

DOI: 10.19666/j.rlfed.202406175

# 超超临界机组协调控制系统误差 自抗扰控制与优化

黄从智<sup>1,2</sup>, 蒋茜茜<sup>1</sup>, 谭祥帅<sup>3</sup>

(1. 华北电力大学控制与计算机工程学院, 北京 102206;  
2. 电站能量传递转化与系统教育部重点实验室, 北京 102206;  
3. 西安热工研究院有限公司, 陕西 西安 710054)

**[摘要]** 超超临界机组由于其高参数、高效率的优点, 已成为电网深度调频调峰、抑制电网波动的有力支撑, 超超临界机组控制策略的优化改造对于电网安全稳定运行具有重要意义。针对超超临界机组协调控制系统优化问题, 提出一种基于误差自抗扰控制策略和强化学习算法的智能控制策略。首先, 在基于误差的自抗扰控制策略框架中, 根据单元机组机炉耦合过程运行特性, 简化机炉耦合过程被控对象模型, 设计扩张状态观测器来实时估计并补偿机组的未建模动态特性和外部扰动。其次, 构建奖励函数并利用柔性演员-评论家算法来实现控制器参数自适应调整。最后, 基于某超超临界 1 000 MW 二次再热机组实际历史运行数据仿真验证了所提控制策略的有效性。

**[关键词]** 超超临界机组; 误差自抗扰控制; 强化学习; 参数自适应调整

**[引用本文格式]** 黄从智, 蒋茜茜, 谭祥帅. 超超临界机组协调控制系统误差自抗扰控制与优化[J]. 热力发电, 2025, 54(4): 104-116. HUANG Congzhi, JIANG Xixi, TAN Xiangshuai. Error-based active disturbance rejection control and optimization of coordinated control system for ultra-supercritical unit[J]. Thermal Power Generation, 2025, 54(4): 104-116.

## Error-based active disturbance rejection control and optimization of coordinated control system for ultra-supercritical unit

HUANG Congzhi<sup>1,2</sup>, JIANG Xixi<sup>1</sup>, TAN Xiangshuai<sup>3</sup>

(1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China;  
2. Key Laboratory of Power Station Energy Transfer Conversion and System, Beijing 102206, China;  
3. Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China)

**Abstract:** Due to its high parameters and high efficiency, ultra supercritical units have become a powerful support for deep frequency regulation, peak shaving, and suppression of power grid fluctuations. The optimization and transformation of control strategies for ultra supercritical units are of great significance for the safe and stable operation of the power grid. Aiming at the optimization problem of coordinated control system for ultra supercritical units, an intelligent control strategy based on error self-disturbance rejection control strategy and reinforcement learning algorithm is proposed. Firstly, in framework of the error-based self-disturbance rejection control strategy, the controlled object model of the machine furnace coupling process is simplified according to operating characteristics of the unit's turbine-boiler coupled process, and an extended state observer is designed to estimate and compensate for the unmodeled dynamic characteristics and external disturbances of the unit in real time. Secondly, a reward function is constructed and the flexible actor-critic algorithm is used to achieve self-adaptive adjustment of controller parameters. Finally, the effectiveness of the proposed control strategy is verified through simulation based on actual historical operating data of a certain ultra supercritical 1 000 MW secondary reheating unit.

收稿日期: 2024-06-13

基金项目: 国家自然科学基金项目(62373149); 中央高校基本科研业务费学科交叉创新专项(2023JC001)

Supported by: National Natural Science Foundation of China (62373149); Fundamental Research Funds for the Central Universities (2023JC001)

第一作者简介: 黄从智(1982), 男, 博士, 教授, 主要研究方向为智能控制理论及其新型电力系统应用, hcz190@ncepu.edu.cn.

**Key words:** ultra-supercritical unit; error-based active disturbance rejection control; reinforcement learning; self-adaptive adjustment of controller parameters

能源电力低碳转型是实现碳达峰、碳中和的关键,事关我国经济社会发展全局<sup>[1]</sup>。然而由于新能源发电具有间歇性、波动性等特性,在并网过程中,电网不得不承担消纳新能源所造成的冲击,导致其稳定性降低<sup>[2]</sup>。在这一过程中,超超临界发电机组是能源高效利用的先锋<sup>[3]</sup>,一般采用直流锅炉<sup>[4]</sup>,由于其具有发电效率高<sup>[5]</sup>、出力形式稳定等优势,已成为电网深度调频调峰、抑制电网波动的有力支撑。因此,为了更好地配合电网负荷指令并维持机组内部的安全和稳定,迫切需要优化改造超超临界机组协调控制系统控制策略<sup>[6]</sup>。

协调控制系统是满足机组运行内外能量平衡的关键,承担着协调锅炉和汽轮机以响应电网调度指令的重要任务<sup>[7]</sup>。超超临界机组协调控制系统具有大惯性、大延迟、非线性的特点,传统的PID控制方法难以满足新型电力系统背景下的控制要求<sup>[8]</sup>。因此,采用智能控制方法来弥补这些缺点,提升机组运行的效率、稳定性和可靠性显得更加迫切和必要。

在现代工程领域,复杂性、非线性和外部扰动的挑战日益凸显<sup>[9]</sup>,由韩京清提出的自抗扰控制(active disturbance rejection control, ADRC)已经在诸多领域引起了广泛的关注<sup>[10]</sup>。ADRC的核心思想是以简单的积分串联型为标准型,把系统动态中不同于标准型的部分视为总扰动,利用扩张状态观测器(extended state observer, ESO)实时估计并补偿总扰动,从而把充满扰动、不确定性和非线性的被控对象还原为标准的积分串联型,使控制系统的设计从复杂到简单,从抽象到直观。但这种非线性控制的结构较为复杂,理论分析较难,控制参数过多,不利于工程应用。因此,高志强应用频率尺度的概念将ADRC线性化,其线性ESO和比例-微分(PD)控制分别估计和补偿干扰,使ADRC整定参数降低为2个<sup>[11-12]</sup>。最近,基于误差的自抗扰控制(error-based active disturbance rejection control, EADRC)被提出<sup>[13]</sup>,EADRC采用参考跟踪误差作为其直接输入,综合ADRC和PID的优点,将二自由度自抗扰控制设计重构为一自由度形式,使其在工业应用中成为一种有吸引力的方案。EADRC的控制性能已在风力涡轮机功率设定值跟踪俯仰控制系统<sup>[14]</sup>和风电机组恒功率控制系统中得到验证<sup>[15]</sup>。

控制器参数调整是确保控制系统正常运行,实现期望性能并适应不确定性和变化的关键步骤,通过合适的参数设置,可以使控制系统的性能达到最优。参数整定方法分为固定和自动,第一种类型包括群智能优化算法,如粒子群算法<sup>[16]</sup>和遗传算法<sup>[17]</sup>。然而,这些算法可能存在收敛速度慢和易陷入局部最优等缺点。第二种类型包括模糊自适应控制<sup>[18]</sup>和神经网络自适应控制<sup>[19]</sup>等算法来实现参数自适应调整。然而,这些方法可能会受到噪声、估计误差、收敛性等因素的影响,不能有效应对环境的突发变化。因此,需要将控制器参数调整智能化,使其能够在复杂多变的工作环境中自动地调整控制器参数,以提高控制系统性能和适应性。

强化学习旨在让智能体通过尝试不同的行动来与环境互动,根据环境的反馈来调整其行动策略,以达到最优的长期累积奖励,其在许多领域已有广泛的应用,如机器人控制<sup>[20]</sup>、过程控制<sup>[21]</sup>等。此外,强化学习也广泛应用于电力系统中,Hou等人<sup>[22]</sup>通过采用TD3算法获得控制器的最优参数,提升热电联产电厂的运行灵活性。Meng等人<sup>[23]</sup>通过采用DDPG算法智能调节控制器参数,提高直流微电网直流固态变压器系统运行效率,Duan等人<sup>[24]</sup>通过采用DQN算法和DDPG算法实现电网运行自主电压控制,Dong等人<sup>[25]</sup>通过融合一种复合体验回放策略的DDPG算法,优化风电场的总发电量。尽管这些算法在电力系统控制及优化问题中表现出色,但在训练的稳定性、样本效率以及超参数调整等方面都存在一些挑战。然而,柔性演员-评论家(soft actor-critic, SAC)算法通过采用最大熵原则来提高算法探索能力,其效率更高,收敛性更好,性能更稳定<sup>[26-27]</sup>。

因此,针对超超临界1000 MW机组协调控制系统的性能优化问题,本文提出一种基于EADRC和SAC算法的解决方案。首先,在EADRC框架中,根据机炉耦合过程运行特性,确定系统的输入输出变量,简化机炉耦合过程被控对象模型并设计其EADRC控制策略。其次,在奖励函数设计中主要考虑系统快速跟踪性能、运行过程各项重要参数的稳定和燃料量的投入。然后,利用SAC算法最大化奖励函数获得EADRC最优参数。最后,通过仿真

实验验证了所提方案的有效性。

# 1 协调控制系统模型及其自抗扰控制器设计

## 1.1 超超临界机组协调控制系统被控对象

超超临界机组协调控制系统是由直流锅炉、汽轮机和发电机共同构成的多输入多输出的被控对象。超超临界机组的工作原理是：煤炭在磨煤机中被磨成颗粒大小均匀的煤粉，煤粉经空气预热器加热到一定的温度后由一次风吹入炉膛内部燃烧，通过煤粉的燃烧将水加热成水蒸气；水蒸气经过热器

进一步加热之后形成过热蒸汽进入汽轮机中，通过蒸汽压力带动汽轮机转动；汽轮机转动之后通过连轴带动发电机转子转动，从而使发电机产生电能；最后，由主变压器将各发电机发出的电能整合升压后进入电网进行传输。

因此，基于机组整体运行机理和动态特性，超超临界机组协调控制系统模型可简化为图 1 所示的三输入三输出系统，其控制变量分别为燃料量 ( $M$ , t/h)、阀门开度 ( $\mu_T$ , %) 和给水流量 ( $W$ , t/h)，被控变量分别为主蒸汽压力 ( $p_T$ , MPa)、有功功率 ( $P_E$ , MW) 和分离器出口温度 ( $T$ ,  $^{\circ}C$ )。

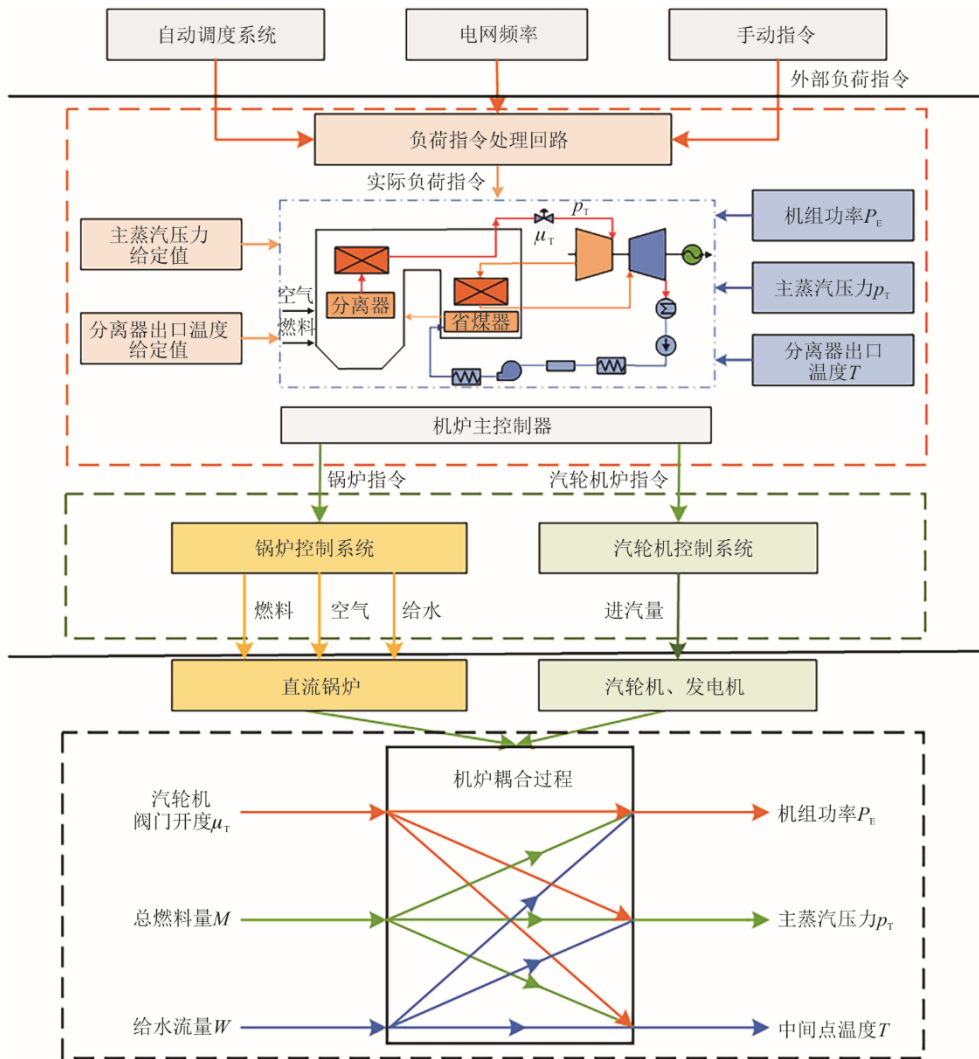


图 1 超超临界机组协调控制系统模型简化结构

Fig.1 Simplified structure of coordinated control system modeling for ultra-supercritical unit

## 1.2 基于误差的自抗扰控制器设计及稳定性分析

### 1.2.1 控制器设计

基于 EADRC 策略的超超临界机组协调控制系统如图 2 所示。结合实际工程应用，综合考虑限幅、

限速等因素并根据专家经验设定前馈控制环节<sup>[30-31]</sup>。EADRC 将系统内部干扰和外部干扰的总和等效为一种总扰动，通过 ESO 来实时在线观测并采用 PD 控制部分将其补偿掉。

由图2可以看出, 该控制系统包括EADRC1、EADRC2和EADRC3共3个控制器, 这3个控制器分

别对应有功功率通道、主蒸汽压力通道和分离器出口温度通道。

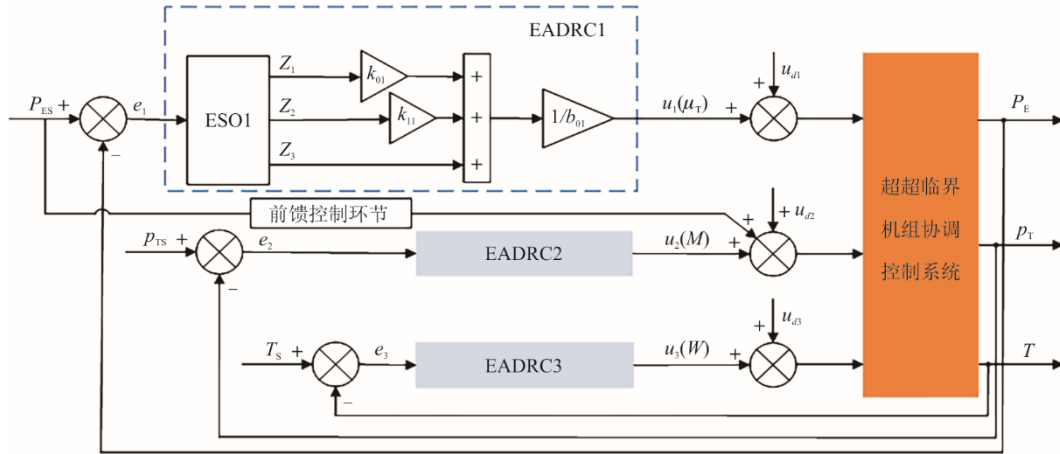


图2 基于EADRC策略的超超临界机组协调控制系统  
Fig.2 Coordinated control system for ultra-supercritical unit based on EADRC strategy

以有功功率通道的EADRC1控制器为例, EADRC的设计过程如下。根据实际工程应用, 有功功率子系统动态调节过程可以描述为一个二阶系统:

$$\ddot{y} = d + bu \quad (1)$$

式中:  $y$ 为系统输出变量;  $d$ 为系统不确定性扰动;  $b$ 为系统控制输入增益;  $u$ 为系统控制输入变量。

将系统的跟踪误差定义为 $e=r-y$ ,  $r$ 为系统参考输入, 则系统的误差动力学为:

$$\begin{aligned} \ddot{e} &= \ddot{r} - \ddot{y} = \ddot{r} - bu - d = \\ &= -b_0u - \underbrace{bu + b_0u + \ddot{r} - d}_f = \\ &= -b_0u + f \end{aligned} \quad (2)$$

式中:  $b_0$ 为控制输入增益 $b$ 的最佳估计;  $f$ 为扰动项。

为了生成控制律, 在式(2)的两端进行增项, 则跟踪误差的动力学为:

$$\ddot{e} + k_1\dot{e} = -b_0u + \underbrace{k_1\dot{e} + f}_F = -b_0u + F \quad (3)$$

式中:  $F$ 为系统总扰动, 其值可通过ESO估计。

假设, 总扰动 $F$ 可以通过ESO估计和控制律补偿来得到抑制, 即 $\hat{F} = F$ , 则控制律可以表示为:

$$u = \frac{u_0 + F}{b_0}, \quad u_0 = k_0e \quad (4)$$

则系统的跟踪误差动力学可以表示为:

$$\ddot{e} + k_1\dot{e} + k_0e = F - \hat{F} = 0 \quad (5)$$

将控制器极点统一配置在 $s$ 平面左半实轴 $-\omega_c$  (控制器带宽) 处<sup>[28]</sup>, 得到控制器增益为:

$$k_0 = \omega_c^2, \quad k_1 = 2\omega_c \quad (6)$$

将系统的跟踪误差定义为新的状态向量 $z = [z_1 \ z_2 \ z_3]^T \triangleq [e \ \dot{e} \ F]^T$ , 则式(3)可以用状态空间

表达式表示为:

$$\dot{z} = Az + Bu + EF \quad (7)$$

$$\text{其中, } A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -k_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ -b_0 \\ 0 \end{bmatrix}, \quad E = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

则相应的基于误差的三阶ESO可以表示为:

$$\dot{\hat{z}} = A\hat{z} + Bu + L(e - \hat{e}) \quad (8)$$

其中,  $\hat{z} = [\hat{z}_1 \ \hat{z}_2 \ \hat{z}_3]^T \triangleq [\hat{e} \ \hat{\dot{e}} \ \hat{F}]^T$ , 代表状态变量 $z$ 的观测值,  $L = [l_1 \ l_2 \ l_3]^T$ 代表观测器的增益。

结合式(4)和式(8), 得到闭环系统误差方程:

$$\dot{\hat{z}}_e = A_e \hat{z}_e + Le \quad (9)$$

其中,  $\hat{z}_e = [e - \hat{e} \ \dot{e} - \hat{\dot{e}} \ F - \hat{F}]^T$ ,

$$A_e = \begin{bmatrix} -l_1 & 1 & 0 \\ -l_2 - k_0 & -k_1 & 1 \\ -l_3 & 0 & 0 \end{bmatrix}.$$

将观测器的极点统一配置在 $s$ 平面的左半实轴 $-\omega_o$  (观测器带宽) 处, 将得到观测器增益为:

$$\begin{aligned} l_1 &= 3\omega_o - k_1 \\ l_2 &= 3\omega_o^2 - k_0 - l_1k_1 \\ l_3 &= \omega_o^3 \end{aligned} \quad (10)$$

因此, 超超临界机组协调控制系统需要整定的控制器参数为 $\omega_{o1}$ 、 $\omega_{c1}$ 、 $b_{01}$ 、 $\omega_{o2}$ 、 $\omega_{c2}$ 、 $b_{02}$ 、 $\omega_{o3}$ 、 $\omega_{c3}$ 、 $b_{03}$ 。

### 1.2.2 稳定性分析

为验证 EADRC 下的闭环系统的稳定性, 基于奇异摄动理论对其进行理论分析<sup>[13]</sup>。根据 1.2.1 中的描述, 观测器估计误差方程可表示为:

$$\dot{\mathbf{q}} = \mathbf{H}\mathbf{q} + \mathbf{h}\eta \quad (11)$$

其中,  $\mathbf{q} = [e - \hat{e} \quad \dot{e} - \dot{\hat{e}} \quad F - \hat{F}]^T$ ,  $\eta = \dot{F}$ ,

$$\mathbf{H} = \begin{bmatrix} -l_1 & 1 & 0 \\ -l_2 - k_0 & -k_1 & 1 \\ -l_3 & 0 & 0 \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

通过极点配置, 式(11)的特征方程为:

$$\det(s\mathbf{I} - \mathbf{H}) = s^3 + p_1s^2 + p_2s + p_3 = (s + \omega_0)^3 \quad (12)$$

当  $[p_1 \ p_2 \ p_3] = [3\omega_0 \ 3\omega_0^2 \ \omega_0^3]$  时, 估计误差矩阵  $\mathbf{H}$  的特征值为:

$$\lambda_i\{\mathbf{H}\} = -\omega_0, \quad \forall i = 1, 2, 3 \quad (13)$$

因此, 式(12)的系数可以表示为:

$$\begin{cases} s_1 = k_1 + l_1 = 3\omega_0 \Rightarrow l_1 = 3\omega_0 - k_1 \\ s_2 = k_1l_1 + k_0 + l_2 = 3\omega_0^2 \Rightarrow l_2 = 3\omega_0^2 - k_1l_1 - k_0 \\ s_3 = l_3 + 0 = \omega_0^3 \Rightarrow l_3 = \omega_0^3 \end{cases} \quad (14)$$

因此, 可以将估计误差矩阵  $\mathbf{H}$  重新描述为:

$$\mathbf{H} = \begin{bmatrix} -l_1 & 1 & 0 \\ -l_2 - k_0 & -k_1 & 1 \\ -l_3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -3\omega_0 & 1 & 0 \\ -3\omega_0^2 & 0 & 1 \\ -\omega_0^3 & 0 & 0 \end{bmatrix} + \begin{bmatrix} k_1 & 0 & 0 \\ k_1l_1 & -k_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (15)$$

对误差方程引入坐标变换:  $\mathbf{q} = \Lambda\xi$ ,  $\xi = [\xi_1 \ \xi_2 \ \xi_3]$ ,  $\Lambda = \text{diag}(\omega_0^{-2}, \omega_0^{-1}, 1)$ , 可将其重新描述为:

$$\Lambda\dot{\xi} = \mathbf{H}\Lambda\xi + \mathbf{h}\eta \quad (16)$$

式(16)左乘  $\Lambda\xi^{-1}$  可得:

$$\frac{1}{\omega_0}\dot{\xi} = \mathbf{H}_q\xi + \frac{1}{\omega_0}\mathbf{h}\eta \Rightarrow \varepsilon\dot{\xi} = \mathbf{H}_q\xi + \varepsilon\mathbf{h}\eta \quad (17)$$

其中,  $\varepsilon = \frac{1}{\omega_0}$ ,  $\mathbf{H}_q = \mathbf{A}_q + \varepsilon\mathbf{H}_\varepsilon$ ,

$$\mathbf{A}_q = \begin{bmatrix} -3 & 1 & 0 \\ -3 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}, \quad \mathbf{H}_\varepsilon = \begin{bmatrix} k_1 & 0 & 0 \\ \varepsilon k_1 l_1 & -k_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

转换后的误差方程(17)及其初始条件为:

$$\varepsilon\dot{\xi} = \mathbf{H}_q\xi + \varepsilon\mathbf{h}\eta, \quad \xi(0) = \xi_0 \quad (18)$$

式(18)是一个标准的奇异摄动线性微分方程初值问题, 该方程中包含一个取值较小的正参数  $\varepsilon$ ,  $\varepsilon = 1/\omega_0$ . 根据文献[29], 可得到稳定性定理 1.

定理 1: 如果  $\|\xi_0\| = O(\varepsilon)$  和  $\dot{F} = \eta$  存在, 那么:

1) 系统的渐近解在所有有限时间  $L$  ( $0 \leq t \leq L < \infty$ ) 内均有效, 并且渐近解的具体形式为:

$$\xi(\varepsilon, t) = \exp(\mathbf{A}_q \frac{t}{\varepsilon})\xi_0 + O(\varepsilon^2) +$$

$$\varepsilon[\mathbf{A}_q^{-1}\mathbf{h} \cdot \eta + \exp(\mathbf{A}_q \frac{t}{\varepsilon})\{\mathbf{A}_q^{-1}\mathbf{h} \cdot \eta(0)\} - \quad (19)$$

$$\mathbf{H}_0 \int_0^t \exp[\mathbf{A}_q (\frac{t-s}{\varepsilon})]\mathbf{A}_q^{-1}\mathbf{h} \cdot \eta(s)ds]$$

2) 存在  $\varepsilon^* > 0$ , 使得对于所有  $\varepsilon \in [0, \varepsilon^*]$ , 系统(18)在  $\varepsilon \rightarrow 0$  时是指数稳定的;

3) 对于所有  $\xi \in \Omega = [-\rho_1, \rho_1]^3 \subset \mathbb{R}^3$ ,  $\gamma > 0$ ,  $\rho_1 = \gamma \max\{\omega_c, \varepsilon, \|\mathbf{H}_q\|\}$ , 则存在独立于  $\varepsilon$  的正常数  $C_2$ , 使系统(18)的解满足:

$$\begin{cases} \|\xi(\varepsilon, t)\| \leq \varepsilon C_1 \exp\left[-\left(\frac{1}{2} - \varepsilon C_1 C_2\right)\frac{t}{\varepsilon}\right] \\ \|\eta\| \cdot \|\xi(\varepsilon, t)\|^{-1} \leq C_2 \end{cases} \quad (20)$$

其中,  $C_1 = \sqrt{3} + \frac{1}{\omega_c} \sum_{j=1}^2 \frac{L^j}{j!} \left\| (\mathbf{H}_q + \mathbf{I}_3)^j \right\|$ ,  $\mathbf{I}_3$  是三阶

单位矩阵,  $\mathbf{H}_0 = \begin{bmatrix} 2\omega_c & 0 & 0 \\ 0 & -2\omega_c & 0 \\ 0 & 0 & 0 \end{bmatrix}$ .

由定理 1 可知, 当  $\|\xi_0\| = O(\varepsilon)$  和  $\dot{F} = \eta$  存在时, 系统(18)是指数稳定且是一致渐近稳定的.

## 2 基于强化学习的自抗扰控制器参数整定

### 2.1 SAC 算法

SAC 算法是一种基于最大熵强化学习的策略优化方法, 其目标是在最大化累积奖励的同时, 最大化策略的熵, 最大化熵可以鼓励策略更好地探索环境, 有助于提高算法的稳定性和收敛速度.

熵是随机变量随机程度的度量, 可描述为:

$$H(X) = \mathbb{E}_{x \sim p}[-\lg p(x)] \quad (21)$$

式中:  $x$  为随机变量;  $p$  为  $x$  的概率密度函数. 在强化学习中使用  $H(\pi(\cdot|s))$  表示策略  $\pi$  在状态  $s$  下的随机程度.

最大熵强化学习的思想是除了最大化累积奖励, 还要使得策略更加随机. 因此, 强化学习的目标中就加入了一项熵的正则项, 可以描述为:

$$\pi^* = \arg \max_x \mathbb{E}_x [\sum_t r(s_t, a_t) + \alpha H(\pi(\cdot|s_t))] \quad (22)$$

式中:  $\alpha$  为熵正则化系数,  $\alpha$  越大, 探索性越强, 越有助于加速后续的策略学习, 并减少策略陷入较差的局部最优的可能性.

在 SAC 算法中, Soft 贝尔曼方程可以表示为:

$$Q(s_t, a_t) = r(s_t, a_t) + \gamma \mathbb{E}_{s_{t+1}} [V(s_{t+1})] \quad (23)$$

其中，状态价值函数表示为：

$$V(s_t) = \mathbb{E}_{a_t \sim \pi} [Q(s_t, a_t) - \alpha \lg \pi(a_t | s_t)] = \mathbb{E}_{a_t \sim \pi} [Q(s_t, a_t)] + H(\pi(\cdot | s_t)) \quad (24)$$

可根据以下Soft策略提升公式改进策略：

$$\pi_{\text{now}} = \arg \min_{\pi} D_{KL} \left( \pi(\cdot | s), \frac{\exp\left(\frac{1}{\alpha} Q^{\pi \text{old}}(s, \cdot)\right)}{Z^{\pi \text{old}}(s, \cdot)} \right) \quad (25)$$

图3为SAC算法的结构示意。由图3可见，SAC算法主要由缓冲区、评论家网络和演员网络组成。演员网络负责将状态映射到动作，评论家网络负责估计状态和状态-动作的值，缓冲区负责存储经验。在SAC算法中有2个Q网络，每次使用Q网络时挑选一个Q值较小的网络，可缓解Q值过高估计的问题。

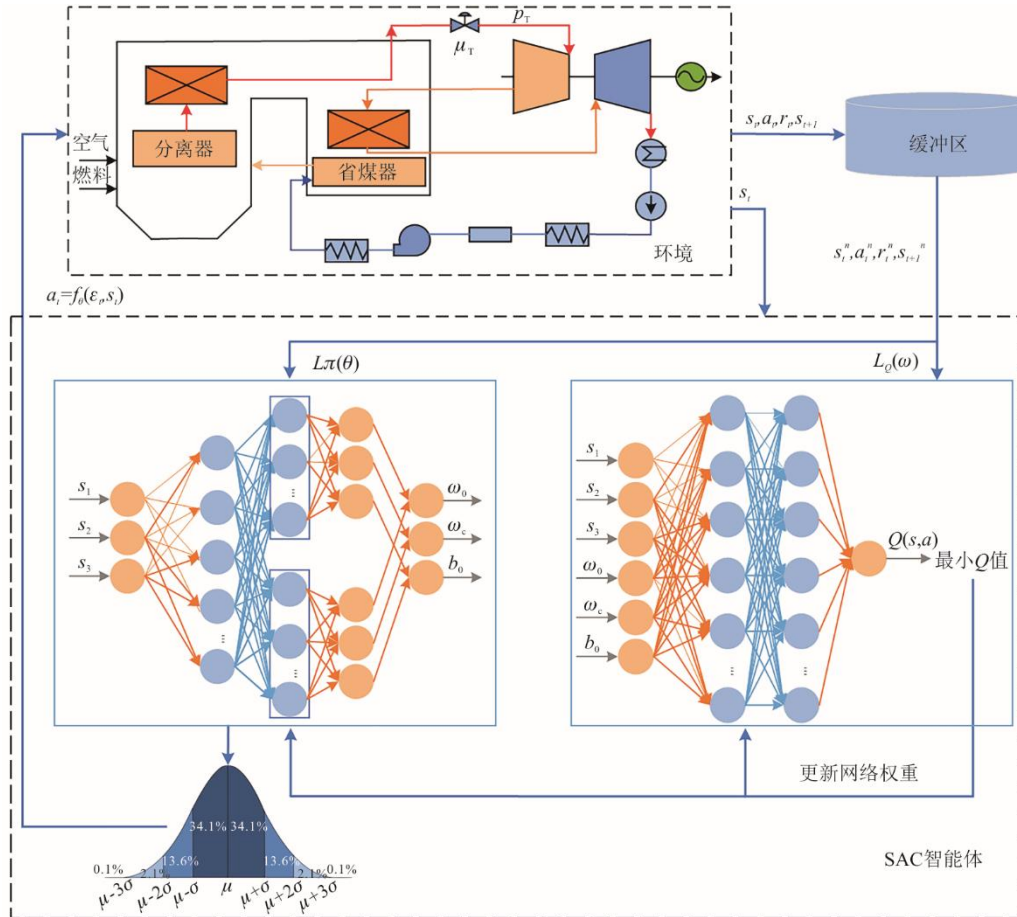


图3 SAC 算法的结构示意  
Fig.3 Schematic structure of the SAC algorithm

任意一个Q函数的损失函数为：

$$L_Q(\omega) = \mathbb{E}_{(s_t, a_t, r_t, s_{t+1}) \sim R, a_{t+1} \sim \pi_{\theta}(\cdot | s_{t+1})} \left[ \frac{1}{2} (Q_{\omega}(s_t, a_t) - (r_t + \gamma (\min_{j=1,2} Q_{\omega_j^-}(s_{t+1}, a_{t+1}) - \alpha \lg \pi(a_{t+1} | s_{t+1}))))^2 \right] \quad (26)$$

式中： $R$ 为策略先前收集的数据； $Q_{\omega^-}$ 为目标Q网络，训练中不断更新目标软Q值网络参数。

$$\omega \leftarrow \tau \omega + (1 - \tau) \omega \quad (27)$$

策略  $\pi$  的损失函数由KL散度得到：

$$L_{\pi}(\theta) = \mathbb{E}_{s_t \sim R, a_t \sim \pi_{\theta}} [\alpha \lg \pi_{\theta}(a_t | s_t) - Q_{\omega}(s_t, a_t)] \quad (28)$$

最大化状态价值函数  $V$ ：

$$V(s_t) = \mathbb{E}_{a_t \sim \pi} [Q(s_t, a_t) - \alpha \lg \pi(a_t | s_t)] \quad (29)$$

为了自动调整熵正则项  $H_0$ ，即最大化期望回报，同时约束熵的均值大于  $H_0$ ，可将强化学习的目标改写为一个带约束的优化问题：

$$\begin{aligned} \max_{\pi} & E_{\pi} \left[ \sum_t r(s_t, a_t) \right] \\ \text{s.t.} & E_{(s_t, a_t) \sim \rho_{\pi}} [-\lg \pi(a_t | s_t)] \geq H_0 \end{aligned} \quad (30)$$

化简后  $\alpha$  的损失函数为：

$$L(\alpha) = E_{s_t \sim R, a_t \sim \pi(\cdot | s_t)} [-\alpha \lg \pi(a_t | s_t) - \alpha H_0] \quad (31)$$

当策略的熵低于目标值  $H_0$  时，训练目标  $L(\alpha)$  会使  $\alpha$  增大，进而在上述最小化损失函数  $L_{\pi}(\theta)$  的过程中增加策略熵对应项重要性；反之， $\alpha$  会减小，进

而使得策略训练时更专注于价值提升。SAC算法的具体流程见表1。

表 1 SAC 算法的具体流程  
Tab.1 Specific process of the SAC algorithm

算法 1: SAC 算法伪代码	
1	初始化 critic 网络, 目标 critic 网络和目标 actor 网络权重
2	初始化经验回放池 $R \leftarrow \emptyset$
3	for $e=1 \rightarrow E$ do
4	获取环境初始状态 $s_t$
5	for $t=1 \rightarrow T$ do
6	根据当前策略选择动作 $a_t = \pi_\theta(s_t)$
7	执行当前动作 $a_t$ , 获得当前动作奖励 $r_t$
8	执行下一时刻状态 $s_{t+1}$
9	将 $(s_t, a_t, r_t, s_{t+1})$ 存入经验回放池 $R$
10	end
11	for $k=1 \rightarrow K$ do
12	从 $R$ 中随机提取数据 $\{(s_t, a_t, r_t, s_{t+1})\}$
13	将参数输入 critic 网络生成 $Q_1, Q_2$
14	采用随机梯度下降法更新 $Q$ 网络
15	将参数输入 critic 网络生成目标 $Q_1$ , 目标 $Q_2$
16	更新目标 $Q$ 网络
17	更新策略网络
18	更新熵正则项系数 $\alpha$
19	end
20	end

## 2.2 自抗扰控制器参数整定

为实现机组控制器参数自适应调整, 在保持机组快速跟踪能力下, 同时考虑机组运行安全性和经济性。对于整个控制系统, 根据已知描述, SAC算法需调整的参数是EADRC控制器参数。整个控制系统可以看作是强化学习的环境, 通过环境和智能体之间不断交互学习, 可以找到最优策略, 即最优控制器参数。而动作空间、状态空间和奖励函数的选取作为任务重要组成部分, 会直接影响SAC算法的学习和决策过程, 需精心设计。

动作空间:

$$\mathbf{a} = [\omega_{01} \ \omega_{c1} \ b_{01} \ \omega_{02} \ \omega_{c2} \ b_{02} \ \omega_{03} \ \omega_{c3} \ b_{03}] \quad (32)$$

状态空间:

$$\mathbf{s} = [e(t) \int_0^t e(t) dt \ y_1(t) \ y_2(t) \ y_3(t) \ u_1(t) \ u_2(t) \ u_3(t) \ \dot{u}_1(t) \ \dot{u}_2(t) \ \dot{u}_3(t)] \quad (33)$$

式中:  $e(t) = \alpha_1 e_1(t) + \alpha_2 e_2(t) + \alpha_3 e_3(t)$ ,  $\alpha_1$ 、 $\alpha_2$ 、 $\alpha_3$  分别为3个被控变量有功功率、主蒸汽压力和分离器出口温度的跟踪误差权重;  $y_1(t)$ 、 $y_2(t)$ 、 $y_3(t)$  分别为3个被控变量有功功率、主蒸汽压力和分离器出

口温度的实际输出值;  $u_1(t)$ 、 $u_2(t)$ 、 $u_3(t)$  分别为3个控制变量阀门开度、给煤量和给水量,  $\dot{u}_1(t)$ 、 $\dot{u}_2(t)$ 、 $\dot{u}_3(t)$  分别为3个控制变量阀门开度、给煤量和给水量的变化速率。

奖励函数设计为:

$$r_1 = \begin{cases} -|e(t)| + 1, & |e(t)| \leq 0.1 \\ -\tanh(|e(t)| - 0.1), & |e(t)| > 0.1 \end{cases} \quad (34)$$

$$r_2 = \begin{cases} 1, & 1068 \leq \sum_{i=1}^3 |\Delta u_i| \leq 2205 \\ -10, & \sum_{i=1}^3 |\Delta u_i| < 1068 \text{ 或 } \sum_{i=1}^3 |\Delta u_i| > 2205 \end{cases} \quad (35)$$

$$r_3 = \begin{cases} 0.1, & 0.02 \leq \sum_{i=1}^3 |\Delta \dot{u}_i| \leq 4 \\ -5, & \sum_{i=1}^3 |\Delta \dot{u}_i| < 0.02 \text{ 或 } \sum_{i=1}^3 |\Delta \dot{u}_i| > 4 \end{cases} \quad (36)$$

$$r_4 = \frac{1}{T_{\max}} \int_0^{T_{\max}} t |\Delta M| dt \quad (37)$$

$$r = r_1 + r_2 + r_3 + r_4 \quad (38)$$

式中: 控制变量增量和控制变量增量变化速率的范围约束条件是通过观察机组历史运行数据以及结合现场运行专家经验得到的,  $r_1$ 为跟踪误差惩罚项,  $r_2$ 为控制量超出惩罚项,  $r_3$ 为控制量变化速率超出惩罚项,  $r_4$ 为燃料量利用率超出惩罚项,  $r$ 为最终的奖励函数。

## 3 仿真实验

结合某超超临界 1 000 MW 二次再热机组的真实历史运行数据, 基于文献[32-35]所提出的子空间辨识建模方法, 采样时间为 1 s, 选择大范围变负荷下的 2 000 组数据作为训练集, 不同额定负荷下 600 组数据作为验证集, 建立 90% 额定负荷和 40% 额定负荷下的超超临界机组协调控制系统状态空间模型。2 种额定负荷下的数据辨识结果验证分别如图 4 和图 5 所示, 从图 4 和图 5 可以看出, 该模型能够较好地拟合机炉耦合过程动态特性。

由于 ADRC 并未对被控对象的形式作出具体要求, 在设计使用时可以忽略系统的线性或非线性、时变或时不变等因素, 具有很强的适用性。因此, 本文将非线性系统在某个额定工况下进行线性化处理, 将系统内部非线性、未建模动态特性和外

部干扰等效为一种总扰动,通过 ESO 来实时在线观测并采用 PD 控制部分将其补偿掉。

因此,本文选取 90%额定负荷和 40%额定负荷下超超临界机组机炉耦合过程作为被控对象,其状态空间模型及运行工况如下。

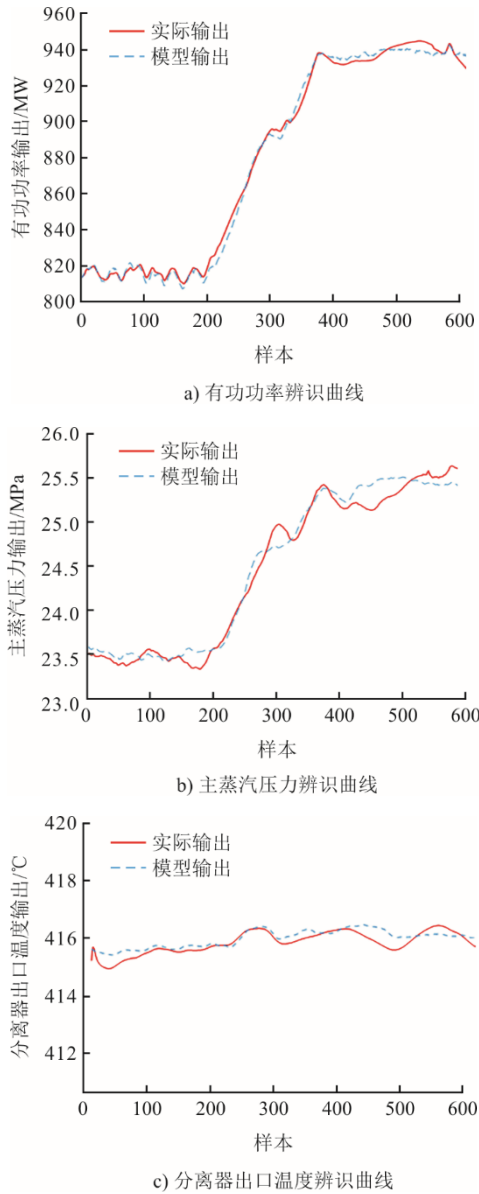


图 4 90%额定负荷下数据辨识结果验证  
Fig.4 Verification of data identification result under 90% rated load

$$A = \begin{bmatrix} -0.044 & 92 & 0.015 & 62 & -0.023 & 8 & -0.091 & 42 & -0.037 & 87 & 0.018 & 97 & 0.201 & 5 \\ 0.055 & 22 & -0.029 & 18 & -0.006 & 576 & 0.021 & 24 & 0.135 & 9 & -0.039 & 27 & -0.278 & 4 \\ 0.102 & 6 & -0.053 & 91 & 0.011 & 63 & 0.142 & 2 & 0.190 & 2 & -0.164 & 8 & -1.044 & \\ 0.045 & 79 & 0.004 & 693 & -0.022 & 59 & -0.127 & 1 & 0.058 & 25 & -0.192 & 7 & -0.185 & \\ -0.005 & 938 & -0.044 & 75 & -0.014 & 82 & 0.169 & 4 & -0.011 & 45 & 0.088 & 18 & -0.457 & 4 \\ 0.023 & 67 & -0.005 & 278 & 0.034 & 08 & 0.015 & 58 & 0.080 & 3 & -0.193 & 7 & -0.462 & 9 \\ 0.085 & 73 & 0.038 & 78 & 0.008 & 42 & 0.263 & 4 & 0.428 & & -0.858 & 4 & -3.222 & \end{bmatrix},$$

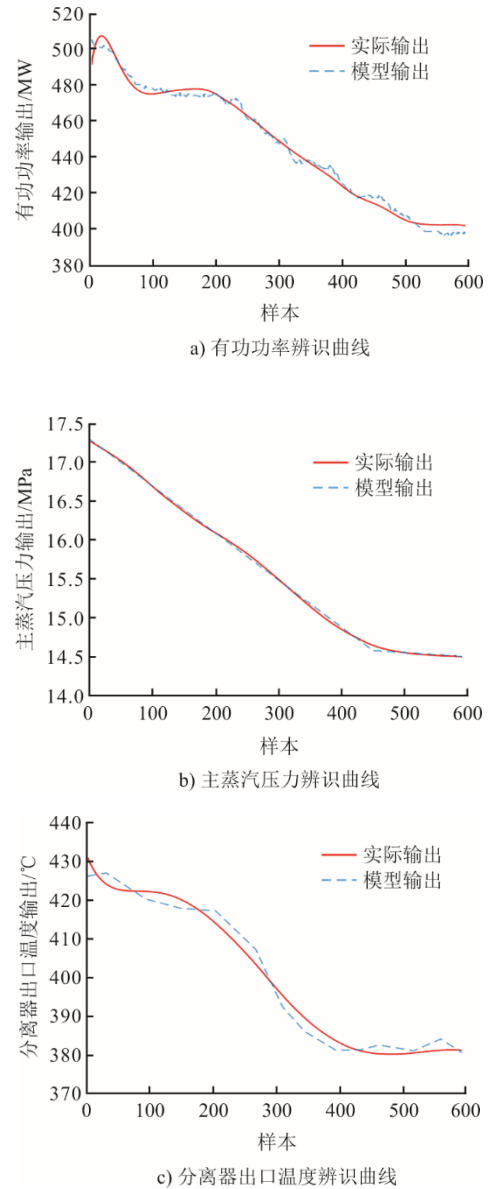


图 5 40%额定负荷下数据辨识结果验证  
Fig.5 Verification of data identification result under 40% rated load

超超临界机组机炉耦合过程状态空间表达式为:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot \mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C} \cdot \mathbf{x}(k) + \mathbf{D} \cdot \mathbf{u}(k) \end{aligned} \quad (39)$$

$$\mathbf{D} = 0$$

90%额定负荷时:

$$B = \begin{bmatrix} 0.047\ 99 & 0.000\ 115\ 6 & 0.000\ 385\ 7 \\ -0.059\ 91 & -0.000\ 337\ 3 & -0.000\ 416\ 9 \\ -0.257\ 4 & -0.000\ 343\ 2 & -0.001\ 634 \\ -0.048\ 98 & -0.000\ 841\ 6 & 0.000\ 028\ 76 \\ -0.119\ 3 & -0.001\ 739 & -0.000\ 998 \\ -0.113\ 8 & -0.001\ 14 & -0.000\ 738\ 9 \\ -0.816\ 6 & -0.010\ 35 & -0.005\ 687 \end{bmatrix}, C = \begin{bmatrix} -85.67 & -1.422 & 24.5 \\ 42.43 & 1.298 & 19.89 \\ -37.48 & 0.153\ 3 & -1.365 \\ 2.737 & 0.023\ 48 & -0.457\ 4 \\ -1.159 & 0.060\ 05 & 0.528\ 8 \\ 0.337 & 0.004\ 997 & 0.284\ 6 \\ -0.048\ 86 & -0.008\ 625 & 0.206\ 1 \end{bmatrix}^T。$$

40%额定负荷时:

$$A = \begin{bmatrix} -0.020 & 0.012 & 0.019 & -0.085 & 0.122 & 0.141 & 0.048 \\ -0.002 & -0.034 & -0.040 & 0.158 & 0.355 & -0.019 & 0.015 \\ 0.000\ 3 & 0.009 & -0.023 & -0.014 & 0.007 & -0.058 & -0.011 \\ 0.021 & 0.033 & 0.067 & -0.319 & -0.072 & 0.661 & -0.053 \\ -0.060 & -0.115 & 0.106 & 0.663 & -1.489 & -0.759 & -0.605 \\ 0.003 & -0.010 & 0.015 & -0.394 & 0.204 & -0.221 & -0.228 \\ 0.025 & -0.004 & -0.067 & -0.039 & 1.322 & 0.657 & 0.162 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.017 & 0.005 & 0 \\ -0.158 & -0.000\ 8 & -0.000\ 3 \\ -0.052 & -0.001 & 0 \\ 0.664 & 0.016 & 0.000\ 1 \\ 0.175 & -0.035 & 0.000\ 6 \\ -0.618 & 0.002 & 0.000\ 2 \\ -0.297 & 0.023 & -0.001 \end{bmatrix}, C = \begin{bmatrix} -47.92 & 133.2 & 46.24 \\ 9.79 & 26.19 & -150.1 \\ -175.6 & -5.91 & -3.45 \\ -12.06 & -2.31 & -10.17 \\ 14.06 & 3.65 & -6.76 \\ -6.27 & -0.20 & 1.73 \\ -6.33 & -4.40 & 2.72 \end{bmatrix}^T。$$

表 2 为机组协调控制系统运行工况。

表 2 机组协调控制系统运行工况  
Tab.2 Operation condition of the coordinated control system

项目	90%额定负荷	40%额定负荷
有功功率被控变量/MW	857.872	450.310
主蒸汽压力被控变量/MPa	29.086	15.574
分离器出口温度被控变量/℃	448.735	399.538
阀门开度控制变量/%	44.749	27.505
给煤量控制变量/(t·h <sup>-1</sup> )	396.946	197.166
给水量控制变量/(t·h <sup>-1</sup> )	2 374.120	1 169.924

将上述模型作为仿真系统中的被控对象，应用所提出的基于 RL+EADRC 的控制策略，基于 MATLAB/Simulink 仿真软件开展设定值跟踪测试和抗扰能力测试。并将经过粒子群算法优化的 EADRC 控制策略和 PID 控制策略作为基准控制策略。

### 3.1 预训练

SAC 算法主要参数见表 3。

表 3 SAC 算法主要参数  
Tab.3 Main parameters of the SAC algorithm

项目	数值
Critic网络学习率	0.000 1
Actor网络学习率	0.001
折扣因子	0.99
经验池大小	10 <sup>6</sup>
目标平滑因子	0.001
最大迭代次数	1 000
平均窗口长度	30
样本学习个数	128

基于 SAC 算法的超超临界机组协调控制系统控制策略的训练过程如图 6 所示。

根据图 6 训练结果可以看出，在智能体训练过程中，智能体通过不断地与环境交互学习使得平均奖励不断增加，当训练轮数达到 200 时，平均奖励开始收敛，达到停止条件。经过训练后的智能体能够根据环境的改变自适应地调整自抗扰控制器的参数。

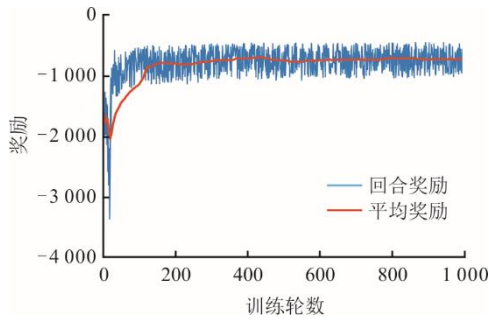


图 6 超超临界机组协调控制系统控制策略训练过程  
Fig.6 Training process of the control strategy of the coordinated control system of ultra-supercritical unit

### 3.2 设定值跟踪测试

为验证所提出的基于 RL+EADRC 的超超临界机组协调控制系统控制策略的性能，令有功功率被控变量在其设定值的基础上以斜坡形式递增 20 MW，仿真时间设定为 5 000 s。图 7—图 12 为 90% 额定负荷和 40% 额定负荷下的超超临界机组协调控制系统的被控变量响应情况。

根据图 7—图 12 结果可知，所提控制策略的响应速度优于 EADRC 控制策略和 PID 控制策略，具有更好的跟踪性能和动态性能，证明了所采用的基于 SAC 算法的参数优化策略比基于粒子群算法的参数优化策略更具优越性。

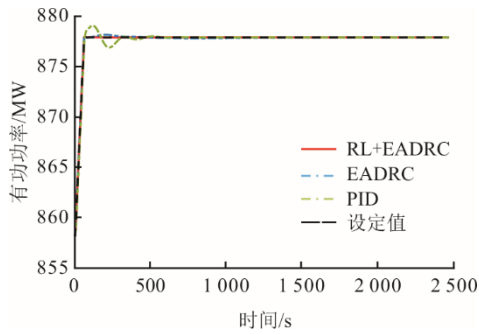


图 7 90% 额定负荷下有功功率被控变量响应  
Fig.7 Response of active power controlled variable under 90% rated load

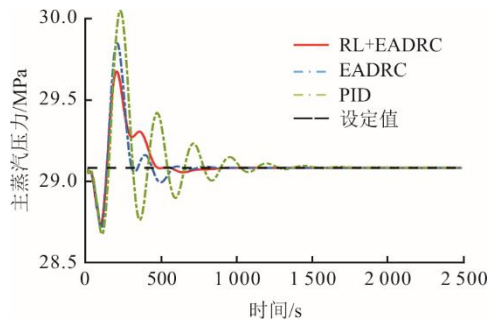


图 8 90% 额定负荷下主蒸汽压力被控变量响应  
Fig.8 Response of main steam pressure controlled variable under 90% rated load

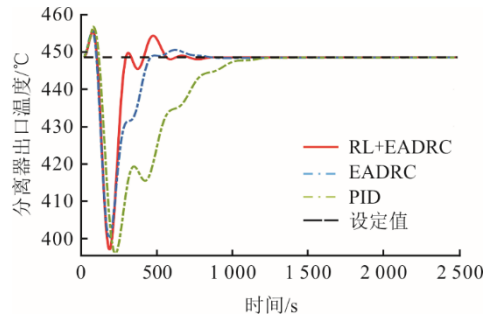


图 9 90% 额定负荷下分离器出口温度被控变量响应  
Fig.9 Response of separator outlet temperature controlled variable under 90% rated load

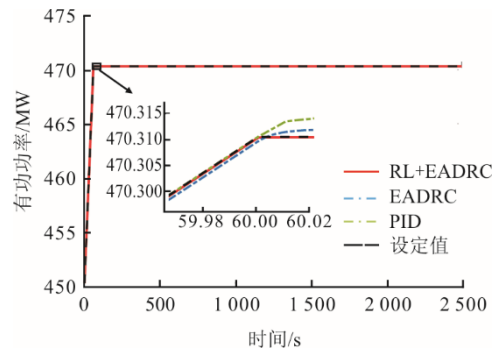


图 10 40% 额定负荷下有功功率被控变量响应  
Fig.10 Response of active power controlled variable under 40% rated load

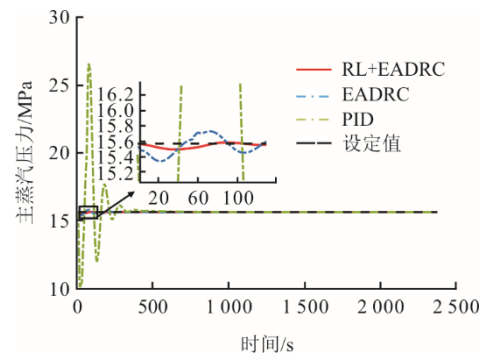


图 11 40% 额定负荷下主蒸汽压力被控变量响应  
Fig.11 Response of main steam pressure controlled variable under 40% rated load

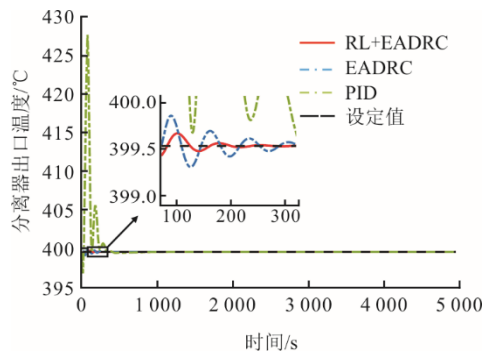


图 12 40% 额定负荷下分离器出口温度被控变量响应  
Fig.12 Response of separator outlet temperature controlled variable under 40% rated load

### 3.3 抗扰能力测试

机组在实际运行过程中会受到各种扰动, 为了测试所提控制策略的抗扰能力, 在  $t=1\ 000\ \text{s}$  时令给煤量阶跃上升 5%。通过仿真, 90% 额定负荷和 40% 额定负荷下的超超临界机组协调控制系统被控变量响应分别如图 13—图 18 所示。

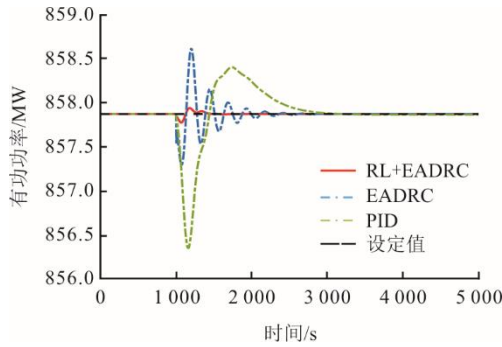


图 13 90% 额定负荷下给煤量控制指令在  $t=1\ 000\ \text{s}$  时阶跃上升 5% 有功功率被控变量响应

**Fig.13 Response of active power controlled variable with 5% step increase of coal feeding rate control law at  $t=1\ 000\ \text{s}$  under 90% rated load**

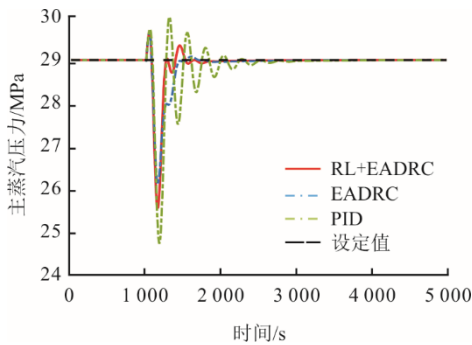


图 14 90% 额定负荷下给煤量控制指令在  $t=1\ 000\ \text{s}$  时阶跃上升 5% 主蒸汽压力被控变量响应

**Fig.14 Response of main steam pressure controlled variable with 5% step increase of coal feeding rate control law at  $t=1\ 000\ \text{s}$  under 90% rated load**

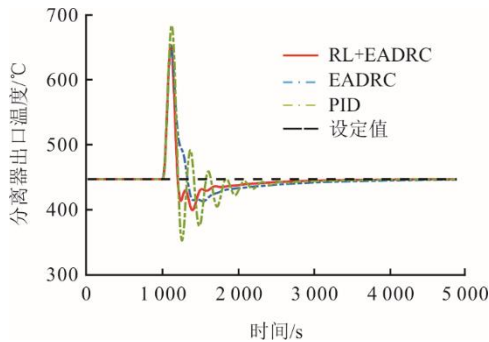


图 15 90% 额定负荷下给煤量控制指令在  $t=1\ 000\ \text{s}$  时阶跃上升 5% 分离器出口温度被控变量响应

**Fig.15 Response of separator outlet temperature controlled variable with 5% step increase of coal feeding rate control law at  $t=1\ 000\ \text{s}$  under 90% rated load**

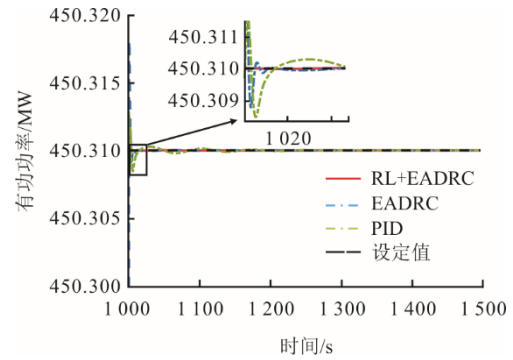


图 16 40% 额定负荷下给煤量控制指令在  $t=1\ 000\ \text{s}$  时阶跃上升 5% 有功功率被控变量响应

**Fig.16 Response of active power controlled variable with 5% step increase of coal feeding rate control law at  $t=1\ 000\ \text{s}$  under 40% rated load**

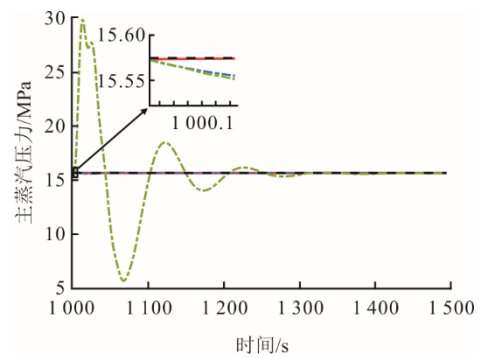


图 17 40% 额定负荷下给煤量控制指令在  $t=1\ 000\ \text{s}$  时阶跃上升 5% 主蒸汽压力被控变量响应

**Fig.17 Response of main steam pressure controlled variable with 5% step increase of coal feeding rate control law at  $t=1\ 000\ \text{s}$  under 40% rated load**

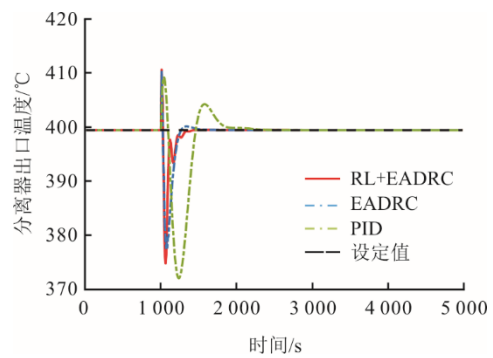


图 18 40% 额定负荷下给煤量控制指令在  $t=1\ 000\ \text{s}$  时阶跃上升 5% 分离器出口温度被控变量响应

**Fig.18 Response of separator outlet temperature controlled variable with 5% step increase of coal feeding rate control law at  $t=1\ 000\ \text{s}$  under 40% rated load**

根据图 13—图 18 结果可知, 基于 RL+EADRC 的超超临界机组协调控制系统控制策略的实际被控变量能在仿真时间 3 500 s 内达到设定值, 且在受到外界扰动时, 所提控制策略在恢复稳态值时间方面优于 EADRC 控制策略和 PID 控制策略。因此,

证明了所提出的 RL+ EADRC 控制策略相比传统的 PID 控制策略具有较好的抗扰能力。

## 4 结 论

针对超超临界机组协调控制系统性能优化问题,提出了一种基于 EADRC 和 SAC 算法的超超临界机组协调控制系统控制策略。根据超超临界机组机炉耦合过程运行特性,确定了系统的输入输出变量,给出了机炉耦合过程被控对象简化模型,并根据此模型设计了 EADRC 控制策略。将 SAC 算法引入 EADRC 参数优化中,通过最大化以系统的跟踪性能、燃料量等为目的的奖励函数,获得了 EADRC 的最优参数。设定值跟踪测试和抗扰能力测试的实验结果表明了所提控制策略具有较好的跟踪性能和抗扰性能。

### [参 考 文 献]

- [1] 舒印彪, 赵勇, 赵良, 等. “双碳”目标下我国能源电力低碳转型路径[J]. 中国电机工程学报, 2023, 43(5): 1663-1672.  
SHU Yinbiao, ZHAO Yong, ZHAO Liang, et al. Study on low carbon energy transition path toward carbon peak and carbon neutrality[J]. Proceedings of the CSEE, 2023, 43(5): 1663-1672.
- [2] 任大伟, 肖晋宇, 侯金鸣, 等. 双碳目标下我国新型电力系统的构建与演变研究[J]. 电网技术, 2022, 46(10): 3831-3839.  
REN Dawei, XIAO Jinyu, HOU Jinming, et al. Construction and evolution of China's new power system under dual carbon goal[J]. Power System Technology, 2022, 46(10): 3831-3839.
- [3] 王倩, 王卫良, 刘敏, 等. 超(超)临界燃煤发电技术发展及展望[J]. 热力发电, 2021, 50(2): 1-9.  
WANG Qian, WANG Weiliang, LIU Min, et al. Development and prospect of(ultra)supercritical coal-fired power generation technology[J]. Thermal Power Generation, 2021, 50(2): 1-9.
- [4] 王玮, 申朋玉, 李雪寒, 等. 提升直流炉供热机组灵活性的电热协调控制策略[J]. 中国电机工程学报, 2023, 43(6): 2100-2109.  
WANG Wei, SHEN Pengyu, LI Xuehan, et al. Electricity-heat coordinated control strategy for improving the flexibility of a once-through CHP unit[J]. Proceedings of the CSEE, 2023, 43(6): 2100-2109.
- [5] 崔晓宁, 赵志刚, 姜宁, 等. 1 000 MW 超超临界二次再热锅炉系统焓分析[J]. 动力工程学报, 2019, 39(8): 605-610.  
CUI Xiaoning, ZHAO Zhigang, JIANG Ning, et al. Exergy analysis of a 1 000 MW ultra-supercritical double reheat boiler unit[J]. Journal of Chinese Society of Power Engineering, 2019, 39(8): 605-610.
- [6] 赵征, 于悦波, 孙昊天. 基于凝结水节流的新型协调优化控制策略[J]. 动力工程学报, 2021, 41(2): 107-112.  
ZHAO Zheng, YU Yuebo, SUN Haotian. Optimization of a new coordinated control strategy based on condensate throttling[J]. Journal of Chinese Society of Power Engineering, 2021, 41(2): 107-112.
- [7] MOHAMMAD E, HAMED M. Robust & nonlinear control of an ultra-supercritical coal fired once-through boiler-turbine unit in order to optimize the uncertain problem[J]. Energy, 2023, 282: 128312.
- [8] 张志豪, 李军, 王林, 等. 基于内模控制和改进自抗扰控制的主汽温串级控制系统[J]. 中国电机工程学报, 2024, 44(23): 9331-9340.  
ZHANG Zhihao, LI Jun, WANG Lin, et al. Cascade control system of main steam temperature based on internal model control and improved active disturbance rejection control[J]. Proceedings of the CSEE, 2024, 44(23): 9331-9340.
- [9] WANG Z X, SONG C, YAN A, et al. Complete synchronization and partial anti-synchronization of complex Lü chaotic systems by the UDE-based control method[J]. Symmetry, 2022, 14(3): 517.
- [10] HAN J Q. From PID to active disturbance rejection control[J]. IEEE Transactions on Industrial Electronics, 2009, 56(3): 900-906.
- [11] GAO Z Q. On the foundation of active disturbance rejection control[J]. Control Theory & Applications, 2013, 12: 1498.
- [12] GAO Z Q. On the centrality of disturbance rejection in automatic control[J]. ISA Transactions, 2014, 53(4): 850-857.
- [13] MADONSKI R, SHAO S, ZHANG H. General error-based active disturbance rejection control for swift industrial implementations[J]. Control Engineering Practice, 2019, 84: 218-229.
- [14] HUANG C Z, ZHUANG J N. Error-based active disturbance rejection control for pitch control of wind turbine by improved coyote optimization algorithm[J]. IEEE Transactions on Energy Conversion, 2022, 37(2): 1394-1405.
- [15] HUANG C Z, ZHAO H T. Error-based active disturbance rejection control for wind turbine output power regulation[J]. IEEE Transactions on Sustainable Energy, 2023, 14(3): 1692-1701.
- [16] 杨延西, 谢宙桦, 邓毅. 过热蒸汽温度量子粒子群优化无模型自适应控制[J]. 热力发电, 2016, 45(4): 84-88.  
YANG Yanxi, XIE Zhouhua, DENG Yi. Model-free adaptive control of superheated steam temperature based on quantum-behaved particle swarm optimization[J]. Thermal Power Generation, 2016, 45(4): 84-88.
- [17] 岑俊, 黄云峰, 方世清, 等. 基于改进 GA-CPSO 融合算法的 1 000 MW 机组给水系统建模[J]. 热能动力工程, 2023, 38(7): 117-127.  
CEN Jun, HUANG Yunfeng, FANG Shiqing, et al. Modeling of feed water system for 1 000 MW unit based on improved GA-CPSO fusion algorithm[J]. Journal of Engineering for Thermal Energy and Power, 2023, 38(7): 117-127.
- [18] SUN C H, LIU M Q, LIU C A, et al. An industrial quadrotor UAV control method based on fuzzy adaptive linear active disturbance rejection control[J]. Electronics, 2021, 10(4): 376.
- [19] LIU W T, ZHAO T, WU Z W, et al. Linear active disturbance rejection control for hysteresis compensation based on backpropagation neural networks adaptive control[J]. Transactions of the Institute of Measurement and Control, 2021, 43(4): 915-924.
- [20] CHEN G, LU Y W, YANG X, et al. Reinforcement learning control for the swimming motions of a

- beaver-like, single-legged robot based on biological inspiration[J]. *Robotics and Autonomous Systems*, 2022, 154: 104116.
- [21] ZHENG Y M, TAO J, SUN Q L, et al. Soft actor-critic based active disturbance rejection path following control for unmanned surface vessel under wind and wave disturbances[J]. *Ocean Engineering*, 2022, 247: 110631.
- [22] HOU G L, HUANG T, ZHENG F M, et al. Application of multi-agent EADRC in flexible operation of combined heat and power plant considering carbon emission and economy[J]. *Energy*, 2023, 263: 125711.
- [23] MENG X Q, JIA Y B, REN C G, et al. DDPG-based back stepping controller for DC solid state transformer in DC microgrid with constant power loads[J]. *IEEE Transactions on Smart Grid*, 2022, 13(6): 4269-4283.
- [24] DUAN J J, SHI D, DIAO R S, et al. Deep-reinforcement-learning-based autonomous voltage control for power grid operations[J]. *IEEE Transactions on Power Systems*, 2020, 35(1): 814-817.
- [25] DONG H Y, ZHAO X W. Composite experience replay-based deep reinforcement learning with application in wind farm control[J]. *IEEE Transactions on Control Systems Technology*, 2021, 30(3): 1281-1295.
- [26] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[J]. *International Conference on Machine Learning*, 2018, 80: 1861-1870.
- [27] HAARNOJA T, ZHOU A, HARTIKAINEN K, et al. Soft actor-critic algorithms and applications[Z/OL]. <https://arxiv.org/pdf/1812.05905>.
- [28] GAO Z Q. Scaling and bandwidth-parameterization based controller tuning[C]//*Proceedings of the 2003 American Control Conference*. Denver: IEEE, 2003, 6: 4989-4996.
- [29] SHAO S S, GAO Z Q. On the conditions of exponential stability in active disturbance rejection control based on singular perturbation analysis[J]. *International Journal of Control*, 2017, 90(10): 2085-2097.
- [30] DAI C B, XUE Y L, WU Z L, et al. Active disturbance rejection control for the reheated steam temperature of a double-reheat boiler[J]. *Journal of Process Control*, 2023, 130: 103078.
- [31] WU Z L, HE T, LIU Y H, et al. Physics-informed energy-balanced modeling and active disturbance rejection control for circulating fluidized bed units[J]. *Control Engineering Practice*, 2021, 116: 104934.
- [32] VAN OVERSCHEE P, DE MOOR B. A unifying theorem for three subspace system identification algorithms[J]. *Automatica*, 1995, 31(12): 1853-1864.
- [33] QIN S J. An overview of subspace identification[J]. *Computers & chemical engineering*, 2006, 30(10/11/12): 1502-1513.
- [34] FAVOREEL W, MOOR B D, OVERSCHEE P V. Subspace state space system identification for industrial processes[J]. *Journal of Process Control*, 2000, 10(2/3): 149-155.
- [35] KADALI R, HUANG B, ROSSITER A. A data driven subspace approach to predictive controller design[J]. *Control Engineering Practice*, 2003, 11(3): 261-278.

(责任编辑 杜亚勤)