Adaptive Fuzzy Optimal Backstepping Control of Twin-Roll Inclined Casting System Based on Actor-Critic Architecture

Dongxiang Gao, Wenhua Cui*, Yujun Zhang

Abstract—**This paper investigates the fuzzy optimal control problem of the twin-roll inclined casting system (TRICS). Firstly, the nonaffine function is decoupled based on the mean value theorem, thereby simplifying the design difficulty of the controller. Secondly, the nonlinear function is approximated by the fuzzy logic method, and combined with the backstepping design framework and reinforcement learning technology, a fuzzy optimal controller based on the actor-critic architecture is proposed. The proposed control strategy not only ensures the boundedness of all signals in the closed-loop system, but also achieves the optimization goal of the control signal. Finally, the effectiveness and feasibility of the method are verified through a simulation example.**

Index Terms—**Twin-roll inclined casting, Optimal backstepping control, Actor-critic architecture, Reinforcement learning**

I. INTRODUCTION

THE two-roll inclined continuous casting and rolling process holds significant application value in modern industrial production [1, 2]. Compared with traditional rolling processes, this process eliminates intermediate steps such as initial rolling and heating, significantly simplifying the production process, enhancing production efficiency, and effectively reducing equipment investment, energy consumption, and environmental pollution [3, 4]. It also notably improves the mechanical properties and surface quality of the plates. Therefore, this process has been widely applied in the steel production field where higher requirements are placed on strength, toughness, and precision [5]. However, due to the complex influence of multiple coupled parameters on the rolling effect, precise control of the two-roll inclined continuous casting and rolling process faces considerable challenges. To address these issues, it is necessary to conduct in-depth research and optimization on the control methods of this process and take effective measures to ensure the controllability and stability of the rolling process.

In recent years, with the rapid development of adaptive control theory, new design ideas have been provided for the optimization of many practical engineering systems. To address the impact of parameter uncertainties and unmeasurable states on the control performance of certain nonlinear systems, some control strategies based on neural networks or fuzzy output feedback have been proposed in [6–10], which have significantly advanced the progress of adaptive control theory. However, it is worth noting that in the current control field, how to effectively save control resources and further improve system performance has become one of the key issues to be solved urgently. For this purpose, optimal control theory has been introduced to provide solutions. Its core objective is to design an optimized control strategy for dynamic systems or motion processes, which not only meets the requirements of control tasks but also minimizes predefined performance indicators and reduces the consumption of control resources.

Optimal control for nonlinear systems is one of the core aspects of modern control theory, aiming to optimize the performance indicators of control systems [11]. It integrates the fundamental conditions and methods distilled from practical problems, with the research object being controlled dynamic systems or motion processes. It seeks the best control scheme among the allowable ones to ensure the system achieves the optimal performance when transitioning from the initial state to the target state [12]. With the rapid development of digital technology and electronic computers, optimal control has been widely applied in production, military, and economic activities, playing a significant role in the national economy and national defense. The optimal problem is theoretically equivalent to solving the Hamilton-Jacobi-Bellman (HJB) equation [13], but due to its strong nonlinearity and dynamic uncertainty, it is difficult to solve directly through analysis. To overcome this challenge, reinforcement learning (RL) and adaptive dynamic programming (ADP) have been proven to be effective solutions. RL and ADP were initially proposed by Werbos for discrete systems [14] and later extended to continuous systems [15, 16], but they are only applicable to affine nonlinear systems. For the control problem of nonlinear mismatched systems, in [17] proposed an optimal control method based on the backstepping framework, ensuring the optimization of each subsystem. To simplify complexity and alleviate the continuous excitation condition, in [18–21] further simplified the optimal backstepping control strategy.

Based on the above research, this paper takes the two-roll inclined casting system as the control object, and combines the reinforcement learning algorithm under the actor-critic architecture with the backstepping control theory to design a fuzzy optimal controller. This controller aims to optimize the

Manuscript received June 20, 2025; revised August 15, 2025.

This work was supported by the Key Laboratory of Internet of Things Application Technology on Intelligent Construction, Liaoning Province (2021JH13/10200051).

Dongxiang Gao is a doctoral student at the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, China (e-mail: lkdgao1997@163.com).

Wenhua Cui is a Professor of School of Computer and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China (Corresponding author, e-mail: lkdcwh@126.com).

Yujun Zhang is a Professor of School of Computer and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China (e-mail: 1997zyj@163.com).

melt pool liquid level height, roll gap and multiple coupled control parameters of the system, thereby enhancing the stability and robustness of the overall system.

II. MODEL DESCRIPTION AND PRELIMINARIES

In this section, the schematic diagram of the entire twin-roll inclined casting process is shown in Figure 1. Subsequently, based on the mathematical model of the molten steel liquid level during the twin-roll inclined continuous casting and rolling process, it can be concluded that:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{Q_{in} - Q_{out} - L\Phi(y, x_g, \beta)}{LF(y, x_g, \beta)} \tag{1}$$

where y is the height of molten metal, L is the surface width of the casting mill roll, x_g is the roll gap, and β is the inclination angle. The metal inflow quantity Q_{in} is a nonlinear function of the ladle stopper rod height h_s, that is, $Q_{in} = ah_s(b - ch_s)$, where $a = 0.2466\pi$, $b = 0.01585$, and $c = 0.2165$. The metal outflow quantity $Q_{out} = Lx_g\omega R$, where ω represents the angular velocity of the roll's rotation and R is the radius of roll.

Meanwhile, $F(y, x_g, \beta)$ and $\Phi(y, x_g, \beta)$ are defined as the following two nonlinear functions:

$$\begin{aligned} F(y, x_g, \beta) =& (2R + x_g)\cos\beta - \sqrt{R^2 - y^2} \\ & - \sqrt{R^2 - \big(y - (2R + x_g)\sin\beta\big)^2} \end{aligned} \tag{2}$$

$$\begin{aligned} \Phi(y, x_g, \beta) =& \Bigg(y\cos\beta + \sin\beta\sqrt{R^2 - (R + \frac{x_g}{2})^2 \sin\beta^2} \\ & - (2R + x_g)\cos\beta\sin\beta\Bigg)\frac{\mathrm{d}x_g}{\mathrm{d}t} \end{aligned} \tag{3}$$

Overall, dynamic equation (1) can be rewritten as

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{ah_s(b - ch_s) - Lx_g\omega R - L\Phi(y, x_g, \beta)}{LF(y, x_g, \beta)} \tag{4}$$

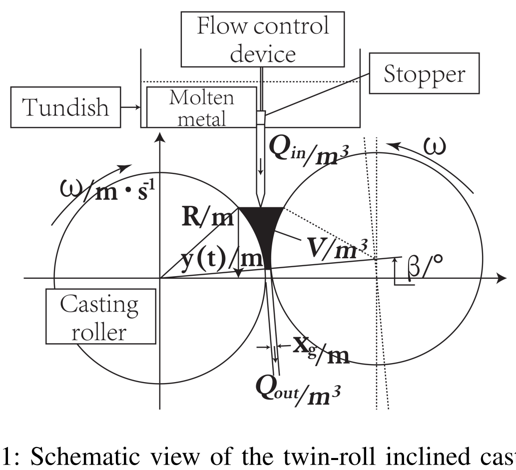

Fig. 1: Schematic view of the twin-roll inclined casting.

To develop the control scheme for the subsequent phase, we need to incorporate the coordinate transformations $\xi_1 = y$, $\xi_2 = \frac{\mathrm{d}y}{\mathrm{d}t}$, and $u = h_s$, which allow us to convert the dynamic equation (4) of the molten steel level into a nonaffine nonlinear system in the following format:

$$\begin{cases} \dot{\xi_1} = \xi_2 \\ \dot{\xi_2} = f(\xi, u) \\ y = \xi_1 \end{cases} \tag{5}$$

where $\xi = [\xi_1, \xi_2]^{\mathrm{T}}$ is the state vector, symbols u and y represent the control input and output of the system, respectively. $f(\xi, u)$ is an nonaffine nonlinear function.

Taking the derivative of the molten pool liquid level equation (4) with respect to time t yields a strongly coupled smooth non-affine nonlinear function $f(\xi, u)$. This function $f(\xi, u)$ includes several important time-varying parameters such as the molten pool liquid level height y and the gap opening of the rolls x_g, as well as several adjustable parameters such as the height of the cast-rolling bar h_s, the angular velocity of the rolls ω, and the tilt angle β. Subsequently, in order to simplify the design difficulty of the subsequent controller, the non-affine nonlinear function $f(\xi, u)$ in (5) can be decoupled as follows by means of the mean value theorem:

$$\begin{aligned} f(\xi, u) &= f(\xi, 0) + u\frac{\partial f(\xi, v)}{\partial v}|_{v=u\varrho} \\ &= u + H(\xi, u) \end{aligned} \tag{6}$$

where $0 < \varrho < 1$, and $H(\xi, u) = u\frac{\partial f(\xi,v)}{\partial v}|_{v=u\varrho} + f(\xi, 0)$. By using (6), (5) can be rewritten in the following form:

$$\begin{cases} \dot{\xi_1} = \xi_2 \\ \dot{\xi_2} = u + H(\xi, u) \\ y = \xi_1 \end{cases} \tag{7}$$

Lemma 1 [22] Let $H(\xi)$ be a continuous function defined on a compact set Ω_ξ. Then for $\forall\varepsilon > 0$, there exist the FLS $\zeta^T\Psi(\xi)$ such that

$$\sup_{\xi\in\Omega_\xi} |H(\xi) - \zeta^T\Psi(\xi)| \le \varepsilon \tag{8}$$

where $\zeta = [\zeta_1, \zeta_2, \ldots, \zeta_m]^T \in \mathbb{R}^m$ is the weight vector and $\Psi(\xi) = [\psi_1(\xi), \psi_2(\xi), \ldots, \psi_m(\xi)]^T$ is the FLS basis function with $m > 1$ is the number of FLS rules. $\psi_i(\xi) = exp[-\|\xi - \xi_i\|^2/\vartheta_i^2], i = 1, 2, \ldots, m$ is the Gaussian function, where ϑ_i and $\xi_i = [\xi_{i1}, \xi_{i2}, \ldots, \xi_{im}]^T$ represent the width and center, respectively. The optimal parameter vector ζ^* of FLS is defined as

$$\zeta^* = \arg\min_{\zeta\in\mathbb{R}^m}\{\sup_{\xi\in\Omega_\xi} |H(\xi) - \zeta^T\Psi(\xi)|\} \tag{9}$$

Therefore, the continuous function $H(\xi)$ can be expressed as

$$H(\xi) = {\zeta^*}^T\Psi(\xi) + \varepsilon(\xi) \tag{10}$$

where $\varepsilon(\xi)$ is the FLS approximation error, which can be bounded by $|\varepsilon(\xi)| \le \overline{\varepsilon}$, where $\overline{\varepsilon}$ is a positive constant. It should be pointed out that since ζ^* is an analytical quantity, it needs to be estimated later for practical use.

In order to achieve the desired tracking control objectives, we adopt the following assumptions

Assumption 1: The reference signal y_r and its first-order derivative $\dot{y}_r$ satisfy $|y_r| \le \overline{r}$ and $|\dot{y}_r| \le \overline{\dot{r}}$ where $\overline{r}$ and $\overline{\dot{r}}$ are unknown positive constants.

Assumption 2: For the nonaffine nonlinear term in (6), there exist unknown positive constants $\underline{f}$ and $\overline{f}$ such that

$$0 < \underline{f} \leq \frac{\partial f(\xi, u)}{\partial u} \leq \overline{f} \tag{11}$$

III. MAIN RESULT

In this section, we combine the reinforcement learning algorithm with the optimal backstepping control method to design a backstepping control strategy under the critic-actor architecture, thereby constructing an optimal controller.

A. Optimized backstepping design

First, consider the following tracking error coordinate transformation:

$$\begin{aligned} z_1 &= \xi_1 - y_r \\ z_2 &= \xi_2 - \hat{\alpha}_1^* \end{aligned} \tag{12}$$

where y_r is selected as the reference signal and set to $0.2\sin(t)$. α_1 and $\hat{\alpha}_1^*$ represent the virtual control and actual optimal virtual control correspondingly.

Step 1: From (4) and (12), the derivative of z_1 can be calculated

$$\dot{z}_1 = \xi_2 - \dot{y}_r \tag{13}$$

The optimal performance index function is chosen as

$$J_1(z_1) = \int_t^{\infty} h_1\Big(z_1(v), \alpha_1\big(z_1(v)\big)\Big) dv \tag{14}$$

where $h_1(z_1, \alpha_1) = z_1^2 + \alpha_1^2$ is the cost function, and let the optimal virtual control α_1^* replace α_1 in (13), the optimal performance index function can be obtained

$$\begin{aligned} J_1^*(z_1) &= \int_t^{\infty} h_1(z_1(v), \alpha_1^*(z_1(v)))dv \\ &= \min_{\alpha_1 \in \Omega_{z_1}} \{\int_t^{\infty} h_1\Big(z_1(v), \alpha_1\big(z_1(v)\big)\Big) dv\} \end{aligned} \tag{15}$$

Replace ξ_2 in (12) with the optimal virtual control α_1^*, and subsequently define the HJB equation associated with (12) and (14) as

$$H_1(z_1, \alpha_1^*, \frac{\mathrm{d}J_1^*}{\mathrm{d}z_1}) = z_1^2 + {\alpha_1^*}^2 + \frac{\mathrm{d}J_1^*}{\mathrm{d}z_1}(\alpha_1^* - \dot{y}_r) = 0 \tag{16}$$

The optimal virtual control α_1^* can be computed by solving $\partial H_1 / \partial \alpha_1^* = 0$ as

$$\alpha_1^* = -\frac{1}{2}\frac{\mathrm{d}J_1^*(z_1)}{\mathrm{d}z_1} \tag{17}$$

Then, $\frac{\mathrm{d}J_1^*(z_1)}{\mathrm{d}z_1}$ is decomposed into

$$\frac{\mathrm{d}J_1^*(z_1)}{\mathrm{d}z_1} = 2\chi_1 z_1 + J_1^o(\xi_1, z_1) \tag{18}$$

where $\chi_1 > \frac{5}{4}$ is design parameter. $J_1^o(\xi_1, z_1) = -\chi_1 z_1 + \frac{\mathrm{d}J_1^*(z_1)}{\mathrm{d}z_1} \in \mathbb{R}$ is a continuous function, and substituting (18) into (17) has

$$\alpha_1^* = -\chi_1 z_1 - \frac{1}{2}J_1^o(\xi_1, z_1) \tag{19}$$

Since $J_1^o(\xi_1, z_1)$ is continuous unknown function, it can be approximated by FLS as follows:

$$J_1^o(\xi_1, z_1) = \zeta_{J1}^{*T}\Psi_{J1}(\xi_1, z_1) + \varepsilon_{J1}(\xi_1, z_1) \tag{20}$$

where ζ_{J1}^* represents the ideal weight vector, $\Psi_{J1}(\xi_1, z_1)$ is the basis function vector, and $\varepsilon_{J1}(\xi_1, z_1)$ represents the approximation error bounded by $|\varepsilon_{J1}(\xi_1, z_1)| \leq \overline{\varepsilon}_{J1}$ as arbitrarily small. Then, (18) and (19) can be reorganized as

$$\frac{\mathrm{d}J_1^*(z_1)}{\mathrm{d}z_1} = 2\chi_1 z_1 + \zeta_{J1}^{*T}\Psi_{J1} + \varepsilon_{J1} \tag{21}$$

$$\alpha_1^* = -\chi_1 z_1 - \frac{1}{2}\zeta_{J1}^{*T}\Psi_{J1} - \frac{1}{2}\varepsilon_{J1} \tag{22}$$

Since ζ_{J1}^* is unknown constant vector, the optimal virtual control (22) is not available for the controlled system. To derive the effective optimized virtual control, the following RL algorithm with critic and actor is performed.

$$\frac{\mathrm{d}\hat{J}_1^*(z_1)}{\mathrm{d}z_1} = 2\chi_1 z_1 + \hat{\zeta}_{c1}^T\Psi_{J1} \tag{23}$$

$$\hat{\alpha}_1^* = -\chi_1 z_1 - \frac{1}{2}\hat{\zeta}_{a1}^T\Psi_{J1} \tag{24}$$

where $\frac{\mathrm{d}\hat{J}_1^*(z_1)}{\mathrm{d}z_1}$ and $\hat{\alpha}_1^*$ are the estimates of $\frac{\mathrm{d}J_1^*(z_1)}{\mathrm{d}z_1}$ and α_1^*, respectively. $\hat{\zeta}_{c1}^T\Psi_{J1}$ and $\hat{\zeta}_{a1}^T\Psi_{J1}$ are the FLS weight vectors of critic and actor, respectively.

Following this, the weight vectors of the neural networks for both the critic and actor are trained according to the respective adaptive laws outlined below.

$$\dot{\hat{\zeta}}_{c1} = -\kappa_{c1}\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1} \tag{25}$$

$$\dot{\hat{\zeta}}_{a1} = -\Psi_{J1}\Psi_{J1}^T\big(\kappa_{a1}(\hat{\zeta}_{a1} - \hat{\zeta}_{c1}) + \kappa_{c1}\hat{\zeta}_{c1}\big) \tag{26}$$

where $\kappa_{c1} > 0$ and $\kappa_{a1} > 0$ represent critic and actor design parameters, while κ_{c1} and κ_{a1} satisfy $\kappa_{a1} > \frac{1}{2}$, $\kappa_{a1} > \frac{\kappa_{c1}}{2}$.

Using (24), (13) can be rewritten as

$$\dot{z}_1 = z_2 - \chi_1 z_1 - \frac{1}{2}\hat{\zeta}_{a1}^T\Psi_{J1} - \dot{y}_r \tag{27}$$

For the first backstepping step, the Lyapunov function V_1 is designed as follows:

$$V_1 = \frac{1}{2}z_1^2 + \frac{1}{2}\tilde{\zeta}_{c1}^T\tilde{\zeta}_{c1} + \frac{1}{2}\tilde{\zeta}_{a1}^T\tilde{\zeta}_{a1} \tag{28}$$

where $\tilde{\zeta}_{c1} = \zeta_{J1}^* - \hat{\zeta}_{c1}$ and $\tilde{\zeta}_{a1} = \zeta_{J1}^* - \hat{\zeta}_{a1}$ are the estimation errors of the critic and the actor, respectively.

Then, the derivative of V_1 is

$$\begin{aligned} \dot{V}_1 =& z_1\big(z_2 - \chi_1 z_1 - \frac{1}{2}\hat{\zeta}_{a1}^T\Psi_{J1} - \dot{y}_r\big) + \kappa_{c1}\tilde{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1} \\ &+ \tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\big(\kappa_{a1}(\hat{\zeta}_{a1} - \hat{\zeta}_{c1}) + \kappa_{c1}\hat{\zeta}_{c1}\big) \end{aligned} \tag{29}$$

The Young's inequality yields the following results

$$\begin{aligned} z_1 z_2 &\leq \frac{1}{2}z_1^2 + \frac{1}{2}z_2^2 \\ -z_1\dot{y}_r &\leq \frac{1}{2}z_1^2 + \frac{1}{2}\dot{y}_r^2 \\ -\frac{1}{2}z_1\hat{\zeta}_{a1}^T\Psi_{J1} &\leq \frac{1}{4}z_1^2 + \frac{1}{4}\hat{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{a1} \end{aligned} \tag{30}$$

Along with (29) and (30), we can calculate:

$$\begin{aligned} \dot{V}_1 \leq& -(\chi_1 - \frac{5}{4})z_1^2 + \kappa_{c1}\tilde{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1} \\ &+ \kappa_{a1}\tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{a1} + \frac{1}{2}z_2^2 + \frac{1}{2}\dot{y}_r^2 \\ &+ (\kappa_{c1} - \kappa_{a1})\tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1} \\ &+ \frac{1}{4}\hat{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{a1} \end{aligned} \tag{31}$$

Based on $\tilde{\zeta}_{c1}=\zeta^*_{J1}-\hat{\zeta}_{c1}$, $\tilde{\zeta}_{a1}=\zeta^*_{J1}-\hat{\zeta}_{a1}$ and Young's inequality, we have

$$\begin{aligned}\tilde{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1}=&\frac{1}{2}\zeta_{J1}^{*T}\Psi_{J1}\Psi_{J1}^T\zeta_{J1}^*-\frac{1}{2}\tilde{\zeta}_{c1}^T\Psi_{J1}\\&\times\Psi_{J1}^T\tilde{\zeta}_{c1}-\frac{1}{2}\hat{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1}\\\tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{a1}=&\frac{1}{2}\zeta_{J1}^{*T}\Psi_{J1}\Psi_{J1}^T\zeta_{J1}^*-\frac{1}{2}\tilde{\zeta}_{a1}^T\Psi_{J1}\\&\times\Psi_{J1}^T\tilde{\zeta}_{a1}-\frac{1}{2}\hat{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{a1}\\\tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1}\le&-\frac{1}{2}\tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\tilde{\zeta}_{a1}\\&-\frac{1}{2}\hat{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1}\end{aligned}\tag{32}$$

Subsequently, we can acquire

$$\begin{aligned}\dot{V}_1\le&-(\chi_1-\frac{5}{4})z_1^2-\frac{\kappa_{c1}}{2}\tilde{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\tilde{\zeta}_{c1}\\&-(\kappa_{a1}-\frac{\kappa_{c1}}{2})\tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\tilde{\zeta}_{a1}\\&-\frac{\kappa_{a1}}{2}\hat{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{c1}-(\frac{\kappa_{a1}}{2}-\frac{1}{4})\\&\times\hat{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\hat{\zeta}_{a1}+\frac{1}{2}z_2^2+\frac{1}{2}\dot{y}_r^2\\&+\frac{\kappa_{c1}+\kappa_{a1}}{2}\zeta_{J1}^{*T}\Psi_{J1}\Psi_{J1}^T\zeta_{J1}^*\end{aligned}\tag{33}$$

The following inequality holds when $\lambda_{\Psi_{J1}}^{\min}$ is the minimum eigenvalue of $\Psi_{J1}\Psi_{J1}^T$.

$$\begin{aligned}&-\tilde{\zeta}_{c1}^T\Psi_{J1}\Psi_{J1}^T\tilde{\zeta}_{c1}\le-\lambda_{\Psi_{J1}}^{\min}\tilde{\zeta}_{c1}^T\tilde{\zeta}_{c1}\\&-\tilde{\zeta}_{a1}^T\Psi_{J1}\Psi_{J1}^T\tilde{\zeta}_{a1}\le-\lambda_{\Psi_{J1}}^{\min}\tilde{\zeta}_{a1}^T\tilde{\zeta}_{a1}\end{aligned}\tag{34}$$

According to the design parameters $\kappa_{a1}>\frac{\kappa_{c1}}{2}$ and $\kappa_{a1}>\frac{1}{2}$, as well as (34), it can yield

$$\begin{aligned}\dot{V}_1\le&-(\chi_1-\frac{5}{4})z_1^2-\frac{\kappa_{c1}}{2}\lambda_{\Psi_{J1}}^{\min}\tilde{\zeta}_{c1}^T\tilde{\zeta}_{c1}\\&-(\kappa_{a1}-\frac{\kappa_{c1}}{2})\lambda_{\Psi_{J1}}^{\min}\tilde{\zeta}_{a1}^T\tilde{\zeta}_{a1}+\frac{1}{2}z_2^2+\sigma_1\end{aligned}\tag{35}$$

where $\sigma_1=\frac{1}{2}\dot{y}_r^2+\frac{\kappa_{c1}+\kappa_{a1}}{2}\zeta_{J1}^{*T}\Psi_{J1}\Psi_{J1}^T\zeta_{J1}^*$. Since all the terms in σ_1 are bounded, there exists a positive constant $\overline{\sigma}_1$ such that $|\sigma_1|\le\overline{\sigma}_1$.

Step 2 : The derivative of z_2 is calculated in a similar manner.

$$\dot{z}_2=u+H(\xi,u)-\dot{\alpha}_1^*\tag{36}$$

Among them, $H(\xi,u)$ can be approximated by FLS as $\zeta_{f2}^{*T}\Psi_{f2}+\varepsilon_{f2}$, there exists a positive constant $\overline{\varepsilon}_{f2}$ such that $|\varepsilon_{f2}|\le\overline{\varepsilon}_{f2}$. Then, the selection of the most suitable integral cost function is detailed as follows:

$$\begin{aligned}J_2^*(z_2)=&\int_t^\infty h_2\Big(z_2(v),u^*\big(z_2(v)\big)\Big)dv\\=&\min_{u\in\Omega_{z_2}}\{\int_t^\infty h_2\Big(z_2(v),u\big(z_2(v)\big)\Big)dv\}\end{aligned}\tag{37}$$

where $h_2(z_2,u)=z_2^2+u^2$ is the cost function, u^* represents the optimal controller.

Based on (37), the HJB equation is constructed as

$$\begin{aligned}H_2(z_2,u^*,\frac{\mathrm{d}J_2^*}{\mathrm{d}z_2})=&z_2^2+u^{*2}+\frac{\mathrm{d}J_2^*}{\mathrm{d}z_2}\big(u^*+\zeta_{f2}^{*T}\Psi_{f2}(\xi)\\&+\varepsilon_f(\xi)-\dot{\alpha}_1^*\big)=0\end{aligned}\tag{38}$$

The same as before, we can solve for $\partial H_2/\partial u^*=0$ as

$$u^*=-\frac{1}{2}\frac{\mathrm{d}J_2^*(z_2)}{\mathrm{d}z_2}\tag{39}$$

Then, $\frac{\mathrm{d}J_2^*(z_2)}{\mathrm{d}z_2}$ can be factored as

$$\frac{\mathrm{d}J_2^*(z_2)}{\mathrm{d}z_2}=2\chi_2z_2+2\zeta_{f2}^{*T}\Psi_{f2}+2\varepsilon_{f2}+J_2^o(\xi_2,z_2)\tag{40}$$

where $\chi_2>0$ is design parameter. $J_2^o(\xi_2,z_2)=-2\chi_2z_2-2\zeta_{f2}^{*T}\Psi_{f2}-2\varepsilon_{f2}+\frac{\mathrm{d}J_2^*(z_2)}{\mathrm{d}z_2}$ is a continuous function, and the u^* can be expressed as

$$u^*=-\chi_2z_2-\zeta_{f2}^{*T}\Psi_{f2}-\varepsilon_{f2}-\frac{1}{2}J_2^o(\xi_2,z_2)\tag{41}$$

Since $J_2^o(\xi_2,z_2)$ is unknown continuous term, it can also be approximated using FLS as follows:

$$J_2^o(\xi_2,z_2)=\zeta_{J2}^{*T}\Psi_{J2}+\varepsilon_{J2}\tag{42}$$

where ζ_{J2}^* is the ideal weight vector, Ψ_{J2} is the FLS basis function vector, and the FLS approximation error ε_{J2} is bounded.

Similarly, we can derive the following conclusion

$$\frac{\mathrm{d}J_2^*(z_2)}{\mathrm{d}z_2}=2\chi_2z_2+2\zeta_{f2}^{*T}\Psi_{f2}+\zeta_{J2}^{*T}\Psi_{J2}+\varepsilon_2\tag{43}$$

$$u^*=-\chi_2z_2-\zeta_{f2}^{*T}\Psi_{f2}-\frac{1}{2}\zeta_{J2}^{*T}\Psi_{J2}-\frac{1}{2}\varepsilon_2\tag{44}$$

where $\varepsilon_2=2\varepsilon_{f2}+\varepsilon_{J2}$.

The optimal control (44), however, remains unattainable, necessitating the execution of an RL algorithm featuring both a critic and an actor to acquire viable control signal.

$$\frac{\mathrm{d}\hat{J}_2^*(z_2)}{\mathrm{d}z_2}=2\chi_2z_2+2\hat{\zeta}_{f2}^T\Psi_{f2}+\hat{\zeta}_{c2}^T\Psi_{J2}\tag{45}$$

$$\hat{u}^*=-\chi_2z_2-\hat{\zeta}_{f2}^T\Psi_{J2}-\frac{1}{2}\hat{\zeta}_{a2}^T\Psi_{J2}\tag{46}$$

where $\frac{\mathrm{d}\hat{J}_2^*(z_2)}{\mathrm{d}z_2}$ and $\hat{\alpha}_2^*$ are the estimate of $\frac{\mathrm{d}J_2^*(z_2)}{\mathrm{d}z_2}$ and u^*, respectively. $\hat{\zeta}_{c2}^T\Psi_{J2}$ and $\hat{\zeta}_{a2}^T\Psi_{J2}$ are the FLS weight vectors of critic and actor, respectively.

Same as the first step, the corresponding three adaptive update laws are designed as follows:

$$\dot{\hat{\zeta}}_{f2}=\Gamma_{f2}\Psi_{J2}z_2-\kappa_{f2}\hat{\zeta}_{f2}\tag{47}$$

$$\dot{\hat{\zeta}}_{c2}=-\kappa_{c2}\Psi_{J2}\Psi_{J2}^T\hat{\zeta}_{c2}\tag{48}$$

$$\dot{\hat{\zeta}}_{a2}=-\Psi_{J2}\Psi_{J2}^T\big(\kappa_{a2}(\hat{\zeta}_{a2}-\hat{\zeta}_{c2})+\kappa_{c2}\hat{\zeta}_{c2}\big)\tag{49}$$

where $\Gamma_{f2}>0$, $\kappa_{f2}>0$, $\kappa_{c2}>0$ and $\kappa_{a2}>0$ are design parameters, while κ_{c2} and κ_{a2} satisfy $\kappa_{a2}>\frac{1}{2}$, $\kappa_{a2}>\frac{\kappa_{c2}}{2}$.

According to (46), the $\dot{z}_2$ can be expressed as follows

$$\dot{z}_2=-\chi_2z_2-\frac{1}{2}\hat{\zeta}_{a2}^T\Psi_{J2}+\tilde{\zeta}_{f2}^T\Psi_{f2}+\varepsilon_{f2}-\dot{\alpha}_1^*\tag{50}$$

Subsequently, the Lyapunov function V_2 is established as

$$V_2=\frac{1}{2}z_2^2+\frac{1}{2\Gamma_{f2}}\tilde{\zeta}_{f2}^T\tilde{\zeta}_{f2}+\frac{1}{2}\tilde{\zeta}_{c2}^T\tilde{\zeta}_{c2}+\frac{1}{2}\tilde{\zeta}_{a2}^T\tilde{\zeta}_{a2}\tag{51}$$

where $\tilde{\zeta}_{f2}=\zeta_{f2}^*-\hat{\zeta}_{f2}$, $\tilde{\zeta}_{c2}=\zeta_{J2}^*-\hat{\zeta}_{c2}$ and $\tilde{\zeta}_{a2}=\zeta_{J2}^*-\hat{\zeta}_{a2}$.

Then, the $\dot{V}_2$ can be calculated as

$$\begin{aligned}\dot{V}_2 =& z_2\big(-\chi_2 z_2 - \frac{1}{2}\hat{\zeta}_{a2}^T\Psi_{J2} + \tilde{\zeta}_{f2}^T\Psi_{f2} + \varepsilon_{f2} - \dot{\hat{\alpha}}_1^*\big)\\ &+ \tilde{\zeta}_{a2}^T\Psi_{J2}\Psi_{J2}^T\big(\kappa_{a2}(\hat{\zeta}_{a2} - \hat{\zeta}_{c2}) + \kappa_{c2}\hat{\zeta}_{c2}\big)\\ &+ \frac{\kappa_{f2}}{\Gamma_{f2}}\tilde{\zeta}_{f2}^T\hat{\zeta}_{f2} + \kappa_{c2}\tilde{\zeta}_{c2}^T\Psi_{J2}\Psi_{J2}^T\hat{\zeta}_{c2}\end{aligned} \tag{52}$$

Using the Young's inequality, we have

$$\begin{aligned} z_2\varepsilon_{f2} &\le \frac{1}{2}z_2^2 + \frac{1}{2}\bar{\varepsilon}_{f2}^2\\ -z_2\dot{\hat{\alpha}}_1^* &\le \frac{1}{2}z_2^2 + \frac{1}{2}\dot{\hat{\alpha}}_1^{*2}\\ -\frac{1}{2}z_2\hat{\zeta}_{a2}^T\Psi_{J2} &\le \frac{1}{4}z_2^2 + \frac{1}{4}\hat{\zeta}_{a2}^T\Psi_{J2}\Psi_{J2}^T\hat{\zeta}_{a2}\end{aligned} \tag{53}$$

Substituting (53) into (52) yields

$$\begin{aligned}\dot{V}_2 \le& -(\chi_2 - \frac{5}{4})z_2^2 - \frac{\kappa_{f2}}{2\Gamma_{f2}}\tilde{\zeta}_{f2}^T\tilde{\zeta}_{f2} - \frac{\kappa_{c2}}{2}\tilde{\zeta}_{c2}^T\Psi_{J2}\Psi_{J2}^T\tilde{\zeta}_{c2}\\ &-(\kappa_{a2} - \frac{\kappa_{c2}}{2})\tilde{\zeta}_{a2}^T\Psi_{J2}\Psi_{J2}^T\tilde{\zeta}_{a2} - \frac{\kappa_{a2}}{2}\hat{\zeta}_{c2}^T\Psi_{J2}\Psi_{J2}^T\hat{\zeta}_{c2}\\ &-(\frac{\kappa_{a2}}{2} - \frac{1}{4})\hat{\zeta}_{a2}^T\Psi_{J2}\Psi_{J2}^T\hat{\zeta}_{a2} + \frac{\kappa_{c2}+\kappa_{a2}}{2}\zeta_{J2}^{*T}\Psi_{J2}\\ &\times\Psi_{J2}^T\zeta_{J2}^* + \frac{1}{2}\bar{\varepsilon}_{f2}^2 + \frac{1}{2}\dot{\hat{\alpha}}_1^{*2} + \frac{\kappa_{f2}}{2}\zeta_{f2}^{*T}\zeta_{f2}^* + \frac{1}{2}z_3^2\\ \le& -(\chi_2 - \frac{5}{4})z_2^2 - \frac{\kappa_{f2}}{2\Gamma_{f2}}\tilde{\zeta}_{f2}^T\tilde{\zeta}_{f2} - \frac{\kappa_{c2}}{2}\lambda_{\Psi_{J2}}^{\min}\tilde{\zeta}_{c2}^T\tilde{\zeta}_{c2}\\ &-(\kappa_{a2} - \frac{\kappa_{c2}}{2})\lambda_{\Psi_{J2}}^{\min}\tilde{\zeta}_{a2}^T\tilde{\zeta}_{a2} - \frac{1}{2}z_2^2 + \frac{1}{2}z_3^2 + \sigma_2\end{aligned} \tag{54}$$

where $\sigma_2 = \frac{1}{2}\bar{\varepsilon}_{f2}^2 + \frac{\kappa_{c2}+\kappa_{a2}}{2}\zeta_{J2}^{*T}\Psi_{J2}\Psi_{J2}^T\zeta_{J2}^* + \frac{\kappa_{f2}}{2}\zeta_{f2}^{*T}\zeta_{f2}^* + \frac{1}{2}\dot{\hat{\alpha}}_1^{*2}$ is bounded, and there exists a positive constant $\bar{\sigma}_2$ that ensures the existence of $|\sigma_2| \le \bar{\sigma}_2$. Additionally, $\lambda_{\Psi_{J2}}^{\min}$ represents the minimum eigenvalue of $\Psi_{J2}\Psi_{J2}^T$.

B. Stability analysis

Theorem 1 The optimal backstepping control strategy proposed in this paper is applied to TRICS, where the adaptive laws of the fuzzy parameters, the critic and the actor are respectively (47) and (25), (48) and (26), (49), and the optimal virtual control and the optimal controller are respectively (24) and (46). Based on this, it can be concluded that this optimal control strategy can ensure that all signals in the closed-loop system remain bounded, and achieve the precise decoupling and performance optimization between the molten pool liquid level height and the roll gap of TRICS.

Proof: Construct a Lyapunov function $V = \sum_{i=1}^{2} V_i$, we can compute

$$\begin{aligned}\dot{V} \le& -\sum_{i=1}^{2}(\chi_i - \frac{5}{4})z_i^2 - \frac{\kappa_{f2}}{2\Gamma_{f2}}\tilde{\zeta}_{fi}^T\tilde{\zeta}_{fi} - \sum_{i=1}^{2}\frac{\kappa_{ci}}{2}\lambda_{\Psi_{Ji}}^{\min}\tilde{\zeta}_{ci}^T\tilde{\zeta}_{ci}\\ &-\sum_{i=1}^{2}(\kappa_{ai} - \frac{\kappa_{ci}}{2})\lambda_{\Psi_{Ji}}^{\min}\tilde{\zeta}_{ai}^T\tilde{\zeta}_{ai} + \sum_{i=1}^{2}\bar{\sigma}_i\\ \le& -\Theta V + \Delta\end{aligned} \tag{55}$$

where $\Theta = \min\{2\chi_i - \frac{5}{2}, \frac{\kappa_{f2}}{\Gamma_{f2}}, \kappa_{ci}, (\kappa_{ai} - \frac{\kappa_{ci}}{2})\lambda_{\Psi_{Ji}}^{\min}, i=1,2\}, \Delta = \sum_{i=0}^{2}\bar{\sigma}_i$.

■

The proof of Theorem 1 is completed.

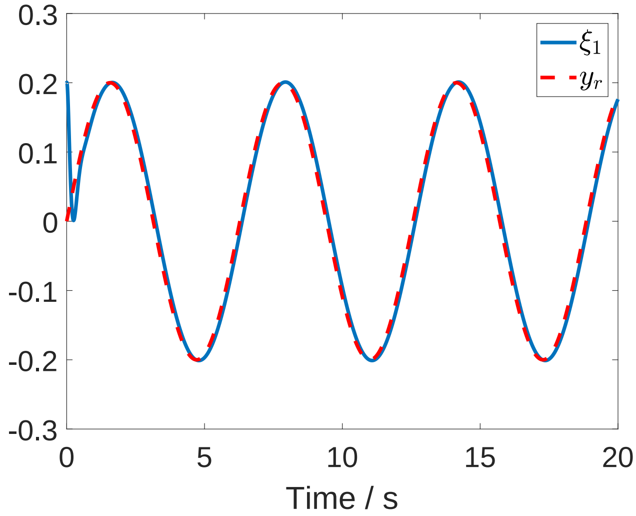

Fig. 2: State ξ_1 and reference signal y_r.

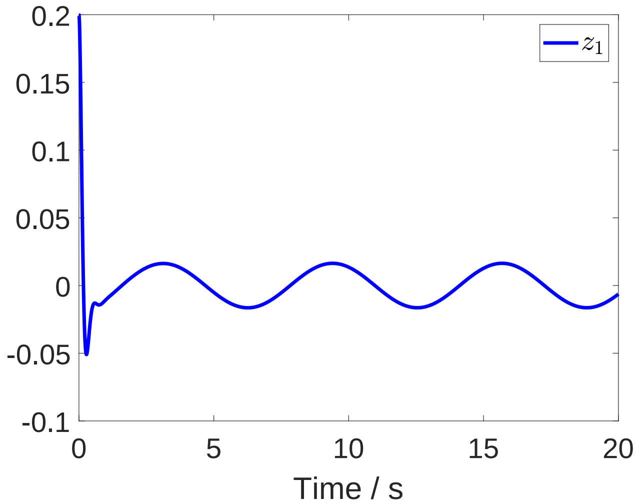

Fig. 3: Tracking error z_1.

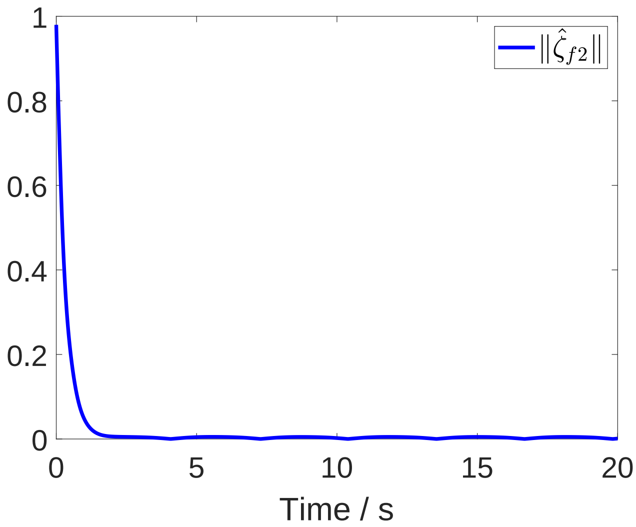

Fig. 4: The norm of the $\hat{\zeta}_{f2}$.

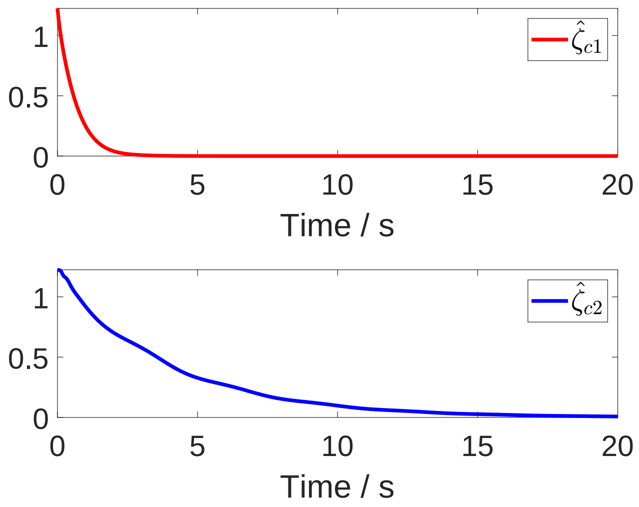

Fig. 5: The norms of the $\hat{\zeta}_{c1}$ and $\hat{\zeta}_{c2}$.

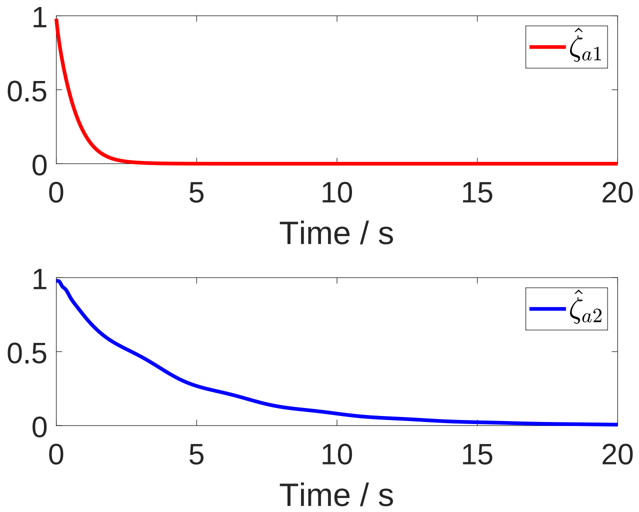

Fig. 6: The norms of the $\hat{\zeta}_{a1}$ and $\hat{\zeta}_{a2}$.

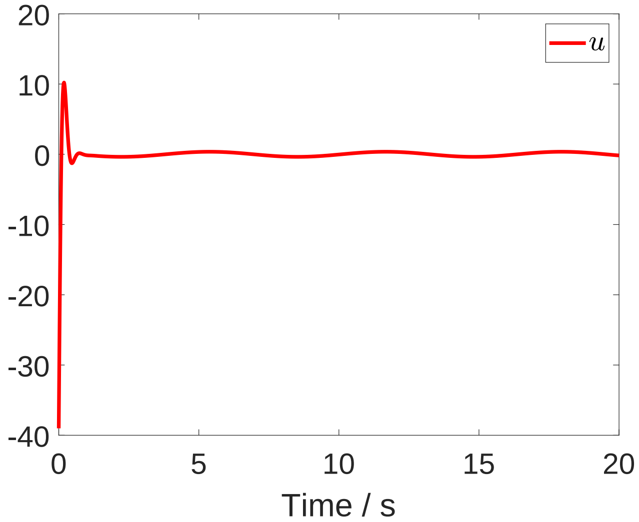

Fig. 7: Control input u.

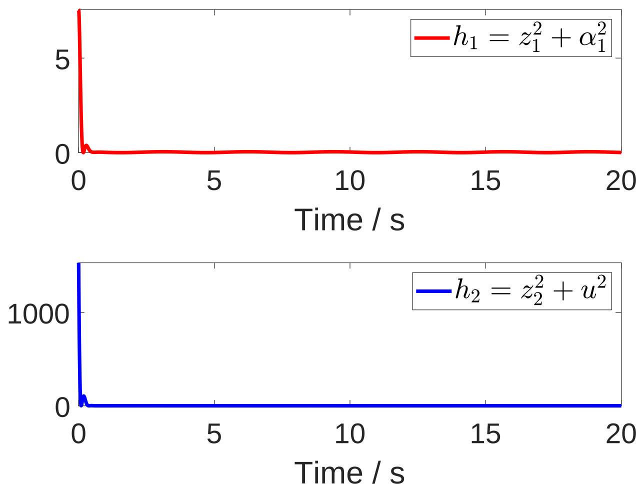

Fig. 8: Cost functions h_1 and h_2.

IV. SIMULATION EXAMPLE

The effectiveness of the control algorithm proposed in this paper was verified through simulation examples. The parameters used in the simulation process are summarized as follows:

The corresponding process parameters in the TRICS are $R = 150\ mm$, $L = 200\ mm$, $\omega = 170\ mm/s$ and $\beta = 5°$. The reference signal is selected as $y_r = 0.2\sin(t)$. The control parameters are designed as $\chi_1 = 8$, $\chi_2 = 12$, $\kappa_{f2} = 15$, $\kappa_{c1} = \kappa_{c2} = 10$, $\kappa_{a1} = \kappa_{a2} = 12$. Furthermore, the initial values are set as $\xi_1(0) = 0.2$, $\xi_2(0) = 0.2$, $\hat{\zeta}_{f2}(0) = [0.4, \ldots, 0.4]^T \in \mathbb{R}^{6\times1}$, $\hat{\zeta}_{c1}(0) = \hat{\zeta}_{a1}(0) = [0.5, \ldots, 0.5]^T \in \mathbb{R}^{6\times1}$, $\hat{\zeta}_{c2}(0) = \hat{\zeta}_{a2}(0) = [0.4, \ldots, 0.4]^T \in \mathbb{R}^{6\times1}$.

The simulation results show that the dual neural network structure based on the critic-actor framework proposed in this paper can efficiently evaluate the value function of the current control strategy, generate adaptive compensation terms to optimize the control law, and ensure that the output error of the TRICS is always within the preset range by dynamically adjusting the system control gain online. Figures 2 and 3 verify the excellent tracking performance of the system under this strategy. Figures 4-8 further demonstrate that the critic adaptive law, the actor adaptive law, and the optimal controller designed in this paper all exhibit fast convergence characteristics and maintain stability, thereby effectively achieving the optimal control state of the system.

V. CONCLUSION

This study demonstrates that the fuzzy optimal control framework based on the actor-critic architecture established for TRICS has significant advantages in terms of theoretical innovation and method integration. This approach achieves a deep integration of nonlinear dynamic characteristics and multi-objective optimization through an adaptive mechanism. The internal stability analysis and the accessibility demonstration of performance indicators provide a new theoretical perspective for the control of complex industrial processes. Future research can further explore the deployment of the algorithm to physical entity systems in the context of intelligent manufacturing, verify its multi-condition robustness and real-time performance, and explore the extended application of multi-agent coordination mechanisms in the group control of heterogeneous equipment, providing technical support for the full-process intelligent casting.

REFERENCES

[1] C. Santos, J. Spim Jr, and A. Garcia, "Modeling of solidification in twin-roll strip casting," *Journal of Materials Processing Technology*, vol. 102, no. 1-3, pp. 33–39, 2000.

[2] Z. Yu, Z. S. Lei, Q. S. Li, K. Deng, and Z. M. Ren, "Physical simulation of nozzle electromagnetic brake in twin-roll strip casting," *Steel Research International*, vol. 79, no. 11, pp. 839–842, 2008.

[3] M. Vidoni, M. Daamen, and G. Hirt, "Advances in the twin-roll strip casting of strip with profiled cross section," *Key Engineering Materials*, vol. 554, pp. 562–571, 2013.

[4] M. Daamen, C. Haase, J. Dierdorf, D. A. Molodov, and G. Hirt, "Twin-roll strip casting: A competitive alternative for the production of high-manganese steels with advanced mechanical properties," *Materials Science and Engineering: A*, vol. 627, pp. 72–81, 2015.

[5] T. Haga and Y. Kurahashi, "Casting process at roll bite in strip cast using vertical-type high-speed twin-roll caster," *Metals*, vol. 12, no. 7, p. 1169, 2022.

[6] Y. J. Zhang, L. B. Wu, H. Y. Zhao, X. D. Hu, W. Y. Zhang, and D. Y. Ju, "Robust adaptive fuzzy output tracking control for a class of twin-roll strip casting systems," *Mathematical Problems in Engineering*, vol. 6742630, 2017.

[7] D. Gao, Y. Zhang, L. Wu, and S. Liu, "Adaptive neural command filtered fault-tolerant control for a twin roll inclined casting system," *Metalurgija*, vol. 62, no. 3-4, pp. 355–358, 2023.

[8] D. Gao, Y. Zhang, L. Wu, and S. Liu, "Mathematical modeling and command filter adaptive fuzzy control based on twin-roll inclined strip casting process," *Journal of Control, Automation and Electrical Systems*, vol. 34, no. 6, pp. 1220–1230, 2023.

[9] Y. Lian, J. Xia, J. H. Park, W. Sun, and H. Shen, "Disturbance observer-based adaptive neural network output feedback control for uncertain nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 10, pp. 7260–7270, 2022.

[10] M. Liu, L. Wu, H. Wang, L. Guo, and S. Huang, "Disturbance observer-based nonsingular fixed-time fuzzy adaptive event-triggered output feedback control of uncertain nonlinear systems," *Fuzzy Sets and Systems*, vol. 498, p. 109132, 2025.

[11] X. Yang, D. Liu, and Q. Wei, "Online approximate optimal control for affine non-linear systems with unknown internal dynamics using adaptive dynamic programming," *IET Control Theory & Applications*, vol. 8, no. 16, pp. 1676–1688, 2014.

[12] Y. Zhu, D. Zhao, and Z. Zhong, "Adaptive optimal control of heterogeneous cacc system with uncertain dynamics," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 4, pp. 1772–1779, 2018.

[13] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems," *Automatica*, vol. 49, no. 1, pp. 82–92, 2013.

[14] P. Werbos, "Approximate dynamic programming for real-time control and neural modeling," *Handbook of intelligent control*, 1992.

[15] D. Liu, D. Wang, F. Wang, H. Li, and X. Yang, "Neural-network-based online hjb solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2834–2847, 2014.

[16] D. Wang, D. Liu, and H. Li, "Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 2, pp. 627–632, 2014.

[17] G. Wen, S. S. Ge, and F. Tu, "Optimized backstepping for tracking control of strict-feedback systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 8, pp. 3850–3862, 2018.

[18] G. Wen, C. P. Chen, and S. S. Ge, "Simplified optimized backstepping control for a class of nonlinear strict-feedback systems with unknown dynamic functions," *IEEE Transactions on Cybernetics*, vol. 51, no. 9, pp. 4567–4580, 2020.

[19] G. Wen, B. Li, and B. Niu, "Optimized backstepping control using reinforcement learning of observer-critic-actor architecture based on fuzzy system for a class of nonlinear strict-feedback systems," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 10, pp. 4322–4335, 2022.

[20] S. Liu, H. Yan, L. Zhao, and D. Gao, "Fault estimate and reinforcement learning based optimal output feedback control for single-link robot arm model," *Engineering Letters*, vol. 33, no. 1, pp. 21–28, 2025.

[21] D. Gao, W. Cui, L. Wu, Y. Zhang, and Y. Tao, "Fixed-time optimized control for nonlinear strict-feedback systems based on reinforcement learning and disturbance observer," *Applied Mathematics and Computation*, vol. 508, p. 129628, 2026.

[22] Q. Yu, J. Ding, L. Wu, and X. He, "Event-triggered prescribed time adaptive fuzzy fault-tolerant control for nonlinear systems with full-state constraints," *Engineering Letters*, vol. 32, no. 8, pp. 1577–1584, 2024.