

DOI: 10.13718/j.cnki.xdzk.2025.05.016

邹玉意, 陈勇, 刘越智, 等. 基于 Q 学习的源荷扰动下交直流微电网负荷频率控制方法 [J]. 西南大学学报(自然科学版), 2025, 47(5): 188-198.

# 基于 Q 学习的源荷扰动下 交直流微电网负荷频率控制方法

邹玉意<sup>1</sup>, 陈勇<sup>1</sup>, 刘越智<sup>1</sup>,  
AHMED Lotfy Haridy<sup>1,2</sup>, ESAM H. Abdelhameed<sup>2</sup>

1. 电子科技大学 自动化工程学院, 成都 611731; 2. 阿斯旺大学 能源工程学院, 埃及 阿斯旺 81528

**摘要:** 交直流微电网是实现可再生能源高效利用的有效手段, 为了保证电能质量, 交直流微电网系统的频率应该在源荷扰动的条件下保持稳定。常见的负荷频率控制方法存在易受模型精度影响和对扰动考虑不完善等问题, 控制效果不容易达到预期。提出了基于 Q 学习的交直流微电网负荷频率控制方法: 在微电网模型上, 分析了交直流微电网的特点, 并建立了交直流微电网基准系统; 在控制方法上, 采用的 Q 学习算法能够无需系统动力学知识而求解系统的控制问题, 提高了方法的实用性; 仿真表明, 通过所提方法, 在源荷扰动下交直流微电网系统的频率可达到期望的扰动抑制效果, 从而保证自身的安全运行。

**关 键 词:** 交直流微电网; Q 学习算法; 负荷频率控制; 源荷扰动

中图分类号: TM46

文献标志码: A

文章编号: 1673-9868(2025)05-0188-11

开放科学(资源服务)标识码(OSID):



## Load Frequency Control Method for AC-DC Microgrid with Power and Load Disturbance Based on Q-learning

ZOU Yuyi<sup>1</sup>, CHEN Yong<sup>1</sup>, LIU Yuezhi<sup>1</sup>,  
AHMED Lotfy Haridy<sup>1,2</sup>, ESAM H. Abdelhameed<sup>2</sup>

1. School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China;

2. Faculty of Energy Engineering, Aswan University, Aswan 81528, Egypt

**Abstract:** AC-DC microgrid is an effective means to realize the efficient utilization of renewable energy. In

收稿日期: 2024-11-30

基金项目: 国家重点研发计划“政府间国际科技创新合作”重点专项(2022YFE0120700); 埃及科技创新基金项目(44236)。

作者简介: 邹玉意, 硕士研究生, 主要从事微电网控制、隐私保护研究。

通信作者: 陈勇, 博士, 教授, 博士研究生导师。

order to ensure power quality, the frequency of AC-DC microgrid system should be kept stable under power and load disturbance. The common load frequency control methods are easy to be affected by the model accuracy and imperfect consideration of disturbance, so the control effect sometimes fails to reach the expectation. Therefore, the load frequency control method of AC-DC microgrid based on Q-learning is proposed. On the microgrid model, the characteristics of AC-DC microgrid were analyzed, and the benchmark system of AC-DC microgrid was established. In the control method, Q-learning algorithm can solve the control problem of the system without the knowledge of system dynamics, which improves the practicability of the method. The simulation results show that the proposed method can still achieve the desired disturbance suppression effect and ensure the safe operation of the AC-DC microgrid system under the source load disturbance. The proposed method is helpful for microgrid control.

**Key words:** AC-DC microgrid; Q-learning algorithm; load frequency control; power and load disturbance

大力发展战略风、光等可再生能源, 提高其在能源结构中的比重成为实现“双碳”战略的重要途径<sup>[1]</sup>。风、光等可再生能源具有污染小、储量大、分布广等优点, 但也存在随机性、波动性大且易受到外界环境影响等缺点<sup>[2]</sup>。微电网是实现风、光等可再生能源高效利用的有效手段<sup>[3]</sup>, 其中交直流微电网结合了交流微电网和直流微电网的优点<sup>[4]</sup>, 在顺应交流系统的主导地位的同时又兼顾直流负荷的需求, 因此交直流微电网技术的发展与研究受到前所未有的关注和重视<sup>[5]</sup>, 是未来微电网的重要研究方向<sup>[6]</sup>。分布式电源的功率扰动以及负荷变化都会造成交直流微电网系统频率的偏差, 给其正常运行带来巨大挑战<sup>[7]</sup>。交直流微电网的频率稳定是保障其运行的关键<sup>[8]</sup>, 而且频率也是评价微电网电能质量的重要指标。负荷频率控制(Load Frequency Control, LFC)方法就是通过微电网电源和负荷之间的功率平衡来实现其频率稳定的。

$H_{\infty}$ 控制因其在抑制扰动方面的鲁棒性而倍受关注<sup>[9]</sup>, 被认为与零和博弈问题密切相关。求解博弈代数黎卡提方程(Game Algebraic Riccati Equation, GARE)可以解决线性系统的零和博弈问题, 但是直接求解 GARE 需要完整的系统动力学知识, 而在实际生产中很难建立精确的系统动力学模型, 因此无模型方法的研究就显得非常重要<sup>[10]</sup>。强化学习(Reinforcement Learning, RL)能够通过定期实施性能评估和策略更新, 在不确定甚至未知的环境中学习最优控制策略<sup>[11]</sup>。重要的是, 强化学习可以应用于解决  $H_{\infty}$  控制问题<sup>[12]</sup>。近年来强化学习在动力学系统的控制上取得了许多成果<sup>[13]</sup>, 李臻等<sup>[14]</sup>针对线性离散系统在扰动存在条件下的控制问题, 提出了一种利用历史输入输出数据实现最优输出反馈策略的学习算法。在实际应用中, 整个系统的状态可能无法完全测量, 因而所提出的输出数据驱动方法比状态数据驱动方法更强大。

本研究提出了一种基于 Q 学习的交直流微电网负荷频率控制方法, 通过采集微电网的输入输出数据用最小二乘法求解 GARE 且无需系统动力学知识, 避免了在线算法需要频繁调节控制和扰动策略的缺点。仿真结果对本文所提出算法的有效性进行了验证。

## 1 交直流微电网的负荷频率控制问题模型

结合柴油发电机特点设计的交直流微电网结构如图 1。随着对可再生分布式能源的利用与发展, 不同类型的负荷迅速增加, 交直流微电网的发展受到了重点关注。交直流微电网包含直流母线和交流母线, 可以同时给交流负载和直流负载供电, 能够减少电能变换器的数量。交直流微电网也结合了交流微电网和直流微电网的优点, 降低了电能变换的损耗, 提高了微电网的可靠性, 成本大大降低。

风力发电机组的发电功率取决于风速  $v$ , 本研究风力发电机组的输出功率  $P_{\text{WT}}$  满足:

$$P_{\text{WT}} = 0.5\rho A_r C_p v^3 \quad (1)$$

式中:  $\rho$  为空气密度;  $A_r$  为风力发电机组叶片面积;  $C_p$  是风力发电机组的功率系数。

本研究光伏发电系统的输出功率  $P_{PV}$

满足:

$$P_{PV} = \eta S \Phi [1 - 0.005(T_a + 25)] \quad (2)$$

式中:  $\eta$  为光伏阵列的转换效率;  $S$  为光伏阵列的面积;  $\Phi$  为太阳辐射强度;  $T_a$  为环境温度。

定义  $\mathbf{x}(t) = [\Delta f, \Delta P, \Delta p]^T \in \mathbb{R}^3$  为状态向量,  $\mathbf{u}(t) \in \mathbb{R}^m$  为交直流微电网系统的控制信号,  $\mathbf{w}(t) = \mathbf{w}_{WT}(t) + \mathbf{w}_{PV}(t) + \mathbf{w}_{load}(t) \in \mathbb{R}^m$  是源荷扰动, 那么源荷扰动下的交直流微电网系统模型可以被描述为:

$$\dot{\mathbf{x}}(t) = \mathbf{A}_c \mathbf{x}(t) + \mathbf{B}_c \mathbf{u}(t) + \mathbf{D}_c \mathbf{w}(t) \quad (3)$$

其中具体的状态空间方程参数为:

$$\mathbf{A}_c = \begin{bmatrix} -\frac{1}{T_p} & \frac{k_p}{T_p} & 0 \\ 0 & -\frac{1}{T_t} & \frac{1}{T_t} \\ -\frac{1}{s_p T_g} & 0 & -\frac{1}{T_g} \end{bmatrix} \quad \mathbf{B}_c = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{T_g} \end{bmatrix} \quad \mathbf{D}_c = \begin{bmatrix} \frac{k_p}{T_p} \\ 0 \\ 0 \end{bmatrix}$$

式中:  $\Delta f$  表示频率偏差;  $\Delta P$  表示涡轮功率;  $\Delta p$  表示调速器位置值;  $\mathbf{w}_{WT}$ 、 $\mathbf{w}_{PV}$ 、 $\mathbf{w}_{load}$  分别表示由风机、光伏以及负载引起的功率变化;  $T_p$  表示交直流微电网系统的时间常数;  $k_p$  表示交直流微电网系统的增益;  $T_t$  表示涡轮机的时间常数;  $T_g$  表示调速器的时间常数;  $s_p$  表示调速系数。

随后对交直流微电网模型进行离散化, 得到其离散模型:

$$\mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k + \mathbf{B} \mathbf{u}_k + \mathbf{D} \mathbf{w}_k \quad (4)$$

从交直流微电网的模型中可以发现, 风能、光伏发电以及负载功率的变化都会引起系统频率的变化。频率的稳定性是评价交直流微电网系统质量和性能的重要指标, 因此需要设计一种鲁棒的 LFC 方案来实现交直流微电网系统频率的稳定。

$H_\infty$  控制广泛用于减弱扰动对动力系统性能的影响, 对于式(4)中的交直流微电网离散模型, 当其满足:

$$\sum_{k=0}^{\infty} [\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k] \leq \gamma^2 \sum_{k=0}^{\infty} \mathbf{w}_k^T \mathbf{w}_k \quad (5)$$

时, 交直流微电网系统的  $L_2$  增益小于等于  $\gamma$ 。

定义交直流微电网系统的  $H_\infty$  控制目标: 找到一个控制信号  $u$ , 使得:

- ① 当  $w=0$  时, 系统渐进稳定。
- ② 满足扰动衰减调节, 即系统的  $L_2$  增益小于等于  $\gamma$ 。

## 2 基于 Q 学习的负荷频率控制方法

$H_\infty$  控制问题与零和博奕问题有很大的联系, 而求解线性系统的博奕代数黎卡提方程(GARE)可以找到零和博奕问题的解。目前大部分求解 GARE 的方法需要完整的系统动力学知识, 而在实际应用中, 很多情况下我们无法得到完整的系统动力学知识, 故而需要无模型的算法来求解  $H_\infty$  控制问题。

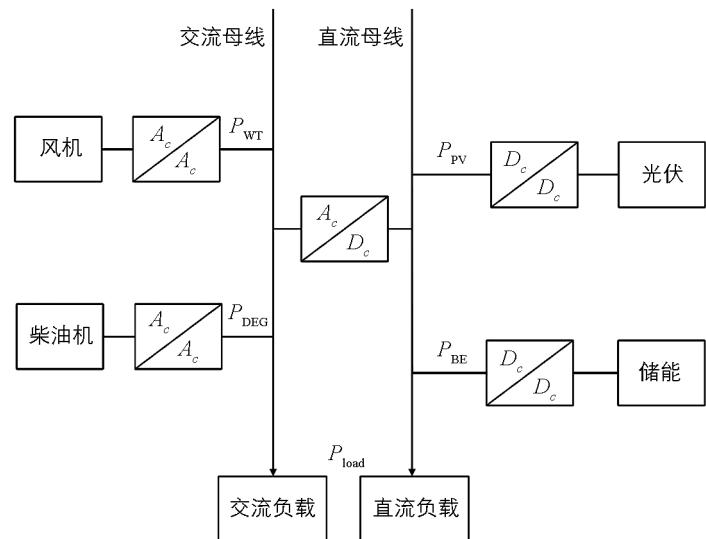


图 1 交直流微电网结构图

常规 Q 学习算法虽然在求解最优控制策略和最坏扰动时无需系统动力学知识, 但算法中每次迭代过后都需要通过更新后的控制策略来产生输出和扰动, 在实际情况下难以保证在迭代过程中每一次更新的策略都是稳定的, 而且常规的 Q 学习算法要求扰动按照期望进行调节并应用于系统动力学以收集数据, 这在实际应用中不符合物理条件, 扰动是独立的, 无法按照指定的方式进行更新, 因此考虑对常规 Q 学习算法进行改进, 源荷扰动下的交直流微电网离散模型式(4)重写为:

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}(\mathbf{K}_1^j \mathbf{x}_k + \mathbf{u}_k) + \mathbf{D}(\mathbf{K}_2^j \mathbf{x}_k + \mathbf{w}_k) \quad (6)$$

其中  $\mathbf{A}_k = \mathbf{A} - \mathbf{B}\mathbf{K}_1^j - \mathbf{D}\mathbf{K}_2^j$ 。

依据式(6)定义目标策略:  $\mathbf{u}_k^j = -\mathbf{K}_1^j \mathbf{x}_k$ ,  $\mathbf{w}_k^j = -\mathbf{K}_2^j \mathbf{x}_k$ , 目标策略作为算法正在更新和迭代, 而  $\mathbf{u}_k$ 、 $\mathbf{w}_k$  则作为行为策略, 实际应用于系统动力学以收集数据, 因此作用于系统的扰动  $\mathbf{w}_k$  无需按照期望进行调节。

那么贝尔曼方程为:

$$\mathbf{V}^{j+1}(\mathbf{x}_k) - \mathbf{V}^{j+1}(\mathbf{x}_{k+1}) = \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + (\mathbf{u}_k^j)^T \mathbf{R} \mathbf{u}_k^j - \gamma^2 (\mathbf{w}_k^j)^T \mathbf{w}_k^j \quad (7)$$

考虑函数  $\mathbf{V}(\mathbf{x}_k)$  在  $\mathbf{x}_{k+1}$  处进行泰勒展开:

$$\mathbf{V}(\mathbf{x}_k) = \mathbf{V}(\mathbf{x}_{k+1}) + (\nabla \mathbf{V}(\mathbf{x}_{k+1}))^T (\mathbf{x}_k - \mathbf{x}_{k+1}) + \frac{1}{2} (\mathbf{x}_k - \mathbf{x}_{k+1})^T \nabla^2 \mathbf{V}(\mathbf{x}_{k+1}) (\mathbf{x}_k - \mathbf{x}_{k+1}) \quad (8)$$

考虑式(8)并结合式(7)得到:

$$\begin{aligned} \mathbf{x}_k^T \mathbf{P}^{j+1} \mathbf{x}_k - \mathbf{x}_{k+1}^T \mathbf{P}^{j+1} \mathbf{x}_{k+1} &= \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{x}_k^T (\mathbf{K}_1^j)^T \mathbf{R} \mathbf{K}_1^j \mathbf{x}_k - \\ &\quad \gamma^2 \mathbf{x}_k^T (\mathbf{K}_2^j)^T \mathbf{K}_2^j \mathbf{x}_k - (\mathbf{K}_1^j \mathbf{x}_k + \mathbf{u}_k)^T \mathbf{B}^T \mathbf{P}^{j+1} (\mathbf{x}_{k+1} + \mathbf{A}_k \mathbf{x}_k) - \\ &\quad (\mathbf{K}_2^j \mathbf{x}_k + \mathbf{w}_k)^T \mathbf{D}^T \mathbf{P}^{j+1} (\mathbf{x}_{k+1} + \mathbf{A}_k \mathbf{x}_k) \end{aligned} \quad (9)$$

然后结合式(6), 将式(9)重写为 Kronecker 积的形式:

$$\begin{aligned} (\mathbf{x}_k^T \otimes \mathbf{x}_k^T) \text{vec}(\mathbf{P}^{j+1}) - (\mathbf{x}_{k+1}^T \otimes \mathbf{x}_{k+1}^T) \text{vec}(\mathbf{P}^{j+1}) &= \\ \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{x}_k^T (\mathbf{K}_1^j)^T \mathbf{R} \mathbf{K}_1^j \mathbf{x}_k - \gamma^2 \mathbf{x}_k^T (\mathbf{K}_2^j)^T \mathbf{K}_2^j \mathbf{x}_k - 2(\mathbf{x}_k^T \otimes (\mathbf{K}_1^j \mathbf{x}_k + \mathbf{u}_k)^T) \text{vec}(\mathbf{B}^T \mathbf{P}^{j+1} \mathbf{A}) + \\ ((\mathbf{K}_1^j \mathbf{x}_k - \mathbf{u}_k)^T \otimes (\mathbf{K}_1^j \mathbf{x}_k + \mathbf{u}_k)^T) \text{vec}(\mathbf{B}^T \mathbf{P}^{j+1} \mathbf{B}) - 2(\mathbf{x}_k^T \otimes (\mathbf{K}_2^j \mathbf{x}_k + \mathbf{w}_k)^T) \text{vec}(\mathbf{D}^T \mathbf{P}^{j+1} \mathbf{A}) + \\ ((\mathbf{K}_1^j \mathbf{x}_k - \mathbf{u}_k)^T \otimes (\mathbf{K}_2^j \mathbf{x}_k + \mathbf{w}_k)^T) \text{vec}(\mathbf{D}^T \mathbf{P}^{j+1} \mathbf{B}) + ((\mathbf{K}_2^j \mathbf{x}_k - \mathbf{w}_k)^T \otimes (\mathbf{K}_1^j \mathbf{x}_k + \mathbf{u}_k)^T) \text{vec}(\mathbf{B}^T \mathbf{P}^{j+1} \mathbf{D}) + \\ ((\mathbf{K}_2^j \mathbf{x}_k - \mathbf{w}_k)^T \otimes (\mathbf{K}_2^j \mathbf{x}_k + \mathbf{w}_k)^T) \text{vec}(\mathbf{D}^T \mathbf{P}^{j+1} \mathbf{D}) \end{aligned} \quad (10)$$

令:

$$\begin{aligned} \mathbf{L}_1^{j+1} &= \mathbf{P}^{j+1}, \quad \mathbf{L}_2^{j+1} = \mathbf{B}^T \mathbf{P}^{j+1} \mathbf{A}, \quad \mathbf{L}_3^{j+1} = \mathbf{B}^T \mathbf{P}^{j+1} \mathbf{B} \\ \mathbf{L}_4^{j+1} &= \mathbf{D}^T \mathbf{P}^{j+1} \mathbf{A}, \quad \mathbf{L}_5^{j+1} = \mathbf{D}^T \mathbf{P}^{j+1} \mathbf{B} \\ \mathbf{L}_6^{j+1} &= \mathbf{B}^T \mathbf{P}^{j+1} \mathbf{D}, \quad \mathbf{L}_7^{j+1} = \mathbf{D}^T \mathbf{P}^{j+1} \mathbf{D} \\ \mathbf{A}_1^j(i) &= [\mathbf{H}_{xx, i} \quad \mathbf{H}_{xu, i} \quad \mathbf{H}_{uu, i} \quad \mathbf{H}_{xw, i} \quad \mathbf{H}_{uw, i} \quad \mathbf{H}_{wu, i} \quad \mathbf{H}_{ww, i}] \\ \mathbf{A}_2^j(i) &= \mathbf{x}_{k+i-1}^T \mathbf{Q} \mathbf{x}_{k+i-1} + \mathbf{x}_{k+i-1}^T (\mathbf{K}_1^j)^T \mathbf{R} \mathbf{K}_1^j \mathbf{x}_{k+i-1} - \gamma^2 \mathbf{x}_{k+i-1}^T (\mathbf{K}_2^j)^T \mathbf{K}_2^j \mathbf{x}_{k+i-1} \end{aligned}$$

其中:

$$\begin{aligned} \mathbf{H}_{xx, i} &= \mathbf{x}_{k+i-1}^T \otimes \mathbf{x}_{k+i-1}^T - \mathbf{x}_{k+i}^T \otimes \mathbf{x}_{k+i}^T \\ \mathbf{H}_{xu, i} &= 2(\mathbf{x}_{k+i-1}^T \otimes (\mathbf{K}_1^j \mathbf{x}_{k+i-1} + \mathbf{u}_{k+i-1})^T) \\ \mathbf{H}_{uu, i} &= -(\mathbf{K}_1^j \mathbf{x}_{k+i-1} - \mathbf{u}_{k+i-1})^T \otimes (\mathbf{K}_1^j \mathbf{x}_{k+i-1} + \mathbf{u}_{k+i-1})^T \\ \mathbf{H}_{xw, i} &= 2(\mathbf{x}_{k+i-1}^T \otimes (\mathbf{K}_2^j \mathbf{x}_{k+i-1} + \mathbf{w}_{k+i-1})^T) \\ \mathbf{H}_{uw, i} &= -(\mathbf{K}_1^j \mathbf{x}_{k+i-1} - \mathbf{u}_{k+i-1})^T \otimes (\mathbf{K}_2^j \mathbf{x}_{k+i-1} + \mathbf{w}_{k+i-1})^T \\ \mathbf{H}_{wu, i} &= -(\mathbf{K}_2^j \mathbf{x}_{k+i-1} - \mathbf{w}_{k+i-1})^T \otimes (\mathbf{K}_1^j \mathbf{x}_{k+i-1} + \mathbf{u}_{k+i-1})^T \\ \mathbf{H}_{ww, i} &= -(\mathbf{K}_2^j \mathbf{x}_{k+i-1} - \mathbf{w}_{k+i-1})^T \otimes (\mathbf{K}_2^j \mathbf{x}_{k+i-1} + \mathbf{w}_{k+i-1})^T \end{aligned}$$

而且有：

$$\mathbf{A}_1^j = \begin{bmatrix} \mathbf{A}_1^j(1) \\ \mathbf{A}_1^j(2) \\ \cdots \\ \cdots \\ \mathbf{A}_1^j(N_3) \end{bmatrix} \quad \mathbf{A}_2^j = \begin{bmatrix} \mathbf{A}_2^j(1) \\ \mathbf{A}_2^j(2) \\ \cdots \\ \cdots \\ \mathbf{A}_2^j(N_3) \end{bmatrix}$$

那么可以求解  $\mathbf{L}_1 \sim \mathbf{L}_7$ ：

$$\begin{bmatrix} \text{vec}(\mathbf{L}_1^{j+1}) \\ \text{vec}(\mathbf{L}_2^{j+1}) \\ \text{vec}(\mathbf{L}_3^{j+1}) \\ \text{vec}(\mathbf{L}_4^{j+1}) \\ \text{vec}(\mathbf{L}_5^{j+1}) \\ \text{vec}(\mathbf{L}_6^{j+1}) \\ \text{vec}(\mathbf{L}_7^{j+1}) \end{bmatrix} = ((\mathbf{A}_1^j)^T \mathbf{A}_1^j)^{-1} (\mathbf{A}_1^j)^T \mathbf{A}_2^j \quad (11)$$

需要注意的是，使用最小二乘法求解  $\mathbf{L}_1 \sim \mathbf{L}_7$  需要至少  $N_3 \geq n^2 + m_1^2 + m_2^2 + 2m_1m_2 + n(m_1 + m_2)$  数量的数据集。

依据  $\mathbf{L}_1 \sim \mathbf{L}_7$  求解目标策略  $\mathbf{K}_1^{j+1}、\mathbf{K}_2^{j+1}$ ：

$$\begin{aligned} \mathbf{K}_1^{j+1} &= (\mathbf{R} + \mathbf{L}_3^{j+1} + \mathbf{L}_6^{j+1} (\gamma^2 \mathbf{I} - \mathbf{L}_7^{j+1})^{-1} \times \mathbf{L}_5^{j+1})^{-1} [\mathbf{L}_2^{j+1} + \mathbf{L}_6^{j+1} (\gamma^2 \mathbf{I} - \mathbf{L}_7^{j+1})^{-1} \mathbf{L}_4^{j+1}] \\ \mathbf{K}_2^{j+1} &= (\mathbf{L}_7^{j+1} - \gamma^2 \mathbf{I} - \mathbf{L}_5^{j+1} (\mathbf{R} + \mathbf{L}_3^{j+1})^{-1} \times \mathbf{L}_6^{j+1})^{-1} [\mathbf{L}_4^{j+1} - \mathbf{L}_5^{j+1} (\mathbf{R} + \mathbf{L}_3^{j+1})^{-1} \mathbf{L}_2^{j+1}] \end{aligned} \quad (12)$$

离线策略无模型的 Q 学习算法求解目标策略无需系统动力学知识，且将目标策略与行为策略分离开，行为策略直接作用于系统用于收集数据而不参与更新，目标策略并不直接作用于系统也可根据收集的数据进行迭代更新，那么系统扰动即行为策略无需按照期望进行调节。

离线策略无模型的 Q 学习算法的具体流程如图 2。

算法流程：

① 初始化行为策略  $\mathbf{K}_1、\mathbf{K}_2$ ，目标策略  $\mathbf{K}_1^{j+1}、\mathbf{K}_2^{j+1}$  设置  $j = 0$ 。

② 将行为策略  $\mathbf{K}_1、\mathbf{K}_2$  作用于系统，收集数据。

③ 策略评估：依据式(11)用最小二乘法求解  $\mathbf{L}_1 \sim \mathbf{L}_7$ 。

④ 策略改进：依据式(12)用  $\mathbf{L}_1 \sim \mathbf{L}_7$  得到目标策略  $\mathbf{K}_1^{j+1}、\mathbf{K}_2^{j+1}$ ，更新控制和扰动策略。

⑤ 终止条件：满足  $|\mathbf{K}_1^{j+1} - \mathbf{K}_1^j| \leq \varepsilon, |\mathbf{K}_2^{j+1} - \mathbf{K}_2^j| \leq \varepsilon$  时停止迭代，否则返回 ②。

### 3 结果与分析

测试平台为 MATLAB R2024a，测试系统采用式(4)的交直流微电网离散模型，基准系统具体的参数设置为  $T_t = 5; T_g = 0.2; T_p = 2; k_p = 0.5; s_p = 0.5^{[15]}$ 。根据具体的测试系统参数以及采样时间  $T = 0.05$  s 可以得到交直流微电网的离散模型：

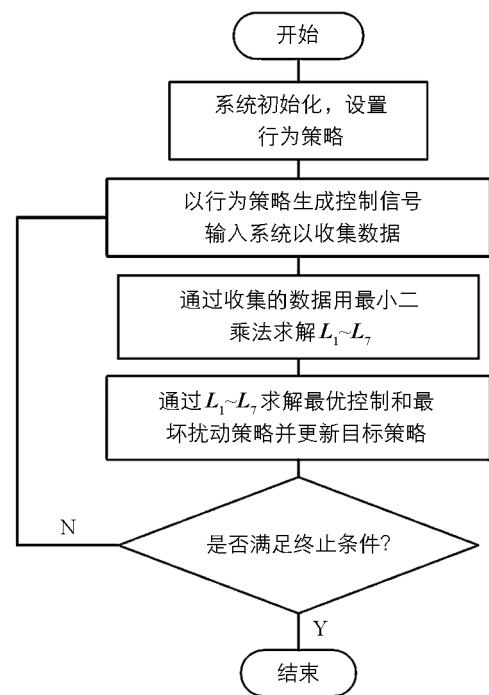


图 2 Q 学习算法流程图

$$\mathbf{x}_{k+1} = \begin{bmatrix} 0.9753 & 0.0123 & 5.6921 \times 10^{-5} \\ -9.1073 \times 10^{-5} & 0.9900 & 0.0088 \\ -0.0175 & -1.1384 \times 10^{-4} & 0.7788 \end{bmatrix} \mathbf{x}_k + \\ \begin{bmatrix} 4.8553 \times 10^{-6} \\ 0.0011 \\ 0.2212 \end{bmatrix} \mathbf{u}_k + \begin{bmatrix} 0.0123 \\ -3.8842 \times 10^{-7} \\ -1.1423 \times 10^{-4} \end{bmatrix} \mathbf{w}_k \quad (13)$$

采用  $Q = \text{diag}(1, 0, 1)$ ,  $R = 0.1$ ,  $\gamma = 0.8$  作为算法的参数, 根据离散模型以及参数值, 在算法开始前先用 GARE 求解得到理论的最优控制策略( $\mathbf{K}_1^*$ )和最坏扰动策略( $\mathbf{K}_2^*$ )为:

$$\mathbf{K}_1^* = [0.1390 \ 0.2659 \ 1.5112]$$

$$\mathbf{K}_2^* = [-0.4305 \ -0.1793 \ -0.0019]$$

设置初始状态和策略为:

$$\mathbf{x}_0 = [0.1 \ -0.2 \ 0.2]^T \quad \mathbf{K}_1 = [3 \ 2.5 \ 1] \quad \mathbf{K}_2 = [1 \ 1 \ 1]$$

设置总周期数为 800 个, 分为训练周期和验证周期, 在训练周期内向控制信号添加探测噪声, 在验证周期内训练得到的策略进行系统控制。利用本文提出的改进 Q 学习算法与常规 Q 学习算法分别进行最优控制策略和最坏扰动策略的求解。对于常规 Q 学习算法每一轮迭代至少需要收集 15 组数据用于求解, 对于改进 Q 学习算法每一轮迭代至少需要收集 19 组数据用于求解  $\mathbf{L}_1 \sim \mathbf{L}_7$ , 本研究设置两种算法每一轮迭代都收集 20 组数据用于求解, 那么总迭代轮次为 40 轮, 设置训练周期为 400 个, 即训练轮次为 20 轮, 验证周期为 400 个, 验证轮次也为 20 轮。

基于式(13)中的交直流微电网离散模型, 并考虑两种不同的探测噪声:

情况 1:

$$e(n) = 10 * (\text{rand} - 0.5)$$

情况 2:

$$e(n) = 0.2 \sin(1.009n) + \cos^2(0.538n) + \sin(0.9n) + \cos(100n)$$

在探测噪声为情况 1 的条件下, 常规的 Q 学习算法得到的结果如图 3。图 3a 的参数收敛图表示策略  $\mathbf{K}_1$ 、 $\mathbf{K}_2$  的迭代过程; 图 3b 中的系统收敛图则表示系统状态随周期的收敛过程; 图 3c 中的扰动示意图则表示系统扰动随周期的变化过程。

参数收敛图表示每一轮迭代通过策略改进计算得到的控制策略与扰动策略收敛到理论的最优控制策略和最坏扰动策略的过程; 系统收敛图与扰动示意图则表示每一个周期的系统状态与扰动的值。

在探测噪声为情况 1 的条件下, 改进 Q 学习算法得到的结果如图 4。

通过图 3 与图 4 中的参数收敛图可以看到, 在同样的参数下, 改进 Q 学习算法的收敛速度更快且收敛值更为精确。通过扰动示意图可以看到, 在图 3c 中常规 Q 学习算法的扰动情况是根据参数迭代情况进行调节的, 而在图 4c 中的改进 Q 学习算法是预设的真实扰动情况, 无需根据参数迭代情况进行调节。

在探测噪声为情况 2 的条件下, 常规 Q 学习算法得到的结果如图 5, 改进 Q 学习算法得到的结果如图 6。

通过图 3 和图 5、图 4 和图 6 中的参数收敛图可以看出, 参数的收敛曲线探测噪声改变时几乎不会发生变化, 说明算法不会受到探测噪声影响而导致错误的结果。通过两种算法的参数收敛图以及系统收敛图可以看出, 改进 Q 学习算法后的收敛速度更快而且收敛精度也更高, 且系统收敛后在扰动作用下可以保证系统频率偏差  $|\Delta f| < 0.2$ , 满足性能的要求。通过两种算法的扰动示意图可以看出, 当探测噪声改变后, 常规 Q 学习算法的扰动需要根据参数迭代情况而进行调节, 因而其扰动也发生了改变, 而改进 Q 学习算法的扰动是预设的真实情况, 不会随探测噪声改变而改变, 说明改进 Q 学习算法的性能更好而且更实用。

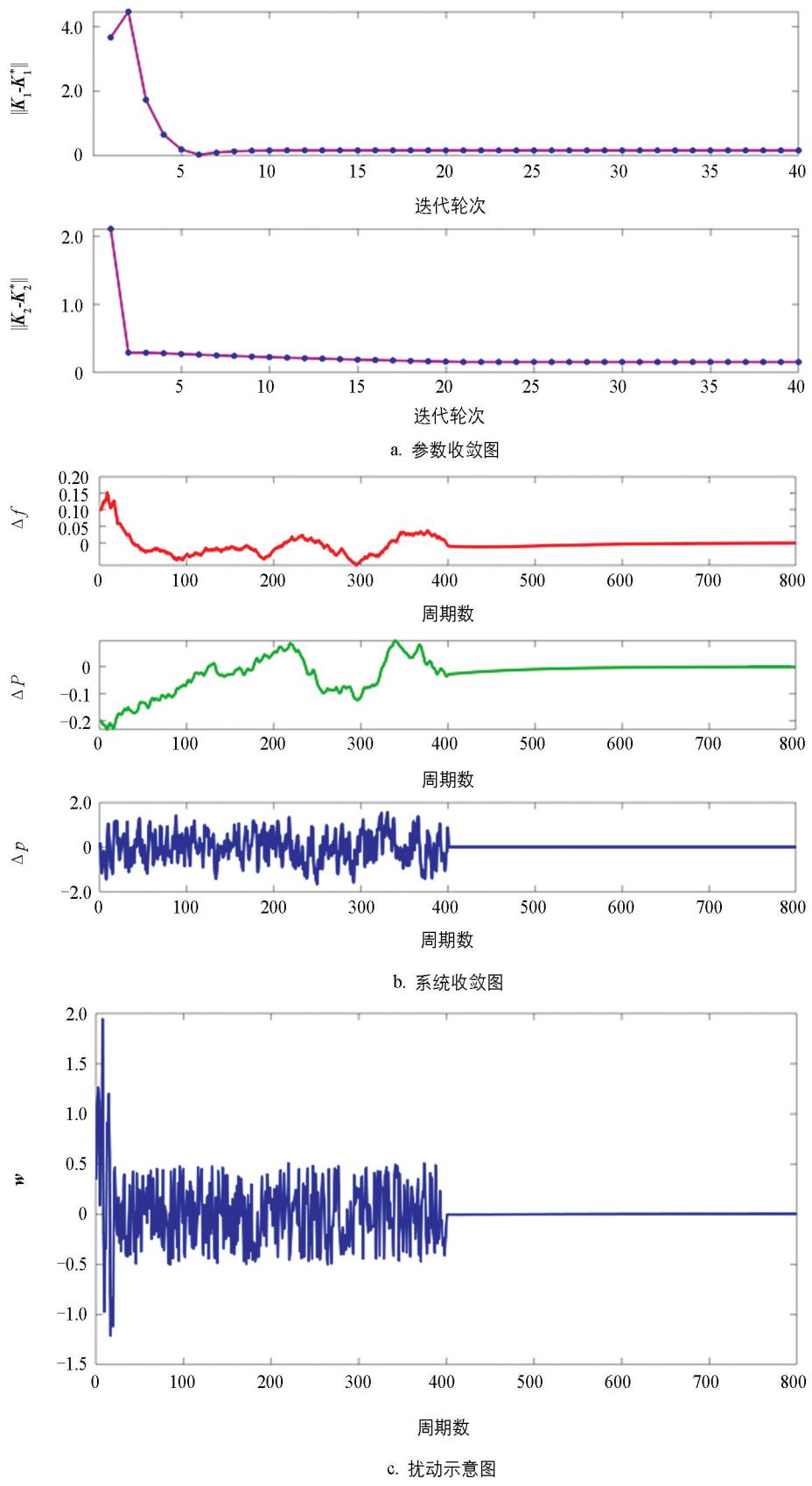


图 3 在情况 1 下常规 Q 学习算法的结果图

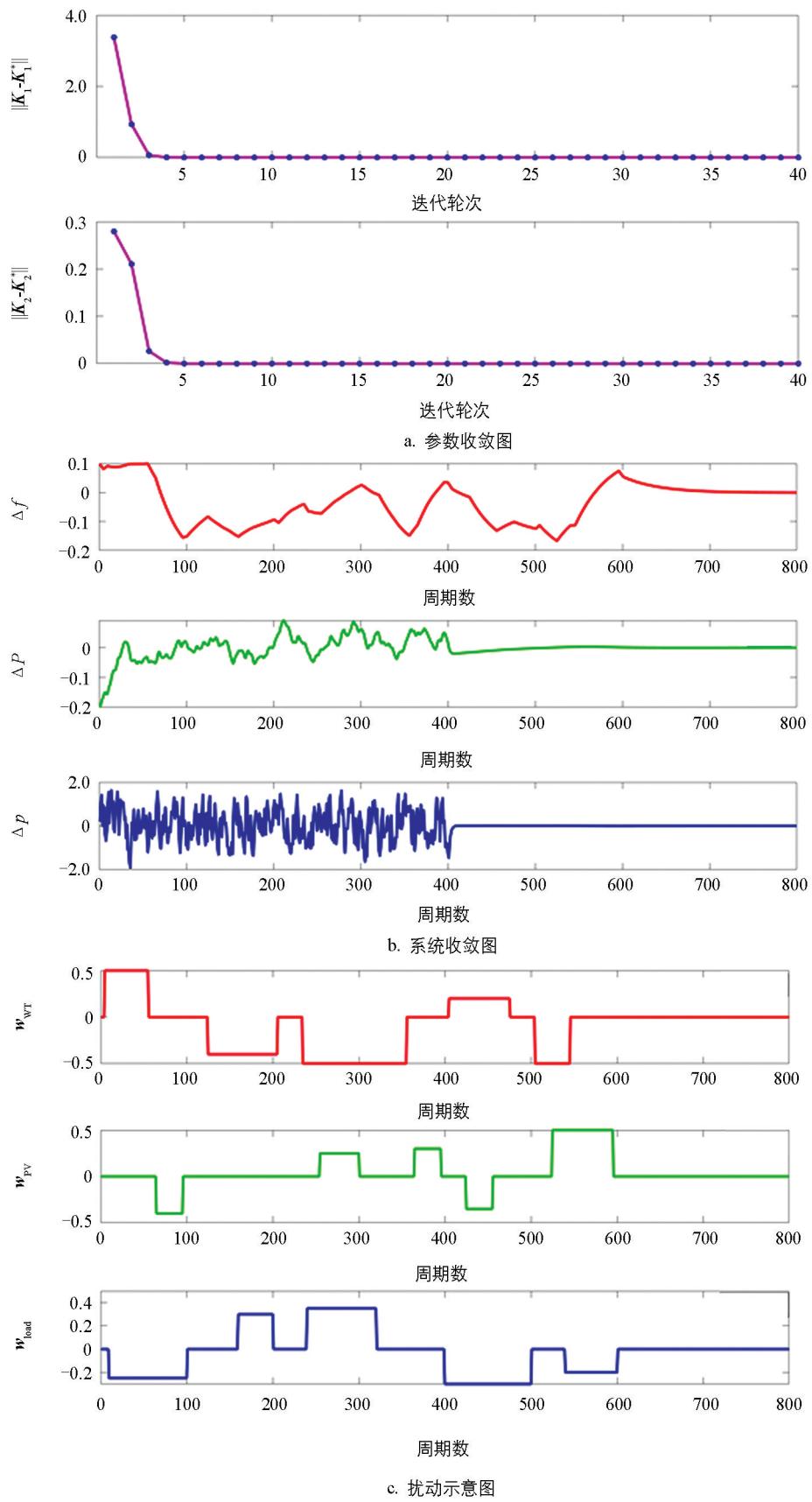


图 4 在情况 1 下改进 Q 学习算法的结果图

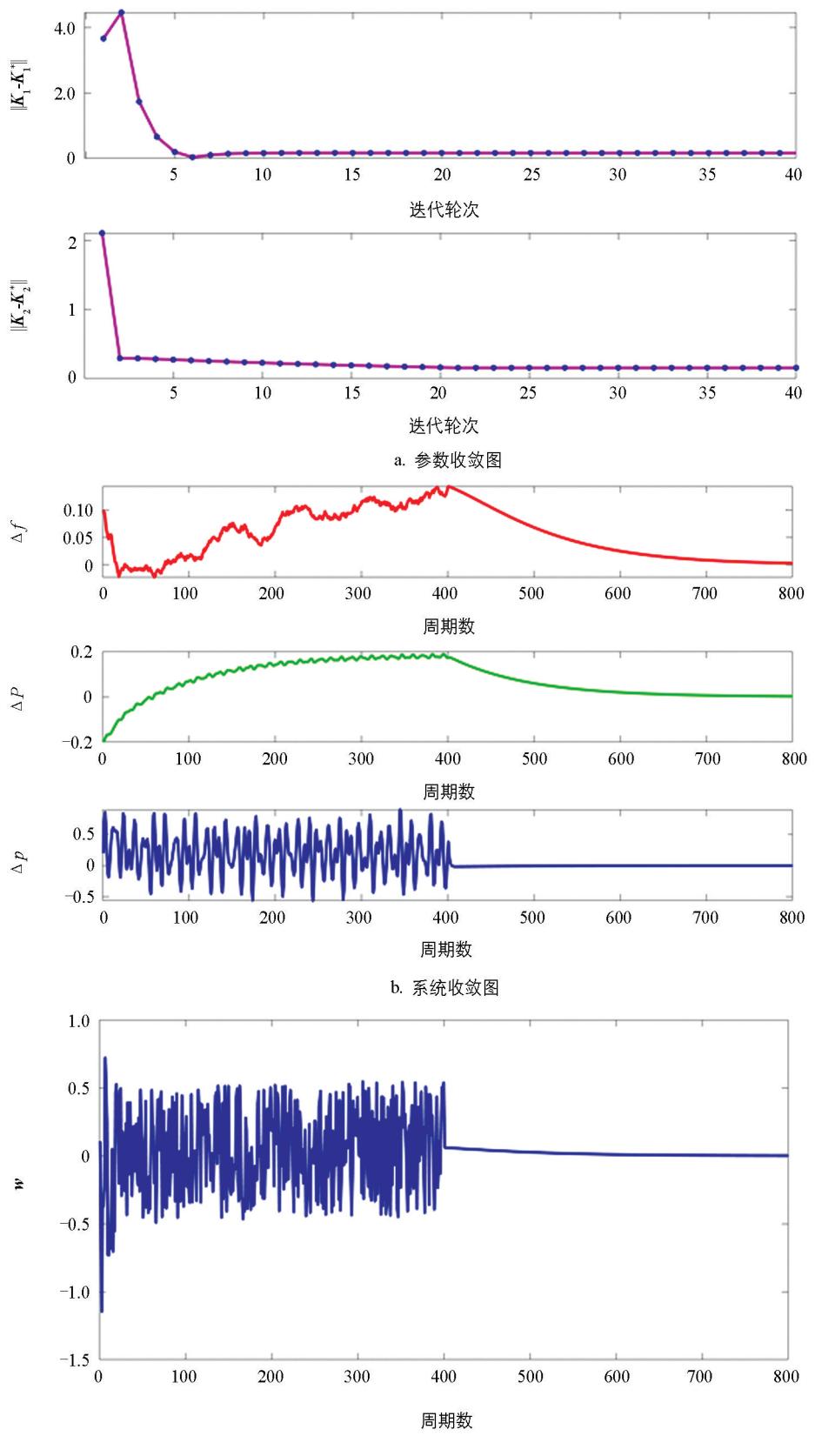


图 5 在情况 2 下常规 Q 学习算法的结果图

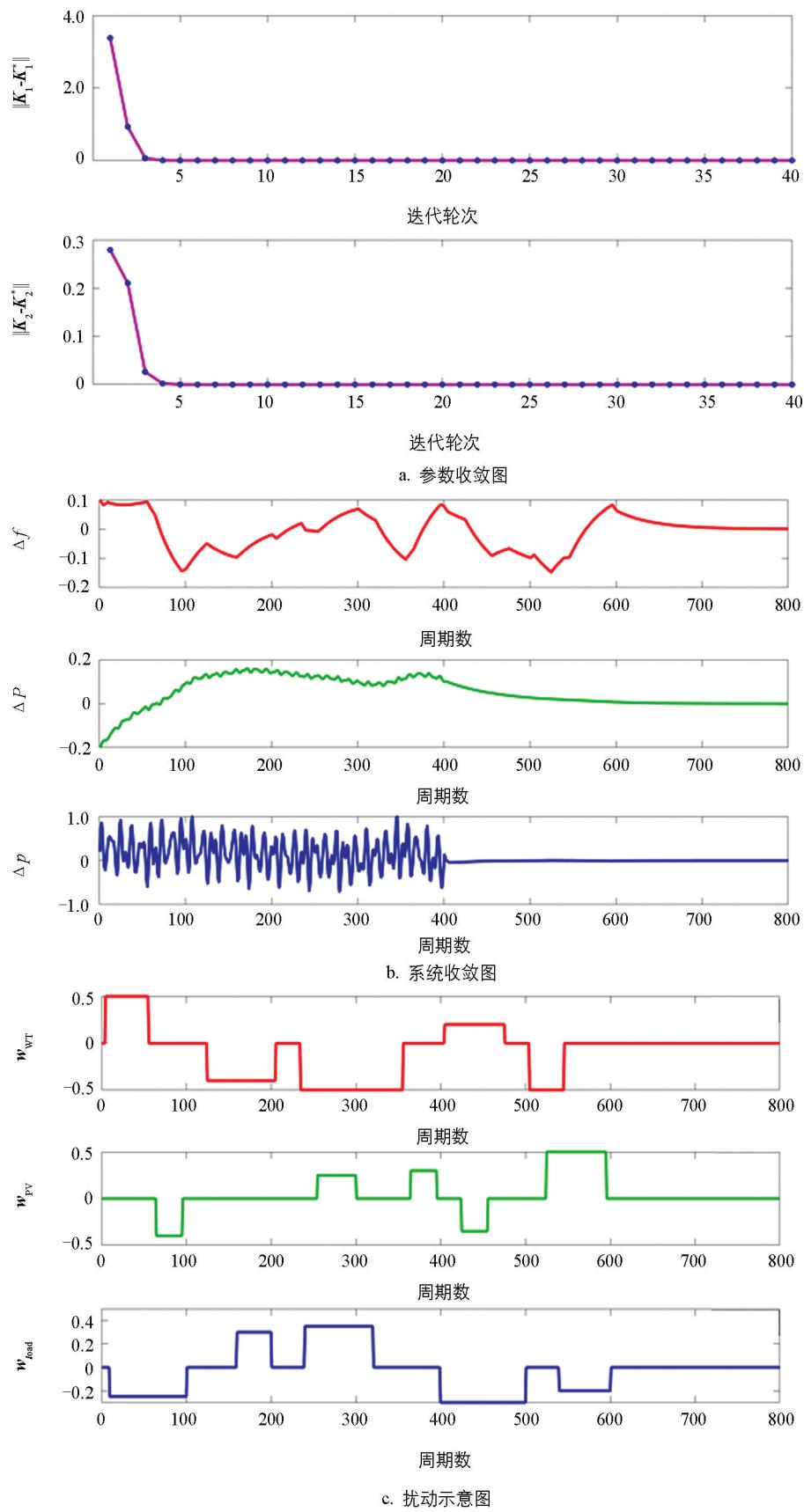


图 6 在情况 2 下改进 Q 学习算法的结果图

## 4 结论

本研究建立了在源荷扰动下的交直流微电网的系统模型,提出了基于 Q 学习的负荷频率控制方法,考虑到可能存在的源荷扰动情况,提出了基于离线策略 Q 学习的负荷频率控制策略,主要结论为:①通过对交直流微电网结构以及各分布式设备的建模与分析,得到了源荷扰动下的交直流微电网的基准系统模型。②本研究设计的方法用于解决交直流微电网的  $H_{\infty}$  控制问题,不仅无需系统动力学知识,而且可以使扰动以真实情况进行调节。③在源荷扰动下交直流微电网系统的频率能够达到期望的性能,保证其安全运行。后续研究将考虑交直流微电网频率和电压的关系并进行电压的稳定控制,以及将 Q 学习算法与隐私保护法相结合设计加密控制器后对系统的影响。

### 参考文献:

- [1] 成煜铃,李帅,孟高军,等. 基于微电网储能控制系统的电能质量综合治理研究 [J]. 可再生能源, 2023, 41(6): 810-817.
- [2] 夏栋,徐耀良,郭卫民,等. 基于实时功率判别的直流微电网协调控制策略研究 [J]. 可再生能源, 2017, 35(7): 1038-1046.
- [3] 范培潇,杨军,柯松,等. 基于改进 EDRL 的含 V2G 孤岛微电网频率综合控制策略 [J]. 电力系统自动化, 2023, 47(20): 23-32.
- [4] 岳应娟,凤林,蔡艳平,等. 交直流混合微电网运行控制技术 [J]. 科学技术与工程, 2022, 22(28): 12242-12252.
- [5] 刘奕彤,李正烁,樊森. 基于量测反馈的交直流混合微电网分布式指令追踪方法 [J]. 电力系统自动化, 2022, 46(24): 85-94.
- [6] 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.
- [7] LIUW, SHEN J, ZHANG S C, et al. Distributed Secondary Control Strategy Based on Q-learning and Pinning Control for Droop-controlled Microgrids [J]. Journal of Modern Power Systems and Clean Energy, 2022, 10(5): 1314-1325.
- [8] 范培潇,杨军,肖金星,等. 基于深度 Q 学习的含电动汽车孤岛微电网负荷频率控制策略 [J]. 电力建设, 2022, 43(4): 91-99.
- [9] YANGY J, WAN Y, ZHU J H, et al.  $H_{\infty}$  Tracking Control for Linear Discrete-time Systems: Model-free Q-learning Designs [J]. IEEE Control Systems Letters, 2021, 5(1): 175-180.
- [10] LINS W, CHU C C, TUNG C F. Data-driven Distributed Q-learning Droop Control for Frequency Synchronization and Voltage Restoration in Isolated AC Micro-grids [J]. IEEE Transactions on Industry Applications, 2023, 59(6): 7306-7317.
- [11] LIUH T, WU W C. Federated Reinforcement Learning for Decentralized Voltage Control in Distribution Networks [J]. IEEE Transactions on Smart Grid, 2022, 13(5): 3840-3843.
- [12] YANGY L, GUO Z S, XIONG H Y, et al. Data-driven Robust Control of Discrete-time Uncertain Linear Systems via Off-policy Reinforcement Learning [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(12): 3735-3747.
- [13] 赵斐然,游科友. 数据驱动的策略优化控制律设计最新研究综述 [J]. 中国科学: 信息科学, 2023, 53(6): 1027-1049.
- [14] 李臻,范家璐,姜艺,等. 一种基于 Off-Policy 的无模型输出数据反馈  $H_{\infty}$  控制方法 [J]. 自动化学报, 2021, 47(9): 2182-2193.
- [15] WANG D, HE H B, MU C X, et al. Intelligent Critic Control with Disturbance Attenuation for Affine Dynamics Including an Application to a Microgrid System [J]. IEEE Transactions on Industrial Electronics, 2017, 64(6): 4935-4944.