedance identification for reactive power sharing in microgrid[J]. High Voltage Engineering, 2017, 43(4): 1271-1279.
- [9] 颜丽, 米阳, 孙威, 等. 基于改进下垂控制的孤岛交流电网无功分配研究[J]. 太阳能学报, 2021, 42(8): 7-15.
YAN Li, MI Yang, SUN Wei, et al. Reactive power distribution control strategy in islanded AC microgrid based on improved droop control[J]. Acta Energetica Solaris Sinica, 2021, 42(8): 7-15.
- [10] DING Lei, HAN Qinglong, ZHANG Xianming. Distributed secondary control for active power sharing and frequency regulation in islanded microgrids using an event-triggered communication mechanism[J]. IEEE Transactions on Industrial Informatics, 2019, 15(7): 3910-3922.
- [11] 吕振宇, 苏晨, 吴在军, 等. 孤岛型微电网分布式二次调节策略及通信拓扑优化[J]. 电工技术学报, 2017, 32(6): 209-219.
LÜ Zhenyu, SU Chen, WU Zaijun, et al. Distributed secondary control strategy and its communication topology optimization for islanded microgrid[J]. Transactions of China Electrotechnical Society, 2017, 32(6): 209-219.
- [12] LIAN Zhijie, DENG Chao, WEN Changyun, et al. Distributed event-triggered control for frequency restoration and active power allocation in microgrids with varying communication time delays[J]. IEEE Transactions on Industrial Electronics, 2021, 68(9): 8367-8378.
- [13] CHEN Meng, XIAO Xiangning, GUERRERO J M. Secondary restoration control of islanded microgrids with a decentralized event-triggered strategy[J]. IEEE Transactions on Industrial Informatics, 2018, 14(9): 3870-3880.
- [14] 范士雄, 李立新, 王松岩, 等. 人工智能技术在电网调控中的应用研究[J]. 电网技术, 2020, 44(2): 401-411.
FAN Shixiong, LI Lixin, WANG Songyan, et al. Application analysis and exploration of artificial intelligence technology in power grid dispatch and control[J]. Power System Technology, 2020, 44(2): 401-411.
- [15] FORUZAN E, SOH L, ASGARPOOR S. Reinforcement learning approach for optimal distributed energy management in a microgrid[J]. IEEE Transactions on Power Systems, 2018, 33(5): 5749-5758.
- [16] LIU Weiyong, ZHUANG Peng, LIANG Hao, et al. Distributed economic dispatch in microgrids based on cooperative reinforcement learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(6): 2192-2203.
- [17] 刘洪, 李吉峰, 葛少云, 等. 基于多主体博弈与强化学习的并网型综合能源微网协调调度[J]. 电力系统自动化, 2019, 43(1): 40-48.
LIU Hong, LI Jifeng, GE Shaoyun, et al. Coordinated scheduling of grid-connected integrated energy microgrid based on multi-agent game and reinforcement learning[J]. Automation of Electric Power Systems, 2019, 43(1): 40-48.
- [18] DU Yan, LI Fangxing. Intelligent multi-microgrid energy management based on deep neural network and model-free reinforcement learning[J]. IEEE Transactions on Smart Grid, 2020, 11(2): 1066-1076.
- [19] 杨挺, 赵黎媛, 刘亚闯, 等. 基于深度强化学习的综合能源系统动态经济调度[J]. 电力系统自动化, 2021, 45(5): 39-47.
YANG Ting, ZHAO Liyuan, LIU Yachuang, et al. Dynamic economic dispatch for integrated energy system based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2021, 45(5): 39-47.
- [20] 华和安, 方勇纯, 钱辰, 等. 基于线性滤波器的四旋翼无人机强化学习控制策略[J]. 电子与信息学报, 2021, 43(12): 3407-3417.
HUA He'an, FANG Yongchun, QIAN Chen, et al. Reinforcement learning control strategy of quadrotor unmanned aerial vehicles based on linear filter[J]. Journal of Electronics & Information Technology, 2021, 43(12): 3407-3417.
- [21] 杨忱, 赵红, 王宁, 等. 基于强化学习的指定性能轨迹跟踪最优控制[J]. 大连海事大学学报, 2021, 47(2): 1-10, 25.
YANG Chen, ZHAO Hong, WANG Ning, et al. Optimal control of specified performance trajectory tracking based on reinforcement learning[J]. Journal of Dalian Maritime University, 2021, 47(2): 1-10, 25.
- [22] 乔骥, 王新迎, 张擎, 等. 基于柔性行动器-评判器深度强化学习的电-气综合能源系统优化调度[J]. 中国电机工程学报, 2021, 41(3): 819-833.
QIAO Ji, WANG Xinying, ZHANG Qing, et al. Optimal dispatch of integrated electricity-gas system with soft actor-critic deep reinforcement learning[J]. Proceedings of the CSEE, 2021, 41(3): 819-833.
- [23] LAI Jingang, ZHOU Hong, LU Xiaoqing, et al. Droop-based distributed cooperative control for microgrids with time-varying delays[J]. IEEE Transactions on Smart Grid, 2016, 7(4): 1775-1789.

收稿日期: 2022-01-22; 修回日期: 2022-06-09

作者简介:

汪超(1999—), 男, 硕士研究生, 研究方向为微电网网络化控制、无线传感器网络、优化控制、机器学习; E-mail: wchao_super@163.com

赵婵娟(1989—), 女, 通信作者, 博士, 讲师, 研究方向为微电网网络化控制、无线传感器网络、优化控制、机器学习; E-mail: jojo20061864@126.com

程志友(1972—), 男, 博士, 教授, 博士生导师, 研究方向为电能质量分析、检测和评估。E-mail: czy@ahu.edu.cn

(编辑 许威)