

DOI: 10.19783/j.cnki.pspc.221771

# 基于深度确定性策略梯度与模糊 PID 的直流微电网 VRB 储能系统就地功率控制

陆鹏<sup>1</sup>, 付华<sup>1</sup>, 卢万杰<sup>2</sup>

(1. 辽宁工程技术大学电气与控制工程学院, 辽宁 葫芦岛 125105;

2. 辽宁工程技术大学机械工程学院, 辽宁 阜新 123000)

**摘要:** 针对直流微电网全钒液流电池(vanadium redox flow battery, VRB)储能系统在实际运行时就地控制层中的功率控制器存在时滞、精度低及抗干扰能力差等问题, 提出了一种基于深度确定性策略梯度与模糊 PID 的功率跟踪控制策略。首先, 建立 VRB 的等效电路模型来描述功率传输特性, 并设计了由模糊 PID 与深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法组成的复合控制器。将模糊 PID 作为主控制器对功率环进行控制, DDPG 作为辅助控制器来补偿功率跟踪误差。然后, 设计了 VRB 储能系统就地功率跟踪控制器, 采用麻雀搜索算法(sparrow search algorithm, SSA)对 PID 参数和模糊规则进行优化, 并通过阶跃信号对优化后的系统输出响应进行测试。同时将分配指令功率与储能单元给定功率偏差作为数据集在 DDPG 中进行训练, 以提高主控制器的响应速度和抗干扰能力。最后, 通过在 3 种不同场景的算例下进行仿真, 验证了控制策略的有效性及稳定性。结果表明: 所提控制策略在电池充放电时, 能够快速跟踪到功率指令值; 实时跟踪时, 跟踪功率值与调度指令值偏差小于 $\pm 2\%$ ; 受到扰动时, 能准确修正功率偏差, 满足实际要求。

**关键词:** DDPG 算法; 模糊 PID; 全钒液流电池; 储能系统; 功率控制; 微电网

## Power control of the ground layer of a VRB energy storage system in a DC microgrid based on depth deterministic policy gradient and fuzzy PID

LU Peng<sup>1</sup>, FU Hua<sup>1</sup>, LU Wanjie<sup>2</sup>

(1. Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China;

2. School of Mechanical Engineering, Liaoning Technical University, Fuxin 123000, China)

**Abstract:** A power controller has problems of time delay, low control accuracy, and poor anti-interference ability in the actual operation of a vanadium redox flow battery (VRB) energy storage system in the ground layer of a DC microgrid. Thus a power cooperative control strategy based on the depth deterministic policy gradient and fuzzy PID is proposed. First, the equivalent circuit model of the VRB is established to describe the power transmission characteristics, and a composite controller composed of a fuzzy PID and a deep deterministic policy gradient (DDPG) algorithm is designed. In addition, the fuzzy PID is used as the main controller to control the power loop, and the DDPG is used as the auxiliary controller to compensate for the power tracking error. Then, the power tracking controller of the VRB energy storage system in the ground layer is designed, the PID parameters and fuzzy rules are optimized by the sparrow search algorithm (SSA) algorithm, and the output response of the optimized system is further tested by step signal. At the same time, the deviation between the command power of the distribution layer and the given power of the energy storage unit is used as a data set for training in the DDPG to improve the response speed anti-interference capability of the main controller. Finally, the effectiveness and stability of the control strategy are verified by simulation in three different scenarios. The results show that the proposed control strategy can quickly track the power command value during battery charging and discharging. The deviation between tracking power value and dispatching instruction value is less than  $\pm 2\%$  during real-time tracking; the power deviation can be corrected accurately to meet the actual requirements when the power controller is disturbed. This is in line with the actual requirements.

This work is supported by the National Natural Science Foundation of China (No. 51974151).

**Key words:** deep deterministic policy gradient algorithm; fuzzy PID; vanadium redox battery; energy storage system; power control; microgrid

**基金项目:** 国家自然科学基金项目资助(51974151); 辽宁省高等学校创新团队项目资助(LT2019007); 辽宁省重点实验室项目资助(LJZS003)

## 0 引言

由分布式能源、储能系统、柔性负载及控制器等组成的微电网系统是未来智能电网的重要组成部分。直流微电网具有控制结构简单、灵活性强、损耗小、不存在相位和无功功率的优点,受到许多学者的关注<sup>[1-2]</sup>。直流微电网系统能够稳定运行的前提可归结为直流母线电压的稳定<sup>[3]</sup>,而全钒液流电池(vanadium redox flow battery, VRB)储能系统对直流母线电压的稳定具有重要作用<sup>[4-5]</sup>。在VRB储能系统中,随着控制方式的多元、智能化,电池管理系统(battery management system, BMS)、变换器、传感器已成为深度融合的信息物理系统<sup>[6]</sup>,有效提升了储能系统运行控制的精度和灵活性。在VRB储能系统中,控制结构通过信息层、功率分配层和就地层之间的协调控制来完成系统的控制目标。因在直流微电网实际运行中,VRB储能系统需平抑受光伏发电、风力发电及负载等引起的功率波动问题,所以VRB储能系统就地控制层在收到功率分配层的指令时,控制器需满足快速准确地对功率分配层发出的功率需求指令进行跟踪,以确保整个直流微电网系统的功率平衡。

良好的控制策略可以使VRB储能系统就地层中功率控制器在受到外界扰动时,仍具有较好的自适应能力。在经典控制理论中常用的有PID控制器、模糊控制、神经网络及智能算法控制等。文献[7]针对风电机组最大功率跟踪控制,提出一种二阶滑模-PID的最大功率跟踪控制策略,有效提高了风电机组的发电效率和稳定性。文献[8]为了实现光储系统中的功率跟踪控制,提出了一种模糊自适应控制策略,从而提高微电网系统的调频能力。文献[9]提出了一种应用于直驱式波浪发电系统捕获功率的粒子群优化神经网络算法,克服了传统控制方法的不足,提高了预测精度。但上述控制算法在局部泛化、收敛速度上有一定的局限性,如PID控制器其参数不确定、非线性问题等缺点导致系统无法达到预期控制效果,神经网络在线辨识时计算复杂度高,响应速度慢,易局部收敛。

因此一些新方法被提出,主要分为两类:一类是对控制器自身进行改进,另一类是将经典控制器与智能算法融合形成新的复合控制器。文献[10]提出一种基于双重驱动的果蝇优化算法对PID参数进行优化,优化后的PID控制器响应速度快,稳定性高。文献[11]提出一种基于分数阶PID控制器参数优化整定方法,并采用灰狼优化算法获得最优的分数阶PID控制器参数,结果表明分数阶PID控制器比

传统整数阶PID控制器的控制性能更优。文献[12]提出了一种基于模糊PID算法的动力电池组直接式液体冷却策略,与传统PID策略相比,缩短了系统的调节时间、提高了抵抗电流扰动的能力,但模糊规则库存在一定的局限性,需进一步优化。文献[13]提出一种基于粒子群算法优化模糊PID隶属度函数和规则库的自适应多目标控制策略,以提高插电式混合动力汽车燃油经济性。文献[14]采用多经验池概率回放的双延迟深度确定性策略梯度算法对自动发电控制的发电功率指令分配算法进行优化,以减小区域控制偏差。文献[15]为了进一步提高深度确定性策略梯度(deep deterministic policy gradient, DDPG)方法中经验回放机制的效率,提出分类经验回放方法,并采用两种方式对经验样本进行分类。文献[16]提出一种基于改进多智能体深度确定性策略梯度(multi-agent deep deterministic policy gradient, MA-DDPG)算法的多微网负荷频率控制策略,设计了多智能体控制器的状态、动作空间与奖励函数,验证了DDPG控制器在强随机性扰动、系统网络拓扑参数改变等复杂运行工况下能够快速稳定系统频率。但上述控制策略仅从PID参数优化及DDPG自身进行改进,没有将两种控制器进行结合组成新的复合控制器,对其改进控制精度和系统响应时间。

综上,在VRB储能系统就地控制层中,其控制目标就是快速对功率分配层发出的功率需求指令进行跟踪。因国家标准规定其电化学储能系统中跟踪功率给定值的响应时间应小于2s,且功率偏差值始终控制在额定功率的 $\pm 2\%$ <sup>[17]</sup>,所以在保证系统对响应时间及功率控制精度要求的前提下,需要提高控制器的控制精度与响应速度。但VRB储能系统中因功率指令误差和外部功率扰动等因素导致功率控制系统中存在频繁的干扰以及大滞后,很难对VRB储能系统中功率进行精确跟踪。对此,本文提出了一种基于DDPG与模糊PID组成的复合控制策略,在VRB储能系统就地层控制结构中将模糊PID作为主控制器,并采用麻雀搜索算法(sparrow search algorithm, SSA)对模糊PID控制器中的参数进行优化,有利于算法在实际应用中的可操作性。将DDPG作为辅助控制器,利用DDPG算法在线训练学习,提高其主控制器的响应速度和抗干扰能力,通过主控制器与辅助控制器联合控制,实现当功率分配层发出功率需求指令时,控制器能够快速准确地跟踪功率指令值。

## 1 VRB储能系统就地层结构与建模

### 1.1 VRB储能系统分层控制结构

在VRB储能系统中,控制结构通过信息层、

功率分配层和就地层之间的协调控制来完成系统的控制目标。信息层主要用于信息数据采集、管理决策,反馈给直流微电网能量调度管理系统。功率分配层主要根据上层信息层调度指令完成各个 VRB 储能单元的功率分配,在各约束条件下,设置 VRB 储能系统总优化目标函数,通过协调控制算法完成 VRB 储能单元的功率给定值。就地层主要在所设计的功率控制器的基础上,去实现 VRB 储能系统中电池单元能够快速准确地跟踪分配层下发的功率给定值,保证电池组出力可靠、安全稳定运行。VRB 储能系统分层控制结构如图 1 所示。

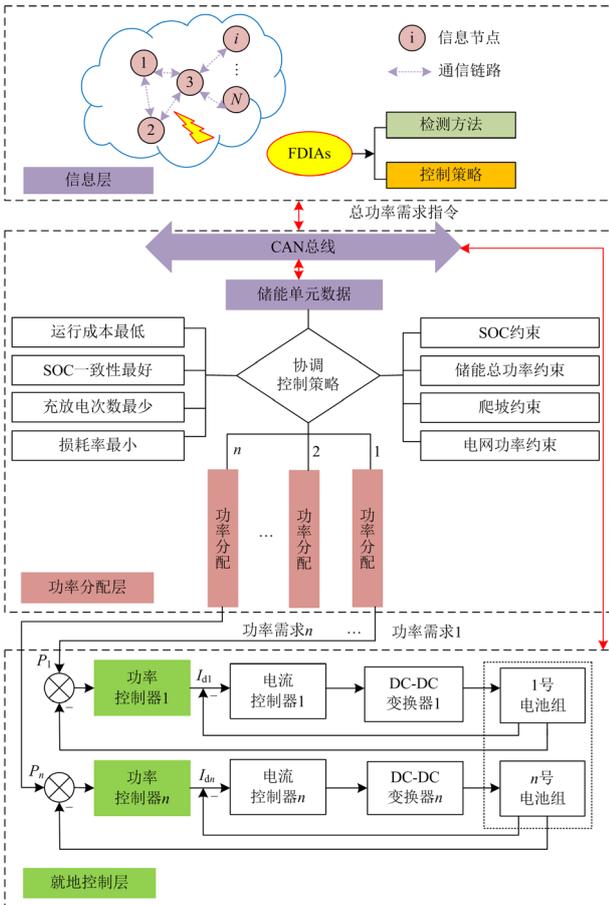


图 1 VRB 储能系统控制结构

Fig. 1 VRB energy storage system control structure

### 1.2 VRB 储能系统内部结构及等效电路模型

VRB 储能系统结构如图 2 所示,内部结构由电极材料、液钒储罐、离子交换膜、管路及循环泵等组成。

为了进一步研究 VRB 储能系统就地层的功率跟踪控制,首先建立了 VRB 的等效电路模型来描述其功率动态特性。本文采用 VRB 的一阶 RC 等效电路模型进行研究。电路结构如图 3 所示<sup>[18]</sup>。图中:

$U_d$  为 VRB 的端电压;  $I_d$  为 VRB 的放电电流;  $V_s$  为开路电压;  $I_p$  为泵损电流;  $R_f$  为寄生电阻;  $R_a$  为反应电阻;  $R_b$  为欧姆内阻;  $I_c$  为流过电容的电流;  $U_c$  为电容两端电压;  $C_e$  为极化电容。

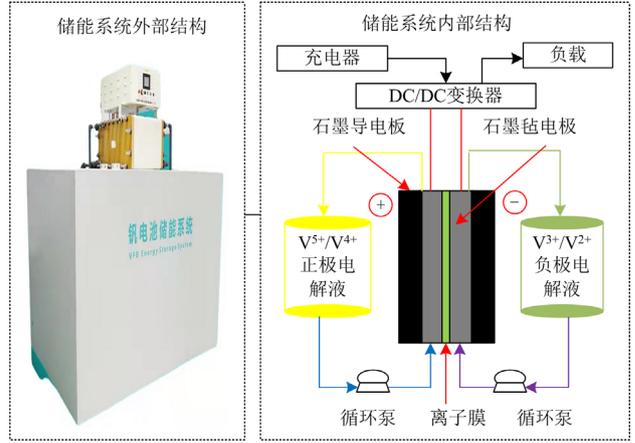


图 2 VRB 储能系统结构

Fig. 2 Structure of VRB energy storage system

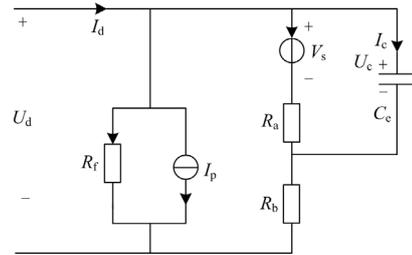


图 3 VRB 等效电路模型

Fig. 3 VRB equivalent circuit model

下一步根据 VRB 的等效电路模型建立其状态方程,其中将  $U_c$  和 SOC 作为状态变量,  $I_d$  为输入标量,  $U_d$  为输出变量。

利用 KCL 和 KVL 等效电路原理,建立运算分析电路得出其传递函数,电路中参数均为复变量  $s$  的函数,通过戴维南等效电路得出输入与输出的关系如下:

$$U_d(s) = E_q(s) + I_{eq}Z_{eq}(s) \quad (1)$$

$$E_q(s) = K(s)V_s(s) \quad (2)$$

$$I_{eq}(s) = I_d(s) - I_p(s) \quad (3)$$

$$Z_{eq}(s) = \frac{k_1(s + z_1)}{s + p_1} \quad (4)$$

$$K(s) = \frac{k_2}{s + p_1} \quad (5)$$

公式中相关参数如式(6)所示。

$$\begin{cases} k_1 = \frac{R_b R_f}{R_b + R_f} \\ k_2 = \frac{R_f}{R_a C_1 (R_b + R_f)} \\ z_1 = \frac{R_a + R_b}{R_a R_b C_1} \\ p_1 = \frac{R_a + R_b + R_f}{R_a C_1 (R_b + R_f)} \end{cases} \quad (6)$$

由式(1)~式(6)可推导出等效传递函数 $Y(s)$ , 如式(7)所示。

$$Y(s) = \frac{U_d(s) - E_{eq}(s)}{I_{eq}(s)} \quad (7)$$

式中: 所涉及的字母物理量含义与复变前的一致,  $E_{eq}(s)$  为等效电压;  $I_{eq}(s)$  为等效电流;  $Z_{eq}(s)$  为等效阻抗;  $k_1$ 、 $k_2$ 、 $z_1$ 、 $p_1$  为相关系数;  $C_1$  为电容值。

## 2 基于 DDPG-模糊 PID 复合控制器的设计

### 2.1 模糊 PID 主控制器

为了更加精确地跟踪 VRB 储能系统的功率值, 将采用模糊自适应理论与 PID 控制理论组成的模糊自适应 PID 控制器作为主控制器实现功率的跟踪。模糊 PID 控制器是利用模糊规则将 VRB 储能系统功率给定值与测量反馈值的偏差  $e(t)$  和偏差信号变化率  $e_c(t)$  的值作为输入量, PID 系数  $k_p$ 、 $k_i$ 、 $k_d$  作为输出量并作为 PID 控制的修正因子, 从而实现控制性能的最优化, 其控制原理如图 4 所示。

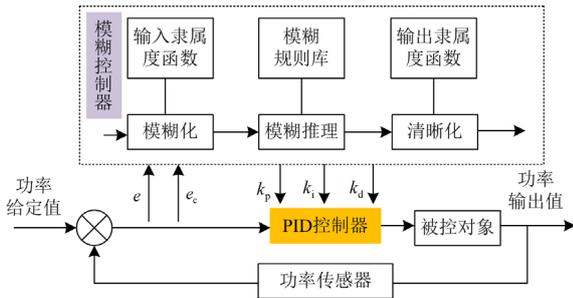


图 4 模糊 PID 控制原理图

Fig. 4 Schematic diagram of fuzzy PID control

### 2.2 PID 参数优化

因模糊 PID 隶属函数与模糊规则的合理性将影响系统的控制性能, 文中采用麻雀搜索算法对模糊 PID 控制器参数进行整定, 以使得 VRB 储能系统达到更好的功率跟踪效果。

文献[19]提出 SSA 算法, 该算法比较新颖, 具有寻优能力强、收敛速度快的优点。其主要思想是通过模仿麻雀觅食和反捕食行为进行局部和全局搜

寻, 而麻雀群觅食过程遵循发现者-跟随者模型, 同时还叠加了侦查预警机制。其发现者位置更新如式(8)所示。

$$X_{i,j}(t+1) = \begin{cases} X_{i,j}(t) \cdot \exp(-\frac{i}{\alpha \cdot t}), R < S_v \\ X_{i,j}(t) + Q \cdot L, R \geq S_v \end{cases} \quad (8)$$

式中:  $X_{i,j}$  为种群第  $t$  代中第  $i$  个个体的第  $j$  维位置信息值;  $t$  为当前迭代次数;  $\alpha$  为[0,1]的随机数;  $T$  为最大迭代次数;  $Q$  为服从正态分布的随机数;  $R$  为[0,1]的预警值;  $S_v$  为[0.5,1]的安全值; 当  $R < S_v$  时, 发现者可以进行大规模搜索; 当  $R \geq S_v$  时, 侦查者发现捕食者, 立即发出预警信号。

跟随者位置更新如式(9)所示。

$$X_{i,j}(t+1) = \begin{cases} Q \cdot \exp(\frac{x_{\text{worst}}(t) - X_{i,j}(t)}{i^2}) \\ X_p(t+1) + |X_{i,j}(t) - X_p(t+1)| \cdot A^+ \cdot L \end{cases} \quad (9)$$

式中:  $X_p(t+1)$  为当前发现者中适应度值最优的位置;  $x_{\text{worst}}(t)$  为麻雀当前全局适应度最差的位置; 其中  $A^+ = A^T(AA^T)^{-1}$ , 表示一个与麻雀个体同维度的向量矩阵。

侦查者位置更新如式(10)所示。

$$X_{i,j}(t+1) = \begin{cases} x_{\text{best}}(t) + \beta \cdot |X_{i,j}(t) - x_{\text{best}}(t)| \\ X_{i,j}(t) + K \cdot \frac{|X_{i,j}(t) - x_{\text{worst}}(t)|}{(f_i - f_w) + \varepsilon} \end{cases} \quad (10)$$

式中:  $x_{\text{best}}(t)$  为当前的全局最优位置;  $\beta$  为步长控制参数, 服从正态分布;  $K$  为[-1,1]随机数;  $f_i$  为当前麻雀个体的适应度值;  $f_i$  和  $f_w$  分别为当前全局最佳和最差的适应度值;  $\varepsilon$  是较小的常数。

其 PID 参数优化具体步骤如下:

1) 求解模糊时, 先根据不同逻辑的模糊子集、隶属度函数、模糊规则表求出系数  $k_p$ 、 $k_i$ 、 $k_d$  的隶属度, 然后反模糊化求出  $k_p$ 、 $k_i$ 、 $k_d$ , 再将  $k_p$ 、 $k_i$ 、 $k_d$  参数代入 PID 控制器中求解。

2) 对模糊域进行编码, 因储能系统在充放电过程中的功率偏差不能超过额定功率的  $\pm 2\%$ , 故偏差  $e(t)$  的论域定义为[-12, 12], VRB 储能系统功率的变化率  $e_c(t)$  设置为[-600, 600],  $k_p$ 、 $k_i$ 、 $k_d$  实际论域设置为[-1, 1]、[-0.06, 0.06]、[-0.3, 0.3]。

3) 根据模糊规则系统模糊子集设定为负大 NB、负中 NM、负小 NS、零 ZO、正小 PS、正中 PM、正大 PB 共 7 个语言变量值。

4) 麻雀算法参数设置如下, 优化变量的维度为 3, 将 PID 控制的 3 个参数作为麻雀的位置向量, 种群规模为 100, 安全阈值  $S_v = 0.8$ , 发现者  $P_{\text{find}} = 0.18$ , 警戒者  $S_d = 0.1$ ,  $k_p$ 、 $k_i$ 、 $k_d$  三个参数的搜索范围为  $[-1, 1]$ , 步长为 0.01, 迭代次数为 100。基于 SSA 算法优化 PID 参数流程如图 5 所示。

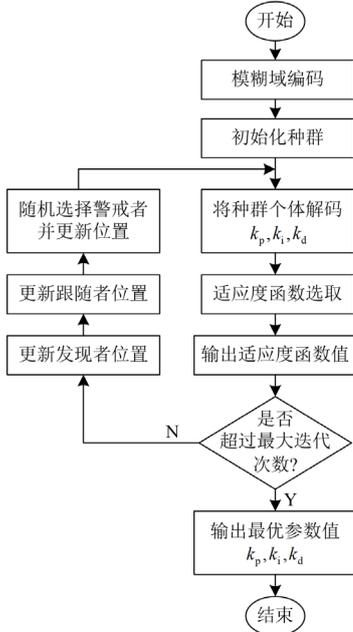


图 5 SSA 算法优化 PID 流程

Fig. 5 Process of PID optimization using SSA algorithm

### 2.3 模糊 PID 控制器有效性验证

为验证整定后 PID 参数的有效性, 采用 SSA 算法对模糊 PID 中的参数优化, 并与优化前的模糊 PID 对比其控制器的有效性及寻优能力, 通过阶跃信号对系统输出响应进行测试。通过 SSA 仿真迭代后, 迭代过程图如图 6 所示, 其  $k_p$ 、 $k_i$ 、 $k_d$  参数的隶属度函数如图 7 所示, 优化后的模糊 PID 系统输出如图 8 所示。

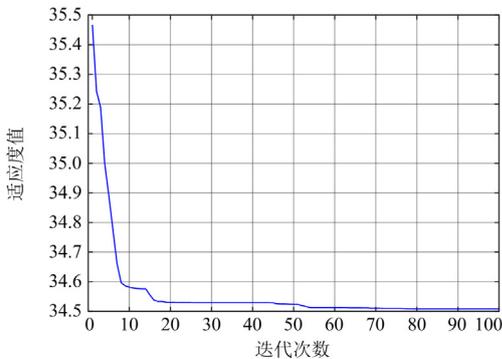


图 6 麻雀搜索算法迭代过程曲线

Fig. 6 Iterative process curve of SSA

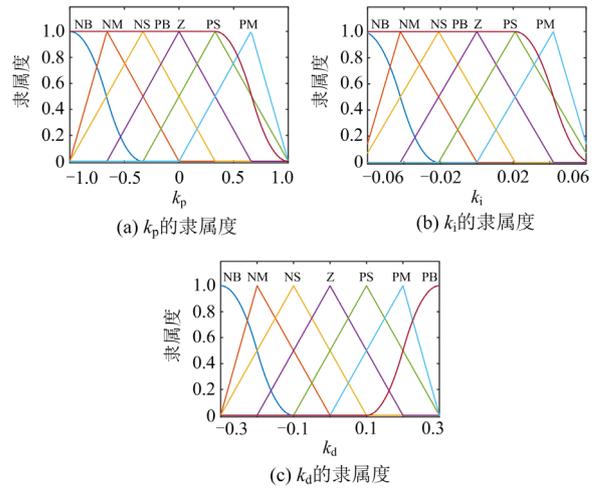


图 7 模糊 PID 参数的隶属度函数

Fig. 7 Membership function of fuzzy PID parameters

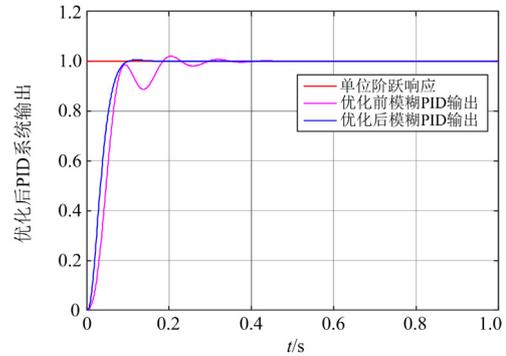


图 8 优化后的 PID 参数系统输出

Fig. 8 Output of optimized PID parameter system

通过图 8 可知, 优化后的模糊 PID 较优化前响应速度更快、稳定性更好, 从而验证了模糊 PID 主控制器的有效性。

### 2.4 DDPG 辅控制器

在 VRB 储能系统就地控制层中, 为了实现功率跟踪快速响应、减小控制器的超调量与误差, 文中采用 DDPG 算法作为辅助控制器对其主控制器进行补偿, 其框架采用动作-评价(Actor-Critic)网络。Actor 网络是策略网络, 负责输出控制指令, Critic 网络为评价网络, 其目标是选取最佳动作值。其中评价网络通过损失函数为最小化目标进行网络参数更新, 定义如式(11)和式(12)<sup>[20-22]</sup>所示。

$$J(\theta^Q) = \frac{1}{n} \sum_{k=1}^n \omega_k (y_i - Q(s_i, a_i / \theta^Q))^2 \quad (11)$$

$$y_i = r_{t+1} + \gamma Q(s_{t+1}, a_{t+1} / \theta^Q) \quad (12)$$

式中:  $y_i$  为更新后的目标动作价值;  $n$  为样本量;  $\omega_k$  为样本权重;  $Q$  为在样本状态量  $s_i$  和动作  $a_i$  下的目

标价值函数;  $r_{t+1}$  为奖励参数;  $\gamma$  为折扣因子。

DDPG 算法在训练过程中数据不断更新, 策略网络更新后为现实网络, 评价网络更新后为目标网络, 更新后的网络定义为

$$\theta^Q = \tau\theta^Q + (1-\tau)\theta^{Q'} \quad (13)$$

$$\theta^\mu = \tau\theta^\mu + (1-\tau)\theta^{\mu'} \quad (14)$$

式中:  $\theta^Q$  为评价网络参数;  $\theta^{Q'}$  为目标网络参数;  $\theta^\mu$  为策略网络参数;  $\theta^{\mu'}$  为现实网络参数;  $\tau$  为学习效率; 函数  $\mu$  为最优动作策略。

DDPG 训练过程中执行动作  $a_t$  定义为

$$a_t = \mu(s_t | \theta^\mu) + n_t \quad (15)$$

式中,  $n_t$  为自适应噪声参数。

DDPG 算法通过 VRB 储能系统中实时功率值获取样本参数  $(s_t, a_t, r_{t+1}, s_{t+1})$ , 其中  $s_t$  为  $t$  时刻储能系统的功率需求值,  $s_{t+1}$  为  $t+1$  时刻在执行动作  $a_t$  后的功率跟踪值。基于 DDPG 控制器训练过程如图 9 所示。

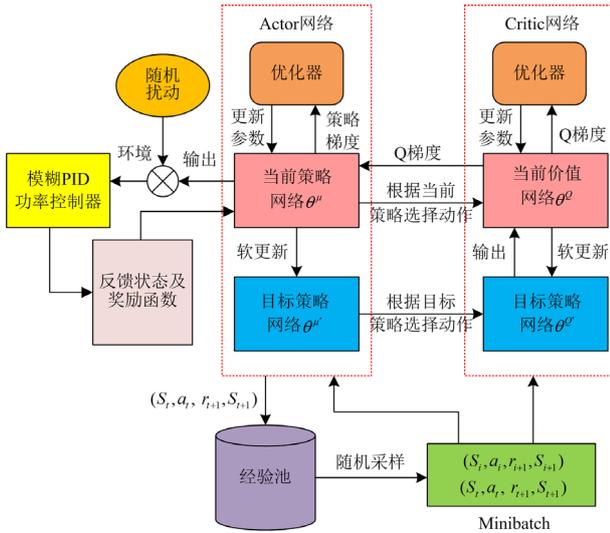


图 9 DDPG 控制器训练过程

Fig. 9 Training process of DDPG controller

### 3 VRB 储能系统就地功率控制策略

#### 3.1 功率调节器设计

在 VRB 储能系统中功率跟踪控制时, 功率调节器(automatic power regulator, APR)的设计将采用在强化学习中的 DDPG 算法与模糊 PID 组成的复合控制器, 其中将模糊 PID 控制器作为主控制器, 当受到外部扰动时, 利用 DDPG 辅助控制器在线学习训练, 输出最优控制, 其 APR 的控制结构如图 10 所示。

#### 3.2 VRB 储能系统就地功率控制实现

模糊 PID 控制算法具有简单、稳定、不需要精

确的数学模型等优点, 被广泛应用于解决储能系统中功率控制问题<sup>[23-24]</sup>。但传统模糊 PID 在非线性系统、时变不确定性系统中, 容易出现计时精度低、抗干扰能力差等问题, 且缺乏一定的在线学习能力, 控制器不能对设置过的 PID 参数进行更改, 对于不确定非线性控制系统中, 控制效果还需进一步改善提升。因此本节首先将 DDPG 算法作为辅助控制器, 对模糊 PID 主控制器进行辅助, 进一步减小误差和超调量, 提升控制器的抗干扰能力, 其次对 VRB 储能系统中功率给定值进行跟踪控制, 以满足实际要求。其 DDPG-模糊 PID 控制器的控制结构如图 11 所示。

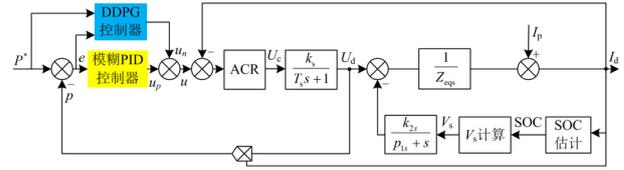


图 10 功率调节器等效结构

Fig. 10 Equivalent structure of APR

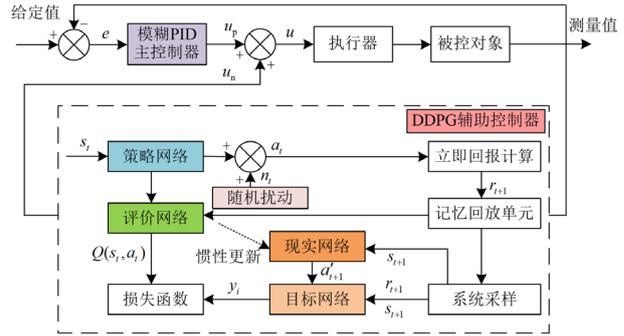


图 11 DDPG-模糊 PID 控制结构

Fig. 11 DDPG-fuzzy PID control structure

在 VRB 储能系统中, 首先将模糊控制器通过所设计的功率控制环进行充放电功率调节。模糊控制器的输入量为系统当前功率分配层指令功率  $P_{ref}$ , 其功率偏差变化率为  $dP/dt$ , 输出量为储能电池单元功率跟踪值  $P_0$ 。本文所提 VRB 储能系统由 6 个 100 kW 的储能电池单元组成, 故单个电池单元的指令功率  $P_{ref}$  作为输入量, 其范围在  $[-100, 100]$ , 整个 VRB 储能系统的功率范围在  $[-600, 600]$ ; 因国家标准中规定储能系统在充放电过程中的功率偏差不能超过额定功率的  $\pm 2\%$ , 故单个电池单元的偏差率论域定义为  $[-2, 2]$ , 整个储能系统的偏差率论域定义为  $[-12, 12]$ ; 储能电池单元的输出功率  $P_0$  与分配层指令功率  $P_{ref}$  变换范围一致。

针对 VRB 储能系统功率跟踪控制时, 偏差功率值通过传感器回采的实时功率跟踪值与设定指令

功率值进行获取, 然后按照模糊规则经过推理得出模糊控制量, 再通过 SSA 算法对模糊 PID 的隶属度函数和模糊规则进行优化, 具体优化过程及参数设置已在 2.2 节中给出, 其所建立模糊规则库如表 1 所示。

表 1 模糊 PID 控制器规则表

Table 1 Fuzzy PID controller rule

功率偏差变化率 $dP/dt$	VRB 储能系统功率跟踪值 $P_b$						
	$P_b$	$P_b$	$P_b$	$P_b$	$P_b$	$P_b$	$P_b$
NB	NB	NB	NB	NM	NM	NS	ZO
NM	NB	NB	NM	NM	NS	ZO	PS
NS	NB	NM	NM	NS	ZO	PS	PM
ZO	NM	NM	NS	ZO	PS	PM	PM
PS	NM	NS	ZO	PS	PM	PM	PB
PM	NS	ZO	PS	PM	PM	PB	PB
PB	ZO	PS	PM	PM	PB	PB	PB

根据模糊规则子集, 将系统实时输入量指令功率  $P_{ref}$ 、功率偏差变化率  $dP/dt$  以及输出功率跟踪值  $P_b$  设定为负大 NB、负中 NM、负小 NS、零 ZO、正小 PS、正中 PM、正大 PB 共 7 个语言变量值。其中当跟踪值  $P_{ref}$  为零时, 储能系统既不充电也不放电。其输入与输出变量的隶属度函数如图 12 所示。

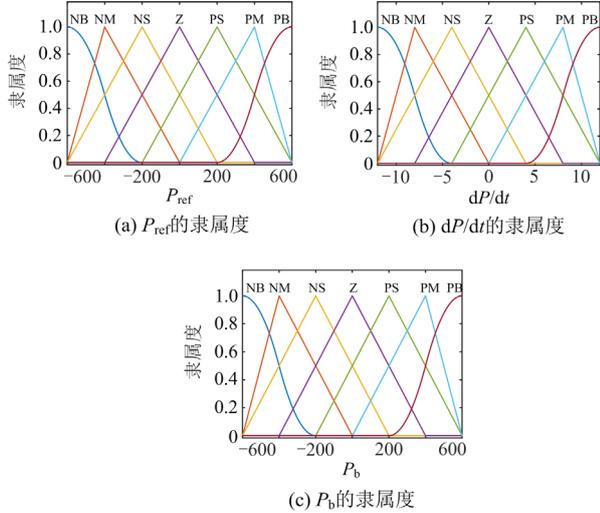


图 12 输入与输出变量的隶属度函数

Fig. 12 Membership function of input and output variables

在 VRB 储能系统功率控制过程中, 模糊 PID 主控制器和 DDPG 辅助控制器会同时输出执行器的控制量, 在系统功率反馈值发生较大偏差或受到外部随机因素扰动时, 此时为了使给定功率跟踪值能够实现响应速度快、控制精准, 则在模糊 PID 的基础上采用 DDPG 控制器对功率偏差进行修正, 其模糊 PID 主控制器与相邻 DDPG 组成的控制过程如图 13 所示。



图 13 DDPG 与模糊 PID 联合过程

Fig. 13 Combined process of DDPG and fuzzy PID

VRB 储能系统中分配层指令功率与跟踪电池单元给定功率偏差集  $\Delta P = [\Delta P_1, \Delta P_2, \dots, \Delta P_i]$  以及各电池单元给定值的输出动作集  $P_b = [P_{b1}, P_{b2}, \dots, P_{bi}]$  将作为输入反馈给模糊 PID 控制中。在 VRB 储能系统中 DDPG 控制器通过协调控制层和功率跟踪控制层去实现, 其协调控制层将根据功率偏差集  $\Delta P$ 、储能系统实时功率上下限向功率控制层中的各个电池单元输送实时控制信号, 在主控制器功率偏差较大和受到外界扰动时, 能够快速、准确地实现 VRB 储能系统电池单元功率的跟踪。

模糊 PID 主控制器中的状态输出变量为实时功率偏差变化  $\Delta P$ 、VRB 储能系统分配层指令功率  $P_{ref}$  以及各电池单元功率跟踪值  $P_b$ 。故在 DDPG 算法中, 定义状态空间  $S$  如下:

$$S = [\Delta P(t), P_{ref}(t), P_b(t)] \quad (16)$$

模糊 PID-DDPG 训练过程中联合动作集  $A$ , 即控制器的输出, 通过指令发送给各储能电池单元集, 故定义动作空间如下:

$$A = [P_{b1}(t), P_{b2}(t), \dots, P_{bi}(t)] \quad (17)$$

通过上述状态空间与动作空间定义后, 即可确定 DDPG 控制器中网络的输入输出变量的维数, 即输入为 VRB 储能系统功率分配层状态信息  $S$ , 其中 Actor 网络是策略网络, 负责输出控制指令, Critic 网络为评价网络, 其目标是选取最佳动作值。

因国家标准中规定电化学储能系统在充放电过程中的功率偏差不能超过额定功率的  $\pm 2\%$ , 故本文以额定功率 600 kW 的 VRB 储能系统为例, 其单节储能单元电池额定功率为 100 kW, 在 VRB 储能系统实际运行中, 考虑其一定的调节死区, 故设计实时功率偏差  $\Delta P$  的集合区间为  $(-\infty, -2)$ 、 $[-2, -1.5]$ 、 $[-1.5, -1.0]$ 、 $[-1.0, -0.5]$ 、 $[-0.5, -0.1]$ 、 $[-0.1, 0.1]$ 、 $(0.1, 0.5)$ 、 $(0.5, 1)$ 、 $(1, 1.5)$ 、 $(1.5, 2)$ 、 $(2, +\infty)$ , 单位为 kW。故奖励函数如式(18)所示。

$$r_i(k) = \begin{cases} 0 \\ -\mu_1 |\Delta P_1| - \delta_1 (|\Delta P_2| + \dots + |\Delta P_i|) \\ -\mu_2 |\Delta P_1| - \delta_2 (|\Delta P_2| + \dots + |\Delta P_i|) \\ -\mu_3 |\Delta P_1| - \delta_3 (|\Delta P_2| + \dots + |\Delta P_i|) \\ -\mu_4 |\Delta P_1| - \delta_4 (|\Delta P_2| + \dots + |\Delta P_i|) \\ -\mu_5 |\Delta P_1| - \delta_5 (|\Delta P_2| + \dots + |\Delta P_i|) \end{cases} \quad (18)$$

式中:  $\mu_1$ 、 $\mu_2$ 、 $\mu_3$ 、 $\mu_4$  与  $\mu_5$  为 VRB 储能系统中功率偏差的各控制区域奖励函数所对应的权值;  $\delta_1$ 、 $\delta_2$ 、 $\delta_3$ 、 $\delta_4$  与  $\delta_5$  为针对功率随机扰动的各控制区域奖励函数所对应的权值, 而奖励值会影响 DDPG 的学习能力。

通过大量仿真实验, 在 DDPG 控制器预训练时, 其各参数设置如下, 奖励函数中各权值分别取:  $\mu_1 = 5$ 、 $\mu_2 = 10$ 、 $\mu_3 = 30$ 、 $\mu_4 = 40$  与  $\mu_5 = 100$ ;  $\delta_1 = 5$ 、 $\delta_2 = 10$ 、 $\delta_3 = 15$ 、 $\delta_4 = 25$  与  $\delta_5 = 40$ ; Actor 学习率为 0.001, Critic 学习率为 0.02, 奖励因子为 0.9, 记忆回放单元存放容量最大值为 10 000, 步长设置为 0.01 s, 迭代次数为 500; 并通过加入随机扰动函数模拟功率扰动对 DDPG 控制器进行训练的效果, 同时根据 VRB 储能系统跟踪功率与指令功率偏差数据设置一组随机变化的功率偏差约束函数, 并截取上述各函数在某次迭代中的过程, 其预训练阶段功率扰动随机函数与功率约束函数的输入如图 14 所示。

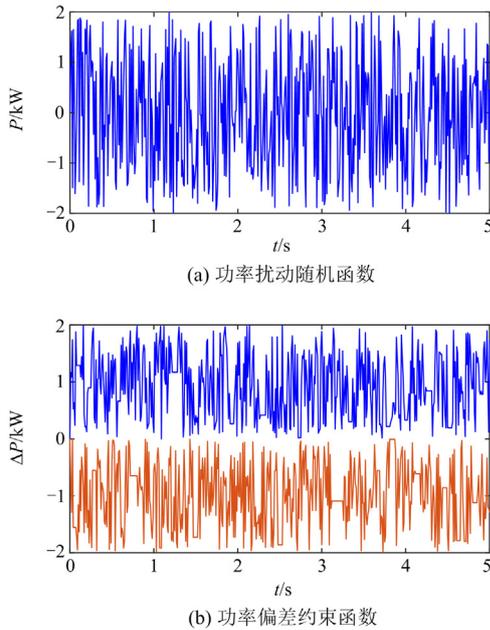


图 14 预训练阶段输入函数

Fig. 14 Pre training stage input function

DDPG 控制器在预训练完成后, 以模糊 PID 主控制的输入与反馈偏差为例进行训练, 其迭代 500 次的训练过程如图 15 所示。

根据图 15 可知, 在完成大约 300 次的迭代后, DDPG 控制器基本收敛, 并在训练 460 次时, 自动停止了训练, 此时模糊 PID-DDPG 联合动作集  $A$  的平均奖励值为 -10.8635, 最终实时奖励值为 -1.0809, 具有较高的在线学习能力。

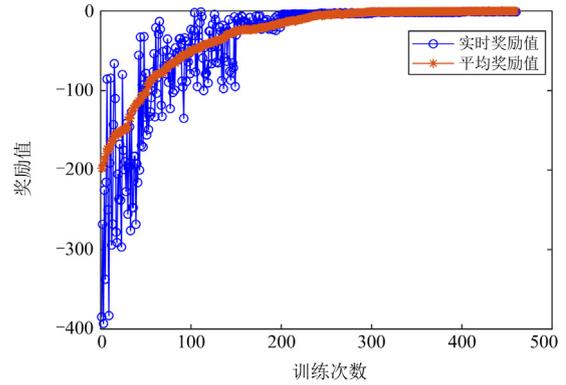


图 15 训练过程中奖励函数趋势

Fig. 15 Trend of reward function during training

综上, 可见 DDPG 控制器通过训练后的收敛性较好, 对主控制器进行辅助, 能进一步提升其响应速度和抗干扰能力, 最后对 VRB 储能系统中功率给定值进行跟踪控制, 以满足实际要求。

## 4 算例分析

为了进一步验证所提 DDGP 算法与模糊 PID 组成的复合控制器在实际应用中的响应速度和抗干扰能力, 本文通过 3 个场景进行验证:

1) 由于 VRB 储能系统由多个 VRB 单元联接组成, 仿真中忽略单个电池的性能差异, 通过跟踪单个电池充放电过程验证其控制器的响应速度;

2) 为了保证直流微电网中功率的平衡, VRB 储能系统实时充放电, 故对 VRB 储能系统实时功率变化进行跟踪, 验证所提复合控制器的精确度及动态特性;

3) 假设由 VRB 储能系统组成的直流微电网初始状态为稳态, 因风电与负载功率随时间变化时, 会对分配层功率指令值产生一定的随机功率偏差扰动, 故通过观察受到扰动后控制器对偏差功率的控制效果来验证所提控制器的抗干扰性及准确性。

通过在 Matlab/Simulink 中搭建仿真模型, 仿真中模糊 PID 控制器的相关参数以 2.2 节中 SSA 算法整定后的 PID 参数进行设置, DDPG 控制器相关参数以 3.2 节中训练后的参数进行设置; VRB 储能电池单元相关参数设置如表 2 所示。

### 4.1 场景 1: 电池充放电瞬间功率跟踪

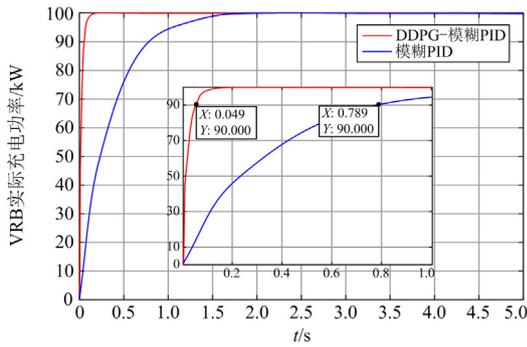
通过跟踪 VRB 储能系统单个电池充放电过程中功率控制效果进行验证。

1) VRB 储能系统单元以额定功率 100 kW 进行充电 5 s, 对比分析复合控制器 DDPG-模糊 PID 与模糊 PID 控制器的跟踪效果, 充电过程中电池的初始 SOC 值为 0.1, 其控制器的实际跟踪结果如图 16 所示。

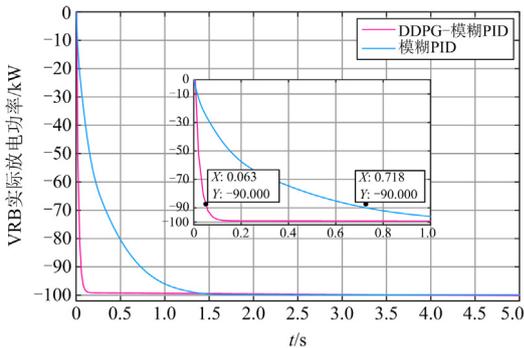
表 2 VRB 储能单元相关参数

Table 2 Relevant parameters of VRB energy storage unit

VRB 参数	数值
额定功率 $P/kW$	100
反应内阻 $R_s/\Omega$	0.22
欧姆电阻 $R_b/\Omega$	0.13
寄生电阻 $R_f/\Omega$	45.33
极化电容 $C_p/F$	0.02
变换器开关周期 $T_s/s$	0.000 125
初始 SOC 值	0.1



(a) VRB 充电功率曲线



(b) VRB 放电功率曲线

图 16 VRB 充放电功率曲线

Fig. 16 Curves of VRB charge and discharge power

对图 16(a)分析可知,在 VRB 储能单元收到充电信号后,通过 DDPG-模糊 PID 复合控制器快速跟踪到了给定值,充电响应时间为 0.049 s,而模糊 PID 充电响应时间为 0.789 s。

2) 同上以额定功率 100 kW 进行放电 5 s,对比分析复合控制器 DDPG-模糊 PID 与模糊 PID 控制器的跟踪效果,放电过程中电池的初始 SOC 值为 0.9,其控制器的实际跟踪结果如图 16(b)所示。

对图 16(b)分析可知,在 VRB 储能单元收到放电信号后,DDPG-模糊 PID 复合控制器快速跟踪到了给定值,放电响应时间为 0.063 s,而模糊 PID 放电响应时间为 0.718 s。

通过上述对 VRB 储能电池单元充放电瞬时功率跟踪分析可知,在充电控制过程中采用 DDPG-模糊 PID 复合控制器比模糊 PID 的响应时间快 0.74 s;在放电控制过程中采用 DDPG-模糊 PID 复合控制器比模糊 PID 的响应时间快 0.655 s,其功率跟踪响应时间符合国家标准 GBT36547-2018《电化学储能系统接入电网技术规定》中的响应时间,验证了所提控制器的响应速度满足控制要求。

#### 4.2 场景 2: VRB 储能系统实时需求功率跟踪

为了验证 VRB 储能系统动态功率跟踪,以某地区典型的微电网系统功率需求数据作为测试算例,VRB 储能系统由 6 个 100 kW 的单体电池单元组成,选取某一调度周期内的功率变化数据,采样时间为 1 min,采用复合控制器 DDPG-模糊 PID 与模糊 PID 控制器对其功率进行跟踪,其中 VRB 储能需求总功率计算如式(19)所示。

$$P_{VRB}(t) = P_{PV} + P_{WT} - P_{Grid} \quad (19)$$

式中:  $P_{VRB}$  为 VRB 储能需求总功率;  $P_{PV}$  为光伏出力;  $P_{WT}$  为风电功率;  $P_{Grid}$  为直流微电网需求总功率。

其各微电源负荷变化如图 17 所示,根据式(19)计算出 VRB 储能系统总功率需求,如图 18 所示。

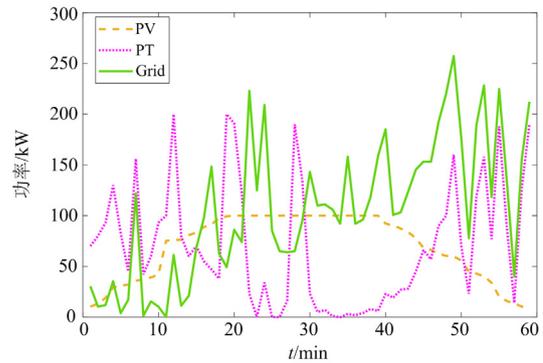


图 17 各微电源功率输出

Fig. 17 Power output of each micro power supply

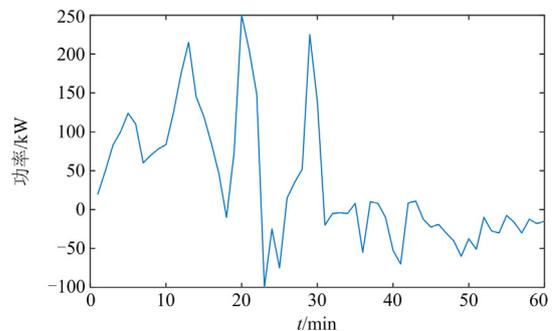


图 18 VRB 储能系统总功率需求

Fig. 18 Total power demand of energy storage system

通过所设计的系统控制策略, 其复合控制器 DDPG-模糊 PID 与模糊 PID 控制器对其瞬时功率跟踪结果如图 19(a)所示, 两种控制器相对误差结果如图 19(b)所示。

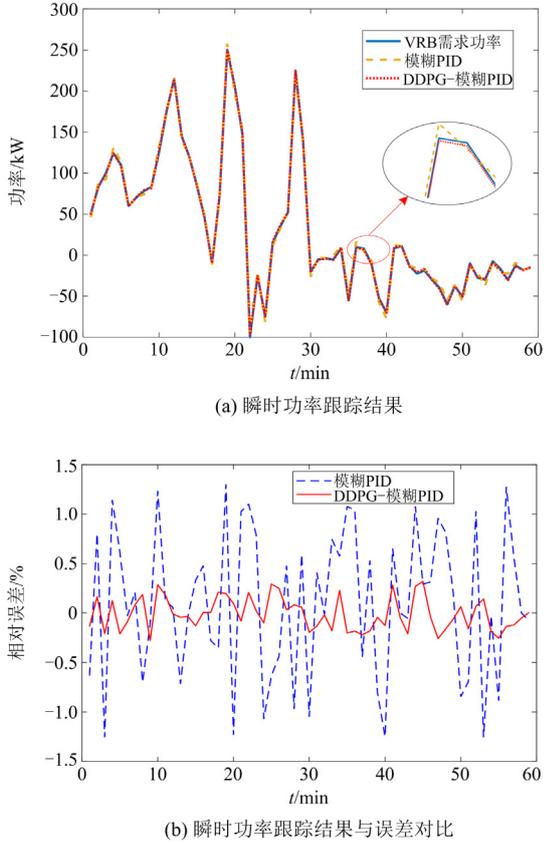


图 19 两种控制器对其瞬时功率跟踪结果与误差对比  
Fig. 19 Comparison of instantaneous power tracking results and errors between two controllers

从图 19 可知, DDPG-模糊 PID 复合控制器对 VRB 储能系统的瞬时需求功率跟踪比模糊 PID 有一定的优势, 系统实际出力与调度指令功率偏差在额定功率的 $\pm 2\%$ 范围内, 符合 GBT36547-2018 国家标准中的要求。

#### 4.3 场景 3: 系统受风电与负载扰动时功率跟踪

在场景 2 直流微电网稳态的基础上, 需要进一步考虑当风电与负载功率随时间变化时, 会给分配层功率指令值产生一定的随机功率偏差扰动。故通过观察受到扰动后两种控制器对功率偏差  $\Delta P$  的控制效果进行验证。将由风电和负载变化后引起的随机功率偏差视为对 VRB 储能系统功率控制器的总扰动, 其风电和负载变化产生的扰动引起的功率偏差如图 20 所示。

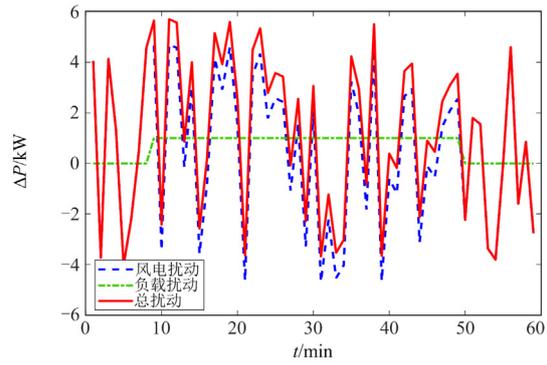


图 20 扰动引起的功率偏差变化  
Fig. 20 Power deviation change caused by disturbance

为了验证 DDPG-模糊 PID 复合控制器和模糊 PID 对 VRB 储能系统受到总扰动引起的功率偏差  $\Delta P_t$  的跟踪修正效果, 分别与两种控制器跟踪修正后的功率偏差  $\Delta P_{1t}$  和  $\Delta P_{2t}$  作差比较, 差值结果  $e_{p1t}$  和  $e_{p2t}$  可分别表示为

$$e_{p1t} = \Delta P_t - \Delta P_{1t} \quad (20)$$

$$e_{p2t} = \Delta P_t - \Delta P_{2t} \quad (21)$$

式中:  $\Delta P_t$  为 VRB 储能系统受扰动后的总功率偏差值;  $\Delta P_{1t}$  和  $\Delta P_{2t}$  分别为通过 DDPG-模糊 PID 复合控制器和模糊 PID 跟踪修正后的功率偏差值, 差值结果越接近零值附近, 跟踪修正效果越好。

两种控制器通过对 VRB 储能系统受总扰动后的功率偏差值进行跟踪修正后, 并与跟踪修正前的功率偏差作差比较后的差值结果如图 21 所示。

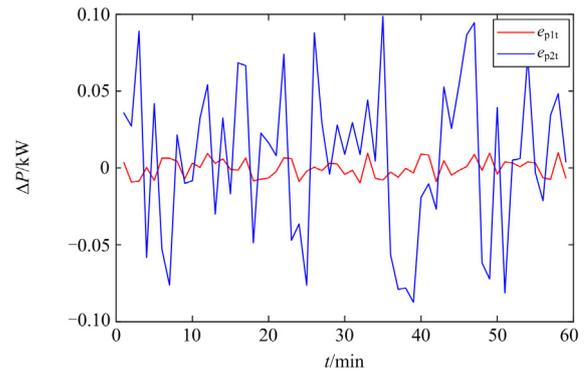


图 21 两种控制器跟踪修正前后差值结果对比  
Fig. 21 Comparison of difference results between two controllers before and after tracking correction

为了进一步说明所提 DDPG-模糊 PID 复合控制器的效果, 在场景 3 的基础上, 在相同的实验环境下, 本节选择了与改进粒子群优化的神经网络<sup>[9]</sup>和改进多智能体深度确定性策略梯度<sup>[16]</sup>进行对比分析, 计算出控制器对随机变化引起的 VRB 储能

系统总功率偏差扰动控制效果, 其消除功率偏差  $e_p$  的各指标对比如表 3 所示。

表 3 随机扰动下的结果对比

Table 3 Comparison of results under random disturbance

控制器	平均值/kW	最大值/kW	标准差/kW
模糊 PID	0.007 03	0.0917	0.050 21
DDPG-模糊 PID	0.000 31	0.0087	0.005 96
PSO-BP	0.010 08	0.011 66	0.082 12
MA-DDPG	0.004 16	0.0228	0.006 18

通过图 21 与表 3 分析可知, DDPG-模糊 PID 复合控制器用于跟踪 VRB 储能系统功率时, 在受到扰动时的消除功率偏差的平均值为 0.000 31, 最大值为 0.0087, 能够趋于零值附近, 满足控制精度要求, 并与 PSO-BP、MA-DDPG 算法相比, 有一定的优势, 从而验证了所提控制器的抗干扰能力和准确性。

## 5 结论

为了解决直流微电网 VRB 储能系统就地层中功率跟踪控制器的响应速度和抗干扰能力问题, 分别通过场景 1、场景 2 及场景 3 仿真验证所提控制策略的有效性, 主要结论如下:

1) 提出了一种基于 DDPG 与模糊 PID 组成的复合控制器, 将模糊 PID 控制器作为主控制器, 利用 DDPG 辅助控制器在线学习训练, 补偿功率控制器的外部扰动, 实现 VRB 储能系统就地控制层中指令功率的快速响应, 并能准确跟踪到电池单元;

2) 分别对 VRB 储能单元充电与放电瞬时功率进行跟踪, 采用 DDPG-模糊 PID 复合控制器比模糊 PID 控制器在充电与放电时的响应时间分别快 0.74 s 和 0.655 s, 其功率跟踪响应时间符合《电化学储能系统接入电网技术规定》中的响应时间, 验证了所提控制策略在响应速度上较快;

3) 对 VRB 储能系统实时需求功率进行跟踪, 并对比所提复合控制器与模糊 PID 控制器的控制性能, 得出功率跟踪值与调度指令曲线偏差均小于额定功率的  $\pm 2\%$ , 且 DDPG-模糊 PID 复合控制器误差更小, 验证了所提控制策略的准确性;

4) 系统受到随机扰动时, 对产生的功率偏差进行跟踪修正, 所提 DDPG-模糊 PID 复合控制器差值结果的平均值较模糊 PID 小 0.006 72 kW, 且差值结果能够趋于零值附近, 满足控制精度要求, 验证了在系统发生扰动时, 加入 DDPG 控制策略, 能够有效地提升主控制器的抗干扰能力。

## 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联络线功率控制策略[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] 张勤进, 张瀚文, 刘彦呈, 等. 基于参数自适应的直流微源虚拟发电机控制策略[J]. 电力系统保护与控制, 2021, 49(18): 90-97.  
ZHANG Qinjin, ZHANG Hanwen, LIU Yancheng, et al. Control strategy for a DC micro source virtual generator based on adaptive parameters[J]. Power System Protection and Control, 2021, 49(18): 90-97.
- [3] MUTARRAF M U, TERRICHE Y, NASIR M, et al. A communication-less multi-mode control approach for adaptive power-sharing in ships-based seaport microgrid[J]. IEEE Transactions on Transportation Electrification, 2021, 7(4): 3070-3081.
- [4] 张宇涵, 杜贵平, 雷雁雄, 等. 直流微网混合储能系统控制策略现状及展望[J]. 电力系统保护与控制, 2021, 49(3): 177-187.  
ZHANG Yuhuan, DU Guiping, LEI Yanxiong, et al. Current status and prospects of control strategy for a DC micro grid hybrid energy storage system[J]. Power System Protection and Control, 2021, 49(3): 177-187.
- [5] GABER M, ABUALKASIM B, MOHAMMED A. Superconducting energy storage technology-based synthetic inertia system control to enhance frequency dynamic performance in microgrids with high renewable penetration[J]. Protection and Control of Modern Power Systems, 2021, 6(4): 460-472.
- [6] SIOSHANSI R, ARTEAGA J, BHATTACHARJEE S, et al. Energy-storage modeling: state-of-the-art and future research directions[J]. IEEE Transactions on Power Systems, 2022, 37(2): 860-875.
- [7] 李浩, 朱才朝, 樊志鑫, 等. 基于二阶滑膜-PID 控制的风力发电机最大功率跟踪策略研究[J]. 太阳能学报, 2022, 43(3): 304-314.  
LI Hao, ZHU Caichao, FAN Zhixin, et al. Research on maximum power point tracking strategy of wind turbine based on second order sliding model-PID control[J]. Acta Energiae Solaris Sinica, 2022, 43(3): 304-314.
- [8] 李忠文, 吴龙, 程志平, 等. 光储系统参与微电网频率调节的模糊自适应滑模控制[J]. 高电压技术, 2022, 48(6): 2065-2076.  
LI Zhongwen, WU Long, CHENG Zhiping, et al. Fuzzy adaptive sliding mode control of photovoltaic and storage

- systems for providing frequency regulation of microgrid[J]. *High Voltage Engineering*, 2022, 48(6): 2065-2076.
- [9] 黄宝洲, 杨俊华, 卢思灵, 等. 基于改进粒子群优化神经网络算法的波浪捕获功率预测[J]. *太阳能学报*, 2021, 42(2): 302-308.  
HUANG Baozhou, YANG Junhua, LU Siling, et al. Wave capture power forecasting based on improved particle swarm optimization neural network algorithm[J]. *Acta Energetica Solaris Sinica*, 2021, 42(2): 302-308.
- [10] 凤丽洲, 王友卫, 韩琳琳, 等. 双重驱动的果蝇优化算法及其在 PID 控制器中的应用[J]. *控制与决策*, 2021, 36(9): 2221-2233.  
FENG Lizhou, WANG Youwei, HAN Linlin, et al. Double drive fruit fly optimization algorithm and its application in PID controller[J]. *Control and Decision*, 2021, 36(9): 2221-2233.
- [11] 左剑, 王子琪, 李银红, 等. 计及时滞的互联电网负荷频率控制最优分数阶 PID 控制器设计[J]. *控制理论与应用*, 2017, 34(9): 1151-1160.  
ZUO Jian, WANG Ziqi, LI Yinhong, et al. Optimal fractional-order PID controller design for interconnected power grid load frequency control considering time-delay[J]. *Control Theory & Applications*, 2017, 34(9): 1151-1160.
- [12] 马彦, 丁浩, 牟洪元, 等. 基于模糊 PID 算法的动力电池液体冷却策略[J]. *控制理论与应用*, 2021, 38(5): 549-560.  
MA Yan, DING Hao, MU Hongyuan, et al. Liquid cooling strategy of power battery based on fuzzy PID algorithm[J]. *Control Theory & Applications*, 2021, 38(5): 549-560.
- [13] 林歆悠, 王召瑞. 应用粒子群算法优化模糊规则的自适应多目标控制策略[J]. *控制理论与应用*, 2021, 38(6): 842-850.  
LI Xinyou, WANG Zhaorui. Adaptive multi-objective control strategy based on particle swarm optimization algorithm optimized fuzzy rules[J]. *Control Theory & Applications*, 2021, 38(6): 842-850.
- [14] 李嘉文, 余涛, 张孝顺, 等. 基于改进深度确定性梯度算法的 AGC 发电功率指令分配方法[J]. *中国电机工程学报*, 2021, 41(21): 7198-7212.  
LI Jiawen, YU Tao, ZHANG Xiaoshun, et al. AGC power generation command allocation method based on improved deep deterministic policy gradient algorithm[J]. *Proceedings of the CSEE*, 2021, 41(21): 7198-7212.
- [15] 时圣苗, 刘全. 采用分类经验回放的深度确定性策略梯度方法[J]. *自动化学报*, 2022, 48(7): 1816-1823.  
SHI Shengmiao, LIU Quan. Deep deterministic policy gradient with classified experience replay[J]. *Acta Automatica Sinica*, 2022, 48(7): 1816-1823.
- [16] 范培潇, 柯松, 杨军, 等. 基于改进多智能体深度确定性策略梯度的多微网负荷频率协同控制策略[J]. *电网技术*, 2022, 46(9): 3504-3514.  
FAN Peixiao, KE Song, YANG Jun, et al. Load frequency coordinated control strategy of multi-microgrid based on improved MA-DDPG[J]. *Power System Technology*, 2022, 46(9): 3504-3514.
- [17] 李建林, 武亦文, 王楠, 等. 吉瓦级电化学储能电站研究综述及展望[J]. *电力系统自动化*, 2021, 45(19): 2-14.  
LI Jianlin, WU Yiwen, WANG Nan, et al. Review and prospect of gigawatt-level electrochemical energy storage power station[J]. *Automation of Electric Power Systems*, 2021, 45(19): 2-14.
- [18] 邵军康, 李鑫, 邱亚, 等. 全钒液流电池多场耦合建模研究[J]. *高电压技术*, 2021, 47(5): 1881-1891.  
SHAO Junkang, LI Xin, QIU Ya, et al. Multi-field coupling modeling of vanadium redox battery[J]. *High Voltage Engineering*, 2021, 47(5): 1881-1891.
- [19] XUE Jiankai, SHEN Bo. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. *Systems Science & Control Engineering*, 2020, 8(1): 22-34.
- [20] ZHANG Zhizheng, CHEN Jiale, CHEN Zhibo, et al. Asynchronous episodic deep deterministic policy gradient: toward continuous control in computationally complex environments[J]. *IEEE Transactions on Cybernetics*, 2021, 51(2): 604-613.
- [21] CAI Mingyu, MOHAMMADHOSEIN H, XIAO Shaoping, et al. Modular deep reinforcement learning for continuous motion planning with temporal logic[J]. *IEEE Robotics and Automation Letters*, 2021, 6(4): 7973-7980.
- [22] MOHAMMAD H K, MEYSAM G. A novel deep reinforcement learning controller based type-II fuzzy system: frequency regulation in microgrids[J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2021, 5(4): 689-699.
- [23] DAVOUD F, MEHDI R. Novel PID-fuzzy video rate controller for high-delay applications of the HEVC standard[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(6): 1379-1389.
- [24] CHEN Jianqi, MA Dan, XU Yong, et al. Delay robustness of PID control of second-order systems: pseudo concavity, exact delay margin, and performance tradeoff[J]. *IEEE Transactions on Automatic Control*, 2022, 67(3): 1194-1209.

收稿日期: 2022-11-08; 修回日期: 2023-02-23

作者简介:

陆鹏(1990—), 男, 博士研究生, 研究方向为微电网与储能协调运行及控制技术; E-mail: lupeng1913@163.com

付华(1962—), 女, 博士, 教授, 研究方向为微电网理论与技术。E-mail: fxfuhua@163.com

(编辑 魏小丽)