

DOI: 10.19783/j.cnki.pspc.250857

# 基于 SAC 算法的混储微网观测修正自抗扰稳压策略

周雪松<sup>1</sup>, 马培铭<sup>1</sup>, 练继建<sup>2</sup>, 李苏扬<sup>1</sup>, 陶珑<sup>1</sup>, 刘晓林<sup>1</sup>

(1. 天津市新能源电力变换传输与智能控制重点实验室(天津理工大学), 天津 300384;

2. 水利工程智能建设与运维全国重点实验室(天津大学), 天津 300350)

**摘要:** 新型电力系统中储能双向 DC-DC 变换器作为新能源高比例消纳的柔性枢纽, 在微电网的多重不确定性影响下, 面临着输出侧电压失稳风险。因此, 提出一种基于柔性动作-评价(soft actor-critic, SAC)算法辅助寻优的观测修正自抗扰控制技术。首先, 引入二维扰动信息至待补偿项中进行扰动状态量的协同观测, 同时设计迟滞函数修正微分环节的固有缺陷, 从而精准重构总和扰动。随后, 量化参数整定准则, 并借助 SAC 算法的最大熵学习框架与随机策略探索, 实现控制器参数在多频域扰动下的柔性匹配, 使储能系统能更加充分发挥“削峰填谷”的调控作用。最后, 在不同工况的仿真对比下, 验证了所提策略在多种内外不确定扰动下均具备良好的动态性能。

**关键词:** DC-DC 变换器; 混储微电网; SAC 算法; 迟滞函数; 二维扰动

## Observation-corrected active disturbance rejection voltage stabilization strategy for hybrid storage microgrid based on SAC algorithm

ZHOU Xuesong<sup>1</sup>, MA Peiming<sup>1</sup>, LIAN Jijian<sup>2</sup>, LI Suyang<sup>1</sup>, TAO Long<sup>1</sup>, LIU Xiaolin<sup>1</sup>

(1. Tianjin Key Laboratory of New Energy Power Conversion, Transmission and Intelligent Control (Tianjin University of Technology), Tianjin 300384, China; 2. National Key Laboratory of Intelligent Construction and Operation and Maintenance of Hydraulic Engineering, Tianjin University, Tianjin 300350, China)

**Abstract:** In new power systems, bidirectional DC-DC converters for energy storage serve as flexible hubs for accommodating high-penetration of renewable energy. However, under multiple uncertainties in microgrids, they face the risk of voltage instability. To address this issue, an observation-corrected active disturbance rejection control (ADRC) strategy assisted by the soft actor-critic (SAC) algorithm is proposed. First, two-dimensional disturbance information is incorporated into the compensation term to enable coordinated observation of disturbance state variables. Meanwhile, a hysteresis function is designed to compensate for the inherent shortcomings of the differential component, thereby accurately reconstructing the total disturbance. Subsequently, parameter tuning criteria are quantified, and the maximum entropy learning framework and stochastic policy exploration of the SAC algorithm are employed to achieve flexible matching of controller parameters under multiple disturbances, enabling the energy storage system to better perform peak shaving and valley filling. Finally, simulation comparisons under various operating conditions verify that the proposed strategy exhibits excellent dynamic performance under multiple internal and external uncertainties.

This work is supported by the Major Program of National Natural Science Foundation of China (No. U24B6011).

**Key words:** DC-DC converters; hybrid storage microgrid; SAC algorithm; hysteresis function; two-dimensional disturbance

## 0 引言

在环境问题严峻和能源结构变革的时代下, 大规模可再生能源发电和分布式清洁能源成为系统主

导<sup>[1]</sup>。母线电压作为衡量供需平衡与直流微网正常运行的关键因素, 其扰动偏离不仅会恶化电能质量、破坏功率平衡并诱发振荡, 严重时还会导致微电网停运<sup>[2]</sup>。因此, 以蓄电池和超级电容所组成的混合储能系统的引入可以弥补间歇性新能源输出, 而考虑储能本身的特性设计协调控制框架, 并构造合适的内外环控制器, 则成为其通过调控双向 DC-DC 变

**基金项目:** 国家自然科学基金重大项目资助(U24B6011); 国家自然科学基金重点项目资助(U23B20142)

换器(bidirectional DC-DC converter, BDC)的模式切换, 实现储能单元与直流母线间能量交换的关键<sup>[3]</sup>。

面对新能源的强不确定性以及微网的多变量、强耦合等复杂特征。现有控制策略, 如模型预测控制<sup>[4]</sup>虽便于处理多变量与约束, 但参数漂移与寄生效应也会引起性能退化; 滑模控制<sup>[5]</sup>虽对匹配扰动具备鲁棒性, 但其固有抖振表现出高频纹波与附加损耗。这些控制策略对改善系统动态表现有限。而近年来强化学习如近端策略优化、深度 Q 网络等算法凭借其自适应强以及无需精确模型便可探索学习的能力适配非线性与时变环境, 在系统控制领域得到广泛应用<sup>[6-7]</sup>, 同时传统自抗扰控制<sup>[8]</sup>(traditional linear active disturbance rejection control, TLADRC)所展现出的强抗扰、强鲁棒性以及快速抑制非线性系统波动的能力而备受关注。但在实际微电网运行下, 受寄生参数、外部噪声扰动等因素制约, BDC 电路模型既非线性也非二阶<sup>[9]</sup>。而传统低阶自抗扰的扩张状态观测器(linear extended state observer, LESO)假定被控对象为低阶模型, 基于一维扰动观测输出对模型的内外扰动总和进行表征是不够全面且准确的。当 LESO 对被控系统的估计模型与实际系统阶次存在差异时, 会导致扰动观测量的收敛速度降低, 观测偏差使控制表现出过冲和时滞等非理想特征, 致使 TLADRC 补偿环节呈现时间滞后现象<sup>[10]</sup>。

针对上述问题, 新能源发电领域现有研究大多以改进观测器自身结构为主使之匹配高阶系统, 文献[11]针对 LESO 环节引入二级观测器中形成级联结构用于高阶系统控制, 达到了对总扰动更为精准的描述。文献[12]引入模型信息到 LESO 的系数矩阵改善了低阶自抗扰阶数不匹配的时延特性, 减轻了观测负担。这种“变结构”思想虽在一定程度上改善了电能质量, 但仍未脱离用一维扩张描述总扰动, 欠缺考虑仅从观测结构的角度优化 TLADRC, 难以突破既有框架实现更强的鲁棒控制。此外, 以“变参数”思想弥补低阶观测器不匹配高阶对象主要分为两类主流措施。第一种为模糊自适应、神经网络等智能算法<sup>[13-14]</sup>的引入, 实现了 TLADRC 参数在多频扰动耦合下的实时整定。然而这类智能算法具有初值敏感性且缺少理论依据。第二种为深度 Q 网络、深度确定性策略梯度的基础强化学习<sup>[15-16]</sup>实现参数策略整定, 但是前者因离散化参数空间面临策略振荡, 后者因确定性策略与外部噪声的探索机制导致学习效率低下, 均难以在混储微电网中实现

兼具高效、稳定与强鲁棒性的参数自适应。

为此本文融合“变结构”和“变参数”思想, 聚焦扰动观测输出通道, 引入扩张出的二维状态变量予以二阶补偿, 并增设迟滞修正环节。旨在不改变原有 LESO 结构基础上, 达到不加重观测负担并精确描述内外总扰动并予以补偿的目的, 由此构成扰动观测重构自抗扰控制(dynamic observation reconfiguration liner active disturbance rejection control, DLADRC)。并基于现有研究算法面对微电网非线性、强不确定性存在的策略振荡和探索效率低下等问题, 将遵循最大熵原则的柔性动作-评价(soft actor-critic, SAC)算法嵌入到 DLADRC 中, 实现时变扰动下参数自适应学习与联合寻优。最后通过内外扰动多工况数字仿真的控制效果, 验证了所提控制策略在微电网中对各类扰动的抑制性能。

## 1 系统结构及被控对象模型建立

典型混合储能微电网运行结构如图 1 所示。经由 BDC 电路, 蓄电池与超级电容所构成的混合储能系统可与直流母线之间建立桥梁, 通过多模式协同控制策略实现源-储-荷功率柔性互补<sup>[17]</sup>。其中,  $L_b$  和  $L_{sc}$  分别为蓄电池侧和超级电容侧的电感;  $S_{a1}$ 、 $S_{a2}$  和  $S_{b1}$ 、 $S_{b2}$  为两种储能设备的开关管。

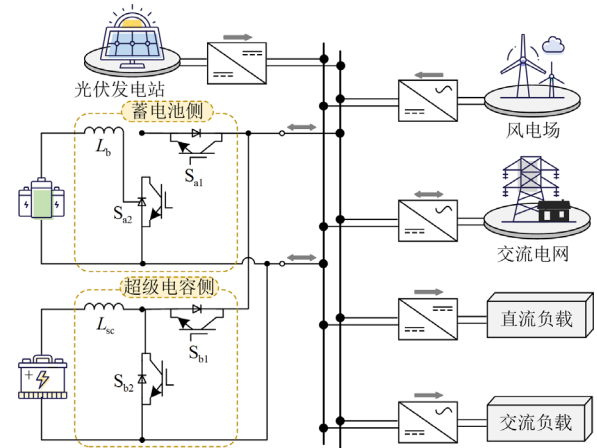


图 1 直流微电网结构图

Fig.1 Structure diagram of a DC microgrid

储能装置在放电模式下, 以蓄电池储能装置为例, 当  $S_{a1}$  关断、 $S_{a2}$  处于导通时间内, 储能工作在放电状态一, 在两开关管状态交换后动态切换为放电状态二。两个阶段的状态空间方程如式(1)所示。

$$\begin{cases} C \frac{du_{dc}}{dt} = \frac{u_{dc}}{R} \\ L \frac{di_{L_b}}{dt} = u_0 \end{cases}, \begin{cases} C \frac{du_{dc}}{dt} = i_{L_b} - \frac{u_{dc}}{R} \\ L \frac{di_{L_b}}{dt} = u_0 - u_{dc} \end{cases} \quad (1)$$

由式(1)得开关周期内的平均值为

$$\begin{cases} \langle u_{L_b}(t) \rangle_{T_s} = \int_t^{t+T_s} u_{L_b}(t) dt = d \langle u_{dc}(t) \rangle_{T_s} + \langle u_0(t) \rangle_{T_s} \\ \langle i_{dc}(t) \rangle_{T_s} = -\langle u_{dc}(t) \rangle_{T_s} / R + d \langle i_{L_b}(t) \rangle_{T_s} \end{cases} \quad (2)$$

在系统中引入小信号扰动，并分离消去稳态分量和交流分量，可得 BOOST 模式下小信号矩阵为

$$\frac{d}{dt} \begin{bmatrix} \hat{i}_{L_b} \\ \hat{u}_{dc} \end{bmatrix} = \begin{bmatrix} 0 & -D/L_b \\ D/C & -1/RC \end{bmatrix} \begin{bmatrix} \hat{i}_L \\ \hat{u}_{dc} \end{bmatrix} + \begin{bmatrix} u_{dc}/L_b \\ i_{L_b}/C \end{bmatrix} \hat{d} + \begin{bmatrix} L_b^{-1} \\ 0 \end{bmatrix} \hat{u}_0 \quad (3)$$

式中： $u_{dc}$  为母线电压； $u_0$  为储能侧电压； $C$  为等效电容； $R$  为等效负载； $L$  为等效电感； $i_{L_b}$  为储能侧电感电流； $i_{dc}$  为母线侧电流； $T_s$  为开关周期； $d$  为占空比； $D$  为占空比稳态值， $d = D + \hat{d}$ ；忽略变换器两侧扰动，得到其传递函数为

$$\begin{cases} G_{i_{L_b}, D}(s) = \left. \frac{\hat{i}_{L_b}(s)}{\hat{d}(s)} \right|_{\hat{u}_0(s)=0} = \frac{u_{dc}(Cs + 2R^{-1})}{L_b Cs^2 + L_b Rs + D^2} \\ G_{u_{dc}, i_L}(s) = \left. \frac{\hat{u}_{dc}(s)}{\hat{i}_{L_b}(s)} \right|_{\hat{u}_0(s)=0} = \frac{D - L_b(RD)^{-1}s}{Cs + 2R^{-1}} \end{cases} \quad (4)$$

式中： $\hat{i}_{L_b}(s)$ 、 $\hat{u}_{dc}(s)$ 、 $\hat{u}_0(s)$ 、 $\hat{d}(s)$  为对应状态量小信号模型。

同理，在储能设备充电模式下，接口单元以 Buck 模式实现能量由外部电源向储能系统定向传输。该模式传递函数可由相似的运算过程得到。

$$\begin{cases} G_{i_{L_b}, D}(s) = \left. \frac{\hat{i}_{L_b}(s)}{\hat{d}(s)} \right|_{\hat{u}_{dc}(s)=0} = \frac{u_{dc}(C's + R'^{-1})}{L_b C's^2 + LR'^{-1}s + 1} \\ G_{u_0, D}(s) = \left. \frac{\hat{u}_0(s)}{\hat{i}_{L_b}(s)} \right|_{\hat{u}_{dc}(s)=0} = \frac{R'}{R'C's + 1} \end{cases} \quad (5)$$

式中： $C'$ 、 $R'$  分别为储能侧等效电容与等效负载。

## 2 混合储能改进自抗扰控制策略

### 2.1 TLADRC 策略控制模式

在自抗扰范式<sup>[8]</sup>下，状态函数  $y$  和外部扰动  $\eta$  和控制误差  $(b - b_0)u$  等实际模型与理想模型间结构的偏差量统一定义为总扰动函数  $f = f(y, \eta, u)$ ，系统状态空间可表述<sup>[21]</sup>为

$$\begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_0 \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} f \\ y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{cases} \quad (6)$$

基于式(6)，可设计扩张状态观测如式(7)所示。

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} b_0 \\ 0 \end{bmatrix} u - L_0(z_1 - y) \quad (7)$$

式中：矩阵满足  $L_0 = [\beta_1 \quad \beta_2]^T$ ， $\beta_1$ 、 $\beta_2$  为观测器增益； $b_0$  为控制量增益； $u$  为自抗扰输出的控制量。

借助式(8)线性状态误差反馈控制(linear state error feedback, LSEF)使系统从被动抗扰转换至主动补偿总扰动的模式。

$$u = [K_p(r - z_1) - z_2] \cdot b_0^{-1} \quad (8)$$

式中： $K_p$  为控制律增益； $r$  为给定值。

基于上述设计，考虑参考信号  $e_r$  追溯误差满足  $e_r = r - x_1$ ，根据式(7)得到

$$\dot{e}_r = \dot{r} - K_p e_r + K_p(z_1 - x_1) + z_2 - f \quad (9)$$

则  $e_r$  对应的复频域表达结果为

$$E_r(s) = \frac{sR(s)}{s + K_p} + \frac{K_p[\hat{X}_1(s) - X_1(s)]}{s + K_p} + \frac{\hat{F}(s) - F(s)}{s + K_p} \quad (10)$$

由式(10)可知，模型观测误差分量  $\hat{X}_1(s) - X_1(s)$  由建模时未考虑结构偏差造成，难以消除。参考信号分量  $sR(s)$  的改进又受诸多因素限制。因此跟踪误差更多取决于扰动观测误差分量  $\hat{F}(s) - F(s)$ ，在新能源并网的复杂波动环境下，应尽可能改善由低阶 LESO 造成的观测偏差，提高系统控制性能。

### 2.2 混储微电网外环 DLADRC 策略设计

根据式(7)可得 LESO 扰动观测传递函数。

$$H_1(s) = \frac{Z_2(s)}{F(s)} = \frac{\beta_2}{s^2 + \beta_1 s + \beta_2} \quad (11)$$

由于低阶 LESO 受到光伏混储微电网的多变量、高维度、干扰复杂等特征的制约。因此参考增阶原理，并考虑“观测器阶次越高，其相位滞后越严重”这一规律<sup>[18]</sup>，在不改变原本观测器阶次的基础上，引入高阶信息于扰动观测输出通道，依靠二维扰动变量观测机制驱动收敛，可得

$$Z_3(s) = Z_2(s) + m_0 \dot{Z}_2(s) \quad (12)$$

式中： $Z_3(s)$  为高阶观测输出； $m_0$  为高阶增益系数。

通过拉氏变换，基于扰动模型高阶状态引入可将  $H_1(s)$  加以修正为同/高阶扰动观测传函  $H_2(s)$ 。

$$H_2(s) = \frac{Z_3(s)}{F(s)} = \frac{m_0 \beta_2 s + \beta_2}{s^2 + \beta_1 s + \beta_2} \quad (13)$$

相较  $H_1(s)$  与  $H_2(s)$ ，高阶分量传递函数新引入自由零点  $-1/m_0$ ， $m_0$  增大可以使观测器根轨迹向左偏移，提升观测收敛性能。但  $m_0$  过大，零点越靠近原点便越会呈现出微分特性使高频噪声抑制能力衰减<sup>[19]</sup>，则改善高频鲁棒性是进一步提升扰动观测

精度的必要条件。针对上述不足, 引入迟滞校正函数作用于观测输出通道, 得到扰动观测重构型线性扩张状态观测器(dynamic observation reconfiguration linear extended state observer, DLESO), 以其频域衰减特性降低噪声频段增益的形式进一步增强扰动抑制能力。基于式(9), 将其原有观测特性修正为

$$H_3(s) = \frac{Z_4(s)}{F(s)} = \frac{m_0\beta_2s + \beta_2}{(s^2 + \beta_1s + \beta_2)(\tau s + 1)} \quad (14)$$

式中:  $Z_4(s)$  为最终扰动观测重构的总和扰动高阶观测输出;  $\tau$  为迟滞时间常数, 保证观测通道中千赫兹等级扰动观测幅值衰减到 10% 以下。针对高频噪声抑制能力衰减的改善, 令  $s = j\omega$ , 因考虑高频段  $\omega$  值很大, 代入式(13)与式(14)取模。

$$\begin{cases} |H_2(j\omega)| = \left| \frac{m_0\beta_2(j\omega) + \beta_2}{(j\omega)^2 + \beta_1(j\omega) + \beta_2} \right| \sim \left| \frac{jm_0\beta_2}{\omega} \right| = \frac{m_0\beta_2}{\omega} \\ |H_3(j\omega)| = |H_2(j\omega)| \cdot \left| \frac{1}{\tau\omega} \right| = \frac{m_0\beta_2}{\tau\omega^2} \end{cases} \quad (15)$$

由于迟滞环节的加入, 扰动观测通道幅值特性的斜率从每十倍频程下降 20 dB 变为下降 40 dB。通过观测重构后所增加的每十倍频程 20 dB 的滚降, 使得改进后观测器对高频扰动的衰减能力增强, 进一步提高了扰动抑制精度。取观测器带宽  $\omega_0 = 100 \text{ rad/s}$ , 绘制以上 3 种表征扰动估计性能的传递函数频率特性曲线, 如图 2 所示。

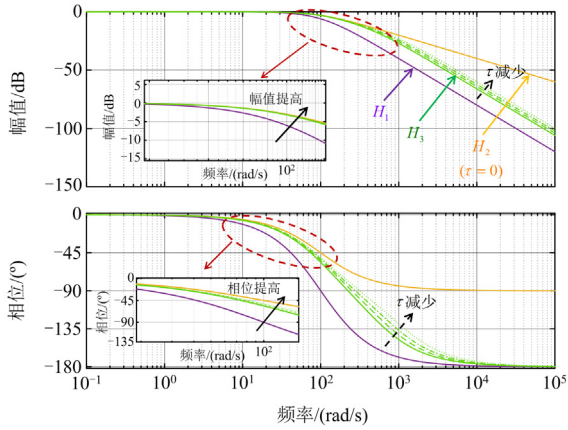


图 2 扰动观测特性对比曲线

Fig. 2 Comparison curves of perturbation observation characteristic

由图 2 可知, 在低频段三者随频率升高相位一致性特征显著; 中频段  $H_2(s)$ 、 $H_3(s)$  相较于  $H_1(s)$  相位滞后得到明显改善, 提高了观测器的扰动观测能力; 高频段  $H_3(s)$  基于  $H_2(s)$  的优劣取长补短, 使高频衰减能力显著增强。令  $\alpha = m_0/\tau$ , 经式(15)

得到最大相位超前角  $\phi = \sin^{-1}(\alpha - 1/\alpha + 1)$ , 以及对应频率  $\omega = 1/\sqrt{m_0\tau}$ , 即  $\alpha = (1 + \sin(\Delta\phi))/(1 - \sin(\Delta\phi))$ 。随后选取中频段相位得到明显改善的中间频率  $\omega = 316.2 \text{ rad/s}$ , 由  $H_1(s)$ 、 $H_2(s)$  对应的相位约为  $-163^\circ$  与  $-102^\circ$  得知, 可增加相位量  $\Delta\phi$  约为  $65^\circ$ , 代入得到  $\tau \approx 0.00015$ , 可利用此频域法进行  $\tau$  值选取。

基于式(12)和上述分析可知, 此时, DLADRC 观测方程数学表达如式(16)所示。

$$\begin{cases} \dot{z}_1 = \beta_1(y - z_1) + z_2 + b_0u, & \dot{z}_2 = \beta_2(y - z_1) \\ \dot{z}_3 = m_0\beta_2(\dot{y} - \dot{z}_1) + \dot{z}_2, & \dot{z}_4 = \tau^{-1}(z_3 + z_4) \end{cases} \quad (16)$$

基于式(8), 新的 LSEF 单元模型构造如式(17)所示。

$$u = [K_p(r - z_1) - z_4] \cdot b_0^{-1} \quad (17)$$

基于 DLADRC 策略的混合储能控制如图 3 所示。因电压外环控制器主导系统全局对不确定扰动的抑制能力, 故 DLADRC 策略主要针对电压外环进行设计。同时电流内环部署 PI 控制器, 驱动蓄电池与超级电容实现功率指令的快速动态响应。

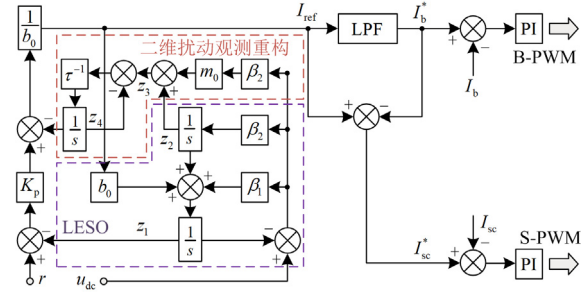


图 3 DLADRC 策略的混合储能控制

Fig. 3 Hybrid energy storage control of DLADRC strategy

图 3 中:  $I_{ref}$  为总参考电流;  $I_b$ 、 $I_b^*$  分别为蓄电池电流及其参考值;  $I_{sc}$ 、 $I_{sc}^*$  分别为超级电容电流及其参考值;  $k_p = \omega_c$ ,  $\omega_c$  为 LSEF 单元的带宽参数。

## 2.3 性能分析

### 2.3.1 扰动抑制性能对比分析

根据式(11)、式(12)可以获得控制变量相对于设定值和输出信号的传递函数。

$$U(s) = \frac{G(s)}{b_0N(s)}R(s) - \frac{H(s)}{b_0N(s)}Y(s) \quad (18)$$

其中:

$$\begin{cases} G(s) = \omega_c(\tau s + 1)(s^2 + 2\omega_0s + \omega_0^2) \\ H(s) = (m_0\omega_0^2 + 2\tau\omega_0\omega_c)s^2 + (\omega_0^2 + \tau\omega_0^2\omega_c + 2\omega_0\omega_c)s + \omega_0^2\omega_c \\ N(s) = \tau s^3 + (1 + 2\tau\omega_0 + \tau\omega_c)s^2 + (2\omega_0 - m_0\omega_0^2 + \tau\omega_0^2 + \omega_c)s \end{cases} \quad (19)$$

DLADRC 下含扰系统等效结构如图 4 所示。

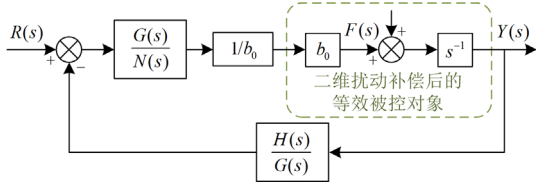


图 4 DLADRC 系统含扰结构图

Fig. 4 Disturbance structure diagram of DLADRC system

由图 4 可以得到输出  $Y(s)$  的传递函数为

$$Y(s) = \frac{G(s)}{sN(s) + H(s)} R(s) + \frac{N(s)}{sN(s) + H(s)} F(s) \quad (20)$$

根据式(13)得到输出  $Y(s)$  与总扰动  $F(s)$  之间关系式为  $\Phi_f(s)$ ，同理可得 TLADRC 下的  $\Phi_c(s)$ 。

$$\Phi_f(s) = \frac{\tau s^3 + c_1 s^2 + c_2 s}{\tau s^4 + c_3 s^3 + c_4 s^2 + c_5 s + \omega_c \omega_0^2} \quad (21)$$

$$\Phi_c(s) = \frac{s^2 + (2\omega_0 + \omega_c)s}{s^3 + (2\omega_0 + \omega_c)s^2 + (\omega_0^2 + 2\omega_0\omega_c)s + \omega_c\omega_0^2} \quad (22)$$

其中：

$$\begin{cases} c_1 = 1 + 2\tau\omega_0 + \tau\omega_c, c_2 = 2\omega_0 + (\tau - m_0)\omega_0^2 + \omega_c \\ c_3 = 1 + 2\tau\omega_0 + \tau\omega_c, c_4 = \tau\omega_0^2 + 2\omega_0 + 2\tau\omega_0\omega_c + \omega_c \\ c_5 = \omega_0^2 + \tau\omega_0^2\omega_c + 2\omega_0\omega_c \end{cases} \quad (23)$$

因此，控制系统抗扰频率特性曲线如图 5 所示。

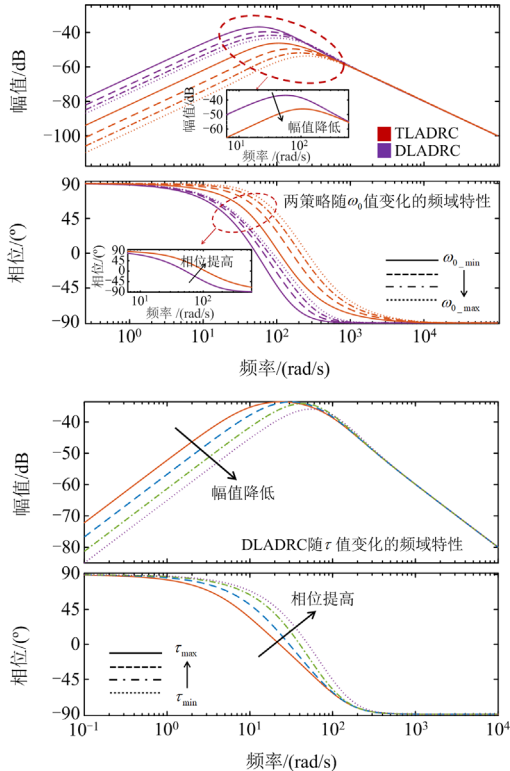


图 5 抗扰性能对比分析

Fig. 5 Comparative analysis of immunity performance

由图 5 可知，在中低频带，TLADRC 控制策略具有较大的响应幅度，表明对干扰更为敏感。相比之下，DLADRC 响应幅度均小于传统控制策略，具有更强的抑制干扰能力。在高频段，两种控制策略的响应幅值相近，抗干扰效果更依赖于  $\omega_0$  的取值。同时可知， $\tau$  值减小可使扰动抑制具备更好的效果，但随着  $\tau$  减小，抗扰能力对于其敏感程度呈现弱化趋势。基于图 4 与式(19)得到噪声频域响应为

$$G_L(s) = \frac{Y(s)}{N_0(s)} = \frac{H(s)}{N(s) - H(s)} \quad (24)$$

基于式(24)，通过系统对噪声响应的高频增益来评价系统的降噪性能，令  $K_d$  为降噪指标，得到

$$\lim_{s \rightarrow \infty} G_L = -K_d/s, K_d = \tau^{-1}m_0\omega_0^2 + 2\omega_0\omega_c \quad (25)$$

将式(28)代入式(25)可得： $K_d = \omega_0^2 + 2\tau^{-1}\omega_0 + 2\omega_0\omega_c$ ， $K_d$  值越小对噪声的高频鲁棒性越好， $\tau$  值的改变可以对降噪指标产生影响， $\tau$  值越大降噪指标  $K_d$  越小，对噪声抑制性能越好。因此该值选取存在一定权衡。

### 2.3.2 扰动估计偏差与参数约束分析

因为相位滞后存在，传统单维度 LESO 难以精确估计时变扰动。由式(6)、式(13)得到改进前后观测偏差  $E_2(s)$  与扰动  $F(s)$  的频域函数关系为

$$\Psi_c(s) = -\frac{\tau s^3 + (1 + 2\tau\omega_0)s^2 + [2\omega_0 + (\tau - m_0)\omega_0^2]s}{\tau s^3 + (1 + 2\tau\omega_0)s^2 + (\tau\omega_0^2 + 2\omega_0)s + \omega_0^2} \quad (26)$$

$$\Psi_f(s) = -\frac{s^2 + 2\omega_0 s}{s^2 + 2\omega_0 s + \omega_0^2} \quad (27)$$

对这两种控制系统引入 4 种不同类型的扰动函数信号进行分析， $M$  为扰动函数增益系数。4 种扰动下的观测函数与跟踪误差曲线如表 1、图 6 所示。

表 1 观测误差函数

Table 1 Observation error function		
$f(t)$	策略	观测偏差信号
$M$	T	$ME_1q_1(t)$
	D	$ME_1q_{c1} + ME_2q_{c2}$
$Mt$	T	$ME_1q_2(t) - 2M\omega_0^{-1}$
	D	$M(E_1q_{c3} + E_2q_{c4}) + M(m_0 - \tau - 2\omega_0^{-1})$
$Mt^2$	T	$M(6\omega_0^2 - 4\tau\omega_0^{-1}) + ME_1q_3(t)$
	D	$M(E_1q_{c5} + E_2q_{c6}) - 2M[(2\omega_0^{-1} + \tau)(m_0 - \tau) - 3\omega_0^{-2}] + 2Mt(m_0 - \tau - 2\omega_0^{-1})$
$M \sin(\omega_0 t)$	T	$ME_1q_4(t) + M[a_1(\omega_0) \cdot \sin(\omega_0 t) + a_2(\omega_0) \cdot \cos(\omega_0 t)]$
	D	$ME_1q_{c7} + ME_2q_{c8} + Mb_1(\omega_0, m_0, \tau) \cdot \sin(\omega_0 t) + Mb_2(\omega_0, m_0, \tau) \cdot \cos(\omega_0 t)$

表 1 中： $E_1 = e^{-\omega_0 t}$ ； $E_2 = e^{-\tau t}$ ； $q_i(t)$  与  $q_{c_j}$  为两种策略下的不同参数变量函数， $i=1,2,3,4$ ， $j=1,2,\dots,8$ ； $a_1(\cdot)$ 、 $a_2(\cdot)$ 、 $b_1(\cdot)$ 、 $b_2(\cdot)$  表示不同的函数；

策略 T 和策略 D 表示 TLADRC 策略与 DLADRC 策略。策略 D 扰动观测偏差函数除随时间收敛到零的部分外, 构成干扰信号最高阶项的表达式均与  $m_0 - \tau - 2\omega_0^{-1}$  及其整数倍相关联。

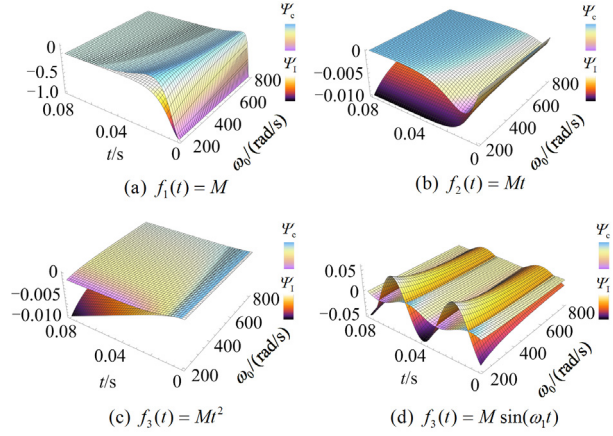


图 6 不同扰动下的跟踪偏差

Fig. 6 Tracking bias under different perturbations

综合图表分析, TLADRC 因低阶扰动补偿, 仅限于在阶跃扰动下实现无差跟踪, 而 DLADRC 策略在面对斜坡与阶跃扰动时观测均可收敛。对于加速度以及周期扰动虽不会随时间收敛到零, 但重构网络补偿了传统观测误差中指数与三角函数项的偏差幅度, 使 DLADRC 能够更加精准识别扰动信号。

$$m_0 = \tau + 2\omega_0^{-1} \quad (28)$$

同时, 当式(28)成立时, 调整三者之间的参数耦合关系可使得观测偏差最小化。为考虑扰动典型性, 令式(21)总扰动  $F(s) = s^{-1}$ , 可将式(21)重新表述为

$$\Phi_l'(s) = \frac{\tau s^2 + c_1 s + c_2}{\tau s^4 + c_3 s^3 + c_4 s^2 + c_5 s + \omega_c \omega_0^2} \quad (29)$$

令  $\omega_0 = \omega_c = \omega$ , 并结合式(29), 得到储能变换器的输出电压信号为

$$v(t) = -\frac{2e^{-t/\tau} \tau^2 \omega}{D^3} + \frac{2Q\tau^2 \omega}{D^3} + \frac{Qt^2 \tau \omega^2}{D} + \frac{Qt(1 + \tau^2 \omega^2)}{D^2} \quad (30)$$

式中:  $Q = e^{-t\omega}$ ,  $D = \tau\omega - 1$ 。由表 1 可知该输出电压最终趋于稳定, 因此, 对式(30)求导所得到  $\dot{v}(t) = 0$  的点, 即为所受外扰下最大输出波动电压时间点  $t'$ , 所对应的数值为最大电压波动幅值  $v(t')$ , 由图 2 分析  $\tau$  取较小的值, 令  $\tau = 10^{-4}$ , 如式(31)所示。

$$\begin{cases} t' = \omega^{-1} + 315.07\omega^{-2} \\ v(t') = \frac{2 \times 10^{-8} \omega (e^{-10^4/\omega - 3.15/\omega^2} - e^{-1-315.07/\omega})}{(1-10^{-4}\omega)^3} + \frac{e^{-1-315.1/\omega} \times M}{\omega^2 (1-10^{-4}\omega)^2} \\ M = (\omega + 313.07)(10^{-12}\omega^2 - 9.68 \times 10^{-9}\omega + 1 + 3.15 \times 10^{-6}) \end{cases} \quad (31)$$

依托上述参数解析式, 令波动峰值时间小于 20 ms, 得到控制参数约束范围:  $\omega > 152.98 \text{ rad/s}$ 。并从  $\dot{v}(t)$  可知随着参数大于 152.98, 波动幅值逐渐降低, 将  $\omega$  边界代入幅值式得  $v(t') = 0.96\% < 5\%$ , 具备良好控制性能。式(28)的量化与参数约束范围的界定, 既作为先验调节理论依据, 也为后续引入 SAC 算法所需空间的设计奠定了约束条件。

### 2.3.3 控制参数鲁棒性分析

考虑到被控对象模型内部参数摄动的影响, 以及控制策略在该不确定性影响下的响应差异, 可定义被控对象模型为  $\dot{y} + \Delta c \cdot y = b_0 u + f$ ,  $\Delta c$  其物理意义表示输出参数摄变量的影响程度, 则控制模型修改为含不确定性参数的形式  $P(s)$ , 如式(32)所示。

$$P(s) = \frac{Y(s)}{F(s) + b_0 U(s)} = \frac{1}{s + \Delta c} \quad (32)$$

基于图 4 的等效结构图, 可得到含观测修正网络下改进前后系统输出量  $Y(s)$  与参考信号  $R(s)$  的响应传递函数分别为  $\Psi_D(s)$ 、 $\Psi_T(s)$ , 如式(33)所示。

$$\Psi_D(s) = \frac{G(s)P(s)}{N(s) + H(s)P(s)}, \Psi_T(s) = \frac{G_T(s)P(s)}{N_T(s) + H_T(s)P(s)} \quad (33)$$

依据式(15)推得  $Y(s)$ 、 $R(s)$ , 剩余参数如式(34)所示。

$$\begin{cases} Y_T(s) = \omega_c (s^2 + 2\omega_0 s + \omega_0^2) \\ R_T(s) = s^3 + (c + 2\omega_0 + \omega_c) s^2 + \\ (2\omega_0 c + c\omega_c + 2\omega_0 \omega_c + \omega_0^2) s + \omega_0^2 \omega_c \end{cases} \quad (34)$$

当参数摄动量  $\Delta c$  变化时, 两控制策略的零极点分布趋势基于式(33)得到图 7, 两控制策略的动态性能分别以蓝色及红色极点主导。

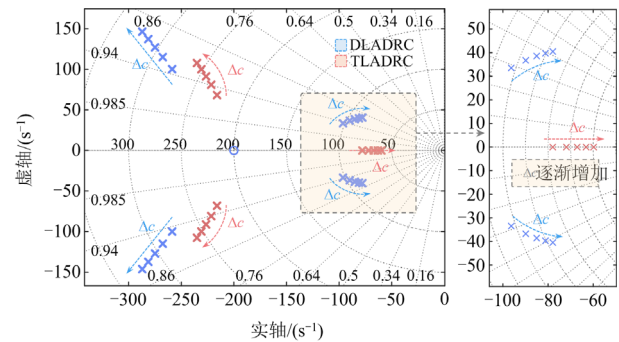


图 7 零极点分布走势图

Fig. 7 Pole zero distribution map

在  $\Delta c$  取值按递增规律变化时, DLADRC 修正了主导极点的位置, 极点分布均向远离虚轴的更优区域迁移, 使系统响应速度相比 TLADRC 得到改善。当  $\Delta c$  参数摄动时, DLADRC 相较于 TLADRC

具备更小的系统阻尼, 进而实现更好的振荡抑制效果与更短的调节过渡时间。但与自抗扰联合整定时, 控制参数与观测重构参数间协同设计复杂, 为寻求两网络间的理想配置, 引入 SAC 算法优化协同整定该过程, 使得所提控制策略能够更好地适应不确定性下的环境, 具备更优越地系统动态响应性能。

### 3 SAC\_DLADRC 控制策略的设计

#### 3.1 SAC\_DLADRC 算法与原理

混储微电网多可控元件的高维异构特性, 以及多时间尺度动态的耦合对超参数配置形成强约束条件<sup>[20]</sup>。因此, 本文采用 SAC 算法, 使其在复杂电力环境中能够探索到更优的策略。SAC 算法伪代码如表 2 所示。

表 2 算法伪代码

Table 2 Algorithmic pseudo-code

算法: SAC 算法伪代码	
01	初始化: Q 网络参数 $\theta_i$ 、策略网络参数 $\phi$ 、温度系数 $\alpha$ , $D \leftarrow \emptyset$ ;
02	设置训练次数 $M'$ 、折扣 $\gamma$ 、学习率 $\lambda$ 和批量 $m$
03	for episode 为 1: $M$ do:
04	for $t = 1: T$ do:
05	观测状态 $s_t$ 从策略 $\pi_\theta(a_t s_t)$ 中采用动作 $a_t$ ;
06	执行动作 $a_t$ 并获得奖励 $r_t$ 和下一状态 $s_{t+1}$ ;
07	储存交互经验样本 $(s_t, a_t, r_t, s_{t+1})$ 到缓冲经验池 $D$ ;
08	从缓冲经验池 $D$ 中随机小批量采样 $m$ 个样本;
09	for $k = 1: K$ do:
10	对 $i = 1, 2$ 最小化损失函数 $J_Q(\theta_i)$ , 更新 Q 网络参数 $\theta_i$ ;
11	根据 KL 散度得策略损失 $J_\pi(\phi)$ , 更新策略网络参数 $\phi$ ;
12	调整温度系数 $\alpha$ ;
13	软更新目标 Q 网络参数 $\tilde{\theta}_i$ ;
14	end for
15	end for
16	end for

在 SAC 算法中, 任意时隙  $t$  下 SAC 的目标函数  $J(\pi)$  和最优策略  $\pi^*$  分别定义为

$$J(\pi) = \sum_{t=0}^T E_{(s_t, a_t) \sim \rho_\pi} [r(s_t, a_t) + \alpha H(\pi_\theta(\cdot|s_t))] \quad (35)$$

$$\pi^* = \arg \max_{\pi_\theta} \sum_{t=0}^T E_{(s_t, a_t) \sim \rho_\pi} [r(s_t, a_t) + \alpha H(\pi_\theta(\cdot|s_t))] \quad (36)$$

式中:  $r(s_t, a_t)$  为  $t$  时刻奖励函数;  $\rho_\pi$  为策略  $\pi$  状态动作轨迹;  $\alpha$  为温度系数用来权衡奖励与熵值;  $H(\pi_\theta(\cdot|s_t)) = E_{a \sim \pi_\theta(\cdot|s_t)} [-\log \pi_\theta(a_t|s_t)]$  为策略  $\pi$  在状态  $s_t$  下的动作熵;  $\theta$  为策略  $\pi$  的网络参数。

算法目标网络所用的软更新方法如式(37)所示, 其余神经网络利用 Adam 优化器进行参数更新。

$$\tilde{\theta}_i = (1 - \xi)\tilde{\theta}_i + \xi\theta_i \quad (37)$$

SAC\_DLADRC 的算法网络结构如图 8 所示。

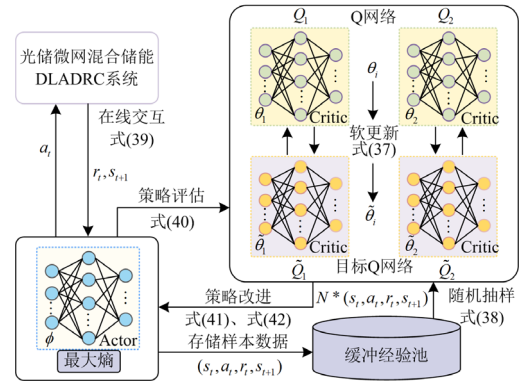


图 8 SAC\_DLADRC 整体算法框架

Fig. 8 SAC\_DLADRC overall algorithm framework

本文 SAC 算法两个网络均设有两层隐藏层, 前后级设有 128 与 64 个神经元, 并用 ReLU 作为激活函数。Actor 网络输入四维状态  $s_t$  输出随机动作性  $a_t$ , Critic 网络根据输入的  $s_t$  和  $a_t$  输出价值。参数  $\theta_i$  通过梯度下降法进行迭代更新, 如式(38)所示。

$$\theta_i \leftarrow \theta_i - \lambda_Q \hat{\nabla}_{\theta_i} J_Q(\theta_i), i \in \{1, 2\} \quad (38)$$

式中:  $\lambda_Q$  为价值网络学习率;  $\hat{\nabla}_{\theta_i}$  为对  $\theta_i$  求导。

SAC 算法用策略迭代方法交替执行策略评估。

$$a_t = \pi(s_t), s_{t+1} = s_t + a_t \quad (39)$$

式中:  $\pi(s_t)$  为策略函数。

$$Q^{\pi_{\text{new}}}(s_t, a_t) = r(s_t, a_t) + \gamma E_{s_{t+1} \sim p} [V^{\pi_{\text{old}}}(s_{t+1})] \quad (40)$$

策略网络  $\pi(s|\phi)$  则利用最小化期望 KL 散度, 以实现网络参数的更新, 可表示为

$$\pi_{\text{new}} = \arg \min_{\pi_\phi} D_{\text{KL}}(\pi_\phi(\cdot|s_t) \parallel \frac{\exp(\alpha^{-1} Q_\theta(s_t, \cdot))}{Z_\theta(s_t)}) \quad (41)$$

$$J(\phi) = E_{s_t \sim D, a_t \sim \pi_\phi} [\alpha \log \pi_\phi(a_t|s_t) - Q_\theta(s_t, a_t)] \quad (42)$$

式中:  $\phi$  为策略网络参数;  $Q_\theta$  为更新后的函数;  $D_{\text{KL}}$  为 KL 散度;  $Z_\theta(s_t)$  为分配函数。

#### 3.2 SAC\_DLADRC 控制策略设计

将深度强化学习 SAC 算法引入 DLADRC 策略的模型如图 9 所示。整体混储微电网系统作为环境, 以电压控制效果作为评估智能体调参的优劣, 通过与微电网的多模态环境交互实现奖励最大化。

状态空间与动作空间如图 9 所示。奖励函数  $R$  以混合储能单元输出电压偏差  $e$  的累积折扣回报最大化作为优化目标, 其表达式如式(43)所示。

$$R = \begin{cases} 130 - 50|e| & 0 \leq |e| \leq 2 \\ 120 - 60|e| & 2 < |e| \leq 10 \\ 80 - 70|e| & 10 < |e| \leq 15 \\ -80|e| & \text{other} \end{cases} \quad (43)$$

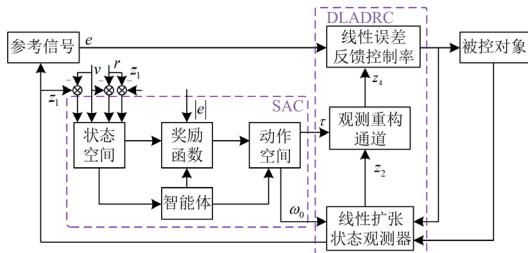


图 9 SAC\_DLADRC 控制策略模型

Fig. 9 SAC\_DLADRC control policy model

### 3.3 SAC\_DLADRC 稳定性分析

定义观测偏差为  $\tilde{z}_i = x_i - \hat{x}_i$ , 其中  $i=1,2$ 。通过带宽  $\omega_0$  对误差的缩放可定义为  $\varepsilon_j = \tilde{z}_j / \omega_0^{j-1}$ , 其中  $j=1,2$ 。观测器偏差估计可重新表述为

$$\begin{bmatrix} \dot{\varepsilon}_1 \\ \dot{\varepsilon}_2 \\ \dot{\varepsilon} \end{bmatrix} = \omega_0 \begin{bmatrix} -2 & 1 \\ -1 & 0 \end{bmatrix} \varepsilon + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{[\dot{f}(x, \omega) - \dot{f}(\hat{x}, \omega)]}{\omega_0} \quad (44)$$

式中:  $f(x, \omega)$  和  $f(\hat{x}, \omega)$  分别为实际与估计的扰动项。

由于  $A_\varepsilon$  是 Hurwitz 稳定的, 即存在一个正定的赫尔米特矩阵  $P$  使  $A_\varepsilon^T P + P A_\varepsilon = -Q$ , 有

$$\dot{V}(\varepsilon) = \varepsilon^T P \dot{\varepsilon} + \dot{\varepsilon}^T P \varepsilon = -\omega_0 \varepsilon_1^2 - \omega_0 \varepsilon_2^2 + \{\omega_0^{-1} [\dot{f}(x, \omega) - \dot{f}(\hat{x}, \omega)] (3\varepsilon_2 - \varepsilon_1)\} \quad (45)$$

由于  $\dot{f}(x, \omega)$  在定义域内满足全局 Lipschitz 条件, 则存在常数  $c$  可满足式(46)。

$$|\dot{f}(x, \omega) - \dot{f}(\hat{x}, \omega)| \leq c \|x - \hat{x}\| \quad (46)$$

由式(44)和式(45)可得

$$3\varepsilon_2 - \varepsilon_1 = 2\varepsilon^T P B_\varepsilon \quad (47)$$

相应地, 基于式(46)和式(47)可得

$$2\varepsilon^T P B_\varepsilon \omega_0^{-1} [\dot{f}(x, \omega) - \dot{f}(\hat{x}, \omega)] \leq 2\varepsilon^T P B_\varepsilon c \omega_0^{-1} \|\tilde{x}\| \quad (48)$$

考虑到  $\|P B_\varepsilon c\|^2 - 2\|P B_\varepsilon c\| + 1 \geq 0$ , 可得

$$\dot{V}(\varepsilon) \leq -\omega_0 \varepsilon_1^2 - \omega_0 \varepsilon_2^2 + (\|P B_\varepsilon c\|^2 + 1) \|\varepsilon\|^2 \quad (49)$$

因此, 当  $\omega_0 > \|P B_\varepsilon c\|^2 + 1$  时, 则存在  $\dot{V}(\varepsilon) < 0$ ,

根据李雅普诺夫渐进稳定性的意义, 有  $\lim_{t \rightarrow \infty} \tilde{z}_i = 0$ 。

基于此, 引入扰动观测重构环节, 可得

$$\lim_{t \rightarrow \infty} \tilde{z}_i \cdot [\tau^{-2}(\tau - m_0)e^{-t\tau^{-1}} + \tau^{-1}m_0\delta(t)] = 0 \quad (50)$$

由式(50)可知, 观测修正网络的重构偏差收敛。随后基于式(14)得到闭环特征方程, 分析 DLESO 整体在系统抗扰性能上的稳定。

$$\lambda_0 s^3 + \lambda_1 s^2 + \lambda_2 s + \lambda_3 = 0 \quad (51)$$

式中:  $\lambda_0 = \tau; \lambda_1 = 2\omega_0\tau + 1; \lambda_2 = \omega_0^2\tau + \omega_0; \lambda_3 = \omega_0^2$ 。

由于  $\tau, m_0, \omega_0$  均为正数, 则  $\lambda_0 - \lambda_3$  都为正数,

满足特征方程系数大于零的条件, 计算可得

$$\lambda_1 \lambda_2 - \lambda_0 \lambda_3 = \tau \omega_0^3 + 3\tau^2 \omega_0^4 + 2\tau^3 \omega_0^5 > 0 \quad (52)$$

因此, 判据表第 1 列系数均为正, 由劳斯判据可知, DLESO 整体对扰动的观测是收敛的, 即  $\exists M_{\tilde{z}} > 0$ , 使得  $\|\tilde{z}\| \leq M_{\tilde{z}}$ 。

假设控制设计目标是使双向 DC-DC 变换器的输出遵循给定的有界参考信号  $v$ , 其导数  $\dot{v}, \dots, v^{(n)}$  也是有界的。设  $e_r = v - \hat{x}_1$ , 结合式(17)可重新表述为

$$\dot{e}_r = \dot{v} - \dot{y} = \dot{v} - K_p(e + \tilde{z}_1) - \tilde{z}_4 \quad (53)$$

在本文中, 为避免线性跟踪微分器对噪声的放大作用未加以应用, 故式(44)可写为

$$\dot{e}_r = A_{\tilde{z}} e_r + B_{\tilde{z}} \tilde{z} \quad (54)$$

式中:  $A_{\tilde{z}} = [-K_p]$ ;  $B_{\tilde{z}} = [-K_p \quad -1]$ ;  $\tilde{z} = [\tilde{z}_1 \quad \tilde{z}_4]^T$ 。

考虑到  $K_p$  使特征多项式  $s + K_p$  满足 Hurwitz 稳定性判据, 同时再由  $\lim_{t \rightarrow \infty} \tilde{z}_i = 0$ , 可知  $\lim_{t \rightarrow \infty} \|B_{\tilde{z}} \tilde{z}\| = 0$ , 最后得到  $\lim_{t \rightarrow \infty} e_r = 0$ 。根据李雅普诺夫理论可知,

DLADRC 是渐进稳定的。由于控制参数改变可视为系统时间尺度的线性映射, 该尺度缩放关系使得稳定性判据不变<sup>[21]</sup>。故 SAC\_DLADRC 在李雅普诺夫意义上渐近稳定。

## 4 数值仿真实验及对比分析

为验证所提控制策略的有效性与优越性, 在 MATLAB/Simulink 平台搭建含光伏储能单元的直流微网系统。混合储能接口 BDC 电路基于 TLADRC、DLADRC、SAC\_DLADRC、DDPG\_DLADRC 4 种策略加以控制, 其参数设置如表 3 和表 4 所示。为保证对比的公平性, 对 4 种控制系统施加多种相同的内外扰动, 利用多项性能指标对比测试 4 种策略的稳定性、抗扰性以及鲁棒性。

表 3 系统模型仿真参数

参数	数值	参数	数值
母线参考电压/V	650	开关频率/kHz	10
直流母线电容/mF	5	等效负载/kW	35
蓄电池端电压/V	200	蓄电池侧电感/mH	1
超级电容端电压/V	150	超级电容侧电感/mH	1

表 4 控制器参数

控制策略	电压环	电流环
TLADRC	$\omega_0 = 550 \text{ rad/s}, \omega_c = 200 \text{ rad/s}$	
DLADRC	$\omega_0 = 550 \text{ rad/s}, \omega_c = 200 \text{ rad/s}, \tau = 2 \times 10^{-4}$	$K_{pi} = 2$ $K_{ii} = 0.01$
DDPG_DLADRC	参数自适应调整	
SAC_DLADRC	参数自适应调整	
LPF 时间常数	0.015	

基于频域法设定准则，按照期望相角裕度为 $60^\circ$ 要求，设计PI控制参数；根据2.3.2节约束条件与带宽法及倍数关系设计 $\omega_0$ 、 $\omega_c$ ；按2.2节的分析确定滞后时间常数 $\tau$ ；LPF时间常数基于低通滤波截止频率进行确定<sup>[22]</sup>。

#### 4.1 SAC\_DLADRC 预训练过程

在智能体预训练前，加入多组不确定性干扰模拟实际工况复杂多变的环境。为保证收敛速度，经多次调试，设置动作临界值为 $\omega_{0\_min} = 200$ 、 $\tau_{min} = 1 \times 10^{-4}$ 。SAC算法智能体参数经综合参考文献超参数取值<sup>[15,20-21]</sup>，并考虑关键参数的敏感性<sup>[7]</sup>，经测试设置如表5和图10所示。

表5 SAC智能体参数

Table 5 SAC agent parameters

参数	数值	参数	数值
Actor 学习率	0.0001	Critic 学习率	0.001
温度系数学习率	0.0001	折扣因子	0.99
经验回放池容量	1 000 000	批量采样容量	128
软更新因子	0.001	优化器	Adam

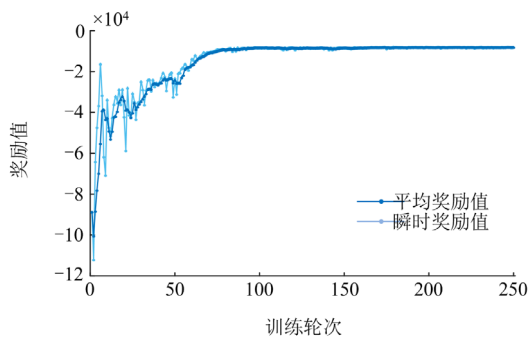


图10 训练奖励曲线及参数自适应

Fig. 10 Training reward curve and parameter adaptation

由图10可知，随训练迭代次数递增，SAC智能体呈现出由探索学习阶段(0~95次)过渡至策略收敛阶段(95~250次)动态演化特性。SAC智能体的输出动作值 $\omega_0$ 由初始的 $[200, +\infty)$ 收敛至 $[200, 1200]$ ， $\tau$ 从 $[10^{-4}, +\infty)$ 收敛至 $[10^{-4}, 2.5 \times 10^{-3})$ ，实现对扰动的动态观测与补偿，以提升控制系统在不确定工况扰动下的适应性与鲁棒性。

#### 4.2 外界因素变化性能分析对比

##### 4.2.1 分频响应验证及负荷功耗波动工况对比

保持光照幅度为 $1000 \text{ W/m}^2$ 恒定，并在0.3s和0.5s分别施加负载跳降与跳增20%投切扰动。储能变换器输出电压对比、混合储能系统控制信号以及评估指标分别如图11、表6所示，表中DDPG\_C与SAC\_C为基于两种算法下的DLADRC。

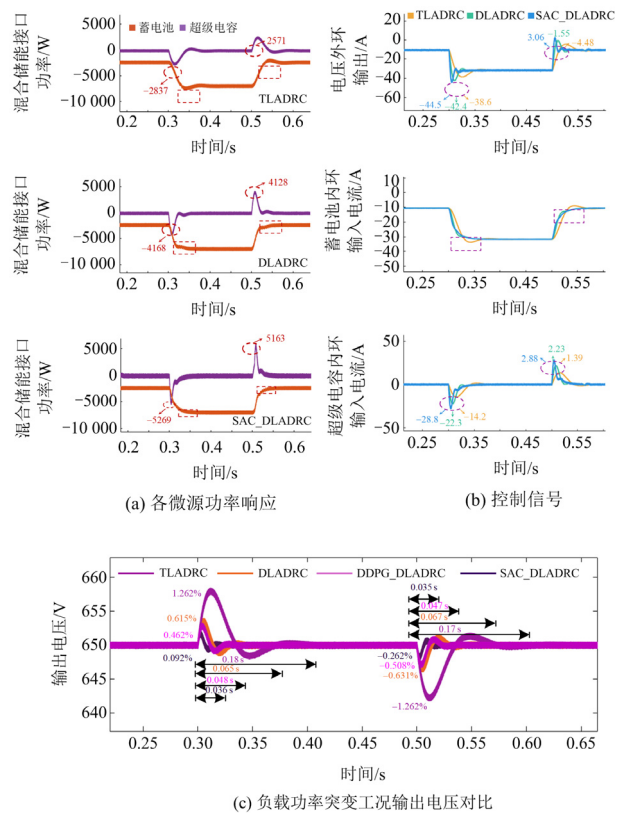


图11 分频响应及输出电压对比

Fig. 11 Comparison of frequency response and output voltage

表6 负载功率突变指标

Table 6 Load power step change indicator

控制策略	电压最大偏差/V	调节时间/ms	IAE/(V·s <sup>-1</sup> )
TLADRC	(8.62, 7.91)	(184.9, 171.4)	4.84
DLADRC	(4.26, 3.79)	(65.7, 67.1)	3.55
DDPG_C	(3.14, 3.04)	(48.2, 47.1)	3.51
SAC_C	(2.39, 1.85)	(36.2, 35.8)	3.46

光伏混储系统的功率分配以及双闭环控制信号如图 11(a)和图 11(b)所示。本文所提 SAC\_DLADRC 策略的蓄电池暂态电流最为平缓达到目标值, 且电池功率变化峰值也最小, 同时超级电容承担了更多的瞬时功率, 补偿了快速变化的高频分量以及因蓄电池响应滞后而残余的波动功率。结果有效验证了图 3 中 LPF 对不同频段信号分离策略的有效性。

由表 6 与图 11(c)的指标综合分析, 横向比较 4 种控制策略可知, 负载功率变化时, SAC\_DLADRC 相比另外 3 种控制策略, 在最大电压偏差、调节时间、绝对误差积分 IAE 等指标上均表现出更优越的效果。与 TLADRC 相比, 双变思想的引入使电压偏差减少了 6.23 V 与 6.06 V, 调节时间缩短了 148.7 ms 与 135.6 ms, 有效抑制了电压暂态过冲现象。

#### 4.2.2 噪声影响下控制策略对比

鉴于微电网的高集成环境下元器件受固有振动与高频噪声污染, 分析噪声抑制有效性如图 12 所示。

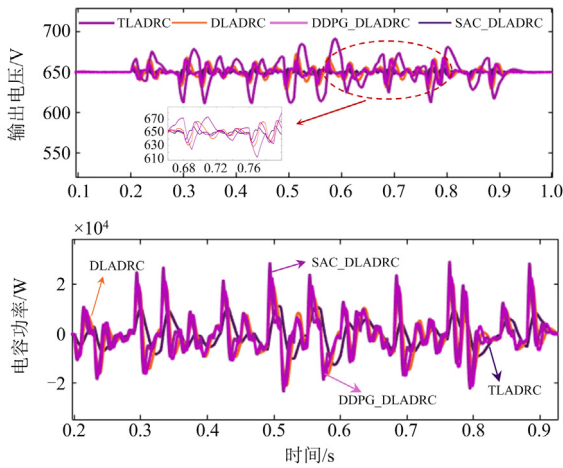


图 12 噪声影响下电压对比

Fig.12 Voltage comparison under noise influence

由图 12 可知, SAC\_DLADRC 下电压仅出现小幅度波动, 最大电压偏差相较于其他 3 种控制策略减少了 27.7 V、15.4 V、7.2 V。从电容功率对比中可知, 噪声的存在加剧了电容充放电的不稳定性。依托对噪声的快速跟踪, SAC\_DLADRC 及时补偿了储能瞬时功率的突变部分, 使电池充放电状态更加迅速平稳, 凸显了“变结构”与“变参数”设计的柔性特征。

#### 4.2.3 光伏输出功率波动

为定量评估系统在多辐照度变化场景下的动态鲁棒性, 本节模拟多种复杂型工况, 对直流母线电压进行仿真分析。

1) 针对阶跃、斜坡型外扰的系统输出波形  
考虑光照强度对微电网影响, 在 0.2 s、0.4 s 与 0.7 s 施加不同程度的阶跃与斜坡型外扰, 不同控制策略光伏输出功率及变换器输出电压如图 13 所示。

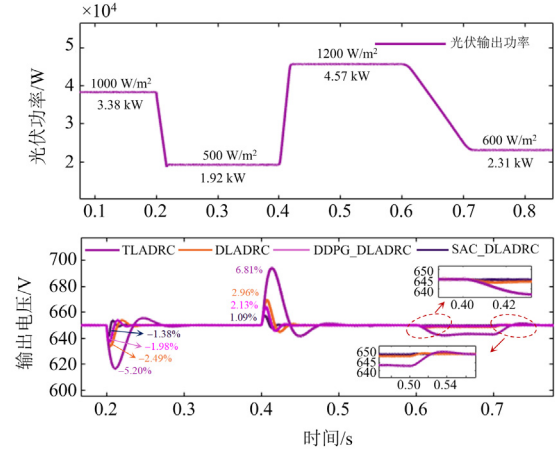


图 13 阶跃、斜坡外扰下电压响应曲线

Fig. 13 Voltage response under step and ramp disturbances

通过表 7 与图 13 可以看出, 在斜坡外扰时段下, TLADRC 无法实现期望值的无差跟踪, 而其余 3 种控制策略仅在起始和撤销时产生波动。同时 SAC\_DLADRC 最大电压偏差与调节时间均最小, 相较 TLADRC 偏差减少 6.77 V, 时间缩短 70 ms。此外, 辐照度在阶跃与斜坡变化时, SAC\_DLADRC 策略下的储能变换器 IAE 值最小, 输出电压偏差量可在最短过渡时间内收敛至允许误差区间。

表 7 阶跃、斜坡外扰下电压指标

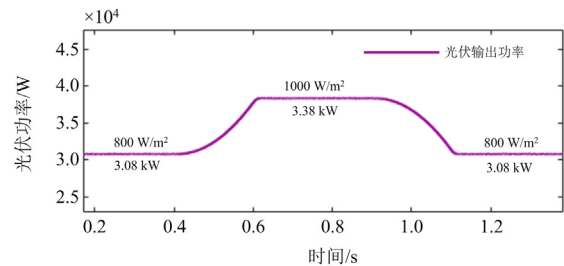
Table 7 Step and ramp disturbance voltage indicators

控制策略	电压最大偏差/V	调节时间/ms	IAE/(V·s <sup>-1</sup> )
TLADRC	(34.1, 44.5, 6.81)	(101.6, 103.6, 183.1)	7.12
DLADRC	(16.3, 19.9, 0.99)	(47.8, 49.6, 125.8)	3.99
DDPG_C	(13.5, 14.6, 0.39)	(37.3, 38.4, 119.3)	3.75
SAC_C	(9.6, 7.6, 0.04)	(28.5, 27.7, 113.1)	3.72

2) 针对加速度型外扰的系统输出波形

引入辐照度加速度型外扰等效表征微电网所受时变激励, 光伏功率变化与母线电压如图 14 所示。

通过表 8 与图 14 可以看出, 在加速度上升型外扰阶段, 其最大输出电压偏差为 0.5 V, 相较另外 3 种



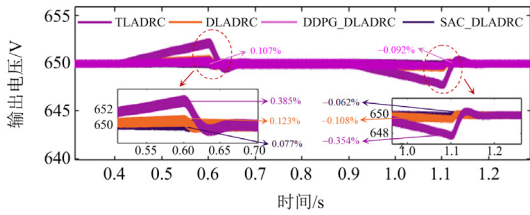


图 14 加速度外扰下电压响应曲线

Fig. 14 Voltage response curve under acceleration disturbance

表 8 加速度外扰下电压指标

Table 8 Acceleration external disturbance voltage index

控制策略	电压最大偏差/V	调节时间/ms	IAE/(V·s <sup>-1</sup> )
TLADRC	(2.51, 2.30)	(295.2, 278.4)	5.09
DLADRC	(0.81, 0.71)	(228.3, 226.7)	3.65
DDPG_C	(0.69, 0.59)	(216.2, 221.4)	3.59
SAC_C	(0.50, 0.42)	(205.6, 215.6)	3.47

策略分别减少了 2.01 V、0.31 V 和 0.19 V。调节时间各缩短了 89.6 ms、22.7 ms 和 10.6 ms。同时在加速度上升型外扰下，同样能在更小时间和偏差量下恢复稳定。因此 SAC\_DLADRC 控制下更适用于此类复杂功率扰动，具备更优越的扰动抑制性能。

### 3) 针对正弦型外扰的系统输出波形

对于控制系统引入辐照度正弦型外扰等效表征微电网所受的正弦时变激励，光伏功率变化以及储能变换器输出电压如图 15 所示。

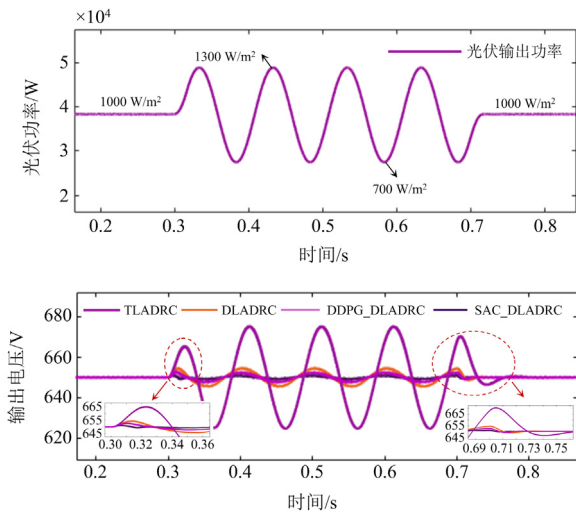


图 15 正弦外扰下电压响应曲线

Fig. 15 Voltage response curve under sinusoidal disturbance

由表 9 与图 15 的指标综合分析可知，混储变换器输出侧呈现正弦周期性暂态波动与电压幅值偏移现象。相较另外三种控制策略，SAC\_DLADRC 基于全局收益最优导向的策略框架，电压波动最为平缓，中部以及左右两侧峰谷电压偏差相比较基础的

TLADRC 策略分别减少了为 23.75 V、14.24 V 和 19.26 V。同时，SAC\_DLADRC 在正弦时变扰动下的调节速度最快，历时 0.4 s 扰动得到平抑且 IAE 性能指标最小，凸显了该控制策略高精度扰动观测与快速动态补偿的协同控制优势。

表 9 正弦外扰下电压指标

Table 9 Voltage indicators under sinusoidal disturbance

控制策略	两侧峰谷偏差/V	中部峰谷偏差/V	IAE/(V·s <sup>-1</sup> )
TLADRC	(15.84, 20.54)	25.14	10.83
DLADRC	(4.78, 4.81)	4.86	4.54
DDPG_C	(2.88, 2.82)	2.83	3.98
SAC_C	(1.60, 1.28)	1.39	3.79

综合工况误差平方积分值(composite working condition integral of square error, CISE)能够综合表征多种工况下系统的响应速度及抗扰能力。该指标数值越小，表明系统受扰后过渡时间越短，扰动抗击打能力越强。不同工况 CISE 分布如图 16 所示。

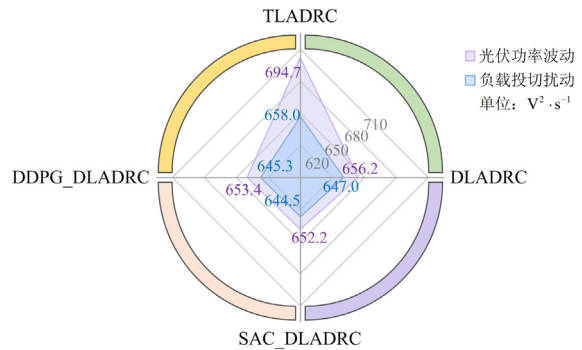


图 16 不同扰动工况下 CISE 分布

Fig. 16 CISE distribution under different working conditions

## 4.3 内部参数摄动下动态性能分析对比

### 4.3.1 控制器参数的随机摄动

保持 BDC 内部参数不变，对自抗扰增益参数在所选值  $\pm 10\%$  内进行摄动。同时令  $\Delta\omega = \omega_0/\omega_c$ ，得到改进前后自抗扰参数摄动 CISE 指数曲面与衡量控制策略的调节时间指标如图 17 所示。由于深度强化学习随机探索策略的引入，可以赋予自抗扰本身控制器参数变化下的主动抗扰特性，故不再示意 SAC\_DLADRC 与 DDPG\_DLADRC 两种控制策略。

由图 17 可知，在控器参数的摄动下，DLADRC 策略倾角相比 TLADRC 减少了 8°，且距离底部明显小于传统控制策略，表明观测修正网络的引入使控制策略性能得到提升。当  $\Delta\omega$  最小时，两种控制系统响应均较慢。与传统策略相比较，DLADRC 策略整体调节时间更短，且波动更为平缓。则 DLADRC 策略的鲁棒性和抗扰要比 TLADRC 策略更加优越。

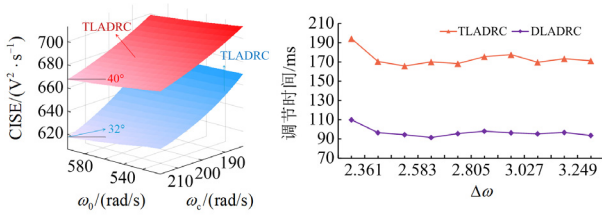


图 17 控制器参数摄动鲁棒性及性能指标曲线

Fig. 17 Controller parameter perturbation robustness and performance indicator curves

#### 4.3.2 储能变换器参数的随机摄动

为说明 SAC 算法对控制器鲁棒性的改善,因此考虑变换器内部器件参数会因材料劣化、长期运行应力等因素产生时变漂移现象,保证控制器参数恒定,对混合储能变换器内部电感  $L(L_b、L_{sc})$  和母线电容  $C_{dc}$  在标称值  $\pm 15\%$  内进行变化。记录 4 种系统各工况下参数摄动 CISE 指数曲面和平均最大电压偏差与参数摄动量间数值分布,如图 18 所示。

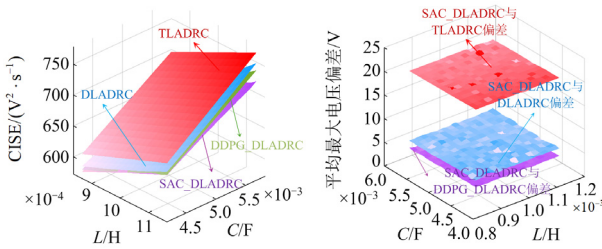


图 18 变换器参数摄动鲁棒性及性能指标分析

Fig. 18 Analysis of converter parameter perturbation robustness and performance indicators

由图 18 可知,相较于 TLADRC 与 DLADRC,在参数不规则摄动下,具备自趋优机制和观测网络的 SAC\_DLADRC 系统数值分布距离底部水平面最近,平面倾角最小为  $40^\circ$ 。同时在参数摄动范围内, SAC\_DLADRC 与 TLADRC 的电压偏差最大在 20 V 附近波动,改善效果最为良好,凸显了“变结构与变参数”双维驱动优势。相比 DDPG\_DLADRC 策略与 DLADRC 策略,本文所提控制策略在系统内部参数摄动下表现出更强的抗扰性和鲁棒性特质。

## 5 结论

本文以储能侧变换器为研究对象,在传统自抗扰控制基础上融入“变结构、变参数”思想双维驱动,提出一种 SAC\_DLADRC 控制策略。通过理论与仿真的多维度对比分析,得到以下结论。

1) SAC\_DLADRC 策略因其扰动观测重构网络的嵌入与参数自趋优的控制架构,相比传统控制策略对扰动的平均观测精度偏差提高 20.6%且平均响

应时间提升 2.3 倍。并应用于混储侧变换器中,突破耗能端应用场景,面对多类复杂函数扰动相较 TLADRC 更具优势,显著改善了自抗扰系统性能。

2) 在光伏功率波动、低压负载投切以及内部参数摄动的内外部扰动工况下, SAC\_DLADRC 策略可有效输出令不确定波动更为有效衰减的控制信号且调节系统功率平衡,相比 TLADRC 策略、DLADRC 策略和 DDPG\_LADRC 策略展现出更优越的稳压性能与更广泛的适用范围。

综上,本文所提 SAC\_DLADRC 策略更适合于多耦合变量、强非线性特性、存在严重未知扰动交织的新能源微电网运行场景,可以为各类储能微电网智能化的应用与分析提供更多思路。

## 参考文献

- [1] 李鹏, 刘念, 胡秦然, 等. “新型电力系统数字化关键技术综述”专辑评述[J]. 电力系统自动化, 2024, 48(6): 1-12. CHEN Fufeng, YANG Liming, SONG Guobing, et al. Overview and prospect of key technologies for active detection protection[J]. Power System Protection and Control, 2023, 51(15): 175-186.
- [2] 李霞林, 郭力, 王成山, 等. 直流微电网关键技术研究综述[J]. 中国电机工程学报, 2016, 36(1): 2-17. LI Xialin, GUO Li, WANG Chengshan, et al. Key technologies of DC microgrids: an overview[J]. Proceedings of the CSEE, 2016, 36(1): 2-17.
- [3] 岳改丽, 张哲, 杜光辉, 等. 储能电池双向 DC-DC 变换器控制策略[J]. 电工技术学报, 2025, 40(20): 6618-6629. YUE Gaili, ZHANG Zhe, DU Guanghui, et al. Control strategies of bi-directional DC-DC converters for energy-storage batteries[J]. Transactions of China Electrotechnical Society, 2025, 40(20): 6618-6629.
- [4] GORBACHEV S, GUO J, MANI A, et al. MPC-based LFC for interconnected power systems with PVA and ESS under model uncertainty and communication delay[J]. Protection and Control of Modern Power Systems, 2023, 8(4): 1-17.
- [5] 侯传羽, 李岚, 赵楠, 等. DFIG 网侧变换器反馈线性化与滑模控制[J]. 南方电网技术, 2021, 15(3): 47-54. HOU Chuanyu, LI Lan, ZHAO Nan, et al. Feedback linearization and sliding mode control of DFIG grid-side converter[J]. Southern Power System Technology, 2021, 15(3): 47-54.
- [6] 孙永辉, 孟云帆, 葛磊蛟, 等. 人工智能赋能微电网运行优化的应用及展望[J]. 高电压技术, 2023, 48(6): 2239-2252. SUN Yonghui, MENG Yunfan, GE Leijiao, et al. Application and prospect of artificial intelligence empowering microgrid operation optimization[J]. High

- Voltage Engineering, 2023, 48(6): 2239-2252.
- [7] 李若青. 基于深度强化学习的列车节能驾驶控制方法[D]. 北京: 北京交通大学, 2024.  
LI Ruoqing. Application of train energy-saving driving control method based on deep reinforcement learning[D]. Beijing: Beijing Jiaotong University, 2024.
- [8] GAO Z. Scaling and bandwidth-parameterization based controller tuning[C] // Proceedings of the 2003 American Control Conference, June 4-6, 2003, Denver, CO, USA: 4989-4996.
- [9] 周雪松, 王博, 马幼捷, 等. 含二阶扰动补偿的交错并联变换器自抗扰控制[J]. 电机与控制学报, 2023, 27(12): 159-170.  
ZHOU Xuesong, WANG Bo, MA Youjie, et al. Active disturbance rejection control of interleaved parallel converters with second-order disturbance compensation[J]. Electric Machines and Control Journal, 2023, 27(12): 159-170.
- [10] CHEN S, XUE W, ZHONG S, et al. On comparison of modified ADRCs for nonlinear uncertain systems with time delay[J]. Science China Information Sciences, 2018, 61.
- [11] TAO Long, WANG Ping, WANG Yifeng, et al. Variable structure ADRC-based control for load-side buck interface converter: formation, analysis, and verification[J]. IEEE Transactions on Industrial Electronics, 2022, 96(6): 6236-6246.
- [12] YANG Z, JIA J, SUN X, et al. An enhanced linear ADRC strategy for a bearingless induction motor[J]. IEEE Transactions on Transportation Electrification, 2022, 8(1): 1255-1266.
- [13] HE W, ZHENG Y, WEN J, et al. Path following control of parafoil system based on SGLOS and improved ADRC tuned by MSMPA[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2024, 71(8): 3895-3899.
- [14] ZHAO C, ZUO Y, WANG H, et al. Online-trained radial basis function neural network compensator for current harmonics suppression of electric drives[J]. IEEE Transactions on Industrial Electronics, 2024, 71(12): 15488-15498.
- [15] 周雪松, 韩静, 马幼捷, 等. 基于 DQN 算法的直流微电网负载接口变换器自抗扰控制策略[J]. 电力系统保护与控制, 2025, 53(1): 95-103.  
ZHOU Xuesong, HAN Jing, MA Youjie, et al. Active disturbance rejection control strategy for load interface converter in DC microgrid based on DQN algorithm[J]. Power System Protection and Control, 2025, 53(1): 95-103.
- [16] 周雪松, 王馨悦, 马幼捷, 等. 基于信息熵量化评价的深度强化学习储能微网柔性补偿自抗扰稳压技术[J]. 电力系统保护与控制, 2025, 53(23): 127-138.  
ZHOU Xuesong, WANG Xinyue, MA Youjie, et al. A flexible compensation active-disturbance-rejection voltage stabilization technology for energy storage microgrid based on deep reinforcement learning with information entropy quantitative evaluation[J]. Power System Protection and Control, 2025, 53(23): 127-138.
- [17] WANG Y, TAO L, WANG P, et al. Improved linear ADRC for hybrid energy storage microgrid output-side converter[J]. IEEE Transactions on Industrial Electronics, 2022, 69(9): 9111-9120.
- [18] LIU F, LI Y, CAO Y, et al. A two-layer active disturbance-rejection controller design for load frequency control of interconnected power system[J]. IEEE Transactions on Power Systems, 2015, 31(4): 3320-3321.
- [19] 周杰, 皇金锋, 黄红杰, 等. T 型三电平储能变流器直流母线电压改进自抗扰控制[J]. 电力系统保护与控制, 2024, 52(21): 81-91.  
ZHOU Jie, HUANG Jinfeng, HUANG Hongjie, et al. Improved active-disturbance-rejection control for DC bus voltage of T-type three-level energy storage converter[J]. Power System Protection and Control, 2024, 52(21): 81-91.
- [20] 李柏楠, 赵津蔓, 韩肖清, 等. 基于双智能体深度强化学习的电力系统静态安全预防控制方法[J]. 中国电机工程学报, 2023, 43(5): 1818-1831.  
LI Boyu, ZHAO Jinman, HAN Xiaoqing, et al. Static security preventive control method for power systems based on double-agent deep reinforcement learning[J]. Proceedings of the CSEE, 2023, 43(5): 1818-1831.
- [21] 马幼捷, 刘熠铭, 周雪松, 等. 微网储能侧 DC-DC 变换器的强化学习自抗扰控制策略[J]. 太阳能学报, 2025, 46(3): 63-72.  
MA Youjie, LIU Yiming, ZHOU Xuesong, et al. Reinforcement learning active disturbance rejection control strategy for DC-DC converter on the energy storage side of microgrid[J]. Acta Energetica Solaris Sinica, 2025, 46(3): 63-72.
- [22] LATRECHE, ABDERREZZAK, et al. Control and integration of a hybrid PV-battery-supercapacitor system for grid stability[J]. Engineering, Technology & Applied Science Research, 2025, 15(4): 24334-24343.

收稿日期: 2025-08-06; 修回日期: 2025-12-06

作者简介:

周雪松(1964—), 男, 博士, 教授, 研究方向为新能源发电、新型电力系统控制; E-mail: zxsmyj@126.com

马培铭(2002—), 男, 硕士研究生, 研究方向为新能源发电与控制技术; E-mail: m1436290178@163.com

练继建(1965—), 男, 博士, 教授, 研究方向为水利水电、新能源发电。

(编辑 石晋美)